

Research Article

Immunoinformatics Prediction of Epitope Based Peptide Vaccine against *Clostridium perfringens* Fructose Bisphosphate Aldolase Protein

Mustafa Elhag^{1,6*}, Tebyan Ameer Abdelhameed², Moaaz Mohammed Saadaldin^{2,3}, Abdelrahman Hamza Abdelmoneim^{2,4}, Taha Salah Taha^{2,3}, Israa Abdulrahman Mohammed^{2,5}, Mohammed A.Hassan^{2,6}

¹Department of Medicine, University of Seychelles-American Institute of Medicine, Anse Royale, Seychelles;²Department of Technology, Al-Neelain University, Sudan;³Department of Biotechnology, Omdurman Islamic University, Sudan;⁴Department of Medicine, Al-Neelain University, Sudan;⁵Department of Medical Laboratory Sciences, University of Science and Technology, Sudan;⁶Department of Bioinformatics, Detagen Genetics Diagnostic Center, Kayseri, Turkey

ABSTRACT

Clostridium perfringens is a member of the gastrointestinal tract (GIT) microbial community for both diseased and healthy humans and animals. Also, this bacterium is responsible for around 5%–15% of all circumstances of antibiotic-associated diarrhea, which develops in 5%–40% of all patients receiving antibiotic therapy. In addition, it causes enteritis necroticans; an often-fatal human disease. C. perfringens is clear and defined the underlying factors responsible for specific aspects of pathology remains uncertain. This study predicts an effective epitope-based vaccine against fructose 1,6-biphosphate aldolase (FBA) enzyme of *Clostridium perfringens* using immunoinformatics tools. The sequences were retrieved from NCBI and several prediction tests were conducted to analyze possible epitopes for B-cell, T-cell MHC class I and II. Tertiary structure of the most promising epitopes was obtained. 48 epitopes showed high binding affinity for B-cells, while five epitopes showed high binding affinity for MHC I and MHC II. The results were promising to formulate a vaccine with more than 98% population coverage. We hope that these promising epitopes serves as a preventive measure for the disease in the future and recommend in vivo and in vitro studies.

Keywords: Immunoinformatics; Clostridium perfringens, fructose-1,6-bisphosphate aldolase; Peptide vaccine; Epitope

INTRODUCTION

Clostridium perfringens is a gram positive, rod-shaped, sporeforming, anaerobic bacterium [1-6] associated with miscellaneous environments including: soils, food, sewage, and is regarded as a member of the gastrointestinal tract (GIT) microbial community of both diseased and healthy humans and animals [1,3,4,6] with various significant systemic and enteric diseases, in both humans and animals [1,3-8]. C. perfringens is the principal cause of traumatic gas gangrene. It is also considered as a major cause of food-borne illness, classified as the second most common bacterial cause of food poisoning in the USA. Also, this bacterium is responsible for around 5-15% of all circumstances of antibiotic-associated diarrhea, which develops in 5%-40% of all patients receiving antibiotic therapy. In addition, it causes enteritis necroticans; an often-fatal human disease. As an animal pathogen, C. perfringens is responsible for several serious diseases, including necrotic enteritis (NE). Furthermore, widespread vaccination is practiced to protect livestock from C. perfringens-induced enteritis and enterotoxaemia, the latter characterized by intestinally produced toxins that are absorbed into the circulation and then affect other organs such

as the brain [4]. Despite that, C. *perfringens* is clear and defined the underlying factors responsible for specific aspects of pathology remains uncertain. Decoding the genes that are involved in the virulence factors could lead to more targeted clinical preventions in C. *perfringens*-associated intestinal diseases, whether it is humans or animals [1,5]. Significantly, C. *perfringens* strains are known to secrete more than twenty identified toxins or enzymes that could possibly be the principal virulence factors involved in the pathophysiology [1]. Fructose-1,6-bisphosphate aldolase may function as an adhesion enzyme that may serve an important step in C. *perfringens* pathogenesis [9].

Fructose-1,6-bisphosphate aldolase (FBA) catalyzes the cleavage of fructose 1,6-bisphosphate (FBP) into glyceraldehyde 3-phosphate and dihydroxyacetone phosphate (DHAP) or the reverse aldol condensation reaction. This is an essential enzyme for glycolysis, gluconeogenesis and the Calvin cycle. FBAs can be divided into two groups with different catalytic mechanisms, named Class-I and Class-II respectively; Class-I FBAs utilize an active site lysine residue to stabilize a reaction intermediate via Schiff-base formation, and mainly found in animals, plants and green algae and occasionally in

Correspondence to: Mustafa Elhag, Faculty of Medicine, University of Seychelles-American Institute of Medicine, Anse Royale, Seychelles, E-mail: mustafa.alfakih@gmail.com

Received: February 11, 2021; Accepted: February 26, 2021; Published: March 05, 2021

Citation: Elhag M, Abdelhameed TM, Saadaldin MM, Abdelmoneim AH, Taha TS, Mohammed IA, et al. (2021) Immunoinformatics Prediction of Epitope Based Peptide Vaccine against Clostridium perfringens Fructose Bisphosphate Aldolase Protein. Immunome Res. 17:184.

Copyright: © 2021 Elhag M, et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

bacteria. Class-II FBAs have an absolute requirement for a divalent ion, usually zinc and mainly found in bacteria. Most organisms contain only one class of FBAs, although a small number possess enzymes of both classes [10,11]. In recent studies, FBA not only contributed to energy production in the glycolytic pathway but also had non-glycolytic functions. It is exciting that glycolytic enzymes can be located at the surface of bacteria, where these enzymes can exert functions associated with virulence. Due to their functions in virulence, glycolytic enzymes may be used as vaccine candidates. An advantage of the glycolytic enzymes as vaccines is that they are highly conserved among different pathogens, which is a premise for broad-spectrum vaccine development. The fact that the FBA Class-II of prokaryotes has little homology with the FBA Class-I of eukaryotes provided an advantage for a possible therapeutic or vaccine target [10-12].

FBA was firstly identified as one of six proteins that could be used in immunization of some animal against necrotic enteritis and tested to be found significantly protect against it [13-15]. In Mahmood study, they found that the antigenic proteins (including FBA) and alpha-toxin conferred significant protection to broiler chicks against mild and severe infections with virulent C. perfringens. Besides alpha-toxin immunization which imparted the highest protection to severely challenged birds, certain proteins secreted can play a role in immunizing broiler chickens against necrotic enteritis [16]. On other hand, Wilde study mentioned that there is currently no necrotic enteritis vaccine commercially available for use in broiler birds, the most important target population. Salmonella-vectored vaccines represent a convenient and effective option for controlling this disease [9]. They used a single attenuated Salmonella vaccine strain, to deliver up to three C.perfringens antigens; two of the antigens were toxoids and the third was FBA. Oral immunization with a single Salmonella vaccine strain producing one of the three antigens or all of them was immunogenic, generating humoral, cellular and mucosal responses against these antigens. The strains delivering FBA only or all three antigens provided the best protection. They also demonstrated that both toxins and FBA are present on the C. perfringens cell surface. In several organisms, FBA is recognized as a "moonlighting protein", one that can perform two or more autonomous functions. However, it may seem to be a usual choice for inclusion in a vaccine.

The aim of this study is to predict an effective epitope-based vaccine against fructose 1,6-biphosphate aldolase (FBA) enzyme of Clostridium perfringens using immunoinformatics tools. Immunoinformatics may be placed at the junction point between experimental and computational approaches. It can be described as a branch of bioinformatics concerned with in silico analysis and modeling of immunological data. More advanced analyses of the immune system using computational models typically involve conversion of an immunological question to a computational problem, followed by solving of the computational problem and translation of these results into biologically meaningful answers. Immunoinformatics research stresses mostly on the design and study of algorithms for mapping potential B- and T-cell epitopes, which speeds up the time and lowers the cost needed for laboratory analysis of pathogen gene products. Using such tools and information, the sequence areas with potential binding sites could be analyzed, which in turn leads to the development of new vaccines [17-19]. This study is regarded as the first study using computational approach to design an epitope-based vaccine.

MATERIALS AND METHODS

Protein sequence retrieval

A total of 94 Clostridium perfringes FBA strains were retrieved from National Center for Biotechnology Information (NCBI) database on July 2019 in FASTA format. These strains were gathered from different parts of the world for immunoinformatics analysis. The retrieved protein strains had length of 288 that hold the name fructose-1,6-bisphosphate aldolase.

Determination of conserved regions

The retrieved sequences of Clostridium perfringes FBA were subjected to multiple sequence alignment (MSA) by blasting them against reference sequence (WP_ 124041924.1) using ClustalW tool of BioEdit Sequence Alignment Editor Software version 7.2.5 to determine the conserved regions. Molecular weight and amino acid composition of the protein were also obtained [20].

Sequenced-based method

The reference sequence of Clostridium perfringes FBA was submitted to different prediction tools at the Immune Epitope Database (IEDB) analysis resource to predict various B and T cell epitopes. Conserved epitopes would be considered as candidate epitopes for B and T cells [21].

B cell epitope prediction

B cell epitopes is the part of the vaccine that interacts with B-lymphocytes. Candidate epitopes were analysed using several B cell prediction methods from IEDB, to identify the surface accessibility, antigenicity and hydrophilicity with the aid of computerized algorithm. The Bepipred Linear Epitope Prediction 2 was used to predict linear B cell epitope with default threshold value 0.533. The EMINI Surface Accessibility Prediction tool was used to detect the surface accessibility with default threshold value 1.000. The Kolaskar and Tongaonkar antigenicity method was used to identify the antigenicity sites of candidate epitope with default threshold value 1.032. Parker Hydrophilicity Prediction tool was used to identify the hydrophilic, accessible, or mobile regions with default threshold value 1.695 [22-25].

AllerTOP 2

This method is used for allergenicity predictions; it is based on auto cross covariance (ACC) transformation of protein sequences into uniform equal-length vectors. The reference protein sequence was inserted in the appropriate site in the software page. The principal Characteristics of the amino acids were represented by five E descriptors, which indicate amino acid hydrophobicity, molecular size, helix-forming propensity, relative abundance of amino acids, and β -strand forming propensity [26].

T cell epitope prediction MHC class I binding

T cell epitopes is the part of the vaccine that interacts with T lymphocytes. Analysis of peptide binding to the MHC (Major Histocompatibility complex) class I molecule was assessed by the IEDB MHC I prediction tool. Artificial Neural Network (ANN) 4.0 prediction method was used to predict the binding affinity. Before the prediction, all human allele lengths were selected and set to9 amino acids. The half-maximal inhibitory concentration (IC50) value required for all conserved epitopes to bind at score less than 100 were selected [27-29].

T cell epitope prediction MHC class II binding

Prediction of T cell epitopes interacting with MHC class II was assessed by the IEDB MHC II prediction tool for helper T cells. There are six tools for prediction: SMM_align, NN-align, Compinatorial Libraries, Sturniolo's method, Net MHC II pan and consensus method. Human allele references set were used to determine the interaction potentials of T cell epitopes and MHC Class II allele (HLA DR, DP and DQ). NN-align method was used to predict the binding affinity. IC50 values at score less than 500 were selected [30,31].

Population Coverage

In IEDB, the population coverage link was selected to analyse the epitopes. This tool calculates the fraction of individuals predicted to respond to a given set of epitopes with known MHC restrictions. The appropriate checkbox for calculation was checked based on MHC I, MHC II separately and combination of both which is set against the whole world population [32].

Homology Modelling

The 3D structure was obtained using raptorX which is a protein structure prediction server developed by Xu group, excelling at predicting secondary and tertiary structure for protein sequences without close homologs in the Protein Data Bank (PDB). Obtained 3D protein structure was visualized by USCF chimera (version 1.8) which was also used for visualization and analysis of molecular structure of the promising epitopes [33-35].

Molecular docking analysis

In silico molecular docking was performed to explore the binding affinity between the promising peptides and the target HLA-A02:01. The latter has been selected due to its involvement in many immunological and pathological diseases. AutoDock Vina was used to perform the docking analysis and the binding energies were calculated and ranked [36-38]. The model with the least RMSD value denotes a high binding affinity.

RESULTS

Multiple sequence alignment

The amino acid composition for the reference sequence of class II fructose-1,6-bisphosphate aldolase *Clostridium perfringens* and their conserved regions are represented in Figures 1 and 2 individually. Alanine and Glycine were the most frequent amino acids (Table 1). **Table 1:** Molecular weight and amino acid frequency distribution of the protein.

Amino Acid	Number	Mol%
Ala	37	12.85
Cys	3	1.04
Asp	10	3.47
Glu	26	9.03
Phe	9	3.13
Gly	32	11.11
His	5	1.74
Ile	22	7.64
Lys	21	7.29
Leu	23	7.99
Met	11	3.82
Asn	19	6.6
Pro	8	2.78
Gln	5	1.74
Arg	4	1.39
Ser	14	4.86
Thr	10	3.47
Val	21	7.29
Trp	2	0.69
Tyr	6	2.08

B-cell epitope prediction

Class II fructose-1, 6-bisphosphate aldolase *Clostridium perfringens* reference sequence was subjected to Bepipred linear epitope 2, Kolaskar and Tongaonkar antigenicity, EMINI surface accessibility, and Parker hydrophilicity prediction methods to test for different immunogenicity parameters. Epitopes successfully passed the three tests were subjected to AllerTOP 2.0 server to test their allergenicity (Table 2 and Figures 3-7). The four tests of 48 epitopes have been successfully passed. 3D tertiary structure of the proposed B cell

epitopes is shown (Figure 7).

Table 2: List of conserved peptides (Epitopes) with their antigenicity, EMINI surface accessibility, Parker, Hydrophilicity scores and their allergencity that successfully passed the four tests.

Peptide	Start	End	Length	Kolaskar and Tongaonkar antigenicity score (TH: 1.018)	EMINI surface accessibility score (TH: 1)	Parker Hydrophilicity prediction score (TH: 1.447)	Allertop
TAQENNSPVILGVSE	37	51	15	1.025	1.011	2.607	non-allergen
HLDHGSYQ	83	90	8	1.048	1.972	2.662	non-allergen
LDHGSYQG	84	91	8	1.02	1.434	3.113	non-allergen
DPAECKQI	155	162	8	1.044	1.263	3.387	non-allergen
HLDHGSYQG	83	91	9	1.029	1.637	3	non-allergen
LDHGSYQGA	84	92	9	1.025	1.215	3	non-allergen
YSIEENIVK	110	118	9	1.03	1.331	1.467	non-allergen
PAECKQIAE	156	164	9	1.045	1.11	3	non-allergen
VSKINVNTE	226	234	9	1.019	1.206	2.644	non-allergen
HLDHGSYQGA	83	92	10	1.033	1.399	2.91	non-allergen
SHYSIEENIV	108	117	10	1.045	1.027	1.61	non-allergen
HYSIEENIVK	109	118	10	1.037	1.533	1.53	non-allergen

OPEN OACCESS Freely available online

YSIEENIVKT	110	119	10	1.018	1.626	1.84	non-allergen
SHYSIEENIVK	108	118	11	1.035	1.697	1.982	non-allergen
HYSIEENIVKT	109	119	11	1.026	1.828	1.864	non-allergen
DPAECKQIAEL	155	165	11	1.047	1.069	2.527	non-allergen
HLDHGSYQGAID	83	94	12	1.029	1.102	2.592	non-allergen
GSHYSIEENIVK	107	118	12	1.022	1.367	2.292	non-allergen
SHYSIEENIVKT	108	119	12	1.024	1.994	2.25	non-allergen
HYSIEENIVKTK	109	120	12	1.018	2.976	2.183	non-allergen
CQLSFAEATRKY	235	246	12	1.053	1.427	1.725	non-allergen
EIADPAECKQIAE	152	164	13	1.026	1.07	2.992	non-allergen
CQLSFAEATRKYIE	235	248	14	1.045	1.191	1.464	non-allergen
VNTECQLSFAEATRK	231	245	15	1.026	1.562	2.593	non-allergen
CQLSFAEATRKYIEA	235	249	15	1.047	1.006	1.507	non-allergen
SHYSIEENIVKTKEII	108	123	16	1.024	1.621	1.531	non-allergen
VNTECQLSFAEATRKY	231	246	16	1.034	2.034	2.313	non-allergen
NTECQLSFAEATRKYI	232	247	16	1.02	1.921	2.044	non-allergen
VNTECQLSFAEATRKYI	231	247	17	1.041	1.179	1.706	non-allergen
ECQLSFAEATRKYIEAG	234	250	17	1.025	1.185	2.124	non-allergen
CQLSFAEATRKYIEAGK	235	251	17	1.03	1.369	2	non-allergen
VNTECQLSFAEATRKYIE	231	248	18	1.031	1.684	2.044	non-allergen
TECQLSFAEATRKYIEAG	233	250	18	1.018	1.41	2.294	non-allergen
ECQLSFAEATRKYIEAGK	234	251	18	1.02	1.955	2.322	non-allergen
CQLSFAEATRKYIEAGKD	235	252	18	1.02	1.885	2.444	non-allergen
VNTECQLSFAEATRKYIEA	231	249	19	1.033	1.361	2.047	non-allergen
CQLSFAEATRKYIEAGKDL	235	253	19	1.033	1.243	1.832	non-allergen
VSKINVNTECQLSFAEATRK	226	245	20	1.032	1.293	2.32	non-allergen
VNTECQLSFAEATRKYIEAG	231	250	20	1.025	1.097	2.23	non-allergen
CQLSFAEATRKYIEAGKDLE	235	254	20	1.023	1.754	2.13	non-allergen
FSSVMFDGSHYSIEENIVKTK	100	120	21	1.018	1.134	1.671	non-allergen
GVSKINVNTECQLSFAEATRK	225	245	21	1.025	1.049	2.481	non-allergen
VSKINVNTECQLSFAEATRKY	226	246	21	1.038	1.661	2.119	non-allergen
VNTECQLSFAEATRKYIEAGK	231	251	21	1.02	1.798	2.395	non-allergen
GVSKINVNTECQLSFAEATRKY	225	246	22	1.031	1.321	2.282	non-allergen
VSKINVNTECQLSFAEATRKYIE	226	248	23	1.035	1.325	1.926	non-allergen
SLGVSKINVNTECQLSFAEATRKY	223	246	24	1.039	1.002	1.979	non-allergen
GVSKINVNTECQLSFAEATRKYIE	225	248	24	1.028	1.101	2.083	non-allergen
VSKINVNTECQLSFAEATRKYIEA	226	249	24	1.036	1.123	1.933	non-allergen



Figure 1: Amino acid composition for class II fructose-1,6-bisphosphate aldolase *Clostridium perfringens* using BioEdit software.



Figure 2: Multiple Sequence Alignment using BioEdit software.



Figure 3: Bepipred Linear Epitope Prediction 2.0. Yellow areas above the threshold (red line) are proposed to be a part of B cell epitopes. While green areas are not.



Figure 4: Kolaskar & Tongaonkar Antigenicity. Yellow areas above the threshold (red line) are proposed to be a part of B cell epitopes. While green areas are not.



Average: 1.000 Minimum: 0.132 Maximum: 5.033

Figure 5: Emini Surface Accessibility Prediction. Yellow areas above the threshold (red line) are proposed to be a part of B cell epitopes. While green areas are not.





OPEN OACCESS Freely available online

Prediction of Cytotoxic T-lymphocyte epitopes and modelling

Class II fructose-1, 6-bisphosphate aldolase *Clostridium perfringens* reference protein was analyzed using (IEDB) MHC-I binding prediction tool to predict T cell epitopes suggested interacting with different types of selected Human MHC Class I alleles, based on Artificial Neural Network (ANN) method with half-maximal inhibitory concentration (IC50) less than 100 nm. 41 peptides were predicted to interact with different MHC-I alleles. Epitopes and their corresponding MHC I alleles are shown in Table 3 followed by the three-dimensional structure of the proposed T cell epitope (Figure 8).



Figure 8: Proposed T cell epitopes that interact with MHC I. The arrow shows position of (SVMFDGSHY) in green colour at structural level using Chimera software.

 Table 3: The most promising T cell epitopes and their corresponding MHC I alleles.

Peptide	MHC I alleles
SVMFDGSHY	HLA-A*29:02, HLA-B*15:01, HLA-A*30:02, HLA-A*11:01
EELKIAVPV	HLA-B*40:02, HLA-B*18:01, HLA-B*40:01
KYMCGFKTI	HLA-A*23:01, HLA-C*14:02, HLA-A*24:02
MLAAGIGNI	HLA-A*02:01, HLA-A*02:06, HLA-A*68:02
YMCGFKTIV	HLA-C*12:03, HLA-A*02:01, HLA-A*02:06
AMDAGFSSV	HLA-A*02:06, HLA-A*02:01
KLLNPGFEA	HLA-A*02:06, HLA-A*02:01
MALVNAKEM	HLA-C*03:03, HLA-B*35:01
MVNGMLEEL	HLA-A*68:02, HLA-A*02:06
YSIEENIVK	HLA-A*68:01, HLA-A*11:01

Prediction of the Helper T-lymphocyte epitopes and modelling

Class II fructose-1,6-bisphosphate aldolase *Clostridium perfringens* reference sequence was analyzed using (IEDB) MHC-II binding prediction tool based on NN-align with half-maximal inhibitory concentration (IC50) less than 500 nm. 222 predicted epitopes were found to interact with MHC-II alleles. The most promising epitopes and their corresponding alleles are shown in Table 4 along

with the 3D structure of the proposed epitope (Figure 9). **Table 4:** The most promising T cell epitopes and their corresponding MHC-II alleles.

Peptide	HLA Alleles		
INNLEWTKAILLTAQ	HLA-DQA1*01:02/DQB1*06:02, HLA-DRB1*07:01, HLA-DPA1*03:01/ DPB1*04:02, HLA-DRB1*01:01, HLA-DPA1*02:01/DPB1*01:01, HLA- DRB1*09:01, HLA-DPA1*01:03/ DPB1*02:01, HLA-DRB4*01:01, HLA- DRB1*15:01, HLA-DRB1*04:04, HLA-DQA1*05:01/DQB1*03:01, HLA- DRB1*13:02, HLA-DRB5*01:01, HLA- DRB1*13:02, HLA-DRB5*01:01, HLA- DRB1*13:02, HLA-DPA1*01/DPB1*04:01, HLA-DRB1*11:01, HLA-DRB1*04:01, HLA- DQA1*03:01/DQB1*03:02,		
NNLEWTKAILLTAQE	HLA-DQA1*01:02/DQB1*06:02, HLA-DRB1*07:01, HLA-DPA1*03:01/ DPB1*04:02, HLA-DRB1*09:01, HLA-DPA1*02:01/DPB1*01:01, HLA- DRB1*01:01, HLA-DPA1*01:03/ DPB1*02:01, HLA-DQA1*03:01/ DQB1*03:02, HLA-DRB4*01:01, HLA-DQA1*05:01/DQB1*03:01, HLA-DRB1*11:01, HLA-DRB1*04:05, HLA-DRB1*04:04, HLA-DRB1*15:01, HLA-DRB1*04:02, HLA-DRB1*15:01, HLA-DRB1*04:02, HLA-DRB5*01:01, HLA- DRB1*04:01		
NLEWTKAILLTAQEN	HLA-DQA1*01:02/DQB1*06:02, HLA-DRB1*09:01, HLA-DPA1*02:01/ DPB1*01:01, HLA-DPA1*03:01/ DPB1*04:02, HLA-DRB1*07:01,HLA- DRB1*01:01, HLA-DQA1*03:01/ DQB1*03:02, HLA-DRB1*11:01, HLA-DRB4*01:01, HLA-DQA1*04:01/ DQB1*04:02, HLA-DRB1*04:05, HLA- DRB1*04:04, HLA-DQA1*05:01/ DQB1*03:01, HLA-DQA1*05:01/ DQB1*03:01, HLA-DRB1*04:01, HLA- DRB1*15:01, HLA-DRB1*04:01, HLA-		
FKTIVGMVNGMLEEL	HLA-DRB1*01:01, HLA-DRB1*04:04, HLA-DRB1*04:05, HLA-DRB1*09:01, HLA-DRB1*07:01, HLA-DRB5*01:01, DPB1*01:01, HLA-DRB5*01:01, HLA-DQA1*05:01/DQB1*02:01, HLA-DQA1*05:01/DQB1*03:01, HLA-DRB1*04:01, HLA-DRB1*15:01, HLA-DQA1*04:01/DQB1*04:02, HLA-DQA1*01:02/DQB1*06:02, HLA-DRB4*01:01, HLA-DPA1*01:03/ DPB1*02:01, HLA-DPA1*03:01/ DPB1*04:02		
NWAGLNFEALANIKA	HLA-DRB1*01:01, HLA-DRB1*09:01, HLA-DRB1*04:04, HLA-DRB1*11:01, HLA-DRB1*04:01, HLA-DQA1*05:01/ DQB1*03:01, HLA-DPA1*01:03/ DPB1*02:01, HLA-DQA1*01:02/ DQB1*06:02, HLA-DPA1*03:01/ DPB1*04:02, HLA-DPA1*02:01/ DPB1*01:01, HLA-DPA1*01/DPB1*04:01, HLA-DRB1*07:01, HLA-DQA1*05:01/ DQB1*02:01, HLA-DRB5*01:01, HLA- DQA1*03:01/DQB1*03:02, HLA- DRB1*04:05		

OPEN OACCESS Freely available online



Figure 9: Proposed T cell epitopes that interact with MHC II. The arrow shows position of (INNLEWTKAILLTAQ) in green colour at structural level using Chimera software.

Population coverage analysis

All MHC I and MHC II epitopes were evaluated for population coverage against the whole world using the IEDB population coverage tool. Epitopes for MHC I with highest population coverage were YMCGFKTIV (46.73%) and MLAAGIGNI (42.53%) (Figure 10 and Table 5). For MHC class II, the epitopes that showed the highest population coverage was INNLEWTKAILLTAQ (71.51%), NNLEWTKAILLTAQE & NLEWTKAILLTAQEN (69.46%) (Figure 11 and Table 6). For MHC class I and II combined, the epitopes that showed the highest population coverage were INNLEWTKAILLTAQ (71.51%), and NNLEWTKAILLTAQEN (69.46%) NLEWTKAILLTAQEN (69.46%) (Figure 12 and Tables 7 and 8).





Table 5: Population coverage of peptides interaction with MHC class I.

Epitope	Coverage class 1 (%)	Total HLA hits
YMCGFKTIV	0.4673	3
MLAAGIGNI	0.4253	3
AMDAGFSSV	0.406	2
KLLNPGFEA	0.406	2
MLEELKIAV	0.3908	1
KYMCGFKTI	0.2843	3
SVMFDGSHY	0.2793	4
KYAVGQFNI	0.2138	1
YSIEENIVK	0.2088	2
EELKIAVPV	0.1802	3
MALVNAKEM	0.1585	2
LSFAEATRK	0.1553	1
FAEATRKYI	0.1031	1
KAREGKYAV	0.1031	1
TAQENNSPV	0.1031	1
KAATGDMPL	0.0812	1
WAGLNFEAL	0.0812	1
AIDAMDAGF	0.0785	1
LEWTKAILL	0.0781	1
QENNSPVIL	0.0781	1

 Table 6: Population coverage of proposed peptides interaction with MHC class II.

Epitope	Coverage Class II (%)	Total HLA hits
INNLEWTKAILLTAQ	0.7151	18
CGFKTIVGMVNGMLE	0.6946	12
GFKTIVGMVNGMLEE	0.6946	13
MCGFKTIVGMVNGML	0.6946	11
NLEWTKAILLTAQEN	0.6946	17
NNLEWTKAILLTAQE	0.6946	18
AKYMCGFKTIVGMVN	0.6437	9
AAKYMCGFKTIVGMV	0.6254	8
FKTIVGMVNGMLEEL	0.6174	16
PGFEAIKATVKEKME	0.6047	9
FNINNLEWTKAILLT	0.5831	15
NINNLEWTKAILLTA	0.5831	15
QFNINNLEWTKAILL	0.5831	14
NPGFEAIKATVKEKM	0.5783	10
WAGLNFEALANIKAA	0.5783	16

OPEN OACCESS Freely available online

DMIAEAISLGVSKIN	0.5672	8
MIAEAISLGVSKINV	0.5672	8
PSDMIAEAISLGVSK	0.5672	8
SDMIAEAISLGVSKI	0.5672	9
LEWTKAILLTAQENN	0.563	15
NWAGLNFEALANIKA	0.563	16
FEAIKATVKEKMELF	0.5584	7
GFEAIKATVKEKMEL	0.5584	7
PLVLHGGTGIPSDMI	0.556	7
ATGDMPLVLHGGTGI	0.5355	6
DMPLVLHGGTGIPSD	0.5355	6
GDMPLVLHGGTGIPS	0.5355	6
LVLHGGTGIPSDMIA	0.5355	6
MPLVLHGGTGIPSDM	0.5355	6
TGDMPLVLHGGTGIP	0.5355	6
YMCGFKTIVGMVNGM	0.535	9
KYMCGFKTIVGMVNG	0.5189	8
CQLSFAEATRKYIEA	0.5094	8
ECQLSFAEATRKYIE	0.5094	8
QLSFAEATRKYIEAG	0.5094	6
TECQLSFAEATRKYI	0.5094	9
LNPGFEAIKATVKEK	0.5051	9
	1 . 1 .	1.

 Table 7: Population coverage of proposed peptide interaction with MHC

 class I and II combined.

Epitope	Coverage Class I and II	Total HLA hits
INNLEWTKAILLTAQ	0.7151	18
CGFKTIVGMVNGMLE	0.6946	12
GFKTIVGMVNGMLEE	0.6946	13
MCGFKTIVGMVNGML	0.6946	11
NLEWTKAILLTAQEN	0.6946	17
NNLEWTKAILLTAQE	0.6946	18
AKYMCGFKTIVGMVN	0.6437	9
AAKYMCGFKTIVGMV	0.6254	8
FKTIVGMVNGMLEEL	0.6174	16
PGFEAIKATVKEKME	0.6047	9
FNINNLEWTKAILLT	0.5831	15
NINNLEWTKAILLTA	0.5831	15
QFNINNLEWTKAILL	0.5831	14
NPGFEAIKATVKEKM	0.5783	10
WAGLNFEALANIKAA	0.5783	16
DMIAEAISLGVSKIN	0.5672	8
MIAEAISLGVSKINV	0.5672	8
PSDMIAEAISLGVSK	0.5672	8
SDMIAEAISLGVSKI	0.5672	9
LEWTKAILLTAQENN	0.563	15
NWAGLNFEALANIKA	0.563	16
FEAIKATVKEKMELF	0.5584	7
GFEAIKATVKEKMEL	0.5584	7
PLVLHGGTGIPSDMI	0.556	7
ATGDMPLVLHGGTGI	0.5355	6
DMPLVLHGGTGIPSD	0.5355	6
GDMPLVLHGGTGIPS	0.5355	6
LVLHGGTGIPSDMIA	0.5355	6

MPLVLHGGTGIPSDM	0.5355	6
TGDMPLVLHGGTGIP	0.5355	6
YMCGFKTIVGMVNGM	0.535	9
KYMCGFKTIVGMVNG	0.5189	8
CQLSFAEATRKYIEA	0.5094	8
ECQLSFAEATRKYIE	0.5094	8
QLSFAEATRKYIEAG	0.5094	6
TECQLSFAEATRKYI	0.5094	9
LNPGFEAIKATVKEK	0.5051	9
GAAKYMCGFKTIVGM	0.499	6
YMCGFKTIV	0.4673	3
AECKQIAELGVTMLA	0.4665	10
ECKQIAELGVTMLAA	0.4665	10
AGLNFEALANIKAAT	0.4448	12
EWTKAILLTAQENNS	0.4448	13
GLNFEALANIKAATG	0.4448	10
LNFEALANIKAATGD	0.4448	10
ALANIKAATGDMPLV	0.4356	7
EALANIKAATGDMPL	0.4356	7
GIPSDMIAEAISLGV	0.4306	7
IPSDMIAEAISLGVS	0.4306	7
AIKATVKEKMELFGS	0.4277	5
EAIKATVKEKMELFG	0.4277	3
MLAAGIGNI	0.4253	3
FEALANIKAATGDMP	0.422	8
NFEALANIKAATGDM	0.422	9
VTMLAAGIGNIHGKY	0.4194	5
CKQIAELGVTMLAAG	0.4179	9
EGAAKYMCGFKTIVG	0.4139	6
AMDAGFSSV	0.406	2
KLLNPGFEA	0.406	2
KTIVGMVNGMLEELK	0.404	14
TIVGMVNGMLEELKI	0.404	13
KLLNPGFEAIKATVK	0.4025	8
LLNPGFEAIKATVKE	0.4025	7
AGFSSVMFDGSHYSI	0.4002	9
WTKAILLTAQENNSP	0.3953	10
VMFDGSHYS	0.3908	1
NSPVILGVSEGAAKY	0.3527	6
SPVILGVSEGAAKYM	0.3527	7
KAILLTAQENNSPVI	0.3448	6
AIDAMDAGFSSVMFD	0.3418	6
GAIDAMDAGFSSVMF	0.3418	4
IDAMDAGFSSVMFDG	0.3418	8
IAEAISLGVSKINVN	0.3402	5
NTECQLSFAEATRKY	0.3402	7
PAECKQIAELGVTML	0.3402	6

Molecular docking analysis

Molecular docking analysis for the MHC I epitope YMCGFKTIV with HLA allele (HLA-A*02:01) was promising. The model with the least RMSD value was chosen for visualization (Figure 13).

OPEN ORCESS Freely available online



Figure 13: Molecular docking of MHC I peptide YMCGFKTIV with allele (HLA-A*02:01).

DISCUSSION

The main goal of a vaccine development is to elicit immune response against particular pathogen by selectively stimulating antigen specific B and T cells [39]. Peptide vaccines overcome the side effects of conventional vaccines and are characterized by easy production, stimulating effective immune response and no potential infection possibilities [19]. Many studies had predicted peptide vaccines for different microorganisms such as, Rubella virus, Ebola virus, Dengue virus, Zika virus, Human Papilloma virus, Lagos rabies virus, mycetoma, pseudomona aeruginosa using immunoinformatics tools [40-50].

Current study provides many different peptides for *Clostridium perfringens* vaccine development against FBA protein for the first time, including twelve promising peptides that can be recognized by B cells and T cells (QENNSPVIL, KYIEAGKDL, EIADPAECK, FAEATRKYI, SVMFDGSHY, TAQENNSPV, AECKQIAEL, QLSFAEATR, LSFAEATRK, SFAEATRKY, YSIEENIVK, NNSPVILGV). These peptides had passed all B cell prediction tests as well as MHC I and MHC II tests scoring the lowest rates of IC50 with their corresponding HLA alleles, indicating strong interaction between the peptide and allele. The resulting peptide vaccine is expected to be more antigenic and less allergic than the conventional biochemical vaccine.

Clostridium Perfringens FBA protein sequence was subjected to B cell epitope prediction test in IEDB. Out of the nine predicted epitopes using Bepipred 2 test, eight epitopes were the most promising. A peptide length greater than twenty four is not preferred for vaccine design as it will turn pathogenic.

B cell immunity stands for short time, while T cell immunity is lasts longer. Therefore designing of peptide vaccine against T cell is more promising and effective. The protein sequence was analyzed to predict T cell epitopes. 41 epitopes were predicted to interact with MHC I alleles. Five of them were most promising and had the affinity to bind the highest number of MHC I alleles (SVMFDGSHY, YMCGFKTIV, MLAAGIGNI, KYMCGFKTI, EELKIAVPV). 222 predicted epitopes have interacted with MHC II alleles. Six of them had the affinity to bind to the highest number of alleles and were the most promising (INNLEWTKAILLTAQ, NNLEWTKAILLTAQE, NLEWTKAILLTAQEN,FKTIVGMVN GMLEEL, WAGLNFEALANIKAA, NWAGLNFEALANIKA).

All epitopes were tested using population coverage tool of IEDB which measures the percentage of people in whole world who have potential to develop immune response to vaccine containing this epitope. The best epitope with the highest population coverage for MHC I were YMCGFKTIV with 46.73% in three HLA hits and MLAAGIGNI with 42.53% in three hits, and the coverage of population set for whole MHC I epitopes was 90.46%. Excluding certain alleles for MHC II, the best epitope were INNLEWTKAILLTAQ scoring 71.51% with eighteen HLA hits and NNLEWTKAILLTAQE scoring 69.46% with eighteen HLA hits, and the coverage of population set was 73.08% for the whole MHC II epitope set. When combined, the peptide INNLEWTKAILLTAQ had the highest population coverage percent 71.51 in eighteen HLA hits for both MHC I and MHC II.

Limitations include the exclusion of certain HLA alleles for the MHC II population coverage tests and unavailability of MHC II alleles models for docking analysis. We hope that the world will benefit from this epitope-based vaccine upon its successful development following in vivo and in vitro studies to prove its effectiveness.

CONCLUSION

Vaccination is a method to protect and minimize the possibility of infection. Design of vaccines using in silico prediction methods is highly appreciated due to the significant reduction in cost, time and effort. Peptide vaccines overcome the side effects of conventional vaccines. We presented different peptides that can produce antibodies against FBA of *Clostridium perfringens* for the first time. Nine B cell epitopes passed the antigenicity, accessibility and hydrophilicity tests. Five MHC I epitopes was the most promising ones, while six for MHC II. For the population coverage, the epitopes covered 90.46% and 73.08% of the alleles worldwide for MHC I and II respectively excluding certain alleles.

DATA AVAILABILITY

The data which support our findings in this study are available from the corresponding author upon reasonable request.

COMPETING INTEREST

The authors declare that they have no competing interests.

ACKNOWLEDGEMENT

Many thank goes to National Center for Biotechnology Information (NCBI), BioEdit, Immune Epitope Database Analysis Resource (IEDB), RaptorX server, and UCSF Chimera team. We are grateful to Editor-in-Chief and all the unknown reviewers for their valuable comments on the previous version of this manuscript.

REFERENCES

- Kiu R, Hall LJ. An update on the human and animal enteric pathogen Clostridium perfringens. Emerg Microbes Infect. 2018;7:141.
- 2. Hunter JGL, Wilde S, Tafoya AM, Horsman J, Yousif M, Diamos AG, et al., Evaluation of a toxoid fusion protein vaccine produced in plants to protect poultry against necrotic enteritis. PeerJ. 2019;7:6600.
- Sakanoue H, Nakano T, Sano K, Yasugi M, Monma C, Miyake M. Adherence of *Clostridium perfringens* spores to human intestinal epithelial Caco-2 cells. FEMS Microbiol Lett. 2018;365.
- 4. Uzal FA, Freedman JC, Shrestha A, Theoret JR, Garcia J, Awad MM, et al. Towards an understanding of the role of *Clostridium perfringens* toxins in human and animal disease. Future Microbiol. 2014;9:361-377.
- 5. Uzal FA, McClane BA, Cheung JK, Theoret J, Garcia JP, Moore RJ, et al. Animal models to study the pathogenesis of human and animal *Clostridium perfringens* infections. Vet Microbiol. 2015;179:23-33.

OPEN OACCESS Freely available online

Elhag M, et al.

- 6. Yonogi S, Matsuda S, Kawai T, Yoda T, Harada T, Kumeda Y, et al. BEC a novel enterotoxin of *Clostridium perfringens* found in human clinical isolates from acute gastroenteritis outbreaks. Infect Immun. 2014;82:2390-2399.
- Li J, Evans DR, Freedman JC, McClane BA. NanR Regulates nanI Sialidase Expression by Clostridium perfringens F4969, a Human Enteropathogenic Strain. Infect Immun. 2017;85.
- 8. Li J, McClane BA. Contributions of NanI sialidase to Caco-2 cell adherence by *Clostridium perfringens* type A and C strains causing human intestinal disease. Infect Immun. 2014;82: 4620-4630.
- Wilde S, Jiang Y, Tafoya AM, Horsman J, Yousif M, Vazquez LA, et al. Salmonella-vectored vaccine delivering three *Clostridium perfringens* antigens protects poultry against necrotic enteritis. PLoS One. 2019;14:0197721.
- Tunio SA, Oldfield NJ, Berry A, Ala'Aldeen AA, Wooldridge KG, Turner DPJ. The moonlighting protein fructose-1, 6-bisphosphate aldolase of Neisseria meningitidis: surface localization and role in host cell adhesion. Mol Microbiol. 2010;76:605-615.
- Shams F, Oldfield NJ, Wooldridge KG, Turner DPJ. Fructose-1,6bisphosphate aldolase (FBA)-a conserved glycolytic enzyme with virulence functions in bacteria: 'ill met by moonlight'. Biochem Soc Trans. 2014;42:1792-1795.
- 12. Sun Z, Shen B, Wu H, Zhou X, Wang Q, Xiao J, et al. The secreted fructose 1,6-bisphosphate aldolase as a broad spectrum vaccine candidate against pathogenic bacteria in aquaculture. Fish Shellfish Immunol. 2015;46:638-647.
- Kulkarni RR, Parreira VR, Sharif S, Prescott JF. Clostridium perfringens antigens recognized by broiler chickens immune to necrotic enteritis. Clin Vaccine Immunol. 2006;13:1358-1362.
- Kulkarni RR, Parreira VR, Sharif S, Prescott JF. Immunization of broiler chickens against *Clostridium perfringens*-induced necrotic enteritis. Clin Vaccine Immunol. 2007;14:1070-1077.
- Kulkarni RR, Parreira VR, Sharif S, Prescott JF. Oral immunization of broiler chickens against necrotic enteritis with an attenuated Salmonella vaccine vector expressing *Clostridium perfringens* antigens. Vaccine. 2008;26:4194:4203.
- Mahmood K, Rahman SU, Hussain I, Abbas RZ, Khaliq T, Arif J, et al. Non-antibiotic strategies for the control of necrotic enteritis in poultry. World's Poultry Science Journal. 2014;70:865-879.
- 17. Brusic V, Petrovsky N. Immunoinformatics--the new kid in town. Novartis Found Symp. 2003;254:3-13.
- Tomar N, De RK. Immunoinformatics: An integrated scenario. Immunology. 2010;131:153-168.
- Patronov A, Doytchinova I. T-cell epitope vaccine design by immunoinformatics. Open Biol. 2013;3:120-139.
- Islam SU, Qasim M, Lin W, Islam W, Arif M, Ali H, et al. Genetic interaction and diversity of the families Libellulidae and Gomphidae through COI gene from China and Pakistan. Acta Trop. 2018;182:92-99.
- 21. Vita R. The Immune Epitope Database (IEDB): 2018 update. Nucleic Acids Res. 2019;47:339-343.
- 22. Jespersen MC, Peters B, Nielsen M, Marcatili P. BepiPred-2.0: improving sequence-based B-cell epitope prediction using conformational epitopes. Nucleic Acids Res. 2017;45:W24-W29.
- Emini EA, Hughes JV, Perlow DS, Boger J. Induction of hepatitis A virus-neutralizing antibody by a virus-specific synthetic peptide. J Virol. 1985;55:836-839.
- Kolaskar AS, Tongaonkar PC. A semi-empirical method for prediction of antigenic determinants on protein antigens. FEBS Lett. 1990;276:172-174.

- 25. Parker JM, Guo D, Hodges RS. New hydrophilicity scale derived from high-performance liquid chromatography peptide retention data: correlation of predicted surface residues with antigenicity and X-rayderived accessible sites. Biochemistry. 1986;25:5425-5432.
- Dimitrov I, Flower DR, Doytchinova I. AllerTOP v.2-a server for in silico prediction of allergens. J Mol Model. 2014;20:2278.
- Jorgensen KW, Rasmussen M, Buus S, Nielsen M. NetMHCstab predicting stability of peptide-MHC-I complexes; impacts for cytotoxic T lymphocyte epitope discovery. Immunology. 2014;141:18-26.
- Lundegaard C, Nielsen M, Lund O. The validity of predicted T-cell epitopes. Trends Biotechnol. 2006;24:537-538.
- 29. Nielsen M, Lundegaard C, Worning P, Lauemoller SL, Lamberth K, Buus S, et al. Reliable prediction of T-cell epitopes using neural networks with novel sequence representations. Protein Sci. 2003;12:1007-1017.
- Nielsen M, Lund O. NN-align: An artificial neural network-based alignment algorithm for MHC class II peptide binding prediction. BMC Bioinformatics. 2009;10:296.
- 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 Bioinformatics. 2010;11:568.
- 32. Bui HH, Sidney J, Dinh K, Southwood S, Newman MJ, Sette A. Predicting population coverage of T-cell epitope-based diagnostics and vaccines. BMC bioinformatics. 2006;7:153.
- 33. Pettersen EF, Goddard TD, Huang CC, Couch GS, Greenblatt DM, Meng EC, et al. UCSF Chimera-a visualization system for exploratory research and analysis. J Comput Chem. 2004;25:1605-1612.
- 34. Peng J, Xu J. RaptorX: exploiting structure information for protein alignment by statistical inference. Proteins. 2011;79:161-171.
- 35. Trott O, Olson AJ. AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. J Comput Chem. 2010;31:455-461.
- Davenport MP, Hill AV. Reverse immunogenetics: from HLA-disease associations to vaccine candidates. Mol Med Today. 1996;2:38-45.
- 37. Burton PR, Tobin MD, Hopper JL. Key concepts in genetic epidemiology. Lancet. 2005;366:941-951.
- Purcell AW, McCluskey J, Rossjohn J. More than one reason to rethink the use of peptides in vaccine design. Nat Rev Drug Discov. 2007;6:404-414.
- 39. Elhag M, Alaagib RM, Ahmed NM, Abubaker M, Haroun EM, Albagi SOA, et al. Design of Epitope-Based Peptide Vaccine against Pseudomonas aeruginosa Fructose Bisphosphate Aldolase Protein Using Immunoinformatics. J Immunol Res. 2020;2020: 9475058.
- 40. Elhag MM, Saadaldin MM, Abdelmoneim AH, Taha TS, Abdelrhman FA, Hassan MA. Immunoinformatics Approach for Designing an Epitope-Based Peptide Vaccine against Treponema pallidum Outer Membrane Beta-Barrel Protein. Immunome Research. 2020;16: 1-12.
- 41. Mohammed AA, ALnaby AMH, Sabeel SM, AbdElmarouf FM, Dirar AI, Ali MM, et al. Epitope-Based Peptide Vaccine Against Fructose-Bisphosphate Aldolase of Madurella mycetomatis Using Immunoinformatics Approaches. Bioinform Biol Insights. 2018;12:1177932218809703.
- 42. Elgenaid S, Al-Hajj AM, Ibrahim AA, Essa MEA, Hamdi A, Abdelrahman KA, et al. Prediction of Multiple Peptide Based Vaccine from E1, E2 and Capsid Proteins of Rubella Virus: An In-Silico Approach. Immunome Research. 2018;14:2.
- 43. Bolis S, Omer WA, Abdelhamed MA, Shambal MA, Adam EA, Abass MA, et al. Immunoinformatics Prediction of Epitope Based Peptide Vaccine Against Madurella mycetomatis Translationally Controlled Tumor Protein. BioRxiv. 2018;441881.

OPEN OACCESS Freely available online

- 44. Mohammed AA, Hashim O, Elrahman KAA, Hamdi A and Hassan MA. Epitope-based peptide vaccine design against Mokola rabies virus glycoprotein G utilizing in silico approaches. Immunome Res. 2017;13:2.
- 45. Albagi S, Ahmed OH, Gumaa MA, Abd_elrahman, Abu-Haraz AH, Hassan MA. Immunoinformatics-Peptide Driven Vaccine and In silico Modeling for Duvenhage Rabies Virus Glycoprotein G. J Clin Cell Immunol. 2017;8:2.
- 46. Albagi S, Ahmed OH, Gumaa MA, Abd_elrahman, Abu-Haraz AH, Hassan MA. Immunoinformatic approach for epitope-based peptide vaccine against Lagos rabies virus glycoprotein G. Immunome Res. 2017;13:2.
- Abu-haraz A. Multi epitope peptide vaccine prediction against Sudan Ebola virus using immuno-informatics approaches. Adv Tech Biol Med. 2017;5:2379-1764.1000203.
- 48. Mohamed SS, Sami SM, Elgailany HA, Bushara MK, Hassan HA, Mustafa GF, et al. Immunoinformatics Approach for Designing Epitope-Based Peptides Vaccine of L1 Major Capsid Protein against HPV Type 16. Int J Multidiscipl Curr Res. 2016.
- 49. Badawi MM. Highly conserved epitopes of Zika envelope glycoprotein may act as a novel peptide vaccine with high coverage: immunoinformatics approach. Am J Biomed Res. 2016; 4:46-60.
- 50. Badawi MM. Immunoinformatics predication and in silico modeling of epitope-based peptide vaccine against virulent Newcastle disease viruses. Am J Infe Dis Microbiol. 2016;4:61-71.