IJBPAS, September, 2015, 4(9): 5781-5791

ISSN: 2277–4998

International Journal of Biology, Pharmacy and Allied Sciences (IJBPAS)

'A Bridge Between Laboratory and Reader'

www.ijbpas.com

RETICULON-4 ANTIGEN OF *ECHINOCOCCUS GRANULOSUS*- A PROMISING VACCINE CANDIDATE, AN *IN-SILICO* ANALYSIS

VARUN CHAUHAN AND UMAR FAROOQ*

Molecular and Immuno–Parasitology Research Laboratory (MIPL), Department of Biotechnology, Shoolini University, Solan (H.P.), India

***Corresponding Author: E Mail: ufarooq8@gmail.com**

ABSTRACT

Hydatidosis, a cystic disease, caused by a tapeworm *Echinococcus granulosus*, is a chronic zoonotic disease responsible causing considerable human morbidity and mortality. Diagnosis, treatment as well as control of this cosmopolitan disease isvery difficult. *Reticulon-4*,a transmembrane protein of *E. granulosus*, is highly expressed during all stages of its life cycle, thus indicating it to be an important protein for the parasite survival. The present study was aimed to identify promiscuous peptides of HLA Class-I, II and B cell epitopestargeting *Reticulon-4* protein using *in-silico* tools. The identified T and B cell epitopes were confirmed by visualizing their locations on the respective 3D modeled protein. In order to ascertain the binding pattern of the identified peptides with HLA alleles, the class I peptides were docked with HLA*A 02:01 allele and class II peptides with HLA*DRB1 01:01 allele. The identified B cell and T cell epitopes in the present study offers insightsin better understanding the humoral and adaptive immune responses generated in an intermediate host in response to *Reticulon-4* protein of the parasiteand could provide a platform to facilitate the development of subunit vaccine.

Keywords: *Echinococcus granulosus*, Reticulon-4 protein,T cell epitope, Docking, HLA allele

INTRODUCTION

Cystic echinococcosis (CE) caused by taeniid tapeworm *E. granulosus*is recognised as one of the world's major zoonotic diseasethat infects at least 50 million people globally [1,

2].Up to 3 million people are infected currently with CEand, even in some areas, abdominal ultrasound and chest X-ray analysis revealed about 10% of the population with detectable hydatid cysts [3, 4, 5].Thus, the disease is recognised as one of the major re-emerging infection among helminths causing considerable infections in human population and is a point of concern. Thus a vaccine for *E. granulosus* is needed.

With the accelerating growth in the field of Immunoinformatics, the epitope based vaccines have shown their potential in immune prevention against several infections [6]. A major challenge in developing an epitope based vaccines is to establish the immunogenic sites of the protein/antigen that exhibit the greatest immunogenic response [7]. By using certain Immunoinformatics tools an immunologist can analyse the sequence areas in a protein/antigen with potential binding sites for B and T cell epitopes, which in turn may help in development of subunit vaccines. Through constant improvements in the field of bioinformatics it has now become possible to synthesize peptides of varying lengths [8].

There are few reports revealing that the prevention of CE is quite feasible by a molecular vaccine. Kouguchi *et al*, and Katoh *et al*, synthesized vaccine against the Emy162 recombinant antigen and Em95 antigen from *E. multilocularis*in order to protect against the larval-stage infection, the main infective stage in intermediate host [9, 10].There results suggested that designing a molecular vaccine targeting *E. granulosus* is also possible. No human vaccine has been reported till now for treatment of CE. Immunization against the larval stage (main infective stage in human) of the parasite may provide protection. Reticulon-4 is a transmembrane protein which is associated with endoplasmic reticulum [11]. Though its function is not clearly understood but recently it is reported to be an important player in malarial and Schistosomal infections [12, 13].The expression of Reticulon-4 protein was also found high in larval stages of Schistosoma [13].Several vaccine candidates of *E. granulosus* have been identified, among which the gene encoding Reticulon-4 protein is one of the highly expressed protein in all stages- adult worms, oncospheres, scolesces and hydatid cyst membranes [14].Thus keeping these facts in view, the present study was aimed to identify the immunogenic regions in Reticulon-4 protein of *E. granulosus* that could elicit the cytotoxic T cell and B cell response in host, a critical step in the development of subunit vaccines.

METHODOLOGY

Retrieval of amino acid sequence of the protein

The amino acid sequence of Reticulon-4 protein of *E. granulosus* was obtained from National Centre for Biotechnology Information (NCBI) with gene ID-EUB60463.1

MASDTVAPKNDIVYICPAIERLSPRVISLIYWK KPVESAIVMATGFTLLFSIGCLSLISVFAYVCL AILCCTGAARVYYDLITSKRDETASPPFSDWFI KDKEALRQRVHDYVNEKFDKTESTFQDLRH YLLIENYIDSLKAIRSTHVSPVNSGWLLQFDH PRSSRLYADLHDPDDLRLVQIDKFTEKPKRE WRKIWSKAKAQFEKIPYLKKEKQN

Figure 1: The amino acid sequence of Reticulon-4 protein

Prediction of the secondary structure

The secondary structure of Reticulon-4 protein was predicted by the improved self-optimized prediction method (SOPMA) software (http://npsa-pbil.ibcp.fr/cgi-bin/npsa_automat.pl?page=/NPSA/npsa_sopma.html) [15].The parameters of similarity threshold and window width were set to 8 and 17, respectively, whilst the remaining parameters were not adjusted.

Prediction of the tertiary structure

The target protein was modeled using I-Tasser [16].A template model was obtained for protein by submitting FASTA format protein sequence and modeling it.

Identification of T-cell epitopes

For the identification of peptides of HLA class-I T cell epitopes, the servers NetMHCpan 2.4, NetMHC3.4, IEDB-ANN and IEDB-SMM servers were used [17, 18, 19, 20].The alleles selected for HLA class-I were, HLA-A*01:01, HLA-A *02:01, HLA-A *11:01, HLA-A*30:01, and HLA-B*35:01. For identification of HLA Class-II peptides, the alleles selected were: HLA-DRB1*01:01, HLA-DRB1*07:01, HLA-DRB1*11:01, HLA-DRB1*12:01 and HLA-DRB1*15:01. The servers used for the prediction of class-II alleles specific peptides were: NetMHCIIpan 3.0, ProPred, MultiPred, IEDB-ANN and IEDB-SMM. T cell epitopes were classified based on their binding affinity for Human Leukocyte Antigen (HLA) alleles, using the half-maximal inhibitory concentration of a biological substance (IC_{50}) as the unit of measure, as follows: high-affinity binding, IC_{50s} of <50 nM; intermediate-affinity binding, IC50s of <500 nM; and low-affinity binding, IC_{50s} of <5,000 nM.

Identification of Linear B cell epitopes

The Linear B cell epitopes were identified using ABCpred, Ellipro, BCPREDS and BepiPred 1.0 online web servers [21, 22, 23]. In ABCpred, prediction was done at a window length of 16 and the threshold was

set from 0.8 to 1.0. In BCPREDS the classifier specificity was set at 80% and the epitope length was set 20. Using BepiPred server, the threshold for epitope assignment was set at 0.35. In ellipro, prediction was made at a minimum level of 0.5 to the most stringent level of 1.0 and the maximum distance for residue clustering was kept at 6.0 Å. Selection of diagnostic and vaccine proteins as B cell epitope candidates was based on the number of epitopes predicted with a minimum cutoff score of 0.8. The regions which were predicted to be B cell epitopes by atleast three servers were selected.

Characterization of the predicted peptides

The predicted T cell epitopes were characterized by using different parameters including 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 calculated using ExPASy ProtParam tool (http://web.expasy.org/protparam/).

Identification of transmembrane regions

The transmembrane helices in Reticulon-4 protein were predicted using TMHMM server (www.cbs.dtu.dk/services/TMHMM/).

Modeling of peptides and HLA alleles

The tertiary structure of the peptides was modeled using the online PEP-FOLD server [24]. The X-ray crystallographic model of HLA*A 02:01 and HLA*DRB1 01:01 alleles were retrieved from Protein Data Bank (PDB) bearing PDB ID-1QEW and 2g9h. The bounded ligand was successfully removed from the alleles.

Molecular Docking

The receptor and the ligand files were prepared using AutoDock 4.2. The energy minimization was also done for the receptor after successfully removing the previously bound ligand. Then, molecular docking of the promiscuous peptides with both the HLA alleles was carried out using AutoDockVina [25].

RESULTS

Prediction of secondary structure of Antigens

In order to assess the antigenic features of the protein, secondary structure was predicted using SOPMA Server software. A greater proportion of extended strands and random coils present in the structure of the protein corresponded with an increased likelihood of the protein forming an antigenic epitope. The predicted secondary structure results are demonstrated in Figure 2.

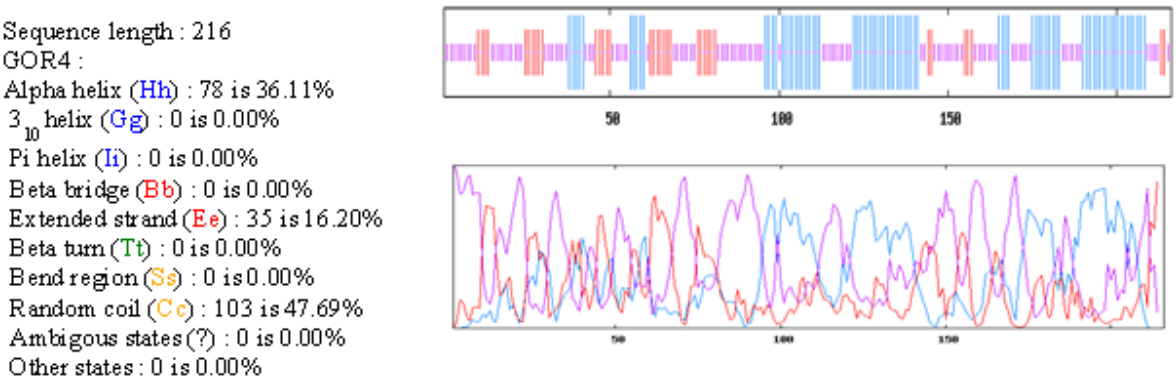

Figure 2. Secondary structure prediction results using SOPMA server. Lines in different colors represent different secondary structures: Blue, α helix; green, β turn; red, extended strand; and purple, random coil.

Identified T cell epitopes

The criteria for selecting the promising T cell epitopes, was that it should have least IC_{50} value, should be present on the surface of the protein, should be charged and predicted by at least three out of four different servers used. Five HLA Class I and four HLA Class II peptides were identified to be most promising based on the above mentioned criteria (Table 1). The position of each predicted epitope was confirmed by visualizing on its 3D modeled protein using Pymol viewer (Figure 4).

Identified B cell epitopes

The B-cell epitopes were predicted using Ellipro, BCEPREDS, BepiPred and ABCpred online application tools. All the epitopes identified were present on the surface of protein. Hydrophobic content of the identified epitopes were also analyzed in order to identify regions with higher probability to interact with immunoglobulins. The polar, non-polar and charged residues of each predicted epitope is shown in Table 1. The position of each predicted epitope was confirmed by visualizing on its 3D modeled protein using Pymol viewer (Figure 3).

Table 1: Promising regions within Reticulon-4 protein of *E. granulosus* bearing HLA-Class I, II T cell and Linear B cell epitopes

Peptide	Sequence[a]	Mol. wt	pI	Position	Type of immune response generated	Tools[b]
T1	CLAILCCtgAA**R**Vyy	1619.9	7.95	64	T_H	5,6,8,9
T2	**E**nyl**D**sL**K**AI**R**stHV	1745.9	6.85	133	T_H	5,6,7,8,9
T3	nyl**D**sL**K**AI**R**stHVs	1703.9	8.60	134	T_H	5,6,7,8,9
T4	**D**L**R**HyLLI**E**nyl**D**sL	1877.1	4.54	125	T_H	5,6,8,9
C1	sLIsVFAyV	998.1	5.24	55	T_C	10,11,5,6
C2	AqF**EK**IPyL	1108.3	6.05	201	T_C	10,11,5,6
C3	VMAtgFtLL	952..1	5.49	40	T_C	10,11,5,6
C4	**R**LsP**R**VIsL	1040.2	12	20	T_C	10,5,6

C5	sLIywKKPV	1133.4	9.70	27	T_C	10,5,6
B1	MAsDtVAPKnDIVylCPAIE	2150.4	4.03	1	B	2,3,4
B2	KRDEtAsPPFsDwFIKDKEA	2367.6	4.86	84	B	1,2,3,4
B3	FtEKPKREwRKIwsKAKAqF	2565.0	10.56	184	B	1,2,3

a- Non polar residues are shown in upper case, polar residues in lower case and charged residues are depicted in bold.

b- Tools that recognized the corresponding B or T cell epitope or a part of the epitope are numbered as follows: (1) BCPREDS, (2) BepiPred 1.0b (3) ABCpred, (4) Ellipro, (5) IEDB-NN (6) IEDB-SMM, (7) NetMHCIIpan 3.0, (8) MultiPred, (9) ProPred, (10) NetMHC 3.4, (11) NetMHC2.4 .

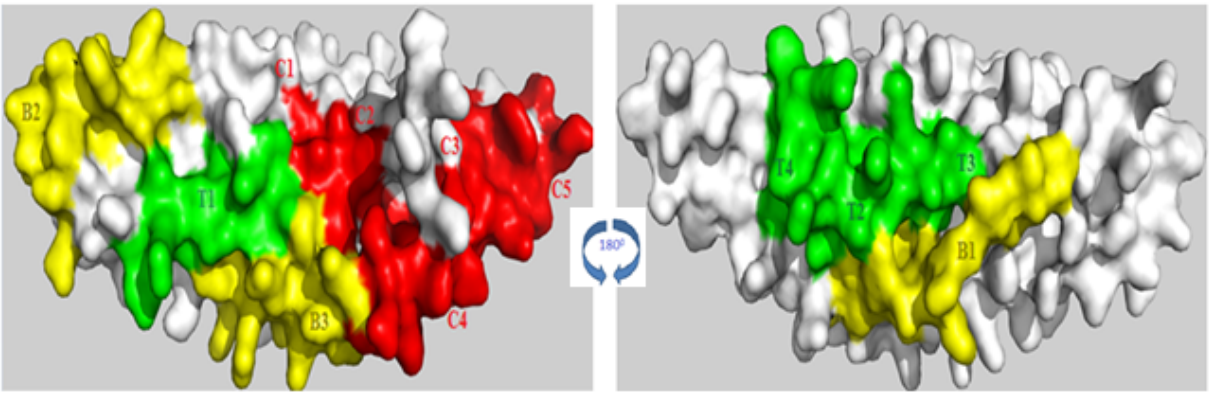

Figure 3: Surface view of the Reticulon-4 protein showing the location of identified T and Linear B cell epitopes. B cell epitopes are shown red in color and T cell epitopes are shown green in color. All the identified epitopes are located on the surface of the protein.

Prediction of transmembrane regions in protein

The Reticulon-4 protein was observed to possess a transmembrane region at position 48-70 as predicted byTMHMM server (Figure 4).

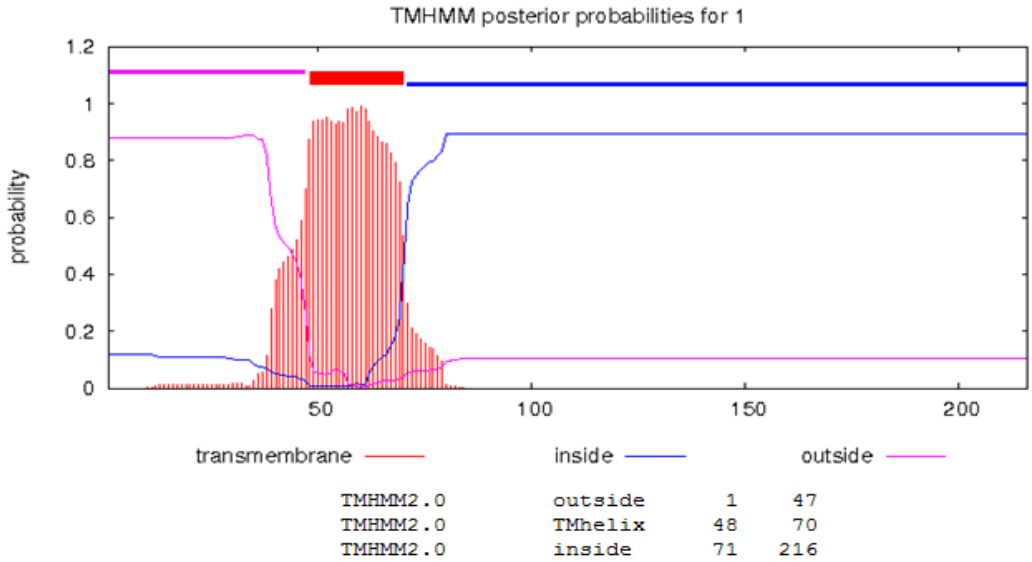

Figure 4: Transmembrane region in Reticulon-4 protein at position 48-70.

Molecular Docking

The docking between the predicted T cell epitopes and HLA alleles was carried out using AutodockVina. The peptides C1, C2 and C5 showed the best binding patters with HLA-A*02:01 allele (1QEW) with least docking energies. Similarly, T1 and T3 peptides C5 showed the best binding patters with HLA-DRB1*01:01 allele (2G9H) with least docking energies. The interaction pattern of the peptides with their respective alleles (receptor) is shown in Figure 5.

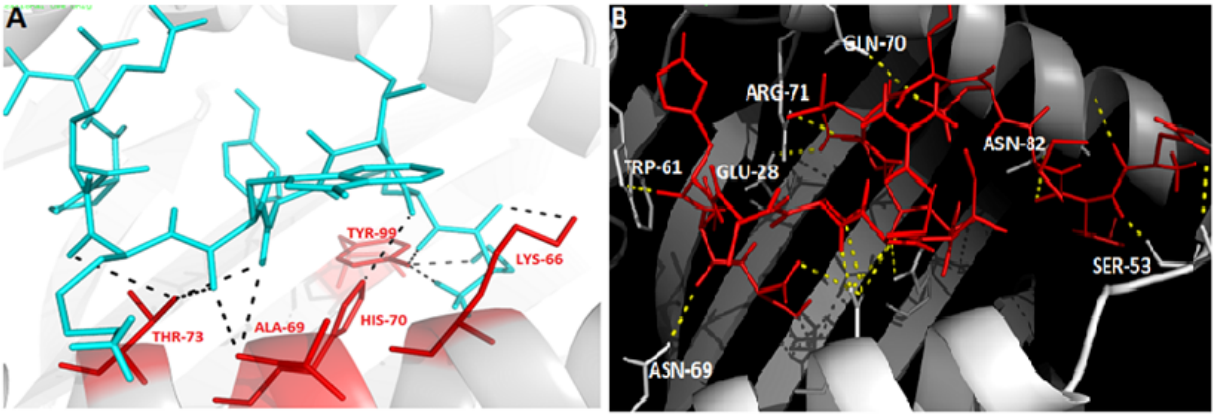

Figure 5: Binding interaction between peptides and HLA molecules. A-Binding pattern of peptide C1 with HLA Class I allele HLA-A* 02:01 (1QEW), B-Binding pattern of peptide T1 with HLA Class II allele HLA-DRB1* 01:01 (2G9H).

DISCUSSION

During the course of infection, the immune system reacts to a multitude of foreign antigens for which T cells are crucial for generating an efficient immune response. For the development of subunit vaccine against any pathogen, knowledge of the interactions between HLA alleles, peptides, and host immune cells is crucial. The identification and use of immunodominant peptides/epitopes can be potentially employed as vaccine as they couldhelp in efficient priming of the host immune systemas the immune response is always generated by exposure of such regions.With the accelerating growth of bioinformatics techniques, an immunologist can analyze the immunodominant areas in the protein sequence with potential binding sites for B and T cells, which in turn leads to the development of new vaccines. Molecular docking is a key structure-based method of immunoinformatics and has proved to be a rapid and accurate method for evaluating peptide binding to MHCs [26].

Keeping these points under consideration, the present study was designed to find the Cytotoxic T cell and B cell epitopes of Reticulon-4 protein of *E. granulosus*. Though the function of this protein is not clearly understood, its expression level is found to be high in helminths like *Schistosoma* [27]

and *Echinococcus*[15]during their larval stages, the infective stage of the parasites, thus *implying* its importance in parasite survival and growth. Such points were the keen of interest for us, thus we identified the immunogenic regions of the protein using several in-silico tools which could develop the adaptive and humoral immunity in host in response to the antigenic protein. First of all, the secondary structure of Reticulon-4 protein was predicted in order to obtain the antigenic features of the protein. The primary factors involved in an epitope formation like hydrophilicity, antigenicity, flexibility, the exposed surface area were analyzed. The tertiary structure is a three-dimensional conformation of the naturally folded protein formed by further coiling and folding. It was a useful supplement to the prediction of the Reticulon-4 protein epitopes. The 3D structure of the protein was designed using I Tasser server and matched to the most relevant templates in the PDB. The template identified by the server was considered and designed using pymol. The protein taken in this study showed above 40% sequence identity, which is in accordance with the findings from other studies where sequence identity ≥40% proved to be significantly accurate [28].

The epitopes predicted to be most promising were docked with the MHC class I allele, HLA-A 02:01. The criterion for choosing the most promising epitopes was based on the combined results of all the identification servers used, as well as IC_{50} value<50 nM. The docking results suggested the higher affinity of T1 and T2 peptides towards the HLA-A 02:01allele. The location of the identified epitopes was viewed on the designed 3-D modeled template.Brinda and Vishveshwara, demonstrated that hydrophobic residues can function as weak interface hubs and arginine presence in any peptide increases the solvent accessibility of that peptide and also may interact with several amino acid residues facilitating subunit interaction [29].The identified epitopes were found to possess various degrees of polar and non-polar residues and also most of them possessed Arginine residues as well, thus implying high solvent accessibility of the predicted peptides.

In conclusion, the following study led to the identification of potential immunogenic epitopes present in the Reticulon-4 protein which should further be tested for their immunoreactivity using *in vitro* and *in vivo* approaches to support the *in-silico* findings that may have an enhanced safety and efficacy.

REFERENCES

[1] McManus D, Smyth J. Hydatidosis: changing concepts in epidemiology and speciation. Parsitol. Today.2, 1986, 163–167.

[2] Garcia HH, Del B, Neurocysticercosis OH. Updated concepts about an old disease. Lancet Neurol.4, 2005, 653–661.

[3] McManus DP, Zhang W, Li J, Bartley PB. Echinococcosis. Lancet 362, 2003, 1295–1304.

[4] Craig Philip S, Donald P, McManus, Marshall W,Lightowlers, Jose A,Chabalgoity Hector H, Garcia C M. Gavidia RH, Gilman et al. Prevention and control of cystic echinococcosis. Lancet Infect. Dis. 7, 2007, 385–394.

[5] Li T, Akira I, Renqing P, Yasuhito S, Xingwang C, Dongchuan Q, Ning X, Philip S. Craig. Post-treatment follow-up study of abdominal cystic echinococcosis in tibetan communities of northwest Sichuan Province, China. PLoSNegl. Trop. Dis., 5, 2011, e1364.

[6] Ben-Yedidia T,Arnon R. Towards an epitope-based human vaccine for influenza. Hum Vaccin. 1, 2005, 95-101.

[7] Pfaff E, Thiel HJ, Beck E, Strohmaier K and Schaller H. Analysis of neutralizing epitopes on foot-and-mouth disease virus. J Virol. 62,1988, 2033-2040.

[8] Corradin G, Villard V, Kajava AV. Protein structure based strategies for antigen discovery and vaccine development against malaria and other pathogens. Endocrine Metabolic & Immune Disorders. Drug Targets. 7, 2007, 259-265.

[9] Kouguchi H, Matsumoto J, Yamano K, Katoh Y, Oku Y, Suzuki T and Yagi K.*Echinococcus multilocularis*: purification and characterization of glycoprotein antigens with serodiagnostic potential for canine infection. ExpParasitol 128,2011, 50-56.

[10] Katoh Y, Kouguchi H, Matsumoto J, Goto A, Suzuki T, Oku Y and Yagi K. Characterization of emY162 encoding an immunogenic protein cloned from an adult worm-specific cDNA library of *Echinococcus multilocularis*. BiochimBiophysActa. 1780,2008, 1-6.

[11] van de Velde HJ, Senden NH, Roskams TA, Broers JL, Ramaekers

FC, Roebroek AJ, Van de Ven WJ. NSPencodedreticulons are neuroendocrine markers of a novel category in human lung cancer diagnosis. Cancer Res. 54,1994, 4769-4776.

[12] Manjurano A, Clark TG, Nadjm B, Mtove G, Wangai H, SepulvedaN, Campino SG, Maxwell C, OlomiR, RockettKR, Jeffreys A, Consortium M, Riley EM, Reyburn H, DrakeleyC. Candidate Human Genetic Polymorphisms and Severe Malaria in a Tanzanian Population. PLoS ONE. 7(10), 2012,e47463.

[13] Prosdocimi F, Faria-Campos AC, Peixoto FC, Pena SDJ, Ortega JM, Franco GR. Clustering of Schistosoma mansoni mRNA Sequences and Analysis of the Most Transcribed Genes: Implications in Metabolism and Biology of Different Developmental Stages,MemInstOswaldo Cruz, Rio de Janeiro. 97(I), 2002, 61-69

[14] Zheng H, Zhang W, Zhang L, Zhang Z et al. The genome of the hydatid tapeworm Echinococcus granulosus. Nature genetics. 45(10), 2013, 1168-1175.

[15] Geourjon C,Deleage G. SOPMA: significant improvements in protein secondary structure prediction by consensus prediction from multiple alignments. ComputApplBiosci. 11,1995, 681-684.

[16] Ambrish R, Alper K, Yang Z. I-TASSER: a unified platform for automated protein structure and function prediction. Nature Protocols. 5,2010,725-738.

[17] Hoof I, Peters B, Sidney J, Pedersen LE, Sette A, Lund O, Buus S, Nielsen M. NetMHCpan, a method for MHC class I binding prediction beyond humans. Immunogenetics. 61, 2009, 1-13.

[18] Lundegaard C, Lamberth K, Harndahl M, Buus S, Lund O, and Nielsen M.. NetMHC-3.0: Accurate web accessible predictions of Human, Mouse, and Monkey MHC class I affinities for peptides of length 8-11. *NAR*, 36,2008, W509-512.

[19] Lundegaard C, Nielsen M, Lund O. The validity of predicted T-cell epitopes. *Trends Biotechnol*. 24,2006,537-538.

[20] Peters B, Sette A. Generating quantitative models describing the sequence specificity of biological processes with the stabilized matrix method. *BMC Bioinformatics*. 6: 2005, 132.

[21] Saha S,Raghava GPS. Prediction of Continuous B-cell Epitopes in an Antigen Using Recurrent Neural Network. Proteins. 65(1), 2006, 40-48.

[22] EL-Manzalawy Y, Dobbs D, Honavar V. Predicting linear B-cell epitopes using string kernels. J MolRecognit. 21,2008, 243-255.

[23] Jens VK, Claus L, Ole L, Morten N. Reliable B Cell Epitope Predictions: Impacts of Method Development and Improved Benchmarking. Plos Computational Biology. 8(12), 2012, e1002829.

[24] Thevenet P, Shen Y, Maupetit J, Guyon F, Derreumaux P, Tufféry P. PEP-FOLD: an updated de novo structure prediction server for both linear and disulfide bonded cyclic peptides. Nucleic Acids Res. 40, 2012, W288-293.

[25] Trott O., OlsonA. J., AutoDockVina: improving the speed and accuracy of docking with a new scoring function, efficient optimization and multithreading. Journal of Computational Chemistry. 31,2010, 455-461.

[26] Patronov A, Doytchinova I. T-cell epitope vaccine design by immunoinformatics. Open Biol 3, 2010, 120139.

[27] Prosdocimi F, Faria-Campos AC, Peixoto FC, Pena SDJ, Ortega JM, Franco GR. Clustering of Schistosoma mansoni mRNA Sequences and Analysis of the Most Transcribed Genes: Implications in Metabolism and Biology of Different Developmental Stages, MemInstOswaldo Cruz, Rio de Janeiro. 97(I),2002, 61-69.

[28] Vijayasri, S. and Agrawal S. Domain-based homology modeling and mapping of the conformational epitopes of envelope glycoprotein of West Nile virus. *J Mol Model*11:2005,248 – 255.

[29] Brinda KV, Vishveshwara S.Oligomeric protein structure networks: insights into protein-protein interactions. BMC Bioinformatics. 6,2005,296.