Computer Aided Drug Designing (CADD) of *Pneumocystis Carinii* Pneumonia inhibitors using Pharmacophore, Molecular Docking and QSAR

Researcher

Yusra Sajid Kiani

32-FBAS/MSBI/F09

Supervisors

Dr. Naveeda Riaz

Mrs. Saima Kalsoom

Department of Environmental Sciences
Faculty of Basic and Applied Sciences
International Islamic University Islamabad

(2012)

1 - Environmental protection

2 - Environmental studies

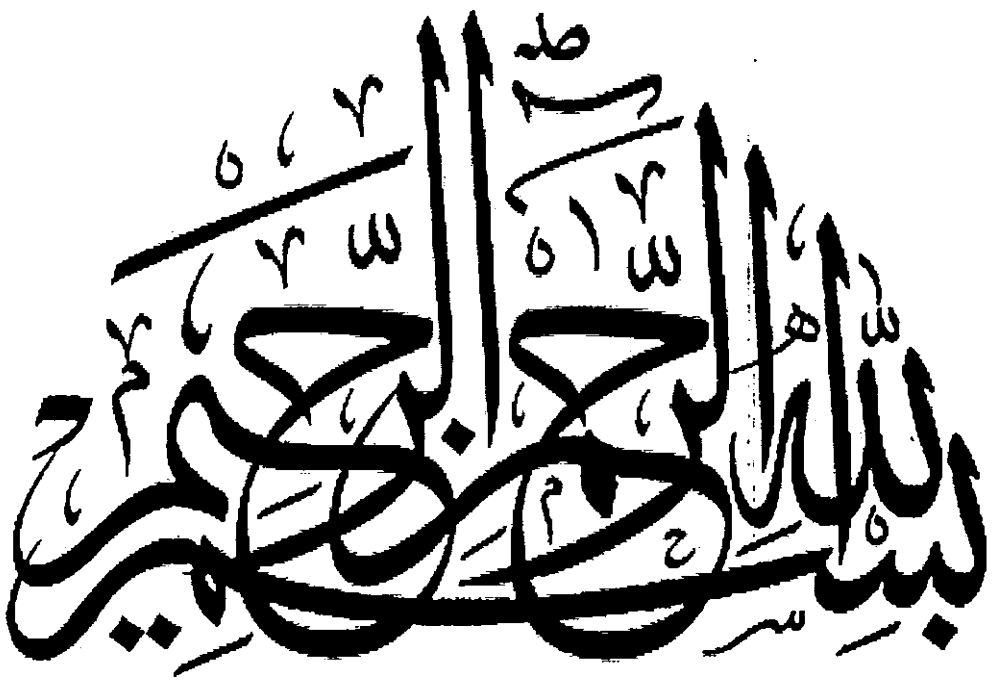
Computer Aided Drug Designing (CADD) of *Pneumocystis Carinii* Pneumonia inhibitors using Pharmacophore, Molecular Docking and QSAR

Researcher:

Yusra Sajid Kiani

32-FBAS/MSBI/F09

Department of Environmental Sciences
Faculty of Basic and Applied Sciences
International Islamic University Islamabad
(2012)



In the name of Allah Most Gracious and Most Beneficial

Department of Environmental Sciences
International Islamic University Islamabad

Dated: 10-02-2012

FINAL APPROVAL

It is certificate that we have read the thesis submitted by Ms. Yusra Sajid Kiani and it is our judgment that this project is of sufficient standard to warrant its acceptance by the International Islamic University, Islamabad for the M.S Degree in Bioinformatics.

COMMITTEE

External Examiner

Dr. Mahmood Kiani

Professor, Bioscience

CIIT Islamabad, Pakistan

10/2/11

Internal Examiner

Dr. Asma Gul

Assistant professor, Dept. of Environmental Science

International Islamic University, Islamabad

Supervisor

Dr.Naveeda Riaz

Assistant Professor, Dept. of Environmental Science

International Islamic University, Islamabad

Co-Supervisor

Mrs.Saima Kalsoom

PhD Scholar, Quaid-e- Azam University, Islamabad

Dean, FBAS

Dr. Irfan Khan

International Islamic University, Islamabad

A thesis submitted to Department of Environmental Sciences,
International Islamic University, Islamabad as a partial
fulfillment of requirement for the award of the
degree of MS in Bioinformatics.

Dedicated to my loving parents, the august and imperial people in my thoughts. Without their patience, understanding, support and love the completion of this work would have not been possible.

Finally, this thesis is also dedicated to all those who believe in the richness of learning

DECLARATION

I hereby declare that the work presented in the following thesis is my own effort, except where acknowledged otherwise, and that the thesis is my own composition. No part of the thesis has been previously presented for any other degree.

Date _____

Yusra Sajid Kiani

TABLE OF CONTENTS

ACKNOWLEDGMENTS

First of all I express gratitude to Allah (subhanawataala) for endowing me with the strength, devotion, health and knowledge to complete this thesis. It is His unlimited blessings that I have been able to come this far. I am indebted to His unlimited blessings that He showered upon me. All praises for His Holy Prophet Muhammad (S.A.W) who enabled us to recognize our Lord and Creator and brought to us the real source of knowledge from Allah, The Quran and who is the role model for us in every aspect of life.

I express deepest thanks and profound regards to my parents whose affection has been a source of encouragement for me and who have provided immense support throughout my life.

I acknowledge, with deep gratitude and appreciation, the inspiration, encouragement, valuable time and guidance given to me by **Mrs. Saima Kalsoom,** PhD Scholar, who served as my major advisor, enabled me to develop an understanding of the project and led to successful completion of the project. I express my sincerest gratitude to **Dr. Naveeda Riaz,** Assistant Professor, International Islamic University Islamabad for her extensive guidance, continuous support, and personal involvement in all phases of this research.

Lastly I offer my regards blessings to all those who supported me in any respect during the completion of the project. Special thanks to the department of Environmental Sciences (IIUI). My time at IIUI was made enjoyable in large part due to my friends Javeria Ashraf, Mehrin Gul, Saba Munawar, Syeda Uzma Ali and Zurah Bibi and all my

classmates that became a part of my life. Heartiest thanks to my classmates Asma Abro and Sumra Wajid for their help and support. I am grateful for the time spent with my friends.

The support and assistance of Pakeeza Akram student of Department of Environmental Science International Islamic University is acknowledged and appreciated. I must say, this work could not be accomplished without her assistance.

May Allah reward them all abundantly.I pray to Allah (SWT) that may He bestow all of us with true success in all fields in both worlds and shower His blessed knowledge upon us for the betterment of all Muslims and whole Mankind.

(Ameen)

Yusra Sajid Kiani

LIST OF ABBREVIATIONS

Abbreviations	Acronyms
2D	Two Dimensional
3D	Three Dimensional
Å	Angstrom
AIDS	Acquired Immune Deficiency Syndrome
Ala	Alanine
Ar	Aromatic
Arg	Arginine
Asn	Asparagine
Asp	Aspartic acid
CADD	Computer Aided Drug Design
cgDHFR	*C.glabrata* Dihydrofolate reductase
COMFA	Comparative Molecular Field Analysis
COMSIA	Comparative Molecular Similarity Indices Analysis
CV	Critical Volume
DHF	Dihydrofolate
DHFR	Dihydrofolate reductase
DHPS	Dihydropteroate synthase
DNA	Deoxyribose Nucleic Acid
dTMP	Deoxythymidylate
dUMP	Deoxyuridine monophosphate
FDA	Food and Drug Administration
GA	Genetic Algorithm

Gly	Glycine
HBD	Hydrogen Bond Donor
His	Histidine
HIV	Human Immunodeficiency virus
HOF	Heat of formation
HOMO	Highest Occupied Molecular Orbital
HBA	Hydrogen Bond Acceptor
HP	Hydrophobic
IC	Incremental Construction
IC_{50}	Half Maximal Inhibitory Concentration
Ileu	Isoleucine
IUPAC	International Union of Pure and Applied Chemistry
Leu	Leucine
LogP	Partition Coefficient
LUMO	Lowest Unoccupied Molecular Orbital
Lys	Lysine
MC	Monte Carlo
MOE	Molecular Orbital Environment
MR	Molecular refractivity
NADPH	Nicotinamide adenine phosphate
pcDHFR	*Pneumocystis carinii* Dihydrofolate reductase
PCP	Pneumocystis Carinii Pneumonia
PDB	Protein Data Bank

Phe	Phenylalanine
Pro	Proline
QSAR	Quantitative Structure Activity Relationship
QSPR	Quantitative Structure Property Relationship
RNA	Ribonucleic Acid
Ser	Serine
SHMT	Serine hydroxyl methyl transferase
SMX	Sulphamethoxazole
TE	Total Energy
THF	Tetrahydrofolate
Thr	Threonine
TMP	Trimethoprim
Trp	Tryptophan
TS	Thymidylate synthase
UNICEF	Unite Nations Children's Fund
Val	Valine
VMD	Visual Molecular Dynamics
WHO	World Health Organization

LIST OF FIGURES

LIST OF TABLES

ABSTRACT

Pneumonia continues to pose serious threats to world health mainly due to the emergence and spread of drug resistant strains. The advent of Computer aided drug designing and discovery leads to the betterment of mankind by the amalgamation of science and technology. More over it holds the promise to unlock the mysteries of disease mechanism and its cure. Ligand-based pharmacophore modeling is carried out on a set of 10 compounds and the standard drug Trimethoprim that were superimposed and merged into single pharmacophore showing three common features: one aromatic volume, two hydrogen bond acceptor and two hydrogen bond donors. The pharmacophore triangle is determined using the *In-silico* approaches. Molecular docking was brought in to use to determine the Lead compound as the *P.carinii* dihydrofolate reductase inhibitor. AutoDock Vina was used for docking studies of data set and the target protein used was PDB ID: 1DYR. However VMD was employed to identify the binding interactions of the active conformations of the ligands and the target protein. Lead compound revealed strong ligand-protein interaction which includes 7 ionic and 19 hydrogen bonds interactions. Four analogues of the lead compound were designed and they were also docked in order to predict their bioactivity. Quantitative structure-activity relationship was established in order to attain the information useful for the design of new compounds acting on a specific target. The activity parameter used in the QSAR analysis is the IC-50 value. The correlation reveals that Critical volume (CV), Molar refractivity (MR) and LogP proved to be good descriptors for the activity. This research study revealed novel anti *P.carinii* agents for clinical trials.

CHAPTER NO. 1

INTRODUCTION

1. Introduction

The fungal opportunistic infection, *Pneumocystis carinii* pneumonia (PCP) is a major area of research and thus requires consistent efforts for further clinical investigations. PCP remains a major cause of morbidity and mortality throughout the world among patients with compromised immune systems and as a result PCP remains a leading AIDS (Acquired Immune Deficiency Syndrome) defining opportunistic infection in Human Immunodeficiency Virus (HIV) infected individuals.

PCP has remained the point of interest for the research hub since its identification and the major reason behind its selection is that *Pneumocystis* has several unique features that have proven *Pneumocystis* study to be exciting but quite difficult in part due to the limitation that it cannot be reliably cultured outside the host lungs. Secondly the organism's source in nature has not been identified, as a result the question regarding its transmission has remained difficult to examine. Thirdly the organism is host specific however the mechanisms that are responsible for its specificity have not been elucidated extensively. Lastly this type of infection is essentially limited to lungs at all times but it has also been revealed that the characteristics of the host that allow progression of the infection are not understood completely. The life cycle of *Pneumocystis carinii* is also not understood completely. The development and application of various molecular techniques made over the past several years has helped in understanding the organism and the disease, in order to achieve significant advances (Huang *et al.*, 2006).

The selection of Dihydrofolate reductase (DHFR) as a drug target for the cure of PCP is based on the fact that DHFR is considered to be a key enzyme in terms of treatment and has been successfully used as a target against a variety of pathogenic

microorganisms to aid the process of antimicrobial drug discovery. DHFR has the potential to be used in the treatment of Pneumocystosis. The use of DHFR inhibitors in this study as a research issue is significant in the sense, that it is the one of the best known drug type for folic acid synthesis inhibition and yet this type of information is important in developing new forms of therapy to cure PCP (Da Cunha *et al.*, 2010). There are few Food and Drug Administration (FDA) approved drugs available for the cure of PCP as the first and second line treatment namely Trimethoprim, Sulfamethoxazole, Dapsone, Atovaquone and Pentimidine. Trimethoprim is a DHFR inhibitor, but if the worldwide trends are brought into consideration it is obvious that there exists a strong need for the identification of new potent compounds that possess important drug features including effectiveness, selectivity and efficiency in order to act as a potential drug.

Using the Computer aided drug design (CADD) or in-silico drug designing approaches the identification of a novel drug for the treatment of *Pneumocystis carinii* Pneumonia is promised to be developed in short time span. Research is conducted with the sole purpose to identify a lead compound that possesses the potential to act as candidate *Pneumocystis carinii* DHFR inhibitor. This can be achieved by the involvement of the in-silico techniques by reducing the time required to develop new drugs mainly focusing on drug bioavailability by increasing binding interactions. In the present study, the research was focused on deciphering the following areas:

- In this research study pharmacophore models are generated by extracting information from various anti PCP compounds included in the data set, as yet not much focus has been laid on pharmacophore model identification for *Pneumocystis carinii* DHFR inhibitors thus it will serve as a valuable

contribution.

- Moreover, a protocol is designed to aid the in-silico drug development by including approaches such as pharmacophore modeling, molecular docking and quantitative structure activity relationship (QSAR) for the identification of potential lead compound followed by the analogue design using the lead to discover the next potential drug candidates.
- 2D QSAR analysis was also performed by calculating the molecular descriptors. The correlation was determined which helped in finding a relationship between a biological activity and pharmacological descriptors.
- The effectiveness of the research work can be justified, since the most active compound the lead was identified that was involved in a greater number of binding interactions which enhanced the therapeutic ability. The research study conducted would prove to be an extremely valuable and helpful source for the cure of *Pneumocystsis carinii* Pneumonia by increasing the binding interactions and the bioavailability of the lead compound.

CHAPTER NO. 2

LITERATURE REVIEW

2. Literature Review

Disease, a harmful deviation from normal structural or functional state can have highly devastating effects if it remains untreated. The advancements in science and technology have a great impact on the society when it comes to medicine and issues related to the health of humanity. Devising a cure for a particular disease is not an easy task since it requires consistent efforts for the amalgamation of modern science and technology. However it cannot be denied that Pneumonia continues to represent a major threat to world health and it is one of the major causes of childhood mortality, as a consequence of which the World Health Organization (WHO) and the United Nations Children's Fund (UNICEF) have identified it as the major "forgotten killer of children". According to the statistics provided by WHO, almost 99.9% of childhood deaths due to pneumonia occur in developing and least developed countries, with a maximum death toll of 1022 000 cases per annum in sub-Saharan Africa followed by a death toll of 702000 cases per annum in South Asia. However considering all the deaths caused by pneumonia about 47.7% occur in the least developed countries (Madhi *et al.*, 2008). The mortality rate however decreases with age until adulthood but the elderly individuals are at greater risk of pneumonia. The WHO estimates show that one in three newborn infant deaths are due to the onset of pneumonia (Garenne *et al.*, 1992).

The HIV epidemic has also contributed largely to the increase in incidence and childhood mortality from pneumonia in the recent years. The bacterial infection remains a main cause of childhood mortality in children with the incidence of pneumonia and HIV both, but in the case of HIV infected children other pathogens also exist including the pathogen *Pneumocystis jiroveci* (Rudan *et al.*, 2008). Figure

2.1 demonstrates the Global distribution of Acute Respiratory Infections including pneumonia and influenza both.

According to the Annual Report of Infectious Disease Epidemiology Sections (2004) from the Louisiana Office of Public Health, Pneumonia is particularly caused by a wide number of infections including the viral, bacterial and fungal infections. The viral agents that are capable of causing pneumonia commonly are the respiratory syncytial virus, adenoviruses, influenza viruses and the parainfluenza viruses. However the bacterial agents including *Haemophilus influenza* type B, *Streptococcus pneumoniae*, *Streptococcus pyogenes, Staphylococcus aureus, Klebsiella pneumonia, Neisseria meningitides* can cause pneumonia. Moreover other organisms, including *Escherichia coli* are less commonly considered as causes of pneumonia. However the pathogen *P.carinii* is a major cause of pneumonia in immunosuppressed hosts.

2.1 *Pneumocystis Carinii* Pneumonia

The fungal opportunistic infection, *Pneumocystis Carinii* Pneumonia (PCP) is caused by the pathogen *Pneumocystis jirovecii* that is formerly known *Pneumocystis carinii* (Redhead *et al.,* 2006; Hawksworth *et al.,* 2007). The opportunistic pathogen, *Pneumocytis carinii* causes an interstitial pneumonia. A lethal type of pneumonia is caused by the organism *P.carinii* in immunosuppressed hosts especially in AIDS patients (Cushion, 1994). PCP also targets other patients such as those undergoing organ and bone marrow transplants, cancer therapies and possessing a weak and compromised immune system (Fishman, 1998). Originally, this disease was associated to premature and marasmic infants, but now it is being reported in immunosuppressed and immunodeficient patients with an increasing frequency (Goodell *et al.,* 1970; Hughes *et al.,* 1973). The primary mode of transmission is the

respiratory route that has been documented in the animal studies however it is unknown in case of humans thus the respiratory route is likely to be important (Cushion, 1994). Pneumocystis pneumonia patients have fewer pneumocystis organisms in their lungs but possess a greater number of neutrophils as compared to patients with AIDs who have pneumocystis pneumonia (Limper *et al.*, 1989). In case of patients with the onset of AIDs, pneumocystis pneumonia occurs when the T-helper cell count (CD4+) becomes less than 200 cells per cubic millimeter (Phair *et al.*, 1990; Masur *et al.*, 2002).

2.2 Biology and epidemiology of *P.carinii*

2.2.1 Historical perspective

In terms of a historical perspective *P.carinii* was first identified by Chagas (Chagas, 1909) and Carinii (Carini, 1910) in trypanosome-infected lungs of animals in the years 1909 and 1910, who thought that it was a form of trypanosome. However three years later it was reported that the organism was quiet distinct from trypanosome hence the genus Pneumocystis was chosen by Delanoe and Delanoe which was descriptive enough of the form of the microbe possessing small but highly refractive and densely staining spherical cyst form (Delanoe and Delanoe, 1912). However it was named in honor of Carinii after being recognized as a new genus by Delanoe and Delanoe in the year in 1912. The organism *P.carinii*, for nearly 50 years has been recognized as a human pathogen (Kovacs *et al.*, 1984). During the 1970s, *P.carinii* was thought to be a protozoan but its molecular studies have helped in its reclassification as a fungus (Edman *et al.*, 1988).

2.2.2 Incidence of PCP

Pneumocystis pneumonia is a major cause of morbidity and mortality among patients with suppressed and weak immune systems (Huang *et al.*, 2006). However the individuals with impaired immune systems are at a greater risk of Pneumocystis pneumonia. The incidence of pneumocystis infection has increased dramatically in the developing countries with mortality rates in the range of 20% to 80% (Fisk *et al.*, 2003). As far as the mortality rate among the non-HIV patients with the onset of pneumocystis pneumonia is considered, it remains 30% to 60% depending on the population at risk, since the patients with cancer are at a greater risk of death as compared to the patients undergoing transplantation or those suffering from connective-tissue disease (Pareja *et al.*, 1998; Sepkowitz, 2002).

2.2.3 Host specificity

Molecular as well as immunologic studies have revealed that there are multiple *P.carinii* strains that are unique to single host specie in terms of causing infection (Kovacs *et al.*, 1989; Stringer, 1996). Amongst many different types of *Pneumocystis* organisms some show sufficient differences to be declared as separate species. The *P. carinii* isolates obtained from different host species indicate that they differ both phenotypically and genetically. The variation of phenotypic characteristics among the *P.carinii* organisms includes factors such as (i) host species specificity, (ii) electrophoretic mobility of enzymes, (iii) antigens, and (iv) morphology (Stringer, 1996).

2.2.4 Genome of *P.carinii*

The genome of *P.carinii* that infects rats has been identified to be about 7.7 Mb in size, and consists of 13 to 15 linear chromosomes that range in size from 300 to 700

kb. Amongst the genes identified to date most genes are found to have numerous short introns (Thomas *et al.*, 1999). Pneumocystis Genome Project (PGP) has resulted in the identification of new targets however the sequence of mitochondrial genome has shown implications for its use as a potential target for future therapies. The therapies including agents targeted to gene products within the *Pneumocystis* mitochondrion have already been shown to be effective including the use of drug atovaquone (Sesterhenn *et al.*, 2006). In order to explore and identify *Pneumocystis* gene function an active *Pneumocystis* genome project exists which will highlight the homologies between *Pneumocystis* to infer gene function.

2.2.5 Proposed Cell cycle of Pneumocystis

In the 1970s *P.carinii* could not be cultured as a result of which the life cycle of the organism was unknown and the metabolism of the organism was also poorly understood. The life cycle remains unknown currently, however some recent reports show a greater degree of success which is encouraging in this particular area (Merali *et al.*, 1999). The advances in having a better understanding of the metabolism have been achieved by the application of molecular biology approaches (Kovacs *et al.*, 2001). Currently it is not possible to propagate *pneumocystis* in culture media outside the infected host, because of this limitation heterologous expression in related fungi is brought into use to perform the functional analysis regarding the pneumocystis cell cycle and of signal-transduction molecules. The signal-transduction cascades for the regulation of cellular responses in fungi including the responses for mating, filamentous growth, environmental stress and cell-wall integrity make use of mitogen-activated protein kinase. The mitogen-activated protein kinases identified in *pneumocystis* are found to be homologous to the ones that are found in mating and

cell wall integrity pathways (Fox and Smulian, 1999; Vohra *et al.*, 2003). Because of the difficulties with in vitro culturing the life cycle of *Pneumocystis* is found to consist of a trophic form, a precystic form, and a cystic form that have been identified with the aid of morphologic criteria (Huang *et al.*, 2006).

In fact the life cycle of *pneumocystis* is very complex as several forms are observed after the progression of infection. Figure 2.2 shows the proposed life cycle of *P.carinii*. During the infection stage the trophic forms are more abundant as compared to cysts usually in a ratio 9:1. During the normal growth majority of the trophic forms are thought to be haploid along with a small fraction that is diploid. The electron micrograph in the panel A shows a trophic form. This form is tightly adherent to the alveolar epithelium by apposition of its cell membrane with the cell membrane of the host. The electron micrograph panel B shows that trophic forms attach to one another as a result of which clusters of clumped trophic forms are witnessed during infection. The events that lead to the formation of the cyst that is shown in electron micrograph panel C are unclear but it is hypothesized that the trophic forms conjugate and mature into cysts which on maturation contains two, four or eight nuclei (Thomas and Limper, 2004). The infection is established when the pneumocystis trophic forms adhere to the alveolar epithelium tightly. Specific signaling pathways in the organisms are activated when the pneumocystis binds to lung epithelial cells including gene encoding PCSTE20 kinase which involves in signaling responses in fungal organisms for the mating and proliferation (Kottom *et al.*, 2003). Other signaling molecules that have been identified so far include the receptors such as the putative pheromone, along with the transcription factors and the heterotrimeric G-protein subunit (Smulian *et al.*, 1996; Vohra *et al.*, 2003).

2.3 Clinical features of Pneumocystis Pneumonia

Clinical investigations reveal that a clinically apparent pneumonia is witnessed in patients with suppressed immune systems. PCPs clinical finding suggest that it is characterized by symptoms such as fever, shortness of breath, other symptoms include substernal tightness and a nonproductive cough. However the symptoms can be mild and progressing slowly specially in the case of HIV-infected patients, which is a major cause of delay in diagnosis (Kovacs *et al.*, 1984). Pneumocystis pneumonia patients without the incidence of AIDs typically show an onset of respiratory insufficiency that is abrupt which may correlate to a narrow or increased dosage of immunosuppressant medications (Sepkowitz *et al.*, 1992; Yale and Limper, 1996). PCP appears with the onset of common symptoms including suffering from low-grade fever and the onset of the subtle progressive dyspnea. Pneumothorax can develop that can be indicated by acute dyspnea with pleuritic chest pain (Thomas and Limper, 2004).

2.4 Drug targets for PCP

The possible target enzymes due to their therapeutic importance have been cloned, sequenced and characterized. These include the genes encoding the drug targets of sulfamethoxazole and trimethoprim, the dihydropteroate synthase (DHPS) and dihydrofolate reductase (DHFR) (Edman *et al.*, 1989; Volpe *et al.*, 1993). More recent research studies have identified that *Pneumocystis* contains a set of several specific enzymes including the DHFR (Allegra *et al.*, 1987; Kovacs *et al.*, 1990), thymidylate synthase (Edman *et al.*, 1989), DHPS (Kovacs *et al.*, 1989), ornithine decarboxylase (Lipschick *et al.*, 1991), topoisomerase (Dykstra *et al.*, 1991) and β-1, 3-glucan synthetase (Schmatz *et al.*, 1990) can be used as the possible drug targets. Due to the

immense importance of DHFR as a drug target in the treatment of PCP, pcDHFR was selected as the target protein in this particular study.

2.5 Role of DHFR

Since the 1960s DHFR had been recognized and selected as a valid drug target with the successful discovery of methotrexate (Bertino, 1993). DHFR is a ubiquitous enzyme that catalyzes the reduction process of dihydrofolate to tetrahydrofolate in the presence of nicotine adenine dinucleotide phosphate (NADPH) as a co-factor. This process yields reduced folates that are capable of functioning as carbon carriers in the biosynthesis of nucleotides and amino acids which are essential for cell proliferation. Thus if inhibition of DHFR is carried out, it leads to limitations in cell growth and division. This reason accounts for the use of DHFR inhibitors as cytostatic agents in cancer therapy.

The importance of DHFR cannot be neglected since it is a key enzyme in the treatment of Pneumocystosis. Its role is to reduce 7, 8-dihydrofolate (DHF) to 5, 6, 7, 8-tetrahydrofolate (THF) with the aid of the cofactor nicotinamide adenine phosphate (NADPH) in the thymidine biosynthesis as represented in Figure 2.3. After the process of reduction, regeneration of 5, 10-methylene-tetrahydrofolate is catalysed by serine hydroxyl methyl transferase (SHMT) and methylation of deoxyuridine monophosphate (dUMP) to give deoxythymidylate (dTMP) in a reaction catalysed by thymidylate synthase (TS) occurs. This reaction causes the conversion of methylene-tetrahydrofolate back to dihydrofolate, thus completing the cycle (Da Cunha *et al.*, 2010). Therefore, the entire process results in the inhibition of DHFR preventing the biosynthesis of thymidine and, as a result the inhibition of DNA biosynthesis (Da Cunha *et al.*, 2005).

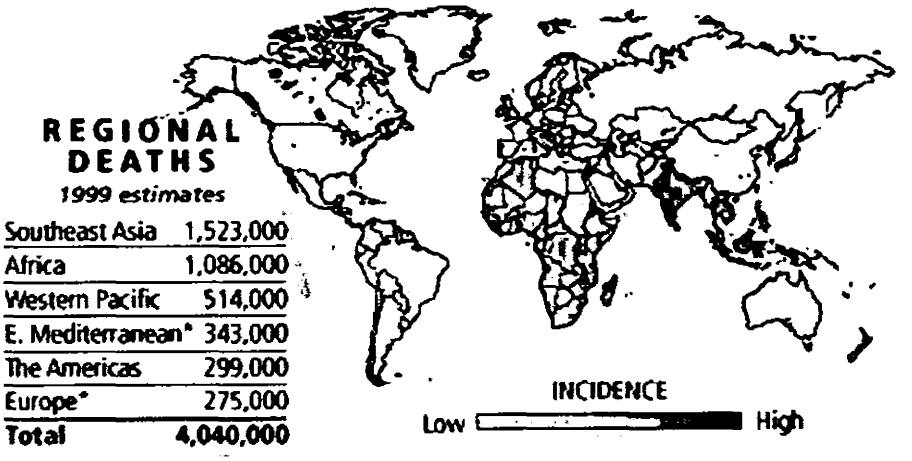


Figure 2.1: Global distribution of Acute Respiratory Infections (Emily *et al.*, 2008)

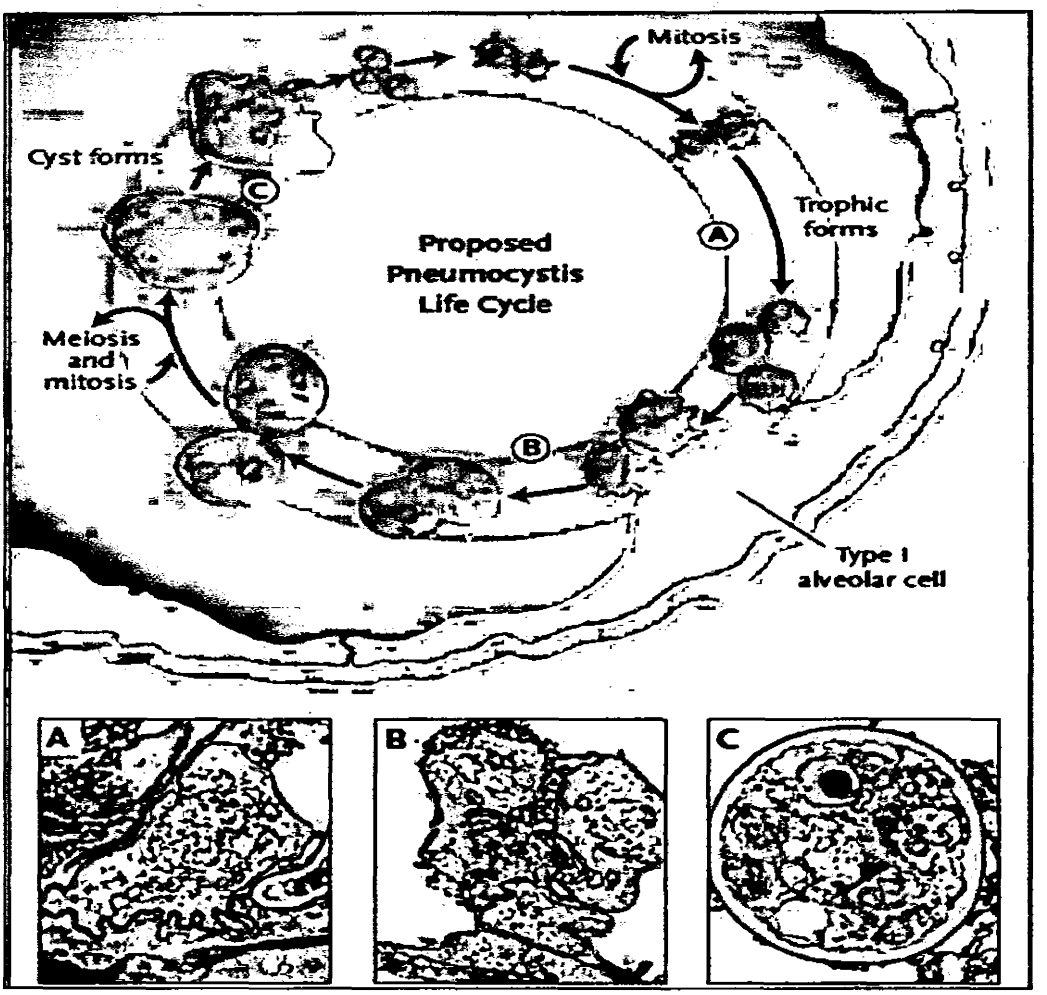


Figure 2.2: Proposed life of Pneumocystis Carinii (Thomas *et al.*, 2004)

2.6 Structure of pcDHFR

Comparing the structures of the fungal enzymes from *C. glabrata* and *P. carinii* reveal that they are very similar to each other and a superimposition with a root mean square deviation (RMSD) of 3.6 Å over the 161 C-α atoms is shown. The residues composing the active site exhibit a greater degree of structural similarity thus highlighting the similarity between the two enzymes. The cgDHFR and pcDHFR differ by only one residue, located at Met33 in the cgDHFR and Ile33 in case of pcDHFR. On the basis of the active site similarity and appeared fit of the lead to the pcDHFR after the docking, the results show that the propargyl-linked antifolates are more likely to act as the effective inhibitors. The key loop at the active site maintains the same position and conformation in both enzymes and the residues Pro 66 to Phe 69 are observed (Liu *et al.*, 2008). Figure 2.4 demonstrates the structures of both enzymes the cgDHFR presented in black and pcDHFR shown in gray bound to antifolates and NADPH. The ligands are shown in stick form.

In this study the structure of pcDHFR with PDB ID: 1DYR with a resolution of 1.9 Å was selected (Champness *et al.*, 1994). John and coworkers solved the structure of pcDHFR composed of 206 amino acids with reduced NADPH and trimethoprim and the pcDHFR ribbon diagram is shown in Figure 2.5. The structures are helpful to show the interactions of two drugs with the fungal DFHR (Champness *et al.*, 1994). However other pcDHFR structures have also been solved to date.

2.7 Importance of DHFR

The human and pathogen forms of DHFR reveal several critical differences in terms of their sequence, fortunately as a consequence of this it has been possible to design antifolates that are species selective for the treatment of infectious diseases including

the urinary tract infections, toxoplasmosis and malaria (Liu *et al.,* 2008). DHFR has been more prominently selected as a pharmacologic target and the protein chemists are attracted towards the study of DHFR enzyme structure/function relationships mainly because of two reasons, firstly due to its small size (18-22 kDa) and secondly due to the availability of purified enzyme. For molecular biologists it has immense importance as a model for gene amplification studies and in terms of gene transfer studies it can be used as a selectable marker (Schweitzer *et al.,* 1990). DHFR inhibitors can be used in several other disease conditions including rheumatoid arthritis (Bannwarth *et al.,* 1994), inflammatory bowel diseases (Peppercorn, 1990), psoriasis (Weinstein and Frost, 1971) and asthma (Mullarkey *et al.,* 1988). In addition to its application in the above mentioned diseases DHFR has the potential to be used as an effective target protein to combat infectious diseases (Queener, 1995).

2.8 DHFR Inhibitors

Some anti-pneumocystis drugs have a well-established mechanism of action. However in vitro DHFR inhibition can be achieved by Trimethoprim, pyrimethamine, methotrextate along with other investigational agents including piritrexim and trimetrexate (Allegra *et al.,* 1987). An important class of therapeutic compounds is the folate antagonists, which is evident from their usage as antineoplastic, anti-infective and anti-inflammatory drugs. Thus it is evident that all drugs that are clinically useful belonging to this class have been inhibitors of DHFR. The folic acid inhibitors have a major role in PCPs treatment. Major focus has been placed on the use of inhibitors of pcDHFR primarily trimethoprim that is being used in combination with a sulfonamide mainly sulfamethoxazole (Hughes *et al.,* 1974; Kluge *et al.,* 1978; Winston *et al.,* 1980; Siegel *et al.,* 1984; D'Antonio *et al.,* 1986).

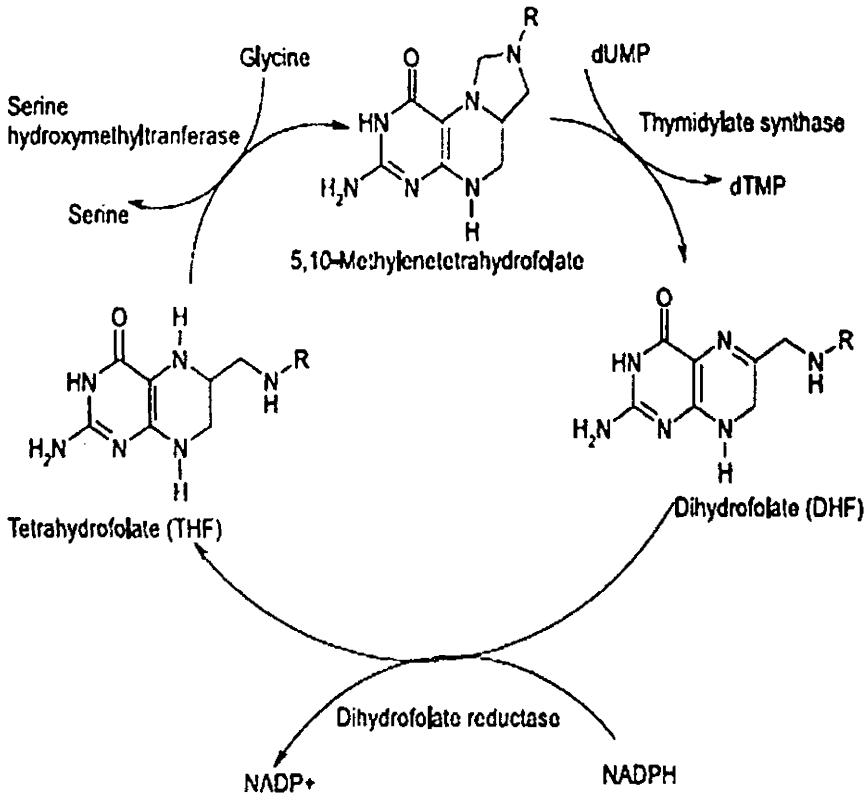


Figure 2.3: Role of DHFR in the synthesis of thymine (Da Cunha *et al.*, 2010)

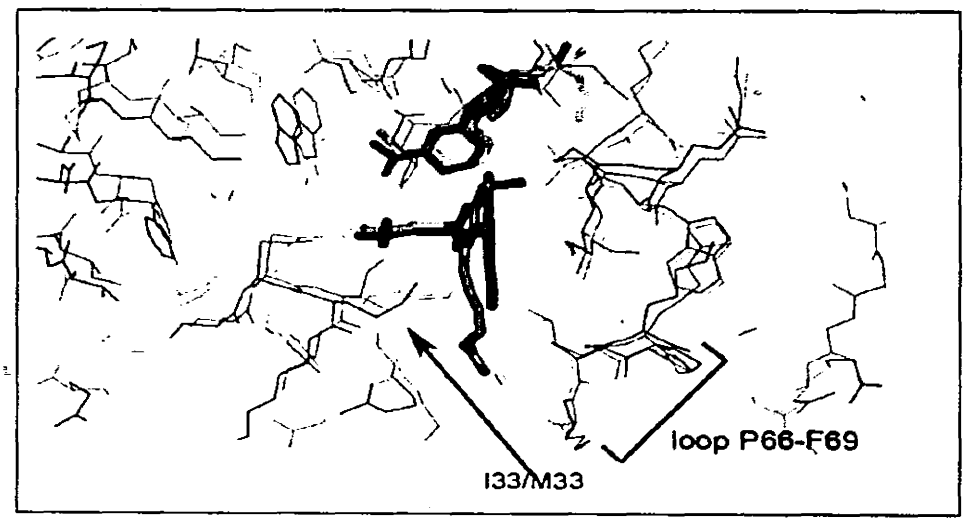


Figure 2.4: The structures of cgDHFR shown in black and pcDHFR represented in gray bound to antifolates and NADPH shown in stick form (Liu *et al.*, 2008)

Other DHFR inhibitors have also been employed into use for the treatment of clinical and experimental pneumocystosis which include inhibitors such as pyrimethamine, tetroxoprim, pitrexim, and trimetrexate, sulfonamides and sulfones including the sulfadiazine, sulfadoxine, sulfamonomethoxine, dapsone, and sulfonyl bisformanilide (Kirby *et al.,* 1971; Post *et al.,* 1971; Whisnant *et al.,* 1976; Young *et al.,* 1976; Yoshida *et al.,* 1977; Hughes and Smith, 1984 ; Hussain *et al.,* 1985; Yamada *et al.,* 1985; Leoung *et al.,* 1986; Allegra *et al.,* 1987; Devita *et al.,* 1987; Pearson *et al.,* 1987; Queener *et al.,* 1987). Not much is known about the about the toxicity of the different folic acid inhibitors and their relative efficacy in the treatment of *P.carinii,* yet such type of information is of extreme importance in the process of developing new forms of therapy (Walzer *et al.,* 1987).

Compounds which effectively exhibit anti-*P.carinii* activity in various models specifically the rat models include antifolate drugs that can be used alone effectively or in combination with various other drugs. This group of antifolate drugs includes DHFR inhibitors, sulfonamides, sulfones, and sulfonylureas (Frenkel *et al.,* 1966; Hughes *et al.,* 1974; Yoshida *et al.,* 1977; Kluge *et al.,* 1978; Ulrich *et al.,* 1984; Hussain *et al.,* 1985; Yamada *et al.,* 1985; D'Antonio *et al.,* 1986; Hughes, 1987; Queener *et al.,* 1987; Kovacs et al, 1988; Hughes *et al.,* 1990). Sulfonamides are often used in combination with other drugs and can also be used in combination with macrolides, 8-aminoquinolones can be used effectively alone or in combination with other agents (Queener *et al.,* 1988; Bartlett *et al.,* 1991), diamidines and related cationic compounds (Debs *et al.,* 1987; Tidwell *et al.,* 1990) other compounds with anti-*P. carinii* activity include nitrofurans (Walzer *et al.,* 1991), 3-glucan inhibitors (Schmatz *et al.,* 1990; Matsumoto *et al.,* 1991; Schmatz *et al.,* 1991), hydroxyl

naphthoquinones (Hughes *et al.*, 1990), fluoroquinolones (Brun-Pascaud *et al.*, 1992.), iron chelators (Clarkson *et al.*, 1990) and immunological agents like cytokines and antibodies (Gigliotti *et al.*, 1988 ; Shear *et al.*,1990).

The task of evaluating anti-*P.carinii* drugs is a very expensive, time-consuming and labor demanding process. In a given research experiment only a few compounds can be tested. Various types of treatment studies have been performed by incorporating several different experimental protocols also making use of various methods of evaluating drug efficacy. Thus there is a lack of standardization. This deficiency has prevented the results of one investigator from being compared directly with the outcomes of another investigator. The set of compounds that have not shown activity against *P.carinii* have few published sources of information (Walzer *et al.*, 1992).

2.9 Prophylaxis and Treatment of PCP

In terms of regime of prophylaxis, the most effective treatment of PCP is the use of the drug combination trimethoprim–sulfamethoxazole (TMP–SMX; Bactrime™, Septra®). However it is unfortunate to witness that almost half of the HIV-positive patients develop an allergic reaction on the intake of sulpa drugs thus necessitating a move towards other less effective drugs (Sattler *et al.*, 1988) including pentamidine (NebuPent®, Pentacarinat®), dapsone (Avlosulfon®), and atovaquone (Mepron®). Figure 2.6 shows the regime of prophylaxis for the treatment of PCP. However important factors associated with these drugs such as the deleterious side effects, limited efficacies of these drugs, and emerging mutations and drug resistance cannot be neglected and thus provide a justification for the search of more effective and less toxic medicinal agents (Eynde *et al.*, 2004).

Figure 2.5: The structure of pcDHFR to 1.9Å resolution (Champness *et al.*, 1994).

Drugs for Prophylaxis against Pneumocystis Pneumonia.

Drug	Dose	Route	Comments
Trimethoprim–sulfamethoxazole	1 double-strength tablet daily or 1 single-strength tablet daily	Oral	First choice
	1 double-strength tablet 3 times per week		Alternate choice
Dapsone	50 mg twice daily or 100 mg daily	Oral	Ensure patient does not have glucose-6-phosphate dehydrogenase deficiency
Dapsone plus pyrimethamine plus leucovorin	50 mg daily 50 mg weekly 25 mg weekly	Oral	
Dapsone plus pyrimethamine plus leucovorin	200 mg weekly 75 mg weekly 25 mg weekly	Oral	
Pentamidine	300 mg monthly	Aerosol	
Atovaquone	1500 mg daily	Oral	Give with high-fat meals, for maximal absorption

Figure 2.6: Regimes of Chemoprophylaxis of PCP (Thomas *et al.*, 2004)

2.10 Computer Aided Drug Design

The task of introducing therapeutically beneficial and relatively new solutions requires expenses and is in fact a time consuming process (Song *et al.*, 2009). Rapid expansion in this area has been achieved by advances in features such as computational power, improvements in software and hardware sophistication, identification of important molecular targets, and an increased availability of target protein structure databases (Kapetanovic, 2008). The typical drug discovery cycle initiating from the lead identification towards the clinical trials takes an estimated time span of about fourteen years (Myers and Baker, 2001) with a cost up to 800 million US dollars (DiMasi *et al.*, 2003). The production of compounds in vast quantities and the requirement to examine these corresponding huge libraries within short time intervals are the characteristics of modern drug discovery. The essential need for the storage, management and analysis of these rapidly accumulating resources has given a major rise to the field of computer-aided drug design (CADD). CADD is representation and embodiment of the computational methods and resources that are utilized in the therapeutic design and discovery of new solutions in order to facilitate the process (Song *et al.*, 2009). In this biomedical arena, CADD or *in silico* design is being used to accelerate and facilitate drug design phenomena including the hit identification, selection in terms of hit to lead, optimizing the absorption, distribution, excretion, metabolism and toxicity profile and avoid safety issues. CADDD entails:

1. Utilizing computational power to streamline the drug discovery and development process
2. Effective use of chemical and biological properties of ligands and or targets

for the identification and optimization of new drugs (Kapetanovic, 2008).

CADD is a stimulating and highly diverse discipline where various aspects of applied and basic research amalgamate and stimulate each other. With latest technological advances by utilization of structure-based design, QSAR, combinatorial library design, cheminformatics and bioinformatics, the ever increasing number of chemical and biological databases and an exponential growth in currently available software tools are providing a much upgraded and stronger basis for the ligands and inhibitors design with desired and required specificity (Bajorath *et al.*, 1999).

Computer aided drug design development process is being used with a major role to identify active drug candidates as hits, select leads that are most likely the prominent candidates for further evaluation and carry out the optimization of leads by the transformation of biologically active compounds into therapeutically appropriate drugs by posing considerable improvement in their physicochemical, pharmacokinetic and pharmaceutical properties (Kapetanovic, 2008). The major role of computational models is to make use of existing knowledge to increase prediction. Computational methods are actively participating in drug discovery and development process by playing an increasingly greater and extremely important role in drug (Kumar *et al.*, 2006; Stahl *et al.*, 2006) and they are thought to offer ways of improved efficiency for the industry (Kumar *et al.*, 2006).

Virtual screening is brought in to use for the discovery of novel drug candidates from different chemical scaffolds by involvement of the database searching including the commercial, public or private 3-D chemical structure databases with a major goal to enhance and supplement various sets of molecules with desirable properties including activity, drug-like, lead-like properties and side by side elimination of compounds

possessing undesirable properties like inactiveness, reactive nature, toxic effects, poor ADMET/PK. In alternative form it can be stated that *in silico* modeling is brought into use to minimize time and resource constraints significantly that are usually employed in the process of chemical synthesis and biological testing of drugs.

The most frequently employed computational approaches include ligand-based drug design by involvement of a 3-D spatial arrangement of chemical features, the pharmacophore that is essential for biological activity, however the structure-based drug design approach involves drug-target docking phenomena and use of other computational methodologies such as quantitative structure activity relationships (QSAR) and quantitative structure property relationships (QSPR) (Kapetanovic, 2008).

2.11 Pharmacophore Modeling

Three-dimensional (3D) pharmacophore modeling is a technique for elucidation and description of the interactions of small molecule ligands with a macromolecular protein target (Wolber *et al.*, 2008). Considering the early pharmacophore modeling techniques including the Active Analog Approach scheme proposed by Garland Marshall and coworkers (Marshall *et al.*, 1979) demonstrates that the features constituting a pharmacophore could contain any fragment or atom type. Since 1998 a more precise official definition in accordance with the IUPAC (Wermuth, 1998) defines pharmacophore as: "an ensemble of steric and electronic features that is necessary to ensure the optimal supramolecular interactions with a specific biological target structure and to trigger or block its biological response" (Wermuth, 2006). A pharmacophore model can be interpreted as a spatial arrangement of atoms or functional groups that are considered to be responsible for biological activity (Ghose

and Wendoloski, 1998). In order to assist the process of drug design, a pharmacophore model (Wermuth and Langer, 1993) should provide effective and valid information for the investigation of structure–activity relationships by the medicinal chemists. The pharmacophore model that have emerged as the most effective type of models are the feature-based pharmacophores and these models have been reviewed recently for their utility as queries for searching 3D database (Kurogi and Guner, 2001; Langer and Krovat, 2003).

Pharmacophore modeling is conceptually simple and in many aspects it is capable of producing results that would be intuitive and helpful for an experienced medicinal chemist. This technique used is in a particular binding situation to rigidly formulate and model the interactions between a ligand and its binding site. Pharmacophore can be generated either in a structure based fashion by determination of complementarities between a ligand and its corresponding binding site, or a ligand-based approach can be used by flexible overlay of a particular set of active molecules and determining those conformations that possess a maximum number of significant chemical features that geometrically overlap when they are overlaid (Wolber *et al.*, 2008).

The most crucial and significant elements in terms of a pharmacophore might range from a group of atoms to a part of the molecules volume, however other very 'classical' pharmacophoric features include Hydrogen bond acceptors and donors, hydrophobic and/or aromatic rings, charged or ionizable groups along with geometrical constraints such as distances, angles and dihedral angles. The set of these properties is referred to as pharmacophoric 'model' or 'hypothesis'. During the process of pharmacophore generation if ligand information is solely available, principally the pharmacophore identification involves two steps, firstly the analysis of

molecules itself that are included in the training set in order to identify the pharmacophoric features and secondly by determination of the best overlay of the corresponding features by performing alignment of the molecules assumed bio-active conformations (Wolber and Langer, 2005).

For construction of pharmacophore models using ligand information various software programs are available these include Catalyst (Barnum *et al.*, 1996) that is available from Accelrys Inc, it is by far the most frequently used program because large flexibility with an integration of high-speed 3D database searching capability is offered during this process. Other successful programs also exist, including the DiscoTech (Martin, 2000), and Gasp (Jones and Willet, 2000) both from Tripos Inc (Wolber and Langer, 2005). Another software package for pharmacophore modeling is GALAHAD (Richmond *et al.*, 2004) that was developed at the University of Sheffield and also marketed by the company Tripos effectively makes use of a modified genetic algorithm. More recent and currently utilized pharmacophore modeling software packages like Catalyst (Barnum *et al.*, 1996), Phase (Dixon *et al.*, 2006), MOEs pharmacophore module, and LigandScout (Wolber and Kosara, 2006) always have to face a trade-off in the design of a generally applicable universal feature set and at the same time it needs to be selective enough to reflect all types of ligand–receptor interactions that are relevant.

Using the software **LigandScout** (Wolber and Langer, 2005) available from Inte:Ligand GmbH is one of the possibilities for the automatic generation of a feature-based pharmacophore model based on a ligand target complex structure. The program initiates the process of pharmacophore generation by the assignment of ligand information on the basis of features such as hybridization status and bond

characteristics that are absent in the input data files retrieved from the Protein Databank (Berman *et al.*, 2000) by use of two approaches together the extended heuristic method along with template based numeric analysis. Then the feature based pharmacophores are then generated by determination of interactions on the basis of charge, the hydrophobic contact and hydrogen bond formation between ligand and target atoms, and a corresponding model refinement can be then be achieved with reference to the binding data or one common feature pharmacophore can be obtained by combination of several models (Wolber and Langer, 2005). The concept of pharmacophore modeling is a fruitful and highly recognized approach for both ligand and structure based drug design as well as for virtual screening process. For the pharmacophore elucidation and virtual screening different alignment approaches which constitute the computational methodology has significantly improved with the ease in availability of newer software packages. Over the past few years performance optimizations have been achieved by the application of new algorithms thus leading to modern pattern recognition approaches that are capable enough to perform the superpositioning of pharmacophores and molecules in a smaller portion of the time as compared to the time required by earlier approaches (Wolber *et al.*, 2008).

2.12 Molecular Docking

Molecular docking is defined as a computer simulation process that is used for the prediction of the conformation of a receptor-ligand complex where ligand is a small molecule or a protein and the receptor is often a protein or a nucleic acid that can be a DNA or RNA molecule (Dias *et al.*, 2008). In order to execute the drug design process by discovering the new lead compounds, the virtual compound screening through the application of molecular docking is widely used. Molecular docking is

helpful in predicting the conformation of a protein-ligand complex and calculating the binding affinity (Okimoto *et al.*, 2009). Most of the docking programs (Morris *et al.*, 1996; Rarey *et al.*, 1996) involve the use of two operations firstly the **docking** and secondly **scoring** procedure. In the first operation, docking process involves the poses generated by multiple protein-ligand conformations or deals with the sampling of the ligand's most likely conformations in the binding pocket of the macromolecular target protein. However in the second operation a scoring function is brought in to use to calculate the affinity for each pose between the target protein and the ligand (Okimoto *et al.*, 2009). A broader classification of the scoring functions reveals that they are classified into three distinct categories the empirical, knowledge-based and force field-based (Mohan *et al.*, 2005).

Assuming that the structure of receptor is available the major challenge is the prediction of both the orientation and binding affinity of ligand during the lead discovery and optimization process; the former is often referred to as molecular docking approach (Lybrand, 1995). The number of algorithms available for the task of assessing and rationalizing ligand protein interactions, is considerably large and constantly increasing (Taylor *et al.*, 2002). The possibility to dock thousands of ligands in a timeline has increased owing to the increase in the power of computers and algorithm performance, which is valuable and of immense use to the pharmaceutical industry (Abagyan and Totrov, 2001). The docking process employs the use of the following principal techniques that are currently available these include approaches such as the Monte Carlo methods, molecular dynamics, fragment-based methods, genetic algorithms, distance geometry methods, point complementarity methods, tabu searches and systematic searches (Taylor *et al.*, 2002).

Kuntz et al developed the pioneer program DOCK about two decades ago in the year 1982 (Kuntz *et al.,* 1982) and during the span of these two decades development of numerous docking programs took place (Morris *et al.,* 1998; Li *et al.,* 2004). Molecular docking is rather a typical type of optimization problem since it is hard to achieve a solution that is globally optimal (Boehm and Stahl, 2002; Ferrara *et al.,* 2004). It is unfortunate to reveal that no scoring function has been developed so far that is effective enough to predict a ligand-protein binding mode and the binding affinity at the same time reliably and consistently (Bissantz *et al.,* 2000; Boehm and Stahl, 2002; Ferrara *et al.,* 2004).

Several docking programs are being incorporated in to docking studies of biological molecules these include softwares such as DOCK (Ewing *et al.,* 2001), AUTODOCK (Goodsell *et al.,* 1996; Morris *et al.,* 1998), GOLD (Verdonk *et al.,* 2003; Joy *et al.,* 2006), FLEXX (Rarey *et al.,*1996; Kramer *et al.,* 1999), ZDOCK (Chen *et al.,* 2003), M-ZDOCK (Pierce *et al.,* 2005), MSDOCK (Sauton *et al.,* 2008), Surflex (Jain, 2003), MCDOCK (Liu and Wang, 1999) and still others are available. Each of the docking application is based on a particular search algorithm such as Monte Carlo (MC) (Liu and Wang, 1999), Incremental Construction (IC) (Rarey *et al.,* 1996; Kramer *et al.,* 1999), Genetic Algorithm (GA) (Jones *et al.,* 1997; Morris *et al.,* 1998), etc. The extensively used docking methods make use of search algorithms based on Monte Carlo, genetic algorithm, fragment-based and molecular dynamics (Mohan *et al.,* 2005). The Monte Carlo methods are one of the most established and extensively used stochastic optimization techniques (Clark *et al.,* 2002). Evolutionary programing algorithms are used with the purpose of discovering or approximating solution in case of molecular docking and in search associated problems with an effort

to identify the exact or closest conformation of the global energy minimum and genetic algorithms belong to this respective class. The GOLD algorithm has adopted the GA and it has a requirement of the approximate size, the coordinates of protein and a ligand conformation along with the location of the receptors active site as an input (Jones *et al.*, 1997; Verdonk *et al.*, 2003). Another example of GA's implementation is the program DOCK which is capable of docking the entire ligand inside active site or docking a ligands rigid fragment (Ewing *et al.*, 2001). AUTODOCK is amongst the most recognized molecular docking programs that make use of "Lamarckian" GA (Goodsell *et al.*, 1996; Morris *et al.*, 1998). Many improvements have been introduced and added to this algorithm since its development in 1990s and applied to the virtual screening process for speed and accuracy optimization (Morris *et al.*, 1996; Vaque *et al.*, 2006).

AUTODOCK is a software package that can be successfully applied for automated docking of small molecules including peptides, enzyme inhibitors and drugs to macromoleculer proteins, enzymes, antibodies, DNA, and RNA. In this software three search methods are involved and tested the simulated annealing, genetic algorithm, and Lamarckian genetic algorithm but the research study concludes that the most efficient, reliable, and successful is the Lamarckian genetic algorithm (Morris *et al.*, 1998). AutoDock4 makes use of a semi empirical free energy force field in order to predict the binding free energies of small molecules the ligands to macromolecular target proteins and it allows the fully flexible modeling of specific portions of the protein in a similar way as the ligand. The major and most prominent limitation of AutoDock4 is its dependence on a grid-based energy evaluation (Morris *et al.*, 2009).

AutoDockVina is an upgraded version of Auto Dock 4 and both work in a roughly

similar manner by pairing an empirically weighted scoring function by means of a global optimization algorithm. The major difference lies between the parameterization of the scoring function and local search function of the two. AutoDockVina is compatible with the Auto Dock PDBQT file format and it offers the several advantages over Auto Dock 4. One advantage is that Vina is designed to function considerably faster and its accuracy in redocking protein-ligand complexes is greater than AutoDock4 which is shown by the authors (Trott and Olson, 2010). Other advantages are that using AutoDockVina grid computation is not necessary which was a complex process and it gives higher accuracy of binding mode, furthermore it is available for each operating system and it makes use of iterated local search algorithm (Chang *et al.*, 2010).

2.13 Quantitative Structure Activity Relationship

Three-dimensional quantitative structure-activity relationship (3D-QSAR) techniques are very beneficial and useful approaches involved in the ligand-based drug design by the correlation of a set of related compounds physicochemical descriptors to their known molecular activity or the values of their molecular properties (Hansch *et al.*, 1972; Yang *et al.*, 2007). 3D-QSAR is a mathematical model for a series of compounds in which correlation between the variation on chemical structure and the activity profile is found and is statistically validated (Da Cunha *et al.*, 2009). Nowadays the most commonly and routinely used 3D-QSAR techniques used in the field of modern drug design are the comparative molecular similarity analysis abbreviated as CoMSIA and comparative molecular field analysis abbreviated as CoMFA (Cramer *et al.*, 1988). QSAR is a very useful approach for the analysis of different potential drug molecules. It is basically employed into use to study the

biological activities with various structure associated properties and is a helpful source to elucidate how structural features in a drug molecule influence the biological activities of molecules. Other softwares that can be used are the ChemDraw (Zielesny, 2005), HyperChem (Tsuji, 2010) and many more are used for finding molecular descriptors.

A successful in silico-based QSAR analysis assures certain advantages like the higher speed and lower costs for bioactivity evaluation in comparison to experimental testing procedures. Hansch and Fujita first developed the QSAR about 40 years ago and it has remained a most valuable source in the lead discovery and optimization process for understanding the drug activity relations. This particular method involves recognition of organic, protein, peptide and still other molecules to be really three dimensional. In terms of a QSAR analysis the most important properties are the steric properties that include shape and volume along with the electronic properties such as charge and electrostatic potential last but not the least lipophilic properties, i.e. degree of polarity. Whole molecule descriptors are some of the earliest descriptors that are most dense in terms of information and are extremely informative. This particular type of descriptors captures information in relevance to molecular size and through the use of lipophilic properties such as molar weight or volume and logP that is the log of octanol-water partition coefficient (Winkler, 2002). The ChemDraw software package is a drawing tool for chemical structure and is capable of enabling several features including boiling point, melting point, critical volume, heat of formation, Log P and molar refractivity (MR) upon drawing the structure. HyperChem enables energy minimization of the compounds which lowers the systems energy by altering its molecular geometry, and helps in yielding a more stable conformation by generating a

log file using computational chemistry techniques such as semi-empirical formula, molecular mechanics etc (hypercube *et al.*, 2002).

Over several decades QSAR has been applied successfully and comprehensively to bioactive agents in order to predict activity models. It has been extensively applied to areas associated with the field of bioactive molecules discovery and development with the sole purpose to gain therapeutic benefits. These areas include drug resistance (Wiese and Pajeva, 2001), distinguishing drug like from non-drug like molecules (Ajay *et al.*, 1998), prediction of toxicity (Schultz and Seward, 2000; Andreas and Joop, 2001), physicochemical property prediction (Gombar and Enslein, 1996), peptide activity (Brusic *et al.*, 2001), datamining (Burden and Winkler, 2000), drug metabolism (Lewis, 2000), other pharmacokinetic and ADME properties prediction (Vedani and Dobler, 2000) and gastrointestinal absorption (Agatonovic *et al.*, 2001).

CHAPTER NO. 3

METHODOLOGY

3. Methodology

In-silico drug designing and development process makes use of screening huge set of compounds. Large sets of compounds were screened and correspondingly selected on the basis of particular selection criteria. The methodology adopted for this research study involves the application of ten basic steps to a set of anti-*Pneumocystis carinii* test compounds that have already been reported in several research studies. The protocol that compiles the entire methodology brought in to use for lead identification and devising a cure for *Pneumocystis carinii* Pneumonia is summarized in the Figure 3.1.

The in-silco drug designing approach adopted in this research study begins with the initial step of disease identification and followed by the identification of the potential protein target. The next step involves the collection of appropriate compounds to compose the data sets that meet the selection criteria in relevance to this research study, followed by drawing of selected compounds using the 2D view. The process of Pharmacophore generation was brought into use by the aid of LigandScout software which proceeded to the docking of compounds included in the data set. Molecular docking and QSAR study was performed as a next step in this study. The molecular docking results were interpreted and binding interactions were studied, as a result of which the lead compound was identified. The identification of the lead compound helped in the design of the respective analogues for further investigation. The steps involved in this protocol along with the challenges faced are discussed.

3.1 Disease Identification

One of the leading causes of pneumonia is the pathogen *Pneumocystis carinii* especially in patients with suppressed immune systems. PCP was selected as a target disease, since it is a fatal illness if it remains untreated. Despite advances in the treatment of AIDS the incidence of Pneumocystosis has increased markedly over the past decade (Walzer *et al.*, 1987). The incidence of PCP has also increased dramatically in the developing countries with mortality rates in the range of 20% to 80% (Fisk *et al.*, 2003). Moreover the anti-pneumocystis drugs that are currently available have a limited use due to the advent of various problems such as emerging resistance, efficacy and toxicity (Fishman, 1998).

3.2 Protein Target

Multiple set of targets can be brought into consideration in order to cure PCP. These targets include various specific enzymes found in *Pneumocystis* that can be used effectively as drug targets. DHFR was selected as the drug target in this study and the protein structure was extracted from Protein Data Bank. The protein structure used for this study is pcDHFR with the PDB ID: **1DYR** (Champness *et al.*, 1994).

3.3 Data Set Formation

Various *Pneumocystis carinii* inhibitors specifically the pcDHFR inhibitors have been identified and selected for study along with anti *P. carinii* drugs. Large sets of compounds were screened and correspondingly selected to compose the data set on the basis of particular selection criteria. The considerations for the selection of compounds include the fact that all compounds included in the data set should have a reported IC_{50} value not greater than 100 μM and these test compounds had already

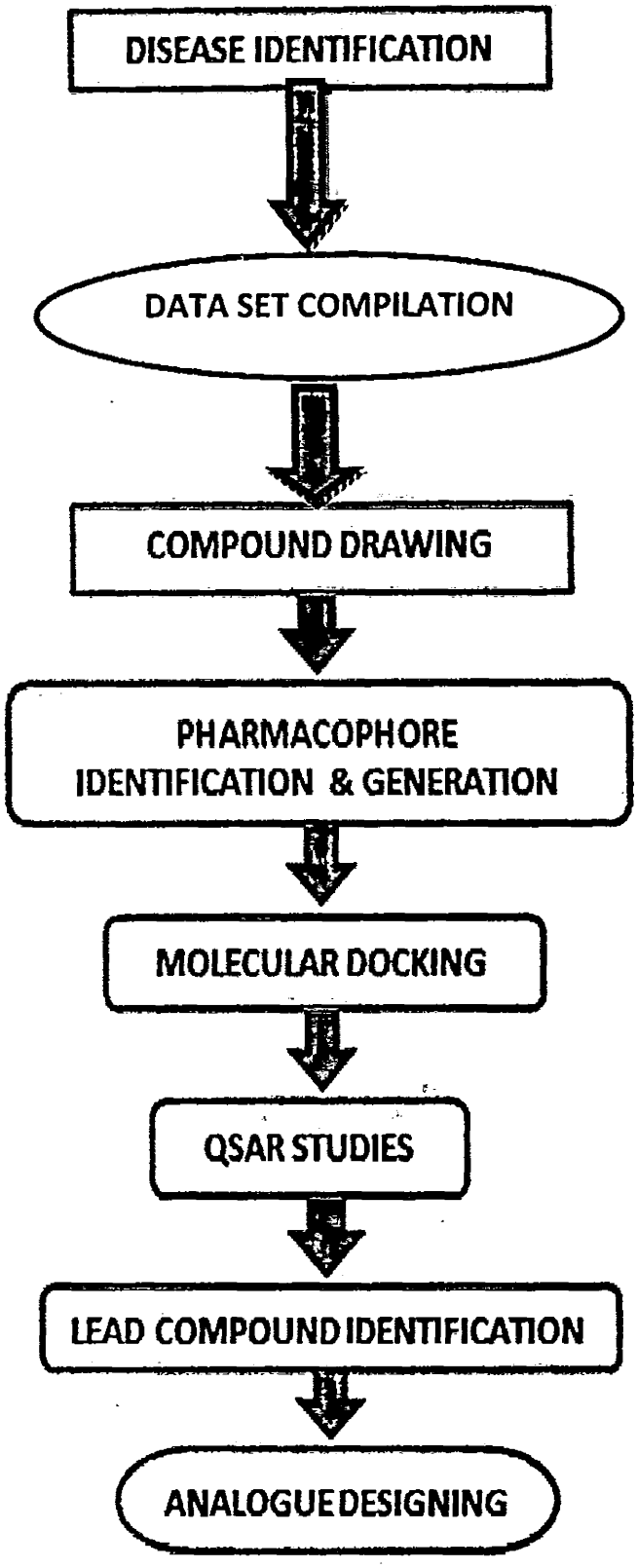


Figure 3.1: Protocol for the *Insilco* drug designing and development.

passed through bioassay. More over the compounds must belong to different classes that have numerous functional groups. Lastly all the compounds in this data set must not be reported earlier than 1990. The drugs should show pcDHFR inhibition activity and selected compounds should also be capable of taking part in combinatorial drug therapy.

The data set includes 46 compounds along with 4 drugs making a total of 50 compounds in the data set (Queener *et al.*, 1993; Bartlett *et al.*, 1995; Rosowsky *et al.*, 1995; Jackson *et al.*, 1996; Patrick *et al.*, 1997; Cushion *et al.*, 2004; Cushion *et al.*, 2006; Gangjee *et al.*, 2006; Da Cunha *et al.*, 2010). The FDA approved drugs incorporated into this study were namely Atovaquone, Primaquine, Sulfamethoxazole (SMX) and Trimethoprim (TMP) with IC_{50} values of 4.2 uM, 2.5 uM, 0.104 uM and 12 uM respectively (Broughton and Queener, 1991; Cushion *et al.*, 1999; Cushion *et al.*, 2004; Vale et al., 2009). The compounds included in the data set are shown in Table 3.1.

3.4 Compounds Drawing

In order to draw compounds many types of softwares are available. Compounds included in the data set were drawn using Chem Draw Ultra Version 8.0 (Cambridgesoft.com) (Mendelsohn *et al.*, 2004). ChemDraw provides user friendly drawing environment where one can draw structures by simple drag and drop method. It is basically a molecule editor that can be used to draw the chemical structures of compounds. It is the part of the ChemOffice suite of programs and is available for Macintosh and Microsoft Windows operating systems. Chemdraw supports a wide range of features such as chemical properties, logP values, molar refractivity, volume, melting point, chemical structure to name conversion and many other features.

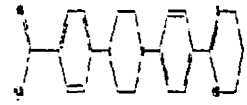
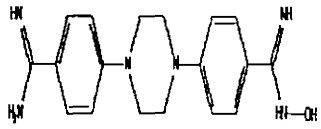
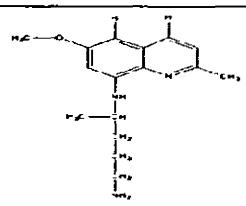
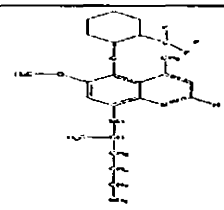
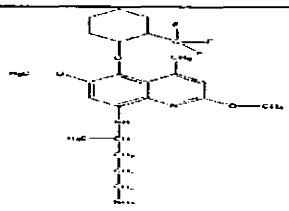
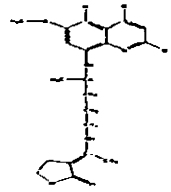
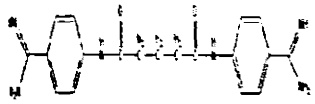
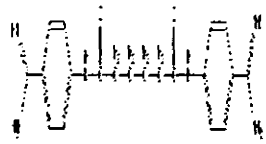
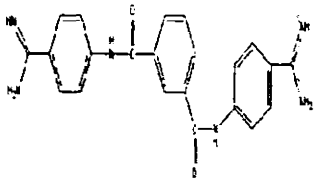
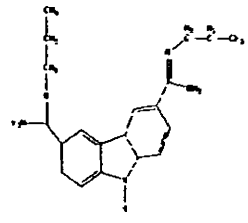
Chemical structures of all compounds were drawn using this tool and saved in 'cdx' format and then saved correspondingly in 'pdb' format using Chem3D Ultra 8.0 from the Chem Office suit of programs for pharmacophore modeling.

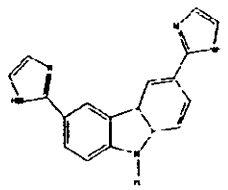
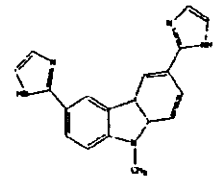
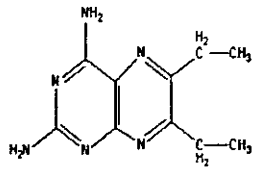
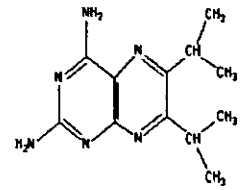
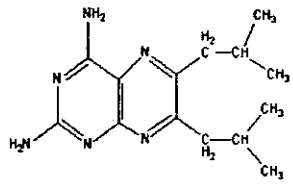
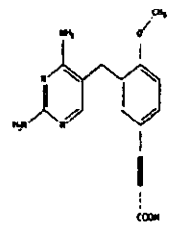
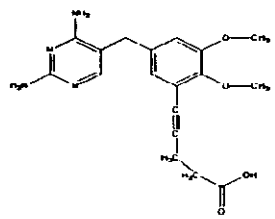
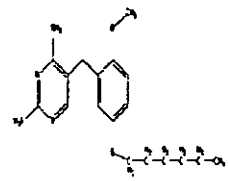
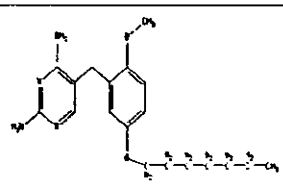
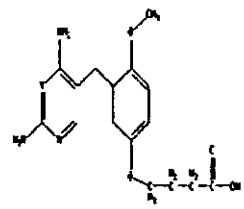
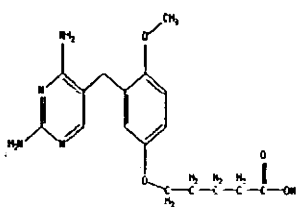
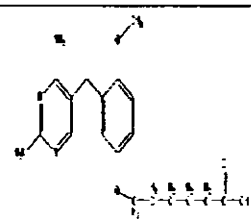
3.5 Pharmacophore Modeling

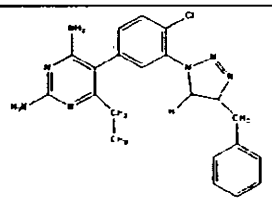
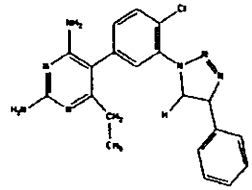
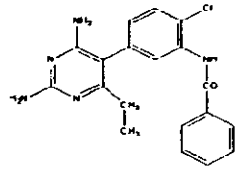
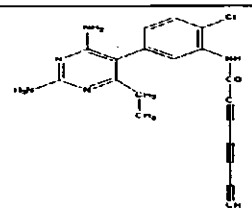
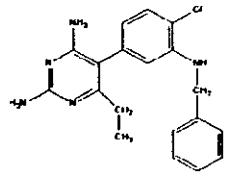
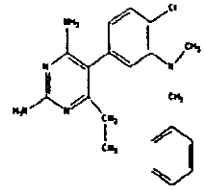
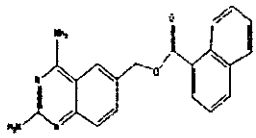
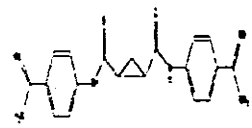
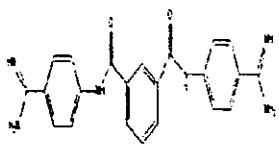
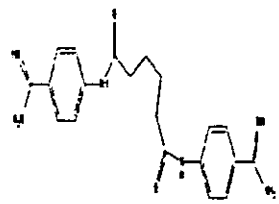
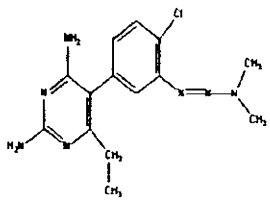
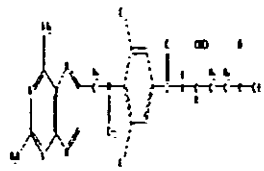
The 2D and 3D pharmacophores of the compounds were generated by the effective use of LigandScout Version 3.0, (Cambridge *et al.*, 2009) which is a powerful structure and ligand based pharmacophore model generator that is based on sophisticated algorithms for performing alignments, interpreting and customizing ligand-macromolecule interactions.

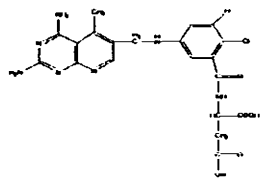
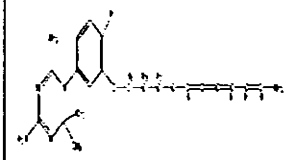
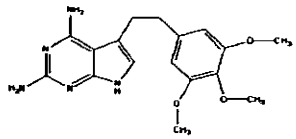
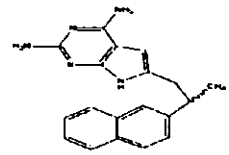
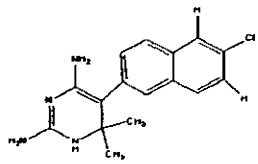
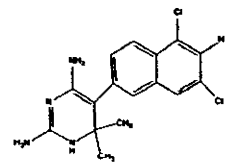
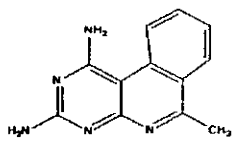
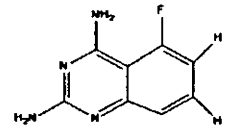
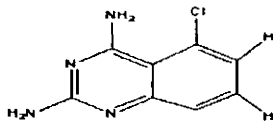
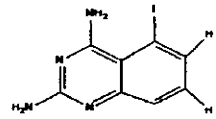
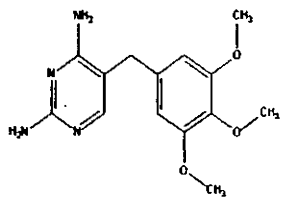
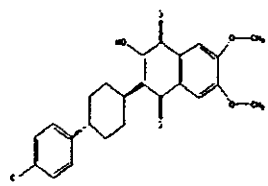
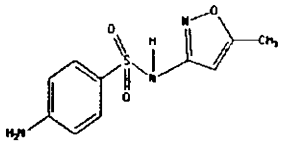
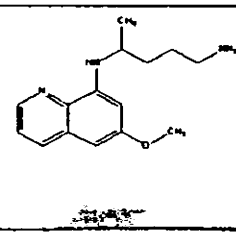
The main aim of this particular study was ligand based drug designing therefore ligand based pharmacophore models were generated. A 'pdb' file is taken as an input by the software. It extracts and interprets ligands and their macromolecular environment from 'pdb' files and visualizes advanced 3D pharmacophore models with multiple features. Pharmacophore was generated by using simple commands. Hence 'pdb' files were created for all compounds included in the data set followed by visualization of each compound using LigandScout. The command Create pharmacophore with default parameters was used to generate 3D pharmacophore for ligands. After the process of pharmacophore generation the common features of ligands were predicted. All the pharmacophore features i.e. aromatic rings; hydrophobic region, hydrogen bond acceptor and hydrogen bond donor were studied and analysed. Shared pharmacophore model was generated using the command 'generate shared feature pharmacophore'. As a result, the common feature pharmacophore was identified and obtained by superimposition of ligands. At the end a unique or shared pharmacophore was predicted for inhibitors of pcDHFR.

Table 3.1: Molecular Structures along with their IC_{50} Values of the dataset

Comps	STRUCTURE	IC_{50} (μM)	Comps	STRUCTURE	IC_{50} (μM)
YPCP1		1.31	YPCP2		1.71
YPCP3		3.3	YPCP4		6.48
YPCP5		0.028	YPCP6		0.18
YPCP7		0.17	YPCP8		0.027
YPCP9		0.001	YPCP10		0.0008
YPCP11		0.578	YPCP12		5-10

Comps	STRUCTURE	IC_{50} (μM)	Comps	STRUCTURE	IC_{50} (μM)
YPCP13		100	YPCP14		5-10
YPCP15		0.62	YPCP16		0.2
YPCP17		0.082	YPCP18		1.0
YPCP19		1.0	YPCP20		14
YPCP21		5.6	YPCP22		0.25
YPCP23		0.049	YPCP24		0.80

Comps	STRUCTURE	IC_{50} (μM)	Comps	STRUCTURE	IC_{50} (μM)
YPCP25		5.18	YPCP26		3.53
YPCP27		10.8	YPCP28		1.3
YPCP29		0.12	YPCP30		1.07
YPCP31		110	YPCP32		5.3
YPCP33		1.2	YPCP34		0.002
YPCP35		2.8	YPCP36		0.000035

Comps	STRUCTURE	IC_{50} (μM)	Comps	STRUCTURE	IC_{50} (μM)
YPCP37		0.0054	YPCP38		0.17
YPCP39		0.77	YPCP40		0.859
YPCP41		4.4	YPCP42		0.84
YPCP43		1.0	YPCP44		0.41
YPCP45		0.42	YPCP46		0.52
TMP		12	Atova-quone		4.2
SMX		0.104	Prima-quine		2.5

3.6 Molecular Field Analysis

FieldTemplater makes use of the molecules electrostatic and hydrophobic fields for finding common patterns by comparison of molecules. FieldTemplater helps to determine the bioactive conformations and relative alignments of molecules without requiring any receptor or binding site information when it is applied to several different molecular structures that have a common activity. It provides the complete pictorial layout defining how the active molecules bind, which features they make use of, their shape along with the view that how different series can be compared. Most important binding regions of a molecule are summarized by the Fields. FieldTemplater can be used for visual comparison of fields of two molecules and accurate determination of the overall similarity of the fields of two different molecules. It works with the common fields including van der Waals effects, the positive and negative electrostatic fields, and hydrophobic effects on and near the surface of a molecule. Highly similar interactions are displayed by molecules which bind to a common active site and hence they possess highly similar field point patterns. FieldTemplater is a tool that explores the conformational space of a set of ligands inspecting for commonality and it also searches for common field patterns.

For this particular study two most active ligands along with one standard Trimethoprim were taken as the input set for FieldTemplater. After the processing of these compounds by FieldTemplater, a template was generated. After the generation of the template, it was used as a reference for the alignment of further representative ligands of remaining classes using FieldAlign. For this purpose a set of about 19 pcDHFR inhibitors were aligned to the selected template. Thus a common template

was found which reveals the field points that are common in all structurally-distinct ligands against PCP.

3.7 Molecular Docking

In the field of molecular modeling, docking is the computer aided simulation of a candidate ligand binding to a receptor. The two important ingredients for the docking phase are the target protein and the ligand. In order to enter the docking phase the basic requirement is the identification of specific target protein hence a suitable target protein was chosen for anti-*Pneumocystis carinii* activity i.e. the pcDHFR with PDB ID: **1DYR** (Champness *et al.,* 1994) was selected which met all the requirements to be a suitable target protein. Docking was performed using software AutoDock (Goodsell *et al.,* 1996; Morris *et al.,* 1998) and its patch AutoDock Vina (Trott and Olson, 2010).

3.7.1 Steps for Molecular Docking

Docking of the chosen pcDHFR test compounds and standard drugs was performed using Autodock 4.0. The 3D structures of all compounds were placed in the same directory that contained the installed software. The 'pdb' file of the macromolecule was taken as input by software, modifications were made and the resultant macromolecule was saved in '.pdbqt' format. Kollman and Gasteiger charges were computed for the macromolecule. The macromolecule was then repaired for missing atoms after the charges were computed. The next step involved the addition of polar hydrogens to the macromolecule. After the processing the macromolecule was saved as a 'pdb' file in the Autodock folder of the software's directory as RH.pdb.

After modifying the protein the Kollman and Gasteiger charges were computed for the ligands. Similarly 'pdb' file of ligand was given as input to modify ligand file

which allowed the program to calculate its parameters such as non-polar hydrogen, aromatic carbons and rotatable bonds along with the ligand torsions. The Torsion count option was used to adjust the number of rotatable and non-rotatable bonds. For all the ligands Torsions were defined by the use of Torsion Tree option from the Autodock software. The Torsion Tree option was used to Choose Torsions, Set Number of Torsions to most atoms and to Detect Root which is the rigid part of the ligand. The number of active torsions was set to the maximum number of atoms. After the modifications the ligand was saved as a new file with the extension .pdbqt with the file name L.pdbqt. The rigid macromolecule was saved with an extension of .pdbqt after computing the charges and merging the hydrogens. This rigid protein file was brought into use for the preparation of grid parameter file which was prepared by using the ligand L.pdb. In order to provide the appropriate docking area and dimensions the properties of the Grid Box were set and saved later. The grid settings were saved as a grid parameter file with the extension '.gpf'. After this processing a text file was created that contained majorly three features: (I) macromolecular file in 'pdbqt' format, (II) ligand file in 'pdbqt' format and (III) 3D location of the grid from where docking algorithm was incorporated in order to search for docked site.

Grid location is important in the sense that it depends upon the active site of the protein. If the active site of the protein is the central pocket then grid must point at the center and so on in the specified direction accordingly. In case of pcDHFR active site the grid pointer was set at the center where center points of x axis = 17.32, y axis= 6.071, z axis = 1.694 respectively and size of the affinity grid was set to 60 x 60 x 60 and then the docking was performed.

For docking purpose AutoDockVina (Trott and Olson, 2010) was used which is an

open source program. Vina uses a 'conf' file that was helpful in referring the pdbqt files of macromolecule and ligands that were already prepared using AutoDock. Log parameter files was generated by running the command ("\Program Files\The Scripps Research Institute\Vina\vina.exe" --config conf.txt --log log.txt) on the command prompt. The outputs generated by Vina are the log files and pdbqt files of different energy models in accordance with the data set. For each ligand, the model with the lowest energy was selected amongst the list of models generated by Vina and was appended at the end of original protein file. The entire procedure outlined for one ligand was applied to all the 50 compounds. The docking parameter files for the entire set of ligands docked into the 1DYR macromolecule were obtained and analyzed.

3.7.2 Ligand-Protein Interactions

The ligand-protein interactions were visualized using Visual Molecular Dynamics (VMD) software. VMD (Humphrey *et al.*, 1996) is a molecular visualization and analysis program designed for biological systems such as proteins, nucleic acids, lipid bilayer assemblies, etc. The active conformation of ligand along with the target protein after being docked successfully was provided as input to the VMD. After this the interactions including ionic, hydrogen and hydrophobic between the ligand and the active site of the target were drawn by selecting atoms within 5 Å.

3.7.3 Lead Identification

After the extensive and detailed analysis of the ligand-protein interactions the most active pcDHFR inhibitor was identified. Binding interactions of all the compounds were observed thoroughly and the compound showing the best interactions was identified as the potential lead among all compounds in the data set. Lead was identified keeping in mind three considerations these include properties such as the

IC_{50} values, the Energy values of the model generated through docking and last but not the least the number of interactions, most prominently Hydrogen and ionic bonding.

3.7.4 Analogue Designing

After the identification of the lead four structural analogues of the lead were designed by introduction or elimination of various functional groups with the main aim of increasing the activity. Docking studies were then performed on the analogues using the same procedure already mentioned with the aid of the softwares, AutoDock 4.0 and its patch Vina.

3.8 Quantitative Structure Activity Relationship

A quantitative structure activity relationship (QSAR) analysis was performed for a set of 21 analogues derived from the pentamidine (Cushion *et al.*, 2004). These analogues were obtained by addition of linker, 1, 4-piperazinediyl parent compound with alkyl or cycloalkyl groups introduced on one side of the nitrogen atoms of the amidine moieties. QSAR was performed by computing a number of electronic and steric descriptors. To accomplish the task of descriptors computation ChemDraw (Zielesny, 2005) and HyperChem (Tsuji, 2010) were brought in to use. Hyper Chem generates a log file that was used to study electronic parameters whereas ChemDraw was used to study and calculate the steric parameters that were computed on several mouse clicks. More over the QSAR equation was calculated along with estimation of different statistical parameters. The values for IC-50 value are predicted and a graph is generated by plotting the actual and predicted IC-50 values.

CHAPTER NO. 4

RESULTS AND DISCUSSION

4. Results and Discussions

The research work is based on the extensive study and analysis of a series of compounds that exhibit anti *Pneumocystis carinii* activity in order to identify the most active compound. Computer aided drug designing (CADD) was incorporated to identify the most active compound i-e the Lead from the set of 46 compounds. Analogue designing and lead identification was accomplished by the effective use of CADD.

4.1 Data set formation

With reference to this research study a wide range of pcDHFR inhibitors were taken into consideration and a comprehensive analysis was conducted on the data set of compounds belonging to different classes. The data set was extended by including 4 FDA approved drugs along with the rest 46 compounds as the potential hits for this study.

These compounds belong to following classes: Piperazine-Linked and Alkanediamide or phenylenediamide-Linked Bisbenzamidines, the 8-Aminoquinolines, 2,4-diaminoquinazolines, the dicationic carbazoles, disubstituted Diaminopteridines, disubstituted pyrimidines, benzanilides and benzylamines derivatives , 1-naphthyl derivatives, diamide derivatives, triazolyl analogues, trimethoprim derivatives, methotrexate analogs ,trimetrexate analogs, Pyrimethamine analogs, Triazines and1,3-diamino-7,8,9,10-tetrahydropyrimido[4,5-*c*]isoquinolines (Queener *et al.*, 1993; Bartlett *et al.*, 1995; Rosowsky *et al.*, 1995; Jackson *et al.*, 1996; Patrick *et al.*, 1997; Cushion *et al.*, 2004; Cushion *et al.*, 2006; Gangjee *et al.*, 2006; Da Cunha *et al.*, 2010).

4.2 Rule of Five

Although the compounds and drugs included in the study have already undergone through the bioassay but in order to counter check their drug-likeness properties in-silico techniques such as the rule of 5 or Lipinski rule was brought in to use to incorporate the pharmacokinetics of the drugs. Lipinski rule or the rule of 5 should obey the following characteristics:

- Hydrogen bond acceptors should not be more than 10
- Similarly Hydrogen bond donors should not exceed 5
- The molecular weight should be under 500Da
- Log P should be less than 5

The data represented in the Table 4.1 reveals the results of Lipinski's rule (Rule of Five). The results show that all most all the compounds follow the HBA, molecular weight and Log P constraints but it is also revealed that some compounds deviate from the regular rule of five required range of HBD. The compounds that deviate from the required HBD criteria for the rule of five are the compounds namely YPCP10, YPCP11, YPCP32, YPCP33 and YPCP34. Only two compounds deviate from the logP criteria namely YPCP26 and YPCP30. The percentages computed against the Rule of five features that are fulfilled by the compounds in the data set are shown in Table 4.2 and demonstrated in Figure 4.1 by means of a bar chart. Thus the elucidation of the Lipinski rule or the Rule of five properties revealed that the HBD feature is not being followed by certain compounds. Hence the result is verified and compatible with the selected standard drugs thus it can be concluded that all the potential hits included in this specific study possess the required druggable properties.

Table 4.1: Lipinski's rule (Rule of Five) applied to complete data set

Compound No.	HBA ≤10	HBD ≤5	Molecular Weight (Da)	Log P
YPCP 1	1	3	396.49	4.12
YPCP 2	1	3	348.44	2.95
YPCP 3	2	3	338.19	2.83
YPCP 4	1	4	338.41	1.95
YPCP 5	2	2	273.37	2.17
YPCP 6	6	2	439.51	3.39
YPCP 7	7	2	469.54	3.63
YPCP 8	3	3	369.46	1.57
YPCP 9	2	4	366.42	0.56
YPCP 10	2	6	380.44	0.98
YPCP 11	2	6	400.44	1.93
YPCP 12	1	4	337.46	2.23
YPCP 13	2	3	301.35	0.8
YPCP 14	2	2	315.38	1.59
YPCP 15	4	2	218.26	2.48
YPCP 16	4	2	246.31	3.28
YPCP 17	4	2	274.36	3.97
YPCP 18	5	2	298.3	1.51
YPCP 19	6	2	356.38	1.75
YPCP 20	4	2	330.42	3.64
YPCP 21	4	2	344.45	4.05
YPCP 22	6	2	332.35	1.4
YPCP 23	6	2	346.38	1.82
YPCP 24	6	2	360.41	2.23
YPCP 25	2	4	407.9	2.35
YPCP 26	2	4	393.87	5.14
YPCP 27	3	3	367.83	3.95
YPCP 28	3	3	363.8	2.97

Compound No.	HBA ≤10	HBD ≤5	Molecular Weight (Da)	Log P
YPCP 29	2	3	353.85	4.37
YPCP 30	2	2	367.88	5.16
YPCP 31	4	3	344.37	3.69
YPCP 32	2	6	364.4	0.36
YPCP 33	2	6	400.43	1.93
YPCP 34	2	6	380.44	0.98
YPCP 35	2	4	319.79	3.67
YPCP 36	9	3	476.45	2.06
YPCP 37	8	4	473.87	1.41
YPCP 38	2	5	328.46	2.43
YPCP 39	3	5	343.38	2.14
YPCP 40	2	3	317.39	3.76
YPCP 41	0	4	300.79	2.93
YPCP 42	0	4	335.23	3.48
YPCP 43	3	2	178.17	1.3
YPCP 44	2	2	194.62	1.7
YPCP 45	2	2	286.07	2.5
YPCP 46	3	2	225.25	2.71
Trimethoprim	5	2	290.32	1.43
Atovaquone	5	2	426.89	3.42
Sulfamethoxazole	4	2	253.28	0.86
Primaquine	2	2	259.35	1.47

4.3 Pharmacophore Modeling

In order to interpret the interactions of ligand molecules with the macromolecular target protein, three-dimensional (3D) pharmacophore modeling techniques were brought in to use. The pharmacophore model of anti-PCP agents has not been reported yet therefore it is an attempt to generate the general pharmacophore model. The dataset selected from literature review was employed in to use in order to outline electronic interactions of bio-active compounds that are capable of producing comparable biological effects in a comparable situation within the same binding site (Wolber *et al.*, 2008).

LigandScout was used to generate the pharmacophore model of compounds included in the dataset. It was also used to identify a common pharmacophore that was shared among the ligands in the data set and the standard drugs. It essentially focused on the common features such as Hydrophobic, Hydrogen Bond Acceptors, Hydrogen Bond donors and positive ionizable that were present in compounds of the data set. The 2D and 3D pharmacophores of a few selective compounds from different classes of compounds were identified during the study along with the drugs. The 2D and 3D pharmacophores of YPCP11 and YPCP21 are shown in Figure 4.2 and 4.3. The 2D and 3D pharmacophores of drugs Sulfamethoxazole and Trimethoprim are shown in Figure 4.4 and 4.5 respectively. Figures 4.2- 4.5 indicate that the compounds consist of hydrophobic unit, hydrogen bonds (acceptor and donor) and aromatic rings. In accordance with LigandScout scheme of features every single compound contains hydrophobic patch shown by yellow sphere, hydrogen bond donors represented as Green spheres and hydrogen bond acceptors as red spheres and the rare aromatic rings are shown by means of Blue Circles.

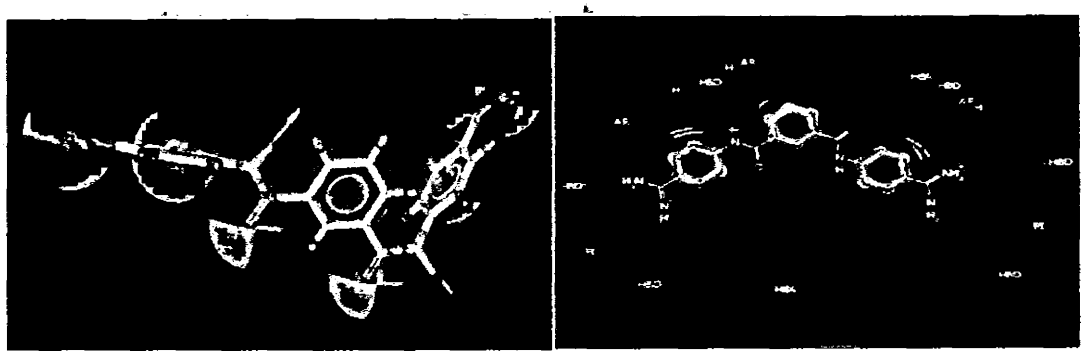


Figure4.2: 3D and 2D Pharmacophore Model of ligand YPCP11

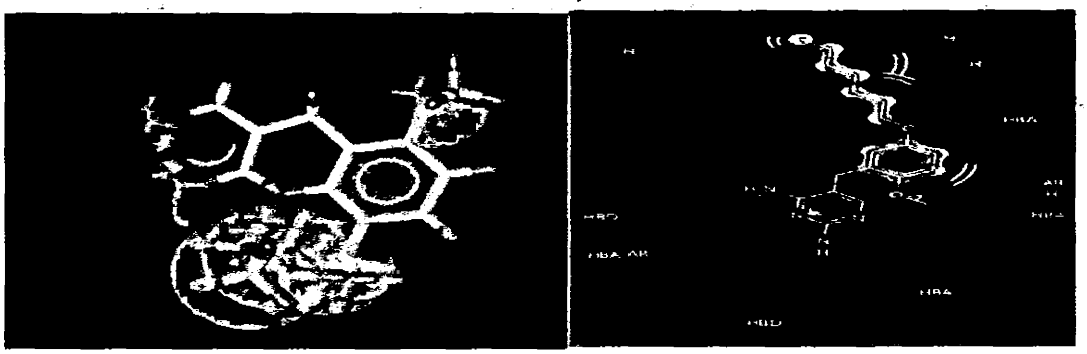


Figure 4.3: 3D and 2D Pharmacophore Model of ligand YPCP21

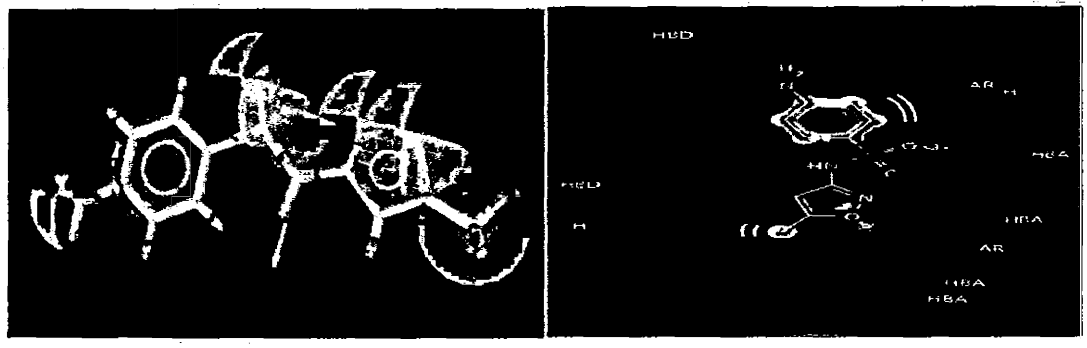


Figure 4.4: 3D and 2D Pharmacophore Model of drug Sulfamethoxazole

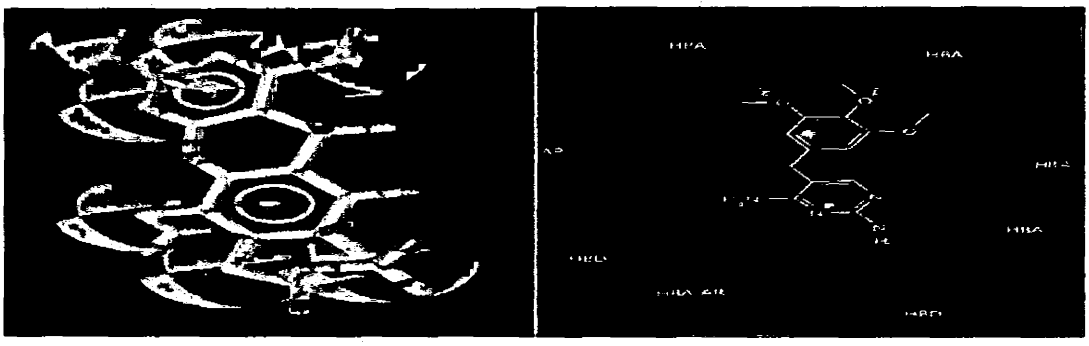


Figure 4.5: 3D and 2D Pharmacophore Model of standard drug Trimethoprim

4.3.1 Shared Pharmacophore Generation

In order to develop a pharmacophore model, ligands including YPCP17, YPCP18, YPCP20, YPCP21, YPCP24, YPCP25, YPCP27, YPCP29,YPCP31, YPCP44 and the standard drug Trimethoprim were superimposed. The merged pharmacophore is shown in Figure 4.6a. This shared pharmacophore indicates that every compound qualifying to have activity similar to the active pcDHFR inhibitors must have one Aromatic volume along with two hydrogen bond acceptors (HBA) and two hydrogen bond donors (HBD). The common featured pharmacophore predicted for the various classes of DHFR inhibitors against the PCP is represented in the Figure 4.6b; it indicates one Aromatic volume shown by yellow circle, two Hydrogen Bond Acceptors shown by red and two Hydrogen Bond donors shown by green. The calculation of these pharmacophoric features help in the identification of more effective, active and improved anti *Pneumocystis carinii* drugs. Pharmacophoric features of each class of compounds superimposed along with Trimethoprim to generate the shared pharmacophore are summarized in the Table 4.3.

4.3.2 Pharmacophore Triangle

Pharmacophore distance triangle was constructed which basically covers three featuresi.e. one HP/Ar feature, one HBD and one HBA. The distance ranges represented in the distance triangle were calculated with the aid of VMD. The distance range from HBD to HBA lieswithin 2.25–2.83Å whereas the distance from HBA to HP/Aris between 2.36–3.08Å and the distance from HP/Arto HBD is between 3.62–3.90Å.The pharmacophore triangle is shown in Figure 4.7.Table 4.4 shows the distance of the compounds used to identify the general pharmacophore model.

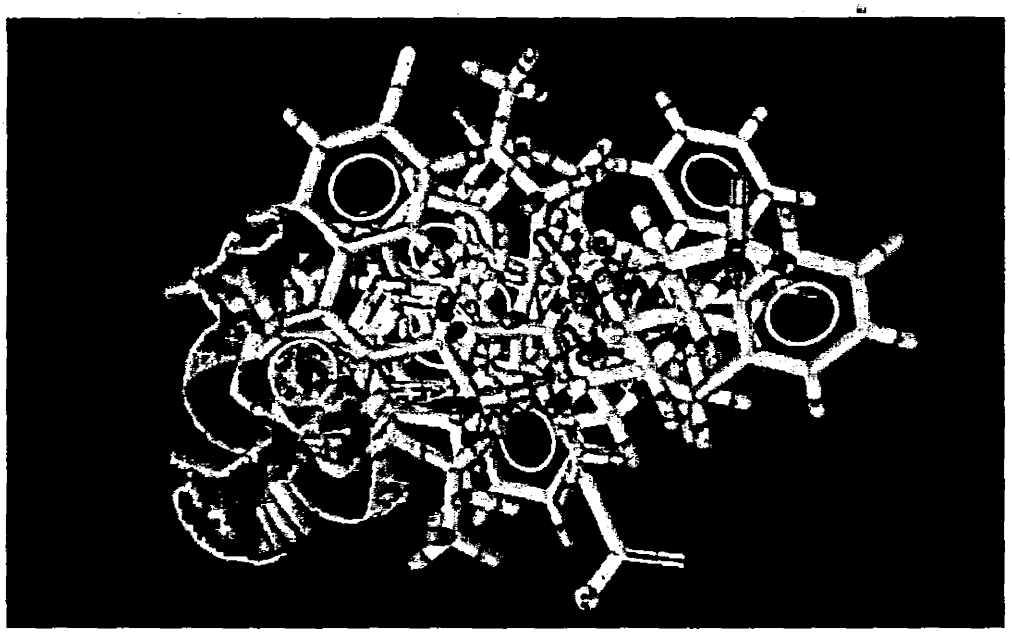


Figure 4.6a: Merged Pharmacophore of compounds YPCP17, YPCP18, YPCP20, YPCP21, YPCP24, YPCP25, YPCP27, YPCP29, YPCP31, YPCP44 and Trimethoprim generated by LigandScout

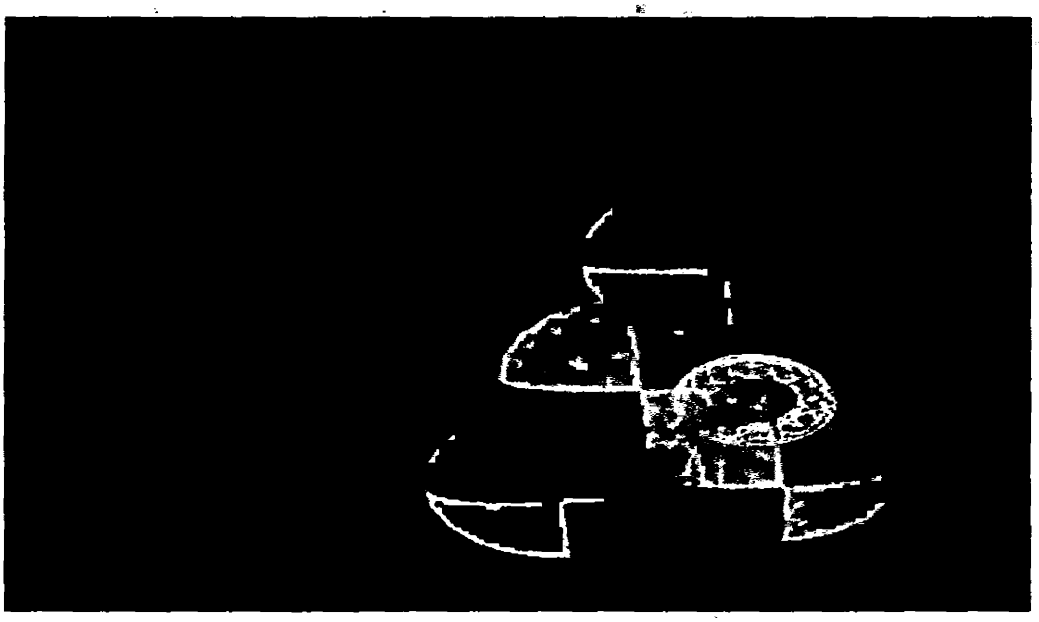


Figure 4.6b: Shared Pharmacophore showing two Hydrogen Bond Acceptors, two Hydrogen Bond Donors and one Aromatic volume

Table 4.3: Pharmacophoric Features of Data Set used for generation of a shared pharmacophore.

Compounds	HBDs	HBAs	Ar	HP	Positive ionizable	Negative ionizable
6,7-dibutyl-2-4-diaminopteridine	Two	Four	One	One	None	None
3-(3-((2,4-diaminopyrimidin-5-yl)-4-methoxyphenyl) propiolic acid	Two	Five	Two	One	None	One
5-(5-(hexyloxy)-2-methoxybenzyl)pyrimidine-2,4-diamine	Two	Four	Two	Three	None	None
5-(5-(heptyloxy)-2-methoxybenzyl)pyrimidine-2,4-diamine	Two	Four	Two	Four	None	None
6-(3-((2,4-diaminopyrimidin-5-yl)methyl)-4-methoxyphenoxy) hexanoic acid	Two	Six	Two	Two	None	One
5-(3-(4-benzyl-4,5-dihydro-1,2,3-triazol-1-yl)-4-chloro phenyl)-6-ethylpyrimidine-2,4-diamine	Four	Two	Three	Four	One	None
N-(2-chloro-5-(2,4-diamino-6-ethylpyrimidin-5-yl)phenyl) benzamide	Three	Three	Three	Four	None	None
5-(3-(benzylamino)-4-chlorophenyl)-6-ethylpyrimidine -2,4-diamine	Three	Two	Three	Four	None	None
(2,4-diaminoquinazoline-6-yl) methyl 1-naphthoate	Three	Four	Four	Three	None	None
5-chloroquinazoline-2,4-diamine	Two	Two	Two	Two	None	None
Trimethoprim	Two	Five	Two	None	None	None

4.4 Molecular field based similarity analysis

Cheeseright and coworkers have described the molecular fields in a form that enables similarity comparisons across molecules in three dimensions in a series of publications in the years 2004, 2006 and 2007 respectively. Cheeseright *et al* have also demonstrated the use of molecular fields as non-structural templates for defining similar biological behavior. It is also revealed and publicized that the field patterns can be used for the alignment of molecules that possess the capability of acting at the same target by means of their common field pattern and the field patterns derive the biologically active conformation of a ligand without access to any protein structural data. These principles are being incorporated in Field Templating however Field Aligning encompasses the virtual screening process of all the synthesized compounds using field patterns for potential hits. Both Field Templating and Field Aligning rely on the assumption that those molecules should be chosen for further investigation whose field patterns are most similar to those of an active search molecule and these molecules are most likely to show the same patterns of biological activity.

One FDA approved standard drug Trimethoprim and two Ligands were used by FieldTempler to generate templates with their own field pattern. Figure 4.8a shows the template generated through FieldTemplater. Of all the highest ranked template sets generated by FieldTemplater, the top-ranking template set was chosen as the master template on the basis of which calculation of field similarities across the whole data set was to be computed in this particular study. The scores for field alignment of the top-rank template set are given in Table 4.5.

Once a template has been generated, multiple molecules can be aligned using it as part of lead optimization process. The tool FieldAlign is capable of generating the 3D

conformers from a 2D molecule followed by the calculation of a field similarity score for each conformer. Figure 4.8b shows all the 19 ligands aligned with the template. Figure 4.9 (a-d) show the different fields; including hydrophobic, Van Der Waal, positive and negative fields, represented in different colors. Table 4.6 represents the values for similarity score against each 2D molecule given as input and aligned with the template. The software FieldAlign identifies, scores, and displays the conformers of the ligands that best match the template and the most probable conformer of each database molecules was selected on the basis of pair wise matching and close observation of molecular field point alignment.

Table 4.5: Scores for Top-Rank Template

S. No.	Template score	Value of score
1	Molecular similarity	0.630
2	Field similarity	0.563
3	Shape similarity	0.697

Table 4.6: Similarity scores after alignment of YPCPs with template

Compound	Similarity score	Compound	Similarity score
YPCP4	0.49	YPCP29	0.601
YPCP6	0.541	YPCP34	0.514
YPCP10	0.515	YPCP35	0.617
YPCP11	0.513	YPCP37	0.54
YPCP14	0.597	YPCP39	0.663
YPCP17	0.588	YPCP39	0.663
YPCP24	0.555	YPCP41	0.565
YPCP25	0.595	YPCP44	0.557
YPCP26	0.58	YPCP46	0.582
YPCP27	0.632		

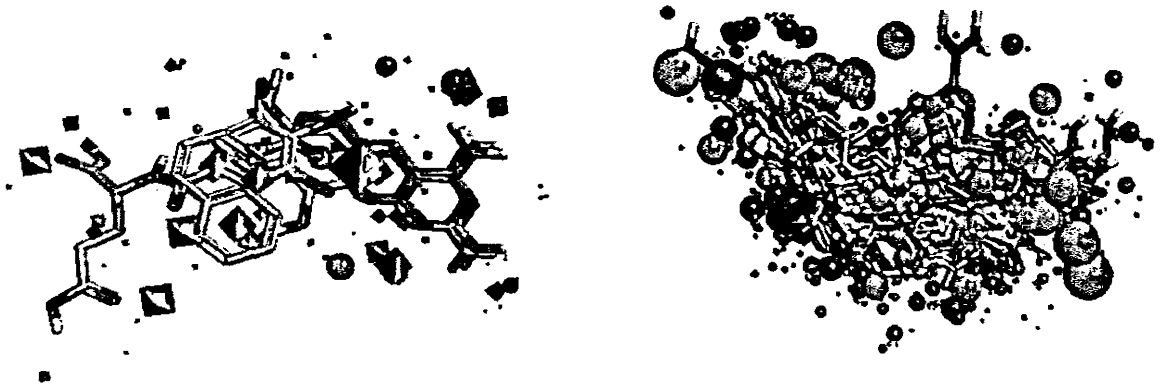


Figure 4.8a & b: (a) field alignment of template molecules in lowest energy conformations **(b)** field alignment of all pcDHFR inhibitors to the template ((the positive (red),negative (Blue), Van der Waals (yellow), and hydrophobic (orange) field point are represented as balls or cubes or polygons)

Figure 4.9a & b: Hydrophobic field points (Gold) & van der Waals surface field points (yellow)

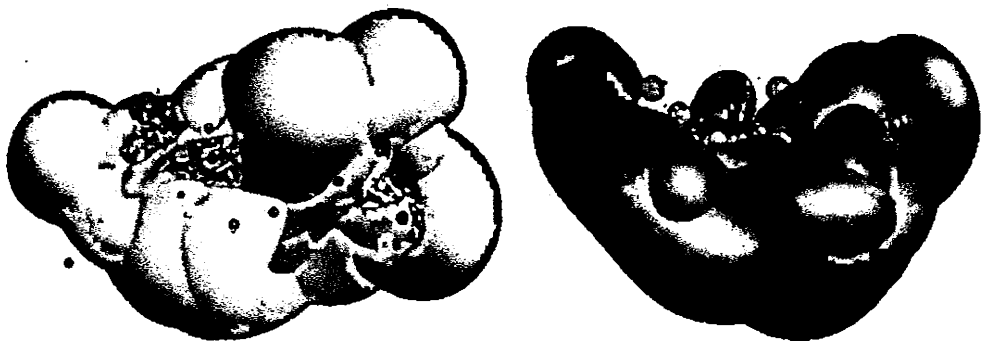


Figure 4.9c & d: Negative field points likely to interact with HBD (blue) & positive field points likely to interact with HBA (red)

4.5 Molecular Docking

Molecular docking studies were performed on the entire set compounds comprising of the pcDHFR inhibitors along with the drugs. The compounds were docked in order to obtain different conformations of the ligands docked into the protein 1DYR. The software AutoDock Vina was used to predict how ligands bind to the target and it was also helpful in indicating the best conformation in ligand protein binding interactions.

4.5.1 Active site of *P.carinii* DHFR

With reference to the pdb structure 1DYR enzyme complex, the active site was ranked second in an already reported study by the use of softwares OPLS and Amber94. It was also concluded that the best site for binding interactions is located approximately 20Å away. The active site is composed of the residues namely Asp139, Leu7, Ile20, Ser6, Lys5, Thr8, Trp180, Arg140, Arg119 and Asn118 as shown in Figure 4.10. This active site is shallow and the atoms of pterin that are accessible to the solvent are N1 and N8. This particular combination of residues that structure the active site seems to be unique to pcDHFR (Bliznyuk and Gready, 1997).

In this particular research study the compounds in the data set were docked into the active site one after the other, as a result it was identified that most of the compounds bind at the same active site position with a few deviations. The amino acids located within the active site were identified by observing the amino acids in the 5Å vicinity. The residues found within 5Å vicinity of active site that are also significantly involved in the binding interactions are namely: Leu7, Leu25, Leu72, Leu128, Ile10, Ile19, Ile33, Ile36, Ile64, Ile65, Ile80, Ile117, Ile123, Phe36, Phe69, Ser6, Ser24, Ser51, Ser64, Ser97, Thr8, Thr61, Thr81, Thr144, Thr183, Gly20, Gly58, Gly124, Gly125, Val11, Val154, Val181, Val185, Ala12, Ala126, Ala131, Asn23, Asn118,

Asp99, Asp134, Asp139, His135, Glu32, Lys5, Lys28, Lys37, Lys60, Lys96, Lys184, Tyr129, Arg21, Arg38, Arg59, Arg75, Arg82, Arg119, Trp27 and Pro66. Table 4.7 shows the list of amino acids within 5Å of the ligand docked with pcDHFR. The '+' indicates the presence of interaction with the amino acid whereas the '–' indicates the absence of interaction and the drugs Trimethoprim, Atovaquone, Sulfamethoxazole and Primaquine are denoted by the following alphabets, T for Trimethoprim, A for Atovaquone, S for Sulfamethoxazole and P for Primaquine.

4.5.2 Docking of compounds in the data set

The compounds added to the data set were docked into the active site of pcDHFR. In order to have a deeper understanding of binding activity of compounds the docking studies were brought into use. The docked files were thus used for visualizing all three types of binding interactions including the hydrogen bonding, hydrophobic and ionic interactions. The active conformations were used for identifying the ligand protein interactions using the software VMD. A detailed and comprehensive 3D study of the docked files revealed that all the compounds of dataset have the same amino acids within 5 Å of the ligand and all the interactions were computed after selecting the best conformation based on energy values. It was also observed during the analysis phase that the amino acids such as Ile10, Ile19, Ile33, Ile65, Ile123, Ser24, Ser64, Thr61, Leu25, Phe36, Glu32, Ala12, Val1, Val54, Gly124, Gly125, Tyr129 and Asn23, were most important in terms of involvement in the binding interactions. Figure 4.11 displays the amino acids within the active site of pcDHFR. Amongst all the conformations the best conformation with minimum binding energy were considered and their values recorded in Table 4.8a, however the interactions for all the compounds along with the drugs are displayed in Table 4.8b.

Table 4.7: Amino acids Present within the 5 Å Vicinity of the Ligand

A.A	LIGANDS (YPCP1 – YPCP18)																	
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
ILE10	-	-	-	+	-	+	+	-	-	+	+	-	+	-	-	-	+	-
ILE19	-	-	+	-	-	-	+	-	-	+	+	+	-	-	-	-	-	-
ILE33	+	-	-	-	-	-	+	-	+	+	-	+	-	-	+	+	-	+
ILE36	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
ILE64	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
ILE65	+	-	-	-	-	-	-	+	-	+	-	+	-	-	-	-	-	+
ILE80	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
ILE117	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
ILE123	-	-	+	+	-	-	-	-	-	+	-	-	-	-	+	+	+	+
SER6	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
SER24	-	-	-	-	-	-	-	-	+	-	-	+	-	+	-	-	-	-
SER51	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
SER64	-	-	-	-	-	+	+	+	-	+	-	-	-	+	+	+	-	+
SER97	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
THR8	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
THR61	+	+	-	-	-	+	+	+	-	+	-	-	-	+	+	+	+	-
THR81	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
THR144	-	-	-	-	-	-	-	-	-	-	+	-	-	-	-	-	-	-
THR183	-	-	-	-	+	-	-	-	-	-	-	-	-	-	-	-	-	-
LEU7	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
LEU25	+	+	-	-	-	-	+	+	+	+	+	+	+	+	+	-	+	+
LEU72	-	-	-	-	-	-	-	-	+	-	-	+	-	-	-	-	-	+
LEU128	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
PHE36	+	+	+	-	-	+	+	-	+	+	-	+	-	-	+	+	-	+
PHE69	-	-	-	-	-	-	-	-	-	-	-	+	-	-	-	-	-	-
PRO66	-	-	-	-	-	-	-	+	-	+	-	-	-	-	-	-	-	-
GLU32	-	-	+	-	-	-	-	-	-	-	+	+	-	-	-	-	-	+
LYS5	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
LYS28	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
LYS37	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	+
LYS60	+	+	+	+	-	-	-	+	+	-	+	-	+	-	+	+	-	-
LYS96	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
LYS184	-	-	-	-	+	-	-	-	-	-	-	-	-	-	-	-	-	-
ASP99	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
ASP130	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
ASP139	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
ALA12	-	+	+	-	-	-	+	-	-	+	+	+	-	-	+	+	-	+
ALA126	-	-	-	-	-	-	-	-	+	-	-	-	-	-	-	-	-	-
ALA131	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
ARG21	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
ARG38	-	-	-	-	+	-	-	-	-	-	-	-	-	-	-	-	-	-
ARG59	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
ARG75	-	-	-	-	-	-	-	-	+	-	-	-	-	-	-	-	-	+
ARG82	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
ARG119	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
VAL11	-	+	+	+	-	+	-	-	-	+	+	+	-	-	-	-	-	+
VAL154	-	-	+	+	-	-	-	+	+	-	+	-	+	-	-	-	+	-
VAL181	-	-	-	-	+	-	-	-	-	-	-	-	-	-	-	-	-	-
VAL185	-	-	-	-	+	-	-	-	-	-	-	-	-	-	-	-	-	-
GLY20	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
GLY58	-	-	-	-	-	-	-	-	-	-	-	-	+	-	-	-	-	-
GLY124	-	-	-	-	-	-	-	-	-	-	-	-	-	-	+	+	+	-

A.A	LIGANDS (YPCP19 – YPCP36)																	
	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36
ILE10	+	-	-	-	+	-	-	-	+	+	-	-	-	+	-	+	-	-
ILE19	+	+	+	-	+	-	+	-	-	+	-	+	-	+	+	+	-	+
ILE33	+	-	+	+	-	+	-	-	+	-	+	-	+	-	+	+	-	+
ILE36	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
ILE64	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	+	-
ILE65	+	+	-	+	+	+	+	-	+	-	+	+	+	-	-	+	-	-
ILE80	-	-	-	-	-	-	-	+	-	-	-	-	-	-	-	-	-	-
ILE117	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
ILE123	+	-	-	-	+	-	-	-	+	-	+	-	-	+	-	+	-	+
SER6	-	-	-	-	-	-	-	-	-	+	-	-	-	-	-	-	-	-
SER24	-	-	+	+	-	+	-	-	-	-	-	-	-	-	-	-	-	+
SER51	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
SER64	+	+	+	+	-	+	+	-	+	+	+	+	+	-	-	+	+	-
SER97	-	-	-	-	-	-	-	+	-	-	-	-	-	-	-	-	-	-
THR8	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
THR61	-	+	+	-	+	+	+	-	-	+	+	-	+	+	+	-	+	+
THR81	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
THR144	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	+	-
THR183	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
LEU7	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
LEU25	+	-	+	+	-	+	+	-	-	-	-	-	-	-	+	+	+	+
LEU72	-	-	-	-	-	-	-	-	-	-	+	-	-	-	-	-	-	-
LEU128	-	-	-	-	-	-	-	-	-	-	-	-	-	+	-	-	-	-
PHE36	-	-	+	+	+	-	+	-	+	+	+	-	+	-	-	+	-	-
PHE69	+	-	-	+	-	-	-	-	-	-	-	-	-	-	-	-	-	-
PRO66	+	-	-	-	-	-	-	-	-	+	+	-	-	-	-	-	-	-
GLU32	-	-	-	+	+	+	-	-	-	-	-	-	-	-	-	+	+	+
LYS5	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
LYS28	-	-	-	-	-	+	-	-	-	-	-	-	-	-	-	-	-	-
LYS37	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
LYS60	-	-	-	-	-	-	+	-	-	-	+	-	+	+	+	-	-	+
LYS96	-	-	-	-	-	-	-	+	-	-	-	-	-	-	-	-	-	-
LYS184	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
ASP99	-	-	-	-	-	-	-	+	-	-	-	-	-	-	-	-	-	-
ASP130	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
ASP139	-	-	-	-	-	-	-	+	-	-	-	-	-	-	-	-	-	-
ALA12	+	+	+	+	+	+	-	-	+	-	-	+	-	+	-	+	+	+
ALA126	-	-	-	-	-	-	+	-	-	-	-	-	-	+	+	-	-	-
ALA131	-	-	-	-	-	-	-	+	-	-	-	-	-	-	-	-	-	-
ARG21	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
ARG38	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
ARG59	-	-	-	-	-	-	-	-	-	-	-	-	-	+	-	-	-	+
ARG75	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
ARG82	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
ARG119	-	-	-	-	-	-	-	-	-	+	-	-	-	-	-	-	-	-
VAL11	+	-	+	-	+	-	-	-	+	-	-	-	-	+	-	+	+	-
VAL154	-	-	-	-	-	-	-	-	-	-	-	-	-	+	+	-	-	-
VAL181	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
VAL185	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
GLY20	-	+	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	+
GLY58	-	-	-	-	-	-	-	-	-	-	-	-	-	+	-	-	-	-
GLY124	-	-	-	-	+	-	-	-	+	-	-	-	-	-	-	-	-	-

A.A	LIGANDS (YPCP37 – YPCP46)																
	37	38	39	40	41	42	43	44	45	46	T	A	S	P			
ILE10	-	-	-	-	-	-	-	-	-	-	+	-	-	+			
ILE19	-	+	+	+	-	+	+	+	+	-	+	+	-	+			
ILE33	-	-	+	-	-	-	-	-	+	-	-	+	-	-			
ILE36	-	-	+	-	-	-	-	-	-	-	-	-	-	-			
ILE64	-	-	-	-	-	-	-	-	-	-	-	-	-	-			
ILE65	-	-	+	-	-	+	-	-	-	-	-	+	-	+			
ILE80	-	+	-	-	-	-	-	-	-	-	-	-	-	-			
ILE117	+	-	-	-	-	-	-	-	-	-	-	-	-	-			
ILE123	-	-	+	-	-	-	-	+	+	-	+	-	-	+			
SER6	+	-	-	-	-	-	-	-	-	-	-	-	-	-			
SER24	-	-	-	+	-	+	-	-	-	-	-	+	-	-			
SER51	+	-	-	-	-	-	-	-	-	-	-	-	-	-			
SER64	-	+	-	-	+	+	-	-	-	-	-	-	+	-			
SER97	-	-	-	-	-	-	-	-	-	-	-	-	-	-			
THR8	+	-	-	-	-	-	-	-	-	-	-	-	-	-			
THR61	-	+	-	+	+	+	+	-	+	-	-	-	+	+			
THR81	-	-	-	-	-	-	-	-	-	+	-	-	-	-			
THR144	-	-	-	-	-	-	+	-	-	-	-	-	-	-			
THR183	-	-	-	-	-	-	-	-	-	-	-	-	-	-			
LEU7	+	-	-	-	-	-	-	-	-	-	-	-	-	-			
LEU25	-	-	-	+	+	+	+	-	-	-	+	+	-	-			
LEU72	-	-	-	-	-	-	-	-	-	-	-	-	-	-			
LEU128	-	+	-	-	-	-	-	-	-	-	-	-	-	-			
PHE36	-	-	+	+	-	+	+	+	+	-	-	+	+	+			
PHE69	-	-	-	-	-	-	-	-	-	-	-	+	-	-			
PRO66	-	-	-	-	-	-	-	-	-	-	-	-	-	-			
GLU32	-	-	+	+	-	-	+	+	-	-	+	-	-	+			
LYS5	+	-	-	-	-	-	-	-	-	-	-	-	-	-			
LYS28	-	-	-	-	-	-	-	-	-	-	-	-	-	-			
LYS37	-	-	-	-	-	-	-	-	-	-	-	-	-	-			
LYS60	-	+	-	-	+	-	-	-	-	+	-	-	+	-			
LYS96	-	-	-	-	-	-	-	-	-	-	-	-	-	-			
LYS184	-	-	-	-	-	-	-	-	-	-	-	-	-	-			
ASP99	-	-	-	-	-	-	-	-	-	-	-	-	-	-			
ASP130	-	-	-	-	-	-	-	-	-	-	-	-	-	-			
ASP139	+	-	-	-	-	-	-	-	-	-	-	-	-	-			
ALA12	-	-	+	+	-	-	+	+	+	-	-	-	+	+			
ALA126	-	+	-	-	-	-	-	-	-	-	-	+	-	-			
ALA131	-	-	-	-	-	-	-	-	-	-	-	-	-	-			
ARG21	-	-	-	-	-	+	-	-	-	-	-	-	-	-			
ARG38	-	-	-	-	-	-	-	-	-	-	-	-	-	-			
ARG59	-	+	-	-	-	-	-	-	-	+	-	-	-	-			
ARG75	-	-	-	-	-	-	-	-	-	-	-	-	-	-			
ARG82	-	-	-	-	-	-	-	-	-	+	-	-	-	-			
ARG119	+	-	-	-	-	-	-	-	-	-	+	-	-	-			
VAL11	-	-	-	-	-	-	-	-	-	-	-	+	-	+			
VAL154	-	-	-	-	+	-	-	-	-	-	-	-	-	-			
VAL181	-	-	-	-	-	-	-	-	-	-	-	-	-	-			
VAL185	-	-	-	-	-	-	-	-	-	-	-	-	-	-			
GLY20	-	+	-	-	+	-	-	-	-	-	-	-	-	-			
GLY58	-	+	-	-	-	-	-	-	-	-	-	-	-	-			
GLY124	-	-	+	+	-	-	-	-	-	-	-	-	-	-			

A.A	LIGANDS (YPCP1 – YPCP17)																
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
GLY125	-	+	-	-	-	-	+	-	-	+	-	-	-	+	+	+	-
GLN127	-	-	-	-	-	-	-	+	+	-	+	-	-	-	-	-	-
TRP27	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
HIS135	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
TYR129	-	-	+	+	-	-	-	-	-	+	-	-	+	-	+	+	+
ASN23	-	+	+	+	-	+	-	-	-	-	+	-	+	+	-	-	-
ASN118	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
A.A	LIGANDS (YPCP19 – YPCP35)																
	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35
GLY125	-	-	-	-	+	-	-	-	+	-	-	+	-	+	-	-	-
GLN127	-	-	-	-	-	-	-	-	-	-	-	-	-	+	+	-	-
TRP27	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
HIS135	-	-	-	-	-	-	-	+	-	-	-	-	-	-	-	-	-
TYR129	+	+	-	-	+	+	-	-	+	-	+	-	-	+	-	+	-
ASN23	-	-	-	+	-	-	+	-	-	+	+	-	-	+	-	-	-
ASN118	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
A.A	LIGANDS (YPCP37 – YPCP46)																
	37	38	39	40	41	42	43	44	45	46	T	A	S	P			
GLY125	-	+	-	-	+	-	-	-	-	-	+	-	+	-			
GLN127	-	+	-	-	-	-	-	-	-	+	-	-	-	-			
TRP27	-	-	-	-	-	-	-	-	-	-	-	-	-	-			
HIS135	-	-	-	-	-	-	-	-	-	-	-	-	-	-			
TYR129	-	-	+	+	-	-	-	+	+	-	+	+	+	+			
ASN23	-	-	-	-	+	+	-	-	-	-	-	-	+	-			
ASN118	+	-	-	-	-	-	-	-	-	-	-	-	-	-			

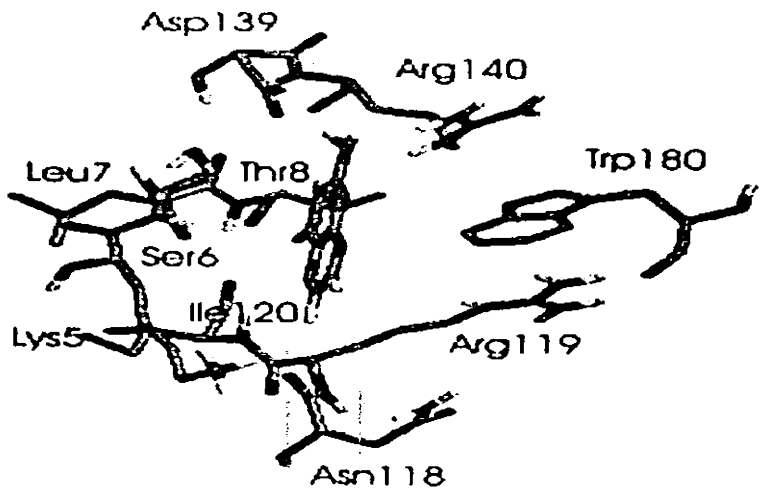


Figure 4.10: The shallow active site of *P.carinii* for pterin binding.

4.5.3 Molecular Docking of Standard drugs

Docking of standard drugs was performed using the same parameters and molecular docking tools. Standard drugs were docked in order to compare their binding activity with the entire set of compounds to determine the use of these compounds as future drugs. The docking analysis indicated that the standard drugs bind to the same active site. Figure 4.12 shows the interactions of the standard drug Trimethoprim.

While considering the binding activity of **Trimethoprim** it was revealed that it was involved in seventeen active binding interactions with amino acids Glu32, Ile10, Ile19, Ile123, Val11, Tyr129 and Gly125 respectively. Three Hydrogen bonds were formed between O of Glu32 and H of the ligand with distances 3.093 Å, 3.913 Å and 2.532Å. Whereas the remaining four hydrogen bonds were formed between O of Ile10 and H of ligand, N of Val11 and H of ligand, O of Ile123 and H of ligand and OH of Tyr129 and H of ligand with the distances 2.595 Å,3.786 Å,2.365 Å and 3.172 Å respectively. However amino acid involved ionic interactions were Glu32, Ile10, Ile123 and Tyr129. Five ionic bonds were formed between O of Glu32 and N of ligand, O of Ile10 and N of ligand, O of Ile123 and N of ligand and lastly between OH of Tyr129 and N of ligand with the following distances 2.884 Å,2.759 Å,2.916 Å,3.564 Å and 2.800 Å. Five Hydrophobic interactions were found between C of Gly125 and C of ligand, C of Ile19 and C of ligand, C of Leu25 and C of ligand with the distances 3.993 Å,3.894 Å,3.757 Å,3.99 Å and 3.99 Å respectively.

In case of **Atovaquone** all three types of interactions Ionic, hydrogen and Hydrophobic were observed. One Hydrogen bond with a distance of 3.112 Å was formed between O of Ile19 and H of ligand. Two ionic interactions with a distance of 3.061Å, 3.999Å were observed between N of Ala126 and O of ligand and N of Ser24

and the ligand. These interactions were found between C of ligand and C of Ile65, C of Phe36, C of Glu32, C of Ile10, C of Tyr129, C of Ile19 and C of Val11 with the following distances 3.946Å, 3.858Å, 3.734Å, 3.935Å, 3.992Å, 3.572Å, 3.942Å, 3.960Å, 3.537Å, 3.692Å, 3.823Å, 3.932Å, 3.991Å, 3.771Å, 3.864Å, 3.813Å and 3.875Å. The detailed binding interactions of the drugs are represented in Table 4.8b.

4.6 Lead compound identification

The lead identification process was accomplished on the basis of the binding interactions, the binding energy value and the IC_{50} value of the ligands. It is evident from the analysis that compounds YPCP1, YPCP9, YPCP11, YPCP25, YPCP28, YPCP30, YPCP31, YPCP33 and the drug Atovaquone have lower binding energies. The compound YPCP6 has the lowest value i.e. -11.5 Kcal/mol among all the compounds in the data set and hence it must be more potent than other compounds included in the data set but it was biased to take this decision by the considering binding affinity solely. To predict a compounds activeness IC_{50} value and binding interactions were also important.

It is clearly evident from the Table 4.8b that compounds YPCP3, YPCP10, YPCP11, YPCP15, YPCP16, YPCP37, YPCP31, YPCP33, YPCP36, YPCP41, YPCP42 and YPCP44 were involved in a good and considerable number of all the three types of interactions including the hydrogen, ionic and hydrophobic interactions. However three drugs namely Trimethoprim, Sulfamethoxazole and Primaquine were also involved in a notable number of interactions and Atovaquone was involved with the minimum number of interactions. The compounds having least activity in terms of all three types of interactions are the YPCP1, YPCP2, YPCP6, YPCP7, YPCP8, YPCP14, YPCP30, YPCP32 and YPCP40.

Table 4.8a: Inhibition Concentration and Energy Value of the Data Set

Compound Number	IC_{50} (μM)	Energy Value (Kcal/mol)
YPCP1	1.31	-10.5
YPCP2	1.71	-9.3
YPCP3	3.3	-9.8
YPCP4	6.48	-9.0
YPCP5	0.028	-8.9
YPCP6	0.18	-11.5
YPCP7	0.17	-8.7
YPCP8	0.027	-9.6
YPCP9	0.001	-10.5
YPCP10	0.0008	-8.3
YPCP11	0.578	-11.4
YPCP12	5-10	-8.0
YPCP13	100	-9.5
YPCP14	5-10	-8.4
YPCP15	0.62	-6.3
YPCP16	0.2	-8.1
YPCP17	0.082	-8.1
YPCP18	1.0	-8.1
YPCP19	1.0	-9.5
YPCP20	14	-9.1
YPCP21	5.6	-7.4
YPCP22	0.25	-9.5
YPCP23	0.049	-6.9
YPCP24	0.80	-9.3
YPCP25	5.18	-10.4

Compound Number	IC_{50} (μM)	Energy Value (Kcal/mol)
YPCP26	3.53	-8.3
YPCP27	10.8	-8.8
YPCP28	1.3	-10.4
YPCP29	0.12	-9.0
YPCP30	1.07	-10.5
YPCP31	110	-11.3
YPCP32	5.3	-8.9
YPCP33	1.2	-11.0
YPCP34	0.002	-7.9
YPCP35	2.8	-8.4
YPCP36	0.000035	-9.3
YPCP37	0.0054	-8.9
YPCP38	0.17	-8.7
YPCP39	0.77	-7.8
YPCP40	0.859	-8.9
YPCP41	4.4	-9.7
YPCP42	0.84	-9.5
YPCP43	1.0	-8.4
YPCP44	0.41	-6.8
YPCP45	0.42	-6.4
YPCP46	0.52	-6.8
Trimethoprim	12	-7.2
Atovaquone	4.2	-11.4
Sulfamethoxazole	0.104	-7.9
Primaquine	2.5	-7.1

Table 4.8b: Binding Interactions and distances of Data Set showing all the three kinds of interactions including Hydrogen Bonding, Ionic and Hydrophobic Interactions

LIGAND NO	HYDROGEN BONDING		IONIC BONDING		HYDROPHOBIC INTERACTIONS		IC_{50}
	Amino Acids	Dist	Amino Acids	Dist	Amino Acids	Dist	
YPCP1	No hydrogen bonds		No ionic bonds		PHE36:CG UNKO:C PHE36:CD2 UNKO:C PHE36:CE2 UNKO:C ILE33:CD1 UNKO:C ILE33:CD1 UNKO:C ILE65:CD1 UNKO:C ILE65:CG1 UNKO:C ILE65:CG1 UNKO:C SER64:CB UNKO:C SER64:CB UNKO:C THR61:CG2 UNKO:C LEU25:CD1 UNKO:C LYS60:CB UNKO:C	3.771 3.344 3.501 3.303 3.694 3.728 3.512 3.891 3.788 3.718 3.676 3.788 3.945	1.31
YPCP2	ASN23:OD1 UNKO:H ASN23:ND2 UNKO:H LYS60:NZ UNKO:H	2.233 2.965 3.083	THR61:OG1 UNKO:N ASN23:OD1 UNKO:N	3.565 3.210	ALA12:CB UNKO:C VAL11:C UNKO:C PHE36:CD1 UNKO:C PHE36:CB UNKO:C PHE36:CG UNKO:C LEU25:CD1 UNKO:C LEU25:CD1 UNKO:C LEU25:CD1 UNKO:C LEU25:CD1 UNKO:C LEU25:CD1 UNKO:C THR61:CA UNKO:C GLY125:CA UNKO:C	3.783 3.770 3.243 3.647 3.601 3.993 3.863 3.865 3.422 3.578 3.949 3.299	1.71
YPCP3	VAL11:O UNKO:H ALA12:N UNKO:H GLU32:OE1 UNKO:H GLU32:OE1 UNKO:H GLU32:OE2 UNKO:H GLU32:OE2 UNKO:H GLU32:OE1 UNKO:H LYS60:NZ UNKO:H ASN23:OD1 UNKO:H ASN23:ND2 UNKO:H ASN23:N UNKO:H	3.536 3.100 3.473 3.769 3.328 3.075 3.417 3.127 2.761 3.366 3.579	GLU32:OE1 UNKO:N GLU32:OE2 UNKO:N GLU32:OE1 UNKO:N ILE123:O UNKO:N LYS60:NZ UNKO:O ASN23:OD1 UNKO:N ASN23:OD1 UNKO:N	3.321 3.229 2.754 3.837 3.191 2.800 3.833	PHE36:CD1 UNKO:C GLU32:CD UNKO:C ALA12:CB UNKO:C ASN23:CA UNKO:C VAL154:CG2 UNKO:C VAL154:CG' UNKO:C VAL154:CG' UNKO:C ILE19:CD1 UNKO:C TYR129:CE2 UNKO:C TYR129:CZ UNKO:C TYR129:CE2 UNKO:C	3.828 3.988 3.905 3.878 3.211 3.993 3.547 3.765 3.778 3.934 3.931	3.3
YPCP4	ILE10:O UNKO:H VAL11:N UNKO:H ILE10:O UNKO:H ILE123:O UNKO:H ILE123:O UNKO:H LYS60:NZ UNKO:H ASN23:OD1 UNKO:H ASN23:ND2 UNKO:H TYR129:OH UNKO:H	2.156 3.409 2.946 2.513 3.221 3.151 2.447 2.434 1.942	ALA12:N UNKO:O ILE10:O UNKO:N ILE10:O UNKO:N ASN23:OD1 UNKO:N ILE123:O UNKO:N ILE123:O UNKO:N	3.006 2.849 3.144 3.424 3.272 3.166	VAL11:CA UNKO:C VAL154:CG2 UNKO:C VAL154:CG2 UNKO:C VAL154:CG2 UNKO:C ASN23:CG UNKO:C LYS60:CE UNKO:C ASN23:CA UNKO:C	3.775 3.866 3.396 3.238 3.571 3.887 3.623	6.48

LIGAND NO	HYDROGEN BONDING			IONIC BONDING			HYDROPHOBIC INTERACTIONS			IC_{50}
	Amino Acids		Dist	Amino Acids		Dist	Amino Acids		Dist	
YPCP10	ILE65:N	UNKO:H	3.222	SER64:O	UNKO:N	3.181	PHE36:CE2	UNKO:C	3.967	0.000
	SER64:O	UNKO:H	2.144	ILE123:O	UNKO:N	3.181	PHE36:CZ	UNKO:C	3.886	8
	PRO66:N	UNKO:H	3.913	THR61:OG1	UNKO:N	3.212	PHE36:CE2	UNKO:C	3.770	
	SER64:O	UNKO:H	3.849	ALA12:N	UNKO:O	3.074	ILE33:CD1	UNKO:C	3.588	
	LEU25:N	UNKO:H	2.234	VAL11:N	UNKO:O	3.972	PHE36:CG	UNKO:C	3.828	
	THR61:OG1	UNKO:H	2.213				PHE36:CD2	UNKO:C	3.763	
	GLY125:N	UNKO:H	3.789				PHE36:CE2	UNKO:C	3.901	
	ILE10:O	UNKO:H	3.468				PHE36:CD2	UNKO:C	3.712	
	ILE123:O	UNKO:H	2.454				PHE36:CE1	UNKO:C	3.788	
	TYR129:OH	UNKO:H	3.618				PHE36:CZ	UNKO:C	3.905	
							SER64:CB	UNKO:C	3.971	
							SER64:CB	UNKO:C	3.793	
							GLY125:CA	UNKO:C	3.808	
							ILE19:CD1	UNKO:C	3.776	
YPCP11	THR144:OG	UNKO:H	3.550	ILE10:O	UNKO:N	3.731	VAL11:CA	UNKO:C	3.901	0.578
	GLU32:OE2	UNKO:H	2.432	GLU32:OE2	UNKO:N	2.918	VAL11:C	UNKO:C	3.804	
	GLU32:OE1	UNKO:H	3.179	GLU32:OE1	UNKO:N	2.907	VAL11:CA	UNKO:C	3.869	
	VAL11:O	UNKO:H	3.458	ILE19:O	UNKO:N	3.972	VAL11:C	UNKO:C	3.979	
	ALA12:N	UNKO:H	3.270	ASN23:OD1	UNKO:N	2.800	ALA12:CA	UNKO:C	3.890	
	GLU32:OE1	UNKO:H	2.085				ALA12:CB	UNKO:C	3.452	
	GLU32:OE2	UNKO:H	2.822				LEU25:CD1	UNKO:C	3.371	
	VAL11:N	UNKO:H	3.849				VAL154:CG'	UNKO:C	3.587	
	ILE19:O	UNKO:H	3.227				VAL154:CG'	UNKO:C	3.984	
	ASN23:OD1	UNKO:H	1.837				LYS60:CE	UNKO:C	3.811	
	GLN127:NE2	UNKO:H	2.859				LYS60:CE	UNKO:C	3.637	
							LYS60:CD	UNKO:C	3.734	
							LYS60:CB	UNKO:C	3.638	
							VAL154:CG'	UNKO:C	3.584	
							VAL154:CG'	UNKO:C	3.951	
YPCP12	GLU32:OE1	UNKO:H	2.155	GLU32:OE1	UNKO:N	3.126	LEU72:CD2	UNKO:C	3.723	5-10
	GLU32:OE2	UNKO:H	2.611	GLU32:OE2	UNKO:N	3.323	ILE65:CD1	UNKO:C	3.672	
	GLU32:OE1	UNKO:H	3.720				LEU72:CD2	UNKO:C	3.533	
	GLU32:OE2	UNKO:H	3.316				PHE69:CD2	UNKO:C	3.721	
	ALA12:N	UNKO:H	3.475				PHE36:CD2	UNKO:C	3.691	
	VAL11:O	UNKO:H	3.740				ILE65:CG1	UNKO:C	3.756	
	VAL11:N	UNKO:H	3.945				ILE33:CD1	UNKO:C	3.151	
							ILE33:CD1	UNKO:C	3.599	
							ILE33:CD1	UNKO:C	3.969	
							ILE19:CB	UNKO:C	3.930	
							ALA12:CB	UNKO:C	3.872	
							VAL11:CG2	UNKO:C	3.828	
							SER24:CA	UNKO:C	3.835	
							SER24:CA	UNKO:C	3.812	
							LEU25:CD1	UNKO:C	3.863	
YPCP13	ILE19:O	UNKO:H	3.015	ILE19:O	UNKO:N	3.132	LEU25:CD1	UNKO:C	3.567	100
	ASN23:OD1	UNKO:H	2.306	ASN23:OD1	UNKO:N	2.993	ILE19:CD1	UNKO:C	3.756	
	ASN23:ND2	UNKO:H	3.679				TYR129:CE2	UNKO:C	3.677	
	ASN23:N	UNKO:H	3.804				VAL154:CG'	UNKO:C	3.728	
							VAL154:CG'	UNKO:C	3.552	
							VAL154:CG'	UNKO:C	3.652	
							ASN23:CA	UNKO:C	3.762	
							LYS60:CE	UNKO:C	3.966	
							LYS60:CD	UNKO:C	3.909	
							LYS60:CE	UNKO:C	3.724	
							LYS60:CE	UNKO:C	3.551	
							LYS60:CE	UNKO:C	3.959	
							GLY58:CA	UNKO:C	3.321	
							GLY58:C	UNKO:C	3.746	

LIGAND NO	HYDROGEN BONDING			IONIC BONDING			HYDROPHOBIC INTERACTIONS			IC_{50}
	Amino Acids		Dist	Amino Acids		Dist	Amino Acids		Dist	
YPCP19	ILE19:O	UNKO:H	2.635	LEU25:N	UNKO:O	3.081	PHE69:CE2	UNKO:C	3.929	1.0
	ILE10:O	UNKO:H	2.843	TYR129:OH	UNKO:N	3.156	PHE69:CD2	UNKO:C	3.537	
	VAL11:N	UNKO:H	3.186	ILE123:O	UNKO:N	3.830	PHE69:CD2	UNKO:C	3.619	
	ALA12:N	UNKO:H	3.171	ILE10:O	UNKO:N	3.045	PRO66:CD	UNKO:C	3.903	
	ILE10:O	UNKO:H	2.662	TYR129:OH	UNKO:N	2.862	ILE65:CG2	UNKO:C	3.384	
	LEU25:N	UNKO:H	3.945	ILE123:O	UNKO:N	3.366	ILE65:CB	UNKO:C	3.972	
							ILE65:CG1	UNKO:C	3.539	
							ILE33:CD1	UNKO:C	3.974	
							ILE33:CG2	UNKO:C	3.666	
							ILE33:CD1	UNKO:C	3.527	
							ILE33:CD1	UNKO:C	3.886	
							SER64:C	UNKO:C	3.733	
							SER64:CB	UNKO:C	3.741	
YPCP20	GLU32:OE1	UNKO:H	2.988	ALA12:O	UNKO:N	2.997	SER64:C	UNKO:C	3.790	14
	GLU32:OE2	UNKO:H	3.671	ILE19:O	UNKO:N		SER64:CB	UNKO:C	3.425	
	ALA12:N	UNKO:H	3.165			3.883	ILE65:CG1	UNKO:C	3.417	
	ALA12:O	UNKO:H	2.020				SER64:CB	UNKO:C	3.483	
	ILE19:O	UNKO:H	3.612				THR61:CG2	UNKO:C	3.392	
	TYR129:OH	UNKO:H	3.466				THR61:CG2	UNKO:C	3.644	
	ALA12:O	UNKO:H	3.601				THR61:CB	UNKO:C	3.802	
							THR61:CA	UNKO:C	3.654	
							GLY20:CA	UNKO:C	3.580	
YPCP21	THR61:OG1	UNKO:H	2.651	ALA12:N	UNKO:O	2.983	VAL11:CG2	UNKO:C	3.753	5.6
	THR61:N	UNKO:H	3.803	THR61:OG1	UNKO:N	2.977	ILE19:CB	UNKO:C	3.450	
	SER64:OG	UNKO:H	3.941				ILE19:CB	UNKO:C	3.982	
	THR61:OG1	UNKO:H	3.049				ILE19:CD1	UNKO:C	3.479	
	LEU25:N	UNKO:H	2.655				PHE36:CZ	UNKO:C	3.891	
	SER24:N	UNKO:H	3.868				PHE36:CG	UNKO:C	3.881	
	SER64:O	UNKO:H	3.202				PHE36:CE2	UNKO:C	3.767	
							PHE36:CZ	UNKO:C	3.734	
							PHE36:CE2	UNKO:C	3.916	
							PHE36:CD1	UNKO:C	3.607	
							PHE36:CB	UNKO:C	3.799	
							ALA12:CB	UNKO:C	3.776	
							ILE33:CD1	UNKO:C	3.757	
							ILE33:CD1	UNKO:C	3.823	
							ILE33:CO1	UNKO:C	3.774	
							PHE36:CD2	UNKO:C	3.726	
YPCP22	LEU25:N	UNKO:H	3.420	ALA12:N	UNKO:O	3.633	SER24:C	UNKO:C	3.893	0.25
	SER24:O	UNKO:H	2.914	SER24:O	UNKO:N	3.861	SER24:CA	UNKO:C	3.946	
	LEU25:N	UNKO:H	3.518	SER64:O	UNKO:N	3.674	LEU25:CD2	UNKO:C	3.834	
	SER24:N	UNKO:H	3.372	ANS23:O	UNKO:N	3.873	SER64:CB	UNKO:C	3.976	
	SER64:O	UNKO:H	2.843				ALA12:CB	UNKO:C	3.961	
	GLU32:OE1	UNKO:H	2.474				PHE69:CE2	UNKO:C	3.765	
	GLU32:OE2	UNKO:H	2.876				ILE65:CG1	UNKO:C	3.888	
	ALA12:N	UNKO:H	3.519				ILE65:CG1	UNKO:C	3.740	
							ILE65:CO1	UNKO:C	3.867	
							ILE65:CD1	UNKO:C	3.726	
							PHE36:CZ	UNKO:C	3.645	
							PHE36:CE2	UNKO:C	3.585	
							ILE33:CD1	UNKO:C	3.483	

LIGAND NO	HYDROGEN BONDING			IONIC BONDING			HYDROPHOBIC INTERACTIONS			IC_{50}
	Amino Acids		Dist	Amino Acids		Dist	Amino Acids		Dist	
YPCP23	GLU32:OE1	UNKO:H	3.114	GLU32:OE1	UNKO:N	3.244	PHE36:CD1	UNKO:C	3.522	0.049
	GLU32:OE2	UNKO:H	2.493	GLU32:OE2	UNKO:N	3.097	PHE36:CE1	UNKO:C	3.729	
	VAL11:O	UNKO:H	3.464	GLU32:OE1	UNKO:N	3.460	PHE36:CG	UNKO:C	3.671	
	ALA12:N	UNKO:H	2.693	ILE10:O	UNKO:N	3.573	PHE36:CD1	UNKO:C	3.675	
	GLU32:OE2	UNKO:H	3.201				ILE65:CG1	UNKO:C	3.715	
	GLU32:OE1	UNKO:H	3.450				ILE123:CD1	UNKO:C	3.906	
							ILE65:CG1	UNKO:C	3.990	
							THR61:CG1	UNKO:C	3.719	
							GLY124:CA	UNKO:C	3.707	
							THR61:CG2	UNKO:C	3.743	
							GLY125:CA	UNKO:C	3.580	
							GLY125:CA	UNKO:C	3.724	
							TYR129:CE2	UNKO:C	3.696	
							ILE19:CB	UNKO:C	3.696	
							ILE19:CD1	UNKO:C	3.851	
							ILE19:CB	UNKO:C	3.489	
							ILE19:CD1	UNKO:C	3.968	
							ILE19:CG2	UNKO:C	3.798	
							VAL11:CG2	UNKO:C	3.777	
YPCP24	GLU32:OE1	UNKO:N	2.193	GLU32:OE1	UNKO:N	3.093	ILE33:CD1	UNKO:C	3.833	0.80
	GLU32:OE2	UNKO:N	2.518	GLU32:OE2	UNKO:N	3.224	ILE33:CD1	UNKO:C	3.952	
	GLU32:O	UNKO:N	3.965	TYR129:OH	UNKO:N	3.949	SER64:CB	UNKO:C	3.905	
	GLU32:OE1	UNKO:N	3.908	ALA12:O	UNKO:N	3.645	LYS28:CA	UNKO:C	3.967	
	GLU32:OE2	UNKO:N	3.479				ILE33:CO1	UNKO:C	3.553	
	TYR129:OH	UNKO:H	3.060				ILE65:CD1	UNKO:C	3.948	
	ALA12:O	UNKO:H	3.676				ILE65:CG1	UNKO:C	3.636	
	ALA12:N	UNKO:H	3.163				ILE65:CG1	UNKO:C	3.009	
							ILE65:CD1	UNKO:C	3.683	
							THR61:CG2	UNKO:C	3.956	
							SER24:C	UNKO:C	3.949	
							LEU25:CD1	UNKO:C	3.244	
YPCP25	GLY20:N	UNKO:H	3.941	ILE19:O	UNKO:N	3.957	PHE36:CE1	UNKO:C	3.513	5.18
	ILE19:O	UNKO:H	2.909	ASN23:OD1	UNKO:N	3.564	PHE36:CZ	UNKO:C	3.590	
	ALA126:N	UNKO:H	3.496	SER64:OG	UNKO:N	2.884	PHE36:CD1	UNKO:C	3.741	
	ASN23:OD1	UNKO:H	3.401	LYS60:O	UNKO:N	3.220	PHE36:CG	UNKO:C	3.932	
	THR61:N	UNKO:H	3.435	THR61:OG1	UNKO:N	3.384	PHE36:CE2	UNKO:C	3.859	
	LYS60:O	UNKO:H	3.097				PHE36:CD2	UNKO:C	3.980	
	SER64:OG	UNKO:H	3.551				PHE36:CZ	UNKO:C	3.741	
	SER64:OG	UNKO:H	2.170				PHE36:CE2	UNKO:C	3.647	
	ANS23:O	UNKO:H	3.224				PHE36:CD2	UNKO:C	3.991	
							ILE65:CD1	UNKO:C	3.758	
							ILE65:CG1	UNKO:C	3.285	
							ILE65:CG1	UNKO:C	3.441	
							SER64:C	UNKO:C	3.802	
							SER64:CA	UNKO:C	3.858	
							SER64:CB	UNKO:C	2.943	
							THR61:CB	UNKO:C	3.954	
							THR61:CG2	UNKO:C	3.706	
							LEU25:CD1	UNKO:C	3.903	
							ALA12:CB	UNKO:CL	3.827	
							ALA12:CA	UNKO:CL	3.888	
							VAL11:CA	UNKO:CL	3.643	
							VAL11:C	UNKO:CL	3.684	
YPCP26	ASP134:O	UNKO:H	3.780	LYS96:O	UNKO:N	3.900	ILE80:CG2	UNKO:C	3.595	3.53
	ASP99:OD2	UNKO:H	3.252	ASP134:OD2	UNKO:N	3.306	ILE80:CG2	UNKO:C	3.524	
	HIS135:ND1	UNKO:H	2.710	ASP99:OD2	UNKO:N	3.435	ILE80:CG2	UNKO:C	3.837	
	ASP134:OD2	UNKO:H	2.626	ASP99:OD2	UNKO:N	3.880	ALA131:CB	UNKO:C	3.486	
	ASP134:OD2	UNKO:H	3.470				ALA131:CB	UNKO:C	3.868	
							ALA131:CA	UNKO:C	3.924	
							ALA131:CB	UNKO:C	3.555	
							ASP134:CB	UNKO:C	3.913	
							SER97:CB	UNKO:C	3.798	

LIGAND NO	HYDROGEN BONDING			IONIC BONDING			HYDROPHOBIC INTERACTIONS			IC_{50}
	Amino Acids		Dist	Amino Acids			Amino Acids		Dist	
YPCP27	SER64:O	UNKO:H	2.600	ILE123:O	UNKO:N	3.125	PHE36:CE1	UNKO:C	3.885	10.8
	ILE65:N	UNKO:H	3.649	TYR129:OH	UNKO:N	3.731	PHE36:CD1	UNKO:C	3.825	
	TYR129:OH	UNKO:H	3.359	SER64:O	UNKO:N	3.576	PHE36:CZ	UNKO:C	3.872	
	GLY124:N	UNKO:H	3.831				PHE36:CE2	UNKO:C	3.771	
	ILE123:O	UNKO:H	2.236				PHE36:CD2	UNKO:C	3.620	
							PHE36:CG	UNKO:C	3.649	
							ILE33:CD1	UNKO:C	3.672	
							ILE65:CG1	UNKO:C	3.733	
							GLY125:CA	UNKO:CL	3.742	
							ALA12:CB	UNKO:C	3.559	
							ALA12:CB	UNKO:C	3.948	
							VAL11:CA	UNKO:C	3.703	
							ILE10:C	UNKO:C	3.961	
YPCP28	ILE19:O	UNKO:H	2.617	ILE19:O	UNKO:N	3.006	PRO66:CG	UNKO:C	3.923	1.3
	ILE19:O	UNKO:H	3.443				PRO66:CG	UNKO:C	3.998	
	ANS23:O	UNKO:H	3.026				PHE36:CZ	UNKO:C	3.850	
	SER64:OG	UNKO:H	3.648				ILE10:C	UNKO:C	3.964	
	SER24:N	UNKO:H	2.874				VAL11:CA	UNKO:CL	3.921	
	SER24:O	UNKO:H	3.872				VAL11:CA	UNKO:C	3.999	
							SER64:CB	UNKO:C	3.478	
							THR61:CG2	UNKO:C	3.651	
							THR61:CG2	UNKO:C	3.617	
YPCP29	TYR129:OH	UNKO:H	3.141	TYR129:OH	UNKO:N	3.995	ILE33:CD1	UNKO:C	3.860	0.12
	THR61:OG1	UNKO:H	3.061	THR61:OG1	UNKO:N	3.076	PHE36:CE1	UNKO:C	3.889	
	LYS60:O	UNKO:H	3.956	THR61:OG1	UNKO:N	3.277	PHE36:CZ	UNKO:C	3.779	
	THR61:N	UNKO:H	3.663	LYS60:O	UNKO:N	3.925	PHE36:CE2	UNKO:C	3.756	
	SER64:OG	UNKO:H	2.598	SER64:OG	UNKO:N	3.543	PHE36:CD2	UNKO:C	3.789	
	ASN23:O	UNKO:H	3.768	THR61:OG1	UNKO:N	3.731	PHE36:CG	UNKO:C	3.916	
	ILE123:O	UNKO:H	3.232				PHE36:CE2	UNKO:C	3.633	
							PHE36:CD2	UNKO:C	3.911	
							LEU72:CD2	UNKO:C	3.911	
							PRO66:CD	UNKO:C	3.779	
							ILE65:CG2	UNKO:C	3.731	
							ILE65:CG1	UNKO:C	3.741	
							THR61:CB	UNKO:C	3.824	
							THR61:CG2	UNKO:C	3.964	
							THR61:CA	UNKO:C	3.887	
							THR61:CG2	UNKO:C	3.759	
YPCP30	SER64:O	UNKO:H	3.162	SER64:O	UNKO:N	3.702	ILE65:CG1	UNKO:C	3.710	1.07
	ILE65:N	UNKO:H	3.972				ILE19:CD1	UNKO:C	3.769	
	SER64:O	UNKO:H	3.933				ALA12:CB	UNKO:C	3.997	
							ALA12:CB	UNKO:C	3.765	
							GLY125:CA	UNKO:C	3.650	
YPCP30	SER64:O	UNKO:H	3.162	SER64:O	UNKO:N	3.702	ILE65:CG1	UNKO:C	3.710	1.07
	ILE65:N	UNKO:H	3.972				ILE19:CD1	UNKO:C	3.769	
	SER64:O	UNKO:H	3.933				ALA12:CB	UNKO:C	3.997	
							ALA12:CB	UNKO:C	3.765	
							GLY125:CA	UNKO:C	3.650	
YPCP31	ILE123:O	UNKO:H	3.095	ALA12:O	UNKO:N	2.984	ILE19:CB	UNKO:C	3.808	110
	ILE10:O	UNKO:H	3.591	ILE10:O	UNKO:N	3.666	ILE19:CD1	UNKO:C	3.729	
	ILE10:O	UNKO:H	3.274	ALA126:N	UNKO:O	2.874	TYR129:CE2	UNKO:C	3.795	
	VAL11:N	UNKO:H	3.366	TYR129:OH	UNKO:N	3.217	ILE19:CD1	UNKO:C	3.577	
	ALA12:N	UNKO:H	2.924	THR61:OG1	UNKO:N	2.903	GLY58:CA	UNKO:C	3.821	
	ALA12:N	UNKO:H	2.197	THR61:N	UNKO:O	3.958	LYS60:CB	UNKO:C	3.413	
	ALA12:O	UNKO:H	2.713				LYS60:CA	UNKO:C	3.846	
	TYR129:OH	UNKO:H	3.228				LYS60:CB	UNKO:C	3.661	
	TYR129:OH	UNKO:H	3.486				LYS60:CB	UNKO:C	3.655	
	THR61:OG1	UNKO:H	2.385				VAL154:CG'	UNKO:C	3.638	
	GLY125:N	UNKO:H	3.162				VAL154:CG2	UNKO:C	3.874	
	LYS60:N	UNKO:H	2.422				VAL154:CG'	UNKO:C	3.871	
	ARG59:N	UNKO:H	3.442				ASN23:CA	UNKO:C	3.568	
	GLN127:NE2	UNKO:H	3.998				ASN23:C	UNKO:C	3.974	
	LEU128:N	UNKO:H	3.170							
	GLN127:N	UNKO:H	3.803							

UGAND NO	HYDROGEN BONDING			IONIC BONDING			HYDROPHOBIC INTERACTIONS			IC_{50}
	Amino Acids		Dist	Amino Acids		Dist	Amino Acids		Dist	
YPCP32	ALA126:N	UNKO:H	3.422	No ionic bonds			ILE33:CD1	UNKO:C	3.767	5.3
	GLN127:NE2	UNKO:H	3.539				ILE33:CD1	UNKO:C	3.875	
	LYS60:NZ	UNKO:H	2.773				LEU25:CD2	UNKO:C	3.433	
							LEU25:CD2	UNKO:C	3.993	
							THR61:CG2	UNKO:C	3.752	
							ILE19:CD1	UNKO:C	3.957	
							LYS60:CE	UNKO:C	3.887	
							VAL154:CG'	UNKO:C	3.607	
							VAL154:CG2	UNKO:C	3.972	
							VAL154:CG2	UNKO:C	3.466	
YPCP33	GLU32:OE1	UNKO:H	3.542	ILE10:O	UNKO:N	2.985	ILE33:CD1	UNKO:C	3.594	1.2
	ILE10:O	UNKO:H	3.949	TYR129:OH	UNKO:N	3.041	PHE36:CE2	UNKO:C	3.868	
	ALA12:N	UNKO:H	2.606	ILE19:O	UNKO:N	3.326	PHE36:CD2	UNKO:C	3.751	
	VAL11:N	UNKO:H	3.569	ILE123:D	UNKO:N	3.027	PHE36:CD1	UNKO:C	3.651	
	VAL11:O	UNKO:H	3.709	LEU25:N	UNKO:O	3.201	PHE36:CE1	UNKO:C	3.825	
	ILE123:O	UNKO:H	2.082				PHE36:CG	UNKO:C	3.792	
	TYR129:OH	UNKO:H	3.412							
	SER64:O	UNKO:H	3.524							
	SER64:OG	UNKO:H	3.947							
	ILE65:N	UNKO:H	3.775							
YPCP34	LYS60:O	UNKO:H	2.600	LYS60:O	UNKO:N	3.244	ILE33:CD1	UNKO:C	3.209	0.002
	THR61:N	UNKO:H	3.705				ILE33:CD1	UNKO:C	3.438	
							ILE33:CG2	UNKO:C	3.832	
							PHE36:CE2	UNKO:C	3.795	
							PHE36:CD2	UNKO:C	3.629	
							ILE33:CD1	UNKO:C	3.603	
							PHE36:CE1	UNKO:C	3.815	
							PHE36:CD1	UNKO:C	3.674	
							ILE65:CG2	UNKO:C	3.990	
							ILE65:CG1	UNKO:C	3.770	
							ILE65:CG1	UNKO:C	3.279	
							ILE65:CD1	UNKO:C	3.765	
							SER64:CB	UNKO:C	3.795	
							THR61:CG2	UNKO:C	3.558	
							THR61:CG2	UNKO:C	3.762	
YPCP35	GLU32:OE1	UNKO:H	2.394	GLU32:OE1	UNKO:N	3.489	VAL11:CG2	UNKO:C	3.376	2.8
	GLU32:OE2	UNKO:H	2.009	GLU32:OE1	UNKO:N	3.272	VAL11:C8	UNKO:C	3.980	
	GLU32:OE2	UNKO:H	3.452	GLU32:OE2	UNKO:N	3.005	VAL11:CA	UNKO:C	3.527	
	VAL11:N	UNKO:H	3.459	VAL11:O	UNKO:N	3.971	SER64:CB	UNKO:C	2.896	
	ALA12:N	UNKO:H	3.411	THR61:OG1	UNKO:N	3.137	SER64:CA	UNKO:CL	3.992	
	VAL11:O	UNKO:H	3.302	THR61:OG1	UNKO:N	3.243	ILE64:CG1	UNKO:C	3.822	
	THR144:OG	UNKO:H	3.829	THR61:OG1	UNKO:N	3.649	LEU25:CD1	UNKO:C	3.659	
							THR61:CA	UNKO:CL	3.700	
YPCP36	GLN127:NE2	UNKO:H	3.167	ALA12:N	UNKO:O	3.921	ILE33:CD1	UNKO:C	3.872	0.000
	GLN127:N	UNKO:H	2.686	THR61:OG1	UNKO:N	3.084	LEU25:CD2	UNKO:C	3.838	035
	GLN127:NE2	UNKO:H	3.887	THR61:OG1	UNKO:N	3.116	LEU25:CD1	UNKO:C	3.719	
	ALA12:O	UNKO:H	3.436				VAL11:CG2	UNKO:C	3.909	
	TRP27:NE1	UNKO:H	3.998				VAL11:CA	UNKO:C	3.871	
	GLU32:OE1	UNKO:H	3.563				GLY20:CA	UNKO:C	3.854	
	ARG59:N	UNKO:H	3.913				GLY125:CA	UNKO:C	3.644	
	LYS60:N	UNKO:H	3.497				LYS60:CB	UNKO:C	3.992	
	THR61:N	UNKO:H	3.553				GLY125:CA	UNKO:C	3.894	
	THR61:OG1	UNKO:H	3.803				SER24:C	UNKO:C	3.642	
	THR61:OG1	UNKO:H	2.075				SER64:CB	UNKO:C	3.560	
	THR61:N	UNKO:H	2.729				SER24:CA	UNKO:C	3.988	
	GLY125:N	UNKO:H	3.127							
	GLY125:N	UNKO:H	3.925							
	ILE123:O	UNKO:H	3.583							
	ILE19:O	UNKO:H	3.385							

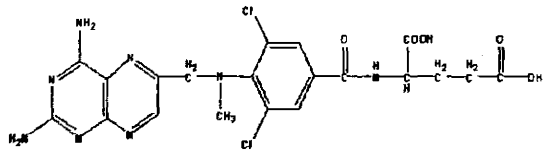
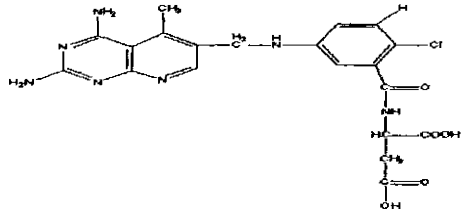
LIGAND NO	HYDROGEN BONDING			IONIC BONDING			HYDROPHOBIC INTERACTIONS			IC_{50}
	Amino Acids		Dist	Amino Acids		Dist	Amino Acids		Dist	
YPCP41	ASN23:OD1	UNKO:H	1.925	ASN23:OD1	UNKO:N	3.401	VAL154:CG'	UNKO:C	3.559	4.4
	ASN23:ND2	UNKO:H	3.621	ASN23:OD1	UNKO:N	2.800	GLY125:CA	UNKO:C	3.380	
	LYS60:NZ	UNKO:H	3.657	SER64:OG	UNKO:N	3.311	GLY125:C	UNKO:C	3.719	
	LYS60:NZ	UNKO:H	3.632	THR61:OG1	UNKO:N	3.167	LEU25:CD1	UNKO:C	3.914	
	ASN23:OD1	UNKO:H	3.530	LYS60:O	UNKO:N	3.808	LEU25:CD1	UNKO:C	3.667	
	THR61:OG1	UNKO:H	3.474				LEU25:CD1	UNKO:C	3.796	
	SER64:OG	UNKO:H	2.404				VAL154:CG'	UNKO:C	3.968	
	THR61:O	UNKO:H	3.878				VAL154:CB	UNKO:C	3.787	
	LYS60:O	UNKO:H	2.840				VAL154:CG'	UNKO:C	3.760	
	THR61:N	UNKO:H	3.301				GLY20:CA	UNKO:C	3.318	
YPCP42	ASN23:OD1	UNKO:H	3.811	ILE19:O	UNKO:N	3.650	ILE65:CO1	UNKO:CL	3.471	0.84
	ASN23:O	UNKO:H	3.087	ILE19:O	UNKO:N	3.022	ILE65:CG1	UNKO:CL	3.936	
	SER24:N	UNKO:H	3.015				PHE36:CZ	UNKO:CL	3.901	
	SER64:OG	UNKO:H	3.581				PHE36:CE2	UNKO:CL	3.708	
	ILE19:O	UNKO:H	2.993				PHE36:CE1	UNKO:C	3.389	
	ILE19:O	UNKO:H	3.340				PHE36:CZ	UNKO:C	3.511	
	ARG21:O	UNKO:H	3.976				PHE36:CD1	UNKO:C	3.825	
	ASN23:N	UNKO:H	2.978				SER64:CB	UNKO:C	3.525	
	ARG21:N	UNKO:H	3.508				THR61:CG2	UNKO:C	3.786	
	SER24:O	UNKO:H	3.341				THR61:CB	UNKO:C	3.819	
	SER24:N	UNKO:H	2.996				LEU25:CD1	UNKO:C	3.798	
YPCP43	GLU32:OE1	UNKO:H	2.467	GLU32:OE1	UNKO:N	3.283	PHE36:CD1	UNKO:C	3.734	1.0
	GLU32:OE2	UNKO:H	1.938	GLU32:OE2	UNKO:N	2.945	ILE19:CB	UNKO:C	3.566	
	GLU32:OE2	UNKO:H	3.534	GLU32:OE1	UNKO:N	3.336	LEU25:CD1	UNKO:C	3.790	
	THR144:OG	UNKO:H	3.974	THR144:OH	UNKO:N	3.820	LEU25:CD1	UNKO:C	3.620	
				ALA12:O	UNKO:N	3.873	THR61:CG2	UNKO:C	3.887	
YPCP44	GLU32:OE1	UNKO:H	2.823	ILE123:O	UNKO:N	3.368	PHE36:CD1	UNKO:C	3.955	0.41
	GLU32:OE2	UNKO:H	3.113	TYR129:OH	UNKO:N	3.853	PHE36:CE1	UNKO:C	3.800	
	GLU32:OE1	UNKO:H	2.216	TYR129:OH	UNKO:N	2.909	PHE36:CZ	UNKO:C	3.754	
	GLU32:OE2	UNKO:H	3.028	GLU32:OE1	UNKO:N	2.809	PHE36:CD1	UNKO:C	3.893	
	TYR129:OH	UNKO:H	1.962	GLU32:OE2	UNKO:N	3.438	PHE36:CG	UNKO:C	3.842	
	ILE19:O	UNKO:H	3.999				PHE36:CD2	UNKO:C	3.778	
	TYR129:OH	UNKO:H	3.730				PHE36:CE2	UNKO:C	3.773	
	ALA12:O	UNKO:H	3.309				PHE36:CD1	UNKO:C	3.631	
	ALA12:N	UNKO:H	3.637				PHE36:CG	UNKO:C	3.723	
	ALA12:N	UNKO:H	3.938				PHE36:CE1	UNKO:C	3.901	
YPCP45	ILE19:O	UNKO:H	2.835	ILE123:O	UNKO:N	2.966	PHE36:CE1	UNKO:C	3.760	0.42
	ALA12:O	UNKO:H	2.272	TYR129:OH	UNKO:N	3.735	PHE36:CD1	UNKO:C	3.686	
	ALA12:N	UNKO:H	2.680	ILE19:O	UNKO:N	3.760	PHE36:CG	UNKO:C	3.800	
	ILE123:O	UNKO:H	3.774	ALA12:O	UNKO:N	2.957	ILE33:CD1	UNKO:C	3.738	
	THR61:OG1	UNKO:H	3.959	ALA12:O	UNKO:N	3.928	ILE123:CD1	UNKO:CL	3.895	
	ILE123:O	UNKO:H	2.179				PHE36:CD1	UNKO:C	3.731	
							PHE36:CG	UNKO:C	3.753	
							PHE36:CB	UNKO:C	3.996	
YPCP46	ARG82:N	UNKO:H	3.467	ILE80:O	UNKO:N	3.891	LYS60:CD	UNKO:C	3.979	0.52
	THR81:OG1	UNKO:H	3.766	ILE80:O	UNKO:N	3.692	LYS60:CG	UNKO:C	3.793	
	ARG59:NH2	UNKO:H	3.737	THR81:OG1	UNKO:N	3.160	LYS60:CG	UNKO:C	3.606	
	ARG59:NE	UNKO:H	3.249				LYS60:CD	UNKO:C	3.947	
							GLN127:CB	UNKO:C	3.878	

DRUGS	HYDROGEN BONDING			IONIC BONDING			HYDROPHOBIC INTERACTIONS			IC_{50}
	Amino Acids		Dist	Amino Acids		Dist	Amino Acids		Dist	
Trimeth-oprim	GLU32:OE1	UNKO:H	3.093	ILE10:O	UNKO:N	2.884	GLY125:CA	UNKO:C	3.993	12
	GLU32:OE2	UNKO:H	3.913	TYR129:OH	UNKO:N	2.759	ILE19:CD1	UNKO:C	3.894	
	GLU32:OE1	UNKO:H	2.532	ILE123:O	UNKO:N	2.916	LEU25:CD1	UNKO:C	3.757	
	ILE10:O	UNKO:H	2.595	GLU32:OE1	UNKO:N	3.564	LEU25:CD1	UNKO:C	3.99	
	VAL11:N	UNKO:H	3.786	GLU32:OE1	UNKO:N	2.800	ILE19:CB	UNKO:C	3.99	
	ILE123:O	UNKO:H	2.365							
	TYR129:OH	UNKO:H	3.172							
Atova-quone	ILE19:O	UNKO:H	3.112	ALA126:N	UNKO:O	3.061	ILE33:CD1	UNKO:C	3.954	2.5
				SER24:N	UNKO:O	3.999	ILE33:CD1	UNKO:C	3.686	
							PHE69:CD2	UNKO:C	3.807	
							PHE69:CD2	UNKO:C	3.040	
							PHE36:CD2	UNKO:C	3.274	
							LEU25:CO2	UNKO:C	3.901	
							LEU25:CD2	UNKO:C	3.883	
							LEU25:CD2	UNKO:C	3.590	
							LEU25:CD1	UNKO:C	3.527	
							ILE65:CD1	UNKO:C	3.844	
Sulpha-methox azole	ASN23:O	UNKO:H	2.977	SER64:OG	UNKO:N	2.804	PHE36:CD1	UNKO:C	3.999	0.104
	SER64:OG	UNKO:H	2.352	LYS60:O	UNKO:N	3.251	SER64:CB	UNKO:C	3.572	
	SER64:OG	UNKO:H	3.185	ASN23:O	UNKO:N	3.977	THR61:CG2	UNKO:C	3.999	
	THR61:N	UNKO:H	3.452	GLY125:N	UNKO:O	3.842				
	ASN23:OD1	UNKO:H	2.966	ALA12:N	UNKO:O	3.112				
	ASN23:OD1	UNKO:H	3.155	ASM23:OD1	UNKO:N	3.289				
	LYS60:O	UNKO:H	2.775	TYR129:OH	UNKO:N	3.015				
	LYS60:O	UNKO:H	3.591							
Prima-quine	TYR129:OH	UNKO:H	2.849	THR61:OG1	UNKO:N	2.980	ILE65:CD1	UNKO:C	3.946	4.2
	ILE123:O	UNKO:H	2.278	TYR129:OH	UNKO:N	3.235	ILE65:CD1	UNKO:C	3.858	
	THR61:OG1	UNKO:H	2.094	ILE123:O	UNKO:N	3.353	PHE36:CE1	UNKO:C	3.734	
	THR61:N	UNKO:H	3.963	ALA12:O	UNKO:N	3.184	PHE36:CD1	UNKO:C	3.935	
	THR61:OG1	UNKO:H	3.675	ILE123:O	UNKO:N	3.116	PHE36:CD1	UNKO:C	3.992	
							PHE36:CD1	UNKO:C	3.572	
							PHE36:CE1	UNKO:C	3.942	
							GLU32:CD	UNKO:C	3.960	
							ILE10:CG1	UNKO:C	3.537	
							PHE36:CZ	UNKO:C	3.692	
							PHE36:CE2	UNKO:C	3.823	
							PHE36:CD2	UNKO:C	3.932	
							PHE36:CG	UNKO:C	3.991	
							TYR129:CE'	UNKO:C	3.771	
							ILE19:CG2	UNKO:C	3.864	
							VAL11:CG2	UNKO:C	3.813	
							VAL11:CA	UNKO:C	3.875	

It must be noted that all the interactions were taken with the side chain functional groups of the target protein amino acids as rest are bonded by means of a peptide bond therefore none of the interaction is taken with those functional groups. The most active compounds possessing strong binding affinity were namely YPCP3, YPCP10, YPCP11, YPCP31, YPCP36, YPCP37 and YPCP42. The interactions of these compounds were as follow: **YPCP3** had a total of about 7 ionic interactions, 11 hydrogen bonding and 11 hydrophobic interactions. **YPCP10** had 11 hydrogen bonds, 5 ionic and 14 hydrophobic respectively. **YPCP11** had 5 ionic, 11 hydrogen and 15 hydrophobic interactions. **YPCP31** had 6 ionic, 16 hydrogen and 14 hydrophobic interactions. **YPCP36** had 16 hydrogen bonds, 3 ionic and 12 hydrophobic interactions. **YPCP37** had 19 hydrogen bonds, 7 ionic and no hydrophobic interaction. Last but not the least compound **YPCP42** had a total of 11 hydrogen bonds, 2 ionic along with 11 hydrophobic interactions. But the binding affinity of the compounds YPCP1, YPCP6, YPCP9, YPCP11, YPCP25, YPCP28, YPCP30, YPCP31, YPCP33, YPCP37 was least in the entire data set i.e. -10.5, -11.5, -10.5, -11.4, -10.4, -10.4, -10.5, -11.3, -11.0 and -9.3 Kcal/mol respectively.

By bringing into consideration all these factors the hits were reduced to three i.e. YPCP31, YPCP36 and YPCP37. Although IC_{50} value has 30% role in identifying the lead but when the IC_{50} value of YPCP31, YPCP36 and YPCP37 were compared there was a remarkable difference because IC_{50} value of YPCP31 was 110 μM, however YPCP36 had an IC_{50} value of 0.000035 and that of YPCP37 is 0.0054 μM. The analysis revealed that YPCP36 had a lowest IC_{50}, but YPCP37 had the lowest binding affinity and a greater number of interactions as compared to YPCP36. All these observations led to the decision that YPCP37 was the lead having a binding affinity of

Table 4.9: Table showing the two hits for lead identification along with their energy values, IC50 and the number of binding interactions.

COMP NO	STRUCTURE	ENERGY VALUE	IC_{50} µM	H BOND	IONIC BOND	HP
YPCP36	 2-(3,5-dichloro-4-(((2,4-diaminopteridin-6-yl)methyl) (methyl) amino)benzamido)pentanedioic acid	-8.9	0.000 035	16	3	12
YPCP37	 2-(2-chloro-5-((2,4-diamino-5-methylpyrido[2,3-d]prrimidine-6-yl)methylamino)benzamido)succinic acid	-9.3	0.005 4	19	7	0

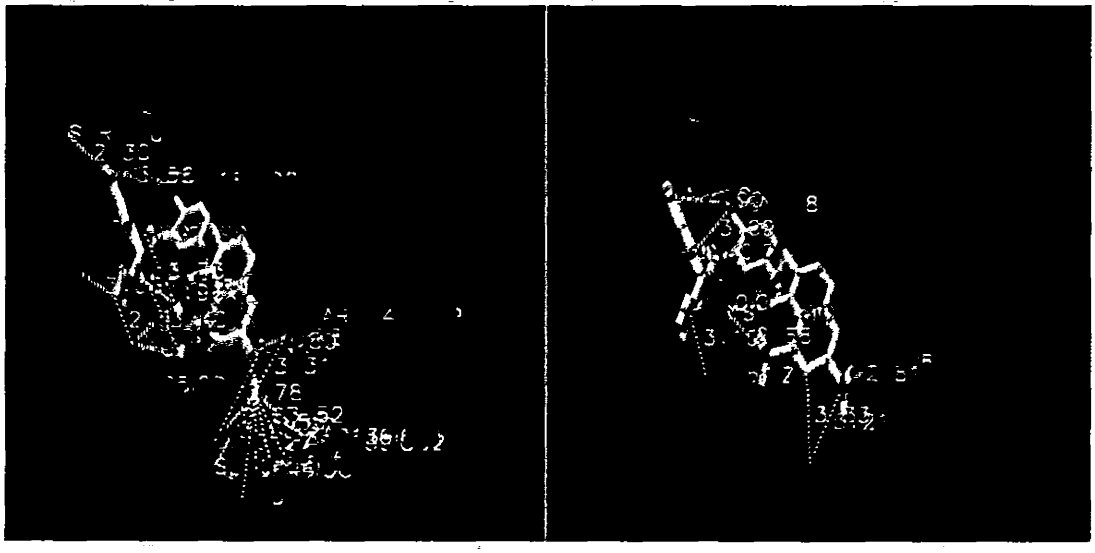


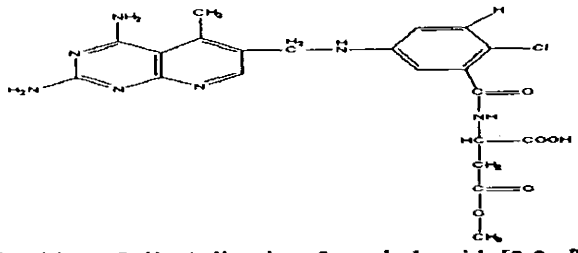
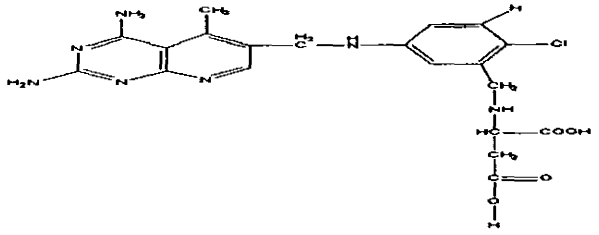
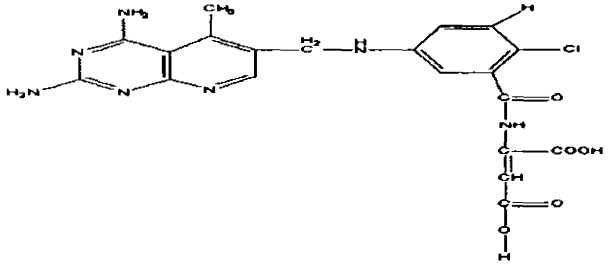
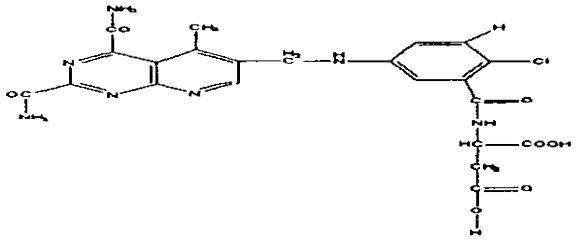
Figure 4.13a& b: (a) Binding interactions of YPCP37 the potential lead compound showing 19hydrogen bonds **(b)** Binding interactions of YPCP37 the potential lead compound showing 7 ionic interactions

4.7 Analogues of the Lead compound

On the basis of binding interactions, binding affinity and IC_{50} value the compound YPCP37 had been selected as the potential lead. Using the lead analogues were formulated correspondingly in order to obtain the most active compound to be used as a pcDHFR inhibitor. Table 4.10 shows the analogues of the lead compound with their IUPAC names obtained from ChemDraw. Analogues were designed by changing the functional groups which either increased the hydrophobicity or hydrophilicity of the analogue formulated thus reworking the efficacy of the compound.

The first analogue was designed by carrying out the process of ester formation, by replacing H with the methyl group resulting in an increase in hydrophobicity by conversion of COOH at the side chain of ring to $COOCH_3$. A second analog was prepared by performing reduction. The O of the carbonyl group was replaced by the hydrogen at the side chain attached to the ring structure. To obtain the third analog dehydrogenation was applied by adding double bond at C of the side chain of the ring structure. The fourth analogue was prepared by amide formation. The two amines on the first ring structure were converted to amide and owing to the greater electronegativity of oxygen, the carbonyl (C=O) act as a stronger dipole as compared to the N-C dipole. The presence of a C=O and to a lesser extent presence of N-C allows amides to act as H-bond acceptors. In amides the presence of N-H allows amides to function as H-bond donors as well. Thus amides can participate in hydrogen bonding since the N-H hydrogen atoms can donate H-bonds whereas the oxygen atom can accept hydrogen bonds. Thus amide formation would increase the hydrogen bonding character. The pharmacophore features were also computed for all the four analogues using LigandScout and the results are shown in Table 4.11.

Table 4.10: Analogues formed from lead compound along with their IUPAC names

Compound	Structure	Energy Value
Ester formation	 2-(2-chloro-5-((2,4-diamino-5-methylpyrido[2,3-*d*] prrimidin-6-yl)methylamino)benzamido)-4-methoxy-4-oxobutanoic acid	-8.5
Reduction	 2-(2-chloro-5-((2,4-diamino-5-methylpyrido[2,3-*d*] prrimidin-6-yl) methylamino) phenyl)) methylamino) succinic acid	-9.5
Dehydrogenation	 2-(2-chloro-5-((2,4-diamino-5-methylpyrido[2,3-*d*] prrimidin-6-yl) methylamino)benzamido)but-2-enedioic acid	-8.4
Amide formation	 2-(2-chloro-5-((2,4-dicarbamoyl-5-methylpyrido[2,3-*d*] prrimidin-6-yl) methylamino) benzamido) succinic acid	-10.5

4.7.1 Docking and Interactions of Analogues with the Target Protein

All the analogues were docked within the active site of pcDHFR with the already mentioned approach. The best conformation was selected and visualized using the VMD software in order to calculate binding interactions for each of the analogues. It is also revealed that the 1st , 2nd and 4th analogue was docked at the same position where the rest of the compounds in the dataset were docked i.e. same active site position including amino acids Ile10, Ile19, Ile33, Ile65, Ile123, Ser24, Ser64, Thr61, Leu25, Phe36, Glu32, Ala12, Val1, Val54, Gly124, Gly125, Tyr129 and Asn23.

The docking studies revealed that the first analogue had a binding energy of -8.5 Kcal/mol and it was involved in a lesser number of interactions as compared to the lead YPCP37 with a total of 3 hydrogen bonds, 1 ionic bond but a greater number of about 15 hydrophobic interactions that were absent in the lead but it is also evident from the docking analysis that analogues 1 has a binding energy greater than the lead. Hydrogen bonds were formed between the H of the ligand and O of Ser24, Leu25 with distances 3.858Å, 3.735Å and one was formed with N of Leu25 with distance 2.642Å. One ionic bond was observed between O of the ligand and N of Leu25 of the target. Moreover 15 hydrophobic interactions were observed between the C of ligand and C of amino acids Ile33, Ile65, Ser64, Pro66, Phe36 and Leu256.

Second analogue designed by reduction showed 15 hydrogen bonds formed between the H of ligand and N of Gln127, Ala12, Ala126, Leu25 and Gly125 and O of amino acids Thr61, Tyr129, Ala12 and Glu32. 11 ionic interactions were observed out of which 5 formed between the N of ligand and O of Thr61, Glu32, Tyr129 and Arg59 and the remaining 6 formed between O of ligand and N of Lys60, Gln127, and

Ala126. 9 hydrophobic interactions were observed between the C of ligand and C of Leu25, Lys60, Gly125 and Val154 respectively. It possesses a binding energy of -9.5Kcal/mol which is lower than that of the lead.

However the third analogue designed by dehydrogenation process was found to bind to relatively different set of residues including Leu88, Trp62, Leu86, Glu63, Ser85, Asn83, Glu84 and Asp87. It showed 11 hydrogen bonds formed between H of ligand and O of Leu86, Leu88, Ser85, Asn83, Glu63, Glu84 and Asp87 and N of amino acid Trp62 of the target. 5 ionic interactions were observed between O of ligand and N of residues Leu86, Leu88, Asn83 and Glu84. However the 6 hydrophobic interactions were observed between C of ligand and C of Leu88. Third analogue has a binding affinity of -8.4 Kcal/mol, which is greater in comparison to that of the lead.

The fourth analogue produced through amide formation yielded 13 hydrogen bonds that were observed between the H of ligand and O of Lys60, Ile80, Ile10, Ile123, Thr61, Ser64 and Tyr129 and N of Arg59, Gln127, Leu128 and Lys60. 15 ionic bonds were formed between O of ligand and N of Gln127, Ala126, Leu128, Thr61, Lys60, Arg59, Gly125 and Gly127. The remaining ionic interactions were found between N of ligand and O of residues Ile19, Ile123, Thr61 and Ser64. However 10 hydrophobic interactions were found between C of ligand and C of Ser64, Gly125, Gly58, Gln127, Phe36, Ile19 and Val154. It has a binding energy of about -10.5 Kcal/mol. The binding interactions of all the four analogues along with the residues and distances are shown in Table 4.12.

The binding energies of all the four analogues were far better than the lead and the analogues 2 and 4 have lower binding affinities in comparison to the lead that is a plus

point for the further investigations of these two analogues in terms of drug designing. In terms of activity and binding interactions it can be clearly noticed that that analogues produced by reduction and amide formation show a greater number of interactions and along with this the possess lower binding energies even lower than the lead. Thus on the basis of the active binding interactions of the above mentioned analogues it can be concluded that they have a potential to be tested, to undergo trail and further investigation to achieve effective anti *P.carinii* drugs that specifically target the pcDHFR. All three types of binding interactions of the two analogues produced by reduction and amide formation are represented in Figure 4.14(a-c) and 4.15(a-c). Figure 4.14(a-c) Hydrogen, ionic and hydrophobic interactions of the Analogue 2 with the target pcDHFR and Figure 4.15(a-c) Hydrogen, ionic and hydrophobic interactions of the Analogue 4 with the target pcDHFR.

Table 4.11: Pharmacophoric Features of the four analogues designed from the lead compound using LigandScout.

Compounds	**HBDs**	**HBAs**	**Ar**	**HP**	**Positive ionizable**	**Negative ionizable**
ESTER FORMATION	Five	Eight	Three	Three	None	One
REDUCTION	Four	Seven	Three	Three	Two	Two
DEHYDROGENATION	Six	Eight	Three	Two	None	One
AMIDE FORMATION	Five	Nine	Three	Two	Two	None

Table 4.12: Binding interactions of the analogues which include hydrophobic, hydrogen bonding and ionic interactions along with distances in Angstrom

Ana no	Hydrogen Bonding	dist	Ionic Bonding	dist	Hydrophobic	dist
1.	LEU25:N UNKO:H	2.642	LEU25:N UNKO:O	3.924	ILE33:CG2 UNKO:C	3.624
	LEU25:O UNKO:H	3.858			ILE33:CB UNKO:C	3.736
	SER24:OG UNKO:H	3.735			ILE33:CG1 UNKO:C	3.620
					SER64:C UNKO:C	3.773
					SER64:CB UNKO:C	3.826
					ILE65:CG1 UNKO:C	3.301
					PRO66:CG UNKO:C	3.950
					PRO66:CB UNKO:C	3.999
					PHE36:CE2 UNKO:C	3.759
					PHE36:CE2 UNKO:C	3.999
					PHE36:CE2 UNKO:C	3.751
					PHE36:CZ UNKO:C	3.881
					LEU256:CD2 UNKO:C	3.775
					LEU256:CD2 UNKO:C	3.999
					LEU256:CD2 UNKO:C	3.817
2.	GLN127:N UNKO:H	3.229	THR61:OG1 UNKO:N	3.240	LEU25:CD1 UNKO:C	3.446
	GLN127:N UNKO:H	3.377	GLU32:OE1 UNKO:N	2.975	LEU25:CD2 UNKO:C	3.962
	ALA126:N UNKO:H	3.703	GLU32:OE2 UNKO:N	3.324	LYS60:CD UNKO:C	3.753
	ALA126:N UNKO:H	3.663	TYR129:OH UNKO:N	3.006	LYS60:CG UNKO:C	3.985
	THR61:OG1 UNKO:H	2.908	LYS60:N UNKO:O	3.276	LYS60:CB UNKO:C	3.469
	TYR129:OH UNKO:H	3.426	ARG59:N UNKO:N	3.855	LYS60:CB UNKO:C	3.909
	TYR129:OH UNKO:H	2.474	GLN127:N UNKO:O	3.952	GLY125:C UNKO:C	3.847
	ALA12:O UNKO:H	3.918	ALA126:N UNKO:O	3.930	GLY125:CA UNKO:C	3.720
	ALA12:N UNKO:H	3.257	ALA126:N UNKO:O	3.928	VAL154:CG2 UNKO:C	3.446
	GLU32:OE2 UNKO:H	3.333	ALA126:N UNKO:O	3.073		
	GLU32:OE1 UNKO:H	3.559	GLN127:N UNKO:O	2.924		
	GLU32:OE1 UNKO:H	1.956				
	GLU32:OE2 UNKO:H	2.613				
	LEU25:N UNKO:H	3.886				
	GLY125:N UNKO:H	3.779				
3.	LEU88:O UNKO:H	1.804	LEU86:O UNKO:N	3.611	LEU88:CB UNKO:C	3.760
	LEU88:O UNKO:H	3.680	LEU88:O UNKO:N	2.778	LEU88:CG UNKO:C	3.987
	TRP62:NE1 UNKO:H	2.592	LEU88:O UNKO:N	3.734	LEU88:CG UNKO:C	3.989
	TRP62:NE1 UNKO:H	3.446	GLU84:O UNKO:N	3.945	LEU88:CG UNKO:C	3.732
	LEU86:O UNKO:H	3.373	ASN83:O UNKO:N	3.674	LEU88:CB UNKO:C	3.754
	LEU86:O UNKO:H	3.719			LEU86:C UNKO:C	3.959
	GLU63:OE2 UNKO:H	2.860				
	SER85:OG UNKO:H	3.695				
	ASN83:O UNKO:H	3.097				
	GLU84:O UNKO:H	2.969				
	ASP87:OD1 UNKO:H	3.144				
4.	LYS60:O UNKO:H	3.320	GLN127:N UNKO:O	3.103	SER64:CB UNKO:C	3.995
	ILE123:O UNKO:H	3.408	ALA126:N UNKO:O	2.915	GLY125:CA UNKO:C	3.894
	ILE80:O UNKO:H	3.186	GLN127:NE2 UNKO:O	3.951	GLY58:CA UNKO:C	3.575
	ARG59:N UNKO:H	2.916	LEU128:N UNKO:O	3.696	GLN127:CA UNKO:CL	3.809
	GLN127:N UNKO:H	3.736	THR61:N UNKO:O	2.906	PHE36:CE1 UNKO:C	3.904
	LEU128:N UNKO:H	3.115	LYS60:N UNKO:O	3.310	ILE19:CB UNKO:C	3.780
	LYS60:NZ UNKO:H	3.853	LYS60:N UNKO:O	3.775	ILE19:CD1 UNKO:C	3.909
	THR61:OG1 UNKO:H	2.493	ARG59:N UNKO:O	2.993	VAL154:CG2 UNKO:C	3.666
	SER64:OG UNKO:H	3.469	ILE19:O UNKO:N	3.247	VAL154:CG2 UNKO:C	3.984
	LYS60:O UNKO:H	3.026	THR61:OG1 UNKO:N	3.154	VAL154:CG2 UNKO:C	3.791
	ILE10:O UNKO:H	3.320	ILE123:O UNKO:N	3.013		
	ILE10:O UNKO:H	2.434	ILE123:O UNKO:N	3.685		
	TYR129:OH UNKO:H	3.201	SER64:OG UNKO:N	3.712		
		2.764	GLY127:N UNKO:O	3.891		
			GLY125:N UNKO:O	2.882		

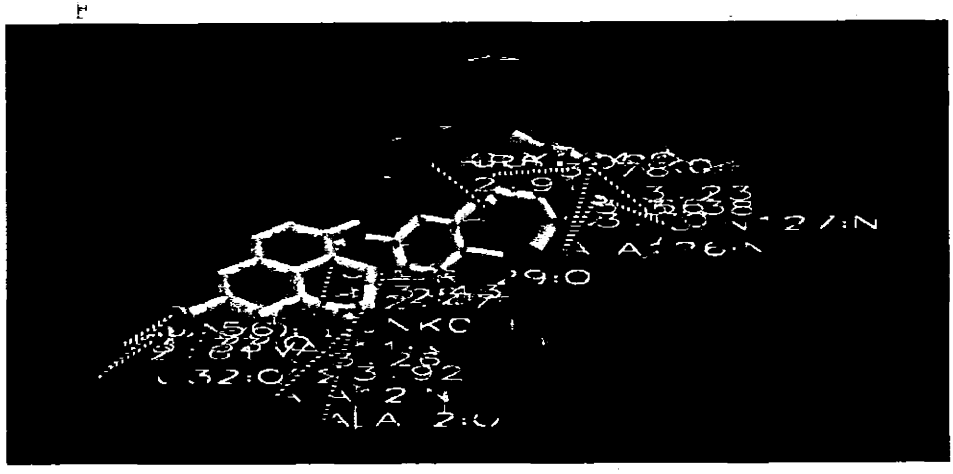


Figure 4.14a: Hydrogen Bonds of the Analogue 2 (Reduction) and the target pcDHFR

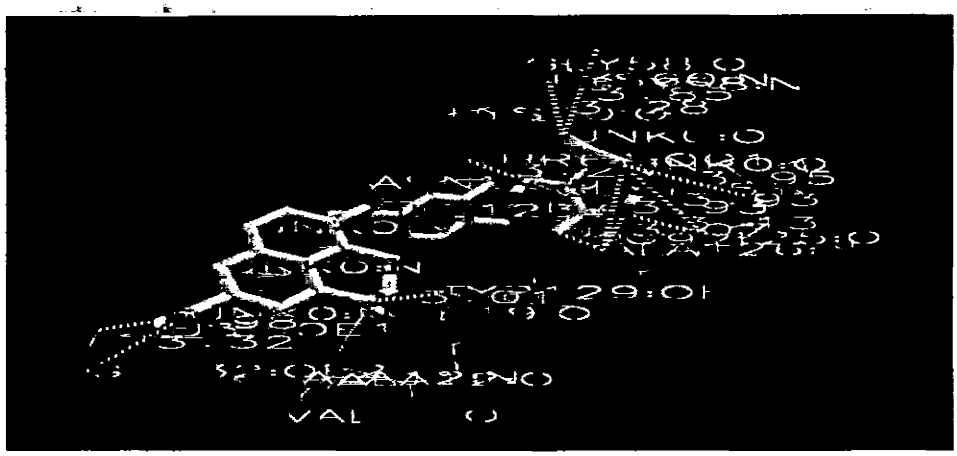


Figure 4.14b: Ionic Bonds of the Analogue 2 (Reduction) and the target protein pcDHFR

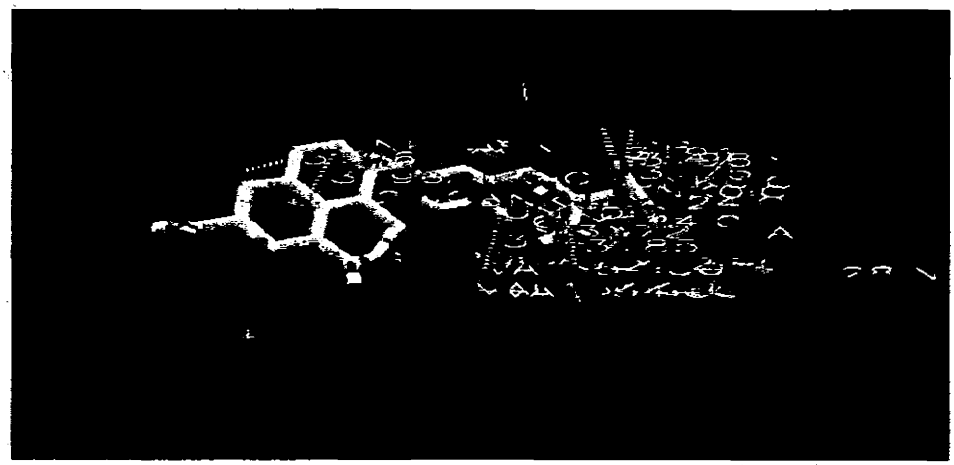


Figure 4.14c: Hydrophobic interactions of the Analogue 2 (Reduction) and the target pcDHFR

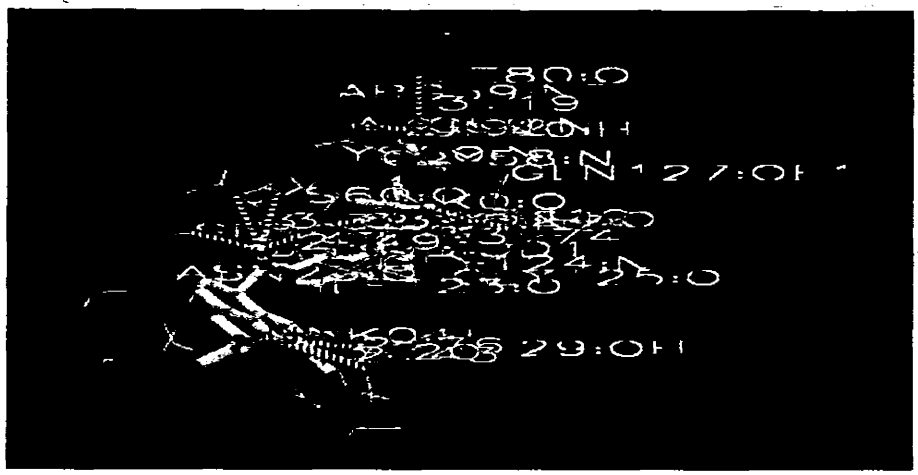


Figure 4.15a: Hydrogen bonds of the Analogue 4 (Amide formation) and the target pcDHFR

Figure 4.15b: Ionic bonds of the Analogue 4 (Amide formation) and the target protein pcDHFR

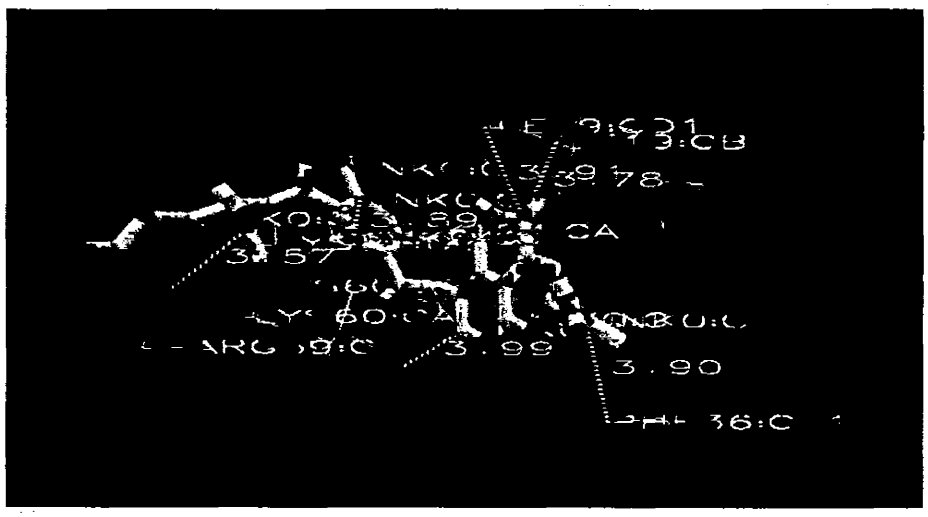


Figure 4.15c: Hydrophobic interactions of the Analogue 4 (Amide formation) and the pcDHFR

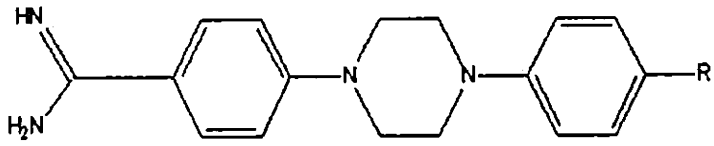
4.8 Quantitative Structure Activity Relationship

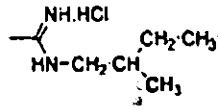
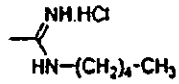
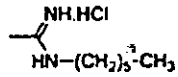
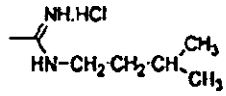
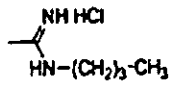
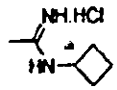
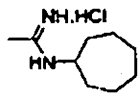
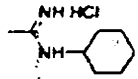
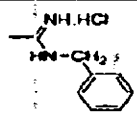
In this particular study one of the major areas of focus is *in-silico* QSAR (Quantitative structure-activity relationship). The relationship between molecular structure and change in biological activity, when similar molecules with minute difference or variations show different biological activity is main area of focus of QSAR Modeling. QSAR is defined as the process "by which chemical structure is quantitatively correlated with a well-defined process, such as biological activity or chemical reactivity" .QSAR are usually based on a comparison executed between a particular type of activity and the chemical structure or a comparison executed between the physicochemical properties of a series of chemical compounds (Mishra *et al.*, 2010).

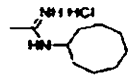
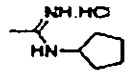
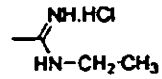
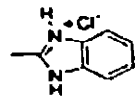
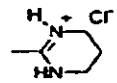
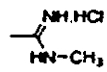
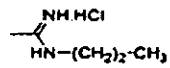
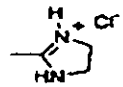
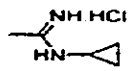
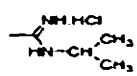
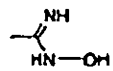
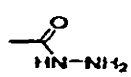
For the selected data set QSAR model was built in order to describe the direct and indirect relation of descriptors to the biological activity i.e. pharmacokinetics of a compound. A selected set of descriptor was chosen and then applied to data set. It was assumed that the descriptors influence the fact whether a given compound will successfully bind to the target protein or fail in binding to the target. 21 compounds were included in this set data. These analogues were obtained by addition of linker, 1, 4-piperazinediyl parent compound with alkyl or cycloalkyl groups introduced on one side of the nitrogen atoms of the amidine moieties (Cushion *et al.*, 2004). The data set for the QSAR analysis is shown in table 4.13.

The descriptors selected for QSAR analysis of the ligands are the partition coefficient i.e. Log P, Molar refractivity (MR), Critical volume (CV), Highest occupied molecular orbital (HOMO), Lowest unoccupied molecular orbital (LUMO), Heat of formation (HF) and Total binding energy of the ligand (TE) respectively. These descriptors were computed using the tools Hyper Chem and Chem Draw.

Table 4.13: Chemical Structure and IC_{50} values of compounds for QSAR studies.



COMPOUNDS	R	IC_{50} (μM)
YQSAR 1		0.002
YQSAR 2		0.002
YQSAR 3		0.007
YQSAR 4		0.016
YQSAR 5		0.091
YQSAR 6		0.115
YQSAR 7		0.117
YQSAR 8		0.208
YQSAR 9		0.242

COMPOUNDS	R	IC_{50} (μM)
YQSAR 10		0.317
YQSAR 11		0.425
YQSAR 12		1.20
YQSAR 13		1.31
YQSAR 14		3.03
YQSAR 15		1.53
YQSAR 16		3.34
YQSAR 17		1.71
YQSAR 18		2.5
YQSAR 19		3.01
YQSAR 20		3.3
YQSAR 21		6.48

The values of LogP and the molar refractivity were obtained using Chem Draw. The descriptors including total binding energy, CV, E_{HOMO}, E_{LUMO}, and heat of formation were computed using the software Hyper Chem. The values of the calculated descriptors have been shown in Table 4.14.

While performing QSAR the first step was the calculation of QSAR equation. The activity parameter used in the QSAR analysis is the IC-50 value. The following QSAR equation was obtained.

IC50 = 3.951227499760E+000 + -2.996849988100E-001*(LogP) + -3.542425594621E-001*(MR) + 9.952824647556E-003*(CV) + -3.713228694800E+000*(EHomo) + 8.138843835032E-002*(ELumo) + -2.915664271882E-004*(T.E) + 1.004416345322E-002*(H.O.F)

QSAR equation calculation was followed by the next step the statistical analysis of the data set. The RSQ value must be high to obtain a good QSAR equation. RSQ is the indication of the high degree fitting of the calculated QSAR equation to the given the data. This would be helpful to get much better predictions for new test data. According to the data set, the RSQ value is adjusted to give a new value. If there is a significant difference in actual and adjusted values, it gives an indication that the overall prediction is weaker. The F statistics show the measure of strength of regression. It can be stated that the QSAR equation is not effective or good if the critical F is greater than F statistics. Table 4.15 shows the values for the above statistical parameters.

The next step involves the establishment of correlation of descriptors with activity. The descriptors including the E_{LUMO}, E_{HOMO}, Heat of formation and E_{TOTAL} values were found to have no correlation with activity so these descriptors are discarded from the set of descriptors.

Table 4.14: QSAR, Steric and Electronic descriptors of Ligands along with IC_{50} value.

Compounds	IC_{50} μM	LogP	MR cm^3/mol	CV cm^3/mol	E_{Homo} kcal/mol	E_{Lumo} kcal/mol	T.E kcal/mol	H.O.F
QPCP1	0.001	4.6	122.94	1093.97	-0.07402	0.064422	-96551	177.1719
QPCP2	0.001	4.62	122.76	1104.17	-0.07539	0.062977	-96553.2	174.9624
QPCP3	0.004	5.03	127.36	1148.25	-0.07096	0.060613	-100001	169.9704
QPCP4	0.009	4.53	122.94	1097.08	-0.07502	0.062179	-96552.3	175.815
QPCP5	0.046	4.2	118.16	1059.32	-0.07558	0.062295	-93104.5	180.4537
QPCP6	0.058	3.67	115.97	1023.6	-0.05854	0.032016	-92354.3	223.5575
QPCP7	0.069	4.92	129.77	1127.72	-0.26382	0.003473	-102756	151.533
QPCP8	0.116	4.51	125.17	1091.25	-0.00493	0.086812	-99273.2	190.9217
QPCP9	0.139	4.83	128.7	1126.97	-0.00198	0.055304	-100556	229.8447
QPCP10	0.195	5.34	134.37	1161.55	-0.26383	0.004742	-106203	146.9504
QPCP11	0.226	4.09	120.57	1058.73	-0.24746	0.013035	-95828.2	192.788
QPCP12	0.524	3.3	108.96	971.04	-0.00694	0.085902	-86208.9	189.8288
QPCP13	0.711	4.12	122.51	1073.15	-0.07481	0.052841	-96430.8	204.661
QPCP14	1.44	3.06	111.88	982.25	-0.25997	0.007357	-88984	150.7096
QPCP15	1.53	2.96	104.16	925.62	-0.00848	0.078141	-82758.9	196.6962
QPCP16	1.6	3.78	113.56	1015.27	-0.00581	0.085371	-89656.8	185.1109
QPCP17	1.71	2.95	107.28	939.8	-0.03008	0.242141	-85463.5	228.1163
QPCP18	2.5	3.25	111.37	999.91	-0.01415	0.054839	-88913.1	221.6328
QPCP19	3.01	3.61	113.65	1005.73	-0.25971	0.006893	-89646.3	195.6194
QPCP20	3.3	2.83	101.69	902.06	-0.02538	0.02737	-86078.5	200.9702
QPCP21	6.48	1.95	101.79	902.32	-0.0776	0.080731	-96551	166.0862

Table 4.15: Statistical parameters and their values

SS_R	44.94
SS_E	7.38
SS_T	52.32
RSQ	85.89 %
Adjusted RSQ	78.30 %
F statistics	11.31
Critical F	2.51

The correlation value ranges from -1 to +1 through 0. -1 indicates ideal negative correlation. 0 as correlation value indicates no correlation at all and +1 gives an indication of perfect positive correlation. The correlation value is best measure for witnessing the tendency of relevance between descriptors and activity (IC-50). To be notable as a good predictor the descriptor should contribute more than half to the activity, that is more than 50%. The values that indicate the percentage contribution are the independent RSQ values of each descriptor if they were used alone. The correlation reveals that Critical volume (CV), Molar refractivity (MR) and LogP proved to be good descriptors for the activity. The table 4.16 shows the values of correlation and the percentage contribution.

Further a plot was generated for the training data that was generated on the basis of the QSAR equation. This shows the relationship between the actual IC-50 values and those predicted by the QSAR model. This will provide a proof of how much correctly the equation is fitting the data. Table 4.17 shows the actual IC-50 values and the predicted values. Figure 4.16 shows the plot between the actual and predicted values.

Table 4.16: Correlation of descriptors with activity and the percentage contribution of each descriptor to activity

Descriptors	Correlation	Percentage
LogP	-0.80	64.69 %
CV	-0.75	55.93 %
MR	-0.73	53.78 %
E_{HOMO}	0.06	0.33 %
E_{LUMO}	0.11	1.15 %
E_{TOTAL}	0.37	13.70 %
H.O.F	0.05	0.21 %

Table 4.17: Actual and predicted IC-50 values of QSAR data set

No.	ACTUAL	PREDICTED	No.	ACTUAL	PREDICTED
1.	0.00	0.12	12.	0.52	1.10
2.	0.00	0.26	13.	0.71	0.45
3.	0.00	-0.11	14.	1.44	1.60
4.	0.01	0.16	15.	1.53	1.52
5.	0.05	0.62	16.	1.60	0.72
6.	0.06	1.35	17.	1.71	1.76
7.	0.07	0.19	18.	2.50	1.68
8.	0.12	0.01	19.	3.01	1.69
9.	0.14	-0.23	20.	3.30	3.27
10.	0.20	-0.27	21.	6.48	6.40
11.	0.23	1.35			

CONCLUSION AND FUTURE INVESTIGATIONS

Conclusion and Future Investigations

The present study was aimed at the identification of novel drugs for the treatment of *Pneumocystis carinii* pneumonia that possessed the best pharmacophore features and binding interactions.

The Ligand based pharmacophore modeling approach was brought in to use and pharmacophore models were generated. The shared pharmacophore was identified using the training set composed of 10 compounds belonging to different classes along with one standard drug Trimethoprim. All the compounds along with the standard drug were superimposed for the purpose of generating the shared pharmacophore model. The selected compounds were different groups of disubstituted diaminopteridines (Jackson *et al.*, 1996), O-alkyl derivatives (Rosowsky and Forsch, 2003), triazolyl analogues (Chan *et al.*, 2002), benzanilides and benzylamines (Da Cunha *et al.*, 2010), substituted 2, 4-diamino-5-benzylpyrimidine (Forsch *et al.*, 2004), 1-naphthyl derivative (Hallberg *et al.*, 2004) and Diaminoquinazolines (Rosowsky *et al.*, 1995). The identified pharmacophore consisted features including two hydrogen bond acceptors, two hydrogen bond donors and one aromatic volume. No such model has been predicted or reported in literature till date for pcDHFR inhibitors.

A pharmacophore model specific for the pcDHFR inhibitors was identified. The generated pharmacophore can be utilized effectively to predict the activity of a wide variety of chemical scaffolds. Moreover it has the capability to can be used as a 3D query for database searches. This helps in determining a variety of compounds that can prove as potent pcDHFR inhibitors.

Molecular docking studies were performed on the entire data set and the potential lead compound was identified on the basis of best energy score i.e. minimum energy score and greater number of interactions by the effective use of Auto Dock Vina. Lead compound was selected and the corresponding four analogues were designed. Amongst the four analogues, two showed the potential to for further investigations since they had more binding interactions thus it can be concluded that they will have more bioavailability. The two analogues that show the significant number of interactions are proposed for clinical trials and synthesis in laboratory with the sole purpose to bring forth a better drugs having high bioavailability than the already available drugs which may treat pneumocystis infection.

Molecular docking was used as an important tool through which the important interactions between the potent inhibitors and the active site residues were uncovered. Using a combination of pharmacophore modeling, virtual screening, and molecular docking, putative novel pcDHFR analogues were successfully identified, which can be further evaluated by in vitro and in vivo biological tests.

For the QSAR studies a selected set of descriptor was chosen and then applied to data set of 21 compounds. These analogues were obtained by addition of linker, 1, 4-piperazinediyl parent compound with alkyl or cycloalkyl groups introduced on one side of the nitrogen atoms of the amidine moieties (Cushion *et al.*, 2004). The analysis showed that descriptors including the E_{LUMO}, E_{HOMO}, Heat of formation and E_{TOTAL} values were found to have no correlation. The correlation reveals that Critical volume (CV), Molar refractivity (MR) and LogP proved to be good descriptors for the activity. The descriptor which was found to be critical for *P.carinii* inhibitor in this study may be evaluated for other classes of compounds to get a general view.

In terms of recommendations for future directions the research can be split into two branches firstly research can be conducted in terms of *Pneumocystis* basic science and secondly in terms of drug designing for the cure of PCP. The following aspects should be considered to be of immense importance for *Pneumocystis* basic science research:

- Improvement in understanding *Pneumocystis* interactions with a range of cells including the epithelial cells, the dendritic cells, macrophages and proteins of the lung principally surfactant protein should be made.
- Further development of vaccines and novel methods to therapeutically deal with the lung inflammation during the infection should be focussed. A better and improved understanding of life cycle along with the signaling pathways and the particular culture requirements are basic necessity for its future investigation.
- An improved understanding of the current epidemiology of PCP throughout the world including the middle income countries needs to be developed. More over an appropriate examination of the PCP prophylaxis is required. The continuation to the development and validation of new methods for PCP diagnosis is also a major area of research. Extensive research study should be conducted on putative drug resistance in human including the particular association between the mutations in the *P.carinii* dihydropteroate synthase gene and trimethoprim-sulfamethoxazole. Further exploration of issues concerning PCP treatment failure and death should be considered. Comprehensive research is required for studying the further clarification of the role of PCP prophylaxis for the non-HIV immune compromised hosts. Future research should be conducted in order to further investigate the risk of PCP or

transmission of PCP (Huang et al., 2006).

- However important factors associated with the drugs for the cure of PCP such as the deleterious side effects, limited efficacies of these drugs, and emerging mutations and drug resistance cannot be neglected and thus provide a justification for the search of more effective and less toxic medicinal agents (Eynde *et al.*, 2004).

Last but not the least in terms of drug designing for the *Pneumocystis carinii* pneumonia future investigations can be focused on the study of the energy landscape of the two analogues suggested for clinical trials. Simulation studies can also be performed on the analogues. Molecular dynamics may assist in the exploration of the energy landscape and the free energy simulations can also be used to compute the relative binding free energies of a series of putative drugs.

REFERENCES

REFERENCES

Abagyan R, and Totrov M (2001). High-throughput docking for lead generation, *Curr Opin Chem Biol*, 5(4): 375-382.

Agatonovic-Kustrin S, Beresford R and Yusof AP (2001). Theoretically-derived molecular descriptors important in human intestinal absorption, *J Pharm Biomed Anal*, 25(2): 227-237.

Ajay A, Walters WP, and Murcko MA (2001). Can we learn to distinguish between "drug-like" and "nondrug-like" molecules?, *Curr Med Chem*, 8(6): 685-713.

Allegra CJ, Kovacs JA, Drake JC, Swan JC, Chabner BA and Masur H (1987). Activity of antifolates against *Pneumocystis carinii* dihydrofolate reductase and identification of a potent new agent, *J Exp Med*, 165: 926-931.

Bajorath J (1999). Computer-Aided Drug Discovery: from target proteins to drug candidates, *Pacific Symposium on Biocomputing*, 4: 413-414.

Bannwarth B, Labat L, Moride Y and Schaeverbeke T (1994). Methotrexate in rheumatoid-arthritis - an update, *Drugs*, 47(1): 25-50.

Barnum D, Greene J, Smellie A and Sprague P (1996). Identification of common functional configurations among molecules, *J Chem Inf Comput Sci*, 36(3): 563-571.

Bartlett MS, Queener SF, Tidwell RR, Milhous WK, Berman JD, Ellis WY and Smith JW (1991). 8-Aminoquinolines from Walter Reed Army Institute for Research for treatment and prophylaxis of Pneumocystis carinni in rat models. *Antimicrob Agents Chemother.* 35(2): 277-282.

Bartlett MS, Shaw M, Navaran P, Smith JW, Queener SF (1995). Evaluation of potent inhibitors of dihydrofolate reductase in a culture model for growth of

Pneumocystis carinii, *Antimicrob Agents Chemother*, 39(11): 2436-2441.

Berman HM, Westbrook J, Feng Z, Gilliland G, Bhat TN, Weissig H, Shindyalov IN, and Bourne PE (2000). The Protein Data Bank, *Nucleic Acids Res*, 28(1): 235-242.

Bertino J (1993). Karnof sky Memorial Lecture: ode to methotrexate, *J Clin Oncol*, 11(1): 5-14.

Bissantz C, Folkers G, Rognan D (2000). Protein-based virtual screening of chemical databases. 1. Evaluation of different docking/scoring combinations, *J Med chem*, 43(2): 4759-4767.

Boehm HJ and Stahl M (2002). The Use of Scoring Functions in Drug Discovery Applications, *Rev Comp Chem*, 18: 41-87.

Broughton MC, Queener SF (1991). Pneumocystis carinii dihydrofolate reductase used to screen potential anti-pneumocystis drugs, *Antimicrob Agents Chemother*, 35(7): 1348-1355.

Brun-Pascaud M, Fay M, Zhong M, Bauchet J, Dux-Guyot A and Poeidalo J. J (1992). Use of fluoroquinolones for prophylaxis of murine Pneumocystis carinii pneumonia, *Antimicrob Agents Chemother*, 36(2): 470-472.

Brusic V, Bucci K, Schönbach C, Petrovsky N, Zeleznikow J and Kazura JW (2001). Efficient discovery of immuneresponse targets by cyclical refinement of QSAR models of peptide binding, *J Mol Graph Model*, 19(5): 405-411.

Burden FR and Winkler DA (2000). The computer simulation of high throughput screening of bioactive molecules, *Mol Model Predict Bioact*, [Proceedings of the 12th European Symposium on Quantitative Structure-Activity Relationships], 175-180.

Carini A (1910). Formas de eschizogonia de *Trypanosomalewisii*, *Arch Soc Med Ci Sao Paulo,* 38: 204.

Chagas C (1909). Nova tripanomiazehumana. Estudossobrea morfolojia e o cicloevolutivo do Schizotrypanumcruzi n. gen., n. sp., ajenteetiolojico de nova entidade morbida do homen, *Mem Inst Oswaldo Cruz,* 1: 159-218.

Champness JN, Achari A, Ballantine SP, Bryant PK, Delves CJ and Stammers DK (1994). The structure of Pneumocystis carinii dihydrofolate reductase to 1.9Å resolution, *Structure*, 2(10): 915-24.

Chang MW, Ayeni C, Breuer S and Torbett BE (2010). Virtual Screening for HIV Protease Inhibitors: A Comparison of AutoDock 4 and Vina, *PLoS ONE*, 5(8): e11955.

Chen R, Li L and Weng Z (2003). ZDOCK: an initial-stage protein-docking algorithm, *Proteins*, 52(1): 80-87.

Clark RD, Strizhev A, Leonard JM, Blake JF and Matthew JB (2002). Consensus scoring for ligand/protein interactions, *J Mol Graph Model*, 20(4): 281-295.

Clarkson AB, Saric M, and Grady RW (1990). Deferoxamine and eflornithine (DL-alpha-difluoromethylornithine) in a rat model of *Pneumocystis carinii* pneumonia, *Antimicrob Agents Chemother*, 34(9): 1833-1835.

Cramer RD, Patterson DE and Bunce JD (1988). Comparative molecular field analysis (comfa). 1. Effect of shape on binding of steroids to carrier proteins, *Eur J Med Chem*, 110(18): 5959–5967.

Cushion M.T (1994). Transmission and epidemiology. In:Walzer D.P., ed. *Pneumocystis carinii Pneumonia.* NewYork, NY: Marcel Dekker, pp122-139.

Cushion MT, Chen F, Kloepfer N (1997). A cytotoxicity assay for evaluation of

candidate anti-Pneumocystis carinii agents, *Antimicrob Agents Chemother*, 41(2): 379-384.

Cushion MT, Walzer PD, Ashbaugh A, Rebholz S, Brubaker R, Vanden Eynde JJ, Mayence A, Huang TL (2006). In vitro selection and in vivo efficacy of piperazine- and alkanediamide-linked bisbenzamidines against Pneumocystis pneumonia in mice, *Antimicrob Agents Chemother*, 50(7): 2337-2343.

Cushion MT, Walzer PD, Collins MS, Rebholz S, Vanden Eynde JJ, Mayence A, Huang TL (2004). Highly active anti-Pneumocystis carinii compounds in a library of novel piperazine-linked bisbenzamidines and related compounds, *Antimicrob Agents Chemother*, 48(11): 4209-4216.

Da Cunha EFF, Ramalho TC, Maia ER and De Alencastro RB (2005). The search for new DHFR inhibitors: a review of patents, January 2001 – February 2005, *Expert Opinion on Therapeutic Patents*, 15: 967.

Da Cunha EFF, Ramalho TC, Mancini DT, Fonseca EMB and Oliveira AA (2010). New Approaches to the Development of Anti-Protozoan Drug Candidates: a Review of Patents, *J Braz Chem Soc*, 21(10): 1787-1806.

Da Cunha EFF, Sippl W, de Castro Ramalho T, Ceva Antunes OA, de Alencastro RB and Albuquerque MG (2009). 3d-qsar comfa/comsia models based on theoretical active conformers of hoe/bay-793 analogs derived from hiv-1 protease inhibitor complexes, *Eur J Med Chem*, 44: 4344-4352.

D'Antonio RG, Johnson DB, Winn RE, Van Dellen AF and Evans ME (1986). Effect of folinic acid on the capacity of trimethoprim-sulfamethoxazole to prevent and treat Pneumocystis carinii pneumonia in rats, *Antimicrobial Agents and Chemotherapy*, 29(2): 327-329.

Debs RJ, Blumenfeld W, Brunette EN, Straubinger RM, Montgomery AB, Lin E, Agabian N and Papahadjopoulos D (1987). Successful treatment with aerosolized pentamidine of Pneumocystis carinii pneumonia in rats, *Antimicrob Agents Chemother*, 31(1): 37-41.

Delanoe P and Delanoe M (1912). Sur les rapports des kystes de Carini du poumon des rats avec le *Trypanosomalewisi*, *C R Acad Sci (Paris)*, 155: 168-170.

Devita VT, Broder S, Fauci AS, Kovacs JA, and Chabner BA (1987). Developmental therapeutics and the acquired immunodeficiency syndrome, *Ann Intern Med*, 106(4): 568-581.

Dias R and de Azevedo WF (2008). Molecular Docking Algorithms, *Current Drug Targets*, 9(12):1040-1047

DiMasi JA, Hansen RW and Grabowski HG (2003). The price of innovation: new estimates of drug development costs, *J Health Econ*, 22(2): 151-185.

Dixon SL, Smondyrev AM, Knoll EH, Rao SN, Shaw DE and Friesner RA (2006). PHASE: a new engine for pharmacophore perception, 3D QSAR model development, and 3D database screening: 1. Methodology and preliminary results, *J Comput Aid Mol Des*, 20(10-11): 647-671.

Dykstra CC and Tidwell RR (1991). Inhibition of topoisomerases from Pneumocystis carinii by aromatic dicationic molecules, *J Protozool*, 38(6): 78S-81S.

Edman JC, Edman U, Cao M, Lundgren B, Kovacs JA and Santi DV (1989). Isolation and expression of the *Pneumocystis carinii* dihydrofolate reductase gene, *Proc Natl Acad Sci USA*, 86:8625-8629.

Edman U, Edman JC, Lundgren B and Santi DV (1989). Isolation and expression of the *Pneumocystis carinii* thymidylate synthetase gene, *Proc Natl Acad Sci U S A*,

86(17): 6503-6507.

Ewing TJ, Makino S, Skillman AG and Kuntz ID (2001). DOCK 4.0: search strategies for automated molecular docking of flexible molecule databases, *J Comput Aided Mol Des*, 15(5): 411-428.

Ferrara P, Gohlke H, Price DJ, Klebe G and Brooks CL (2004). Assessing scoring functions for protein-ligand interactions, *J Med Chem*, 47(12): 3032-3047.

Fishman JA (1998). Prevention of Infection Due to *Pneumocystis carinii*, *Antimicrobial Agents and Chemotherapy*, 42(5): 995-1004.

Fisk TD, Meshnick S, Kazanjian HP (2003). *Pneumocystis carinii* pneumonia in patients in the developing world who have acquired immunodeficiency syndrome, *Clin Infect Dis*, 36: 70-78.

Fox D and Smulian GA (1999). Mitogen-activated protein kinase Mkp1 of *Pneumocystis carinii* complements the slt2Delta defect in the cell integrity pathway of *Saccharomyces cerevisiae*, *Mol Microbiol*, 34(3): 451-462.

Frenkel JK, Good JT and Schultz JA (1966). Latent Pneumocystis infection of rats, relapse, and chemotherapy, *Lab Invest*, 15(10): 1559-1577.

Gangjee A, Yang J and Queener SF (2006). Novel non-classical C9-methyl-5-substituted-2,4-diaminopyrrolo-[2,3-d]pyrimidines as potential inhibitors of dihydrofolate reductase and as anti-opportunistic agents, *Bioorg Med Chem*, 14(24): 8341-8351.

Garenne M, Ronsmans C, Campbell H (1992). "The magnitude of mortality from acute respiratory infections in children under 5 years in developing countries", *World Health Statistics Quarterly*, 45(2-3): 180-91.

Ghose AK and Wendoloski JJ (1998). In Perspective in Drug Discovery and Design,

Escom, 9(10-11): 253.

Gigliotti F and Hughes WT (1988). Passive immunoprophylaxis with specific monoclonal antibody confers partial protection against Pneumocystis carinii pneumonitis in animal models, *J Clin Invest*, 81(6): 1666-1668.

Gombar VK and Enslein K (1996). Assessment of n-octanol/water partition coefficient: when is the assessment reliable?, *J Chem Inf Comput Sci*, 36(6): 1127-1134.

Goodell B, Jaccbs JB, Powell RD, DeVita VT (1970). Pneumocystis carinii: The spectrum of diffuse interstitial pneumonia in patients with neoplastic disease, *Ann Intern Med*, 72: 337-340.

Goodsell DS, Morris GM and Olson AJ (1996). Automated docking of flexible ligands: applications of AutoDock, *J Mol Recognit*, 9(1): 1-5.

Hansch C, Smith N, Engle R and Wood H (1972). Quantitative structure-activity relationships of antineoplastic drugs: nitrosoureas and triazenoimidazoles, *Cancer Chemother Rep*, 56(4): 443-456.

Hawksworth DL (2007). Responsibility in namingpathogens: the case of Pneumocystis jirovecii, the causal agent of pneumocystispneumonia, *Lancet Infect Dis*, 7 (1):3-5.

Huang L, Morris A, Limper AH and Beck JM (2006). An Official ATS Workshop Summary: Recent Advances and Future Directions in *Pneumocystis* Pneumonia (PCP), *Proc Am Thorac Soc*, 3(8): 655-64.

Hughes TW, McNabb CP, Makres DT and Feldman S (1974). Efficacy of trimethoprim and sulfamethoxazole in the prevention and treatment of Pneumocystis carinii pneumonitis, *Antimicrobial Agents Chemotherapy*, 5: 289-

293.

Hughes WT and Smith BL (1984). Efficacy of diaminodiphenylsulfone and other drugs in murine Pneumocystis carinji pneumonitis, *Antimicrobial Agents Chemotherapy*, 26(4): 436-440.

Hughes WT, Gray VL, Gutteridge WE, Latter VS and Pudney M (1990). Efficacy of a hydroxynaphthoquinone, 566C80, in experimental Pneumocystis canniipneumonitis, *Antimicrob Agents Chemother*, 34(2): 225-228.

Hughes WT, McNabb PC, Makres TD and Feldman S (1974). Efficacy of trimethoprim and sulfamethoxazole in theprevention and treatment of Pneumocystis carinii pneumonitis, *Antimicrobial Agents Chemotherapy*, 5(3): 289-293.

Humphrey W, Dalke A, Schulten K (1996). VMD – Visual Molecular Dynamics, *J Mol Graphics*, 14:33-38.

Hussain Z, Carlson ML, Craig ID and Lannigan R (1985). Efficacy of tetroxoprim/sulfadiazone in the treatment of Pneumocystis carinii in rats, *Journal of Antimicrobial Chemotherapy*, 15(5):575-578.

Hypercube, Inc., HyperChem® Release 7 for Windows®, (Jan 2002)

Jackson HC, Biggadike K, McKilligin E, Kinsman OS, Queener SF, Lane A, Smith JE (1996). 6,7-disubstituted 2,4-diaminopteridines: novel inhibitors of Pneumocystis carinii and Toxoplasma gondii dihydrofolate reductase, *Antimicrob Agents Chemother*, 40(6): 1371-1375.

Jain AN (2003). Surflex: fully automatic flexible molecular docking using a molecular similarity-based search engine, *Med Chem*, 46(4): 499-511.

Jones G and Willet P (2000). GASP: genetic algorithm superimposition program. In

Gu¨ner, O.F. (ed.), Pharmacophore Perception, Development and Use in Drug Design, International University Line, La Jolla, CA, pp. 85–106.

Jones G, Willett P, Glen RC, Leach AR and Taylor R (1997). Development and validation of a genetic algorithm for flexible docking. *J Mol Biol*, 267(3): 727-748.

Joy S, Nair PS, Hariharan R and Pillai MR (2006). Detailed comparison of the protein-ligand docking efficiencies of GOLD, a commercial package and ArgusLab, a licensable freeware, *In Silico Biol*, 6(6): 601-605.

Kapetanovic IM (2008). Computer-aided drug discovery and development (CADDD): in silico-chemico-biological approach, *Chem Biol Interact*, 171(2): 165-176.

Kirby HB, Kenamore B, and Guckian JC (1971). Pneumocystis carinii pneumonia treated with pyrimethamine and sulfadiazine, *Ann Intern Med*, 75(4): 505-509.

Kluge RM, Spaulding DM, and Spain AJ (1978). Combination of pentamidine and trimethoprim-sulfamethoxazole in the therapy of Pneumocystis carinii pneumonia in rats, *Antimicrobial Agents Chemotherapy*, 13(6): 975-978.

Kottom TJ, Kohler JR, Thomas CF, Fink GR and Limper AH (2003). Lung epithelial cells and extracellular matrix components induce expression of Pneumocystis carinii STE20, a gene complementing the mating and pseudohyphal growth defects of STE20 mutant yeast, *Infect Immun*, 71(11): 6463-6471.

Kovacs JA, Allegra CJ and Masur H (1990). Characterization of dihydrofolate reductase of Pneumocystis carinii and Toxoplasma gondii, *Exp Parasitol*, 71(1): 60-68.

Kovacs JA, Allegra CJ, Beaver J, Boarman D, Lewis M, Parrillo JE, Chabner B, and Masur H (1989). Characterization of de novo folate synthesis of Pneumocystis

carinii and Toxoplasma gondii: potential for screening therapeutic agents, *J Infect Dis*, 160(2): 312-320.

Kovacs JA, Gill VJ, Meshnick S and Masur H (2001). New insights into transmission, diagnosis, and drug treatment of *Pneumocystis carinii* pneumonia, *JAMA*, 286(19): 2450-60.

Kovacs JA, Halpern JL, Lundgren B, Swan JC, Parrillo JE, and Masur H (1989). Monoclonal antibodies to *Pneumocystis carinii*, *J Infect Dis*, 159(1): 60-70.

Kovacs JA, Hiemenz JW, Macher AM, Stover D, Murray HW, Shelhamer J, Lane HC, Urmacher C, Honig C, Longo DL, Parker MM, Natanson C, Parrillo JE, Fauci AS, Pizzo PA and Masur H (1984). *Pneumocystis carinii* pneumonia: a comparison between patients with the acquired immunodeficiency syndrome and patients with other immunodeficiencies, *Ann Intern Med*, 100(5): 663-671.

Kramer B, Rarey M and Lengauer T (1999). Evaluation of the FLEXX incremental construction algorithm for protein-ligand docking, *Proteins*, 37(2): 228-241.

Kumar N, Hendriks BS, Janes KA, de Graaf D and Lauffenburger DA (2006). Applying computational modeling to drug discovery and development. *Drug Discovery Today*, 11(17–18): 806-811.

Kuntz ID, Blaney JM, Oatley SJ, Langridge R and Ferrin TE (1982). A geometric approach to macromolecule-ligand interactions. *J Mol Biol*, 161(2): 269-288.

Kurogi Y and Gu¨ner OF (2001). Pharmacophore modeling and three dimensional database searching for drug design using catalyst, *Curr Med Chem*, 8(9): 1035-1055.

Lahger T and Krovat EM (2003). Chemical feature-based pharmacophores and virtual library screening for discovery of new leads, *Curr Opin Drug Discov Devel*, 6(3):

370-376.

Leoung GS, Mills J, Hopewell PC, Hughes W and Wofsy C (1986). Dapsone-trimethoprim for *Pneumocystis carinii* pneumonia in the acquired immunodeficiency syndrome, *Ann Intern Med*, 105(1): 45-48.

Lewis DF (2000). Structural characteristics of human P450s involved in drug metabolism: QSARs and lipophilicity profiles, *Toxicology*, 144(1-3): 197-203.

Li H, Li C, Gui C, Luo X, Chen K, Shen J, Wang X and Jiang H (2004). GAsDock: a new approach for rapid flexible docking based on an improved multi-population genetic algorithm, *Bioorg Med Chem Lett* , 14(18): 4671-4676.

Limper AH, Offord KP, Smith TF, Martin WJ (1989). Pneumocystis carinii pneumonia. Differences in lung parasite number and inflammation in patients with and without AIDS, *Am Rev Respir Dis*, 140(5): 1204-1209.

Lipschick GY, Masur H and Kovacs JA (1991). Polyamine metabolism in Pneumocystis carinii, *J Infect Dis*, 163(5): 1121-1127.

Liu J, Bolstad DB, Bolstad ES, Wright DL, Anderson AC (2008). Towards New Antifolates Targeting Eukaryotic Opportunistic Infections, *Eukaryot Cell*, 8(4):483-486.

Liu M and Wang S (1999). MCDOCK: a Monte Carlo simulation approach to the molecular docking problem, *J Comput Aided Mol Des*, 13(5): 435-51.

Lybrand TP (1995). Ligand-protein docking and rational drug design, *Current Opinion in Structural Biology*, 5(2): 224–228.

Madhi SA, Levine OS, Hajjeh R, Mansoor OD, Cherian T (2008). Vaccines to prevent pneumonia and improve child survival, *Bulletin of the World Health Organization*, 86: 321-416.

Marshall GR et al. (1979). The conformational parameter in drug design: the active analog approach. In Olson, E.C. and Christoffersen, R.E,(eds), Computer-Assisted Drug Design, American Chemical Society, 112: 205-225.

Martin, Y.C. (2000) DISCO: what we did right and what we missed. In Gu¨ner, O.F., (ed), Pharmacophore Perception, Development and Use in Drug Design, International University Line, La Jolla, CA, pp49-68.

Matsumoto Y, Yamada M and Amagai T (1991). Yeast glucan of Pneumocystis carinii cyst wall: an excellent target for chemotherapy, *J Protozool*, 38(6): 6S-7S.

Mendelsohn LD (2004). "ChemDraw 8 Ultra: Windows and Macintosh Versions", *J Chem Inf Comput Sci*, 44 (6): 2225–2226.

Mendelsohn LD (2004). "ChemDraw 8 Ultra: Windows and Macintosh Versions", *J. Chem Inf Comput. Sci.*, 44 (6): 2225–2226

Merali S, Frevert U, Williams JH, Chin K, Bryan R and Clarkson AB (1999). Continuous axenic cultivation of *Pneumocystis carinii*, *Proc Natl Acad Sci U S A*, 96(5): 2402-2407.

Mishra P, Tripathi V, Yadav BS (2010). Insilco QSAR modeling and drug development process, *GERF Bulletin of Biosciences*, 1(1): 37-40.

Mohan V, Gibbs AC, Cummings MD, Jaeger EP and DesJarlais RL (2005). Docking: Successes and Challenges, *Curr Pharm Des*, 11(3): 323-333.

Morris GM, Goodsell DS, Halliday RS, Huey R, Hart WE, Belew RK, and Olson AJ (1998). Automated Docking Using a Lamarckian Genetic Algorithm and an Empirical Binding Free Energy Function, *J Comp Chem*, 19(14): 1639-1662.

Morris GM, Goodsell DS, Huey R and Olson AJ (1996). Distributed automated docking of flexible ligands to proteins: parallel applications of AutoDock 2.4, *J*

Comput Aided Mol Des, 10(4): 293-304.

Morris GM, Goodsell DS, Huey R, Olson AJ (1996). Distributed automated docking of flexible ligands to proteins: parallel applications of AutoDock 2.4, *J Comput Aided Mol Des*, 10(4): 293-304.

Morris GM, Huey R, Lindstrom W, Sanner MF, Belew RK, Goodsell DS and Olson AJ (2009). AutoDock4 and AutoDockTools4: Automated Docking with Selective Receptor Flexibility, *J Comput Chem,* 30(16): 2785-291.

Mullarkey MF, Blumenstein BA, Andrade WP, Bailey GA, Olason I, Wetzel CE (1988). Methotrexate in the treatment of corticosteroid-dependent asthma. A double-blind crossover study, *New England Journal of Medicine*, 318(10): 603-607.

Myers S and Baker A (2001). Drug discovery—an operating model for a new era, *Nat Biotechnol*, 19(8): 727-730.

Okimoto N, Futatsugi N, Fuji H, Suenaga A, Morimoto G, Yanai R, Ohno Y, Narumi T, Taiji M (1996). High-performance drug discovery: computational screening by combining docking and molecular dynamics simulations, *PLoS Comput Biol*, 5(10):e1000528.

Pareja JG, Garland R and Koziel H (1998). Use ofadjunctive corticosteroids in severe adultnon-HIV Pneumocystis carinii pneumonia, *Chest*, 113(5): 1215-1224.

Patrick DA, Boykin DW, Wilson WD, Tanious FA, Spychalaz J, Bender BC, Hall JE, Dykstra CC, Ohemeng KA, Tidwell RR (1997). Anti-Pneumocystis carinii pneumonia activity of dicationic carbazoles Eur J Med Chem, 32, 78 l-793.

Pearson RD and Hewlett EL (1987). Use of pyrimethamine sulfadoxine (Fansidar) in prophylaxis against chloroquine-resistant Plasmodium falciparum and

Pneumocystis carinii, *Ann Intern Med*, 106(5): 714-718.

Peppercorn MA (1990). Advances in drug therapy for inflammatory bowel disease, *Ann Intern Med*, 112(1): 50-60.

Pierce B, Tong W and Weng Z (2005). M-ZDOCK: a grid-based approach for Cn symmetric multimer docking, *Bioinformatics*, 21(8): 1472-1478.

Post C, Fakouhi T, Dutz W, Bandarizadeh B and Kohut E (1971). Prophylaxis of epidemic infantile pneumocytosis with a 20:1 sulfadoxine plus pyrimethamine combination, *Curr Ther Res Clin Exp*, 13(5): 273-279.

Queener SF (1995). New drug developments for opportunistic infections inimmunosuppressed patients: Pneumocystis carinii, *J Med Chem*, 38(24): 4739-4759.

Queener SF, Bartlett MS, Jay MA, Durkin MM and Smith JW (1987). Activity of lipid-soluble inhibitors of dihydrofolate reductase against Pneumocystis carinii in culture and in a rat model of infection, *Antimicrob Agents Chemother*, 31(9): 1323-1327.

Queener SF, Bartlett MS, Nasr M, Smith JW (1993). 8-aminoquinolines effective against Pneumocystis carinii in vitro and in vivo, *Antimicrob Agents Chemother*, 37(10): 2166-2172.

Queener SF, Bartlett MS, Richardson JD, Durkin MM, Jay MA and Smith JW (1988). Activity of clindamycin with primaquine against Pneumocystis carinii in vitro and in vivo, *Antimicrob Agents Chemother*, 32(6): 807-813.

Rarey M, Kramer B, Lengauer T and Klebe G (1996). A fast flexible docking method using an incremental construction algorithm, *J Mol Biol*, 261(3): 470-489.

Rarey M, Kramer B, Lengauer T, Klebe G (1996). A fast flexible docking method

using an incremental construction algorithm, *J Mol Biol*, 261: 470-489.

Redhead SA, Cushion MT, Frenkel JK, Stringer JR (2006) *Pneumocystis* and *Trypanosomacruzi*: nomenclature and typifications, *J Eukaryot Microbiol*, 53:2-11.

Richmond NJ, Willett P and Clark RD (2004). Alignment of three-dimensional molecules using an image recognition algorithm, *Journal of Molecular Graphics and Modelling*, 23(2): 199-209.

Rosowsky A, Hynes JB, Queener SF (1995).Structure-Activity and Structure-Selectivity Studies on Diaminoquinazolines and Other Inhibitors of *Pneumocystis carinii* and *Toxoplasma gondii* Dihydrofolate Reductase, *Antimicrob Agents Chemother*, 39(1): 79-86.

Rudan I, Boschi-Pinto C, Biloglav Z, Mulholland K, Campbell H (2008). Epidemiology and etiology of childhood pneumonia, *Bulletin of the World Health Organization*, 86: 321-416.

Sattler FR, Cowan R, Nielsen DM and Rushkin J (1988). Trimethoprim-sulfamethoxazole compared with pentamidine for treatment of Pneumocystis carinii pneumonia in the acquired immunodeficiency syndrome. A prospective, noncrossover study, *J Ann Int Med*, 109(4): 280-287.

Sauton N, Lagorce D, Villoutreix BO and Miteva MA (2008). MS-DOCK: accurate multiple conformation generator and rigid docking protocol for multi-step virtual ligand screening, *BMC Bioinformatics*, 9: 184-196.

Schmatz D.M, Powles M, McFadden D, Pittarelli L, Liberator PA and Anderson JW (1991). Treatment and prevention of Pneumocystis cannii pneumonia and further elucidation of the Pneumocystis carinii life cycle with 1, 3, -beta-glucan synthesis

inhibitor, *J Protozool*, 38:151S-153S.

Schmatz DM, Romancheck MA, Pittarelli LA, Schwartz RE, Fromtling RA, Nollstadt KH, Vanmiddlesworth FL, Wilson KE and Turner MJ (1990). Treatment of Pneumocystis carinii with 1, 3- beta-glucan synthesis inhibitors, *Proc Natl Acad Sci USA*, 87(15): 5950-5945.

Schultz TW and Seward JR (2000). Health effects related structure-toxicity relationships: A paradigm for the first decade of the new millennium, *Sci Total Environ*, 249(1-3): 73-84.

Schweitzer BI, Dicker AP and Joseph R (1990). Dihydrofolate reductase as a therapeutic target, *FASEB J*, 4(8): 2441-2452.

Sepkowitz KA (2002). Opportunistic infections in patients with and patients without acquired immunodeficiency syndrome, *Clin Infect Dis*, 34(8): 1098-107.

Sepkowitz KA, Brown AE, Telzak EE, Gottlieb S and Armstrong D (1992). Pneumocystis carinii pneumonia among patients without AIDS at a cancer hospital, *JAMA*, 267(6): 832-837.

Sesterhenn TM, Cushion MT, Slaven BE and Smulian AG (2006). Sequence of the Mitochondrial Genome of Pneumocystis carinii: Implications for Biological Function and Identification of Potential Drug Targets, *J Eukaryot Microbiol*, 53(1): 154–155.

Shear HL, Valladares G and Narachi MA (1990). Enhanced treatment of Pneumocystis carinii pneumonia in rats with interferon--y and reduced doses of trimethoprim/sulfmethoxazole, *J Acquired Immune Defic Syndr*, 3(10): 943-948.

Siegel SE, Wolff LJ, Baehner RL and Hammond D (1984). Treatment of Pneumocystis carinii pneumonitis, *Am J Dis Child*, 138(11): 1051-1054.

Smulian AG, Ryan M, Staben C and Cushion M (1996). Signal transduction in Pneumocystis carinii: characterization of the genes (pcg1) encoding the alpha subunit of the G protein (PCG1) of Pneumocystis carinii carinii and Pneumocystis carinii ratti, *Infect Immun*, 64(3): 691-701.

Song CM, Lim SJ and Tong JC (2009). Recent advances in computer-aided drug design, *Brief Bioinform*, 10(5): 579-591.

Stahl M, Guba W and Kansy M (2006). Integrating molecular design resources within modern drug discovery research: the Roche experience, *Drug Discovery Today*, 11(7–8): 326-333.

Stringer JR (1996). *Pneumocystis carinii*: what is it, exactly?, *Clin Microbiol Rev*, 9(4): 489-498.

Taylor RD, Jewsbury PJ and Essex JW (2002). A review of protein-small molecule docking methods, *Journal of Computer-Aided Molecular Design*, 16(3): 151–166.

Thomas CF and Limper AH (2004). Pneumocystis Pneumonia, *N Engl J Med*, 350(24): 2487-2498.

Thomas CF, Leof EB and Limper AH (1999). Analysis of *Pneumocystis carinii* introns, *Infect Immun*, 67(11): 6157–6160.

Tidwell RR, Jones SK, Geratz JD, Ohemeng KA, Cory M and Hall JE (1990). Analogues of 1, 5-bis(4-amidinophenoxy) pentane (pentamidine) in the treatment of experimental Pneumocystis cannii pneumonia, *J Med Chem*, 33(4): 1252-1257.

Trott O and Olson AJ (2010). AutoDockVina: Improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading, *J Comput Chem*, 31(2): 455-461.

Tsuji M (2010), "Homology Modeling for HyperChem", Revision F1; Saitama,

JAPAN, *Brief Bioinform,* 3(1):73-86.

Ulrich P and Cerami A (1984). Trypanocidal 1, 3 arylenediketonebis (guanylhydrazone)s. Structure-activity relationships among substituted and heterocyclic analogues, *J Med Chem*, 27(1): 35-40.

Vale N, Moreira R, Gomes P (2009). Primaquine revisited six decades after its discovery, *Eur J Med Chem*, 44(3):937-953.

Vanden Eynde JJ, Mayence A, Huang TL, Collins MS, Rebholz S, Walzer PD, and Cushion MT (2004). Novel bisbenzamidines as potential drug candidates for the treatment of Pneumocystis carinii pneumonia, *Bioorg Med Chem Lett*, 14(17): 4545-4548.

Vaque M, Arola A, Aliagas C and Pujadas G (2006). BDT: an easy-to-use front-end application for automation of massive docking tasks and complex docking strategies with AutoDock, *Bioinformatics*, 22(14): 1803-1804.

Vedani A and Dobler M (2000). Multi-dimensional QSAR in drug research. Predicting binding affinities, toxicity and pharmacokinetic parameters, *Prog Drug Res*, 55:105-135.

Verdonk ML, Cole JC, Hartshorn MJ, Murray CW and Taylor RD (2003). Improved protein–ligand docking using GOLD, *Proteins*, 52(4): 609-623.

Vohra PK, Puri V, Kottom TJ, Limper AH and Thomas CF (2003). Pneumocystis carinii STE11, an HMG-box protein, is phosphorylated by the mitogen activated protein kinase PCM, *Gene*, 312:173-179.

Volpe F, Ballantine SP and Delves CJ (1993). The multifunctional folic acid synthesis fasgene of *Pneumocystis carinii* encodes dihydroneopterinaldolase, hydroxymethyldihydropterin pyrophosphokinase and dihydropteroate synthase,

Eur J Biochem, 216(2): 449-458.

Walzer PD, Foy J, Steele P and White M (1992). Treatment of Experimental Pneumocystosis: Review of 7 Years of Experience and Development of a New System for Classifying Antimicrobial Drugs, *Antimicrobial Agents And Chemotherapy*, 36(9):1943-1950.

Walzer PD, Kim CK and Foy J (1991). Furazolidone and nitrofurantoin in the treatment of experimental Pneumocystis carinii pneumonia, *Antimicrob Agents Chemother*, 35(1): 158-163.

Walzer PD, Kim CK, Foy JM, Linke MJ and Cushion MT (1987). Inhibitors of folic acid synthesis in the treatment of experimental Pneumocystis carinii Pneumonia, *Antimicrob Agents Chemother*, 32(1) :96-103.

Weinstein DG and Frost P (1971). Methotrexate for psoriasis, *Archives of Dermatology,* 103(1): 33-38.

Wermuth CG and Langer T (1993). Pharmacophore identification. In 3D QSAR in Drug Design. Theory, Methods and Applications (Kubinyi, H., ed.), pp. 117–136.

Wermuth CG et al. (1998). Glossary of terms used in medicinal chemistry (IUPAC Recommendations 1997), *Annu Rep Med Chem*, 33: 385-395.

Whisnant KJ and Buckley HR (1976). Successful pyrimethamine-sulfadiazine therapy of Pneumocystis pneumonia in infantswith x-linked immunodeficiency with hyper-IgM, *Natl Cancer Inst Monogr*, 43: 211-216.

Wiese M and Pajeva IK (2001). Structure-activity relationships of multidrug resistance reversers, *Curr Med Chem*, 8(6): 685-713.

Winkler DA (1998). The role of quantitative structure--activity relationships (QSAR) in biomolecular discovery, *J Med Chem*, 41(18): 3314-3324.

Winston JD, Lau KW, Gale PR, and Young SL (1980). Trimethoprim-sulfamethoxazole for treatment of Pneumocystis carinii infection, *Ann Intern Med,* 92(2): 762-769.

Wolber G and Kosara R (2006) Pharmacophores from macromolecular complexes with LigandScout. In Pharmacophores and Pharmacophore Searches, (Langer, T. and Hoffmann, R.D., eds), Wiley-VCH, 32: pp131–150.

Wolber G and Langer T (2005). LigandScout: 3D Pharmacophores derived from protein-bound ligands and their use as virtual screening filters, *J Chem Inf Model,* 45(1): 160-169.

Wolber G, Seidel T, Bendix F and Langer T (2008). Molecule-pharmacophore superpositioning and pattern matching in computational drug design, *Drug Discovery Today*, 13(1-2): 23-29.

Yale SH and Limper AH (1996), Pneumocystis carinii pneumonia in patients without acquired immunodeficiency syndrome: associated illness and prior corticosteroid therapy, *Mayo Clin Proc*, 71(1): 5-13.

Yamada M, Takeuchi S, Shiota T, Matsumoto Y, Yoshikawa H, Okabayashi K, Tegoshi T, Yoshikawa T and Yoshida Y (1985). Experimental studies on the chemoprophylaxisfor Pneumocystis carinii pneumonia with intermittent administration of trimethoprim-sulfamethoxazole and pyrimethamine sulfamonomethoxine, *Jpn J Trop Med Hyg*, 13: 287-294.

Yang Z and Sun P (2007). 3D-QSAR study of potent inhibitors of phosphodiesterase-4 using a CoMFA Approach, *Int J Mol Sci*, 8:714–722.

Yoshida Y, Takeuchi S, Ogino K, Ikai T and Yamada M (1977). Studies on Pneumocystis carinii and Pneumocystis carinii pneumonia. III. Therapeutic

experiment of the pneumonia with pyrimethamine + sulfamonomethoxine and trimethoprim + sulfamethoxazole, *Jpn J Parasitol*, 26: 367-375.

Young CR and DeVita TV (1976). Treatment of Pneumocystis carinii pneumonia: current status of the regimens of pentamidine isethionate and pyrimethamine-sulfadiazine, *NatI Cancer Inst Monogr*, 43: 193-198.

Zielesny A (2005). Chemistry Software Package ChemOffice Ultra 2005, *J Chem Inf Model*, 1474–1477.