International Journal of Advanced Education and Research
ISSN: 2455-5746
Impact Factor: RJIF 5.34
www.alleducationjournal.com
Volume 2; Issue 5; September 2017; Page No. 36-42

3d QSAR, molecular docking studies and virtual screening of *pc* DHFR inhibitors

[1] Ajmal Hashemi, [2] Niaz Mohammad Zahidi, [3] M Vijjulatha
Department of Chemistry, University College of Science, Osmania University, Hyderabad, Telangana, India

Abstract
DHFR is the most intensively studied enzyme in the folate pathway. It is an essential and ubiquitous enzyme. It generates Tetra Hydro Folate (THF), various cofactors which are involved in the transfer reactions of one carbon unit used in the biosynthesis of nucleic and amino acids, including methylation of dUMP to dTMP. DHFR inhibitors act by halting synthesis of DNA, RNA, and proteins, thereby arresting cell growth. DHFR is an important target for drug development against cancer and a variety of infectious diseases caused by bacteria, protozoa, and fungi. DHFR from *Pneumocystis carinii* has received significant attention because this pathogen is often associated with AIDS and other immune deficiencies. Molecular modeling techniques are now widely used in medicinal chemistry for design and optimization of inhibitors for a specific target. A pharmacophore model was generated using reported *Pc* DHFR inhibitors in PHASE module of Schrodinger suite, the pharmacophore model was able to accurately predict *P*c DHFR activity. Further the molecules were subjected to molecular docking to understand the binding mode, the docking results also provide additional confidence in the proposed Pharmacophore model. Results suggested that the proposed 3D QSAR model and docking analysis can be useful to rationally design new inhibitors. A combined pharmacophore based and docking based virtual screening was performed to obtain a possible lead molecule for further optimization.

Keywords: 3D-QSAR Molecular Docking and Virtual Screening of *Pc* DHFR, Schrodinger suite

Introduction

Pneumonia: Pneumonia is an inflammatory condition of the lung affecting primarily the microscopic air sacs known as alveoli. Typical signs and symptoms include a cough with phlegm, chest pain, fever, and trouble breathing. Symptoms can vary from mild to severe. People who are old or very young may not have typical symptoms. Usually people begin improving within three days of starting treatment; however, they may feel tired for more than a month afterwards.

Pneumonia is usually caused by infection with viruses or bacteria and less commonly by other microorganisms, certain medication sand conditions such as autoimmune diseases.

Pneumocystis pneumonia {PCP}

Pneumocystis pneumonia (PCP) is a form of pneumonia, caused by the yeast-like fungus *Pneumocystis carinii* [5]. *Pneumocystis* pneumonia is not commonly found in the lungs of healthy people, but, being a source of opportunistic infection, it can cause a lung infection in people with a weak immune system. *Pneumocystis* pneumonia is especially seen in people with cancer undergoing chemotherapy, HIV/AIDS, and the use of medications that suppress the immune system.

Methodology

Development of new drugs is undoubtedly one of the most challenging tasks of today's science. Human genome project that gave 30,000 or so genes encoded within the human genome, it was expected that a large number of new drug targets would be found expeditiously. But it did not turn out to offer a direct source for drug development, because it is the proteins encoded by the genes that are usual drug targets. This much larger proteome is far more complex than the collection of genes, as proteins may undergo post-translational modifications, associations with other molecules and prosthetic groups, and formation of multimeric complexes. [59] Today, the field of drug development may seem more fertile than ever before, with vast amounts of information from genomic and proteomic studies facilitating the finding of new targets, usage of rational combinatorial chemistry for the production of libraries of compounds, generation of genetically modified animal models for the development and testing of new drugs, and the possibility of using ultra-high-throughput test techniques for screening of large libraries. Drug research and development is comprehensive, expensive and time consuming. It is estimated that a drug from concept to market would take approximately 12 years and exceeding US$800 million on an average [60]. Several new technologies have been developed, applied in drug research to shorten research processes and to reduce expenses. In post genomic era, Computer-Aided Drug Design (CADD) has considerably extended its range of applications, spanning almost all stages in drug discovery pipeline, from target identification to lead discovery, from lead optimization to pre-clinical or clinical trials. [61] In early 1960s, Quantitative Structure–Activity Relationship (QSAR) analysis emerged as the first computer aided drug design technique. In recent decade the concept of CADD has evolved very quickly, with an unprecedented development of structural biology and computer capabilities. However, despite all these advances, revolutionary era of drug design has not yet arrived. [62-64] there is no unique solution to

a drug design problem, appropriate experimental techniques or computational methods to use will depend on characteristics of the system itself and information available. A variety of computational approaches can be applied at different stages of drug-design process: in an early stage, these focuses on reducing number of possible ligands, while at the end, during lead-optimization stages, emphasis is on decreasing experimental costs and reducing time. CADD now plays a critical role in search for new molecular entities. [65-68] Current focus includes improved design and management of data sources, creation of computer programs to generate huge libraries of pharmacologically interesting compounds, development of new algorithms to assess the potency and selectivity of lead candidates and design of predictive tools to identify potential ADME/Tox liabilities.

There are two major types of drug design, first is referred to as ligand-based drug design and second, structure-based drug design. These are two distinctapproaches used in the area of computer-aided drug design. When only lead is a set ofknown active compounds or knowledge of a biochemical transformation which is to beinterrupted, then the path is less direct, this approach is referred as Ligand (analogue) Based Drug Design (LBDD). Currently favored tactics include use of molecularsimilarity methods and employment of neural networks. Recent advances include theprediction of relative potency of different chiral forms of drugs. If molecular structure ofthe target macromolecule is known methods are direct and can give a high level ofsophistication, this approach is referred as Structure Based Drug Design (SBDD).

Results and Discussion

The importance of dihydrofolatereductase (DHFR) (EC: 1.5.1.3) in parasiticchemotherapy arises from its function in DNA biosynthesis and cell replication. Inhibitors of DHFR block the function resulting in inhibition of DNA synthesis. 3D-QSAR model was performed on 165 previously reported *Pc* DHFR inhibitors. [115] there structures are shown in figure 4

3D QSAR Pharmacophore Generation: The present QSAR model was developed using the 'pharmacophore based' option of PHASE. To find the common pharmacophore hypothesis, the dataset was divided into actives and inactive set. Molecules with pIC_{50} values higher than 6.80 were considered to be active, and those with pIC_{50} values less than 5.00 were considered to be inactive, whereas those in-between were considered to be moderately active. One hypothesis was identified from the set of 11 actives (Table 1 and 2).

Table 1: Best pharmacophore hypotheses according to scoring values.

Hypotheses	Survival	Surv-inactive	Post-hoc	# matches
DDRRR.110	3.331	1.49	3.331	11

Table 2: Statistic parameters for the best pharmacophore hypotheses DDRRR

PLS factor	SD	R^2	Q^2	F	RMSE
1.	0.7839	0.486	0.368	81.3	0.7599
2.	0.5501	0.7499	0.6342	127.4	0.5781
3.	0.4201	0.8558	0.6352	166.2	0.5773

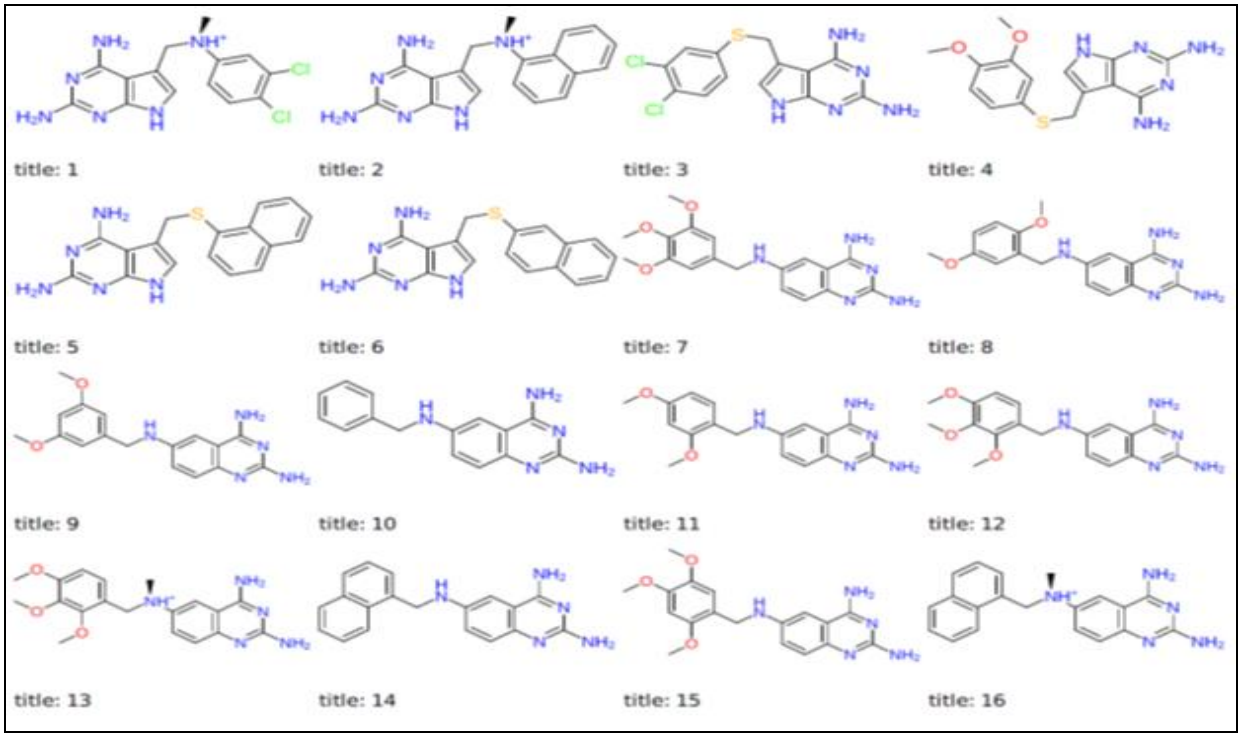


Fig 1: Structure of *Pc*DHFR inhibitors

This hypothesis survived three different phases of PHASE scoring procedure (Survival, Surv-inactive, Post-hoc), and therefore this was used for the generation of QSAR models. For the QSAR models generation, non-modeled (inactive or moderately active) molecules in the dataset were then aligned, based on matching with at least three pharmacophore features. The dataset was randomly divided into training set of 89compounds and 76 in the test set, in order to create, at least, the standard 2:1 training and test set ratio needed for QSAR study. The best pharmacophore model resulted DDRRR.11 (R^2 = 0.8558). The goodness of the model was validated by Q^2 for test set (Table 2). Plot of predicted vsexperimental pIC_{50} for training and test set is shown in Fig2.

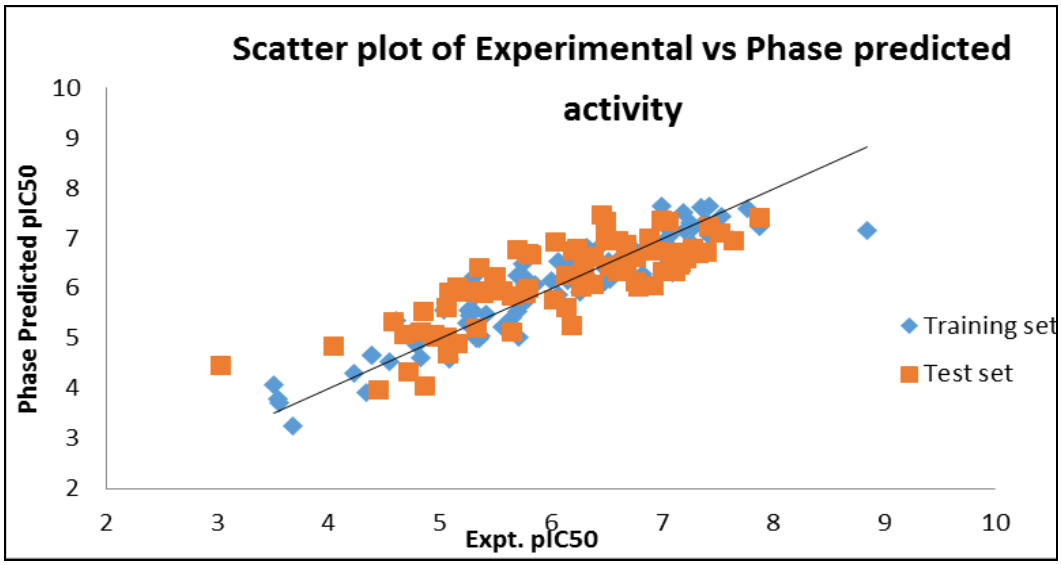
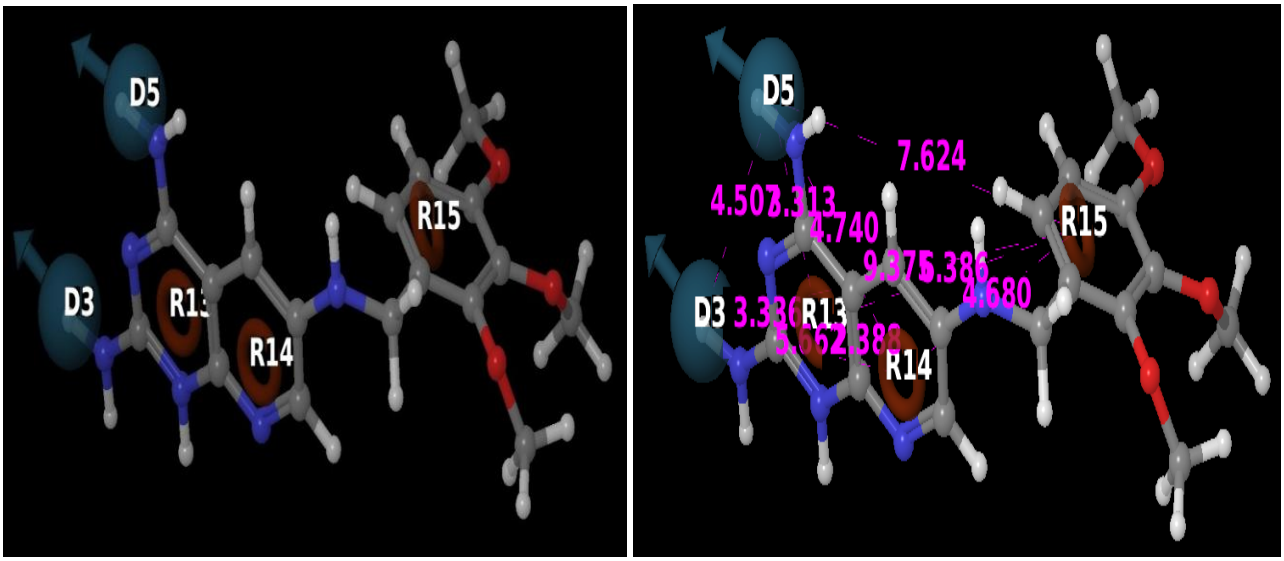


Fig 2: Scatter Plot of Experimental vs Phase predicted pIC_{50} for training and test sets.

Pharmacophore sites spatial distribution of DDRRR.11 model (Figure 7) shows that two donor site (D3, D5) and three aromatic ring(R13, R14, and R15) are found to be in the space of about 3.3 to 7.6 Å. In the pharmacophore mapping study, it was found that the major structural factors, affecting the potency of these compounds, are related to the basic skeleton. The pharmacophore hypothesis shows distance between pharmacophoric sites are depicted in figure -- and values are given in table 3 & 4.



(a) **(b)**

Fig 3: (a) Pharmacophore model aligned with the best active compound 150 (b) Pharmacophore model DDRRR.11, all distances are expressed in Å.

Table 3: Distances between different sites of model DDRRR

Site 1	Site 2	Distance(å)
D3	D5	4.507
D3	R13	3.336
D3	R14	5.662
D3	R15	9.375
D5	R13	3.313
D5	R14	4.74
D5	R15	7.624
R13	R14	2.388
R13	R15	6.386
R14	R15	4.68

Table 4: Angles between different sites of model DDRRR

Site 1	Site 2	Site 3	Angle (å)
D5	D3	R13	47.1
D5	D3	R14	54.1
D5	D3	R15	53.7
R13	D3	R14	7.1
R13	D3	R15	21.4
R14	D3	R15	22.6
D3	D5	R13	47.5
D3	D5	R14	75.5
D3	D5	R15	97.9
R13	D5	R14	28
R13	D5	R15	55.9
R14	D5	R15	35.7
D3	R13	D5	85.4

The pharmacophore map and QSAR contour maps can be used to design new and more active analogues. A descriptive representation of the contours generated in the QSAR is shown infigure 8. The major advantages of 3D-QSAR techniques are the cubes generated using PLS regression which could be visualized in 3D space.

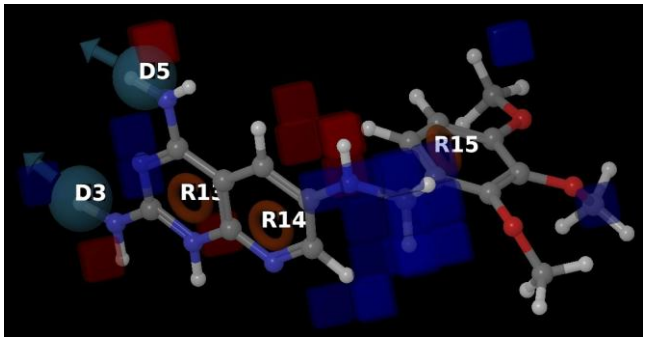


Fig 4: Pictorial representation of the cubes generated using the QSAR model of most active molecule. Blue cubes indicate favorable regions, while red cubes indicate unfavorable region for the activity.

The activity cubes can be generated for different properties such as hydrogen bond acceptor, hydrogen bond donor, hydrophobic, positive and negative ionic features, which define the no covalent interactions with receptor. In these generated cubes, blue cubes indicate the favorable features and red cubes indicate the unfavorable features for the biological activity spectrum.

Table 5: QSAR set, Experimental activity, predicted activity, Phase fitness and Docking score of P*c* DHFR inhibitors.

Molecule	QSAR Set	Activity	Predicted Activity	Fitness	Docking score in kcal/mol
1	training	4.548	4.52	1.14	-5.91174
2	training	3.679	3.24	1.9	-6.02551
3	training	4.232	4.31	2.15	-9.61434
4	test	4.954	5.06	0.7	-8.15706
5	test	4.974	5.01	0.71	-9.83589
6	test	3.032	4.46	2.06	-10.4548
7	test	5.167	4.89	1.15	-9.21118
8	training	5.337	4.99	2.02	-8.44865
9	training	5.657	5.41	2.42	-9.32969
10	test	5.06	5.61	2	-8.52724

Molecular Docking

For the docking analysis of 165 molecules, crystal structure of P*c*DHFR was collected from the RCSB protein data bank (PDB ID:3NZB). The protein was prepared in 'protein preparation wizard' of Maestro. Protein preparation included addition of hydrogen atoms, deletion of solvent molecules except for active site waters, completion of bond order. Finally, the protein complex was prepared which was considered for docking analysis. Glide extra precision (XP) was utilized for docking.

Docking studies were carried out on all molecules and the most active molecule was analysed in the receptor ligand binding region.Figure 9shows docked model of best active compound within the active site of 3NZB. Best active molecule showed hydrogen bonding interactions with Glu 32, Ile 10 and Ile 123. Dock score of all the Pc DHFR inhibitors is given in table 5.

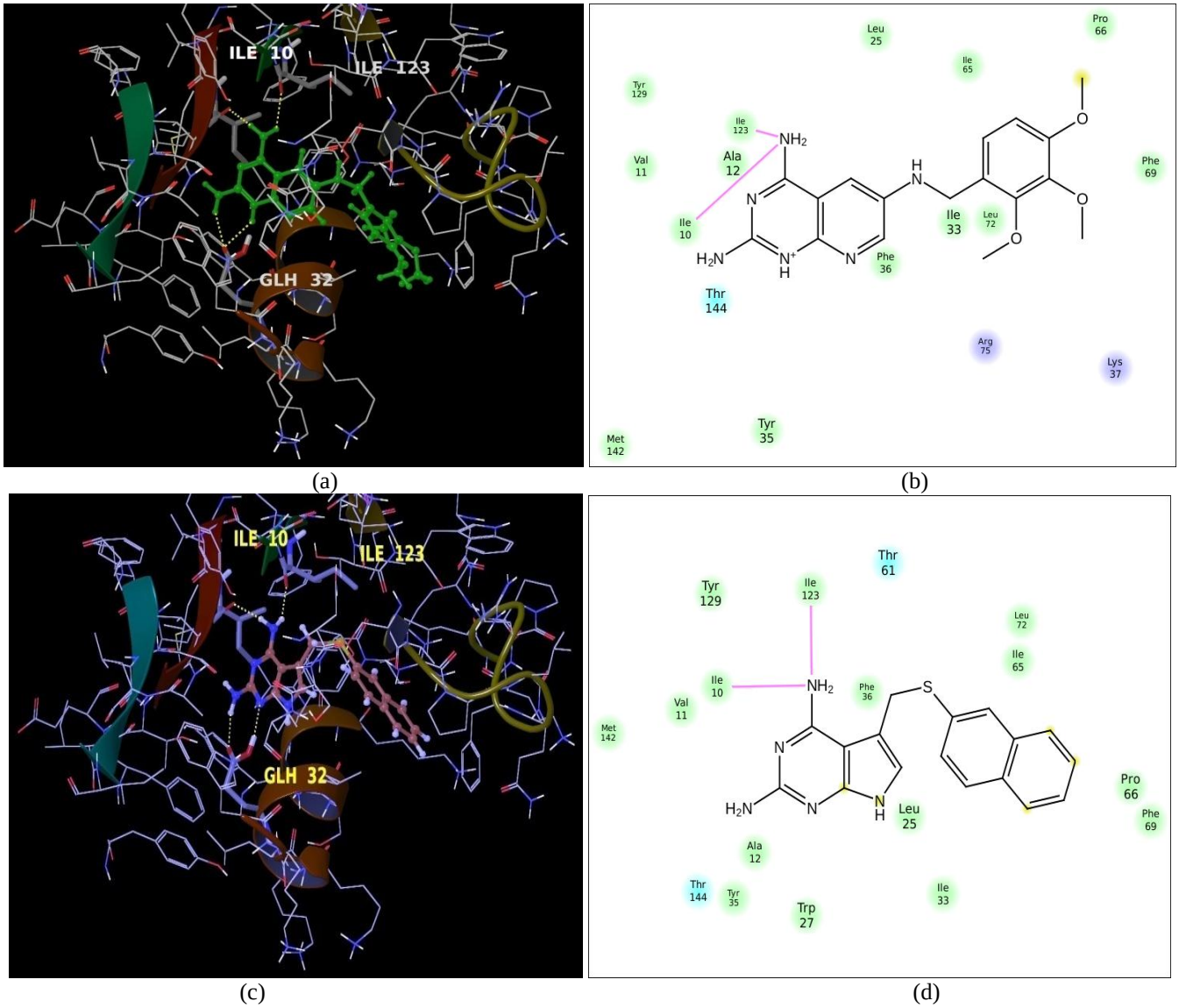


Fig 5: Dock pose and ligand interaction diagram of (a) & (b) most active molecule 150 and (c)&(d) best docked molecules with hydrogen bond interactions.

Virtual Screening

The pharmacophore model has given an idea about how to get the necessary andsufficient molecular required for anew candidate. Thus the well validated pharmacophore DDRRR.11 was used as a 3D query for retrieving potential new scaffoldsagainst P*c*DHFR in the chemical databases. TOSLab containing 7,054 molecules were used for screening process by the Find matchesfor hypothesis method. Initial screening and retrieved hits were filtered by usingLipinski's rule of five for the refinement of drug likeness andapplying ADME properties, a total of 131 molecules were passedout these filtrations. These hit molecules were taken for molecular docking studies, Different docking studies have been applied to identify suitable orientation of a ligand, ability to interact with the active site of the protein and also to check the retrieved molecules from the database whether the pharmacophore chemical features are mapped with structure-based interaction mode or not.

One hitwas obtained from the virtual screen that carried common structural features for behaving as pharmacophoric features i.e. the donor groups being NH group which are forming the hydrogen bond interactions with the essential active site amino acid residues Ser 64. From the above analysis, we conclude that the small molecule, which was retrieved from TOSLab databases have satisfied all necessary conditions such as binding affinity, calculated drug-like properties(figure 6) and thus could be treated as good leads in the design of potent inhibitors of P*c*DHFR. Hence, these filtering results have given a confidence of the validity and robustness of the quantitative pharmacophore model DDRRR.11.The dock pose and ligand interaction diagram of the virtual hit is shown in the figure 7.

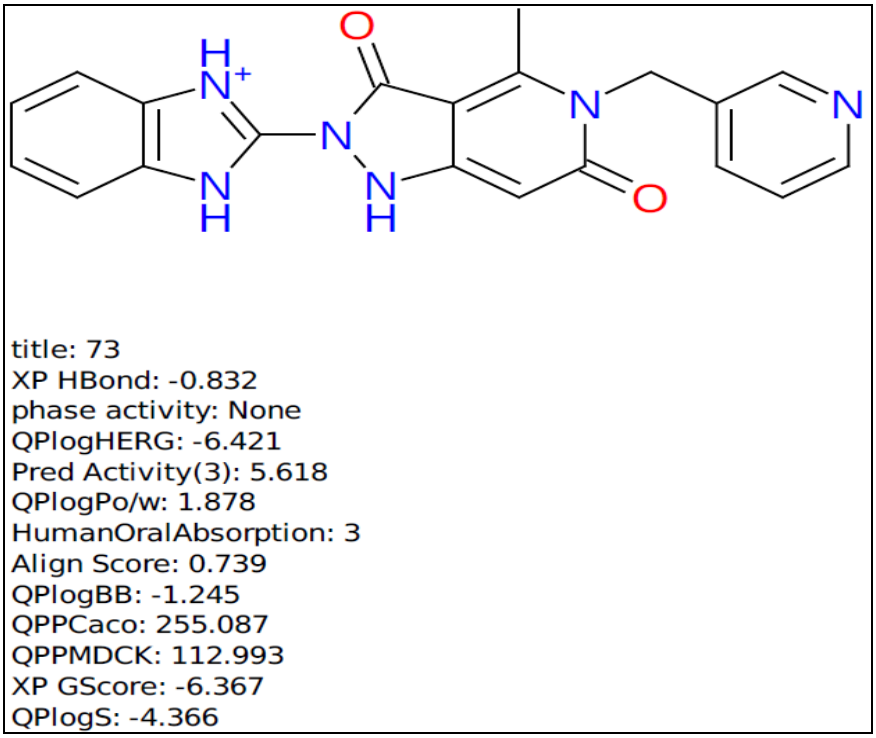


Fig 6: Structure of Virtual screening hit along with the calculated ADME and docking parameters

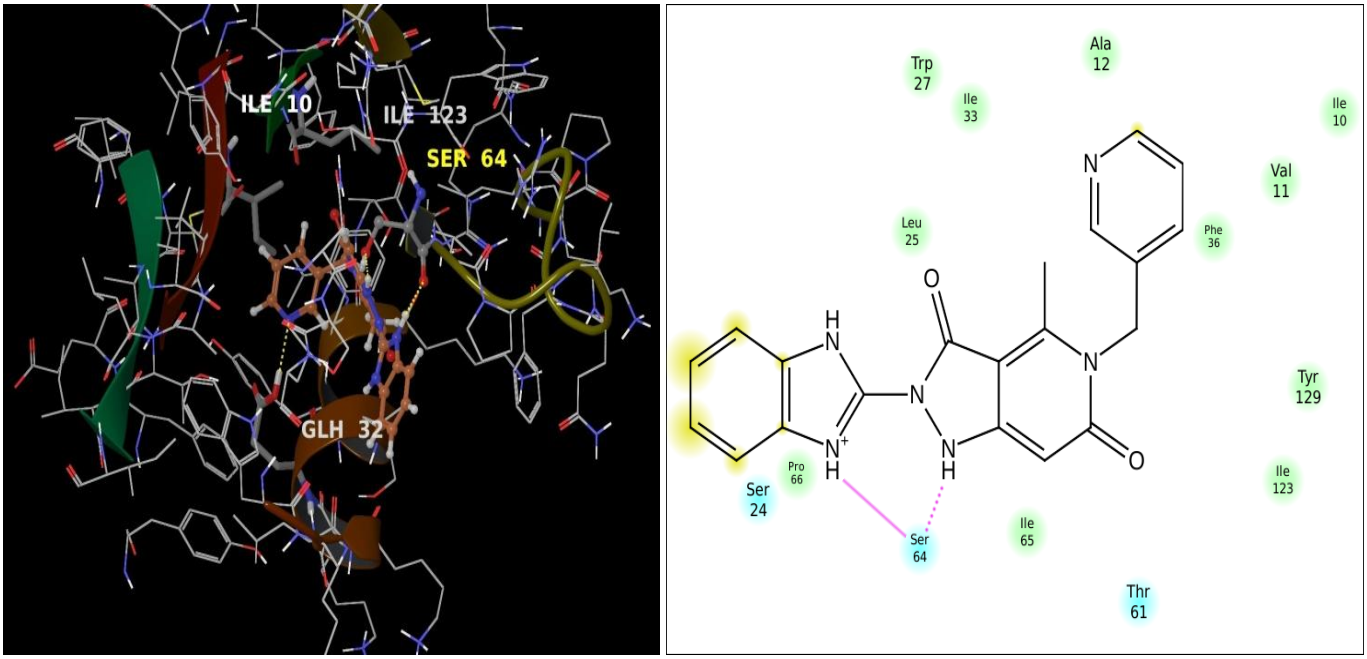


Fig 7: Dock pose and Ligand interaction diagram of VS hit molecules showing hydrogen bond interaction with Glu32 and Ser 64.

Conclusion

This study shows the generation of a Pharmacophore model DDRRR for reported molecules as potent *Pc*DHFR inhibitors. Pharmacophore modeling correlates activities with the spatial arrangement of various chemical features. Hypothesis DDRRR represents the best Pharmacophore model for determining *Pc*DHFRactivity. This Pharmacophore model was able to accurately predict *Pc*DHFR activity, the validationand the docking results also provide additional confidence in the proposed Pharmacophore model. Results suggested that the proposed 3D, QSAR model and docking analysis can be useful to rationally design new inhibitors. A combined pharmacophore based and docking based virtual screening was performed to obtain a possible lead molecule for further optimization.

References

1. McLuckie A. ed. Respiratory disease and its management. New York: Springer, 2009, 51.
2. Leach Richard E. Acute and Critical Care Medicine at a

Glance (2nd ed.). Wiley-Blackwell, 2009. ISBN 1-4051-6139-6
3. Ashby Bonnie, Tarkington, Carol The encyclopaedia of infectious diseases (3rd ed.). New York: Facts on File. 2007, 242. ISBN 0-8160-6397-4.
4. Jeffrey Pemberville C. Alcamo's Fundamentals of Microbiology (9th ed.). Sudbury MA: Jones & Bartlett, 2010, 323. ISBN 0-7637-6258-X.
5. Hoover DR, Saah AJ, Bacellar H, Phair J, Detels R, Anderson R, *et al.* Clinical manifestations of AIDS in the era of pneumocystis prophylaxis. New England Journal of Medicine. 1993; 329:1922-6.
6. Stern A, Green H, Paul M, Vidal L, Leibovici L Prophylaxis for Pneumocystis pneumonia (PCP) in non-HIV immunocompromised patients". Cochrane Database Syst Rev, 2014, 10(CD005590).
7. Aliouat-Denis CM, *et al.* Pneumocystis species, co-evolution and pathogenic power". Infection, Genetics and Evolution. 2008; 8(5):708-726.
8. Puzio J, Kucewicz E, Sioła M, *et al.* Atypical and opportunistic pulmonary infections after cardiac surgery. Anestezjologia Intensywna Terapia (in Polish). 2009; 41(1):41-5.
9. Morris A, Lundgren JD, Masur H, *et al.* Current epidemiology of Pneumocystis pneumonia. Emerging Infect. Dis. 2004; 10(10):1713-20.
10. Masur H, Michelis MA, Greene JB, *et al.* An outbreak of community-acquired Pneumocystis carinii pneumonia. N Engl J Med. 1981; 305(24):1431-8.
11. Foley NM, Griffiths MH, Miller RF. Historically atypical Pneumocystis cariniipneumonia. Thorax. 1993; 48:996-1001
12. Kovacs JA, Halpern JL, Suran JC, Moss J, Parillo JE, Masur H. Identification of antigens and antibodies specific for Pneumocystis carinii. Journal of Immunology. 1988; 140:2023-31