International Journal of Chemical Studies

P-ISSN: 2349–8528
E-ISSN: 2321–4902
IJCS 2018; 6(2): 1426-1429

Received: 03-01-2018
Accepted: 07-02-2018

Pravas Ranjan Sahoo
Department of Veterinary Biochemistry, CVSc & A.H. Orissa University of Agriculture and Technology, Bhubaneswar, Odisha, India

Ramesh D
Department of Veterinary Physiology and Biochemistry, CVSc, KVAFSU, Gadag, Karnataka, India

Bijayendranath Mohanty
Department of Veterinary Parasitology, CVSc & A.H. Orissa University of Agriculture and Technology, Bhubaneswar, Odisha, India

Prakash Chandra Behera
Department of Veterinary Biochemistry, CVSc & A.H. Orissa University of Agriculture and Technology, Bhubaneswar, Odisha, India

Correspondence
Pravas Ranjan Sahoo
Department of Veterinary Biochemistry, CVSc & A.H. Orissa University of Agriculture and Technology, Bhubaneswar, Odisha, India

Insilico physiochemical and structural characterization of hypopothetical proteins of *Theileria annulata*

Pravas Ranjan Sahoo, Ramesh D, Bijayendranath Mohanty and Prakash Chandra Behera

Abstract
Theileria annulata is one important apicomplexon parasite which causes a serious disease such as tropical theileriosis in cattle. With advancement of bioinformatics and genetic sequencing, although lots of important protein of this parasite have been well characterized, but there are some hypopothetical proteins still existing needs to be physiochemically and structurally analyzed due to their high importance, through *insilico* platform. In this study, six hypothetical proteins were randomly selected and analyzed through in Pfam and CD Blast to find conserved domain in their structure. The physiochemical characters were characterized through EXPASY Protparam application programme and secondary, three dimensional structures were predicted through GORIV and SWISS homology modeling software tool respectively. It was found that the conserved domains from families like Der1, WD 40, AP2, and PA14 are present in this organism and maximum numbers of these proteins are found as stable, soluble and hydrophilic in nature that will provide a suitable therapeutic target against this parasite. The structural organization study revealed that these proteins contain abundant number of random coils followed by alpha helix and exist in monomer with low QMEAN Z score upon SWISS homology modeling. So this study would provide a better platform for the researchers to develop a suitable therapeutic strategy against *Theileria annulata* in docking studies in nearest future.

Keywords: *Theileria annulata*, hypopothetical proteins, swiss homology modeling, conserved domains

Introduction
Theileria annulata is one important blood protozoan parasite which causes tropical theileriosis of cattle in the regions of Mediterranean and Middle East area, from Morocco to Western parts of India and China [1]. This Tropical theileriosis is a *hyalomma* tick borne diseases most prominently characterized with fever and lymh node enragement [2] causing a severe constraint in bovine productivity among the dairy farmers. It is a lymphoproliferative disease with clinical signs of high rise in temperature, weakness, anemia, jaundice, heart failure, emaciation, destruction of the lymphoid system, pulmonary oedema and even if death which causes a greater economic loss to the farmers [3]. At present, although there is existing of suitable therapeutic regimen against this disease, but there is need of new drugs which can be designed by exploiting different proteins of this organism through molecular docking phenomena [4]. The structural and functional characterization of maximum proteins of this protozoon has already been done with advancement of whole genome sequencing and bioinformatics tools but there is presence of some uncharacterized hypothetical proteins (HP) with unknown functions in this organism [5]. These HP proteins can provide a better pathway to design drug molecules against this disease in modern biological research. As there is no evidence of experimental study about these HP proteins till now, so there is extensive need of to study and characterize these HP proteins both by physiochemically and structuarally. Many potential antiprotozoal drugs may be developed after improved understanding of these HP proteins [6]. So this present study was carried to characterize some hypopothetical proteins of *Theileria annulata* for better drug discovery in nearest future.

Materials and Methods
Retrieval of HP proteins
Six HP proteins of *Theileria annulata* were randomly selected from Kyoto Encyclopedia of Gene and Genomes (KEGG) data base.

The amino acid sequences of these proteins with accession number gi/84994126, gi/84994190, gi/84996643, gi/85000171, gi/84994240 and gi/84997095 were retrieved from National Centre for Biotechnological information (NCBI) under FASTA format.

Estimation of conserved domain

HP proteins were searched for the conserved domain by using Basic local alignment search tool (psi-BLAST) from protein family data base (Pfam) [7]. A batch CD search was done by using multiple sequence alignment (MSA) and 3D structures for the homologous domains were observed on Pfam [8].

Physiochemical characterization of HP proteins

The physiochemical parameters such as molecular weight, theoretical pI, instability index, aliphatic index, grand average of hydropathicity (GRAVY), extinction coefficient, absorbance of the these six HP proteins were estimated under ProtParam characterization tools on the Expert Protein Analysis System (Expasy) server[9]. The solubility of these proteins was also checked to find out whether the protein is transmebrane or cytoplasmic in nature through SOSUI programme [10].

Secondary structure prediction

The secondary structures such as alpha helix, extended strand, and random coils of HP proteins were found out through GOR IV application programme [11] method.

Tertiary (3D) structural characterization of HP proteins

The 3D structure of these six proteins was predicted from a suitable template through Swiss homology modeler [12] and visualized by Rasmol molecular graphics.

Result and Discussion

In this study, the physiochemical properties of all six HP proteins were given in Table No I. It has been found that all six proteins have amino acids varied from 1070 (gi/84997095) to 1577 **(**gi/84996643). The value of pI revealed that all six proteins are acidic in nature except gi/ 84994126 and gi/84996643 which are basic in nature due presence of more amount of arginine, lysine and histidine amino acids similar to the findings of [13]. The extinction coefficient was found higher in gi/ 84994126 and gi/84996643 HP protein than other proteins inferred that these two proteins have higher lambda max than other. The instability index, indicator of stability of protein, showed that all six proteins are stable in nature except the gi/ 84994190 proteins having a higher value above 40[14]. The aliphatic index, which indicates relative volume of aliphatic side chain, was found higher in all HP proteins except gi/ 84994190 inferred that this protein is thermo unstable in nature. However the GRAVY value was found less than zero in all proteins suggested that all six proteins are hydrophilic in nature due to presence more amount of polar amino acids in their structure [15].

Table 1: Showing the physiochemical parameters of selected Hypopothetical proteins of *Theileria annulata*

	gi/ 84994126	gi/ 84994190	gi/84996643	gi/85000171	gi/84994240	gi/84997095
No. of Amino acids	1337	1477	1577	1158	1642	1070
Mol. Weight	152472.66	172051.78	179837.62	127631.50	177929.37	121879.81
pI	8.78	5.29	8.53	5.31	4.75	5.53
Ext. coefficient	199940	127065	191430	99200	94200	120125
Absorbance	1.305	0.739	1.064	0.777	0.529	0.986
Instability index	39.85	47.90	37.10	26.86	28.42	37.23
Aliphatic index	82.06	66.36	93.66	65.19	76.58	75.36
GRAVY	- 0.314	-0.923	-0.194	-0.707	-0.369	-0.541

The SOSUI results of this study showed that all six proteins are cytoplasmic in nature except gi/ 84994126 which has two transmembrane helices in its structure as shown in Table No 2. Due to presence of transmembrane helices in the first protein, it can be inferred that this protein may act as gateways to permit the transport of specific substances across the biological membrane [16].

Table 2: Showing the SOSUI results Hypopothetical proteins

	Types of Protein	No of TM helices	N terminal	Transmembrane region	C terminal	length
gi/ 84994126	Transmembrane	2	4	NSLITNTWKLCFLCMLSSKFSL	25	22
			1120	IPPLTGFYVLLSTITAFVSYFFN	1142	
gi/ 84994190	cytoplasmic	-------	-------	-------	-------	-------
gi/ 84996643	cytoplasmic	-------	-------	-------	-------	-------
gi/ 85000171	cytoplasmic	-------	-------	-------	-------	-------
gi/ 84994240	cytoplasmic	-------	------	-------	-------	-------
gi/ 84997095	cytoplasmic	-------	-------	-------	-------	-------

The CD Blast revealed that all the protein except gi/84994190 and gi/ 84994240 have putative conserved domain shown in Table no 3. The result showed that the hypopothetical protein gi/ 84994126 was found to be a member of Der1-like family. So this protein may be involved in the degradation process of some misfolded endoplasmic reticulum (ER) luminal proteins [17]. The gi/ 84996643 protein was found a member of WD 40 super family indicating a wide range of functions ranging from signal transduction, transcription regulation, cell cycle control, autophagy and apoptosis in the protozoa [18].The psi BLAST results showed that gi/ 85000171 HP is AP2 domain containing protein acting as a transcription factor, not only plays an important role in the ABC model of flower development but also helps in disease process in this protozoa [19]. However, the HP protein gi/ 84997095 has PA14 domain suggesting a binding function, rather than a catalytic role, may helps in carbohydrate binding in the protozoa [20].

Table 3: Conserved Domain of Hypopothetical proteins

. Sl. No.	Name of Protein	Type of Conserved Domain
1.	gi/ 84994126	Der1-like family
2.	gi/ 84994190	No putative domain
3.	gi/ 84996643	WD 40 Superfamily
4.	gi/ 85000171	AP2 domain-containing protein
5.	gi/ 84994240	No putative domain
6.	gi/ 84997095	PA14 domain

The secondary structural organization of these HP proteins revealed that all six proteins have more abundance of random coils, may be due to absence of specific, stabilizing interactions [21]. On other hand, there is less variation in the composition the alpha helix and extended strands which make them in significant regarding secondary structure of six HP proteins. The result of secondary structural organization was shown in Table No 4.

Table 4: Showing the composition of different secondary structure of HP proteins in *Theileria annulata*

	gi/ 84994126	gi/ 84994190	gi/84996643	gi/85000171	gi/84994240	gi/84997095
Alpha helix (%)	22.66	35.95	18.90	14.94	18.33	21.59
Extended Strand (%)	22.89	10.36	26.95	26.68	23.14	23.36
Random coils(%)	54.45	53.69	54.15	58.38	58.53	55.05

The Swiss homology modeling report showed that all the hypopothetical proteins may be predicted in exist as monomer in 3 Dimensional organizations which is similar to the finding of [22]. All the best model of proteins was shown in Fig 1, taken according to the suitable template with lowest QMEAN Z score.

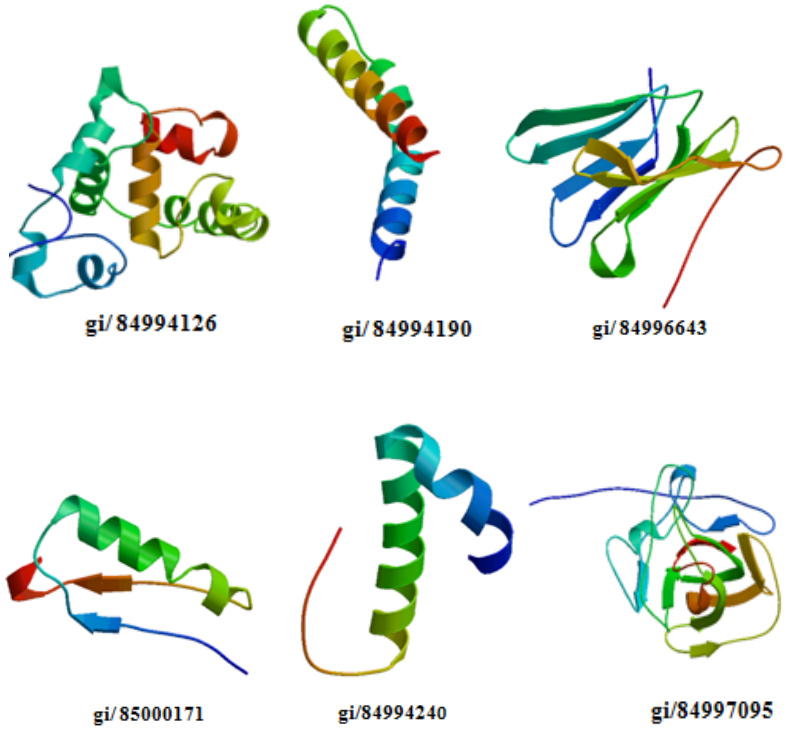

Fig 1: Showing 3D structure of Hypopothetical proteins of *Theileria annulata* in cartoon model as displayed in Rasmol

Conclusion

This study can be concluded that these six hypopothetical proteins may have important function such as virulence, channel transport and cellular processes in *Theileria annulata.* The homology modeling and the physiochemical prediction made these HP proteins as a suitable therapeutic target against this protozoa infection. It can be concluded that, *in silico* study would be a better platform for characterization of these proteins due to development of potential bioinformatics tools and databases. So this study will help for the development of potential and effective targets against the *Theileria annulata* through molecular docking in nearest future.

Acknowledgment

The authors are thankful to the Dean, College of Veterinary Science and Animal Husbandry, OUAT, Odisha, India, for providing the necessary facilities to undertake this study.

References

1. Lempereur L, Larcombe SD, Durrani Z, Karagenc T, Bilgic HB, Bakirci S, *et al*. Identification of candidate transmission blocking antigen genes in *Theileria annulata* and related vector-borne apicomplexan parasites. BMC Genomics. 2017; 18:438.
2. Gharbi M, Souidi K, Boussaadoun MA, Rejeb A, Jabloun S, Gnaoui A *et al*. Dermatological symptoms in tropical theileriosis (*Theileria annulata* infection), a review. Scientific and Technical Review of the Office International des Epizooties. 2017; 36(3):225-229.
3. Weir W, Karagenç T, Gharbi M, Simuunza M, Aypak S, Aysul N *et al*. Population diversity and multiplicity of infection in *Theileria annulata*. International Journal of Parasitology. 2011; 41(2):193-203.
4. El-Deeb WM, Younis EE. Clinical and biochemical studies on *Theileria annulata* in Egyptian buffaloes (Bubalus bubalis) with particular orientation to oxidative

stress and ketosis relationship. Veterinary Parasitology, 2016; 36(3):301-305.
5. Adinarayana KP, Sravani TS, Hareesh C. A database of six eukaryotic hypothetical genes and proteins. Bioinformation, 2016, 128-130.
6. Galperin MY, Koonin EV. Searching for drug targets in microbial genomes. Current Opinion in Biotechnology, 1999; 10:571-578.
7. Marchler-Bauer A *et al*. CDD: NCBI's conserved domain database. Nucleic Acids Research, 2015; 43(D):222-6.
8. Punta M, Coggill PC, Eberhardt RY, Mistry J, Tate J *et al*. The Pfam protein families database. Nucleic Acids Research, 2012; 40(D):290-301.
9. Gasteiger E, Hoogland C, Gattiker A, Duvaud S, Wilkins MR, Appel RD *et al*. Protein Identification and Analysis Tools on the ExPASy Server*;* (In) John M. Walker (ed): The Proteomics Protocols Handbook, Humana Press, 2005, 571-607.
10. Mitaku S, Hirokawa T, Tsuji T. Amphiphilicity index of polar amino acids as an aid in the characterization of amino acid preference at membrane-water interfaces. Bioinformatics. 2002; 18:608-616.
11. Garnier J, Gibrat JF, Robson B. GOR secondary structure prediction method version IV. Methods in Enzymology, R.F. Doolittle Ed. 1996; 266:540-553.
12. Biasini M, Bienert S, Waterhouse A, Arnold K, Studer G, and Schmidt T *et al*. Swiss-Model: modeling protein tertiary and quaternary structure using evolutionary information. Nucleic Acids Research. 2014; 42:W252-W258.
13. Bidkar AP. In silico Structural and Functional analysis of Hypothetical proteins of Leptospira Interrogans. Biochemical Pharmacology, 2014; 3:136.
14. Sahoo PR, Mohapatra S, Sahoo G, Behera PC. Deciphering physiochemical and structural characterization of p53 tumor suppressor protein in domestic animals through in silico approaches. International Journal of Chemical Studies. 2018; 6(2):254-257.
15. Ikai AJ. Thermostability and aliphatic index of globular proteins. Journal of biochemistry, 1980; 88:1895-1898.
16. Carpenter EP, Beis K, Cameron AD, Iwata S. Overcoming the challenges of membrane protein crystallography. Current Opinion in Structural Biology. 2008; 18(5):581-586.
17. Hitt R, Wolf DH. Der1p, a protein required for degradation of malfolded soluble proteins of the endoplasmic reticulum: topology and Der1-like proteins**.** FEMS yeast research. 2004; 4:721-729.
18. Stirnimann CU, Petsalaki E, Russel RB, Muller CE. WD40 proteins propel cellular networks. Trends in Biochemical Science. 2015; 35(10):565-74.
19. Balaji S, Babu MM, Iyer LM, Aravind L. Discovery of the principal specific transcription factors of Apicomplexa and their implication for the evolution of the AP2-integrase DNA binding domains. Nucleic Acids Research. 2005; 33(13):3994-4006.
20. Rigden DJ, Mello LV, Galperin MY. The PA14 domain, a conserved all-beta domain in bacterial toxins, enzymes, adhesins and signaling molecules. Trends in Biochemical Sciences. 2004; 29(7):335-339.
21. Smith LJ, Fiebig KM, Schwalbe H, Dobson CM. The concept of a random coil: Residual structure in peptides and denatured proteins. Folding and Design. 1996; 1(5):R95-R106.
22. Habibi G. In silico analysis of Ta9 gene polymorphism in an Iranian *Theileria annulata* schizont-infected cell line S15 vaccine strain and native isolates. Archives of Razi Institute, 2016; 71(2):71-80.