Contents lists available at SciOpen

Food Science and Human Wellness

journal homepage: https://www.sciopen.com/journal/2097-0765

China Food Publishing Co.

Molecular characteristics analysis and B-cell linear epitopes, key amino acids identification of the sesame allergen Ses i 3

Ning Yu[a,1], Yufei Qin[a,b,1], Wenhan Kang[a], Jiukai Zhang[a], Lijia Chen[c], Hongtian Wang[c], Xiaoyan Wang[c,*], Ying Chen[a,*]

[a] Chinese Academy of Quality and Inspection & Testing, Beijing 100176, China
[b] College of Food Science and Engineering, Nanjing University of Finance and Economics, Nanjing 210023, China
[c] Food Allergy Department, Beijing Shijitan Hospital, Capital Medical University, Beijing 100038, China

ARTICLE INFO

Article history:
Received 22 December 2023
Received in revised form 4 March 2024
Accepted 29 April 2024

Keywords:
Sesame allergen
Ses i 3
Molecular characteristics
Epitope identification
Slot blot immuno-microarrays
Key amino acids

ABSTRACT

Sesame is one of the eight major allergens that cause food allergies. Study of the epitopes of sesame allergens is important for understanding their sensitization mechanisms. Currently, less information is available on the epitope studies of sesame allergens. In this study, we analyzed the molecular characteristics, structure and homology of Ses i 3, one of the important sesame allergens. We predicted the B-cell linear epitopes of Ses i 3 using bioinformatics tools and characterized them by slot blot immuno-microarrays technology. Eight peptides as B-cell linear epitopes of Ses i 3 were identified, in addition, key amino acids in these epitopes were predicted and leucine 422 was identified as a key amino acid. The present work will contribute to further understanding of the sesame allergen and provide some help in the prevention and treatment of sesame allergy.

1. Introduction

Sesame has a high nutritional value which is rich in fat, protein and other nutrients and is one of the main oil crops[1]. In 2019, global sesame cultivation was 12.8 million hectares and production was 6.5 million tons[2]. However, sesame seeds can also cause severe allergic reactions, the main symptoms of sesame allergy include: rash, urticaria, asthma, etc., which can seriously lead to anaphylaxis or even death[2-3]. According to the 2022 recommendations of the Food and Agriculture Organization of the United Nations and the World Health Organization (FAO/WHO), sesame is listed as one of the eight new categories of allergens[4].

Sesame seeds mainly contain seven allergens, which are named Ses i 1-Ses i 7 by International Union of Immunological Societies (IUIS)[3]. Among them, Ses i 3 is recognized by the serum of 75% of sesame allergic patients and is a very important sesame allergen[4]. Ses i 3 is a 7S-like pea globulin[5], belonging to the Cupin superfamily. Pea globulin is a flattened trimer of three individual peptide chains bound into a triangle and the Cupin superfamily is characterized by a three-dimensional structure with a β-folded barrel-shaped protein supersecondary structure. Ses i 3 consists of 585 amino acid residues, of which amino acids 1–23 are signal peptides. Ses i 3 has a molecular weight of 45 kDa and is resistant to trypsin digestion[6-7].

Sesame allergy is a type I hypersensitivity reaction mediated by immunoglobulin E (IgE)[8], which is divided into a sensitizing phase and an effector phase[9]. When the sesame allergen comes into contact with the body for the first time, the body is stimulated to produce specific immunoglobulin E (sIgE), which is the sensitization phase[10]. When the same allergenic food re-enters the organism, it is recognized

[1] These authors contributed equally to this work.
[*] Corresponding authors.
E-mail address: chenyingcaiq@163.com (Y. Chen); wxy100083@163.com (X.Y. Wang)
Peer review under responsibility of Beijing Academy of Food Sciences.

SciOpen Publishing services by Tsinghua University Press

http://doi.org/10.26599/FSHW.2024.9250265

and bound by IgE antibodies on the surface of mast cells and basophils, this leads to the release of pro-inflammatory factors, causing an allergic reaction, a phase known as the elicitation phase. sIgE binds to the receptor on the surface of mast cells or basophils, and when the body is exposed to the allergen again, it will stimulate the release of various inflammatory mediators from this binding body, thus leading to an allergic reaction in the body, which is the effector phase[11-13].

Allergen sensitization is closely related to epitope[14]. Epitopes are antigenic determinants on the surface of a protein, usually consisting of 5–17 amino acid residues or 5–7 polysaccharide residues[15]. Epitopes can be classified into T-cell epitopes and B-cell epitopes according to immunological properties[16], and linear epitopes and conformational epitopes according to structural properties, respectively[17]. The main techniques for localizing antigenic epitopes include bioinformatics technology, peptide scanning techniques and bacterial surface display techniques, etc[17-18]. Among these methods, bioinformatics techniques are used to predict antigenic epitopes based on the physicochemical properties of proteins such as hydrophilicity, flexibility and secondary structure, and have been gaining more and more attention in recent years due to their rapidity, accurate and easy to operate features[19]. Fu et al.[20] predicted B-cell linear epitopes in tropomyosin and arginine kinase in Chinese shrimp by bioinformatics tools, and identified them by dot blot, enzyme-linked immunosorbent assay (ELISA), etc. In addition, epitopes of allergens were screened using bioinformatics tools, such as shrimp allergen Lit v 1[21], Exo m 1[22], walnut allergen Jug r 1[23], etc. So far, the epitope identification of the sesame allergen Ses i 3 has not been reported.

In this study we predicted the secondary structure, hydrophilicity, surface accessibility, flexibility, and antigenicity of the sesame allergen Ses i 3 and analyzed its homology, as well as predicted and mapped the spatial structure of Ses i 3. Furthermore, we used bioinformatics methods to predict B-cell linear epitopes and key amino acids of Ses i 3. Finally, we identified the predicted epitopes by slot blot immuno-microarrays and key amino acid was selected in these epitopes. These studies will provide a reference for the resolution of sesame allergen sensitization mechanisms.

2. Materials and methods

2.1 Reagents and instruments

Horseradish peroxidase (HRP) labeled mouse anti-human IgE, antibody diluent and enhanced chemiluminescence (ECL) were obtained from BioHuge BioTechnology Co., Ltd. (Beijing, China). 3,3'-diaminobenzidine (DAB) chromogenic kit was purchased from Solarbio Co., Ltd. (Beijing, China). Other reagents were all analytical grade. Sesame allergy protein Ses i 3 with purity ≥ 85% (1 mg/mL) was synthesized by Huaan Biotechnology Co., Ltd. (Hangzhou, China). All peptides with purity ≥ 95% were provided by Guoping Pharmaceutical Co., Ltd. (Anhui, China). Sesame allergy patient serum and healthy adult serum were provided by Beijing Shijitan Hospital (Table S1). ChemiDoc™ imaging system and electrophoresis instrument (Bio-Rad, USA), molecular hybridization spotter were purchased from Popular Science Instrument Factory (Beijing, China).

2.2 Sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) and IgE binding capacity analysis of Ses i 3

The molecular weight of synthetic Ses i 3 was characterized by SDS-PAGE. The protein was diluted 5 times with double distilled water and mixed with dithiothreitol (DTT) at a volume ratio of 1:1, and the loading volume was 10 μL per well. Electrophoresis time was first 80 V, 20 min, and then 120 V, 70 min. Subsequently stained with coomassie brilliant blue staining solution for 2 h, after that, decolorized for 24 h with 2–4 changes of decolorization solution in between.

The IgE binding capacity of Ses i 3 was characterized by Western blot. Serum pools were made with 10 sera from sesame allergic patients as primary antibody (serum and bovine serum albumin (BSA) diluted at a ratio of 1:100, *V/V*) and incubated the polyvinylidene fluoride (PVDF) membrane with primary antibody at 4 °C overnight. After that, remove the PVDF membrane and wash it with phosphate buffered saline with Tween-20 (PBST) 3 times for 5 min each time. Mouse anti-human IgE (HRP) diluted in skim milk at a volume ratio of 1:2 000 was used as secondary antibody and the PVDF membrane was incubated with the secondary antibody at 37 °C for 2 h. Then the PVDF membrane was removed and washed 3 times with PBST for 5 min each time. DAB staining solution was added dropwise to the PVDF membrane and left at room temperature for 15 min. Finally, took pictures with ChemiDoc™ imaging system.

2.3 Analysis of the properties and structure of Ses i 3

The amino acid sequence of sesame allergen Ses i 3 were downloaded through the uniprot website. Hydrophilicity, surface accessibility, antigenic index and flexibility of Ses i 3 were analyzed by the subroutine protean of DNAStar software (DNASTAR, Inc., Madison, WI, USA). The secondary structure of Ses i 3 was predicted by the SOPMA secondary structure prediction website (https://npsa-prabi.ibcp.fr/cgi-bin/npsa_automat.pl?page=npsa%20_sopma.html). Homology modeling of Ses i 3 was performed through the SWISS-MODEL online website (https://swissmodel. expasy. org) and its 3D structural model was drawn through pymol software (DeLano Scientific, San Carlos, CA, USA). The stability of the model was assessed by Ramachandran plots. Proteins for species with high homology to Ses i 3 were retrieved from the NCBI website (https://www.ncbi.nlm.nih.gov/), and the amino acid sequences of these species were analyzed using Bioedit 7.2 software (Borland Limited, USA). In addition, phylogenetic tree was constructed by MEGA 7.0 software (Mega Limited, Auckland, New Zealand).

2.4 Epitope prediction of Ses i 3 by bioinformatics tools

Antigenic epitopes of Ses i 3 were predicted by 3 bioinformatics tools, including the DNAStar software, SOPMA secondary structure prediction website and BCPreds antigen epitope prediction website (http://ailab-projects1.ist.psu.edu:8080/bcpred/). Prediction of protein properties of Ses i 3 was used protean, a subroutine of DNAStar software, predicted hydrophilicity by Kyte-Doolittle scheme, flexible regions by Karplus-Schulz scheme, antigenic index by Jameson-Wolf scheme, surface accessibility with the Emini solution. Meanwhile, we used SOPMA secondary structure prediction website and BCPreds online sites to predict Ses i 3 antigen epitopes. Finally, the prediction results of the 3 tools were combined to analyze the possible antigenic epitope regions of Ses i 3.

2.5 Epitope identification of Ses i 3 by slot blot immuno-microarrays

The protein and carbonate buffer solutions (CBS) were mixed in 1:2 (*V*/*V*) ratio and the peptides were dissolved to 2 mg/mL with CBS. Nitrocellulose membrane was moistened in double-distilled water for 15 min, and installed in the extraction molecular hybridization spotter. Spotted 10 μL per well in extraction molecular hybridization spotter and incubated at constant temperature for 2 h. Subsequently, 20 μL of PBST was added to each well to wash off the proteins that were not bound to the nitrocellulose membrane. Then 10 μL BSA was add into per well for closure and leave at room temperature for 2 h. The serum of patients with sesame allergy was diluted with BSA in a ratio of 1:10 (*V*/*V*) as primary antibody (serum pools were made by mixing three healthy adult sera in equal proportions, and were mixed with BSA at a ratio of 1:10 as a negative control), added 10 μL per well and incubated at 37 °C for 2 h. Afterwards, 20 μL of PBST was added to each well to wash off the primary antibody that was not bound to the protein. HRP-labeled mouse anti-human IgE was diluted at 1:2 000 with antibody diluent as secondary antibody, added 10 μL per well and incubated at 37 °C for 2 h. Subsequently, the nitrocellulose membrane was removed and washed for 3 times with PBST for 5 min each time, stained with ECL. Finally, the nitrocellulose membrane was photographed with the ChemiDoc™ Imaging System and its grayscale values were analyzed by the ImageJ software (National Institute of Health, Bethesda, MD, USA). All spiking and cleaning operations were carried out under vacuum filtration. The maximum grayscale value of the protein bound to serum was considered as 100%, while 0%–10% of the maximum grayscale value was considered as 0 point, 10%–40% as 1 point, 40%–70% as 2 points, and 70%–100% as 3 points. In addition, some peptides with a grayscale value greater than that of the protein after binding to serum were considered as 3 points.

2.6 Identification of key amino acids of Ses i 3

Amino acid sequences for 10 allergens with high homology to Ses i 3 were searched on the UniProt website, sequence alignment was performed by MEGA software to identify conserved amino acids. At the same time, amino acids appeared more frequently in epitopes than in the whole allergen were researched[24]. The amino acids that are conserved in the epitopes and occur with high frequency were considered as key amino acids, which were mutated to alanine and verified by slot blot[25-26]. Incubated the peptide before and after mutation on the nitrocellulose membrane and 10 parts of sesame-allergic patient serum were diluted with antibody diluent at a ratio of 1:100 to form a pool of serum as the first antibody, and the rest of the steps were the same as in 2.5, trials were conducted in triplicate.

3. Results and discussion

3.1 SDS-PAGE and IgE binding capacity analysis of Ses i 3

The molecular mass of Ses i 3 was analyzed by SDS-PAGE, and the results were shown in Fig. 1A, the relative molecular mass of the synthesized Ses i 3 is approximately 60 kDa, which is somewhat larger than its theoretical molecular weight, probably due to the vector and tag in the process of prokaryotic expression. The IgE binding ability of recombinant Ses i 3 was characterized by Western blot (Fig. 1A), the recombinant Ses i 3 had a pronounced immunoblot against the sesame allergic patient serum, which indicated that recombinant Ses i 3 was with a strong IgE binding ability to the serum pool of sesame allergic patients.

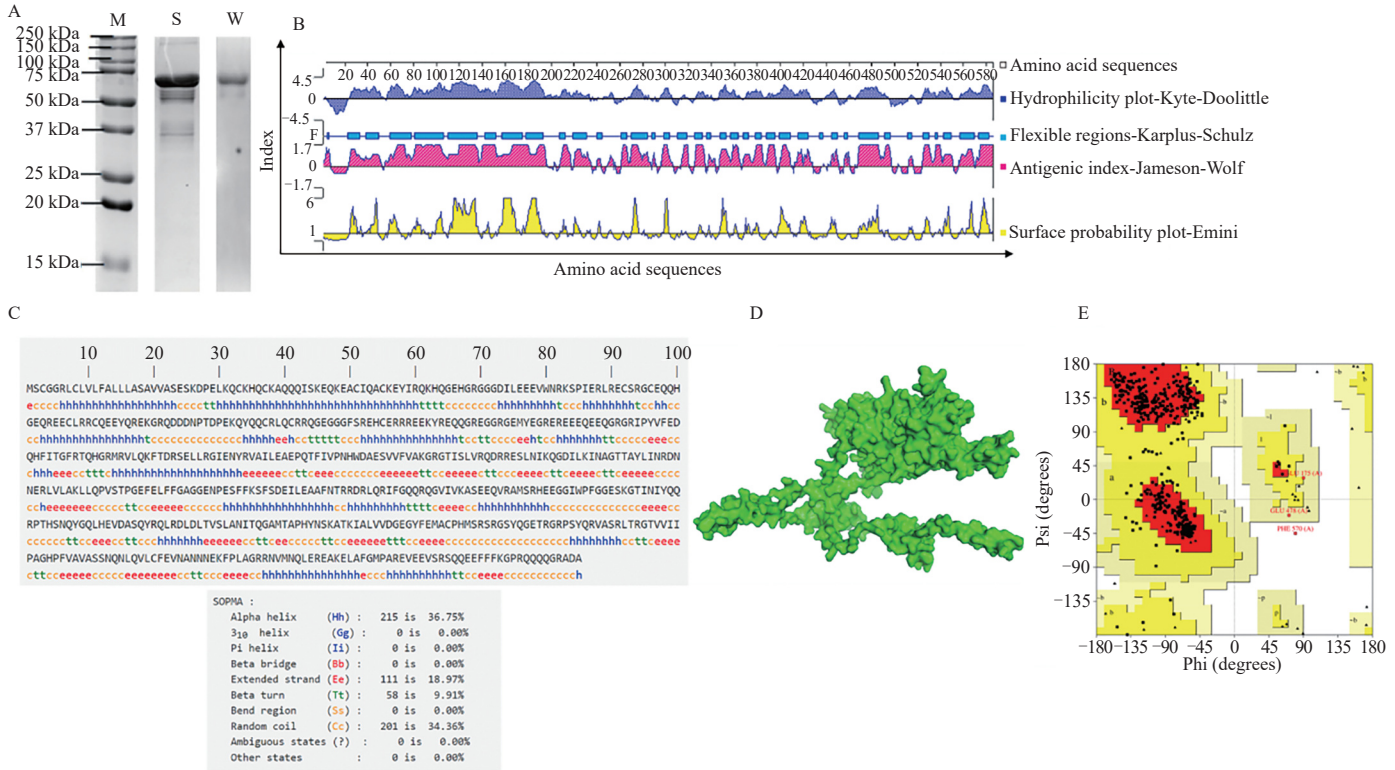


Fig. 1 Structural and property analysis of Ses i 3. (A) Results of SDS-PAGE and Western blot. M, Protein marker. S, SDS-PAGE. W, Western blot. (B) Analysis results of DNAStar. (C) Predicted results for the secondary structure of Ses i 3. (D) Spatial structure model of Ses i 3. (E) Ramachandran plot for the Ses i 3 model.

3.2 Property, structural and cross-reactivity analysis of Ses i 3

The properties of Ses i 3 were analyzed with DNAStar software, and the results were shown in Fig. 1B. The hydrophilicity represents the ability of the protein to attract water molecules or dissolve in water[27], the hydrophilic region of Ses i 3 was more evenly distributed, and the longer continuous hydrophobic region was mainly concentrated in the first half of the amino acid sequence. Protein flexibility refers to the ability of a protein to stretch physical traits by physical means[28], and the analysis of DNAStar software showed that Ses i 3 was with a strong flexibility. The antigenic index is the region in the protein predicted by DNAStar with the Jameson-Wolf algorithm where antigenic epitopes may exist, and the results indicated that the antigenic index was greater than 0 and the regions that may form longer epitopes were mainly concentrated at the ends of the protein. Surface accessibility refers to the possibility of amino acid residues in protein antigens being contacted by solvent molecules[29], and it reflects the distribution of amino acid residues in the internal and external layers of the protein. The results of DNAStar analysis indicated that the regions with surface accessibility greater than 1 in Ses i 3 were distributed in the whole protein, however, the distribution was relatively dispersed. In general, sequences with strong hydrophilicity and surface accessibility are more likely to form epitopes when exposed to the protein surface[19]. Meanwhile, sequences with flexibility and high antigenicity index are more likely to form epitopes[23]. The secondary structure of Ses i 3 was predicted through the SOPMA, the results exhibited that among the secondary structures of Ses i 3, the α-helices accounted for the highest proportion at 36.75%, followed by random coils with 34.36%, β-folds and β-turn accounted for 18.97% and 9.91%, respectively (Fig. 1C).

By retrieving protein models with highly similar sequences to Ses i 3 on SWISS-MODEL online website as a template for 3D modeling, the 3D models were then drawn with PyMOL software. The three-dimensional structure of Ses i 3 was shown in Fig. 1D. Meanwhile, Ramachandran plot was used to assess the stability of the plotted tertiary structure. Ramachandran denotes the dihedral angle adjacent to the α carbon, and the plotted tertiary structure is stable only when the position of the dihedral angle of the α carbon is within a reasonable range[30]. As shown in Fig. 1E, the dihedral angle of α carbon is basically in the red acceptable region, indicating that the tertiary structure map of Ses i 3 is basically stable and reliable.

The presence of the same or similar epitopes between different allergens can easily cause allergic cross-reactivity. Allergens with highly similar amino acid sequences or tertiary structures are also prone to cross-reactivity. In general, if the homology between two allergens is higher than 35%, then it can be assumed that these two allergens may cause cross-reactivity[31]. In this work, 27 proteins of species with high homology to Ses i 3 were retrieved and screened on NCBI. The results were shown in Fig. 2A, it was found that Ses i 3 was homologous with some proteins in perilla, matsumoto, tobacco and azalea with higher homology of 56.9%, 58.7%, 52.5%, 50%, respectively. People who are allergic to sesame seeds may also be allergic to the above species. At the same time, the phylogenetic tree was constructed to further illustrate the homology between Ses i 3 and the other species (Fig. 2B).

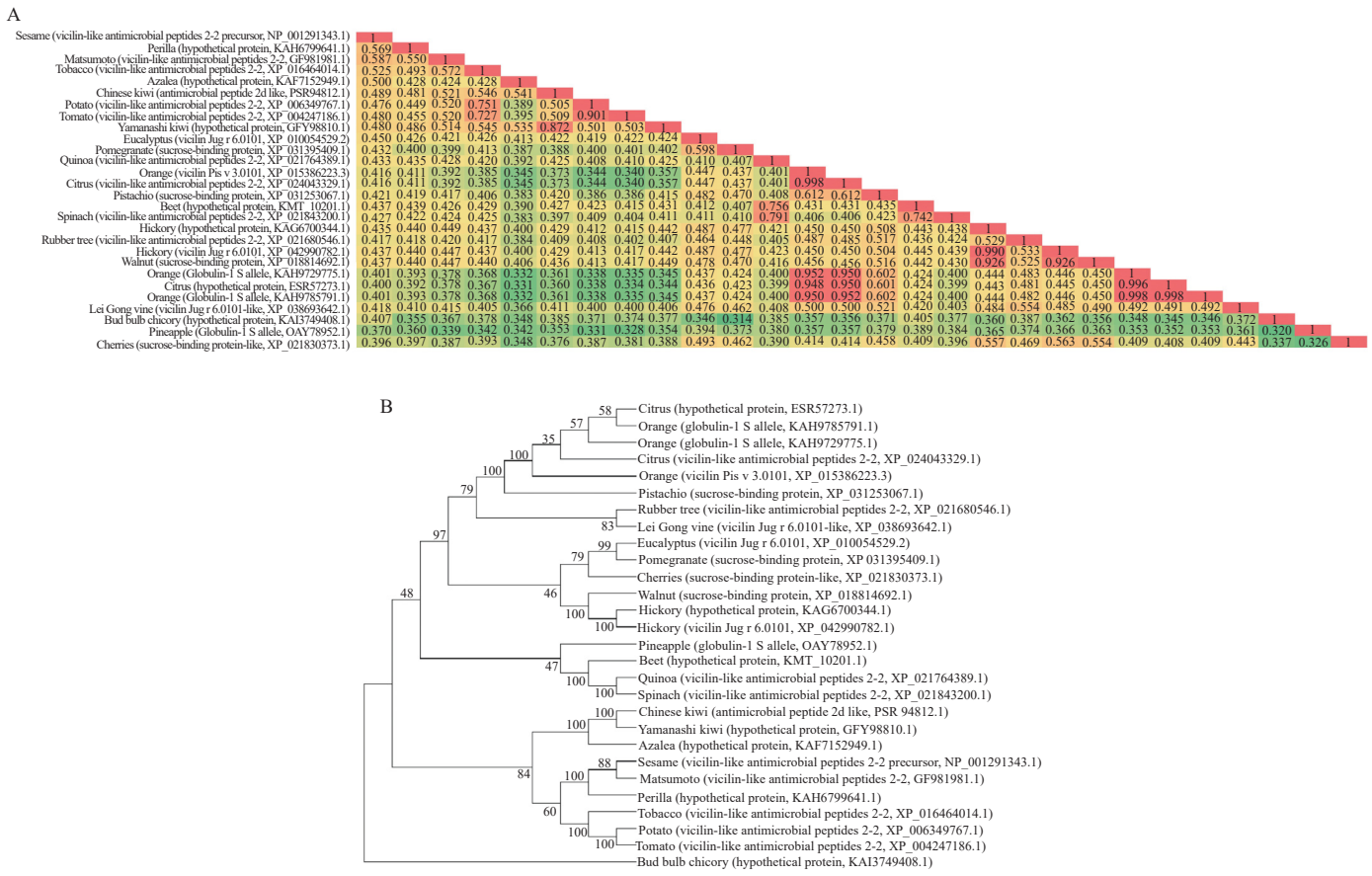


Fig. 2 Homology analysis of proteins from 27 species with Ses i 3 (A) and phylogenetic tree (B).

3.3 Epitope prediction results by bioinformatics tools

Prediction of antigenic epitopes of Ses i 3 was through 3 bioinformatics tools. In the analysis results of DNAStar, the regions with hydrophilicity > 0, surface accessibility index > 1, flexibility > 0, and antigenic index > 0 were used as the predicted basis of antigenic epitopes[24]. Since in the secondary structure of proteins, α-helices and β-folds are more stable and not easily deformed, so it is difficult to bind with antibodies to form epitopes[31]; while β-turns and random coils are easier to bind with antibodies due to their prominent structure, so it is easy to form epitopes, which are selected as the basis for epitope prediction in SOPMA secondary structure prediction. Submitted the amino acid sequence of Ses i 3 to the BCPreds online website, which gives antigenic epitope prediction results directly.

The prediction results of 3 bioinformatics tools were shown in Table S2. Since the length of epitopes is usually 5–17 amino acid sequences, predictions that were too long or too short were excluded. Sequences identified by all three software simultaneously were considered as the final predicted epitopes. The final 10 predicted epitopes were screened, which were peptide 1 with 11 amino acids (AA38-48, KAQQQISKEQK), peptide 2 with 12 amino acids (AA61-72, QKHQGEHGRGGG), peptide 3 with 12 amino acids (AA110-121, RCQEEYQREKGR), peptide 4 with 12 amino acids (AA122-133, QDDDNPTDPEKQ), peptide 5 with 10 amino acids (AA142-151, RRQGEGGGFS), peptide 6 with 16 amino acids (AA157-172, RRREEKYREQQGREGG), peptide 7 with 16 amino acids (AA178-193, EGREREEEQEEQGRGR), peptide 8 with 10 amino acids (AA399-408, QQRPTHSNQY), peptide 9 with 10 amino acids (AA414-423, VDASQYRQLR), and peptide 10 with 18 amino acids (AA469-486, SRSRGSYQGETRGRPSYQ) (Table 1).

Table 1
Predictive results of B-cell linear epitopes of Ses i 3.

Number	Location	Amino acid sequence
1	38-48	KAQQQISKEQK
2	61-72	QKHQGEHGRGGG
3	110-121	RCQEEYQREKGR
4	122-133	QDDDNPTDPEKQ
5	142-151	RRQGEGGGFS
6	157-172	RRREEKYREQQGREGG
7	178-193	EGREREEEQEEQGRGR
8	399-408	QQRPTHSNQY
9	414-423	VDASQYRQLR
10	469-486	SRSRGSYQGETRGRPSYQ

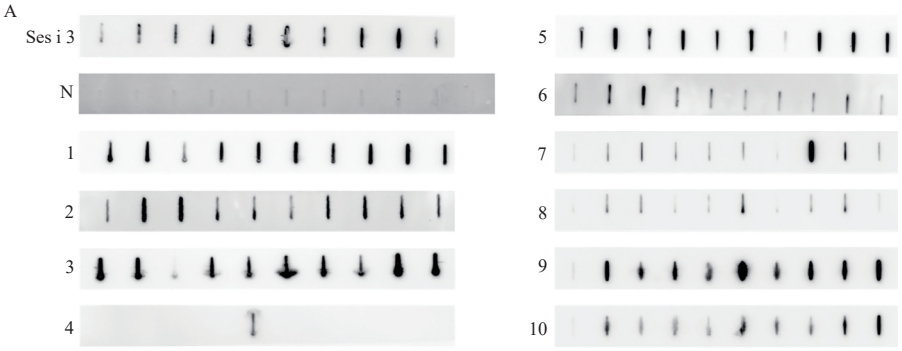
3.4 Epitope identification results by slot blot immuno-microarrays

The IgE binding ability of the predicted epitopes was detected by slot blot immuno-microarrays, and the grayscale values were scored using ImageJ software. Peptides that bound to more than half of the serum or scored greater than the mean plus standard deviation were considered as epitopes. The results were shown in Fig. 3A, the sesame allergen Ses i 3 was recognized by all sera, the binding bands of serum No. 1 and No. 10 were lighter, while serum No. 5, No. 6 and No. 9 had darker binding bands with protein, the binding of different sera to protein and peptides varied widely. Peptides 1, 2, 3, 5, 6, 7, 9, and 10 could be recognized by more than half of the sera. However, peptide 4 was barely bound to serum and only mildly bound to serum No. 5; in addition, peptide 8 was bound to serum No. 2, No. 3, No. 6, and No. 9 with the shallow band, indicated a low binding capacity to serum. Therefore, peptide 4 and peptide 8 were not considered as epitopes of Ses i 3. It is worth noting that some of the sera were weakly bound to the IgE of the protein and more strongly bound to the peptides, such as serum No. 1, No. 2 and No. 10. However, these serums did not bind to all peptides. And the negative control (N) had no visible binding bands, indicating that neither the sesame allergen Ses i 3 nor the peptides produced IgE binding to the serum of healthy adults. Subsequently, the grayscale values of these strips were scored, in which the maximum grayscale value was 121 542.000 and considered as 100%. An immunization scoring matrix was created, where the score of peptide 3 was greater than the mean plus the standard deviation, indicated the strongest ability to bind IgE from serums. Peptide 10 bound weakly to IgE from serums, but still bound to more than half of the serum. According to the scoring criteria described in 2.5, peptides 1, 2, 3, 5, 6, 7, 9 and 10 were finally identified as B-cell linear epitopes of Ses i 3 (Fig. 3B). The results of epitope identification showed that the B-cell linear epitopes of Ses i 3 were mainly concentrated in the amino acid sequences of 1–200 and 400–500. In the prediction results of DNAStar, these regions were the more antigenic regions, indicating that the prediction results of DNAStar were accurate. Predicting the spatial structure of allergens aimed to better understand the correlation between antigenic epitopes and structure. The spatial structure of an allergen is critical to its immunogenicity, and binding of antibodies to epitopes usually requires a stable spatial structure of the allergen. In addition, in order to better understand the link between epitopes and structures, we labeled the exact location of epitopes in the spatial structure of Ses i 3. We found that these epitopes are on the surface of the spatial structure of Ses i 3, which may contribute to the specific binding of these epitopes to the antibody. The results were shown in Fig. 4.

3.5 Identification of key amino acids of Ses i 3

Amino acids that occur more frequently in epitopes than in the entire protein was screened (Fig. 5A). At the same time, through the conservative analysis of amino acid sequences of 11 allergens (Fig. 5B), the amino acids with conserved nature were identified. The part where the results of the two methods coincide was used as the prediction results of key amino acids. The results showed that most of the amino acids appeared more frequently in epitopes than in protein, conserved amino acids appeared essentially after amino acid sequence 212. At last, leucine 422 in peptide 9, tyrosine 485 in peptide 10 were identified as the candidates for key amino acids. These two amino acids were replaced with alanine, and the replaced amino acid sequences were shown in Table S3. Key amino acids were verified by slot blot immuno-microarrays, and grayscale values were also analyzed. The analysis revealed that when leucine 422 was mutated to alanine, a significant decrease in IgE binding capacity of peptide 9 occurred, indicating it was the key amino acid of peptide 9. When tyrosine 485 was mutated to alanine, there was no significant change in the IgE binding capacity of peptide 10 (Fig. 6). In summary, leucine 422 was identified as a key amino acid of Ses i 3.

A



B

Sera	Protein	Peptide									
		1	2	3	4	5	6	7	8	9	10
1	1	2	1	3	0	2	1	0	0	0	0
2	1	3	3	3	0	3	2	1	1	3	1
3	1	1	3	0	0	2	3	1	1	2	1
4	2	2	2	3	0	3	1	1	0	3	1
5	3	3	2	3	1	2	1	0	0	2	1
6	3	3	1	3	0	3	1	1	1	3	2
7	1	2	2	3	0	0	1	0	0	2	1
8	2	3	3	2	0	3	1	3	0	3	1
9	3	3	2	3	0	3	1	1	1	3	2
10	1	2	2	3	0	2	1	0	0	3	3
Positive subjects, *n*	10	10	10	9	1	9	10	6	4	9	9
Intensity	18	24	21	26	1	23	13	8	4	24	13
> 50% of subjects	+	+	+	+		+	+	+		+	+
> Mean intensity + SD				+							

Fig. 3 Epitope identification results for Ses i 3. (A) Slot blot-immunomicroarray results for Ses i 3. Ses i 3, Slot blot-immunoblotting of the sesame allergen Ses i 3 with 10 sesame allergy sera. N, Slot blot-immunoblotting of sesame allergen Ses i 3 and 10 predicted epitope peptides (from left to right were Ses i 3, peptides 1-10) with a pool of three healthy adult sera. 1-10, Slot blot-immunoblotting of 10 predicted epitope peptides with 10 sera from sesame-allergic patients. (B) Immunization score matrix of sesame allergen Ses i 3 and ten predicted epitopes.

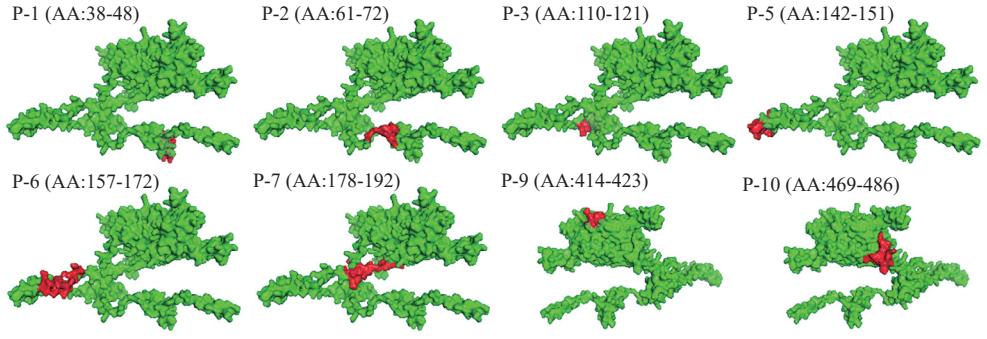


Fig. 4 Position of the epitopes of Ses i 3 in its three-dimensional structure.

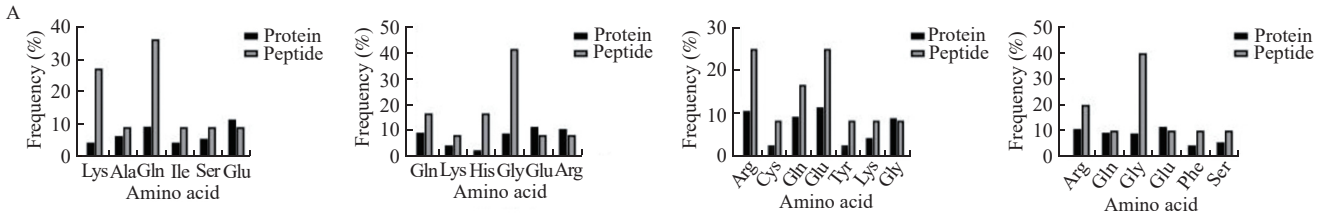
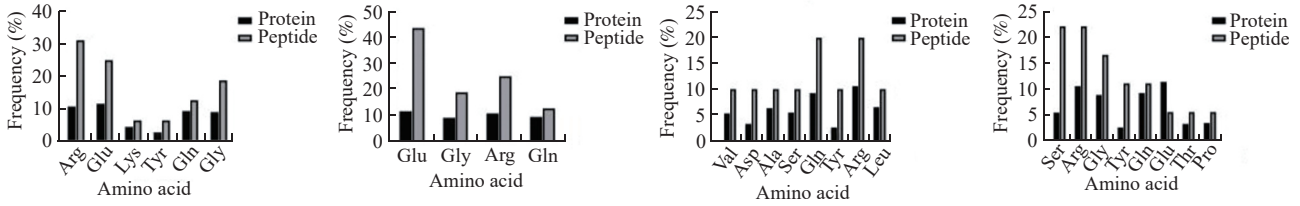


Fig. 5 Predicted results of key amino acids. (A) Comparison of the frequency of occurrence of amino acids in proteins and epitopes. (B) Conservative analysis of amino acid sequences of selected allergens. The UniProt login number is in parentheses.



B

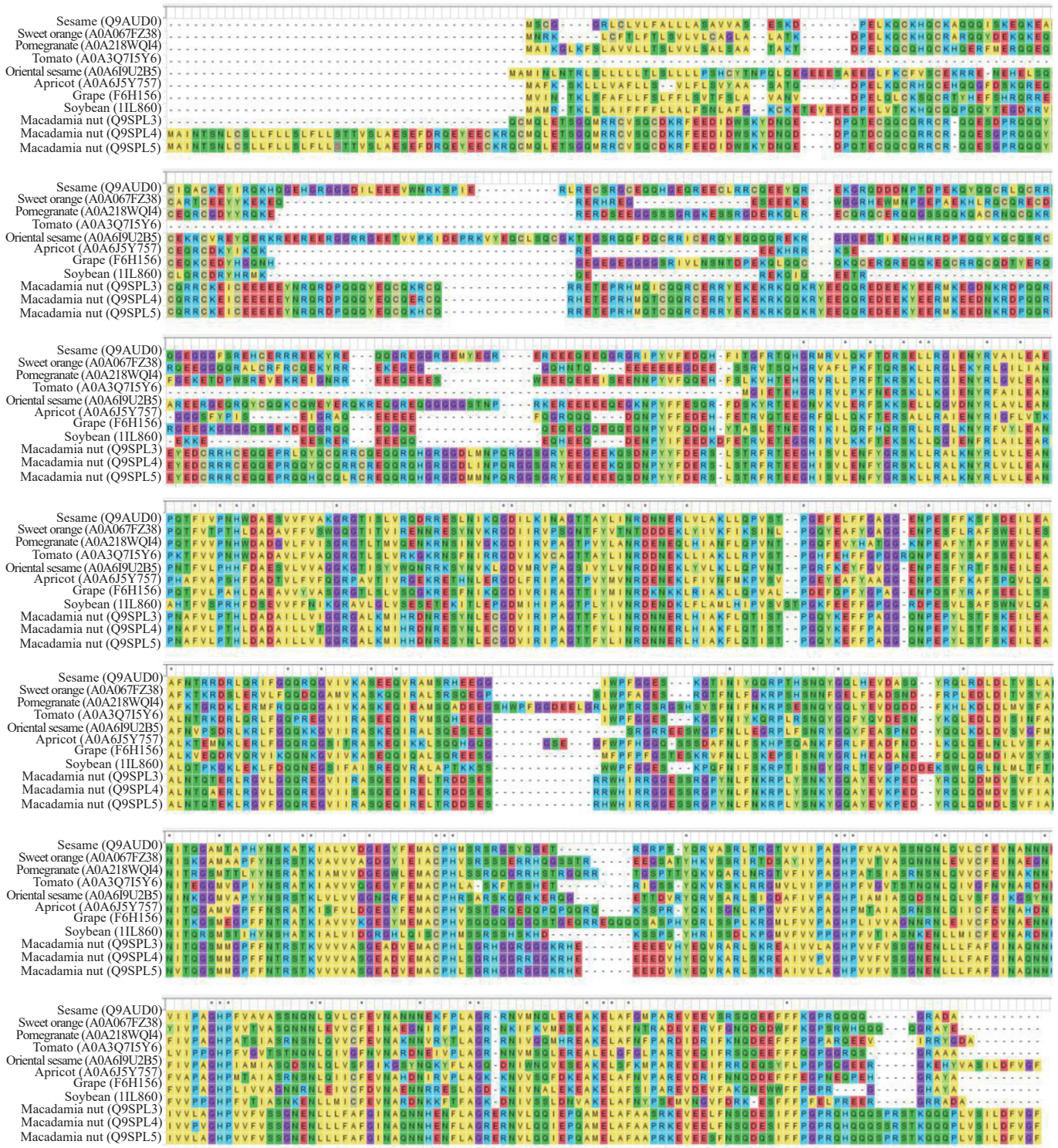


Fig. 5 (*Continued*)

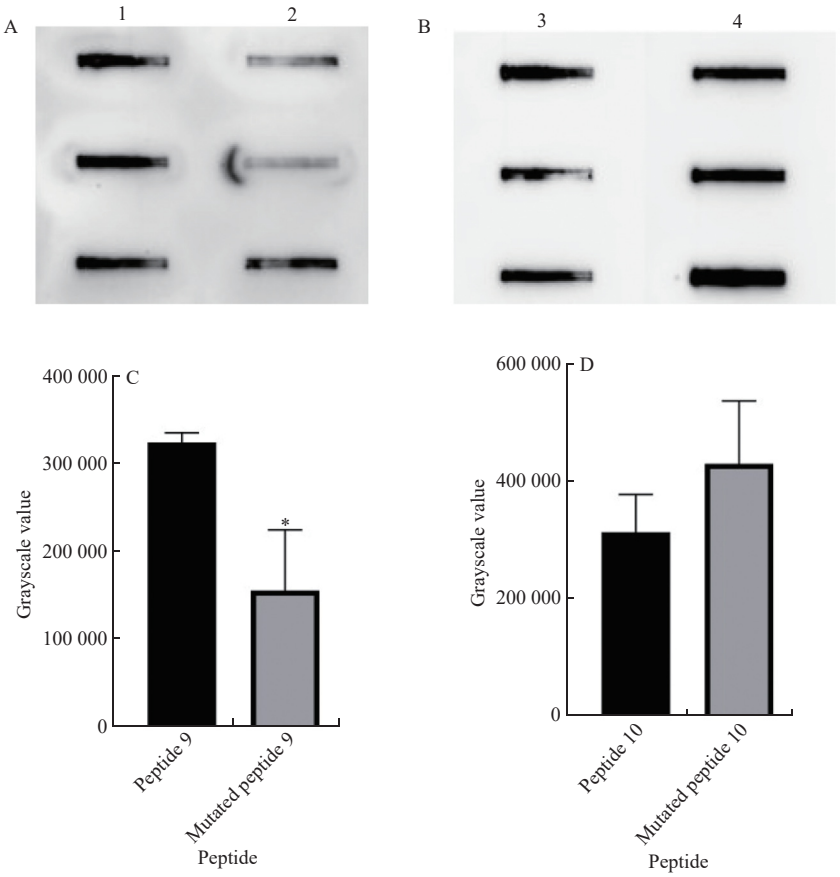


Fig. 6 Results of identification of key amino acids. (A) Results of immunoblotting changed in epitope 9 after mutation of leucine 422 to a lanine. 1, premutation. 2, post mutation. (B) Results of immunoblotting changed in epitope 10 after mutation of tryptophan 485 to alanine. 3, premutation. 4, postmutation. (C, D) Grayscale value changes before and after amino acid mutations. Marked with '*' indicated a significant difference ($P < 0.05$).

4. Conclusions

In the present work, we have analyzed the physicochemical properties, structure and homology of the sesame allergen Ses i 3. The amino acid sequences of Ses i 3 had a long hydrophobic region and with high antigenic index and easy epitope formation. Also, Ses i 3 is highly flexible, which may be one of the important reasons for the allergen formation. Combined with bioinformatics methods and immune-microarray techniques, 8 peptides were identified as B-cell antigenic epitopes of Ses i 3. By analyzing the amino acid conservativeness, the frequency of amino acid occurrence, and the further verification by slot blot immuno-microarrays, leucine 422 was identified as the key amino acid for the epitope. These findings will help further improve our understanding of sesame allergens.

Conflict of interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

We would like to thank the Beijing Shijitan Hospital for providing human sera. We are grateful for the financially supported by the Fundamental Research Funds for the Public Research Institutes of Chinese Academy of Inspection and Quarantine (2022JK04).

Appendix A. Supplementary information

Supplementary data associated with this article can be found, in the online version, at http://doi.org/10.26599/FSHW.2024.9250265.

References

[1] M.H. Li, J.N. Luo, M.A. Nawaz, et al., Phytochemistry, bioaccessibility, and bioactivities of sesame seeds: an overview, Food Rev. Int. 40 (2024) 309-335. https://doi.org/10.1080/87559129.2023.2168280.

[2] X. Pi, Z. Peng, J. Liu, et al., Sesame allergy: mechanisms, prevalence, allergens, detectionresidue, effects of processing and cross-reactivity, Crit. Rev. Food Sci. Nutr. 64 (2024) 1-16. https://doi.org/10.1080/10408398.2022.2128031.

[3] F.A. Farraj, R. Kteish, A. Salmons, et al., Sesame allergy: characteristics and outcomes, J. Allergy Clin. Immun. 149 (2022) AB318. https://doi.org/10.1016/j.jaci.2021.12.018.

[4] K. Beyer, L. Bardina, G. Grishina, et al., Identification of sesame seed allergens by 2-dimensional proteomics and Edman sequencing: seed storage proteins as common food allergens, J. Allergy Clin. Immun. 110 (2002) 154-159. https://doi.org/10.1067/mai.2002.125487.

[5] S.S. Tai, T.T. Lee, C.C. Tsai, et al., Expression pattern and deposition of three storage proteins, 11S globulin, 2S albumin and 7S globulin in maturing sesame seeds, Plant Physiol. Biochem. 39 (2001) 981-992. https://doi.org/10.1016/S0981-9428(01)01314-6.

[6] E. Orruño, M.R. Morgan, Resistance of purified seed storage proteins from sesame (*Sesamum indicum* L.) to proteolytic digestive enzymes, Food Chem. 128 (2011) 923-929. https://doi.org/10.1016/j.foodchem.2011.03.120.

[7] H. Zhang, L. Wei, H. Miao, et al., Development and validation of genic-SSR markers in sesame by RNA-seq, BMC Genomics 13 (2012) 1-11. https://doi.org/10.1186/1471-2164-13-316.

[8] C. Irani, G. Maalouly, M. Germanos, et al., Food allergy in Lebanon: is sesame seed the "middle eastern" peanut, World Allergy Organ. J. 4 (2011) 1-3. https://doi.org/10.1097/wox.0b013e318204b745.

[9] T. Ando, J. Kitaura, Tuning IgE: IgE-associating molecules and their effects on IgE-dependent mast cell reactions, Cells 10 (2021) 1697. https://doi.org/10.3390/cells10071697.

[10] M.B. Bilo, M. Martini, C. Tontini, et al., Lessons from peculiar cases of anaphylaxis: why allergists should be prepared for the unexpected, Eur. Ann. Allergy Clin. Immunol. 53 (2021) 4-17. https://doi.org/10.23822/EurAnnACI.1764-1489.198.

[11] K.D. Stone, C. Prussin, D.D. Metcalfe, IgE, mast cells, basophils, and eosinophils, J. Allergy Clin. Immun. 125 (2010) S73-S80. https://doi.org/10.1016/j.jaci.2009.11.017.

[12] W. Yu, D.M.H. Freeland, K.C. Nadeau, Food allergy: immune mechanisms, diagnosis and immunotherapy, Nat. Rev. Immunol. 16 (2016) 751-765. https://doi.org/10.1038/nri.2016.111.

[13] R.S. Chinthrajah, J.D. Hernandez, S.D. Boyd, et al., Molecular and cellular mechanisms of food allergy and food tolerance, J. Allergy Clin. Immun. 137 (2016) 984-997. https://doi.org/10.1016/j.jaci.2016.02.004.

[14] K.Y. Jeong, Physical and biochemical characteristics of allergens, Allergy Asthma Respi. 4 (2016) 157-166. https://doi.org/10.4168/aard.2016.4.3.157.

[15] Z.Y. Zhang, X.M. Li, H. Wang, et al., Seafood allergy: allergen, epitope mapping and immunotherapy strategy, Crit. Rev. Food Sci. 63 (2023) 1314-1338. https://doi.org/10.1080/10408398.2023.2181755.

[16] C. Liu, S.K. Sathe, Food allergen epitope mapping, J. Agric. Food Chem. 66 (2018) 7238-7248. https://doi.org/10.1021/acs.jafc.8b01967.

[17] F.L. Zhou, S.D. He, H.J. Sun, et al., Advances in epitope mapping technologies for food protein allergens: a review, Trends Food Sci. Tech. 107 (2021) 226-239. https://doi.org/10.1016/j.tifs.2020.10.035.

[18] K.A. Galanis, K.C. Nastou, N.C. Papandreou, et al., Linear B-cell epitope prediction for *in silico* vaccine design: a performance review of methods available via command-line interface, Int. J. Mol. Sci. 22 (2021) 3210. https://doi.org/10.3390/ijms22063210.

[19] E. Raoufi, M. Hemmati, S. Eftekhari, et al., Epitope prediction by novel immunoinformatics approach: a state-of-the-art review, Int. J. Pept. Res. Ther. 26 (2020) 1155-1163. https://doi.org/10.1007/s10989-019-09918-z.

[20] L. Fu, J. Wang, S. Ni, et al., Identification of allergenic epitopes and critical amino acids of major allergens in Chinese shrimp (*Penaeus chinensis*) by immunoinformatics coupled with competitive-binding strategy, JAFC 66 (2018) 2944-2953. https://doi.org/10.1021/acs.jafc.7b06042.

[21] L.L. Xu, M. Gasset, H. Lin, et al., Identification of the dominant T-cell epitopes of Lit v 1 shrimp major allergen and their functional overlap with known B-cell epitopes, J. Agric. Food Chem. 69 (2021) 7420-7428. https://doi.org/10.1021/acs.jafc.1c02231.

[22] Z.Y. Zhang, X.M. Li, H. Xiao, et al., IgE-binding epitope mapping of tropomyosin allergen (Exo m 1) from *Exopalaemon modestus*, the freshwater Siberian prawn, Food Chem. 309 (2020) 125603. https://doi.org/10.1016/j.foodchem.2019.125603.

[23] X.Y. Guo, S.S. Jiang, X.R. Li, et al., Sequence analysis of digestion-resistant peptides may be an efficient strategy for studying the linear epitopes of Jug r 1, the major walnut allergen, Food Chem. 322 (2020) 126711. https://doi.org/10.1016/j.foodchem.2020.126711.

[24] C. Yu, X. Gao, H. Lin, et al., Identification and amino acid analysis of allergenic epitopes of a novel allergen paramyosin (Rap v 2) from *Rapana venosa*, J. Agric. Food Chem. 69 (2021) 5381-5391. https://doi.org/10.1021/acs.jafc.1c00775.

[25] S. Ikagawa, S. Matsushita, Y.Z. Chen, et al., Single amino acid substitutions on a Japanese cedar pollen allergen (Cry j 1)-derived peptide induced alterations in human T cell responses and T cell receptor antagonism, J. Allergy Clin. Immun. 97 (1996) 53-64. https://doi.org/10.1016/S0091-6749(96)70283-X.

[26] T.T. Cheng, M.F. Tam, H. Chou, et al., Lys89, Lys90, and Phe91 are critical core amino acid residues of the Pen ch 18 major fungal allergen recognized by human IgE antibodies, Biochem. Bioph. Res. Co. 375 (2008) 671-674. https://doi.org/10.1016/j.bbrc.2008.08.097.

[27] C.N. Kopoin, A.K. Atiampo, B.G. N'Guessan, et al., Prediction of protein-protein interactions from sequences using a correlation matrix of the physicochemical properties of amino acids, International Journal of Computer Science and Network Security 21 (2021) 41-47. https://doi.org/10.22937/ijcsns.2021.21.3.6.

[28] Y. Vander Meersche, G. Cretin, A.G. de Brevern, et al., MEDUSA: prediction of protein flexibility from sequence, J. Mol. Biol. 433 (2021) 166882. https://doi.org/10.1016/j.jmb.2021.166882.

[29] C. Savojardo, M. Manfredi, P.L. Martelli, et al., Solvent accessibility of residues undergoing pathogenic variations in humans: from protein structures to protein sequences, Front. Mol. Biosci. 7 (2021) 626363. https://doi.org/10.3389/fmolb.2020.626363.

[30] B.K. Ho, R. Brasseur, The Ramachandran plots of glycine and pre-proline, BMC Structural Biology 5 (2005) 1-11. https://doi.org/10.1186/1472-6807-5-14.

[31] R.E. Goodman, Practical and predictive bioinformatics methods for the identification of potentially cross-reactive protein matches, Mol. Nutr. Food Res. 50 (2006) 655-660. https://doi.org/10.1002/mnfr.200500277.