Annals of R.S.C.B., ISSN: 1583-6258, Vol. 25, Issue 7, 2021, Pages. 1588-1598
Received 05 May 2021; Accepted 01 June 2021.

Computational Approach for Constraining the Endemic Malarial Parasite

. S. Jeevan Amos*, Nasreen Sulthana and K. Kasturi
Dept. of Biotechnology, Acharya Nagarjuna University, Guntur.

Human intervention into the environment led to pollution. One such kind of pollution is water pollution which has led to the occurrence of several health hazards. The impact of water pollution and regular increase in the usage of non degradable materials led to clogging of water resources, this inturn resulted in the increase of mosquito population, these are the vectors for the transmission of malarial parasite, this leads to occurrence of different types of malaria. According to the WHO statistics, Malaria stands as 2nd major health problem. *Plasmodium falciparum* is one of the malarial parasite, which is a causative of cerebral malaria. In the present study, potential drug ATCases were screened for multidrug resistant *P.falciparum*. The screening of the putative ATCases from metabolic database has lead to identification of specific enzyme sites. These were proven as drug ATCases for antibacterial activity. In the present work we propose that they can also be used as an antimalarial drug ATCases. The work involves the modelling and validation of the drug ATCase. This is an innovative approach to overcome the resistance of this parasite towards the presently available drugs, being made with Bioinformatics tools using the available genome sequence of this parasite. So the modified drugs can inhibit this ATCase in the parasite growth.

Keywords: Malaria, *Plasmodium falciparum*, Homology modelling, Aspartate carbamoyl – transferase

.

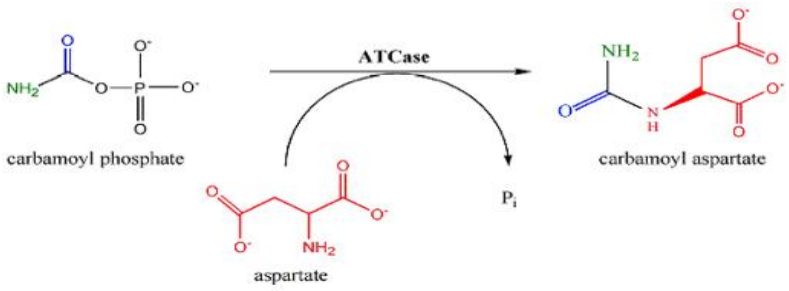
INTRODUCTION

Malaria is the 2nd major health problem as per WHO reports. The rapid urbanization has led to dumping of sewage and non biodegradable materials into the local fresh water irrigation canals. This has led clogging of water resources and their by have turned to be breeding sits for mosquito larva. The sudden increase has led to the infestation of malaria parasite mediated diseases. *Plasmodium falciparum* is one of the species of the *Plasmodium* that causes cerebral malaria. The occurrence is alarming and the drug dosage is complex as the parasite posses' multi drug resistance.

The innovative approach in the present work is to identify potential drug target by utilizing the computational tools, generate 3D models and validation of the generated models. The biologically fit model of the drug target is further taken up for the docking analysis using *P. falciparum 3D7* ATCase the available antimalarial drugs. The aim of the approach is to decrease the cost of the drug and the dosage levels.

Aspartate carbamoyl transferase (ATCase) is a highly regulated enzyme that catalyses the first committed step in pyrimidine biosynthesis, the condensation of aspartate and carbamyl phosphate to form N-carbamyl-L-aspartate and inorganic phosphate. Reaction of this enzyme is as follows:

Annals of R.S.C.B., ISSN: 1583-6258, Vol. 25, Issue 7, 2021, Pages. 1588-1598
Received 05 May 2021; Accepted 01 June 2021.



Reaction catalysed by Aspartate carbamoyltransferase

Reaction source: http://www.absoluteastronomy.com/topics/Aspartate_carbamoyltransferase

METHODOLOGY

Screening of ATCases

P. falciparum 3D7 was selected as the candidate organism for the present study whose complete genome sequence is available. The potential ATCases were identified from the literature on the basis of its importance for the growth of parasite and might not similar to the human proteins. Blastp was performed by taking the parameters like E-value is ≤ 0.0005, Query coverage is ≥ 80%, identity is ≥ 35% and other parameters like Threshold, word size and substitution matrix kept default for all the potential ATCases. Aspartate carbamoyl transferase or ATCase identified as potential one, which is the major component in tropozoite stage and involved in Pyramidine Biosynthesis pathway, to inhibit the parasite growth. The amino acid sequence of ATCase (UniprotKB ID: Q81DP8) was retrieved from the protein database of National Center of Biotechnology Information (NCBI, http://www.ncbi.nlm.nih.gov/). The three-dimensional structure of the protein was not yet available in Protein Data Bank and Modbase Database. Hence the present exercise of developing the 3D model of the ATCase (Q81DP8) was undertaken.

Sequence analysis

The protein sequence of *P. falciparum 3D7* ATCase **(**UniprotKB ID: Q81DP8**)** was submitted to p-Blast against PDB database for similarity search by keeping default parameters like E- value threshold 10, word size 3 and Blosum 62 Matrix to get the template for comparative modeling. Since the BLAST algorithm detects local as well as global alignments, regions of similarity embedded in otherwise unrelated proteins can be detected (Altschul *et al*, 1997).The derived homologues sequences of ATCase protein were collected and best BLAST hit was selected as template for further analysis (**Table 1)**. Template selection is a critical step in homology modelling. Multiple sequence alignment was performed using CLUSTALX2 (Higgins *et al.*, 1994). Regions of conservation and variation were detected from CLUSTALX2 results (Fig. 1).

Physico-chemical characterization of ATCase protein

The basic physico-chemical properties of the ATCase protein sequence were calculated using the ProtParam tool (http://expasy.org/tools/protparam.html). The parameters computed by ProtParam are molecular weight, theoretical pI, amino acid composition, atomic composition, extinction coefficient, estimated half-life, instability index, aliphatic index and grand average of hydropathicity (GRAVY) (Kyte and Doolittle, 1982). All the results are tabulated in (Table 2,3,4).

Annals of R.S.C.B., ISSN: 1583-6258, Vol. 25, Issue 7, 2021, Pages. 1588-1598
Received 05 May 2021; Accepted 01 June 2021.

Function prediction

Functional characterization of ATCase protein sequence was done by finding motif using ELM tool (Eukaryotic Linear Motif) (**http://elm.eu.org**) (Puntervoll *et al.,* 2003) and domain analysis was carried out by using PROTSCAN (**http://npsa-pbil.ibcp.fr/cgi-bin/npsa_ automat.pl?page=npsa_prosite.html**). Results of ELM tool and PROTSCAN was given in (Table 5 and 6)

Secondary structure prediction

GOR-IV tool (Garnier-Osguthorpe-Robson, 1978) (**http://npsa-pbil.ibcp.fr/cgi-bin/npsa_ automat.pl?page=npsa_gor4.html)** was used to obtain the secondary structure of ATCase protein (Table 7 and Fig.2). The secondary structure prediction is the definition of each residue into either alpha helix, beta sheet or random coil secondary structures.

3D Model

A comparative 3D structure analysis of *P. falciparum 3D7,* ATCase protein was performed using Modeller 9.10, (Tosatto, 2011) (Fig 3). The MODELLER software employs probability density functions (PDFs) as the spatial restraints rather than energy. The 3D model of a protein is obtained by optimization of the molecular PDF such that the model violates the input restraints as possible. The molecular PDF was derived as a combination of PDFs restraining individual spatial features of the whole molecule.

Validation

The result was evaluated using an web based validation servers such as PROCHECK.

PROCHECK (http://swissmodel.expasy.org/workspace/index.php?func=tools_structure assessment1&userid=USERID&token=TOKEN).This programs assess the "stereo-chemical quality" of a given protein structure. The aim of PROCHECK is to assess how normal or how unusual, the arrangement of residues geometry in a given protein structure is, as compared with stereo-chemical parameters derived from well-refined, high-resolution structures. Results were tabulated.

RESULTS AND DISCUSSION

>tr|Q8IDP8|Q8IDP8_PLAF7 Aspartate carbamoyltransferase OS=Plasmodium falciparum (isolate 3D7) GN=atcasE PE=3 SV=1
MIEIFCTAIVVITILIVGVFVYMIIRTKKKKLKLDNMFYINSKYKIDLDKIMT
KMKNKSVINIDDVDDEELLAILYTSKQFEKILKNNEDSKYLENKVFCSVFLEPSTRTRCSFDA
AILKLGSKVLNITDMNSTSFYKGETVEDAFKILSTYVDGIIYRDPSKKNVDIAVSSSSKPIINA
GNGTGEHPTQSLLDFYTIHNYFPFILDRNINKKLNIAFVGDLKNGRTVHSLSKLLSRYNVSFN
FVSCKSLNIPKDIVNTITYNLKKNNFYSDDSIKYFDNLEEGLEDVHIIYMTRIQKERFTDVDEY
NQYKNAFILSNKTLENTRDDTKILHPLPRVNEIKVEVDSNPKSVYFTQAENGLYVRMALLYL
IFSSTS

ATCase sequence from uniprot

The amino acid sequence of ATCase protein was retrieved from NCBI and P-BLAST program was used to find out the sequences that shared structure and sequence similarity against PDB database (Table 1). The results suggest that protein sequence with 1ML4 which belongs to The Pala-Liganded Aspartate Transcarbamoylase Catalytic Subunit From *Pyrococcus Abyssi*, showed the highest degree

Annals of R.S.C.B., ISSN: 1583-6258, Vol. 25, Issue 7, 2021, Pages. 1588-1598
Received 05 May 2021; Accepted 01 June 2021.

of similarity (38% identity) to ATCase protein as query sequence, indicated by E-value cut off<0.0005 and score>100.

Property	**Value**
Template Id	1ML4
Organism	*Pyrococcus Abyssi*
% Identity	38
No. of Gap	14
E value	9e-76
Bit Score	240

Table 1: Summary of P- BLAST 3 (www.ncbi.nlm.nih.gov/BLASTp/) of ATCase protein sequence when searched against PDB database

Further, ATCase protein was subjected to CLUSTALX2 for multiple sequence alignment (Fig. 1). The result suggests that ATCase protein of *P. falciparum 3D7* was very much similar to Pala-Liganded Aspartate Transcarbamoylase Catalytic Subunit from *Pyrococcus Abyssi*

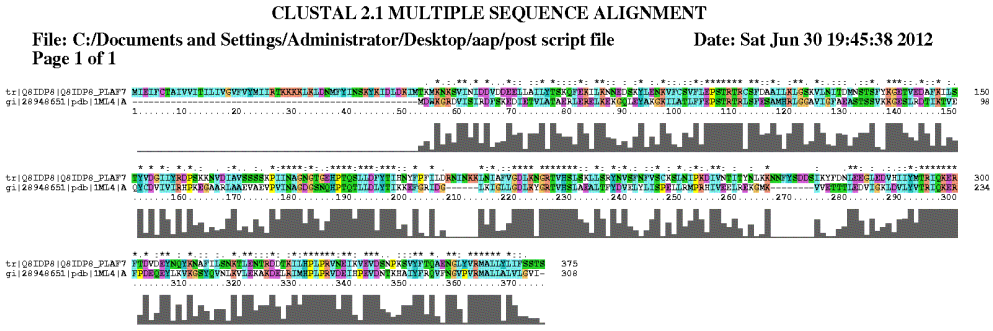


Figure 1: Multiple Sequence Alignment using CLUSTALX2 as visualized by Jalview.

Amino acid	**Number**	**Percentage**
Ala (A)	**11**	**2.9%**
Arg (R)	**12**	**3.2%**
Asn (N)	**32**	**8.5%**
Asp (D)	**27**	**7.2%**
Cys (C)	**4**	**1.1%**
Gln (Q)	**5**	**1.3%**
Glu (E)	**19**	**5.1%**
Gly (G)	**11**	**2.9%**
His (H)	**5**	**1.3%**
Ile (I)	**38**	**10.1%**
Leu (L)	**34**	**9.1%**
Lys (K)	**38**	**10.1%**
Met (M)	**8**	**2.1%**
Phe (F)	**21**	**5.6%**

Annals of R.S.C.B., ISSN: 1583-6258, Vol. 25, Issue 7, 2021, Pages. 1588-1598
Received 05 May 2021; Accepted 01 June 2021.

Pro (P)	**9**	**2.4%**
Ser (S)	**31**	**8.3%**
Thr (T)	**24**	**6.4%**
Trp (W)	**0**	**0.0%**
Tyr (Y)	**20**	**5.3%**
Val (V)	**26**	**6.9%**
Pyl (O)	**0**	**0.0%**
Sec (U)	**0**	**0.0%**

Table 2: Amino acid composition of ATCase protein of *P. falciparum 3D7*

Element	Symbol	Number of atoms
Carbon	C	1958
Hydrogen	H	3098
Nitrogen	N	496
Oxygen	O	580
Sulfur	S	12

Table 3: Atomic composition of *P. falciparum 3D7* ATCase protein

PROPERTY	VALUE
Number of amino acids	375
Molecular weight	43251.8
Theoretical Pi	8.55
Total number of negatively charged residues (Asp + Glu)	46
Total number of positively charged residues (Arg + Lys)	50
Total number of atoms	6144
Aliphatic index	97.92
Grand average of hydropathicity (GRAVY)	-0.218
Extinction coefficients at 280 nm#	$30050M^{-1}\ cm^{-1}$
Extinction coefficients at 280 nm*	$29800M^{-1}\ cm^{-1}$
Estimated half-life	30 hours
Instability index	32.15

Table 4: Physico-chemical properties of ATCase protein of *P. falciparum 3D7*

The function of ATCase protein of *P. falciparum 3D7* was analyzed by submitting the amino acid sequence to Eukaryotic Linear motif (ELM) server (www.elm.eu.org). Based on EML and PROSCAN results, the following motif and domains were assigned to ATCase protein amino acid sequence(Table 5 and Table 6).

Annals of R.S.C.B., ISSN: 1583-6258, Vol. 25, Issue 7, 2021, Pages. 1588-1598
Received 05 May 2021; Accepted 01 June 2021.

Elm Name	Instances (Matched Sequence)	Positions	Cell Compartment	Pattern	Probability
LIG_CYCLIN_1	KKLKL	30-34 [A]	cytosol, nucleus	[RK].L.{0,1}[FYLIVMP]	0.0053
LIG_MAPK_1	KKKKLKL	28-34 [A]	nucleus, cytosol	[KR]{0,2}[KR].{0,2}[KR].{2,4}[ILVM].[ILVF]	0.0043
TRG_NLS_MonoExtC_3	KKKKLK KKKLKLD	28-33 [A] 29-35 [A]	nucleus, Nuclear pore, NLS-dependent protein nuclear import complex	[^DE]((K[RK])|(RK))(([^DE][KR])|([KR][^DE]))(([PKR])|([^DE][DE]))	0.0007
TRG_NLS_MonoExtN_4	KKKKLKL KKKLKL	28-34 [A] 29-34 [A]	nucleus, Nuclear pore, NLS-dependent protein nuclear import complex	(([PKR].{0,1}[^DE])|([PKR]))((K[RK])|(RK))(([^DE][KR])|([KR][^DE]))[^DE]	0.0013

Table 5: Motif analysis of P. falciparum 3D7 ATCase

Domine name	Prosite Accession Number	Randomized probability	Start position	End position	Sequence of prosite
N-glycosylation site	PS00001	5.138e-03	57	60	NKSV
			127	130	NITD
			132	135	NSTS
			182	185	NGTG
			239	242	NVSF
			318	321	NKTL
Protein kinase C phosphorylation site	PS00005	1.423e-02	27	29	TKK
			77	79	TSK
			107	109	STR
			162	164	SKK
			173	175	SSK

Annals of R.S.C.B., ISSN: 1583-6258, Vol. 25, Issue 7, 2021, Pages. 1588-1598
Received 05 May 2021; Accepted 01 June 2021.

			246	248	SCK
			274	276	SIK
			317	319	SNK
Casein kinase II phosphorylation site	PS00006	1.482e-02	141	144	TVED
			151	154	TYVD
			191	194	SLLD
			302	305	TDVD
			324	327	TRDD
			354	357	TQAE
Tyrosine kinase phosphorylation site	PS00007	min=4.074e-04 max= .083e-04	86	93	KNNEDSKY
			300	307	RFTDVDEY
			86	93	KNNEDSKY
			31	39	KLKLDNMFY
N-myristoylation site	PS00008	1.397e-02	181	186	GNGTGE
Aspartate and ornithine carbamoyltransferases signature	PS00097	1.340e-07	103	110	FLEPSTRT
Glutamine synthetase putative ATP-binding region signature	PS00181	min=2.844e-09 max=2.850e-09	175	191	KPIINAGNGTGEHPTQS

Table 6: Domains obtained in *P. falciparum 3D7* ATCase protein Protein using PROSCAN

GORIV program that was used to predict secondary structures in *P. falciparum 3D7* suggest that it contained more helices than beta sheets. Window width was kept at 17, similarity threshold: 8 and number of states: 4 were used as parameters for analysis.

Annals of R.S.C.B., ISSN: 1583-6258, Vol. 25, Issue 7, 2021, Pages. 1588-1598
Received 05 May 2021; Accepted 01 June 2021.

Secondary Structure		Stretch	Percentage
Alpha helix	(Hh)	176	46.93%
3_{10} helix	(Gg)	0	0.00%
Pi helix	(Ii)	0	0.00%
Beta bridge	(Bb)	0	0.00%
Extended strand	(Ee)	55	14.67%
Beta turn	(Tt)	29	7.73%
Bend region	(Ss)	0	0.00%
Random coil	(Cc)	115	30.67%
Ambigous states	(?)	0	0.00%
Other states		0	0.00%

Table 7: Percentage and type of Secondary structures present in *P. falciparum 3D7* ATCase protein

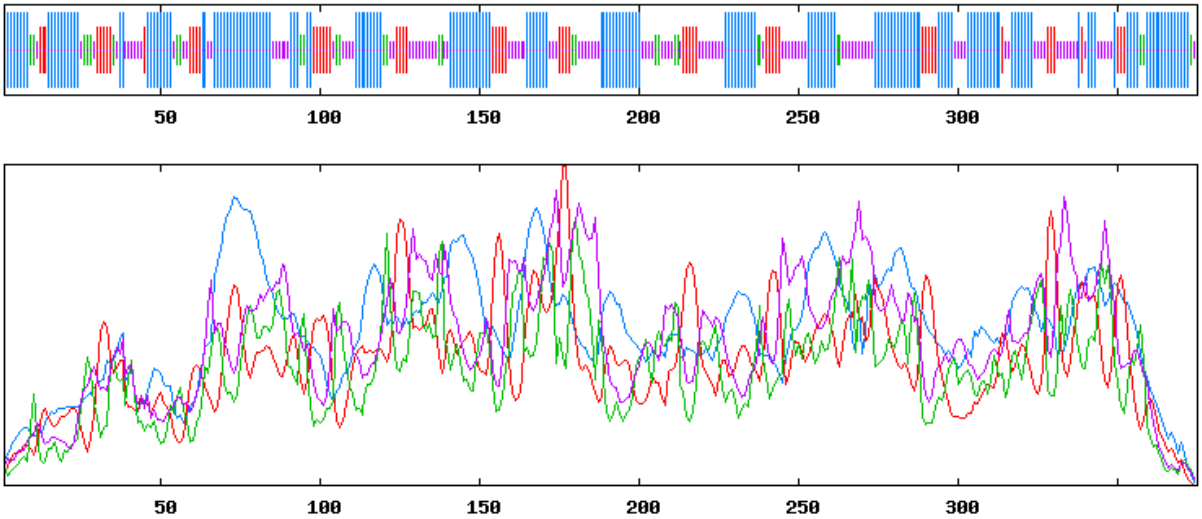


Figure 2: Secondary structure prediction of ATCase protein of *P. falciparum 3D7* by GORIV tool.

Three dimensional structure of ATCase of *P. falciparum 3D7* was predicted by MODELLER 9.10 using template of Pala-Liganded Aspartate Transcarbamoylase Catalytic Subunit From *Pyrococcus Abyssi* (PDBID 1ML4*)* keeping default parameters

Annals of R.S.C.B., ISSN: 1583-6258, Vol. 25, Issue 7, 2021, Pages. 1588-1598
Received 05 May 2021; Accepted 01 June 2021.

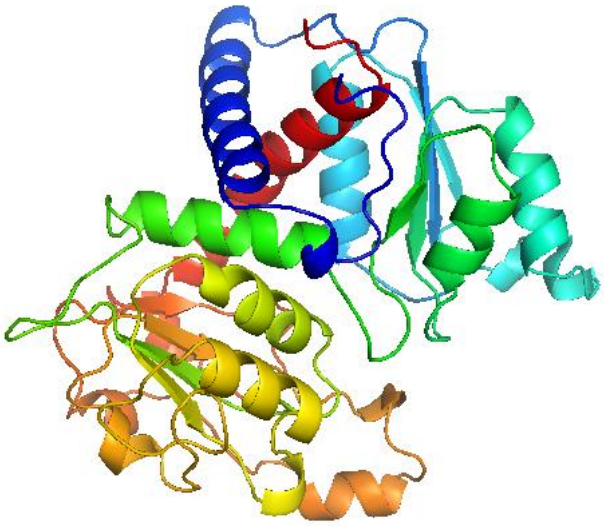


Figure 3: 3D structure of *P. falciparum* 3D7 ATCase protein model

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, 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 (Fig: 4, Table: 8)

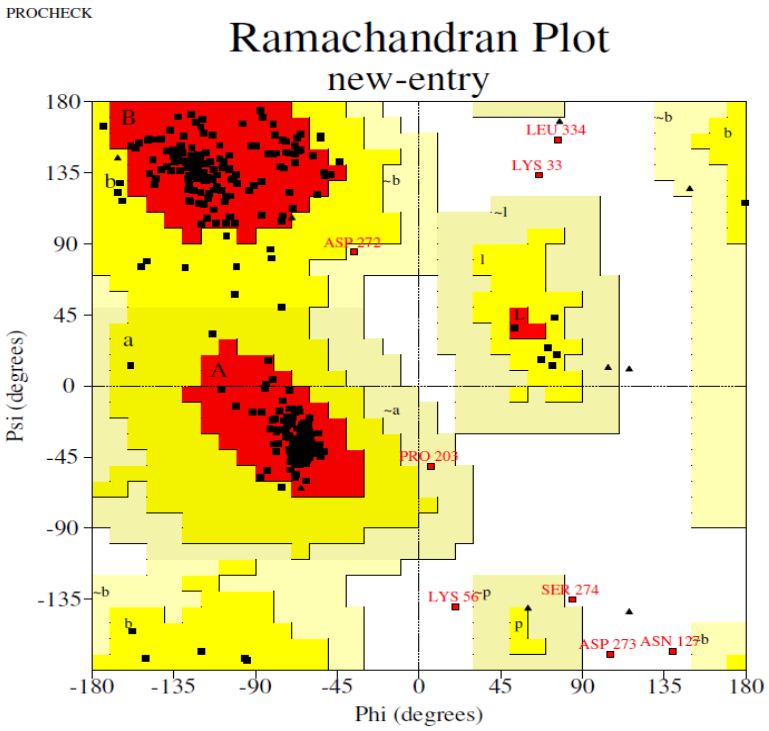


Figure 4: Ramachandran plot of modeled protein obtained using PROCHECK

Annals of R.S.C.B., ISSN: 1583-6258, Vol. 25, Issue 7, 2021, Pages. 1588-1598
Received 05 May 2021; Accepted 01 June 2021.

Amino acid position	Modeled protein (%)	1ML4 template Structure (%)
Most favoured region	89.5	88.3
Additional allowed region	8.5	11.3
Generously allowed region	0.3	0.0
Disallowed region	1.7	0.4

Table 8: Comparison of validation reports of plasmodium falciparum 3D7 Aspartate carbamoyl transferase modeled protein and crystal structure 1ML4

SUMMARY AND CONCLUSION

In 2010, 106 countries and territories were reported to show vulnerability to malaria transmission. These 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. Anti-malarial drug resistance is recognized to be 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.

Selection and validation of novel molecular targets have become of paramount importance in light of the plethora of new potential therapeutic drug targets that have emerged from genomics revolution where we visage an avalanche of data but only flakes of information. With the increasing drug resistance in Plasmodium, there is an imperative need for exploring novel drugs to reduce the impending impact of the emergence of multidrug-resistant *P. falciparum*. Pyramidine biosynthetic pathway is critical and specific for *P. falciparum* and other apicomplexan parasites. The description of metabolic pathways, particularly those differing from humans, provides new targets for drug development. Computational methods play a crucial role in accelerating the drug development process and among them; comparative protein modelling is of great assistance during the rational design of drug molecules. In the dearth of experimental data, model-building on the basis of the known three dimensional structure of a homologous protein is the only unfailing method to obtain structural information. Hence in the present study we have focused to characterize the ATCase protein of *P. falciparum 3D7* from sequence and structure and in elucidating function using bioinformatics tools. This study will provide a new insight into the structure of *P. falciparum 3D7* Aspartate carbamoyl transferase for rational designing of inhibitors in our crusade against this disease.

ADDITIONAL MATERIAL

Additional file 1 :Table 5 Motif analysis of *P. falciparum 3D7* ATCase .

REFERENCES

1. Altschul, S.F., Madden, T.L., Schäffer, A.A., Zhang, J., Zhang, Z., Miller, W. & Lipman, D.J. (1997) "Gapped BLAST and PSI-BLAST: a new generation of protein database search programs." Nucleic Acids Res. 25:3389-3402
2. Higgins D., Thompson J., Gibson T. Thompson J. D., Higgins D. G., Gibson T. J. (1994). CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through

sequence weighting,position-specific gap penalties and weight matrix choice. Nucleic Acids Res. 22:4673-4680

3. Kyte, J. and Doolittle, R.F. (1982) A simple method for displaying the hydropathic character of a protein. J. Mol. Biol. 157, 105-132. [PubMed: 7108955]
4. Puntervoll, P., Linding, R., Gemünd, C., Chabanis-Davidson, S., Mattingsdal, M., Cameron, S., Martin, D. M. A., 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., & Gibson, T. J. (2003) ELM server: a new resource for investigating short functional sites in modular eukaryotic proteins.. Nucleic Acids Res. 31: 3625-3630.
5. Silvio C.E. Tosatto. (2005) The Victor/FRST Function for Model Quality Estimation. Journal of Computational Biology; 12(10):1316-1327
6. Wiederstein & Sippl (2007) ProSA-web: interactive web service for the recognition of errors in three-dimensional structures of proteins. *Nucleic Acids Research* 35, 407-410.
7. Sippl, M.J. (1993) Recognition of Errors in Three-Dimensional Structures of Proteins. *Proteins* 17, 355-362
8. Garnier J, Osguthorpe DJ, Robson B. (1978). Analysis of the accuracy and implications of simple methods for predicting the secondary structure of globular proteins. *J Mol Biol* 120:97-120