Development of inhibitors to target glyoxylate and methylcitrate cycles essential for persistence of *Mycobacterium tuberculosis*

Submitted By

URVASHI

122501

Under the guidance of

DR. CHITTARANJAN ROUT

Thesis Submitted in Partial Fulfilment of the Degree of

MASTERS OF TECHNOLOGY

in

COMPUTATIONAL BIOLOGY

to the

DEPARTMENT OF BIOTECHNOLOGY AND BIOINFORMATICS

JAYPEE UNIVERSITY OF INFORMATION TECHNOLOGY

WAKNAGHAT, SOLAN (H.P.)

MAY-2014

CERTIFICATE

This is to certify that the work titled, **"Development of inhibitors to target glyoxylate and methylcitrate cycles essential for persistence of *Mycobacterium tuberculosis*"** submitted by **Urvashi** in partial fulfilment for the award of degree of Masters of Technology in Computational Biology of Jaypee University of Information Technology has been carried out under our supervision. This work has not been submitted partially or wholly to any other University or Institute for the award of this or any other degree or diploma.

24/05/14

Dr.Chittaranjan Rout

Assistant professor

Dept. of Biotechnology and Bioinformatics

Jaypee University of Information Technology

Waknaghat, Solan (H. P.)-173234

ACKNOWLEDGEMENT

With the grace and will of almighty it is my privilege to take the opportunity to thank all those who have directly or indirectly helped me in the completion of this project work.

I wish to express deep regards and gratitude to**Dr.Chittaranjan Rout**, Department of Biotechnology & Bioinformatics, Jaypee University of Information Technology, Waknaghat, Solan for their immense help and valuable guidance, constant encouragement and inspiration throughout the progress of this work.

I would like to express my deep sense of gratitude to Prof. R.S. Chauhan, Head of Department of Biotechnology & Bioinformatics for his constant encouragement and suggestions for motivating us to put in our best effort.

I am also thankful to all other faculty members of bioinformatics department for their support and time to time guidance.

I will feel incomplete if I do not express my thanks to Mrs. Somlata, Miss. Charu & Miss. Nupur, who always helped me in my whole project work.

My words fail to express my humble gratitude and profound regards to my parents and my brother for their affectionate encouragement, cooperation and blessings throughout my life, which always remained a source of inspiration for me.

Last but not least I thank all my classmates, friends without whose advices and love I would not have been completed this project.

I wish to gratefully acknowledge the library staff and several authors of various text books and research papers which have been referred in this study.

Place: Waknaghat

Urvashi

Date: 24.05.14

Urvashi

ABSTRACT

Mycobacterium tuberculosis is one of the most significant pathogen since its discovery in 1882. It causes one of the most deadly disease TB. It can occur in two stages in the body of the host cell, namely active stage and the persistent stage. Active stage in which the pathogen is active and damages the host system whereas in persistent stage, the pathogen is inactive and doesn't cause any damage to the host system but remain dormant in the host system and becomes active whenever the immunity of the host is compromised. So, it becomes important to discover drugs which not only kills the pathogen in the active state but also in the persistent stage.

Our target is GLYOXYLATE and METHYLCITRATE CYCLES which play a pivotal role in both the active and the persistent stage of the pathogen, as it catalyses an alternative pathway i.e. glyoxylate shunt pathway in which pathogen adopts an alternative pathway for quickly generating ATP whenever the pathogen is under stress. Through literature and our docking studies, we found that the active site of the target is small and only inhibitors with M.WT<200 are able to dock in the active site of target. This poses a challenge in the drug discovery process against this particular target. Lot of *in vitro* and *in vivo* studies has been done but none of them has reached the market, this raises the need to develop more inhibitors against ICL.

Our study included the docking of various ligand libraries in the active site of the target protein, which included manually designed ligands and ligands from various databases. Docking score validation by CDOCK, ADME/TOX prediction and studying the interactions of the top ranking inhibitors by LIGPLUS.

Our results show some promising drug candidates which are showing good interactions with the conserved residues in the active site and good ADME/TOX properties and we further propose these lead molecules for chemical synthesis and activity prediction (*in vitro* and *in vivo*) against *M. tuberculosis* ICL.

LIST OF FIGURES

LIST OF TABLES

LIST OF TOOLS, DATABASES AND SOFTWARES USED

- Discovery Studio 3.5
- Pymol
- VMD (*Visual Molecular Dynamics*)
- ClustalW: http://www.ebi.ac.uk/Tools/msa/clustalw2/
- ZINC Database: zinc.docking.org
- Saves Server: http://nihserver.mbi.ucla.edu/SAVES/
- Symyx Draw 4.0
- Argus Lab
- LIGPLUS
- ZINCPharmer
- GROMACS

LIST OF ABBREVIATIONS

MTB-*Mycobacterium tuberculosis*

ICL-Isocitrate lyase

TB-Tuberculosis

ZINC-Zinc Is Not Commercial

ADMET- Absorption, Distribution, Metabolism and Excretion/T oxicity

HIV-Human Immunodeficiency Virus

AIDS- Acquired ImmunoDeficiency Syndrome

CONTENTS

CHAPTER 1

INTRODUCTION

1.1 GENERAL

Tuberculosis (TB) is a disease caused by a bacterium called *Mycobacterium tuberculosis*. The bacteria usually attack the lungs, but TB bacteria can attack any part of the body such as the kidney, spine, brain, etc. If not treated properly, TB disease can be fatal. TB can also pass from a mother to her unborn child before and after birth. It is more common in children and people with weak immune systems, including people living with HIV.

Mycobacterium tuberculosis is a highly successful pathogen that hides in the macrophages of its host [1]. Its success can be attributed directly to its ability to manipulate the phagosome that it resides in and to prevent the normal maturation of this organelle into an acidic & hydrolytic compartment. As the macrophage is key to clearing the infection, the interplay between the pathogen and its host cell reflects a constant battle for control.

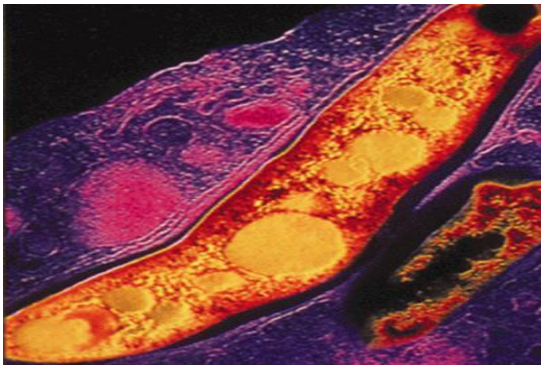


Fig. 1: The *Mycobacterium tuberculosis* strain

1.2 *Mycobacterium tuberculosis* BACTERIA

Mycobacterium tuberculosis is the etiologic agent of tuberculosis in humans. It's a fairly large non motile rod shaped bacterium distantly related to the actinomycetes. The rods are 2-4 aerometers in length and 0.2-0.5 µm in width. It is an obligate aerobe. For that reason, *M. tuberculosis* complexes are always found in the well aerated upper lobes of the lungs. The

bacterium is a facultative intracellular parasite, usually of macrophages and has a slow generation time, 15-20 hours, and a physiological characteristic that may contribute to its virulence. *M. tuberculosis* is not classified as either gram positive or gram negative because it does not have chemical characteristics of either, although the bacteria do contain peptidoglycan in their cell wall. If a gram stain is performed on *M. tuberculosis*, it stains very weakly gram positive or not at all.

1.3 GENETICS OF *M. tuberculosis*

The complete genome of *M. tuberculosis* was sequenced in 1998. Tuberculosis structure consortium has a collection of structures of over 400 proteins from *M. tuberculosis* many review articles and publications have analyzed these structures in the context of functional information [2].

1.4 TYPES OF TUBERCULOSIS

Persons with latent TB do not feel sick and do not have any symptoms, but usually have a positive reaction to the tuberculin skin test. Latent TB infections are not infectious and cannot spread TB infection to others. In some people, TB bacteria overcome the defense of the immune system and begin to multiply, resulting in the progression from latent to active infection.

Table1.1: The difference between latent TB infection and TB disease	
A person with latent TB infection	**A person with TB disease**
• Has no symptoms	• Has symptoms that may include: - a bad cough that lasts 3 weeks or longer - pain in the chest - coughing up blood or sputum -weakness or fatigue -weight loss -no appetite -chills -fever - sweating at night
• Does not feel sick	• Usually feels sick
• Cannot spread TB bacteria to others	• May spread TB bacteria to others
• Usually has a skin test or blood test result indicating TB infection	• Usually has a skin test or blood test result indicating TB infection
• Has a normal chest x-ray and a negative sputum smear	• May have an abnormal chest x-ray, or positive sputum smear or culture
• Needs treatment for latent TB infection to prevent active TB disease	• Needs treatment to treat active TB disease

Source: http://www.cdc.gov/TB/topic/basics/default.htm

1.5 HISTORY OF INFECTION

The natural history of *M. tuberculosis* infection follows these 5 stages [3]:

Stage 1: The droplet of nuclei is inhaled. One droplet nuclei contains more than 3 bacilli. These droplet nuclei reach the alveoli where infection begins.

Stage 2: It begins 7-21 days after the initial infection. *M. tuberculos*is multiplies virtually unrestricted within the inactivated macrophages until the macrophages burst.

Stage 3: At this stage lymphocytes begin to infiltrate the lymphocytes specifically T-cells recognize, processed and presented *M. tuberculosis*. The liberation of IFN causes in the activation of macrophages. These activated macrophages are now capable of destroying *M. tuberculosis*. Extracellular bacteria are resistant to complement killing due to the high lipid concentration in its cell wall [4].

Stage 4: Although many activated macrophages can be found surrounding the tubercles, these macrophages replicate and hence the tubercle grows.

Stage 5: For unknown reasons, the gaseous centers of the tubercles liquefy. This liquid is very conductive to *M. tuberculosis* growth and hence the organism begins to rapidly multiply extracellular. During the initial infection disease is most readily transmitted.

1.6 HOW DOES TB SPREAD

(http://www.cdc.gov/TB/topic/basics/default.htm)

TB is spread through the air from one person to another. The TB bacteria are put into the air when a person with active TB disease of the lungs or throat coughs, sneezes, speaks, or sings. People nearby may breathe in these bacteria and become infected.

1.7 SYMPTOMS OF TB

The symptoms of TB include:

- a persistent cough for more than three weeks that brings up phlegm, which may be bloody
- weight loss
- night sweats
- high temperature (fever)
- tiredness and fatigue
- loss of appetite

1.8 CURRENT WORLD SCENARIO

Tuberculosis is second only to HIV/AIDS as the greatest killer worldwide due to a single infectious agent. In 2012, 8.6 million people fell ill with TB and 1.3 million died from TB. In 2012, an estimated 530 000 children became ill with TB and 74 000 HIV-negative children died of TB. [5]

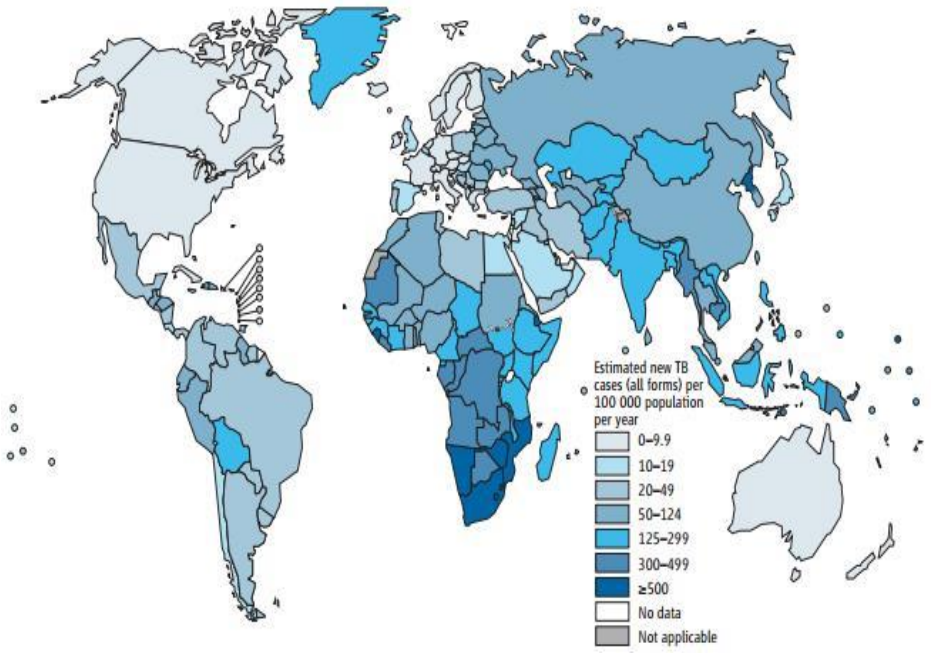


Fig 1.2: World population suffering from TB

1.9 TB AND HIV

(http://www.uphs.upenn.edu/bugdrug/antibiotic_manual/TBHIVclinicalmanual.pdf)

For many people, initial HIV symptoms are not clearly visible. But when they go for TB diagnosis, they realize that they are also suffering from HIV which probably increases susceptibility to infection with *M. tuberculosis*. HIV increases the risk of progression of *M. tuberculosis* infection to TB disease. This risk increases with increasing immunosuppression. HIV infection also interferes with TB diagnosis. HIV increases not only the risk but also the rate of progression of recent or latent *M. tuberculosis* infection to disease.

TB in the course of HIV progression

TB can occur at any point in the course of progression of HIV infection. The risk of developing TB rises sharply with worsening immune system.

Consequences of HIV/*M. tuberculosis* co-infection

Compared with an individual who is not affected with HIV, a person infected with HIV has a 10 times increased risk of developing TB. TB notifications have increased in populations where both HIV infection and *M. tuberculosis* infection are common. For example, some parts of sub Saharan Africa have seen a 3-5 fold increase in the number of TB case.

Impact of HIV on TB control

The principles of TB control are the same even when there are many HIV/TB patients. However, in populations where HIV/TB is common, health services struggle to cope with large and rising numbers of TB patients.

Practical Point

HIV is the most powerful factor known to increase the risk of TB. The consequences include the following:

1. overdiagnosis of sputum smear-negative pTB (due to difficult in diagnosis)

2. underdiagnosis of sputum smear-positive pTB (due to excess laboratory workload)
3. inadequate supervision of anti-TB chemotherapy
4. low cure rates
    - high morbidity during treatment
    - high mortality rates during treatment
    - high default rates because of adverse drug reactions
    - high rates of TB recurrence
    - increased transmission of drug

1.10 TESTING FOR THE DISEASE

A skin test is used to detect TB infection. A small amount of liquid (ppd tuberculin) is injected under the skin of the arm. A hard swelling at the site larger than 10mm means the patient in infected with TB. A smaller swelling of 5mm used to detect TB infection in people with HIV. Chest X-rays are used to detect active TB disease and to check for damage in the lungs. In a TB smear test, a sample of sputum is studied under microscope to see if it contains the bacteria [6].

1.11 THERAPY (http://en.wikipedia.org/wiki/tuberculosistreatment)

The standard "short" course treatment for TB is isoniazid, rifampicin(also known as rifampin in the United States), pyrazinamide, and ethambutol for two months, then isoniazid and rifampicin alone for a further four months. The patient is considered cured at six months (although there is still a relapse rate of 2 to 3%). For latent tuberculosis, the standard treatment is six to nine months of isoniazid alone.

First line

All first-line anti-tuberculosis drug names have a standard three-letter and a single-letter abbreviation:

- Ethambutol as emb or e,
- Isoniazid as inh or h,
- Pyrazinamide as pza or z,
- Rifampicin as rmp or r,

- Streptomycin as stm or s.

US commonly uses abbreviations and names that are not internationally recognized such as rifampicin is called rifampin and abbreviated as rif and streptomycin is commonly abbreviated sm.

Most regimens have an initial high-intensity phase, followed by a continuation phase (also called a consolidation phase or eradication phase): the high-intensity phase is given first, then the continuation phase, the two phases divided by a slash which means isoniazid, rifampicin, ethambutol, pyrazinamide daily for two months, followed by four months of isoniazid and rifampicin given three times a week.

Second line

The six classes of second-line drugs used for the treatment of TB are:

1 aminoglycosides: e.g.amikacin (amk), kanamycin (km)
2 polypeptides: e.g. capreomycin, viomycin, enviomycin
3 fluoroquinolones: e.g.ciprofloxacin (cip), levofloxacin, moxifloxacin (mxf)
4 thioamides: e.g. ethionamide, prothionamide
5 cycloserine (the only antibiotic in its class)
6 *p*-aminosalicylic acid (pas or p).

A drug may be classed as second-line instead of first-line because of these possible reasons:

a. it may be less effective than the first-line drugs (e.g. *p*-aminosalicylic acid)

b. it may have toxic side-effects (e.g. cycloserine)

c. it may be unavailable in many developing countries (e.g., fluoroquinolones)

1.12 ISOCITRATE LYASE

Isocitrate lyase (ICL) plays a pivotal role in the persistence of *Mycobacterium tuberculosis*. The enzyme allows net carbon gain by diverting acetyl-CoA from β-oxidation of fatty acids into the glyoxylate shunt pathway [7].

Bacteria and some species of higher plants are able to obtain a net increase in malate or oxaloacetate through expression of enzymes of the glyoxylate cycle or glyoxylate shunt. The two additional enzymes that permit the glyoxylate shunt are *isocitrate lyase* and *malate synthase,* which convert isocitrate to succinate or to malate via glyoxylate under construction.

Isocitrate lyase (ICL) catalyses the first step of the glyoxylate bypass pathway, which reversibly catalysis isocitrate into succinate and glyoxylate. In normal condition, isocitrate is converted into keto-glutarate through *isocitrate dehydrogenase* enzyme, but when the enzyme is under nutrient stress condition, it follows an alternative pathway and isocitrate is converted into glyoxylate by enzyme, is *citrate lyase*. It is the key enzyme in the glyoxylate bypass pathway. Glyoxylate shunt was reported to be pivotal for microbes to survive and thrive on the fatty acid and two-carbon compound as sole carbon source ICL existed extensively in organisms such as bacteria, nematodes, fungi and plants. No human being homologs have been documented. This made ICL a promising target for broad spectrum antibiotics.

The net reaction for the above explained cycle is as follows:-

2 acetyl-CoA + NAD^+ → succinate + 2 CoA + NADH + H^+

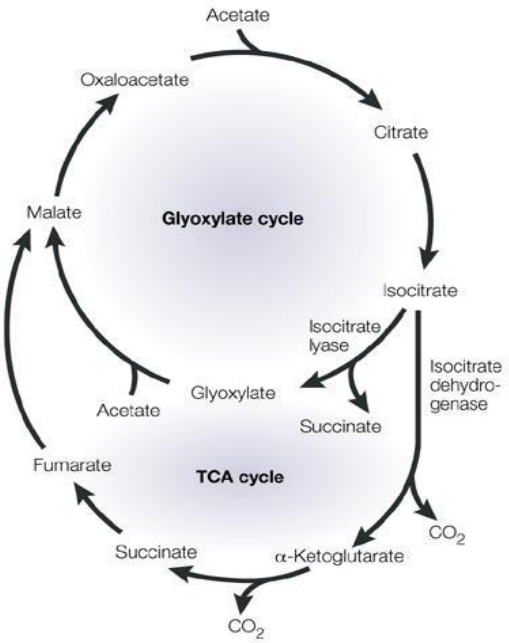


Fig 1.2: Diagram showing glyoxylate shunt pathway and TCA cycle. Isocitrate lyase helps bacteria to follow shunt pathway in the absence of nutrient supply by thriving on the fatty acid and two-carbon compound as sole carbon source. Source: Nature

1.13 ROLE IN PERSISTENT STAGE [8]

The ability of *Mycobacterium tuberculosis* to persist in its human host despite extensive chemotherapy is thought to be based on subpopulations of Non-replicating phenotypically drug-resistant bacilli. Nutrient-starved bacilli lacking the glyoxylate shunt enzyme *isocitrate lyase* failed to reduce their intracellular ATP level and died, thus establishing a link between ATP control and intermediary metabolism. This conclude that ICL is plays an important role in the persistent stage of MTB.

1.14 ICL AS A DRUG TARGET

ICL is an enzyme of glyoxylate cycle that allows bacteria to grow using acetate, propionate or certain fatty acids as a carbon source, the substrates presumably available to MTB in granulomas. The genome of MTB (H37Rv) strain has three genes encoding for ICL:(i) Rv0467 also known as ICL1 that is closely related to isocitrate lyases in other eubacteria, (ii) Rv1915 and(iii) Rv1916 labeled aceAa and AceAb respectively.ICL1 is more active than ICL2.In *M. tuberculosis*, ICL also performs MCL activity with additional methyl group of methyl-isocitrate lyase as compared to isocitrate is accommodated in a hydrophobic depression formed by three residues: Thr 347, Trp283 and Phe345 of ICL1 from MTB in place of Pro236, Leu234 and Phe186 of MCL from *E. coli*.[9]

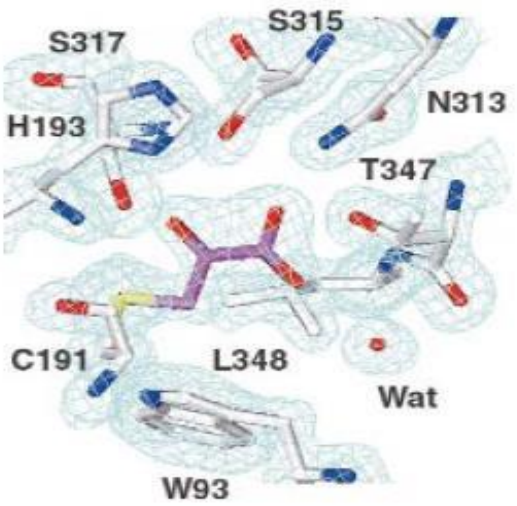


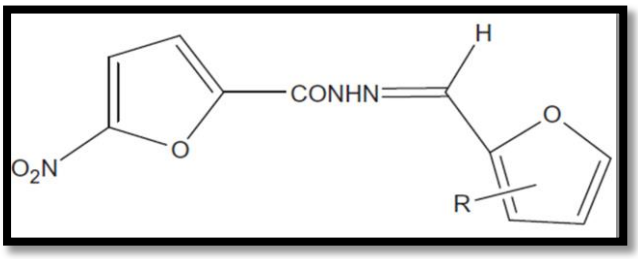
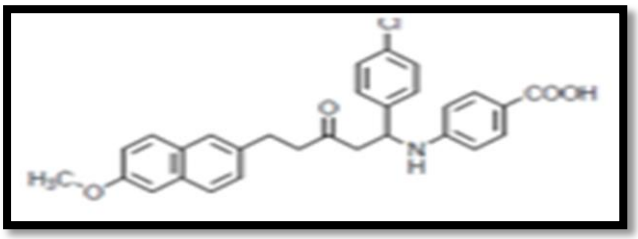
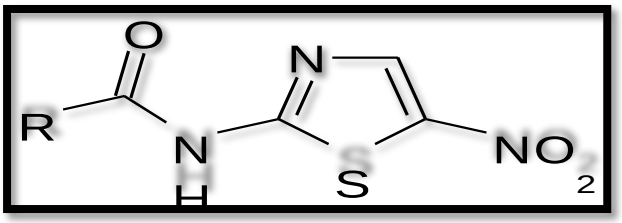
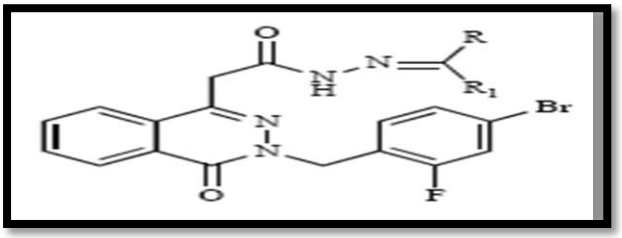
Fig 1.3: Active site of ICL

1.15 CHALLENGES AGAINST ICL AS A DRUG TARGET

ICL is a computationally challenging target because of small size of the active site. It limits the scope of inhibitor design and only inhibitors with M.WT. < 200 are able to dock properly in the active site.

1.16 NEED OF INHIBITORS

- There is a dire need to intensify our fight against this deadly disease by promoting rapid diagnosis and development of new vaccines & drugs against all forms of MTB.
- The challenge is to eliminate the persistent organism as most of the lengthy and costly standard drug treatment is just to get rid of the roughly 1% of bacteria that persist after the first week of a patient's treatment.
- Novel active molecules should shorten the duration of therapy, prevent resistance development and eliminate latent disease.

Table 1.2: Known ICL inhibitors		
Name of Compound	**Mol.Wt.**	**Experimental System**
Hydrazone derivative 	280	MICs of 2.65 and 10.64 µM respectively against log- and starved-phase culture of MTB
Ydmc67 	487	IC50 values of this molecule to ICL was 0.0535 mg/ml *in vitro* enzyme assay
R=5-Nitrofuran-2-yl 	284	*In vitro* with MICs of 5.48µM against log-phase culture of MTB
(R=H, R_1=5-Nitrofuran-2-yl) 	518	*In vitro* with MIC's of 0.18and <0.09 µM against log-phase cultures of MTB and MDR MTB

1.17 LIMITATIONS OF INHIBITORS

Though few inhibitors of ICLs have been reported against MTB, however, none has been developed as an anti-mycobacterial drug probably because of being:neurotoxic, cytotoxic or having higher MICs.

1.18 PROBLEM STATEMENT

- Design molecules which have broad-specificity: Inhibition of ICL1, ICL2 & Active site of ICL with methyl-isocitrate binding specificity.
- MTB ICL is having two categories of structures: Apo- and closed-structure. The active form is in between both. The primary challenge is to identify active form.
- Docking results also corroborated that active form is in between both.
- Docking ligands with active form to get correlation between docking score and IC50 values.

1.19 OBJECTIVE OF THE PROJECT

M. tuberculosis infects about 32%of the world's population [10].Very less new compounds may show suitable biological activities in the laboratories and none of these compounds result in successful clinical trials and reaches the market place. Even many antibacterial molecules are available in market but none of the drugs are targeting ICL. Since noncompliance to therapy is due to longer period of treatment which is again arising due to persistence of MTB. Targeting to ICL, a persistent target, through drug molecules is expected to reduce the treatment time of the disease which may ameliorate the problem of MDR and XDR TB. Therefore a drug against ICL can be used as combination therapy with other anti *M. tuberculosis* drugs to reduce the treatment period.

Our project aims at determination of potential lead small molecules for inhibiting *Mycobacterium tuberculosis* isocitrate lyase (ICL1/ICL2) and MCL by employing computational methods (virtual screening, docking and simulation techniques).

- Comparative analysis of active sites of ICL1 and ICL2.

- Common fragments in inhibitor of ICL targeting three different forms.
- Identifying active form.
- Broad-range specific inhibitors.
- Designing of inhibitors & their analysis.
- Ligand library docking.

CHAPTER 2

METHODS & METHDOLOGY

2.1 *IN SILICO* HOMOLOGY MODELLING FOR ICL2

Crystal structure of the protein (PDB ID: 1F8I) was downloaded from protein data bank. (http://www.pdb.org/pdb/results/results.do?outformat=&qrid=b7e953ca&tabtoshow=current) . Though many structures of ICL are available in PDB but the above structure was preferred as it is a high resolution enzyme-inhibitor complex. For homology modeling of ICL2, PDB ID: 1DQU was taken as a template.

Step 1: Sequence alignment

In order to construct a homology model, a sequence alignment must be done. However, prior to aligning the two sequences, the template structure was cleaned. Cleaning of the template structure was done using the Discovery Studio 3.5. The alignment was done using Align Sequences to Templates protocol in Create Homology Models under Macromolecules utilizes which uses Accelry's Align123 program.

Step 2: Building of Homology Model

Now we had a template structure and a good sequence alignment. The next step was to build a series of homology models and the model with highest quality was selected. For doing this, Create Homology Model Protocol based on MODELER was used in Discover Studio.

Step 3: Analysis of Homology Model

Once the structure was modeled, it was validated using Saves Server (http://nihserver.mbi.ucla.edu/SAVES/) and Ramachandran plot, Verify 3D and Errat Score were calculated.

Now structure was further refined using GROMACS. A 5ns MD simulation was done using OPLS-AA/L as a force field. It included the following steps:

i. pdb2gmx –f modelled.pdb –o conf.pdb
ii. editconf –f conf.peb –o box.pdb –d 1.2

iii. genbox –cp box.pdb –csspc 216 –o water.pdb –p topol.top
iv. grompp -f em.mdp -c water.pdb -p topol.top -o em.tpr -maxwarn 2
v. mdrun -v -s em.tpr -c em.pdb
vi. genion -s em.tpr -o ions.pdb -np 31
vii. grompp -f em.mdp -c ions.pdb -p topol.top -o em.tpr -maxwarn 2
viii. mdrun -v -s em.tpr -c em.pdb
ix. grompp -f pr.mdp -c em.pdb -p topol.top -o pr.tpr -maxwarn 2
x. mdrun -v -s pr.tpr -c pr.pdb
xi. grompp -f md.mdp -c pr.pdb -p topol.top -o md.tpr -maxwarn 2
xii. mdrun -v -s md.tpr -c md.pdb

After the MD simulation, the structure was again verified using Saves Server (http://nihserver.mbi.ucla.edu/SAVES/) and Ramachandran plot, Verify 3D and Errat Score were calculated.

2.2 PREPARATION OF PROTEIN

Crystal structure of the proteins (PDB IDs: 1F8I, 1F61 & 1DQU) were downloaded from protein data bank. Preparation of protein is very first and important step in docking study. Protein X-ray crystallographic structures present in the PDB does not contain all structural information to perform proper docking studies. This information is hydrogen atom, protonation states, tautomers, partial charges, etc. This was done using protein preparation wizard in Discovery Studio and steps included the following jobs:

- Standardize atom names, insert missing atoms in residues and remove alternate conformations.
- Insert missing loop regions.
- Optimize short and medium size loop regions with the LOOPER algorithm.
- Minimize the remaining loop regions.
- Calculate the pK and protonate the structure.

The X-ray structure (1F8I) of the ICL is a 4-mer (A, B, C & D and chains) formed by same protein chain. Active site of protein formed by two dimeric symmetric structures: i. Chains A & B and ii. C & D. Isocitrate binding site is formed by residues from two chains (either A &

B or C & D) so chains A & B have been considered for our docking studies and taking all the chains was unnecessary. All the water atoms were deleted and succinate was also deleted for our studies but glyoxylate was kept as since it was a cofactor.

2.3 PREPARATION OF LEAD COMPOUNDS

Before docking the ligand molecules, ligand library was prepared by Discovery Studio 'Prepare Ligands' module. In this step, all ligand were prepared and for docking studies ligand should be in 3D formats. The following processes were carried out in stepwise manner in 'Prepare Ligands':

- 2D to 3D conversion.
- Enumerate ionization states, tautomers and isomers.
- Multiple rapid and exhaustive conformational generation methods.
- Molecular properties and fingerprints for filtering.
- Filter poor candidates with undesirable functional groups and Lipinski and Veber rules.
- Enumerate libraries for screening using reactions or core and R-groups.

Optimize geometries: The geometry optimization was carried out in ArgusLab by using Hamiltonian's Quantum Mechanics(QM) Parametric Method 3 (PM3) and the Hartree-Fock Self –Consistent-Field (HC SCF) method which is a *ab initio* approach to solving the Schroedinger equation for molecules.

Minimize Ligands: The minimization of ligands was performed by Full Minimization using CHARMm in Discovery Studio.

Libraries of ligands were generated and were converted to .sd format.

Set1: Manually design of ligands of hydroquinone derivatives, pthalazinyl derivatives on the basis of residues present in active site to provide better interactions. Some groups of succinate were modified on the basis more delocalization of electronic distribution as it provide more stable ligands. All modifications provided 60 analogs and all the molecules were drawn and optimized (Symyx Draw).

2.4 ACTIVE SITE RESIDUES PREDICTION

With the help of literature (Ref) and manual visualization of the protein in VMD software (keeping 5 Å as the cut-off), active site residues were identified. All the residues mentioned in the literature and few other residues where we confirmed interactions were used as active site residues in the grid generation while doing docking process in Schrödinger's. In the literature, His 193,Asn 313,Ser 315,Ser 317,Thr 347,Asp 108,Asp 153,Glu 155,Glu 182 were found as important residues whose interactions are important for the inhibition of enzyme and apart from these, residues that were manually considered by us were His 393(B Chain), Gln 394(B Chain), Thr 422(B Chain), Leu 194,Trp 93,Ser 91,Ser 92,Ser 191,Lys 190,Glu 285,Gln 184,His 193,Glu 182,Gln 394,Asp 108,Lys 197,Arg 228,Phe 318,Ser 317,Pro 316,Cys 314,Leu 348,(those not specified are from A chain).

2.5 PREPARING A RECEPTOR BINDING SITE

The receptor binding site is important for docking and scoring applications. It defines a region in a protein cavity where binding interactions can occur. The binding site is defined by selecting active site residues using Define a Binding Site Sphere from a Selection module in Discovery Studio.

2.6 LIGAND DOCKING

LibDock

LibDock is an algorithm for docking small molecules into an active receptor site. Initially, a hotspot map is calculated for the receptor active site which contains polar and apolar groups. This hotspot map is subsequently used to rigidly align the ligand conformations to form favorable interactions. After a final energy-minimization step (allowing the ligand poses to be flexible), the top scoring ligand poses are saved.

The following steps are performed:

- Calculates ligand conformations

- Docks the conformations using LibDock
- Minimizes docked poses using CHARMm

All ligands selected from the ligand.sd file and were docked with the active site using the HTVS (high throughput virtual screening) only for fragments. LibDock was performed against structures having PDB IDs: 1F8I, 1F61, 1DQU and modeled ICL2 structure. Then with a certain cut off molecules were selected (depending on the source) which were subjected further to CDOCKER.

CDOCKER

CDOCKER is a powerful CHARMm-based docking method that generates highly accurate docked poses. CDOCKER was performed against structures having PDB IDs: 1F8I, 1F61, 1DQU and modeled ICL2 structure.

Then with a certain cut off molecules were selected (depending on the source) which were refined and optimized.

2.7 SIMILARITY SEARCH

ZincPharmer

ZINCPharmer (http://zincpharmer.csb.pitt.edu) is an online interface for searching the purchasable compounds of the ZINC database using the Pharmer pharmacophore search technology. A pharmacophore describes the spatial arrangement of the essential features of an interaction. Compound that match a well-defined pharmacophore serve as potential lead compounds for drug discovery. ZINCPharmer provides tools for constructing and refining pharmacophore hypotheses directly from molecular structure.[11]

2.8 ADMET SCREENING

ADMET refers to the Absorption, Distribution, Metabolism, Excretion, and Toxicity properties of a molecule within an organism. Optimizing these properties during early drug discovery is crucial for reducing ADMET problems later in the development process. ADMET descriptors allow you to eliminate compounds with unfavorable ADMET

characteristics early on to avoid expensive reformulation later, and to evaluate proposed structural refinements that are designed to improve ADMET properties.

ADMET using Discovery Studio

It will generate an ADMET plot which is a 2D chart of ADMET_PSA_2D versus ADMET_AlogP98.

FAF Drugs

The main goal is the computational prediction of some ADME-Tox properties *(adsorption, distribution, metabolism, excretion and toxicity)* in order to assists hit selection before chemical synthesis or ordering. FAF-*drugs* employ pre-defined filters, but user can also customize its own filtering parameters by using the *filter-editor service*.[12]

2.9 INTERACTION STUDIES

LIG PLUS

It is used to study the interaction between protein ligand complexes. It is available by PDBSum and its online tool generates 2D diagram of hydrogen and hydrophobic bonding interactions between protein and ligand. Also there are some limitations with this tool as it displays only 2D picture.

CHAPTER 3

RESULTS AND DISCUSSSION

3.1 IN SILICO HOMOLOGY MODELED ICL2 STRUCTURE

ICL2 structure was modeled using homology modeling in Discovery Studio. It was validated using Saves Server.

Superimposed Structure of Modeled ICL2 with template (PDB ID: 1DQU)

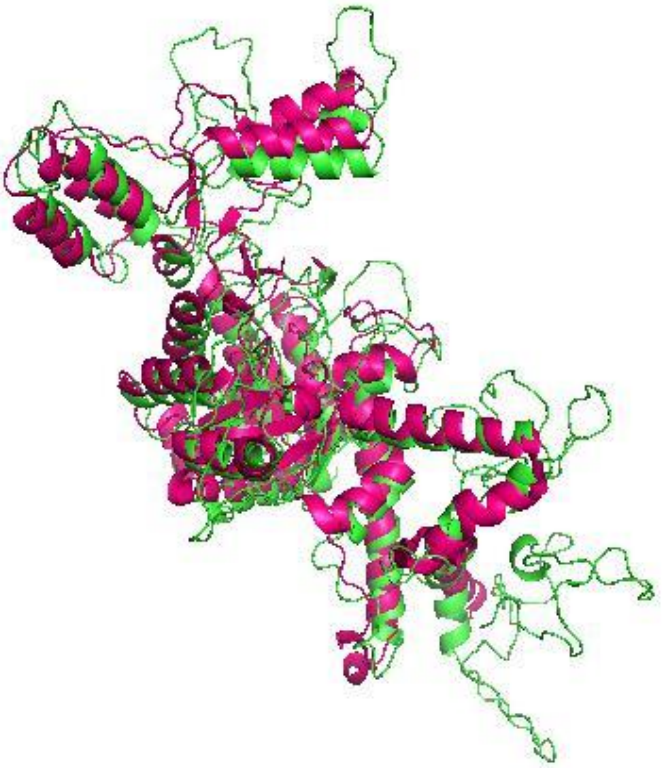


Fig. 3.1: Superimposed structure of ICL2 with 1 DQU

Before MD Simulation

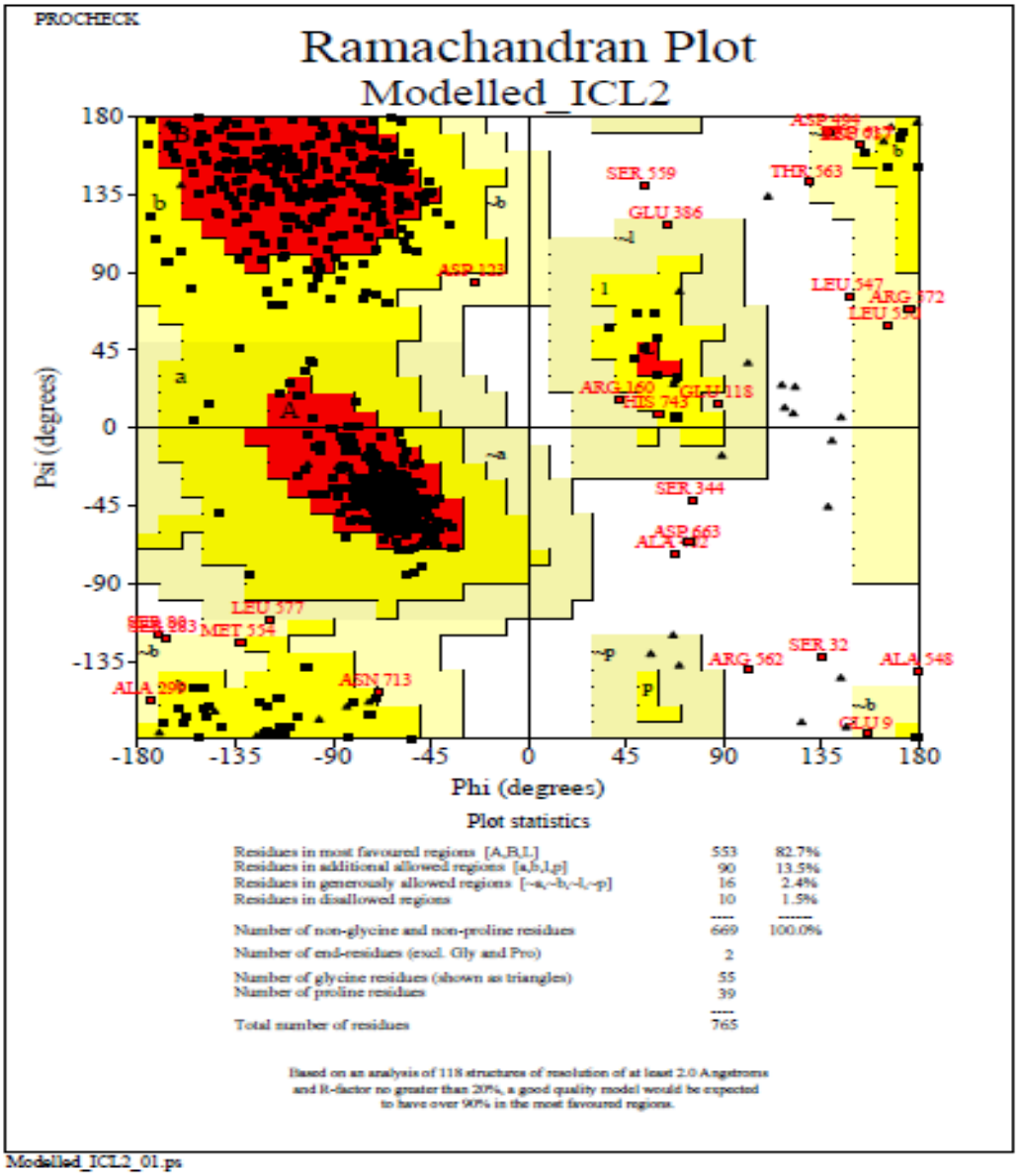


Fig. 3.2: Ramachandran plot showing 82.7% core 13.5% allow 2.4% gener 1.5% disallowed regions.

The Verify_3D was 53.39% of the residues had an averaged 3D-1D score > 0.2 and the Errat Score was 45.879 of the modeled score. The overall score of the modeled structure was very poor so, it was MD simulated using GROMACS and again validated using Saves Server.

After MD Simulation

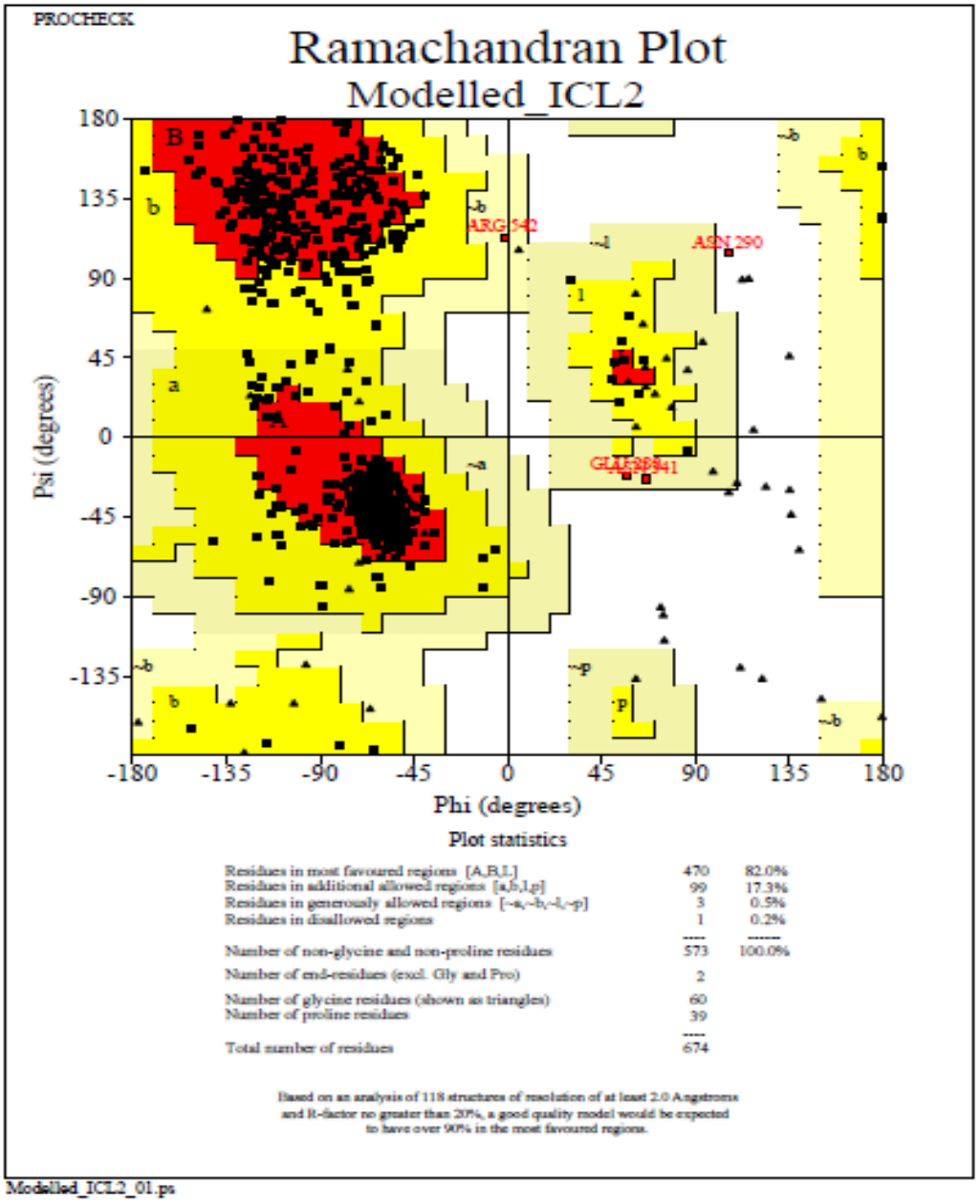


Fig. 3.3: Ramachandran plot: 82.0% core 17.3% allow 0.5%gener 0.2% disallowed region.

The Verify_3D score was 97.19% of the residues had an averaged 3D-1D score > 0.2 and Errat Score was 93.232 of the modeled structure after MD Simulation.

3.2 PREPARED PROTEIN TARGET STRUCTURE

Protein structures were downloaded from the PDB and were subjected to protein minimization wizard in Discovery Studio.

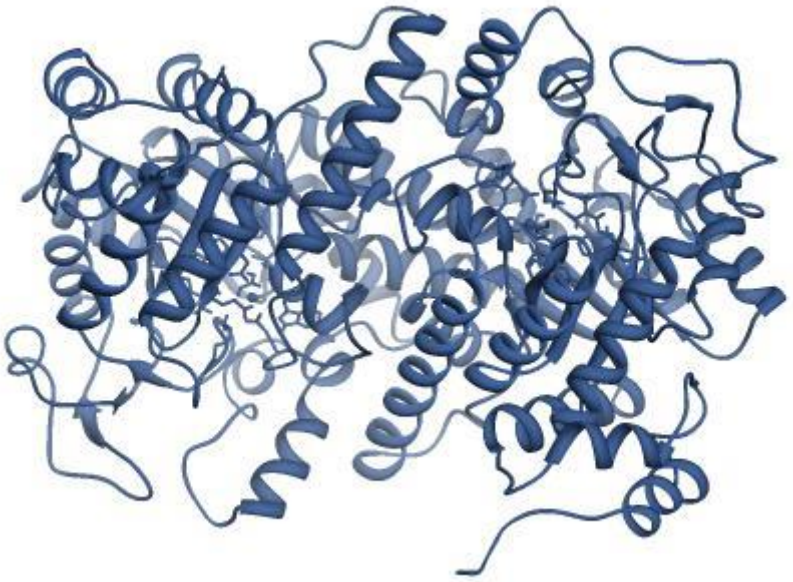


Fig. 3.4: Minimized structure of 1F8I consisting of A & B chains.

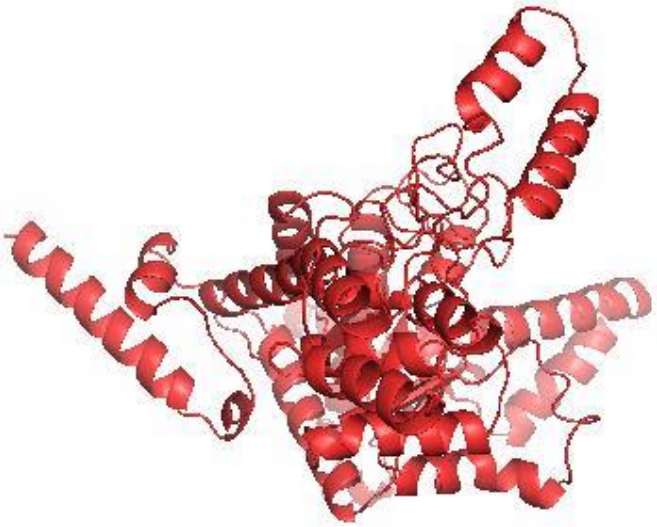


Fig 3.5: Minimized structure of 1DQU

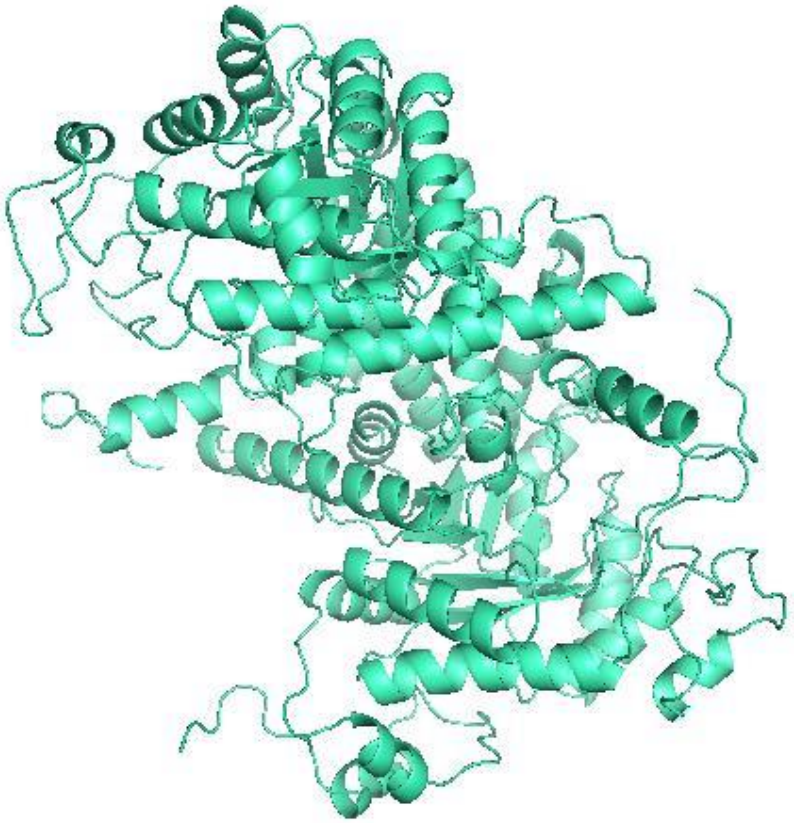


Fig 3.6: Minimized structure of 1F61.

3.3 DOCKING RESULTS

Source 1: Manually Designed Ligands

Manual design of ligands: By analyzing various electrochemical interactions in 1F8I, 1F61, 1DQU and modeled ICL2 between ICL and pthalazinyl & hydroquinone derivatives initially around 60 analogs were designed which were having 4-C and were docked and further refinement was done.

- 1st stage: All the analogs were generated and docked.

Top 15 molecules and their LibDock scores for 1F8I, 1F61, 1DQU and modeled ICL2 are shown in the Table 3.1, Table 3.2, Table 3.3 & Table 3.4 respectively.

Molecule Name	LibDock Score
Hydroquinone4	82.0569
Hydroquinone7	82.0569
Pthalazinyl_j	81.5684
Pthalazinyl_b	81.0363
Pthalazinyl_d	80.9872
Pthalazinyl_g	78.9673
Pthalazinyl_h	77.7432
Hydroquinone5	74.6585
Hydroquinone6	74.6585
Pthalazinyl_i	72.5678
Hydroquinone12	70.9077
Hydroquinone8	69.9023
Hydrouinone13	68.0823
Hydroquinone15	67.0634
Pthalazinyl_y	65.0894

Table 3.1: LibDock Score of Top 15 molecules for 1F8I.

Molecule Name	LibDock Score
hydroquinone12	85.1258
hydroquinone4	83.7568
pthalazinyl_j	81.0246
pthalazinyl_a	80.1675
pthalazinyl_d	79.4567
pthalazinyl_x	76.2456
pthalazinyl_h	74.0259
hydroquinone5	74.0259
hydroquinone8	66.2961
pthalazinyl_i	66.0937
hydroquinone13	64.7089
hydroquinone8	64.0128
hydrouinone16	62.0823
hydroquinone17	62.0634
pthalazinyl_y	61.9872

Table 3.2: LibDock Score of Top 15 molecules for 1F61.

Molecule Name	LibDock Score
hydroquinone12	88.5671
hydroquinone4	86.0582
pthalazinyl_j	83.0612
pthalazinyl_o	83.0612
pthalazinyl_d	80.1672
pthalazinyl_p	75.7861
pthalazinyl_h	74.6742
hydroquinone5	74.0259
hydroquinone8	70.7491
pthalazinyl_i	69.0719
hydroquinone13	69.0719
hydroquinone8	66.7159
hydrouinone16	64.7311
hydroquinone17	62.9856
pthalazinyl_y	62.0872

Table 3.3: LibDock Score of Top 15 molecules for 1DQU.

Molecule Name	LibDock Score
hydroquinone4	84.6901
pthalazinyl_j	82.7812
pthalazinyl_o	80.6431
hydroquinone12	80.2976
pthalazinyl_d	79.0167
pthalazinyl_y	74.0259
hydroquinone8	70.7491
hydroquinone17	70.7491
pthalazinyl_i	69.5916
hydroquinone13	69.1739
hydroquinone5	65.8162
hydroquinone8	65.0659
hydrouinone16	64.7311
pthalazinyl_h	62.7915
pthalazinyl_p	61.8539

Table 3.4: LibDock Score of Top 15 molecules for modeled ICL2.

- 2nd Stage: Then we took these top scored molecules obtained by LibDock (1F8I, 1F61, 1DQU & modeled ICL2) and these were further optimized by CDOCK.

Top 5 molecules and their CDOCKER energy for 1F8I, 1F61, 1DQU and modeled ICL2 are represented in the Table 3.5, Table 3.6, Table 3.7 & Table 3.8 respectively.

Molecule Name	-CDOCKER_Energy
hydroquinone4	28.3927
pthalazinyl_j	27.3858
hydroquinone5	27.2104
pthalazinyl_d	26.9909
pthalazinyl_i	26.6506

Table 3.5: CDOCKER energy of Top 5 molecules for 1F8I.

Molecule Name	-CDOCKER_Energy
pthalazinyl_j	26.5669
hydroquinone8	26.4652
hydroquinone4	25.8112
pthalazinyl_d	25.5586
pthalazinyl_a	25.203

Table 3.6: CDOCKER energy of Top 5 molecules for 1F61.

Molecule Name	-CDOCKER_Energy
pthalazinyl_i	25.1185
hydroquinone12	25.0996
hydroquinone4	24.8928
pthalazinyl_j	24.7099
pthalazinyl_o	24.5199

Table 3.7: CDOCKER energy of Top 5 molecules for 1DQU.

Molecule Name	-CDOCKER_Energy
pthalazinyl_o	22.9823
hydroquinone4	22.7708
hydroquinone8	22.6789
pthalazinyl_j	21.935
hydroquinone5	21.8233

Table 3.8: CDOCKER energy of Top 5 molecules for modeled ICL2.

- 3rd Stage: Since the site is not completely filled, so further modifications were done to occupy the active site. ZincPharmer database searching was done by taking good interaction fragment of best ligand based on docking score. Many structures were retrieved and some of them fit effectively in the active site. But unfortunately docking score couldn't increase considerably indicating lack of improvement in stereoelectronic interactions.

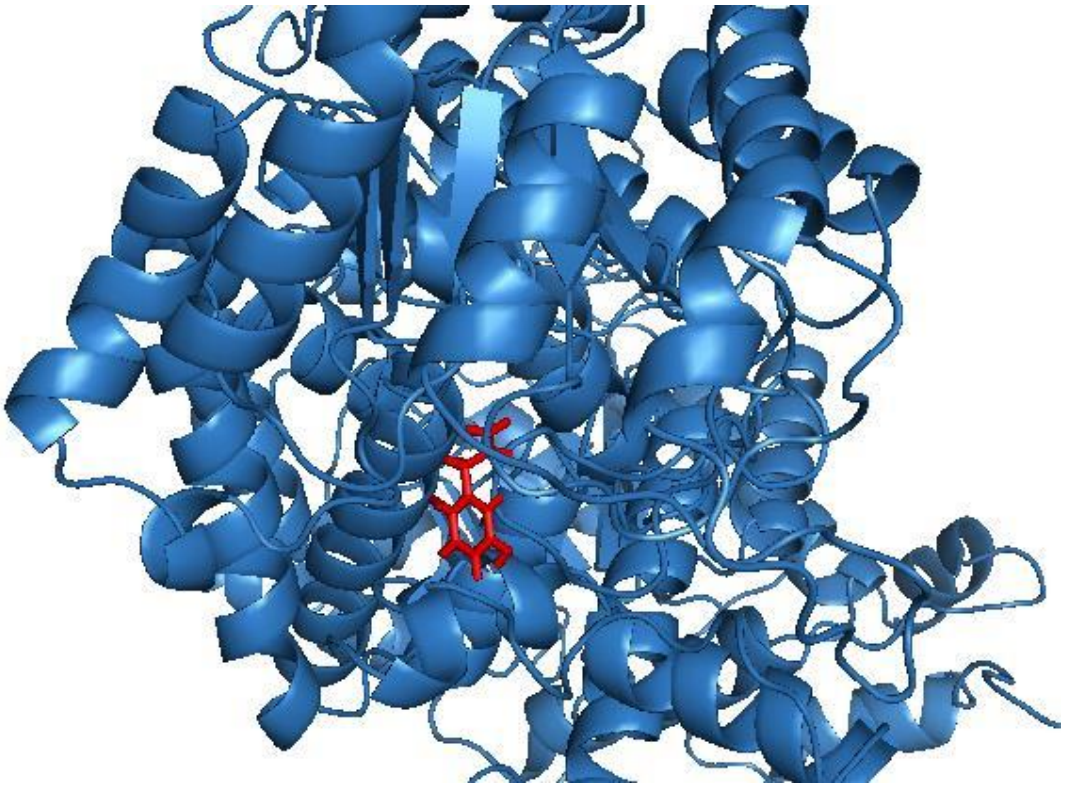


Fig. 3.7: Docked structure of best scored hydroquinone derivative with 1F8I.

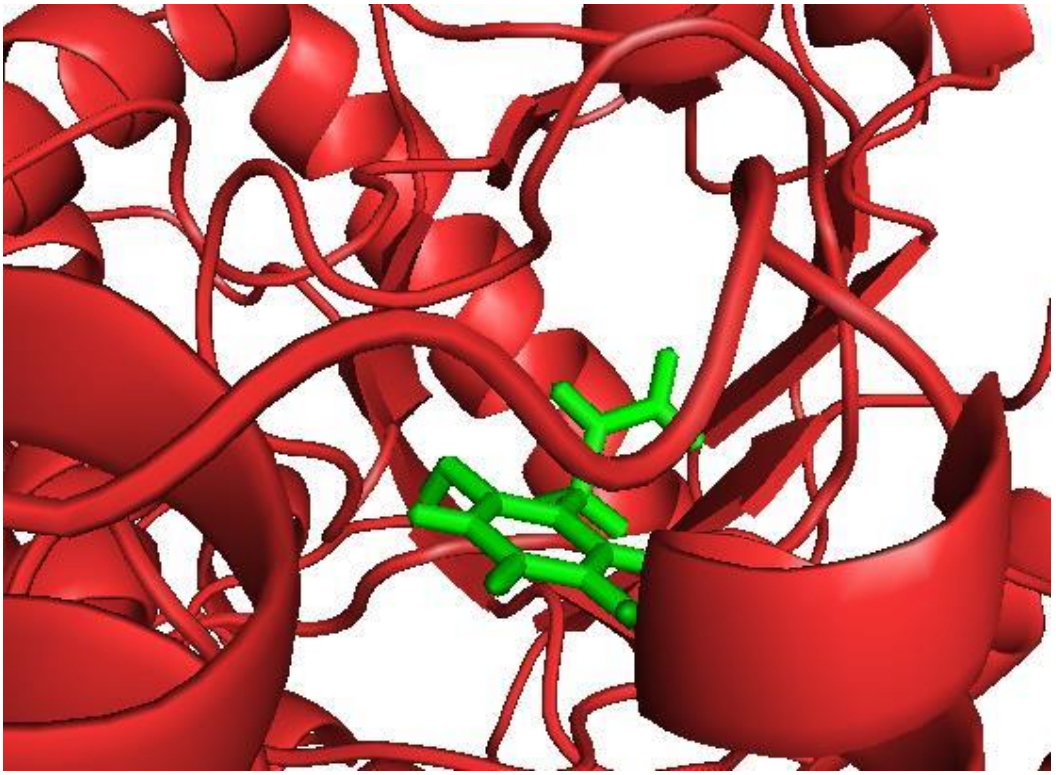


Fig. 3.8: Docked structure of best scored hydroquinone derivative with 1DQU.

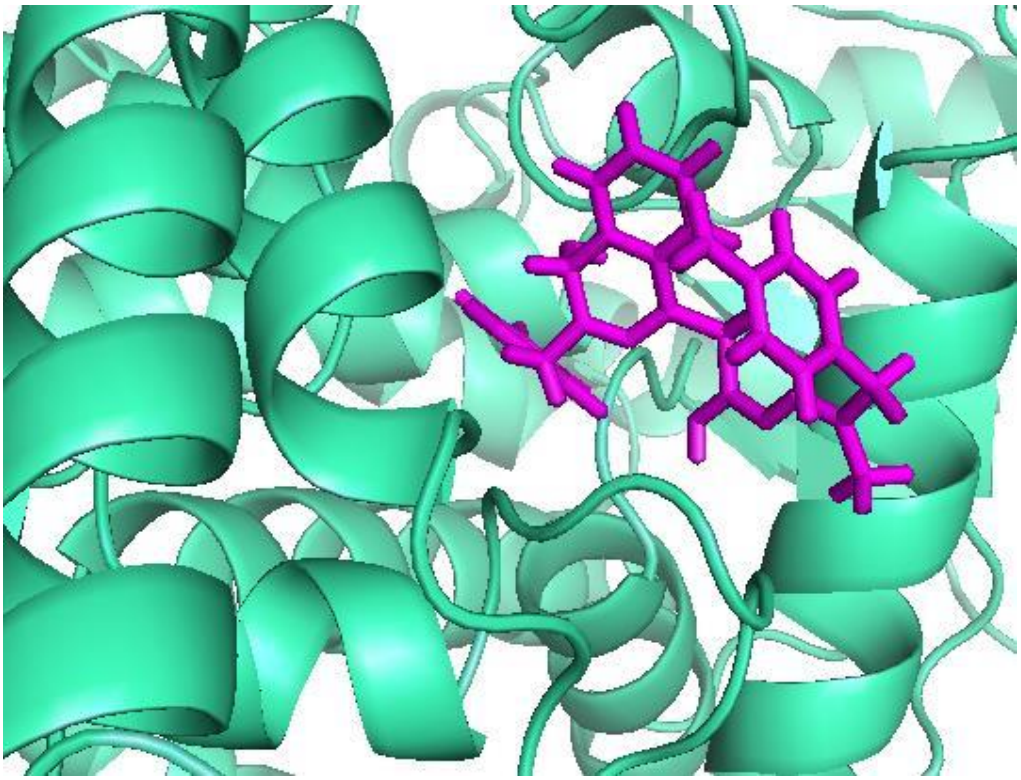


Fig. 3.9: Docked structure of best scored pthalazinyl derivative with 1F61.

3.4 ADMET RESULTS

Source 1: Discovery Studio

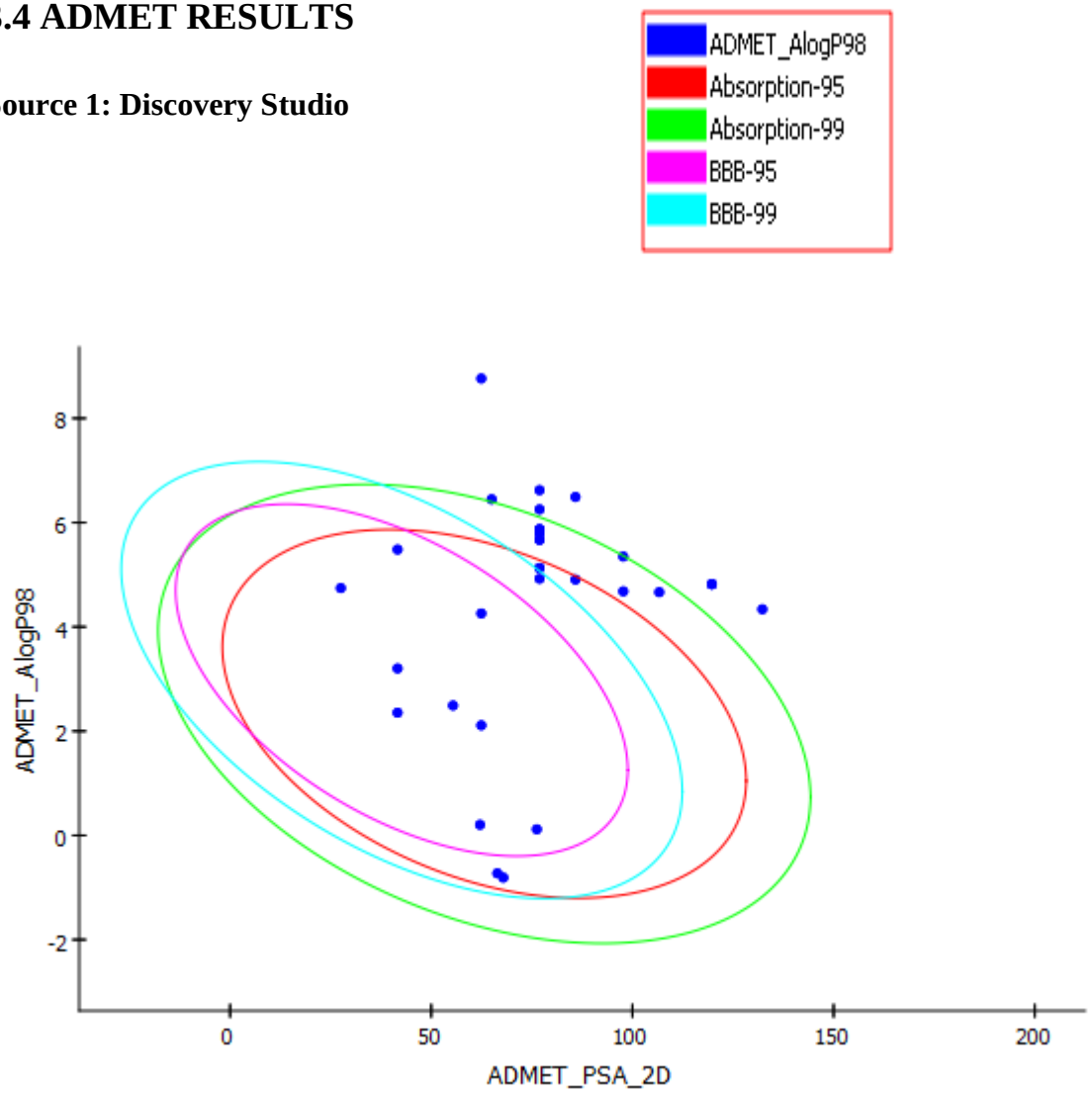


Fig. 3.10: An ADMET plot of manually designed ligands using Discovery Studio.

Source 2: FAF Drugs

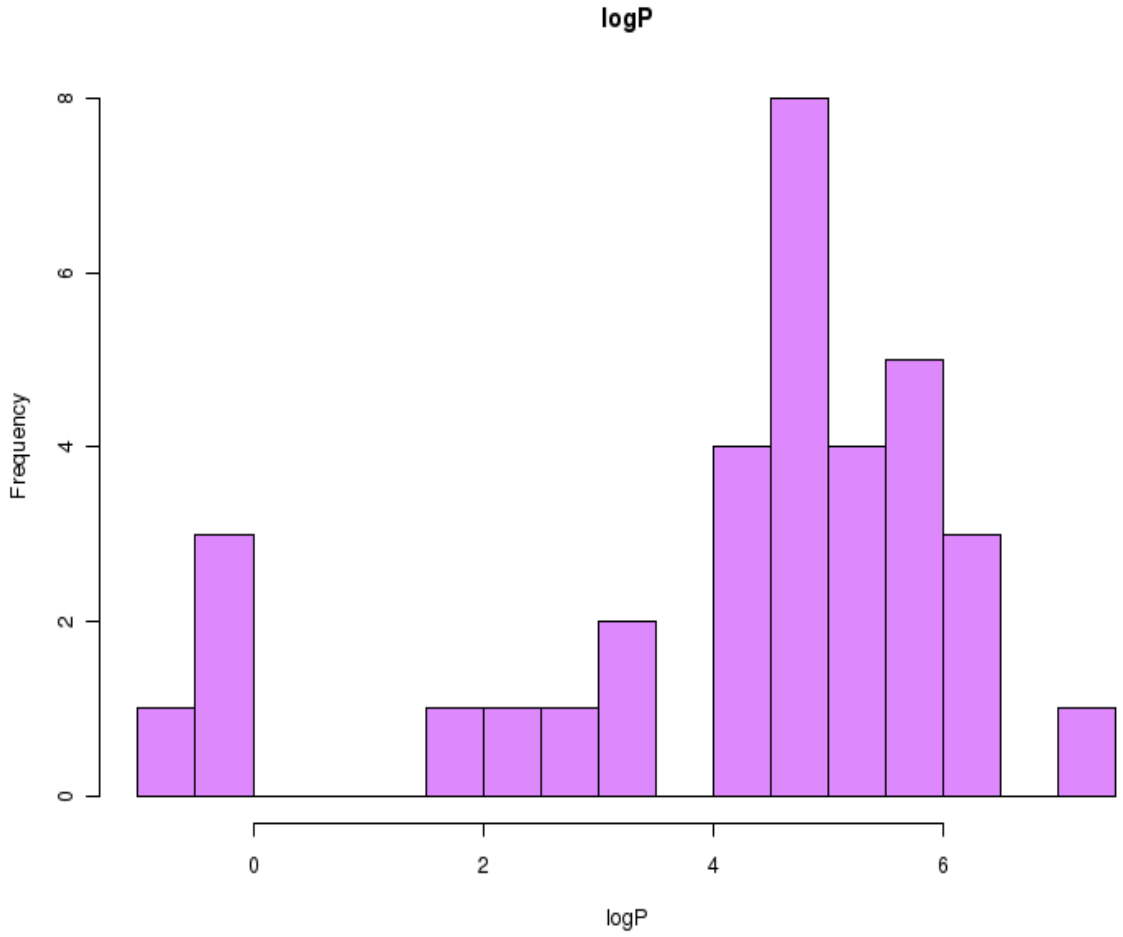


Fig. 3.11(a)

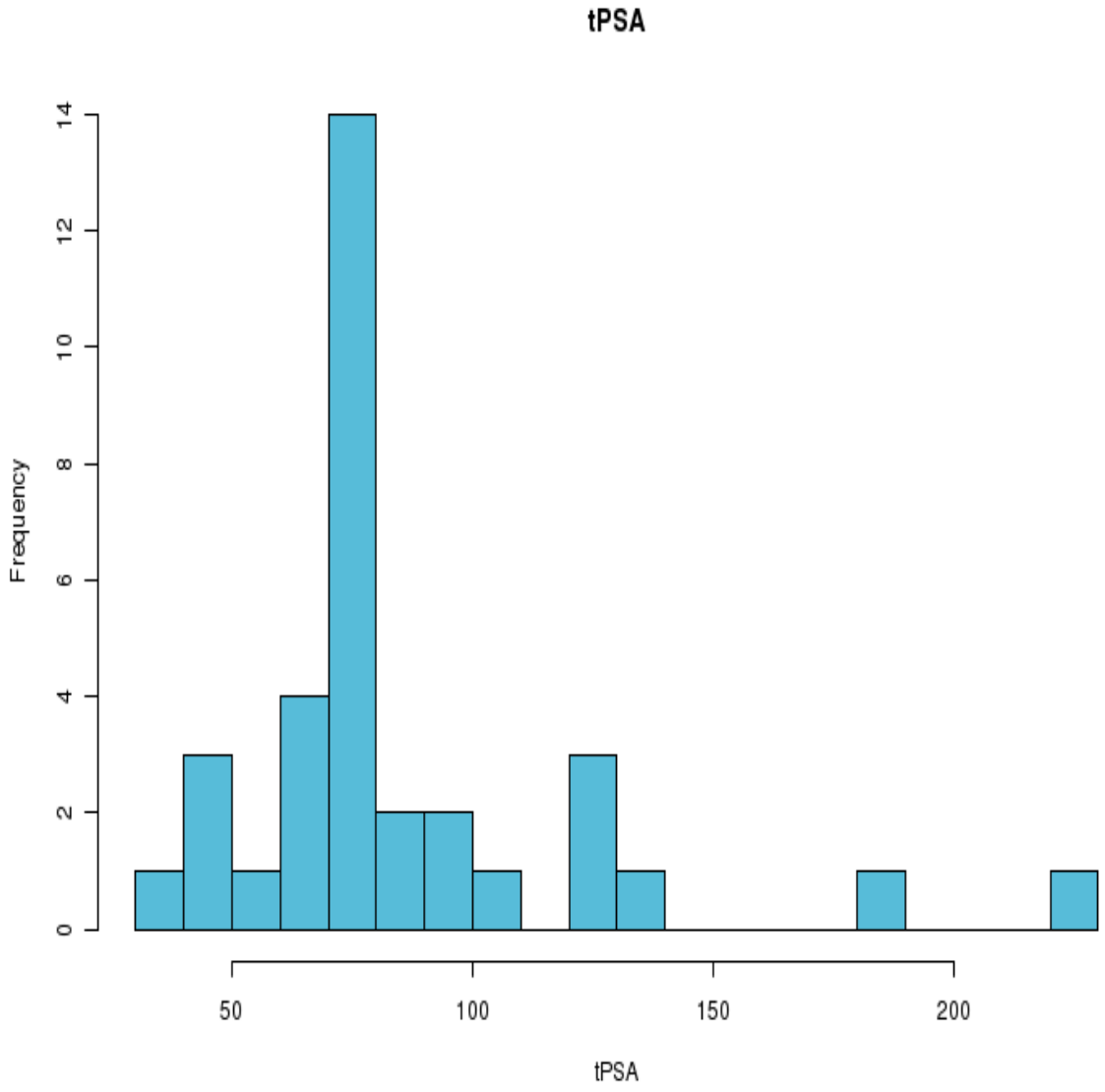


Fig. 3.11(b)

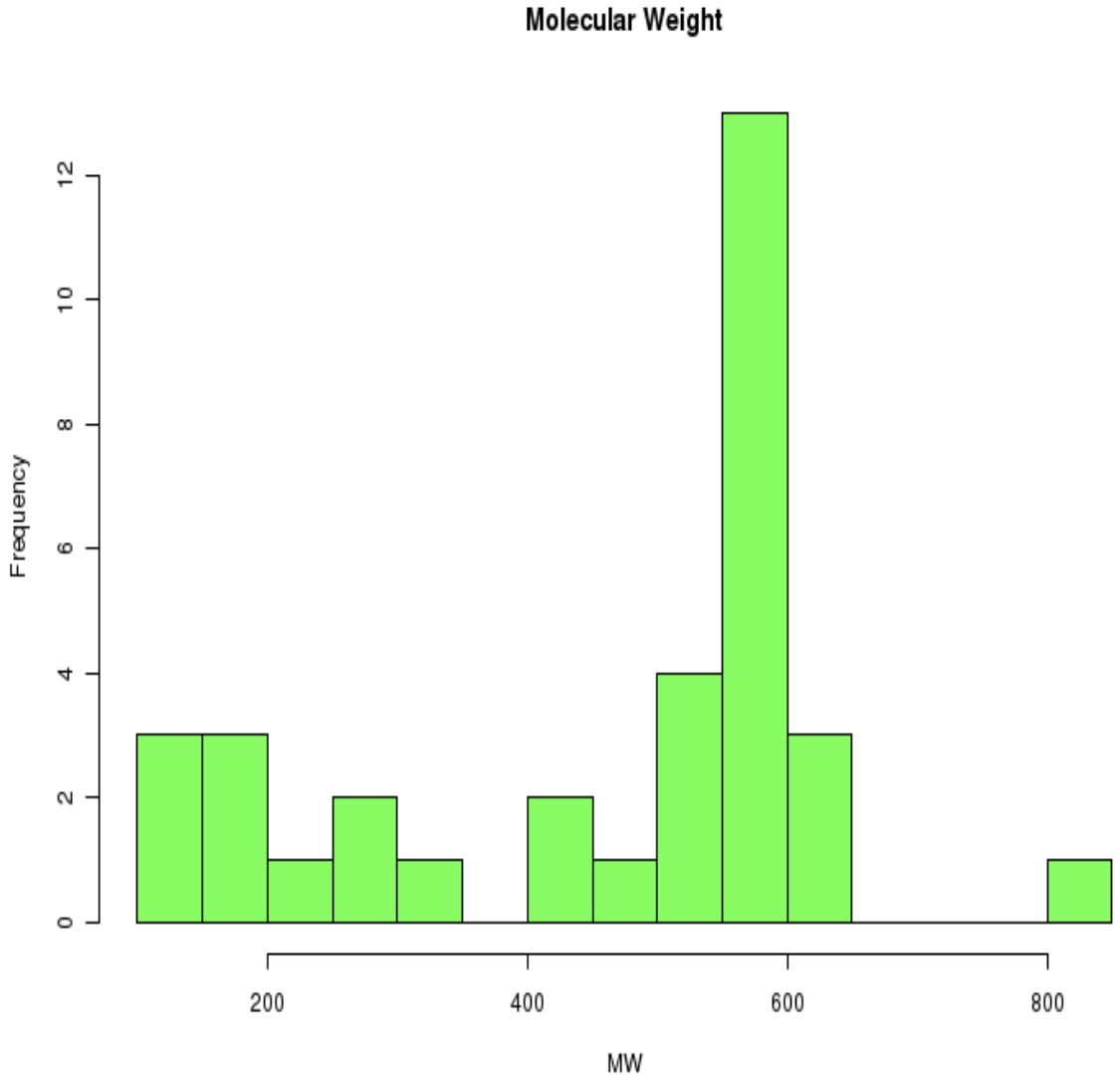


Fig. 3.11(c)

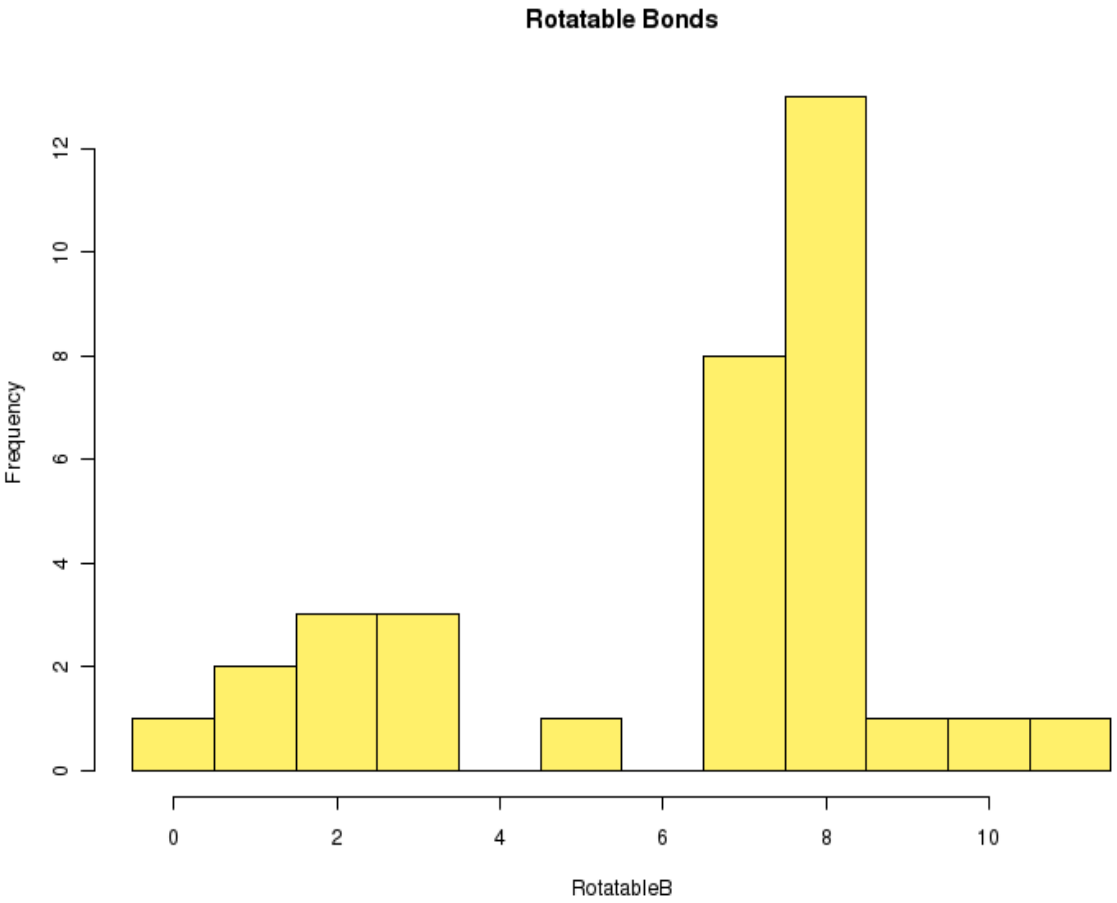


Fig. 3.11(d)

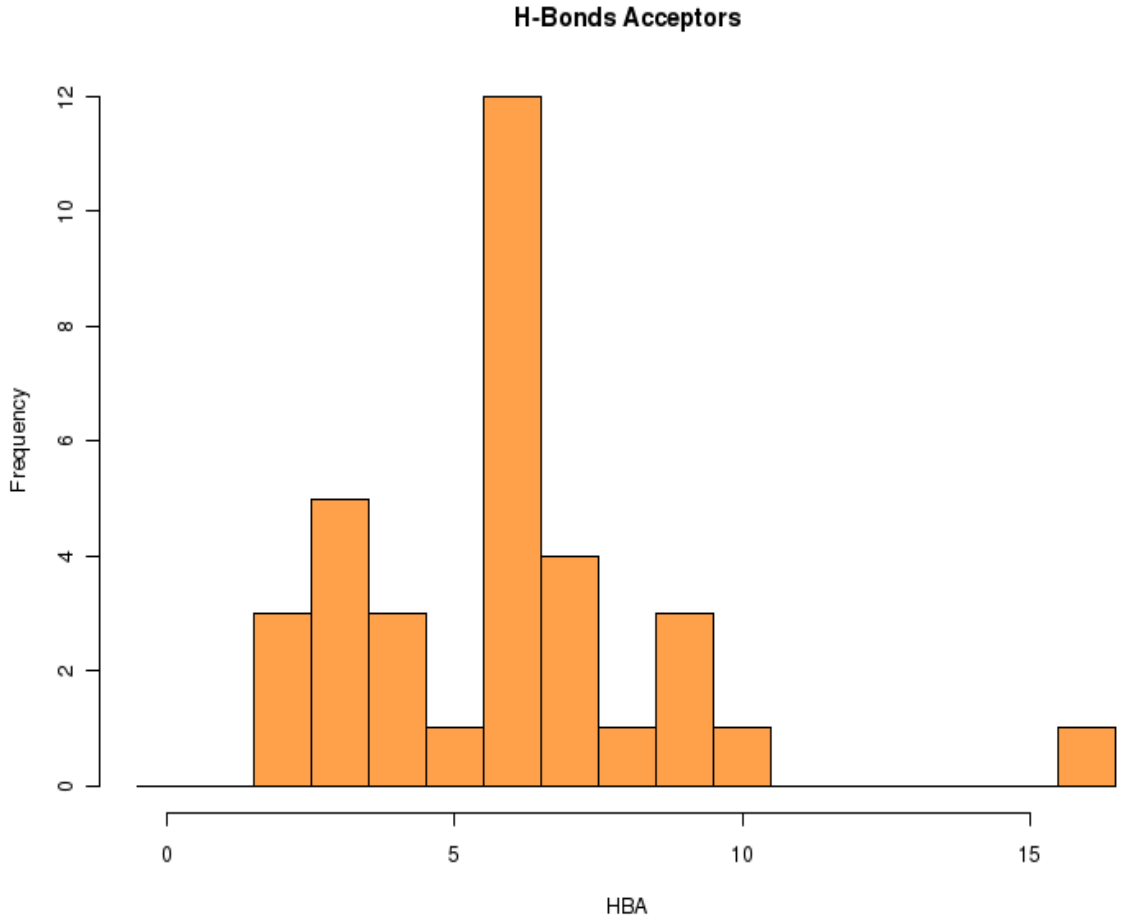


Fig. 3.11(e)

Fig. 3.11: ADMET properties of the predicted lead molecules showing properties (a) logP (b) tPSA (c) Molecular Weight (d) Rotatable Bonds & (e) H-Bonds Acceptor, respectively.

3.5 INTERACTION RESULTS

LigPlus

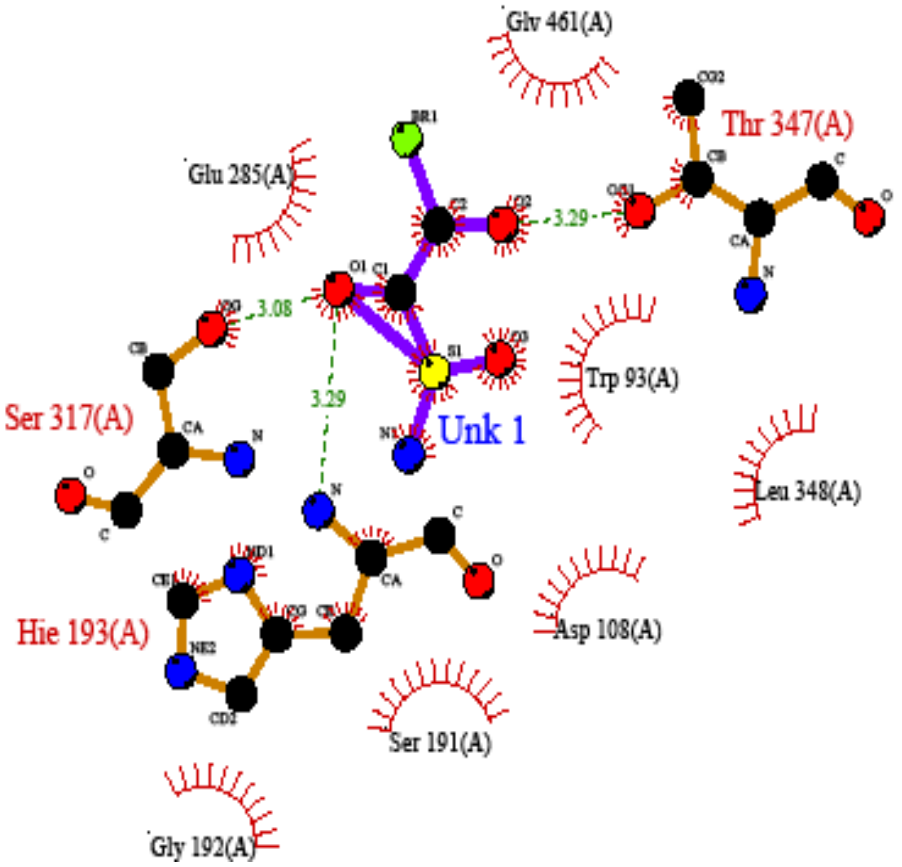


Fig. 3.12: Ligand interaction with active sites in 1F8I.

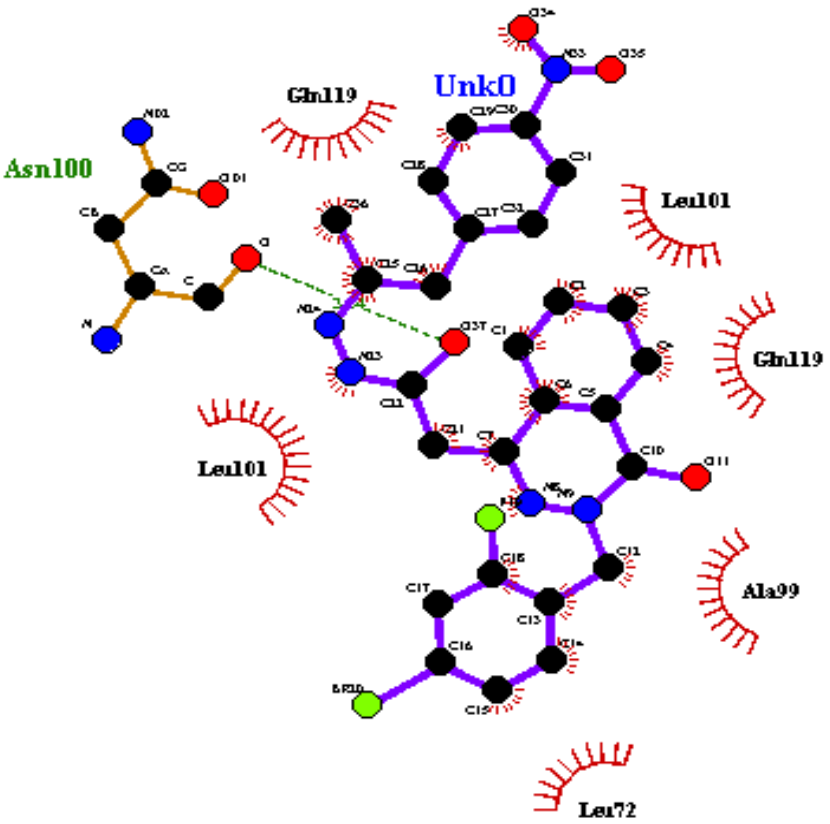


Fig. 3.13: Ligand interaction in 1F61.

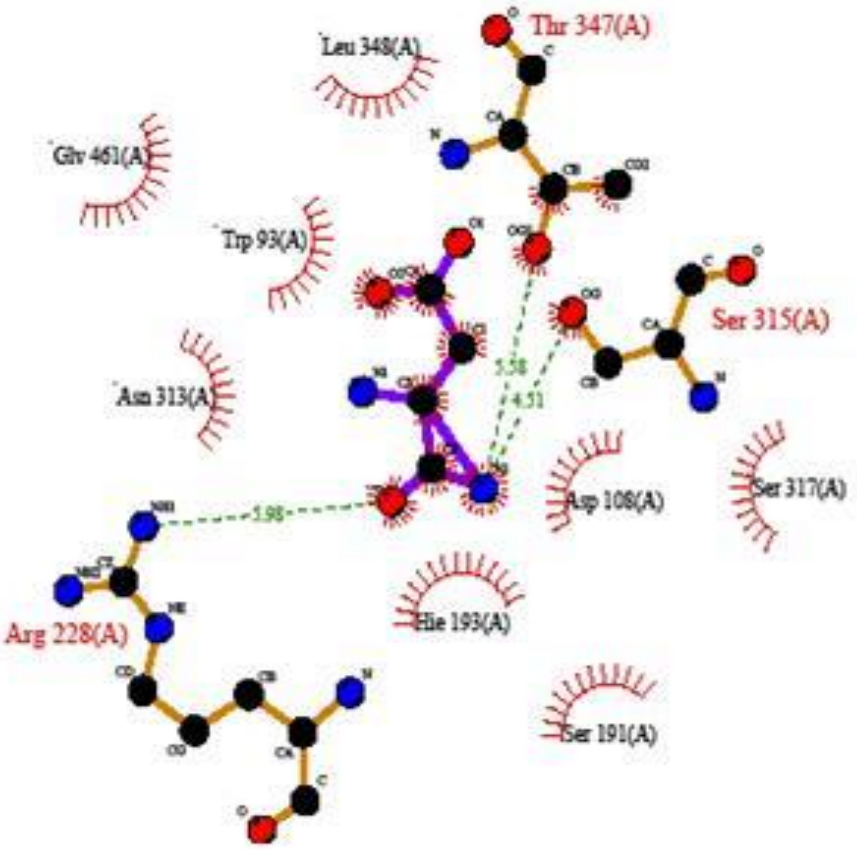


Fig. 3.14: Fragment interaction in 1F8I.

CHAPTER 4

CONCLUSION

TB appears as a global threat to humanity and persistence of *M. tuberculosis* is cure of current morbidity in TB treatment. Protein ICL involved in key process of persistence considered to be a good drug target. Since it is having a small active site, development of drug poses a real challenge. Therefore computational potential lead identification approach is followed to identify diverse core of compounds which will facilitate drug development process.

These 3 PDB IDs namely 1F8I, 1F61 & 1 DQU were preferred for docking as active sites are conserved for both ICL1 & ICL2 in these proteins. The *in vitro* lead molecules were inhibiting both ICL1 and ICL2 but *in vivo* lead molecules were inhibiting ICL1 or ICL2. The pharmacophore search using ZincPharmer was done for *in vivo* lead molecules as they showed better docking score than *in vitro* lead molecules.

The binding structures of ligands from various sources were predicted by LibDock and further validated by CDOCKER and CDOCK score showed satisfactory energy scores. Only 5 of the manually designed ligands show satisfactory CDOCK energy scores and thus were taken for further analysis.

Then toxicity and ADMET calculations were done to find out the ADMET properties of the ligands, then screening was further done.

After the ADME filtering of the ligands and finding the conserved residues, interaction study was carried out for top fragment and manually designed ligands and also of results after ADME screening of fragments, in total of 34 ligands were studied and results were promising as 5 ligands were showing good hydrogen and hydrophobic bonding with conserved residues in the active site. We propose synthesis and activity testing of these ligands against ICL of *Mycobacterium tuberculosis* for second generation of drug discovery and hope to one day contribute in eradicating this disease.

REFERENCES

1. D. G. Russell, "*Mycobacterium tuberculosis*: here today, and here tomorrow," Mol cell biol. 2001, pp. 569-77
2. C W Goulding, I.J. Perry, D. Anderson, M. R. Sawaya, D Cascio, S. Chan, A Parseghian, S. S. Wang, and V. Cassano, "Structural genomics of *Mycobacterium tuberculosis*: a preliminary report of progress at UCLA,".Biophys Chem. 2003, pp. 361-70.
3. A Stanley Sarah, Raghavan Sridharan, W. Wlliam, and S. Cox Jeffery, "Acute infection and macrophage subversion by *Mycobacterium tuberculosis* require a specialized secretion system," 2003, pp. 13001-13006.
4. B. R. Bloom, and E. Rubin, "Genome-wide requirements for *Mycobacterium tuberculosis* adaptation and survival in macrophages," Proc Natl Acad Sci USA. 2005, pp. 8327-32.
5. WHO report 2012 Global tuberculosis control (http://www.who.int/tb /publications/global_report/ 2012/pdf/ full.pdf)
6. J. Irwin John and Brian K. Shoichet, "Zinc − a free database of commercially available compounds for virtual screening," Journal of chemical information and modeling, 2005, pp. 177-182.
7. Vivek Sharma, Sujata Sharma, K. Hoener, Z. U. Bentrup, John D. Mckinney, and David G. Russell, "Structure of isocitrate lyase,a persistence factor of *Mycobacterium tuberculosis.*" Nature Structural & Molecular Biology, 2000, pp. 663-668.
8. C. V Smith, V. Sharma, and J.C. Sacchettini, "TB drug discovery: addressing issues of persistence and resistance," Tuberculosis, 2004, pp. 45-55.
9. R. Kumar, and V. Bhakuni, "*Mycobacterium tuberculosis* isocitrate lyase: role of divalent cations in modulation of functional and structural properties," Proteins, 2008, pp.892-900.
10. D. Sriram, P. Yogeeswari, D.R.Vyas, P. Senthilkumar, P. Bhat, and M. Srividya, "5-nitro-2-furoic acid hydrazones: design, synthesis and in vitro antimycobacterial evaluation against log and starved phase cultures," Bioorg med chem let, 2010, pp. 4313-6.
11. D. Sriram, P. Yogeeswari, P. Senthilkumar, G. Naidu, and P. Bhat, "7,5-nitro-2,6-dioxohexahydro-4-pyrimidinecarboxamides: synthesis, *in vitro* antimycobacterial activity, cytotoxicity, and isocitrate lyase inhibition studies in enzyme inhibition," J med chem, 2010 , pp. 765-72.
12. D. Sriram, P. Yogeeswari, P. Senthilkumar, D. Sangaraju, R. Nelli, D. Banerjee, P.Bhat, and T.H. Manjashetty, "Synthesis and antimycobacterial evaluation of novel phthalazin-4-ylacetamides against log- and starved phase cultures,"Chem biol drug des, 2010, pp. 381-91.

13. D. Sriram, P. Yogeeswari, P. Senthilkumar, S. Dewakar, N. Rohit, B. Debjani, P. Bhat B. Venugopal, V. Pavan, H. M. Thimmappa, "Novel phthalazinyl derivatives: synthesis, antimycobacterial activities, and Inhibition of *Mycobacterium tuberculosis* isocitrate lyase enzyme," Med chem. 2009, pp.422-33.

14. C. Y. Hyeong, Yu Jisu, Oh Ki-Bong, S. Dong-Sun, C. Won-Jea, S. Jongheon, and K. Sanghee, "Symthesis and evaluation of hydroquinone derivatives as inhibitors of isocitrate lyase," Arch Pharm Res, vol.30, 2007, pp. 955-961.
15. R. A. Carr, M. Congreve, C. W. Murray, and D. C. Rees, "Fragment-based lead discovery: leads by design," Drug Discov Today.2005, pp. 987-92.
16. M. A. Miteva, S. Violas, M. Montes, D. Gomez, P. Tuffery, and B. O. Villoutreix, "FAF-drugs: free ADME/tox filtering of compound collections," Nucleic acids res, 2006, pp 738-44.
17. P. A. Smith, M. J. Sorich, R. A. Mckinnon, and J. O. Miners, "Towards integrated ADME prediction: past, present and future directions for modelling metabolism by UDP-glucuronosyl transferases," Journal of Molecular Graphics and Modelling, 2004, pp. 507–517.

18. R. K. David, and J. C. Carlos, "ZINCPharmer: pharmacophore search of the ZINC database," Nucleic Acids Research, vol. 40, 2012, pp.W409-W414.
19. U. C. Kumar, H. Munipalli, and S. Mahmood, "2D QSAR, Pharmacophore and docking studies of Mycobacterium tuberculosis enoyl acyl carrier protein reductase inhibitors," pp.74-89.

BRIEF BIO DATA OF STUDENT

Urvashi is pursuing Master of Technology in Computational Biology and will be completing her degree in June 2014. Her areas of interests are in industrial biotechnology and computational drug designing. She is planning to go for PhD in future after working for few years and acquire programming and various other skills required in bioinformatics. Email-urvi19@gmail.com