In Silico Pharmacology (2023) 11:16
https://doi.org/10.1007/s40203-023-00153-5

ORIGINAL RESEARCH

In silico analysis of sporozoite surface antigen 1 of *Theileria annulata* (TaSPAG1) for multi-epitope vaccine design against theileriosis

Mojtaba Azimi-Resketi[1] · Mehdi Akbari[2] · Saeed Heydaryan[3] · Amirreza Eftekhari[4] · Javad Balali[5] · Morteza Shams[6] · Dariush Sargazi[7]

Received: 17 February 2023 / Accepted: 5 July 2023

Abstract
Tropical theileriosis is a protozoan infection caused by *Theileria annulata*, which significantly affects cattle worldwide. This study was aimed to analyze the TaSPAG1 protein and design a novel multi-epitope vaccine candidate. Online tools were employed for the prediction of Physico-chemical properties, antigenicity, allergenicity, solubility, transmembrane domains and signal peptide, posttranslational modification (PTM) sites, secondary and tertiary structures as well as intrinsically disordered regions, followed by identification and screening of potential linear and conformational B-cell epitopes and those peptides having affinity to bind bovine major histocompatibility complex class I (MHC-I) molecules. Next, a multi-epitope vaccine construct was designed and analyzed. This 907-residue protein was hydrophilic (GRAVY: -0.399) and acidic (pI: 5.04) in nature, with high thermotolerance (aliphatic: 71.27). Also, 5 linear and 12 conformational B-cell epitopes along with 8 CTL epitopes were predicted for TaSPAG1. The 355-residue vaccine candidate had a MW of about 35 kDa and it was antigenic, non-allergenic, soluble and stable, which was successfully interacted with cattle MHC-I molecule and finally cloned into the pET28a(+) vector. Further wet studies are required to assess the vaccine efficacy in cattle.

Keywords *Theileria annulata* · SPAG1 · Bioinformatics · Immunogenic epitopes

Abbreviations

TaSPAG1	*Theileria annulata* sporozoite surface antigen 1
MEME	Multiple Em for Motif Elicitation
MAST	Motif Alignment and Search Tool
MW	Molecular weight
pI	Isoelectric point
GRAVY	Grand average of hydropathicity
I-TASSER	Iterative Treading ASSEmbly Refinement
GDT-HA	global distance test-high accuracy
RMSD	root mean square deviation
SVM	Support vector machine
SSK	Subsequent kernel
IEDB	Immune Epitope Database
MHC	Major histocompatibility complex
BoLA	Bovine Leukocyte Antigen
PTM	Post-translational modification

✉ Morteza Shams
Shamsimorteza55@gmail.com

1 Department of Parasitology and Mycology, School of Medicine, Isfahan University of Medical Sciences, Isfahan, Iran

2 Department of Parasitology, School of Medicine, Hamadan University of Medical Sciences, Hamadan, Iran

3 Department of Internal Medicine, Faculty of Veterinary Medicine, University of Tehran, Tehran, Iran

4 Faculty of Veterinary Medicine, Garmsar Branch, Islamic Azad University, Garmsar, Iran

5 Doctor of Veterinary Medicine student, Sanandaj Branch, Islamic Azad University, Sanandaj, Iran

6 Zoonotic Diseases Research Center, Ilam University of Medical Sciences, Ilam, Iran

7 Doctorate in Veterinary Medicine, Head of Zabol Veterinary Network, Zabol, Baluchistan, Sistan, Iran

Published online: 19 July 2023

Introduction

Theileriosis is a serious vector-borne disease caused by *Theileria* parasites with significant morbidity, mortality and clinical disorders among affected animals (Majidiani et al. 2016). Hard ticks transmit the parasitic agents to the domestic and wild ruminants in tropical and subtropical countries as the main host species (Gharbi et al. 2020). Several *Theileria* spp. may infect ruminants, but only a few cause severe or malignant clinical disease and major economic losses, including *Theileria lestoquardi* (small ruminants), *T. annulata* and *T. parva* (cattle and buffalo) (Stuen 2020). *Theileria annulata* is the causative agent of tropical theileriosis, which is transmitted by various species of *Hyalomma*, mostly in the Mediterranean Basin, the Middle East, northeast Africa, southern Asia and India (Bishop et al. 2004). The life cycle includes a host animal (ruminant) in which asexual replication of blood stages occurs and a tick vector within which sexual multiplication happens (Bishop et al. 2004). The infective sporozoites are inoculated while blood-feeding of infected ticks, which subsequently invade host leukocytes. At this stage, some *Theileria* species (*T. annulata*, *T. parva*, *T. lestoquardi*) transform the cell and force it to multiply indefinitely while harboring parasites. These cellular structures are the so-called schizonts, being frequently seen in the circulating blood (Dobbelaere and Heussler 1999). Upon leukocyte lysis, merozoites are released and invade erythrocytes to undergo merogony. Finally, tick vectors suck up the blood stages, including merozoites and gametes, to complete the life cycle (Sivakumar et al. 2014).

The disease, theileriosis, is defined by leukocyte lymphoproliferation, anemia, fever and lymph node enlargement, low milk productivity and loss of carcass quality. Due to the fatal, acute nature of the infection in vulnerable stocks, untreated animals may succumb to death 3–4 weeks post-infection (Gharbi et al. 2006; Irvin and Morrison 1987; Norval et al. 1991). Utilization of acaricides for vector control, chemotherapeutic agents to combat the parasites and vaccination to prevent the infection are three major control measures against theileriosis. The application of acaricides is only efficacious at the community level and for long periods (Kar and Srivastava 2018). Buparvaquone, parvaquone and halofuginone are the main treatment options against theileriosis; however, they are only effective against initial stages (McHardy et al. 1985) of infection and the development of resistant parasites is a matter of concern (Mhadhbi et al. 2010). Only a live attenuated vaccine is available for theileriosis, which requires cold-chain storage. Also, some vaccinated immunocompromised cattle may play as a provenance for the transmission of *Theileria* to other cattle (Maritim et al. 1989; Pipano and Tsur 1966). Because of the shortcomings of such measures, vaccination seems essential for improved disease control (Nene and Morrison 2016).

Today, next-generation vaccine design is a pioneer procedure to develop more efficacious and affordable vaccines, based on identifying novel antigenic proteins/epitopes in the organism and engineering those immunodominant regions in the context of multi-epitope vaccine constructs (Asghari et al. 2022; Kallerup and Foged 2015). This kind of vaccine research is desirable to produce a highly efficient immunization approach against tropical theileriosis in cattle herds. *Theileria annulata* sporozoite surface antigen 1 (TaSPAG1) is encoded by single-copy genes, and it is a principal component of the sporozoite membrane (Williamson et al. 1989). This protein plays a critical role in host cell recognition and invasion (Nene and Morrison 2016). Hence, in the present *in silico* study, TaSPAG1 was selected to reveal further its immunogenic epitopes and biochemical characteristics, employing them towards multi-epitope vaccine design.

Methods

Retrieval of TaSPAG1 protein sequence

The amino acid sequence of TaSPAG1 protein was obtained using the UniProtKB database, available at https://www.uniprot.org/ (Consortium 2015), under the accession number of Q26675.

Prediction of antigenicity, allergenicity, solubility and physico-chemical properties

Prediction of antigenicity is the main step in the identifying of a promising vaccine candidate. For this purpose, two web servers, including ANTIGENpro (http://scratch.proteomics.ics.uci.edu/) and VaxiJen v2.0 (http://www.ddg-pharmfac.net/vaxijen/), were employed, which apply different machine-learning methods to identify antigens (Asghari et al. 2021c; Asghari et al. 2021d; Doytchinova and Flower 2007; Magnan et al. 2010). A proper vaccine candidate should not represent allergenicity. Hence three web servers were used to predict allergenicity in the protein sequence, including AllerTOP v2.0 (http://www.ddgpharmfac.net/AllerTOP), AllergenFP v1.0 (https://ddgpharmfac.net/AllergenFP/) and AlgPred (http://crdd.osdd.net/raghava/algpred/). In order to spot the IgE epitopes, MEME (Multiple Em for Motif Elicitation)/MAST (Motif Alignment and Search Tool) allergen motifs were utilized by the AlgPred webserver to predict allergens (Sharma et al. 2020). Moreover, an alignment-free approach with Mathews correlation coefficient of 0.759 is employed by AllergenFP v1.0 server, while AllerTOP v2.0 exploits several machine learning

methods, comprising *k*-nearest neighbors, cross variance transformation and E descriptors (Dimitrov et al. 2014a, b). Finally, ExPASy ProtParam server (https://web.expasy.org/protparam/) was used to estimate some important Physicochemical properties of TaSPAG1 such as molecular weight (MW), the number of negatively and positively-charged residues, aliphatic and instability indices, isoelectric point (pI), half-life and grand average of hydropathicity (GRAVY) (Gasteiger et al. 2005; Shams et al. 2021b).

Prediction of post-translational modification (PTM) sites of TaSPAG1

A number of PTM sites were predicted for TaSPAG1 protein, including Phosphorylation, O-glycosylation, N-glycosylation and GPI anchor prediction. For this purpose, the DTU Health Tech Services (Denmark) online tools, including NetPhos, NetOGlyc, NetNGlyc and NetGPI, were used, respectively (https://services.healthtech.dtu.dk/). The "All Asn residues" option was used for NetNGlyc 1.0 prediction, while default parameters were applied to NetOGlyc 4.0 server.

2.4. Signal peptide and transmembrane domain prediction

Prediction of the putative signal peptide in the protein sequence was done using the SignalP web server, available at http://www.cbs.dtu.dk/services/SignalP/. Moreover, transmembrane domains were predicted by TMHMM 2.0 web tool of DTU Health Tech Services (https://services.healthtech.dtu.dk/).

Prediction of TaSPAG1 secondary structure and disordered regions

The secondary structure of the TaSPAG1 protein was predicted by utilization of the NetSurfP-2.0 tool of DTU Health Tech Services, available at https://services.healthtech.dtu.dk/. This Server predicts the surface accessibility, secondary structure, disordered regions, and phi/psi dihedral angles of amino acids in an amino acid sequence (Klausen et al. 2019).

Prediction of three-dimensional (3D) model, refinement, validations and solubility analysis

An Iterative Treading ASSEmbly Refinement (I-TASSER) modality was employed for a fine-tuned, automated homology modelling of the TaSPAG1 protein (https://zhanglab.dcmb.med.umich.edu/I-TASSER/) (Yang et al. 2015). Subsequently, the best-fit 3D model provided by I-TASSER was subjected to refinement analysis using GalaxyRefine web server, being accessible at http://galaxy.seoklab.org/cgi-bin/submit.cgi?type=REFINE. Based on the submitted pdb file as the initial model, the server provides five refined models based on several parameters, including global distance test-high accuracy (GDT-HA), root mean square deviation (RMSD), MolProbity, Clash score, Poor rotamers and Rama favored (Heo et al. 2013). Validation of the refining process was done using Prosa-web (https://prosa.services.came.sbg.ac.at/prosa.php) (Wiederstein and Sippl 2007) and PROCHECK (https://saves.mbi.ucla.edu/) servers, in order to determine the Z-score and Ramachandran plot analysis of the refined model in comparison with the initially submitted one (Laskowski et al. 1993). The protein solubility was calculated using the Protein-Sol web server available at https://proteinsol.manchester.ac.uk/, which outputs above 0.45 indicates higher solubility (Hebditch et al. 2017).

Prediction of continuous and conformational B-cell epitopes

Linear B-cell epitopes were predicted using a multi-step approach (Asghari et al. 2021a; Shams et al. 2021a). For this aim, 14-mer epitopes were initially predicted for TaSPAG1 using BCPREDS server with 80% specificity (http://ailab.ist.psu.edu/bcpred/predict.html). This server applies support vector machine (SVM) and subsequent kernel (SSK) to predict epitopes (El-manzalawy et al. 2008). In the next step, cross-validation of the predicted epitopes was accomplished with the outputs of two other web servers, including ABCpred (http://crdd.osdd.net/raghava/abcpred/ABC_submission) (Saha and Raghava 2006) and SVMTriP (http://sysbio.unl.edu/SVMTriP/prediction.php) (Yao et al. 2012). Those epitopes being shared among outputs of the above servers were selected for further screening with respect to antigenicity, allergenicity and water solubility using VaxiJen v2.0, AllerTOP v2.0 and PepCalc web servers, respectively. Another server was, also, employed for the prediction of continuous antibody epitope (http://tools.iedb.org/bcell/). "Parameters such as hydrophilicity, flexibility, accessibility, turns, exposed surface, polarity and the antigenic propensity of polypeptides chains have been correlated with the location of continuous epitopes. Those residues with higher probability to be a part of a linear epitope are colored in yellow" (Vita et al. 2019). Conformational B-cell epitopes are highly essential for improved antigen-antibody interactions. Hence, they were also predicted using the ElliPro tool of the Immune Epitope Database (IEDB) server (http://tools.iedb.org/ellipro/) (Ponomarenko et al. 2008).

Prediction and screening of epitopes capable to bind cattle major histocompatibility complex (MHC)-I molecule

All epitope predictions were made using the MHC-I binding epitope prediction tool of the IEDB server (http://tools.iedb.org/mhci/), and the cow was selected as the main host species for *T. annulata*. The lower percentile rank demonstrated a higher capacity for the epitope to bind the respective MHC molecule. Based on this, nine bovine leukocyte antigen (BoLA) alleles including BoLA-1:00901, BoLA-2:00501, BoLA-3:05002, BoLA-4:02401, BoLA-5:00301, BoLA-6:01501, BoLA-AW10, BoLA-T2a and BoLA-JSP.1 were used with subsequent screening in terms of antigenicity, allergenicity and toxicity using VaxiJen v2.0, AllergenFP v1.0 and ToxinPred (https://webs.iiitd.edu.in/raghava/toxinpred/algo.php) web servers, respectively.

Molecular docking analysis of the TaSPAG1 protein and bovine MHC-I

The online computerized docking procedure was performed using ClusPro2 web server using refined TaSPAG1 protein and cattle MHC-I (PDB ID: 2XFX) in order to estimate the binding affinity between both molecules. The server output provides top-rank clusters, among which the most populated docking pose with highest energy was selected for visualization.

Multi-epitope vaccine design

After a computer-based multi-step prediction and screening of specific epitopes, a multi-epitope vaccine was designed using selected B-cell and bovine MHC-I epitopes. To enhance the immunogenicity of the vaccine, 50 S ribosomal protein L7/L12 of *Mycobacterium tuberculosis* (Accession No: P9WHE3) was added as genetic adjuvant to the N-terminal of the vaccine sequence. Also, a 6×histidine (His-tag) sequence (CATCACCATCACCATCAC) was added C-terminally for future protein purification. It is noteworthy that "EAAAK", "GPGPG" and "AAY" linkers were used to connect adjuvant, B-cell and MHC-I epitopes together, respectively.

Basic characteristics and structural analysis of the designed vaccine candidate

The antigenicity, allergenicity and solubility of the multi-epitope vaccine were predicted using VaxiJen v2.0, AllergenFP and AllerTOP as well as Protein-Sol web servers, respectively. Physico-chemical properties of the vaccine model were, also, forecasted using ProtParam web tool. In addition, NetSurfP-2.0 and I-TASSER servers were employed to analyze the secondary and tertiary structures of the designed vaccine candidate. Upon refinement using GalaxyRefine server, the most appropriate refined 3D vaccine model was further validated through Prosa-Web and PROCHECK online tools.

Conformational B-cell epitope prediction and molecular docking with bovine MHC-I

Conformational B-cell epitopes are important for specific antibody responses against vaccine molecule, so they were predicted using a highly-reliable web tool, ElliPro of IEDB server. In the following, a protein-protein docking was performed to assess possible interactions between the designed multi-epitope vaccine model and bovine MHC-I molecule as receptor (PDB ID: 2XFX). The most populated cluster with high energy was chosen for visualization of the residue-by-residue interactions.

Reverse translation, codon adaptation and *in silico* cloning

Higher protein yields produced by the *E. coli* expression system is a vital step for subunit vaccine production. Accordingly, reverse translation and codon optimization were accomplished using reverse translate tool of sequence manipulation suite (https://www.bioinformatics.org/sms2/rev_trans.html) and JCat server (http://www.jcat.de/), respectively. The latter demonstrates the codon adaptation index (CAI) and GC content of a submitted DNA sequence, which are necessary for high-throughput expression in the respective host. Here, we optimized the codons for enhanced protein expression in *E. coli* K12 strain using pET28a(+) vector. Finally, upon checking the potential cutting sites of commercially-available restriction enzymes through NEBcutter 2.0 server (https://nc2.neb.com/NEBcutter2/), start/stop codons and restriction sites of *Eco53KI* and *EcoRV* enzymes were added to the 5′ and 3′-OH of the codon adapted vaccine sequence, respectively. The *in-silico* cloning process was performed using SnapGene® v6.2 software.

Results

General antigenic, allergenic and biochemical characteristics of TaSPAG1

Theileria annulata SPAG1 protein was shown to be highly antigenic; this was substantiated by the VaxiJen score of 0.8560 and the ANTIGENpro score of 0.938922. Moreover,

the protein neither was allergenic nor had IgE epitopes and/or MEME/MAST motifs. Based on the ProtParam server output, this 907-residue protein had 12,903 atoms, a molecular weight of 91885.12 Dalton, with most residues (104) as negatively charged (Asp + Glu), whereas only 81 residues were positively-charged (Arg + Lys). The estimated half-life was 30 h (mammalian reticulocytes, in vitro), > 20 h (yeast, in vivo) and > 10 h (*Escherichia coli*, in vivo). Based on the computed theoretical pI of 5.04, TaSPAG1 was relatively acidic. The instability index of 41.68 classified the protein as unstable. Furthermore, the aliphatic index (71.27) and GRAVY score (-0.399) demonstrated that this protein is highly thermotolerant and hydrophilic.

Prediction of transmembrane domain, signal peptide and PTM sites

Based on TMHMM server output, there was one transmembrane domain in the sequence. Also, the SignalP server demonstrated that a signal peptide is present within the protein sequence, with a high probability (0.8123). The cleavage site for the signal peptide was embedded somewhere between 18 and 19 positions. With respect to PTM sites, there were 138 O-glycosylation regions and 3 N-glycosylation sites. Also, it was a GPI-anchored protein, based on the NetGPI server result, with the likelihood of 0.509. Of note, the protein had several phosphorylation sites (121 regions); in detail, serine, tyrosine and threonine phosphorylation sites were calculated to be 80, 39 and 2, respectively (**Supplementary File 1**).

Prediction of secondary structures and disordered regions

NetSurfP-2.0 server was used for a fine-tuned secondary structure prediction. The output revealed that over half of the protein sequence, particularly between 1 and 545 amino acid residues, was structurally disordered. Such disordered areas were relatively surface accessible, while other areas were buried. Most of the sequence was constituted by random coils (mostly between 1 and 547 residues) followed by helices, whereas few areas were predicted to be extended strand (**Supplementary File 2**).

3D structure homology modelling, refinement, validations and solubility analysis

I-TASSER server-generated top-five structural conformations, among which the model with the highest C-score (high confidence) was selected as the final TaSPAG1 tertiary structure. Among top 10 threading templates used by I-TASSER, template 7qfpA (Iden1: 0.09, Iden2: 0.19, Cov: 0.93, Norm Z-score: 2.66) was chosen. Also, the selected model had a C-score of -1.77, an estimated TM-score of 0.50 ± 0.15 and an estimated RMSD of 13.1 ± 4.2 Å. GalaxyRefine server provided five refined models, and based on several criteria mentioned before, we selected model number one for further validations. This refined model possessed GDT-HA of 0.8881, RMSD of 0.580, MolProbity of 2.592, Clash score of 23.4, Poor rotamers of 0.7 and Rama favored of 80.2. Finally, the quality of the refined model, as compared with the crude model, was evaluated using two servers. The Z-score of the crude model (-1.31) was significantly increased in the refined model (-2.35). Moreover, Ramachandran plot analysis showed that in crude model 328 (47.3%), 270 (39.0%), 67 (9.7%) and 28 (4.0%) of residues are appointed to the most favored, additional allowed, generously allowed and disallowed regions, respectively. Upon refinement, the arrangement of the residues was changed as 471 (68.0%), 175 (25.3%), 17 (2.5%) and 30 (4.3%) in the most favored, additional allowed, generously allowed and disallowed regions, respectively. A considerably high solubility score was calculated for TaSPAG1, with predicted scaled solubility of 0.652 (Fig. 1).

Linear and conformational B-cell epitopes

Multi-step approach showed 15 shared epitopes predicted using three web servers (ABCpred, BCPREDS and SVMTriP) (**Supplementary File 3**). Among these, five epitopes were potentially antigenic, non-allergenic with acceptable water solubility, including "TTTTTTSSGKPSDQ" (VaxiJen score: 1.8820), "SEGEDDDDEEEEEE" (VaxiJen score: 2.1164), "DQTSGSGSKGTE" (VaxiJen score: 1.7398), "SSASPTSPT" (VaxiJen score: 1.5176) and "GPIPSPGD" (VaxiJen score: 0.5559) (Table 1). Also, ElliPro tool demonstrated 12 conformational B-cell epitopes in TaSPAG1 with the following lengths and scores: (1) 129 residues, score: 0.759; (2) 143 residues, score: 0.735; (3) 59 residues, score: 0.715; (4) 80 residues, score: 0.698; (5) 33 residues, score: 0.602; (6) 6 residues, score: 0.581; (7) 8 residues, score: 0.57; (8) 23 residues, score: 0.557; (9) 7 residues, score: 0.553; (10) 3 residues, score: 0.542; 11) 6 residues, score: 0.526, and 12) 10 residues, score: 0.509 (**Supplementary File 4**). Moreover, linear antibody epitope prediction was done using B-cell tool of the IEDB server, with regard to beta-turn, accessibility, flexibility, antigenicity and hydrophilicity (**Supplementary File 5**).

Prediction of cattle MHC-I binding epitopes

The top three MHC-I binding epitopes that were strong binders (lower percentile ranks) were predicted for several BoLA molecules. Among these, 8 epitopes were shown

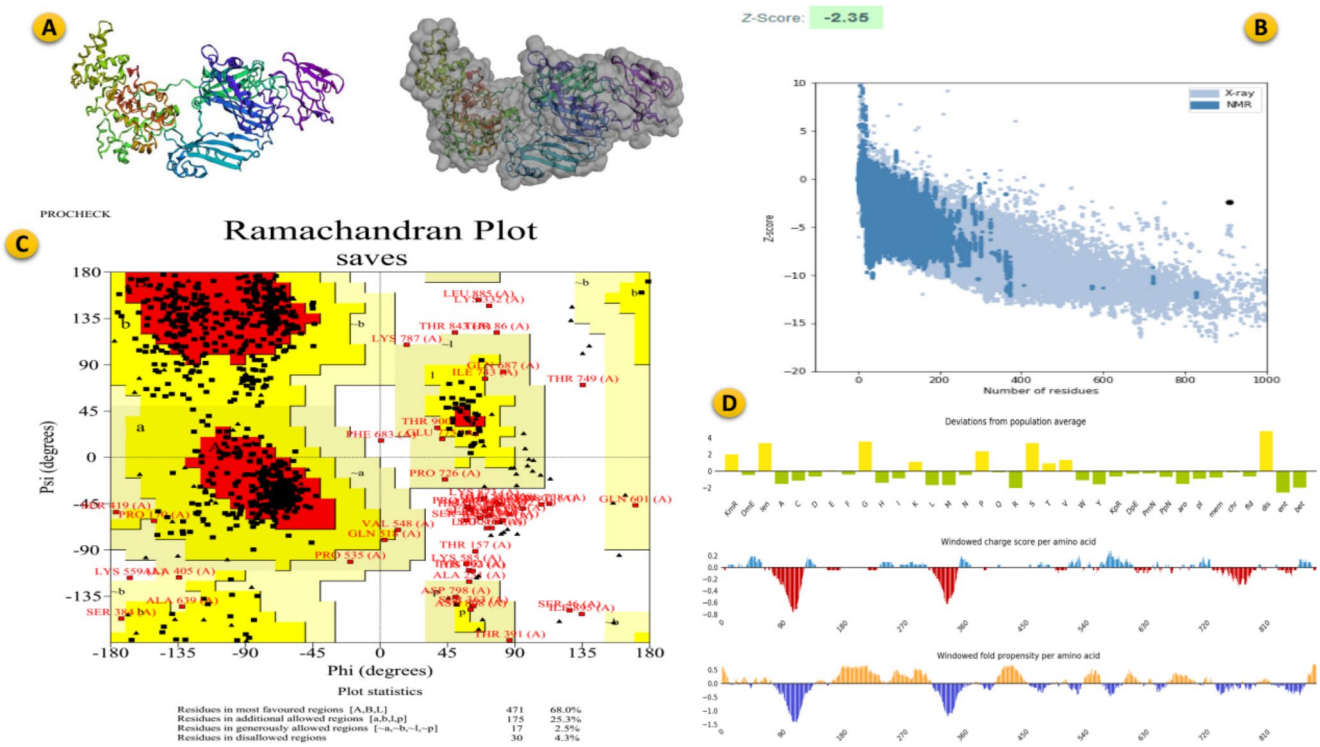

Fig. 1 Homology modelling (**A**), 3D model validations (**B, C**) and solubility analysis (**D**) of TaSPAG1

Table 1 The final screening of shared linear B-cell epitopes from *T. annulata***SPAG1**.

No.	Shared B-cell epitopes	VaxiJen antigenicity score	AllergenFP allergenicity prediction	PepCalc water solubility prediction
1	GVGVPGVGGVPGVG	0.1574	Allergen	Poor
2	PGVGVPGVGVPG	1.8473	Non-allergen	Poor
3	TTTTTTSSGKPSDQ*	1.8820	Non-allergen	Good
4	SEGEDDDDEEEEEE*	2.1164	Non-allergen	Good
5	DEDDDDDEEEEEDD	1.8020	Allergen	Good
6	SKNGKGSPKAQPGV	1.1911	Allergen	Good
7	DQTSGSGSKGTE*	1.7398	Non-allergen	Good
8	LNTIPDKVGREF	-1.4691	Non-allergen	Good
9	QQDPAPSKP	0.4336	Non-allergen	Good
10	LVKDVSEE	0.5285	Allergen	Good
11	VGVAPGVGV	-0.6451	Allergen	Poor
12	LSSITNSVYSLIK	0.5128	Non-allergen	Poor
13	SASQTSPTTT	0.8422	Non-allergen	Poor
14	SSASPTSPT*	1.5176	Non-allergen	Good
15	GPIPSPGD*	0.5559	Non-allergen	Good

(*) indicates antigenic, non-allergenic epitopes with potential good water solubility

to be antigenic, non-allergenic and non-toxic, comprising "LSQTGLGPSGSH" (VaxiJen score: 0.9975), "SQTGLGPSGSHA" (VaxiJen score: 1.1970), "SSASPTSPTTTL" (VaxiJen score: 0.9973), "SFQEPVSQELEF" (VaxiJen score: 0.5648), "KLLGSGFEVASI" (VaxiJen score: 0.9138), "TSPTTTLSQTGL" (VaxiJen score: 0.5101), "VTLESAVTQPSK" (VaxiJen score: 0.7541) and "SAVTQPSKDPFK" (VaxiJen score: 0.9265) (Table 2).

Protein-protein docking analysis between TASPAG1 and cattle MHC-I

The piper-based ClusPro 2.0 web server provided a number of docking poses between the cattle MHC-I molecule (PDB code: 2XFX) as receptor and the refined protein structure. The first ranked, highly populated (70 members) docking cluster with the highest binding score (-973.3) was chosen for further visualization of the molecular interactions and

Table 2 Prediction of bovine MHC-I-binding epitopes of *T. annulata***SPAG1** using IEDB server followed by antigenicity, allergenicity and toxicity screening

Bovine MHC-I alleles	Position	T-cell peptide	Percentile rank	VaxiJen antigenicity score	AllergenFP allergenicity prediction	ToxinPred toxicity prediction
BoLA-1:00901	43–54	LSQTGLGPSGSH*	1.5	0.9975	Non-Allergen	Non-Toxin
	44–55	SQTGLGPSGSHA*	2.0	1.1970	Non-Allergen	Non-Toxin
	2–13	LDELVKDVSEEH	3.2	0.3472	Non-Allergen	Non-Toxin
BoLA-2:00501	32–43	SSASPTSPTTTL*	1.0	0.9973	Non-Allergen	Non-Toxin
	31–42	KPSPLVTLESAV	1.4	0.3746	Allergen	Non-Toxin
	23–34	SSRTPNAKPAEL	2.2	1.1019	Allergen	Non-Toxin
BoLA-3:05002	28–39	NAKPAELGPSLV	7.3	0.8086	Allergen	Non-Toxin
	32–43	SSASPTSPTTTL	8.0	0.9973	Non-Allergen	Non-Toxin
	27–38	SRTSKPSPLVTL	11	0.4847	Non-Allergen	Non-Toxin
BoLA-4:02401	49–60	FLGDFKPKPRRY	1.8	2.0571	Allergen	Non-Toxin
	32–43	VNKPEELAEFLW	3.0	0.4296	Allergen	Non-Toxin
	28–39	SFQEPVSQELEF*	3.4	0.5648	Non-Allergen	Non-Toxin
BoLA-5:00301	26–37	AQYATNDILSSI	0.68	0.4415	Non-Allergen	Non-Toxin
	11–22	KQFIFEEVKSLV	1.6	0.2293	Non-Allergen	Non-Toxin
	43–54	KLLGSGFEVASI*	1.9	0.9138	Non-Allergen	Non-Toxin
BoLA-6:01501	34–45	SQELEFQSDTEI	0.99	1.3244	Allergen	Non-Toxin
	26–37	AQYATNDILSSI	1.2	0.4415	Non-Allergen	Non-Toxin
	10–21	KELEEAFNSIGL	1.4	-0.2356	Non-Allergen	Non-Toxin
BoLA-AW10	37–48	TSPTTTLSQTGL*	0.44	0.5101	Non-Allergen	Non-Toxin
	32–43	SSASPTSPTTTL	0.6	0.9973	Non-Allergen	Non-Toxin
	27–38	SRTSKPSPLVTL	0.64	0.4847	Non-Allergen	Non-Toxin
BoLA-T2a	35–46	SSITNSVYSLIK	0.33	0.5926	Non-Allergen	Non-Toxin
	36–57	VTLESAVTQPSK*	0.39	0.7541	Non-Allergen	Non-Toxin
	40–51	SAVTQPSKDPFK*	0.71	0.9265	Non-Allergen	Non-Toxin
BoLA-JSP.1	37–48	TSPTTTLSQTGL	0.12	0.5101	Non-Allergen	Non-Toxin
	27–38	SRTSKPSPLVTL	0.6	0.4847	Non-Allergen	Non-Toxin
	32–43	SSASPTSPTTTL	0.86	0.9973	Non-Allergen	Non-Toxin

(*) indicates potential high-ranked, antigenic, non-allergenic and non-toxic epitopes

binding conformation. As illustrated in Fig. 2, the molecular interaction was observed among chains C (TaSPAG1) and both chains (A and B) of the cattle MHC-I molecule. Moreover, the interactions at the amino acid level are provided.

Designing the multi-epitope vaccine candidate

In the present study, a multi-epitope vaccine candidate was designed using five potent B-cell and eight cattle MHC-I binding epitopes. To enhance the adjuvancity and purification rate of the polypeptide, *M. tuberculosis* 50 S ribosomal L7/L12 and H6-tag were added to the N- and C-termini of the vaccine sequence. Of note, the vaccine had 355 residues in length (Fig. 3A).

Basic and structural properties of the vaccine candidate

Based on VaxiJen server, this vaccine model possessed a high antigenicity score of 0.7871, while it was not shown to be allergenic in nature by AllergenFP and AllerTOP web servers. According to the Protein-Sol web tool, the vaccine demonstrated considerable solubility (0.802). ProtParam server revealed that the vaccine had a MW of 35214.85 Dalton, with a relatively acidic theoretical pI (4.36). Moreover, the vaccine model possessed considerable half-life in mammalian reticulocytes (30 h), was stable (Instability index: 38.51), highly thermotolerant (aliphatic index: 64) and hydrophilic in nature (GRAVY: -0.203).

Secondary structure analysis of the vaccine model using NetSurfP-2.0 server showed the predominance of exposed areas and disordered regions, and coils followed by helices were abundant in the sequence. Next, I-TASSER server was used for 3D model prediction, and the output showed model number 1 as the best predicted model for the designed vaccine, with the highest C-score (-2.31), estimated TM-score of 0.44 ± 0.14 and RMSD of 12.0 ± 4.4 Å (Fig. 3B**&C**). Based on Prosa-Web, the Z-score was improved from -2.78 (crude) to -3.04 (refined). In addition, Ramachandran plot analysis showed improvements in those residues in favored (77.9%), additional allowed (15.8%), generously-allowed (1.8%) and disallowed regions (4.6%) in the refined model,

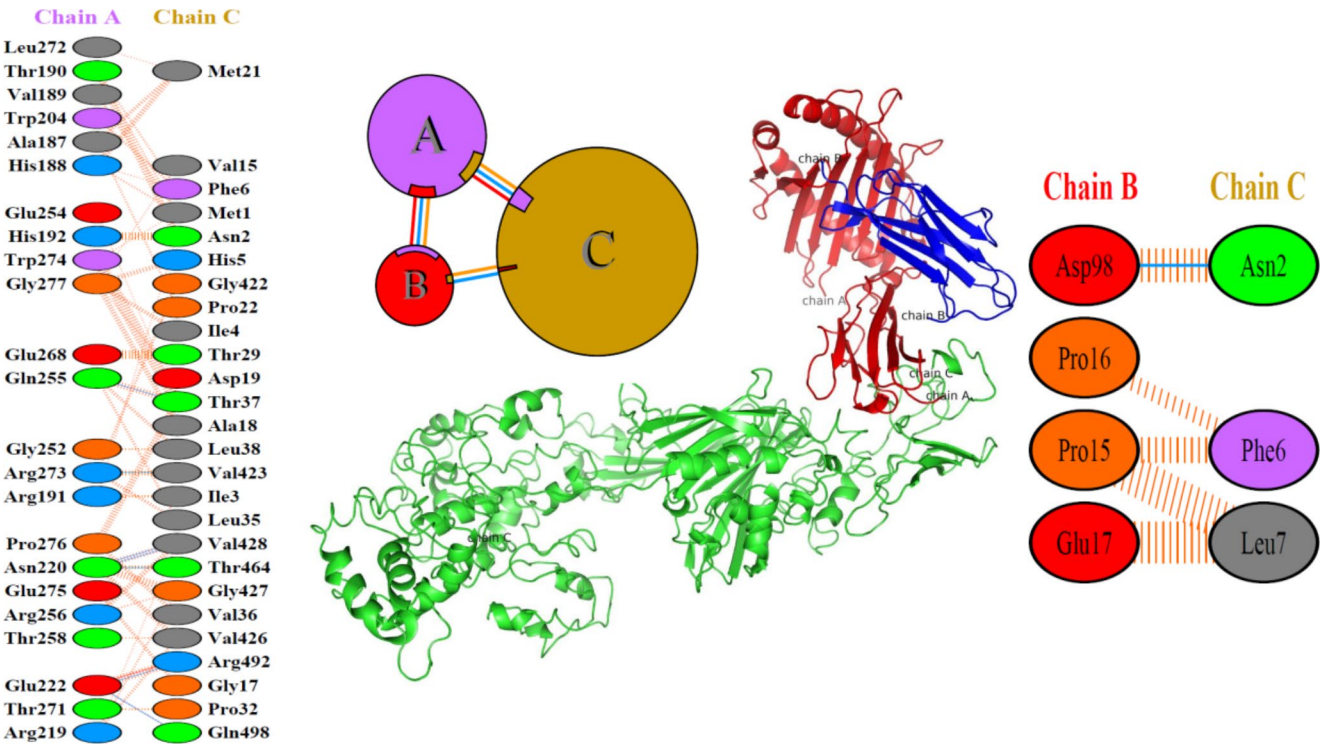

Fig. 2 Schematic diagram of the type and the extent of interactions between TaSPAG1 protein and cattle MHC-I (chains **A** and **B**). Blue line denotes hydrogen bond, red denotes salt bridge and orange striped lines denote non-bonded contacts. In cases which the non-bonded contacts are plentiful, the width of the striped line increases depending on the number of the non-bonded contacts. The size of each circle is proportional to the surface area of the corresponding protein chain

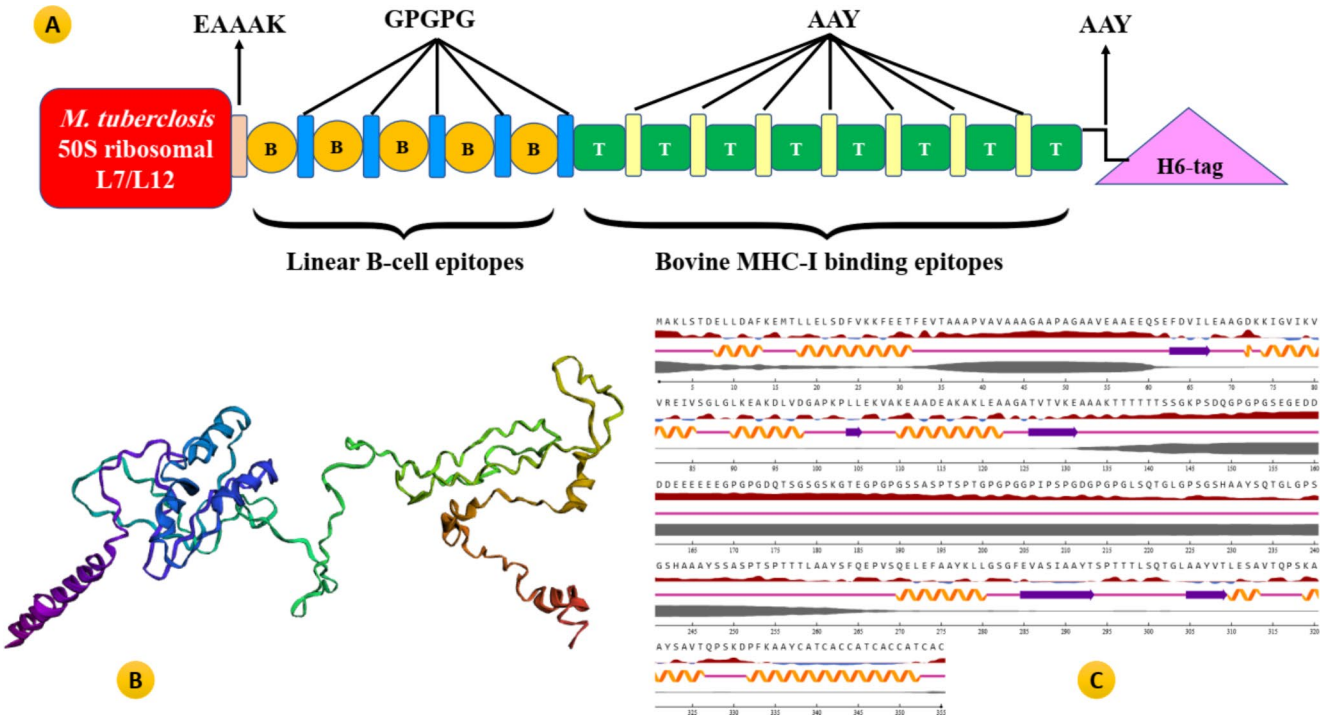

Fig. 3 The engineered vaccine construct and structural analyses. (**A**) Multi-epitope vaccine construct; (**B**) tertiary structure prediction of the vaccine candidate using I-TASSER server; and (**C**) Secondary structure prediction using NetSurfP-2.0 server

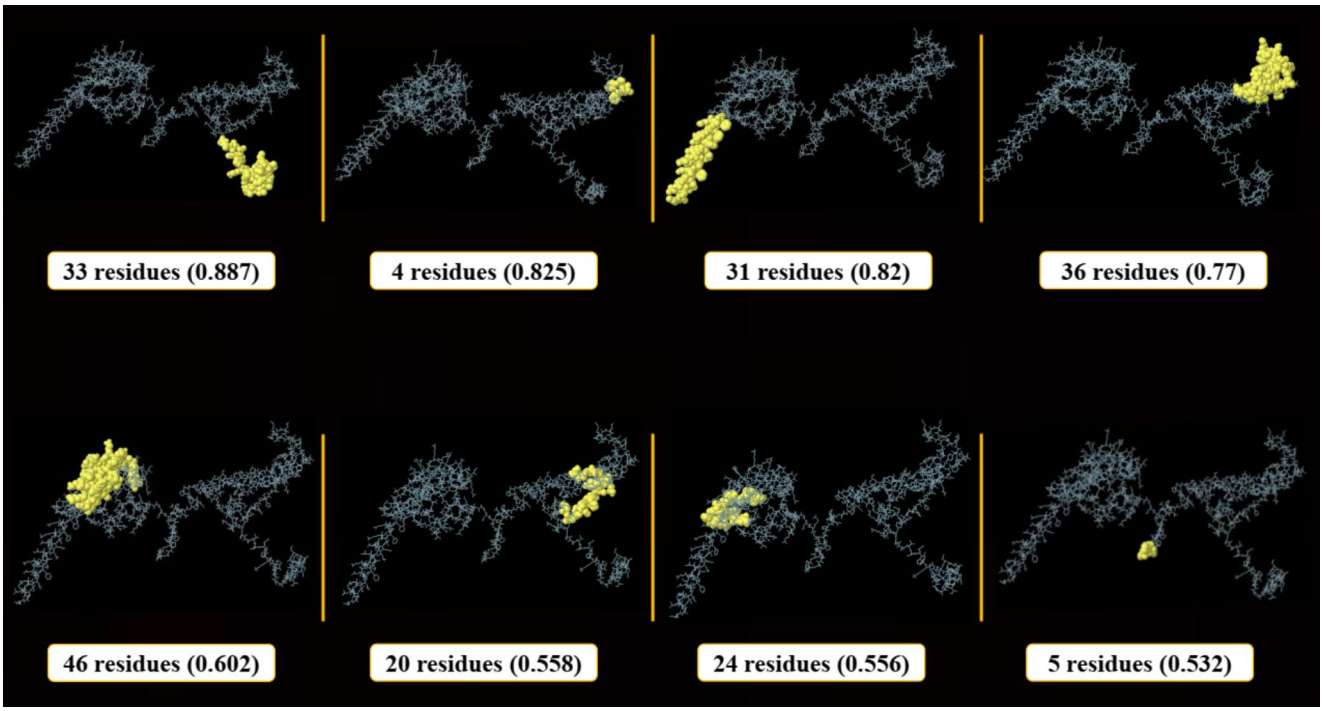

Fig. 4 Conformation B-cell epitopes present in the final vaccine model, predicted by ElliPro tool

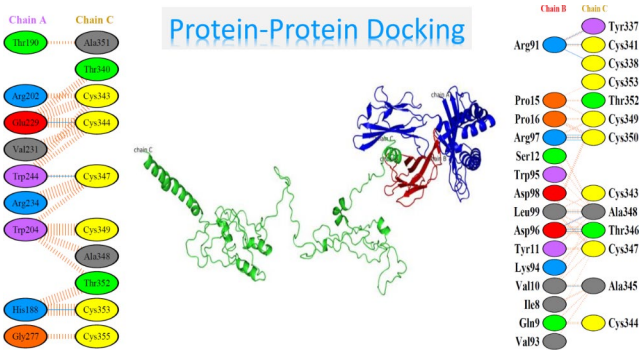

Fig. 5 Schematic diagram of the type and the extent of interactions between vaccine candidate protein and cattle MHC-I (chains **A** and **B**). Blue line denotes hydrogen bond, red denotes salt bridge and orange striped lines denote non-bonded contacts. In cases which the non-bonded contacts are plentiful, the width of the striped line increases depending on the number of the non-bonded contacts. The size of each circle is proportional to the surface area of the corresponding protein chain

in comparison with those (53%, 37.9%, 5.6% and 3.5%) in the crude model.

Non-linear B-cell epitopes and protein-protein docking

Based on ElliPro tool, eight conformational B-cell epitopes were predicted for the vaccine model, with the following residues and scores: (1) 33 residues, 0.887; (2) 4 residues, 0.825; (3) 31 residues, 0.82; (4) 36 residues, 0.77; (5) 46 residues, 0.602; (6) 20 residues, 0.558; (7) 24 residues, 0.556; and (8) 5 residues, 0.532 (Fig. 4).

Based on the ClusPro server, cluster 0 had the highest number of members (n = 160) and the highest binding energy score (-1002.2), which was selected as the best docked complex for visualization. Significant molecular interactions were observed between the vaccine and bovine MHC-I molecule, prominently with respect to the chain B of MHC-I. The interactions at the molecular and amino acid levels are illustrated in Fig. 5.

In silico cloning

The reverse-translated vaccine sequence was further codon-adapted for expression in *E. coli* K12 strain using JCat server. The output showed substantial improvements in the CAI-value and GC%, so that these values were changed from 0.62 to 65.35% in the crude sequence, respectively, to 1 and 54.74% in the refined model. Ultimately, the start (ATG) and stop (TAA) codons were added, along with *Eco53KI* (GAGCTC) and *EcoRV* (GATATC) restriction sites, and the complete DNA sequence of the multi-epitope vaccine model was successfully cloned into the pET28a(+) plasmid using SnapGene® v6.2 software (Fig. 6).

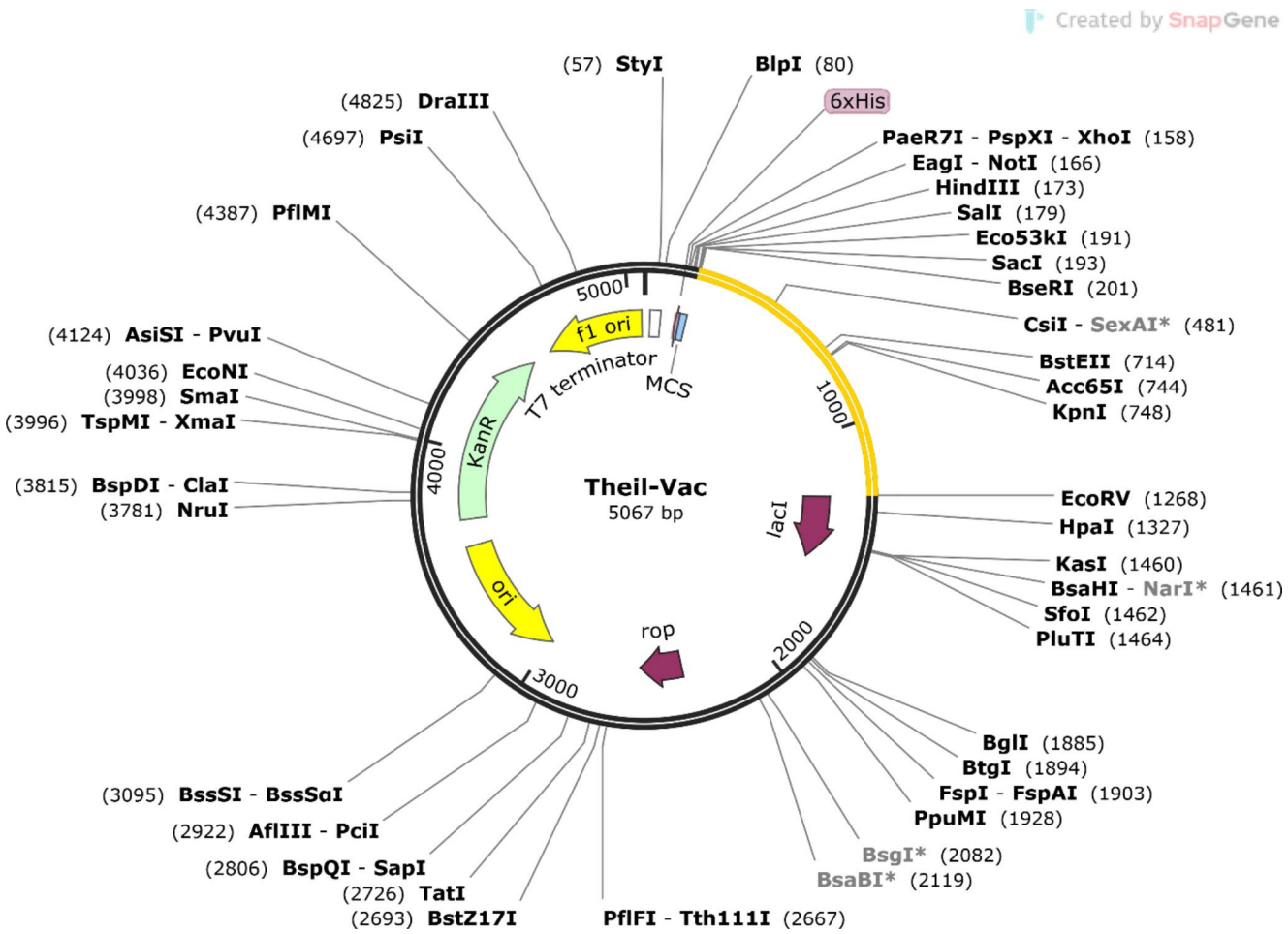

Fig. 6 Illustration of the *in silico* cloning process of the final vaccine DNA sequence into the pET28a(+) vector, using *Eco53KI* and *EcoRV* restriction sites

Discussion

Tropical theileriosis significantly affects the cattle production industry in tropics and subtropical regions of the world (Sivakumar et al. 2014). The disease is difficult to control because of insidious tick vectors in the life cycle and drug resistance phenomenon in both ticks and *Theileria* parasites (Gharbi et al. 2020). Therefore, the application of vaccination strategies directed towards ticks and parasitic agents seems rational to harness the disease effectively (Agina et al. 2020). Live vaccination strategy is a principal process to prevent *Theileria* infections in cattle herds. This procedure was initiated in Israel during the 1970s using *Theileria*-infected cell lines, which was further developed in several other countries such as Iran, Tunisia, India and Turkey using native *Theileria* isolates (Darghouth et al. 1996; Hashemi-Fesharki 1988, 1998; Morrison and McKeever 2006). However, prolonged cell-culture passage, cryopreservation, dose-dependent effects and the risk of disease development by vaccine strains are the major concerns for live vaccines (Morrison and McKeever 2006). These practical disadvantages force us to improve the vaccine design approaches to control theileriosis more successfully. Over the past few decades, vaccine development technologies have changed so that characterization of antigenic fragments and/or epitopic regions in a protein that could fruitfully stimulate the desired immune responses is the major challenge in the post-genomic era. In this sense, computational immunoinformatics approaches could assist us in virtually analyzing the virulence factors of a given pathogen, e.g. a protein compound, for a rational vaccine design in a cost- and time-effective manner (Parvizpour et al. 2020).

It is said that cattle may survive a primary *Theileria* infection, with or without treatment; this fact may suggest only low-levels of immunological control of infection at sporozoite stage, after further challenge with sporozoites (Musoke et al. 1984). Nevertheless, it is obvious that the frequent challenge using sporozoites would result in the induction of neutralizing antibodies (Musoke et al. 1982). Hence, the investigation of sporozoite-stage antigens deserves

further attention to develop of novel vaccines. *Theileria annulata* SPAG1 is a sporozoite-stage protein homologous to its counterpart in *T. parva* (p67 protein), particularly in N- and C-terminals (Hall et al. 1992; Knight et al. 1996). Previously, the protective efficacy of TaSPAG1 was mostly focused on the full-length and/or C-terminal fragment (SR1), including neutralizing epitopes, both in recombinant form (Morrison and McKeever 2006). Boulter et al. (1999) showed reduced early parasitemia, increased animal survival upon challenge, delay in macroschizont formation and partial protection following cattle vaccination using TaSPAG1 formulated with different adjuvants (Boulter et al. 1999). Another vaccination trial in calves using TaSPAG1 and attenuated *T. annulata*- infected cell-line showed a high rate of recovery in sporozoite-challenged animals vaccinated with both TaSPAG1 and infected cell-lines, showing substantial synergistic effects (Darghouth et al. 2006). Immunoblotting evidence showed that neutralizing constituents elicited by SPAG1 or p67 appear to be cross-protective against challenge with the reciprocal parasite (Hall et al. 2000). This arises from highly-conserved epitopes between two molecules, in spite of only 47% sequence similarity (Knight et al. 1996). Due to the potential capacity of TaSPAG1 to be a vaccine candidate, we further investigated its immunogenic B- and MHC-I binding epitopes in this *in silico* study.

In the first step, biochemical characteristics of the TaSPAG1 was evaluated using a set of bioinformatics web servers. This 907-residue, heavy protein had a MW of about 92 kDa, and most of its residues were negatively-charged (Asp + Glu). It is said that charged residues in a protein play a significant role in protein orientation/position (Kiefer et al. 1997; Monné et al. 1998). The charge at which the pH turns zero is called pI, which was 5.04 for this protein (Asghari et al. 2021a, b). The protein instability index (41.68) showed that the protein is only borderline unstable. Moreover, GRAVY score of about −0.400 and aliphatic index of 71.27 showed that TaSPAG1 is hydrophilic and highly-thermotolerant, respectively. Hydrophilic proteins could better interact in the water-based milieu, while thermotolerant proteins could withstand higher temperatures. In addition, this protein was proved to be highly antigenic, non-allergenic with high solubility. Understanding such preliminary chemical and biophysical properties is essential for basic wet-lab experiments. Of the predicted PTM sites, O-glycosylation was the most frequent with 138 regions, followed by phosphorylation sites with 121 regions and N-glycosylation with only three regions. Reportedly, these PTM regions are crucial in the recombinant production process of the proteins, so that eukaryotic expression systems (yeast, insect or mammalian) are more preferred than bacterial hosts (Rai and Padh 2001). The presence of a signal peptide demonstrates that a synthesized protein could be destined towards several pathways, including excretory-secretory, virulence factor or surface proteins (Antelmann et al. 2001). On this basis, a signal peptide sequence was detected with high probability in TaSPAG1. Using the NetSurfP-2.0 server, surface accessibility, secondary structure and disordered regions were predicted in the submitted protein sequence. The output showed that half of the protein is structurally disordered and surface accessible in nature. Disordered proteins are highly abundant, mostly dedicated to regulatory functions and molecular signaling. Supposedly, these regions are likely immunological targets for antibodies, hence they seem to be important in vaccination studies (MacRaild et al. 2018). Also, random coils and helices were the most frequent secondary structures within the protein. the random coil is a randomly-oriented polymer conformation bonded to nearby units (Smith et al. 1996). Together, the protein conformation is maintained and protected during molecular interactions using such internally-located structures (Shaddel et al. 2018). Based on the I-TASSER server, pair-wise structure similarity reported five models, among which the first model with the highest C-score (-1.77) was selected with a TM-score of 0.50 ± 0.15 and estimated RMSD of 13.1 ± 4.2 Å. The 3D model was further subjected to refinement and validation. According to Prosa-Web and PROCHECK analyses, the quality of the refined model was enhanced after refinement, compared to the crude model.

In general, innate and adaptive immunity play a substantial role in response to different stages (sporozoite, schizonts and merozoites) of *Theileria* infection. Despite a low-level production of antibodies against *T. annulata* sporozoites, repeated sporozoite challenge of animals is seemingly capable of increasing the levels of fully-neutralizing antibodies in the bovine circulating blood (Gray and Brown 1981). Monoclonal neutralizing antibodies have been produced by vaccinating mice using sporozoites; such antibodies recognize SPAG1 protein in *T. annulata* (Williamson et al. 1989). On this basis, we predicted linear and conformational B-cell epitopes for TaSPAG1 protein. A multi-step approach was conducted to screen linear B-cell epitopes using six web servers; three web servers were employed for the identification of shared epitopes (BCPREDS, ABCpred and SVMTriP), and another three were used for screening in terms of antigenicity, allergenicity and water solubility (VaxiJen, AllerTOP and PepCalc). The output showed five potentially antigenic, non-allergenic epitopes having good water solubility, comprising "TTTTTTSSGKPSDQ" (VaxiJen score: 1.8820), "SEGEDDDDEEEEEE" (VaxiJen score: 2.1164), "DQTSGSGSKGTE" (VaxiJen score: 1.7398), "SSASPTSPT" (VaxiJen score: 1.5176) and "GPIPSPGD" (VaxiJen score: 0.5559). Moreover, 12 conformational B-cell epitopes were predicted for this protein, and most of

the involved residues were located C-terminally. Such similar neutralizing epitopes were found in *T. parva* p67 protein, which showed to be equivalently protective, comparable to the full-length p67 protein (Musoke et al. 2005). Previously, Boulter et al. (1994) remarked that "the C-terminus of TaSPAG1 is an immunologically-relevant region that can be recognized by the bovine immune response, which could be implicated in subunit vaccine design" (BOULTER et al. 1994).

A large body of evidence demonstrates that natural killer cells, γδ and CD_8^+ T cells play a crucial role in IFN-γ production, which stimulates macrophages to produce nitric oxide, which is essential for parasite clearance (Agina et al. 2020). Also, adoptive transfer of established immunity could be done using CD_8^+, but not CD_4^+ T cells from immunized animals. This represents the central role of cytotoxic T-lymphocyte (CTL) in developing immune responses against *Theileria* agents (Kar and Srivastava 2018). Here, we predicted the potential epitopes with higher affinity to bind with bovine MHC-I molecules using the IEDB web server. In total, 8 different epitopes were shown to be strong binders, with high antigenicity scores and without allergenicity and toxicity, including "LSQTGLGPSGSH" (VaxiJen score: 0.9975) and "SQTGLGPSGSHA" (VaxiJen score: 1.1970) for **BoLA-1:00901**, "SSASPTSPTTTL" (VaxiJen score: 0.9973) for **BoLA-2:00501**, "SFQEPVSQELEF" (VaxiJen score: 0.5648) for **BoLA-4:02401**, "KLLGSGFEVASI" (VaxiJen score: 0.9138) for **BoLA-5:00301**, "TSPTTTLSQTGL" (VaxiJen score: 0.5101) for **BoLA-AW10**, "VTLESAVTQPSK" (VaxiJen score: 0.7541) and "SAVTQPSKDPFK" (VaxiJen score: 0.9265) for **BoLA-T2a**. While predicting several BoLA molecules in the IEDB server, an interesting finding was that about all CTL epitopes were determined N-terminally, between 1 and 55 residues, mostly between 25 and 50 residues.

In the final step of this study, a 355-residue, 31.25 kDa multi-epitope vaccine candidate was designed using highly antigenic, non-allergenic B-cell and bovine MHC-I-binding epitopes connected together with appropriate adjuvants and *M. tuberculosis* 50 S ribosomal protein L7/L12 as a genetic adjuvant. This vaccine model was shown to be highly antigenic (score: 0.7871), hydrophilic (GRAVY: -0.203), stable (index: 38.51), soluble (0.802) and thermophilic (64) in nature, without allergenic traits. Numerous coils and helices existed in the vaccine sequence, as shown by NetSurfP-2.0 server, and the best predicted 3D model by I-TASSER server possessed a C-score of -2.31, which further refined by GalaxyRefine server, then validated by Z-score (Prosa-web server) and Ramachandran plot (PROCHECK tool) comparison between the crude and refined models. Similar to the naïve TaSPAG1 protein, several conformational B-cell epitopes were predicted in the vaccine construct, being critical for antigen-antibody interactions. The ClusPro 2.0 output, also, demonstrated that significant interactions existed between cattle MHC-I molecule and the designed multi-epitope vaccine candidate, particularly between the vaccine and chain A of MHC molecule. Finally, the vaccine candidate was codon-adapted for *E. coli* K12 strain and successfully cloned into the pET28a(+) vector for efficient protein expression.

Conclusion

Due to the importance of tropical theileriosis in the tropical and subtropical regions of the globe and its impact on the cattle industry, preventive measures such as vaccination seems necessary. Next-generation vaccine design using strictly-screened, highly-antigenic epitopic fragments of known *T. annulata* antigens in the context of unprecedented immunization platforms provides novel insights into the vaccination against tropical theileriosis. The present *in silico* study highlighted the most important biophysical characteristics and novel B-cell and MHC-I binding epitopes of TaSPAG1 protein using a set of immunoinformatics servers. It was found that conformational antibody determinants are mostly located C-terminally, whereas CD_8^+ epitopes were located at the N-termini of the protein. As well, a remarkable molecular interaction between the TaSPAG1 protein and cattle MHC class I molecule was estimated. Altogether, the information provided here could be further directed towards immunization studies alone or combined with other immunodominant *T. annulata* antigens.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s40203-023-00153-5.

Acknowledgements Not applicable.

Author contributions M. Azimi-Resketi and M. Shams conceived the study protocol; M. Akbari, S. Heydarian, A. Eftekhari performed the bioinformatics analyses; J. Balali and D. Sargazi performed the molecular docking process; M. Azimi-Resketi and M. Akbari drafted the manuscript. M. Shams supervised the manuscript. All authors read and approved the final version of the manuscript.

Funding This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Data Availability Data are available from the corresponding author on reasonable request.

Declarations

Competing interests The authors declare no competing interests.

Ethical approval Not applicable.

Consent to participate Not applicable.

Consent for publication Not applicable.

References

Agina OA, Shaari MR, Isa NMM, Ajat M, Zamri-Saad M, Hamzah H (2020) Clinical pathology, immunopathology and advanced vaccine technology in bovine theileriosis: a review. Pathogens 9:697. https://doi.org/10.3390/pathogens9090697

Antelmann H, Tjalsma H, Voigt B, Ohlmeier S, Bron S, van Dijl JM, Hecker M (2001) A proteomic view on genome-based signal peptide predictions. Genome Res 11:1484–1502. https://doi.org/10.1101/gr.182801

Asghari A, Majidiani H, Fatollahzadeh M, Nemati T, Shams M, Azizi E (2021a) Insights into the biochemical features and immunogenic epitopes of common bradyzoite markers of the ubiquitous Toxoplasma gondii infection. Genet Evol 105037. https://doi.org/10.1016/j.meegid.2021.105037

Asghari A, Nourmohammadi H, Majidiani H, Shariatzadeh A, Shams M, Montazeri F (2021b) In silico analysis and prediction of immunogenic epitopes for pre-erythrocytic proteins of the deadly Plasmodium falciparum infection. Genet Evol 104985. https://doi.org/10.1016/j.meegid.2021.104985

Asghari A et al (2021d) Development of a chimeric vaccine candidate based on Toxoplasma gondii major surface antigen 1 and apicoplast proteins using comprehensive immunoinformatics approaches european. J Pharm Sci 105837. https://doi.org/10.1016/j.ejps.2021.105837

Asghari A, Kordi B, Maleki B, Majidiani H, Shams M, Naserifar R (2022) Neospora caninum SRS2 protein: essential vaccination targets and biochemical features for Next-Generation Vaccine Design BioMed. https://doi.org/10.1155/2022/7070144. Research International 2022:7070144

Bishop R, Musoke A, Morzaria S, Gardner M, Nene V (2004) Theileria: intracellular protozoan parasites of wild and domestic ruminants transmitted by ixodid ticks parasitology. 129:S271–S283. https://doi.org/10.1017/S0031182003004748

Boulter N et al (1994) Theileria annulata sporozoite surface antigen (SPAG-1) contains neutralizing determinants in the C terminus. Parasite Immunol 16:97–104. https://doi.org/10.1111/j.1365-3024.1994.tb00328.x

Boulter N et al (1999) Evaluation of recombinant sporozoite antigen SPAG-1 as a vaccine candidate against Theileria annulata by the use of different delivery systems. Tropical Med Int Health 4:A71–A77. https://doi.org/10.1046/j.1365-3156.1999.00453.x

Consortium U (2015) UniProt: a hub for protein information nucleic acids research 43:D204-D212. https://doi.org/10.1093/nar/gku989

Darghouth M, Miled LB, Bouattour A, Melrose T, Brown C, Kilani M (1996) A preliminary study on the attenuation of tunisian schizont-infected cell lines of. Theileria annulata Parasitology research 82:647–655. https://doi.org/10.1007/s004360050179

Darghouth MA, Boulter NR, Gharbi M, Sassi L, Tait A, Hall R (2006) Vaccination of calves with an attenuated cell line of Theileria annulata and the sporozoite antigen SPAG-1 produces a synergistic effect Veterinary parasitology. 142:54–62. https://doi.org/10.1016/j.vetpar.2006.06.010

Dimitrov I, Bangov I, Flower DR, Doytchinova I (2014a) AllerTOP v. 2—a server for in silico prediction of allergens. J Mol Model 20:2278. https://doi.org/10.1007/s00894-014-2278-5

Dimitrov I, Naneva L, Doytchinova I, Bangov I (2014b) AllergenFP: allergenicity prediction by. Descr fingerprints Bioinf 30:846–851. https://doi.org/10.1093/bioinformatics/btt619

Dobbelaere D, Heussler V (1999) Transformation of leukocytes by Theileria parva and T. annulata Annual Reviews in Microbiology 53:1–42. https://doi.org/10.1146/annurev.micro.53.1.1

Doytchinova IA, Flower DR (2007) VaxiJen: a server for prediction of protective antigens, tumour antigens and subunit vaccines. BMC Bioinformatics 8:1–7. https://doi.org/10.1186/1471-2105-8-4

El-manzalawy Y, Dobbs D, Honavar V (2008) Predicting flexible length linear B-cell epitopes. Computational Systems Bioinformatics: (volume 7). Imperial College Press, USA, pp 121–132. https://doi.org/10.1142/9781848162648_0011

Gasteiger E, Hoogland C, Gattiker A, Wilkins MR, Appel RD, Bairoch A (2005) Protein identification and analysis tools on the ExPASy server. The proteomics protocols handbook. Springer, pp 571–607. https://doi.org/10.1385/1-59259-890-0:571

Gharbi M, Sassi L, Dorchies P, Darghouth MA (2006) Infection of calves with Theileria annulata in Tunisia: economic analysis and evaluation of. the potential benefit of vaccination Veterinary Parasitology 137:231–241. https://doi.org/10.1016/j.vetpar.2006.01.015

Gharbi M et al (2020) Current status of tropical theileriosis in Northern Africa: a review of recent epidemiological investigations and implications for control Transboundary and emerging diseases. 67:8–25. https://doi.org/10.1111/tbed.13312

Gray M, Brown C (1981) InVitro neutralization of Theilerial Sporozoite Infectivity with Immune serum. Advances in the control of Theileriosis. Springer, pp 127–131. https://doi.org/10.1007/978-94-009-8346-5_22

Hall R, Hunt PD, Carrington M, Simmons D, Williamson S, Mecham RP, Tait A (1992) Mimicry of elastin repetitive motifs by Theileria annulata sporozoite surface antigen Molecular and biochemical parasitology. 53:105–112. https://doi.org/10.1016/0166-6851(92)90012-9

Hall R et al (2000) Reciprocal cross-protection induced by sporozoite antigens SPAG-1 from Theileria annulata and p67 from Theileria parva. Parasite Immunol 22:223–230. https://doi.org/10.1046/j.1365-3024.2000.00302.x

Hashemi-Fesharki R (1988) Control of Theileria annulata. Iran Parasitol Today 4:36–40. https://doi.org/10.1016/0169-4758(88)90062-2

Hashemi-Fesharki R (1998) Recent development in control of Theileria annulata in Iran Parasite. 5:193–196. https://doi.org/10.1051/parasite/1998052193

Hebditch M, Carballo-Amador MA, Charonis S, Curtis R, Warwicker J (2017) Protein–Sol: a web tool for predicting protein solubility. Seq Bioinf 33:3098–3100. https://doi.org/10.1093/bioinformatics/btx345

Heo L, Park H, Seok C (2013) GalaxyRefine: protein structure refinement driven by side-chain repacking nucleic acids res 41:W384-W388. https://doi.org/10.1093/nar/gkt458

Irvin A, Morrison WI (1987) Immunopathology, immunology, and immunoprophylaxis of Theileria infections Immune responses in parasitic infections: immunology, immunopathology, and immunoprophylaxis Volume III: Protozoa:223–274

Kallerup RS, Foged C (2015) Classification of vaccines. Subunit vaccine delivery. Springer, pp 15–29. https://doi.org/10.1007/978-1-4939-1417-3_2

Kar PP, Srivastava A (2018) Immuno-informatics analysis to identify novel vaccine candidates and design of a multi-epitope based vaccine candidate against Theileria parasites frontiers in immunology 9:2213. https://doi.org/10.3389/fimmu.2018.02213

Kiefer D, Hu X, Dalbey R, Kuhn A (1997) Negatively charged amino acid residues play an active role in orienting the Sec-independent Pf3 coat protein in the Escherichia coli inner membrane. EMBO J 16:2197–2204. https://doi.org/10.1093/emboj/16.9.2197

Klausen MS et al (2019) NetSurfP-2.0: Improved prediction of protein structural features by integrated deep learning Proteins: Structure, Function, and Bioinformatics 87:520–527. https://doi.org/10.1002/prot.25674

Knight P et al (1996) Conservation of neutralizing determinants between the Sporozoite Surface Antigens ofTheileria annulata-andTheileria. parva Experimental parasitology 82:229–241. https://doi.org/10.1006/expr.1996.0030

Laskowski RA, MacArthur MW, Moss DS, Thornton JM (1993) PROCHECK: a program to check the stereochemical quality of protein structures. J Appl Crystallogr 26:283–291. https://doi.org/10.1107/S0021889892009944

MacRaild CA, Seow J, Das SC, Norton RS (2018) Disordered epitopes as peptide vaccines. Pept Sci 110:e24067. https://doi.org/10.1002/pep2.24067

Magnan CN, Zeller M, Kayala MA, Vigil A, Randall A, Felgner PL, Baldi P (2010) High-throughput prediction of protein antigenicity using protein. microarray data Bioinformatics 26:2936–2943. https://doi.org/10.1093/bioinformatics/btq551

Majidiani H, Nabavi R, Ganjali M, Saadati D (2016) Detection of Theileria annulata carriers in holstein–friesian (Bos taurus taurus) and Sistani (Bos taurus indicus) cattle breeds by polymerase chain reaction in Sistan region. Iran J Parasitic Dis 40:1184–1188. https://doi.org/10.1007/s12639-015-0646-4

Maritim A, Young A, Lesan A, Ndungu S, Mutugi J, Stagg D (1989) Theilerial parasites isolated from carrier cattle after immunization with Theileria parva by the infection and treatment method parasitology. 99:139–147. https://doi.org/10.1017/S0031182000061126

McHardy N, Wekbsa L, Hudson A, Randall A (1985) Antitheilerial activity of BW720C (buparvaquone): a comparison with parvaquone. Res Vet Sci 39:29–33. https://doi.org/10.1016/S0034-5288(18)31765-X

Mhadhbi M, Naouach A, Boumiza A, Chaabani MF, BenAbderazzak S, Darghouth MA (2010) In vivo evidence for the resistance of Theileria annulata to buparvaquone veterinary parasitology. 169:241–247. https://doi.org/10.1016/j.vetpar.2010.01.013

Monné M, Nilsson I, Johansson M, Elmhed N, Von Heijne G (1998) Positively and negatively charged residues have different effects on the position in the membrane of a model transmembrane helix. J Mol Biol 284:1177–1183. https://doi.org/10.1006/jmbi.1998.2218

Morrison W, McKeever D (2006) Current status of vaccine development against. Theileria parasites Parasitology 133:S169–S187. https://doi.org/10.1017/S0031182006001867

Musoke A, Nantulya V, Buscher G, Masake R, Otim B (1982) Bovine immune response to Theileria parva: neutralizing antibodies to sporozoites Immunology 45:663

Musoke A, Nantulya V, Rurangirwa F, Buscher G (1984) Evidence for a common protective antigenic determinant on sporozoites of several Theileria. parva strains Immunology 52:231

Musoke A et al (2005) Subunit vaccine based on the p67 major surface protein of Theileria parva sporozoites reduces severity of infection derived from field tick challenge. 23:3084–3095. https://doi.org/10.1016/j.vaccine.2004.09.039

Nene V, Morrison WI (2016) Approaches to vaccination against Theileria parva and theileria annulata. Parasite Immunol 38:724–734. https://doi.org/10.1111/pim.12388

Norval R, Lawrence J, Young A, Perry BD, Dolan T, Scott J (1991) Theileria parva: influence of vector, parasite and host relationships on the epidemiology of theileriosis in southern Africa Parasitology. 102:347–356. https://doi.org/10.1017/S0031182000064295

Parvizpour S, Pourseif MM, Razmara J, Rafi MA, Omidi Y (2020) Epitope-based vaccine design: a comprehensive overview of bioinformatics approaches. Drug Discovery Today 25:1034–1042. https://doi.org/10.1016/j.drudis.2020.03.006

Pipano E, Tsur I (1966) Experimental immunization against Theileria annulata with a tissue culture vaccine Refuah vet 23:186–194

Ponomarenko J, Bui H-H, Li W, Fusseder N, Bourne PE, Sette A, Peters B (2008) ElliPro: a new structure-based tool for the prediction of antibody epitopes BMC bioinformatics 9:514. https://doi.org/10.1186/1471-2105-9-514

Rai M, Padh H (2001) Expression systems for production of heterologous proteins current science:1121–1128

Saha S, Raghava GPS (2006) Prediction of continuous B-cell epitopes in an antigen using recurrent neural network Proteins: Structure, Function, and Bioinformatics 65:40–48. https://doi.org/10.1002/prot.21078

Shaddel M, Ebrahimi M, Tabandeh MR (2018) Bioinformatics analysis of single and multi-hybrid epitopes of GRA-1, GRA-4, GRA-6 and GRA-7 proteins to improve DNA vaccine design against Toxoplasma gondii. J Parasitic Dis 42:269–276. https://doi.org/10.1007/s12639-018-0996-9

Shams M et al (2021a) Bioinformatics features and immunogenic epitopes of Echinococcus granulosus Myophilin as a promising target for vaccination against cystic echinococcosis infection. Genet Evol 89:104714. https://doi.org/10.1016/j.meegid.2021.104714

Shams M, Nourmohammadi H, Basati G, Adhami G, Majidiani H, Azizi E (2021b) Leishmanolysin gp63: Bioinformatics evidences of immunogenic epitopes in Leishmania major for enhanced vaccine design against zoonotic cutaneous leishmaniasis Informatics. Med Unlocked 24:100626. https://doi.org/10.1016/j.imu.2021.100626

Sharma N, Patiyal S, Dhall A, Pande A, Arora C, Raghava GP (2020) AlgPred 2.0: an improved method for predicting allergenic proteins and mapping of IgE epitopes briefings in Bioinformatics. https://doi.org/10.1093/bib/bbaa294

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

Smith LJ, Fiebig KM, Schwalbe H, Dobson CM (1996) The concept of a random coil: residual structure in peptides and denatured proteins folding and design 1. https://doi.org/10.1016/S1359-0278(96)00046-6. :R95-R106

Stuen S (2020) Haemoparasites—Challenging and Wasting Infections in small ruminants. Rev Anim 10:2179. https://doi.org/10.3390/ani10112179

Vita R et al (2019) The immune epitope database (IEDB): 2018 update nucleic acids research 47:D339-D343. https://doi.org/10.1093/nar/gky1006

Wiederstein M, Sippl MJ (2007) ProSA-web: interactive web service for the recognition of errors in three-dimensional structures of proteins nucleic acids research. 35:W407–W410. https://doi.org/10.1093/nar/gkm290

Williamson S et al (1989) Theileria annulata sporozoite surface antigen expressed in Escherichia coli elicits neutralizing antibody Proceedings of the National Academy of Sciences 86:4639–4643. https://doi.org/10.1073/pnas.86.12.4639

Yang J, Yan R, Roy A, Xu D, Poisson J, Zhang Y (2015) The I-TASSER suite: protein structure and function prediction. Nat Methods 12:7–8. https://doi.org/10.1038/nmeth.3213

Yao B, Zhang L, Liang S, Zhang C (2012) SVMTriP: a method to predict antigenic epitopes using support vector machine to integrate tri-peptide similarity and propensity. PLoS ONE 7:e45152. https://doi.org/10.1371/journal.pone.0045152