

In Silico Sequence Analysis and Homology Modeling of Dihydropteroate Synthase in Pneumocystis Pneumonia

Yogesh Joshi^{1*}, Chakre G.D², Vinod P.S³ and Anilkumar S.Katti⁴

Dept. of PG studies and Research in Bioinformatics, Walchand centre for Biotechnology,
Walchand College of Arts and Science, Solapur, Maharashtra, India. 413006.

*Correspondence author: Email: ynjoshi0307@gmail.com

Abstract

Pneumocystis pneumonia (PCP) remains a leading opportunistic infection in patients with weakened immune system. The fungus causing the infection belongs to the genus, *Pneumocystis*, and its members are found in a large variety of mammals. Although *Pneumocystis carinii* pneumonia is one of the leading causes of morbidity and mortality among patients with the acquired immunodeficiency syndrome, many questions about its epidemiology and transmission remain unanswered. *Pneumocystis* is classified as a fungus and is host-species specific, but an understanding of its reservoir, mode of transmission, and pathogenesis is incomplete. Dihydropteroate synthase has been targeted in *Pneumocystis* pneumonia (PCP) disease. DHPS is part of a trifunctional protein along with dihydroneopterin aldolase and hydroxyl methyl dihydro pterinpyrophospho kinase, two other enzymes in the folic acid biosynthesis pathway. Sulfa resistance has evolved in numerous bacterial pathogens as well as in the malaria parasite, *Plasmodium falciparum*. In this present study, structural characterization carried out the by various *In silico* tools. The predicted three dimensional structure depicts the structure based drug design.

Keywords:- *Pneumocystis*, dihydropteroate synthase, structure and *In silico*

despite concerted efforts towards identification of new chemotherapeutic agents, trimethoprim-sulfamethoxazole (TMP-SMX) remains the standard prophylactic and therapeutic modality in use today, with clindamycin-primaquine, atovaquone and pentamidine being standard secondary PCP treatments.[2] With such a limited repertoire of therapeutic options, there is a great potential for developing resistance to these compounds by the pathogen. In this regard, mutations identified in the *P. jirovecii* dihydropteroate synthase and cytochrome bc1 genes, the targets of SMX and atovaquone, have already been associated with resistance in other infections.

The AIDS epidemic, however, marked the beginning of the disease's impact on a substantial number of patients. PCP has long been the most common serious AIDS-defining opportunistic infection in the United States. The introduction of highly active antiretroviral therapy (HAART) for the treatment of HIV infection has been accompanied by substantial reductions in mortality and the incidence of opportunistic infections, including PCP [3]. *Pneumocystis* has also been found in immunosuppressed patients without PCP [4]. Sulfa drugs play a key role in *P. carinii* pneumonia treatment and prophylaxis. Cotrimoxazole, a combination of sulfamethoxazole and trimethoprim (Bactrim), is the most widely used drug for therapy and prophylaxis. The antipneumocystis activity of cotrimoxazole is due almost entirely to sulfamethoxazole. Other sulfa drugs, such as dapsone and sulfadoxine, have also been used [5]. The enzymatic target for sulfa drugs is dihydropteroate synthase (DHPS). In *P. carinii*, DHPS is part of a trifunctional protein along with dihydroneopterin aldolase and hydroxymethyl dihydropterin pyrophosphokinase, two other enzymes in the folic acid biosynthesis pathway [6]. Prior to the AIDS epidemic, PCP was an infrequent infection and attracted little attention from the scientific or clinical communities, subsequently there was a poor understanding of its basic biology,

1. Introduction

The disease known as *Pneumocystis carinii* pneumonia (PCP) is one of the leading causes of illness and death in persons with impaired immunity. The disease has been described in immunocompromised patients for many years, including outbreaks in malnourished young children in orphanages in Iran in the 1950s[1]. *Pneumocystis jirovecii* is a fungus that causes *Pneumocystis* pneumonia (PCP) in humans, which remains a leading opportunistic infection associated with AIDS patients, even in the era of Highly Active Anti-Retroviral Therapy (HAART). Moreover, PCP increasingly targets new groups of patients with underlying chronic diseases states, such as Chronic Obstructive Pulmonary Disorder (COPD). However,

pathogenic mechanisms, and few treatment options. There were considerable impediments to standard experimental approaches for investigation of this unique pathogen, the most problematic being the lack of a long term culture system. This problem remains today, and many fundamental scientific questions about the basic biology, metabolism, and life cycle of *Pneumocystis* remain unanswered. At present, *Pneumocystis* species have been found in a large variety of mammalian species, each of which has its own species of the fungus, i.e. species of *Pneumocystis* do not proliferate when transferred from the host in which they are found to a different host. When a host immune system is compromised, oftentimes the fungi grow to fill the alveolar lumens and effectively block oxygenation, leading to host death. [2]

2. MATERIALS AND METHODS

2.1 Retrieval of Dihydropteroate synthetase sequence

Dihydropteroate synthetase protein was responsible for *Pneumocystis pneumonia* sequence was retrieved from UniprotKB database [7]. UniprotKB (www.uniprotkb.org) is the central hub for the collection of functional information of proteins.

2.2 Primary structure analysis

The primary structure of dihydropteroate synthetase (DHPS) was analyzed using ProtParam tool [8]. ProtParam calculated physio-chemical characterization, theoretical isoelectric point (pI), molecular weight, total number of positive and negative residues, extinction coefficient [9] instability index [10] aliphatic index and grand average hydropathy (GRAVY) [11] were computed (<http://expasy.org/cgi-bin/protparam>).

2.3 Secondary structure prediction

The secondary structure of dihydropteroate synthetase was predicted using FASTA sequence of by SOPMA method [11]. It was employed for calculating the alpha helix, beta sheet and coil or turns of the selected protein sequence. The SOPMA method available on (http://npsa-pmat.pl?page=/NPSA/npsa_sopma.html)

2.4 Protein functional sites

InterPro scan and Fingerprint scan are the tools used to predict the signatures and the motif regions in the dihydropteroate synthetase sequence.

2.5 Homology Modeling

The dihydropteroate synthetase sequence was subjected for comparative homology modeling via SWISS-MODEL server according to the Arnold method [12] which generate putative 3D model. SWISS-MODEL is fully automated protein structure homology modeling server to make the protein modeling accessible to all biotechnologist. The SWISS MODEL performs the sequence alignments and searches for the putative template protein for generating the 3D model.

2.6 Protein Structure validation

The predicted protein structure of dihydropteroate synthetase was validate by PROCHECK method.[13]. PROCHECK generate the ramchandran plot and checks the stereo chemical quality of a protein structure, producing a number of PostScript plots analyzing its overall residue-by residue geometry. It includes PROCHECK NMR for checking the quality of structures solved by NMR. The structure was visualized and analyzed in Chimera visualization tool.

3. RESULTS and DISCUSSION

3.1 Protein primary sequence analysis:

The primary structure of Dihydropteroate synthetase was predicted using Expasy's ProtParam server (<http://expasy.org/cgi-bin/protparam>). The result revealed that Dihydropteroate synthetase from *Plasmodium falciparum* (uniprot id-Q8IAU3)had 706 amino acid residues and molecular weight 83373.9. The maximum number of amino acids present in the sequence was found to be Asn (12.5%) and least was that of Trp (0.3%). The total number of negatively charged residues (Asp+Glu) was 100 and the total number of positively charged residues (Arg+Lys) was 98. The isoelectric point pI was 6.73, protein is acidic in nature. The high aliphatic index

(99.15). While instability index 36.96. It classifies that protein is stable. The grand average hydropathicity (GRAVY) is very low -0.486.

3.2 Secondary structure prediction:

The secondary structure is composed of alpha helix, beta sheets and coil. It is predicted by SOPMA as shown in Table-1. The secondary structure prediction was done and random coil was found to be 40.98% followed by extended strand 26.33% [Fig-1].

3.3 Homology modeling and structure validation:-

The protein 3D structure was built using SWISS-MODEL automated homology or comparative modeling server. The template 2bmb (Folic acid synthesis FOL1) is used which shows 28.12% sequence identity with target sequence. Target and template sequence alignment was done alignment was done. The energy was minimized using Swiss-PDBviewer. The predicted protein Dihydropteroate synthetase model validate by PROCHECK analysis. The Ramchandran plot in PROCHECK [Fig-2] it shows model quality or most favoured region is 83.2%. The predicted protein model was visualized by UCSF Chimera visualization tool as shown in [Fig-3]. The domain analysis was done using Pfam database and functional domains were obtained shown in [Fig-4]. As figure shows HPPK and pterin binding enzyme functional or domain site in dihydropteroate synthetase protein sequence. The predicted three dimensional structure of dihydropteroate synthetase deposited in Protein model database with PMDB id:PM0080376.

3.4 Protein functional sites

Interpro scan and Fingerprint scan were the tools used to predict the signatures and the motif regions in the protein as shown in the Table-2.

4. CONCLUSION

The *In silico* approach enables rapid potential drug target identification, thereby greatly facilitating the search for new antibiotics. The mechanism of dihydropteroate synthetase it reveal to design novel inhibitors in *Pneumocystis Pneumonia*. In this present study we tried to model three dimensional structure of dihydropteroate synthetase for the structure based drug design.

[V] REFERENCE:

1. James R. Stringer,* Charles B. Beard,† Robert F. Miller,‡ and Ann E. Wakefield§1 A New Name (*Pneumocystis jiroveci*) for *Pneumocystis* from Humans Emerging Infectious Diseases Vol. 8, No. 9, September 2002
2. Aleksey Porollo1,* , Jaroslaw Meller1,2, Yogesh Joshi1, Vikash Jaiswal1, A. George Smulian1,3, and Melanie T. Cushion1,3 Analysis of Current Antifungal Agents and Their Targets within the *Pneumocystis carinii* Genome *Curr Drug Targets*. 2012 November ; 13(12): 1575–1585.
3. Kaplan JE, Sepkowitz K, Masur H, Sirisanthana T, Russo M, Chapman L. Opportunistic infections in persons with HIV or other immunocompromising conditions. *Emerg Infect Dis* 2001;7(3 Suppl):541.
4. Nevez G, Raccurt C, Jounieaux V, Dei-Cas E, Mazars E. *Pneumocystis* versus pulmonary *Pneumocystis carinii* colonization in HIV-negative and HIV-positive patients. *AIDS*. 1999;13:535–56.
5. Brian R. Lane, Jennifer C. Ast, Paul A. Hossler, David P. Mindell, Marilyn S. Bartlett, James W. Smith, and Steven R. Meshnick Dihydropteroate Synthase Polymorphisms in *Pneumocystis carinii* *The Journal of Infectious Diseases* 1997;175:482-5.
6. Volpe F, Dyer M, Scaife JG, et al. The multifunctional folic acid synthesis fas gene of *Pneumocystis carinii* appears to encode dihydropteroate synthase and hydroxymethyldihydropterin pyrophosphokinase. *Gene* 1992; 112:213-8.
7. Magrane, M, and the UniProt consortium, (2009), UniProt Knowledgebase: a hub of integrated protein data Database, 2011: bar009.
8. Gasteiger, E, Hooglan, C, Gattiker, A, Duvaud, S, Wilkins, M. R, Appel R. D, Bairoch, A, John, M, Walker, (2005), *The Proteomics Protocols Handbook*, Human Press. 571-607.
9. Gill, S. C, and VonHippel, P. H, (1989), Calculation F Protein Extinction Coefficients from Amino acid sequence data, *Analytical Biochem*. 182:319-326
10. Guruprasad, K, Reddy, B. V, Pandit, M.W, (1990), Correlation between Stability of a Protein and its dipeptide composition: a novel approach for prediction in vivo stability of a protein from its Primary sequence, *Protein Engineering*. 4:155-161.
11. Kyte, J, and Doolittle, R. F, (1982), A simple method for displaying the hydrophobic character of a protein, *Journal of molecular Biology*. 157:105-132.
12. Kiefer, F, Arnald, K, Kunzli, M, Bordoli, L, Schwede, T, (2009), The SWISS-MODEL repository and associated resources, *Nucleic acids resources*. 37:387-392
13. Laskowski, R. A, MacArthur, M. W, Moss, D. S. Thornton, J. M, (1993), PROCHECK: program to check the stereochemical quality of protein structures, *J. Appl. Cryst*. 26:283-291.

Table 1:-Secondary structure prediction of Dihydropteroate synthetase by SOPMA

Secondary Structure	SOPMA
Alpha Helix	45.61%
310 Helix	0.00%
Pi Helix	0.00%
Beta Bridge	0.00%
Extended Strand	23.94%
Beta Turn	7.79%
Bend Region	0.00%
Random Coil	22.66%
Ambiguous States	0.00%
Other States	0.00%

Fig.1:-Secondary structure of Dihydropteroate synthetase.

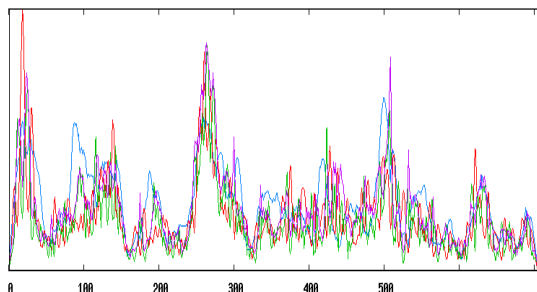
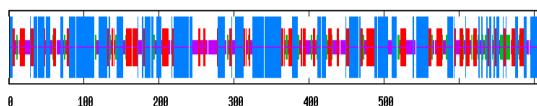
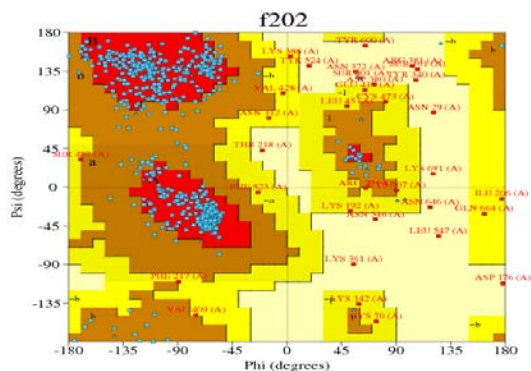


Fig. 2: Ramchandran Plot analysis of Dihydropteroate synthetase



	No. of residues	%-tage
Most favoured regions [a,b,l]	546	83.2%*
Additional allowed regions [a,b,l,p]	77	11.7%
Generously allowed regions [-a,-b,-l,-p]	18	2.7%
Disallowed regions [XX]	15	2.3%*
Non-glycine and non-proline residues	656	100.0%
End-residues (excl. Gly and Pro)	2	
Glycine residues	18	
Proline residues	15	
Total number of residues	691	

Fig. 3:-The three dimensional structure of Dihydropteroate synthetase

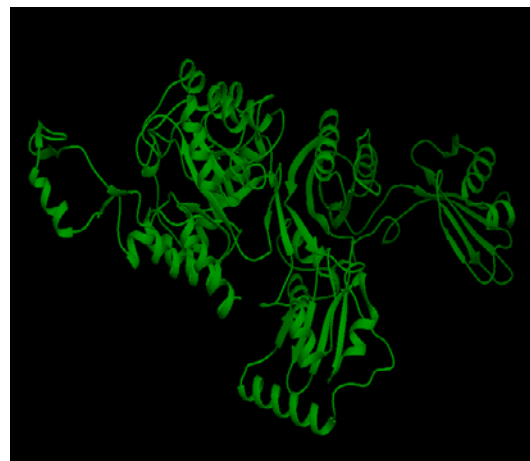


Fig.

4:-

Domain analysis of Dihydropteroate synthetase from pfam database.

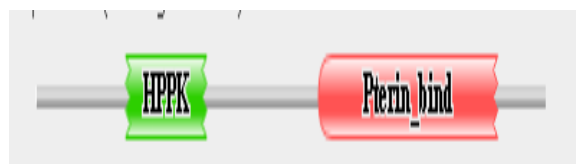


Table 2:-Fingerprint scan result of Dihydropteroate synthetase.

Finger PRINT	No. of motifs
VACCYTOXIN	15
LVDCCALPHA1C	13
MG045FAMILY	9
NISCPROTEIN	12
INTERLEUKIN	4
FOLLOCULNIP1	5
TRPCHANNEL6	5
VEGFRECPTR2	7
GLYCHROMONER	8
CYTOFMRINTP	12