*Research Journal of Recent Sciences* _____ ISSN 2277-2502
Vol. **1(9),** 79-82, September **(2012)** *Res.J.Recent Sci.*

***Short Communication***

# Computational Studies on Calpain from Plasmodium falciparum

**Bhatt T.K.**
Department of Biotechnology, Central University of Rajasthan, Kishangarh, INDIA

**Available online at: www.isca.in**
Received 20th April 2012, revised 13th May 2012, accepted 18th May 2012

## Abstract

*Malaria is one of the most devastating diseases prevalent in the world. It is caused by the parasite Plasmodium falciparum. Many species of Plasmodium are shown to infect human host. The increased resistant in malaria parasite against drugs remain the major concern. Hence identification of new and effective drug targets against Plasmodium is a regular process. In the same line, proteases are the major group of the proteins in the parasite which plays crucial role in various processes like migration, evasion and cell cycle etc. Therefore, in this study, we have performed structural studies on cysteine proteases called 'calpain' from malaria parasite (Pfcalpain). In addition, phylogenetic analysis was also performed on Pfcalpain. We believe that these results will help in understanding various biological processes of parasite and will be instrumental in discovering effective chemotherapy against malaria.*

**Keywords:** Calpain, cysteine proteases, molecular modelling, phylogeny, malaria, drug discovery.

## Introduction

Millions of death occurred every year due to disease of malaria. The causative agent of malaria is an apicomplexan, *Plasmodium*. Four species of *Plasmodium* are responsible for disease in human, P. falciparum, P. vivax, P. malariae and P. ovale. Out of four species, P. falciparum is the most prevalent and responsible for the most of the malaria pathology. Different drugs are available in market against parasite but growing resistant towards existing drugs compels scientific community to discover more drug target in *Plasmodium*. Proteases of the parasite play major roles in different metabolic pathways including cell cycle regulation, differentiation and development, parasite invasion and evasion, migration and nutrition[1-7]. There are almost 35 members of cysteine protease are known but there is only single copy of cysteine protease, Calpain, is found in the parasite genome[1]. Pfcalpain has unusually very long N-terminal compared to counterparts in other species. The protein can be divided into different domains like nuclear localization domain, palmitoylation domain etc. It has been already shown that Pfcalpain is essential for intraerythrocytic cycle of the parasite, specifically in cell cycle progression during trophozoite development[8]. In this work, we have performed homology modeling to solve three-dimensional structure of Pfcalpain Also deciphered the probable active site of the protein. In addition, phylogenetic tree was constructed to study evolutionary relationship of protein with other species. Taking all together, we hope that this study would be instrumental in enhancing the process of drug discovery against malaria parasite.

## Material and Methods

The sequence of Pfcalpain was extracted from PlasmoDB with accession number of Mal13P1.310. The protein sequence of Pfcalpain was pasted in NCBI blast column to run against referenced sequences of various proteomes. Total of 25 sequences of calpain from various phyla of living organisms were selected for sequence alignment. ClustalW online server was used for generating multiple sequence alignment of above selected sequences. The output file of the multiple sequence alignment was further submitted to MEGA5 program for generation of phylogenetic tree. Test neighbour-joining method was used for making phylogenetic tree. Images were created with MEGA5 program in pdf format. Using NCBI Blast against protein data bank (PDB), the template for homology modeling was identified. 2NQA pdb structures were used as a template for homology modeling. Modeller[9] and Swiss Model Server's online facility was used to submit the protein sequence for homology modeling to build the in-silico structure of Pfcalpain. CASTp server was utilized for active site prediction using modelled structure of Pfcalpain[10-11]. Images were produced using CHIMERA [12]. Images were processed at higher resolution in PNG format.

## Results and Discussion

Multiple sequence alignment of Pfcalpain was performed using ClustalW, where 25 selected sequences were taken including sequence of Calpain from *Plasmodium* falciparum. Interestingly, we found that P. falciparum calpain has a unique extension at N-terminal of protein which was absent in all other sequence homolog (figure1). Insertion of small sequence or motif in protein coding gene is a common phenomenon in malaria proteins which leads to the increase in the size of the protein. Though, the insertion in the case of calpain was very long. Although, the significance of this N-terminal extension is not yet known but could be involved in making protein-protein interactions. Figure 2 showing the alignment of 25 sequences of

Pfcalpain from various organisms with different colour coding based on the degree of conservation of amino acid residue at that particular position. In addition, evolutionary tree of selected Pfcalpain sequences was constructed using MEGA5 program. Various methods of phylogenetic analysis were employed including maximum-likelihood, neighbour-joining. The P. falciparum calpain sequence was tagged with green colour Square in the constructed tree (figure 3). Phylogenetic tree revealed that Pfcalpain makes a separate branch with apicomplexan along with it. Interestingly, human calpain break out earlier and makes a separate branch, away from the Pfcalpain. This evolutionary distance between the human and parasite calpain could be utilized for better chemotherapy. Further, the three-dimensional structure of Pfcalpain was obtained with homology modeling. Structure is predominantly beta stranded in nature (figure 4). Loops are scattered throughout the structure might be involved in the interactions between protein. Positively charged residues are shown in blue on hydrophobic surface representation of Pfcalpain structure (figure 4). Structure of Pfcalpain also revealed a cavity of size capable of holding one or more substrates for catalysis. Further prediction of active site using CASTp (computed atlas of surface topography of proteins) also revealed the same pocket for possible active site of the enzyme (figure 4). The active site amino acids are shown in Figure 4 in green colour. Taking into consideration all above results, we think that present structural characterization of Pfcalpain would be helpful in deciphering proteolytic activity of enzyme in *Plasmodium* and would be crucial in accelerating process of drug discovery against malaria parasite.

## Conclusion

Construction of phylogenetic tree tells you about the evolutionary closeness of given sequence among available sequences chosen for studies. In the analysis of Pfcalpain in terms of its phylogenetic relationship with other species provided the direct evidence of distant nature of *Plasmodium* sequence. In addition, structural analysis of Pfcalpain also provided key insight into the active site of the enzyme. Hence, we conclude that this study will not only pave the way for drug development but also helps in the understanding of evolutionary path of present day malaria parasite.

## Acknowledgement

I would like to thank Central University of Rajasthan, Department of Biotechnology for providing resources to conduct these studies.

## References

**1.** Wu Y., Wang X., Liu X. and Wang Y., Data-mining approaches reveal hidden families of proteases in the genome of malaria parasite, *Genome Res*., **13**, 601–616 **(2003)**

**2.** Shenai B.R., Sijwali P.S., Singh A. and Rosenthal P.J., Characterization of native and recombinant falcipain-2, a principal trophozoite cysteine protease and essential hemoglobinase of Plasmodium falciparum, *J Biol Chem*., **275,** 29000–29010 **(2000)**

**3.** Rosenthal P.J., Cysteine proteases of malaria parasites, *Int J Parasitol.,* **34**, 1489–1499 **(2004)**

**4.** Klemba M., Gluzman I. and Goldberg D.E., A Plasmodium falciparum dipeptidyl minopeptidase 1 participates in vacuolar hemoglobin degradation, *J Biol Chem*., **279**, 43000–43007 **(2004)**

**5.** Sijwali P.S. and Rosenthal P.J., Gene disruption confirms a critical role for the cysteine protease falcipain-2 in hemoglobin hydrolysis by Plasmodium falciparum, *Proc Natl Acad Sci U S A*., **101**, 4384–4389 **(2004)**

**6.** Hadley T., Aikawa M. and Miller L.H., Plasmodium knowlesi: Studies on invasion of rhesus erythrocytes by merozoites in the presence of protease inhibitors, *Exp Parasitol*., **55**, 306–311 **(1983)**

**7.** Croall D.E. and Ersfeld K., The calpains: Modular designs and functional diversity, *Genome Biol*., **8**, 218 **(2007)**

**8.** Ilario R. et al., A calpain unique to alveolates is essential in Plasmodium falciparum and its knockdown reveals an involvement in pre-S-phase development, *Proc Natl Acad Sci U S A*., **106**, 1554–1559 **(2004)**

**9.** Lovell S.C., Davis I.W., Arendall W.B. , de Bakker P.I., Word J.M., Prisant M.G., Richardson J.S. and Richardson D.C., Structure validation by Calpha geometry: phi,psi and Cbeta deviation, *Proteins.*, **15;50(3)**, 437-50 **(2000)**

**10.** Dundas J., et. Al., CASTp: computed atlas of surface topography of proteins with structural and topographical mapping of functionally annotated residues, *Nucleic Acids Res.*, **34,** 116–118 **(2006)**

**11.** Bhatt T.K., In-Silico Structure Determination of Protein Falstatin from Malaria Parasite Plasmodium Falciparum, *Res. J. Recent Sci.,* **1(4)**, 68-71 **(2012)**

**12.** Pettersen E.F., Goddard T.D., Huang C.C., Couch G.S., Greenblatt D.M., Meng E.C. and Ferrin TE, UCSF Chimera - A Visualization System for Exploratory Research and Analysis, *J Comput Chem.,* **25**,1605-1612, **(2004)**

```
Gallus          ------------------------------------------------------------
Meleagris       ------------------------------------------------------------
Xenopus         ------------------------------------------------------------
Monodelphis     ------------------------------------------------------------
Oreochromis     ------------------------------------------------------------
Danio           ------------------------------------------------------------
Anopheles       ------------------------------------------------------------
Culex           ------------------------------------------------------------
Aedes           ------------------------------------------------------------
Solenopsis      ------------------------------------------------------------
Acromyrmex      ------------------------------------------------------------
Camponotus      ------------------------------------------------------------
Harpegnathos    ------------------------------------------------------------
Megachile       ------------------------------------------------------------
Apis            ------------------------------------------------------------
Nasonia         ------------------------------------------------------------
Tribolium       ------------------------------------------------------------
Pediculus       ------------------------------------------------------------
Trichoplax      ------------------------------------------------------------
Hydra           ------------------------------------------------------------
Caenorhabditis  ------------------------------------------------------------
Loa             ------------------------------------------------------------
Clonorchis      ------------------------------------------------------------
Plasmodium      KVKEKRKIKKRKKEECNLIENVEGNNVGNKNVSSYVMKEKKNNEKDDENNNIDCNNNDNN
Toxoplasma      CSVFLSSPSPFRGNYLERGVPDCGRDSSPAVVAFGTAPYTRRLSPWSPLRSPNSASACAQ
```

**Figure-1**

**Showing the multiple sequence alignment of PfAdT, where N-terminal extension of several amino acids is found only in apicomplexan including Plasmodium**

```
Gallus          VTQAFDEDDKGNAEE-AIELYTEAVELCLKTA-TETSEAGLQSKLKQLARQALDRAEALK 323
Meleagris       VTQAFDEDDKGNAEE-AIELYTEAVELCLKTA-TETSEAGLQAKLKQLARQALDRAEALK 149
Xenopus         VTQAFDEDAKGNAEE-AIELYSEAVELCINTS-NETVDQNLQAKLKQLARQALDRAESLK 107
Monodelphis     VTQAFDEDDKGNAEE-AIELYTEAVDLCLKTS-NETSDQALQSKLKVLARQALDRAEALK 151
Oreochromis     VTQAFEEDEKGNDDE-AIELYTQAVELCIKTS-NETSEQVLQNKLKQLARQALDRAEGLK 151
Danio           VTQAFEEDEKENADE-AIELYTQAVELCIQAS-NETSDPALQAKLKQLARQALDRAEGLK 151
Anopheles       LGRALDADEAGRKDE-AIDLYGQAVEKILR-----LEDREKREKLNKFAKQALDRAEELK 64
Culex           LSRALDADESGQKEL-AIELYGQTVETILR-----IENRESREKLHRFAMQALERAEELK 64
Aedes           LTRALDADEAGQKEL-AIDLYGQAVESVLR-----IENREKRDKLNKFAKQALERAEVLK 64
Solenopsis      MNQALDADEAGLKDI-AVKLYTDAAELGLS---TKTVDPEVKAKLTNLVRVAVERAEDLK 145
Acromyrmex      MNQALDADEAGLKDI-AIKLYTDAAELGLS---AKTADTEVKAKLTNLVRVAVERAESLK 145
Camponotus      MNQALDADEAGLKDI-AIKLYTDAAELGLS---AKTVDTDVKAKLTNLVRVAVERAESLK 145
Harpegnathos    ISQALDADEAGLKDI-AVKLYTDAAELGLS---TKTSDVEIKAKLTDLVRVAVERAESLK 144
Megachile       INQAQDADEAGLKDI-AVKLYTNAAEFGLS---IKTTDTELKGKITALVRLALDRAESLK 142
Apis            MNQAQDADEAGLKNI-AVKLYTDAAELGLN---MKIIDAEAKIKLTDLIKLALDRAESLK 142
Nasonia         LNQALDADEAGFKEE-AIKLYTNAAELGLK---AKSTSNE-KQKITNLVRHALDRAESLK 144
Tribolium       LQEAIEEDESGDKSD-AIELYAQAIEF------ITKNPDLMQGELKQLALQALERAEALK 144
Pediculus       FSQALDADERDHKDI-AVELYSQTAEYALTQ--KGECDEIVQQKIVTRAKQAIERAEEIK 152
Trichoplax      MKQAL-LEDERDRSDDAEPLYMDAAELCLR----VKSTCSDPTAIKKLALLANQAVDRKS 146
Hydra           ADTYLGTVHTVQRTVDTVQRAVEFIQGAVD----IVQDAVD-TVFKKQEDDATKYLLNNN 142
Caenorhabditis  MIKASVLQNYGNKLEESRSLYENVVEQCLGVSRNSNLSQETLKKLRQTAESALKCIEELV 151
Loa             LYQALDQDEAGNTG-EAIMLYSLATELCINSS-NASSDAAMAGKLRQLAKKALDRAEVLK 129
Clonorchis      INAQTAGSTTPNIRDIAHRYIKRAEDLKAMSS-TRDTDVGVEGRRPSVLQNSLSRAKFIF 165
Plasmodium      SSSHLNENHDKKSENFNLNLIYQKKKENNNNKKKENNDNKKKENNNNKKKENNNNKKKEN 726
Toxoplasma      QQKLLPSHLSGSLAVSEILLESSIVGRYVMLPWNEEDGDIRQNFYVHCPPPVALCS-PLS 1069
```

**Figure-2**

**Showing the multiple sequence alignment of PfAdT along with other sequence homolog from 25 different species from all the three domains of life**

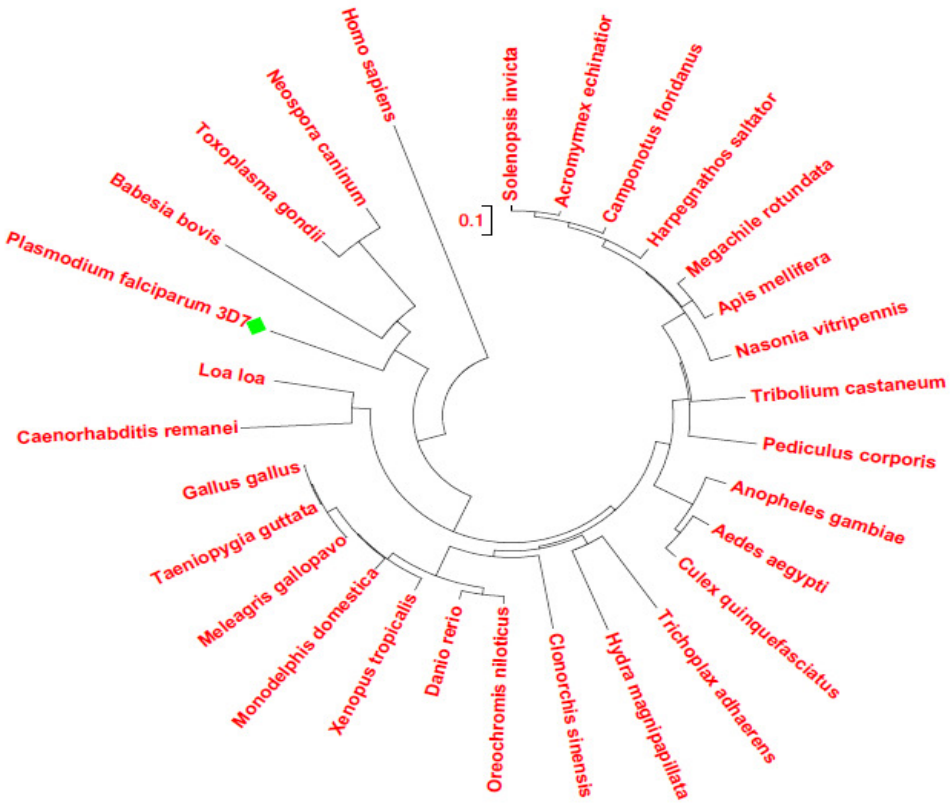


**Figure-3**
**Showing the phylogenetic tree constructed with Pfcalpain along with other sequence homolog from 25 different species from all the three domains of life, using neighbour-joining method**

**Chain A**

1731- H L S L Y E I P P L L P D N Y S S L Y F K G M W T N K S A G G C S N N L W S Y F R N P H I R L Y V P
1781- E C T R F Y I F L E C S Q E H S V N L R I F K G N T S S P R S L K K G D I I S S G P Y K A G C C Y I
1831- E C T L E S G I Y C L L P S L Y R A N V T G N Y Q I C V H Y

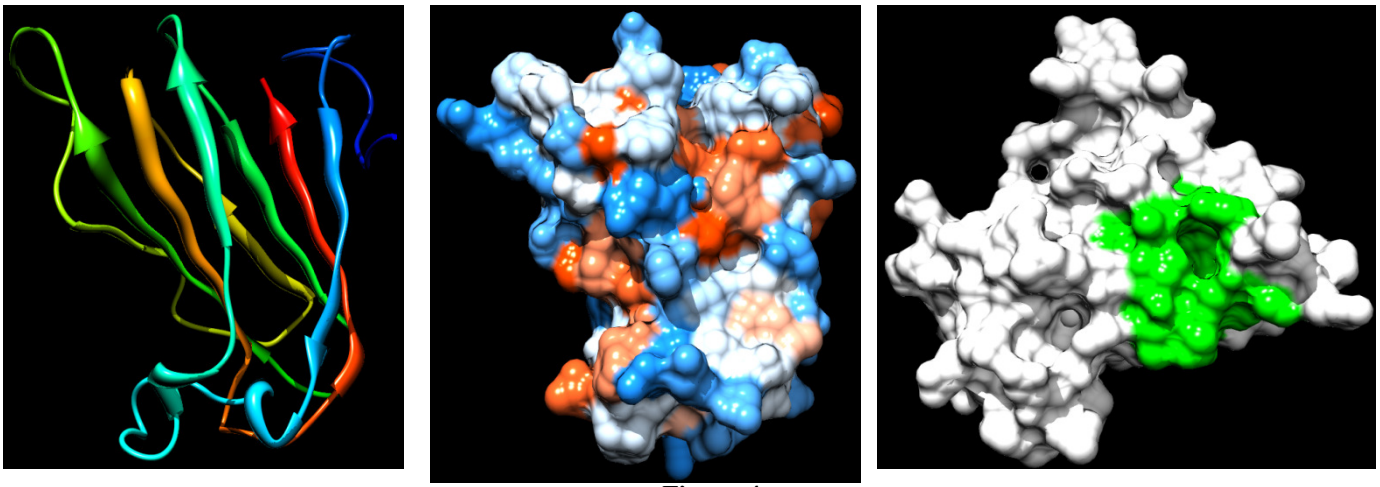

**Figure-4**
**Prediction of active site of Pfcalpain using CASTp. Upper panel showing active site prediction using CASTp where active site residues shown in green colour in between amino acid sequence of protein, Lower panel shows three-dimensional structure of Pfcalpain along with predicted active site in green**