Designing and Modeling of Multi-epitope Proteins for Diagnosis of *Toxocara canis* Infection

Maryam Ebrahimi[1] · Seyyed Javad Seyyedtabaei[2] · Mohammad Mehdi Ranjbar[3] · Farid Tahvildar-biderouni[2] · Amirreza Javadi Mamaghani[2]

Abstract

Serological investigation is the main method to achieve satisfactory results in *Toxocara canis* diagnosis. The accuracy of the native antigen used in the current diagnostic kits has proven to be insufficient as well as difficult to standardize. Therefore significant efforts have been made to find alternative reagents as capture antigens. Multi-epitope peptides are potential diagnostic markers to improve the accuracy of diagnostic kits. The main aim of this study is the prediction and design of a novel synthetic protein consisting of multiple immunodominant B cell epitopes by Use of three proteins TES-120, TES-30, and TES-26 of *T. canis*. Primary, secondary and tertiary structures of the proteins were analyzed by using several various online software (ExPasy, IEDB, ABCpred, SVMTriP). Then, B cell construct was assessed by machine learning and Physico-chemical approaches and finally, homology modeling of 3D structure of protein was evaluated. The results of in silico analyses indicated that regions with high immunogenicity for TES-120 protein are located at between residues 97–167, for TES-30 protein are in the residues 52–102, 172–207 and for TES-26 are in the residues 33–83, 130–180. These regions could have good potential features for designing the Multi-epitopes. Finally, selected epitopes were linked to each other by linkers. The average length of the constructs was 342 bp. Also, the high proportion of random coils and extended strands in construct suggest that the protein form antigenic epitopes. Expasy ProtParam classified the constructs with moderate stability and 56.2% residues of constructs were located in favored regions of the Ramachandran plot. In conclusion, immunoinformatics analysis indicated that this multi-epitope peptide can laid a theoretical basis to develop an appropriate diagnostic kit for human toxocariasis.

Keywords Multi-epitope · Synthetic gene · *Toxocara canis* · Bioinformatics softwar

✉ Seyyed Javad Seyyedtabaei
  Seyyedtabaei@gmail.com; seyyedtabaei@sbmu.ac.ir

1 Department of Medical Parasitology and Mycology, School of Medicine, Student Research Committee, Shahid Beheshti University of Medical Sciences, Tehran, Iran

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

3 Department of Immunology, Razi Vaccine and Sera Research Instute, Karaj, IR, Iran

Introduction

Toxocariasis is one of the most distributed zoonotic infectious diseases worldwide with a wide range of clinical manifestations (Baneth et al. 2016; Aghaei et al. 2018; Mohammadzadeh et al. 2018). The disease is mostly caused by the larvae of *T. canis* and *T. cati*, which are intestinal nematodes found in canids and felids as definitive hosts (Despommier 2005; Fakhri et al. 2018; Aghamolaie et al. 2018). Children are at higher risk for infection with *Toxocara* spp., due to more direct contact with the soil contaminated with eggs (Baneth et al. 2016; Siyadatpanah et al. 2013). Based on the migration of larvae in different organs (such as the liver, brain, and eyes), clinical manifestations vary in different patients (Despommier 2005; Smith et al. 2009).

Human patients are diagnosed based on serological tests alongside of clinical symptoms and signs, and laboratory

findings such as eosinophilia, leukocytosis, and hyperglobulinemia (Rubinsky-Elefant et al. 2010). Direct detection of larvae in tissue by microscopy from biopsies is rarely successful (Rubinsky-Elefant et al. 2010). Serological diagnosis of toxocariasis is usually performed by commercial immunoglobulin G-enzyme-linked immunosorbent assay (IgG-ELISA) kits using *T. canis* excretory–secretory (TES) antigens of second-stage larvae, although this method could sometimes be laborious and time-consuming with limited production capacity (Overgaauw 1997). Moreover, in the other developing countries, due to the high rate of soil-borne worms, serum samples from patients with ascariasis and strongyloidiasis cross-react with the native TES antigen in immunoassay (Noordin et al. 2005; Fong et al. 2003). Therefore, the development of highly specific and sensitive methods for timely and accurate detection of anti-Toxocara antibodies is critical to improve the diagnosis of human toxocariasis (Rubinsky-Elefant et al. 2010). For this purpose, the employment of recombinant antigens was widely investigated in recent years.

Advantages of recombinant antigens include unlimited production, highly sensitive and specific results, and minimized possibility of cross-reaction with antigens of the other parasites (Norhaida et al. 2008; Yamasaki et al. 2000). Previous studies indicated several recombinant antigens suitable for the serodiagnosis of toxocariasis, namely TES-26, TES-30, and TES-120 (Norhaida et al. 2008; Fong and Lau 2004; Mohamad et al. 2009). The surface of parasite molecules contains many overlapping antibody-binding sites called epitopes (Frank 2002). All epitopes have fuzzy regions that can be exactly recognized by the antibodies paratope (i e, antigen binding site). Antigenic cross-reactivity is a common occurrence, since at any time antibodies can recognize a large number of epitopes (Frank 2002). This place can help to reduce the specificity of antibodies in binding to the antigen. The antigenicity is the capacity of an epitope to react with an antibody in terms of the recognition, immunogenicity, and the ability to induce the immune system to produce antibodies in a competent vertebrate host (Frank 2002; Dai et al. 2013; Sela-Culang et al. 2013). If it is not done correctly, it could cause problems in the design of vaccines and diagnostic kits. Therefore, the prediction of epitopes can be beneficial in the development of vaccines and diagnostic tests (Karpenko et al. 2014; Dipti et al. 2006).

Recently, bioinformatics software is employed to predict antigenic epitopes by researchers in biotechnology and immunology research centers (Karpenko et al. 2014; Kulkarni et al. 2012). The employment of the software could be helpful for a good prediction and the exact identification of epitopes in terms of high availability and antigenicity (Pruess and Apweiler 2003). It also designs and constructs the recombinant multi-epitope antigen; using such epitopes may provide a new tool and an alternative to reach

cost-effective and more accurate diagnostic kits (Dipti et al. 2006). Moreover, experimental studies revealed that the use of peptide-based antigens can meet the need for serological test standardization and increase the sensitivity and specificity of ELISA (Dai et al. 2013; Lv et al. 2016). Multi-epitope approach as a potential antigen-capture is assessed in different studies for a range of pathogens, but such studies are not conducted to determine antigenic epitopes of *T. canis* (Hajissa et al. 2015). According to the above statements, the present study aimed at predicting and designing a novel synthetic protein consisting of multiple immunodominant B-cell epitopes of several *T. canis* antigens, TES-26, TES-30, TES-120, and analyzing its immunogenicity and preliminarily evaluation to improve the accuracy of serodiagnostic kits for human toxocariasis.

Methods

A flowchart for the creation of new synthetic construct is presented in Fig. 1. The different main steps performed in the methodology are shown.

Protein Sequence

Nucleotide sequences were retrieved from the National Centre for Biotechnology Information (NCBI) Nucleotide Database. TES-120 (GenBank: U39815), TES-30 (GenBank: AB009305) and TES-26 (GenBank: U29761), sequence of protein was recovered from UniProt (Universal Protein resource) that is an easily accessible database which comprises data of proteins. The protein sequence was recovered through their accession number and it was in FASTA format.

Membrane Protein Topology and Signal Peptide Prediction

The transmembrane structure of proteins were predicted by TMHMM Server v. 2.0 (https://www.cbs.dtu.dk/services/TMHMM/). The sequence of proteins were input, and three regions, including outside, transmembrne and inside regions, were studied (Krogh et al. 2001).

The signal peptide prediction of proteins were predicted by SignalP 4.1 Server (https://www.cbs.dtu.dk/services/SignalP/) (Geourjon and Deleage 1995).

Prediction of Linear and Conformational B Cell Epitopes

For prediction of linear B-cell epitopes, the IEDB, ABCpred, Kirloskar, Bcepred, LBTope, SVMTriP, Bepipred and Emini surface accessibility web servers were employed. These tools evaluates the epitopes, on the basis of Chou
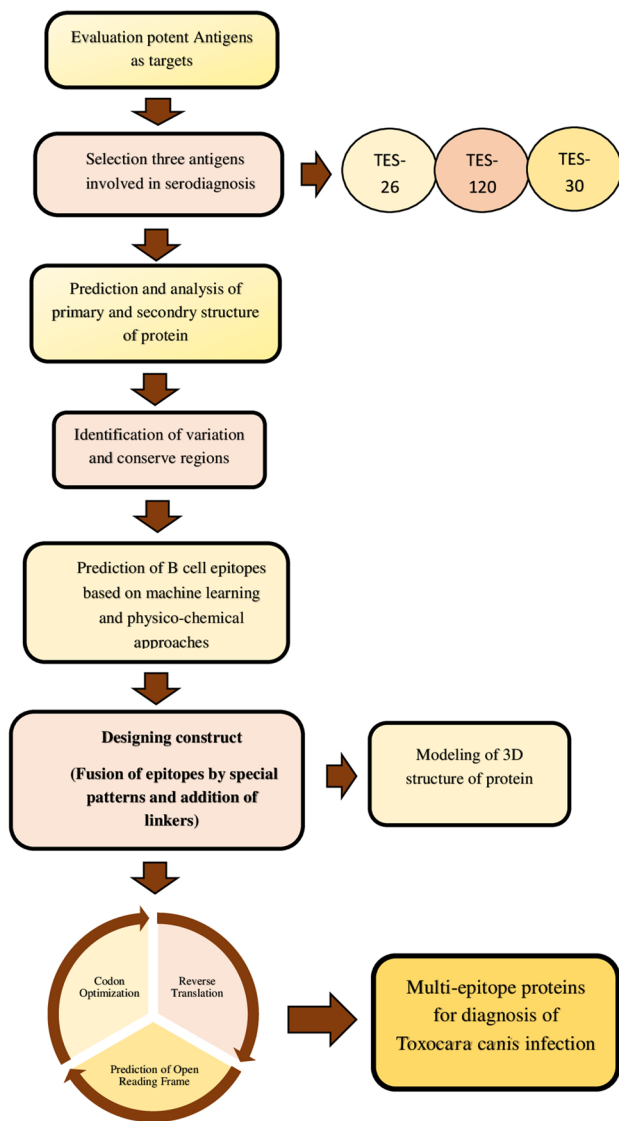
Fig. 1 Flowchart summarizing the steps of a multi-epitope *Toxocara canis* design for diagnosis of human infections

and Fasman beta-turn, Emini surface accessibility prediction, Karplus & Schulz Flexibility Prediction, Kolaskar & Tongaonkar Antigenicity, Parker Hydrophilicity Prediction (Kolaskar and Tongaonkar 1990; Singh et al. 2013; Yang 2004; Saha et al. 2005). Also, for prediction of conformational B-cell epitopes, the CBTOPE server was used. This server can predict conformational B cell epitope by antigen primary sequence in the lack of any homology with the known structures (Zhang et al. 2011).

Designing and Modeling Multi-epitope Antigenic (Construct) and Improving Immunogenicity

Primary, secondary structures of the proteins were analyzed by using several various online software. Finally the peptides

which included most high score B-cell epitopes, possessing higher antigenicity and corresponding to most prediction results were chosen. To achieve the best immunization in construct, first, incorporated two repeats of TES-120 epitope in the beginning and the end of constructs and put linkers between epitopes. Finally, more distances between epitopes were created by adding two or more amino acids in the constructs. These tactics can help to achieve the better immunization in construct.

Reverse Translation and Codon Optimization & Prediction of Open Reading Frame (ORF)

The B cell protein construct was backtranslated into nucleotide sequences using backtranseq program of mEMBOSS 6.0.1 (https://www.ebi.ac.uk/Tools/st/emboss_ backtranseq/). The degeneracy of the genetic code cause backtranslation potentially obscure since most amino acids are encoded by multiple codons. Backtranseq was restricted to codon uses within the Bos tau. Codon optimization is a method for higher gene expression of vectors to arrive optimum expression of a foreign gene (Sandhu et al. 2008). Large numbers of C–G sequences in the messenger RNA (mRNA) can prevent protein translation from increased constitution of secondary structures, So, increase mRNA stability that considerably improves immune responses (Kalwy et al. 2006). Constructs were optimized by the codon adaptation tool server (http: //www. jcat.de/Start.jsp). At the –NH2 and –COOH terminus of construct, an initiation codon ATG and a termination codon TAA were added (Ramakrishna et al. 2004) For confidence of ORF, we used the gorf tool available at the NCBI server (https://www.ncbi.nlm.nih.gov/gorf/). This server distinguishes all ORF by the standard or alternative genetic codes.

Primary, Secondary & Tertiary Structure Prediction of Construct

Protein sequence statistics for constructs including length, molecular weight, isoelectric point (IEP), total number of positive and negative residues, instability index, grand average hydropathicity (GRAVY), aliphatic index, and amino acids distribution were calculate by using the ExPASy ProtParam server (https://expasy.org/cgi-bin/protpraram) (Gasteiger et al. 2003). The secondary structure of Construct was assessed by the Self-Optimized Prediction method With Alignment (SOPMA) Server (https://npsa-prabi.ibcp.fr/cgi-bin/secpred_sopma.pl). This method predicts 69.5% of amino acids for a three-state description of the secondary structure (a-helix. (3-sheet and coil) in a whole database containing 126 chains of non-homologous (less than 25% identity) proteins (Geourjon and Deleage 1995).

The tertiary structure of construct was predicted by the online prediction server I-TASSER server (https://zhang lab.ccmb.med.umich.edu/I-TASSER/). I-TASSER (Iterative Threading ASSEmbly Refinement) is one of the top tools for automatic protein structure prediction. This server is in active development with the aim to provide the most accurate structural and function predictions using state-of-the-art algorithms (Yang et al. 2015). For confirmation of the predicted structures, the ProSA-web at https://prosa .services.came.sbg.ac.at/prosa.php was applied to recognize the potential errors in modeled structure before and after minimization process and Ramachandran plot was studied through PROCHECK analyses in the PSVS server v. 1.5 (http://psvs.nesg.org/) too (Bhattacharya et al. 2007; Wiederstein and Sippl 2007). Lastly, the 3D model result generated in a PDB format was analysed and visualized using PyMOL version 1.3 available at https://pymol.org/2/ (DeLano 2002).

Results

Analysis of Transmembrane Topology & Signal Peptide Properties Proteins

The results showed that outside regions of TES-120, TES-30 and TES-26 were located at positions 1–176, 1–225 and 1–262 respectively, thus all the proteins in this study are in the outer part of the membrane. The results are displayed in Supplementary Fig. 1. The signal peptide prediction results for proteins are displayed in Supplementary Fig. 2, that due to their interference role, these sequences should be remove from the epitope.

Designing and Modeling B Cell Construct

A combination of the results predicted for every three proteins by the different parameters and bioinformatics software for predicting antigenic epitopes TES-120, TES-30, and TES-26 are presented in Table 1. Five epitopes were found to have a consensus and used for in silico concatenation, two consensus epitopes were selected for TES-30 and TES-26 and for TES-120 an epitope was selected which was repeated at the beginning and the end of the multi-epitope

sequence (Table 2).The individual epitope lengths varied) from 35 to 65 amino acids(and each epitope with a flexible linker sequence (Gly-Ser-Gly-Ser-Gly) was connected to another epitope (Fig. 2).

Reverse Translation and Codon Optimization

Protein sequences of make were reverse translated into nucleotide sequences using the backtranseq program. The numbers of C–G sequences were optimized. That, it may cause increase mRNA stability and significantly improves immune response (Besse and Ephrussi 2008). Also, the position of the relevant restriction enzymes) BamH I, Hind III) was designed at the beginning and end of the B cell construct.

Final Construct & ORF Checking

A schematic view of the final pattern and location of consensus epitopes in the B cell construct is shown in Supplementary Fig. 3. So, ORF examination show no errors, thus exhibit optimal expression of the construct.

Evaluation of the Primary, Secondary & Tertiary Structure of Designed Construct

A summary of the obtained results of the primary of B cell construct is show in Table 3. The average length of the constructs is 342 bp. IEP is the pH point that the surface of protein is covered with charge, but the net charge of protein is zero. IEP is important for the evaluation of solubility and the mobility in an electric field. The amount IEP was calculate to be 5.78 for B cell construct. The computed value for B cell construct is lower than 7 demonstrating that acidic nature of the protein. The aliphatic index for B cell construct is 55.64. So, indicates that this construct is stable for a wide range of temperature. However the instability index (44.39) offers the estimation of the stability of protein in vitro, and the results classified the constructs with moderate stability. Grand Average Hydropathicity (GRAVY) was negative (− 0.180) for B construct. The GRAVY values indications the degree of protein hydrophilicity and increasing positive score indicates greater hydrophobicity. This one is clear that
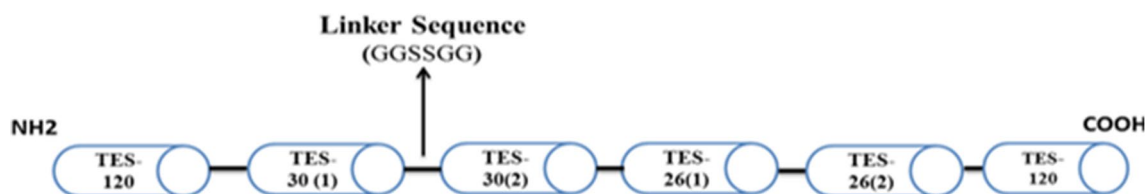

Fig. 2 Schematic illustration of Connect epitopes to each other with flexible linker sequence (Gly-Ser-Gly-Ser-Gly)

Table 1 Predicted epitopes and selected final consensus epitope (Antigen regions were selected)

Pr. name	A	B	C
	TES-120	Tes-30	Tes-26
Accession	U39815	AB009305	U29761
Definition	Surface coat glycoprotein	Proteoglycan core protein	Excretory-secretory antigen
Signal peptide	1–25	1–28	1–22
Type of prediction			
Linear epitope			
LBTope	137–159	33–35, 52–84, 116, 154, 176–188, 217–222	42–70, 135–150, 164–168, 183–186, 200–203, 240–241
SVMTriP	139–158, 97–116	185–204, 150–169	55–74, 150–169, 104–123, 206–225
ABCpred	66–82, 88–104, 25–39	79–95, 282–298, 174–190	21–37, 144–160, 51–67, 34–50
Bepipred	20–120	66–95, 178–188	130–175, 183–189, 231–241
Kloskar	115–127, 150–172	10–20, 35–70, 134–140, 215–232	90–110, 210–217, 210–232
Emini surface accessbility	20–75, 158–165	78–95, 159–165, 177–186	44–54, 76–82, 102–110, 128–148, 159–168, 225–236, 252–260
Conformational epitope			
CBTope	98–133	83–199, 216–226	35–82, 113–178, 207–262
Consensus selected epitope (antigen regions were selected)	97–160	172–207 52–102	33–83 130–180

Table 2 Number of predicted epitopes for TES-120, TES-30 and TES-26 and selected final consensus epitope

Pr. name	Number of epitopes	Consensus selected epitope	
TES-120	1	VmtttpacidtandcqlftplcfvQpysraiqgrcrrtcnicscqdsandcanfvsvclnpt	
TES-30	2	(Epitope 1)	Cvppqvcvapmcvapppaatttaapgvtttrpracppnwtpfnnncyia
		(Epitope 2)	vifsnwrpsqpdgccgsnvtcafvnyanflgqwdd
TES-26	2	(Epitope 1)	nagscftrpvsqvlqnrcqrtcntcdcrdeanncaasinlcqnptfeplvr
		(Epitope 1)	Ntlttaqvanqptvtweaqpndrytlimvdpdfpsaangqqgqrlhwwvin

Table 3 Parameters computed for B cell construct predicted by Expasy ProtParam tool

	B cell construct
No. of amino acids	342
Molecular weight	36, 101.43
Instability index	44.39
Theoretical pI	5.78
Aliphatic index	55.64
Grand average of hydropathicity (GRAVY)	−0.180

construct has high hydrophilicity and tendencies to interact with surrounding water molecules.

Secondary structure B cell construct was predicted using SOPMA Server, and the results is given in Fig. 3. Our results shown that α-helixes, β-turns, random coils and extended strands account for 8.19, 6.14, 60.53, and 25.15% of the secondary structures, respectively. The high proportion of random coils and extended strands in the structure of B

cell construct suggest that the protein may form antigenic epitopes.

For epitopes prediction, it is essential to determine three-dimensional structure of the proteins. Therefore, using I-TASSER servers, tertiary structures of the proteins were predicted. For each target, I-TASSER simulations generate a large ensemble of structural conformations, called decoys. To select the final models, I-TASSER uses the SPICKER program to cluster all the decoys based on the pair-wise structure similarity and reports up to five models which corresponds to the five largest structure clusters. The confidence of each model is quantitatively measured by C-score that is calculated based on the significance of threading template alignments and the convergence parameters of the structure assembly simulations. C-score is typically in the range of (−5, 2), where a C-score of a higher value signifies a model with a higher confidence (Roy et al. 2010).These results are displayed in Fig. 4. The modelled structure was evaluated by RAMPAGE and was used to generate a Ramachandran plot.

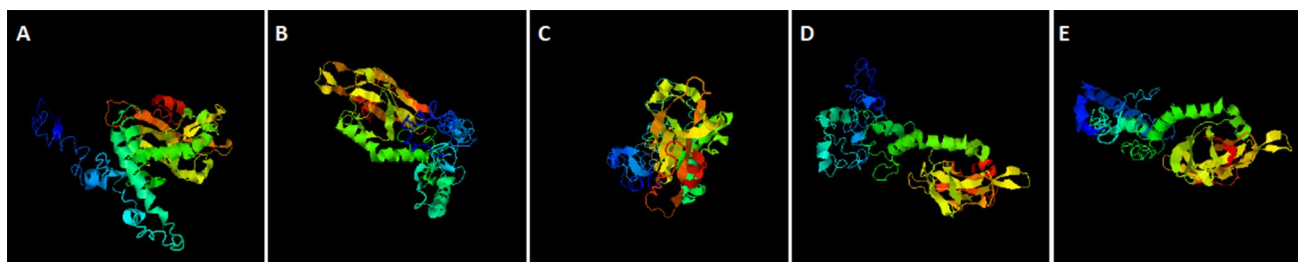**Fig. 3** Secondary structure of B Cell Construct predicted by using SOPMA software

Fig. 4 Tertiary structure B cell construct predicted by I-TASSER server. **a** Model 1: C-score = − 3.03 **b** model 2: C-score = − 3.86 **c** model 3: C-score = − 3.73 **d** model 4: C-score = − 4.09 **e** model 5: C-score = − 3.95

The results of the Ramachandran plot showed 56.2% of residues were found in the favored regions of the plot and the number of residues in the allowed region was 41.3%. Conversely, 2.5% of the residues were found in the outlier region (Fig. 5). Also, ProSA-web zscore plot for 3D structure for B Cell Construct predicted before and after minimization showed that the z-score of the initial model was − 2.92 and the z-score of the model after minimization processes was − 2.94 (Fig. 6a, b). These results indicate that the 3D modeled structure is reliable.

Discussion

Toxocariasis is a serious zoonotic parasitic disease which affects vertebrates with high impact on public health worldwide (Despommier 2005). The development of a highly specific, sensitive and reliable assay to detect the presence of anti-*Toxocara* antibodies is an key aim toward improving the diagnosis (Roldan and Espinoza 2009). Commercial diagnostic kits are known to have
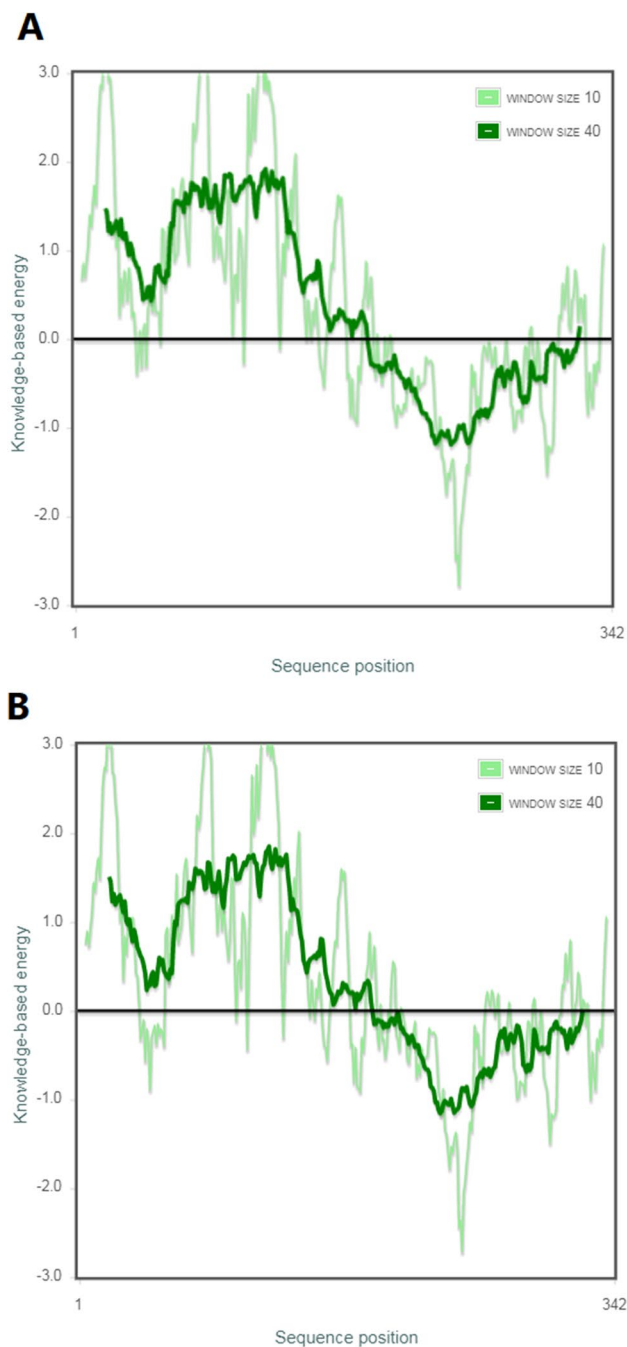
Fig. 5 Ramachandran plot models for B cell construct predicted

Ramachandran Plot
epitope_noHs (1 models)

Plot statistics		
Residues in most favoured regions [A,B,L]	159	56.2%
Residues in additional allowed regions [a,b,l,p]	95	33.6%
Residues in generously allowed regions [~a,~b,~l,~p]	22	7.8%
Residues in disallowed regions	7	2.5%
	---	---
Number of non-glycine and non-proline residues	283	100.0%
Number of end-residues (excl. Gly and Pro)	2	
Number of glycine residues (shown as triangles)	29	
Number of proline residues	28	

Total number of residues	342	

issues with specificity when used in countries endemic with soil-transmitted helminthiasis, which is because of the non-specific nature of components of native TES antigens that cross-react with other helminth antigens. Therefore, native TES antigen is suitable only for differential diagnosis, and test interpretation is difficult when the result is positive. To control this limitation, recombinant antigen can standardize diagnostic methods and increase the sensitivity and specificity of these tests. One description for the high specificity of the recombinant antigen is that it is a molecule with a molecular mass with high immunogenicity, while TES consists of multiple components with a wide range of molecular masses (Maizels

et al. 1984). Also, in contrast to glycosylated TES the recombinant antigen produced in bacteria is not glycosylated. This can also lead to a decrease in cross-reactivity with antibodies that recognize the sugar moieties of the TES produced in *T. canis* larvae (Maizels et al. 1987). Therefore, the study of proteins can be useful for diagnosis and therapeutic purposes. Also, the use of tools for in silico analysis are needed to predict structural and functional features of proteins (Pruess and Apweiler 2003). Recently, with the development of bioinformatics studies, epitope prediction has drastically developed (Frank 2002; Karpenko et al. 2014). Performing predictions with a multi-parameter and method analysis greatly enhances

A

B

Fig. 6 ProSA-web z-score plot for 3D structure of B cell construct predicted before and after minimization. **a** The z-score of the initial model is −2.92 and **b** the z-score of the model after minimization processes is −2.94

the accuracy of the epitope prediction (Saha et al. 2005). So, the aim of the present study was prediction and design of B cell multi-epitope antigenic *T. canis* and evaluation of its immunogenicity. For this purpose, a number of online prediction software applications, including IEDB, Bcepred and ABCpred, were used. Making predictions using a multi-parameter and multi-method analysis improves the

accuracy of epitope prediction significantly. For example the flexibility parameter prediction show the ability to bend of protein. By a greater flexibility, protein has a high capacity to bend, thus helping the formation of a secondary structure. The hydrophilicity parameter prediction describes the location of hydrophilic residues in the amino acid sequence of the protein. The hydrophilic residues are located on the out of the protein and are suitable for ligand binding. The dominant epitopes are more likely to be in regions with a high hydrophilicity.

According to the results of this study, the distribution of amino acids & the number of hydrophilic residues in selected epitopes are more than hydrophobic amino acids. Therefore the most of them are located outside the membrane. Also, the transmembrane structure of proteins was predicted using the online CBS prediction software TMHMM Server. The results of transmembrane topology can be useful in the selection of proper epitope indications. These analysis of transmembrane structure, show that multi-epitope is more outside regions. Thus, it demonstrations that the proteins have suitable solubility and have a decent condition of exposure for immune system in body.

So, the secondary structure of protein is predicted by SOPMA which evaluates the percentage of alpha (α) helices, extended strand, random coils and beta (β) turn. The secondary structure of protein was closely related with antigenic features. So, it show that the selected multi-epitope in this study has suitable resistance. The tertiary structure is a three-dimensional globular structure composed of additional coiling and folding of secondary structure elements, such as α-helixes, β-turns, random coils and extended strands therefore, it can account the degree of similarity in sequence protein can be modeled using this tool. Alignment of sequence can be ameliorated manually, and homology methodology is utilized to build the structure of protein.. The modelled structure was evaluated by RAMPAGE and was used to generate a Ramachandran plot. This tool was used in determining the energetically stable conformations of the psi (ψ) and phi (Φ) torsion or dihedral bond angles for each amino acid in the structure. These results indicate that the modeled structure is reliable. Moreover, ProSA-web plot assessment after minimization showed to resulted in modeling of a high quality 3D model. Therefore we conclude that for TES-120 protein, the residues in the region between 97–167 & for TES-30 in the residues 52–102, 172–207 and for TES-26 regions among 33–83, 130–180 have the most immunogenic potential. These regions have decent potential for designing epitopes. These epitopes were fused to each other with proper linkers. Proper linkers have an important role in functional and structural features of a B cell construct and applying various linkers may result in generation of a new protein with different characteristics. So, furthermore to

immunological function of linkers, they prevent the formation of new epitopes (Dorosti et al. 2019; Saadi et al. 2017).

Conclusion

This study aimed to obtain the bioinformatic characteristics the B cell epitopes for three proteins TES-120, TES-30, TES-26 and analyze its immunogenicity. The results of the secondary structure prediction demonstrated that there are potential epitopes in these proteins. So, it can provide a theoretical basis for investigating its antigenicity and provides a theoretical foundation for epitope diagnosis development for human toxocariasis.

Compliance with Ethical Standards

Conflict interest All authors declare that they have no conflict interest.

References

Aghaei S, Riahi SM, Rostami A, Mohammadzadeh I, Javanian M, Tohidi E et al (2018) *Toxocara* spp. infection and risk of childhood asthma: a systematic review and meta-analysis. Acta Trop 182:298–304

Aghamolaie S, Seyyedtabaei SJ, Behniafar H, Foroutan M, Saber V, Hanifehpur H et al (2018) Seroepidemiology, modifiable risk factors and clinical symptoms of *Toxocara* spp. infection in Northern Iran. Trans R Soc Trop Med Hyg 113:116–122

Baneth G, Thamsborg SM, Otranto D, Guillot J, Blaga R, Deplazes P et al (2016) Major parasitic zoonoses associated with dogs and cats in Europe. J Comp Pathol 155(1 Suppl 1):S54–S74

Besse F, Ephrussi A (2008) Translational control of localized mRNAs: restricting protein synthesis in space and time. Nat Rev Mol Cell Biol 9(12):971–980

Bhattacharya A, Tejero R, Montelione GT (2007) Evaluating protein structures determined by structural genomics consortia. Proteins 66(4):778–795

Dai JF, Jiang M, Qu LL, Sun L, Wang YY, Gong LL et al (2013) *Toxoplasma gondii*: enzyme-linked immunosorbent assay based on a recombinant multi-epitope peptide for distinguishing recent from past infection in human sera. Exp Parasitol 133(1):95–100

DeLano WL (2002) The PyMOL molecular graphics system. San Carlos, DeLano Scientific, p 700

Despommier D (2005) Toxocariasis: clinical aspects, epidemiological, medical ecology and molecular aspects. Clin Microbiol Rev 16:265–272

Dipti CA, Jain SK, Navin K (2006) A novel recombinant multiepitope protein as a *hepatitis* C diagnostic intermediate of high sensitivity and specificity. Protein Expr Purif 47(1):319–328

Dorosti H, Eslami M, Negahdaripour M, Ghoshoon MB, Gholami A, Heidari R et al (2019) Vaccinomics approach for developing multi-epitope peptide *pneumococcal* vaccine. J Biomol Struct Dyn 37(13):3524–3535

Fakhri Y, Gasser R, Rostami A, Fan C, Ghasemi S, Javanian M et al (2018) *Toxocara* eggs in public places worldwide—a systematic review and meta-analysis. Environ Pollut 242:1467–1475

Fong MY, Lau YL (2004) Recombinant expression of the larval excretory-secretory antigen TES-120 of *Toxocara canis* in the *methylotrophic* yeast *Pichia pastoris*. Parasitol Res 92(2):173–176

Fong MY, Lau YL, Init I, Jamaiah I, Anuar AK, Rahmah N (2003) Recombinant expression of *Toxocara canis* excretory-secretory antigen TES-120 in *Escherichia coli*. Southeast Asian J Trop Med Public Health 34(4):723–726

Frank SA (2002) Immunology and evolution of infectious disease. Princeton, Princeton University Press

Gasteiger E, Gattiker A, Hoogland C, Ivanyi I, Appel RD, Bairoch A (2003) ExPASy: the proteomics server for in-depth protein knowledge and analysis. Nucleic Acids Res 31(13):3784–3788

Geourjon C, Deleage G (1995) SOPMA: significant improvements in protein secondary structure prediction by consensus prediction from multiple alignments. Comput Appl Biosci 11(6):681–684

Hajissa K, Zakaria R, Suppian R, Mohamed Z (2015) Design and evaluation of a recombinant multi-epitope antigen for serodiagnosis of *Toxoplasma gondii* infection in humans. Parasit Vectors 8:315

Kalwy S, Rance J, Young R (2006) Toward more efficient protein expression: keep the message simple. Mol Biotechnol 34(2):151–156

Karpenko LI, Bazhan SI, Antonets DV, Belyakov IM (2014) Novel approaches in polyepitope T-cell vaccine development against HIV-1. Expert Rev Vaccines 13(1):155–173

Kolaskar AS, Tongaonkar PC (1990) A semi-empirical method for prediction of antigenic determinants on protein antigens. FEBS Lett 276(1–2):172–174

Krogh A, Larsson B, von Heijne G, Sonnhammer EL (2001) Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes. J Mol Biol 305(3):567–580

Kulkarni R, Sapkal G, Mahishi L, Shil P, Gore MM (2012) Design and characterization of polytope construct with multiple B and TH epitopes of Japanese *encephalitis* virus. Virus Res 166(1–2):77–86

Lv C, Hong Y, Fu Z, Lu K, Cao X, Wang T et al (2016) Evaluation of recombinant multi-epitope proteins for diagnosis of goat schistosomiasis by enzyme-linked immunosorbent assay. Parasit Vectors 9:135

Maizels RM, de Savigny D, Ogilvie BM (1984) Characterization of surface and excretory-secretory antigens of *Toxocara canis* infective larvae. Parasit Immunol 6(1):23–37

Maizels RM, Kennedy MW, Meghji M, Robertson BD, Smith HV (1987) Shared carbohydrate epitopes on distinct surface and secreted antigens of the parasitic nematode *Toxocara canis*. J Immunol 139(1):207–214

Mohamad S, Azmi NC, Noordin R (2009) Development and evaluation of a sensitive and specific assay for diagnosis of human toxocariasis by use of three recombinant antigens (TES-26, TES-30USM, and TES-120). J Clin Microbiol 47(6):1712–1717

Mohammadzadeh I, Riahi SM, Saber V, Darvish S, Amrovani M, Arefkhah N et al (2018) The relationship between *Toxocara* species seropositivity and allergic skin disorders: a systematic review and meta-analysis. Trans R Soc Trop Med Hyg 112(12):529–537

Noordin R, Smith HV, Mohamad S, Maizels RM, Fong MY (2005) Comparison of IgG-ELISA and IgG4-ELISA for *Toxocara* serodiagnosis. Acta Trop 93(1):57–62

Norhaida A, Suharni M, Liza Sharmini AT, Tuda J, Rahmah N (2008) rTES-30USM: cloning via assembly PCR, expression, and evaluation of usefulness in the detection of toxocariasis. Ann Trop Med Parasitol 102(2):151–160

Overgaauw PA (1997) Aspects of *Toxocara* epidemiology: human toxocarosis. Crit Rev Microbiol 23(3):215–231

Pruess M, Apweiler R (2003) Bioinformatics Resources for in silico proteome analysis. J Biomed Biotechnol 2003(4):231–236

Ramakrishna L, Anand KK, Mohankumar KM, Ranga U (2004) Codon optimization of the tat antigen of human immunodeficiency virus type 1 generates strong immune responses in mice

following genetic immunization. J Virol 78(17):9174–9189 **Epub 2004/08/17**

Roldan WH, Espinoza YA (2009) Evaluation of an enzyme-linked immunoelectrotransfer blot test for the confirmatory serodiagnosis of human toxocariasis. Mem Inst Oswaldo Cruz 104(3):411–418

Roy A, Kucukural A, Zhang Y (2010) I-TASSER: a unified platform for automated protein structure and function prediction. Nat Protoc 5(4):725–738

Rubinsky-Elefant G, Hirata CE, Yamamoto JH, Ferreira MU (2010) Human toxocariasis: diagnosis, worldwide seroprevalences and clinical expression of the systemic and ocular forms. Ann Trop Med Parasitol 104(1):3–23

Saadi M, Karkhah A, Nouri HR (2017) Development of a multi-epitope peptide vaccine inducing robust T cell responses against *brucellosis* using immunoinformatics based approaches. Infect Genet Evol 51:227–234

Saha S, Bhasin M, Raghava GP (2005) Bcipep: a database of B-cell epitopes. BMC Genom 6:79

Sandhu KS, Pandey S, Maiti S, Pillai B (2008) GASCO: genetic algorithm simulation for codon optimization. In Silico Biol 8(2):187–192

Sela-Culang I, Kunik V, Ofran Y (2013) The structural basis of antibody-antigen recognition. Front Immunol 4:302

Singh H, Ansari HR, Raghava GP (2013) Improved method for linear B-cell epitope prediction using antigen's primary sequence. PLoS ONE 8(5):e62216

Siyadatpanah A, Tabatabaei F, Zeydi AE, Spotin A, Fallah-Omrani V, Assadi M et al (2013) Parasitic contamination of raw vegetables in Amol North of Iran. Arch Clin Infect Dis 8(2):e15983

Smith H, Holland C, Taylor M, Magnaval JF, Schantz P, Maizels R (2009) How common is human toxocariasis? Towards standardizing our knowledge. Trends Parasitol 25(4):182–188

Wiederstein M, Sippl MJ (2007) ProSA-web: interactive web service for the recognition of errors in three-dimensional structures of proteins. Nucleic Acids Res 35:W407–W410

Yamasaki H, Araki K, Lim PK, Zasmy N, Mak JW, Taib R et al (2000) Development of a highly specific recombinant *Toxocara canis* second-stage larva excretory-secretory antigen for immunodiagnosis of human toxocariasis. J Clin Microbiol 38(4):1409–1413

Yang ZR (2004) Biological applications of support vector machines. Brief Bioinform 5(4):328–338

Yang J, Yan R, Roy A, Xu D, Poisson J, Zhang Y (2015) The I-TASSER suite: protein structure and function prediction. Nat Methods 12(1):7–8

Zhang W, Xiong Y, Zhao M, Zou H, Ye X, Liu J (2011) Prediction of conformational B-cell epitopes from 3D structures by random forests with a distance-based feature. BMC Bioinform 12:341