In Silico Analysis of CadF Epitope-based Vaccine Design Against Campylobacter Jejuni

Mona Moballegh Naseri
Cellular and Molecular Research Center, Qom University of Medical Sciences, Qom, Iran
Saeed Shams (sshamsmed@gmail.com)
Cellular and Molecular Research Center, Qom University of Medical Sciences, Qom, Iran
https://orcid.org/0000-0002-7552-0938
Mohammad Moballegh Naseri
Department of computer and IT,Shahab-Danesh University, Qom, Iran
Bita Bakhshi
Department of Bacteriology, Faculty of Medical Sciences, TarbiatModares University, Tehran, Iran

Research note

Keywords: Campylobacter jejuni, vaccine, CadF, in silico, bioinformatics

Posted Date: August 13th, 2020

DOI: https://doi.org/10.21203/rs.3.rs-45952/v1

Version of Record: A version of this preprint was published on November 10th, 2020. See the published version at https://doi.org/10.1186/s13104-020-05364-z.

Abstract

Objective: To eradicate infectious diseases, vaccination is an important strategy. CadF protein of *Campylo bacter jejuni* is one of the important factors in the process of the pathogenesis of the bacterium. So, the purpose of the work was to perform a bioinformatics study for identifying an epitope-based CadF vaccine, as a subunit vaccine. CadF sequences were extracted from the NCBI database. In silico analysis including all ergeni city, antigenicity, epitope conservancy assessment, molecular docking, etc.was done by different servers.

Results:The results showed that CadF is an antigenic and non-allergenic protein and provides a suitable structure for vaccine design. Among epitopes, LSDSLALRL has been confirmed to stimulate both B and T cells. This 9-mers peptide is located in 135-143 segment of the CadF protein and interacted with HLA-A0101 and HLA-DRB1 0101 with energies of docking -26.18kcal/mol and -109.89kcal/mol. The peptide is not an allergen and as an antigen, it has the ability for motivating the immune system. Hence, the analyses are performed on that the epitope structure could verify the design of a vaccine against *C. jejuni*. The obtained theoretical results showed that CadF protein could be used for designing and evaluating a new vaccine in humans.

Introduction

Campylobacter jejuni (*C. jejuni*) is one of the significant pathogens belongs to the genus *Campylobacters*, which are gram-negative, curved, and rod-shaped bacteria [1-3]. They can be transmitted to humans through direct contact with animals, consumption of contaminated food and water, and unpasteurized milk [1, 4].

Some gastrointestinal problems, called campylobacteriosis, especially in children have observed affected by *Campylobacter jejuni* [5-7]. CadF is one of the important proteins, a 37kDa protein as a conserved and genus-specific outer membrane protein, that binds to fibronectin and the bacterium attaches to the host cell, lead to facilitate colonization. The reports have been shown that CadF can induce massive immune responses, including humoral- and cellular- immunities [8-10].

Subunit vaccines contain a part of the target microorganism, which are known as safe and effective vaccines for humans and animals. They induce both humoral- and cell-mediated immune mechanisms in protection against the pathogens. To develop an effective subunit vaccine, the identification and prediction of the antigenic epitopes by bioinformatics tools are useful [11, 12]. Although there are some studies on the evolution of the outer membrane proteins of *C. jejuni* as vaccine candidates, CadF can be considered independently for the design of a protective vaccine [13, 14].

Our aim was to analysis CadF protein for the identification of the epitope-based peptide candidates and to evaluate its proteomic database by In silico tools for developing a new vaccine candidate.

Methods

Protein analysis and identification of conserved regions

The CadF protein sequences were acquired from NCBI Protein Data Bank (https://www.ncbi.nlm.nih.gov/protein) in FASTA format. To evolutionary analysis, multiple alignments and phylogenetic tree sequences were used the clustalw2 tool (https://www.ebi.ac.uk/Tools/msa/clustalw2) and Molecular Evolutionary Genetics Analysis version 7 (MEGA 7) software package. Besides, the amino acid position belonging to CadF protein was obtained by UniProt (https://www.uniprot.org).

Allergenicity and antigenicity assessment

The Allertop (www.ddg-pharmfac.net/AllerTOP) and AllergenFP (ddg-pharmfac.net/AllergenFP) web servers were used to determine the allergenicity of the protein. The Allertop server was programmed based on amino acid features such as hydrophobicity, size, and helix forming which could classify some allergen and non-allergen targets. The AllergenFP was databased to set options for predicting allergens. The conserved CadF antigens were also forecast by the Vaxijen server.

Epitope conservancy assessment

To evaluate the MHC-I and MHC-II (Major Histocompatibility Complex), epitopes were used the IEDB (crdd.osdd.net/raghava/propred), NetCTL (www.cbs.dtu.dk/services/NetCTL), NETMHC (www.cbs.dtu.dk/services/NetMHC), NHLApred (crdd.osdd.net/raghava/nhlapred), SYFPEITHI (www.syfpeithi.de), and MHC2Pred (http://crdd.osdd.net/raghava/mhc2pred/) online servers . Each above-mentioned data bases were employed to determine the T-cell epitopes scoring systems.

The B-cell epitopes were identified using the IEDB (crdd.osdd.net/raghava/propred), BCPREDS (http://ailab.ist.psu.edu/bcpred/) servers with the default setting specificity 75% and ABCpred (http://crdd.osdd.net/raghava/abcpred) server with considering threshold value 0.5. Linear and discontinuous B cell epitopes were also predicted by the Bepipre server (http://www.cbs.dtu.dk/services/BepiPred). This server was predicted for B-cell epitopes through Hidden Markov Model (HMM) model. Each epitope obtained from these servers was checked to determine the antigenicity property.Finally, the identified common and repetitive epitopes were selected and analyzed as predicted epitopes.

Design and evaluation of molecular docking

To recognize the three-dimensional structures and biological functions, Phyre2 (www.sbg.bio.ic.ac.uk/phyre2), was used as an online protein fold recognition server. The secondary structure of the protein was also analyzed by Psipred (http://bioinf.cs.ucl.ac.uk/psipred) server.

Two TMHMM v.2.0 (http://www.cbs.dtu.dk/services/TMHMM) and Protprom (https://web.expasy.org/protparam) servers were used to predict sequences of exo-membrane amino acid and physico-biochemical characteristics of CadF protein. To gain the best epitope structure, we used from PyMOL software.

Molecular Docking of adoptive epitope and alleles

PyMOL Graphics system was used to analyze the three-dimensional structure of the LSDSLALRL epitope, while the three-dimensional structures of HLA –A 0101 and HLA-DRB1*0101 alleles were extracted from the Protein Data Bank (https://www.rcsb.org) with the Uniprot KB ID: Q5SUL5 and P01911. The receptors mentioned was prepared by Notepad^{++} and the Molecular Docking of the epitope and HLA-A 0101 allele of MHC class I, HLA-DRB1 0101 of MHC class II were performed with the help of Molecular Virtual Docker and Molecular Virtual Viewer software. The interface between the epitope and the alleles were selected based on a grid, computed on three axes x: -0.16, y: -17.63, z: -15.67.

Results

The complete sequences of the CadF protein were contained 319 amino acids and multiple sequence alignment showed that this protein was a highly conserved protein among Campylobacters. It belonged to the outer member proteins superfamily (ompA) and an ompA-like domain was identified in the 193-287 position of the protein. The result of the phylogenic tree was also confirmed that CadF was classified in outer membrane protein superfamily (data not shown).

Antigenicity and Allergenicity protein analysis

The score of the antigenic prediction was calculated by about 0.79 by the Vaxijen server. The results showed that the protein was probably an antigen and could be used for further analysis. The obtained data from the AllergenFP server was indicated a 0.82 similarity for the protein; hence it could not be an allergen. Analysis of the Allertop server was also confirmed that CadF protein was not an allergen.

The physicochemical characterizes

Using the ProtParam server, the MV (molecular weight) and PI (isoelectric point) parameters were 35979.04 Da and 5.89, respectively. The aliphatic index was 69.12 and the GRAVY(grand average of hydropathicity) of protein was -0.679. As a result, these amino acids of CadF protein had hydrophobicity and acidity properties (PI≤7.35). The aliphatic index included alanine, valine, isoleucine, and leucine amino acids, indicating the thermostability of the protein. Moreover, TMHMM server analysis was confirmed that CadF was an outer membrane protein.

Prediction of secondary and tertiary structures

By the PSIPRED server, graphical results of the secondary structures of protein were obtained that indicated sheet, helix, and extracellular transmembrane structures (Additional file 1). Besides, The Phyre

online servers were reported a three-dimensional structure of the modeled CadF with a 97% confidence score and 192 known-domains aligns. The structural contents were included 16% alpha-helix, 41% beta strands, and 16% disordered regions. Also, the prediction of the CadF protein showed a binding site at GIU-HIS-LYS residue and a large amount of metallic heterogenic sections in its structure Also, the three-dimensional (3D) structure belonged to selected epitopes was drawn by PyMOL software (Supplementary file, A).

Forecasted of antigenic T cell epitopes

To predict T cell epitopes, the best score of the epitopes were selected from SYFPEITH, IEDB, NetCTL, NHLAPred, NETMHC I, and MHCPred II online servers. Except for IEDB, which showed a high value for the lowest number, the other servers score a high value for the highest number. The epitopes of the MHC I (A-0101, A-0201, and B-2705) and the MHCII (DR1-0101 and DRB1-0401) were the most common epitopes in Iranian alleles that have been considered in this study.According to achieved data from the above-mentioned servers, the predicted epitopes of MHC I and MHC II were presented in Tables 1 and 2, respectively.

Among selected epitopes, LLCLGLASV, RRVDAKFIL, FSADNNVKF, and LSDSLALRL (belong to MHC class I); and EGHFGFDKTTINPTF, QINFNHANH, LSDSLALRL, ASVLFSADNNVKFEI, and QINFNHANHNWVSTL (belong to MHC II) were showed in Tables 1 and 2, respectively. Due to achieving a high score on multiple servers and being antigen and the lack of allergenicity, we estimated that they could act as a proper epitope.

B-cell epitopes prediction

According to the data, WVSTLGISFG, LETRDQINFN, VGEKFYFYGL, and NPRSSNDTKEGRADNRRVDA peptides were found which could be analyzed as predicted B-cell epitopes. The graph was plotted by Bepipred server which serves the yellow areas as B-cell epitopes with a suitable threshold (0.5) and Y-axes showed scores related to the amino acids and X-axes defined positions related to the protein regions. (Supplementary file, B). More results are showed in Table 3.

The Overall result of above-mentioned epitopes

The retrieved results of tables have identified the favored residues from T- and B-cells called LSDSLALRL because it was a common epitope with antigenicity and allergenicity properties. So, we suggest it as a candidate vaccine for the next analysis.

Analysis of Docking

The results of the binding of the best epitope with the desired HLA molecules were observed by the Molecular Virtual Docker software and five models were estimated. The proposed models showed the interaction of the epitope side chains bound with the cavities in the groove of MHC I and MHC II. The energies of the bonding models acquired of linking LSDSLALRL peptide with HLA-0101 that consist of

-26.18 kcal/mol, -18.62 kcal/mol, -12.77 kcal/mol, and -12.56 kcal/mol, respectively. As shown in supplementary file, C, the best scores of docking the peptide to HLA-DRB1*0101 was computed -109.86 kcal/mol, -99.52 kcal/mol, -98.40 kcal/mol, -85.79 kcal/mol. According to the principles of docking energy evaluation, the most negative model obtained from docking results was selected as the best model and had an energy of -26.18 kcal/mol which is related to HLA-0101 MHC class I and -109.86 kcal/mol of HLA-DRB1 0101 MHC class II.

Discussion

We focused on the immunogenic protein of the CadF to design vaccines through bioinformatics tools which can dramatically reduce the number of In vitro tests [15]. The past studies have been declared some efforts to suggest an effective vaccine against *C. jejuni* [14, 16, 17]. Despite many efforts to make the vaccine, there is no approved vaccine against *C. jejuni* that is useful in humans so far [8, 11, 18].

We collected T-cell and B-cell epitopes from different servers and the highest epitopes were elicited to make an effective vaccine against *C. jejuni [19-21]*.

The present study showed the correct topology model based on the phyre2 server that predicts CadF is a stable target. This analysis was done with bioinformatics methods and helps to design novel vaccines according to sequence profile and spatial structure and dimension of protein.

LSDSLALRL epitope was selected as the best potential vaccine candidate and wasn't an allergen. The epitope was located in 135-143 regions and can be interacted with HLA-A0101 according to collected results from many above-mentioned servers. In a partial contrary by Yasmin study, who gained its knowledge based on just IEDB and SYFPEITHI servers on CadF protein, suggested that FRLSDSLAL epitope from the protein can be as good chosen for deigning vaccine [18, 22].

It is clear that our epitope has fairly matched with the epitope presented by Yasmin et al. (77.77% of amino acids are matched, LSDSLALRL and FRLSDSLAL, which are marked by underline) and this similarity can be a higher claim for designing an effective vaccine against *C. jejuni*. Based on the Allertop server, the presented epitope by Yasmin et al. was an allergen, while no allergenicity was observed for our epitope "LSDSLALRL" in this study.

Beside, CadF is a significant protein for the colonization and maximum connection can be detected in the regions of the fibronectin-binding domain including phenylalanine-arginine-leucine-serine (FRLS) residues of CadF. Although only fifty percent of the amino acids of our epitope was identified as the binding site to host cells, multiple servers confirmed that this reign has a high score for developing the vaccine.

According to the aliphatic index, alanine, valine, isoleucine, and leucine amino acids were detected in the structure of the protein, which proposed it as a thermostable protein. These amino acids in thermophilic bacteria, e.g. *C. jejuni*, are significantly higher than that of ordinary proteins [23]. This is another advantage of CadF that can be proposed for the development of the vaccine [24].

Conclusion

We suggest that the CadFprotein of *C. jejuni* can be used to prepare an effective vaccine to prevent the disease. However, to predict an actual vaccine without any side effect, we need to improve our knowledge of the pathogenesis and molecular structure of *C. jejuni* on both in vivo and in vitro studies in association with In silico researches.

Limitations

We have some limitations in the use of some servers.

Abbreviations

OmpA: Outer member proteins; CadF: Campylobacter adhesion to Fibronectin; Fn: Fibronectin;

MHC: Major histocompatibility complex.

Declarations

Authors' Contributions

MMN, SS, MMN, and BB involved in the management of the project, the analysis of data, and writing up the paper. All authors read and approved the final manuscript.

Acknowledgment

We wish to thank the Research Council of Qom University of Medical Sciences.

Conflict of interest

No conflict of interest.

Availability of data and materials

The datasets used and/or analyzed during the current study are available from the corresponding author on reasonable request.

Consent for publication

Not applicable.

Ethics approval and consent to participate

The study was reviewed and approved by Medical Ethics Committee of Qom University of Medical Sciences(Code: IR.MUQ.REC.1398.027).

References

1. Parkhill J, Wren B, Mungall K, Ketley J, Churcher C, Basham D, Chillingworth T, Davies R, Feltwell T, Holroyd S: **The genome sequence of the food-borne pathogen Campylobacter jejuni reveals hypervariable sequences**. *Nature* 2000, **403**(6770):665.

2. Boosinger TR, Powe TA: **Campylobacter jejuni infections in gnotobiotic pigs**. *American journal of veterinary research* 1988, **49**(4):456-458.

3. Parkhill J, Wren BW, Mungall K, Ketley JM, Churcher C, Basham D, Chillingworth T, Davies RM, Feltwell T, Holroyd S: **The genome sequence of the food-borne pathogen Campylobacter jejuni reveals hypervariable sequences**. *Nature* 2000, **403**(6770):665.

4. Kaakoush NO, Castaño-Rodríguez N, Mitchell HM, Man SM: **Global epidemiology of Campylobacter infection**. *Clinical microbiology reviews* 2015, **28**(3):687-720.

5. Altekruse SF, Stern NJ, Fields PI, Swerdlow DL: **Campylobacter jejuni—an emerging foodborne pathogen**. *Emerging infectious diseases* 1999, **5**(1):28.

6. Riddle MS, Guerry P: **Status of vaccine research and development for Campylobacter jejuni**. *Vaccine* 2016, **34**(26):2903-2906.

7. Shams S, Bakhshi B, Moghadam TT: **In silico analysis of the cadF gene and development of a duplex polymerase chain reaction for species-specific identification of Campylobacter jejuni and Campylobacter coli**. *Jundishapur journal of microbiology* 2016, **9**(2).

8. Meunier M, Guyard-Nicodème M, Vigouroux E, Poezevara T, Beven V, Quesne S, Bigault L, Amelot M, Dory D, Chemaly M: **Promising new vaccine candidates against Campylobacter in broilers**. *PloS one* 2017, **12**(11):e0188472.

9. Krause-Gruszczynska M, Van Alphen LB, Oyarzabal OA, Alter T, Hänel I, Schliephake A, König W, Van Putten JP, Konkel ME, Backert S: **Expression patterns and role of the CadF protein in Campylobacter jejuni and Campylobacter coli**. *FEMS microbiology letters* 2007, **274**(1):9-16.

10. Riddle MS, Guerry P: **Status of vaccine research and development for Campylobacter jejuni**. *Vaccine*, **34**(26):2903-2906.

11. Meunier M, Guyard-Nicodème M, Hirchaud E, Parra A, Chemaly M, Dory D: **Identification of novel vaccine candidates against Campylobacter through reverse vaccinology**. *Journal of immunology research* 2016, **2016**.

12. Soria-Guerra RE, Nieto-Gomez R, Govea-Alonso DO, Rosales-Mendoza S: **An overview of bioinformatics tools for epitope prediction: implications on vaccine development**. *Journal of biomedical informatics*, **53**:405-414.

13. Neal-McKinney JM, Samuelson DR, Eucker TP, Nissen MS, Crespo R, Konkel ME: **Reducing Campylobacter jejuni colonization of poultry via vaccination**. *PloS one* 2014, **9**(12):e114254.

14. Tribble DR, Baqar S, Carmolli MP, Porter C, Pierce KK, Sadigh K, Guerry P, Larsson CJ, Rockabrand D, Ventone CH: **Campylobacter jejuni strain CG8421: a refined model for the study of Campylobacteriosis and evaluation of Campylobacter vaccines in human subjects**. *Clinical infectious diseases* 2009, **49**(10):1512-1519.

15. Delfani S, Fooladi I, Ali A, Mobarez AM, Emaneini M, Amani J, Sedighian H: **In silico analysis for identifying potential vaccine candidates against Staphylococcus aureus**. *Clinical and experimental vaccine research* 2015, **4**(1):99-106.

16. Scott DA: **Vaccines against Campylobacter jejuni**. *Journal of Infectious Diseases* 1997, **176**(Supplement_2):S183-S188.

17. Meunier M, Guyard-Nicod□¨me M, Hirchaud E, Parra A, Chemaly M, Dory D: **Identification of novel vaccine candidates against Campylobacter through reverse vaccinology**. *Journal of immunology research*, **2016**.

18. Soria-Guerra RE, Nieto-Gomez R, Govea-Alonso DO, Rosales-Mendoza S: **An overview of bioinformatics tools for epitope prediction: implications on vaccine development**. *Journal of biomedical informatics* 2015, **53**:405-414.

19. Liu L, Lai W, Hu N, Zhang W, Chen T, Wang F, Gu R: **Evaluation of immunological efficiency induced by Campylobacter jejuni PEB1 DNA combined with PEB1 protein in mice**. *Xi bao yu fen zi mian yi xue za zhi= Chinese journal of cellular and molecular immunology*, **30**(6):576-580.

20. Nazifi N, Mousavi SM, Moradi S, Jaydari A, Jahandar MH, Forouharmehr A: **In Silico B Cell and T Cell Epitopes Evaluation of lipL32 and OmpL1 Proteins for Designing a Recombinant Multi-Epitope Vaccine Against Leptospirosis**. *International Journal of Infection*, **5**(2).

21. Ranjbar m, Mosavi Nasab d, ghelian chian a, nazoktabar a, ahmadi n, khoshnevisan r, esfandiari s, vafaee manesh j, akbari a: **Immnoinformatics and epitope prediction methods dynamic science with promising achievements**. *scientific journal of ilam university of medical sciences*, **21**(6):300-309.

22. Yasmin T, Akter S, Debnath M, Ebihara A, Nakagawa T, Nabi AN: **In silico proposition to predict cluster of B-and T-cell epitopes for the usefulness of vaccine design from invasive, virulent and membrane associated proteins of C. jejuni**. *In silico pharmacology* 2016, **4**(1):5.

23. Ikai A: **Thermostability and aliphatic index of globular proteins**. *The Journal of Biochemistry* 1980, **88**(6):1895-1898.

24. Chen D, Kapre S, Goel A, Suresh K, Beri S, Hickling J, Jensen J, Lal M, Preaud J, Laforce M: **Thermostable formulations of a hepatitis B vaccine and a meningitis A polysaccharide conjugate vaccine**

produced by a spray drying method. *Vaccine* 2010, **28**(31):5093-5099.

Tables

Table 1. The list of high scored predicted T cell epitopes using online software and their Vaxijen and Allertop score. The purple highlighted epitopes were repeated in some servers but aren't selected because of being problems in their antigenicity and allergenicity. The blue-colored epitopes are suitable about their traits and the green epitope was considering the most common peptide with correct abilities.

MHCI					
Position Sequence	Allele	Server	Score	Vaxigen score	AllerTOP Score
93 GIDVGEKFY	HLA-A01	SYFPEITHI	26	0.017 (Probable NON-ANTIGEN)	NON-Allergen
110 YEDFSNAAY			26	0.39 (Probable NON-ANTIGEN)	NON-Allergen
61 QLEFGLEHY			25	1.2 (Probable ANTIGEN)	Allergen
25 ITPTLNYNY			22	1.2 (Probable ANTIGEN)	Allergen
79 KTTDITRTY			21	0.5 (Probable ANTIGEN)	Allergen
5 LLCLGLASV	HLA02:01		32	0.43 (Probable ANTIGEN)	NON-Allergen
9 GLASVLFSA			23	0.33 (Probable NON-ANTIGEN)	NON-Allergen
13 VLFSADNNV			23	0.05 (Probable NON-ANTIGEN)	NON-Allergen
247 ILEGHTDNI			23	0.05 (Probable ANTIGEN)	Allergen
42 NRYAPGIRL	HLA-B2705		26	1.68 (Probable ANTIGEN)	NON-Allergen
310 RRVDAKFIL			26	1.45 (Probable ANTIGEN)	NON-Allergen
134 FRLSDSLAL			24	1.35 (Probable ANTIGEN)	Allergen
146 TRDQINFNH			24	0.34 (Probable NON-ANTIGEN)	Allergen
84 TRTYLSAIK			23	0.27 (Probable NON-ANTIGEN)	Allergen
13 FLCLGLASV	HLA-A*02:01	IEDB	0.3	0.30 (Probable NON-ANTIGEN)	NON-Allergen
49 HTDNIGSRA	HLA-A*01:01		0.4	1.49 (Probable ANTIGEN)	Allergen
38 RRVDAKFIL	HLA-B*27:05		0.4	1.45 (Probable ANTIGEN)	NON-Allergen
17 GLASVLFGA	HLA-A*02:01		0.5	0.19 (Probable NON-ANTIGEN)	NON-Allergen
48 YEDFSNAAY	HLA-A*01:01		0.55	0.39 (Probable NON-ANTIGEN)	NON-Allergen

$_{15}$FSADNNVKF	HLA-A0101	NetCTL	1.3198	0.66 (Probable ANTIGEN)	NON-Allergen
$_{25}$ITPTLNYNY			2.376	1.2 (Probable ANTIGEN)	Allergen
$_{17}$QLEFGLEHY			1.380	1.2 (Probable ANTIGEN)	Allergen
$_{79}$KTTDITRTY			1.8915	0.56 (Probable ANTIGEN)	Allergen
$_{80}$TTDITRTYL			1.4347	0.13 (Probable NON-ANTIGEN)	NON-Allergen
$_{110}$YEDFSNAAY			1.7677	0.39 (Probable NON-ANTIGEN)	NON-Allergen
$_{136}$LSDSLALRL			2.0179	1.82 (Probable ANTIGEN)	NON-Allergen

Table2. The selected MHC II class binding epitopes were summarized according to most scores predicted by several servers and assess their antigenicity and allergenicity. The purple highlighted epitopes were repeated in some servers but aren't selected because of being problems in their antigenicity and allergenicity. The blue- colored epitopes are suitable concerning their traits and the green epitope was considering the most common peptide with correct abilities.

MHC II					
Position Sequence	Allele	Server	Score	Vaxigen score	AllertopScore
17 EGHFGFDKTTINPTF	HLA-DRB1*04:01	IEDB	1.70	0.42 (Probable ANTIGEN)	NON-Allergen
16 LEGHFGFDKTTINPT			1.74	0.2 (Probable NON-ANTIGEN)	NON-Allergen
18 GHFGFDKTTINPTFQ			1.76	0.31 (Probable NON-ANTIGEN)	NON-Allergen
23 GLASVLFGADNNVKF			1.77	0.47 (Probable ANTIGEN)	NON-Allergen
24 LASVLFGADNNVKFE			1.77	0.67 (Probable ANTIGEN)	Allergen
25 ASVLFGADNNVKFEI			1.77	0.71 (Probable ANTIGEN)	Allergen
216 FGFDKTTIN	HLA-DRB1*0101	MHC2Pred	1.517	0.33 (Probable NON-ANTIGEN)	Allergen
149 QINFNHANH			1.415	1.08 (Probable ANTIGEN)	NON-Allergen
87 YLSAIKGID			1.248	0.065 (Probable NON-ANTIGEN)	Allergen
305 GRADNRRVD			1.157	2.78 (Probable ANTIGEN)	NON-Allergen
260 YNQKLSERR	HLA-DRB1*0101		1.117	1.60 (Probable ANTIGEN)	NON-Allergen
187 PQAKCPVEP			0.901	0.053 (Probable NON-ANTIGEN)	Allergen
236 KVLDENERY			0.786	0.15 (Probable NON-ANTIGEN)	Allergen
36 GNLDMDNRY			0.785	0.23 (Probable NON-ANTIGEN)	Allergen
85 RTYLSAIKGIDVGEK	HLA DRB10101	SYFPEITHI	30	0.13 (Probable NON-ANTIGEN)	NON-Allergen
99 KFYFYGLAGGGYEDF			28	0.72 (Probable ANTIGEN)	NON-Allergen
131 GVKFRLSDSLALRLE			27	2.11 (Probable ANTIGEN)	Allergen
156 NHNWVSTLGISFGFG			27	0.86 (Probable ANTIGEN)	Allergen

$_{37}$NLDMDNRYAPGIRLG			26	1.30 (Probable ANTIGEN)	Allergen
$_{149}$QINFNHANHNWVSTL	HLA DRB1-0401		28	0.63 (Probable ANTIGEN)	NON-Allergen
$_{213}$EGHFGFDKTTINPTF			28	0.42 (Probable NON-ANTIGEN)	NON-Allergen
$_{276}$LEKYGVEKSRIKTVG			28	0.50 (Probable ANTIGEN)	NON-Allergen
$_{11}$ASVLFSADNNVKFEI			26	0.58 (Probable ANTIGEN)	NON-Allergen
$_{18}$DNNVKFEITPTLNYN			26	1.36 (Probable ANTIGEN)	Allergen
$_{135}$LSDSLALRL	HLA-A0101	NETMCH	0.15	1.82 (Probable ANTIGEN)	NON-Allergen
$_{250}$HTDNIGSRA			0.17	1.49 (Probable ANTIGEN)	Allergen
$_{109}$YEDFSNAAY			0.40	0.39 (Probable NON-ANTIGEN)	NON-Allergen

Table3. The list of high scored predicted B cell epitopes using online software and their Vaxijen and Allertopscores.The purple highlighted epitopes were repeated in some servers but aren't selected because of being problems in their antigenicity and allergenicity. The blue- colored epitopes are suitable with regard to their traits and the green epitope was considering the most common peptide with correct abilities.

Position Sequence	Server	Score	Vaxigen score	Allertop Score
16 IFLCLGLASVLFG	IEDB	13	0.36 (Probable NON-ANTIGEN)	NON-Allergen
56 APGVRLGYHFDD		12	0.81 (Probable ANTIGEN)	NON-Allergen
113 FSNAAYDNKS	ABCpred	0.81	0.29 (Probable NON-ANTIGEN)	NON-Allergen
19 NNVKFEITPT		0.79	1.33 (Probable ANTIGEN)	Allergen
159 WVSTLGISFG		0.79	0.62 (Probable ANTIGEN)	NON-Allergen
118 YDNKSGGFGH		0.78	0.79 (Probable ANTIGEN)	NON-Allergen
144 LETRDQINFN		0.76	0.89 (Probable ANTIGEN)	NON-Allergen
96 VGEKFYFYGL		0.75	0.52 (Probable ANTIGEN)	Allergen
122 SGGFGHYGAG		0.75	0.82 (Probable ANTIGEN)	NON-Allergen
170 GGKKEKAVEEVADTRPAPQA	BCPREDS	0.3	0.30 (Probable NON-ANTIGEN)	Allergen
295 NPRSSNDTKEGRADNRRVDA		0.4	1.49 (Probable ANTIGEN)	NON-Allergen
214 GHFGFDKTTINPTFQEKIKE		0.4	1.45 (Probable ANTIGEN)	NON-Allergen
70 SDVKYTNTNKTTDITRTYLS		0.5	0.19 (Probable NON-ANTIGEN)	NON-Allergen
93 GIDVGEKFYFYGLAGGGYED		0.55	0.39 (Probable NON-ANTIGEN)	NON-Allergen
115 NAAYDNKSGGFGHYGAGVKF		0.55	0.39 (Probable NON-ANTIGEN)	Allergen
144 LETRDQINFNHANHNWVSTL		0.55	0.39 (Probable NON-ANTIGEN)	Allergen
39 DMDNRYAPGVRLGYHFDDFW		0.55	0.39 (Probable NON-ANTIGEN)	NON-Allergen
237 VLDENERYDTILEGHTDNIG		0.55	0.39 (Probable NON-ANTIGEN)	NON-Allergen
15 FGADNNVKFEITPTLNYNYF		0.55	0.39 (Probable NON-ANTIGEN)	Allergen

$_{319}$NDTKEGRADNRRVDAKFILR	SVMTrip	1.000	2.04 (Probable ANTIGEN)	Allergen
$_{270}$HTDNIGSRAYNQKLSERRAK		0.822	1.43 (Probable ANTIGEN)	NON-Allergen
$_{21}$KKIFLCLGLASVLFGADNNV		0.548	0.1 (Probable NON-ANTIGEN)	NON-Allergen

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- Supplementaryfile.docx