A Computational Study of B-Cell Epitopes of Macadamia (Macadamia Integrifolia) Allergens and Identification of its IgE Binding Residues

Shivani Verma (verma.shivani6991@gmail.com)
Banaras Hindu University

Research

Keywords: Computational biology, Macadamia integrifolia, Allergy, Epitopes

DOI: https://doi.org/10.21203/rs.3.rs-750754/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License. Read Full License
Abstract

In contemporary research, biological computational tools have emerged to play a pivotal role in facilitating both cost and time-efficient research. One such domain is addressing the prevailing food allergy issues, where these computational tools have been proven of vital importance. The present study is done as part of an internship at Jozbiz technologies. In the present study, we discuss the identification of IgE(Immunoglobulin E) binding allergy-causing B-Cell epitopes of macadamia (Macadamia integrifolia) allergens, namely AMP23_MACIN, AMP1_MACIN, AMP21_MACIN, MATK_MACIN and 11S1_MACIN. Using seven web servers (ABCPred, Ellipro, Bepipred 1.0b, BcePred, BCPred, CBTOPE and Disco Tope 2.0) twenty-one epitopes and seventeen conformational epitopes were predicted in the present study. The predicted epitopes are analysed in terms of residues having hydrophilic, polar nature and having exposed surfaces. The unavailable 3-d structure of proteins was developed by homology modelling. Cross-reactivity of Macadamia integrifolia with other food items has also been listed.

Introduction

In the growing age of digital advancement and fast-paced life, things are getting personalised with an increase in its specificity. Every day the field of computational biology is also making advancements in easing the experimenting in wet labs and increasing the pace of development of the desired products. With the boon of in silico methods advancements, this study aims to identify and list the B-cell epitopes of macadamia and its IgE binding residues.

Before moving further with this it is important to understand the concept of allergy. Allergy is an immune response that occurs when any foreign material gets introduced into our bodies. These are generally perceived as harmful to organisms. Several common symptoms of allergy are rashes, swollen bronchi, stomach aches and anaphylaxis which might be fatal. Some common allergens are animal dander, pollen, dust and certain foods too. Food allergens can be categorized into two parts a)Immunoglobulin E (IgE) mediated and b) Non-IgE mediated.

Wheat allergy is an IgE mediated allergy where the antibody IgE is produced by the human immune system in response to certain allergens present in macadamia. The person having IgE mediated allergy will also be allergic to foods with a similar amino acid sequence as the macadamia.

Macadamia the allergen for the study is from the family of nuts and has its benefits as well as certain specialities. Scientifically known as Macadamia integrifolia is indigenous to Australia with a shelf life of a few months. It is widely used in delicacies. It’s not only a decorative food but also has great nutritional benefits too. It is rich in vitamin B6 and antioxidants to get rid of free radicals. Regular intake of macadamia nuts lowers the risk of metabolic syndromes, diabetes and aids in weight loss too. The flavonoids and tocotrienols present in them are considered anti-cancer properties.

The macadamia nuts are not always a blessing to all. A certain population has shown an allergic response to it. The cases reporting macadamia nut allergy are quite low. In 2000, two allergy cases were reported and later the patients were tested in the clinic[2,7]. The skin prick test upon induced allergy proved the cause of it being the macadamia nuts. Along with the reaction on the skin the other effects include decreased gut health with abdominal cramps and vomiting in some cases.

The manually curated proteins namely AMP23_MACIN, AMP1_MACIN, AMP21_MACIN, MATK_MACIN and 11S1_MACIN from the UniProt database are used for the identification of the epitopes. These are processed using web servers that work on different principles. The epitopes identified are linear and conformational.

Materials And Methods

The Uniprot database(https://www.uniprot.org/) is used in retrieving the different amino acid sequences of macadamia nuts namely AMP23_MACIN, AMP1_MACIN, AMP21_MACIN, MATK_MACIN and 11S1_MACIN respectively. Several web servers are benchmarked for the prediction of linear and conformational epitopes. These are ABCpred[10], Bepipred[4], CBTope[1], DiscoTope2.0, Bcepred[9] and Ellipro[8]. These web servers work on principles of Artificial Neural Network, Hydrophilicity scale with Hidden Markov Model, Support Vector Machine, Surface Probability, Antigenic Index and Flexibility which process the amino acid sequence for predicted allergens.

The linear epitopes are obtained from servers like ABCpred, Bepipred and Bcepred. The results obtained for macadamia nuts are later compiled using python pipeline and assessed in SDAP(Structural Database of Allergic Proteins)(https://fermi.utmb.edu/).

The 3-d structures of proteins were not available in the PDB(protein data bank) and are homologically modelled using the SWISS-model server(https://swissmodel.expasy.org/). The obtained homologous model went under editing using the Chimera tool(https://www.cgl.ucsf.edu/chimera/). The Conformational epitopes are obtained using CBTope, DiscoTope2.0 and Ellipro. The sets of continuous and non-continuous epitopes are obtained. The properties for the allergens are computed using the BCEpred server based on hydrophilicity, polarity and antigenic propensity along with surface area exposed.

Results And Discussion

Computational biology has been like a blessing to the fast-paced world with numerous data generated now and then. The field not only increases the speed of process and research but also greatly merges the interdisciplinary. This saves time and resources for better and precise work in future. In this study, we focused on the prediction of the epitopes and the cross-reactive species with similar amino acid sequences potent to cause an allergic reaction. The epitope recognition is of utmost importance for the precise study of antigen-antibody interaction. This not only serves as the base but also helps in the development of personalised medicine hence, preventing the reaction. The servers used in the study work on the pre-existing epitopes derived from experimental evidence increasing the accuracy. The present study comprises predicted epitopes of AMP23_MACIN, AMP1_MACIN, AMP21_MACIN, MATK_MACIN and 11S1_MACIN using the servers ABCpred, Bepipred, CBTOpe, DiscoTope2.0, Bcepred and Ellipro. The conformational epitopes and the prone sites for the highest Ag-Ab
binding require the 3-d model. The unavailability of structures was resolved using the modelling server. The modelled and edited structure of the AMP1_MACIN is validated using the Rampage web servers. The residues are found to be in favourable regions making it an efficient model for use and refer to the Fig. 1.

Table I

| Allergen  | No. of epitopes | Start position | Predicted epitope residues | No. of residues |
|-----------|----------------|----------------|----------------------------|-----------------|
| AMP23_MACIN | 5             | 10             | EEEENEYNQRDPQQ              | 14              |
|           |                | 33             | QRRTEEPR                   | 8               |
|           |                | 61             | QKRYEEQQREDEEKYE            | 16              |
|           |                | 269            | STPGQYKEFPAGQN             | 16              |
|           |                | 539            | RQHQQQPRSTKQQP             | 16              |
| AMP1_MACIN | 4             | 14             | MMLIAMASEMVMNGSAF          | 17              |
|           |                | 31             | VWSGPGCNIERAERY            | 14              |
|           |                | 51             | AIHQKGGYDFSYTGQT           | 15              |
|           |                | 84             | GSSARACNPFGWKSIF           | 15              |
| AMP21_MACIN | 6             | 63             | FEEDIDWSKYDNQEDP           | 15              |
|           |                | 89             | CRQQES                     | 6               |
|           |                | 111            | EEEENEYNQRDPQQY            | 15              |
|           |                | 134            | QRRTEEPHRM                 | 10              |
|           |                | 161            | QKRYEEQQREDEEKYE           | 16              |
|           |                | 226            | GGDHHNPQRGGSGREEEEEQSD     | 24              |
| MATK_MACIN | 3             | 54             | MEKLQEYLEIDRWSQQ           | 16              |
|           |                | 319            | PGRIDINQLSNSFDL           | 15              |
|           |                | 342            | MARGTPIIMMNKWKY            | 15              |
|           |                | 466            | CKNISHYHSGSSKKS           | 15              |
| 11S1_MACIN | 3             | 30             | LNNQANQLDQK                | 11              |
|           |                | 45             | LLPQGHA                    | 7               |
|           |                | 155            | VAHWCLNGKHYLDNPR           | 17              |

**TABLE II: LIST OF CONFORMATIONAL IGE BINDING EPITOPES PREDICTED BY ONLINE WEB - SERVERS**
| Allergen       | No. of epitopes | Start position | Predicted epitope residues       | No. of residues |
|---------------|----------------|----------------|----------------------------------|----------------|
| AMP23_MACIN   | 5              | 10             | MRRCVSQCDKRFEEIDWSKYDNQEDPQ      | 27             |
|               |                |                | RDPQQRE                         | 8              |
|               |                |                | EKQSDN                          | 6              |
|               |                |                | IAKFL                           | 5              |
|               |                |                | FFPAGGQNPEPY                     | 12             |
| AMP1_MACIN    | 3              | 26             | GSAFTVWSGPGCNNAERY               | 18             |
|               |                |                | TGQTAALY                        | 8              |
|               |                |                | GSSARAC                          | 7              |
| AMP21_MACIN   | 4              | 51             | MRRCVSQCDKRFEEIDWSKYDNQEDPQ      | 27             |
|               |                |                | DPQQRE                          | 6              |
|               |                |                | FFPAGGQNPEPY                     | 12             |
|               |                |                | ESSRGYPYN                       | 8              |
| MATK_MACIN    | 4              | 97             | MISEG                           | 5              |
|               |                |                | KDPFTHY                         | 7              |
|               |                |                | PLMMNKWKY                       | 9              |
|               |                |                | IDINQL                          | 6              |
| 11S1_MACIN    | 1              | 93             | DGKHYLDNPR                      | 10             |

**TABLE III: CONSENSUS OF LINEAR AND CONFORMATIONAL IGE BINDING EPITOPIES OF MACADAMIA ALLERGENS**

The predicted linear and conformational epitopes are listed below. The use of computational tools gives the in-silico results which are asked to be proved using in vivo and in vitro experiments. These in silico results aid as the first hand-processed data. The linear and the conformational epitopes were assessed by consensus for the most potent epitopes listed in Table III.

**TABLE IV: CROSS-REACTIVE SPECIES**
### Table V

| Protein       | %Hydrophilicity | %Flexibility | %Accessibility | %Surface area exposed | %Polarity | %Antigenic propensity |
|---------------|-----------------|--------------|----------------|-----------------------|-----------|----------------------|
| AMP23_MACIN   | 27.2            | 16.48        | 49.76          | 22.56                 | 38.72     | 7.68                 |
| AMP1_MACIN    | 5.88            | 1.96         | 6.86           | 1.96                  | 5.88      | 1.96                 |
| AMP21_MACIN   | 27.17           | 15.1651      | 49.39          | 21.02                 | 36.33     | 9.45                 |
| MATK_MACIN    | 2.554           | 5.10806      | 22.39          | 17.87                 | 6.86      | 15.68                |
| 11S1_MACIN    | 1.9607          | 0            | 17.647         | 9.803921             | 6.8627    | 15.6862             |

The physicochemical properties also give an interesting insight into the binding of the Ab molecules. The hydrophilicity is proportional to the power of IgE binding.

The hydrophilicity is recorded highest in 'AMP23_MACIN' possesses the maximum percentage of hydrophilic residues with 27.2%, followed by 'AMP21_MACIN' with 27.17%, 'AMP1_MACIN' with 5.88%, 'MATK_MACIN' with 2.55% and '11S1_MACIN' with 1.96%(Fig. 2).

On comparing percentage polarity in the epitopic regions, the value observed for 'AMP23_MACIN' was found to be the highest (38.72%), followed by 'AMP21_MACIN'(36.33%), 'MATK_MACIN'(17.87%), '11S1_MACIN'(6.86%) and 'AMP1_MACIN'(5.88%) (Fig. 2).

Percentage of Surface Exposed Residues was also estimated for all five macadamia allergens. The highest percentage was observed for 'AMP23_MACIN' being 22.56% followed by 'AMP21_MACIN'(21.06%), 'MATK_MACIN'(4.91%), 'AMP1_MACIN'(1.96%) and '11S1_MACIN'(0.98%). The results have been summarised in Fig. 2.

### Conclusion

Foods are not always good for the body. Some foods are perceived as foreign material and our immune system reacts to it causing an allergic reaction. The study is about the in silico identification of the potent epitopes and their cross-reactivities. For this experiment web servers like ABC pred, Bepipred, CBTope, DiscoTope2.0, Becpred and Ellipro were used. The linear and conformational epitopes were recorded. The macadamia (*Macadamia integrifolia*) proteins used were manually curated namely AMP23_MACIN, AMP1_MACIN, AMP21_MACIN, MATK_MACIN and 11S1_MACIN.

The percentage of hydrophilic residues was maximum in 'AMP23_MACIN' (27.2%). The percentage of polar residues was maximum (38.72%) in 'AMP23_MACIN' followed by 'AMP21_MACIN'(36.33%) whereas the percentage of surface-exposed residues were highest for 'AMP23_MACIN' being 22.56%, followed by 'AMP21_MACIN'(21.06%).

The identification of these potent epitopes gets us closer to the development of a method or any hack to avoid the allergic reaction or its treatment. It eases the work in vivo and gives a better view of specificity to work in the development of personalised medicine. The cross-reactive provides the warning of staying away from similar foods that cause the same reaction.

### Declarations

Ethics approval and consent to participate: Not applicable

Consent for publication: Not applicable

Availability of data and materials: All data generated or analysed during this study are included in this published article and in supplementary material attached.

Competing interests: The authors declare that they have no competing interests
Funding: No funding from any institute aided this study

Authors' contributions: It is a single author article and all the work is completed by Shivani Verma.

Acknowledgements: Shivani Verma is thankful to Jozbiz technologies for providing me this opportunity to work on and explore my potential.

References

1. Ansari, Hifzur Rahman, and Gajendra PS Raghava. "Identification of conformational B-cell Epitopes in an antigen from its primary sequence." *Immunome research* 6.1 (2010): 1-9.

2. De Knop, K. J., et al. "Macadamia nut allergy: 2 case reports and a review of the literature." *Acta Clinica Belgica* 65.2 (2010): 129-132.

3. Herbst, R. A., R. Wahl, and P. J. Frosch. "Specific IgE reactivity and identification of potential allergens in macadamia allergy." *Journal of the European Academy of Dermatology and Venereology* 24.11 (2010): 1361-1363.

4. Jespersen, Martin Closter, et al. "BepiPred-2.0: improving sequence-based B-cell epitope prediction using conformational epitopes." *Nucleic acids research* 45.W1 (2017): W24-W29.

5. Johri, Amogh, and Meenakshi Srivastava. "A computational study of B-cell epitopes of wheat allergens and identification of its IgE binding residues." *International Journal of Information Technology* (2021): 1-8.

6. Johri, Amogh, and Meenakshi Srivastava. "Prediction and Analysis of Allergenic Epitopes of Tree-Nuts and its Cross-Reactivity." *Asian Journal For Convergence In Technology (AJCT)* ISSN-2350-1146 (2019).

7. Pallares, David E. "Allergy to macadamia nut." *Annals of allergy, asthma & immunology: official publication of the American College of Allergy, Asthma, & Immunology* 85.5 (2000): 385-386.

8. Ponomarenko, Julia, et al. "ElliPro: a new structure-based tool for the prediction of antibody epitopes." *BMC bioinformatics* 9.1 (2008): 1-8.

9. Saha, Sudipto, and Gajendra Pal Singh Raghava. "BcePred: prediction of continuous B-cell epitopes in antigenic sequences using physico-chemical properties." *International Conference on Artificial Immune Systems*. Springer, Berlin, Heidelberg, 2004.

10. Saha, Sudipto, and Gajendra Pal Singh Raghava. "Prediction of continuous B-cell epitopes in an antigen using recurrent neural network." *Proteins: Structure, Function, and Bioinformatics* 65.1 (2006): 40-48.

11. Yoshida, Koichi, et al. "Macadamia nut allergy in children: clinical features and cross-reactivity with walnut." *Pediatric Allergy and Immunology* (2021).

Figures

Figure 1

Depicting homology modelled structure of AMP1_MACIN
Figure 2

Depicts the physio-chemical properties of the predicted continuous/linear epitopes of AMP23_MACIN, AMP1_MACIN, AMP21_MACIN, MATK_MACIN and 11S1_MACIN. X-axis: Different properties for each allergen. Y-axis: Percentage number of residues

Supplementary Files

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

- [macadamia.fasta](macadamia.fasta)
- [macadamia.ods](macadamia.ods)