Appl Biochem Biotechnol (2015) 175:323–339
DOI 10.1007/s12010-014-1245-z

Designing and Modeling of Complex DNA Vaccine Based on Tropomyosin Protein of *Boophilus* Genus Tick

Mohamamd Mahdi Ranjbar · Shishir K. Gupta · Khodayar Ghorban · Sedigheh Nabian · Alireza Sazmand · Mohammad Taheri · Sahar Esfandyari · Maryam Taheri

Received: 3 May 2014 / Accepted: 10 September 2014 /
Published online: 1 October 2014

Abstract *Boophilus* tick is a bloodsucking ectoparasite that transfers some pathogens, reducing production and thus leading to economical losses in the cattle industry. Tropomyosin (TPM) protein is a salivary protein, has actin regulator activity, and plays an important role in immune reactions against parasites. In the current study, besides developing a safe, effective, and broad spectrum protective measure against *Boophilus* genus tick based on TPM protein, we attempted to minimize possible problems occurring in the design of polytopic vaccines. Briefly, the steps that were followed in the present study were as follows: retrieving sequences and finding the mutational/conservative regions, selecting consensus and high immunogenic

M. M. Ranjbar · S. Esfandyari
Department of Immunology, Faculty of VeterinaryMedicine, University of Tehran, Tehran, Iran

S. K. Gupta
Department of Bioinformatics, Biocenter, Am Hubland, University of Würzburg, 97074 Würzburg, Germany

K. Ghorban
Department of Immunology, School of Medicine, AJA University of Medical Sciences, Tehran, Iran

S. Nabian (✉) · M. Taheri
Department of Parasitology, Faculty of Veterinary Medicine, University of Tehran, Tehran, Iran
e-mail: nabian@ut.ac.ir

S. Nabian
Tick and Tick-Borne Diseases Center, Faculty of Veterinary Medicine, University of Tehran, Tehran, Iran

A. Sazmand
Department of Agriculture, Payame Noor University, Tehran, Iran

M. Taheri
Department of Chemistry, Shahid Beheshti University, Tehran, Iran

Present Address:
A. Sazmand
Department of Pathobiology, Institute of Parasitology, University of Veterinary Medicine Vienna (Vetmeduni Vienna), Veterinärplatz 1, 1210 Vienna, Austria

epitopes of B and $CD4^+$ T cells by different approaches, three-dimensional structure (3D structure) prediction and representation of epitopes and highly variable/conserve regions, designing vaccinal construct by fusion of B and T cell epitopes by special patterns and improving immunogenicity, evaluation of the constructs' primary structure and posttranslational modification, calculation of hydrophobic regions, reverse translation, codon optimization, open reading frame checking, insertion of start/end codon, Kozak sequence, and finally constructing the DNA vaccine. Variation plot showed some shared epitopes among the ticks' and mites' species that some might be effective only in some species. Finally, by following the steps mentioned above, two constructs for B and T cells were achieved. Checking constructs revealed their reliability and efficacy for in vitro production and utilization. Successful in silico modeling is an essential step of designing vigorous vaccines. We developed a novel protective and therapeutic vaccine against *Boophilus* genus (based on TPM protein). At the next step, constructed DNA vaccine would be produced in vitro and administrated to cattle, and its potency to induction of immune response and protection against *Boophilus* genus as well as other ticks and mites will be evaluated.

Keyword *Boophilus* · Tropomyosin · DNA vaccine · Epitope · In silico

Introduction

Ticks are bloodsucking ectoparasites of human and animals. Tick paralysis, weight loss, anemia, damage to skin and leather, and reduction in animal production are the effects of host tick infestation. They can also transmit viral, bacterial, and protozoan pathogens [1, 2].

Exposure of some cattle breeds to *Boophilus* (syn. *Rhipicephalus*) *microplus* results in partial protective immunity. It has been shown that inoculating extracts of salivary glands, intestine, interior organs, and cement materials of ticks to hosts induce a strong immunity against some genera of ticks [3]. There are also several lines of evidence of immunity induction against ticks using protein antigens. Vaccines such as Bm86 and Bm95 are available in the market for the partial control of *B. microplus*; however, producing more effective vaccines needs exploration of new antigens and making novel recombinants [3]. Using Bm86 protein peptides' parts as a peptidal vaccine has been studied [4, 5].

Tropomyosin (TPM) protein is among the antigens present in the ticks' saliva. This protein has interspecies conservation, binder, and actin regulator properties that are important in the contraction of muscles. It has also actin regulator activity and plays an important role in immune and allergic reactions [6]. TPM has been introduced as a vaccine candidate in several parasites [7], which causes cross-reactivity due to conservation. Also, some of the epitopes of this protein are produced due to posttranslational glycosylation [8, 9]. On the other hand, immunity against this protein induces IL10 and increase in suppressor cells, so it can be used as an effective antigenic vaccine against filarial nematodes by induction of Th1 responses [9, 10]. Wright showed that injecting a mixture of recombinant proteinal paramyosin, tropomyosin, cathepsin L, cathepsin D, and a histamine-releasing factor to birds, as well as a mixture of antitropomyosin and antiparamyosin antibodies, caused an increase in the death of *Dermanyssus gallinae* mites; hence, tropomyosin and paramyosin are potential vaccine candidates [10]. Also, Huntley and colleagues showed that *Psoroptes ovis* tropomyosin produces IgG and IgE antibodies in infested sheep [11]. Finally, injecting troponin-like P27/30 recombinant protein to mouse and rat has induced immunity responses against *Haemaphysalis longicornis*; therefore, muscular proteins can be used as tick vaccines [12].

DNA vaccines are able to trigger both cellular and humoral immune responses. DNA immunization based on polyepitope-based DNA immunization is powerful in the induction of T cell responses as well as in B cells. T lymphocytes play a central role in the generation of a protective immune response in many microbial infections; hence, the identification of peptides that stimulate T cell responses is a critical requirement for the development of successful epitopic vaccines [13, 14]. Polytope vaccine codes a series of T and/or B cells which induces cytotoxic T cells (CTL), helper T cells (Th), or B to a specific epitope [15–19].

Vaccines are usually made from killed pathogens or from their reduced pathogenic forms. Desirable vaccines have properties such as safety, assured nondisease making, flexibility, improvability and ability of choosing the type of immunity, stronger immunity response, and covering a wider spectrum of pathogens. These vaccines can consist of peptides that introduce B cell linear epitopes from a pathogen protein. With immunization of animals with synthetic peptides containing linear B epitopes and $CD8^+$ T, we cause the production of an antibody against a specific protein that can be used as a prevention tool [20–22]. Epitope-based immunizations have shown significant results in animal models in clinical tests, as well as have shown prevention and treatment properties [23].

Previous researches showed that immunogenicity of recombinant proteins or polytopic vaccines is significantly lower than classical inactive or attenuated in clinical challenges [16, 24, 25] due to rapid clearance of recombinant protein, small size of synthetic peptides, and lack of suitable and effective epitopes for Th cell [26–28].

Some defined approaches for improving the immunogenicity and effectiveness of antigenic epitopes in DNA vaccines are increasing the number of antigenic epitopes (redundancy) concerning multiple Th cell epitopes and enable the insertion of antigenic epitopes into immunogenic adjuvants or a carrier protein [17, 25].

In the current study, besides designing and developing a safe and effective vaccine that could help protect cattle against the *Boophilus* genus tick based on TPM protein, we attempted to minimize possible problems that occur in designing peptide (polytopic) vaccines before, which caused low effectiveness, by modification and innovation of the designing process to create a broad spectrum vaccine that covers some ticks and mites.

Materials and Methods

Retrieving Sequences and Entropy Plot for Finding the Mutational/Conservative Regions

Complete protein sequences of TPM of *B. microplus* and TPM sequences of other ticks and mites were retrieved from the UniProtKB (http://www.expasy.org/uniprot) database. Obtained sequences were aligned, analyzed, and trimmed in BioEdit software version 7.7.9. Then, Shannon entropy values were measured to define the conserved regions since the epitopes from highly conserved regions are likely to elicit more immune responses [29, 30]. This Shannon entropy analysis measures variation at each amino acid position in the set of aligned sequences.

Prediction of B and $CD4^+$ T Epitopes

Recently, Ranjbar et al. [31] studied immunoinformatic features of the TPM protein in *B. microplus* with respect to different physicochemical profiles and machine learning approaches for B and $CD4^+$ T cells. In the study, finally, six peptides were found to have a consensus and used for in silico concatenation: $AA_{62-80}(NH_2$-

EQLSQANSKLEEKDKALQA-COOH), $AA_{170–185}$(NH_2-AMVEADLERAEERAETG-COOH), and $AA_{200–218}$(NH_2-VGNNLKSLEVSEEKALQKEET-COOH) for B cell and epitope $AA_{192–207}$(NH_2-LEEELRVVGNNLKSL-COOH), $AA_{218–233}$(ETYEMQIRQMTNRLQ), and $AA_{221–236}$(EMQIRQMTNRLQEAE) for T cell. Epitope $AA_{192–207}$ predicted for T cell was shared between some alleles, so it was selected as a major epitope.

3D Structure Prediction Using Homology Modeling and Representation of Epitopes and Highly Variable/Conserved Regions

Local alignment search using BLASTP was employed against Protein Data Bank (PDB) to select a suitable template—homologous sequence—to create a putative 3D model of *Boophilus* TPM protein [32]. Comparison of homology models with the template proteins reveals similarities which allow biochemical and biological functions to be inferred. Homology modeling was used to build a 3D model of the chosen sequence using Modeller software version 9v8 [33]. The 3D structure of protein representation of the location of epitopes and highly variable/conserved regions in this structure was visualized with CCP4mg software [34].

Designing and Modeling Vaccinal Construct, Fusion of B and T Cell Epitopes by Special Patterns, and Improving Immunogenicity

To obtain maximum yield in immunization, it is crucial to understand and respect immunology system response rules. Besides, it is important to consider placing patterns of epitopes in the right positions near each other, their antigenicity scores, and way of their composition, number of each epitope in the construct, spatial limitation, and immunodominancy of epitopes. To overcome these challenges, we incorporated two repeats of each epitope in construct, linkers, and more distances between epitopes by insertion of two more amino acids to the termini of B cell epitope so that these strategies have several benefits.

Also, in the NH_2 and COOH termini, different epitopes are considered to improve the immunogenicity of the construct. For noninterfering and immunodominancy reasons of epitopes or even the formation of a new epitope, we inserted a linker sequence of NH_2-GGSSGG-COOH. For releasing of construct from cells in the host body (signal peptide), controlling metabolism, and rapid removal of B cell construct from the body due to small size, a constant heavy chain region of cattle immunoglobulin G (scIgG) (secretory signals) was fused to B cell RE from NH_2 end to obtain RE-scIgG.

In case of $CD4^+$ T epitopes, specific motifs were inserted within the optimized protein sequences to provide target-specific cleavage in proteasomal and lysosomal degradation machineries. NH_2-KFERQ-COOH sequence was placed at both the -NH_2 and -COOH termini of each $CD4^+$ T epitope [35, 36]. Epitopes were arranged randomly in the construct except epitope $AA_{192–207}$ that was shared a epitope between some alleles. Also, one of these repeats was located in the center region of the construct. DNA vaccines have multiple proteins in-frame on a single construct. The addition of an antigenic segment—the sequences related to the adjuvant—can also be added to enhance vaccine potency and efficacy [37].

PADRE (universal T helper Pan DR epitope) is a synthetic immunogenic Th epitope, engineered by introducing anchor residues for the different DR motifs of MHC II into a polyalanine backbone [38] resulting peptide to bind a variety of DR molecules (pan DR-binding epitope) [39]. This sequence (AKFVAAWTLKAAA) was added to the C-terminal of the designed T cell construct. Challenge studies revealed that the parasite load was significantly decreased by Poly/hsp/vector vaccination in comparison to other forms of vaccines [40]. For enhancing T cell construct activity and improving immunogenicity of our DNA vaccine,

polytope fused with small heat shock protein (HSP) (Poly/hsp/pVAX1) of *Boophilus annulatus* (accession number: E0YPC0) was also added to the C-terminal of Th construct after PADRE sequence [41]. This HSP sequence contains a complete coding region of this protein with a length of 180 AA and separated by three lysine spacers (AAA) from PADRE.

Primary Structure Prediction of Constructs and Posttranslational Modification

Protein sequence statistics for B and T constructs including length, molecular weight, isoelectric point (IEP), total number of positive and negative residues, instability index, grand average hydropathicity (GRAVY), aliphatic index, and amino acids distribution was computed using the ExPASy ProtParam server (http://expasy.org/cgi-bin/protpraram). ProtParam results exhibit the physicochemical parameters of uncharacterized proteins. For posttranslational modification analysis, NetNGlyc 1.0 (Gupta et al., in preparation), NetCGlyc 1.0 [42], NetOGlyc 4.0 [43], DictyOGlyc1.1 [44], NetAcet 1.0 [45], and NetPhos 2.0 [46] were applied. The NetNGlyc server (Gupta et al., in preparation) predicts N-glycosylation sites in human proteins using artificial neural networks that examine the sequence context of Asn-Xaa-Ser/Thr sequons. NetCGlyc 1.0 [42] produces neural network predictions of C-mannosylation sites in mammalian proteins. The NetOglyc server [43] generates neural network predictions of mucin-type GalNAc O-glycosylation sites in mammalian proteins. The DictyOGlyc server [44] provides neural network predictions for GlcNAc O-glycosylation sites based on processes known for *Dictyostelium discoideum* proteins. The NetAcet 1.0 server [45] predicts substrates of N-acetyltransferase A (NatA). The method was trained on yeast data, but it obtains similar performance values on other substrates acetylated by NatA orthologs. The NetPhos 2.0 server [46] produces neural network predictions for serine, threonine, and tyrosine phosphorylation sites in eukaryotic proteins. All these services can be accessed at http://www.cbs.dtu.dk/services/.

Calculation of Hydrophobic Regions

By using the algorithm of Kyte and Doolittle [47], we evaluated the B and T cell TPM constructs for hydrophobic and hydrophilic regions. Peak and down regions represent hydrophobicity and hydrophilicity, respectively.

Reverse Translation and Codon Optimization

The B and T cell TPM protein construct was backtranslated into nucleotide sequences using backtranseq program of mEMBOSS 6.0.1 (http://www.ebi.ac.uk/Tools/st/emboss_backtranseq/). The degeneracy of the genetic code makes backtranslation potentially ambiguous because most amino acids are encoded by multiple codons. Backtranseq was limited to codon uses within the *Bos taurus* [48]. Codon optimization is a technique for higher gene expression of vectors used in DNA vaccination to achieve optimum expression of a foreign gene in the host cells' body [49]. Large numbers of C-G sequences in the messenger RNA (mRNA) can inhibit protein translation from increased formation of secondary structures [50] and, therefore, increase mRNA stability that significantly improves immune responses [51]. Codon optimization has improved the production of constructs leading to higher yields of antibodies [51]. Constructs were optimized by the codon adaptation tool server (http://www.jcat.de/Start.jsp). The optimization generates enhanced T cell response [52, 53] and antibody induction [54, 55].

Insertion of Start/End Codon and Kozak Sequence

At the -NH_2 and -COOH terminus of each construct, an initiation codon ATG and a termination codon TAA were added. Furthermore, to enhance in vitro expression and mRNA stability, the Kozak consensus sequence was inserted upstream of the constructs [56].

Prediction of Open Reading Frame (ORF)

For assurance of correct open reading frame, we used the gorf tool available at the NCBI server (http://www.ncbi.nlm.nih.gov/gorf/). This tool identifies all open reading frames using the standard or alternative genetic codes.

Constructing DNA Vaccine

We selected the commercially available vector pVAX1 approved for clinical use as an expression system to design a DNA vaccine. The plasmid backbone and the transcriptional unit are two main parts of a DNA vaccine plasmid. The transcriptional unit contains the promoter and our inserted construct. pVAX1 contains apUC *Escherichia coli* origin of replication. As the pVAX1 vector is a nonfusion vector, it requires insertion genes, so we added Kozak translation initiation sequence (ACC sequence), start codon (ATG), and a termination codon (TAA) to our optimized constructs. pVAX1 is a pUC-based vector that is designed to stimulate cellular as well as humoral immune responses [57]. Kozak sequence is a sequence flanking the initiator codon within the mRNA that helps its recognition by eukaryotic ribosomes and increases the expression level of constructs [58]; therefore, we inserted it to the upstream of our construct.

Results

Entropy Plot

Based on the entropy plot, five highly variable regions (38 and 39, 45, 65–67, 227, and 229–231) were observed along the TPM protein (Fig. 1). Among the abovementioned sites, position 229–231 was observed to be more highly variable. Besides, highly conserved TPM sequence (NH_2-KIVELEEELRVVGNNLKSLEVSEEK-COOH) in ticks and mites was observed (Residue no. 189-213). This conserved sequence was part of epitope $AA_{200–218}$ that was selected for B cell. Also, $AA_{65–67}$ located in the epitope $AA_{62–80}$ was predicted for B cell, and concerning near complete similarity of this protein in ticks [31], the mentioned high variability is very low for ticks and the result of the antropy plot reflects mite cases. Therefore, this epitope may be less effective in some genera of mites. Besides, epitope $AA_{192–207}$ for stimulation of $CD4^+$ T covers a highly conserved region along the TPM sequence. Epitopes 218–233 and 221–236 for stimulation of $CD4^+$ T cell have highly variable regions in their termini, and this case is a subject mentioned about B cell epitope ($AA_{62–80}$) before.

Representation of Epitopes and Highly Variable/Conserve Regions in the 3D Structure of TPM Protein

The predicted 3D structure of the modeled protein was visualized by PyMOL software. The positions of highly conserve/variable regions in B and T cell epitopes are depicted graphically in Fig. 2.

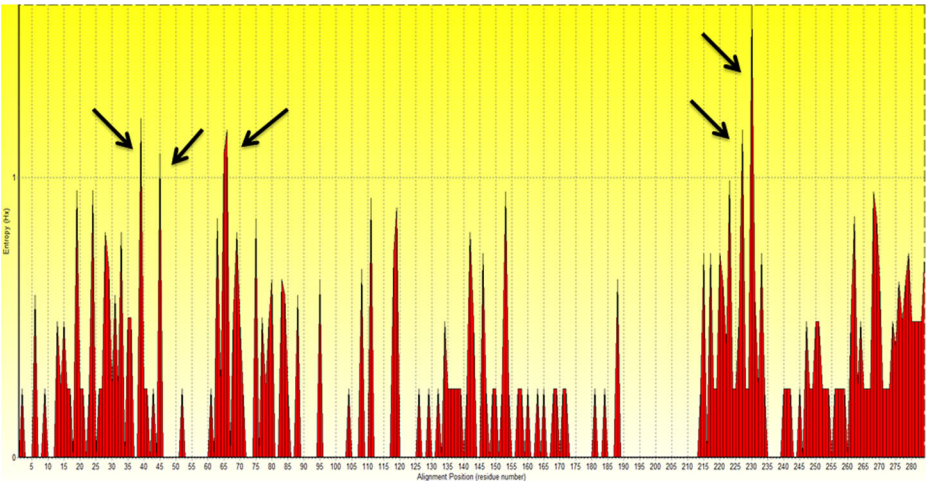


Fig. 1 Variation plot of residues along ticks' and mites' TPM sequence. Regions above threshold 1 (shown by *arrows*) are parts with high variations playing an important role in vaccine designing

B Cell Construct

The final pattern and location of selected epitopes in the B cell construct is demonstrated in Fig. 3; the 3′ terminus was fused to cattle IgG heavy chain constant region. For improving efficacy and convenient antibody bindings to epitopes, two more amino acids in each terminus of epitopes (white color) were inserted from the intact protein of TPM.

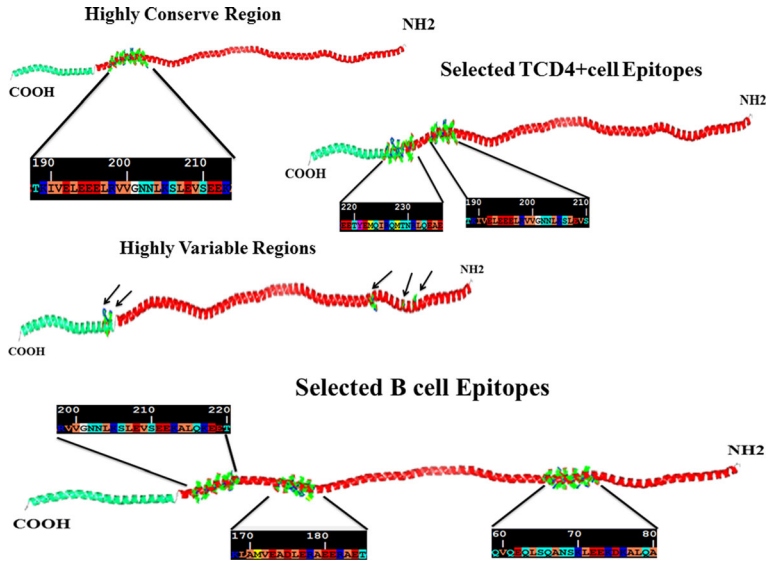
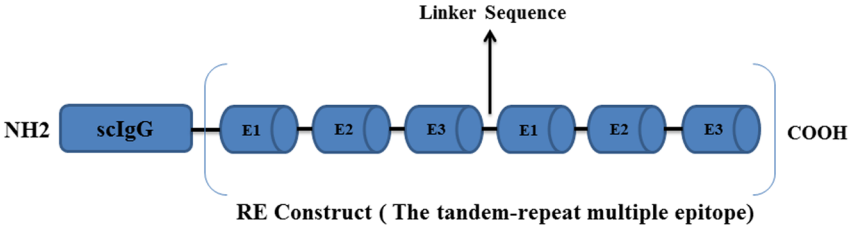


Fig. 2 Graphical representation of conserve/variable regions of B and T cell epitopes in 3D ribbon structure of TPM protein. Amine (-NH_2) and carboxylic acid (-COOH) termini of TPM protein are shown



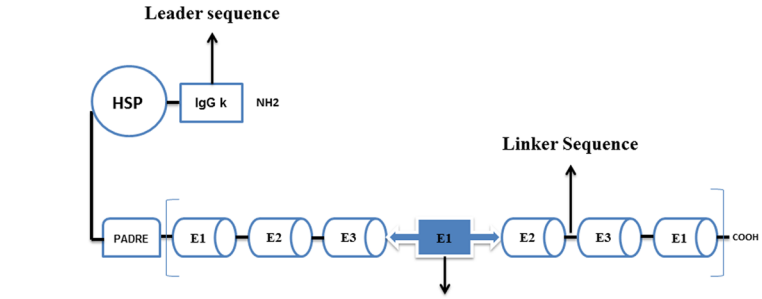
E1: Epitope AA_{62-80}
E2: Epitope $AA_{170-185}$
E3: Epitope $AA_{200-218}$
scIgG: Cattle immunoglobulin G heavy-chain constant region

VQEQLSQANSKLEEKDKALQAAEGGSSGGKLAMVEADLERAEERAETGETGGSSGG
RVVGNNLKSLEVSEEKALQKEETYEGGSSGGVQEQLSQANSKLEEKDKALQAAEGGS
SGGKLAMVEADLERAEERAETGETGGSSGGRVVGNNLKSLEVSEEKALQKEETYE

Fig. 3 Schematic diagram depicting the modeled construct for B cell. Sequence of scIgG and its linker is also shown below the construct where the *black color* shows the epitopes' sequences, *green* the linker, and *white* the inserted amino acids

T Cell Construct

Special pattern and redundancy used for fusion of peptide is illustrated in Fig. 4. E1, E2, and E3 are epitopes that were linked to each other by a linker. PADRE and HSP improving immunogenicity and IgG k work as the Leader sequence (signal peptide).



E1: Epitope $AA_{192-207}$
E2: Epitope $AA_{218-233}$
E3 : Epitope $AA_{221-236}$
PADRE : universal T helper pan DR epitope sequence
HSP : Small heat shock protein (HSP) of *Boophilus annulatus*
IgG k : leader sequence used as a signal peptide

KFERQLEEELRVVGNNLKSLKFERQETYEMQIRQMTNRLQKFERQEMQIRQMTNRLQEA
EKFERQLEEELRVVGNNLKSLKFERQETYEMQIRQMTNRLQKFERQEMQIRQMTNRLQE
AKFERQLEEELRVVGNNLKSLKFERQ

Fig. 4 Schematic illustration for the designed and modeled construct for T cell TPM constructs. In the shown sequence, *black letters* show amino sequence of the construct and *green letters* represent linkers. It should be noticed that PADRE to HSP and HPS to IgG k are linked together with a three alanine linker (AAA)

Evaluation of the Primary Structure of Designed Constructs and Posttranslational Modifications

In this study, the primary structure of B and T cell TPM construct was predicted using the ExPASy ProtParam server (http://expasy.org/cgi-bin/protparam). The input sequence and the obtained results are shown in Table 1. The average length of the protein sequence and the molecular weight of the constructs are mentioned as well. IEP is the pH point that the surface of protein is covered with charge, but the net charge of protein is zero. IEP is important for the evaluation of solubility and the mobility in an electric field. The calculated IEPs were computed to be 9.47 and 4.51 for T and B cell TPM constructs, respectively. The computed value for T cell TPM construct is more than 7 indicating that the protein is basic, but for the B cell TPM construct, it is lower than 7, therefore acidic in nature. The high aliphatic index of 67.64 (T construct) and 65.12 (B construct) indicates that these constructs are stable for a wide range of temperature, while the instability index (45.47 and 43.57) provides the estimation of the stability of protein in vitro, and the results classified the constructs with moderate stability. The GRAVY values were negative (−1.317 for T construct and −0.925 for B construct), which indicate hydrophilicity and better interaction of the constructs with the surrounding water molecules. Also, physicochemical analysis revealed that the most abundant amino acid residues were glutamic acid and glycine.

Also, we investigated the presence and patterns of posttranslational modification in constructs after their expression in mammal cells to improve our understanding of changes undergoing in the host's body (Table 2). Machine learning methods showed promising results, and our constructs are mostly free of much posttranslational modification that could be effective on immunogenicity. Besides, predicted phosphorylation sites by the NetPhos 2.0 server are depicted in Fig. 5.

Experiments provided lines of evidence that glycosylated proteins may be useful as tick vaccine antigens enhancing the protective capacity of tick vaccine, and this fact was demonstrated in recombinant Bm86 and Bm95 proteins as the only commercially available tick vaccines [3, 59–62]. The GalNAc O-glycosylation was only the type of glycosylation posttranslational modification that was predicted in constructs.

Hydrophobicity is one of the major physicochemical characteristics of epitopes, and besides, it is an important challenge in cloning and expression of concerned constructs and helps to evaluate the effectiveness of vaccines. To explore hydrophobic behavior, we used BioEdit software version 7.7.9, and the results are shown in Fig. 6.

Also, the B and T cell TPM construct had a 100 % chance of solubility using logistic regression when overexpressed in *E. coli*. The logistic regression model is a machine learning

Table 1 Calculated parameters for constructs by using the ExPASy ProtParam

Parameters	T cell construct	B cell construct
Number of amino acids	144	168
Mol. wt.	18,048.6	17,520.8
p*I*	9.47	4.51
Overall −R and +R	30, 26	20, 36
Inst.	45.47	43.57
GRAVY, AI	−1.317, 67.64	−0.925, 65.12
Amino acid frequency	Glutamic acid, glycine, and arginine	Glutamic acid, glycine, and alanine-serine

Mol. wt. molecular weight, *pI* theoretical isoelectric point, *−R* number of negative-charged residues (Arg+Lys), *+R* number of positive-charged residues (Asp+Glu), *EC* extinction coefficient at 280 nm, *Inst. II* instability index, *AI* aliphatic index, *GRAVY* grand average hydropathicity

Table 2 Posttranslational modification (PTM) results

PTM type	Construct	Number	Sites/positions
N-glycosylation sites (by NetNGlyc 1.0 server)	B and T cell TPM construct	No	No
GalNAc O-glycosylation sites (by NetOGlyc 4.0 server)	B and T cell TPM construct	2	B cell TPM construct : residue 50 and 54 T cell TPM construct: residue 113
C-mannosylation (by NetCGlyc 1.0 server)	B and T cell TPM construct	No	No
GlcNAc O-glycosylation sites by (DictyOGlyc 1.1 Prediction)	B and T cell TPM construct	No	No
Acetylation sites (by NetAcet 1.0 server)	B and T cell TPM construct	No	No
Phosphorylation sites (NetPhos 2.0 server)	B and T cell TPM construct	15 for B cell and 9 for T cell	B cell construct: Ser: 13, Thr: 0, Tyr: 2 T cell construct: Ser: 3, Thr: 4, Tyr: 2

approach which uses parameters such as molecular weight, amino acid fractions, aliphatic index, alpha-helix propensity, beta-sheet propensity, average p*I*, approximate charge average, and hydrophilicity index for prediction of solubility of recombinant protein [63].

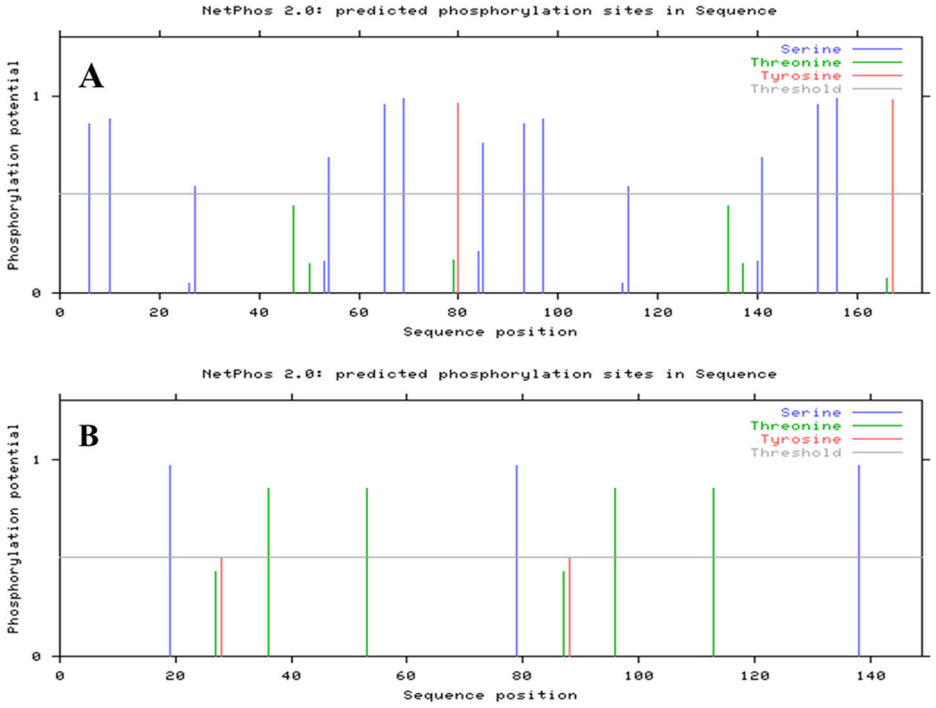


Fig. 5 Prediction of phosphorylation sites in B cell (**a**) and T cell (**b**) constructs. *Each color* represents one phosphorylation site and also related to one amino acid

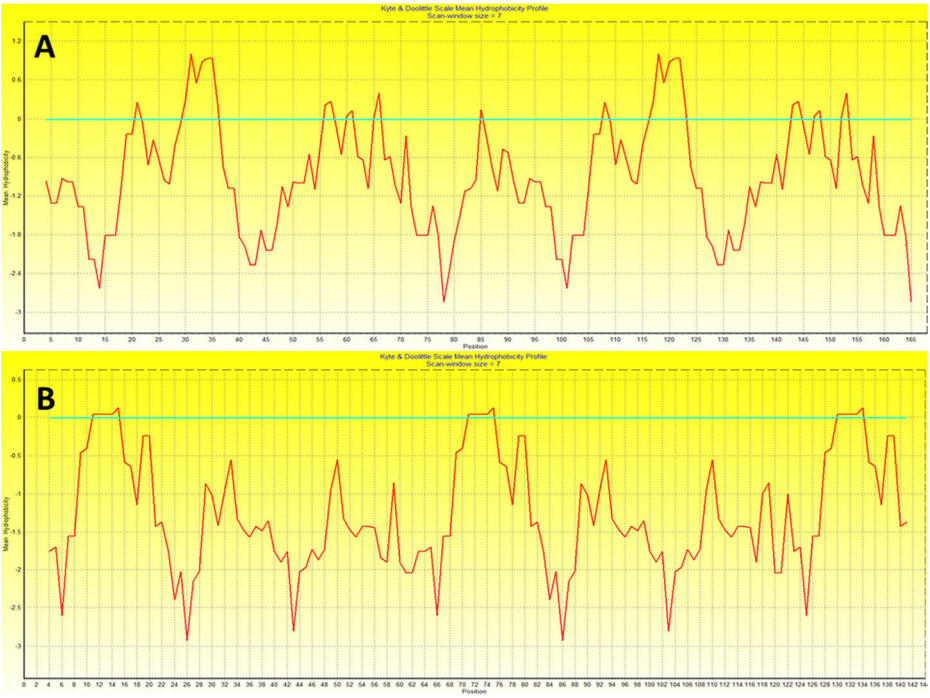


Fig. 6 Hydrophobic profile diagram by using the algorithm of Kyte and Doolittle for TPM constructs. Size of window was 7. In diagram **a** (B cell construct) and **b** (T cell construct), well regions represent hydrophilicity and are antigenic regions. Regions above threshold (0) are predicted to be hydrophobic regions. Usually hydrophobic and hydrophilic regions are near to each other

Reverse Translation and Codon Optimization

Expression is typically higher if the gene is codon optimized to match that of the target organism. Protein sequences of constructs were reverse translated into nucleotide sequences using the backtranseq program. The codon frequency table for cattle was used to replace each amino acid of input sequence with the corresponding most frequently occurring codon [48]. Degeneracy of the genetic code makes backtranslation potentially ambiguous as most amino acids are encoded by multiple codons. From different ways of optimizing the efficiency of a DNA vaccine, one of the most effective is codon optimization which can enhance the efficiency of gene expression. It helps to achieve optimum expression of a foreign gene in the host's cell system [49]. All codons in sequences of B and T cell TPM construct were adapted to cattle codon usage to evaluate the optimal codon for each amino acid encoded by the stretch of DNA. The codons were replaced with the most optimal codon according to the codon usage tables exploited from the Kazusa Codon Usage Database (http://kazusa.or.jp/codon). A graphical view of codon optimization before and after optimization is shown in Figs. 7 and 8.

Final Constructs and ORF Checking

A schematic view of the designed constructs is shown in Fig. 9. ORF examination revealed no errors, hence showing the possible optimal expression of the construct.

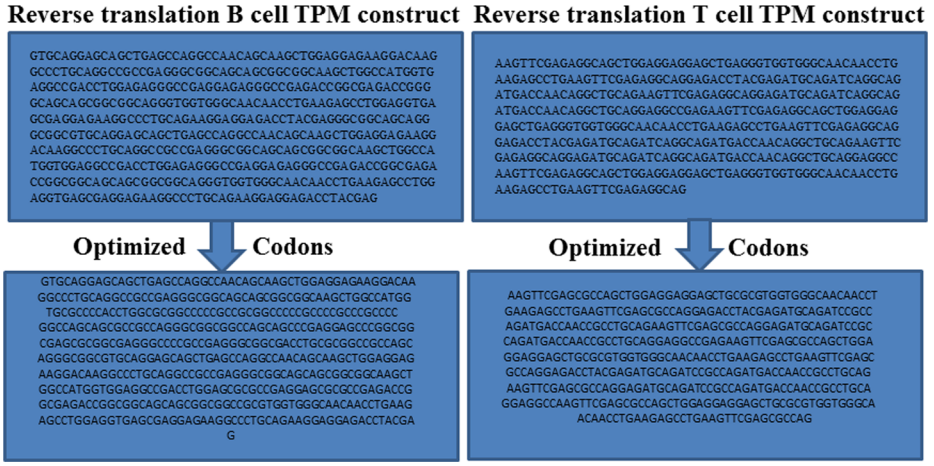


Fig. 7 Reverse translation and codon optimization of constructs in schematic view. GC content of the B cell construct sequence (antigenic part) before and after adaption was evaluated to be 70 and 76.4, respectively. For T cell sequence, GC content before and after adaption was 46.5 and 62.2, respectively

Overall Considerations and Steps Followed in This Study

A flowchart for the creation of new synthetic vaccinal construct base on TPM protein, combining rationale design with immunoinformatics is presented in

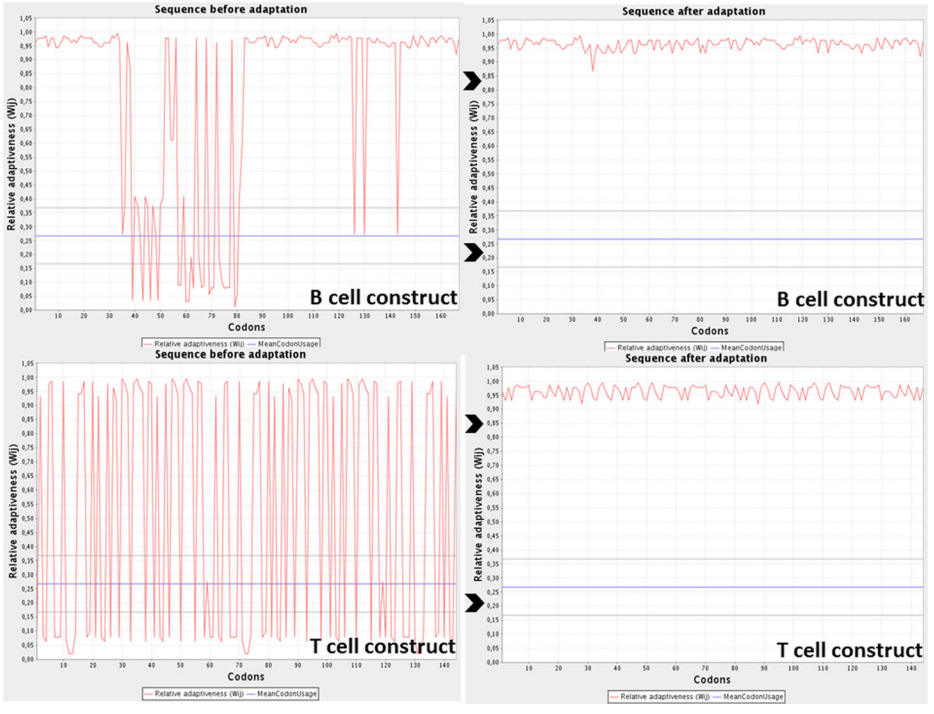


Fig. 8 Graphical view of codon optimization by J CAT server. The codons' statuses before (*left*) and after (*right*) optimization are shown

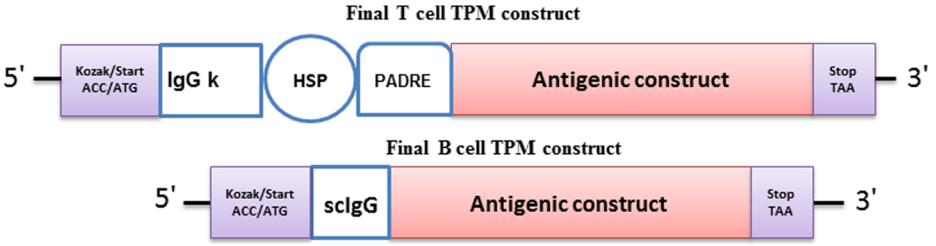


Fig. 9 The graphical representation summarizes the final T and B cell TPM construct

Fig. 10. The different main steps performed in the complete methodology are shown.

Discussion

Vaccines are the most efficacious and cost-effective approaches for reducing morbidity and mortality caused by infectious diseases. Nowadays, it is possible to use a rational step-by-step approach to DNA vaccine design. By evolution of immunoinformatics methods that utilize a wide array of bioinformatics applications, rapid progresses occurred in the field of vaccinology enabling rational design approaches of DNA vaccines [23, 64, 65]. Different advances and

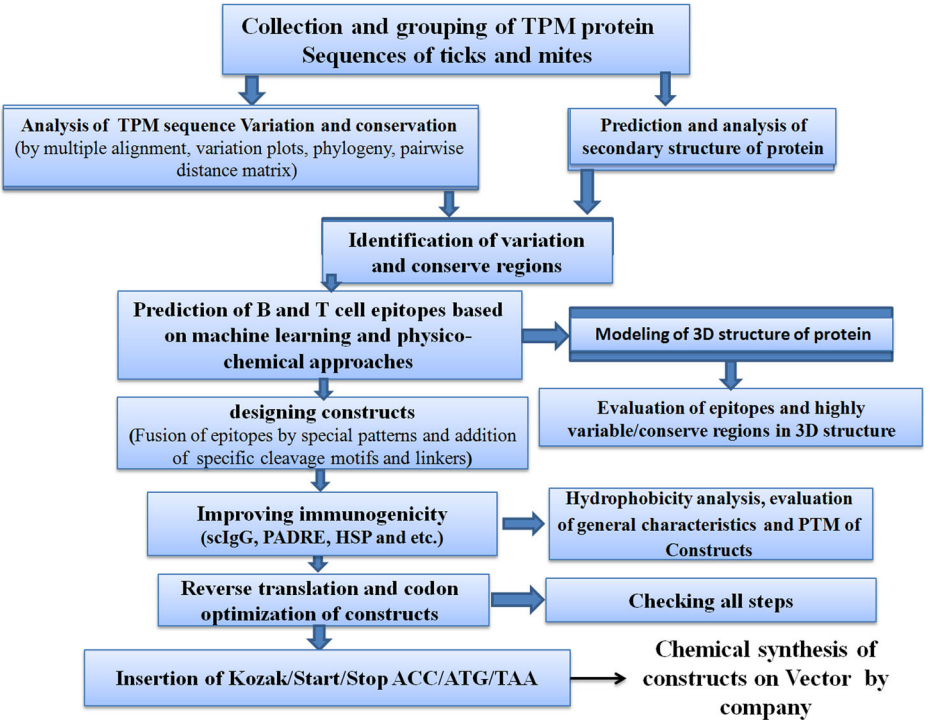


Fig. 10 Flowchart of epitope-driven vaccine design followed in this study. It represents key steps of complete methodology

strategies were considered for the preparation of tick vaccine [3, 61, 66, 67]. Candidate proteins for vaccine against ticks are categorized into two groups as cryptic and exposed antigens. Cryptic antigens predominantly are intestinal, and exposed antigens encompass salivary glands proteins [59]. TPM is one of the exposed antigens and candidates for vaccine development [7, 68, 69]. Identification of the immunogenicity and potency of these proteins for use as a vaccine by bioinformatics methods is one of the major approaches for designing effective antitick vaccines [70]. The antigen for a polytopic DNA vaccine should be selected very carefully since it provokes the production of the protecting antibodies and, thus, affects the quality of the vaccine [71]. With defining a "consensus immunogen," a term used for immunogens that overcome the limitation of serotypes or amino acid variation of a pathogen that causes antibodies only to be effective against some or only one serotype/s or pathogen/s [71], we used tropomyosin as a broad spectrum candidate antigen for constructing a DNA vaccine.

There are several bioinformatics and immunoinformatic researches on immunogenic protein in parasites such as TSOL18 protein in *Taenia solium* worm [72], apical membrane antigen 1 (AMA-1) in *Plasmodium vivax* parasite [73], oncosphere protein of EG95 in *Echinococcus granulosus* cestod [74, 75], and on *Echinococcus multilocularis* proteins [76].

Here, we have attempted to engineer a suitable epitope-driven vaccine by various bioinformatics approaches. An important topic in vaccination is the improvement of immune responses by increasing immunogenicity. Multiple approaches exist to enhance the immunogenicity of DNA vaccines at every level of designing [77], and we used some of these methods such as selection of consensus highly immunogenic peptides based on physicochemical properties and machine learning methods, their pattern of arranging and redundancy in constructs, codon optimization and biased backtranslation, tags such as fragment of cattle antibody, PADRE sequence, HSP, size of construct, or evaluation of posttranslational modification.

There are numerous reports on the significant enhancement of the immune reaction of T and B cells by HSP from different pathways in parasites and nonparasites [39, 78, 79]. On the other side, for the prevention of rapid construct clearance from the host body [27, 28] as the construct may have a small size, sizes increased by redundancy of epitopes and tags such as antibody fragments.

Another major goal followed in this study was defining roles for $CD4^+$ T cell and a special construct for this cell to increase effectiveness and achievement of a vaccine, because one of the important considerations in the stimulation of humeral immunity and outcome immune responses is the interplay of these cells with B cells [16, 27, 28].

Improved polytope expression through the use of optimization of regulatory elements, Kozak sequences, and codon usage [80–82] was the other side of the coin that is mentioned in this study. An essential consideration when optimizing the efficacy and expression of DNA vaccines is the appropriate choice of a plasmid vector [26]; therefore, we selected pVAX1 which is a commercial vector [57].

TPM protein is highly conserved between species and a cross-reactive allergen between mites and other invertebrates. Therefore, it may serve as a wide range vaccine candidate antigen [83]. Thus, vaccine based on TPM of *Boophilus* genus (*B. annulatus* or *B. microplus*) will cover the *Haemaphysalis* genus and *Ixodes* genus and also cross-react with some mites such as *Dermanyssus gallinae*, *P. ovis*, and *Sarcoptes scabiei* although they might have different hosts [83, 84]. These conclusions are helpful in the evaluation of the range of effectiveness of the designed constructs against different parasites.

Taken together, successful in silico designing and modeling is an initial and important step to design a vigorous vaccine, and furthermore, through increased knowledge of the proteins' antigenic nature, it is now possible to produce novel protective and therapeutic vaccines

against parasites. In the current case, at the next step, the constructed DNA vaccine would be produced in vitro and then administrated to cattle for evaluation of its potency toward the induction of immune responses and protection against *Boophilus* tick as well as other ticks and mites.

References

1. Estrada-Peña, A., Bouattour, A., Camicas, J. L., & Walker, A. R. (2004). *Ticks of domestic animals in the Mediterranean region: a guide to identification of species*. Zaragoza: University of Zaragoza Press.
2. Rajput, Z. I., Hu, S., Chen, W., Arijo, A. G., & Xiao, C. (2006). *Journal of Zhejiang University. Science. B, 7*(11), 912–921.
3. Willadsen, P., Riding, G. A., McKenna, R. V., Kemp, D. H., Tellam, R. L., Nielsen, J. N., et al. (1989). *Journal of Immunology, 143*(4), 1346–1351.
4. Peconick, A. P., Sossai, S., Girão, F. A., Rodrigues, M. Q. R. B., Souza e Silva, C. H., Guzman, Q. F., et al. (2008). *Experimental Parasitology, 119*(1), 37–43.
5. Ruiz, L. M., Orduz, S., López, E. D., Guzmán, F., Patarroyo, M. E., & Armengol, G. (2007). *Veterinary Parasitology, 144*(1–2), 138–145.
6. Helfman, D. M., Flynn, P., Khan, P., & Saeed, A. (2008). *Advances in Experimental Medicine and Biology, 644*, 124–131.
7. Jenkins, R. E., Taylor, M. J., Gilvary, N. J., & Bianco, A. E. (1998). *Proceedings of the National Academy of Sciences of the United States of America, 95*(13), 7550–7555.
8. Nisbet, A. J., Huntley, J. F., Mackellar, A., Sparks, N., & McDevitt, R. (2006). *Parasite Immunology, 28*(8), 401–405.
9. Sereda, M. J. (2009). PhD thesis, Humboldt University of Berlin, Berlin, Germany.
10. Wright, H. W. (2011). Edinburgh, UK: The University of Edinburgh
11. Huntley, J. F., Machell, J., Nisbet, A. J., Van den Broek, A., Chua, K. Y., Cheong, N., et al. *Parasite Immunology*, *26*(8–9), 335–342.
12. You, M. J. (2004). *Korean Journal of Parasitology, 42*(4), 195–200.
13. Gupta, S. K., Smita, S., Sarangi, A. N., Srivastava, M., Akhoon, B. A., Rahman, Q., et al. (2010). *Vaccine, 28*(43), 7092–7097.
14. Gupta, S. K., Srivastava, M., Akhoon, B. A., Gupta, S. K., & Grabe, N. (2012). *Infection, Genetics and Evolution, 12*(7), 1513–1518.
15. Fonseca, D. P., Benaissa-Trouw, B., van Engelen, M., Kraaijeveld, C. A., Snippe, H., & Verheul, A. F. (2001). *Infection and Immunity, 69*(8), 4839–4845.
16. Rodriguez, L. L., Barrera, J., Kramer, E., Lubroth, J., Brown, F., & Golde, W. T. (2003). *Vaccine, 21*(25–26), 3751–3756.
17. Yi, J.-Z., Liu, M. Q., Zhu, C. Z., Zhang, Q., Sheng, Z. T., Du, Q. Y., et al. (2004). *Acta Biochimica et Biophysica Sinica, 36*(9), 589–596.
18. Baloria, U., Akhoon, B. A., Gupta, S. K., Sharma, S., & Verma, V. (2012). *Amino Acids, 42*(4), 1349–1360.
19. Akhoon, B. A., Slathia, P. S., Sharma, P., Gupta, S. K., & Verma, V. (2011). *Microbial Pathogenesis, 51*(1–2), 77–87.
20. Nardin, E. H., Calvo-Calle, J. M., Oliveira, G. A., Nussenzweig, R. S., Schneider, M., Tiercy, J. M., et al. (2001). *Journal of Immunology, 166*(1), 481–489.
21. De Groot, A. S. (2006). *Drug Discovery Today, 11*(5–6), 203–209.
22. Suhrbier, A. (2002). *Expert Review of Vaccines, 1*(2), 207–213.
23. Yang, X., & Yu, X. (2009). *Reviews in Medical Virology, 19*(2), 77–96.
24. Sobrino, F., Blanco, E., García-Briones, M., & Ley, V. (1999). *Developments in Biological Standardization, 101*, 39–43.
25. Coban, C., Kobiyama, K., Aoshi, T., Takeshita, F., Horii, T., Akira, S., et al. (2011). *Current Gene Therapy, 11*(6), 479–484.
26. Gurunathan, S., Klinman, D. M., & Seder, R. A. (2000). *Annual Review of Immunology, 18*, 927–974.
27. García-Briones, M. M., Blanco, E., Chiva, C., Andreu, D., Ley, V., & Sobrino, F. (2004). *Virology, 322*(2), 264–275.
28. Gerner, W., Denyer, M. S., Takamatsu, H. H., Wileman, T. E., Wiesmüller, K. H., Pfaff, E., et al. (2006). *Virus Research, 121*(2), 223–228.
29. Sánchez-Burgos, G., Ramos-Castañeda, J., Cedillo-Rivera, R., & Dumonteil, E. (2010). *Virus Research, 153*(1), 113–120.

30. Gupta, S. K., Srivastava, M., Akhoon, B. A., Smita, S., Schmitz, U., Wolkenhauer, O., et al. (2011). *Infection, Genetics and Evolution, 11*(2), 308–319.
31. Ranjbar, M. M., Nabian, S., Nasab, S. D. M., Taheri, M., Ghorban, K., Dadmanesh, M., et al. (2013). *Journal of Paramedical Sciences, 4*(4), 70–72.
32. Altschul, S. F., Madden, T. L., Schäffer, A. A., Zhang, J., Zhang, Z., Miller, W., et al. (1997). *Nucleic Acids Research, 25*(17), 3389–3402.
33. Pilley, H. H., & Atre, N. M. (2011). *Journal of Computational Intelligence in Bioinformatics, 4*(1), 79–86.
34. McNicholas, S., Potterton, E., Wilson, K. S., & Noble, M. E. M. (2011). *Acta Crystallographica. Section D, Biological Crystallography, 67*(Pt 4), 386–394.
35. Chiang, H. L., Terlecky, S. R., Plant, C. P., & Dice, J. F. (1989). *Science, 246*(4928), 382–385.
36. Gupta, S. K., Singh, A., Srivastava, M., Gupta, S. K., & Akhoon, B. A. (2009). *Vaccine, 28*(1), 120–131.
37. Capone, S., Zampaglione, I., Vitelli, A., Pezzanera, M., Kierstead, L., Burns, J., et al. (2006). *Journal of Immunology, 177*(10), 7462–7471.
38. Wei, W. Z., Ratner, S., Shibuya, T., Yoo, G., & Jani, A. (2001). *Journal of Immunological Methods, 258*(1–2), 141–150.
39. Li, X., Yang, X., Jiang, Y., & Liu, J. (2005). *International Immunology, 17*(10), 1293–1302.
40. Sachdeva, R., Banerjea, A. C., Malla, N., & Dubey, M. L. (2009). *PLoS One, 4*(12), e7880.
41. Peng, M., Chen, M., Ling, N., Xu, H., Qing, Y., & Ren, H. (2006). *Vaccine, 24*(7), 887–896.
42. Julenius, K. (2007). *Glycobiology, 17*(8), 868–876.
43. Steentoft, C., Vakhrushev, S. Y., Joshi, H. J., Kong, Y., Vester-Christensen, M. B., Schjoldager, K. T., et al. (2013). *EMBO Journal, 32*(10), 1478–1488.
44. Gupta, R., Jung, E., Gooley, A. A., Williams, K. L., Brunak, S., & Hansen, J. (1999). *Glycobiology, 9*(10), 1009–1022.
45. Kiemer, L., Bendtsen, J. D., & Blom, N. (2005). *Bioinformatics, 21*(7), 1269–1270.
46. Blom, N., Gammeltoft, S., & Brunak, S. (1999). *Journal of Molecular Biology, 294*(5), 1351–1362.
47. Kyte, J., & Doolittle, R. F. (1982). *Journal of Molecular Biology, 157*(1), 105–132.
48. Ferro, A., Giugno, R., Pigola, G., Pulvirenti, A., Di Pietro, C., Purrello, M., et al. (2007). *BMC Bioinformatics, 8*, 58.
49. Sandhu, K. S., Pandey, S., Maiti, S., & Pillai, B. (2008). *In Silico Biology, 8*(2), 187–192.
50. Besse, F., & Ephrussi, A. (2008). Translational control of localized mRNAs: restricting protein synthesis in space and time. *Nature Reviews Molecular Cell Biology, 9*(12), 971–980.
51. Kalwy, S., Rance, J., & Young, R. (2006). *Molecular Biotechnology, 34*(2), 151–156.
52. Frelin, L., Ahlén, G., Alheim, M., Weiland, O., Barnfield, C., Liljeström, P., et al. (2004). *Gene Therapy, 11*(6), 522–533.
53. Ramakrishna, L., Anand, K. K., Mohankumar, K. M., & Ranga, U. (2004). *Journal of Virology, 78*(17), 9174–9189.
54. Cheung, Y. K., Cheng, S. C., Sin, F. W., & Xie, Y. (2004). *Vaccine, 23*(5), 629–638.
55. Yadava, A., & Ockenhouse, C. F. (2003). *Infection and Immunity, 71*(9), 4961–4969.
56. Olafsdóttir, G., Svansson, V., Ingvarsson, S., Marti, E., & Torsteinsdóttir, S. (2008). *Acta Veterinaria Scandinavica, 50*, 44.
57. Montgomery, D. L., & Prather, K. J. (2006). Design of plasmid DNA constructs for vaccines. *Methods in Molecular Medicine, 127*, 11–22.
58. Kozak, M. (1987). *Nucleic Acids Research, 15*(20), 8125–8148.
59. Tellam, R. L., Smith, D., Kemp, D. H., & Willadsen, P. (1992). In W. K. Yong (Ed.), *Animal parasite control utilizing biotechnology* (pp. 303–331). Boca Raton: CRC.
60. García-García, J. C., Montero, C., Redondo, M., Vargas, M., Canales, M., Boue, O., et al. (2000). *Vaccine, 18*(21), 2275–2287.
61. Willadsen, P. (2004). *Parasitology, 129*, S367–S387.
62. De la Fuente, J., & Kocan, K. M. (2006). *Parasite Immunology, 28*(7), 275–283.
63. Diaz, A. A., Tomba, E., Lennarson, R., Richard, R., Bagajewicz, M. J., & Harrison, R. G. (2010). *Biotechnology and Bioengineering, 105*(2), 374–383.
64. Tomar, N., & De, R. K. (2010). *Immunology, 131*(2), 153–168.
65. Korber, B., LaBute, M., & Yusim, K. (2006). *PLoS Computational Biology, 2*(6), e71.
66. Nuttall, P. A., Trimnell, A. R., Kazimirova, M., & Labuda, M. (2006). *Parasite Immunology, 28*(4), 155–163.
67. Willadsen, P. (2006). *Veterinary Parasitology, 138*(1–2), 161–168.
68. Aki, T., Kodama, T., Fujikawa, A., Miura, K., Shigeta, S., Wada, T., et al. (1995). *Journal of Allergy and Clinical Immunology, 96*(1), 74–83.
69. Folkard, S. G., Taylor, M. J., Butcher, G. A., & Bianco, A. E. (1997). *Infection and Immunity, 65*(7), 2846–2851.

70. Grandi, G. (2005). *Genomics, proteomics and vaccines*. Chichester: Wiley.
71. Laddy, D. J., Yan, J., Corbitt, N., Kobasa, D., Kobinger, G. P., & Weiner, D. B. (2007). *Vaccine, 25*(16), 2984–2989.
72. Zimic, M., Gutiérrez, A. H., Gilman, R. H., López, C., Quiliano, M., Evangelista, W., et al. (2011). *Bioinformation, 6*(7), 271–274.
73. Bueno, L. L., Lobo, F. P., Morais, C. G., Mourão, L. C., de Ávila, R. A. M., Soares, I. S., et al. (2011). *PLoS One, 6*(6), e21289.
74. Lightowlers, M. W., Gauci, C. G., Chow, C., Drew, D. R., Gauci, S. M., Heath, D. D., et al. (2003). *International Journal for Parasitology, 33*(11), 1207–1217.
75. Woollard, D. J., Gauci, C. G., Heath, D. D., & Lightowlers, M. W. (2000). *Vaccine, 19*(4–5), 498–507.
76. Gomase, V. S., & Chitlange, N. R. (2012). Prediction of MHC class antigen peptides from Echinococcus multilocularis: application of computer intelligence. *Scientific Reports, 1*(3), 191.
77. Ingolotti, M., Kawalekar, O., Shedlock, D. J., Muthumani, K., & Weiner, D. B. (2010). *Expert Review of Vaccines, 9*(7), 747–763.
78. Cohen-Sfady, M., Nussbaum, G., Pevsner-Fischer, M., Mor, F., Carmi, P., Zanin-Zhorov, A., et al. (2005). *Journal of Immunology, 175*(6), 3594–3602.
79. Nishikawa, M., Takemoto, S., & Takakura, Y. (2008). *International Journal of Pharmaceutics, 354*(1–2), 23–27.
80. Uchijima, M., Yoshida, A., Nagata, T., & Koide, Y. (1998). *Journal of Immunology, 161*(10), 5594–5599.
81. Narum, D. L., Kumar, S., Rogers, W. O., Fuhrmann, S. R., Liang, H., Oakley, M., et al. (2001). *Infection and Immunity, 69*(12), 7250–7253.
82. Stratford, R., Douce, G., Zhang-Barber, L., Fairweather, N., Eskola, J., & Dougan, G. (2000). *Vaccine, 19*(7–8), 810–815.
83. Reese, G., Ayuso, R., & Lehrer, S. B. (1999). Tropomyosin: an invertebrate pan-allergen. *International Archives of Allergy and Immunology, 119*(4), 247–258.
84. Zhang, R., Jise, Q., Zheng, W., Ren, Y., Nong, X., Wu, X., et al. (2012). *Parasites and Vectors, 5*, 176.