Original Article

In silico Analysis of Immunologic Regions of Surface Antigens (Sags) of *Toxoplasma gondii*

Abbas Alibakhshi[1], Mojgan Bandehpour[1,2*], Tina Nafarieh[1], Shivasadat Gheflat[1], Bahram Kazemi[2]

[1] Department of Biotechnology, School of Advanced Technologies in Medicine, Shahid Beheshti University of Medical Sciences, Tehran, Iran
[2] Cellular and Molecular Biology Research Center, Shahid Beheshti University of Medical Sciences, Tehran, Iran

Received: 29 November; 2016; Accepted: 25 January; 2017

Abstract

Background: Surface antigens (SAGs) of Toxoplasma gondii are known candidates for diagnostic tests and vaccines. The present study argues about the main necessary properties for determination and prediction of T-cell agretopes and B-cell epitopes of surface antigens of *Toxoplasma gondii*.

Materials and Methods: Primary, secondary and tertiary structures of the proteins were analyzed by different methods. The three-dimensional structures were determined by use of *ab initio* method for prediction of discontinues epitopes. The agretopes and epitopes were predicted via several various web servers with different methods employed.

Results: The results of *in silico* analyses showed that the regions 129-GAPAGRNNDGSSAPT-143 for protein p22, 234-SENPWQGNASSD-245 for protein p30 and 348-PGTEGESQAGT-358 for protein p43, have the highest immunogenic potential.

Conclusion: We reached to three antigenic epitopes for cloning and protein expression. In following the purified polypeptide will be applied for diagnosis of Toxoplasma gondii.

Keywords: Epitope, Agretope, SAG, *Toxoplasma gondii*, *In silico*

***Corresponding Author:** Mojgan Bandehpour; Cellular and Molecular Biology Research Center, Shahid Beheshti University of Medical Sciences, Tehran, Iran. Tel: (+98) 21 22439957; Email: Bandehpour@gmail.com

Please cite this article as: Alibakhshi A, Bandehpour M, Nafariyeh T, sadat Gheflat Sh, Kazemi B. *In silico* Analysis of Immunologic Regions of Surface Antigens (Sags) of *Toxoplasma gondii*. Novel Biomed. 2017;5(3):109-18.

Introduction

Toxoplasma gondii (*T. gondii*) is an obligate intracellular protozoan that infects virtually all of the warm-blooded animals including humans and hence could be successfully distributed worldwide[1,2]. Furthermore, acquired infection during pregnancy may cause severe damage to fetus[3,4]. Infection mainly is transmitted through food or water contaminated with the parasite oocytes excreted by cats or by eating raw meat containing cysts, and thus disease so-called toxoplasmosis occurs. In fact, after ingesting the parasite invades intestinal epithelial cells and into host cells via some of the factors, is able to inhibit phagosome-lysosome fusion[4].

The factors that are of utmost importance and play a role in the pathogenesis, surface antigens (SAGs) can take into consideration. Namely, SAG1 (p30) plays an essential role in attachment of tachyzoites (one of the several different infectious forms of T. gondii) to the host receptor and in the invasion of tachyzoites into host cells. It has been shown that an anti-SAG1 antibody can partially inhibit the invasion of tachyzoites to host cells[5]. SAG2 (p22) is another protein on the surface of the

parasite, which is a binding ligand and also has a good immunogenicity. SAG3 (p43) is found in both tachyzoites and bradyzoite (other infectious form of T. gondii) and like SAG1, anchors to membrane through Glycosyl Phosphatidyl Inositol (GPI). The role of SAG3 is cellular invasion and attachment, as well6. These factors are used in diagnostic tests and subunit vaccines and are promising candidates for vaccines[4].

One of the new approaches in vaccine design is construction of synthetic polyepitopes using T- and B-cells epitopes that can cause different T-cell responses and induction of neutralizing antibodies[7,8]. Epitopes are generally divided into two types: Linear epitopes (LEs) that comprised of continuous amino acids in the primary structure of proteins and conformational epitopes (CEs) that composed of dispersed amino acids among discontinuous regions but become aggregated on the protein surface. T-cells only recognize linear epitopes. However, B-cell epitopes are linear or conformational. Conformational epitopes have much greater proportion and their prediction is dependent on three-dimensional structural information. Thus, predict of these epitopes has become more difficult and laborious[9-13]. The prediction and applying all of the epitopes can be beneficial in development of vaccines and diagnostic tests. This study aimed to evaluate and predict the potential B-cell epitopes and T-cell agretopes of surface antigens, p22, p30 and p43 of *Toxoplasma gondii* using *in silico* methods.

Methods

Sequence retrieval, translation and alignment: Nucleotide sequences for the surface proteins, p22 (GenBank: JX045478.1), p30 (Reference Sequence: XM_002368164.1) and p43 (GenBank: AF340227.1), were retrieved from the National Centre for Biotechnology Information (NCBI) Nucleotide Database. Basically, the target sequences were selected from well-known RH strain of *Toxoplasma gondii*, which is an extremely virulent strain[14] .The sequences were translated into amino acid (http://web.expasy.org/translate/) and then aligned with protein sequences of different virulent strains of *Toxoplasma gondii*, obtained from UniProt Knowledgebase (www.uniprot.org/help/uniprotkb), using the Clustal algorithm available in the ClustalX2 software[15].

Analyzing primary structure, physicochemical properties and transmembrane topology: For the analysis of primary protein structure, the putative protein sequences of p22, p30 and p43 were submitted to Expasy tools (http://web.expasy.org/protparam/)[16]. Amino acids distribution was evaluated in the lrrfinder server (http://www.lrrfinder.com/lrrfinder.php) and transmembrane topology of the proteins was examined according to a hidden Markov model (HMM) in the TMHMM server v. 2.0[17].

Evaluation of potential N-glycosylation, phosphorylation and palmitoylation sites: For prediction of potential N-glycosylation and phosphorylation sites in the proteins, we used NetNGlyc 1.0 (http://www.cbs.dtu.dk/services/NetNGlyc/) and Phospho.ELM (http://phospho.elm.eu.org)[18] servers, which employ Artificial Neural Networks (ANN) for prediction[19]. Palmitoylation sites were predicted with the help of CSS-Palm 2.0 software. Glycosylation, phosphorylation and palmitoylation are the covalent attachment of a carbohydrate to the sequence context of Asn-Xaa-Ser/Thr sequences, a phosphate (PO4) group to serine, threonine and tyrosine and fatty acids, such as palmitic acid to cysteine residues of the proteins, respectively.

Prediction and analysis of secondary structure: Jpred3, a secondary structure prediction server[20], was used to identify the alpha helices, beta strands and other residues. This server utilizes hybrid method in predicting the structure of a protein. Hybrid methods with accuracy in the range of 70-77% are a combination of alignment-based and single sequence-based methods. For further analysis, the method of GOR4 (http://npsa-pbil.ibcp.fr/cgi-bin/npsa_automat.pl?page=npsa_gor4.html)[21] was used as well.

Prediction and analysis of tertiary structure: Since the parasite proteins are typically unique, *ab initio* method was taken as means of identifying tertiary structure of the p22, p30 and p43 surface proteins of *Toxoplasma gondii*. This method can build structure with no prior information. Thus, we made use of bioinformatics tools available on the Robetta server, a full-chain protein structure prediction server[22], and especially Rosetta 2014.35 software. For further evaluation on tertiary structure, we also use I-TASSER server[23]. These two servers can produce results based on template-based homology modeling (fold-recognition) or can additionally build models in the absence of a template

(*ab initio*)[22].

The best model for each protein was selected by GROMOS96 force field[24] application (with computation of energy) in Swiss-PdbViewer 4.1.0 software. For confirmation of the predicted structures, the Ramachandran plot was studied through PROCHECK[25] analyses in the PSVS server v. 1.5 (http://psvs-1_5-dev.nesg.org)[26]. Finally, 3D models were analyzed by the PyMOL software.

Prediction of T-cell agretopes: The sequences of proteins p22, p30 and p43 were submitted to IEDB (Immuno Epitope Database)[27], MHCpred[28] and SYFPEITHI[29] server websites to predict MHC class I epitopes and IEDB, MHCpred and ProPred[30] to predict MHC class II epitopes. Among all HLA alleles, HLA-A0201 and HLA-DRB10101 are the most available ones in the world[31,32]. Therefore, redictions of agretopes were performed for these alleles.

Prediction of linear and discontinuous B-cell epitopes: For prediction of linear B-cell epitopes, the IEDB, ABCpred[33], Bcepred[34], and Bepipred[35] web servers were employed, so that the epitopes predictions were mainly on the basis of Chou and Fasman beta-turn[36], Emini surface accessibility[37], Karplus and Schulz flexibility[38], Kolaskar and Tongaonkar antigenicity[39], Parker hydrophilicity[40] and Bepipred linear epitope prediction[35] (in the IEDB and in the Bcepred server), Artificial Neural Network (ANN) information-processing paradigm (in the ABCpred server) and Hidden Markov Model (HMM) algorithm (in the Bepipred server).

The discontinuous B-cell epitopes were predicted by online prediction tools in the Discotope[41] and the ElliPro[42] web servers, hosted in IEDB, and also by those given in the SEPPA 2.0[43] web server and Superficial 1.2 software[44].

Results

Sequence retrieval, translation and alignment: After translation of the nucleotide sequences retrieved from NCBI Nucleotide Database, the protein sequences of p22 (186 amino acids), p30 (319 amino acids), and p43 (385 amino acids) were aligned with protein sequences of different strains of *Toxoplasma gondii* obtained from UniprotKB (Fig.1). As seen in figure 1, multiple alignments of the three protein sequences with its equal proteins from other strains represented a very high similarity.

Analysis of physicochemical properties and transmembrane topology: An overview of most important data obtained from Expasy server (ProtParam tool) is given in table 1. Every three proteins especially p30 and p43 have isoelectric point (pI) above 7 and therefore are somewhat basic in nature. Proteins p22 and p43 according to instability index can be classified as unstable and p30 protein considered as a stable protein. Aliphatic index of all three proteins are relatively low and this means that at different temperatures are relatively unstable. Grand Average Hydropathicity (GRAVY) shows the degree of protein hydrophilicity and increasing positive score indicates greater hydrophobicity. Here it is clear that p43 protein compared with two other proteins has more hydrophilicity and more tendencies to interact with surrounding water molecules.

Moreover, the charts of amino acids distribution show that number of amino acids of proteins with hydrophilic positive and negative amino acids in compared with amino acids with hydrophobic R-groups is significant and thus all three proteins contain noteworthy exposed residues.

On the other hand, the results of transmembrane topology that can be helpful in the selection of efficient epitopes, shows that almost all of the proteins are in the outer part of the membrane and only in the p43, amino acid residues between 21 and 40 is located in the region of the transmembrane.

Analysis of potential N-glycosylation, phosphorylation and palmitoylation sites: The results of potential glycosylation using server Net NGlyc 1.0 shows that the protein p30 at positions 178 (NSTL) and 241 (NASS), appears to be prone to glycosylation but in the proteins p22 and p43, no sites predicted. In addition, the analysis of phosphorylation using Phospho.ELM server showed that in protein p30, there is a potential phosphorylation in GSPEKHH sequence at positions 271-274, however this modification was not observed in other proteins. Finally, analysis of potential palmitoylation results by using CSS-Palm 2.0 software indicated that there is only the probability of a palmitoylation at the cysteine residue at position 31 of p43 with a cutoff 10.722.

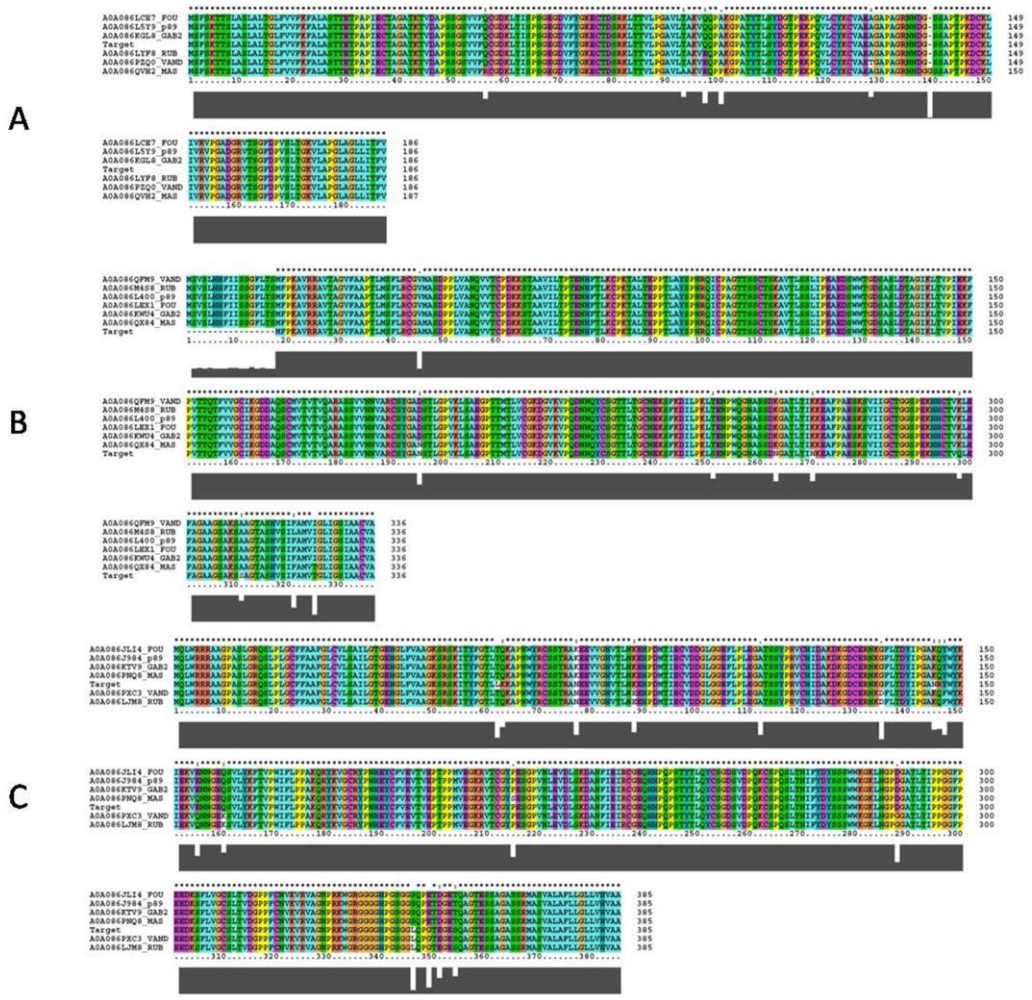

Figure 1. Multiple alignment of protein sequences of surface proteins of different virulent strains of *Toxoplasma gondii.* (A) p22, (B) p30, (C) p43.

Protein secondary and tertiary structures prediction and validation of modeled structures: The prediction results of secondary structure, illustrated the alpha-helices and beta-turns parts of proteins (Fig. 2). For B-cell epitopes prediction, it was necessary that first we determine the spatial and three-dimensional structure of the proteins. Therefore, using the aforementioned servers, tertiary structures of the proteins were predicted. From the presented models, using GROMOS96 force field application that minimizes the energy of models, the best ones for each protein were selected. After all, the quality evaluation of the modeled tertiary structures using PROCHECK analyses, revealed that for the selected model of protein p22, 86.0% of residues are in the favored regions, 13.3% of that are in the allowed regions (additional allowed + generously allowed) and 0.7% are in the disallowed regions. These values for the model of protein p30 are 89.0%, 11.0% and 0.0% and for the model of protein p43, are 81.2%, 18.5% and 0.3%, respectively (Fig. 3). These imply that the modeled structures are reliable.

T-cell agretopes prediction: Table 2 shows T-cell agretopes predicted using different servers. These agretopes were predicted in all relevant servers and were with the highest score in all servers.

Linear and discontinuous B-cell epitopes prediction: Table 3 shows predicted conformational and linear epitopes for B-cells. Overwhelming majority of the epitopes is common at the all servers and besides, only epitopes with the highest scores were selected.

In figure 4, the predicted conformational B-cell epitopes on the 3D structures of the protein is schematically shown.

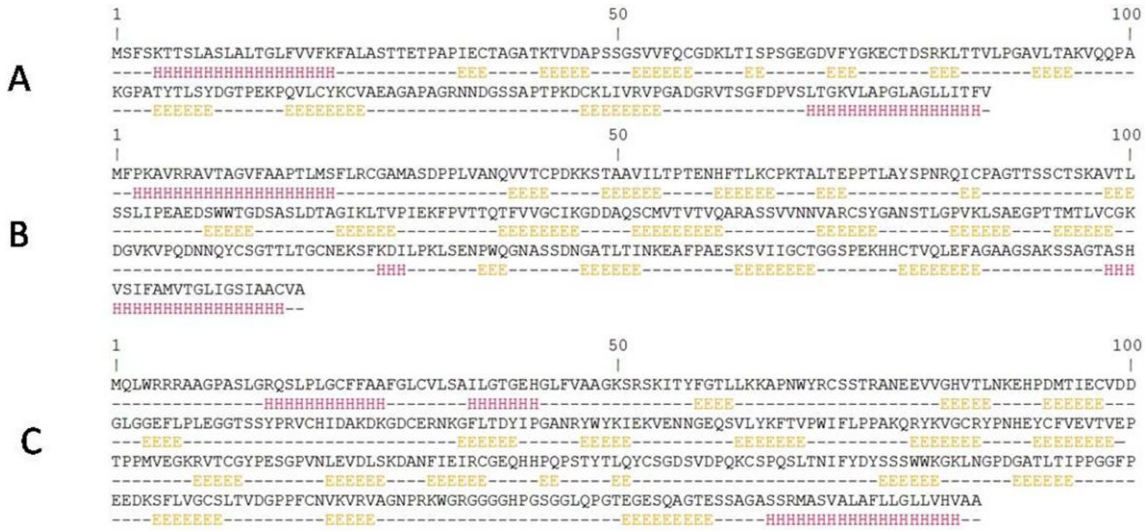

Figure 2. Secondary structure, the alpha-helices and beta-turns parts of the proteins. (A) p22, (B) p30, (C) p43. E: Alpha Helix (α-helix), H: Extended Strand (β-strand).

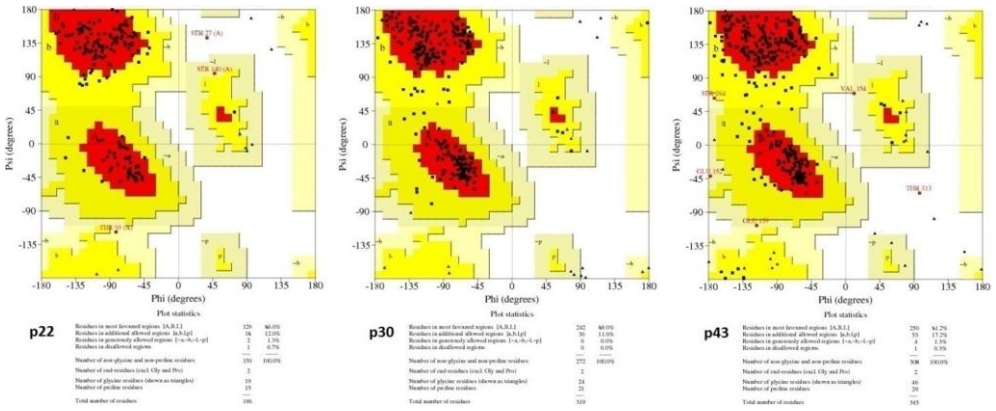

Figure 3. Ramachandran plot of predicted models for the surface proteins p22, p30 and p43.

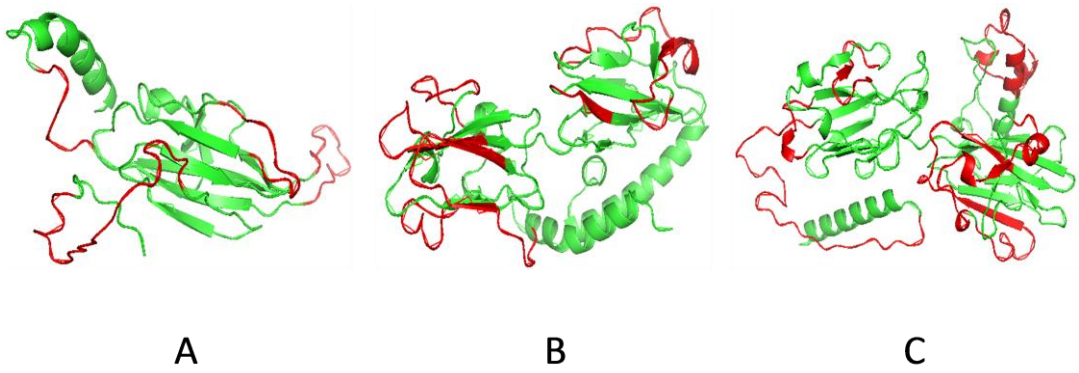

Figure 4 The schematically 3D structures of the proteins. Predicted conformational B-cell epitopes with highest score is shown in red color. (A) p22, (B) p30, (C) p43.

Discussion

In the present study, the protein sequences obtained from different strains of *Toxoplasma gondii* showed a high degree of homology to each other. This highlights that designing of vaccine with using these surface proteins

Table 1: Parameters computed for the proteins by using Expasy ProtParam tool.

	P22	P30	P43
No. of amino acids	186	319	385
Molecular weight	18980.7	32916.5	41785.2
Theoretical pI	8.19	7.89	7.07
Instability index	40.77	37.36	43.73
Aliphatic index	81.29	76.49	69.12
Grand average of hydropathicity (GRAVY)	0.123	0.055	-0.343

can cause similar and universal immune in the all strains. Another important issue that must be considered in the design of a vaccine is the physicochemical properties of proteins. For example, the isoelectric point (pI) can be used to estimate the solubility at a certain pH. Molecules at pH equal to its pI often precipitate in solution. General charge of proteins is determined by the constituent amino acids. Every amino acid has naturally the positive, negative, neutral or polar charge. According to table 1 and figure 1 of the present study, these three proteins, p30 and p43 in particular have the basic nature. Furthermore, in the distribution of amino acids, the number of hydrophilic residues and the residues with positive and negative R group, compared with hydrophobic amino acids, are more and therefore most of that is on the outside of the membrane.

Due to the analysis of trans-membrane topology, it can be also mentioned that all three proteins are completely outside the cells. Previous laboratory studies have shown that these proteins are anchored to the membrane by GPI[45]. Accordingly, it appears that the proteins have significant solubility and have a good level of exposure for immune system. In addition, the grand average of hydropathicity (GRAVY) that measures the hydrophilic amounts based on Kyte-Doolittle scale[46], indicates that these proteins particularly p43 have good hydrophilicity. The other factor, aliphatic index is a measure of the relative volume occupied by the side chain of the alanine, valine, leucine and isoleucine amino acids that shows the protein stability against different *temperatures*. This amounts for three proteins used in this study is relatively low that this condition, for example, can be a sign of high content of amino acids forming hydrogen bonds[47].

Another important feature that should be taken into account is post-translational modifications on proteins such as glycosylation, phosphorylation, palmitoylation etc. The studies show that glycosylation has effects on the *thermodynamics*, folding, hydrophobicity, antigenicity and immunogenicity of proteins. This post-translational modification may disrupt local hydrogen bonding network and hence reduces the solubility of the peptide[48,49]. The addition of a molecule of phosphate (PO4) on residues of amino acids also may have a significant impact on the hydrophobicity of the molecule[50]. Palmitoylation improve surface hydrophobicity of proteins, membrane affinity and several other biochemical changes[51]. In this study, the post-translational modifications have no large share in proteins and it can be considered that this category has the minimal change in immunogenicity of the amino acids of these proteins.

Table 2: Predicted T-cell agretopes by using IEDB, MHCpred and SYFPEITHI servers for HLAA0102 allele and IEDB, MHCpred and ProPred for HLADRB10101 allele. The agretopes were predicted at the all servers with highest score.

HLA type	p22	p30	p43
T cell – HLA – HLAA0102 allele	173-VLAPGLAGL-181 11-SLALTGLFV-19 82-KLTTVLPGA-90 148-KLIVRVPGA-156	28-AMASDPPLV-36 99-TLSSLIPEA-107 302-SIFAMVTGL-310 118-SLDTAGIKL-126 310-LIGSIAACV-318 69-ALTEPPTLA-77	375-FLLGLLVHV-383 306-FLVGCSLTV-314 36-ILGTGEHGL-44 372-ALAFLLGLL-380 136-FLTDYIPGA-144 13-SLGRQSLPL-21
T cell – HLA – DRB10101 allele	17-LFVVFKFALASTTET-31 80-SRKLTTVLPGAVLTA-94 1-MSFSKTTSLASLALT-15 172-KVLAPGLAGLLITFV-186	301-VSIFAMVTGLIGSIA-315 21-MSFLRCGAMASDPPL-35 3-PKAVRRAVTAGVFAA-17 163-ASSVVNNVARCSYGA-177	167-TVPWIFLPPAKQRYK-181 24-FFAAFGLCVLSAILG-38 370-SVALAFLLGLLVHVA-384 102-LGGEFLPLEGGTSSY-116 302-EDKSFLVGCSLTVDG-316 66-PNWYRCSSTRANEEV-80

Success of polytope vaccine approach is dependent on strict criteria for selecting epitopes and also the linkers between these epitopes[7]. Proteins p22, p30, and p43 of T*oxoplasma gondii* were used for diagnostic works and vaccine alone[6,52]. However, some problems related to the expression of proteins and antibodies cross-reaction problems may lead to disturbances in the process. Therefore, the selection of parts of proteins with high immunogenicity can overcome these problems. From the results of the present study, it seems that for p22 protein, especially for induction of antibodies, the residues in the region among 129-143, have the most immunogenic potential. The prediction tools of epitope in p30 protein revealed such region is in the residues 234-245 and likewise for protein p43, region among 348-358, is a region with high immunogenicity.

Conclusion

Finally, in order to build the polytopes that stimulate T cells as well as B cells, it can be made in the following order: N-terminal, Arg-Gly-Asp (RGD), T-cell epitope peptide, di-lysine linker (KK) to B-cell epitope peptide. In this construct, the linker di-lysine is target sequence of one of the important lysosomic proteases in antigen processing and RGD motif improves immunogenicity of peptide by raising the attachment ability[53,54].

Acknowledgment

The present study was extracted from Abbas Alibakhshi's thesis and funded by Vice Dean of Research in Shahid Beheshti University of Medical Sciences (grant no. 4967) and consulted to the doing of the experiment with the ethics code IR.SBMU.RAM.REC.1394.329.

Table 3: The predicted linear and conformational B-cell epitopes for the proteins. The epitopes with highest score and almost shared in the all mentioned servers and software in the text were selected.

	p22	**p30**	**p43**
Predicted linear B-cell epitopes	132-AGRNNDGSSAPT-143 108-LSYDGTP-114 61-LTISPSGEGDV-71 42-TKTVDAPSSGS-52 141-APTPKDCKLIV-151	234-SENPWQGNASSD-245 201-DGVKVPQDNNQY-212 270-GGSPEKHHCT-279 287-AAGSAKSSAGTA-298 185-KLSAEGPTT-193 39-QVVTCPDK-46	348-PGTEGESQAGT-358 297-GGFPEEDKS-305 328-GNPRKWGRGGG-338 209-RVTCGYPE-216 253-CSGDSVDPQKC-263 114-SSYPRVCHID-123 108-PLEGGTSSYP-117 84-VTLNKEHPDM-93
Predicted conformational B-cell epitopes	129-GAPAGRNNDGSSAP-142 156-ADGRVTSGF-164 94-AKVQQPAKGPA-104 10-ASLALTGLFVVFKFAL ASTTETP-32	144-CIKGDDAQSCMVT-156 35-DPPLVANQVVTCPDKKSTAA V-52 215-GTTLTGCNEKSFKDILPKLSE NPWQGNASSDNG-247 87-AGTTSSCTSKAVT-99 113-TGDSASLDTAG-123 273-PEKHHC-278	327-AGNPRKWGRGGGGHPGSGGLQPGTEG ESQAGTES-360 1-MQLWRRRAAGPASLGRQSLPLGCFFAA F-28 162-VLYKFTVPWIFLPPAKQRYKV-182 124-AKDKGDCERNKGFLTDYIPGANRYW- 148 150-KIEKVENNGEQSV-162 295-PPGGFPEED-303 245-PSTYTLQYCSGDSVD-259

References

1. Innes EA. A brief history and overview of Toxoplasma gondii. Zoonoses Public Health. 2010;57(1):1-7.
2. Dubey JP. Chapter 1 - The History and Life Cycle of Toxoplasma gondii. In: Weiss LM, Kim K, editors. Toxoplasma Gondii (Second Edition). Boston: Academic Press; 2014. p. 1-17.
3. Sensini A. Toxoplasma gondii infection in pregnancy: opportunities and pitfalls of serological diagnosis. Clin Microbiol Infect. 2006;12(6):504-312.
4. Montoya JG, Liesenfeld O. Toxoplasmosis. Lancet. 2004;12;363(9425):1965-1976.
5. Wang Y, Yin H. Research progress on surface antigen 1 (SAG1) of Toxoplasma gondii. Parasit Vectors. 2014;7:180.
6. Khanaliha K, Motazedian MH, Kazemi B, Shahriari B, Bandehpour M, Sharifniya Z. Evaluation of recombinant SAG1, SAG2, and SAG3 antigens for serodiagnosis of toxoplasmosis. Korean J Parasitol. 2014;52(2):137-142.
7. Kulkarni R, Sapkal G, Mahishi L, Shil P, Gore MM. Design and characterization of polytope construct with multiple B and TH epitopes of Japanese encephalitis virus. Virus Res. 2012;166(1-2):77-86.
8. Karpenko LI, Bazhan SI, Antonets DV, Belyakov IM. Novel approaches in polyepitope T-cell vaccine development against HIV-1. Expert Rev Vaccines. 2014;13(1):155-173.
9. Barlow DJ, Edwards MS, Thornton JM. Continuous and discontinuous protein antigenic determinants. Nature.

1986;21-27;322(6081):747-8.
10. Van Regenmortel MH. Immunoinformatics may lead to a reappraisal of the nature of B cell epitopes and of the feasibility of synthetic peptide vaccines. J Mol Recognit. 2006;19(3):183-187.
11. Wang HW, Lin YC, Pai TW, Chang HT. Prediction of B-cell linear epitopes with a combination of support vector machine classification and amino acid propensity identification. J Biomed Biotechnol. 2011;2011:432830.
12. Benjamin DC. B-cell epitopes: fact and fiction. Adv Exp Med Biol. 1995;386:95-108.
13. Greenbaum JA, Andersen PH, Blythe M, Bui HH, Cachau RE, Crowe J, et al. Towards a consensus on datasets and evaluation metrics for developing B-cell epitope prediction tools. J Mol Recognit. 2007;20(2):75-82.
14. Current Topics In Microbiology and Immunology. *Toxoplasma gondii*: Springer; 1996.
15. Thompson JD, Gibson TJ, Plewniak F, Jeanmougin F, Higgins DG. The CLUSTAL_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. Nucleic Acids Research. 1997;25(24):4876-82.
16. Gasteiger E, Hoogland C, Gattiker A, Duvaud S, Wilkins MR, Appel RD, et al. Protein Identification and Analysis Tools on the ExPASy Server. In: Walker JM, editor. The Proteomics Protocols Handbook: Humana Press Inc; 2005.
17. Krogh A, Larsson B, von Heijne G, Sonnhammer EL. Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes. J Mol Biol. 2001;305(3):567-780.
18. Dinkel H, Chica C, Via A, Gould CM, Jensen LJ, Gibson TJ, et al. Phospho.ELM: a database of phosphorylation sites--update 2011. Nucleic Acids Research. 2011;39(Database issue):D261-7.
19. Dor O, Zhou Y. Real-SPINE: an integrated system of neural networks for real-value prediction of protein structural properties. Proteins. 2007;68(1):76-81.
20. Cole C, Barber JD, Barton GJ. The Jpred 3 secondary structure prediction server. Nucleic Acids Research. 20081;36(Web Server issue):W197-201.
21. Ed D, J G, J-F G, B R. GOR secondary structure prediction method version IV. Methods in Enzymology. 1996;266:450-3.
22. Kim DE, Chivian D, Baker D. Protein structure prediction and analysis using the Robetta server. Nucleic Acids Research. 2004;32(Web Server issue):W526-31.
23. Yang J, Yan R, Roy A, Xu D, Poisson J, Zhang Y. The I-TASSER Suite: protein structure and function prediction. Nat Methods. 2014;12(1):7-8.
24. Kunz AP, Allison JR, Geerke DP, Horta BA, Hunenberger PH, Riniker S, et al. New functionalities in the GROMOS biomolecular simulation software. J Comput Chem. 2012;33(3):340-53.
25. Laskowski RA, MacArthur MW, Moss DS, Thornton JM. PROCHECK: a program to check the stereochemical quality of protein structures. Journal Of Applied Crystallography. 1993;26(2):283-91.
26. Bhattacharya A, Tejero R, Montelione GT. Evaluating protein structures determined by structural genomics consortia. Proteins. 2007;66(4):778-95.
27. Vita R, Overton JA, Greenbaum JA, Ponomarenko J, Clark JD, Cantrell JR, et al. The immune epitope database (IEDB) 3.0. Nucleic Acids Research. 2014.
28. Guan P, Hattotuwagama CK, Doytchinova IA, Flower DR. MHCPred 2.0: an updated quantitative T-cell epitope prediction server. Appl Bioinformatics. 2006;5(1):55-61.
29. Schuler MM, Nastke MD, Stevanovikc S. SYFPEITHI: database for searching and T-cell epitope prediction. Methods in molecular biology. 2007;409:75-93.
30. Singh H, Raghava GP. ProPred: prediction of HLA-DR binding sites. Bioinformatics. 2001;17(12):1236-7.
31. Ranjbar MM, Ghorban K, Alavian SM, Keyvani H, Dadmanesh M, Roayaei Ardakany A, et al. GB Virus C/Hepatitis G Virus Envelope Glycoprotein E2: Computational Molecular Features and Immunoinformatics Study. Hepat Mon. 2013;13(12):e15342.
32. Pelte C, Cherepnev G, Wang Y, Schoenemann C, Volk HD, Kern F. Random screening of proteins for HLA-A*0201-binding nine-amino acid peptides is not sufficient for identifying CD8 T cell epitopes recognized in the context of HLA-A*0201. J Immunol. 2004;172(11):6783-9.
33. Saha S, Raghava GP. Prediction of continuous B-cell epitopes in an antigen using recurrent neural network. Proteins. 2006;65(1):40-8.
34. S S, G.P.S R. BcePred:Prediction of Continuous B-Cell Epitopes in Antigenic Sequences Using Physico-chemical Properties. LNCS. 2004;3239:197-204.
35. Larsen JE, Lund O, Nielsen M. Improved method for predicting linear B-cell epitopes. Immunome Res. 2006;2:2.
36. Chou PY, Fasman GD. Prediction of the secondary structure of proteins from their amino acid sequence. Adv Enzymol Relat Areas Mol Biol. 1978;47:45-148.
37. Emini EA, Hughes JV, Perlow DS, Boger J. Induction of hepatitis A virus-neutralizing antibody by a virus-specific synthetic peptide. J Virol. 1985;55(3):836-9.
38. PA K, GE S. Prediction of Chain Flexibility in Proteins - A tool for the Selection of Peptide Antigens. Naturwissenschafren. 1985;72:212-3.
39. Kolaskar AS, Tongaonkar PC. A semi-empirical method for prediction of antigenic determinants on protein antigens. FEBS Lett. 1990;276(1-2):172-4.
40. Parker JM, Guo D, Hodges RS. New hydrophilicity scale derived from high-performance liquid chromatography peptide retention data: correlation of predicted surface residues with antigenicity and X-ray-derived accessible sites. Biochemistry. 1986;25(19):5425-32.
41. 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(12):e1002829.
42. Ponomarenko J, Bui HH, Li W, Fusseder N, Bourne PE, Sette A, et al. ElliPro: a new structure-based tool for the prediction of antibody epitopes. BMC Bioinformatics. 2008;9:514.
43. Qi T, Qiu T, Zhang Q, Tang K, Fan Y, Qiu J, et al.

SEPPA 2.0--more refined server to predict spatial epitope considering species of immune host and subcellular localization of protein antigen. Nucleic Acids Research. 201;42(Web Server issue):W59-63.
44. Goede A, Jaeger IS, Preissner R. SUPERFICIAL--surface mapping of proteins via structure-based peptide library design. BMC Bioinformatics. 2005;6:223.
45. Tomavo S, Schwarz RT, Dubremetz JF. Evidence for glycosyl-phosphatidylinositol anchoring of Toxoplasma gondii major surface antigens. Mol Cell Biol. 1989;9(10):4576-80.
46. Kyte J, Doolittle RF. A simple method for displaying the hydropathic character of a protein. J Mol Biol. 1982;157(1):105-32.
47. Ikai A. Thermostability and aliphatic index of globular proteins. J Biochem. 1980;88(6):1858-95.
48. Gavrilov BK, Rogers K, Fernandez-Sainz IJ, Holinka LG, Borca MV, Risatti GR. Effects of glycosylation on antigenicity and immunogenicity of classical swine fever virus envelope proteins. Virology. 2011;420(2):135-45.
49. Cheng S, Edwards SA, Jiang Y, Grater F. Glycosylation enhances peptide hydrophobic collapse by impairing solvation. Chemphyschem. 2010;11(11):2367-74.
50. Ubersax JA, Ferrell JE, Jr. Mechanisms of specificity in protein phosphorylation. Nat Rev Mol Cell Biol. 2007;8(7):530-41.
51. Draper JM, Xia Z, Smith CD. Cellular palmitoylation and trafficking of lipidated peptides. J Lipid Res. 2007;48(8):1873-84.
52. Letscher-Bru V, Pfaff AW, Abou-Bacar A, Filisetti D, Antoni E, Villard O, et al. Vaccination with Toxoplasma gondii SAG-1 protein is protective against congenital toxoplasmosis in BALB/c mice but not in CBA/J mice. Infect Immun. 2003;71(11):6615-9.
53. Yano A, Miwa Y, Kanazawa Y, Ito K, Makino M, Imai S, et al. A novel method for enhancement of peptide vaccination utilizing T-cell epitopes from conventional vaccines. Vaccine. 2013;31(11):1510-5.
54. Yano A, Onozuka A, Asahi-Ozaki Y, Imai S, Hanada N, Miwa Y, et al. An ingenious design for peptide vaccines. Vaccine. 2005;23(17–18):2322-6.