

Received 2019-09-07
Revised 2019-10-06
Accepted 2019-11-10

Construction of A Synthetic Gene Encoding the Multi-Epitope of *Toxoplasma gondii* and Demonstration of the Relevant Recombinant Protein Production: A Vaccine Candidate

Maryam Karimi ^{1,2}, Seyyed Javad Seyyed Tabaei ³, Mohammad Mehdi Ranjbar ⁴, Fardin Fathi ⁵, Ali Jalili ⁶, Ghasem zamini ⁷, Amirreza Javadi Mamaghani ³, Javad Nazari ⁸, Daem Roshani ⁹, Nooshin Bagherani ¹⁰, Mohammad Bagher Khademerfan ^{2,7}✉

¹ Student Research Committee, Kurdistan University of Medical Sciences, Sanandaj, Iran

² Cellular and Molecular Research Center, Kurdistan University of Medical Sciences, Sanandaj, Iran

³ Department of Parasitology and Mycology, School of Medicine, Shahid Beheshti University of Medical Sciences, Tehran, Iran

⁴ Razi Vaccine and Serum Research Institute, Agricultural Research, Education and Extension Organization, Karaj, Iran

⁵ Cellular and Molecular Research Center, Research Institute for Health Development, Kurdistan University of Medical Science, Sanandaj, Iran

⁶ Cancer and Immunology Research Center, Kurdistan University of Medical Sciences, Sanandaj, Iran

⁷ Department of Parasitology and Mycology, Faculty of Medicine, Kurdistan University of Medical Science

⁸ Medical Department, Arak University of Medical Science, Arak, Iran

⁹ Social Determinants of Health Research Center, Research Institute for Health Development, Kurdistan University of Medical Sciences, Sanandaj, Iran

¹⁰ Department of Molecular Medicine, School of Advanced Medical Sciences, Tehran University of Medical Sciences, Tehran, Iran

Abstract

Background: *Toxoplasma gondii* is a widely-distributed parasite all over the world whose attributed severe afflicting complications in human necessitate the development of serodiagnostic tests and vaccines for it. Immunological responses to monovalent vaccines and the application of diagnostic reagents including single antigens are not optimally effective. Bioinformatics approaches were used to introduce these epitopes, predict their immunogenicity and preliminarily evaluate their potential as an effective DNA vaccine and for serodiagnostic goals. **Materials and Methods:** A 3D structure of proteins was predicted by I-TASSER server, and linear and conformational B cell and T cell epitopes were predicted using the online servers. Then, the predicted epitopes were constructed and called Toxoeb, and their expression in the prokaryotic and eukaryotic cells was demonstrated using SDS-PAGE. In the next step, Western blotting with pooled sera of mice infected with *T. gondii* was done. **Results:** The current *in silico* analysis revealed that the B cell epitopes with high immunogenicity for GRA4 protein were located in the residues 34-71, and 230-266, for GRA14 in 308-387, for SAG1 in 182-195, 261-278, and for GRA7 in residues 101-120, 160-176. The T cell epitopes were selected in overlapping regions with the B cell epitopes. The immunogenic region for GRA4 are in the residues 245-253, 50-58, and 40-54, for GRA14 in 307-315, 351-359, and 308-322, for SAG1 261-269, and 259-267, and for GRA7 in the residues 103-112, and 167-175. The results of the western blotting showed that the expressed protein had immunogenicity. **Conclusion:** Our constructed multi-epitope of *T. gondii* could be considered as a candidate for diagnostic and vaccination purposes. [GMJ.2020;9:e1708] DOI:[10.31661/gmj.v9i0.1708](https://doi.org/10.31661/gmj.v9i0.1708)

Keywords: *Toxoplasma gondii*; DNA Vaccine; Multi-Epitope; Bioinformatics

GMJ

Copyright© 2020, Galen Medical Journal. This is an open-access article distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>)
Email: info@gmj.ir



✉ **Correspondence to:**
Mohammad Bagher Khademerfan, Cellular and Molecular Research Center, Kurdistan University of Medical Sciences, Pasdaran Boulevard, Sanandaj, Iran
Telephone Number: +989132292705
Email Address: mb.khademerfan@gmail.com

Introduction

Toxoplasma gondii is a worldwide obligate intracellular parasite that causes zoonotic diseases [1]. It is one of the most well-known protozoa due to its importance in medicine and veterinary and its propriety as a model for cellular biology investigations and molecular studies on unicellular organisms [2]. Almost one-third of people in the world are infected with this protozoa [3]. Its primary infection or recrudescence in immunocompromised individuals such as those with cancer, those undergoing chemotherapy or organ transplantation, and those with untreated HIV infection/AIDS, may lead to ocular damage and severe neurological complications which can be fatal [4, 5]. The primary infections during pregnancy can result in severe complications in fetuses and infants [5]. The infection is mainly transmitted by water or food contaminated with the oocyst of *T. gondii* acquired from cat feces or by eating not cooked meat including cysts [6]. *T. gondii* can cause abortion and neonatal loss in livestock, mainly sheep and goat, that has economic importance [7]. Conventional medicines currently being used for the treatment of toxoplasmosis essentially have therapeutic effects on the tachyzoites, but they cannot eradicate the encysted parasites within infected hosts. Thus, the development of a suitable and effective vaccine with the ability to eradicate all forms of the parasite is of great priority that can immediately inhibit and restrict the infection in both humans and domestic animals [8]. One of the novel methods in vaccine development is the construction of synthetic multi-epitopes using the T cell and B cells epitopes that can cause diverse T cell responses and induction of neutralizing antibodies [9, 10]. Most B cell epitopes (90%) are discontinuous and more complex and difficult to be detected in comparison with linear epitopes [11, 12]. Within the T cell population, the CD8⁺ T cells are the most effective cells against *T. gondii*, while the CD4⁺ T cells have a synergistic role in this field [13]. Studies have revealed that the immunological effects of monovalent vaccines including single antigens are not significant and perfectly effective [14]. In the past, a variety of *T. gondii* antigens have been intro-

duced as appropriate candidates for vaccine development [15]. In different studies, it has been shown that all of the virulence factors are immunostimulatory molecules which can be considered as potential candidates for the toxoplasmosis vaccine [16, 17]. On the tachyzoites, the major surface antigen 1 (SAG1) has a critical role in binding and invading the host cells [15]. One of the secretory dense granule proteins, GRA4, which is present in the bradyzoites and tachyzoites and secreted by the intracellular tachyzoites, plays a key role in the survival of the parasite inside the cells [16]. GRA7 is another secretory protein that induces a strong humoral and cellular response during acute and chronic phases of infection [18, 19]. GRA14 is secreted into the vacuole and traffics to both parasitophorous vacuole membrane and intravacuolar network. This protein has a specific topology with a long structure in the way to be expected that strongly can stimulate the immune system [20]. In the current study, we have exploited an immunoinformatic approach to recognize potential epitopes which could form the basis for the vaccine development against the *T. gondii*. According to this approach, in the current study, the T cell and B cell epitopes of SAG1, GRA4, GRA7, and GRA14 were assigned and considered together for designing a vaccine. Our objective was to introduce a multi-epitope gene with high immunogenicity as an effective candidate for vaccination against the *T. gondii*.

Materials and Methods

Sequence Retrieval of Targets

Nucleotide sequences for the proteins GRA4 (Gene Bank: Q27002.1), GRA7 (Gene Bank: AFO54982.1), GRA14 (Gene Bank: ACH88354.1), and SAG1 (Gene Bank: AFO54882.1) were retrieved from the National Center for Biotechnology Information (NCBI), Nucleotide Database. Basically, the target sequences were selected from the well-known RH strain of *T. gondii*, which is an extremely virulent strain. Complete amino acid sequences of these four proteins were obtained from the UniProt KB database using the proteins' ID (<http://www.uniprot.org/uniprotKB/>).

Analyzing Proteins' Primary Structure and Physicochemical Properties

For the analysis of primary proteins structure, amino acid sequences of the proteins were submitted to the ExPASy tools. Several physicochemical parameters of proteins including the number of amino acids, instability index, molecular weight, theoretical isoelectric point (IP), grand average of hydropathicity (GRAVY), and aliphatic index were predicted using the ExPASy protParam online server (<http://web.expasy.org/protParam>) [21].

Prediction of Proteins' Secondary Structures and Transmembrane Topology

The secondary structure of the proteins was predicted using the Self-Optimized Method with Alignment (SOPMA) server (<http://npsa-prabi.ibep.fr/cgi-bin/npsa-automat>) [22]. The regions containing outside, inside and transmembrane portions of the proteins were predicted by TMHMM server 2.0 (<http://www.cbs.dtu.dk/services/TMHMM>) [23]. The signal peptide region of proteins was predicted by the SignalP-5.0 server (<http://www.cbs.dtu.dk/services/SignalP/>) [24].

Prediction of Proteins' Tertiary Structures

Tertiary structures of the GRA7, GRA14, and GRA4 proteins were predicted by using I-TASSER tool [25, 26]. This tool works based on multiple threading alignments and iterative model fragment assembly simulations. I-TASSER is the best tool for predicting the protein 3D structure in researches for the assessment of modeling servers. In this server, the C-score is explained as the confidence score for evaluating the quality of predicted models and higher values of the C-score illustrate the model with high confidence and vice-versa [27]. The tertiary structure of SAG1 was downloaded from PDB protein data bank (<https://www.rcsb.org/>).

Model Validation

Models were validated by Ramachandran plots (PROCHECK; <https://servicesn.mbi.ucla.edu/PROCHECK>) [28].

Prediction of Conformational and Linear B Cell Epitopes

For prediction of linear B cell epitopes,

ABCpred (<http://crdd.osdd.net/raghava/abcpred/>) [29], LBtope (<http://crdd.osdd.net/raghava/lbtope/protein.php>) [30], SVMtrip (<http://sysbio.unl.edu/SVMTriP/>) [31] and IEDB (<http://tools.immuneepitope.org/bcell/>) web servers were used. Thus, the predictions of epitopes were mostly on the basis of Chou and Fasman beta-turn [32], Karplus and Schulz flexibility [33], Emini surface accessibility [34], Parker hydrophilicity [35], and Kolaskar and Tongaonkar antigenicity [36]. We used Kolaskar and Tongaonkar method which makes 75% accuracy in predicting linear B cell epitopes [36]. For prediction of discontinuous B cell epitopes, DiscoTope 2.0 (<http://www.cbs.dtu.dk/services/DiscoTope/>) [37] and , CBTOPE (<http://crdd.osdd.net/raghava/cbtope>) [38] were administered. DiscoTope calculation and combination connect numbers (equivalent for surface accessibility) with the tendency scores of residues in spatial proximity [37]. The CBTOPE server used the Support Vector Machine (SVM) based on a method by applying the amino acid composition produced from the query sequences with overall accuracy [38]. The result of DiscoTope is more valid than other conformational epitope predictor server [37]. Experimental epitopes available on the IEDB server can also be used. For epitope prediction of GRA7 and SAG1 proteins, available experimental epitopes in IEDB server were used.

Prediction of T Cell Epitopes

Based on using Balb/c mouse as an animal model, the online servers IEDB (<http://tools.iedb.org/main/tcell/>) and SYFPEITHI (<http://www.syfpeithi.de/bin/MHCServer.dll/EpitopePrediction.htm>) were used for predicting alleles in the murine model [39]. The H2-Kd, H2-Ld, and H2-Dd alleles as mouse MHC-I molecules and H2-IA_d and H2-IE_d alleles as mouse MHC-II molecules were chosen. The T cell epitopes which showed the highest scores in relation to the mouse MHC-I and MHC-II were selected. The epitopes were also chosen based upon areas in overlapping with the B cell epitopes.

Construction of Epitopes

The predicted and experimental antigen epitopes were connected by linkers. The most

commonly-used flexible linkers have sequences primarily consisting of stretches of glycine and serine residues [40]. Two oligonucleotide sequences with action sites for *Bam*HI and *Xho*I restriction enzymes were placed at the start and endpoints of the sequence. Additionally, the sequence included a small kozak sequence, ATG start codon, and a TAG stop codon. Then, the sequence was translated into the nucleotide sequence and codon optimization was performed for eukaryotic cells (mice) by the JCAT online software (<http://www.jcat.de/>) (Technical University of Braunschweig, Germany) [41]. Codon optimization is a method for getting the higher expression of the foreign gene introduced by a vector in a host. The different codons of glycine and serine were used in the linkers to achieve the least repetitive codons.

Prediction of Designated Construct's Structure

Physiochemical parameters of construction including the number of amino acids, molecular weight, theoretical IP, instability index, aliphatic index and grand average of hydrophobicity (GRAVY) were analyzed using the ExPasy protparam online server. The secondary structure of the construct was assessed by (SOPMA) server and the tertiary structure was predicted by the online prediction I-TASSER server and was evaluated by using RAMPAGE.

Cloning and Protein Expression of Construct in the Prokaryotic Expression Vector

The pUC57 plasmid harboring the inserted coding named Toxob was constructed and produced by a commercial supplier according to our designation (Bio Basic, Canada). In the next step, using the enzymatic digestion method, the construct was removed from the transfer vector pUC57 and sub-cloned in the expression plasmid pET-32a (+) (Novagen, Germany). The recombinant plasmid pET32-Toxob was transformed into (*E. coli*) BL21 (DE3) cells (Novagen, Darmstadt, Germany) and grown in LB broth medium (Merck, Germany) with 50µg/ml Ampicillin to an optical density at 600nm of 0.6. Expression of the recombinant protein was induced by the addition of 1mM isopropyl

β-D-1 thiogalactopyranoside (IPTG) at 37°C, followed by harvesting the (*E. coli*) BL21 (DE3) cells after 8 hours. After suspending the bacteria in the lysis buffer (Tris-HCl 50 mM, Glycerol 10%, and TritonX-100 0.1%; PH: 8), the cells were disrupted by sonication (Hielscher, UP200H, Germany) for 40 cycle 0.5 S at amplitude 70% and clarified by centrifugation for 20min at 12000×g at 4°C. The supernatant was analyzed by sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE; BioRad, USA).

SDS-PAGE and Western Blot Test

The expression of the protein was confirmed by 12% SDS-PAGE. In this process, the gel was stained with Coomassie Brilliant Blue and analyzed. The recombinant histidine-tagged pET32-Toxob protein was purified by Ni-NTA affinity chromatography (Qiagen, Germantown, MD, USA) according to the manufacturer's protocols. To evaluate immunoreactivity, the purified recombinant protein was analyzed by western blot. Briefly, the protein was separated by 12% SDS-PAGE, then the gel containing protein was electro-transferred on nitrocellulose membrane. The UV Cross Linker (UV Tec, EEC) was used to stabilize the protein on the membrane. The membrane was blocked 16 h at 4°C with 3% bovine serum albumin (BSA; Merck, Germany). After washing with Tris-buffered saline and Tween 20 (TBST) three times, the membrane was incubated with pooled sera from mice infected with *T. gondii* diluted 1:50 in 1% BSA for 2 h in 37°C. Then, the membrane was re-washed three times with TBST, and was incubated with 1:5000 dilution of horseradish peroxidase (HRP) conjugated with goat anti-mouse IgG antibody at 37°C for 2 h. Thereafter, the membrane was washed and incubated with a solution containing DAB (3, 3'-diaminobenzidine; Sigma Aldrich, USA), H₂O₂, Tris and distilled water at RT for 15 min. At the end, the reaction was stopped using distilled water.

Cloning and Protein Expression of Construct in the Eukaryotic Expression Vector

Chinese Hamster Ovary (CHO) cells and plasmid pCDNA3.1(+) were used to confirm the protein expression in the eukaryotic cells.

For this reason, the CHO cells were cultivated in RPMI1640 (BioSera, France) with penicillin/streptomycin (AtoceI, Hungary) and 10% fetal bovine serum (FBS; Gibco, USA). The cells were kept in a humidified atmosphere at 37°C with 5% CO₂. These cells were plated in counts of 1-2×10⁵ in 6 well plates. The cells were transfected using calcium phosphate [42]. After further incubation for 48 hours, the cells were harvested and suspended in lysis buffer (sodium dodecyl sulfate 1%). Then, the cells were disrupted by sonication for 20 cycle 0.5 S at amplitude 70% and clarified by centrifugation for 20 min at 12000×g at 4°C. The expression of the protein was confirmed by 12% SDS-PAGE. The gel was stained with Coomassie Brilliant Blue and analyzed.

Results

Analysis of Physicochemical Properties and Transmembrane Topology

An overview of the most important data collected from ExPasy server is given in Table-1. IP is a PH point in which the net charge of the protein is zero. Its importance is in the evaluation of a proteins' solubility and mobility in an electric field. The IP of the three proteins GRA4, GRA14, and SAG1 was greater than 7, while the IP of GRA7 was less than 7. The instability index shows an estimation of the stability of a given protein in vitro. Based on the instability index, the proteins GRA4, GRA7, and GRA14 were unstable while the SAG1 was a stable protein. The aliphatic index of the three proteins GRA7, GRA14, and SAG1 was relatively high that revealed that

they had been relatively stable at different temperatures. The GRAVY index represents the degree of proteins' hydrophobicity and increasing positive scores show greater hydrophobicity. The results showed that the three proteins GRA4, GRA7, and GRA14 had high hydrophilicity, while the SAG1 compared with the three others had relatively low hydrophilicity. The aliphatic index of a given protein is defined as the relative volume which is occupied by aliphatic side chains including alanine, valine, isoleucine, and leucine. It may be considered as a positive factor for gaining thermo-stability in globular proteins. According to the aliphatic index, the proteins SAG1, GRA7, and GRA14 were stable over a wide temperature range. Assessment of the transmembrane topology revealed that the greater part of all of these four proteins were placed on the outer of the membrane.

Prediction of Proteins' Structures

The prediction results of the secondary structure clarified the alpha helices and beta-turn parts of the proteins. These structures were predicted using the SOPMA server (Figure-1). Tertiary structure models were generated using I-TASSER and PROCHECK server (Figure-2). C-score for the GRA4, GRA7, and GRA14 were 2.36, -3.38, and -1.92, respectively. Additionally, estimated accuracy of model 1s was 0.44±0.14 (TM-score) and 11.8±4.5A° (RMSD) for the GRA4, 0.34±0.11 (TM-score) and 12.7±4.3A°(RMSD) for the GRA7, and 0.48±0.15(TM-score) and 11.2±4.6A°(RMSD) for the GRA14. Hence, the results showed the appropriateness of

Table 1. Parameters Computed for Proteins and constructed gene from Proteins SAG1, GRA4, GRA7, GRA14 *Toxoplasma. gondii* Using by ExPasy ProtParam Tool.

Parameters	SAG1	GRA4	GRA7	GRA14	Toxoeb
No. of amino acid	319	345	236	408	275
Molecular weight (KD)	32.6	36	25.8	44.7	30
Theoretical pI	7.43	7.05	5.09	8.49	6.03
Instability index	36.47	51.84	50.96	52.86	63.96
Aliphatic index	79.84	56.32	69.96	85.59	34.59
GRAVY	0.111	-0.467	-0.615	-0.213	-1.08

GRAVY: Grand average of hydrophobicity



Figure 1. Secondary structures of proteins SAG1, GRA4, GRA7, GRA14 *Toxoplasma gondii* were predicted using SOPMA server. A: GRA4, B: GRA7, C: GRA14, D: SAG1.

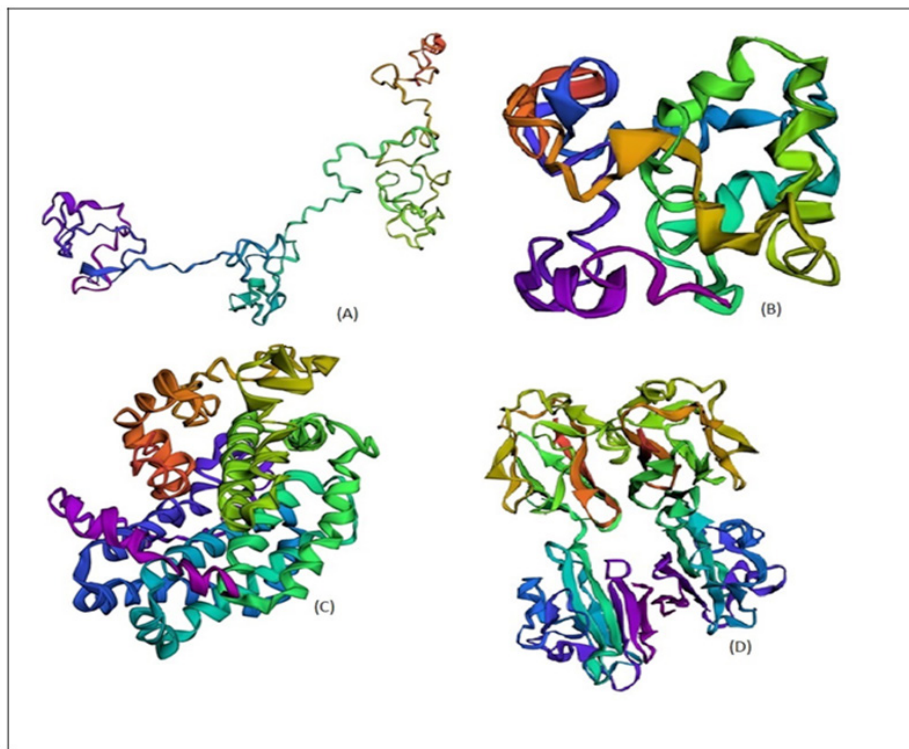


Figure 2. Result of Prediction of 3D Structure of Proteins SAG1, GRA4, GRA7, GRA14 *Toxoplasma gondii*. Using I-TASSER and PRO-CHECK. A: GRA4; B: GRA7; C: GRA14; D: SAG1.

these proteins for modeling.

Model Validation

Evaluation of the quality of the model by PROCHECK is shown in Table-2 All of the four modeled structures were valid.

Prediction of Conformational and Linear B Cell Epitopes

The predicted linear and discontinuous B cell epitopes are showed in Tables-3 and Figure-3. In Figure-4, the predicted B cell epitopes on the 3D structures of the proteins and Toxoeb construct are schematically shown.

Prediction of T Cell Epitopes

The predicted T cells' epitopes are shown in Table-4. The T cell epitopes are inside the regions where the B cell epitopes are selected to simultaneously stimulate cellular and humoral immunity.

Construction of Epitopes

Based on our findings, we designed an 825bp sequence DNA which we named Toxoeb. The presence of the recombinant plasmid was confirmed through digestion by the restriction enzyme, PCR and sequencing.

Prediction of Construct's Structure

The number of amino acids of the Toxoeb construct was 275 (Figure-5). The estimated molecular weight of this construct was 30 kDa and its IP was 6.03. Its computed value was lower than 7, demonstrating the acidic nature of the protein. The aliphatic index for the construct was 34.59 indicating its instability in a wide range of temperatures. The instability index of this construct was 63.96 based upon which the construct was classified as unstable. The construct's GRAVY index was negative (-1.08, Table-1) which revealed that Toxoeb had high hydrophilicity and tended to interact

Table 2. Ramachandran Result of the Predicted model from Proteins SAG1, GRA4, GRA7, GRA14 *Toxoplasma gondii*.

Residue	GRA4	GRA7	GRA14	SAG1
Residue in most favored regions	36.4%	28.7%	70%	87.6%
Residue in most allowed regions	56%	65.2%	27%	12.4%
Residue in most disallowed regions	7.5%	5.9%	2.6%	0%

Table3. Predicted B Cell Epitopes and Selected Final Consensus Epitope from Proteins SAG1, GRA4, GRA7, GRA14 *Toxoplasma gondii*.

Protein ID	Designed B cell epitope	Residue	Status
SAG1 (Gena Bank: AFO54882.1)	SDKGATLTIKKEAFPAES	261-278	Experimental
	SVVNNVARCSYGAD	182-195	Experimental
GRA7 (Gena Bank: AFO54982.1)	RKRGVRSDAEVTDDNIY	101-120	New
	LVPELTEEQQRGDEPLT	160-176	Experimental
GRA4 (Gene Bank: Q27002.1)	VSTEDSGLTGVKDSSSESTVTPA- -DEAASESEEGDKTSRKSQVKKGI	230-266	New
	GLSGRHDKERHQAKKRY- -YHSMYGNQTPYANGQQASP	34-71	New
GRA14 (GenBank: ACH88354.1)	TPRVRAFLERRRMRRDGGGDSGDFGEEGRSKGDVSTSD- -MPREPPPPYSPPMYPFADPEHRWAGTYGTSHGGYRVQPTAP	308-387	New

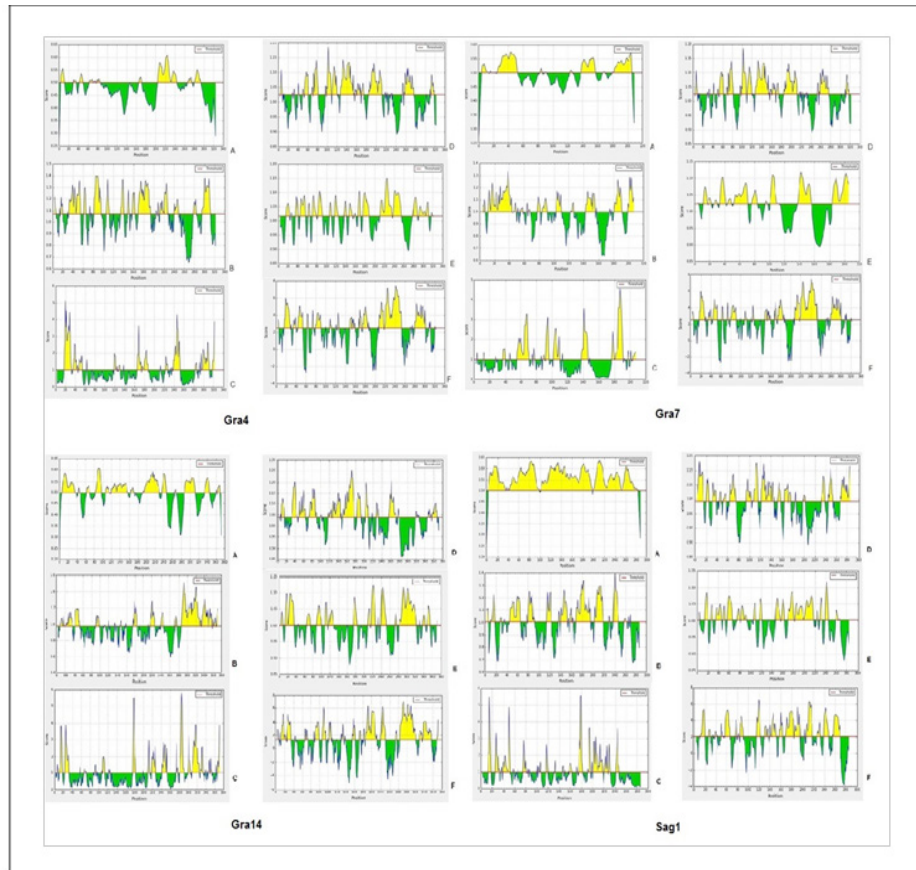


Figure 3. B cell Epitope Prediction for Proteins, GAR4, GRA7, GRA14, and SAG1 *Toxoplasma. gondii* using IEDB software. A: BepiPred linear epitope; B: Chou and Fasman β -turn; C: Emimi surface accessibility; D: Kolaskar and Tongaonkar antigenicity; E: Karplus and Schulz flexibility; F: Parker hydrophilicity.

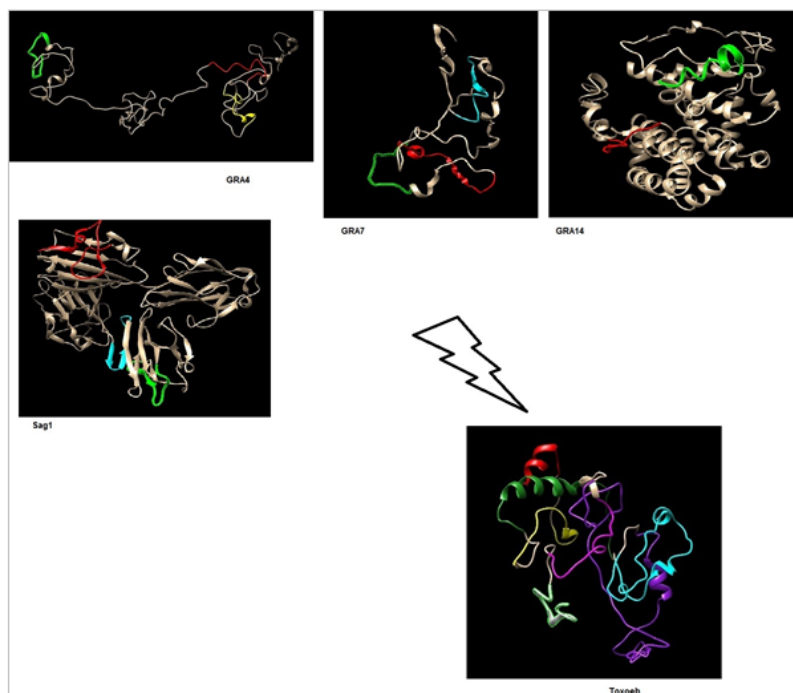


Figure 4. Predicted B cell Epitopes on the 3D Structures of Proteins GAR4, GRA7, GRA14, and SAG1 *Toxoplasma. gondii* and 3D structure of Toxoeb Construct.

Table 4. Predicted T Cell Epitopes from SAG1, GRA4, GRA7, GRA14 *Toxoplasma gondii*

Protein ID	Designed T cell epitope	Residue
SAG1 (Gena Bank:AF054882.1)	SDKGATLTI	261-269
	ASSDKGATL	259-267
GRA7 (Gena Bank:AF054982.1)	RGVRSDAEV	103-112
	EQQRGDEPLT	167-175
	SSESTVTPA	245-253
GRA4 (Gene Bank:Q27002.1)	YYHSMYGNQ	50-58
	DKERHQAKKRYHSM	40-54
	GTPRVRAF	307-315
GRA14 (GenBank:ACH88354.1)	PPPPYSPPM	351-359
	TPRVRAFLERRRMRR	308-322

GSIMVVNNVARCSYGADGGSSGGSDKGATLTIKKEAFPAESGGGSSGGGGL
 SGRHDKERHQAKKRYHSMYGNQTPYANGQQASPPGGSSGGVSTEDSGLT
 GVKDSSSESTVTPADEAASEEGDKTSRKS KVKKIGGGSSGGTPRVRAFL
 ERRRMRRDGGGDSGDFGEEGRSKGDVSTSDMPREPPPPYSPPMYPFADP
 EHRWAGTYGTSHGGYRVQPTAPGGGSSGGGLVPELTEEQQRGDEPLTGGSS
 GGRKRGVRSDAEVTDDNIYstopLE

Figure 5. Amino Acid Sequence of Toxob Construct from Proteins GAR4, GRA7, GRA14, and SAG1 *Toxoplasma gondii*.

with surrounding water molecules. The result of the prediction of Toxob's secondary structure showed that the ratio of α -helixes, β -turn, random coil and extended strand amount were 14.18%, 13.43%, 56.72%, and 15.67%, respectively. A high proportion of random coils in the Toxob construct indicated that the protein might form antigenic epitopes. The 3D structure of the construct was assessed using rampage. The result of a Ramachandran Plot showed 42.6%, 41.6%, 10.2%, and 5.6% of residues found in favored regions, additionally allowed regions, generously-allowed regions, and disallowed regions, respectively.

SDS-PAGE and Western Blot

Results of the SDS-PAGE for the recombinant pET32a-Toxob showed a single protein band with a molecular weight of 30kDa representing the protein expression (Figure-6).

Western blotting with a 1:50 dilution of mice sera demonstrated that the expressed proteins had had immunogenicity, whereas the control intact plasmid pET-32a (+) show no any band (Figure-7). The SDS-PAGE for the recombinant pCDNA3.1-Toxob showed a band representing the protein expression in 30kDa, whereas the intact plasmid pCDNA3.1 shows no any band (Figure- 8).

Discussion

Immunological research has shown that antigens do not function through their entire molecule; their proprietary epitopes are the source of the immune response [43]. Epitopes are chemical groups that introduce the features of a given antigen [44]. The structure of protein antigens is not only made of specific epitopes used by the B, T, and cytotoxic

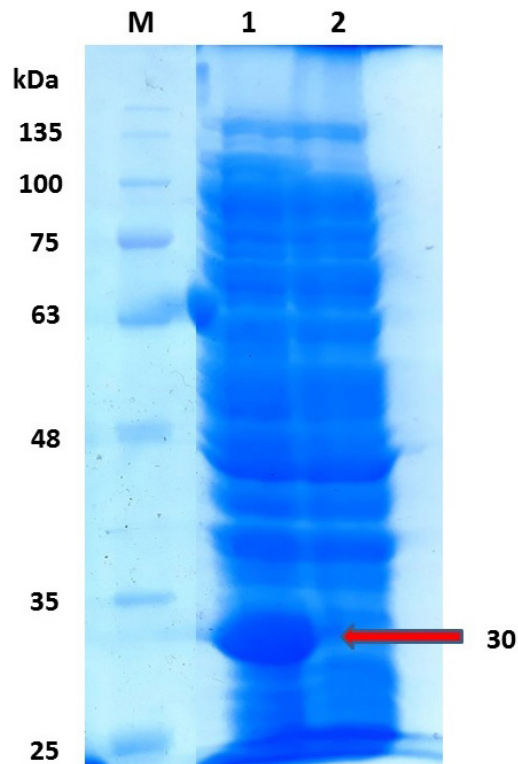


Figure 6. The SDS- PAGE analysis of the expression of rPe-t32a-Toxob, protein in E. Coli BL21(DE3). Lane1: expression of the rPet32a-Toxob protein; Lane2: BL21 containing pET-32a without the insert.

T lymphocytes, and NK cells for triggering immunological responses, but they may also have structures not required for the immune system to induce protective responses. Therefore, studies on these epitopes have increased our knowledge of the structure and activity of antigens, the antigen-antibody interactions, and other immune system activities that play an important role in developing new diagnostic reagents and vaccines [43]. Studies have revealed the immunological responses to monovalent subunit vaccines not to be optimal, and diagnostic reagents consisting of single antigens not to be desirable [45]. The application of immunoinformatic methods for the prediction of immunogenic epitopes has achieved a good place as an essential instrument for epitope localization. Bioinformatics, as an interdisciplinary science, has been widely applied to predict protein function, structure, and epitopes [46]. These methods have decreased blind points in epitope identification, and have improved the accuracy and precision of this identification. Additionally, such methods are efficient, economically ap-

propriate, and able to decrease the need for experimental instruments [47]. Given the subsequent considerable progress in bioinformatics, a diversity of parameters containing a scheme of hydrophilicity, flexibility, accessibility, and antigenicity have been developed [32-34, 48, 49], which have played a significant role in enhancing studies on linear epitopes. To improve the prediction of epitopes, it is often essential to combine a variety of algorithms and findings from the multi-level analysis [50]. Combination of computerized prediction algorithm and experimental tools have resulted in the fast development of conformational epitope analysis and localization, and based on these combinations, several effective prediction programs have been released [43]. The success of polytope vaccines' development is related to rigid criteria for choosing epitopes and also the linkers between these epitopes [9]. Obtaining information regarding antigen epitopes would facilitate the development of epitope vaccines. Previously, the epitope prediction was performed based on only a single parameter, while currently, its accuracy has been limited secondary to progressing development in the bioinformatics field. Nowadays, by using a multi-parameter and multi-method analysis, the accuracy of epitopes' prediction has significantly been improved. Primarily, many researchers all over the world, who are working on the production of vaccines and diagnostic reagents, design and assess them in silico by using bioinformatics tools to reduce the cost for their production, lessen blind spots, and increase the chance of success. Shan Liu *et al.* developed a multi-epitope DNA vaccine expressing the 6 antigen segments from *T. gondii*, which was used for immunizing mice through its introduction by the recombinant plasmid. The authors reported high survival rates in mice when they were challenged with the *T. gondii* RH strain [51]. In another study, Hajissa *et al.* assessed the protective efficacy of a recombinant multi-epitope antigen-expressing nine epitopes from *T. gondii* in mice. They concluded that the multi-epitope recombinant antigen-induced strong immunity against the *T. gondii* RH strain [52]. In the current study, we also firstly designed the vaccine candidate construct by accessible

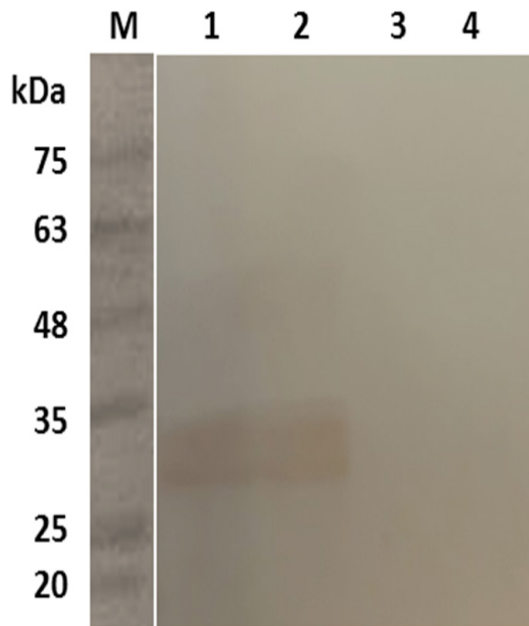


Figure 7. Western blotting analysis of the recombinant rPet32-Toxob with mice pooled sera. Lane 1 and 2: rPet32-Toxob; Lane 3 and 4: BL21 containing pET-32a without the insert.

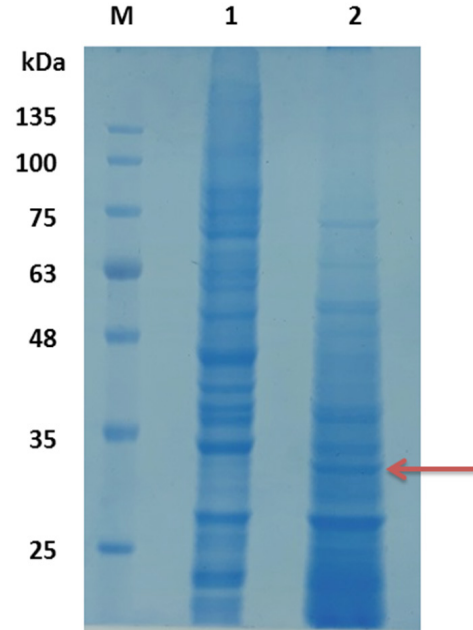


Figure 8. The SDS- PAGE analysis of the expression of rpCDNA3.1-Toxob. M: Marker; Lane 1: CHO cell containing pCDNA3.1 without insert; Lane 2: rpCDNA3.1-Toxob.

software, and then we evaluated its immunogenicity in vitro. Based on our schedule, the next step will be to study the immunogenicity of this gene in the murine model. In the current study, the B cell epitopes GRA4, GRA7, GRA14, and SAG1 were fully analyzed using online services. For the prediction of the conformational B cell epitopes, knowing the spatial structure of proteins is essential. The 3D structure of some proteins has been determined by methods such as crystallography. In the case of proteins whose 3D structure is unknown, their structures would be predicted only based on sequencing and checking the forces, but not using experimental structural information. One of the limitations of our study was the inaccessibility of the actual tertiary structures of GRA7, GRA4, and GRA14 proteins; thus, we had to predict probable tertiary structures using some bioinformatics tools. Inaccessibility to some of the prediction software currently accessible for researchers all over the world, because of our sanctions, was another limitation of our study. Prediction of the linear B cell epitopes was performed using the LBTOPE, SVMtrip, ABCpred, and IEDB; and for conformational B cell epitopes using CBTOPE and Discotope server. Based on the prediction, we concluded

that for the GRA4 protein, the residues in the regions 34-71 and 230-266, for the GRA14 protein in 308-387 and 308-322, for SAG1 in 182-195 and 261-278, and for GRA7 in 101-120 and 160-176 had the most immunogenic potential. Prediction of the T cell epitopes was carried out by the IEDB and SY-FPEITHI servers. The immunogenic regions for the GRA4 protein were the residues of 245-253, 50-58, and 40-54, for the GRA14 protein in the residues of 307-315, 351-359, and 308-322, for the SAG1 in the residues of 261-269, and 259-267, and for the GRA7 in the residues of 103-112 and 167-175. These regions had decent potential for designing epitopes. The T cell and B cell epitopes were selected in the overlapping regions. The selected epitopes were connected by linkers. The flexible linkers are usually used when the attached domains need a certain degree of mobility and interplay. They are commonly composed of small, non-polar (*e.g.*, glycine) or polar (*e.g.*, serine or threonine) amino acids. The small size of these amino acids enables inflexion and permits for movement of the joined functional domains [40]. Using the UCSF Chimera1.8.1 software (Resource for Biocomputing, Visualization, and Informatics University of California, USA), epitopes and

their positions were identified on the 3D structures, and β -turn and α -helices, respectively, and their accessibilities were viewed. Their sequences were translated into the nucleotide one by the JCAT online software. Additionally, the codon optimization was performed for the eukaryotic cells (mice) by this software. Codon usage bias refers to differences in the frequency of synonymous codon occurrence in coding DNA. Although codon optimization was performed for the eukaryotic cells, the sequence was sub-cloned in the prokaryotic plasmid, and finally, its expression was demonstrated. Additionally, its expression in the eukaryotic plasmid, which had already been confirmed by SDS-PAGE, was verified in immunogenicity by the western blotting of the infected mice sera. Western blotting on the mice pooled sera showed that the expressed protein had immunogenicity and can be used for diagnostic purposes and vaccine development.

Conclusion

The current paper describes the primary, secondary, and tertiary structure of the proteins

GRA4, GRA7, GRA14, and SAG1; and their T cell and B cell epitopes using bioinformatics approaches and online servers. An 825 bp sequence DNA was constructed and its immunogenicity was confirmed in vitro. In conclusion, this multi-epitope synthetic gene can be used for the development of DNA vaccine and diagnostic reagents.

Acknowledgment

This paper is a part of the Ph.D. thesis of the first author and was supported by a research grant (code: IR.MUK.REC.1396/116) from Kurdistan University of Medical Science. The authors would like to appreciate very much for kind collaboration of Prof. Bruce R. Smoller, Professor and Chair, Department of Pathology and Laboratory Medicine, Professor, Department of Dermatology University of Rochester School of Medicine and Dentistry.

Conflict of Interest

None declared.

References

1. Raizman RE, Neva FA. Detection of circulating antigen in acute experimental infections with *Toxoplasma gondii*. *J Infect Dis.* 1975;132(1):44-48.
2. Dubey JP. The history of *Toxoplasma gondii*—the first 100 years. *J Eukaryot Microbiol.* 2008;55(6):467-475.
3. Dubey JP. *Toxoplasmosis of animals and humans.* CRC press; 2016.
4. Lindsay D, Dubey J. *Toxoplasma gondii: the changing paradigm of congenital toxoplasmosis.* *Parasitology.* 2011;138(14):1829-1831.
5. Sensini A. *Toxoplasma gondii* infection in pregnancy: opportunities and pitfalls of serological diagnosis. *Clin Microbiol and Infect.* 2006;12(6):504-512.
6. Alibakhshi A, Bandehpour M, Nafarieh T, et al. In silico Analysis of Immunologic Regions of Surface Antigens (Sags) of *Toxoplasma gondii*. *NBM.* 2017;5(3):109-118.
7. Hiszczyńska-Sawicka E, Olędzka G, Holec-Gąsior L, et al. Evaluation of immune responses in sheep induced by DNA immunization with genes encoding GRA1, GRA4, GRA6 and GRA7 antigens of *Toxoplasma gondii*. *Vet Parasitol.* 2011;177(3-4):281-289.
8. Foroutan M, Ghaffarifar F, Sharifi Z, et al. Bioinformatics analysis of ROP8 protein to improve vaccine design against *Toxoplasma gondii*. *Infect Genet Evol.* 2018;62:1-204-93.
9. Kulkarni R, Sapkal G, Mahishi L, et al. Design and characterization of polytope construct with multiple B and TH epitopes of Japanese encephalitis virus. *Virus res.* 2012;166(1-2):77-86.
10. Karpenko LI, Bazhan SI, Antonets DV, et al. Novel approaches in polyepitope T-cell vaccine development against HIV-1. *Expert rev vaccines.* 2014;13(1):155-173.
11. Gershoni JM, Roitburd-Berman A, Siman-

- Tov DD, et al. Epitope mapping. *BioDrugs*. 2007;21(3):145-156.
12. Pomés A. Relevant B cell epitopes in allergic disease. In *arch of allergy immunol*. 2010;152(1):1-11.
 13. Gazzinelli RT, Hakim F, Hieny S, et al. Synergistic role of CD4+ and CD8+ T lymphocytes in IFN-gamma production and protective immunity induced by an attenuated *Toxoplasma gondii* vaccine. *J Immunol*. 1991;146(1):286-292.
 14. Van MR. Structural and functional approaches to the study of protein antigenicity. *Immunol Today*. 1989;10(8):266-272.
 15. Wang Y, Wang M, Wang G, et al. Increased survival time in mice vaccinated with a branched lysine multiple antigenic peptide containing B-and T-cell epitopes from *T. gondii* antigens. *Vaccine*. 2011;29(47):8619-8623.
 16. Dziadek B, Brzostek A. Recombinant ROP2, ROP4, GRA4 and SAG1 antigen-cocktails as possible tools for immunoprophylaxis of toxoplasmosis: what's next?. *Bioengineered*. 2012;3(6):358-364.
 17. Garcia JL, Innes EA, Katzer F. Current progress toward vaccines against *Toxoplasma gondii*. *Vaccine (Auckl)*. 2014;4(1).
 18. Quan J-H, Chu J-Q, Ismail Haha, et al. Induction of protective immune responses by a multiantigenic DNA vaccine encoding GRA7 and ROP1 of *Toxoplasma gondii*. *Clin Vaccine Immunol*. 2012;CVI. 05385-11.
 19. Yang C-S, Yuk J-M, Lee Y-H, et al. *Toxoplasma gondii* GRA7-induced TRAF6 activation contributes to host protective immunity. *Infect Immun*. 2016;84(1):339-350.
 20. Rome ME, Beck JR, Turetzky JM, et al. Intervacuolar transport and unique topology of GRA14, a novel dense granule protein in *Toxoplasma gondii*. *Infect Immun*. 2008;76(11):4865-4875.
 21. Gasteiger E, Hoogland C, Gattiker A, et al. Protein identification and analysis tools on the Expasy server. *The proteomics protocols handbook*: Springer; 2005. p. 571-607.
 22. Garnier J, Gibrat J-F, Robson B. GOR method for predicting protein secondary structure from amino acid sequence. *Methods Enzymol*. Vol. 266: Elsevier; 1996. p. 540-553.
 23. Krogh A, Larsson B, Von Heijne G, et al. Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes. *J Mol Biol*. 2001;305(3):567-580.
 24. Armenteros JJA, Tsirigos KD, Sønderby CK, et al. SignalP 5.0 improves signal peptide predictions using deep neural networks. *Nat Biotechnol*. 2019; 37(4):420-423.
 25. Yang J, Zhang Y. I-TASSER server: new development for protein structure and function predictions. *Nucleic acids research*. 2015;43(W1):W174-W181.
 26. Zhang C, Freddolino PL, Zhang Y. COFACTOR: improved protein function prediction by combining structure, sequence and protein-protein interaction information. *Nucleic Acids res*. 2017;45(W1):W291-W299.
 27. Wu S, Skolnick J, Zhang Y. Ab initio modeling of small proteins by iterative TASSER simulations. *BMC Biol*. 2007;5(1):17.
 28. Laskowski RA, MacArthur MW, Moss DS, et al. PROCHECK: a program to check the stereochemical quality of protein structures. *J Appl Cryst*. 1993;26(2):283-291.
 29. Saha S, Raghava G, editors. BcePred: prediction of continuous B-cell epitopes in antigenic sequences using physico-chemical properties. *International Conference on Artificial Immune Systems*; 2004: Springer.
 30. Singh H, Ansari HR, Raghava GP. Improved method for linear B-cell epitope prediction using antigen's primary sequence. *PloS one*. 2013;8(5):e62216.
 31. Yao B, Zhang L, Liang S, et al. SVMTriP: a method to predict antigenic epitopes using support vector machine to integrate tri-peptide similarity and propensity. *PloS one*. 2012;7(9):e45152.
 32. Chou P, Fasman G. Prediction of the secondary structure of proteins from their amino acid sequence. *Adv Enzymol Relat Areas Mol Biol* 1978; 47: 45-148..
 33. Karplus P, Schulz G. Prediction of chain flexibility in proteins. *Naturwissenschaften*. 1985;72(4):212-213.
 34. Emini EA, Hughes JV, Perlow D, et al. Induction of hepatitis A virus-neutralizing antibody by a virus-specific synthetic peptide. *J Virol*. 1985;55(3):836-839.
 35. Parker J, Guo D, Hodges R. New hydrophilicity scale derived from high-performance liquid chromatography peptide retention data: correlation of predicted surface residues with antigenicity and X-ray-derived accessible sites. *Biochemistry*.

- 1986;25(19):5425-5432.
36. Kolaskar A, Tongaonkar PC. A semi-empirical method for prediction of antigenic determinants on protein antigens. *FEBS letters*. 1990;276(1-2):172-174.
 37. Kringelum JV, Lundegaard C, Lund O, et al. Reliable B cell epitope predictions: impacts of method development and improved benchmarking. *PLoS Comput Biol*. 2012;8(12):e1002829.
 38. Ansari HR, Raghava GP. Identification of conformational B-cell Epitopes in an antigen from its primary sequence. *Immunome Res*. 2010;6(1):6.
 39. Zhao L, Zhang M, Cong H. Advances in the study of HLA-restricted epitope vaccines. *Human vaccines & immunotherapeutics*. 2013;9(12):2577-66.
 40. Argos P. An investigation of oligopeptides linking domains in protein tertiary structures and possible candidates for general gene fusion. *J Mol Biol*. 1990;211(4):943-958.
 41. Grote A, Hiller K, Scheer M, et al. JCat: a novel tool to adapt codon usage of a target gene to its potential expression host. *Nucleic Acids Res*. 2005;33(suppl_2):W526-W531.
 42. Kingston RE, Chen CA, Rose JK. Calcium phosphate transfection. *Current protocols in molecular biology*. 2003;63(1):9.1.1- 9.1. 11
 43. Wang Y, Wang G, Cai J, et al. Review on the identification and role of *Toxoplasma gondii* antigenic epitopes. *Parasitol Res*. 2016;115(2):459-468.
 44. Lu G, Wang L, Zhou A, et al. Epitope analysis, expression and protection of SAG5A vaccine against *Toxoplasma gondii*. *Acta Trop*. 2015;146:66-72.
 45. Wang Y, Wang G, Zhang D, et al. Screening and identification of novel B cell epitopes of *Toxoplasma gondii* SAG1. *Parasit Vectors*. 2013;6(1):125.
 46. Romano P, Giugno R, Pulvirenti A. Tools and collaborative environments for bioinformatics research. *Brief Bioinform*. 2011;12(6):549-561.
 47. Wang Y, Zhang D, Yin H, et al. Advances in predicting methods of antigen epitopes. *Chinese Veterinary Science/Zhongguo Shouyi Kexue*. 2009; 39(10):938-940
 48. Garnier J, Robson B. The GOR method for predicting secondary structures in proteins. *Prediction of protein structure and the principles of protein conformation*: Springer; 1989. p. 417-465.
 49. Hopp TP, Woods KR. Prediction of protein antigenic determinants from amino acid sequences. *Proc. Natl. Acad. Sci*. 1981;78(6):3824-3828.
 50. Saha S, Raghava G. Prediction of continuous B-cell epitopes in an antigen using recurrent neural network. *Proteins : Structure, Function, and Bioinformatics*. 2006;65(1):40-48.
 51. Liu S, Shi L, Cheng Y-b, Fan G-x, Ren H-x, Yuan Y-k. Evaluation of protective effect of multi-epitope DNA vaccine encoding six antigen segments of *Toxoplasma gondii* in mice. *Parasitol.Res* 2009;105(1):267.
 52. Hajissa K, Zakaria R, Suppian R, Mohamed Z. Immunogenicity of Multiepitope Vaccine Candidate against *Toxoplasma gondii* Infection in BALB/c Mice. *Iran J PARASITOL*. 2018;13(2):215.