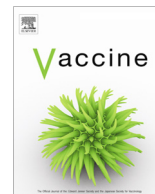




Contents lists available at ScienceDirect

## Vaccine

journal homepage: [www.elsevier.com/locate/vaccine](http://www.elsevier.com/locate/vaccine)

# Examination of antigenic proteins of *Trypanosoma cruzi* to fabricate an epitope-based subunit vaccine by exploiting epitope mapping mechanism

Nazia Khatoon<sup>a</sup>, Rupal Ojha<sup>a</sup>, Amit Mishra<sup>b</sup>, Vijay Kumar Prajapati<sup>a,\*</sup>

<sup>a</sup> Department of Biochemistry, School of Life Sciences, Central University of Rajasthan, Bandarsindri, Kishangarh, Ajmer, Rajasthan 305817 India

<sup>b</sup> Cellular and Molecular Neurobiology Unit, Indian Institute of Technology Jodhpur, Jodhpur, Rajasthan 342011 India

## ARTICLE INFO

## Article history:

Received 4 May 2018

Received in revised form 30 July 2018

Accepted 1 September 2018

Available online xxxx

## Keywords:

Trypanosomiasis

Subunit vaccine

Immunoinformatics

Immune receptor

Molecular dynamics

## ABSTRACT

Chagas disease is a protozoan parasitic disease caused by *Trypanosoma cruzi*. This injurious disease spread by the circulation of the blood sucking triatomine insects and transmitted to humans. Chagas disease is endemic in Latin America and also known as American trypanosomiasis. Currently, 7 million people are infected by *T. cruzi* infection and about 22,000 death cases were reported per year throughout the Americas. Various immunization approaches against *T. cruzi* infection have been examined and some of the developed vaccine showed efficacy in animal models but there is no effective and safe vaccines for humans have been developed yet. Since, the drug resistance is increasing day by day because the developed drug (nifurtimox and benznidazole) to control *T. cruzi* infection, failed to activate a prodrug and still no drug and vaccine have been established. To control the infection of chagas disease, here in this study we use immunoinformatics method to design multi-epitope subunit vaccine against *T. cruzi* infection. Moreover, on the basis of immunogenicity B and T cell epitopes were evaluated. The allergenicity, antigenicity was predicted to ensure the safety of vaccine constructs whereas, the physiochemical property showing the stable nature of final vaccine model. Further, molecular docking was performed to optimize the interaction between TLR-2 and TLR-4 (receptor) and vaccine model (ligand) complex. Molecular dynamics simulation was performed to evaluate the energy minimization; RMSD and RMSF plot which confirm the stability of TLR-2 and TLR-4 (receptor) present on immune cells and vaccine model (ligand) complex. This study needed the experimental validation for the safety and immunogenic behavior of designed vaccine protein and it may be helpful in future to control *T. cruzi* infection.

© 2018 Elsevier Ltd. All rights reserved.

## 1. Introduction

The Chagas disease is one of the injurious, parasitic infections caused by the protozoan *Trypanosoma cruzi* [1]. Trypanosomes are unicellular parasitic protozoa which infect different vertebrates, including animals as well as humans. Approximately, 60–70 million people are estimated to be at risk of this infection. The disease affects about 8 million people in Latin America, among them 30–40% individuals have or will develop cardiomyopathy, digestive mega syndromes, or both [2]. This is a contagious disease and can be transmitted through a number of ways i.e. congenital traits (pregnant women to her baby), blood transmission, and transplantation of organs as well as consumption of uncooked food that has been contaminated with feces from infected bugs. Several

upgrades have been made in the past three decades, regarding to control the administration of chagas diseases [3,4]. After, conducting several programs associated with the screening of blood donors, the illness occurrence and prevalence has diminished significantly. Benznidazole and nifurtimox are drugs that have been used to treat the disease. However, both drugs induce severe side effects. The requirement for new medications against tropical parasites for example, *Plasmodium falciparum* and *Trypanosoma brucei* is persistent since issues with protection and danger are risking the presently accessible pharmaceuticals [5]. About 10,000 cases reported per annum with hostile to trypanosomal drugs, especially against stage 2 disease where the central nervous system is involved, being vital to control. However, protection from melarso-prol, diminazene, and isometamidium (and in addition quinapyrimine) [6] are presently undermining the further utility of these compounds. From the limited accessible data, the resistance phenotypes appear to be stable there certainly is no evidence to the contrary [7]. Current treatment options are inadmissible for

\* Corresponding author at: Department of Biochemistry, Central University of Rajasthan, NH-8, Bandarsindri, Ajmer, Rajasthan 305817, India.

E-mail address: [vkprajapati@curaj.ac.in](mailto:vkprajapati@curaj.ac.in) (V.K. Prajapati).

different reasons, including unsuitable harmfulness, poor adequacy, and unwanted course of the organization, and drug resistant [8,9]. Improvement of new immunizations would now be able to exploit substantially more noteworthy information of the complexities of the resistant of the development of powerful molecular techniques that allow targeted changes to be made in pathogenic organisms and in experimental host [10,11]. Antibody improvement has turned out to be substantially more complex where immunologists working closely with sub-atomic researcher and molecular biologists to plan and deliver exceptionally refined immunizations that are safe and effective [12]. The improvement of an effective antibody to avoid *trypanosomiasis* illness condition has been an objective for very nearly a century, yet there is no such immunization exists. Secretory proteins of pathogens are antigenic in nature and are fit for actuating a noteworthy immunogenic reaction into the host [13]. Therefore, numerous immunization applicants have been derived from secretory proteins of the pathogen. B cell epitopes are a part of proteins which take action and recognized by the human antibodies. Identification of B and T cell epitopes are valuable for the improvement of a novel immunization [14]. The peptide parts of the pathogen and the MHC molecule bind together and show up on the cell surface of infected host cells. A differing scope of peptides is bound to each allelic type of MHC protein [15]. The MHC molecule can tie with peptides firmly as the pathogens attempt to change the MHC receptor. Hence, MHC particles have shown high restricting tendency with an assortment of peptides [16]. Vaccination have been considered as a most secure and useful approach to control the vector-borne parasitic diseases. The investigation of host-parasite co-operations has expanded extensively in the most recent decades, with many examinations concentrating on the recognizable proof of parasite particles (i.e. surface or excretory/secretory proteins (ESP) as potential targets for new specific treatments). Proteins involved in folding and degradation are the major class of secreted proteins. Secretory proteins are known to be able to control the immune system of mammalian hosts to stimulate macrophages or to act as intermediaries for intercellular signaling. Some of the secretory proteins had not yet been discovered and studied in *Trypanosomes*, and some may be new potential and therapeutic targets or diagnostic markers. *T. cruzi* expresses a vast variety of proteins on its surface membrane which involved in different biological processes. Surface membrane proteins were highly involved in adaptation, differentiation, and survival of the parasite during its life cycle [17]. Some membrane proteins families stick out during the host-parasite infection process which helps to prevent infection process and make them as potential targets. Thoroughly, these current advancements can increase the knowledge of host-parasite infection day by day and will help to accelerate the discovery of productive drugs against Chagas diseases. But still it is far from *trypanosomiasis* infection clinical trials. Immunoinformatics approach for the distinguishing resistant of novel antibody is a recent technique which can anticipate a thermodynamically steady and multi-epitope subunit immunization in a less time [18]. Henceforth, here we propose to design a multi-epitope subunit to demonstrate immunization made out of immunogenic epitopes which may be able to activate host humoral and cell-interceded immune system.

## 2. Results and discussion

### 2.1. Sequence retrieval

Overall, 7 secretory proteins of *T. cruzi* was collected from literature survey i.e. Mucin, Trans-sialidase, TcGP63 family, Amastin family, TcTASV family, MASP family and Cruzipain family.

Similarly, 6 membrane proteins were exported from TriTrypDB i.e. RNase PH-like protein, putative ribosomal RNA processing protein 41B, dispersed gene family protein 1, serine-alanine-and proline-rich protein, Mucin-associated surface protein and hypothetical protein, conserved. Next, to sort out them allergenicity and physiochemical property of all 13 proteins were predicted. After, the predictions of allergenicity and physiochemical properties only non-allergen and stable proteins were chosen and rest of the proteins were discarded. Hence, a total of 8 antigenic proteins of *T. cruzi* were selected for further study i.e. **TcCLB.510717** Trypomastigote, Alanine, Serine and Valine rich protein (TASV) [19], **TcCLB.511143.30** peroxidase, **TcCLB.510307.210** trans-sialidase, Group II, [20], **P25779.1** Cruzipain; Cruzaine [21]) (**TcCLB.505037.20** putative ribosomal RNA processing protein 41B, **TcCLB.511317.70** RNA-binding protein, **TcCLB.511487.150** serine-alanine-and proline-rich protein, **TcCLB.511487.170** Mucin-associated surface protein (MASP) [22].

### 2.2. Prediction of CTL and HTL epitopes

For the prediction of CTL epitopes of all 8 antigenic proteins of *T. cruzi* A2, A3 and B7 supertypes were chosen. A total number of 84 epitopes were obtained for A2 supertype, similarly 79 epitopes were predicted for A3 supertype and 86 epitopes have been found for B7 supertype. But, only those epitopes were selected which showing score above 0.75. Therefore, 22 epitopes were found to be as CTL epitopes with high score and hence, showing good binding affinity. Moreover, for the prediction of HTL epitopes we have chosen HLA-DRB1\*01:01 and DPA1\*01/DPB1\*04:01 [23] alleles. The diagrammatic view of HLA's with different colors has shown in Fig. 2(A), (B) and (c). Only low percentile rank epitopes were selected and rest of the epitopes was removed. Therefore, 16 epitopes were found on the basis of low percentile rank. Since, the lower percentile rank of epitopes provides good binding affinity. Predicted CTL and HTL epitopes has shown in Supplementary Table 1 and 2A & 2B.

### 2.3. Prediction of linear and discontinuous B cell epitopes

A total of 8 linear B cell epitopes were selected for vaccine design on the basis of epitope threshold of 0.5. Only, those epitopes were selected which shows score of 1 and rest of the epitopes were left for further analysis shown in Fig. 3(A). Similarly, discontinuous B cell epitopes were also predicted but, this prediction was done only to check which residues found at N-terminal in a folded protein. Therefore, 32 residues were found to be present at N-terminal i.e. V-392 to T-424 shown in Fig. 3(B). Since the conformational or discontinuous epitopes are less stable than linear epitopes. So, only linear B cell epitopes were chosen for vaccine construct. Predicted linear B cell epitopes has shown in Supplementary Table 3.

### 2.4. Scheming of multi-epitope subunit vaccine

All 22 high scored CTL epitopes and 16 lower percentile rank HTL epitopes and 8B-cell epitopes with score 1 were selected for vaccine designing. All epitopes of CTL, HTL and B-cell were linked with proper linkers. For 22 CTL epitopes AAY linker was used. Similarly, for 16 HTL epitope GPGPG linker was used and for 8B-cell epitopes KK linker was used. To make our vaccine protein more potentiate an antigen-specific adjuvant was also added i.e. RS09 (APPHALS) [24] which exaggerated immune response. On the basis of experiments RS09 proved as the capable adjuvant which increases the X-15 antibodies *in vivo* [25]. Hence, the final vaccine constructs were composed of total 645 amino acid residues. The flowchart of designing a multi-epitope subunit vaccine has shown in Fig. 1.

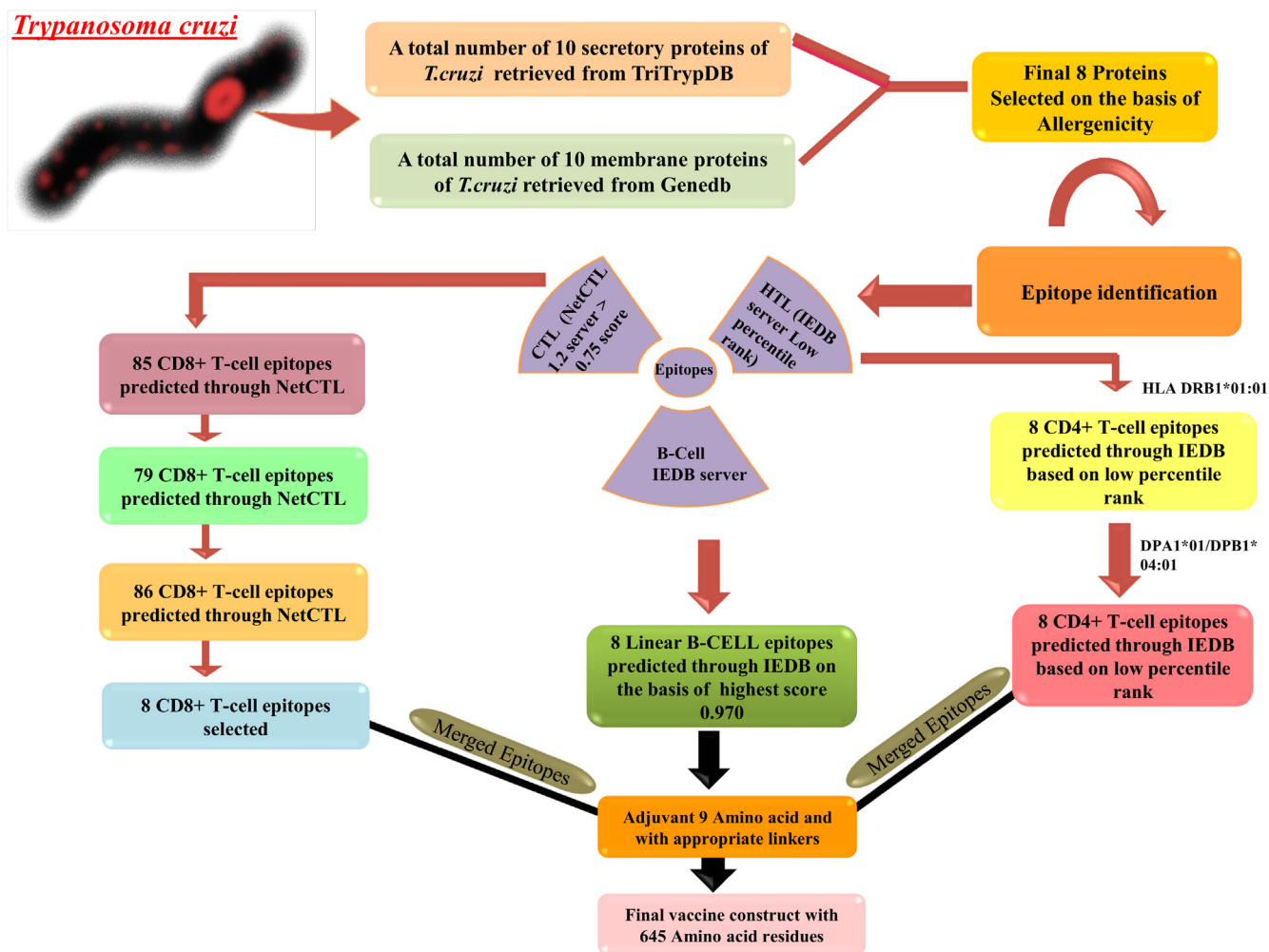


Fig. 1. Flowchart showing the collection of B, T cell epitopes and adjuvant to design vaccine construct.

### 2.5. Optimization of Interferon- $\gamma$ to develop protective immune response

Interferon- $\gamma$  epitopes were predicted, because it is important activators of macrophages and inducers of MHC class II molecules. A total number of 6 positively inducing IFN- $\gamma$  epitopes were found. The presence of INF- $\gamma$  inducing epitopes indicate the protective response of designed vaccine molecule. Predicted Interferon- $\gamma$  epitopes has shown in Supplementary Table 4.

## 3. Vaccine characteristics

### 3.1. Prediction of allergenicity, antigenicity and physiochemical properties

The predicted allergenicity score for final vaccine constructs was 1.25 whereas, the default threshold was 0.4. Hence, the final vaccine constructs are non-allergen in nature. Since, the predicted epitopes with positive value was 15.19% and the epitopes with negative value was 94.18%. Similarly, the antigenicity of final vaccine protein was also predicted. The average antigenic propensity of final vaccine protein was 1.0666 and 24 antigenic determinants epitopes was predicted. Further, physiochemical property of final vaccine construct was also evaluated the molecular weight of final vaccine construct was 68.1 KDa; the theoretical pI of final vaccine protein was 10.06 which shows basic nature of designed vaccine.

Moreover, the estimated half-life was 4.4 h (mammalian reticulocytes, in vitro) which is average half-life period for designing a vaccine, >20 h in yeast and >10 h in *E. coli*. The instability index was computed to be 36.45 which evaluate that our vaccine constructs was stable in nature. The aliphatic index of our vaccine protein was 99.09 which show that our vaccine constructs was thermally stable because higher the value of aliphatic index has more tendencies to be stable in nature. To evaluate the vaccine protein hydrophobicity GRAVY (Grand average of hydropathicity) was predicted. The optimized GRAVY showing the score of 0.576 that means protein was having hydrophobic amino acid in final vaccine protein because the average score of GRAVY is spanning from -0.9 to 1.4.

### 3.2. 3D Structure prediction of final vaccine constructs

The 3D structure of final vaccine construct was generated using RaptorX server. One domain was optimized by the server whereas the best PDB was used as template for homology modeling was 5gpjA. The p-value of modeled structure was 2.06e-05, overall uGDT (GDT) (global distance test); 71(11) which indicates that our protein was good modeled structure because >50 (GDT) is a good indicator. Among, 645 amino acids of residues 100% residues are modeled. Only 17% of residues were disordered in nature. The secondary structure of modeled vaccine protein was 62% H, 2% E and 34% of C, where H, E and C indicate helix, sheet and coil,

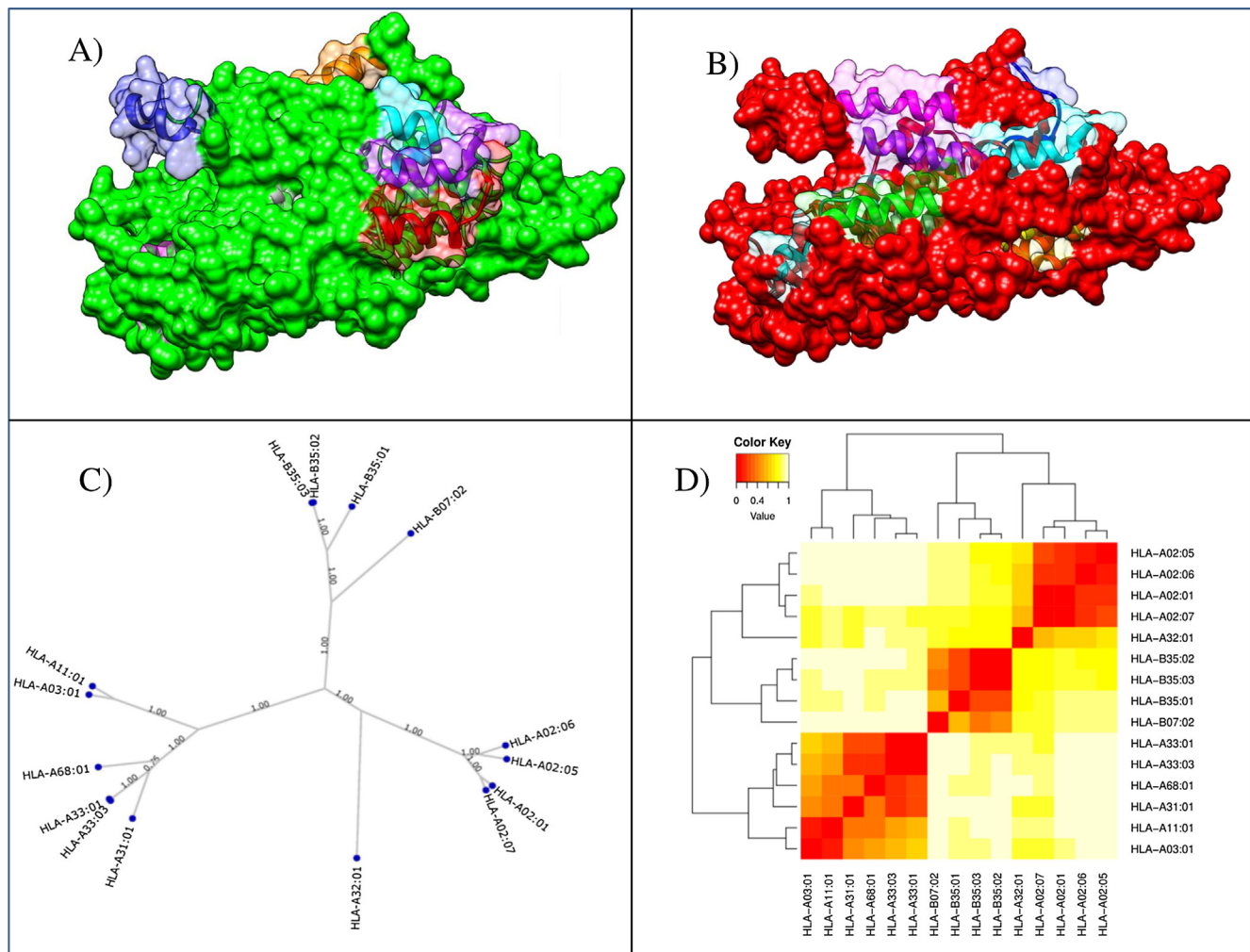


**Fig. 2.** Distribution of HTL epitopes over designed vaccine; (A) the green color showing DPA1/DPB1; (B) the red color shown HLA-DRB1; (C) showing phylogenetic tree of HLA distribution (D) showing heat map of HLA distribution. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

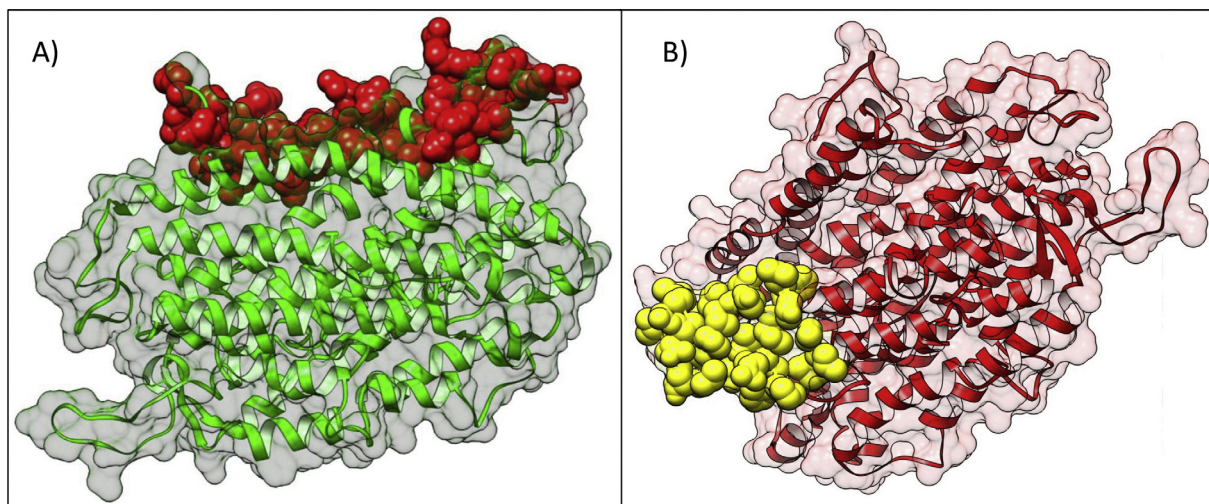


**Fig. 3.** B cell epitopes presence over designed vaccine; (A) the green color surface with highlighted red color sphere showing linear B cell epitopes (B) the red color surface with highlighted yellow color sphere showing discontinuous B cell epitopes. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

respectively. The homology modeling of final vaccine construct has shown in Fig. 4(A).

### 3.3. Validation and refinement of final vaccine constructs

Refinement of final vaccine construct was evaluated to refine the error estimated by the modeled structure. After, refinement of the structure shown as good quality of protein structure because the original modeled structure obtained by homology modeling have Ramachandran plot with 90% of residues in allowed region but after refinement it raised the quality of structure and provides Ramachandran plot with 92% of residues in allowed region. Further, validation of refined model structure was done because validation is one of the important steps after homology modeling; to validate the modeled vaccine protein Ramachandran plot assessment was done. The number of residues in favored region was 92.1%, and number of residues in allowed region was 37% whereas the number of residues in outliers region was predicted to be as 14%. The predicted allowed and disallowed region of modeled vaccine has shown in Fig. 4(C). Further, the overall quality of final vaccine construct and the deviation of the total energy of the modeled structure were evaluated on the basis of Z-score. If ProSA produces positive value, that means the structure is not good, if the Z-scores of the structure are above (−4) then it will be considered as

average structure. Hence, our final protein structure had Z-score value of (−5.44) average energy peaks which was showing that our vaccine construct was more in the range of required Z-score value shown in Fig. 4(B).

### 3.4. Molecular docking of final vaccine constructs (ligand) and immune cell receptor (TLR-2/TLR-4)

The molecular docking analysis was performed to understand the binding efficiency of the final vaccine construct and TLR-2/TLR-4. By, this we identify the most probable orientation of a vaccine protein. Here, we used two TLR's for molecular docking because both play an important role in the agonist development. Both are capable to identify different molecular patterns from invading pathogens and mediated host response against trypanosomiasis. For vaccine development we had chosen cell surface membrane proteins as well as secretory proteins of *T. cruzi* which was identified as effective activators of TLR-2 from both mouse and human origin [26]. Similarly, TLR-4 agonist [27] induces a pro-inflammatory response during trypanosomiasis. Moreover, this server used to evaluate docking study and it provides top ten models as an output, among these models only one model was chosen on the basis of different energy parameters for further study. For TLR-2, top 10 models were predicted, among them only model

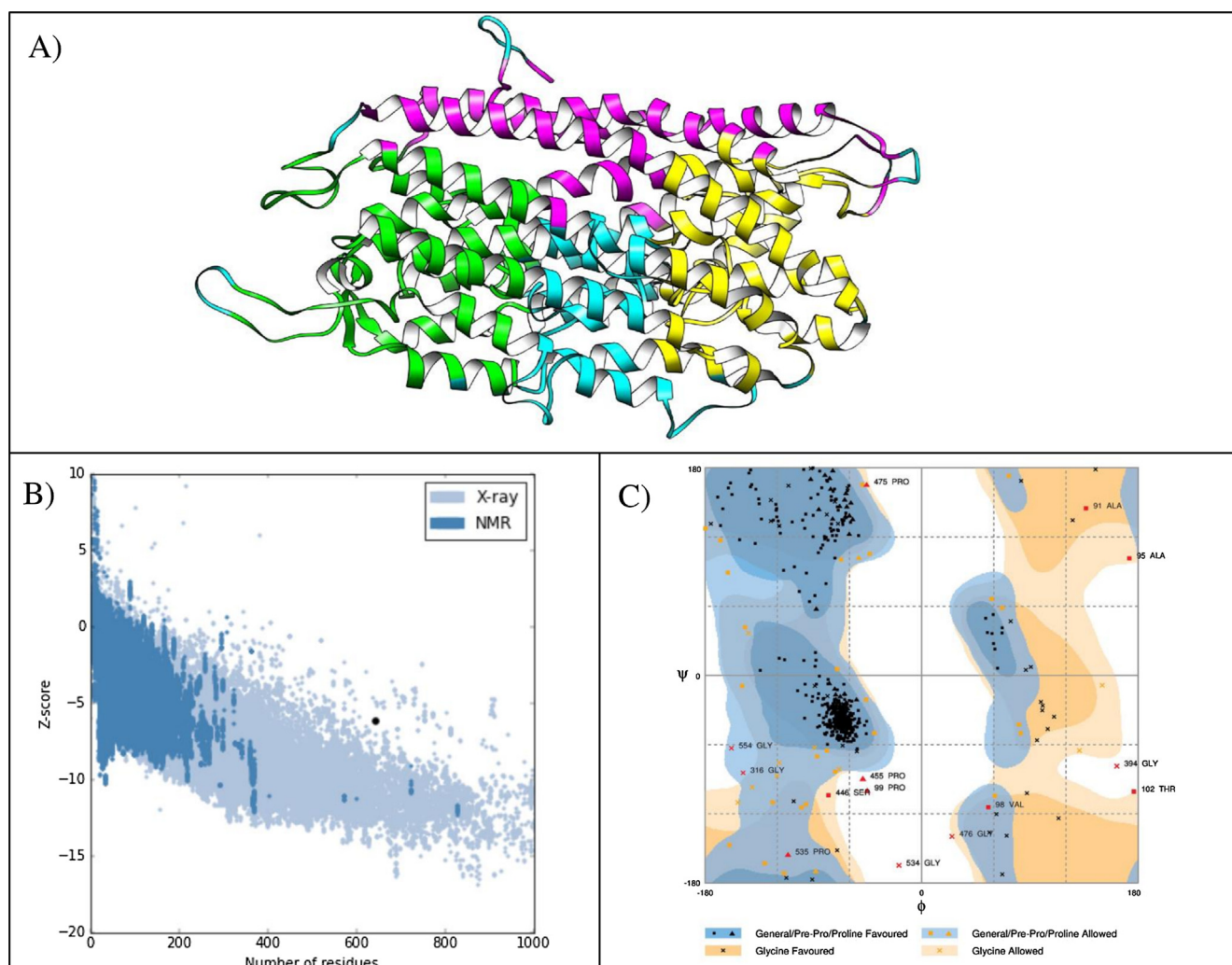


**Fig. 4.** Tertiary structure model development for final vaccine molecule and its validation (A) Showing homology modeling of final vaccine construct; (C) showing Z-score of final vaccine model; (D) showing Ramachandran plot of final vaccine model.

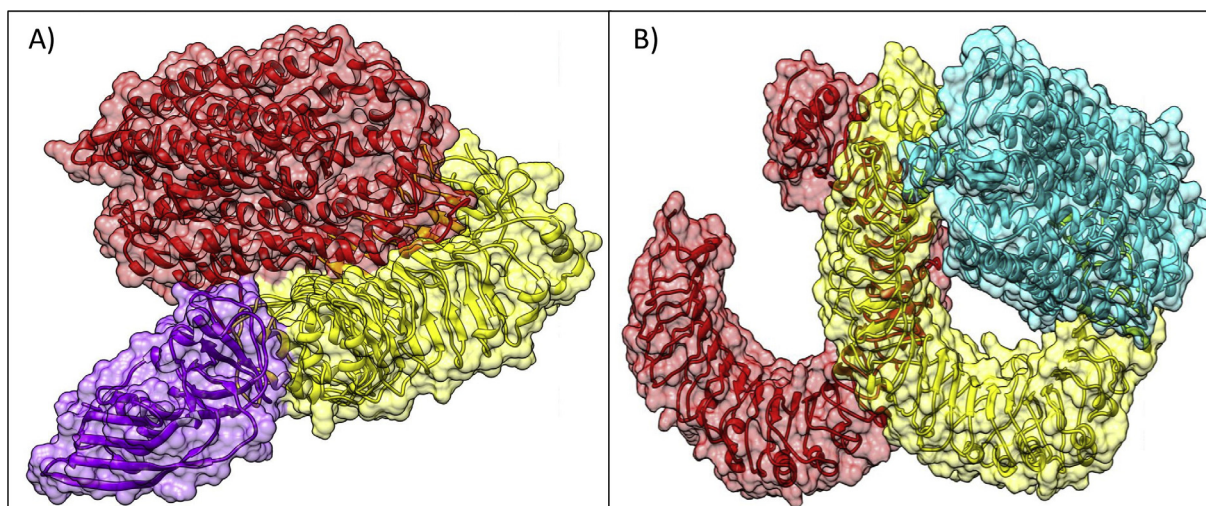


**Fig. 5.** Molecular docking of two different TLR's with final vaccine construct; (A) showing molecular docking with color purple and yellow i.e. TLR-2 (receptor) and final vaccine construct (ligand) red in color; (B) showing molecular docking with color red and yellow i.e. TLR-4 (receptor) and final vaccine construct (ligand) cyan in color. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

number 3 was chosen with lowest energy score of  $-1096.1$ . Further, for TLR-4, among top 10 models only model number 8 was chosen on the basis of lowest energy score i.e.  $-12.674$ . The interaction between immune receptor TLR-2, TLR-4 and vaccine protein has shown in Fig. 5(A) and (B).

### 3.5. Molecular dynamics simulation of final vaccine constructs

Molecular dynamics (MD) simulation of final vaccine constructs was performed to study the motions of protein structure as well as we can easily calculate the protein-folding problem. Here, we optimized the thermodynamic nature of vaccine protein, changes in protein conformations and residues fluctuations. Further, to equilibrate the solvents and ions around the protein two phase were used i.e. NVT and NPT. Here, we conduct 100-ps NVT-equilibration for our vaccine construct. Similarly, to stabilize the pressure 100-ps NPT-equilibration was used. Now, after the equilibration of temperature and pressure MD simulation of our vaccine construct was performed with time duration of 10- ns (nano second). Now, to find out the stability of our vaccine construct RMSD (Root mean square deviation) of protein was calculated. The output plot showing the result of RMSD 0.4 nm with time duration of 5-ns, 0.6 nm with 6-ns and 7–10 ns provides 0.8 nm of RMSD. The average RMSD calculated for vaccine construct was 6 Å. Further, RMSF (Root mean square fluctuation) of vaccine protein was also optimized, it is normal at residue position 150 but the highly fluctuating residues were found at position of 200 after that at 500 residue positions. The 3-Dimensional model superimposition of final vaccine constructs and the simulated model of final vaccine construct has been shown in Fig. 6.

### 3.6. Molecular dynamics simulation of vaccine protein (ligand) and TLR-2(receptor) complex

Molecular dynamics simulation of final vaccine constructs (ligand) and TLR-2 (receptor) complex was performed to evaluate the interaction between vaccine constructs and TLR-2 agonist. The algorithms used for energy minimization was steepest descents converged and the maximum force applied to minimized the potential energy with maximum force ( $F_{\max} < 1000$  in 1494 steps). Further, the potential energy calculated for the complex was  $-5.2423845e+06$ , it rely on maximum force i.e. 9.0253687e

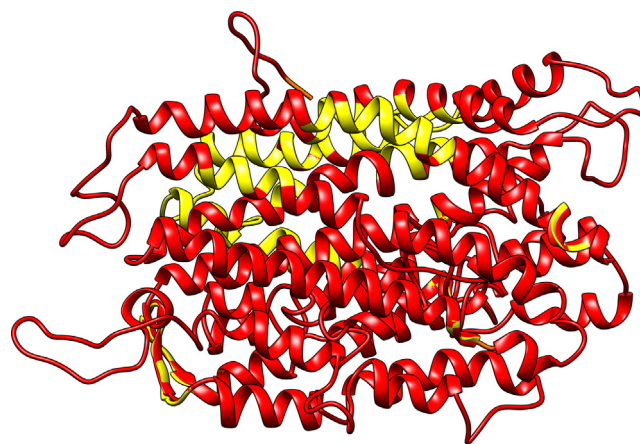


**Fig. 6.** Superimposition model of final vaccine construct (red) in color with molecular dynamics simulation (yellow) in color. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

+02 on atom 8037. Hence, the energy parameter is calculated with total drift of 989(KJ/mol). Next, the temperature was equilibrated for time period of 100-ps with total- drift of 1.19673(K). Similarly, the pressure was equilibrated for 100-ps with total drift of 5.88075 (bar). After, equilibration of temperature and pressure, molecular dynamics simulation of vaccine construct and TLR-2 complex was performed with time duration of 10-ns. To optimize the stability of vaccine construct and TLR-2 complex RMSD was evaluated i.e. ranging from (0.2 to 1 nm). The average RMSD considered for a vaccine construct was 8 Å. RMSF of vaccine construct and TLR-2 was also predicted the fluctuated residues were at position 250 with range of 1.3 nm and at 400 with range of 1.2 nm. The evaluations of NVT, NPT, RMSD and RMSF graph plots have shown in Fig. 7.

### 3.7. Molecular dynamics simulation of vaccine protein (ligand) and TLR-4(receptor) complex

To evaluate the interaction between ligand (vaccine constructs) and TLR-4(receptor) complex molecular dynamics simulation was performed. Further, for energy minimization the steepest descents

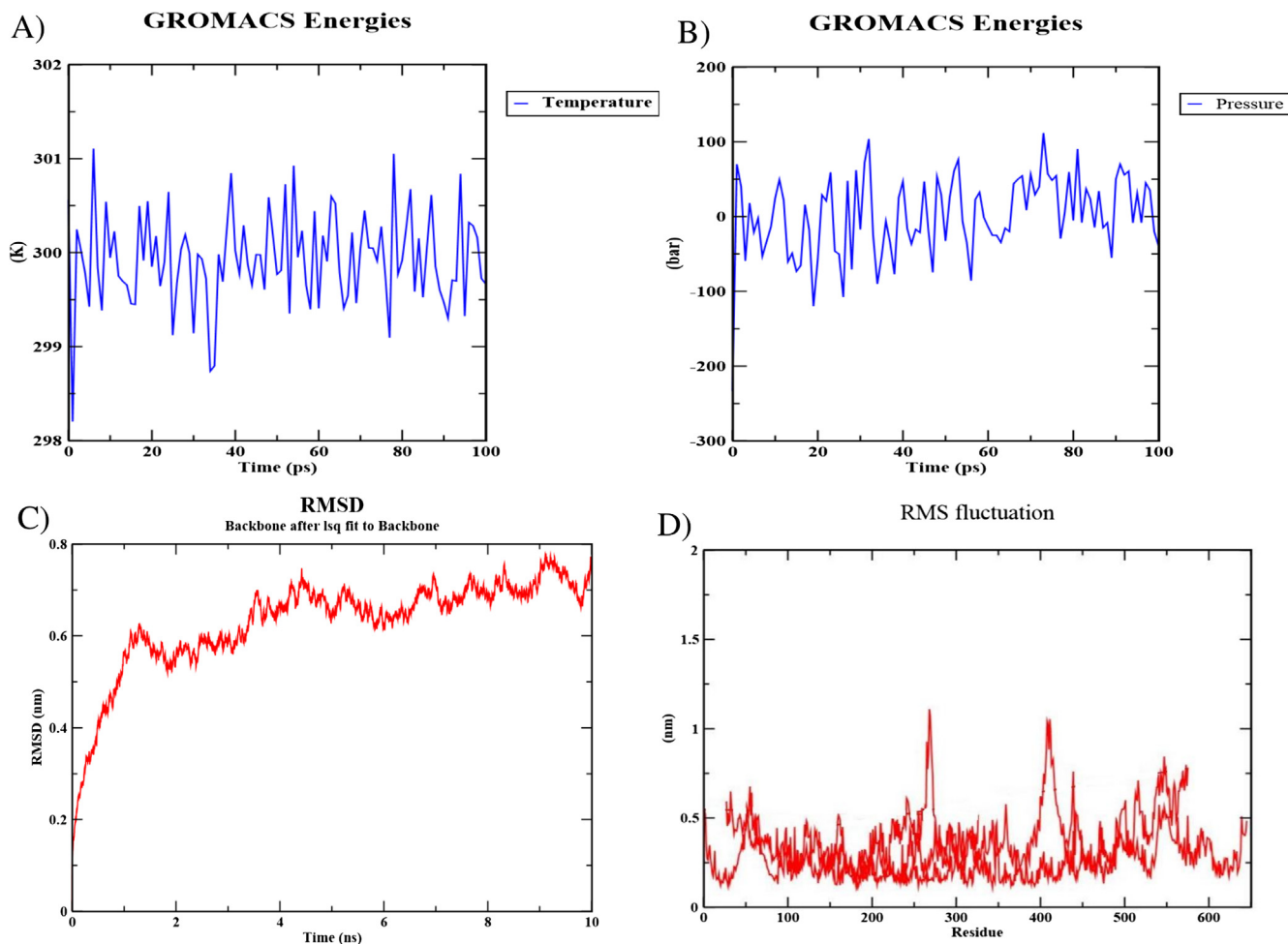


**Fig. 7.** Molecular dynamics of vaccine- immune receptor complex; showing temperature (NVT), pressure (NPT), RMSD (root mean square deviation) and RMSF (root mean square fluctuation) plot of modeled vaccine construct (ligand) and TLR-2 (receptor).

algorithm was utilized to evaluate a potential energy graph by applying maximum force ( $f_{\max}$  1000 in 1339 steps). The potential energy was evaluated  $-9.0769710 \times 10^6$  and the maximum force applied were  $9.7404230 \times 10^2$  on atom 3092. Further, the total drift for potential was estimated with 910,567 (KJ/mol). NVT and NPT were predicted for equilibration of system, NVT was estimated for time period of 100-ps and total drift predicted 1.028 (K). Similarly, NPT was evaluated for time duration of 100-ps and energy parameter predicted 4.78743 (bar). Hence, the molecular dynamics simulation was optimized for time period of 10-ns. The predicted RMSD ranging from time period of 2 to 4 ns was 0.6 nm and after 4 to 6 ns it raised from 0.6 to 0.7 nm and the average RMSD estimated for ligand (vaccine construct) and TLR-4 (receptor) was 0.6 nm i.e. 6 Å. RMSF was fluctuated at residue position 250 with 0.5 nm range and also at residue position 550 with 0.6 nm. The evaluations of NVT, NPT, RMSD and RMSF graph plots have shown in Fig. 8.

### 3.8. Evaluation of insilico cloning of vaccine in microbial system

For prominent protein expression of our final vaccine constructs, we performed in silico cloning. For this study, firstly we analyzed DNA sequences of our vaccine constructs on the basis of CAI (1.0) and GC content (55.5) with host organism *E. coli* (strain /K12) using JCat tool. Further, we optimized reverse of DNA

sequence because to evaluate the expression in vector we want C to N terminal sequences. After, the codon optimization of sequences the cloning processes was start and these optimized sequences were used for further prediction. To clone, the DNA sequences of our vaccine construct we used pET-28a (+) vector with restriction site XhoI and BamHI respectively. This *in-silico* cloning confirms the expression of our vaccine constructs. The restriction clone having a length of 7270 base pair was formed by inserting in pET-28a (+) vector shown in Fig. 9.

## 4. Conclusion

*T. cruzi* is one of the protozoan parasites, causes chagas disease. *T. cruzi* infection endemic in Central and South America, this disease occurs especially in rural area. The two drugs available for chagas disease i.e. nifurtimox and benznidazole have different mechanism of action against *T. cruzi* but both of these drugs are not good to use. Cross- resistance has been found in nitroheterocyclic drugs, similarly mutations in FMN- binding region of TcNTR, which loses the enzyme activity; hence, it creates drug resistant in parasites. In this study, we exported a total of 8 antigenic proteins for *T. cruzi*. Further, these 8 antigenic proteins used to evaluate immunogenic CTL, HTL and linear B cell which is capable to generate humoral and cell-mediated immunity. Moreover, to make our vaccine more potent we combine all predicted epitopes with

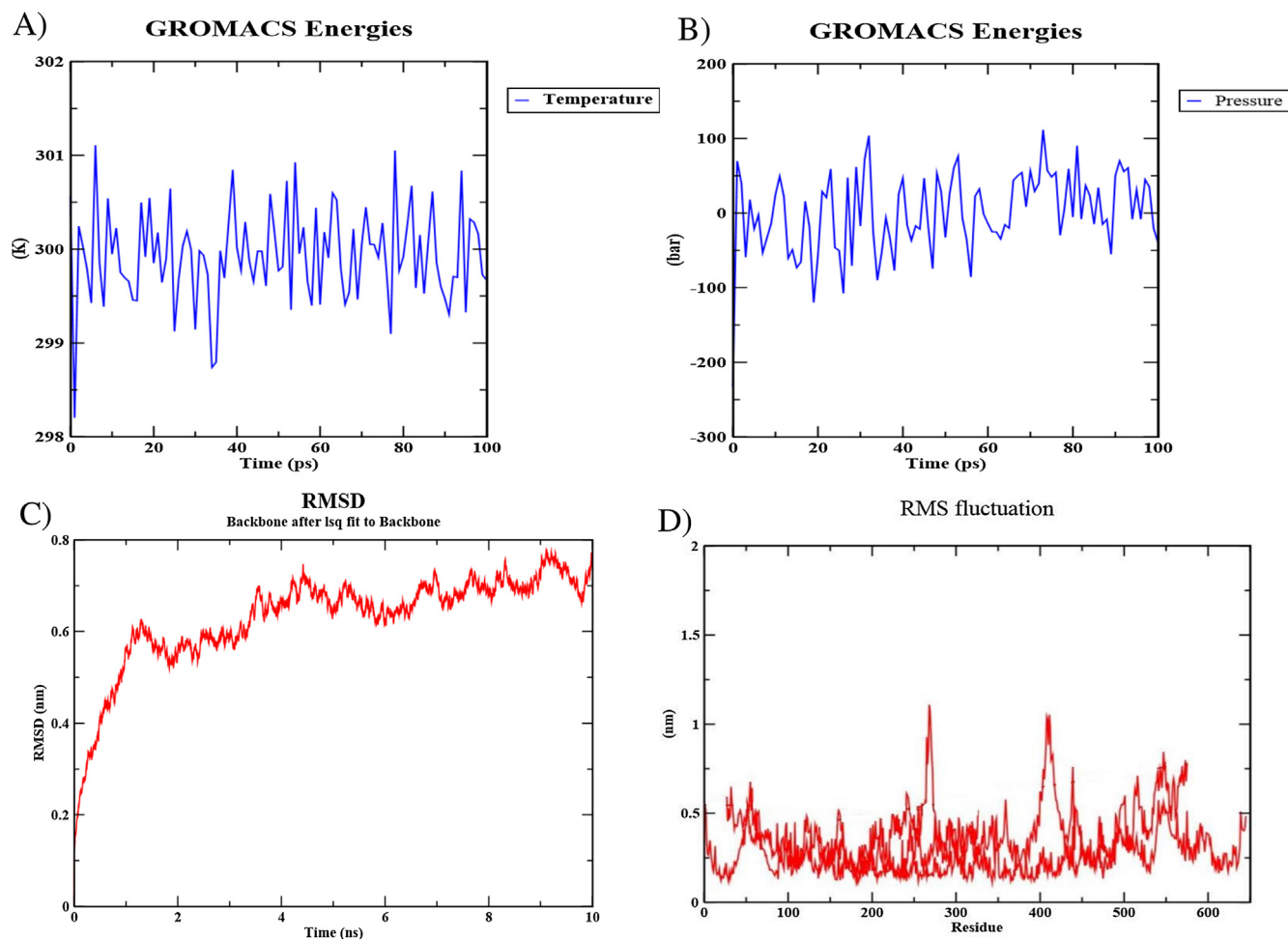


**Fig. 8.** Molecular dynamics of vaccine-immune receptor complex; showing temperature (NVT), pressure (NPT), RMSD (root mean square deviation) and RMSF (root mean square fluctuation) plot of modeled vaccine construct (ligand) and TLR-4 (receptor).

adjuvant to potentiate the immunogenic behavior of vaccine and suitable linkers because they are helpful to divide epitopes in human body. Allergenicity, antigenicity and physiochemical characterization of designed vaccine protein was also confirmed to make our vaccine protein stable, non-allergen and antigenic in nature. TLR-2 and TLR-4 receptors were selected for molecular docking process because they both are good agonist for trypanosomiasis as well as TLR-2 works as effective activators and TLR-4 provides a proinflammatory response during infection. Hence, both used as a receptor to optimized the interaction with final vaccine constructs, which ensures the binding affinity and stability of TLR's and vaccine complex. Further, molecular dynamics simulation was generated by performing energy minimization, NVT, NPT steps. RMSD confirms the configuration of final vaccine proteins (ligand) when bound to TLR's (receptor). Similarly, RMSF calculates the average over time and provides the size for each atom during simulation for TLR-vaccine complex which confirms the flexibility of vaccine protein. Finally, the final vaccine construct was cloned in pET-28a (+) vector which was showing effective expression vaccine model. Consequently, Immunoinformatics approach has been used to designed epitope-based multi subunit vaccine against *T. cruzi* infection. Furthermore, this proposed multi epitope-subunit vaccine model needs to be experimentally validated to confirm immunogenic response against *T. cruzi* infection.

## 5. Methodology

### 5.1. Investigation of secretory and membrane proteins of *T. cruzi*

A total number of 7 secretory and 6 membrane antigenic proteins of *T. cruzi* collected from comprehensive literature survey [19]. All 13 proteins were exported from have been TriTrypDB (The Kinetoplastid genomic resources) <http://tritrypdb.org/tritrypdb/> and then all these exported proteins were subjected to Pred-TMBB <http://bioinformatics.biol.uoa.gr/PRED-TMBB/> sever. It is based on HMM (Hidden Markov Model) approach used for predicting and discriminating  $\beta$ -barrel outer membrane proteins, 2.965 is the prediction score of this server. The sequence scored a value lower than the prediction score indicates the possibility of the protein being outer membrane proteins. Further, TargetP 1.1 server <http://www.cbs.dtu.dk/services/TargetP/> was also used for the prediction of subcellular localization of 6 membrane proteins. Total, 13 secretory and membrane proteins were selected. Further, these 13 proteins all proteins were expose to AlgPred server <http://crdd.osdd.net/raghava/algpred/> and only those proteins were selected for further study which are Non-allergens in nature and rest of the proteins have been avoided by applying this approach only 8 proteins have been chosen for further designing of multi-epitope subunit vaccine [28].

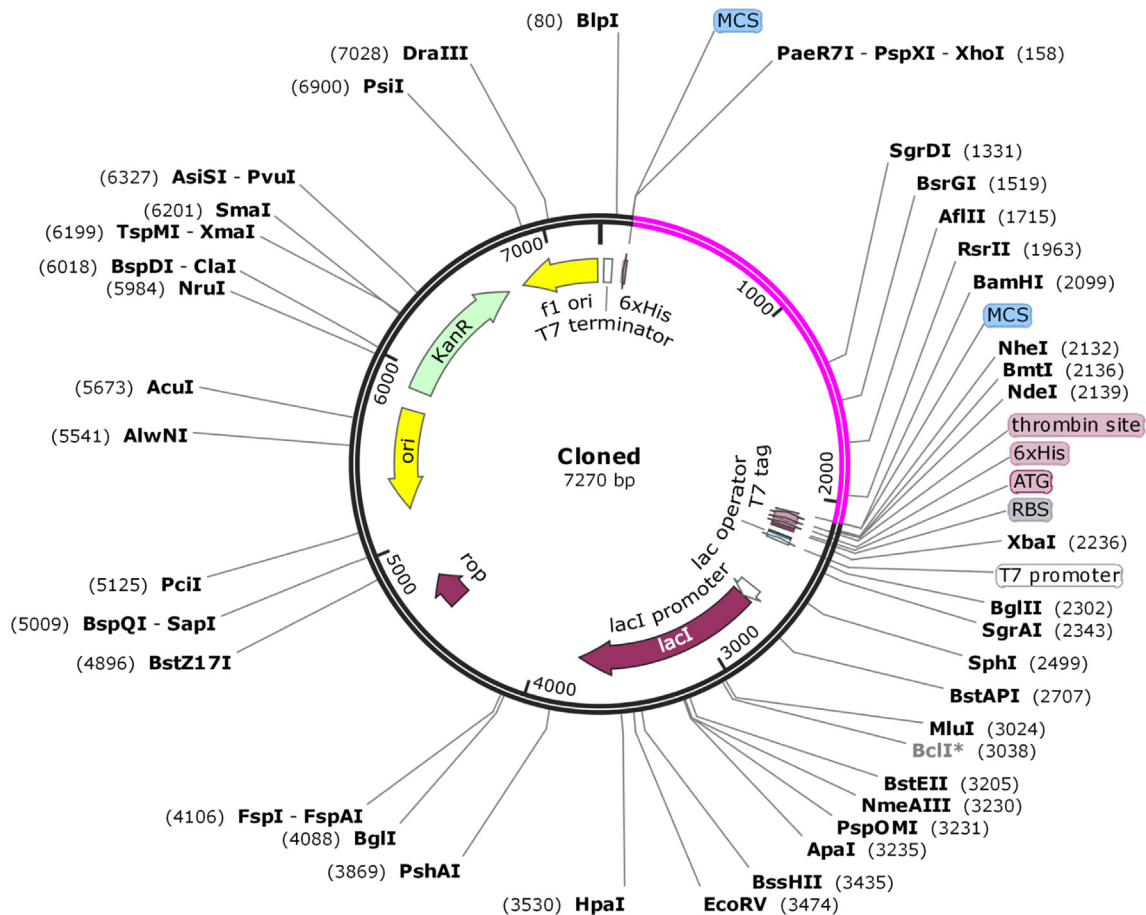


Fig. 9. In-silico cloning map of final vaccine construct.

### 5.2. Assortment of MHC (Major histocompatibility complex) class I (CTL) and MHC class II (HTL)

The MHC class I, CTL (cytotoxic T-lymphocyte) prediction of all 8 antigenic proteins of *T. cruzi* has been evaluated by using NetCTL 1.2 server <http://www.cbs.dtu.dk/services/NetCTL/>. This is online freely available web server, and helpful to predict human CTL epitopes for all given proteins as an input. NetCTL used HIV proteins as its source proteins, and 12 recognized HLA supertypes. The five different methods used for the prediction but only two methods used at a time to compare highest epitope. It predicts the output among 5% top-scoring epitopes with above score of 0.72 [29]. After, the prediction of CTL epitopes, similarly MHC class II, HTL (Helper T-cells) have been evaluated for all 8 antigenic proteins of *T. cruzi* IEDB (Immune Epitope Database and Analysis Resource) <http://www.iedb.org/>. MHC class II binding predictions of IEDB server predicts epitopes on the basis of 17,000 large scale datasets, and a set of 11 HLA DP and DQ alleles as well as, HLA-DR alleles was also present with 40,000 MHC class II binding affinities which cover 26 different alleles. Different types of machine learning methods have been used for better output. All available alleles cover >99% of population. For the prediction of MHC class II binding epitopes, an average of 1713 dataset for each allele and 858 binders have been included. Moreover, different machine learning methods provides output on the basis of IC<sub>50</sub> [30].

### 5.3. Analysis of linear and discontinuous B cell peptides

To analyze the linear B cell epitopes of 8 antigenic proteins of *T. cruzi* BepiPred-2.0 <http://www.cbs.dtu.dk/services/BepiPred/> ser-

ver was used. BepiPred-2.0 is a web server; it is based on random forests or random decision forests algorithms, there are 649 antigen-antibody 3D- structures dataset have been set for the prediction of antigenic linear B-cell epitopes. Thus, the best antigen sequences were chosen and the complexes which have less than 70% identity to any other sequences in dataset have been removed for further prediction. BepiPred-2.0 server provides output with different colors on the basis of epitope threshold 0.5. Thus, to estimate the performance of BepiPred-2.0 server 5-fold cross validation method was used. It provides the output both in tabular and graphical view according to the score of epitopes [31]. Similarly, for the optimization of discontinuous B-cell epitopes ElliPro <http://tools.iedb.org/ellipro/> a structure- based tool used. ElliPro server was established on Thornton's method as well as algorithms used for this method was residue clustering. A dataset used for the ElliPro is a set of discontinuous epitopes from crystal structures of antigen-antibody complexes. The most significant value by which ElliPro provides the good result is (AUC value = 0.732). Hence, ElliPro gives better performance by implementing protrusion index (PI) value for each residue. The best prediction score provides by ElliPro is when the P-value is 0.05. Thus, ElliPro considered as a significant tool for investigating antibody epitopes in 3D structure of proteins [32].

### 5.4. Scheming of Multi-epitope subunit vaccine sequences for *T. cruzi*

Firstly, an appropriate adjuvant was selected to improve the immunogenicity of the vaccine construct. The sequence of RS09 as an adjuvant was chosen for further vaccine design [25]. Next, the exported high scorer CTL and lower percentile of HTL as well

as highest score of B-cell epitopes from antigenic proteins of *T. cruzi* were arranged in linear manner. After, the arrangement of adjuvant, CTL, HTL and B cell epitopes to link them together significant linkers were used for the construction of vaccine [33].

#### 5.5. Prediction of IFN- $\gamma$ epitopes to verify protective immune response

For the prediction of IFN- $\gamma$  epitopes, HTL epitopes of final vaccine construct were used as input in IFNepitope <http://crdd.osdd.net/raghava/ifnepitope/index.php> server. It is freely available server used for the prediction of IFN epitopes which is released from CD4 + T cells. All IFN- $\gamma$  inducing epitopes have been collected from the IEDB server, 3705 inducing and 6728 non-inducing epitopes were found out for further implementation. This server is based on MERCI and SVM as well as on Hybrid (SVM + MERCI) methods. IFNepitope server provides an output with positive and negative epitopes.

#### 5.6. Exploration of allergenicity and antigenicity of designed vaccine constructs

The allergenicity of final vaccine constructs was evaluated using AllgPred <http://crdd.osdd.net/raghava/allgpred/> server. The dataset used for AllgPred is 323 allergens and on non-allergens 101,725 obtained from Swiss-Prot. About, 157 IgE epitopes was obtained from different databases and included for prediction. There are five different methods used for the development of AllgPred server i.e. first method was SVM which achieved an accuracy of 85.02 and 84.00%, respectively. Similarly, a motif-based method used which provides a sensitivity of 93.94 with 33.34% of specificity. Next, IgE epitopes database given a sensitivity of 17.47% with specificity of 98.14% of predicted allergen proteins. Fourth, BLAST search performed against allergen epitopes. At last, hybrid approach developed which provides output on the basis of two or more approaches [28]. Further, antigenicity of designed vaccine constructs has been exported from Immunomedicine Group: Tools > Predicted Antigenic Peptides <http://imed.med.ucm.es/Tools/antigenic.pl>. Predicted Antigenic Peptides tool was based on Kolarik and Tongaonkar method. It is the simplest way to find out the antigenic determinants on the basis of amino acid residues. It gives an output with an accuracy of 75%.

#### 5.7. Physicochemical characterization

The physicochemical parameters of designed vaccine construct were performed using EXPASY- ProtParam tool <https://web.expasy.org/protparam/>. ProtParam determined the length of the amino acids, molecular weight, and theoretical index (pI), and instability index, *in vitro* and *in vivo* half-life, aliphatic index as well as grand average hydropathicity (GRAVY).

#### 5.8. 3-Dimensional structure prediction

Comparative modeling of final vaccine model was obtained from RaptorX web server <http://raptorx.uchicago.edu/>. RaptorX server is not only reliable for prediction of 3 D protein model but also predicts protein secondary structure and alignment assessment. It provides an output on the basis of P-value, GDT (global distance test) and uGDT (un-normalized GDT) for the accurate quality of the 3-D structure. It can also calculate the modeling errors and disordered regions of the modeled protein structure [34].

#### 5.9. 3-Dimensional structure validation and refinement

The modeled vaccine construct was refined using GalaxyRefine server <http://galaxy.seoklab.org/cgi-bin/submit.cgi?type=REFINE>. For the refinement of protein structure this server used *ab initio* method for modeling. It is based on CASP<sub>9</sub> method for protein structure prediction, top 20 homologs and multiple templates were selected to remove disorders and outliers from the protein structure. It enhance the global as well as local structure quality [35]. Furthermore, Ramachandran plot of final vaccine construct was carried out by RAMPAGE server <http://mordred.bioc.cam.ac.uk/~rapper/rampage.php>. ProSA-web <https://prosa.services.came.sbg.ac.at/prosa.php> tool used for recognizing the errors in protein modeled structure. ProSA calculates results on the basis of C $\alpha$  potentials; therefore, it can also calculate a set of low- resolution structures i.e. C $\beta$  potentials. It displayed the results on the basis of Z-score and a plot based on residue energies [36].

#### 5.10. Molecular docking of designed vaccine with the immune cell receptor TLR-2/TLR-4

To explore an optimized confirmation between both final vaccine protein (ligand) and TLR's (receptor) molecular docking method was used. Here, two TLR's used for further study i.e. TLR-2 and TLR-4. The ClusPro 2.0 (protein-protein docking server) <https://cluspro.bu.edu/login.php> was used for recognizing the interaction between designed vaccine protein as a ligand and TLR-2/TLR-4 as a receptor. PIPER, docking algorithm was used for the implementation of ClusPro server, PIPER algorithms calculate the result on the basis of interaction energy of two proteins i.e.  $E = E_{rep} + E_{attr} + E_{elc} + E_{DARS}$ . ClusPro provides an output based on energy parameters of top ten models and low docked energy score. Root mean square deviation of each docked complex was also optimized by clustering of lowest energy structures [37].

#### 5.11. Optimization of molecular dynamics simulation of designed final vaccine model and docked structure of vaccine construct and TLR-2/TLR-4

For molecular dynamics simulation analyses of final vaccine construct and docked structure of vaccine construct and TLR-2/TLR-4 GROMACS 5.x series was used. It is one of the significant platforms to analyze the dynamical simulations of biomolecules. It can calculate the interactions between the particles by differentiating their equations of motion. The interactions calculated on the basis of temperature, pressure and potential energy which correct the geometric restraints and proteins unstable nature. Better force fields applied for the implementation of the GROMACS i.e. AMBER, CHARMM, GROMOS and OPLSS. The algorithms of molecular dynamics simulations were based on Newton's equations of motion, and the Verlet algorithm [38].

#### 5.12. Evaluation of in-silico cloning of vaccine constructs

To precise the final vaccine constructs in a proper expression vector, first of all we used JCat tool i.e. (Java Codon Adaptation Tool) <http://www.jcat.de/>. JCat is a rapid and easy method to find highly expressed genes; it avoids the unwanted cleavage sites for restriction enzymes. It gives output on the basis of CAI values. Secondly, we use Reverse Complement tool [https://www.bioinformatics.org/sms/rev\\_comp.html](https://www.bioinformatics.org/sms/rev_comp.html) for the conversion of DNA sequences into its reverse complement. After, preparation of sequences we use SnapGene tool to ensure the high-level protein expression.

## Acknowledgement

NK is thankful to Science and Engineering Research Board (SERB) for providing fellowship. RO is thankful to UGC for providing university fellowship. VKP is thankful to the Central University of Rajasthan for providing computational facility.

## Author contribution

Protocol designed by NK, VKP.  
Methodology performed by NK, RO, VKP.  
The manuscript was written by NK, VKP.

## Competing financial interests

The Authors have declared no competing interest.

## Appendix A. Supplementary material

Supplementary data associated with this article can be found, in the online version, at <https://doi.org/10.1016/j.vaccine.2018.09.004>.

## References

- [1] Almeida EAd, Ramos Júnior AN, Correia D, Shikanai-Yasuda MA. Co-infection *Trypanosoma cruzi*/HIV: systematic review (1980–2010). *Rev Soc Bras Med Trop* 2011;44:762–70.
- [2] Franco JR, Simarro PP, Diarra A, Jannin JG. Epidemiology of human African trypanosomiasis. *Clinical Epidemiol* 2014;6:257.
- [3] Coura JR, Viñas PA, Junqueira AC. Ecoepidemiology, short history and control of chagas disease in the endemic countries and the new challenge for non-endemic countries. *Mem Inst Oswaldo Cruz* 2014;109:856–62.
- [4] Telleria J, Tibayrenc M. American trypanosomiasis chagas disease: one hundred years of research. Elsevier; 2017.
- [5] Keating J, Yukich JO, Sutherland CS, Woods G, Tediosi F. Human African trypanosomiasis prevention, treatment and control costs: a systematic review. *Acta Trop* 2015;150:4–13.
- [6] Baker N, de Koning HP, Mäser P, Horn D. Drug resistance in African trypanosomiasis: the melarsoprol and pentamidine story. *Trends Parasitol* 2013;29:110–8.
- [7] Tsegaye B, Dagnachew S, Terefe G. Review on drug resistant animal trypanosomes in africa and overseas. *African J Basic Appl Sci* 2015;7:73–83.
- [8] Maguire JH. Treatment of Chagas' Disease—time is running out. *Mass Medical Soc* 2015.
- [9] Brener Z, Gazzinelli RT. Immunological control of *Trypanosoma cruzi* infection and pathogenesis of Chagas' disease. *Int Arch Allergy Immunol* 1997;114:103–10.
- [10] Tarleton R. Depletion of CD8+ T cells increases susceptibility and reverses vaccine-induced immunity in mice infected with *Trypanosoma cruzi*. *J Immunol* 1990;144:717–24.
- [11] Khatoon N, Pandey RK, Prajapati VK. Exploring leishmania secretory proteins to design B and T cell multi-epitope subunit vaccine using immunoinformatics approach. *Sci Rep* 2017;7:8285.
- [12] Ali M, Pandey RK, Khatoon N, Narula A, Mishra A, Prajapati VK. Exploring dengue genome to construct a multi-epitope based subunit vaccine by utilizing immunoinformatics approach to battle against dengue infection. *Sci Rep* 2017;7:9232.
- [13] McGuckin MA, Lindén SK, Sutton P, Florin TH. Mucin dynamics and enteric pathogens. *Nat Rev Microbiol* 2011;9:265–78.
- [14] Narula A, Pandey RK, Khatoon N, Mishra A, Prajapati VK. Excavating chikungunya genome to design B and T cell multi-epitope subunit vaccine using comprehensive immunoinformatics approach to control chikungunya infection. *Infect Genet Evol* 2018. <https://doi.org/10.1016/j.meegid.2018.03.007>.
- [15] Bermejo DA, Jackson SW, Gorosito-Serran M, Acosta-Rodriguez EV, Amezcua-Vesely MC, Sather BD, et al. *Trypanosoma cruzi* trans-sialidase initiates a program independent of the transcription factors ROR [gamma] t and Ahr that leads to IL-17 production by activated B cells. *Nat Immunol* 2013;14:514–22.
- [16] Wilson A. Emerging principles for the recognition of peptide antigens by MHC class I molecules. *Science* 1992;257:927–34.
- [17] Bull PC, Lowe BS, Kortok M, Molyneux CS, Newbold CI, Marsh K. Parasite antigens on the infected red cell surface are targets for naturally acquired immunity to malaria. *Nat Med* 1998;4:358.
- [18] Pandey RK, Ojha R, Aathmanathan VS, Krishnan M, Prajapati VK. Immunoinformatics approaches to design a novel multi-epitope subunit vaccine against HIV infection. *Vaccine* 2018;36:2262–72.
- [19] Ádic Pech-Canul, Monteón V, Solís-Oviedo R-L. A brief view of the surface membrane proteins from *Trypanosoma cruzi*. *J Parasitol Res* 2017.
- [20] Mendonça-Previato L, Regina Todeschini A, Freire de Lima L, Osvaldo Previanto J. The trans-sialidase from *Trypanosoma cruzi* a putative target for trypanocidal agents. *Open Parasitol J* 2010;4.
- [21] Gea S, Guinazu N, Pellegrini A, Carrera Silva E, Giordanengo L, Cano R, et al. Cruzipain, a major *Trypanosoma cruzi* cysteine protease in the host-parasite interplay. *Immunol España* 2006;25:225–38.
- [22] Bartholomeu DC, Cerqueira GC, Leao ACA, daRocha WD, Pais FS, Macedo C, et al. Genomic organization and expression profile of the mucin-associated surface protein (masp) family of the human pathogen *Trypanosoma cruzi*. *Nucleic Acids Res* 2009;37:3407–17.
- [23] Colorado IA, Acquatella H, Cataliotti F, Fernandez MT, Layrisse Z. HLA class II DRB1, DQB1, DPB1 polymorphism and cardiomyopathy due to *Trypanosoma cruzi* chronic infection. *Hum Immunol* 2000;61:320–5.
- [24] Chatterjee N, Ojha R, Khatoon N, Prajapati VK. Scrutinizing mycobacterium tuberculosis membrane and secretory proteins to formulate multipitope subunit vaccine against pulmonary tuberculosis by utilizing immunoinformatic approaches. *Int J Biol Macromol* 2018;118:180–8.
- [25] Shanmugam A, Rajoria S, George AL, Mittelman A, Suriano R, Tiwari RK. Synthetic toll like receptor-4 (TLR-4) agonist peptides as a novel class of adjuvants. *PLoS ONE* 2012;7:e30839.
- [26] Campos MA, Almeida IC, Takeuchi O, Akira S, Valente EP, Procópio DO, et al. Activation of toll-like receptor-2 by glycosylphosphatidylinositol anchors from a protozoan parasite. *J Immunol* 2001;167:416–23.
- [27] Ojha R, Khatoon N, Prajapati VK. Conglomeration of novel culex quinquefasciatus salivary proteins to contrive multi-epitope subunit vaccine against infections caused by blood imbibing transmitter. *Int J Biol Macromol* 2018;118:834–43.
- [28] Saha S, Raghava G. AlgPred: prediction of allergenic proteins and mapping of IgE epitopes. *Nucleic Acids Res* 2006;34:W202–9.
- [29] Larsen MV, Lundegaard C, Lamberth K, Buus S, Lund O, Nielsen M. Large-scale validation of methods for cytotoxic T-lymphocyte epitope prediction. *BMC Bioinf* 2007;8:424.
- [30] Wang P, Sidney J, Kim Y, Sette A, Lund O, Nielsen M, et al. Peptide binding predictions for HLA DR, DP and DQ molecules. *BMC Bioinf* 2010;11:568.
- [31] Jespersen MC, Peters B, Nielsen M, Marcotili P. BepiPred-2.0: improving sequence-based B-cell epitope prediction using conformational epitopes. *Nucleic Acids Res* 2017;45:W24–9.
- [32] Ponomarenko J, Bui H-H, Li W, Fusseder N, Bourne PE, Sette A, et al. ElliPro: a new structure-based tool for the prediction of antibody epitopes. *BMC Bioinf* 2008;9:514.
- [33] Pandey RK, Ali M, Ojha R, Bhatt TK, Prajapati VK. Development of multi-epitope driven subunit vaccine in secretory and membrane protein of *Plasmodium falciparum* to convey protection against malaria infection. *Vaccine* 2018;36:4555–65.
- [34] Peng J, Xu J. RaptorX: exploiting structure information for protein alignment by statistical inference. *Proteins Struct Funct Bioinf* 2011;79:161–71.
- [35] Ko J, Park H, Heo L, Seok C. GalaxyWEB server for protein structure prediction and refinement. *Nucleic Acids Res* 2012;40:W294–7.
- [36] Wiederstein M, Sippl MJ. ProSA-web: interactive web service for the recognition of errors in three-dimensional structures of proteins. *Nucleic Acids Res* 2007;35:W407–10.
- [37] Kozakov D, Hall DR, Xia B, Porter KA, Padhorny D, Yueh C, et al. The ClusPro web server for protein–protein docking. *Nat Protoc* 2017;12:255.
- [38] Allen MP. Introduction to molecular dynamics simulation. *Comput Soft Matter: Syn Polym Proteins* 2004;23:1–28.

**Examination of antigenic proteins of *Trypanosoma cruzi* to fabricate an epitope-based subunit vaccine by exploiting epitope mapping mechanism**

Nazia Khatoon, Rupal Ojha, Vijay Kumar Prajapati\*

Department of Biochemistry, School of Life Sciences, Central University of Rajasthan, Bandarsindri, Kishangarh, Ajmer, Rajasthan-305817 India

**Running title:** Subunit vaccine to control Trypanosomiasis

**\*Corresponding Author**

Prof. Vijay Kumar Prajapati,  
Department of Biochemistry  
Central University of Rajasthan  
NH-8, Bandarsindri, Ajmer  
Rajasthan Pin- 305817 India  
Phone: +91 7597271362  
Email: [vkprajapati@curaj.ac.in](mailto:vkprajapati@curaj.ac.in);  
[Vijay84bhu@gmail.com](mailto:Vijay84bhu@gmail.com)

**Supplementary Table1: Prediction of CTL epitopes using NetCTL server with A2/A3/B7 supertype**

| <b>Protein name</b>                                                        | <b>A2<br/>supertype/Epitopes</b> | <b>A3<br/>supertype/Epitopes</b> | <b>B7<br/>supertype/Epitopes</b> |
|----------------------------------------------------------------------------|----------------------------------|----------------------------------|----------------------------------|
| Trypomastigote, Alanine,<br>Serine and Valine rich protein<br>(TASV)       | RVACYLLVL                        | MMMMATVRR                        | RVACYLLVL                        |
| Peroxidase, putative                                                       | LLMMCIGCL                        | KSWEHLNHK                        | SPASAGSST                        |
| Trans-sialidase, Group II,<br>putative                                     | HVLTFVLL                         | SIVAEIAKK                        | FVPNKTQVL                        |
| Cruzipain;Cruzaine                                                         | ALLLA AVL V                      | LTSQFAEFK                        | RALLLA AVL                       |
| RNase PH-like protein,<br>putative ribosomal RNA<br>processing protein 41B | GLTTVVA AV                       | TQFDGSAWY                        | IPRAGGTDR                        |
| RNA-binding protein,<br>putative                                           | CLFFAAPT V                       | MSSFQLFSR                        | AAPT VKMAM                       |
| Serine-alanine-and proline-<br>rich protein, putative                      | MMRRVFCV                         | RVFCVVLAF                        | RVFCVVLAF                        |
| Mucin-associated surface<br>protein (MASP)                                 | AMMTGRVLL                        | RLQLRQEQR                        | MAMMTGRVL                        |

**Supplementary Table: 2A Prediction of HTL epitopes using IEDB server with HLA-DRB1\*01:01**

| <b>HLA-DRB1*01:01</b>                                                                 | <b>Epitopes</b>    | <b>Score</b> |
|---------------------------------------------------------------------------------------|--------------------|--------------|
| TcCLB.510717 Trypomastigote, Alanine, Serine and Valine rich protein (TASV)           | AWVRMPLLLLLTVLV    | 0.14         |
| TcCLB.511143.30 peroxidase, putative                                                  | EDAASLLAVA AVKFL   | 3.72         |
| TcCLB.510307.210 trans-sialidase, Group II, putative                                  | SRHVLTF AVL LLLLVV | 2.51         |
| P25779.1Cruzipain;Cruzaine                                                            | GWARALLLA AVL VVM  | 0.77         |
| TcCLB.505037.20, RNase PH-like protein, putative ribosomal RNA processing protein 41B | VEMFLTASIQA VVRL   | 1.43         |
| TcCLB.511317.70 RNA-binding protein, putative                                         | CCCLFFAAPT VKMAM   | 5.2          |
| TcCLB.511487.150 serine-alanine-and proline-rich protein, putative                    | CFHYTRLHALPPLVL    | 0.32         |
| TcCLB.511487.170 Mucin-associated surface protein (MASP)                              | PLLVVACAAAAAVVA    | 0.62         |

**Supplementary Table: 2B Prediction of HTL epitopes using IEDB server with DPA1\*01/DPB1\*04:01**

| <b>DPA1*01/DPB1*04:01</b>                                                             | <b>Epitopes</b>    | <b>Score</b> |
|---------------------------------------------------------------------------------------|--------------------|--------------|
| TcCLB.510717 Trypomastigote, Alanine, Serine and Valine rich protein (TASV)           | TAWVRMPLLLLLTVL    | 0.35         |
| TcCLB.511143.30 peroxidase, putative                                                  | YKLSNDYYKFLANE     | 2.79         |
| TcCLB.510307.210 trans-sialidase, Group II, putative                                  | SRHVLTF AVL LLLLVV | 0.03         |
| P25779.1Cruzipain;Cruzaine                                                            | AFRLSVFRENLFAR     | 0.01         |
| TcCLB.505037.20, RNase PH-like protein, putative ribosomal RNA processing protein 41B | LAERAAVSIFSFRK     | 3.88         |
| TcCLB.511317.70 RNA-binding protein, putative                                         | VIRQQATAFAMLQRQ    | 3.82         |
| TcCLB.511487.150 serine-alanine-and proline-rich protein, putative                    | MMRRVFCVVLAFALL    | 1.75         |
| TcCLB.511487.170 Mucin-associated surface protein (MASP)                              | TAVSHTTSPLLPLLV    | 2.84         |

**Supplementary Table: 3 Prediction of B-cell using ABCpred server**

| <b>Protein name</b>                                            | <b>Epitopes</b> | <b>Score</b> |
|----------------------------------------------------------------|-----------------|--------------|
| Trypomastigote, Alanine, Serine and Valine rich protein (TASV) | LLCSSSC         | 1.004        |
| Peroxidase, putative                                           | MWRSLRK         | 1.007        |
| Trans-sialidase, Group II, putative                            | MCCGSGG         | 1.018        |
| P25779.1Cruzipain;Cruzaine                                     | LHAEETL         | 1.005        |
| RNA processing protein 41B, putative                           | MSRASLR         | 1.001        |
| RNA-binding protein, putative                                  | FAAPTVK         | 1.003        |
| TcCLB.511487.150 serine-alanine-and proline-rich protein       | FAVEEGE         | 1.018        |
| TcCLB.511487.170 Mucin-associated surface protein (MASP)       | MMTGRVL         | 1.007        |

**Supplementary Table 4: Prediction of Interferon- $\gamma$  inducing epitopes using IFN- $\gamma$  server**

| <b>Serial no.</b> | <b>Epitopes</b>          | <b>Server</b> | <b>Prediction</b> | <b>Score</b>    |
|-------------------|--------------------------|---------------|-------------------|-----------------|
| 1.                | <b>EDAASLLAVAAVKFL</b>   | <b>SVM</b>    | <b>POSITIVE</b>   | <b>0.060911</b> |
| 2.                | <b>GWARALLLA AVL VVM</b> | <b>SVM</b>    | <b>POSITIVE</b>   | <b>0.945349</b> |
| 3.                | <b>AFRLSVFREN LFLAR</b>  | <b>SVM</b>    | <b>POSITIVE</b>   | <b>0.258771</b> |
| 4.                | <b>VEMFLTASIQA VVRL</b>  | <b>SVM</b>    | <b>POSITIVE</b>   | <b>0.204723</b> |
| 5.                | <b>VIRQQATAFAMLQRQ</b>   | <b>SVM</b>    | <b>POSITIVE</b>   | <b>0.505803</b> |
| 6.                | <b>MMRRVFCVVLAFALL</b>   | <b>SVM</b>    | <b>POSITIVE</b>   | <b>0.163865</b> |