frontiers in Immunology

Immuno-informatics analysis to identify novel vaccine candidates and design of a multi-epitope based vaccine candidate against Theileria parasites

Anand Srivastava[1*], Prajna P. Kar[1]

[1]National Institute of Animal Biotechnology (NIAB), India

Submitted to Journal:
Frontiers in Immunology

Specialty Section:
Vaccines and Molecular Therapeutics

ISSN:
1664-3224

Article type:
Original Research Article

Received on:
12 Dec 2017

Accepted on:
06 Sep 2018

Provisional PDF published on:
06 Sep 2018

Frontiers website link:
www.frontiersin.org

Citation:
Srivastava A and Kar PP(2018) Immuno-informatics analysis to identify novel vaccine candidates and design of a multi-epitope based vaccine candidate against Theileria parasites. *Front. Immunol.* 9:2213. doi:10.3389/fimmu.2018.02213

Title

Immuno-informatics* analysis to identify novel vaccine candidates and design ~~generation~~ of a multi-epitope based vaccine candidate against *Theileria* ~~*annulata*~~*parasites

Running title: Multi-epitope based vaccine candidates against *Theileria* ~~*annulata*~~

Prajna Parimita Kar[1], Anand Srivastava[1*]

[1]National Institute of Animal Biotechnology (NIAB),

[*]Corresponding author:

Anand Srivastava

National Institute of Animal Biotechnology (NIAB)

D. No. 1-121/1, 4th & 5th Floors, Axis Clinicals Building,

Miyapur, Hyderabad, Telangana 500 049, India

Phone: +914023049412

Fax: +914023042740

Email: anand@niab.org.in

Key words: *Theileria*, immuno-informatics, multi-epitope vaccine, subunit vaccine

Abstract:

Theileriosis poses a serious threat to ruminants in tropical and subtropical countries. It is a tick-borne disease, caused by an apicomplexan parasite, *Theileria*. The high disease burden in animals causes huge economic losses to marginal farmers. Further, with increasing cases of resistance to commonly used drugs, it is highly desirable to develop better and cost-effective vaccines against theileriosis. The high disease burden in animals lead to huge economic losses, with increasing cases of resistance to commonly used drugs, there is a strong requisite for a cost-effective vaccine against theileriosis. The only available vaccine, live attenuated parasite vaccine, has many drawbacks and hence is unsuitable for controlling this disease. IThe immunoe-informatics has emerged as a useful tool in down selection of potential molecules for vaccine development. In this study, we haveused an immuno-informatics driven genome-wide screening strategy to identify potential vaccine targets containing important and effective dominant immunogens against *Theileria*. The proteome of *Theileria annulata* was screened for proteins with probability of plasma membrane localization or GPI anchor. The non-homologous proteins non-homologous to the host (bovine) were selected and their antigenicity was analyzed. The B-cell epitopes were identified in the selected proteins and were mapped in the modelled structure of the proteins. A total of nineteen linear epitopes in twelve proteins, which are exposed in the extracellular space and havinge the potential to induce protective antibodies were obtained. Additionally, CTL epitopes which are, peptides with 9-mer core sequence, were also identified, modelled and docked with bovine MHC-I structures. The CTL epitopes showing high binding energy with the bovine MHC-I were further engineered in silico to design a putativeform a multi-epitope vaccine candidate against *Theileria* parasites. The docking studies and molecular dynamics studies with the predicted multi-epitope vaccine candidate and modelled bovine TLR-4

exhibitedshowed strong binding energy, which suggestings that the complex is stable and the putative multi-epitope vaccine candidate can be a potentially good candidate for vaccine development.

Introduction:

Theileria spp. are tick--borne protozoan parasites which belong to the subphylum apicomplexa whichthat includes a number of other important pathogens such as *Plasmodium*, *Babesia*, *Toxoplasma*, *Cryptosporidium* and *Eimeria* species. *Theileria* spp. are predominantly parasites of ruminants and are transmitted transstadially to the host, transstadially, by ticks (Bishop et al., 2004; Sivakumar et al., 2014). These parasites cause huge economic loss to farmers in tropical and subtropical countries. Among various species of these parasites, the two most important ones include *T. parva* and *T. annulata*, which cause economically important diseases in cattle, east coast fever and tropical theileriosis, respectively. *T. parva* is transmitted by *Rhipicephalus appendiculatus* whereas *T. annulata* is transmitted by several species of *Hyalomma*.

The disease, theileriosis, is characterized by the lympho-proliferation of the host leucocytes after invasion by the parasites (Irvin, 1987). Other symptoms include fever, lymph node enlargement and anaemia. The untreated animal usually dies within 3-4 weeks of the infection. This disease kills over million animals every year and significantly reduces the productivity of the cattle (Norval et al., 1991; Gharbi et al., 2006). The cCurrent control measures for theileriosis include the use of acaricides (for controlling vector), chemotherapy (drug such as buparavaquone) and vaccination. Traditionally, control of tick infection has been performed by the application of acaricides. However, it is effective only when performed at the community level and regularly

for a prolonged period. The indiscriminate use of acaricides has beenis leading to resistance in tick, thus adding another dimension to the problem. The chemotherapy agents such as parvaquone, buparvaquone and halofuginone are effective during the initial stages of infection (McHardy et al., 1985). However, these chemotherapy agents are quite expensive and are not a good choice for the marginal farmers. In addition, reports of development of resistance to these chemotherapy agents areis a matter of deep concern (Mhadhbi et al., 2010). The only available vaccine for theileriosis is a live attenuated vaccine ~~which is effective for providing protection to vaccinated animal~~ (Pipano, 1966). However, this vaccine needs a cold chain for transportation and storage which leads to a huge increase in its cost (approximately $8–12 per dose). Further, some vaccinated cattle, especially immunocompromised cattle, remain as a source of transmission of parasites to other cattle (Maritim et al., 1989).

TIn general, the subunit/multi-epitope based vaccines are generally is considered to be safer as compared to a live attenuated vaccines (Kallerup, 2015). Also, the cost of production and distribution of subunit/multi-epitope vaccine is quite comparatively low, hence theseit would be economical for the marginal farmers. Thus, developing a subunit and/or multi-epitope vaccine for theileriosis, is urgently requiredhighly desirable. However, the efforts to develop the subunit vaccines against *T. annulata* and *T. parva* till date have not been successful till date (Nene and Morrison, 2016).

LIt is possible to develop long--lasting immunity against *Theileria* can be developed in cattle by repeated challenge them with *T. parva* and *T. annulata* sporozoites, in cattle by developing neutralizing antibodies (Musoke et al., 1982). TFurther, these neutralizing antibodies recognize the p67 protein of *T. parva* (Dobbelaere et al., 1984). IThe immunization study with recombinant p67 antigen hass been been shown to induce immunity in 50% of vaccinated cattle (Musoke et

al., 2005). Also, ~~in case of *T. annulata*,~~ antibodies against parasite protein SPAG1 and Tams1 in *T. annulata* blocks the invasion of leucocytes by sporozoite (Williamson et al., 1989; Boulter et al., 1994; Boulter et al., 1998).

A large body of evidence indicates that cellular immunity also plays a major role in protection against *T. annulata* and *T. parva*. The immunity can be adoptively transferred by CD8+, but not CD4+ T cells from vaccinated animals. This strongly suggests that cytotoxic T lymphocytes (CTL) play an important role in developing immunity against *Theileria* parasites (Taracha et al., 1995a; Taracha et al., 1995b). Further, major histocompatibility complex (MHC) class I-restricted CD8+ cells, the cytotoxic lymphocytes (CTLs), that target schizont-infected lymphocytes also play a primary role in mediating immunity against *Theileria* in cattle (MacHugh et al., 2009; Morrison, 2009).

The availability of comprehensive genomic, transcriptomic and proteomic datasets of *Theileria* parasite have provided opportunities for *in silico* mining of novel candidates for vaccine design (Gardner et al., 2005; Pain et al., 2005; Witschi et al., 2013). Immuno-informatics, which integrates the transcriptomics and proteomics through advances in computational and molecular immunological tools, has emerged as a new tool for identification of the target antigens for vaccine development (Davies et al., 2007; Tomar and De, 2014; Backert and Kohlbacher, 2015). Since the experimental methods are difficult and time-~~-~~consuming, the immuno-informatics approach can narrow down a vast number of potential molecules to be tested, thus increasing the chance of finding better candidates. Numerous studies have shown that epitopes based vaccines could be effective in elucidating protective immunity against various pathogens such as influenza A, hepatitis B and~~,~~ C virus, *Leishmania* and *Shigella* (Sominskaya et al., 2010; Stanekova and Vareckova, 2010; He et al., 2015; e Silva et al., 2016; Pahil et al., 2017). In this study, using

immunoe-informatics-driven vaccine target screening strategy, the available transcriptomic and proteomic data for *Theileria* parasites were analyzed to predict proteins/epitopes for vaccine development that could elicit protective humoral and cellular immune response. For an effective humoral response, the antigen must be antigenic and should possess B-cell epitopes, and for effective CTL response, the peptide must be presented by MHC-I to the T cell receptor. Thus, the model antigens for developing subunit/multi-epitope vaccine against *Theileria* should either possess B-cell epitopes or be displayed by MHC-I to CTLs. UHere, using immuno-informatics, we have identified 12twelve proteins containing 19nineteen epitopes which are likely to elicit the humoral immune response. We have further designed and constructeda multi-epitope subunit vaccine by combining CTL epitopes which have the potential to elicit the cellular immune response against *Theileria* parasites (*T. annulata* and *T. parva*). The results provide a a new strategy for epitope--based vaccine development for *Theileria* parasites and indicate thathe immuno-informatics could beas an alternative strategy to rapidly discover potentialnew antigens for subunit/multi-epitopes vaccine development against parasites.

Materials and Methods

A systematic workflow (Figure 1A) was generated for the identification of potential vaccine candidates and construction of multi-epitope subunit vaccine against *Theileria* parasites (*T. annulata* and *T. parva*) from the available transcriptomic/proteomic data (Gardner et al., 2005; Pain et al., 2005; Witschi et al., 2013).

Sequence retrieval and sub-cellular localization

All protein sequences of *T. annulata* were retrieved from the National Centre for Biotechnology Information (NCBI) database. The CD-HIT server was used to remove proteins having 90% or

more similarity. All the parameters recommended by the server were set to default (Huang et al., 2010). The CELLO v.2.5 server was used for the identification of subcellular localization of all the selected proteins. The plasma membrane proteins were selected with a confidence score of >=2 (Yu et al., 2006). Further, all the proteins were analyzed for the presence of signal peptide sequence ~~using SignalP 4.1~~ and GPI anchor using SignalP 4.1 and PredGPI server respectively, with default settings (Pierleoni et al., 2008; Petersen et al., 2011).

Screening proteins non-homologous to the host and orthologous to *T. parva*

The selected plasma membrane proteins and proteins having both the signal sequence and GPI anchor were further analyzed for screening the proteins which are non-homologous to the host, *Bos taurus*, using BLASTp server (Altschul et al., 1990) with a cut-off E-value 0.005. Proteins predicted with E-value more than 0.005 were ~~expected~~ likely to be non-homologous to *Bos taurus* and ~~were~~ selected for further analysis. Further, an orthologous search between *T. annulata* and *T. parva* prediction ~~was carried out using OrthoMCL tool to identify conserved orthologous proteins (Chen et al., 2007)~~. The query strategy was designed in such a way that the average identity between two proteins ~~should~~ would be greater than 50%, and ~~is~~ would be specific only to *T. annulata* and *T. parva*.

Antigenicity of proteins

The selected proteins (non-homologous to *Bos taurus* and orthologous to *T. parva*) from the previous step were analyzed for their immunogenicity ~~properties~~ using Vaxijen - v2 (Doytchinova and Flower, 2007). The vaxijen is based on an alignment-free approach for antigen prediction using auto cross--covariance (ACC) transformation of protein sequences into uniform vectors of principal amino acid properties. The classification by V~~v~~axijen is solely based on the

physicochemical properties of proteins. The threshold was kept to 0.5 to predict probable antigenic and non-antigenic proteins.

B-cell epitope prediction

The prediction of B-cell epitopes ~~prediction~~ for all the selected antigenic proteins from the previous step was performed using BCPREDS server with 90% specificity (El-Manzalawy et al., 2008). BCPREDS server (B-cell epitope prediction) is based on Support Vector Machine (SVM) classifiers and the training data-set containing 701 linear B-cell epitopes from Bcipep database and ~~randomly~~ 701 random non-B-cell epitopes from SwissProt sequence database.

Further, the proteins were analyzed for the presence of putative transmembrane domains, exposed and intracellular region(s) using TMHMM server (http://www.cbs.dtu.dk/services/TMHMM/) (Krogh et al., 2001). Finally, those epitopes which were exposed to the extracellular surface of the cells~~,~~ were selected for further analysis.

Model building, refinement and stereochemistry analysis of B-cell protein

The tertiary structure of the selected proteins which possess B-cell epitopes was built using RaptorX which is a web portal for protein structure and function prediction (Kallberg et al., 2012). RaptorX uses a threading-based modelling approach for model building. RaptorX server predicts three-dimensional structures for protein sequences without close homologs in the Protein Data Bank (PDB). RaptorX server can predict secondary and tertiary structures, contacts solvent accessibility, disordered regions and binding sites. It can also assign some confidence scores to indicate the quality of a predicted 3D model such as p-value for the relative global quality, GDT (global distance test) and uGDT (un-normalized GDT) for the absolute global

quality, and modelling error at each residue. The selected epitopes were then mapped on the structure using the information obtained from the TMHMM prediction.

Prediction of T cell epitope(s) ~~prediction~~ and their specific MHC-I restriction element(s)

The prediction of cytotoxic T-cell (CTL) epitopes was done using NetCTL 1.2 server (Larsen et al., 2007). ~~Cytotoxic T-cell epitope (CTL) prediction was done using NetCTL 1.2 server (Larsen et al., 2005).~~ The threshold was set to 0.98 for high sensitivity and accuracy. All antigenic proteins were analyzed for the presence of CTL epitopes. The CTL epitopes identified from the previous step were further screened for bovine MHC–I restriction element identification using HLArestrictor 1.1 (Erup Larsen et al., 2011). All ~~seventy seven~~77 bovine MHC-1 alleles present in the database were analyzed for their ability to bind to identified CTL epitopes and their binding affinity ~~with~~ for respective MHC-I allele was also calculated. The ~~T~~threshold for strong binder (IC_{50}) was set to 50 and percentile rank was set to 0.5. The ~~T~~threshold for weak binder (IC_{50}) was set to 500 and percentile rank was set to 2.

Model building, refinement and stereochemistry analysis of bovine MHC-I

The sequence of bovine MHC-I alleles namely BoLA-3*05101, BoLA-1*00901, and BoLA-1*00902 were downloaded from the Immuno Polymorphism Database (IPD) (Robinson et al., 2013). Since crystal structures for these proteins were not present, a reliable model was built using homology modelling. The best template for the homology modelling was obtained through protein BLAST (pBLAST) against Protein Data Bank (PDB) and the three-dimensional model of BoLA-MHC protein were built using modeller 9.19 (Webb, 2014). The generated model was then refined by ModRefiner (Xu and Zhang, 2011) and stereochemistry analysis was done using PROCHECK, a protein parameters analysis tool ~~PROCHECK~~ (Laskowski, 1993).

Tertiary model prediction for CTL epitopes

PEP-FOLD server was used to predict the three-dimensional structure of allеach CTL epitopes (Shen et al., 2014). Five probable structures were predicted by this server. The structure having the lowest energy model was selected as the final model for the respective epitope.

Interaction studies between selected MHC-1 allele and CTL epitope

The strong bindes pairs of MHC-1 allele and CTL epitope obtained from HLA-restrictors were then used for molecular docking studies using AutoDock vina (Trott and Olson, 2010). The hydrogen atoms were added to the selected MHC-I alleles and the file was then converted to pdbqt format for docking studies with the CTL epitopes. For BoLA-3*05101, grid was set to -23.025, -15.521, -28.935 for X, Y and Z co-ordinates, respectively. The grid box for BoLA-1*00901 was set to 60.708, 61.642, 26.373 for X, Y and Z co-ordinates, respectively. In the case of BoLA-1*00902 the grid box for X, Y and Z waseре set to 58.68, 64.412, 33.956 co-ordinates, respectively. The binding affinity of the ligands with its receptor was measured in kcal/mol.

Building and characterization of the multi-epitope subunit vaccine

All CTL epitopes selected after MHC-I restriction element analysis and interaction studies were fused together with the help of a linker AAY to build a multi-epitope vaccine molecule. The multi-epitope vaccine molecule thus obtained was analyzed for its non-homology to bovine (as described earlier). The antigenic properties of the multi-epitopes vaccine candidate was estimated by Vaxijen server (as described earlier). The allergenicity of the vaccine candidate was evaluated using AlgPred (http://crdd.osdd.net/raghava/algpred/)(Saha and Raghava, 2006). The physiochemical properties such as amino acid composition, theoretical pI, instability index etc.,

of the vaccine candidates were calculated by ProtParam (http://web.expasy.org/protparam/) (Gasteiger, 2005).

Structure characterization of the multi-epitope vaccine candidate and bovine TLR-4

A template-based tertiary structure prediction of the multi-epitope vaccine candidate was carried out by RaptorX server (Kallberg et al., 2012). The Ffurther, the refinement of modelled multi-epitope vaccine candidate was done by using ModRefiner server as described previously (Xu and Zhang, 2011) and stereochemistry analysis was done using protein parameters analysis tool PROCHECK (Laskowski, 1993).

The sequence of bovine TLR-4 was downloaded from UniProt. Since crystal structure for this protein was not present, a reliable model was built using homology modelling. The best template for homology modelling was obtained through BLASTp against PDB database. Using a full length human TLR-4 as a template, Aa three-dimensional model of bovine TLR-4 protein was built on modeller 9.19 using a full-length human TLR4 as a templateusing modeller 9.19. The generated model was then refined by using ModRefiner and stereochemistry analysis was done using protein parameters analysis tool PROCHECK.

Molecular docking of multi-epitope vaccine candidate with bovine TLR-4 receptor

The interaction studies between bovine TLR-4 and final subunit vaccine protein were carried out with the help of a protein-protein docking server PatchDock (Schneidman-Duhovny et al., 2005). The results were refined according to their binding score by FireDock server (Mashiach et al., 2008).

Molecular dynamics simulation of bovine TLR-4 receptor with multi-epitope vaccine candidate

To determine the stability of a protein-ligand stability complex molecular dynamics (MD) simulation study of the complex is an important applicationtool. Molecular dynamics simulation was carried out for the docked complex of the TLR-4 and multi- epitope vaccine protein for 100ns with a time integral of 2fs using Maestreo software. The simulation was done with Single Point Charge (SPC) water as a solvent model and 0.15M of NaCl buffer. The boundary conditions wereas set to orthorhombic (a=b=c=10 Angstrom; alpha=beta=gamma=90^{o}). The isothermal-isobaric pressure (NPT) ensemble weare set at temperature 300K (temperature) and 1 atm pressure. Nose-Hoover Thermostat method was used to maintain the temperature constant. To examine the standard deviation and fluctuation of the backbone of the protein, the root mean square deviation (RMSD) and root mean square fluctuation (RMSF) analysis were performed.

Results

Identification of plasma membrane and GPI anchored proteins of *Theileria* parasite. *annulata*

A systematic workflow based on the immuno-informatics approach was designed to identify the potential vaccine candidates which would elicit an effective antibody response and to designvelop a multi-epitope based vaccine candidate which would generate effective cell-mediated immune response against *T.heileria* parasites *annulata* (Figure 1A). The first step in this workflow was to screen for *T. annulata* proteins which havinge the potential to be localized on the cell surface. For this tThe complete transcriptome/proteome of *T. annulata* comprising of total 3783 protein sequences waserе downloaded from NCBI database for this purpose (Figure 1B). From all the protein sequences, tThe paralogous sequences with more than 90% similarity

were identified using CD-HIT server from all the protein sequences. Twenty- -eight protein sequences were removed as paralogs while remaining 3755 protein sequences were selected for further analysis. The selected 3755 protein sequences were submitted to the CELLO sub-cellular localization servertool, to screen for the proteins predicted to be localized on the plasma membrane. CELLO is based on support vector machine (SVM) system to predict the probability distribution for possible localization of the protein. The output of this software is in the form of different numerical values as gives different confidence scores for predicting the probability of localization of the proteins in the cell. To get better stringency in screening of the plasma membrane proteins, the confidence score of 2 or better was fixed. A total of 711 proteins obtained with the confidence score greater than 2 were selected for further analysis (Figure 1B).

Furthermore,, the criteria chosen for identification of proteins predicted to be localized on the cell surface was that it should have a signal peptide along with a GPI anchor sequence. All the 3755 protein sequences were screened for the presence of a signal peptide sequence and a GPI-anchorr sequence. From SignalP analysis, out of Only 390 out of the 3755 proteins only 390 proteins were found to possess a signal peptide sequence based on SignalP 4.1 analysis. From PredGPI analysis, for GPI anchors signal, only 49 proteins containing a GPI-anchors signal sequence were obtained. Only 24 out of these 3755 proteins were predicted to contain a signal peptide sequence as well as a GPI-anchor signal sequence (Figure 1B).

Ten out of these 24 proteins were also present in the plasma membrane analysis. After adding the results from CELLO for proteins predicted to be localized at the plasma membrane protein, and proteins having both thea signal peptide sequence and athe GPI-anchor signal sequence, finally 726 protein sequences were finally selected for further analysis (Figure 1B).

Identification of antigenic *T. annulata* proteins non-homologous to *Bos taurus* and orthologous to *T. parva* ~~and antigenic proteins of *T. annulata*~~

All 726 protein sequences from previous analysis based on the cell surface localization were analyzed for their non-homology to the host (*Bos taurus*) and orthology to *T. parva*. BLASTp analysis against *Bos taurus* resulted in 443 proteins with E-value above 0.005. It is assumed that these proteins are not present in the host (Figure 1B).

The OrthoMCL server was used~~employed~~ for searching the orthologous or conserved protein between *T. annulata* and *T. parva*. ~~The criteria for orthologous protein selection was based on conservancy of the protein in both the species with an average identity of greater than 50%.~~ Out of the 443 selected proteins ~~which are~~ non-homologous to *Bos taurus*, only 32 proteins were found to be conserved and ~~only~~ present exclusively in *T. annulata* and *T. parva*. These 32 proteins were selected for further screening (Figure 1B).

All 32 protein sequences selected from the previous step were submitted to Vaxijen web server to predict the antigenic properties of each protein. Vaxijen web server is an efficient tool for predicting antigenic properties of proteins with an accuracy of 70% to 89%. ~~With the threshold of 0.5, o~~Out of the ~~thirty two~~32 protein sequences, only ~~twenty one~~21 protein sequences were predicted as probable potent antigen with the threshold of 0.5 (Table S1, S2). F~~Out of 21protein sequences, f~~ifteen out of 21 proteins sequences were found to be hypothetical.

Potential proteins of *Theileria* parasites~~. *annulata* proteins~~ for peptide/subunit vaccine development

B-cell epitopes are the antigenic determinant~~s~~ regions which are recognized by the specific antibodies. The predicted antigenic proteins were then analyzed for 20 amino acid long ~~the~~

linear B-cell epitopes ~~of twenty residue long~~. BCPREDS server was used to predict the probable B-cell epitope and ~~to minimize chances of false positive,~~ the specificity was kept at 90% or more to minimize chances of false positive. Four proteins, namely TA15480, TA21140, TA09756~~,~~ and TA13665, were predicted not to possess any B-cell epitope with the defined cut off value (Table S3). The B-cell epitope must be exposed to the extracellular regions of the cell to be recognized by the antibodies. ~~, h~~Hence the remaining 17 proteins with putative transmembrane domains were analyzed for the ~~transmembrane,~~ exposed and buried ~~inside~~ regions using TMHMM server~~.~~ ~~Hence, rest 17 proteins were analyzed by TMHMM server~~ (Table S4). The tertiary structures of all these proteins were built using RaptorX~~, a web portal for protein structure and function prediction~~ and the exposed epitopes were mapped on the structure (Figure 2, Figure S1). Five proteins, namely TA15965, TA05820, TA14950, TA15795 and TA12045, were found to have no exposed B-cell epitopes, while the remaining~~st~~ 12 proteins ~~were found to~~ contained B-cell epitopes ~~which were~~ exposed to the cell surface (Figure 2, Figure S2). ~~In these 12 proteins a~~A total of ~~nineteen~~ 19 B-cell epitopes were predicted in these 12 proteins (Table 1). Presence of merozoite-piroplasm surface antigen (TA17050) which is one of the leading vaccine candidates for the subunit vaccine against *T. annulata* (Gubbels et al., 2000) in our final list of 12 proteins~~,~~ validates ~~authenticate~~ the *in silico* analysis. TA17050 contains five B-cell epitopes with good peptide score (more than 0.97 for all peptides) and an antigenic score of 0.5708. The remaining 11 proteins also ~~All other eleven proteins~~ are potentially good targets for the development of subunit vaccine, especially TA09755 (FLPQTSRPSMGKKKGSFQLP-with peptide score 0.976) ~~showed~~exhibited the highest antigenic score (0.7931) ~~among all eleven proteins~~. Tpr-related protein family member (TA13385), which is unique to the *Theileria* parasite, could also be considered as a good target for developing a subunit vaccine. Eight other peptides from the list

with scores more than 0.99, could be evaluated~~further tested~~ as potential peptide-~~-~~based vaccine candidates (Table 1).

Identification of cytotoxic T cell epitopes specific to bovine MHC-I restriction element~~(~~s~~)~~

The predicted 21 antigenic proteins were analyzed for the ~~cytotoxic T-cell (~~CTL~~)~~ epitopes using NetCTL server. A total of 179 CTL epitopes were predicted with a threshold value of 0.98 (Table S5). The CTL epitopes work in conjugation with the MHC-I present on the surface of antigen presenting cells. Hence, using HLA-restrictors server, all the 179 CTL epitopes were analyzed for their ability to be recognized by MHC-I restriction elements of *Bos taurus* and ~~to predict~~ the binding affinity of each CTL epitope with bovine MHC-I allele was predicted. Only ~~eleven~~ 11 proteins containing ~~twenty seven~~27 epitopes were predicted to have a~~showed~~ strong binding affinity (IC_{50} value <50 and percentile rank ~~was also~~ below 0.5) with MHC-I molecules (Table S6). Twenty-two ~~O~~out of these 27~~27~~ predicated T-cell epitopes~~, twenty two epitopes~~ were predicted to have a~~showed~~ strong binding affinity with BoLA-3*05101, ~~fourteen~~ 14 out of these 27 epitopes were predicted to have a~~showed~~ strong binding affinity with BoLA-1*00902, and 9~~nine~~ out of these 27 epitopes were predicted to have a~~showed~~ strong binding affinity with BoLA-1*00901. Although many more MHC-I molecules were obtained which were predicted to have a~~showed~~ strong binding to selected CTLs, but only these 3~~three~~ MHC-I molecules (BoLA-3*05101, BoLA-1*00902, and BoLA-1*00901) were predicted to have a~~showed~~ strong binding affinity for multiple epitopes (Table S6).

Molecular modelling and structure refinement of bovine MHC-I allele

The tertiary structure of BoLA-3*05101 was predicted by Modeller 9.19 with bovine MHC-I molecule N*01801 (BoLA-A11) (3PWU) as the template. BoLA-3*05101 showed 84% identity

with N*01801. The root mean square deviation (RMSD) between the structure of modelled BoLA-3*05101 and structure obtained after refinement was found to be 0.489 Å. The predicted tertiary structure of BoLA-3*05101 contain~~s~~ed 47.64% coils, 21.05% helixes and 31.30% beta strands. The Ramachandran plot of this modelled protein obtained by PROCHECK ~~tool~~ suggest~~s~~ed that this was a highly reliable model (Figure S2).

Both BoLA-1*00902 and BoLA-1*00901 showed 76% and 86% identity, respectively with cattle MHC-I N*01301 (2XFX'A') presenting an 11mer peptide from *T. parva* chain A. The RMSD value of modelled BoLA-1*00902 and BoLA-1*00901 protein, before and after refinement, was calculated to be 2.682 Å and 1.713 Å respectively. The predicted structure of BoLA-1*00902 contained 48.74% coils, 23.4% helixes and 27.86% beta strands while the predicted structure of BoLA-1*00901 was found to have 45.54% coils, 24.02% helixes and 30.44% beta strands. The Ramachandran plot of modelled proteins obtained by PROCHECK ~~tool~~ showed that the modelled proteins ~~a~~were stable (Figure S2).

Modelled bovine MHC-I interacts strongly with identified CTL epitopes

~~For generating immune response i~~It is necessary for the CTL epitopes to interact with the respective MHC-I molecule with high binding affinity for generating specific immune response~~.~~. In order to find the affinity between bovine MHC-I allele and identified CTL ~~epitop~~epitope,~~e~~ molecular docking studies were performed using AutoDock Vina. All ~~twenty seven~~27 CTL epitopes modelled using PepFold server were docked with their respective MHC allele. The predicted CTL epitopes "FVAWFYKLY" and "FLYKRDLPY" ~~showed~~exhibited the highest binding affinity of -7.4 kcal/mol while "ILFTISLHY" exhibited~~showed~~ the lowest binding affinity of -5.5 kcal/mol with BoLA-N: 05101 MHC allele (Figure 3, Table 2). The predicted

CTL epitope~~s~~ “FLYKRDLPY” showed highest binding affinity of -6.7 kcal/mol and “IAFCIILYY” showed the lowest binding affinity of -5.1 kcal/mol with BoLA-N: 00901 MHC allele. The predicte~~d~~d CTL epitopes “IAFCIILYY” ~~showed~~displayed the highest binding affinity of -7.2 kcal/mol and “STIAMGLVY” ~~showed~~displayed the lowest binding affinity of -4.4 kcal/mol with BoLA-N: 00902 MHC allele.

Further, TA12045 contains the highest number of CTL epitopes which showed strong affinity to 18~~eighteen~~ bovine MHC-I. Another protein, TA16125, contains 4~~four~~ CTL epitopes which showed strong affinity to 12~~twelve~~ bovine MHC-I (Table S6). Thus these proteins namely TA15965, TA12045 and TA16125 can also be considered as good candidates for~~of~~ T cell-mediated response against *Theileria ~~. annulata~~* parasites.

Design~~Construction~~ of multi-epitope subunit vaccine and prediction of allergenicity, physiochemical and antigenic properties

Based on the predicted binding affinity score of CTL epitope and bovine MHC-I restriction element analysis, all the 27~~27~~ CTL epitopes were fused together with a linker “AAY” to form the multi-epitope subunit vaccine molecule. The final vaccine was composed of 321 amino acid residues. The allergenicity of the vaccine construct was predicted using AlgPred against BLAST search on allergen representative peptides (ARPs). The result showed that the multi-epitope~~s~~ vaccine candidate would not act as ~~is not~~ an allergen. The molecular weight and theoretical pI of this multi-epitope subunit vaccine candidate were calculated as 3.79 kDa and 6.75 respectively. Based on the theoretical ~~According to~~ pI calculation this~~, the multi-epitope~~ vaccine candidate was found to be slightly acidic in nature. Further, the antigenicity of this potential vaccine

candidate ~~construct~~ was predicted to be 0.7286 as analyzed by Vaxijen server, suggesting that ~~the multi-epitope vaccine candidate~~it would be~~is~~ antigenic in nature.

Homology modelling and model refinement of bovine TLR-4 and multi-epitope vaccine candidate

The three dimensional model of the multi-epitope vaccine candidate was generated using the RaptorX server as no template was found ~~against~~ in the PDB database. ~~Only one domain was predicted in the multi-epitope vaccine candidate by~~ T~~t~~he RaptorX server predicted only a single domain for this multi-epitope vaccine candidate. The best template for the multi-epitope vaccine candidate was found to be 4AV3'A' with its p-value $6.61e^{-05}$ (Figure 4A). The Ramachandran plot of this protein revealed that 89.1% amino acids were located in the core region, 8.3% amino acids were located in the allowed region, 1.6% amino acids were located in the generously allowed region and none was located in the disallowed region (Figure 4C). The predicted secondary structure of the multi-epitope vaccine candidate was found to possess only helic~~x~~es (Figure S3A).

As the crystal structure of bovine TLR-4 was not available, the tertiary structure of bovine TLR-4 was modelled using 4G8A (human TLR-4) as a template by Modeller 9.19 software (Figure 5A). Both the query coverage and identity w~~e~~are~~s~~ found to be 72%. The structure was then refined using ModRefiner and validated with PROCHECK analysis. The Ramachandran plot of the modelled bovine TLR-4 revealed that 79.6% amino acids were located in the core region, 17.4% amino acids were located in the allowed region, 1.8% amino acids were located in the generously allowed region and only 1.2% amino acids were located in the disallowed region.

Multi-epitope vaccine candidate ~~does not~~ possesses ~~bovine~~ CTL and B-cell epitopes specific to *Theileria* parasites but not to the bovine host

The analysis for the presence of CTL epitopes in the multi-epitope vaccine protein by the NetCTL server showed the presence of a total of ~~forty nine~~49 CTL epitopes. Out of the ~~forty nine~~49 epitopes, ~~twenty six~~26 epitopes were found to be the same as the previously predicted epitopes in our analysis (Table S7). Remaining~~st~~ ~~twenty three~~23 epitopes were the newly predicted CTL epitopes. All the newly predicted CTL epitopes were subjected to BLASTp against bovine for non-homology search and it was found that none of ~~the~~ ~~these peptide~~epitopes were present in the bovine. Further, the discontinuous B-cell epitopes were also predicted from the structure of the multi-epitope vaccine protein using EliPro server (Ponomarenko et al., 2008). A total of 9~~nine~~ conformational B-cell epitopes were predicted by the server (Figure 4B, Table S8, Figure S3B). Similar to the CTL epitopes, all the predicted B-cell epitopes were also analyzed for their ~~-~~non-homology to the bovine and it was found that none of these B-cell epitopes were present in the bovine proteome database. These results suggest that the~~e~~is potential multi-epitope vaccine protein is non-homologous to bovine and does not possess any predicted B-cell and CTL ~~predicted~~ epitopes from bovine. Hence, the putative multi-epitope vaccine candidate would not generate an autoimmunity or tolerance in bovine.

Molecular docking of the multi-epitope vaccine candidate with bovine TLR~~-~~4 receptor

To ~~identify~~ predict the binding energy of this multi-epitope vaccine candidate with the bovine TLR~~-~~4, docking of this multi-epitope vaccine candidate with TLR~~-~~4 was carried out using PatchDock server ~~which is~~ based on the shape complementarity principle (Figure 5B). PatchDock server generated 20~~twenty~~ complexes which were scored according to the protein

surface, geometry and electrostatic complementarity. ~~In order to refine and re-score the docking solutions of TLR-4 and multi-epitope vaccine candidate, t~~The top ten complexes were then submitted to the FireDock server in order to refine and re-score the docking solutions of TLR4 and multi-epitope vaccine candidate. The ~~F~~final docking model was chosen from the molecular docking studies based on the binding score. The best structure of the multi-epitope vaccine candidate and bovine TLR-4 showed a binding energy of -18.3 kcal/mol (Figure 5B). The attractive van der Waals force was calculated as -36.97 kcal/mol, the repulsive van der Waals force was 49.96 kcal/mol, and the atomic contact energy was -2.10.

Molecular dynamics simulation of TLR-4 receptor with multi-epitope vaccine

The evaluation of the stability of the bovine TLR-4 and multi-epitope vaccine candidate complex using molecular dynamics showed that the complex wa~~i~~s quite stable in SPC water type with 0.15M NaCl. The root mean square deviation (RMSD) of the backbone of the protein and the root mean square fluctuation (RMSF) for all the side chain amino acid residues were analyzed for a time period of 100ns. The interaction between the TLR-4 and the multi-epitope vaccine protein was found to be stable after 10ns (Figure 6A). The RMSF of side chain residues was found to be in between 1Å~~A~~^{0} to 3Å~~A~~^{0} with a little variation (Figure 6B). A total of ~~forty two~~42 hydrogen bonds were formed between the TLR-4 receptor and the multi-epitope vaccine protein. The simulation analysis clearly shows that the multi-epitope vaccine protein can be recognized by the TLR-4 and ~~with~~ form a stable complex.

Discussion:

Various ~~One of the few~~ factors contributing to the ~~of~~ failure of ~~the~~ any antigen as a vaccine candidate under development stages are ~~due to:~~ (a) antigen not being exposed to immune

response ~~and,~~ (b) poor antigenic/immunogenicity response of antigen and, allergenicity to the host. Recently, with the advancement of *in-silico* analysis it has been possible to identify molecule(s) which would likely possess maximum qualities~~y~~/properties to qualify as~~of~~ a good antigen. The primary aim of this study was to identify the candidate gene(s) ~~which~~ having~~e~~ the potential to induce an immune response against macroschizont stage of *T. annulata*. ~~In this study w~~We utilized a systematic immuno-informatics in this study approach to predict potential vaccine candidate(s) in *Theileria~~. annulata~~* parasites using the proteome databases of *T. annulata* and *T. parva*, to design~~velop~~ a multi-epitope based vaccine.

One ~~of the~~ important criteria for selecting a potential antigen for vaccine development is that it should be located on the surface of the cell so that it is more accessible for both the humoral and the cellular immune system. The proteins are translocated to the cell surface using a signal sequence for plasma membrane translocation or could be anchored to the Glycosylphosphatidylinositol (GPI) moieties. These proteins which are exposed to the extracellular surface are easily recognized and are likely~~able~~ to elicit an immune response when used as ~~the~~ a target antigen for ~~a~~ vaccine development. P~~Therefore, the putative p~~roteins from the *Theileria* proteome database were thus examined for their translocation to plasma membrane using CELLO software. CELLO is based on support vector machine (SVM) system to predict the probability distribution for possible localization of the protein w~~in. Prediction of cellular localization by CELLO is much~~hich is more reliable than the other sub-localization prediction tool, ~~such as~~ PSORTb ~~B~~. Furthermore, PSORTb program is trained only for the bacterial proteome (Gardy et al., 2005). We selected a confidence score of 2 or more for the sub localization prediction based on our observation that in many cases the values lesser than 2 resulted in predicting two different sub-localizations for the same protein. ~~Proteins after~~

~~removing duplicates, were predicted for subcellular localization.~~ A total of 711 proteins were predicted to be localized at the cell surface using the CELLO prediction server. To take into account the proteins ~~Further, the proteins~~that may also be present on~~at~~ the cell surface using GPI moieties. ~~. W~~we first looked for the proteins ~~which hh~~avinge signal ~~sequence~~ peptide sequences using the SignalP 4.1 server. Few research groups have also analysed the proteome of *T. annulata* and *T. parva* for the identification of proteins containing signal peptide sequence and GPI anchor using the older version of SignalP (Pain et al., 2005; Weir et al., 2009; Woods et al., 2013; Bilgic et al., 2016). However, number of false positives in SignalP 4.1 is lesser compared to the older versions (www.cbs.dtu.dk/services/SignalIP/performance.php). SignalP 4.1 uses a Hidden Markov Model (HMM) method for predicting the presence of signal peptide in the protein. However, it does not predict the GPI anchors unlike the previous versions. PredGPI server was used to identify the GPI anchored proteins in our study. ~~In order to identify those proteins which may be GPI-anchored, PredGPI server was used.~~ The PredGPI uses a prediction method which is coupled to HMM and SVM methods. PredGPI has been shown to outperform the other methods used for ~~to~~ predicting GPI-anchors as it has a lower rate of false positive predictions with respect to the other available methods (Pierleoni et al., 2008). Only 22~~twenty two~~ proteins were predicted to possess both the signal peptide and GPI anchor. The putative plasma membrane and the GPI anchored proteins obtained from both the analysis were added for further analysis.

In order to generate an immune response, the antigen must be recognized as non-self. Hence, the antigen for which homologous protein is present in the host is not considered a good candidate for vaccine development. Thus, we removed all the proteins homologous to *Bos taurus* in our study. O~~ut of 733, o~~nly 443 out of the 733 proteins were predicted ~~to be~~as non-homologous to

Bos tauras. Further, an ideal vaccine candidate should be capable of stimulating an effective immune response against various species of *Theileria*. Thus, to increase the stringency, the query was set in such a way that only *T. annulata* or *T. parva* proteins ~~ar~~would be~~e~~ screened.

Antigenicity is one of the important features of an antigen for vaccine development. Vaxijen is the only server which uses alignment-independent prediction of protective antigens. It classifies antigen based on the physiochemical properties of the proteins. The prediction efficiency of this method~~s~~ is found to be 70-80%. We finally found 21 ~~Only twenty one~~ proteins which were predicted as antigenic and also ~~were found to be antigenic which~~ qualified other screening conditions.

An effective vaccine should induce a specific immune response against specific pathogens/antigen by selectively stimulating antigen-specific B-cells or CTLs or T helper cells. In order to initiate the humoral response, antigens must be exposed and should be recognized by the naive B-cells. Further, to initiate the cellular response, naive T cells need to be primed by the peptide antigens bound to the major histocompatibility complex~~proteins~~ (MHC) molecules on the antigen presenting cells (APCs). The activation of CTLs (CD8+ cells) is a result of the specific engagement of a 9 mer peptide to the MHC-I proteins. The CTLs eliminate the cells which they recognize as non-self, hence the immune system must be able to discriminate between healthy and infected cells.

Thus, the~~o~~ ~~identify~~ B-cell epitopes were predicted using the ~~,~~ BCPREDS server ~~was used~~. Further, B-cell epitopes must be accessible to the immune response. H~~, h~~ence all the proteins containing B-cell epitopes were modelled and the epitopes ~~were~~ mapped on the structure. Four proteins did not show any B-cell epitopes and epitopes on ~~five~~ 5 other proteins were not

accessible to immune recognition. Finally, 11~~eleven~~ proteins were ~~found~~ predicted to be good targets for the development of subunit vaccine. One of the 11~~eleven~~ proteins was TA17050 which is one of the leading vaccine candidates, thus suggesting the likeliness of the other 10~~ten~~ proteins ~~could also~~as ~~be~~ good candidates for subunit vaccine development.

The CTL epitopes in all the antigenic proteins were identified using NetCTL server. NetCTL is a quite reliable tool for the prediction of cytotoxic T lymphocytes and the~~its~~ predictions done by this ~~is~~are based on proteasomal cleavage, Transporter associated with antigen presentation (TAP) transport efficiency and MHC class I affinity. In comparison to the earlier version NetCTL1.0, the accuracy of prediction of cytotoxic T lymphocytes using NetCTL server is significantly improved (Larsen et al., 2007). Although NetCTL is a tool designed for the prediction of human CTL epitopes in a given protein, it can be used for bovine as well with equal efficiency. A total of 179~~one hundred and seventy nine~~ CTL epitopes were identified using this prediction method. Previously, studies conducted to identify CTL epitopes from *T. parva* specific CD8+ T cell line by transfecting cells with parasite cDNA identified ~~ten~~10 *T. parva* antigens having a total of ~~fifteen~~ 15 CTL epitopes (Graham et al., 2006). It was initially intriguing to observe that ~~N~~none of the antigens identified as CTL epitopes by this group were present in our analysis. However, a~~A~~fter careful analysis, we found that the 10 proteins that this study identified did not possess either a ~~none of the proteins from which these CTL epitopes were identified, did not either possess plasma membrane or~~ signal sequence and GPI anchor together, ~~sequence~~or they were not predicted to be localized to the plasma membrane. Using a similar strategy, three proteins (TA15705, TA17545 and TA14970) were shown to be recognized by CD8 T cells from *T. annulata* infected cattle (MacHugh et al., 2011). TA14970 neither contains a signal sequence nor a GPI anchor. Although, TA15705 (Ta9) and TA17545 possess the signal sequence, but these

proteins lack a GPI anchor. Hence, we did not include these proteins in our further analysis. The CTL epitopes must be presented by MHC-I on the surface of the antigen presenting cells to the T-cell receptor for the activation of cell--mediated response. Hence, we predicted ~~calculated~~ the binding energy of these CTL epitopes with all the bovine MHC-I molecules using HLA restrictor server. The binding pairs which showed the highest energy were selected for further analysis. Three MHC-I molecules, namely BoLA-3*05101, BoLA-1*00902 and BoLA-1*00901, were found to be the best, both in terms of binding energy and recognition to multiple CTL epitopes. Hence these three molecules were further modelled to predict their three--dimensional structures. Protein models having more than 90% of the residues in the core and allowed regions are generally considered as high--quality models. W~~Hence, w~~e found that the predicted model for BoLA-3*05101, BoLA-1*00902, and BoLA-1*00901 were high--quality models based upon the presence of more than 90% of their amino acid residues in the core and allowed regions. The docking of selected CTL epitopes with modelled BoLA-3*05101, BoLA-1*00902 and BoLA-1*00901 further confirmed the stability of the complex. Thus, ~~twenty seven~~27 CTL epitopes were found to be potential peptides for generating effective CTL response. Immunization with all these twenty--seven, 9 mer peptides for the generation of CTL response would be a tedious process. ~~, h~~Hence we designed~~engineered~~ a new protein containing all the CTL epitopes joined with AAY linker. Immunization with a single protein (approximately 3.79 kDa) would be much easier and cost effective if it could generate an effective cellular response. Thus, we re-analyzed our newly designed ~~again analyzed~~ multi-epitope vaccine candidate for its ability to serve as an ideal vaccine candidate. The newly designed ~~engineered~~ molecule was predicted~~found~~ to be non-allergic while it ~~but~~ was antigenic in nature. The multi-epitope vaccine candidate was found to

possess 49~~forty nine~~ CTL epitopes and none of them were~~as~~ found to be present in the bovine, and ~~twenty six~~ 26 were specifically from *Theileria~~. annulata~~* parasites.~~.~~

Since the multi-epitope vaccine candidate is a novel protein, hence no template was found for homologous modelling. Accordingly~~Thus~~, RaptorX~~APTORX~~ was used for generating a model for this protein. The possibility of the generated model being worse than the best of a set of randomly generated models can be calculated by p-value. A~~The~~ lesser p~~--~~value represents the higher ~~-~~quality model. For the alpha helix protein p-value should be lesser than 10^{-3}~~-~~, and for the beta~~--~~sheet protein p-value should be lesser than 10^{-4}. The predicted structure of the multi-epitope subunit vaccine contained all alpha helic~~x~~es with a p-value of $6.61e^{-05}$ which indicate that the structure predicted by RaptorX is a high~~--~~quality model. The un-normalized Global Distance Test (uGDT) also estimate the model quality, like, if a protein contains greater than 100~~hundred~~ residues and if the uGDT score is more than 50, it is a good indicator for high quality model. The final multi-epitope~~s subunit~~ vaccine candidate had the ~~contained~~ uGDT score of 73. Forty~~--~~four percent of the protein residues we~~a~~re exposed in the subunit vaccine protein, while ~~and~~ it is generally assumed ~~was reported t~~that if 15% of the protein residues are exposed and solvent accessible, they can generate effective humoral and cellular immune response.

The modelled tertiary structure of this multi-epitope vaccine candidate suggests that the protein is quite stable and ordered. Furthermore, none of the discontinuous B-cell epitopes of the multi-epitope vaccine protein were from bovine, suggesting that the new molecule would not generate any self-immune response to bovine.

The effective CTL response ~~would be generated~~ against the multi-epitope vaccine candidate would be generated only when it is presented to the CD8+ T cells by the antigen presenting cells.

The Toll-like receptors (TLRs) which are evolutionarily conserved proteins can sense foreign molecules. They are characterized by an extracellular leucine-rich repeat domain and an intracellular Toll/IL-1 receptor-like (TIR) domain. The TLRs play a fundamental role in the initiation of the immune response to the infectious agents through their recognition of the conserved microbial molecular pattern. Out of all the TLRs, only TLR4 and TLR11 have been shown to recognize the proteins (review (Takeuchi and Akira, 2010)). TLR-11 is absent in both human and bovine, thus in our *in silico* analysis we tested the binding ability of multi-epitope vaccine candidate with TLR-4 only. for which We performed homologous modelling offor bovine TLR-4 using human TLR-4 as a template. The binding energy of the designed multi-epitope vaccine candidate and bovine TLR-4 was predicted found to be -18.3 kcal/mol and the binding sites in the protein wereas distributed to the various regions of the proteins, suggesting that the bovine TLR-4 would effectively recognize the vaccine candidate designed herein. Thus overall in this study, we designedengineered a new vaccine candidate *in silico* containing multiple epitopes of *Theileia* parasites. *annulata.*

Conclusions:

TIn conclusion, this study shows that the immuno-informatics driven genome-wide screening of vaccine targets for *Theileria* is a highly promising strategy to accelerate vaccine development for this parasite. Based on this strategy, B-cell epitopes and CTLs were mapped from the proteome of the *Theileria* parasite. This study provides a comprehensive analysis of *Theileria* proteins predicted to be better protective immunogens with high conservancy and, which would have the potential for eliciting both neutralizing antibodies and T-cell responses. Thus, this study opens new avenues for accelerating vaccine development by providing various potential molecules as novel vaccine candidates and a multi-epitope based vaccine candidate.

617 Reference:

618 Altschul, S.F., Gish, W., Miller, W., Myers, E.W., and Lipman, D.J. (1990). Basic local alignment search tool.
619 *J Mol Biol* 215(3)**,** 403-410. doi: 10.1016/S0022-2836(05)80360-2.
620 Backert, L., and Kohlbacher, O. (2015). Immunoinformatics and epitope prediction in the age of genomic
621 medicine. *Genome Med* 7**,** 119. doi: 10.1186/s13073-015-0245-0.
622 Bilgic, H.B., Karagenc, T., Bakirci, S., Shiels, B., Tait, A., Kinnaird, J., et al. (2016). Identification and Analysis
623 of Immunodominant Antigens for ELISA-Based Detection of Theileria annulata. *PLoS One* 11(6)**,**
624 e0156645. doi: 10.1371/journal.pone.0156645.
625 Bishop, R., Musoke, A., Morzaria, S., Gardner, M., and Nene, V. (2004). Theileria: intracellular protozoan
626 parasites of wild and domestic ruminants transmitted by ixodid ticks. *Parasitology* 129 Suppl**,**
627 S271-283.
628 Boulter, N., Knight, P.A., Hunt, P.D., Hennessey, E.S., Katzer, F., Tait, A., et al. (1994). Theileria annulata
629 sporozoite surface antigen (SPAG-1) contains neutralizing determinants in the C terminus.
630 *Parasite Immunol* 16(2)**,** 97-104.
631 Boulter, N.R., Brown, C.G., Kirvar, E., Glass, E., Campbell, J., Morzaria, S., et al. (1998). Different vaccine
632 strategies used to protect against Theileria annulata. *Ann N Y Acad Sci* 849**,** 234-246.
633 Davies, M.N., Guan, P., Blythe, M.J., Salomon, J., Toseland, C.P., Hattotuwagama, C., et al. (2007). Using
634 databases and data mining in vaccinology. *Expert Opin Drug Discov* 2(1)**,** 19-35. doi:
635 10.1517/17460441.2.1.19.
636 Dobbelaere, D.A., Spooner, P.R., Barry, W.C., and Irvin, A.D. (1984). Monoclonal antibody neutralizes the
637 sporozoite stage of different Theileria parva stocks. *Parasite Immunol* 6(4)**,** 361-370.
638 Doytchinova, I.A., and Flower, D.R. (2007). VaxiJen: a server for prediction of protective antigens, tumour
639 antigens and subunit vaccines. *BMC Bioinformatics* 8**,** 4. doi: 10.1186/1471-2105-8-4.
640 e Silva, R.d.F., Ferreira, L.F.G.R., Hernandes, M.Z., de Brito, M.E.F., de Oliveira, B.C., da Silva, A.A., et al.
641 (2016). Combination of In Silico Methods in the Search for Potential CD4(+) and CD8(+) T Cell
642 Epitopes in the Proteome of Leishmania braziliensis. *Front Immunol* 7**,** 327. doi:
643 10.3389/fimmu.2016.00327.
644 El-Manzalawy, Y., Dobbs, D., and Honavar, V. (2008). Predicting linear B-cell epitopes using string kernels.
645 *J Mol Recognit* 21(4)**,** 243-255. doi: 10.1002/jmr.893.
646 Erup Larsen, M., Kloverpris, H., Stryhn, A., Koofhethile, C.K., Sims, S., Ndung'u, T., et al. (2011).
647 HLArestrictor--a tool for patient-specific predictions of HLA restriction elements and optimal
648 epitopes within peptides. *Immunogenetics* 63(1)**,** 43-55. doi: 10.1007/s00251-010-0493-5.
649 Gardner, M.J., Bishop, R., Shah, T., de Villiers, E.P., Carlton, J.M., Hall, N., et al. (2005). Genome sequence
650 of Theileria parva, a bovine pathogen that transforms lymphocytes. *Science* 309(5731)**,** 134-137.
651 doi: 10.1126/science.1110439.
652 Gardy, J.L., Laird, M.R., Chen, F., Rey, S., Walsh, C.J., Ester, M., et al. (2005). PSORTb v.2.0: expanded
653 prediction of bacterial protein subcellular localization and insights gained from comparative
654 proteome analysis. *Bioinformatics* 21(5)**,** 617-623. doi: 10.1093/bioinformatics/bti057.
655 Gasteiger, E.H., C.; Gattiker, A.; Duvaud, S.; Wilkins, MR.; Appel, RD.; Bairoch, A.; (2005). *Protein*
656 *Identification and Analysis Tools on the ExPASy Server.* Humana Press.
657 Gharbi, M., Sassi, L., Dorchies, P., and Darghouth, M.A. (2006). Infection of calves with Theileria annulata
658 in Tunisia: Economic analysis and evaluation of the potential benefit of vaccination. *Vet Parasitol*
659 137(3-4)**,** 231-241. doi: 10.1016/j.vetpar.2006.01.015.
660 Graham, S.P., Pelle, R., Honda, Y., Mwangi, D.M., Tonukari, N.J., Yamage, M., et al. (2006). Theileria parva
661 candidate vaccine antigens recognized by immune bovine cytotoxic T lymphocytes. *Proc Natl Acad*
662 *Sci U S A* 103(9)**,** 3286-3291. doi: 10.1073/pnas.0511273103.

663 Gubbels, M.J., Katzer, F., Hide, G., Jongejan, F., and Shiels, B.R. (2000). Generation of a mosaic pattern of
664 diversity in the major merozoite-piroplasm surface antigen of Theileria annulata. *Mol Biochem*
665 *Parasitol* 110(1)**,** 23-32.
666 He, L., Cheng, Y., Kong, L., Azadnia, P., Giang, E., Kim, J., et al. (2015). Approaching rational epitope vaccine
667 design for hepatitis C virus with meta-server and multivalent scaffolding. *Sci Rep* 5**,** 12501. doi:
668 10.1038/srep12501.
669 Huang, Y., Niu, B., Gao, Y., Fu, L., and Li, W. (2010). CD-HIT Suite: a web server for clustering and comparing
670 biological sequences. *Bioinformatics* 26(5)**,** 680-682. doi: 10.1093/bioinformatics/btq003.
671 Irvin, A.D.a.M., W.I (1987). *Immunopathalogy, immunology and immunoprophylylaxis of Theileria*
672 *infections.* Florida CRC Press, Boca Raton
673 Kallberg, M., Wang, H., Wang, S., Peng, J., Wang, Z., Lu, H., et al. (2012). Template-based protein structure
674 modeling using the RaptorX web server. *Nat Protoc* 7(8)**,** 1511-1522. doi:
675 10.1038/nprot.2012.085.
676 Kallerup, R.S.a.F., C. (2015). *Classification of Vaccines* Springer.
677 Krogh, A., Larsson, B., von Heijne, G., and Sonnhammer, E.L. (2001). Predicting transmembrane protein
678 topology with a hidden Markov model: application to complete genomes. *J Mol Biol* 305(3)**,** 567-
679 580. doi: 10.1006/jmbi.2000.4315.
680 Larsen, M.V., Lundegaard, C., Lamberth, K., Buus, S., Lund, O., and Nielsen, M. (2007). Large-scale
681 validation of methods for cytotoxic T-lymphocyte epitope prediction. *BMC Bioinformatics* 8**,** 424.
682 doi: 10.1186/1471-2105-8-424.
683 Laskowski, R.M., MW.; Moss, DS.; Thornton, JM. (1993). PROCHECK - a program to check the
684 stereochemical quality of protein structures. *J. App. Cryst.* 26**,** 283-291.
685 MacHugh, N.D., Connelley, T., Graham, S.P., Pelle, R., Formisano, P., Taracha, E.L., et al. (2009). CD8+ T-
686 cell responses to Theileria parva are preferentially directed to a single dominant antigen:
687 Implications for parasite strain-specific immunity. *Eur J Immunol* 39(9)**,** 2459-2469. doi:
688 10.1002/eji.200939227.
689 MacHugh, N.D., Weir, W., Burrells, A., Lizundia, R., Graham, S.P., Taracha, E.L., et al. (2011). Extensive
690 polymorphism and evidence of immune selection in a highly dominant antigen recognized by
691 bovine CD8 T cells specific for Theileria annulata. *Infect Immun* 79(5)**,** 2059-2069. doi:
692 10.1128/IAI.01285-10.
693 Maritim, A.C., Young, A.S., Lesan, A.C., Ndungu, S.G., Mutugi, J.J., and Stagg, D.A. (1989). Theilerial
694 parasites isolated from carrier cattle after immunization with Theileria parva by the infection and
695 treatment method. *Parasitology* 99 Pt 1**,** 139-147.
696 Mashiach, E., Schneidman-Duhovny, D., Andrusier, N., Nussinov, R., and Wolfson, H.J. (2008). FireDock: a
697 web server for fast interaction refinement in molecular docking. *Nucleic Acids Res* 36(Web Server
698 issue)**,** W229-232. doi: 10.1093/nar/gkn186.
699 McHardy, N., Wekesa, L.S., Hudson, A.T., and Randall, A.W. (1985). Antitheilerial activity of BW720C
700 (buparvaquone): a comparison with parvaquone. *Res Vet Sci* 39(1)**,** 29-33.
701 Mhadhbi, M., Naouach, A., Boumiza, A., Chaabani, M.F., BenAbderazzak, S., and Darghouth, M.A. (2010).
702 In vivo evidence for the resistance of Theileria annulata to buparvaquone. *Vet Parasitol* 169(3-4)**,**
703 241-247. doi: 10.1016/j.vetpar.2010.01.013.
704 Morrison, W.I. (2009). Progress towards understanding the immunobiology of Theileria parasites.
705 *Parasitology* 136(12)**,** 1415-1426. doi: 10.1017/S0031182009990916.
706 Musoke, A., Rowlands, J., Nene, V., Nyanjui, J., Katende, J., Spooner, P., et al. (2005). Subunit vaccine
707 based on the p67 major surface protein of Theileria parva sporozoites reduces severity of infection
708 derived from field tick challenge. *Vaccine* 23(23)**,** 3084-3095. doi: 10.1016/j.vaccine.2004.09.039.
709 Musoke, A.J., Nantulya, V.M., Buscher, G., Masake, R.A., and Otim, B. (1982). Bovine immune response to
710 Theileria parva: neutralizing antibodies to sporozoites. *Immunology* 45(4)**,** 663-668.

Nene, V., and Morrison, W.I. (2016). Approaches to vaccination against Theileria parva and Theileria annulata. *Parasite Immunol* 38(12), 724-734. doi: 10.1111/pim.12388.

Norval, R.A., Lawrence, J.A., Young, A.S., Perry, B.D., Dolan, T.T., and Scott, J. (1991). Theileria parva: influence of vector, parasite and host relationships on the epidemiology of theileriosis in southern Africa. *Parasitology* 102 Pt 3, 347-356.

Pahil, S., Taneja, N., Ansari, H.R., and Raghava, G.P.S. (2017). In silico analysis to identify vaccine candidates common to multiple serotypes of Shigella and evaluation of their immunogenicity. *PLoS One* 12(8), e0180505. doi: 10.1371/journal.pone.0180505.

Pain, A., Renauld, H., Berriman, M., Murphy, L., Yeats, C.A., Weir, W., et al. (2005). Genome of the host-cell transforming parasite Theileria annulata compared with T. parva. *Science* 309(5731), 131-133. doi: 10.1126/science.1110418.

Petersen, T.N., Brunak, S., von Heijne, G., and Nielsen, H. (2011). SignalP 4.0: discriminating signal peptides from transmembrane regions. *Nat Methods* 8(10), 785-786. doi: 10.1038/nmeth.1701.

Pierleoni, A., Martelli, P.L., and Casadio, R. (2008). PredGPI: a GPI-anchor predictor. *BMC Bioinformatics* 9, 392. doi: 10.1186/1471-2105-9-392.

Pipano, E., Tsur, I. (1966). Experimental Immunization against Theileria annulata with a tissue culture vaccine I. Laboratory trails *Refu Vet.* 23, 186-194.

Ponomarenko, J., Bui, H.H., Li, W., Fusseder, N., Bourne, P.E., Sette, A., et al. (2008). ElliPro: a new structure-based tool for the prediction of antibody epitopes. *BMC Bioinformatics* 9, 514. doi: 10.1186/1471-2105-9-514.

Robinson, J., Halliwell, J.A., McWilliam, H., Lopez, R., and Marsh, S.G. (2013). IPD--the Immuno Polymorphism Database. *Nucleic Acids Res* 41(Database issue), D1234-1240. doi: 10.1093/nar/gks1140.

Saha, S., and Raghava, G.P. (2006). AlgPred: prediction of allergenic proteins and mapping of IgE epitopes. *Nucleic Acids Res* 34(Web Server issue), W202-209. doi: 10.1093/nar/gkl343.

Schneidman-Duhovny, D., Inbar, Y., Nussinov, R., and Wolfson, H.J. (2005). PatchDock and SymmDock: servers for rigid and symmetric docking. *Nucleic Acids Res* 33(Web Server issue), W363-367. doi: 10.1093/nar/gki481.

Shen, Y., Maupetit, J., Derreumaux, P., and Tuffery, P. (2014). Improved PEP-FOLD Approach for Peptide and Miniprotein Structure Prediction. *J Chem Theory Comput* 10(10), 4745-4758. doi: 10.1021/ct500592m.

Sivakumar, T., Hayashida, K., Sugimoto, C., and Yokoyama, N. (2014). Evolution and genetic diversity of Theileria. *Infect Genet Evol* 27, 250-263. doi: 10.1016/j.meegid.2014.07.013.

Sominskaya, I., Skrastina, D., Dislers, A., Vasiljev, D., Mihailova, M., Ose, V., et al. (2010). Construction and immunological evaluation of multivalent hepatitis B virus (HBV) core virus-like particles carrying HBV and HCV epitopes. *Clin Vaccine Immunol* 17(6), 1027-1033. doi: 10.1128/CVI.00468-09.

Stanekova, Z., and Vareckova, E. (2010). Conserved epitopes of influenza A virus inducing protective immunity and their prospects for universal vaccine development. *Virol J* 7, 351. doi: 10.1186/1743-422X-7-351.

Takeuchi, O., and Akira, S. (2010). Pattern recognition receptors and inflammation. *Cell* 140(6), 805-820. doi: 10.1016/j.cell.2010.01.022.

Taracha, E.L., Goddeeris, B.M., Morzaria, S.P., and Morrison, W.I. (1995a). Parasite strain specificity of precursor cytotoxic T cells in individual animals correlates with cross-protection in cattle challenged with Theileria parva. *Infect Immun* 63(4), 1258-1262.

Taracha, E.L., Goddeeris, B.M., Teale, A.J., Kemp, S.J., and Morrison, W.I. (1995b). Parasite strain specificity of bovine cytotoxic T cell responses to Theileria parva is determined primarily by immunodominance. *J Immunol* 155(10), 4854-4860.

Tomar, N., and De, R.K. (2014). Immunoinformatics: a brief review. *Methods Mol Biol* 1184**,** 23-55. doi: 10.1007/978-1-4939-1115-8_3.
Trott, O., and Olson, A.J. (2010). AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. *J Comput Chem* 31(2)**,** 455-461. doi: 10.1002/jcc.21334.
Webb, B., Sali, A. (2014). *Comparative Protein Structure Modeling Using Modeller.* John Wiley & Sons, Inc.
Weir, W., Sunter, J., Chaussepied, M., Skilton, R., Tait, A., de Villiers, E.P., et al. (2009). Highly syntenic and yet divergent: a tale of two Theilerias. *Infect Genet Evol* 9(4)**,** 453-461. doi: 10.1016/j.meegid.2009.01.002.
Williamson, S., Tait, A., Brown, D., Walker, A., Beck, P., Shiels, B., et al. (1989). Theileria annulata sporozoite surface antigen expressed in Escherichia coli elicits neutralizing antibody. *Proc Natl Acad Sci U S A* 86(12)**,** 4639-4643.
Witschi, M., Xia, D., Sanderson, S., Baumgartner, M., Wastling, J.M., and Dobbelaere, D.A. (2013). Proteomic analysis of the Theileria annulata schizont. *Int J Parasitol* 43(2)**,** 173-180. doi: 10.1016/j.ijpara.2012.10.017.
Woods, K.L., Theiler, R., Muhlemann, M., Segiser, A., Huber, S., Ansari, H.R., et al. (2013). Recruitment of EB1, a master regulator of microtubule dynamics, to the surface of the Theileria annulata schizont. *PLoS Pathog* 9(5)**,** e1003346. doi: 10.1371/journal.ppat.1003346.
Xu, D., and Zhang, Y. (2011). Improving the physical realism and structural accuracy of protein models by a two-step atomic-level energy minimization. *Biophys J* 101(10)**,** 2525-2534. doi: 10.1016/j.bpj.2011.10.024.
Yu, C.S., Chen, Y.C., Lu, C.H., and Hwang, J.K. (2006). Prediction of protein subcellular localization. *Proteins* 64(3)**,** 643-651. doi: 10.1002/prot.21018.

Figure ligands:

Figure 1. (A) Systematic workflow for identification of potential vaccine candidates and construction of multi-epitope subunit vaccine against *Theileria annulata* from the available proteomic data. n, represent total number of proteins used for analysis.

[1]https://www.ncbi.nlm.nih.gov/;

[2]http://weizhong-lab.ucsd.edu/cdhit-web-server/cgi-bin/index.cgi?cmd=cd-hit;

[3]http://cello.life.nctu.edu.tw/;

[4]http://gpcr.biocomp.unibo.it/predgpi/pred.htm;

[5]http://www.cbs.dtu.dk/services/SignalP/;

[6] https://blast.ncbi.nlm.nih.gov/Blast.cgi?PAGE=Proteins

[7] http://orthomcl.org/orthomcl/;

[8]http://www.ddg-pharmfac.net/vaxijen/VaxiJen/VaxiJen.html;

[9]http://ailab.ist.psu.edu/bcpred/;

[10] http://raptorx.uchicago.edu/

[11]http://www.cbs.dtu.dk/services/NetCTL/;

[12]http://www.cbs.dtu.dk/services/HLArestrictor/;

[13]http://mobyle.rpbs.univ-paris-diderot.fr/cgi-bin/portal.py#forms::PEP-FOLD3;

[14]Modeller9.19;

[15]AutodockVina;

[16]https://bioinfo3d.cs.tau.ac.il/PatchDock/;

(B) Number of proteins obtained after various analysis. a. total proteins of *T. annulata*, b. proteins after removal of duplicates, c. plasma membrane proteins, d. proteins containing GPI anchor & signal peptide, e. proteins which are non-homologous to bovine, f. orthologous proteins specifically present in *T. annulata* and *T. parva*, g. antigenic proteins

Figure 2. Molecular modelling of antigenic proteins of *Theileria annulata* containing B-cell epitope using RAPTOR X. (A) Ta 09755, (B) Ta 13065, (C) Ta 13810, (D) Ta 13385, and (E) Ta 16735. Identified epitopes mapped onto the modelled proteins are shown in red color, for buried epitopes, and blue color, for exposed epitopes using TMHMM.

Figure 3. Representative model of interaction between predicted CTL epitope through molecular docking studies (A) FLYKRDLPY with BoLA-N:05101, binding energy -7.8 kcal/mol, (B) FLYKRDLPY with BoLA-N:00902, binding energy = -7.7 kcal/mol, (C) WMVFFIVVY with BoLA-N:00901, binding energy = -7.6 kcal/mol. Red represent epitope, blue represent modelled bovine MHC-I and red represent CTL epitope.

Figure 4. Modelled structure of multi-epitope based subunit vaccine candidate: (A) Three dimensional model of multi-epitope based subunit vaccine candidate obtained by homologous modelling and refinement. The red color represents CTL epitopes and gray color represents AAY linker, (B) B-cell epitopes of multi-epitope based subunit vaccine candidate represented in green color, (C) Ramachandran plot analysis of the multi-epitope based subunit vaccine candidate showing favored (91.7%), allowed (4.8%), generously allowed (2.6%) and disallowed (1%) regions.

Figure 5. Docking studies of multi-epitope vaccine candidate and modelled bovine TLR-4: (A) Three dimensional model of bovine TLR-4 obtained by homologous modelling and refinement, (B) Docked complex of TLR-4 and multi-epitope based subunit vaccine candidate, (C) Binding sites of multi-epitope vaccine candidate with bovine TLR-4. The interacting residues of the putative~~both~~ multi-epitope vaccine candidate and bovine TLR-4 are presented as surface model.~~colored in green sphere.~~

Figure 6. Molecular dynamics simulation study of TLR-4 and multi epitope vaccine protein complex. (A) Root Mean Square Deviation (RMSD) with a time duration of 100ns, grey line represent side-chain and black line represent backbone of the protein chain (B) Root mean Square Fluctuation (RMSF) of the docked complex side-chain with a time duration of 100ns.

Table 1. Antigenic proteins with peptide score and antigenic score

Table 2. Interaction studies with modelled bovine MHC-I with CTL epitopes

Supplementary figures

Figure S1 (A) Molecular modelling of antigenic proteins of *Theileria annulata* containing B-cell epitope using RAPTOR X. The red color represent buried epitopes and blue represent exposed using TMHMM, (B) Ramachandran plot analysis of modelled antigenic proteins of *T. annulata* containing B-cell epitope using RAPTOR X.

Figure S2 Ramachandran plot analysis of bovine MHC-I alleles predicted by Modeller 9.19 showing that all the predicted model are high quality model with more than 90% residue are in core region. (A) BoLA-3*05101 (B) BoLA-1*00901 (C) BoLA-1*00902.

Figure S3 (A) Secondary structure of multi-epitope vaccine protein shows 90% residue are helix, 9% residue are coil and no beta sheet are present in the structure **(B)** Conformational B-cell epitope in the multi-epitope vaccine protein predicted by EliPro server. The multi-epitope vaccine protein was represented in red color with ribbon model. There are nine conformational B-cell epitope are present in the final multi-epitope subunit vaccine protein showed in different color with surface model.

Supplementary Table

Table S1: Annotation of antigenic proteins**.** (XLSX)

Table S2: Antigenic protein prediction using Vaxijen server with threshold 0.5. (XLSX)

Table S3: B-cell epitope prediction using BCPREDS with specificity 90%. (XLSX)

852 **Table S4**: Analysis of epitopes present in cytoplasm, transmembrane and extracellular regions

853 using TMHMM server. (XLSX)

854 **Table S5**: T cell epitope prediction using NetCTL with a threshold of 0.98. (XLSX)

855 **Table S6**: Identification of MHC-I restriction elements. (XLSX)

856 **Table S7**: Identification of CTL epitopes generated in multi-epitope vaccine candidate. (XLSX)

857 **Table S8**: Identification of B-cell epitopes generated in multi-epitope vaccine candidate.

858 (XLSX)

859

Table 1: Antigenic proteins with peptide score and antigenic score

S. No.	Gene ID	Position	Peptide sequence	Peptide Score	Antigenic Score
1	TA09755	132	FLPQTSRPSMGKKKGSFQLP	0.976	0.7931
2	TA13065	157	KKPASDFEEQALEEYLKDKD	0.964	0.7595
3	TA13810	193	DEKEETSKKKYVLMVVVVVV	0.972	0.7115
4	TA13385	336	AVGFPSITENWDSTAATGNG	0.993	0.6889
		375	EYKRHDPSKWPTDGMTTRTA	0.985	
5	TA16735	912	VNWEYVWEKEYKYYRMKMSN	0.986	0.6443
6	TA17050	16	VISSVNAANEDEKKKEEKKD	1	
		105	NLHPAQPKMFKKKGDKEYSE	1	
		212	FYTGDSRLKETYFELKDDKW	0.994	0.5708
		243	LNAMNSSWSTDYKPVVDKFS	0.991	
		43	TSCENVTFKNVDSNTTELTV	0.978	
7	TA11900	198	ELTTTKLVNVIVNGTQESIN	0.984	0.578
8	TA13820	24	EKNEQVTIDINKDATNPRKN	0.985	0.5741
9	TA17055	175	HTVERDDESEEEAAITRVCC	0.999	0.5708
		39	LGYTLDTTIITSIGRDKINR	0.986	
10	TA16125	99	NVTEGGTSMYKYFVKVKGKW	0.963	0.5466
11	TA21100	32	YNSTQNDTNTPRTGSYYNAV	0.999	0.5342
12	TA12115	251	FINFICKIIPPGFPGQLLIQ	1	0.5083
		37	LVKKTLFNRETDPSNNNFPF	0.963	

878 Table 2: Interaction studies with modelled bovine MHC-I with CTL epitopes

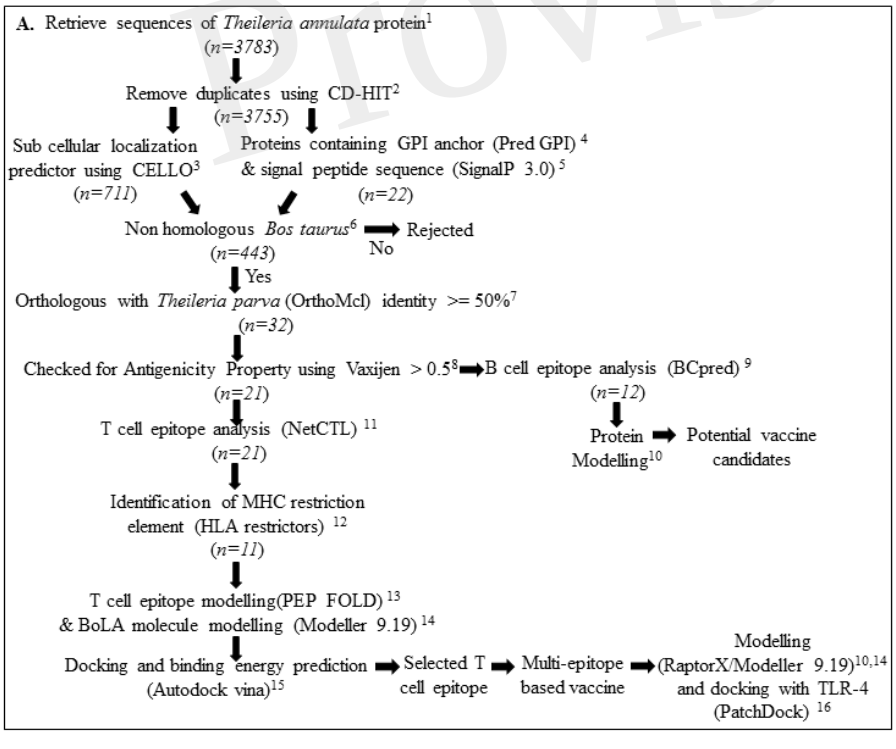
Interaction studies with BoLA-N:05101 and CTL epitopes		
MHC-I	**CTL epitopes**	**Binding energy (Kcal/mol)**
BoLA-N:05101	LYEYVLYLY	-5.8
	FVAWFYKLY	-7.4
	FSNLYSGYY	-6.7
	FLYKRDLPY	-7.4
	IAFCIILYY	-6.9
	FINLTLLTY	-6.4
	ITDVLIYIY	-6.8
	ILLTFNHLY	-7.1
	ILFTISLHY	-5.5
	CSLYFVVLY	-6.3
	KMGHLTIYY	-6.4
	STIAMGLVY	-6.5
	SSFNILLSY	-6.4
	FINLVHYYY	-6.2
	FTEHNSLEY	-6.6
	YSLLFFYLY	-5.8
	SLLFFYLYY	-6.9
	FALDIMTKY	-5.7
	YIKEYFSLY	-7.1
	VVFDYSVKY	-5.5
	VFSIVSSLY	-6.5
	FFELLPSLY	-6.7

Interaction studies with BoLA-N:00901 and CTL epitopes		
MHC-I	**CTL epitopes**	**Binding energy (Kcal/mol)**
BoLA-N:00901	LVLVGSLSY	-5.6
	FLYKRDLPY	-6.7
	WMVFFIVVY	-6.4
	IAFCIILYY	-5.1
	KMGHLTIYY	-6.6
	STIAMGLVY	-5.6
	SSFNILLSY	-6.6
	YSLLFFYLY	-6.3
	VVFDYSVKY	-5.6

Interaction studies with BoLA-N:00902 and CTL epitopes		
MHC-I	**CTL epitopes**	**Binding energy (Kcal/mol)**
BoLA-N:00902	LVLVGSLSY	-6
	FLYKRDLPY	-6.8
	WMVFFIVVY	-6.1
	IAFCIILYY	-7.2
	ILLTFNHLY	-4.8
	KLVNLILLY	-5.6
	MVIAVLSHY	-6.1
	ILFTISLHY	-6.3
	KMGHLTIYY	-6
	STIAMGLVY	-4.4
	SSFNILLSY	-5.6
	YSLLFFYLY	-5.9
	LLVNHTYHY	-5.5
	VVFDYSVKY	-6.5

Figure 01.TIF

Figure 1



B.

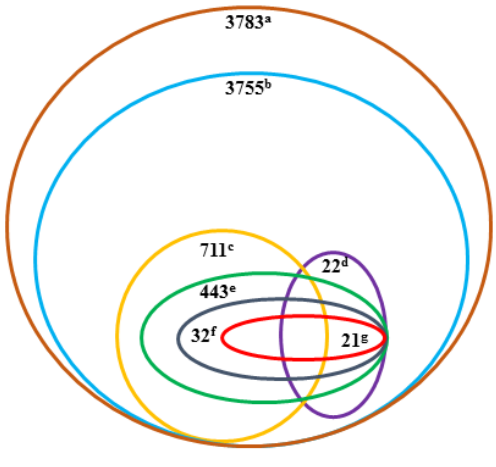


Figure 02.TIF

Figure 2

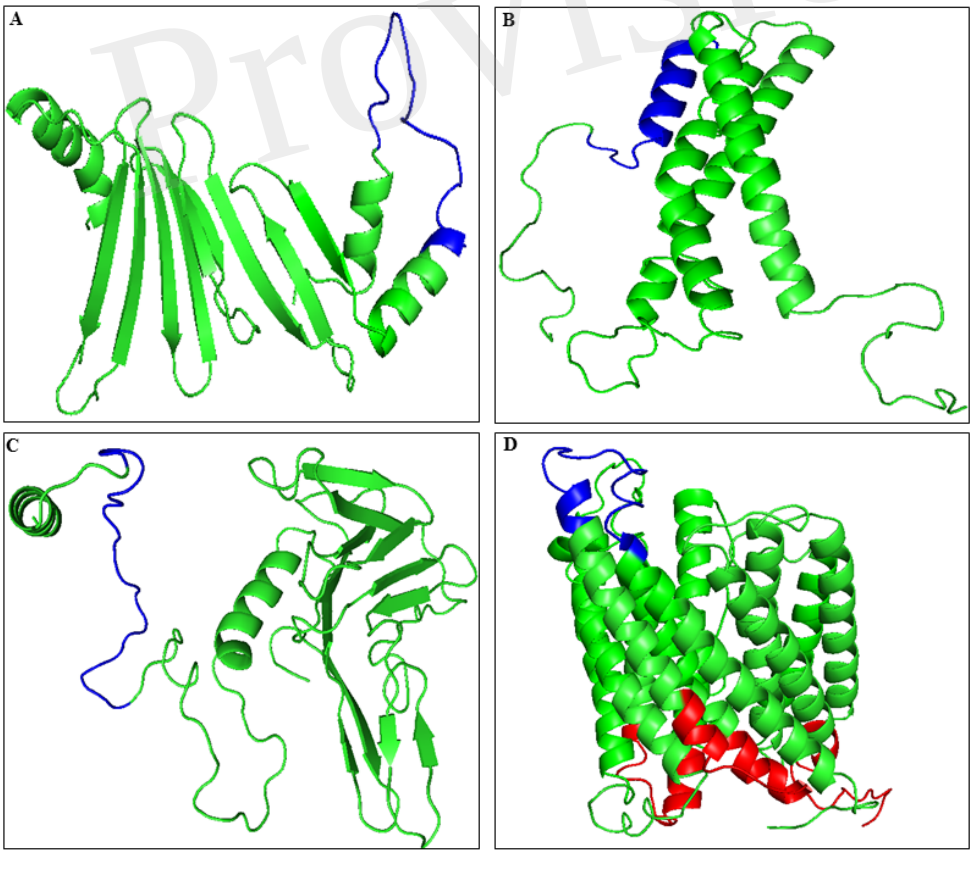


Figure 03.TIF

Figure 3

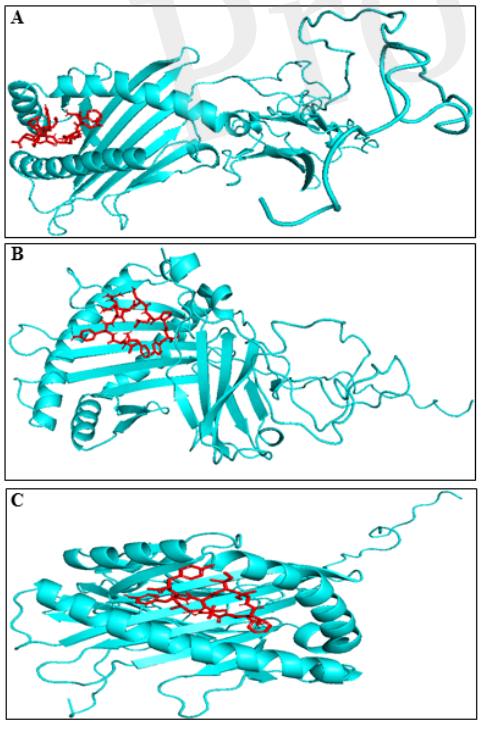


Figure 04.TIF

Figure 4

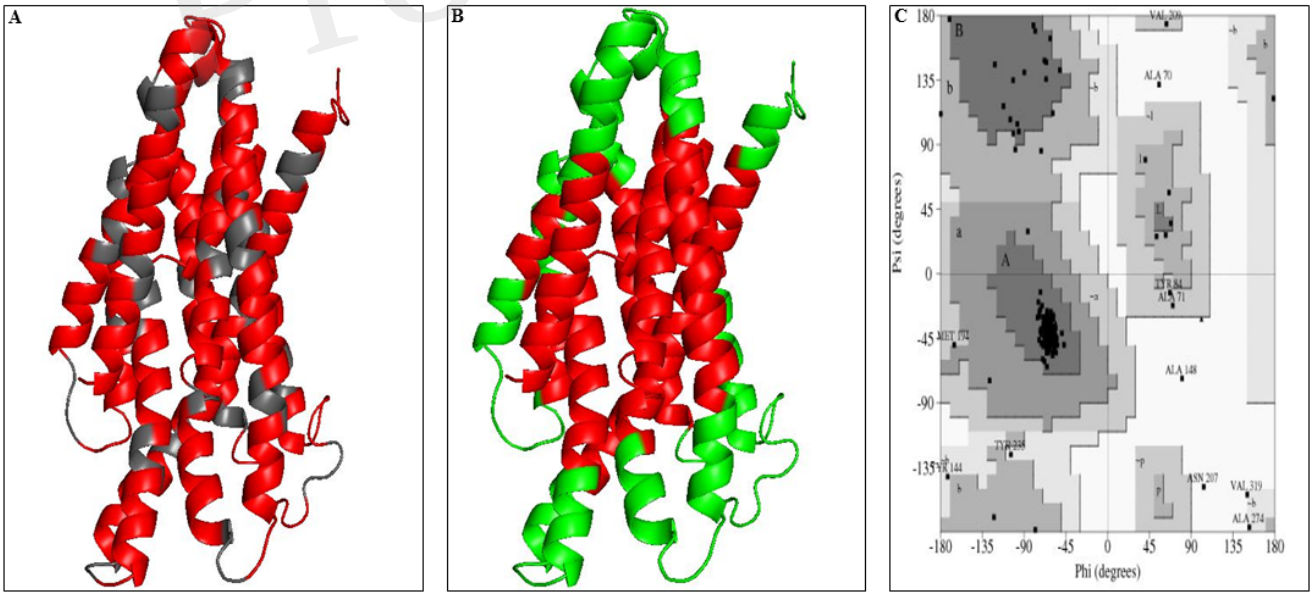


Figure 05.TIF

Figure 5.

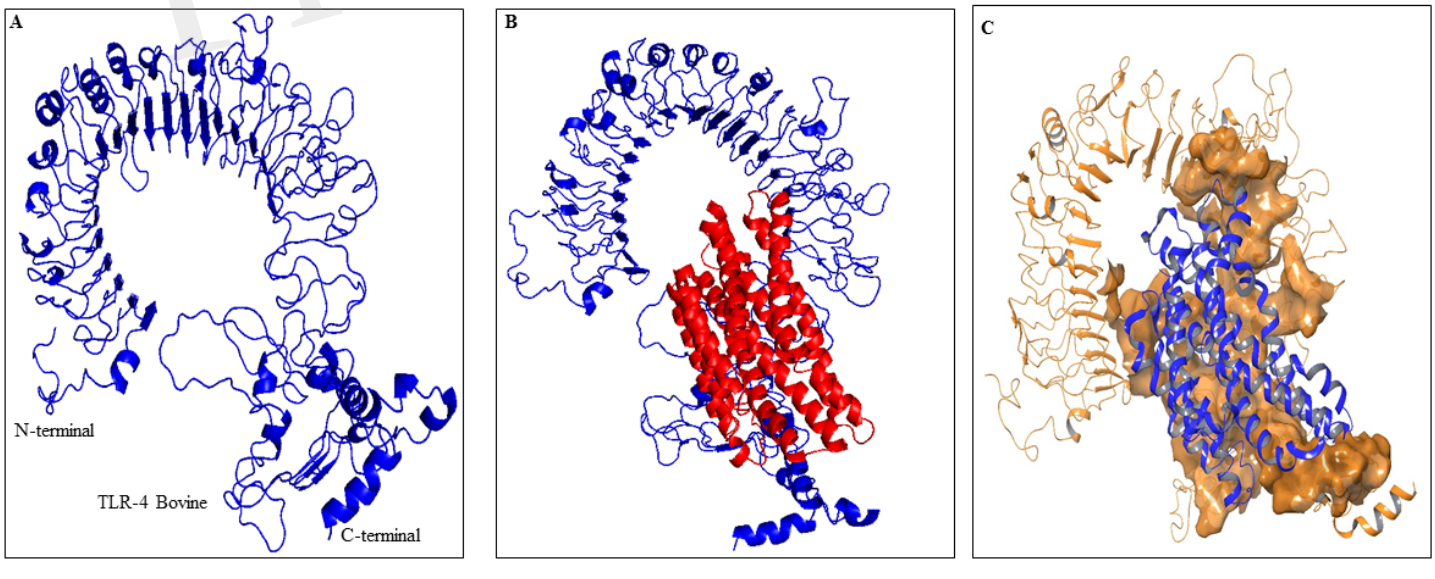


Figure 06.TIF

Figure 6