International Journal of Peptide Research and Therapeutics (2021) 27:1885–1898
https://doi.org/10.1007/s10989-021-10218-8

Designing of a Novel Fusion Protein Vaccine Candidate Against Human Visceral Leishmaniasis (VL) Using Immunoinformatics and Structural Approaches

Amir Atapour[1,4] · Farideh Ghalamfarsa[1] · Samaneh Naderi[2] · Gholamreza Hatam[3]

Accepted: 10 April 2021 / Published online: 24 April 2021

Abstract

Leishmaniasis is caused by an obligate intracellular protozoan parasite. The clinical forms of leishmaniasis differ from cutaneous leishmaniasis, mucocutaneous leishmaniasis and visceral leishmaniasis (VL) which depend on the parasite species and the host's immune responses. There are significant challenges to the available anti-leishmanial drug therapy, particularly in severe forms of disease, and the rise of drug resistance has made it more difficult. Currently, no licensed vaccines have been introduced to the market for the control and elimination of VL. A potential target for use in candidate vaccines against leishmaniasis has been shown to be *leishmania* Kinetoplastid membrane protein-11 (KMP-11) antigen. In this study, we chose KMP-11 antigen as target antigen in our vaccine construct. In addition, B-type flagellin (fliC) was used as an adjuvant for enhancing vaccine immunogenicity. The GSGSGSGSGSG linker was applied to link the KMP-11 antigen and fliC (KMP-11-fliC) to construct our fusion protein. Bioinformatics approaches such as; 3D homology modeling, CTL, B-cell, MHC class I and II epitopes prediction, allergenicity, antigenicity evaluations, molecular docking, fast simulations of flexibility of docked complex and in silico cloning were employed to analysis and evaluation of various properties of the designed fusion construct. Computational results showed that our engineered structure has the potential for proper stimulation of cellular and humoral immune responses against VL. Consequently, it could be proposed as a candidate vaccine against VL according to these data and after verifying the efficacy of the candidate vaccine through in vivo and in vitro immunological tests.

Keywords Visceral leishmaniasis · Vaccine · Computational design · Fusion protein · KMP-11 antigen

Introduction

Leishmaniasis is caused by *Leishmania* parasites and contains various diseases with different clinical manifestations ranging from skin lesions to fatal visceral leishmaniasis. Human visceral leishmaniasis (VL; also known as kala-azar) is the most serious form of leishmaniasis and is estimated to have a high mortality and morbidity rate among the tropical infectious diseases. According to the WHO report, annually about 200,000 to 400,000 new cases of VL happen, among them 20,000 to 30,000 result in deaths (Chappuis et al. 2007; Joshi et al. 2014; Kevric et al. 2015; Heydarpour et al. 2016). Already, chemotherapy-based anti-leishmanial drugs are available, but are costly and toxic with serious side effects and also can lead to drug resistance (Polonio and Efferth 2008). Currently, there is no successful vaccine for prevention or treatment of VL, so the development of an effective and safe vaccine against this disease is required (Khamesipour et al. 2006; Vakili et al. 2018). *Leishmania* produces

✉ Amir Atapour
amir.atapoor58@yahoo.com

1 Department of Medical Biotechnology, School of Advanced Medical Sciences and Technologies, Shiraz University of Medical Sciences, Shiraz, Iran

2 Department of Diagnostic Laboratory Sciences and Technology Research Center, Shiraz University of Medical Sciences, Shiraz, Iran

3 Basic Sciences in Infectious Diseases Research Center, Shiraz University of Medical Sciences, Shiraz, Iran

4 Department of Medical Biotechnology, Faculty of Medical Biotechnology, School of Advanced Medical Sciences and Technologies, Shiraz University of Medical Sciences, 71348-14336 Shiraz, Iran

several secretory proteins with antigenic properties such as KMP-11, LeIF, GP63, p36/LACK and A2 and they are capable of providing an immune response in the host. So far, many of these secretory proteins have been used in various candidate vaccines but the results were not satisfactory and not able to achieve the clinical trials (Okwor and Uzonna 2009; Kumar and Engwerda 2014; Jain and Jain 2015). The KMP-11 antigen, one of the *Leishmania* proteins, might play a role in the mobility of the parasite and also in the binding to the host cell (de Mendonça, Cysne-Finkelstein et al. 2015; Nasiri et al. 2016). Kinetoplastid membrane protein-11 (KMP-11) exists in all species of kinetoplastid family that is completely conserved and the corresponding protein can induce a very intense cellular immune response. Therefore, the KMP-11 antigen has been considered as a potential antigen for the development of vaccine against leishmaniasis (Basu et al. 2005; de Mendonça et al. 2015; Nasiri et al. 2016). Today, in designing a potent vaccine, the efficient stimulation of immune responses by the safe and effective adjuvants such as bacterial toxins, cytokines, bacterial cell wall component, is one of the main goals of vaccination against infection diseases (Coffman et al. 2010; Atapour et al. 2017, 2020). One of the promising molecular adjuvant in the design of candidate subunit vaccines is TLR (toll-like receptors) agonist such as bacterial flagellin (fliC). FliC is a highly conserved bacterial protein that naturally attaches to toll-like receptor 5 (TLR5) and has been mentioned to be an effective adjuvant for vaccination. FliC protein has the ability to stimulate the humoral and cellular immune system and has successful experiments in conjunction with candidate vaccines. FliC binding to toll like receptors 5 (TLR5) on the surface of the immune cells (APCs) leads to the creation of a signaling cascade that eventually results in stimulation of inflammatory responses in the host (Steinhagen et al. 2011; Lockner et al. 2015; Makvandi et al. 2018; Senevirathne et al. 2020). Study by Mizel et al. has showed that the flagellin molecule is a very powerful stimulant for the immune system. It induces cytokines and chemokines TNF-α, GM-CSF, IL-6. In several studies, different forms of flagellin (native, recombinant and truncated) have been applied as immunological adjuvants or carriers in vaccination strategies (Mizel et al. 2003; Lu and Swartz 2016; Wangkahart et al. 2019). In many of these studies, the flagellin has been used as a fusion protein like flagellin-antigen complex (Atapour et al. 2019; Cui et al. 2019).

Both humoral and cellular immune responses play a role in VL control. However, the elimination or progression of VL depends on the interaction between the environment and the genetics of the individual and the pathogen. In human, cell-mediated immunity or CD8 T cells is a predominant immune response to the intracellular parasitic infections such as VL. CD8 T cells lyse the cells infected with pathogens and result in producing important cytokines for the control of infection (Tripathi et al. 2007; Kaye and Aebischer 2011; Nylén and Kumar 2012). Therefore, in order to have an effective immunity against this type of disease, the focus should be on strengthening the adaptive immune response especially the cellular immune system. In addition, due to the complexity of vaccine development, this requires the collaboration of various groups, including immunologists, molecular biologists and chemical engineers. Nowadays, using the immunoinformatic methods, a new thermodynamically stable vaccine can be predicted in very less time. Hence, here we aimed by bioinformatic tools to design and analyze a fusion protein composed of KMP-11 protein from *L.infantum* and fliC protein (as an adjuvant) (KMP-11-fliC) as a candidate vaccine against leishmaniasis.

Methods

Protein Sequence Identification and Designing the Construct

The amino acid sequences of KMP-11 protein of *L. infantum* (A4IBB3) and B-type flagellin (fliC) of *P.aeroginosa* (P72151) were obtained from the UniProtKB database (http://www.uniprot.org/.). A flexible linker with amino acid sequence of (GSGSGSGGSGSG) was used for joining two parts.

The Physico-Chemical Parameters Analysis

The physical and chemical parameters associated with the designed vaccine have a significant impact on the quality of the vaccine. Therefore, different physical and chemical parameters containing theoretical isoelectric point (pI), molecular weight, and total number of positive and negative residues, extinction coefficient, instability index, grand average hydropathy (GRAVY) and aliphatic index of designed vaccine were calculated by ProtParam web server. The primary protein sequence of candidate vaccine construct was applied as input file in ProtParam tool (Gasteiger et al. 2005).

Secondary Structure Prediction

The features of the secondary structure of designed vaccine were evaluated by the online database PSIPRED. This server based on the input data contains the primary amino acid sequences, predicts the secondary structure. PSIPRED is a simple and accurate secondary structure prediction method, incorporating two feed-forward neural networks, which perform an analysis on output obtained from PSI-BLAST (Position Specific Iterated—BLAST). Using a very stringent cross validation method to evaluate the method's

performance, PSIPRED 3.2 achieves an average Q3 score of 81.6% (Buchan and Jones 2019).

Prediction of 3D Structure

Different tools such as SWISS-MODEL, I-TASSER and Phyre2 were used for 3D structure prediction. Finally, based on the obtained data, I-TASSER tool was selected for the tertiary structure modeling of final vaccine candidate. The I-TASSER is based on the protein structure and function, which predicts through iterative threading assembly simulation. The output of I-TASSER server consists of three parameters including the confidence score (C-score), TM-score and RMSD. The more negative C-score is indicative of higher stability and confidence of a model (Roy et al. 2010).

Validation of the 3D Structure

According to the important role of the validation process in 3D modeling, ProSA-web and ERRAT server and RAMPAGE (Ramachandran plot analysis) were utilized. Validation methods help to identify potential errors in the 3D structures. RAMPAGE server was used for Ramachandran plot analysis**.** RAMPAGE (Ramachandran plot) tool shows the number of residues in favored, allowed and outer regions. The overall quality score for each of the input structures is calculated by ProSA-web and the results are displayed in a plot, including the Z-scores of experimentally defined structures deposited in the PDB. The ERRAT server assessed the statistics of non-bonded interactions between different atom types in the reporting structure (compared to a database of highly refined structures) (Colovos and Yeates 1993; Lovell et al. 2003; Wiederstein and Sippl 2007).

Improving Vaccine Stability by Disulfide Engineering

Disulfide by Design 2 server was used for disulfide engineering of final vaccine construct. It is helpful in increasing the protein stability along with the examination of protein interactions and dynamics. The spanning region of high mobility was selected based on the obtained B-factor followed by the creation of four stabilizing mutations to make a disulfide bridge. The B-factor is a measure of dynamic mobility of atom in a given protein (Craig and Dombkowski 2013).

Prediction of B-cell Epitopes

The linear and conformational B-cell epitopes of the designed fusion protein were investigated using BCPRED, ABCPred and CBTOPE web server, respectively. The BCPRED predicts linear B-cell epitopes applying a novel method of a subsequence kernel with 74.57% accuracy was used (EL-Manzalawy et al. 2008). Moreover, the 20-mer B-cell epitopes with the default specificity threshold (75%) was chosen. The ABCPred predicts B cell epitope(s) in an antigen sequence, using artificial neural network. This is the first server developed based on recurrent neural network (machine-based technique) using fixed length patterns (Saha and Raghava 2006a, b). The CBTOPE can predict conformational B cell epitopes using antigen primary sequence in the absence of any homology with the known structures (Ansari and Raghava 2010). In addition, the Discotope server was employed for prediction of discontinuous B-cell epitopes from 3D protein structures (Kringelum et al. 2012).

CTL Epitopes Prediction

The NetCTL-1.2 was used for prediction of CTL epitopes vital in designing of vaccine. The forecasting process is based on the predictions of proteasomal cleavage, TAP transport proficiency, and MHC class I affinity. The NetCTL 1.2 server predicts the MHC class I binding and proteasomal cleavages using artificial neural networks (ANNs) and TAP transport efficiency predicted using weight matrix. The accuracy of the epitope prediction is significantly improved compared to the previous predicting tool, namely EpiJen, MHC-pathway, MAPPP, and WAPP. During production of CTL epitopes, the threshold value was set at 0.75 (Larsen et al. 2007).

MHC-I Binding Epitopes Prediction

NetMHC 4.0 server was employed for the prediction of MHC-I binding epitopes. The NetMHC server predicts the peptide-MHC class I binding peptides using artificial neural networks (ANNs). The peptide will be identified as a strong binder if it is found among the top x% predicted peptides, where x% is the specified threshold for strong binders (by default of 0.5%). The peptide will be identified as a weak binder if the percentage rank is above the threshold of the strongest binders, but below the specified threshold for the weak binders (by default of 2%) (Andreatta and Nielsen 2016).

MHC Class II Epitope Prediction

NetMHC II 2.3 Server was used for identifying MHC-II binding peptides. NetMHC II predicts peptide binding using ANNs (Artificial Neuron Networks) for HLA-DR, HLA-DQ, HLA-DP and mouse MHC class II alleles. The predicted values are defined by NM IC50 values and as a %-Rank to a set of 1,000,000 random natural peptides (Jensen et al. 2018).

Predict the IFN-γ inducing peptides

Interferon-gamma (IFN-γ) generation as an important factor for stimulating cellular immunity, has a significant role in the control of infectious diseases. In here, IFNepitope server was used for prediction of potential epitopes, which could induce the release of interferon gamma. IFNepitope server is based on a dataset, which comprises of IFN-gamma inducing and non-inducing MHC class II binders. The server predicts using various approaches, such as machine learning technique, motifs-based search, and hybrid approach with the 81.39% accuracy (Dhanda et al. 2013).

Allergenicity Evaluation

The allergencity of the final designed construct was evaluated by the AlgPred server. This server can be used with six different approaches. In our study, IgE, PID, and allergen representative peptides (ARPs) approaches were chosen. In IgE and PID approach, the server will search known IgE epitopes in query protein sequence and will define them as allergen if any portion have high similarity with any known epitope. If there is/are a known epitope(s), then mapping of the epitope(s) is done in the query sequence. In ARPs approach, the search is done in the database of 2890 allergen representative peptides (ARPs). If there is a hit, then it will be assigned as allergen and the ARP is shown in the result field. The sensitivity and specificity of this method is 66.56% and 97.97% respectively (Saha and Raghava 2006a, b).

Antigenicity Evaluation

The ability of an antigen for binding to T cell and B cell receptors is defined as an antigenicity. Therefore, it was essential to examine the designed construct vaccine in terms of the nature of the antigenicity. ANTIGENpro server was applied for antigenicity analysis of chimeric construct. ANTIGENpro is a sequence-based, alignment-free, and pathogen independent, using protein antigenicity microarray data for predicting of protein antigenicity. A two-stage architecture based on multiple representations of the primary sequence and five machine learning algorithms used for making of predictions. Antigenic or non-antigenic as well as the corresponding probability of target protein was summarized using a final SVM classifier (Cheng et al. 2005).

Docking Analysis of the Fusion Protein with TLR5

For protein–protein interaction of the flagellin section of our fusion protein as a ligand and the TLR5 as the receptor, SwarmDock server was used. SwarmDock applies the flexible modelling of protein- protein complexes using the SwarmDock algorithm, which includes a normal modes approach. The channel includes three consecutive stages: (1) pre-processing and minimization of input structures, (2) docking, using a hybrid particle swarm optimization/local search, (3) minimizing, re-ranking, and clustering of the docked poses (Torchala and Bates 2014).

Fast Simulations of Flexibility of the Docked TLR5-Vaccine Complex

CABS-flex tool was applied for fast simulation of flexibility of our complex. As input data, only the PDB format (or a protein PDB code) is required. In this tool, indicators such as protein flexibility, contact map and root-mean-square fluctuations (RMSFs) of atoms are reported in protein complex. In nanosecond time, CABS-flex server present RMSF simulation of all amino acid residues provide in a particular protein. Here, the selected docked TLR5-vaccine complex was used as the input and the number of cycles was set at 50, and Temperature range was 1.4 (Kuriata et al. 2018).

Codon Optimization

In order to clone and express the fusion protein in an appropriate expression host (*Escherichia coli (E. coli)* K12 strain) codon optimization of protein sequence was done by Java Codon Adaptation tool. In here, additional options such as rho-independent transcription terminators, prokaryotic ribosome binding sites and cleavage sites of restriction enzymes were avoided during computing. The optimized DNA sequence was introduced to the reverse complement sequence for reverse translation of optimized codon sequence. Reverse translation process confirm the complementation in the direction of the translation of the vector (Grote et al. 2005).

In Silico Cloning

The restriction sites were included to N and C-terminals of the sequence. In the current research, *E. coli* was chosen as the appropriate host. Moreover, for the cloning and expression of the optimized gene sequence in *E. coli* host, *NcoI*, *XhoI* and restriction sites were included in the N and C-terminal sites of final construct, respectively. The SnapGene tool was used to ensure the construct expression by insertion of the adapted sequence (with restriction sites) into the pET-28a (+) vector (Fig. 1).

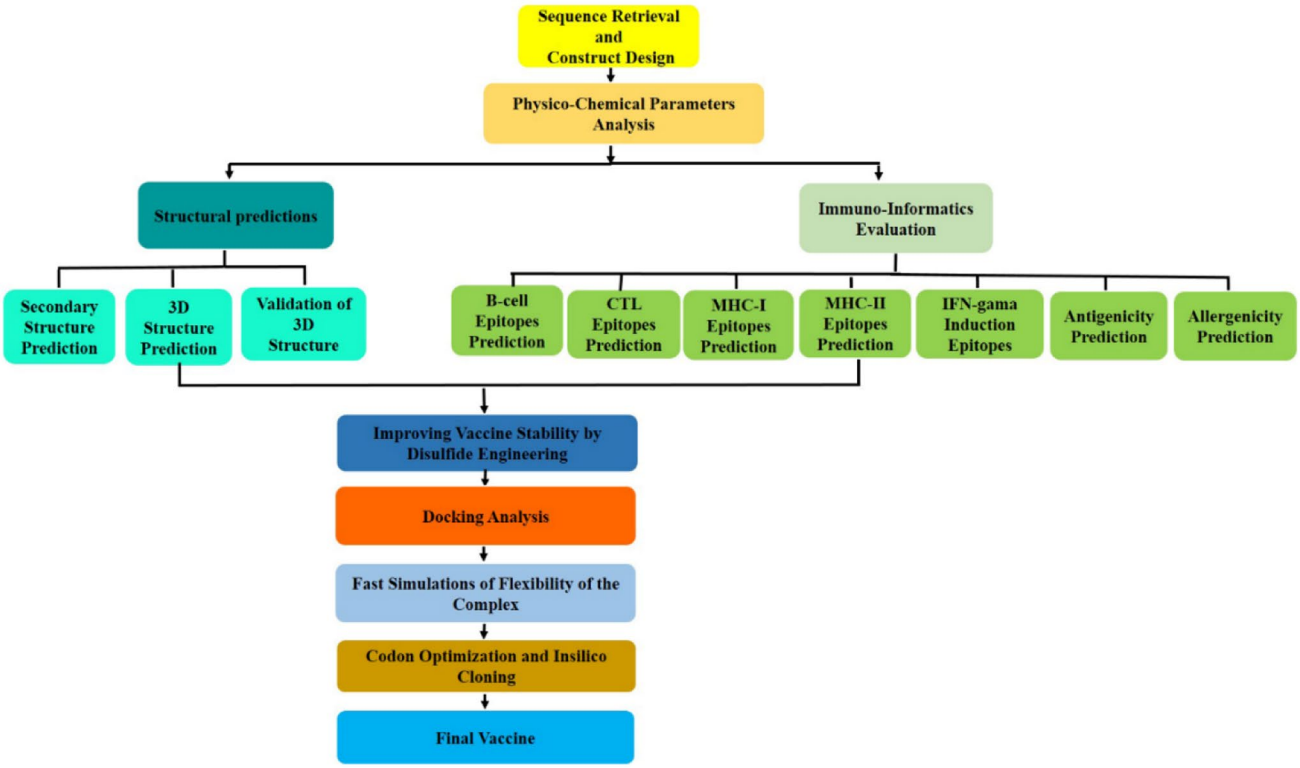

Fig. 1 A summary of the methods that were used for in silico design of fusion protein vaccine

Fig. 2 Schematic view of the final structure of the fusion protein. The first 92 amino acids are belong to KMP-11 protein sequence followed by the GSGGSGGSGGSG linker to bind to the flagellin potein sequence (with 394 amino acids)

Results

Construct Design

The KMP-11 protein sequence was joined by the GSGGSGGSGGSG linker to the fliC protein sequence (Fig. 2).

Physico-Chemical Parameters Analysis

The designed vaccine construct have 498 amino acids consisting of 48 positively (Arg, Lys) residues and 56 negatively charged (Asp, Glu) residues. The molecular weight (MW) of the described construct was about 52 kDa. The isoelectric point (pI) value was 5.40. The extinction coefficient of chimeric protein was 7450 M^{-1} cm^{-1} at 280 nm, and the estimated half-life was > 30 h (mammalian reticulocytes, in vitro), > 20 h (yeast, in vivo) and > 10 h (Escherichia coli, in vivo). Instability index was 26.51. The aliphatic index and GRAVY (Grand average of hydropathcity) of KMP-11-fliC construct were 77.73 and -0.377, respectively.

Prediction of the Secondary Structure

Based on the PSIPRED program, the secondary structure of designing vaccine was predicted (Fig. 3). The structure consists of 51.20% alpha-helix, 10.85% beta strands and 37.95% coils.

Prediction and Validation of the 3D Structure

The best predicted 3D model by I-TASSER was chosen for validation analysis (Fig. 4). The ProSA-web service was applied for the analysis of quality and potential errors in the 3D structure final selected model. Z-score of the final model was -8.91 (Fig. 5a). The result obtained from the ProSA-web showed that the selected model need no refinement and this value is in the range of acceptable scores. Ramachandran plot analysis (RAMPAGE) showed that 99.3%, 0.7%, and 0.0% of the residues were located in the favored, allowed, and outlier regions, respectively

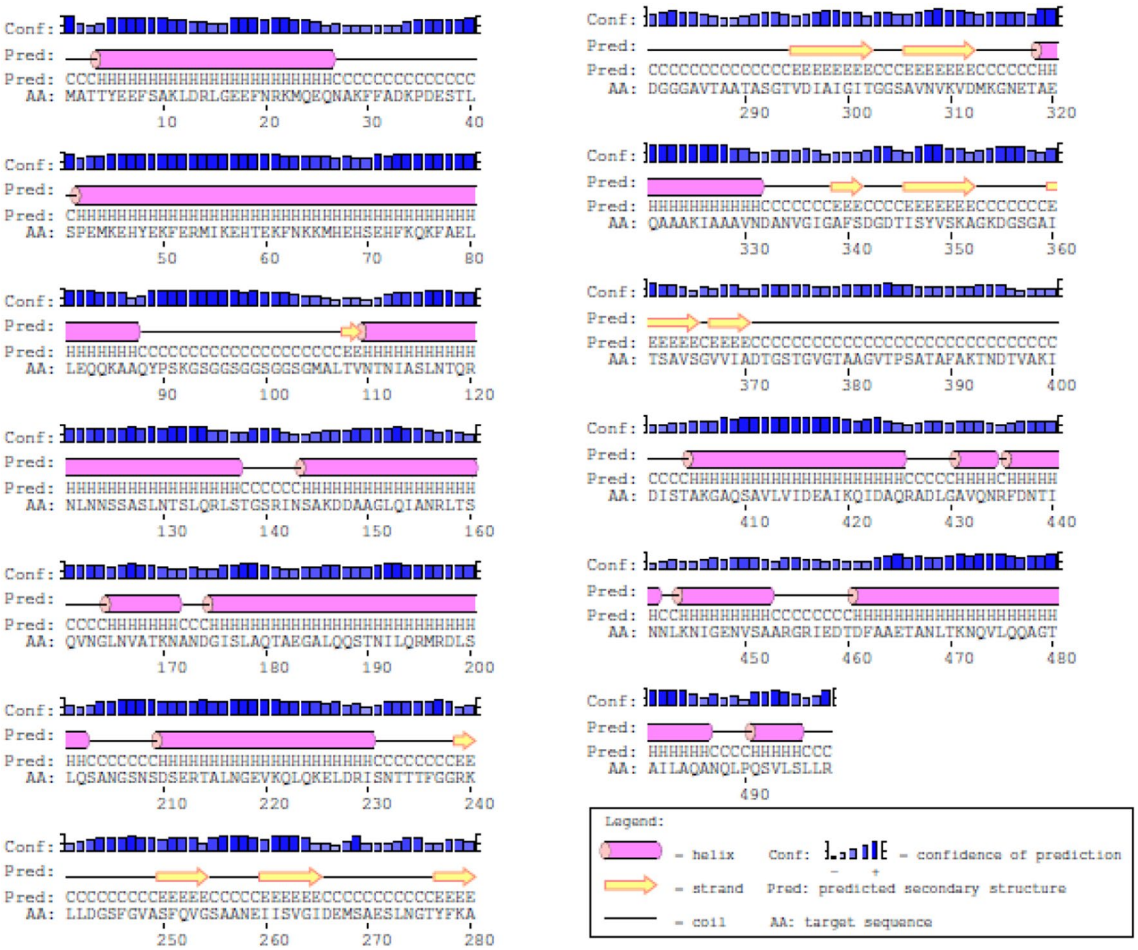

Fig. 3 The secondary structure of the fusion protein construct. The structure consists of 51.20% alpha-helix, 10.85% beta strands and 37.95% coils

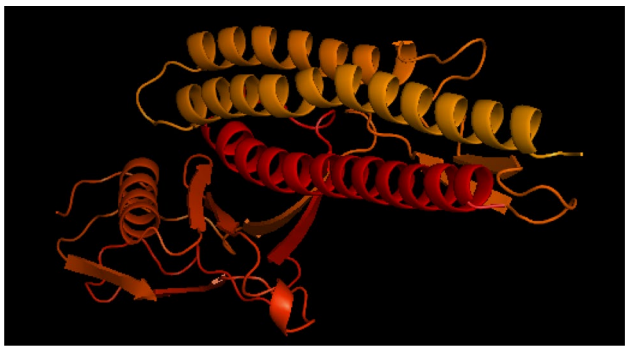

Fig. 4 The final 3D structure model of fusion protein construct obtained by I-TASSER

(Fig. 5b), which represents the high quality of the predicted structure. In the ERRAT server, the overall quality factor of selected model was 97.15% and that is illustrated (Fig. 5c).

Disulfide Engineering of Final Vaccine Construct

Disulfide by Design (DbD) resulted in 142 pairs of amino acid sites for probable disulfide bridges and only two native disulfide bridges were identified based on their high B-factor and their presence in the spanning region of high mobility. By substituting residues GLY175- PHE251, ALA185-THR234 with cysteine in the final vaccine construct, thermal stability was improved (Fig. 6a, b).

B-Cell Epitope Prediction

The high-scored linear and conformational B-cell epitopes that were predicted within the full-length of the designed vaccine by the BCPRED, ABCPred and CBTOPE, respectively (Tables 1, 2 and 3). Applying DiscoTope, conformational B-cell epitopes were predicted in the 3D model of the fusion protein (Table 4).

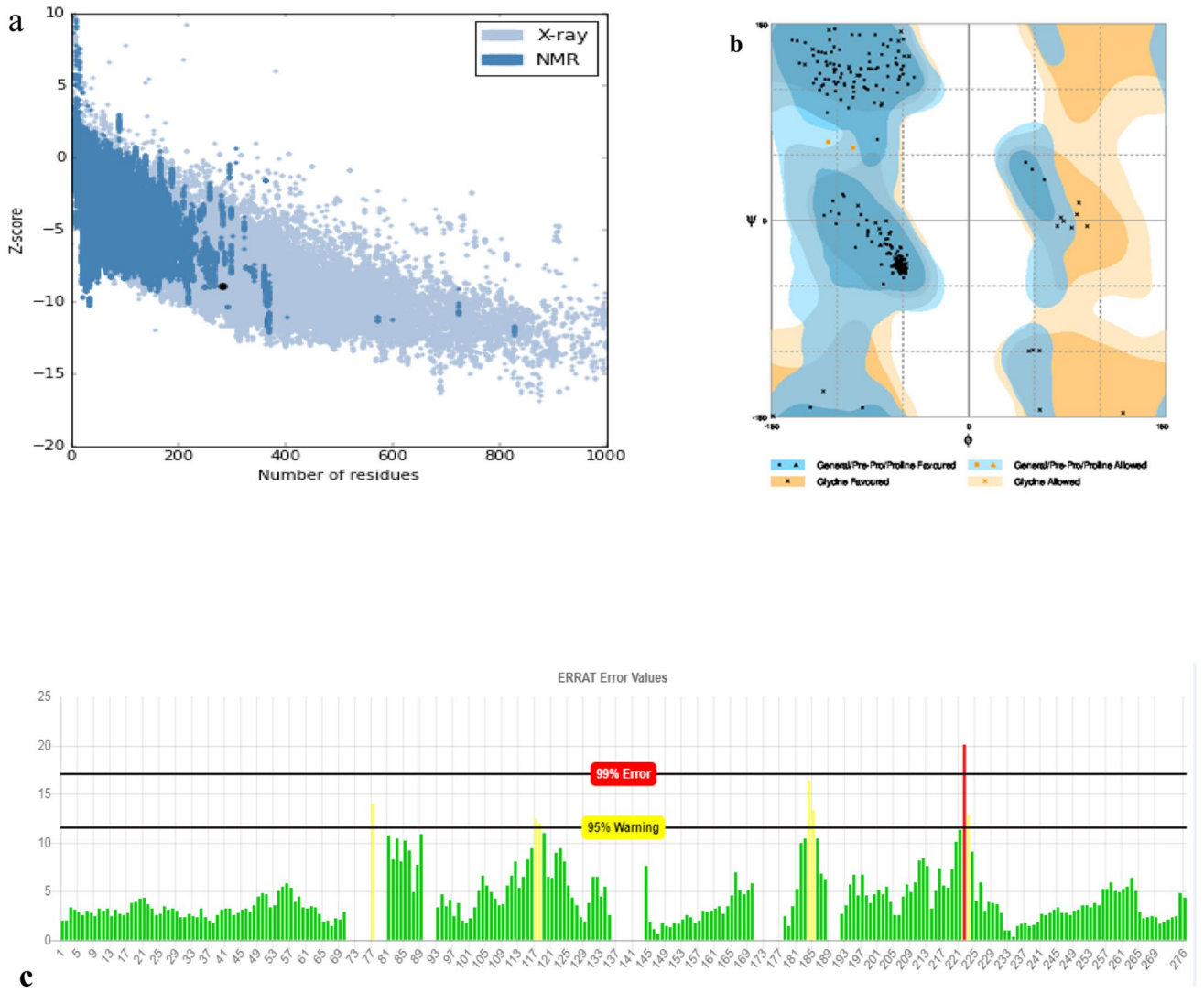

Fig. 5 The validation results of final 3D structure model of fusion protein construct. **a** ProSA-web Z-score plot for 3D structure of fusion protein construct. The Z-score of the best model is − 8.91 (shown in a large black spot), which is in range of native protein conformations. Z-score plot contains the Z-scores of all experimentally protein chains in PDB defined by NMR spectroscopy (dark blue) and X-ray crystallography (light blue). **b** The Ramachandran plot of fusion protein construct. Validation processes using Ramachandran plot displayed that, 99.3%, 0.7%, and 0.0% of residues were placed in the favored, allowed, and outlier regions, respectively. **c** ERRAT plot. The overall quality factor of the final fusion protein model is 97.15%. Values around 95% or higher usually indicate high resolution of structures

CTL Epitopes Prediction

The high-ranked CTL epitopes (9-mer length) with a binding affinity score were selected as final CTL epitopes (Table 5).

MHC-I Peptide Prediction

In this study, peptides with 9-mer length and human HLAs were chosen. Subsequently, three peptides with high binding ranks were selected including Human HLA-A (A0201).The predicted of MHC Class I binding peptides by the NetMHC 4.0 server are shown (Table 6).

MHC-II Peptide Prediction

NetMHC II Server was used for prediction of MHC-II Peptides. The peptides with high rank are shown (Table 7).

IFN-Gamma Inducing Epitopes Prediction

Prediction of the IFN-gamma inducing and non-inducing epitopes from MHC-II binding epitope fragments in the designed vaccine construct (Table 8).

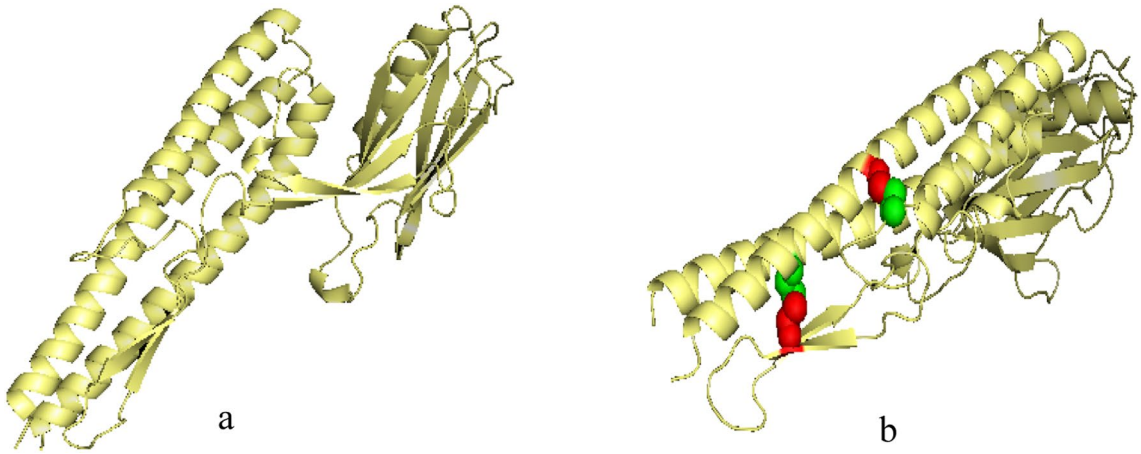

Fig. 6 Improving the stability of the fusion protein construct by Disulfide engineering. Mutated pairs are shown in green and red colour which are selected based on B-factor

Table 1 B-cell epitopes predicted by the BCPRED

Position	BCpred epitope	Score	VaxiJen
86	AAQYPSKGSGGSGGSGGSGM	1	2.5435
31	FADKPDESTLSPEMKEHYEK	0.997	0.9063
200	SLQSANGSNSDSERTALNGE	0.987	1.3220
329	AVNDANVGIGAFSDGDTISY	0.953	0.5784
378	GTAAGVTPSATAFAKTNDTV	0.924	0.5666

Table 2 List of the predicted B-cell epitopes by the ABCPred

Rank	Start position	Sequence	Score
1	103	SGMALTVNTNIASLNTQRNL	0.89
2	33	DKPDESTLSPEMKEHYEKFE	0.88
3	61	KFNKKMHEHSEHFKQKFAEL	0.87
4	336	GIGAFSDGDTISYVSKAGKD	0.87
5	16	GEEFNRKMQEQNAKFFADKP	0.87

Table 3 List of the Predicted Conformational B-cell Epitopes by the CBTOPE

Position	Amino acid	Probability scale
168	A	5
182–183	AE	5
221	K	5
310–313	KVDM	5
465–467	ETA	5

Allergenicity and Antigenicity Properties

Based on the results from AlgPred server, the designed construct is not allergenic. With the score of − 1.06 whereas the default threshold was − 0.4. The antigenicity of chimeric protein was estimated 0.92% by the ANTIGENpro. The obtained data indicated that with a high probability, the construct is antigenic in nature.

Docking Studies

Docking analysis of the final designed construct model with TLR5 (PDB ID- 3J0A) was done by SwarmDock server. In order to select the best-docked complex, the best ten docked models presented by Swarmdock were selected for further analysis. Finally, model number 5 of docked complex showed best interaction (Fig. 7).

Flexibility of the Docked TLR5-Designed Vaccine Complex

In this work, we have assessed the flexibility of the designed vaccine by CABS- Flex tool with 50 cycles' simulation at 1.4 °C temperature. Our complex (TLR5-designed vaccine) gained a high level of fluctuations in the residue positions 444, 20 and 590, were recognized to be 5.09 Å, 3.88 Å and 3.74 Å, respectively (Fig. 8).

In addition, based on parameters such as structural heterogeneity, their optimum free energy, and highly stable configuration, CABS-Flex server offers 10 different models that the first model chosen, finally. Stable protein complex after flexibility fast simulation using CABS-Flex was shown (Fig. 9).

Table 4 Results obtained from the DiscoTope 2.0 server (Conformational B-Cell Epitopes)

Amino acid	Residue number	Contact number	Propensity score	Discotop score
PHE	51	7	8.506	6.723
ARG	53	7	8.574	6.783
ILE	55	1	8.087	7.042
GLU	60	9	9.221	7.126
LYS	61	1	8.518	7.424

Codon Optimization and in Silico Cloning

Reverse translation and codon optimization of fusion protein were performed by Java Codon Adaptation tool, in order to lead to the high-level protein expression in *E. coli* K12. A CAI of our optimized nucleotide sequence was 0.9630; a CAI of > 0.8 is remarked as good for expression in the selected host. CG content between 30 and 70% was considered as the optimum range. Any regions out of this range have a negative impact on the efficiency of gene expression. The mean CG content of our sequence is 49.66%. Finally, in order to clone the gene in pET28a (+) vector by SnapGene tool, the restriction sites *Nco*I and *Xho*I were introduced to the N and C-terminals of the sequence, respectively (Fig. 10).

Discussion

Although various chemotherapy and drug therapy exist for treating visceral leishmaniasis (VL) but they are costly and parasites resistant to first line drugs are observed, for these reasons vaccines are a standout amongst the best strategies for eliminating or controlling the disease (Srivastava, Shankar et al. 2016; Ponte-Sucre et al. 2017; Ghorbani and Farhoudi 2018).

Development and production of vaccine is experimentally cost effective, while the advancements in molecular immunology and the identification of dominant immune epitopes have been making fundamental changes in these experimental methods (Silva-Jardim et al. 2014; Sundar et al. 2019; Oli et al. 2020). Advanced tools based on bioinformatics and immunoinformatic methods are used to develop and design a vaccine that reduces the negative aspects of empirical methods, including preventing predictable ethical aspects of laboratory works and expensive experimental studies and being time consuming. In order to control VL, it is essential to stimulate both humoral and particularly cellular immune responses (Poland and Oberg 2010; Ghorbani and Farhoudi 2018; Farhani et al. 2019; Ahmadi et al. 2020; Ikeogu, Akaluka et al. 2020; Mahboobi, Sedighian et al. 2020; Oli, Obialor et al. 2020). The humoral immune response, providing antibodies, is an essential part of the design of vaccines. To this end, identifying reliable B cell epitopes from protein sequence can be helpful (Irvine and Read 2020). Because visceral leishmaniasis (VL) is an intracellular parasite, cytotoxic T-cell lymphocyte (CTL) can play an important role in regulating a protective response to it; therefore, identification of CTL epitopes can help to develop cell-mediated immunity (Stager and Rafati 2012; Agallou, Margaroni et al. 2020). Bioinformatics software's are able to provide useful information about the characteristics of the candidate vaccine as well as the improvement

Table 5 CTL epitopes predicted by the NetCTL-1.2 server

Residue number	Peptide sequence	Predicted MHC binding affinity	C terminal cleavage affinity	Prediction score
340	FSDGDTISY	0.7344	0.9693	3.4059
269	SAESLNGTY	0.5721	0.9750	2.7214
40	LSPEMKEHY	0.3359	0.8369	1.6999
69	HSEHFKQKF	0.2104	0.7381	1.1128
208	NSDSERTAL	0.1843	0.5796	0.9149

Table 6 MHC-I Peptide Predicted by the NetMHC 4.0 Server

Position	HLA	Peptide sequence	Bind level	Affinity(nM)	%Rank
239	HLA-A0201	KLLDGSFGV	(SB)*	2.87	0.01
339	HLA-A0101	FSDGDTISY	(SB)	7.51	0.01
2	HLA-A1101	TTYEEFSAK	(SB)	10.52	0.03

*Strong Binder

Table 7 MHC-II peptide predicted by the NetMHC II server

Position	HLA	Peptide sequence	Bindlevel	Affinity(nM)	%Rank
190	DRB4_0101	TNILQRMRDLSLQSA	(SB)*	22.3	0.70
191	DRB4_0101	NILQRMRDLSLQSAN	(SB)	23.5	0.80
189	DRB4_0101	STNILQRMRDLSLQS	(SB)	27.7	1.00
417	DRB4_0101	EAIKQIDAQRADLGA	(SB)	31.8	1.30

*Strong Binder

Table 8 IFN-gamma inducing and non-inducing epitopes predicted by IFNepitope sever

Result	Position	Epitope	Score	Method
Positive	Epitope_1	TNILQRMRDLSLQSA	0.036794699	SVM
Positive	Epitope_2	STNILQRMRDLSLQS	0.021853617	SVM
Negative	Epitope_3	EAIKQIDAQRADLGA	−0.11013804	SVM
Negative	Epitope_4	NILQRMRDLSLQSAN	−0.14195861	SVM

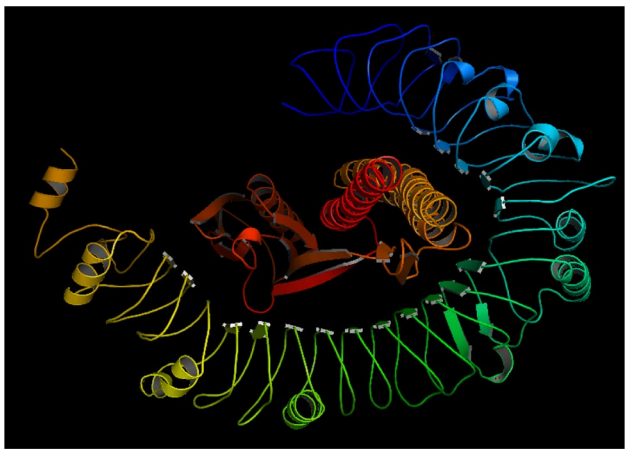

Fig. 7 Docking model of the TLR5 protein and fusion protein obtained by the Swarmdock server. TLR5 molecule is shown in green–blue. The final 3D model of fusion protein construct is shown in red–orange

of its properties (Poland and Oberg 2010; María, Arturo et al. 2017). In the recent decades, the genomic and proteomic information of *leishmania* and other pathogenic microbes is available in bulk. The information from the data is helpful to identify potential antigenic targets using computational approaches. According to recent advances in immunobiology and computer sciences, researchers are able to design the efficient vaccines, particularly in the field of infectious diseases via computational methods (Groot and Rappuoli 2004; Banuls, Hide et al. 2007). In the present study, we intend to use bioinformatics and structural vaccinology approaches to design a fusion protein vaccine consists of *P. aeruginosa* fliC (pafliC) and KMP-11 protein. It is very important to select an appropriate linker for the fusion proteins, especially in terms of maintaining the biological activity of the domains. In the current study, a flexible linker with the sequence of (GSGGSGGSGGSG) (12aa) was applied for separation of the two moieties of the

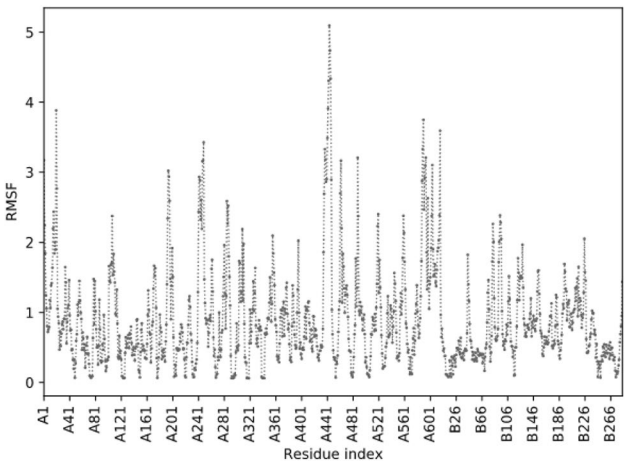

Fig. 8 Root Mean Square Fluctuation (RMSF) plot of the protein complex during the simulation, indicating the fluctuations of MEV residues. A high level of fluctuation in the residue positions 444, 20 and 590, were recognized to be 5.09 Å, 3.88 Å and 3.74 Å, respectively

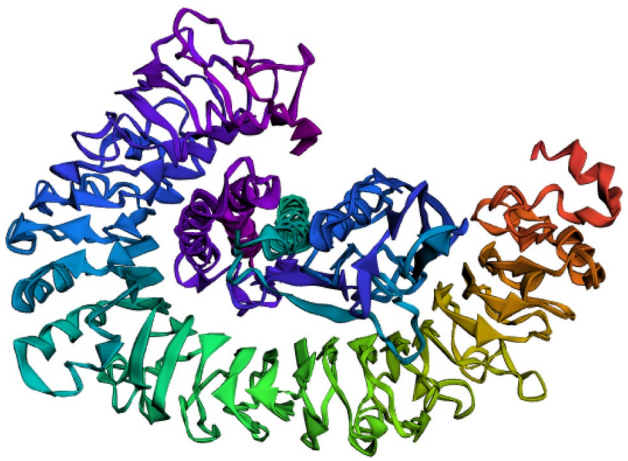

Fig. 9 Stable protein complex structure after flexibility fast simulations using CABS-Flex server

chimeric protein. These types of linkers are suitable, as each domain of the fusion protein requires certain motions or interactions (Zhang et al. 2009; Chen, Zaro et al. 2013). The physicochemical parameters of our chimeric construct were evaluated by using ProtParam. At the isoelectric point, the net charge of the protein is zero, since the positive and negative charges are equal. The calculated pI index of chimeric protein indicated that our vaccine protein is acidic in nature. This indicator can be used to develop buffer systems for protein purification by isoelectric focusing method. The instability index provides an estimate of the stability of the protein in a test tube. A protein with the instability index smaller than 40 is predicted as stable, a value above 40 predicts that the protein may be unstable. The instability index showed that our chimeric protein is a stable protein. Aliphatic index shows that the protein is occupied by aliphatic side chains. Grand average of hydropathicity (GRAVY) shows that vaccine protein is hydrophilic in nature and seems to have a better interaction with water**.** PSIPRED tool is a popular and highly accurate method been used in this research for evaluation of vaccine secondary structure. The biological function of the recombinant proteins is influenced by their three-dimensional structures. Therefore, the details of the 3D structure of our protein can be useful in the study of protein function, dynamics, interaction with ligand and other proteins. Therefore, different tools such as SWISS-MODEL, I-TASSER and Phyre2 were used for 3D modeling of our fusion construct. In this regard, the structure validation tools were used to recognize errors in the 3D modeled structures of vaccine construct. The quality of all fusion construct models was determined by ERRAT, Ramachandran plot and ProSA-Web tools. Based on the validation results, the highest quality 3D structure model of chimeric protein (KMP-11-fliC) was obtained using I-TASSER server as the final model. In the case of the final model, ramachandran plot shows the most of the residue clustered tightly in the most favored regions with very few residues in outliers which depicted that the overall model quality is satisfactory. The result obtained from the ProSA-web showed that the final model was in the range of acceptable scores and the result obtained from ERRAT showed that, the final model has the best overall quality factor. Therefore, validation results indicated that the final model does not require a refinement process. Enhancing the stability of proteins is an imperative target in numerous biomedical and mechanical applications (Khatoon, Pandey et al. 2017). In this study, Disulfide by design v2.0 tool was used in order to improve the thermostability of the final vaccine construct. The increasing stability of protein after disulfide engineering is due to the reduced conformational entropy of unfolded protein state. B and T cells play an important role in induction of a protective immune response in VL; thus, determination of peptides inducing T and B cell responses is a crucial requirement for the design of effective epitope-based vaccines (Khan, Ami et al. 2020). The linear and conformational B-cell epitopes of the designed chimeric protein were analyzed using BCPRED, ABCPred, DiscoTope and CBTOPE servers. The NetCTL-1.2 server was used for prediction of CTL epitopes in our construct. One of the crucial stages in the design and development of a vaccine candidate is the prediction of antigenic epitopes that have a potent binding affinity to different HLA alleles. So that the interaction between epitopes and MHC molecules can cause cellular immunity (Ka, Narsaria et al. 2020; Mahapatra, Sahoo et al. 2020). Therefore, in the current research, NetMHC 4.0 and NetMHC II servers were used to predict immunodominant MHC-I binding epitopes and MHC-II binding epitopes,

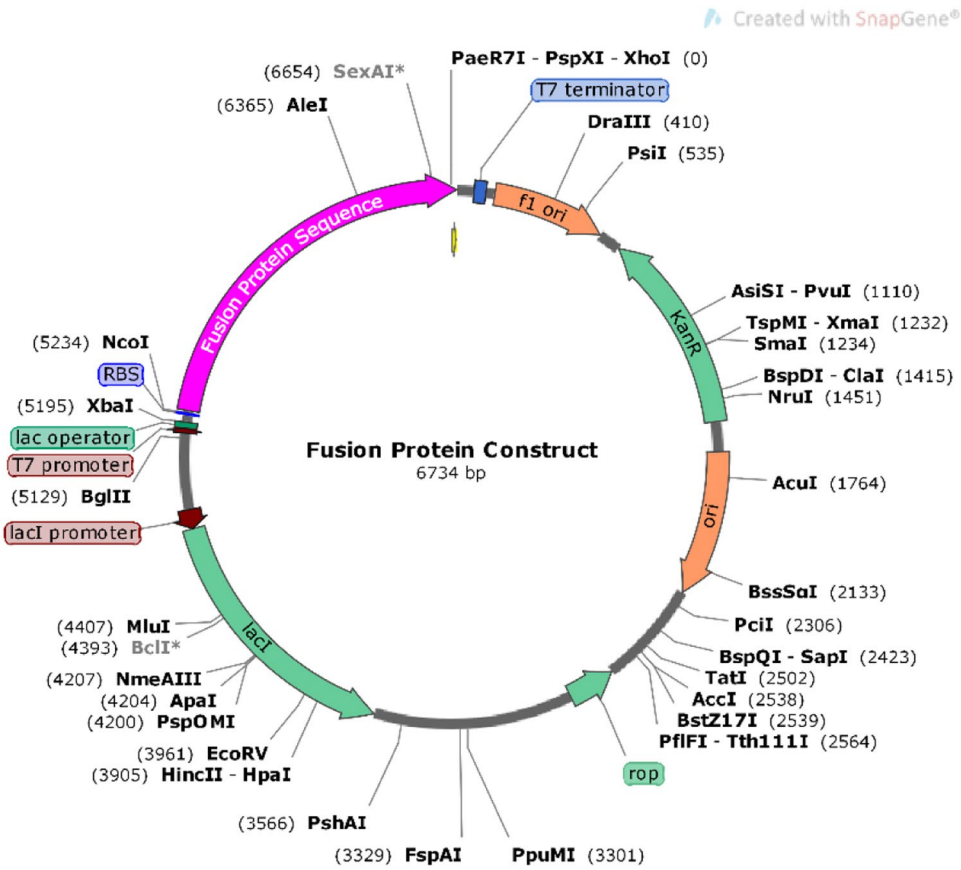

Fig. 10 In silico cloning of fusion protein sequence into pET28a (+) vector. Purple colored semicircle showing fusion protein sequence and Green-colored semicircles indicating backbone of pET28a (+) vector

respectively. The cell response is important in the control of infection, given that the development of a specific T helper (Th) 1 response, based on the production of cytokines, such as interferon-gamma (IFN-γ), can protect mammalian hosts from infection by the parasites. VL exist as obligate intra-macrophage amastigotes in the mammalian host. Activation of macrophages is required for the elimination of the organism in vivo, and IFN-γ has been identified as the major macrophage-activating factor mediating defense against VL (Ikeogu, Akaluka et al. 2020). The ten IFN-γ inducing epitopes presented the ability of our vaccine to incite responses efficiently. Immunoinformatics analyses of KMP-11-fliC fusion protein indicated that our fusion construct is a robust antigen and non-allergenic. In this study, docking analysis between fliC protein of our fusion protein construct as TLR5 agonist and TLR5 receptor was performed via the Swarmdock server. Finally, the best complex model according to hydrophobicity, electrostatic complementarity of protein surfaces was chosen among them. Furthermore, PyMOL and UCSF Chimera tools were used for analysis and visualization of the best model of fusion protein-TLR5 complex. In this study, CABS-Flex 2.0 software was used for flexibility analysis of our docked complex. CABS-Flex introduces the stable arrangement of the TLR5-designed vaccine complex. The highest RMSF value and the lowest value show more fluctuation and low fluctuation of our complex structure during the simulation process, respectively. Fluctuations in the structure of our complex illustrate its high flexibility and validate it as a potential structure of the vaccine. Based on the root mean square fluctuation (RMSF) values, the fluctuation of the individual amino acid residues as reported. Finally, in silico cloning process of the chimeric protein was performed for efficient protein expression in the *E. coli* host. In order to achieve a high-level of recombinant vaccine protein expression in *E. coli* (strain K12), codon optimization was carried out to improve transcriptional and translational efficiency. This was accomplished by analyzing codon adaptation index (CAI) and total GC content of DNA sequence. The CAI and GC indices showed that our fusion construct could be well expressed in the *E. coli* host. Therefore, in this study, we have integrated novel immunoinformatics tools to design a potent, safeand immunogenic subunit vaccine, which may have the ability to control of VL.

Conclusions

One of the strategies for dealing with visceral leishmaniasis (VL) is to activating potent immune responses against the infection. In this study, we designed a fusion protein and

evaluated it through immunoinformatics approaches. The results of the in silico evaluations showed that the designed chimeric protein, KMP-11-fliC, can be applied as a candidate vaccine against VL in the future. Furthermore, the proposed vaccine needs to be experimentally validated to ensure the control of VL by generating an effective immunological memory.

Author Contributions All the authors contributed equally to the work reported. All authors have read and agreed to the published version of the manuscript.

Declarations

Conflict of interest The authors declare that they have no conflict of interest.

References

Agallou M et al (2020) A canine-directed chimeric multi-epitope vaccine induced protective immune responses in BALB/c Mice infected with leishmania infantum. Vaccines 8(3):350

Ahmadi F et al (2020) In silico design of epitope-based allergy vaccine against bellatella germanica cockroach allergens. Int J Pept Res Ther 26(4):1739–1749

Andreatta M, Nielsen M (2016) Gapped sequence alignment using artificial neural networks: application to the MHC class I system. Bioinformatics 32(4):511–517

Ansari HR, Raghava GP (2010) Identification of conformational B-cell Epitopes in an antigen from its primary sequence. Immunome Res 6(1):6

Atapour A et al (2017) Molecular cloning, expression, and purification of a recombinant fusion protein (rNT-gp96-NT300). BioPharm Int 30:38–44

Atapour A et al (2019) Designing a fusion protein vaccine against HCV: an in silico approach. Int J Pept Res Ther 25(3):861–872

Atapour A et al (2020) In silico designing a candidate vaccine against breast cancer. Int J Pept Res Ther 26(1):369–380

Banuls A-L et al (2007) Leishmania and the leishmaniases: a parasite genetic update and advances in taxonomy, epidemiology and pathogenicity in humans. Adv Parasitol 64:1–458

Basu R et al (2005) Kinetoplastid membrane protein-11 DNA vaccination induces complete protection against both pentavalent antimonial-sensitive and-resistant strains of Leishmania donovani that correlates with inducible nitric oxide synthase activity and IL-4 generation: evidence for mixed Th1-and Th2-like responses in visceral leishmaniasis. J Immunol 174(11):7160–7171

Buchan DW, Jones DT (2019) The PSIPRED protein analysis workbench: 20 years on. Nucleic Acids Res 47(W1):W402–W407

Chappuis F et al (2007) Visceral leishmaniasis: what are the needs for diagnosis, treatment and control? Nat Rev Microbiol 5(11):S7–S16

Chen X et al (2013) Fusion protein linkers: property, design and functionality. Adv Drug Deliv Rev 65(10):1357–1369

Cheng, J., et al. (2005). "SCRATCH: a protein structure and structural feature prediction server." Nucleic acids research 33(suppl_2): W72-W76.

Coffman RL et al (2010) Vaccine adjuvants: putting innate immunity to work. Immunity 33(4):492–503

Colovos C, Yeates T (1993) ERRAT: an empirical atom-based method for validating protein structures. Protein Sci 2(9):1511–1519

Craig DB, Dombkowski AA (2013) Disulfide by Design 2.0: a web-based tool for disulfide engineering in proteins. BMC Bioinformatics 14(1):1–7

Cui B et al (2019) Immunogenicity and protective efficacy of recombinant proteins consisting of multiple epitopes of foot-and-mouth disease virus fused with flagellin. Appl Microbiol Biotechnol 103(8):3367–3379

de Mendonça SCF et al (2015) Kinetoplastid membrane protein-11 as a vaccine candidate and a virulence factor in Leishmania. Front Immunol 6:524

Dhanda SK et al (2013) Designing of interferon-gamma inducing MHC class-II binders. Biol Direct 8(1):30

EL-Manzalawy Y et al (2008) Predicting linear B-cell epitopes using string kernels. J Mol Recogni: An Interdisciplinary Journal 21(4):243–255

Farhani I et al (2019) Designing a novel multi-epitope peptide vaccine against pathogenic Shigella spp. based immunoinformatics approaches. Int J Pept Res Ther 25(2):541–553

Gasteiger E et al (2005) Protein identification and analysis tools on the ExPASy server. Springer, The proteomics protocols handbook, pp 571–607

Ghorbani M, Farhoudi R (2018) Leishmaniasis in humans: drug or vaccine therapy? Drug Des Dev Ther 12:25

Groot ASD, Rappuoli R (2004) Genome-derived vaccines. Expert Rev Vaccines 3(1):59–76

Grote A et al (2005) JCat: a novel tool to adapt codon usage of a target gene to its potential expression host. Nucleic Acids Res 33(2):526–531

Heydarpour F et al (2016) Incidence and disability-adjusted life years (Dalys) attributable to leishmaniasis in Iran, 2013. Ethiop J Health Sci 26(4):381–388

Ikeogu NM et al (2020) Leishmania immunity: advancing immunotherapy and vaccine development. Microorganisms 8(8):1201

Irvine DJ, Read BJ (2020) Shaping humoral immunity to vaccines through antigen-displaying nanoparticles. Curr Opin Immunol 65:1–6

Jain K, Jain N (2015) Vaccines for visceral leishmaniasis: A review. J Immunol Methods 422:1–12

Jensen KK et al (2018) Improved methods for predicting peptide binding affinity to MHC class II molecules. Immunology 154(3):394–406

Joshi S et al (2014) Visceral leishmaniasis: advancements in vaccine development via classical and molecular approaches. Front Immunol 5:380

Ka, T., et al. (2020). "A Candidate multi-epitope vaccine against SARS-CoV-2."

Kaye P, Aebischer T (2011) Visceral leishmaniasis: immunology and prospects for a vaccine. Clin Microbiol Infect 17(10):1462–1470

Kevric I et al (2015) New world and old world Leishmania infections: a practical review. Dermatol Clin 33(3):579–593

Khamesipour A et al (2006) Leishmaniasis vaccine candidates for development: a global overview. Indian J Med Res 123(3):423

Khan MAA et al (2020) An immunoinformatic approach driven by experimental proteomics: in silico design of a subunit candidate vaccine targeting secretory proteins of Leishmania donovani amastigotes. Parasit Vectors 13:1–21

Khatoon N et al (2017) Exploring Leishmania secretory proteins to design B and T cell multi-epitope subunit vaccine using immunoinformatics approach. Sci Rep 7(1):1–12

Kringelum JV et al (2012) Reliable B cell epitope predictions: impacts of method development and improved benchmarking. PLoS Comput Biol 8(12):e1002829

Kumar R, Engwerda C (2014) Vaccines to prevent leishmaniasis. Clini Translat Immunol 3(3):e13

Kuriata A et al (2018) CABS-flex 2.0: a web server for fast simulations of flexibility of protein structures. Nucleic Acids Res 46(W1):W338–W343

Larsen MV et al (2007) Large-scale validation of methods for cytotoxic T-lymphocyte epitope prediction. BMC Bioinformatics 8(1):424

Lockner JW et al (2015) Flagellin as carrier and adjuvant in cocaine vaccine development. Mol Pharm 12(2):653–662

Lovell, S. C., et al. (2003) Structure validation by Cα geometry ϕ, ψ and Cβ deviation. Proteins: Structure, Function, and Bioinformatics 50(3): 437–450.

Lu Y, Swartz JR (2016) Functional properties of flagellin as a stimulator of innate immunity. Sci Rep 6(1):1–11

Mahapatra SR et al (2020) Designing an efficient multi-epitope vaccine displaying interactions with diverse HLA molecules for an efficient humoral and cellular immune response to prevent COVID-19 infection. Expert Rev Vaccines 19(9):871–885

Mahboobi, M., et al. (2020). "Harnessing an Integrative In Silico Approach to Engage Highly Immunogenic Peptides in an Antigen Design Against Epsilon Toxin (ETX) of Clostridium perfringens." International Journal of Peptide Research and Therapeutics: 1–8.

Makvandi M et al (2018) Expression of Salmonella typhimurium and Escherichia coli flagellin protein and its functional characterization as an adjuvant. Microb Pathog 118:87–90

María R et al (2017) The impact of bioinformatics on vaccine design and development. Vaccines, InTech, Rijeka, Croatia

Mizel SB et al (2003) Induction of macrophage nitric oxide production by Gram-negative flagellin involves signaling via heteromeric Toll-like receptor 5/Toll-like receptor 4 complexes. J Immunol 170(12):6217–6223

Nasiri V et al (2016) Immunogenicity and efficacy of live L. tarentolae expressing KMP11-NTGP96-GFP fusion as a vaccine candidate against experimental Visceral Leishmaniasis caused by L. infantum. Iran J Parasitol 11(2):144

Nylén S, Kumar R (2012) Immunobiology of visceral leishmaniasis. Front Immunol 3:251

Okwor I, Uzonna J (2009) Vaccines and vaccination strategies against human cutaneous leishmaniasis. Hum Vaccin 5(5):291–301

Oli AN et al (2020) Immunoinformatics and vaccine development: an overview. ImmunoTargets and therapy 9:13

Poland GA, Oberg AL (2010) Vaccinomics and bioinformatics: accelerants for the next golden age of vaccinology. Vaccine 28(20):3509

Polonio T, Efferth T (2008) Leishmaniasis: drug resistance and natural products. Int J Mol Med 22(3):277–286

Ponte-Sucre A et al (2017) Drug resistance and treatment failure in leishmaniasis: A 21st century challenge. PLoS Negl Trop Dis 11(12):e0006052

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

Saha, S. and G. Raghava (2006). "AlgPred: prediction of allergenic proteins and mapping of IgE epitopes." Nucleic acids research 34(suppl_2): W202-W209.

Saha S, Raghava GPS (2006) Prediction of continuous B-cell epitopes in an antigen using recurrent neural network. Proteins: Struct, Funct, Bioinf 65(1):40–48

Senevirathne, A., et al. (2020) Parenteral immunization of Salmonella Typhimurium ghosts with surface-displayed Escherichia coli flagellin enhancesTLR-5 mediated activation of immune responses that protect the chicken against Salmonella infection." Microbial pathogenesis: 104252.

Silva-Jardim I et al (2014) Leishmaniasis and Chagas disease chemotherapy: a critical review. J Braz Chem Soc 25(10):1810–1823

Srivastava S et al (2016) Possibilities and challenges for developing a successful vaccine for leishmaniasis. Parasit Vectors 9(1):277

Stager S, Rafati S (2012) CD8+ T cells in Leishmania infections: friends or foes? Front Immunol 3:5

Steinhagen F et al (2011) TLR-based immune adjuvants. Vaccine 29(17):3341–3355

Sundar S et al (2019) Leishmaniasis: treatment, drug resistance and emerging therapies. Expert Opin Orphan Drugs 7(1):1–10

Torchala M, Bates PA (2014) Predicting the structure of protein–protein complexes using the SwarmDock web server. Springer, Protein structure prediction, pp 181–197

Tripathi P et al (2007) Immune response to leishmania: paradox rather than paradigm. FEMS Immunol Med Microbiol 51(2):229–242

Vakili B et al (2018) Immunoinformatics-aided design of a potential multi-epitope peptide vaccine against Leishmania infantum. Int J Biol Macromol 120:1127–1139

Wangkahart E et al (2019) Studies on the use of flagellin as an immunostimulant and vaccine adjuvant in fish aquaculture. Front Immunol 9:3054

Wiederstein, M. and M. J. Sippl (2007). "ProSA-web: interactive web service for the recognition of errors in three-dimensional structures of proteins." Nucleic acids research 35(suppl_2): W407-W410.

Zhang J et al (2009) Design and optimization of a linker for fusion protein construction. Prog Nat Sci 19(10):1197–1200