

Structural and antigenic analysis and cloning of surface antigen protein 5 in *Toxoplasma gondii*

Lu, G.¹, Wang, L.³, Zhou, A.H.², Han, Y.L.¹, Zhou, J.¹, Guo, J.J.¹, Song, P.X.¹, Zhou, H.Y.¹, Cong, H.¹, Hou, M.⁴ and He, S.Y.^{1*}

¹Department of Parasitology, Shandong University School of Medicine, Jinan, Shandong Province 250012, Peoples Republic of China

²Department of Pediatrics, Provincial Hospital Affiliated to Shandong University, Shandong University School of Medicine, 250021, Jinan, Shandong Province, Peoples Republic of China

³Department of Ji Nan Children's Hospital, 250022, Jinan, Shandong Province, Peoples Republic of China

⁴Department of Hematology, Qilu Hospital, Shandong University, 250012, Jinan, China

*Corresponding author e-mail: shenyi.he@hotmail.com

Received 17 March 2016, received in revised form 14 August 2016; accepted 16 August 2016

Abstract. *Toxoplasma gondii* is identified as an obligate intracellular apicomplexan parasite that infects warm blooded animals and humans worldwide. SAG5 protein includes SAG5A, -5B, -5C, -5D, and -5E five subtypes. SAG5A, -5B, -5C, and -5D are expressed on the surface of *Toxoplasma gondii*. In this study, we used online T-Coffee tool to analyze SAG5 proteins sequence alignment. SMART software was used to predict secondary structures of SAG5A, -5B, -5C, and -5D. The 3D models of SAG5 proteins were constructed and analyzed with SWISS-MODEL server and VMD software. Results indicated that SAG5A, -5B, -5C, and -5D are highly homologous proteins. Furthermore, liner-B cell epitopes and Th-cell epitopes of the four proteins were predicted using DNAMAN software and Epitope Database online service. The bioinformatics analysis of SAG5A, -5B, -5C, and -5D proteins could provide valuable information on prevention and treatment of toxoplasmosis. In addition, the four genes were obtained by PCR and inserted into an eukaryotic expression vector pEGFP-C1 respectively. Identified by restriction enzyme digestion, the four recombinant plasmids were transfected into HEK 293-T cells and tested by RT-PCR. Results showed that the constructed plasmids were all transfected to HEK 293-T cells successfully.

INTRODUCTION

Toxoplasma gondii (*T. gondii*) is an opportunistic intracellular protozoan parasite that can cause human toxoplasmosis (Dubey, 2008; Dubey & Su, 2009). Generally in benign immunocompetent subjects, human toxoplasmosis can lead to serious complications and can be fatal in some cases, clinical manifestations in the developing foetus and in different categories of immunocompromised individuals, including those undergoing chemotherapeutic or immunosuppressive treatments, and acquired immune deficiency syndrome (AIDS) (Montoya & Liesenfeld, 2004; Robert-Gangneux & Darde, 2012). Up to now, there

are no ideal drug available for prevention and treatment of this disease. The study of DNA vaccine against the parasite is particularly important. The surface antigen of *T. gondii* plays decisive roles in contacting with host cell and resistance of host immunity. A large number of studies suggested that SAG1 was the main member of surface antigen and had the best immunogenicity (Nielsen *et al.*, 1999; Siachoque *et al.*, 2006; Liu *et al.*, 2006). SAG5, SAG1-like sequence branch, contains five subtypes (SAG5A-5E). SAG5A protein is not expressed in RH strain tachyzoites (Tinti *et al.*, 2003), while SAG5E is a transcribed pseudogene (Elsheikha & Zhao, 2008; Elsheikha *et al.*, 2008). The remaining subtypes all can express corresponding

proteins in tachyzoites. Previous studies have demonstrated that SAG5 gene cluster plays an important role in tachyzoites and bradyzoites (Furio *et al.*, 2002). All the proteins' structures are still not very clear.

In this study, bioinformatics approaches were used to analyze the structure and function domains of SAG5A, -5B, and -5D. The protein sequence alignments were analyzed by the T-Coffee method. Secondary structural and functional domains were predicted with the PSIPRED v3.0 and SMART software. All the proteins' 3D structure models were mapped by the I-TASSER server. In addition, the epitopes of SAG5A, -5B, and -5D were analyzed and compared. The results of which identified a large number of linear-B cell epitopes and potential Th-cell epitopes on these proteins. This suggested the possibility that SAG5A, -5B, and -5D could be used as vaccines. The result suggested that SAG5 proteins had more potential to become an effective vaccine than SAG1. Furthermore, the four genes were cloned and connected with eukaryotic expression vector pEGFP-C1. After sequencing, the recombinant plasmids were transfected into HEK 293-T cells and identified protein expression.

MATERIALS AND METHODS

Data Resources

All the protein sequences were derived from ToxoDB 10.0 (<http://toxodb.org/toxo/>). *T. gondii* has three common types: type I, *T. gondii* GT1; type II, *T. gondii* ME49; type III, *T. gondii* VEG.

Modular Architecture Identification

T-Coffee (<http://www.tcoffee.org/>) (He *et al.*, 2002; Crawford *et al.*, 2009) was used to analyze the alignment among SAG5A, -5B, -5C, and -5D. The secondary structures were predicted using the software DNASTAR_Lasergene.v7.1 (Madison, WI, USA) by Garnier-Robson method (Zhao *et al.*, 2013).

The 3D models of proteins were constructed by SWISS-MODEL (Guex & Peitsch, 1997; Guex *et al.*, 2009), a protein structure server on the website [http://](http://swissmodel.expasy.org/)

swissmodel.expasy.org/, which is considered to predict protein 3D structures that have lots of amino acids. VMD is a molecular visualization software for displaying, animating, and analyzing large biomolecular systems using 3D graphics. VMD software was used to read standard Protein Data Bank (PDB) files and display the corresponding structure (Humphrey *et al.*, 1996; Schwede *et al.*, 2003).

Prediction of protein secondary structure and linear-B cell epitopes

Epitopes of protein determine antigen specificity and are the foundation of protein antigenicity (Van Regenmortel, 2009; Gao *et al.*, 2012). Many antigen indexes were used to evaluate protein antigenicity, including accessibility, flexibility, hydrophilicity, secondary structure, antigenicity, and charge distribution (Kyte & Doolittle, 1982; Carter & Loomis-Price, 2004; Tong & Tammi, 2008). Although there is no perfect method to predict antigenic epitopes, there are several rules can be followed to analyze which fragments of a protein will be antigenic (Welling *et al.*, 1985). On the one hand, antigenic epitopes should be located in solvent-accessible sections which contain hydrophobic and hydrophilic residues (Gershoni *et al.*, 1997; Subramani & Floudas, 2012). On the other hand, the peptides with long loops connecting with secondary structure motifs will be selected preferably. Given the rules outlined above, linear-B cell epitopes of SAG1, SAG5A, -5B, and -5D were analyzed using DNASTAR software. And the peptides that had good antigenic index and surface probability were chosen. Furthermore, in order to search for linear-B cell epitopes on SAG1, SAG5A, -5B, and -5D amino acid sequences, DNAMAN software was used.

Prediction of Th-cell epitopes

Cellular immunity mediated with T cells plays a decisive role in *T. gondii* infection (El-Kady, 2011), because *T. gondii* is an intracellular parasite. To construct a positive vaccine against *T. gondii*, it is necessary to clarify the type of T cell-mediated immune response. The Immune Epitope Database (<http://tools.immuneepitope.org/analyze/>

html/mhc_II_binding.html) online service was used to predict the half maximal inhibitory concentration (IC50) values of peptides that bind to the major histocompatibility complex (MHC) class II molecules of SAG1, SAG5A, -5B, and -5D.

***T. gondii* strain**

The RH strain of *T. gondii* was maintained in our laboratory. The tachyzoites of *T. gondii* used in this study were harvested from human foreskin fibroblast cells. The tachyzoites were washed by centrifugation, suspended in sterile phosphate-buffered saline (PBS).

Construction of plasmids

The entire SAG5A, -5B, -5C, and -5D open reading frames (ORF) were amplified by PCR from the cDNA of *T. gondii* tachyzoites with designed primers (Table 1). The forward primers contain Kpn I restriction sites and the reverse primers contain BamH I restriction sites. *Trans Tag*TM High Fidelity DNA Polymerase (TransGen Biotech, Beijing, China) was used in PCR amplification. The amplifications were all performed by the selected conditions: 1 cycle of 95°C for 5 min then 30 cycles of 95°C for 30 sec, 53°C for 30 sec, and 72°C for 30 sec. Final primer extension was extended to 5 min at 72°C. PCR products were tested with electrophoresis on 1.0% agarose gel.

The PCR products amplified from cDNA were inserted into pEASY-T1 vectors (TransGen Biotech, Beijing, China) to build recombinant cloning plasmids. After resequencing, SAG5A, -5B, -5C, and -5D were subcloned into eukaryotic expression plasmid pEGFP-C1 (Novagen, Billerica, MA, USA) to obtain pSAG5A, pSAG5B, pSAG5C and pSAG5D. Lastly, the new recombinant plasmids were transfected

into HEK 293-T cells by LipofectamineTM 2000 reagent (Invitrogen, Carlsbad, CA, USA).

Expression of recombinant plasmids in HEK 293-T cells

HEK 293-T cells were maintained in a humidified 5% CO₂ atmosphere at 37°C in 6-well plates (Sigma-Aldrich, St. Louis, MO, USA). Dulbecco's Modified Eagle Medium (DMEM) with streptomycin (100 mg/ml), penicillin (100 IU/ml) and 10% fetal Bovine serum (FBS) was added into the plates. When the density of HEK 293-T cells reached 80%-90%, the recombinant eukaryotic expression plasmids and the original vectors were transfected into cells with the Lipofectamine 2000 reagent (Invitrogen, Carlsbad, CA, USA) according to the manufacturer's guidance. Plasmids were mixed with lipofectamine 2000 reagent at a concentration of 10 µg/ml in DMEM without antibiotics and FBS. The mixtures were incubated at room temperature for 20 min before added into HEK 293-T cells drop by drop. The cells were incubated with the solutions for 6 hr in a humidified 5% CO₂ atmosphere at 37°C. Finally, fresh cell culture fluid was added and the 6-well plates were returned to the cell incubator for 48 hr incubation. The cells from different groups were respectively detected by fluorescence microscope under blue laser after incubation.

To examine the expression of SAG5 genes in HEK 293-T cells, total RNA was extracted from cells after 48 hr incubation by TransZol Up (TransGen, Beijing, China). The reverse transcription reaction was conducted using First Strand cDNA Synthesis Kit (Thermo Scientific, MA, USA) and PCR amplification reaction using *TransTag*TM HiFi PCR SuperMix (TransGen, Beijing, China) with the designed primers. Lastly, RT-PCR

Table 1. The primers of four SAG5 genes

Gene	Forward primer	Reverse primer
SAG5A	GGGGTACCATGGAACGTCAGACTTC	CGGGATCCTCAGTATGCACCGAAAG
SAG5B	GGGGTACCATGGAGCGTACGACTG	CGGGATCCTCAGTATGCACCGAAAG
SAG5C	GGGGTACCATGGAGCGTACGACTGC	CGGGATCCTCAGTATGCACCGAAAG
SAG5D	CGGGGTACCATGGTGCACGGTCTT	CGCGGATCCTCAATATGTGCCAAGA

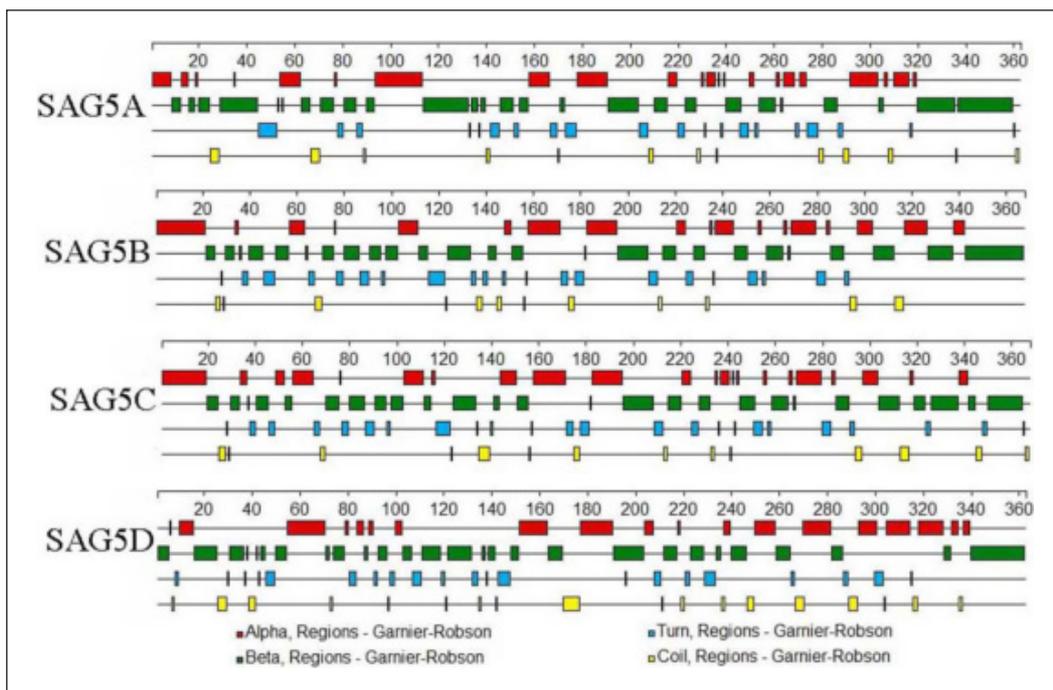


Figure 2. The 2D Structures of SAG5A, -5B, -5C, and -5D proteins. DNASTAR software was used to predict the secondary structure for SAG5A, -5B, -5C, and -5D proteins.

3D Model Constructed for SAG5A, -5B, -5C, and -5D Proteins

The online service, SWISS-MODEL, was used to predict the 3D structures of SAG5A, -5B, -5C, and -5D proteins. Three 3D models were constructed for each protein in corresponding workspace. The model with highest sequence identity was selected to further analyzed by VMD program. The chosen model also had the highest coverage in all templates offered with SWISS-MODEL. The selected templates for the four proteins were 2jks.1.A. SAG5A protein had 30.94% sequence identity with selected model. The accepted template covered 77% SAG5A protein from 55 to 333 amino acids. There was 28.07% consistent sequence between selected model and SAG5B protein. 78% SAG5B protein from 42 to 338 amino acids was covered with selected model. SAG5C protein with 27.18% sequence identity was covered 78% amino acids from 42 to 338 by the selected template. SAG5D protein had 33.33% sequence identity compared with selected model and was covered 74% amino acids from 65 to 333 by

the model. Table 2 showed the parameters discussed above for constructed model of each protein.

The optimized model of every protein was selected and analyzed with VMD software (Figure 3). SAG5A and SAG5B all had obvious three α -helixes, while there were two and one α -helixes in SAG5C and SAG5D respectively. All the four proteins obviously had two main domains which were formed by several β -strands. The β -strands usually form a sheet tube that is a common character of the surface antigen glycoproteins.

As shown in Figure 4, SAG5B and SAG5C fit well to each other on the entire sequence. The analysis result showed that SAG5B and SAG5C had a highly similarity in spatial structure. Besides, the turn regions of protein play a significant part in combining with downstream molecules. The distribution of turn regions on SAG5A, -5B, -5C and -5D were clearly showed in 3D models (Figure 5). The predicted result indicated that lots of turn regions irregularly distributed in the models surface.

Table 2. The parameters of four constructed models

Protein	Template	Seq identity(%)	Range	Coverage	Description
SAG5A	2jks.1.A	30.94	55-333	0.77	BSR4
SAG5B	2jks.1.A	28.07	42-338	0.78	BSR4
SAG5C	2jks.1.A	27.18	42-338	0.78	BSR4
SAG5D	2jks.1.A	33.33	65-333	0.74	BSR4

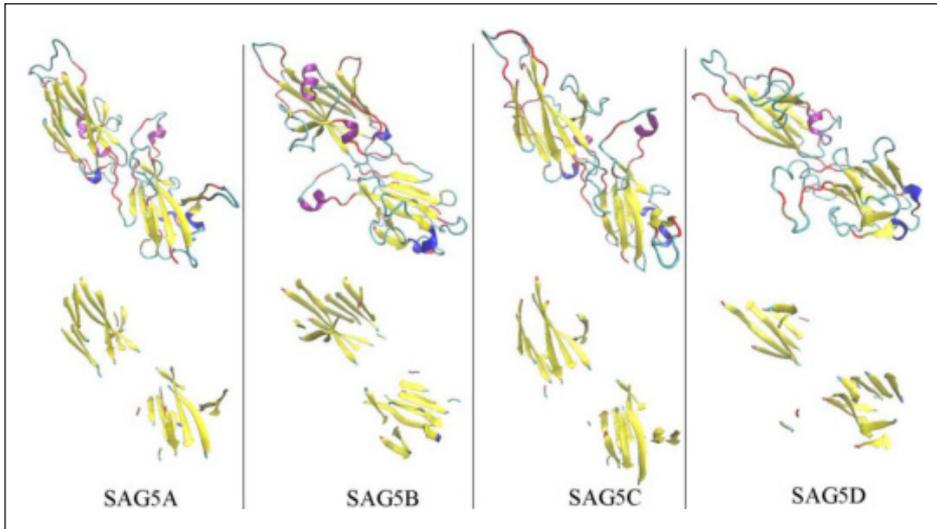


Figure 3. The 3D models of SAG5A, -5B, -5C, and -5D.

The sequences of proteins were sent to SWISS-MODEL from the website <http://swissmodel.expasy.org/>. The 3D models with the highest score for each protein were selected and viewed by VMD software. The color method was secondary structure (yellow: α -strands, purple: α -helix, red: coil, cyan: turn). The domain of each model was shown out in sheet form.

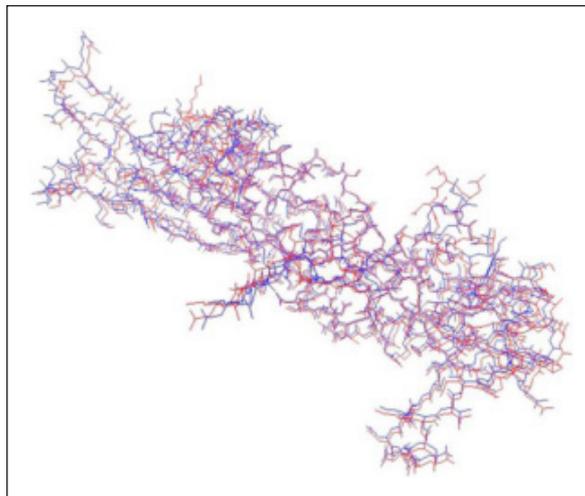


Figure 4. Fit analyses of SAG5B and SAG5C.

VMD software was used to show the fitness. SAG5B is in blue color; SAG5C is in red color.

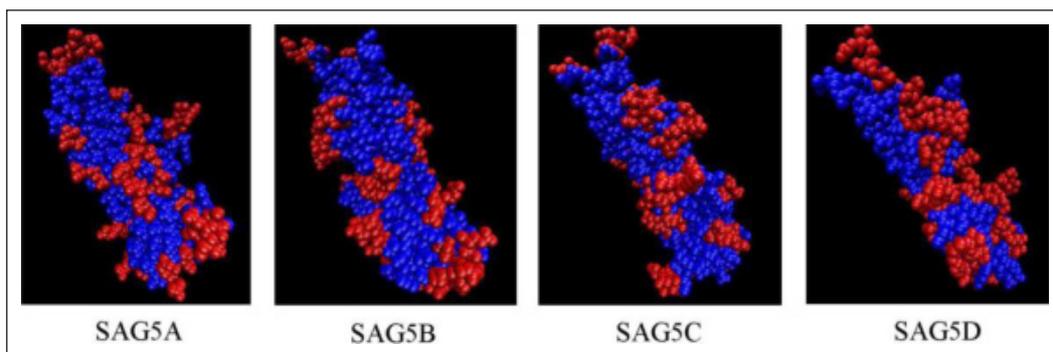


Figure 5. The distribution of turn regions on SAG5A, -5B, -5C and -5D 3D models. The turns are in red color; the remaining regions are in blue color.

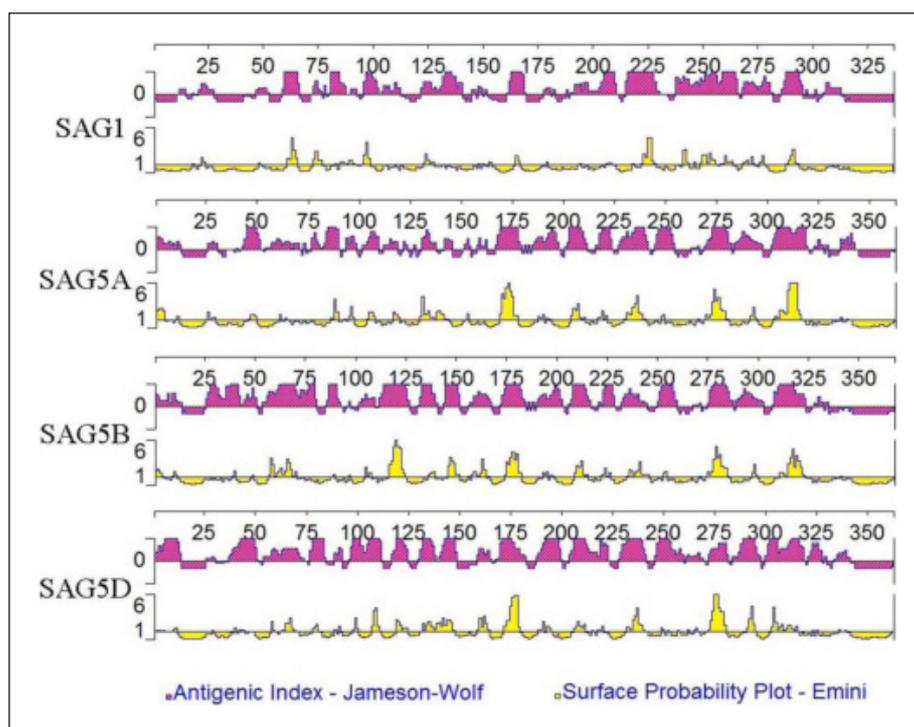


Figure 6. The prediction of linear-B cell epitopes. The linear-B cell epitopes of SAG1, SAG5A, -5B, and -5D predicted by DNASTAR in antigenic index and surface probability rules.

Prediction of linear-B cell epitopes

The linear-B cell epitopes of SAG1, SAG5A, -5B, and -5D were predicted by DNASTAR software (Figure 6). Given the highly similar structures of SAG5B and SAG5C, the linear-B cell epitope of SAG5C was not analyzed. As a good vaccine candidate protein, SAG1 has excellent antigenic index and surface

probability. In the linear-B cell epitopes analysis results, SAG1, SAG5A, -5B, and -5D all had good antigenic index. The prediction analyses indicated that antigenic index of SAG5A, -5B, and -5D was better than SAG1.

In addition, SAG5A, -5B, and -5D all had more significant surface probability than SAG1 in the result. Antigenic index was

used to predict the topological features of a protein directly from its primary amino acid sequence. The output of this algorithm, the antigenic index, is used to create a linear surface contour profile of the protein. Owing to most antigenic sites are located within surface exposed regions of a protein, the algorithm offers a reliable means of predicting potential antigenic determinants. Surface probability referred to the amount of reflection of an antigen's secondary and/or tertiary structure to the outside of the molecule. We used DNAMAN software to search for linear-B cell epitopes in SAG1, SAG5A, -5B, and -5D amino acid sequences. The results of the prediction analyses indicated the presence of 16 potential epitopes on SAG1, SAG5A, SAG5B and 21 on SAG5D. The epitopes which were rated score and got more than 1 point were selected to be potential epitopes.

Prediction of Th-cell epitopes

In order to analyze Th-cell epitopes, the Immune Epitope Database (http://tools.immuneepitope.org/analyze/html/mhc_II_binding.html) online service was used. The half maximal inhibitory concentration (IC50) values of peptides binding to the major histocompatibility complex (MHC) class II molecules of SAG1, SAG5A, -5B, and -5D were predicted with the online service. The Th-cell epitopes on SAG1, SAG5A, -5B, and -5D that were identified by bioinformatic analyses were predicted to have the ability to bind strongly to MHC class II molecules (Table 3). HLA-DRB1*01:01, H2-

Iab, H2-Iad, and H2-Ied four MHC II alleles were selected to analyze the Th-cell epitopes on SAG1, SAG5A, -5B, and -5D. As shown in Table 4, the minimum percentile ranks of each MHC II alleles on SAG1, SAG5A, -5B, and -5D were chosen and listed. The percentile ranks of HLA-DRB1*01:01 and H2-Iad on SAG5D were much lower than the other three proteins, while percentile ranks of the other two alleles were similar to each other. As shown in Table 4, the important epitopes were selected by the predictions of linear-B cell epitopes and Th-cell epitopes.

Identification of the recombinant plasmids

SAG5A, -5B, -5C, and -5D genes were cloned into the eukaryotic expression vectors pEGFP-C1 by suitable restriction enzymes to generate new recombinant plasmids. In order to ensure the accurateness of plasmids, they were detected by restriction enzyme analysis. The result of restriction enzyme digestion of plasmids was shown in Figure 7. Moreover, the constructed plasmids were all sequenced by biotechnology company (Shenggong, Shanghai, China) and had 100% identity to the corresponding sequences in GenBank™.

Expression of the recombinant plasmids in HEK 293-T cells

The recombinant and empty plasmids were successfully transfected into HEK 293-T cells. In order to identify the expression of the plasmids, the transfected cells were detected with fluorescence microscope. In

Table 3. IC50 values for SAG1; SAG5A, -5B, and -5D peptide binding to MHC class II molecules obtained using the immune epitope database

MHC II Allele ^a	Start-Stop ^b				Percentile Rank ^c			
	SAG1	SAG5A	SAG5B	SAG5D	SAG1	SAG5A	SAG5B	SAG5D
HLA-DRB1*01:01	12-26	339-353	8-22	348-362	0.88	0.42	0.09	0.01
H2-IAb	26-40	330-344	332-346	19-33	2.51	1.765	2.09	2.78
H2-IAd	21-35	8-22	330-344	8-22	0.34	0.48	0.735	0.24
H2-IEd	14-28	161-175	323-337	2-16	18.45	17.38	20.23	17.58

^aH2-IAb, H2-IAd and H2-IEd alleles are mouse MHC class II molecules; the HLA-DRB1*01:01 allele is a human MHC class II molecule.

^bWe chose 15 amino acids for analysis each time.

^cLow percentile = high binding.

Table 4. The important epitopes from the results of analysis

Protein	Amino acid position	The amino acid sequence ^a
SAG5A	178–192	DVECLVKVDVKARTS
	345–359	FLGLLAVVFTPVVSF
	12–26	LTAVGLVA AVL FNAC
SAG5B	183–197	DVECLVKVDVKARTS
	330–344	CRVKVTLTAQPAASH
SAG5C	8–22	SNKFRAAAGLLAAVL
	323–337	TTATLP TCRVKVTLT
SAG5D	12–26	RAAVGLVAGVLLFAG
	327–341	VKVTL SAHSSASQAS

^a We chose the amino acids sequences according to the analyzed results of linear-B cell epitopes and Th-cell epitopes.

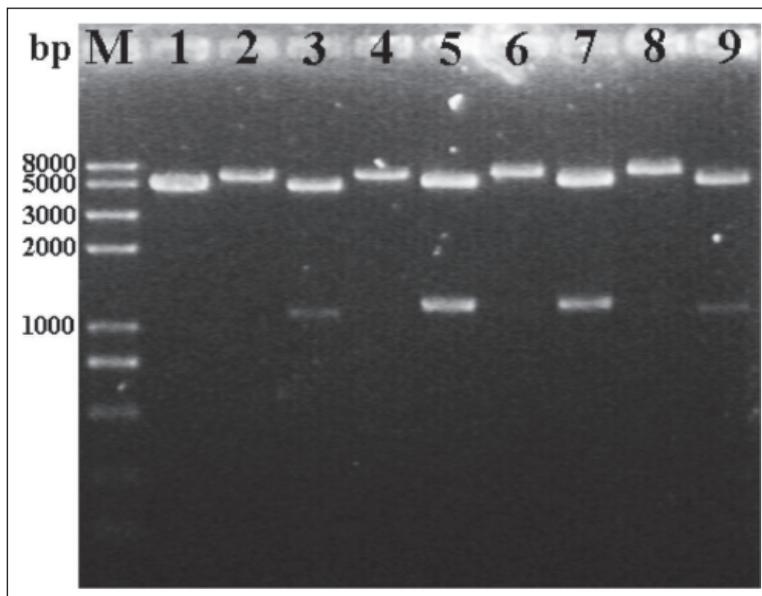


Figure 7. Constructed plasmid maps and identification of the recombinant plasmids with restriction enzyme digestion.

DNA Mark (lane M); pEGFP-C1 digested with BamH I (lane 1); pSAG5A digested with BamH I (lane 2), pSAG5A digested with Kpn I and BamH I (lane 3); pSAG5B digested with BamH I (lane 4), pSAG5B digested with Kpn I and BamH I (lane 5); pSAG5C digested with BamH I (lane 6), pSAG5C digested with Kpn I and BamH I (lane 7); pSAG5D digested with BamH I (lane 8), pSAG5D digested with Kpn I and BamH I (lane 9).

plasmids transfected cells, the proteins were excited green fluorescence under blue laser using fluorescence microscope (Figure 8B-F), whereas no fluorescence was observed in control cells (Figure 8A). To further confirm the expression of SAG5A, -5B, -5C

and -5D genes, RT-PCR was performed with the total RNA of HEK 293-T cells transfected by recombinant plasmids. The result of agarose gel electrophoresis was shown in Figure 9. Bands about 1089bp, 1104bp, 1104bp, and 1089bp were found in cells

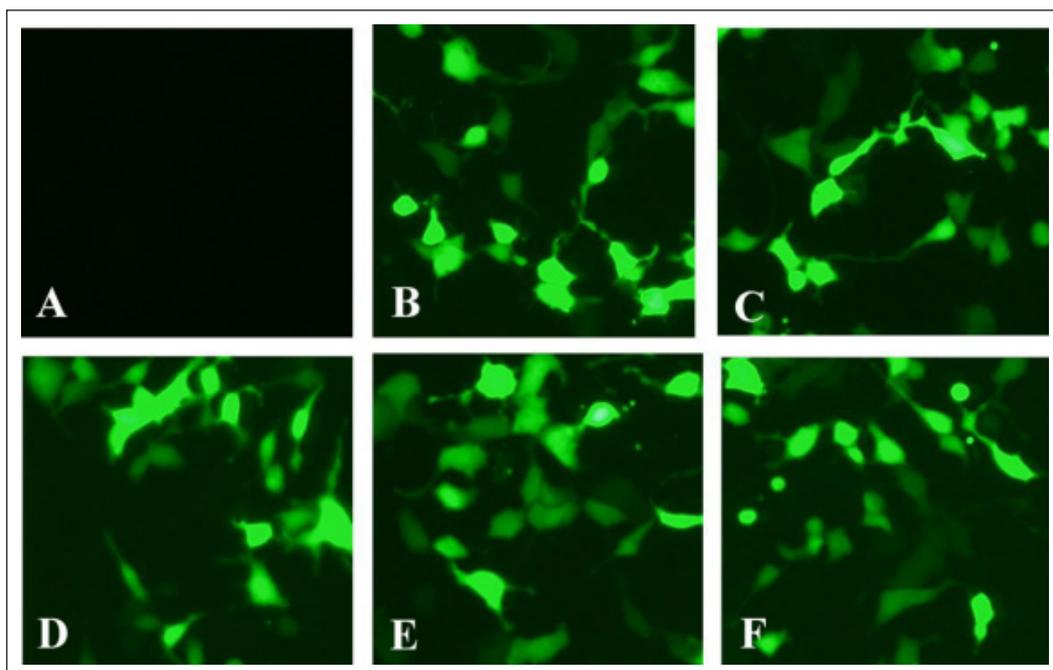


Figure 8. Green fluorescence detection of the fusion protein in transfected HEK 293-T cells. (A) Untransfected cells under blue light. (B) Cells transfected with pEGFP-C1 detected under blue light. (C) Cells transfected with pSAG5A detected under blue light. (D) Cells transfected with pSAG5B detected under blue light. (E) Cells transfected with pSAG5C detected under blue light. (F) Cells transfected with pSAG5D detected under blue light.

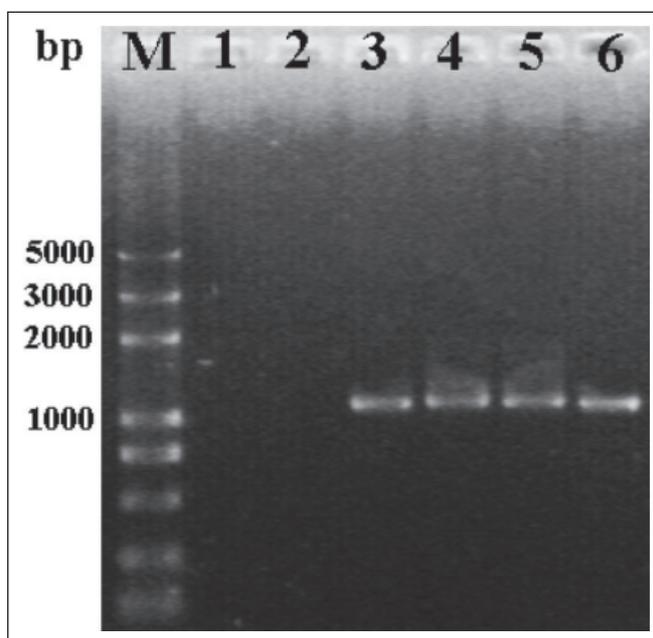


Figure 9. Gene transcription analysis of the recombinant plasmids in HEK 293-T cells by RT-PCR. DNA Mark (lane M), Control cell (lane 1), pEGFP-C1 (lane 2), pSAG5A (lane 3), pSAG5B(lane 4), pSAG5C (lane 5), pSAG5D (lane 6).

transfected with pSAG5A, pSAG5B, pSAG5C, and SAG5D respectively, while no band was found in control cells.

DISCUSSION

In previous study, SAG5D shared 50% amino acid identity with both SAG5A and SAG5C, while SAG5B and SAG5C were 97.5% identical to each other (Elsheikha & Zhao, 2008). It indicated that the difference of amino acid sequence between SAG5D and others did not influence their highly homologous. There was a long β region on the C-terminal of all the proteins. SAG5B and SAG5C had completely similar secondary structures but a small number of differences. All the proteins had the similar number of α regions, β regions, and turn regions, while SAG5D had the more coil regions than others. The secondary structures analysis revealed that a great difference existed between SAG5D and others. The difference in structure usually suggests the difference in function. SAG5D may have role different from others in the membrane of *T. gondii*. The analysis of SAG5A, -5B, -5C, and -5D structures indicated that the 3D space fold of SAG5 was most similar to the *T. gondii* bradyzoite-expressed BSR4 (Crawford *et al.*, 2009). BSR4, a prototypical bradyzoite surface antigen, also is expressed in tachyzoite. In previous study, SAG5A, -5B, -5C, and -5D shared high sequence identity with BSR4. The prediction indicated that the structure might play a decisive part in the process of infection and pathogenesis in *T. gondii*.

In amino acid sequence alignment analysis, SAG5C had a high sequence identity to SAG5B. Their fitness was analyzed with VMD software. Some structure studies suggested that the turn regions, convex structure in the surface of protein, were beneficial to bind with antibody and more likely to become epitopes (Zhang *et al.*, 2015).

The potential epitopes on SAG5D were much more than SAG1, SAG5A, and SAG5B from the analysis result. The result of linear-B cell epitopes analysis showed that

SAG5A, -5B, and -5D were excellent vaccine candidate proteins, especially SAG5D. A lower number indicates higher affinity. And higher affinity indicates better Th-cell epitopes. So SAG5D protein with small numbered percentile ranks had great potential to become a good Th-cell antigen. Furthermore, the binding strength of the interaction is known to influence the direction of Th cell differentiation. As the binding force increases, more cells tend to differentiate into Th-1 cells (Constant *et al.*, 1995; Carter & Loomis-Price, 2004). As such, we speculated that SAG5 proteins, especially SAG5D, were likely to induce Th-1 cell-mediated immune responses. The peptides that have good hydrophilicity, satisfactory flexibility, high accessibility, and strong antigenicity can easily interact with antibodies and generally act as epitopes (Jameson & Wolf, 1988; Wang *et al.*, 2014). The potential epitopes are usually distributed in the turn regions and convex structure, but not the α -helix and β -folded sheets (Zhang *et al.*, 2015). Attempts to develop a peptide-based vaccine for *T. gondii* have been encouraging because they have demonstrated significant protection in murine models (Darcy *et al.*, 1992; Godard *et al.*, 1994; Cong *et al.*, 2008; Wang *et al.*, 2011).

Since SAG5 with outstanding epitopes are potential antigen proteins of strong antigenicity, they may induce a strong humoral and cellular immune response and may be DNA vaccine candidates against *T. gondii* affection. So we conducted a study to clone SAG5A, -5B, -5C, and -5D genes and construct eukaryotic expression plasmids. As expected, SAG5A, -5B, -5C, and -5D genes were detected by agarose gel electrophoresis after PCR and the four plasmids (pSAG5A, pSAG5B, pSAG5C, and pSAG5D) were constructed successfully. The HEK 293-T cells transfected with plasmids showed green fluorescence, which indicated that the recombinant vectors could express proteins successfully. Genes whose size were corresponding to the target genes were transcribed in cells that transfected with recombinant plasmids in RT-PCR. The results showed that SAG5A, -5B, -5C, and -5D genes could effectively transcribed in HEK 293-T

cells, which suggested that the four recombinant eukaryotic expression vectors were constructed successfully.

CONCLUSIONS

In the present study, the bioinformatic analysis fully indicated that SAG5A, -5B, -5C, and -5D are homologous protein members of the SAG1 subfamily. Besides, the fitness analysis showed that SAG5B was similar to SAG5C. The structural prediction demonstrated that SAG5A, -5B, -5C, and -5D were all similar to BSR4 and SAG1. The result further proved that SAG5 proteins had high homology with SAG1. Moreover, the results from antigenic analysis indicated that SAG5 proteins, especially SAG5D with good liner-B cell epitopes and Th-cell epitopes, have great potentiality to become excellent vaccines against *T. gondii*. Furthermore, the bioinformatics analysis of SAG5 proteins could provide some valuable information on study of this parasite. In addition, the eukaryotic expression vectors that could transcribed in cells had been constructed, which provided the basic work for the production of the SAG5 gene vaccines against *T. gondii* infection.

List of Abbreviations

<i>T. Gondii</i>	<i>Toxoplasma gondii</i>
SAG1	<i>Toxoplasma gondii</i> surface antigen 1
SAG5	<i>Toxoplasma gondii</i> surface antigen 5
IC50	Half maximal inhibitory concentration
MHC	Major histocompatibility complex

Conflict of Interests

The authors wish to declare that there is no known conflict of interests associated with this publication.

Acknowledgments. This work was supported, in part, by grants from the National Natural Science Foundation of China (Grant Nos. 81071373 and 81271857), the State Key Laboratory of Veterinary Etiological Biology

(Grant No. SKLVEB2011KFKT005) and the Shandong Provincial Natural Science Foundation (Grant No. ZR2009CM079).

REFERENCES

- Carruthers, V.B. & Sibley, L.D. (1997). Sequential protein secretion from three distinct organelles of *Toxoplasma gondii* accompanies invasion of human fibroblasts. *European Journal of Cell Biology*, **73**: 114-123.
- Carter, J.M. & Loomis-Price, L. (2004). B cell epitope mapping using synthetic peptides. *Current Protoc Immunol Chapter 9: Unit 9.4*, doi: 10.1002/0471142735.
- Chaturvedi, P., Yu, Q., Southwood, S., Sette, A. & Singh, B. (1996). Peptide analogs with different affinities for MHC alter the cytokine profile of T helper cells. *International Immunology*, **8**: 745-755.
- Cong, H., Gu, Q.M., Yin, H.E., Wang, J.W., Zhao, Q.L., Zhou, H.Y., Li, Y. & Zhang, J.Q. (2008). Multi-epitope DNA vaccine linked to the A2/B subunit of cholera toxin protect mice against *Toxoplasma gondii*. *Vaccine*, **26**: 3913-3921.
- Constant, S., Pfeiffer, C., Woodard, A., Pasqualini, T. & Bottomly, K. (1995). Extent of T cell receptor ligation can determine the functional differentiation of naïve CD4+ T cells. *The Journal of Experimental Medicine*, **182**: 1591-1596.
- Crawford, J., Grujic, O., Bruic, E., Czjzek, M., Grigg, M.E. & Boulanger, M.J. (2009). Structural characterization of the bradyzoite surface antigen (BSR4) from *Toxoplasma gondii*, a unique addition to the surface antigen glycoprotein 1-related superfamily. *Journal of Biological Chemistry*, **14**: 9192-9198.
- Dubey, J.P. & Su, C. (2009). Population biology of *Toxoplasma gondii*: what's out and where did they come from. *Memórias do Instituto Oswaldo Cruz*, **104**: 190-195.
- Dubey, J.P. (2008). The history of *Toxoplasma gondii* – the first 100 years. *Journal of Eukaryotic Microbiology*, **55**: 467-475.

- El-Kady, I.M. (2011). T-cell immunity in human chronic toxoplasmosis. *Journal of the Egyptian Society of Parasitology*, **41**: 17-28.
- Elsheikha, H.M. & Zhao, X. (2008). Patterns and role of diversifying selection in the evolution of *Toxoplasma gondii* SAG5 locus. *Parasitology Research*, **103**: 201-207.
- Elsheikha, H., Hafez, A.O., Zaalouk, T.K., Shazly, A.M., Khalil, H.H. & Morsy, T.A. (2008). Phylogenetic evidence for recombination in SAG5 locus in *Toxoplasma gondii*. *Journal of the Egyptian Society of Parasitology*, **38**: 371-384.
- Furio, S., Irene, R., Manlio, D.C., Possenti, A., Tinti, M., Dendouga, N., Tomavo, S. & Crisanti, A. (2002). The SAG5 locus of *Toxoplasma gondii* encodes three novel proteins belonging to the SAG1 family of surface antigens. *International Journal for Parasitology*, **32**: 121-131.
- Gao, J., Faraggi, E., Zhou, Y., Ruan, J. & Kurgan, L. (2012). BEST: improved prediction of B-cell epitopes from antigen sequences. *PLoS One*, **7**: e40104.
- Gershoni, J.M., Stern, B. & Denisova, G. (1997). Combinatorial libraries, epitope structure and the prediction of protein conformations. *Immunology Today*, **18**: 108-110.
- Godard, I., Estaquier, J., Zenner, L., Bossus, M., Auriault, C., Darcy, F., Gras-Masse, H. & Capron, A. (1994). Antigenicity and immunogenicity of P30-derived peptides in experimental models of toxoplasmosis. *Molecular Immunology*, **31**: 1353-1363.
- Guex, N. & Peitsch, M.C. (1997). SWISS-MODEL and the Swiss-PdbViewer: an environment for comparative protein modeling. *Electrophoresis*, **15**: 2714-2723.
- Guex, N., Peitsch, M.C. & Schwede, T. (2009). Automated comparative protein structure modeling with SWISS-MODEL and Swiss-PdbViewer: a historical perspective. *Electrophoresis*, **30**: S162-S173.
- He, X.L., Grigg, M.E., Boothroyd, J.C. & Garcia, K.C. (2002). Structure of the immunodominant surface antigen from the *Toxoplasma gondii* SRS superfamily. *Nature Structural Biology*, **9**: 606-611.
- Humphrey, W., Dalke, A. & Schulten, K. (1996). VMD: visual molecular dynamics. *Journal of Molecular Graphics*, **14**: 33-38.
- Jameson, B.A. & Wolf, H. (1988). The antigenic index: a novel algorithm for predicting antigenic determinants. *Computer Applications in the Biosciences*, **4**: 181-186.
- Kyte, J. & Doolittle, R.F. (1982). A simple method for displaying the hydrophobic character of a protein. *Journal of Molecular Biology*, **157**: 105-132.
- Lekutis, C., Ferguson, D.J.P., Grigg, M.E., Camps, M. & Boothroyd, J.C. (2001). Surface antigens of *Toxoplasma gondii*: variations on a theme. *International Journal for Parasitology*, **31**: 1285-92.
- Liu, K.Y., Zhang, D.B., Wei, Q.K., Li, J., Li, G.P. & Yu, J.Z. (2006). Biological role of surface *Toxoplasma gondii* antigen in development of vaccine. *World J Gastroenterol*, **12**: 2363-2368.
- Montoya, J.G. & Liesenfeld, O. (2004). Toxoplasmosis. *Lancet*, **363**: 1965-1976.
- Nielsen, H.V., Lauemoller, S.L., Christiansen, L., Buus, S., Fomsgaard, A. & Petersen, E. (1999). Complete protection against lethal *Toxoplasma gondii* infection in mice immunized with a plasmid encoding the SAG1 gene. *Infection and Immunity*, **67**: 6358-6363.
- Robert-Gangneux, F. & Darde, M.L. (2012). Epidemiology of and diagnostic strategies for toxoplasmosis. *Clinical Microbiology Reviews*, **25**: 264-296.
- Schwede, T., Kopp, J., Guex, N. & Peitsch, M.C. (2003). SWISS-MODEL: an automated protein homology-modeling server. *Nucleic Acids Research*, **13**: 3381-3385.
- Siachoque, H., Guzman, F., Burgos, J., Patarroyo, M.E. & Gomez Marin, J.E. (2006). *Toxoplasma gondii*: immunogenicity and protection by P30 peptides in a murine model. *Experimental Parasitology*, **114**: 62-65.

- Subramani, A. & Floudas, C.A. (2012). Structure prediction of loops with fixed and flexible stems. *The Journal of Physical Chemistry B*, **116**: 6670-6682.
- Tinti, M., Possenti, A., Cherchi, S., Barca, S. & Spano, F. (2003). Analysis of the SAG5 locus reveals a distinct genomic organisation in virulent and avirulent strains of *Toxoplasma gondii*. *International Journal for Parasitology*, **33**: 1605-1616.
- Tong, J.C. & Tammi, M.T. (2008). Prediction of protein allergenicity using local description of amino acid sequence. *Frontiers in Bioscience-Landmark*, **13**: 6072-6087.
- Van Regenmortel, M.H. (2009). What is a B-cell epitope?. *Methods in Molecular Biology*, **524**: 3-20.
- Wang, Y.H., Wang, G.X., Ou, J.T., Yin, H. & Zhang, D.L. (2014). Analyzing and identifying novel B cell epitopes within *Toxoplasma gondii* GRA4. *Parasites & Vectors*, **7**: 474.
- Wang, Y.H., Wang, M., Wang, G.X., Pang, A.N., Fu, B.Q., Yin, H. & Zhang, D.L. (2011). Increased survival time in mice vaccinated with a branched lysine multiple antigenic peptide containing B- and T-cell epitopes from *T. gondii* antigens. *Vaccine*, **29**: 8619-8623.
- Welling, G.W., Weijer, W.J., van der Zee, R. & Welling-Wester, S. (1985). Prediction of sequential antigenic regions in proteins. *FEBS Letters*, **188**: 215-218.
- Zhang, T.E., Yin, L.T., Li, R.H., Wang, H.L., Meng, X.L. & Yin, G.R. (2015). Protective immunity induced by peptides of AMA1, RON2 and RON4 containing T-and B-cell epitopes via an intranasal route against toxoplasmosis in mice. *Parasites & Vectors*, **8**: 15.
- Zhao, G.H., Zhou, A.H., Lv G., Meng, M., Sun, M., Bai, Y., Han, Y.L., Wang, L., Zhou, H.Y., Cong, H., Zhao, Q.L., Zhu, X.Q. & He, S.Y. (2013). *Toxoplasma gondii* cathepsin proteases are undeveloped prominent vaccine antigens against toxoplasmosis. *BMC Infectious Diseases*, **13**: 207.