*International Research Journal of Biological Sciences* _____ ISSN 2278-3202
Vol. **3(6),** 29-40, June **(2014)** *Int. Res. J. Biological Sci.*

# Computational Approach to Explore *Plasmodium falciparum* 3D7 Seryl tRNA Synthetase Structure

**Kasturi K.*, Mallika D.S., Jeevan Amos S., Pavithra V. and Sambasiva Rao K.R.S.**
Department of Biotechnology, Acharya Nagarjuna University, Guntur, AP, INDIA

**Available online at: www.isca.in, www.isca.me**
Received 8th November 2013, revised 24th January 2014, accepted 10th February 2014

## Abstract

*Malaria remains the leading cause of deaths attributable to a communicable disease globally. The reemergence of drug-resistant to Plasmodium falciparum, which is the most fatal human cerebral malarial parasite, has focused attention on aminoacylation in protein translation. Seryl tRNA Synthetase is the enzyme involved in acylation of Serine amino acid to tRNA. The protein sequence of Seryl tRNA synthetase containing 539 residues was obtained from UniprotKB. The enzyme was characterized using computational tools. The secondary and tertiary (3D) structure of the enzyme was predicted using Seryl tRNA Synthetase from other species and validated by various structural quality validation tools. Results of the current study will provide a deep insight about the structure and function of Seryl tRNA synthetase and aid in drug docking, rational drug designing.*

**Keywords:** *P.falciparum*, aminoacyl tRNase, Seryl tRNA synthetase, 3D structure modelling, cerebral malaria.

## Introduction

Malaria is the significant health problem affecting millions of people and thousands of lives are at risk every year. It is caused by five species of protozoan parasites of the genus *Plasmodium* that affect humans (*P. falciparum, P. ovale, P. vivax, P. malariae and P. knowlesi*). The data from the world malaria report 2011 summarizes that among 106 malaria-endemic countries from 6 regions, *P. falciparum* is the major malarial pathogen causing deaths[1]. The life cycle of the *Plasmodium* is complex with two major phases: liver and blood. As the infected mosquito bites a human host, the *Plasmodium* sporozoites migrate to the liver and infect hepatocytes. After replicating within the hepatocyte, the parasites rupture the cell to release merozoites, a stage specialized to infect erythrocytes. However, two subspecies (*P. vivax and P. ovale*) can assume a dormant state (hypnozoites) within the liver that commonly lasts for months and can last for years[2]. After exiting the liver, the parasite establishes a recurring life cycle in the erythrocytes. During the blood stage, some parasites will differentiate into sexual forms that are transmitted to the next human host by the vector. Each developmental stage is characterized by distinct physiology, and each has varying sensitivity to most drugs[3,4]. To date stage specific antimalarial chemotherapy is applied for the control and treatment[5,6]. But the emergence of multi drug resistant strains of *Plasmodium* towards commercial drugs has made the treatment of disease more critical. The development of resistance occurs in two phases i.e., the development of resistant mutant and the second is multiplication of the resistant parasite resulting in the parasite population that is no longer suitable for treatment. Thus the situations have resulted in the identification of novel drug targets with novel mechanism within the parasite.

Basing on the concepts of Flux balance analysis (FBA) to assess the essentiality of genes for an organism[7-9] and concepts of choke-points and load points for estimating the reactions that are essential for the organism[10,11] *P. falciparum seryl tRNA synthetase* is being considered as a putative target.

*P. falciparum seryl tRNA synthetase* (PF07_0073) belongs to class II aminoacyl-tRNA's synthetase. It catalyses the reaction between amino acid L-serine and tRNA to form L-seryl-tRNA, which is necessary for the synthesis of various proteins. Inhibition of this enzyme is likely to affect the aminoacyl tRNA biosynthesis pathway of the parasite. Unavailability of the 3 - dimensional structure of enzymes is one of the major hindrances in elucidating the interactions of enzymes with possible inhibitors. Comparative homology modeling is promulgated as the most unswerving computer-based technique for deciphering the 3D structures in the absence of the crystal structure of the protein. This article describes the modeling and validation of *seryl tRNA synthetase* of *P. falciparum* which will provide insight into its structure and aid in drug designing.

## Methodology

**Selection of Malarial Parasite:** *P. falciparum 3D7* was selected as the candidate organism for the present study because of 4 reasons. i. Among the five species of *Plasmodium*, *P.falciparum* was reported **as highest** mortality in Africa as well as in South- East Asian regions according to WHO malaria report 2011 (table 1, figure 1). ii. It causes cerebral malaria which is a dreadful disease. iii. Complete genome sequence of *P. falciparum* is available. iv. *P. falciparum* showing high frequency of resistance to available drugs in the market. Resistance towards the presently available drugs was due to

improper identification of *P. falciparum* and abnormal drug dosage without specificity to the *P. falciparum* and finally parasite becomes resistant to the available drugs very rapidly. So there is a need for the emergence of new drugs with new targets. This is the main motive of our present study to choose *P. falciparum* as a target sps.

**Sequence analysis of Seryl tRNA synthetase of *Plasmodium falciparum* 3D7:** Seryl tRNA synthetase protein sequence (Q8IBS3) was collected from UniprotKB, FASTA formatted sequence was retrieved (figure-2) and submitted to pBLAST for similar structural protein against the PDB-RCSB database keeping default parameters like E-value threshold 10, word size 3 and Blosum 62 Matrix. Since the BLAST algorithm detects local as well as global alignments, regions of similarity embedded in otherwise unrelated proteins can be detected[12,13]. The BLAST result page contains template sequences which were showing structural homology to query sequence (table 2). Among them Seryl-tRNA Synthetase (PDB id: 3QNE_a) from *Candida albicans* was taken as a template for the present study because it is having an E-value (2e-114) and query coverage (99%), Max identity (38%) were satisfactory. The template protein sequence was derived from uniprotKB for further analysis. Multiple sequence alignment was performed using CLUSTAL_X2[14.15]. Regions of conservation and variation were detected from CLUSTAL_X2 result (figure 3).

**Functional Characterization of Seryl tRNA synthetase of *Plasmodium falciparum* 3D7:** Functional characterization of SeryltRNA synthetase protein sequence was done by finding motif using Eukaryotic Linear Motif and domain analysis was carried out using PROTSCAN.

**ELM**s are very short motifs nothing but regular expressions of a protein, many of them will overpredict, implying that most matches shown are more likely to be false positives than true matches[16]. To improve the predictive value of ELM, logical filters (or rules) are used based on context information to discriminate between likely true and false positives.

**PROSCAN** is a tool to scan a sequence against the PROSITE database. That is to say, to search for biologically relevant sites and signatures. By its algorithm (shared with PATTINPROT), PROSCAN allows searching with errors. Errors are set by number of mismatches allowed or by a level of significant biological information. This tool is available in NPS (http://npsa-devel.ibcp.fr/NPSA/npsa_proscan.html) [17]. All results were tabulated in table 3 and table 4, figure 4.

**Physico-Chemical Characterization of Seryl tRNA synthetase of *Plasmodium falciparum* 3D7:** The basic physico-chemical properties of the seryl tRNA synthetase protein sequence were calculated using the ProtParam tool[18]. Molecular weight, theoretical pI, amino acid composition, atomic composition, extinction coefficient, estimated half-life, instability index, aliphatic index and grand average of hydropathicity[19] (GRAVY) was computed by ProtParam. Results were tabulated in table 5, 6 and 7.

**Secondary Structure prediction of *P. falciparum 3D7* Seryl tRNA Synthetase:** GOR-IV tool was used to obtain the secondary structure of seryl tRNA synthetase protein (**http://npsa-pbil.ibcp.fr/cgi-bin/npsa_automat.pl? page= npsa_gor4.html**)[20]. Secondary structure prediction is the definition of each residue into either alpha helix, beta sheet or random coil secondary structures. These results were tabulated in table 8 and secondary structural details were shown in figure 5.

**3D structure prediction of *P. falciparum 3D7* Seryl tRNA Synthetase:** 3D structure prediction of *P. falciparum 3D7* seryl tRNA synthetase protein was performed using **Modeller 9.10** (http://salilab.org/modeller/)[21], it generates model based on comparative or homology method using user provided alignment of a sequence to be modeled with known related structures. MODELLER automatically calculates a model containing all non-hydrogen atoms and generates a model by satisfaction of spatial restraints[22,23,24] and other web based servers such as Geno3D, ESYPred 3D, 3DJIGSAW, SWISS MODEL and CPHmodel were also used to analyze the generated model quality (table 9).

**Geno 3D:** It predicts the 3D structure of an amino acid sequence, an automated protein modeling Web server is used to generate protein 3D model. Gene3D is supplementary to the CATH database, contains proteins from complete genomes which have been clustered into protein families and annotated with CATH domains, Pfam domains and functional information from KEGG, GO, COG, Affymetrix and STRINGS[25].

**ESyPred3D:** ESyPred3D is a new automated homology modeling program. The method gets benefit of the increased alignment performances of a new alignment strategy using neural networks. Alignments are obtained by combining, weighting and screening the results of several multiple alignment programs. The final 3D structure is built using the modeling package MODELLER[26].

**3D-JIGSAW:** It is an automated system to build three-dimensional models for proteins based on homologues of known structure from databases (PFAM+PDB+nr) and splits the query sequence into domains. From the template hits, the best covered domain is used to model the query sequence; maximum of 2 models can be generated by this method[27-29].

**SWISS MODEL:** It is a fully automated protein structure homology modeling server accessible via the expasy web server. The purpose of this server is to make protein modeling accessible to all biochemist and molecular biologist's world wide[30,31,32].

**CPH Model:** CPHmodels-3.0 is a web-server predicting protein 3D-structure by use of single template homology modelling[33]. It is a group of high performing 3D-prediction tools. Beside its accuracy, one of the important features of the method is its speed. For most queries, the response time of the server is less than 20 minutes.

**Evaluation of the obtained models of Seryl tRNA Synthetase protein:** The result was evaluated using different web based validation servers such as PROCHECK and ProSA.

**PROCHECK:** This program assesses the "stereo-chemical quality" of a given protein structure. The aim of PROCHECK is to assess how normal or how unusual, the arrangement of the residue geometry in a given protein structure is, as compared with stereo-chemical parameters derived from well-refined, high-resolution structures[34,35]. Results from other web based servers were compared.

**ProSA:** which is frequently employed in protein structure validation, it requires the atomic coordinates (PDB file format) of the model to be evaluated. It uses only the C-alpha atoms of the input structure, hence it can also be applied to low resolution structures and approximate models obtained early in the structure determination process[36].

## Results and Discussion

**Selection of Malarial Parasite:** *P. falciparum* was selected as target parasite because of its existence. Most of the WHO regions had high intensity of *P. falciparum* infections were reported. Statistical data of this was tabulated in table 1, graphical representation was given in figure 1. Generally WHO divides total 101 malaria affected countries in 6 regions, among these regions highest rates of *P. falciparum* infections were reported from the African region and least was reported in the European region.

**Table-1**
**Percentage of existence of Malarial Parasite according to WHO Malaria report- 2011**

| Percentage of Existence of Malarial Parasite | | |
|---|---|---|
| **Regions** | ***P. faiciparum*** | **Others** |
| South east Asia | 44.9 | 56.1 |
| African | 96.795 | 3.315 |
| Eastern Mediterranean | 59.1 | 41.9 |
| Western Pacific | 54.5 | 45.5 |
| European | 0.166 | 99.944 |
| American | 21.428 | 79.682 |

**Selection of Target Protein and Retrieval:** Seryl tRNA synthetase is an enzyme involved in acylation of serine to trna in protein translation. This is a putative drug target to design lead molecules for the inhibition of growth of *P. falciparum.* The protein sequence of target protein was retrieved from Uniprotkb as given in figure 2.

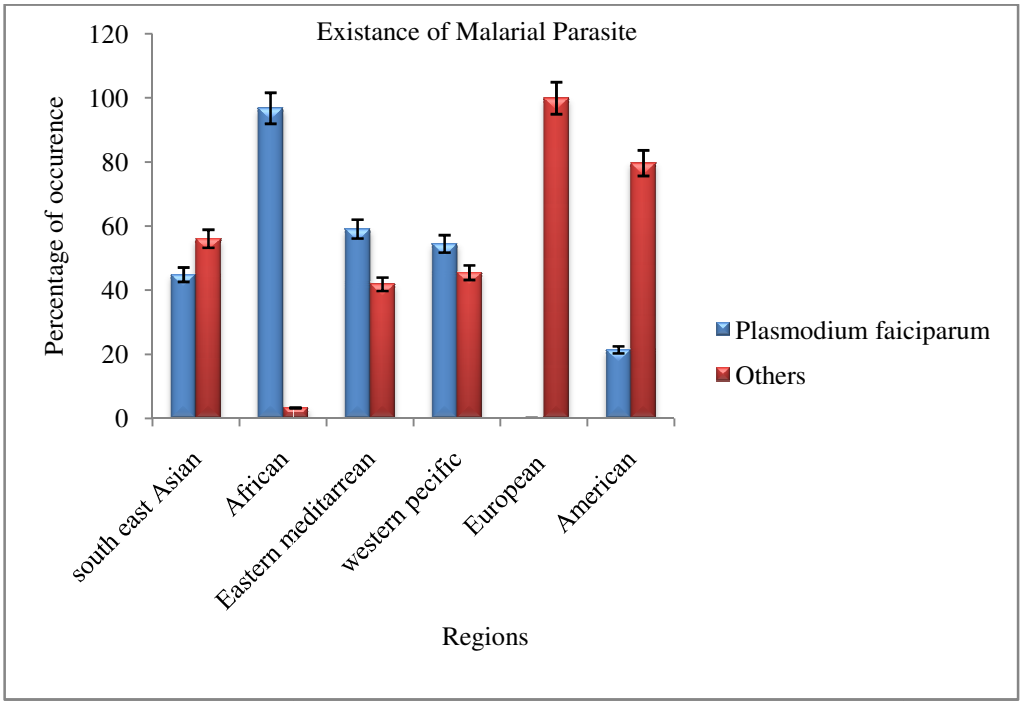


**Figure-1**
**Graphical representation of statistical data of existence of Malarial parasite**

**Sequence analysis of Seryl tRNA synthetase of *Plasmodium falciparum* 3D7: Blast Results:** Upon performing the pBLAST total top 100 hits were listed in the result page as descending order of similarity score. Among them first 10 with the PDB ID of BLAST hits, similarity score, query coverage and e values were tabulated in table 2. The first and best similarity structure is 3VBB which is a human protein not taken as template because the generated model was for human pathogen if human protein is taken as a template then there was chance to inhibit the human protein also when the lead molecule used for target protein. So by considering this even though it shows highest similarity, it was not taken as a template to generate the model.

**Clustal X2 Results:** From the BLAST results 3QNE was taken as a template for the target protein to generate structure model by comparative modeling. This template had queried coverage, Similarity score and e-values 99%, 351 and 2e-114 respectively were quite satisfactory. So the 3QNE protein sequence was retrieved and subjected to Multiple Sequence Alignment with the target or query protein using the Clustal X2 tool. The result of Clustal X2 was reported in figure 3.

>tr|q8ibs3|q8ibs3_plaf7 seryl-trna synthetase, putative os=plasmodium falciparum (isolate 3d7) gn=pf07_0073 pe=3 sv=1
MVLDINLFRKEKGGNPDKIKESERKRYHDENNVDKVIEYDDKWRKCIFELEELKKNINMINKEIGNK
KKVDKNADVEDLKKKSLNIKEEIPKYQLKEKELLKERNKYISKIGNLLNIKVVCSDNEDNNKIVKTW
GECKILPACEENDNSIHDNVVNSNNIKRETLNNEVDNKKKIKYYYHYDLLRKIGGANFKKGIQVAG
HRGYYLTGAGFLLHNAILQYALNFLVNKKYIPVYPPFFMKKNIMEECAELDDFEETLYKIPSTSNSTL
SSQQVSTSPTKISSQADIKDDTTCNSQKKTNIPSNEDLTRDDLFLIATSEQPLCALHKDETIESKRLPLK
YAGFSSCFRKEAGAHGKDIRGILRVHQFDKVEQFCIALPQHSNKIHEEMIQTCEEFYQSLNIPYRIVSI
VSGALNNAASIKYDLEGFFPTSNQYRELVSCSNCTDYQSINLNIRYSDSSIKINDLNKNTNLNDEMDS
EYEHFLTNFNTENKYHVHLLNGTMVAAQRFLCCLLENYQNGEGIVVPEKLRPYMNNMDFIPFME

**Figure-2**
**Sequence of Target protein: Seryl tRNA Synthetase**

**Table-2**
**pBlast result- First 10 hits of the query protein**

| PDB ID | Query Coverage (%) | Similarity Score | E-Value | Max Identity (%) |
|---|---|---|---|---|
| 3VBB | 99 | 381 | 3e-125 | 39 |
| 3QNE | 99 | 351 | 2e-114 | 38 |
| 3QQ5 | 99 | 349 | 1e-113 | 38 |
| 3LSQ | 99 | 338 | 4e-109 | 36 |
| 2DQ0 | 99 | 277 | 5e-86 | 32 |
| 2DQ3 | 98 | 234 | 2e-70 | 31 |
| 1SRY | 61 | 164 | 2e-44 | 28 |
| 3ERR | 61 | 166 | 2e-44 | 28 |
| 1WLE | 65 | 133 | 1e-33 | 26 |
| 3MEY | 13 | 31.6 | 1.3 | 24 |

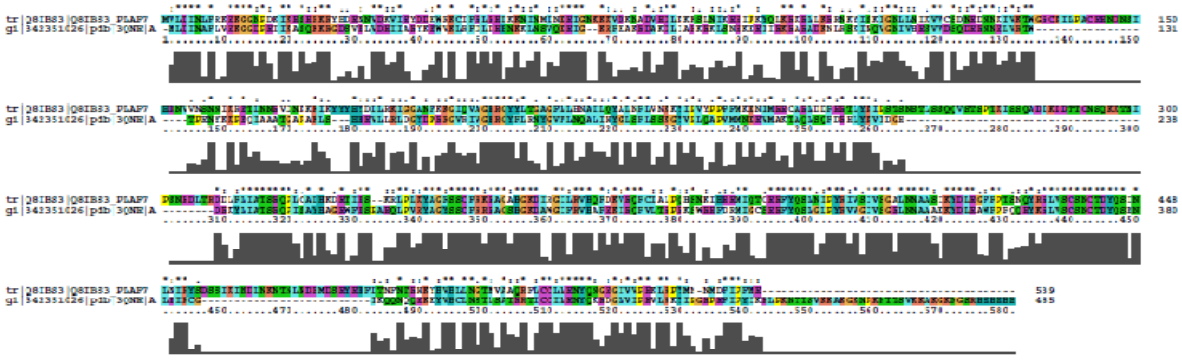

**Figure-3**
**Multiple sequence alignment of Target with 3QNE using Clustal X2 tool**

**Functional Characterization of Seryl tRNA synthetase of *Plasmodium falciparum* 3D7:** The function of Sery tRNA synthetase protein of *P. falciparum 3D7* was analyzed by submitting the amino acid sequence to ELM server and Proscan server. Based on EML and PROSCAN results, the following motif and domains were assigned to target protein amino acid sequence (table 3 and 4, figure 4).

**Table-3**
**Motif analysis of *P. falciparum 3D7* Seryl tRNA synthetase**

| Element Name | Position | Element Description | Pattern |
|---|---|---|---|
| LIG_CYCLIN_1 | 138-141 [A]<br>433-436 [A] | Substrate recognition site that interacts with cyclin and thereby increases phosphorylation by cyclin/cdk complexes. Predicted protein should have the MOD_CDK site. Also used by cyclin inhibitors | [RK].L.{0,1}[FYLIVMP] |
| LIG_eIF4E_1 | 178-184 [A] | Motif binding to the dorsal surface of eIF4E. | Y....L[VILMF] |
| LIG_FHA_2 | 131-137 [A] | Phosphothreonine motif binding a subset of FHA domains that have a preference for an acidic amino acid at the pT+3 position. | ..(T)..[DE]. |
| LIG_TRAF2_1 | 142-145 [A] | Major TRAF2- binding consensus motif. Members of the tumor necrosis factor receptor (TNFR) superfamily initiate intracellular signaling by recruiting the C-domain of the TNFR-associated factors (TRAFs) through their cytoplasmic tails | [PSAT].[QE]E |
| MOD_N-GLC_1 | 439-444 [A]<br>495-500 [A] | Generic motif for N-glycosylation. Shakin-Eshleman et al. showed that Trp, Asp, and Glu are uncommon before the Ser/Thr position. Efficient glycosylation usually occurs when ~60 residues or more separate the glycosylation acceptor site from the C-terminus | .(N)[^P][ST].. |
| MOD_PKA_1 | 161-167 [A] | Main preference for PKA-type AGC kinase phosphorylation. | [RK][RK].([ST])[^P].. |
| MOD_PKA_2 | 161-167 [A]<br>451-457 [A] | Secondary preference for PKA-type AGC kinase phosphorylation. | .R.([ST])[^P].. |
| MOD_PLK | 146-152 [A]<br>454-460 [A] | Site phosphorylated by the Polo-likekinase | .[DE].([ST])[ILFWMVA]. . |
| MOD_SUMO | 160-163 [A] | Motif recognised for modification by SUMO-1 | [VILMAFP](K).E |
| TRG_ENDOCYTIC_ 2 | 181-184 [A]<br>432-435 [A]<br>444-447 [A]<br>477-480 [A] | Tyrosine-based sorting signal responsible for the interaction with mu subunit of AP (Adaptor Protein) complex | Y..[LMVIF] - - |
| TRG_NLS_Bipartite _1 | 161-177 [A] | Bipartite variant of the classical basically charged NLS. | [KR][KR].{7,15} [^DE]((K[RK])\| (RK))(([^DE] [KR])\| ([KR][^DE]))[^DE] |
| TRG_NLS_MonoExt C_3 | 171-176 [A] | Monopartite variant of the classical basically charged NLS. C-extended version. | [^DE]((K[RK])\| (RK))(([^DE] [KR])\| ([KR][^DE])) (([PKR])\|([^DE] [DE])) |
| TRG_NLS_MonoExt N_4 | 172-177 [A] | Monopartite variant of the classical basically charged NLS. N-extended version. | (([PKR]. {0,1}[^DE])\| ([PKR]))((K[RK])\| (RK))(([^DE] [KR])\| ([KR][^DE]))[^DE |

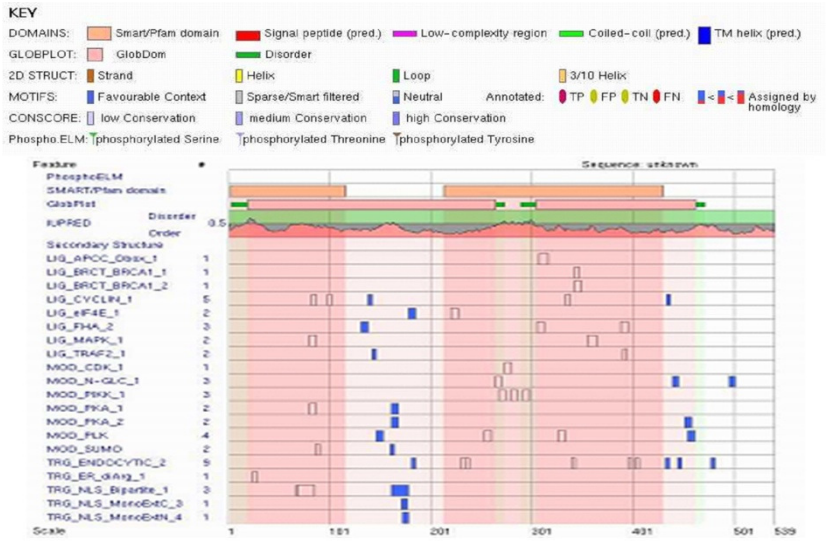


**Figure-4**
**Schematic representation of different types of predicted regions of target protein**

**Table-4**
**Domains obtained in *P. falciparum 3D7* Seryl tRNA synthetase Protein using PROSCAN**

| Domine name | Prosite Accession Number | Randomized probability | Start position | End position | Sequence of Prosite | Pattern |
|---|---|---|---|---|---|---|
| N-Gylcosylation Site | PS00001 | 5.138e-03 | 255 | 268 | NSTL | N-{P}-[ST]-{P} [N is Gylcosylation site] |
| | | | 440 | 443 | NCTD | |
| | | | 496 | 499 | NGTM | |
| cAMP- and cGMP-dependent protein kinase phosphorylation site | PS00004 | 1.572e-03 | 80 | 83 | KKKS | [RK](2)-x-[ST] [S or T is the phosphorylation site] |
| | | | 161 | 164 | KRET | |
| Protein kinase C phosphorylation site | PS00005 | 1.423e-02 | 22 | 24 | SER | [ST]-x-[RK] [S or T is the phosphorylation site] |
| | | | 294 | 296 | SQK | |
| | | | 332 | 334 | SKR | |
| | | | 380 | 382 | SNK | |
| | | | 417 | 419 | SIK | |
| | | | 457 | 459 | SIK | |
| Casein kinase II phosphorylation site | PS00006 | 1.482e-02 | 122 | 125 | SDNE | [ST]-x(2)-[DE] [S or T is the phosphorylation site] |
| | | | 133 | 136 | TWGE | |
| | | | 149 | 152 | SIHD | |
| | | | 282 | 285 | SQAD | |
| | | | 302 | 305 | SNED | |
| | | | 307 | 310 | TRDD | |
| | | | 390 | 393 | TCEE | |
| | | | 475 | 478 | SEYE | |
| Tyrosine kinase phosphorylation site. | PS00007 | 4.074e-04 (min) 4.083e-04 (max) | 20 | 27 | KESERKRY | [RK]-x(2)-[DE]-x(3)-Y or [RK]-x(3)-[DE]-x(2)-Y [Y is the phosphorylation site] |
| N-myristoylation site | PS00008 | 1.397e-02 | 112 | 117 | GNLLNI | G-{EDRKHPFYW}-x(2)-[STAGCN]-{P} [G is the N-myristoylation site] |
| | | | 195 | 200 | GIQVAG | |
| | | | 410 | 415 | GALNNA | |
| | | | 497 | 502 | GTMVAA | |

**Physico chemical characterization of target protein:** ProtParam computes various physico-chemical properties of a protein from its sequence. The molecular weight, theoretical pI, amino acid composition, atomic composition, extinction coefficient, estimated half-life, instability index, aliphatic index and grand average of hydropathicity (GRAVY) were computed by ProtParam listed in table 5, 6, 7.

**Table-5**
**Amino acid composition of Seryl tRNA synthetase of *P. falciparum 3D7***

| Amino acid | Number | Percentage |
|---|---|---|
| Ala (A) | 20 | 3.7% |
| Arg (R) | 18 | 3.3% |
| Asn (N) | 55 | 10.2% |
| Asp (D) | 33 | 6.1% |
| Cys (C) | 14 | 2.6% |
| Gln (Q) | 17 | 3.2% |
| Glu (E) | 46 | 8.5% |
| Gly (G) | 21 | 3.9% |
| His (H) | 13 | 2.4% |
| Ile (I) | 43 | 8.0% |
| Leu (L) | 47 | 8.7% |
| Lys (K) | 58 | 10.8% |
| Met (M) | 10 | 1.9% |
| Phe (F) | 21 | 3.9% |
| Pro (P) | 17 | 3.2% |
| Ser (S) | 34 | 6.3% |
| Thr (T) | 21 | 3.9% |
| Trp (W) | 2 | 0.4% |
| Tyr (Y) | 25 | 4.6% |
| Val (V) | 24 | 4.5% |
| Pyl (O) | 0 | 0.0% |
| Sec (U) | 0 | 0.0% |

**Table-6**
**Atomic composition of *P. falciparum 3D7* Seryl tRNA synthetase**

| Element | Symbol | No. of atoms |
|---|---|---|
| Carbon | C | 2762 |
| Hydrogen | H | 4357 |
| Nitrogen | N | 751 |
| Oxygen | O | 850 |
| Sulfur | S | 24 |

**Secondary structure prediction of target protein:** GOR-IV program that was used to predict secondary structures in *P. falciparum 3D7* suggest that it contained more helices than beta sheets in target protein (figure 5).

Three dimensional structure of target protein of *P. falciparum 3D7* was predicted (figure 6) by modeler 9.10 using a template of Seryl tRNA synthetase of *Candida albicans* (PDB ID: 3QNE*)* keeping default parameters. Other online servers also used to generate the best model. Servers and their templates were listed in table 9. 3vbb is a human protein so in the present study it was not taken as a template.

**Table-7**
**Physico-chemical properties of Seryl tRNA synthetase of *P. falciparum 3D7***

| Property | Value |
|---|---|
| Number of amino acids | **539** |
| Molecular weight | **62453.9** |
| Theoretical pI | **6.53** |
| Total number of negatively charged residues (Asp + Glu) | **79** |
| Total number of positively charged residues (Arg + Lys) | **76** |
| Total number of atoms | **8744** |
| Aliphatic index | **81.74** |
| Grand average of hydropathicity (GRAVY) | **-0.682** |
| Ext. coefficient | **49125** |
| Ext. coefficient | **48250** |
| Estimated half-life | **30 hours** |
| Instability index (II) | **42.72** |
| Formula | **$C_{2762}H_{4357}N_{751}O_{850}S_{24}$** |

**Table-8**
**Percentage and type of Secondary structures present in *P. falciparum 3D7* Seryl tRNA Synthetase**

| Secondary structure | Stretch | Percentage |
|---|---|---|
| Alpha helix (Hh) | 156 | 28.94% |
| 310helix (Gg) | 0 | 0.00% |
| Pi helix (Ii) | 0 | 0.00% |
| Beta bridge (Bb) | 0 | 0.00% |
| Extended strand (Ee) | 105 | 19.48% |
| Beta turn (Tt) | 0 | 0.00% |
| Bend region (Ss) | 0 | 0.00% |
| Random coil (Cc) | 278 | 51.58% |
| Ambiguous states | 0 | 0.00% |
| Other states | 0 | 0.00% |

**Table-9**
**Different servers used to generate best model of the target protein**

| Server | Template |
|---|---|
| Modeller 9.10 | 3QNE |
| Geno3D | 3QNE |
| ESYPred3D | 3QNE |
| 3DJIGSAW | 3ISS |
| Swiss model | 3VBB |
| Cph model | 3VBB |

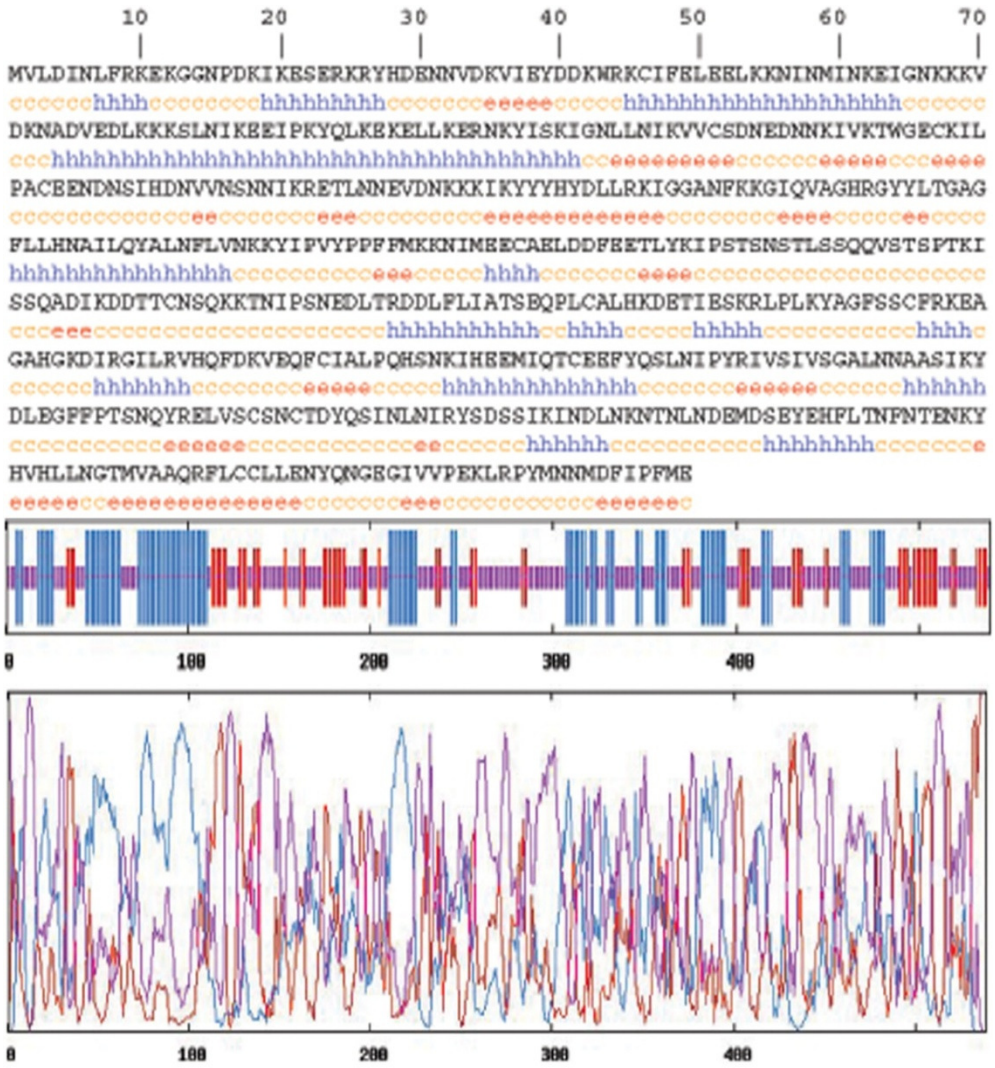


**Figure-5**
**Secondary structure prediction of Target protein of *P. falciparum 3D7* by GOR-IV tool**

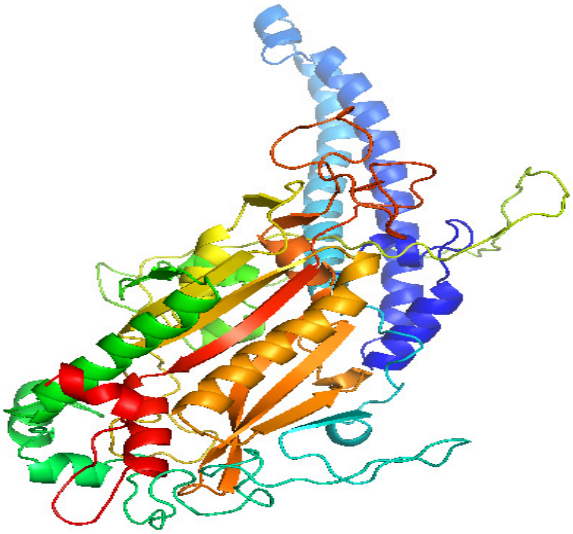

**Figure-6**
**3D structure of *P. falciparum 3D7* Seryl tRNA synthetase**

The model was stereo chemically evaluated using the program PROCHECK. Through the inspection of the Psi/Phi angles of a Ramachandran plot obtained from this analysis (figure 7), the backbone conformation of the model was evaluated. The overall conformation of the backbone was in good agreement with the stereochemistry, which was also found to be reliable (figure 8).

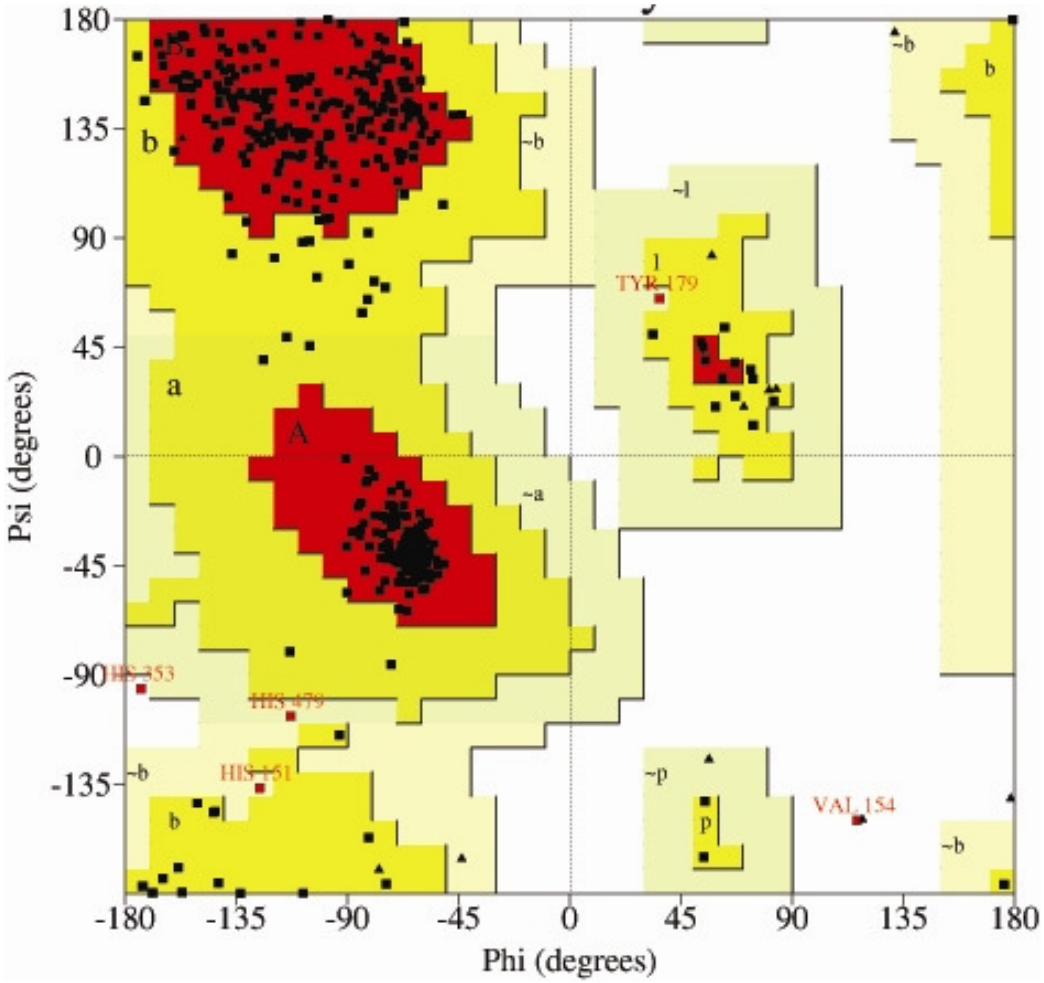


**Figure-7**
**Ramachandran plot of modeled protein obtained from Modeller 9.10 using PROCHECK**

| **Plot statistics** | | |
|---|---|---|
| Residues in most favoured regions [A,B,L] | 441 | 88.4% |
| Residues in additional allowed regions [a,b,l,p] | 53 | 10.6% |
| Residues in generously allowed regions [~a,~b,~l,~p] | 3 | 0.6% |
| Residues in disallowed regions | 2 | 0.4% |
| Number of non-glycine and non-proline residues | 499 | 100.0% |
| Number of end-residues (excl. Gly and Pro) | 2 | |
| Number of glycine residues (shown as triangles) | 21 | |
| Number of proline residues | 17 | |
| Total number of residues | 539 | |
| Based on an analysis of 118 structures of resolution of at least 2.0 Angstroms and R-factor no greater than 20%, a good quality model would be expected to have over 90% in the most favoured regions. | | |

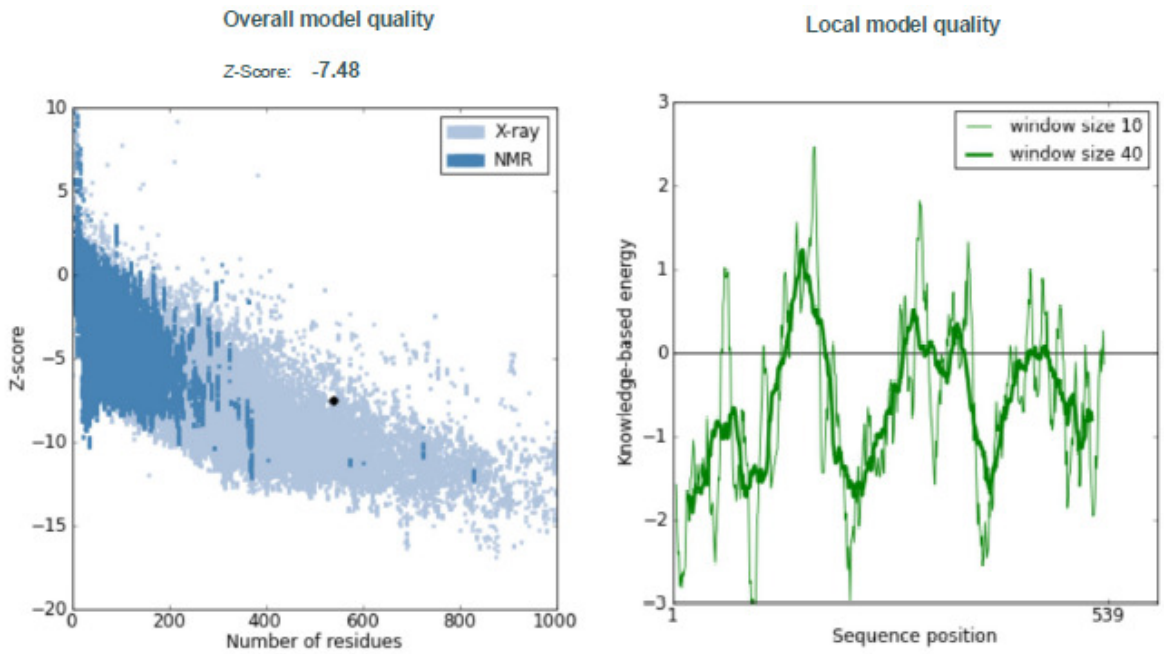


**Figure-8**
**Overall Quality and Local quality of modeled protein obtained from Modeller 9.10 using ProSA with Z- Score**

**Table-10**
**Model quality of Target protein obtained from different servers based on Ramachandran Plot using PROCHECK, Z- Score using ProSA**

| Server | ProSA | PROCHECK | | | |
|---|---|---|---|---|---|
| | Z- Score | Residues in most favored regions (%) | Residues in additional allowed regions (%) | Residues in generously allowed regions (%) | Residues in disallowed regions (%) |
| Modeller 9.10 | -7.48 | 88.4 | 10.6 | 0.6 | 0.4 |
| Geno3D | -7.48 | 64.9 | 25.3 | 5.8 | 4.0 |
| ESYPred3D | -7.48 | 77.1 | 11.6 | 6.7 | 4.6 |
| 3DJIGSAW | -7.48 | 64.9 | 25.3 | 5.8 | 4.0 |

Other web based servers were also employed to generate the 3 dimensional structure of *P. falciparum 3D7* Seryl tRNa synthetase and the results were compared.

Based on these evaluations, we conclude that model obtained using Modeller 9.10 is superior as compared to models generated using other web based servers.

## Conclusion

WHO statistics indicates the severity of malaria, as the pre-eminent tropical disease and it is rated as one of the top three killers among communicable diseases. Emergence of drug resistance strains and the inability of stage specific commercially available anti malarial drugs are one of the greatest coercion to our ability to battle against malaria. The situation continues to be more frightening, with the geographical spread of resistance widening to previously unaffected areas and a ruthless augmentation both in the incidence and degree of drug resistance. Thus there is an urgent need for the development of highly selective and efficacious antimalarial therapies designed against novel Plasmodial targets. The Seryl tRNA synthetase enzyme from P. falciparum represents one such target of interest. A refined model of Seryl tRNA synthetase was generated using the crystal structures of closely related homologues Candida seryl tRNA synthetase. The structural accuracy of the model was extensively validated using protein structure checking tools. This hypothesis describes the modeling and validation of *seryl tRNA synthetase* of *P. falciparum,* which will provide insight into its structure and aid in drug designing.

## Acknowledgement

The authors thank the Department of Biotechnology, Govt. of India for the financial assistance provided to carry out the project work. The authors would also like to thank Acharya Nagarjuna University, Guntur, India for providing the congenial

environment and also for extending the lab facilities to carry out the work.

## References

**1.** World Malaria Report 2011 (Geneva: World Health Organization) WHO, **(2011)**

**2.** White N.J., Determinants of relapse periodicity in Plasmodium vivax malaria, *J. Malaria*., **(10)**, 297 **(2011)**

**3.** Wells T.N., Alonso P.L. and Gutteridge W.E., New medicines to improve control and contribute to the eradication of malaria, *Nat. Rev. Drug Discov*., **(8)**, 879–891 **(2009)**

**4.** Bhatt T.K., Structural Characterization of Histone Deacetylase from Plasmodium falciparum, *ISCA J. Biological Sci*, **(1)**, 65-68 **(2012)**

**5.** Laurence Florens et al., A proteomic view of the *Plasmodium falciparum* life cycle, *Nature*, **(419)**, 520–526 **(2002)**

**6.** Cecile Crosnier, et al., Basigin is a receptor essential for erythrocyte invasion by *Plasmodium falciparum, Nature.,* **(480)**, 534–538 **(2011)**

**7.** Becker, S.A., Feist, A.M., Mo, M.L., Hannum, G., Palsson, B.O. and Herrgard, M.J., 'Quantitative prediction of cellular metabolism with constraint-based models: The COBRA Toolbox', *Nat. Protoc*., **(2)**,727–738**(2007)**

**8.** Edwards J.S. and Palsson B.O., Metabolic flux balance analysis and the in silicoanalysis of *Escherichia coli* K-12 gene deletions, *BMC Bioinf.*, **(1:1)**, **(2000)**

**9.** Kauffman K.J., Prakash P. and Edwards J.S., Advances in flux balance analysis, *Curr. Opin. Biotechnol*., **(14)**, 491–496 **(2003)**

**10.** Puntervoll P., Linding R., Gemünd C., Chabanis-Davidson S., Mattingsdal M., Cameron S., Martin D. M. A., Rahman S.A. and Schomburg D., Observing local and global properties of metabolic pathway - load points and choke points in the metabolic networks, *Bioinformatics,* **10(22)**, 1767–1774 **(2006)**

**11.** Yeh I., Hanekamp T., Tsoka S., Karp P.D. and Altman R.B., Computational analysis of *Plasmodium falciparum* metabolism: organizing genomic information to facilitate drug discovery, *Genome Res*., **(14)**, 917–924 **(2004)**

**12.** Altschul S.F., Madden T.L., Schäffer A.A., Zhang J., Zhang Z., Miller W. and Lipman D.J., Gapped BLAST and PSI-BLAST: a new generation of protein database search programs, *Nucleic Acids Res.*, **(25)**, 3389-3402 **(1997)**

**13.** Katrina C. Martinez, Franco G. Teves and Ma. Reina Suzette B. Madamba, Sequence Analysis of Putative luxS Gene Involved in Prodigiosin Biosynthesis from Philippine Local Strains of Serratia marcescens, *ISCA J. Biological Sci*, **(2)**, 13-19 **(2013)**

**14.** Thompson J.D., Gibson T.J., Plewniak F., Jeanmougin F. and Higgins D.G., The Clustal X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools, *Nucleic Acids Research*, **(25)**, 4876-4882 **(1997)**

**15.** Maithri S.K., Ramesh K.V., Dieudonné Mutangana and Deshmukh Sudha., Molecular Modeling and Docking Studies of PirB Fusion Protein from Photorhabdus Luminescens, *ISCA J. Biological Sci*, **(1)**, 7-18 **(2012)**

**16.** Ausiello G., Brannetti B., Costantini A., Ferrè F., Maselli V., Via A., Cesareni G., Diella F., Superti-Furga G., Wyrwicz L., Ramu C., McGuigan C., Gudavalli R., Letunic, I., Bork, P., Rychlewski, L., Küster, B., Helmer-Citterich, M., Hunter, W. N., Aasland, R. and Gibson, T. J., ELM server: a new resource for investigating short functional sites in modular eukaryotic proteins, *Nucleic Acids Res*, **(31)**, 3625-3630 **(2003)**

**17.** Combet C., Blanchet C., Geourjon C. and Deléage G.NPS@: network protein sequence analysis. *Trends Biochem Sci.,* **(25:3)**,147-150 **(2000)**

**18.** Gasteiger E., Hoogland, C., Gattiker, A., Duvaud, S., Wilkins, M.R., Appel, R.D.and Bairoch, A., 'Protein Identification and Analysis Tools on the ExPASy Server', *The Proteomics Protocols Handbook*, **18**,571-607 **(2005)**

**19.** Kyte J. and Doolittle R.F., A simple method for displaying the hydropathic character of a protein, *J. Mol. Biol*, **(157)**, 105-132 **(1982)**

**20.** Garnier J., Osguthorpe DJ. and Robson B., Analysis of the accuracy and implications of simple methods for predicting the secondary structure of globular proteins, *J Mol Biol*, **(120)**, 97-120 **(1978)**

**21.** Marti-Renom, M.A., Stuart, A., Fiser, A., Sanchez, R., Melo, F. and Sali, A., Comparative protein structure modeling of genes and genomes, *Annu. Rev. Biophys. Biomol. Struct*, **21(29)**,291-325**(2000)**

**22.** Sali A. and Blundell T.L., Comparative protein modelling by satisfaction of spatial restraints, *J. Mol. Biol*.,**(234)**, 779-815**(1993)**

**23.** Fiser, A., Do, R.K. and Sali, A., Modeling of loops in protein structures, *Protein Science*, **(9)**,1753-1773**(2000)**

**24.** Bhatt T.K., Structural Studies on Mitogen Activated Protein Kinase from Plasmodium falciparum, *ISCA J. Biological Sci*, **(1)**, 42-46 **(2012)**

**25.** Combet C., Jambon M., Deleage G. and Geourjon, C.'Geno3D: automatic comparative molecular modelling of protein, *Bioinformatics*, **(18)**, 213-214 **(2002)**

**26.** Lambert, C., Leonard, N., De Bolle, X.and Depiereux, E., 'ESyPred3D: Prediction of proteins 3D structures', *Bioinformatics*, **(18:9)**, 1250-1256 **(2002)**

**27.** Bates, P.A., Kelley, L.A., MacCallum, R.M. and Sternberg, M.J.E.,'Enhancement of Protein Modelling by Human Intervention in Applying the Automatic Programs 3D-JIGSAW and3D-PSSM', *Proteins: Structure, Function and Genetics*, **(5)**,39-46**(2001)**

**28.** Bates, P.A. and Sternberg, M.J.E.,'Model Building by Comparison at CASP3: Using Expert Knowledge and Computer Automation', *Proteins: Structure, Function and Genetics,* **(3)**,47-54**(1999)**.

**29.** Contreras-Moreira, B., Bates, P.A., 'Domain Fishing: a first step in protein comparative modelling', *Bioinformatics*, **(18)**,1141-1142**(2002)**

**30.** Arnold, K., Bordoli, L., Kopp, J. and Schwede, T.'The SWISS-MODEL Workspace: A web-based environment for protein structure homology modelling'', *Bioinformatics*, **(22)**,195-201**(2006)**.

**31.** Schwede, T., Kopp, J., Guex, N. and Peitsch, MC., 'SWISS-MODEL: an automated protein homology-modeling server', *Nucleic Acids Research*., **(31)**,3381-3385**(2003)**

**32.** Guex, N. and Peitsch, M. C., 'SWISS-MODEL and the Swiss-PdbViewer: An environment for comparative protein modeling'., *Electrophoresis.*, **(18)**,2714-2723**(1997)**

**33.** Nielsen, M., Lundegaard, C., Lund, O. and Petersen, TN., 'Remote homology modeling using structure guided sequence profiles', *Nucleic Acids Research*.,**33(38)**,576-81**(2010)**

**34.** Laskowski, R A., MacArthur, M W., Moss, D S. and Thornton, J M. (1993), 'PROCHECK - a program to check the stereochemical quality of protein structures'., *J. App. Cryst*., **(26)**,283-291**(1993)**

**35.** Laskowski, R A., Rullmannn, J A., MacArthur, M W., Kaptein, R. and Thornton, J M., 'AQUA and PROCHECK-NMR: programs for checking the quality of protein structures solved by NMR', *J Biomol NMR*, **35(8)**,477-486**(1996)**

**36.** Sippl, M.J., Recognition of Errors in Three-Dimensional Structures of Proteins, *Proteins*, **(17)**,355-362**(1993)**