Original article

CLINICAL AND EXPERIMENTAL VACCINE RESEARCH

Clin Exp Vaccine Res 2021;10:59-77
https://doi.org/10.7774/cevr.2021.10.1.59
pISSN 2287-3651 • eISSN 2287-366X

Immunoinformatic analysis of immunogenic B- and T-cell epitopes of MIC4 protein to designing a vaccine candidate against *Toxoplasma gondii* through an in-silico approach

Ali Dalir Ghaffari, Abdolhossein Dalimi, Fatemeh Ghaffarifar, Majid Pirestani, Hamidreza Majidiani

Department of Parasitology, Faculty of Medical Sciences, Tarbiat Modares University, Tehran, Iran

Received: October 6, 2020
Accepted: January 25, 2021

Corresponding author: Abdolhossein Dalimi, PhD
Department of Parasitology, Faculty of Medical Sciences, Tarbiat Modares University, Tehran, P.O. Box 14115-331, Iran
Tel: +98-2182883838, Fax: +98-218013030
E-mail: dalimi_a@modares.ac.ir

No potential conflict of interest relevant to this article was reported.

This study was financially supported by Tarbiat Modares University (grant no., 52D/1649, 16 June 2019) and Iran National Science Foundation (INSF) (project no. of 98008896) and Tarbiat Modares University. The authors would like to thank all staff of Department of Parasitology of Tarbiat Modares University as well as Iran National Science Foundation (INSF), Iran. This paper is issued from thesis of Ali Dalir Ghaffari, PhD student of Medical Parasitology.

Purpose: Toxoplasmosis, transmitted by *Toxoplasma gondii*, is a worldwide parasitic disease that affects approximately one-third of the world's inhabitants. Today, there are no appropriate drugs to deter tissue cysts from developing in infected hosts. So, developing an effective vaccine would be valuable to avoid from toxoplasmosis. Considering the role of microneme antigens such as microneme protein 4 (MIC4) in *T. gondii* pathogenesis, it can be used as potential candidates for vaccine against *T. gondii*.

Materials and Methods: In this study several bioinformatics methods were used to assess the different aspects of MIC4 protein such as secondary and tertiary structure, physicochemical characteristics, the transmembrane domains, subcellular localization, B-cell, helper-T lymphocyte, cytotoxic-T lymphocyte epitopes, and other notable characteristic of this protein design a suitable vaccine against *T. gondii*.

Results: The studies revealed that MIC4 protein includes 59 potential post-translational modification sites without any transmembrane domains. Moreover, several probable epitopes of B- and T-cells were detected for MIC4. The secondary structure comprised 55.69% random coil, 5.86% beta-turn, 19.31% extended strand, and 19.14% alpha helix. According to the Ramachandran plot results, 87.42% of the amino acid residues were located in the favored, 9.44% in allowed, and 3.14% in outlier regions. The protein allergenicity and antigenicity revealed that it was non-allergenic and antigenic.

Conclusion: This study gives vital basic on MIC4 protein for further research and also established an effective vaccine with different techniques against acute and chronic toxoplasmosis.

Keywords: *Toxoplasma gondii*, MIC4 protein, Bioinformatics, Vaccine

Introduction

Toxoplasmosis, caused by *Toxoplasma gondii*, is a worldwide parasitic disease that affects approximately one-third of the world's population [1,2]. Cats are the only ultimate hosts, and many warm-blooded animals, such as humans, birds, rodents, etc., act as intermediate hosts [3]. Toxoplasmosis is a life-threatening and severe infectious disease which also leads to economic damage by affecting animals, especially in pigs, goats, and sheep [4]. The primary human infection is generally asymptomatic or

causes just minor symptoms in hosts with intact immunity [5]. Nonetheless, *T. gondii* may be fatal in individuals with compromised immune systems, particularly those with acquired immunodeficiency syndrome, undergoing organ transplantation, and in patients with malignant tumors who undergo cancer treatment due to latent cyst reactivation [6]. Additionally, when transmitted to the embryo through pregnancy, the infection could lead to abortion, congenital abnormalities, or other disorders caused during the growth of the fetus [7,8]. Toxoplasmosis drugs actually constitute a combination of pyrimethamine and sulfadiazine that have some side effects. Additionally, these drugs are ineffective and costly, which would cause dangerous hypersensitivity and teratogenic responses to the embryo, and cannot eradicate bradyzoites in tissue cysts [9]. Thus, a vaccine produced against *T. gondii* could be especially helpful in managing toxoplasmosis in humans and animals [10].

From various research reports, excretory-secretory antigens of *T. gondii* have been recognized as a significant element during the replication and invading of tachyzoites in host cells and are regarded as the main objectives of host immune responses [11]. Considerable progress has been made in recent years to recognize vaccine candidates for both chronic and acute toxoplasmosis which could encourage successful immune response. Most of the *T. gondii* vaccine development research is concentrated on dense granule antigens, surface antigens, microneme antigens (MICs), rhoptry antigens, and other antigens [12-15]. Among these, MICs are small apical organelles with many adhesive proteins that are secreted with the host plasma membrane during initial contact with the apical end of the parasite [16,17]. MICs are not only essential for attachment to the entrance of the host cell, but are also crucial for parasite gliding because their cytoplasmic domains can bind to aldolase linked to the parasite actin-myosin motor, which is the basis for active invasion [18]. Today, the MIC protein is shown to play a significant and predominant role in virulence and pathogenicity [19]. Large numbers of studies on MICs have shown that they are potent antigen targets and vaccine candidates for potent immune responses against toxoplasmosis [6]. Microneme protein 4 (MIC4) locates in all of the invasive types of *T. gondii* including sporozoites, bradyzoites, tachyzoites, and merozoites [20].

Bioinformatics approaches are now extensively being used to identify the potential T- and B-cell epitopes for identification and construction of vaccine candidates; thus, such methods are ideal for selecting immunodominant epitopes [21]. These techniques have been routinely used to analyze protein and gene expression and to determine the structural, immunogenic and general characteristics of proteins. Researching and analyzing the physical, chemical, and immunogenic properties of proteins will improve our awareness of them and allow the researcher determine the correct epitopes for vaccine construction [22]. Bioinformatics has some benefits over traditional approaches, such as relatively cost-effectiveness and time-effectiveness, high accuracy, etc. [23,24]. Hence, the identification of protein epitope features via bioinformatics methods will be beneficial for diagnostic aims and vaccine research [25].

Therefore, we utilized bioinformatics approaches in this study to evaluate the physicochemical characteristics, tertiary and secondary structure, and assess the B- and T-cell epitopes of the MIC4 protein for the design of a suitable *T. gondii* vaccine.

Materials and Methods

Ethics approval

This study was approved by the Ethical Committee of Tarbiat Modares University (IR.MODARES.REC.1398.009).

Retrieval of MIC4 protein sequence of *T. gondii*

In the first step, MIC4's complete amino acid sequence was acquired from a publicly available sequence database, the National Center for Biotechnology Information (NCBI) (https://www.ncbi.nlm.nih.gov/protein/).

Analysis of the MIC4 protein physicochemical properties

To evaluate the physicochemical properties of MIC4 protein (including number of amino acids, molecular weight [MW], aliphatic index, instability index, total number of positive and negative charged residues, theoretical isoelectric point [pI], estimated half-life in mammal's yeast, Escherichia coli, and reticulocytes, extinction coefficients, and grand average of hydropathicity [GRAVY]), the expasy protParam tool was used (https://web.expasy.org/protparam/) [26].

Prediction of post-translational modification sites of MIC4

To analyze the phosphorylation and acylation sites of MIC4 protein, NetPhos 3.1 and CSS-Palm servers were used, respectively [27,28].

 https://doi.org/10.7774/cevr.2021.10.1.59

Prediction of transmembrane domains and subcellular localization of MIC4 protein

The transmembrane domains of the MIC4 protein were predicted by the TMHMM ver. 2.0 at http://www.cbs.dtu.dk/services/TMHMM-2.0/ and PSORT II at http://psort.hgc.jp/form2.html servers, respectively [27].

Secondary structure analysis

In order to construct the secondary structure of the MIC4, Garnier-Osguthorpe-Robson (GOR) and PSIPRED servers were applied [29,30]. We also carried out prediction analyzes using SOPMA online research tools to predict the protein's secondary structure further precisely at https://npsa-prabi.ibcp.fr/cgibin/npsa_automat.pl?page=npsa_sopma.html to enhance the validity of the predicted secondary protein structure [31]. Next, DiANNA online database (unified software for cysteine state and disulfide bond partner prediction) was used (http://clavius.bc.edu/~clotelab/DiANNA/) [32].

Homology modeling and validation

The three-dimensional (3D) structures of the MIC4 protein were built via SWISS-MODEL (https://swissmodel.expasy.org/) [33]. Then the best model (generated via SWISS-MODEL) was selected and improved by the Galaxy refine server. The GalaxyRefine database first rebuilds side chains and conducts side chain repacking and next overall structure relaxation via molecular dynamic simulation [34]. To know the quality and precision of the models, they assessed by the Ramachandran plot utilizing SWISS-MODEL server at https://swissmodel.expasy.org/assess [35]. Also, the overall quality of the constructed model was analyzed via ProSAweb at https://prosa.services.came.sbg.ac.at/prosa.php [36].

Epitope prediction

The inducing epitopes including B- and T-cell were determined using MIC4 protein amino acid sequences.

Linear (continuous) and conformational (discontinuous) B-cell epitopes of the MIC4 protein

Multiple databases were used to analyze the linear epitopes of the B-cells. We initially used a web-based Bcepred (B-cell epitope prediction) server to determine linear B-cell epitopes utilizing physicochemical characteristics (http://crdd.osdd.net/raghava/bcepred/bcepred_submission.html). This database can calculate the greatest accuracy at threshold 2.38 by 58.70% and allows users to predict B-cell epitopes through any of the chemical and physical properties (including hydrophilicity, accessibility, exposed surface, turns, polarity, and flexibility/mobility; http://crdd.osdd.net/raghava/bcepred/) [37]. Besides, an online tool of ABCpred (B-cell epitope prediction based on an artificial neural network) was used to determine B-cell epitopes in an antigen sequence (http://crdd.osdd.net/raghava/abcpred/) [38]. ProtScale online tool was applied to predict linear B-cell epitopes according to hydrophobicity, alpha-helix, beta-turn, average flexibility, and percent of accessible residue (https://web.expasy.org/protscale/) [26]. Moreover, SVMTriP (http://sysbio.unl.edu/SVMTriP/prediction.php) [39] and Bepipred 1.0 servers (http://www.cbs.dtu.dk/services/BepiPred-1.0/) [40] were used to analyze the B-cell linear epitopes. Default parameters have been used for estimation using the Bcepred database. In the ABCpred server, the specificity and linear epitopes were 75% and 20-mer, respectively with the use overlap filter. On the SVMTriP server, linear epitopes were 20-mer length. Also, the epitope assignment value threshold was 0.35 on the Bepipred 1.0 server. Besides, discontinuous B-cell epitopes were estimated using ElliPro (http://tools.iedb.org/ellipro/) from the 3D epitope structure protein data bank file [41]. The default minimum value of 0.5 and the utmost distance (Angstrom) of were used for the analysis.

Prediction of T-cell epitope

CTL epitopes

The IEDB (http://tools.iedb.org/mhci/) [42] and NetMHCcons 1.1 (http://www.cbs.dtu.dk/services/NetMHCcons/) [43] online servers were used to predict peptides binding to major histocompatibility complex (MHC)-I molecule. IEDB suggests making choices based on a percentile value of ≤1% for each (MHC allele, length) mixture to cover most of the immune responses [44,45]. In the IEDB server, the MHC-I prediction was made using the IEDB suggested method with a 10-mer length. NetMHCcons allow the individual to select the MHC molecule from a long list of alleles which upload the MHC protein region of interest [43]. Prediction values are given in the nano molars (nM) IC50 values and are rank percent. A peptide is classified in NetMHCcons as a strong binder if the rank percentage is below 0.5% or the binding affinity (or IC50) is below 50 nM. The peptide is also regarded to be a poor binder if the percentage is below 2% or if the binding affinity (or IC50) is below 500 nM. Six alleles (H2-Db, H2-Dd, H2-Kb, H2-Kd, H2-Kk, and H2-Ld) were chosen as molecules of the MHC class I mouse. Additionally, we predicted cyto-

toxic-T lymphocyte (CTL) epitopes using the CTLpred tool which available at http://www.imtech.res.in/raghava/ctlpred/index.html [46]. CTLpred is a method of determining CTL epitopes which are important to the vaccine design development. The combined approach used for prediction. By default, the Artificial Neural Network (ANN) and Support Vector Machine (SVM) cutoff ratings were set at 0.51 and 0.36, respectively. The precision of the combined technique of prediction was 75.8%.

Helper-T lymphocyte epitopes

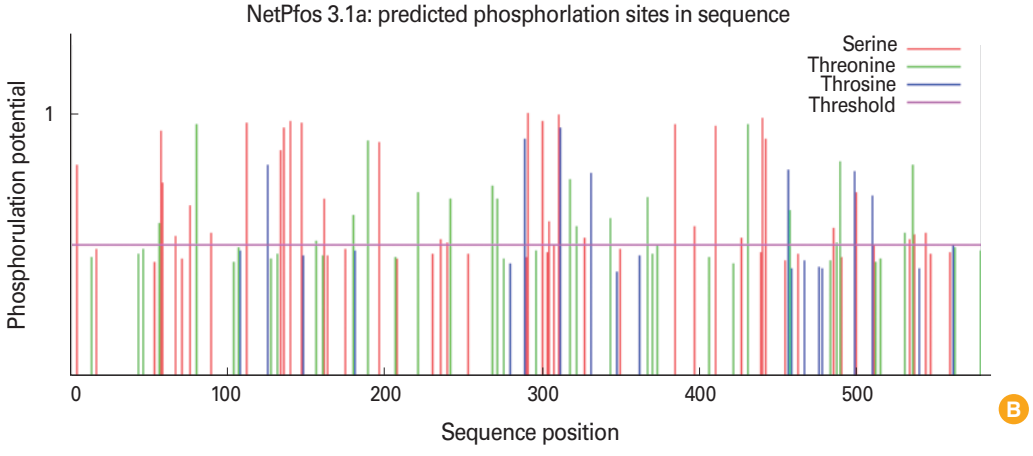
The IEDB (http://tools.immuneepitope.org/mhcii) [47] and NetMHCIIpan 3.2 (http://www.cbs.dtu.dk/services/NetMHCIIpan/) [48] servers were applied to evaluate the 15-mer T-cell epitopes of H-2-IEd, H2IAd, and H2IAb mouse alleles. In the IEDB method, the prediction was performed based on the suggested IEDB method with sort by percentile rank. The NetMHCIIpan 3.2 method classified peptides as powerful, intermediate, and non-binding by percentile rank. For powerful, intermediate, and non-binding, the cutoff was carried at 2%, 10%, and higher than 10%.

Peptide antigenicity, immunogenicity, and solubility evaluation

To evaluate the T- and B-cell epitopes antigenicity, the ANTI-

```
    MRASLPVHLVVCTQLSAVWFGVAKAHGGHRLEPHVPGFLQGFTDITPAGD   #     50
    DVSANVTSSEPAKLDLSCVHSDNKGSRAPTIGEPVPDVSLEQCAAQCKAV   #    100
    DGCTHFTYNDDSKMCHVKEGKPDLYDLTGGKTASRSCDRSCFEQHVSYEG   #    150
    APDVMTAMVTSQSADCQAACAADPSCEIFTYNEHDQKCTFKGRGFSAFKE   #    200
    RGVLGVTSGPKQFCDEGGKLTQEEMEDQISGCIQLSDVGSMTADLEEPME   #    250
    ADSVGACMERCRCDGRCTHFTFNDNTRMCYLKGDKMQLYSSPGDRTGPKS   #    300
    CDSSCFSNGVSYVDDPATDVETVFEISHPIYCQVICAANPLCTVFQWYAS   #    350
    EAKCVVKRKGFYKHRKTGVTGVTVGPREFCDFGGSIRDREEADAVGSDDG   #    400
    LNAEATMANSPDFHDEVECVHTGNIGSKAQTIGEVKHASSLSECRARCQA   #    450
    EKECSHYTYNVKSGLCYPKRGKPQFYKYLGDMTGSRTCDTSCLRRGVDYS   #    500
    QGPEVGKPWYSTLPTDCQVACDAEDACLVFTWDSATSRCYLIGSGFSAHR   #    550
    RNDVDGVVSGPYTFCDNGENLQVLEAKDTE                       #    600
%1  ...S.............................................     #     50
%1  ......TSS.......S........S...T........S...........   #    100
%1  ...........S............Y........S.S...S......S...   #    150
%1  .....T....S..................T........T......S....   #    200
%1  ....................T..............S...S.T........   #    250
%1  .................T..T.................Y.S........S   #    300
%1  ...S..S...SY.....T...T....S...Y...........T.......   #    350
%1  ................T.................S...........S...   #    400
%1  .........S.................S...T........S.S........   #    450
%1  ......YT..........................S.T..T........YS   #    500
%1  .........Y....................T..S.TS......S......   #    550
%1  ...........Y..................
```

A



B

Fig. 1. Bioinformatics analysis of the phosphorylation and acylation locations of the microneme protein 4 (MIC4). (A) If the residue is predicted not to be phosphorylated, either because the score is below the threshold or because the residue is not serine, threonine, or tyrosine, that position is marked by a dot ('.'). Residues having a prediction score above the threshold are indicated by 'S', 'T', or 'Y', respectively. (B) Expected phosphorylation positions in MIC4 sequence.

 https://doi.org/10.7774/cevr.2021.10.1.59

GENpro (http://scratch.proteomics.ics.uci.edu/) [49] and VaxiJen ver. 2.0 (http://www.ddg-pharmfac.net/vaxijen/VaxiJen/VaxiJen.html) [50] tools were utilized. VaxiJen database accuracy differs from 70% to 89% depending on the target organisms (http://www.ddg-pharmfac.net/vaxijen/VaxiJen/VaxiJen_help.html). Moreover, the vaccine's allergic and non-allergic nature was determined by AllerTOP V2.0 database (www.ddg-pharmfac.net/AllerTOP/) [51]. Protein solubility overexpression prediction was also estimated at http://scratch.proteomics.ics.uci.edu/ [52].

The prediction of dominant T- and B-cell epitopes of the MIC4 protein

The overlapping peptides were described as dominant epitopes, based on the epitopes mentioned.

Table 1. The acylation sites of MIC4 sequence

ID	Position	Peptide	Score
ACY68633.1 MIC4 [*T. gondii*]	12	LPVHLVVCTQLSAVW	31.467
ACY68633.1 MIC4 [*T. gondii*]	68	PAKLDLSCVHSDNKG	7.601
ACY68633.1 MIC4 [*T. gondii*]	93	PDVSLEQCAAQCKAV	1.569
ACY68633.1 MIC4 [*T. gondii*]	97	LEQCAAQCKAVDGCT	3.99
ACY68633.1 MIC4 [*T. gondii*]	103	QCKAVDGCTHFTYND	7.277
ACY68633.1 MIC4 [T. gondii]	115	YNDDSKMCHVKEGKP	2.64
ACY68633.1 MIC4 [*T. gondii*]	137	GKTASRSCDRSCFEQ	7.705
ACY68633.1 MIC4 [*T. gondii*]	141	SRSCDRSCFEQHVSY	0.725
ACY68633.1 MIC4 [*T. gondii*]	166	VTSQSADCQAACAAD	2.024
ACY68633.1 MIC4 [*T. gondii*]	170	SADCQAACAADPSCE	2.534
ACY68633.1 MIC4 [*T. gondii*]	176	ACAADPSCEIFTYNE	6.569
ACY68633.1 MIC4 [*T. gondii*]	188	YNEHDQKCTFKGRGF	6.49
ACY68633.1 MIC4 [*T. gondii*]	214	TSGPKQFCDEGGKLT	3.175
ACY68633.1 MIC4 [*T. gondii*]	232	MEDQISGCIQLSDVG	3.863
ACY68633.1 MIC4 [*T. gondii*]	257	EADSVGACMERCRCD	3.423
ACY68633.1 MIC4 [*T. gondii*]	261	VGACMERCRCDGRCT	4.003
ACY68633.1 MIC4 [*T. gondii*]	263	ACMERCRCDGRCTHF	2.077
ACY68633.1 MIC4 [*T. gondii*]	267	RCRCDGRCTHFTFND	9.699
ACY68633.1 MIC4 [*T. gondii*]	279	FNDNTRMCYLKGDKM	2.612
ACY68633.1 MIC4 [*T. gondii*]	301	DRTGPKSCDSSCFSN	8.477
ACY68633.1 MIC4 [*T. gondii*]	305	PKSCDSSCFSNGVSY	5.576
ACY68633.1 MIC4 [*T. gondii*]	332	EISHPIYCQVICAAN	1.072
ACY68633.1 MIC4 [*T. gondii*]	336	PIYCQVICAANPLCT	2.099
ACY68633.1 MIC4 [*T. gondii*]	342	ICAANPLCTVFQWYA	1.86
ACY68633.1 MIC4 [*T. gondii*]	354	WYASEAKCVVKRKGF	0.249
ACY68633.1 MIC4 [*T. gondii*]	380	TVGPREFCDFGGSIR	1.447
ACY68633.1 MIC4 [*T. gondii*]	419	DFHDEVECVHTGNIG	4.389
ACY68633.1 MIC4 [*T. gondii*]	444	HASSLSECRARCQAE	1.364
ACY68633.1 MIC4 [*T. gondii*]	448	LSECRARCQAEKECS	1.188
ACY68633.1 MIC4 [*T. gondii*]	454	RCQAEKECSHYTYNV	5.774
ACY68633.1 MIC4 [*T. gondii*]	466	YNVKSGLCYPKRGKP	0.713
ACY68633.1 MIC4 [*T. gondii*]	488	DMTGSRTCDTSCLRR	4.209
ACY68633.1 MIC4 [*T. gondii*]	492	SRTCDTSCLRRGVDY	9.437
ACY68633.1 MIC4 [*T. gondii*]	517	YSTLPTDCQVACDAE	1.031
ACY68633.1 MIC4 [*T. gondii*]	521	PTDCQVACDAEDACL	8.059
ACY68633.1 MIC4 [*T. gondii*]	527	ACDAEDACLVFTWDS	5.3
ACY68633.1 MIC4 [*T. gondii*]	539	WDSATSRCYLIGSGF	7.001
ACY68633.1 MIC4 [*T. gondii*]	565	VSGPYTFCDNGENLQ	1.197

MIC4, microneme protein 4; *T. gondii*, *Toxoplasma gondii*.

Results

Gene details and the basic features of MIC4 protein

The amino acid sequence of the MIC4 protein was extracted from the NCBI under accession no. ACY68633.1 in the format of FASTA. By using the Expasy ProtParam server, we noticed that the MIC4 protein consists of 580 amino acid residues with the MW of 63,002 KDa, and its theoretical pI is 5.04 which already was known by Brecht et al. [53]. The total number of negatively (Asp+Glu) and positively charged residues (Arg+Lys) were 83 and 56, respectively. The extinction coefficient was 52685 M^{-1} cm^{-1} at 280 nm in water. The predicted half-life was 30 hours in mammalian reticulocytes, >20 hours in yeast, and >10 hours in *E. coli*. The index of instability for this protein was estimated at 42.72, which categorizes the protein as unstable. Furthermore, the aliphatic index was 55.64 and the GRAVY of this protein was -0.451.

Prediction of post-translational modification sites of MIC4

Analysis results of NetPhos 3.1 and CSS-Palm servers revealed that MIC4 protein has 59 sites of phosphorylation (serine: 32, threonine: 19, tyrosine: 8) (Fig. 1A, B) and 38 acylation sites. The acylation sites of the MIC4 sequence are shown in Table 1.

Prediction of transmembrane domains and subcellular localization of MIC4 protein

Since MIC4 is a secreted protein it is expected not to have transmembrane domains. So, the TMHMM ver. 2.0 server findings demonstrated that the MIC4 protein had no transmembrane domain, which is shown in Fig. 2. The protein transmembrane region was larger than 1, indicates that it is an extracellular protein. Moreover, the subcellular localization prediction of MIC4 utilizing PSORT II was as follows: 34.8% mitochondrial, 13.0% cytoplasmic, 4.3% vacuolar, 39.1% nuclear, and 8.7% cytoskeletal.

Secondary structures analysis

The prediction of the secondary and 3D structures of the MIC4 protein has a major effect on its biological function. To identify the secondary structure of MIC4 protein, we used GOR IV, SOPMA, and PSIPRED online servers. The results of the GOR IV server showed that the percentages of the random coil, alpha-helix, and extended strand in the MIC4 sequence were

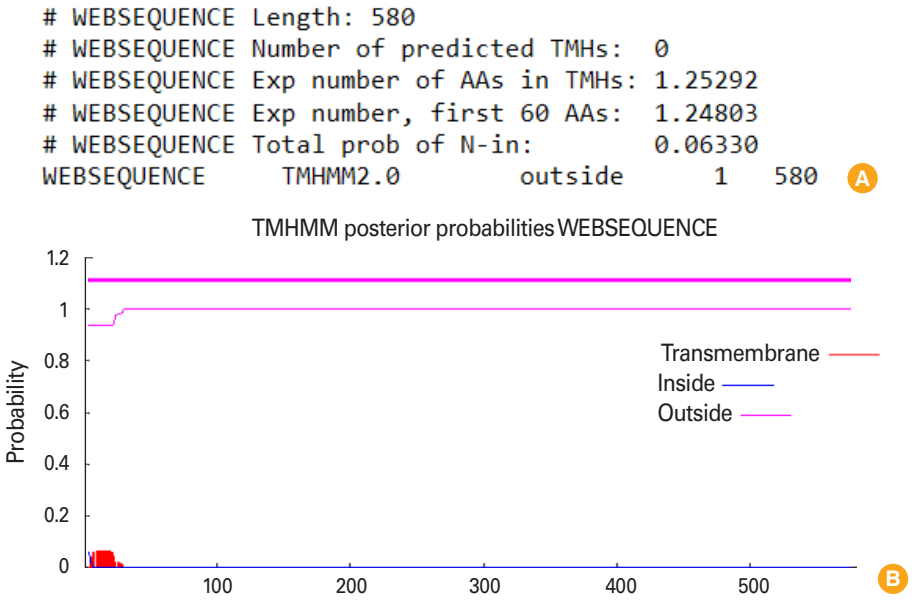


Fig. 2. Transmembrane helices expected in microneme protein 4 (MIC4). (A) Some statistics and a list of the location of the predicted transmembrane helices and the predicted location of the intervening loop regions. Number of predicted TMHs: the number of predicted transmembrane helices; Exp number of AAs in TMHs: the expected number of amino acids in transmembrane helices. If this number is larger than 18, it is very likely to be a transmembrane protein (or have a signal peptide); Exp number, first 60 AAs: the expected number of amino acids in transmembrane helices in the first 60 amino acids of the protein. If this number more than a few, you should be warned that a predicted transmembrane helix in the N-term could be a signal peptide; Total prob. of N-in: the total probability that the N-term is on the cytoplasmic side of the membrane; Possible N-term signal sequence: a warning that is produced when "Exp number, first 60 AAs" is larger than 10. (B) Analysis of MIC4 transmembrane domain.

 https://doi.org/10.7774/cevr.2021.10.1.59

61.72% (358/580), 10.17% (59/580), and 28.10% (163/580), respectively (Figs. 3, 4). In addition to abovementioned software, we employed SOPMA tools to assess the MIC4 secondary structures. The software illustrated graphically the prediction of secondary structure of protein and the percentage of different basic structures including alpha-helix, extended strand, beta-turn, and random coil. MIC4 sequence composition was as follows: 19.14% (111/580) alpha helix, 19.31% (112/580) extended strand, 5.86% (34/580) beta-turn, and 55.69% (323/580) random coil (Fig. 5, Table 2). The estimated number of cysteines utilizing DiANNA software in the MIC4 sequence was 38. More information is indicated in Table 3.

Homology modeling and validation

SWISS-MODEL online server was utilized to determine and construct the 3D structure of MIC4 protein. SWISS-MODEL results for MIC4 showed that 11 models were found to match the target sequence. So, the model with the high sequence identity and coverage which has 100% sequence identity in all templates was chosen. The output of SWISS-MODEL, like the 3D model predicted for MIC4, protein global quality estimate, sequence identity and coverage, model-template alignment, and local quality estimation is shown in Fig. 6. The z-score showing overall model quality was -4.37, and a high number of residues have been found in the favored region. The quality of the 3D structure was enhanced after refine-

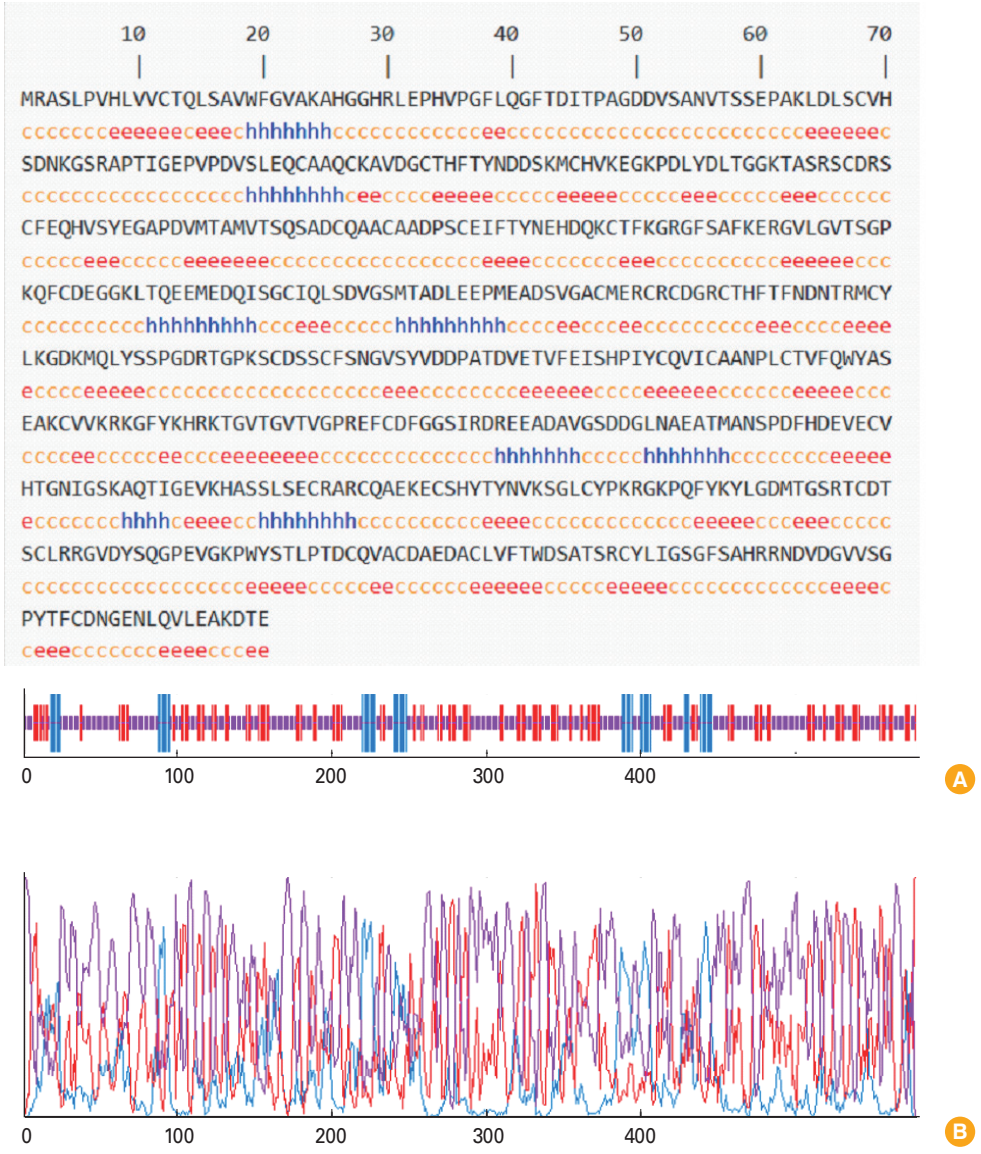


Fig. 3. Microneme protein 4 (MIC4) secondary structural analysis using GOR IV. (A) Predicted secondary structure. (B) Graphical outcomes of secondary structure prediction of MIC4 protein.

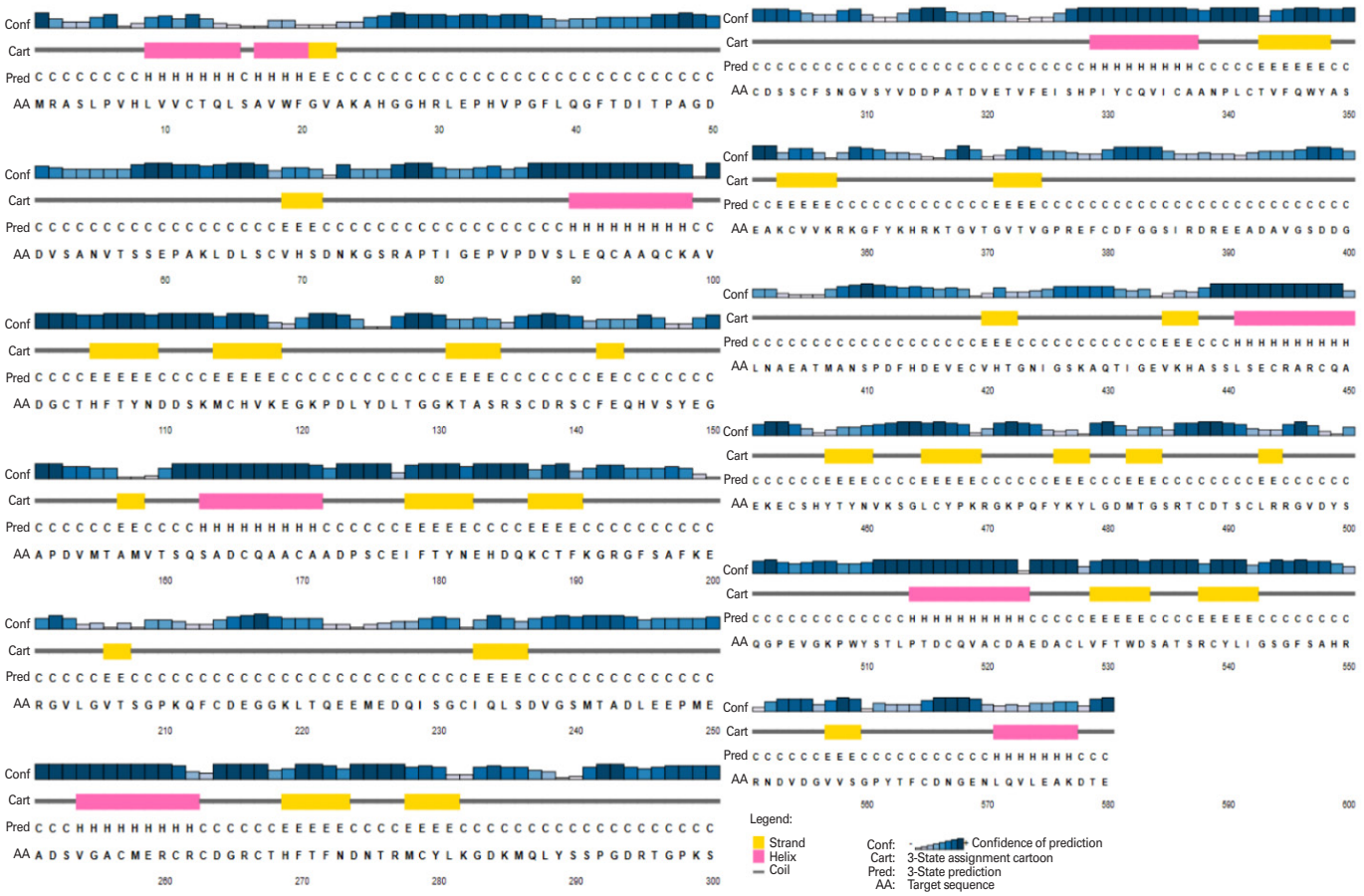


Fig. 4. Graphical outcome of vaccine estimation with secondary structure using PSIPRED.

ment and the z-score was -4.51 (Fig. 7A). Protein validation before refinement showed that in the initial model, 87.42% of residues were located in favored, 9.44% in allowed and 3.14% in outlier regions. After refinement of the 3D model, the results were changed as follows: 96.23% of residues in favored regions, 2.51% in allowed regions, and 1.26% in outlier regions of Ramachandran plot (Fig. 7B).

Epitope prediction

Linear and conformational B-cell epitopes of the MIC4 protein

An epitope analysis can give researchers significant knowledge in order to recognize immunogenic peptides and develop new potential vaccines. The Bcepred, ABCpred, ProtScale, SVMTriP, and Bepipred online server prediction were used to predict the linear epitopes of the MIC4 protein. The predicted epitopes of the Bcepred to predict the linear B-cell epitopes, utilizing physicochemical characteristics (polarity, exposed surface, hydrophilicity, turns, flexibility/mobility exposed surface, and accessibility) are shown in Table 4. These characteristics are very important for the antigenic characteristics of the MIC4 protein. The findings of the ABCpred server are also listed in Table 5 by their scores (only the epitopes over scores of 0.75 are listed in the Table 5). The higher peptide score suggests the greater chance of being an epitope. This server predicted 36 epitopes over 0.75 scores on MIC4 sequence, in which the highest score was for linear epitope KGSRAPTIGEPVPDVSLEQC (0.90). The ProtScale server was employed for graphical prediction of linear B-cell epitopes based on percent of accessible residues, average flexibility, hydrophobicity, alpha-helix, and beta turn (Fig. 8). Also, The SVMTriP and Bepipred findings are presented in Tables 6 and 7, respectively. Also, in the 3D model of the ElliPro, was predicted three discontinuous B-cell epitopes (Table 8).

Prediction of T-cell epitopes

The IEDB, NetMHCcons, and NetMHCIIpan online tools were recruited to estimate the IC50 values for peptide binding to the MHC class I and class II molecules of MIC4. Bioinformatics analysis established T-cell epitopes on MIC4 capable to bind strongly to molecules of MHC class I and class II.

 https://doi.org/10.7774/cevr.2021.10.1.59

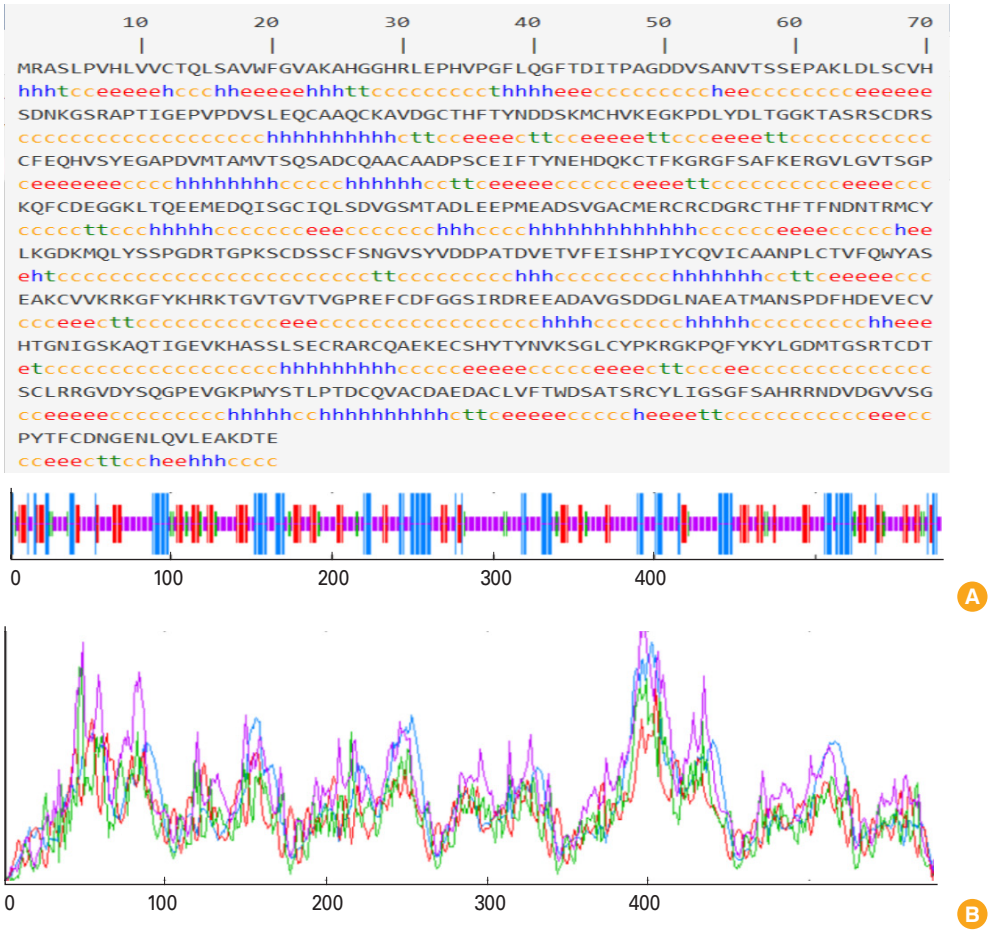


Fig. 5. Microneme protein 4 secondary structural analysis using SOMPA. (A) Visualize the prediction; (B) curves for all predicted states.

Table 2. Secondary structure of the MIC4 protein predicted via GOR IV and SOPMA

Methods	Alpha-helix region	Extended strand	Beta turn	Random coil
GOR IV	10.17 (59/580)	28.10 (163/580)	-	61.72 (358/580)
SOPMA	19.14 (111/580)	19.31 (112/580)	5.86 (34/580)	55.69 (323/580)

Values are presented as % (number/total number).
MIC4, microneme protein 4; GOR, Garnier-Osguthorpe-Robson.

Three peptides with high affinity to MHC molecules were defined for each allele. The minimum percentile ranks for each MHC allele of MIC4 from the IEDB server are shown in Tables 9 and 10. Also, the results of NetMHCcons and NetMHCIIpan are described in Tables 11 and 12, respectively. The information portrayed in the table involving the used alleles, the predicted peptide, the Affinity/IC50 (nM), the percent rank, and the binding level (strong or weak). Ultimately, the results indicated that some MIC4 protein epitopes would bind strongly to the MHC-I and MHC-II molecules. Also, the 10 high-ranking epitopes were selected based on their scores by CTLpred. More details are listed in Table 13.

Peptide antigenicity, immunogenicity, and solubility evaluation

Antigenicity of the MIC4 protein was predicted by ANTIGENpro and VaxiJen v2.0. The antigenicity of the vaccine was 0.9596 and 0.6182 by ANTIGENpro and VaxiJen v.2.0 (threshold for this model was 0.5), respectively, which indicates the antigenic nature of the vaccine. The protein allergenicity was assessed by the AllerTOP V2.0 server, which indicated that the MIC4 is a non-allergen. The SOLpro server estimated the predicted solubility upon overexpression in *E. coli* at 0.7170.

The T- and B-cell dominant epitope predictions

Based on the dominant epitope evaluation, eight dominant

epitopes (five for B-cells and three for T-cells) for MIC4 protein were predicted. More details are shown in Table 14. Among the dominant epitopes, all of the B-cells and one of the T-cells epitopes were expected to be a probable antigen according to VaxiJen v2.0.

Table 3. Disulfide bonds predicted

	Predicted bonds
Disulfide bond scores	
12–115	VHLVVCTQLSA–DDSKMCHVKEG
68–261	KLDLSCVHSDN–ACMERCRCDGR
93–141	VSLEQCAAQCK–SCDRSCFEQHV
97–214	QCAAQCKAVDG–GPKQFCDEGGK
103–380	KAVDGCTHFTY–GPREFCDFGGS
137–301	TASRSCDRSCF–TGPKSCDSSCF
166–267	SQSADCQAACA–RCDGRCTHFTF
170–354	DCQAACAADPS–ASEAKCVVKRK
176–257	AADPSCEIFTY–DSVGACMERCR
188–565	EHDQKCTFKGR–GPYTFCDNGEN
232–454	DQISGCIQLSD–QAEKECSHYTY
263–279	MERCRCDGRCT–DNTRMCYLKGD
305–419	SCDSSCFSNGV–HDEVECVHTGN
332–488	SHPIYCQVICA–TGSRTCDTSCL
336–527	YCQVICAANPL–DAEDACLVFTW
342–466	AANPLCTVFQW–VKSGLCYPKRG
444–492	SSLSECRARCQ–TCDTSCLRRGV
448–539	ECRARCQAEKE–SATSRCYLIGS
517–521	TLPTDCQVACD–DCQVACDAEDA
Predicted connectivity	1–6, 2–16, 3–8, 4–13, 5–26, 7–20, 9–18, 10–25, 11–15, 12–38, 14–30, 17–19, 21–27, 22–32, 23–36, 24–31, 28–33, 29–37, 34–35

Discussion

Despite significant improvements in immune response research that occur after *T. gondii* infection, there is currently just one commercially produced vaccine "Toxovax" which is used to minimize the abortion in sheep for veterinary use but cannot prevent the development of tissue cysts [54]. Nevertheless, it has several disadvantages and cannot be used for humans because these vaccines can regain virulence and even cause iatrogenic infection [55]. Therefore, improving a safe and efficient vaccine can be of global importance to avoid both chronic in immunocompromised patients and fetal infections in pregnant women [56]. Also, the epitope-based vaccines have many advantages in comparison to common vaccines [24]. An increasing number of studies on MICs is shown that they are powerful antigen targets and candidates for vaccines that produce strong immune responses toward toxoplasmosis [6]. Throughout this research, we studied the various elements of MIC4 protein through bioinformatics methods to predict the epitopes to develop a suitable vaccine with the ability to evoke specific B- and T-cell immune responses against *T. gondii* infection. We found different physicochemical characteristics of the MIC4 protein from the results of ProtParam server. The amino acid sequence of MIC4 protein consists of 580 residues with MW of 63,002 KDa, which represents a good antigenic nature 'antigens with MW of <5–10 KDa are regarded as weak immunogens' [57]. We also determined the MIC4 sequence aliphatic index which was 55.64. In short, the high aliphatic index indicates that the target protein is more stable throughout a wide

Table 4. B-cell epitopes predicted from Bcepred server

Prediction parameter	Epitope sequence
Flexibility	VSANVTSSEP; SCVHSDNKGSR; CHVKEGK; DLTGGKTASRSCD; CTFKGRG; LGVTSGP; QLYSSPGDRTGPKSCD; GFYKHRK; DFGGSIRDRE; GNIGSKA; LCYPKRGK; GDMTGSRT; FTWDSAT
Hydrophilicity	TPAGDDVSANVTSSEPAK; VHSDNKGSRAPT; TYNDDSKMCHVKEGKPD; TGGKTASRSCDRSC; SYEGAPD; VTSQSADCQA; AADPSCE; TYNEHDQKCT; TQEEMEDQ; EEPMEADS; ERCRCDGRCT; YSSPGDRTGPKSCDSSC; VDDPATDVE; SIRDREEADAVGSDDGLNAE; RCQAEKECSH; PKRGKPQ; GDMTGSRTCDTSC; DYSQGPE; ACDAEDAC; DSATSRC; SAHRRNDVDG; DNGENLQ; EAKDTE
Accessibility	TSSEPAKLD; HSDNKGSRAPT; THFTYNDDSKM; HVKEGKPDLYDL; GKTASRSCDRS; EQHVSYE; FTYNEHDQKCTFK; SAFKERG; TSGPKQFCDE; GKLTQEEMEDQI; TADLEEPMEADS; TFNDNTRMCYLKGDKMQLYSSPGDRTGPKS; KCVVKRKGFYKHRKTGVT; GSIRDREEADA; NSPDFHDE; RARCQAEKECSHYTYNVKS; CYPKRGKPQFYKYL; RRGVDYSQGPE; GKPWYST; FSAHRRNDVDG; DNGENLQ; LEAKDTE
Turns	SCVHSDNKG; FTYNDDSK; HFTFNDNTRM; SCDSSCF; ANSPDFHDE
Exposed surface	TYNDDSK; VKEGKPDL; YNEHDQK; QEEMEDQ; KCVVKRKGFYKHRKTG; SIRDREEAD; RCQAEKE; CYPKRGKPQFYKY; EAKDTE
Polarity	AKAHGGHRLEPHVP; DDSKMCHVKEGKPDL; RSCFEQHVSYE; TYNEHDQKCTFKGR; FSAFKERG; GKLTQEEMEDQI; DLEEPMEAD; ACMERCRCDGR; AKCVVKRKGFYKHRKTGVT; GSIRDREEADAV; SPDFHDEVECVHT; TIGEVKHASSLSECRARCQAEKECSHYT; CYPKRGKPQF; GFSAHRRNDVDG; LEAKDTE
Antigenic propensity	SLPVHLVVCTQLS; HRLEPHVPGFLQ; KLDLSCVHSD; EPVPDVSLEQC; VDGCTHF; SCDRSCFEQHVSY; GVLGVTSG; QISGCIQLSDV; CMERCRCDGRC; KSCDSSCFS; VSYVDDP; VETVFEISHPIYCQVIC; NPLCTVFQWY; KCVVKRKG; GVTGVTVG; HDEVECVHTGNI; ECSHYTYNVKSGLCYPK; PQFYKYL; TCDTSCLRRGV; LPTDCQV; CLVFTWD; TSRCYLIGSG; DVDGVVSGPYT

 https://doi.org/10.7774/cevr.2021.10.1.59

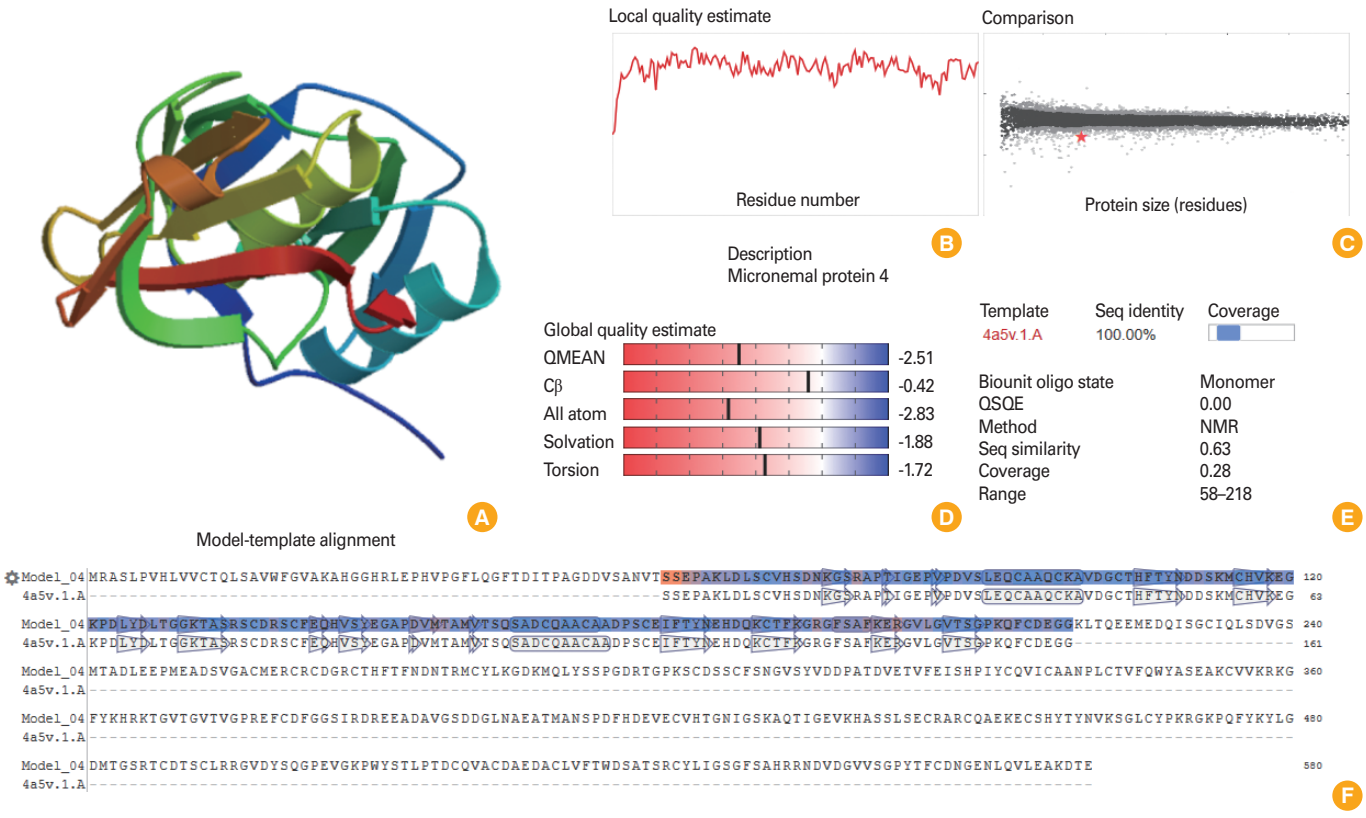


Fig. 6. Analysis of three-dimensional (3D) structure built with SWISS-MODEL for microneme protein 4 (MIC4). (A) 3D structure prediction of MIC4 protein. (B) Local quality estimate. (C) Comparison with non-redundant set of protein data bank structures. (D) Global quality estimate; (E) Sequence coverage and identity. (F) Model-template alignment.

range of temperature. The negative value of GRAVY suggests hydrophilicity of the protein that could bind more with the surrounding water molecules. The average hydrophilic coefficient GRAVY was -0.451, known as a hydrophilic protein. Besides, MIC4's instability indexes were numbered at 42.72. The instability index measures the protein's stability inside the test tube. The MIC4 sequence was classified as unstable, as the value greater than 40 is expected as unstable. The biochemical parameters listed above can help us improve strategies for protein extraction and the subsequent isolation and purification of them in future studies. In the current study, no transmembrane domain was predicted for the MIC4 sequence and could be completely contacted by antigen-presenting cells to induce T- and B-cell priming and powerful immune responses. Post-translational modifications (PTMs) are well known to play an important role in cellular control mechanisms [58]. Awareness of protein phosphorylation sites is also a great tool for evaluating the signaling networks and functional associations between protein signaling [59]. For this purpose, we employed two online servers to predict the MIC4 protein phosphorylation and acylation sites. The findings showed that MIC4 protein contains 97 potential PTM sites (59 phosphorylation and 38 acylation sites) in the sequence indicating that protein functions and activity may be influenced by these sites. The secondary structures of the MIC4 protein were predicted via the three online servers including PSIPRED, GOR IV, and SOPMA. Beta-turn and alpha-helix in the inner parts of the protein with high hydrogen bond ability can sustain a structure of protein and thus cause a strong interaction with antibodies [22]. Precise estimation of the disulfide bond can minimize the conformation space to enhance protein modeling and protein folding in 3D structures [60]. We predicted 38 disulfide bonds in the MIC4 sequence using the DiANNA server. It is well recognized that the principal biological function of proteins relies on their spatial configuration. As apparent, tertiary structure prediction is an important aim of evaluating a protein function. Therefore, understanding the protein structures and recognizing the relationships between structures and functions is very important [24,27]. So, given the importance of tertiary structure in protein biological activity, we used the SWISS-MODEL server to construct the 3D structure of MIC4. To vali-

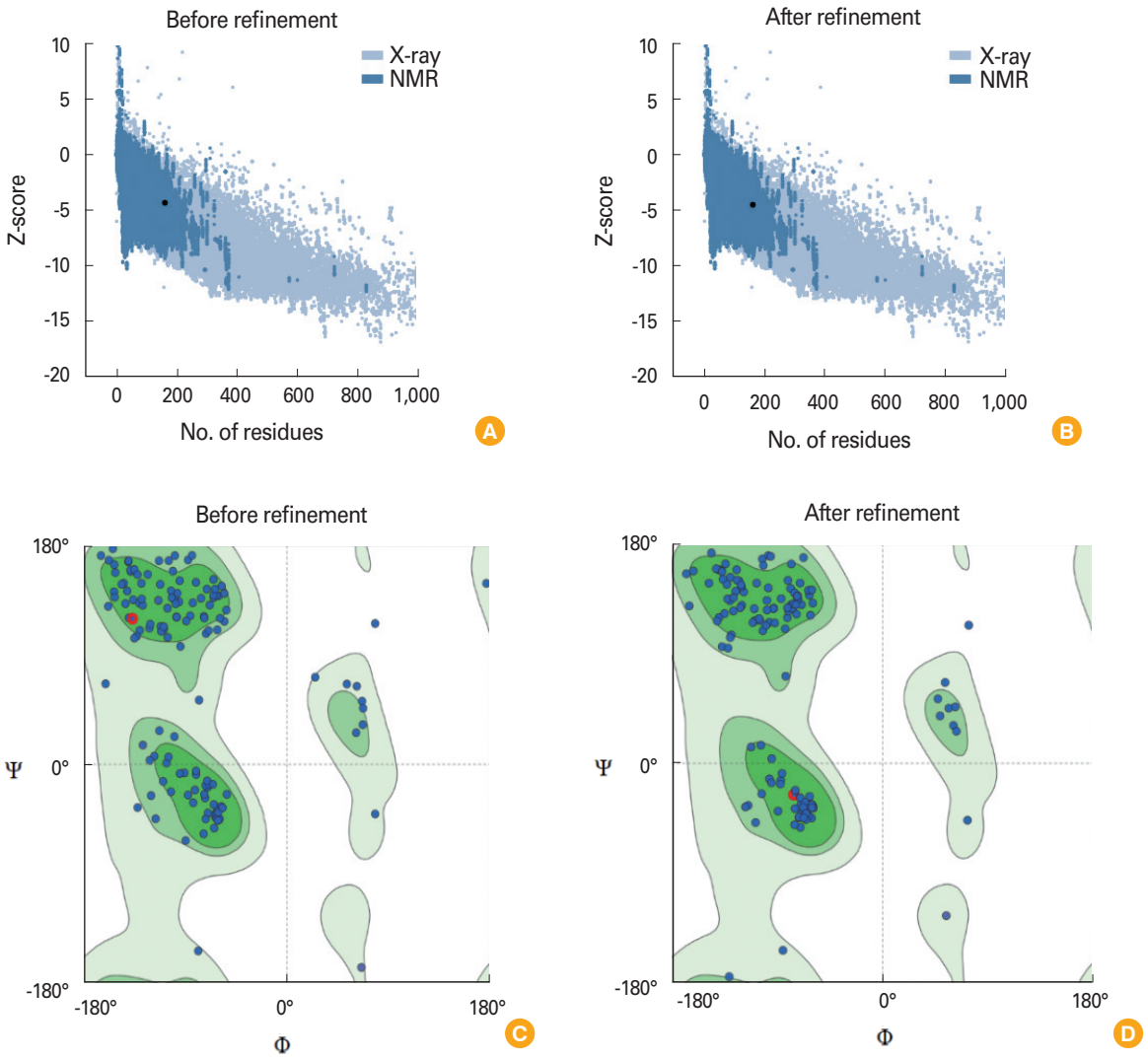


Fig. 7. Validation of three-dimensional (3D) model of microneme protein 4. (A, B) The z-score plot for estimated 3D vaccine structure with ProSA-web server before and after refinement. (C, D) Ramachandran plot analysis of predicted structure. NMR, nuclear magnetic resonance.

date the created 3D model, the Ramachandran plot generation was recruited using the SWISS-MODEL tool. For this purpose, we chosen the best model by SWISS-MODEL, and then refinement was performed using GalaxyRefine. According to the output of Ramachandran plot, protein validation before refinement implied that in the primary model, 87.42% of residues were located in favored, 9.44% in allowed and 3.14% in outlier regions. After refinement of the 3D model, the results were changed as follows: 96.23% of residues in favored regions, 2.51% in allowed regions, and 1.26% in outlier regions of the Ramachandran plot. During *T. gondii* infection, a powerful humoral and cell-mediated immunity is induced [61,62]. Thus, the generation of specific-immunoglobulin G antibodies stopped and limited parasite attachment to the respective host cell receptors. It can also aid immune cells, like macrophages, eliminate *T. gondii* easily and avoid infection reactivation [62]. There is consensus that cell-mediated immune responses, especially associated with $CD8^+$ T cells generating interferon (IFN)-γ, are the crucial mediator of immunity against toxoplasmosis, making it an excellent vaccination development strategy. Besides, $CD4^+$ T cells are important in stimulating immune responses and produce interleukin-2 for the development of $CD8^+$ T cells [63]. Nonetheless, the function of $CD8^+$ T cells and IFN-γ is more important in reducing infection [61,64]. Epitope, component of the antigen, is known by B-cells, T-cells, and host immune system molecules. Just a few residues of amino acids like an epitope (rather than the entire protein) are sufficient to stimulate pro-

 https://doi.org/10.7774/cevr.2021.10.1.59

Table 5. B-cell epitopes predicted from ABCpred tool

Rank	Sequence	Start position	Score
1	KGSRAPTIGEPVPDVSLEQC	74	0.90
2	PYTFCDNGENLQVLEAKDTE	561	0.88
2	FYKHRKTGVTGVTVGPREFC	361	0.88
3	SECRARCQAEKECSHYTYNV	442	0.87
4	SSEPAKLDLSCVHSDNKGSR	58	0.86
4	WYSTLPTDCQVACDAEDACL	509	0.86
4	TIGEVKHASSLSECRARCQA	431	0.86
4	GVTVGPREFCDFGGSIRDRE	371	0.86
4	QWYASEAKCVVKRKGFYKHR	346	0.86
4	LTGGKTASRSCDRSCFEQHV	127	0.86
5	CYLIGSGFSAHRRNDVDGVV	539	0.85
6	FNDNTRMCYLKGDKMQLYSS	272	0.84
6	RSCDRSCFEQHVSYEGAPDV	135	0.84
7	DAVGSDDGLNAEATMANSPD	393	0.83
8	RRNDVDGVVSGPYTFCDNGE	550	0.82
8	VKSGLCYPKRGKPQFYKYLG	461	0.82
8	CRCDGRCTHFTFNDNTRMCY	261	0.82
9	DDPATDVETVFEISHPIYCQ	314	0.81
10	PVPDVSLEQCAAQCKAVDGC	84	0.80
10	PKRGKPQFYKYLGDMTGSRT	468	0.80
10	DFHDEVECVHTGNIGSKAQT	412	0.80
10	PMEADSVGACMERCRCDGRC	248	0.80
10	CDEGGKLTQEEMEDQISGCI	214	0.80
10	CHVKEGKPDLYDLTGGKTAS	115	0.80
11	CSHYTYNVKSGLCYPKRGKP	454	0.79
11	QGFTDITPAGDDVSANVTSS	40	0.79
12	EDACLVFTWDSATSRCYLIG	524	0.78
12	LKGDKMQLYSSPGDRTGPKS	281	0.78
12	DCQAACAADPSCEIFTYNEH	165	0.78
13	LSCVHSDNKGSRAPTIGEPV	66	0.77
13	GNIGSKAQTIGEVKHASSLS	423	0.77
13	MEDQISGCIQLSDVGSMTAD	225	0.77
13	CEIFTYNEHDQKCTFKGRGF	176	0.77
13	CTHFTYNDDSKMCHVKEGKP	103	0.77
14	SAVWFGVAKAHGGHRLEPHV	16	0.76
15	SYEGAPDVMTAMVTSQSADC	147	0.75

Table 6. Linear B-cell epitope of the microneme protein 4 protein by SVMTriP

Rank	Location	Epitope	Score
1	153–172	DVMTAMVTSQSADCQAACAA	1.000
2	475–494	FYKYLGDMTGSRTCDTSCLR	0.723
3	250–269	EADSVGACMERCRCDGRCTH	0.717
4	347–366	WYASEAKCVVKRKGFYKHRK	0.670
5	444–463	CRARCQAEKECSHYTYNVKS	0.583
6	393–412	DAVGSDDGLNAEATMANSPD	0.562
7	174–193	PSCEIFTYNEHDQKCTFKGR	0.528
8	422–441	TGNIGSKAQTIGEVKHASSL	0.518

Table 7. The results of linear B-cell epitopes predicted by the Bepipred 1.0 server

Position	Epitope	Score
287	QLYSSPGDRTGPKSCDSS	1.304
71	SDNKGSRAPTIGEPVPDVSL	1.224
44	DITPAGDDVSANVTSSEPAK	1.182
498	DYSQGPEVGKPWYSTLPT	1.142
311	SYVDDPATDVE	1.105
385	SIRDREEADAVGSDDGLNAEATMANSPDFHD	1.043
147	SYEGAPDV	0.873
119	EGKPDLYDLTGGKTASRSC	0.845
205	GVTSGPKQFCDEGGKLTQEEMEDQ	0.842
467	YPKRGKPQ	0.824
240	SMTADLEEPMEADSV	0.815
547	SAHRRNDVDGVVS	0.687
366	KTGVTGVTVG	0.645
424	NIGSKAQTIGEVK	0.628
161	SQSADCQ	0.623
25	AHGGHRLEPH	0.584
183	EHDQKCT	0.523

tective responses; thus, predicting and identifying this significant segment of amino acid residues can be a key element in understanding the pathogenesis and immune mechanisms of a pathogen and above all in the production of epitope-based vaccines and immunodiagnostic test [65]. Methods of epitope analysis are mostly focused on several (not one) protein properties since researchers believe that only by analyzing one character we cannot access sufficient and reliable epitope knowledge [66]. Secondary structure, beta-turns, surface accessibility, and hydrophilicity are important aspects of amino acids that can provide substantial and useful epitope data for biological research, like DNA vaccine [67]. Thus, a peptide with the above markers can interact effectively with an antibody and generally function as an epitope. So, in the present study, we used several various software services to improve the accuracy of the epitope predictions. The analyses of linear B-cell epitopes demonstrated that the MIC4 protein has good epitopes and satisfactory indexes with Bcepred, ABCpred, ProtScale, SVMTriP, and Bepipred online servers. Predictive accuracy of Bcepred for models varies from 52.92% to 57.53% based on various parameters and this database helps users to predict B-cell epitopes using some of the physicochemical properties (e.g., accessibility, hydrophilicity, polarity, flexibility/mobility, exposed surface,

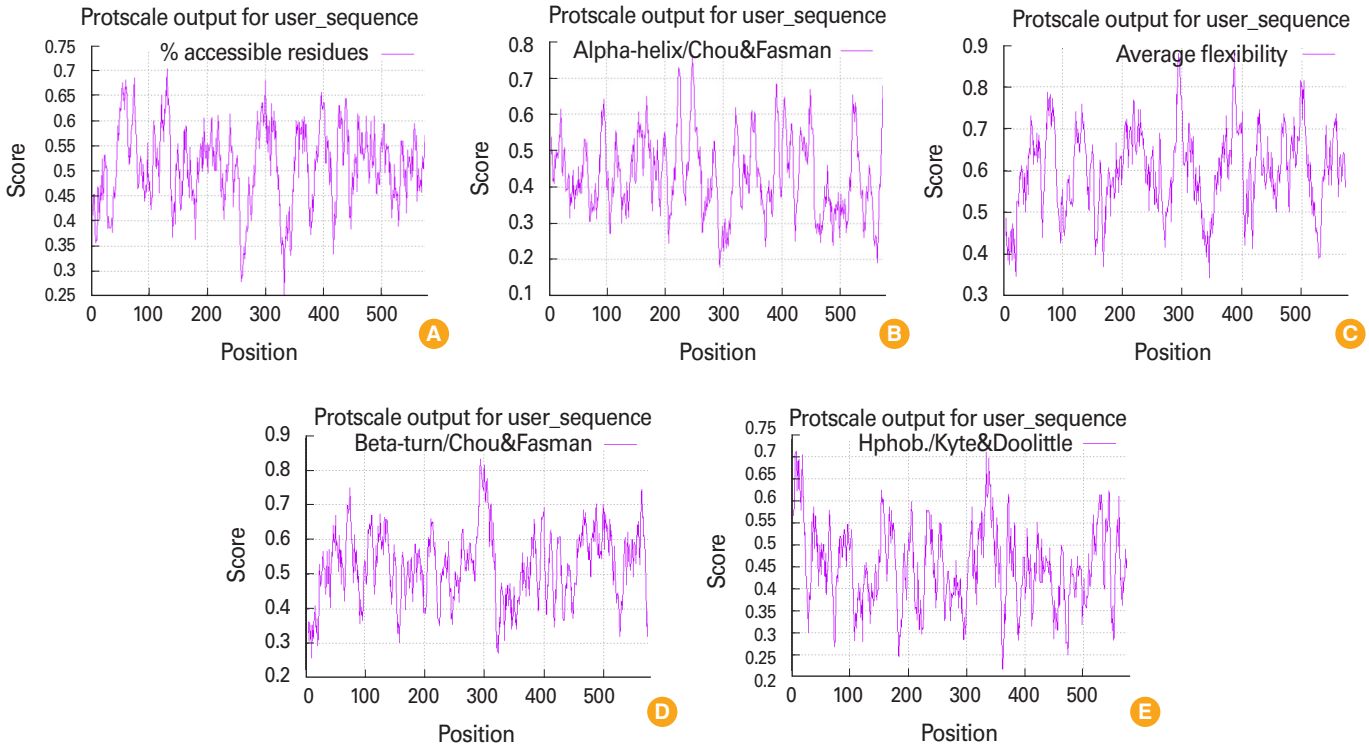
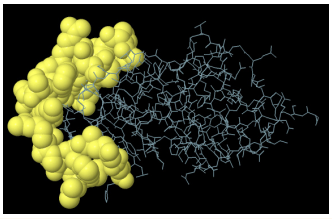
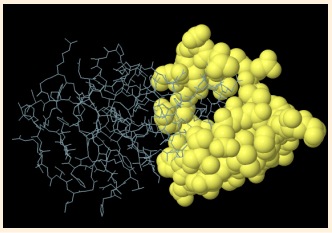
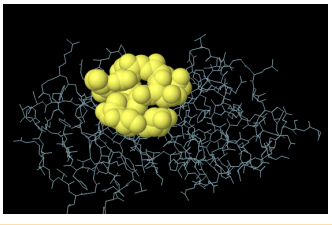


Fig. 8. Linear B-cell epitopes of microneme protein 4 sequence. (A) Percent of accessible residues; (B) alpha-helix; (C) average flexibility; (D) beta-turn; and (E) hydrophobicity.

Table 8. Conformational B-cell epitopes of microneme protein 4 protein predicted by ElliPro server

Residues	No. of residues	Score	Three-dimensional structure
A:Q144, A:H145, A:G150, A:A151, A:P152, A:D153, A:V154, A:D173, A:P174, A:S175, A:C176, A:E177, A:E183, A:G192, A:R193, A:G194, A:F195, A:S196, A:A197, A:F198, A:K199, A:E200, A:R201, A:G202, A:V203, A:L204, A:G205	27	0.768	
A:S58, A:S59, A:E60, A:P61, A:A62, A:K63, A:L64, A:D65, A:L66, A:H70, A:S71, A:D72, A:N73, A:R77, A:A78, A:P79, A:T80, A:I81, A:G82, A:E83, A:P84, A:V85, A:P86, A:D87, A:V88, A:S89, A:E91, A:Q92, A:A99, A:V100, A:D101, A:G102, A:C103, A:D110, A:D111, A:S112, A:K113, A:M114, A:E119, A:G120, A:D123, A:Y125, A:D126, A:L127, A:T128, A:G129	46	0.665	
A:D138, A:S140, A:C141, A:K211, A:Q212, A:C214, A:D215, A:E216, A:G217, A:G218	10	0.534	

and turns) [37,68]. Additionally, ABCpred online server determines B-cell epitopes in an antigen sequence based on an ANN. It is the first server developed using fixed-length patterns based on recurring neural network (machine-based

 https://doi.org/10.7774/cevr.2021.10.1.59

Table 9. Details of selected MHC-I T-cell epitope of *Toxoplasma gondii* microneme protein 4 protein sequence using IEDB server

Allele	Start–stop	Peptide sequence	Percentile rank[a]
H-2-Kb	473–482	PQFYKYLGDM	2.50
	146–155	VSYEGAPDVM	4.70
	195–204	FSAFKERGVL	5.30
H-2-Db	304–313	SCFSNGVSYV	0.35
	146–155	VSYEGAPDVM	1.45
	335–344	ICAANPLCTV	2.25
H-2-Kd	347–356	WYASEAKCVV	0.65
	311–320	SYVDDPATDV	1.05
	124–133	LYDLTGGKTA	1.90
H-2-Kk	224–233	EMEDQISGCI	1.41
	416–425	EVECVHTGNI	4.00
	317–326	ATDVETVFEI	6.20
H-2-Dd	373–382	TVGPREFCDF	0.11
	466–475	CYPKRGKPQF	0.42
	189–198	TFKGRGFSAF	1.10
H-2-Ld	35–44	VPGFLQGFTD	4.40
	36–45	PGFLQGFTDI	7.00
	473–482	PQFYKYLGDM	12.00

MHC, major histocompatibility complex.
[a]Percentile rank=IC50 value; low percentile rank=high level binding

Table 11. Details of selected MHC-I T-cell epitope of *Toxoplasma gondii* microneme protein 4 sequence using NetMHCcons server

Allele	Peptide sequence	IC50 (nM)	% rank	Binding level
H-2-Kb	VSYEGAPDVM	605.05	3.0	WB
	ISHPIYCQVI	1,461.36	5.0	WB
	FSAFKERGVL	1,915.27	6.0	WB
H-2-Db	SCFSNGVSYV	1,044.93	0.8	WB
	SMTADLEEPM	1,085.26	0.8	WB
	YTFCDNGENL	1,915.27	1.0	WB
H-2-Kd	CYLKGDKMQL	226.04	0.4	SB
	WYASEAKCVV	314.41	0.4	SB
	HYTYNVKSGL	316.12	0.4	SB
	SYVDDPATDV	449.33	0.5	SB
H-2-Kk	REFCDFGGSI	47.34	0.4	SB
	MEADSVGACM	273.16	1.5	WB
	LEEPMEADSV	386.17	1.5	WB
H-2-Dd	CYPKRGKPQF	7,366.35	1.0	WB
	IGEVKHASSL	7,989.04	1.5	WB
	QGPEVGKPWY	8,387.64	1.5	WB
H-2-Ld	NPLCTVFQWY	6,719.10	2.0	WB
	GPKSCDSSCF	7,527.49	2.0	WB
	YPKRGKPQFY	10,414.04	2.0	WB

MHC, major histocompatibility complex; SB, strong binders; WB, weak binders.

Table 10. Details of selected MHC-II T-cell epitope of *Toxoplasma gondii* microneme protein 4 sequence using IEDB server

Allele	Start–stop	Peptide sequence	Percentile rank[a]
H2-IAd	151–165	APDVMTAMVTSQSAD	0.61
	152–166	PDVMTAMVTSQSADC	0.62
	150–164	GAPDVMTAMVTSQSA	0.71
H2-IAb	342–356	CTVFQWYASEAKCVV	2.05
	343–357	TVFQWYASEAKCVVK	2.10
	341–355	LCTVFQWYASEAKCV	2.25
H2-IEd	355–369	VVKRKGFYKHRKTGV	0.55
	356–370	VKRKGFYKHRKTGVT	0.69
	354–368	CVVKRKGFYKHRKTG	0.86

MHC, major histocompatibility complex.
[a]Percentile rank=IC50 value; low percentile rank=high level binding.

Table 12. Details of selected MHC-II T-cell epitope of *Toxoplasma gondii* microneme protein 4 sequence using NetMHCIIpan server

Allele	Peptide sequence	IC50 (nM)	% rank	Binding level
H2-IAd	GAPDVMTAMVTSQSA	141.99	1.3	SB
	APDVMTAMVTSQSAD	151.69	1.5	SB
	PDVMTAMVTSQSADC	188.39	2.5	WB
H2-IAb	TVFQWYASEAKCVVK	299.92	1.0	SB
	CTVFQWYASEAKCVV	318.79	1.1	SB
	VFQWYASEAKCVVKR	330.76	1.1	SB
H2-IEd	RKGFYKHRKTGVTGV	1,308.18	3.0	WB
	VKRKGFYKHRKTGVT	1,363.38	3.0	WB
	KRKGFYKHRKTGVTG	1,416.20	3.0	WB

MHC, major histocompatibility complex; SB, strong binders; WB, weak binders.

Table 13. Predicted microneme protein 4 CTL epitopes using CTLpred server

Peptide rank	Start position	Sequence	Score (ANN/SVM)
1	357	KRKGFYKHR	0.94/1.3198907
2	150	GAPDVMTAM	0.64/1.2852105
3	151	APDVMTAMV	0.91/0.85639211
4	305	CFSNGVSYV	0.99/0.58450796
5	409	NSPDFHDEV	0.83/0.73390312
6	171	AADPSCEIF	0.76/0.79240508
7	347	WYASEAKCV	0.98/0.55196569
8	440	SLSECRARC	0.53/0.97958145
9	424	NIGSKAQTI	0.93/0.55425176
10	249	MEADSVGAC	0.93/0.53702617

CTL, cytotoxic-T lymphocyte; ANN, Artificial Neural Network; SVM, Support Vector Machine.

technique). This service can predict epitopes utilizing recurrent neural networks with a precision of 65.93% [38]. The SVMTriP server defined the B-cell epitopes based on similar-

Table 14. Predicted T- and B-cell dominant epitopes of the microneme protein 4

Episode	Methods	Location	Sequence	VaxiJen v2.0 value
B-cell epitope	ABCpred	250–258	EADSVGAC	1.5601 (probable antigen)
	Bcepred	159–166	VTSQSADC	1.1042 (probable antigen)
	SVMTriP	444–458	CRARCQAEKECSHYT	0.9002 (probable antigen)
	Bepipred	289–300	YSSPGDRTGPKS	0.6070 (probable antigen)
	CTLpred	74–80	KGSRAPT	0.5586 (probable antigen)
T-cell epitope	IEDB	357–365	KRKGFYKHR	0.7445 (probable antigen)
	NetMHCcons	151–158	APDVMTAM	0.0556 (probable non-antigen)
	NetMHCIIpan	347–355	WYASEAKCV	0.2876 (probable non-antigen)

ity and propensity scores of the tri-peptide [39]. In this study, we established overlapping peptides of the predicted linear epitopes by comparing the results of the above-mentioned databases. There were five dominant epitopes of 250–258 amino acids, 159–166 amino acids, 444–458 amino acids, 289–300 amino acids, and 74–80 amino acids for the MIC4 protein. It is important to mention that the prediction of discontinuous epitopes is crucial for the interaction between antibodies and antigens [62]. Therefore, the ElliPro tool was used to determine B-cell discontinuous epitopes. We found three conformational B-cell epitopes in the 3D model of the MIC4 protein.

MHC molecules present epitopes to T-cells. Binding the peptides to the MHC is a vital stage in the process of presenting T-cell antigen and also a substantial factor in the selection of potential epitopes. For every MHC class, we used two online databases to analyze the IC50 values of peptides that bind to these molecules (MHC-I and MHC-II) for MIC4. According to findings, the T-cell epitopes on MIC4 were observed to have the ability to stick strongly to the molecules of MHC-I and MHC-II. The results of these databases showed that both tools identified approximately some specific peptides and the results were almost similar. It should be mentioned that the higher percentile levels (or IC50 values) represent the lower affinity rate which is a weaker T-cell epitope and vice versa. Moreover, we determined the CTL epitopes using the CTLpred database and selected the top 10 epitopes for the MIC4 protein. CTLpred is a method for CTL epitopes prediction that is critical in the design of a vaccine. This approach is based on elegant machine learning techniques such as ANN and SVM. Since all MHC binders cannot function as T-cell epitopes, for CTL epitopes need an extremely accurate prediction method. The use of ANN and SVM has been studied to solve the problem. Regarding the above two approaches, the CTLpred server uses combined and consensus prediction. Consensus and combined prediction methods are more accurate and sensitive compared to the individual methods (such as ANN and SVM) [46]. Following a comparison of the different results, one dominant T-cell epitope of the MIC4 protein was predicted: 357–365 amino acids, 151–158 amino acids, and 347–355 amino acids. It must be mentioned that allergenic protein identification is becoming very substantial due to the use of modified proteins in diets (genetically modified foods), bio-pharmaceuticals, therapies, etc. (http://crdd.osdd.net/raghava/algpred/). Finally, the outcomes of antigenicity (ANTIGENpro and VaxiJen servers) and allergenicity evaluation (AllerTOP v2.0 server) demonstrated MIC4 protein is immunogenic and non-allergen.

This paper presented a comprehensive explanation of the fundamental aspects of MIC4 protein, including physicochemical characteristics, a transmembrane domain, subcellular location, secondary and tertiary structure, potential B- and T-cell epitopes and other important features of this protein, using various and accurate bioinformatics methods. Diagnostic assessment findings via different online bioinformatics databases revealed that MIC4 protein had multiple fantastic epitopes of B- and T-cells, implying it could become a remarkable vaccine candidate against *T. gondii*. Also, in parallel with this study previously, Saraav et al. [69] in 2019 indicated that MICs could be useful to control immune status through infection as an adjunct to serological analysis. Also, they showed MIC1, MIC3, and MIC6 were able to induce memory responses from mice which infected with *T. gondii* leading to the development of IFN-γ by T cells. In another study by Sardinha-Silva et al. [70] in 2019, they showed which dendritic cells and macrophages are induced by recombinant MIC1 and recombinant MIC4 to release proinflammatory cytokines, and they do so by involving TLR2 and TLR4. This study provided important basic and theoretical data on MIC4 protein for the development of an effective vaccine

 https://doi.org/10.7774/cevr.2021.10.1.59

against acute and chronic toxoplasmosis for further in vivo investigations.

More studies are required on vaccine development *in vivo* using the MIC4 alone or combined with other antigens in the future.

ORCID

Ali Dalir Ghaffari *https://orcid.org/0000-0001-9635-2876*
Abdolhossein Dalim *https://orcid.org/0000-0001-5591-5513*
Fatemeh Ghaffarifar *https://orcid.org/0000-0003-0891-8214*
Majid Pirestani *https://orcid.org/0000-0003-0046-4772*
Hamidreza Majidiani *https://orcid.org/0000-0001-5568-1366*

References

1. Robert-Gangneux F, Darde ML. Epidemiology of and diagnostic strategies for toxoplasmosis. Clin Microbiol Rev 2012;25:264-96.
2. Ghaffari AD, Dalimi A. Molecular identification of Toxoplasma gondii in the native slaughtered cattle of Tehran province, Iran. J Food Qual Hazards Control 2019;6:153-61.
3. Majidiani H, Dalimi A, Ghaffarifar F, Pirestani M, Ghaffari AD. Computational probing of Toxoplasma gondii major surface antigen 1 (SAG1) for enhanced vaccine design against toxoplasmosis. Microb Pathog 2020;147:104386.
4. Buxton D. Protozoan infections (Toxoplasma gondii, Neospora caninum and Sarcocystis spp.) in sheep and goats: recent advances. Vet Res 1998;29:289-310.
5. Lourenco EV, Bernardes ES, Silva NM, Mineo JR, Panunto-Castelo A, Roque-Barreira MC. Immunization with MIC1 and MIC4 induces protective immunity against Toxoplasma gondii. Microbes Infect 2006;8:1244-51.
6. Dodangeh S, Daryani A, Sharif M, et al. A systematic review on efficiency of microneme proteins to induce protective immunity against Toxoplasma gondii. Eur J Clin Microbiol Infect Dis 2019;38:617-29.
7. Kravetz JD, Federman DG. Toxoplasmosis in pregnancy. Am J Med 2005;118:212-6.
8. Ghaffari AD, Dalimi A, Ghaffarifar F, Pirestani M. Structural predication and antigenic analysis of ROP16 protein utilizing immunoinformatics methods in order to identification of a vaccine against Toxoplasma gondii: an in silico approach. Microb Pathog 2020;142:104079.
9. Majidiani H, Soltani S, Ghaffari AD, Sabaghan M, Taghipour A, Foroutan M. In-depth computational analysis of calcium-dependent protein kinase 3 of Toxoplasma gondii provides promising targets for vaccination. Clin Exp Vaccine Res 2020;9:146-58.
10. Morlon-Guyot J, Pastore S, Berry L, Lebrun M, Daher W. Toxoplasma gondii Vps11, a subunit of HOPS and CORVET tethering complexes, is essential for the biogenesis of secretory organelles. Cell Microbiol 2015;17:1157-78.
11. Ezz Eldin HM, Kamel HH, Badawy AF, Shash LS. A comparative study between excretory/secretory and autoclaved vaccines against RH strain of Toxoplasma gondii in murine models. J Parasit Dis 2015;39:526-35.
12. Wu XN, Lin J, Lin X, Chen J, Chen ZL, Lin JY. Multicomponent DNA vaccine-encoding Toxoplasma gondii GRA1 and SAG1 primes: anti-Toxoplasma immune response in mice. Parasitol Res 2012;111:2001-9.
13. Parthasarathy S, Fong MY, Ramaswamy K, Lau YL. Protective immune response in BALB/c mice induced by DNA vaccine of the ROP8 gene of Toxoplasma gondii. Am J Trop Med Hyg 2013;88:883-7.
14. Pinzan CF, Sardinha-Silva A, Almeida F, et al. Vaccination with recombinant microneme proteins confers protection against experimental toxoplasmosis in mice. PLoS One 2015;10:e0143087.
15. Macedo AG Jr, Cunha JP Jr, Cardoso TH, et al. SAG2A protein from Toxoplasma gondii interacts with both innate and adaptive immune compartments of infected hosts. Parasit Vectors 2013;6:163.
16. Dobrowolski JM, Carruthers VB, Sibley LD. Participation of myosin in gliding motility and host cell invasion by Toxoplasma gondii. Mol Microbiol 1997;26:163-73.
17. Meissner M, Schluter D, Soldati D. Role of Toxoplasma gondii myosin A in powering parasite gliding and host cell invasion. Science 2002;298:837-40.
18. Liu Q, Li FC, Zhou CX, Zhu XQ. Research advances in interactions related to Toxoplasma gondii microneme proteins. Exp Parasitol 2017;176:89-98.
19. Cerede O, Dubremetz JF, Soete M, et al. Synergistic role of micronemal proteins in Toxoplasma gondii virulence. J Exp Med 2005;201:453-63.
20. Liu MM, Yuan ZG, Peng GH, et al. Toxoplasma gondii microneme protein 8 (MIC8) is a potential vaccine candidate against toxoplasmosis. Parasitol Res 2010;106:1079-84.
21. Dodangeh S, Fasihi-Ramandi M, Daryani A, Valadan R, Sarvi S. In silico analysis and expression of a novel chimeric antigen as a vaccine candidate against Toxoplasma

gondii. Microb Pathog 2019;132:275-81.
22. Shaddel M, Ebrahimi M, Tabandeh MR. 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 Parasit Dis 2018;42:269-76.
23. Romano P, Giugno R, Pulvirenti A. Tools and collaborative environments for bioinformatics research. Brief Bioinform 2011;12:549-61.
24. Wang Y, Wang G, Cai J, Yin H. Review on the identification and role of Toxoplasma gondii antigenic epitopes. Parasitol Res 2016;115:459-68.
25. Wang Y, Wang G, Zhang D, Yin H, Wang M. Identification of novel B cell epitopes within Toxoplasma gondii GRA1. Exp Parasitol 2013;135:606-10.
26. Gasteiger E, Hoogland C, Gattiker A, et al. Protein identification and analysis tools on the ExPASy server. In: Walker JM, editor. The proteomics protocols handbook. Totowa, NJ: Humana Press; 2005. p.571-607.
27. Zhou J, Wang L, Zhou A, et al. Bioinformatics analysis and expression of a novel protein ROP48 in Toxoplasma gondii. Acta Parasitol 2016;61:319-28.
28. Foroutan M, Ghaffarifar F, Sharifi Z, Dalimi A, Pirestani M. Bioinformatics analysis of ROP8 protein to improve vaccine design against Toxoplasma gondii. Infect Genet Evol 2018;62:193-204.
29. Garnier J, Gibrat JF, Robson B. GOR method for predicting protein secondary structure from amino acid sequence. Methods Enzymol 1996;266:540-53.
30. McGuffin LJ, Bryson K, Jones DT. The PSIPRED protein structure prediction server. Bioinformatics 2000;16:404-5.
31. Deleage G. ALIGNSEC: viewing protein secondary structure predictions within large multiple sequence alignments. Bioinformatics 2017;33:3991-2.
32. Ferre F, Clote P. DiANNA: a web server for disulfide connectivity prediction. Nucleic Acids Res 2005;33:W230-2.
33. Guex N, Peitsch MC, Schwede T. Automated comparative protein structure modeling with SWISS-MODEL and Swiss-PdbViewer: a historical perspective. Electrophoresis 2009;30 Suppl 1:S162-73.
34. Park H, Seok C. Refinement of unreliable local regions in template-based protein models. Proteins 2012;80:1974-86.
35. Bertoni M, Kiefer F, Biasini M, Bordoli L, Schwede T. Modeling protein quaternary structure of homo- and hetero-oligomers beyond binary interactions by homology. Sci Rep 2017;7:10480.
36. Wiederstein M, Sippl MJ. ProSA-web: interactive web service for the recognition of errors in three-dimensional structures of proteins. Nucleic Acids Res 2007;35:W407-10.
37. Saha S, Raghava GP. BcePred: prediction of continuous B-cell epitopes in antigenic sequences using physico-chemical properties. In: Nicosia G, Cutello V, Bentley PJ, Timmis J, editors. Artificial immune systems. Proceedings of the International Conference on Artificial Immune Systems; 2004 Sep 13-16; Catania, Italy. Berlin: Springer; 2004. p.197-204.
38. Saha S, Raghava GP. Prediction of continuous B-cell epitopes in an antigen using recurrent neural network. Proteins 2006;65:40-8.
39. Yao B, Zhang L, Liang S, Zhang C. SVMTriP: a method to predict antigenic epitopes using support vector machine to integrate tri-peptide similarity and propensity. PLoS One 2012;7:e45152.
40. Larsen JE, Lund O, Nielsen M. Improved method for predicting linear B-cell epitopes. Immunome Res 2006;2:2.
41. Ponomarenko J, Bui HH, Li W, et al. ElliPro: a new structure-based tool for the prediction of antibody epitopes. BMC Bioinformatics 2008;9:514.
42. Wang P, Sidney J, Kim Y, et al. Peptide binding predictions for HLA DR, DP and DQ molecules. BMC Bioinformatics 2010;11:568.
43. Karosiene E, Lundegaard C, Lund O, Nielsen M. NetMHCcons: a consensus method for the major histocompatibility complex class I predictions. Immunogenetics 2012;64:177-86.
44. Kotturi MF, Peters B, Buendia-Laysa F Jr, et al. The CD8+ T-cell response to lymphocytic choriomeningitis virus involves the L antigen: uncovering new tricks for an old virus. J Virol 2007;81:4928-40.
45. Moutaftsi M, Peters B, Pasquetto V, et al. A consensus epitope prediction approach identifies the breadth of murine T(CD8+)-cell responses to vaccinia virus. Nat Biotechnol 2006;24:817-9.
46. Bhasin M, Raghava GP. Prediction of CTL epitopes using QM, SVM and ANN techniques. Vaccine 2004;22:3195-204.
47. Wang P, Sidney J, Dow C, Mothe B, Sette A, Peters B. A systematic assessment of MHC class II peptide binding predictions and evaluation of a consensus approach. PLoS Comput Biol 2008;4:e1000048.
48. Jensen KK, Andreatta M, Marcatili P, et al. Improved methods for predicting peptide binding affinity to MHC

class II molecules. Immunology 2018;154:394-406.
49. Magnan CN, Zeller M, Kayala MA, et al. High-throughput prediction of protein antigenicity using protein microarray data. Bioinformatics 2010;26:2936-43.
50. Doytchinova IA, Flower DR. VaxiJen: a server for prediction of protective antigens, tumour antigens and subunit vaccines. BMC Bioinformatics 2007;8:4.
51. Dimitrov I, Flower DR, Doytchinova I. AllerTOP: a server for in silico prediction of allergens. BMC Bioinformatics 2013;14(Suppl 6):S4.
52. Magnan CN, Randall A, Baldi P. SOLpro: accurate sequence-based prediction of protein solubility. Bioinformatics 2009;25:2200-7.
53. Brecht S, Carruthers VB, Ferguson DJ, et al. The toxoplasma micronemal protein MIC4 is an adhesin composed of six conserved apple domains. J Biol Chem 2001;276:4119-27.
54. Buxton D, Thomson K, Maley S, Wright S, Bos HJ. Vaccination of sheep with a live incomplete strain (S48) of Toxoplasma gondii and their immunity to challenge when pregnant. Vet Rec 1991;129:89-93.
55. Buxton D, Innes EA. A commercial vaccine for ovine toxoplasmosis. Parasitology 1995;110 Suppl:S11-6.
56. Innes EA. Vaccination against Toxoplasma gondii: an increasing priority for collaborative research? Expert Rev Vaccines 2010;9:1117-9.
57. Berzofsky JA. Immunogenicity and antigen structure. In: Paul WE, editor. Fundamental immunology. 3rd ed. New York, NY: Raven Press; 1993. p.235-82.
58. Lee TY, Hsu JB, Chang WC, Wang TY, Hsu PC, Huang HD. A comprehensive resource for integrating and displaying protein post-translational modifications. BMC Res Notes 2009;2:111.
59. Dephoure N, Gould KL, Gygi SP, Kellogg DR. Mapping and analysis of phosphorylation sites: a quick guide for cell biologists. Mol Biol Cell 2013;24:535-42.
60. Yang J, He BJ, Jang R, Zhang Y, Shen HB. Accurate disulfide-bonding network predictions improve ab initio structure prediction of cysteine-rich proteins. Bioinformatics 2015;31:3773-81.
61. El-Kady IM. T-cell immunity in human chronic toxoplasmosis. J Egypt Soc Parasitol 2011;41:17-28.
62. Sayles PC, Gibson GW, Johnson LL. B cells are essential for vaccination-induced resistance to virulent Toxoplasma gondii. Infect Immun 2000;68:1026-33.
63. Nosrati MC, Ghasemi E, Shams M, et al. Toxoplasma gondii ROP38 protein: bioinformatics analysis for vaccine design improvement against toxoplasmosis. Microb Pathog 2020;149:104488.
64. Suzuki Y, Orellana MA, Schreiber RD, Remington JS. Interferon-gamma: the major mediator of resistance against Toxoplasma gondii. Science 1988;240:516-8.
65. Desai DV, Kulkarni-Kale U. T-cell epitope prediction methods: an overview. Methods Mol Biol 2014;1184:333-64.
66. Kringelum JV, Lundegaard C, Lund O, Nielsen M. Reliable B cell epitope predictions: impacts of method development and improved benchmarking. PLoS Comput Biol 2012;8:e1002829.
67. Zhao L, Li J. Mining for the antibody-antigen interacting associations that predict the B cell epitopes. BMC Struct Biol 2010;10(Suppl 1):S6.
68. Kazi A, Chuah C, Majeed AB, Leow CH, Lim BH, Leow CY. Current progress of immunoinformatics approach harnessed for cellular- and antibody-dependent vaccine design. Pathog Glob Health 2018;112:123-31.
69. Saraav I, Wang Q, Brown KM, Sibley LD. Secretory microneme proteins induce T-cell recall responses in mice chronically infected with Toxoplasma gondii. mSphere 2019;4:e00711-18.
70. Sardinha-Silva A, Mendonca-Natividade FC, Pinzan CF, et al. The lectin-specific activity of Toxoplasma gondii microneme proteins 1 and 4 binds Toll-like receptor 2 and 4 N-glycans to regulate innate immune priming. PLoS Pathog 2019;15:e1007871.