

In silico analysis of epitope-based CadF vaccine design against *Campylobacter jejuni*

Mona Moballegh Naseri

Qom University of Medical Sciences and Health Services

Saeed Shams (✉ sshamsmed@gmail.com)

Qom University of Medical Sciences <https://orcid.org/0000-0002-7552-0938>

Mohammad Moballegh Naseri

Shahab Danesh University of Qom

Bitia Bakhshi

Tarbiat Modares University Faculty of Medical Sciences

Research note

Keywords: Campylobacter jejuni, Vaccine, CadF, In silico, Bioinformatics, Epitope

Posted Date: November 6th, 2020

DOI: <https://doi.org/10.21203/rs.3.rs-45952/v3>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

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

Abstract

Objective: Vaccination is an important strategy for the eradication of infectious diseases. CadF protein of *Campylobacter jejuni* is one of the important factors in the pathogenesis of this bacterium. The purpose of this work was to perform a bioinformatics study to identify an epitope-based CadF vaccine, as a subunit vaccine. Full protein sequences of CadF were extracted from the NCBI and UniProt databases and subjected to *in silico* evaluations, including sequence analysis, allergenicity, antigenicity, epitope conservancy, and molecular docking assessments done by different servers.

Results: The results showed that CadF was a highly conserved protein belonging to the outer membrane proteins superfamily. Among the evaluated epitopes, LSDSLALRL was identified as an antigenic and non-allergenic peptide with a suitable structure for vaccine development. It was also able to stimulate both T and B cells. This 9-mer peptide was located in 136-144 segment of CadF protein and interacted with both HLA-A 0101 and HLA-DRB1 0101 alleles. Overall, the obtained theoretical results showed that CadF protein could be used for designing and evaluating a new effective vaccine against *C. jejuni*.

Introduction

Campylobacter jejuni (*C. jejuni*) is one of the significant pathogens belonging to the genus *Campylobacter*. The bacterium is a Gram-negative, curved, flagellated, and rod-shaped pathogen which can be transmitted to humans through direct contact with animals and consumption of contaminated food, water, and unpasteurized milk [1, 2]. A gastrointestinal problem commonly caused by *C. jejuni*, especially in children, is called campylobacteriosis [3].

Vaccination is unanimously accepted as an appropriate strategy to eradicate and prevent infectious diseases [4]. Among different types of vaccines, subunit vaccines usually contain parts of the target microorganisms and are known to be safe and effective vaccines for humans [5]. These vaccines activate both humoral- and cell-mediated immune mechanisms to protect them against pathogens. To obtain novel epitope-based subunit vaccines, the identification and prediction of antigenic epitopes by bioinformatics tools are recommended. These provide new theoretical approaches for the design of the vaccines based on immunological databases [6-11].

CadF (*Campylobacter* adhesion to fibronectin) as one of the virulence factors of *C. jejuni* is a conserved, genus-specific, and 37-kDa outer membrane protein that binds to fibronectin and facilitates bacterial colonization of host cells. The protein is also identified as an antigen and could induce massive immune responses [6, 12, 13].

Although there are some studies on the development of vaccine candidates based on the subunit vaccines of *C. jejuni* [14, 15], little is known about CadF potential to be independently considered in the expansion of a protective vaccine.

This study aimed to analysis CadF protein in order to identify epitope-based peptide candidates and evaluate its proteomic database using bioinformatics tools and servers for developing a new vaccine candidate. Therefore, the work was solely an "*in silico*" study.

Methods

Protein sequences evaluation

CadF protein sequences were obtained from both NCBI (Accession numbers: AGI56319.1, AAD28174.1, CAL35585.1, CAL35585.1, QFZ66584.1, QMX64581.1, EEA6200125.1, WP_004314268.1, EEO9446683.1, ACP52761.1) (<https://www.ncbi.nlm.nih.gov/protein>) and UniProt databases (Accession numbers: Q0P8D9, A0A649PTZ6, A0A430SVE4, A0A5T0DIN6, A0A5T1CIC0, A0A6H8RR06) (<https://www.uniprot.org>) in FASTA format. Evolutionary analysis was performed by multiple alignment and phylogenetic tree of the sequences using the ClustalW2 (<https://www.ebi.ac.uk/Tools/msa/clustalw2>) tool and Molecular Evolutionary Genetics Analysis software Version 7 (MEGA 7).

Protein characterization

The three-dimensional structures and biological functions of the protein were recognized by Phyre2 (www.sbg.bio.ic.ac.uk/phyre2) as an online protein fold recognition server. The protein structures were also analyzed by the PSIPRED (<http://bioinf.cs.ucl.ac.uk/psipred>) server.

Two TMHMM (<http://www.cbs.dtu.dk/services/TMHMM>) and ProtParam (<https://web.expasy.org/protparam>) servers were used to predict exo-membrane amino acid sequences and physico-biochemical characteristics of CadF protein.

Allergenicity and antigenicity assessment

The AllerTOP (www.ddg-pharmfac.net/AllerTOP) and AllergenFP ([ddg-pharmfac.net/AllergenFP](http://www.ddg-pharmfac.net/AllergenFP)) web servers were used to determine the allergenicity of CadF protein and common peptides. The AllergenFP was databased to obtain a set of options for predicting allergens. The VaxiJen (<http://www.ddg-pharmfac.net/vaxijen/VaxiJen/VaxiJen.html>) server was also used to forecast the antigenicity of the sequences.

Epitope conservancy assessment

The MHC I and MHC II (Major Histocompatibility Complex) epitopes were analyzed by the Immune Epitope Database and Analysis Resource or IEDB (<https://www.iedb.org/>), NetCTL (www.cbs.dtu.dk/services/NetCTL), NetMHC (www.cbs.dtu.dk/services/NetMHC), NHLApred (<http://crdd.osdd.net/raghava/nhlaped/>), SYFPEITHI (www.syfpeithi.de), and MHC2Pred (<http://crdd.osdd.net/raghava/mhc2pred/>) online servers.

The B cell epitopes were identified using the IEDB, SVMTriP (<http://sysbio.unl.edu/SVMTriP/>), and BCPREDS (<http://ailab.ist.psu.edu/bcpred/>) servers by setting a default specificity of 75%; the threshold value of 0.5 was considered for ABCpred (<http://crdd.osdd.net/raghava/abcpred>) server. Linear and discontinuous B cell epitopes were also predicted by the BepiPred (<http://www.cbs.dtu.dk/services/BepiPred>) server. This server was applied to predict B cell epitopes through the combination of a Hidden Markov Model (HMM) and a propensity scale method. Each epitope identified by these servers was checked to determine the allergenicity and antigenicity properties. The identified common epitopes were analyzed as predicted epitopes, and finally, the best common peptides were selected.

Molecular docking of adopted epitope and alleles

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 UniProt accession numbers Q5SUL5 and P01911, respectively. The PyMOL Molecular Graphics System (<https://pymol.org>) was used to analyze the three-dimensional structure of the best epitope. Final epitope and alleles were edited by Notepad⁺⁺, and the interaction between them (epitope/ HLA-A 0101 allele of MHC I and epitope/ HLA-DRB1 0101 allele of MHC II) was assessed with the help of Molecular Virtual Docker and Molecular Virtual Viewer software. Finally, the interfaces between the epitope and alleles were selected based on a grid, computed on three axes, including x: -0.16, y: -17.63, and z: -15.67.

Results

Analysis of CadF sequences

The complete sequences of CadF protein contained 319 amino acids, and multiple sequence alignment confirmed that this protein was a highly conserved protein among Campylobacter species. It was shown to belong to the outer membrane proteins superfamily (ompA), and an ompA-like domain was identified in the 193-287 position of the protein. The result of the phylogenetic tree also confirmed CadF classification in the outer membrane proteins superfamily (data not shown).

Characterization of CadF

Using the ProtParam server, the MV (molecular weight) and pI (isoelectric point) parameters were determined as 35979.04 Da and 5.89, respectively. The aliphatic index was 69.12, and the GRAVY (grand average of hydrophobicity) of the protein was -0.679. As a result, the amino acids of CadF protein had hydrophobicity and acidity properties ($pI \leq 7.35$). The aliphatic index included alanine, valine, isoleucine, and leucine amino acids, indicating the thermostability of the protein. Moreover, the TMHMM server data analysis results also confirmed that CadF was an outer membrane protein. According to the obtained result from the Phyre2 server, it was predicted that CadF was a stable target.

The PSIPRED server showed the graphical results of secondary structures of the protein, indicating a sheet, helix, and extracellular transmembrane structure. In addition, the Phyre2 server showed the three-dimensional structure of the modeled CadF with a 97% confidence score and 192 known-domain alignments. The structural content included 16% alpha-helix, 41% beta strands, and 16% disordered regions. Also, the prediction of CadF protein showed a binding site at glutamate-histidine-lysine residues and a large amount of metallic heterogenic sections in its structure.

Evaluation of antigenicity and allergenicity

The score of antigenic prediction of CadF protein was calculated as ~0.79 by the VaxiJen server. The results showed that the protein was probably an antigen and could be used for further analysis. The AllergenFP server data indicated the highest Tanimoto similarity index of 0.82 for the protein; therefore, it could not be an allergen. The AllerTOP server data analysis results also confirmed the finding.

Prediction of T and B cell epitopes

The best score for predicting T cell epitopes was selected from the SYFPEITH, IEDB, NetCTL, NHLAPred, and NetMHC servers. The epitopes of MHC I (A 0101, A 0201, and B 2705) and MHC II (DRB1 0101 and DRB1 0401) were the most frequent epitopes among Iranian alleles that were considered in this study. Using the Kolaskar & Tongaonkar Antigenicity method on the IEDB server, a graph was plotted, suggesting the yellow areas as B cell epitopes (Supplementary File A). According to the obtained results, some epitopes such as VLFGADNNV, GLASVLFGA, LSDSLALRL, etc. were the most common epitopes among T and B cells. Further detailed information about the predicted MHC I, II, and B cell epitopes are presented in Tables 1a, 1b, and 2, respectively. Overall, the results showed that the best epitope was LSDSLALRL located in 136-144 regions of CadF with antigenic properties and no allergenic specifications. The three-dimensional structure of the final epitope painted by the PyMOL is showed in Supplementary File B. Therefore, it was suggested as a candidate vaccine for further analysis.

Analysis of docking

The bindings of the best epitope to the desired HLA molecules were observed by Molecular Virtual Docker software, and five models were estimated. The proposed models showed the interaction of the epitope side chains with the cavities in the groove of MHC I and II (Supplementary File C). The energies of the bonding models resulted from the binding of LSDSLALRL peptide to HLA-A 0101 consisted of -26.18, -18.62, -12.77, and -12.56 kcal/mol. The best scores of the peptide docking to HLA-DRB1 0101 were computed as -109.86, -99.52, -98.40, and -85.79 kcal/mol. According to the principles of docking energy evaluation, the model with the most negative docking results was selected as the best model with energies of -26.18 and -109.86 kcal/mol, which were related to HLA-A 0101 of MHC I and HLA-DRB1 0101 of MHC II, respectively.

Discussion

This study was focused on the immunogenic protein CadF to design a hypothetical vaccine through bioinformatics tools. New candidate epitope-based vaccines identified using these tools could dramatically reduce the number of *in vitro* and *in vivo* tests [16]. The previous studies have reported some candidates to suggest an effective vaccine against *C. jejuni*. Despite many efforts to make a vaccine, no approved vaccine against *C. jejuni* has been developed as suitable so far [17, 18].

T and B cell epitopes were collected from different servers, and the best epitope was elicited to make an effective vaccine against *C. jejuni*. The present study showed the accurate topology model based on the Phyre2 server, predicting CadF as a stable target. This analysis helped design a novel hypothetical vaccine according to the sequence profile, spatial structure, and dimensions of the protein.

LSDSLALRL epitope was selected as the best potential vaccine candidate without any evidence of allergenicity. The epitope was located in 136-144 regions and could interact with HLA-A 0101 according to the results collected from many above-mentioned servers. In a study by Yasmin et al. (2016), gaining their knowledge of CadF protein based on just IEDB and SYFPEITHI servers, FRLSDSLAL epitope of the protein was suggested as a good choice for vaccine development [19].

It is clear that the epitope selected in this study is fairly matched (77.77%) with the epitope presented by Yasmin et al. (LSDSLALRL and FRLSDSLAL, which are marked by underline). This similarity could support the claim of suitability of the selected epitope for designing an effective vaccine against *C. jejuni*. Based on the AllerTOP server, the presented epitope by Yasmin et al. could probably be estimated as an allergen, while no allergenicity was observed for epitope "LSDSLALRL" in this study.

In addition, CadF is a significant protein for colonization, and maximum attachment could be detected in regions of the fibronectin-binding domain, including phenylalanine-arginine-leucine-serine (FRLS) residues of the protein [20]. Although only 50% of the selected epitope amino acids were identified as the binding site to host cells, multiple servers confirmed that this region had a high score for vaccine development.

According to the aliphatic index, alanine, valine, isoleucine, and leucine amino acids were detected in the protein structure, proposing it as a thermostable protein. These amino acids in thermophilic bacteria, e.g. *C. jejuni*, are significantly higher than that of ordinary proteins [21]. This proposes another advantage of CadF for the development of an effective vaccine. Heat stability is an important feature in vaccine production, which can simplify the logistics of vaccine distribution and expand the immunization coverage [22].

Conclusion

It is suggested that CadF protein of *C. jejuni* could be used to prepare an effective vaccine for disease prevention. However, to predict an actual vaccine without any side effect, knowledge of the pathogenesis and molecular structure of *C. jejuni* needs to be improved through *in vitro* and *in vivo* studies in parallel with *in silico* research.

Limitations

There were some limitations in the use of some servers. In addition, due to the limited funding and current facilities of our laboratory, it was not possible to validate the results through in vitro and in vivo projects.

Abbreviations

OmpA: Outer member proteins; CadF: *Campylobacter* adhesion to Fibronectin; MHC: Major histocompatibility complex; MV: molecular weight; pI: isoelectric point; kcal/mol: Kilocalorie/mole; NCBI: National Center for Biotechnology Information.

Declarations

Acknowledgment

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

Authors' Contributions

SS developed and supervised the work. 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.

Funding

The study was supported by Research Council of Qom University of Medical Sciences.

Competing interests

The authors declare that they have no competing interests.

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 to publish

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.1399.027).

References

1. Kaakoush NO, Castao-Rodriguez N, Mitchell HM, Man SM: Global epidemiology of Campylobacter infection. *Clinical microbiology reviews*, 28(3):687-720.
2. Ghorbanalizadgan M, Bakhshi B, Shams S, Najar-Peerayeh S: Pulsed-field gel electrophoresis fingerprinting of Campylobacter jejuni and Campylobacter coli strains isolated from clinical specimens, Iran. *International Microbiology* 2019, 22(3):391-398.
3. Altekrouse SF, Stern NJ, Fields PI, Swerdlow DL: Campylobacter jejuni—an emerging foodborne pathogen. *Emerging infectious diseases* 1999, 5(1):28.
4. Hinman A: Eradication of vaccine-preventable diseases. *Annual review of public health* 1999, 20(1):211-229.
5. Moyle PM, Toth I: Modern subunit vaccines: development, components, and research opportunities. *ChemMedChem* 2013, 8(3):360-376.
6. 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.
7. 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).
8. Ranjbar m, Mosavi Nasab d, ghelian chian a, nazoktabar a, ahmadi n, khoshnevisan r, esfandiari s, vafae manesh j, akbari a: Immunoinformatics and epitope prediction methods dynamic science with promising achievements. *scientific journal of ilam university of medical sciences*, 21(6):300-309.
9. Amani J, Delfani S, Fooladi AA: In silico analysis for identifying potential vaccine candidates against Staphylococcus aureus. *Journal of Cell and Molecular Research* 2014, 6(2):57-64.
10. Gupta N, Kumar A: Designing an efficient multi-epitope vaccine against Campylobacter jejuni using immunoinformatics and reverse vaccinology approach. *Microbial Pathogenesis* 2020, 147:104398.
11. Delfani S, Imani Fooladi AA, 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.
12. 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.
13. 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).
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. Riddle MS, Guerry P: Status of vaccine research and development for *Campylobacter jejuni*. *Vaccine* 2016, 34(26):2903-2906.
16. 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.
17. 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.
18. Poly F, Noll AJ, Riddle MS, Porter CK: Update on *Campylobacter* vaccine development. *Human vaccines & immunotherapeutics* 2019, 15(6):1389-1400.
19. 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.
20. Konkel ME, Christensen JE, Keech AM, Monteville MR, Klena JD, Garvis SG: Identification of a fibronectin binding domain within the *Campylobacter jejuni* CadF protein. *Molecular microbiology* 2005, 57(4):1022-1035.
21. Ikai A: Thermostability and aliphatic index of globular proteins. *The Journal of Biochemistry* 1980, 88(6):1895-1898.
22. 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. Predicted epitopes of MHC I (a) and MHC II (b) and their antigenicity and allergenicity properties. The most common peptides are marked in highlights.

a:

Position	Sequence	Allele	Server	VaxiJen	AllerTOP
93	GIDVGEKFY	HLA-A 0101	SYFPEITHI	0.017 (Probable non-antigen)	Probable non-allergen
110	YEDFSNAAY			-0.39 (Probable non-antigen)	Probable non-allergen
61	QLEFGLEHY			1.28 (Probable antigen)	Probable allergen
25	ITPTLNINY			1.23 (Probable antigen)	Probable allergen
79	KTTDITRTY			0.56 (Probable antigen)	Probable allergen
5	FLCLGLASV	HLA-A 0201		0.30 (Probable antigen)	Probable non-allergen
9	GLASVLFGA			0.19 (Probable non-antigen)	Probable non-allergen
13	VLFGADNNV			0.19 (Probable non-antigen)	Probable non-allergen
247	ILEGHTDNI			0.50 (Probable antigen)	Probable allergen
42	NRYAPGVRL	HLA-B 2705		1.43 (Probable antigen)	Probable non-allergen
310	RRVDAKFIL			1.45 (Probable antigen)	Probable non-allergen
134	FRLSDSLAL			1.35 (Probable antigen)	Probable allergen
146	TRDQINFNH			0.34 (Probable non-antigen)	Probable allergen
84	TRTYLSAIK			-0.27 (Probable non-antigen)	Probable allergen
13	FLCLGLASV	HLA-A 0201	IEDB	0.30 (Probable non-antigen)	Probable non-allergen
251	HTDNIGSRA	HLA-A 0101		1.49 (Probable antigen)	Probable allergen
310	RRVDAKFIL	HLA-B 2705		1.45 (Probable antigen)	Probable non-allergen
9	GLASVLFGA	HLA-A 0201		0.19 (Probable non-antigen)	Probable non-allergen
110	YEDFSNAAY	HLA-A 0101		-0.39 (Probable non-antigen)	Probable non-allergen
15	FGADNNVKF	HLA-A 0101	NetCTL	1.00 (Probable antigen)	Probable non-allergen
25	ITPTLNINY			1.23 (Probable antigen)	Probable allergen
61	QLEFGLEHY			1.28 (Probable antigen)	Probable allergen
79	KTTDITRTY			0.56 (Probable antigen)	Probable allergen
80	TTDITRTRYL			0.13 (Probable non-antigen)	Probable non-allergen
110	YEDFSNAAY			-0.39 (Probable non-antigen)	Probable non-allergen
136	LSDSLALRL			1.82 (Probable antigen)	Probable non-allergen
136	LSDSLALRL	HLA-A 0101	NetMHC	1.82 (Probable antigen)	Probable non-allergen
251	HTDNIGSRA			1.49 (Probable antigen)	Probable allergen

b:

Position	Sequence	Allele	Server	VaxiJen	AllerTOP
213	EGHFGFDKTTINPTF	HLA-DRB1 0401	IEDB	0.42 (Probable antigen)	Probable non-allergen
129	GAGVKFRLSDSLALR			2.15 (Probable antigen)	Probable non-allergen
212	LEGHFGFDKTTINPT			0.29 (Probable non-antigen)	Probable non-allergen
214	GHFGFDKTTINPTFQ			0.31 (Probable non-antigen)	Probable non-allergen
9	GLASVLFGADNNVKF			0.47 (Probable antigen)	Probable non-allergen
10	LASVLFGADNNVKFE			0.67 (Probable antigen)	Probable allergen
11	ASVLFGADNNVKFEI			0.71 (Probable antigen)	Probable allergen
216	FGFDKTTIN	HLA-DRB1 0101	MHC2Pred	0.33 (Probable non-antigen)	Probable allergen
149	QINFNHANH			1.08 (Probable antigen)	Probable non-allergen
87	YLSAIKID			-0.06 (Probable non-antigen)	Probable allergen
305	GRADNRRVD			2.78 (Probable antigen)	Probable non-allergen
260	YNQKLSERR			1.60 (Probable antigen)	Probable non-allergen
187	PQAKCPVEP			0.05 (Probable non-antigen)	Probable allergen
236	KVLDENERY			-0.15 (Probable non-antigen)	Probable allergen
36	GNLMDNRY			0.23 (Probable non-antigen)	Probable allergen
85	RTYLSAIKIDVGEK	HLA-DRB1 0101	SYFPEITHI	0.13 (Probable non-antigen)	Probable non-allergen
99	KFYFYGLAGGGYEDF			0.72 (Probable antigen)	Probable non-allergen
131	GVKFRLSDSLALRLE			2.11 (Probable antigen)	Probable allergen
156	NHNWVSTLGISFGFG			0.86 (Probable antigen)	Probable allergen
37	NLDMNRYAPGVRLG			1.19 (Probable antigen)	Probable allergen
149	QINFNHANHNWVSTL			0.63 (Probable antigen)	Probable non-allergen
213	EGHFGFDKTTINPTF			0.42 (Probable antigen)	Probable non-allergen
276	LEKYGVEKSRIKTVG			0.50 (Probable antigen)	Probable non-allergen
11	ASVLFGADNNVKFEI			0.71 (Probable antigen)	Probable non-allergen
18	DNNVKFEITPTLNYN			1.36 (Probable antigen)	Probable allergen

Table 2. Predicted epitopes of B cell and their antigenicity and allergenicity properties. The most common peptides are marked in highlights.

Position	Sequence	Server	Vaxijen	AllerTOP
4	IFLCLGLASVLFG	IEDB	0.36 (Probable non-antigen)	Probable non-allergen
45	APGVRLGYHFDD		0.81 (Probable antigen)	Probable non-allergen
127	HYGAGVKFRLSDSLALRL		1.89 (Probable antigen)	Probable allergen
136	LSDSLAL	ABCpred	0.55 (Probable antigen)	Probable non-allergen
19	NNVKFEITPT		1.33 (Probable antigen)	Probable allergen
159	WVSTLGISFG		0.62 (Probable antigen)	Probable non-allergen
118	YDNKSGGFGH		0.79 (Probable antigen)	Probable non-allergen
144	LETRDQINFN		0.89 (Probable antigen)	Probable non-allergen
96	VGEKFYFYGL		0.52 (Probable antigen)	Probable allergen
122	SGGFGHYGAG		0.82 (Probable antigen)	Probable non-allergen
170	GGKKEKAVEEVADTRPAPQA		BCPREDS	1.29 (Probable antigen)
295	NPRSSNDTKEGRADNRRVDA	2.42 (Probable antigen)		Probable non-allergen
214	GHFGFDKTTINPTFQEKIKE	0.32 (Probable non-antigen)		Probable non-allergen
70	SDVKYNTNTNKTTDITRITYLS	0.69 (Probable antigen)		Probable non-allergen
93	GIDVGEKFYFYGLAGGGYED	0.65 (Probable antigen)		Probable non-allergen
115	NAAYDNKSGGFGHYGAGVKF	0.91 (Probable antigen)		Probable allergen
144	LETRDQINFNHNHNWVSTL	0.44 (Probable antigen)		Probable allergen
39	DMDNRYAPGVRLGYHFDDFW	0.78 (Probable antigen)		Probable non-allergen
237	VLDENERYDTILEGHTDNIG	0.33 (Probable non-antigen)		Probable non-allergen
300	NDTKEGRADNRRVDAKFILR	SVMTriP		2.04 (Probable antigen)
251	HTDNIGSRAYNQKLSERRAK		1.43 (Probable antigen)	Probable non-allergen
2	KKIFLCLGLASVLFADNNV		0.10 (Probable non-antigen)	Probable non-allergen

Supplementary Files

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

- [Supplementaryfile.docx](#)
- [Supplementaryfile.docx](#)