Journal of Advances in Biology & Biotechnology

Volume 27, Issue 8, Page 280-288, 2024; Article no.JABB.119678
ISSN: 2394-1081

Structural and Functional Characterization of an Amino acid Transporter in *Amaranthus hypochondriacus* using Bioinformatics Tools

Rashmi Chauhan [a], Sharat Prabhakaran [a], Gohar Taj [a], S.K. Guru [a] and Dinesh Pandey [a*]

[a] G. B. Pant University of Agriculture and Technology, Pantnagar, Uttarkhand, India.

Authors' contributions

This work was carried out in collaboration among all authors. Author RC designed the study, wrote first draft of manuscript. Author SP performed in silico analysis. Authors DP, GT and SKG managed the analysis of the study. All authors read and approved the final manuscript.

Article Information

DOI: https://doi.org/10.9734/jabb/2024/v27i81140

Open Peer Review History:
This journal follows the Advanced Open Peer Review policy. Identity of the Reviewers, Editor(s) and additional Reviewers, peer review comments, different versions of the manuscript, comments of the editors, etc are available here: https://www.sdiarticle5.com/review-history/119678

Original Research Article

Received: 10/05/2024
Accepted: 15/07/2024
Published: 18/07/2024

ABSTRACT

Plants produce amino acids by utilizing metabolic pathways, which are essential for metabolism, temporary storage. Despite the existence of various transporter families, there is limited structural information available on the transporters. In the present study, bioinformatics based analysis was performed to decipher the role of amino acid transporter *viz.* amino acid permease which was observed to be associated with high protein content in grain of *Amaranthus hypochondriacus*

**Corresponding author: E-mail: dinesh.pandey@gbpuat-cbsh.ac.in;*

Cite as: *Chauhan, Rashmi, Sharat Prabhakaran, Gohar Taj, S.K. Guru, and Dinesh Pandey. 2024. "Structural and Functional Characterization of an Amino Acid Transporter in Amaranthus Hypochondriacus Using Bioinformatics Tools". Journal of Advances in Biology & Biotechnology 27 (8):280-88. https://doi.org/10.9734/jabb/2024/v27i81140.*

through a genome wide association study (GWAS).The primary, secondary, and tertiary structure of the amino acid permease (an amino acid transporter) was modeled and validated using bioinformatics tools like ExPasy ProtParam, SOPMA, PSIPRED, and SWISS-MODEL. Results were further validated using MolProbity and PROCHECK programs. For phylogenetic tree construction, MEGA11 was used, and the STRING database was utilized to identify its interaction with other proteins. This study is first report of amino acid transporter of *A. hypochondriacus* which enhances our understanding about its structure and function which can be utilized in medical sector for humans benefit.

Keywords: Amino acid permease; homology modelling; primary structure; protein-protein interaction; secondary structure; SWISS-MODEL.

1. INTRODUCTION

Nitrogen (N) is essential for plant growth, with plants absorbing both inorganic (ammonium, nitrate) and organic (amino acids, peptides) nitrogen from the soil [1,2,3,4]. Once absorbed, N assimilation involves converting nitrate to ammonium, which is then transformed into amino acids in the roots or leaves. These amino acids are synthesized in various cellular compartments and transported via the xylem and phloem to support growth and development [5]. Amino acid transporters play a crucial role in ensuring the proper distribution of organic N throughout the plant, moving amino acids in and out of plant cells and across different compartments [6,7].

Amaranthus hypochondriacus, a nutritionally important pseudocereal is known for its high protein content and balanced essential amino acid profile [8,9,10]. However, the specific characteristics of amino acid transporters in *Amaranthus hypochondriacus* are not well understood [11,12].

In our lab, an amino acid transporter was identified through GWAS study for total protein in *Amaranthus hypochondriacus*.

Understanding protein structure is essential for unraveling biological mechanisms on a molecular scale. Determining the three-dimensional shape of proteins is a complicated process typically achieved through NMR spectroscopy or X-ray crystallography. These techniques, while effective, can be time-consuming and challenging, particularly for membrane proteins [13,14]. An alternative method is homology modeling for in silico 3D structure prediction, which requires protein sequences with at least 35% similarity [15].

This research aimed to fill this knowledge gap through simulate the amino acid transporter in *A. hypochondriacus* using in silico analysis, involved examining physicochemical properties, and secondary and tertiary structure with the help of various computational tools. Efforts were also made to validate its role in accumulation in amino acids in grain of Amaranth. Memsat-svm was utilized for conducting predictions on transmembrane helix and topology [16]. Prediction of the three-dimensional structure of a protein using computational methods is a challenging but necessary process to complement and validate NMR or X-ray crystallography data. This study enhances our understanding about amino acid permease (amino acid transporter) in *A. hypochondriacus* by analyzing its structure, physicochemical properties, and structural motifs and also provides an interaction network with other proteins.

2. MATERIALS AND METHODS

2.1 Sequence Retrieval

We analyzed 192 genotypes for protein content and conducted a genome-wide association study (GWAS) using the General Linear Model (GLM) and Mixed Linear Model (MLM) in genome association and prediction integrated tool (GAPIT). Our analysis identified several significant single nucleotide polymorphisms (SNPs). Notably, a significant SNP (Ah12384531) was found near a gene encoding an amino acid permease was found near to SNP, suggesting a potential role for this gene in influencing protein in the genotypes studied. Further The amino acid sequence of amino acid permease of *Amaranthus hypochondriacus* was retrieved in the FASTA format from phytozome having accession no. AH020968.

2.2 Physiochemical Characteristics

The Expasy ProtParam tool was utilized to analyze the physicochemical characteristics of

the amino acid transporter protein. This analysis involved determining the theoretical isoelectric point (pI), molecular weight, atomic composition, extinction coefficient, aliphatic index, amino acid composition, instability index, and the grand average of hydropathy (GRAVY) [17].

2.3 Functional Domain and Family/Superfamily Prediction

The identification of the motifs, family, and superfamily of Ah_AAP was performed using the NCBI Conserved Domain Database, employing default settings for the analysis [18,19].

2.4 Multiple Sequence Alignment and Phylogenetic Analysis

Initially, protein sequences were obtained from the NCBI protein database using BLAST based on their similarity. Further, MEGA 11 software was utilized to conduct the multiple sequence alignment (MSA) and phylogenetic analysis comparing the specified AAP with the retrieved dataset. The MSA analysis utilized the advanced ClustalW algorithm. Additionally, a phylogenetic tree was created using the same sequence alignment to display the evolutionary distance between the related proteins [20]. We utilized the neighbour-joining method with default parameters and 1000 bootstrap replications for this task.

2.5 Secondary Structure Prediction

PSI-blast-based secondary structure PREDiction (PSIPRED) and self-optimized prediction method with alignment (SOPMA) were utilized to analyze the secondary structure features, including the α-helix, β-sheet, and turn, of the amino acid transporter protein's amino acid sequences [21].

2.6 Protein 3D Model Prediction

The protein sequence of the amino acid transporter was used as query sequences in comparative modeling to help predict the transporter's three-dimensional structure via SWISS-MODEL [22].

2.7 Model Evaluation

PROCHECK and MolProbity were used to assess the modeled amino acid permease structure for internal consistency and reliability by examining the stereochemical quality of the model [23].

2.8 Protein-protein Interaction Analysis

The STRING database is a detailed tool to visualize protein-protein interactions (PPIs). It helps in understanding cellular processes, functions, and mechanism on these crucial interactions, which are vital for systems biology. Its extensive coverage made it essential for our analysis [24].

3. RESULTS AND DISCUSSION

3.1 Physiochemical Properties

Through the analysis of 192 *Amaranthus hypochondriacus* genotypes for protein content, we conducted a genome-wide association study (GWAS) using both GLM and MLM models in GAPIT. This led to the discovery of significant SNPs linked to a gene coding for an amino acid permease. This amino acid permease was chosen for the study. indicating it may significantly influence protein homeostatis. The FASTA format from phytozome contained the sequence of amino acid permease of *Amaranthus hypochondriacus*, consisting of 392 amino acids with accession AH020968-RA. To discover features of Ah_AAP, initial physiochemical properties were calculated using Expasy's ProtParam, which offers essential information about the protein. The molecular weight of the protein measured 43581.47 Da, with a calculated isoelectric point (pI) of 7.25 and an aliphatic index of 121.99. Moreover, the instability index was measured at 35.32, with the grand average of hydropathy (GRAVY) was determined to be 0.734. The findings are outlined in Table 1. The findings from analyzing the primary structure indicate that the protein is hydrophobic with the help of GRAVY. The GRAVY value of a protein was calculated by adding up the hydropathy values for each amino acid and then dividing by the total number of residues [25]. The Ah_AAP protein has a positive or higher GRAVY value in Table 1, showing its hydrophobic nature [26]. The existence of cysteine residues signifies the existence of disulfide bridges (S-S bonds) in the Ah_AAP. Furthermore, according to the primary structure analysis, the Ah_AAP is rich in aliphatic residues like Leu (13.6%), Ile (10.7%), Gly (6.9%), Ala (8.7%), Thr (6.4%), and Val (6.4%) while lacking Pyl (0.0%) and Sec (0.0%). The isoelectric point of Ah_AAP was calculated to be 7.25 (Table 1). The pI value calculated for Ah_AAP protein indicates that it is basic due to being above 7. The aliphatic index measures the proportion of a

protein's space taken up by aliphatic side chains. The structural stability of proteins is demonstrated by a higher aliphatic index [27]. This study shows that the high amount of aliphatic amino acids in Ah_AAP protein is evidence of its resistance to high temperatures. An instability index under 40 signifies protein stability, whereas anything above 40 implies instability [22]. According to our data, Ah_AA P is a protein that remains stable in this scenario, as indicated by its instability index value of 35.32 being below 40 (Table 1).

3.2 Domain, Family, and Superfamily Prediction

The results obtained from NCBI Conserved Domain (CD) Search, revealed that the Ah_AAP sequence was found to have solute-binding domain of SLC5 proteins domain. The protein belongs to the SLC5/6 family protein family and the SLC5-6-like_sbd superfamily.

3.3 Phylogenetic Analysis

The BLASTp results for Ah_AAP against non-redundant databases indicated significant similarity to amino acid permease proteins from various species. Phylogenetic analysis was conducted using MEGA11 to confirm the identification of homologous proteins and determine the evolutionary distance between our target protein and aligned amino acid permease proteins. The maximum likelihood method was used in MEGA11 to create a phylogenetic tree analysis that illustrates the evolutionary connections between different plant species amino acid permeases (Fig. 1). Phylogenetic tree which was generated through MEGA11 shows that it is grouped in two group, one group contains only *Arabidopsis thaliana* while second group contains *Amararnthus hypochondriacus*, *Amaranthus tricolor*, *spinacia oleracea*, *Chenopodium quinoa*, *Beta vulgaris*. *Amaranthus hypochondriacus* is closely related to *Amaranthus tricolor* and distantly related to Arabidopsis thaliana in perspective of amino acid permease evolution (Fig. 1).

3.4 Secondary Structure Prediction

The secondary structure was determined using SOPMA, carefully examining beta turns, alpha helices, random coils, and extended strands (Fig. 2). The high percentage (31.20%) of coils found in the Ah_AAP protein signifies the protein's capability to be flexible and modify its shape. Similarly, the greater percentage of alpha helix (49.62%) present in Ah_AAP indicates that the protein is stable at elevated temperatures, similar to thermophilic proteins which usually contain a high amount of alpha helices (Table 2) [28]. MEMSAT-SVM Schematics were utilized in PSIPRED to predict transmembrane presence and secondary structures in target protein sequences through transmembrane predictions. it showed that the Ah_AAP protein has seven transmembrane helices depicted in Fig. 3. Helix 1 spans from residues 12 to 35, helix 2 spans from residues 55 to 78, helix 3 spans from residues 103 to 118, helix 4 spans from residues 127 to 142, helix 5 spans from residues 203 to 221, helix 6 spans from residues 303 to 333, and helix 7 spans from residues 363 to 384 in the protein sequence. The N terminal faces the cytoplasm, whereas the C terminal faces the extracellular space.

Table 1. Physiochemical properties of amino acid permease of *Amaranthus hypochondriacus*

S.No.	Name of the Protein	M.wt. (Da)	No. of amino acid	Pi	GRAVY	AI	Ext. coefficient	Ext. coefficient
1.	Amino acid permease transporter	43581.47	391	7.25	0.734	121.99	89840 assuming all Cys residues are reduced	90340 assuming all pairs of Cys residues form cystines

Table 2. Secondary structure features of amino acid permease transporter of *Amaranthus hypochondriacus* using SOPMA

S.No.	Parameter	Value
1.	α-helix	49.62%
2.	Extended β turn	17.14%
3.	Random coil	31.20%
4.	Ambiguous state	0.00%

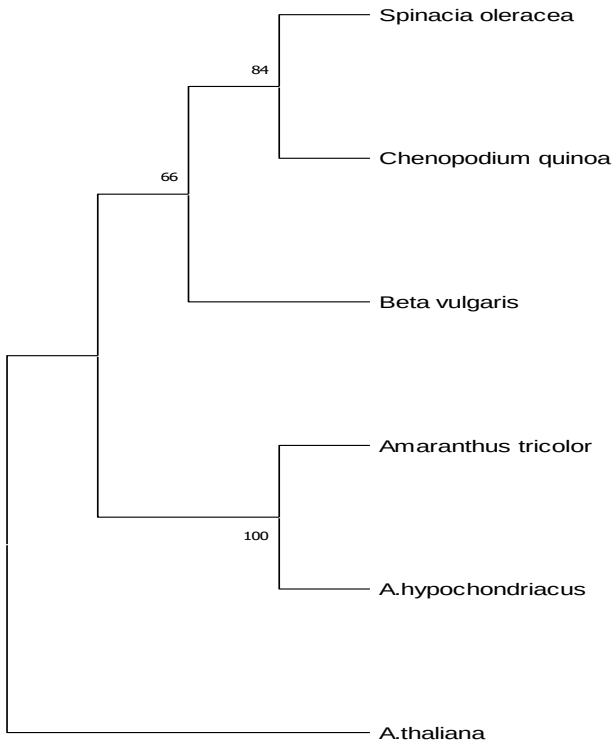

Fig. 1. Phylogenetic tree of six orthologous sequences of amino acid permease transporter

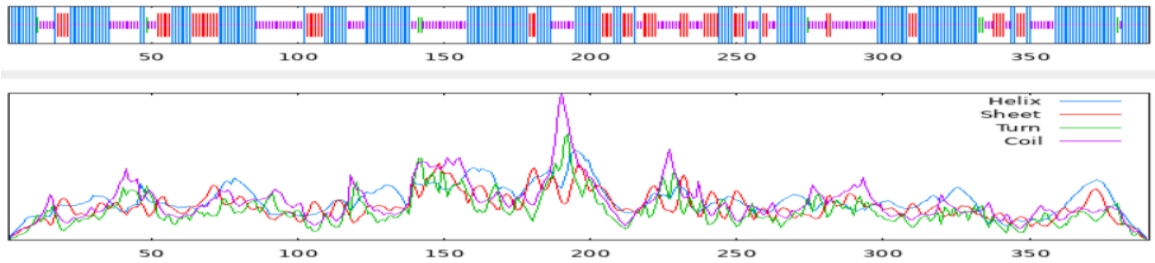

Fig. 2. Secondary structure prediction using SOPMA

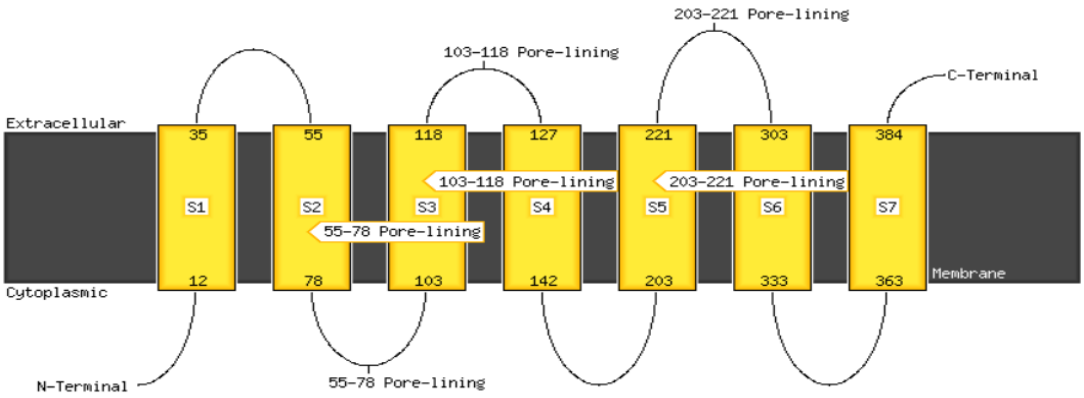

Fig. 3. Transmembrane description of Ah_AAP

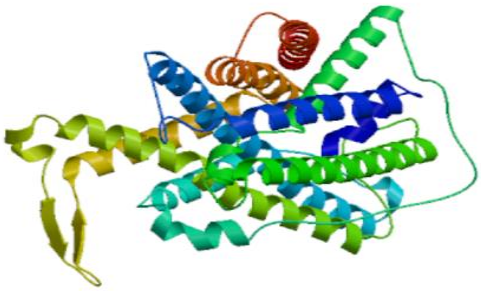

Fig. 4. 3D-model of amino acid permease transporter from *A. hypochondriacus* using SWISS-MODEL workspace

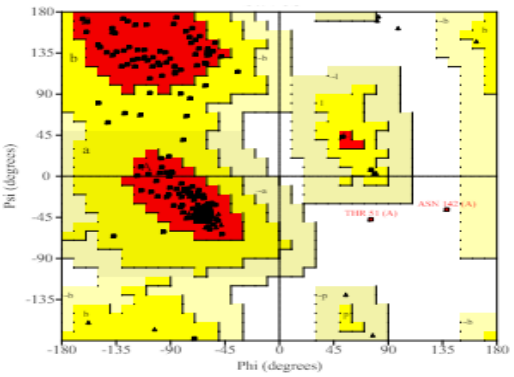

Fig. 5. Ramachandran plot of Ah_AAP

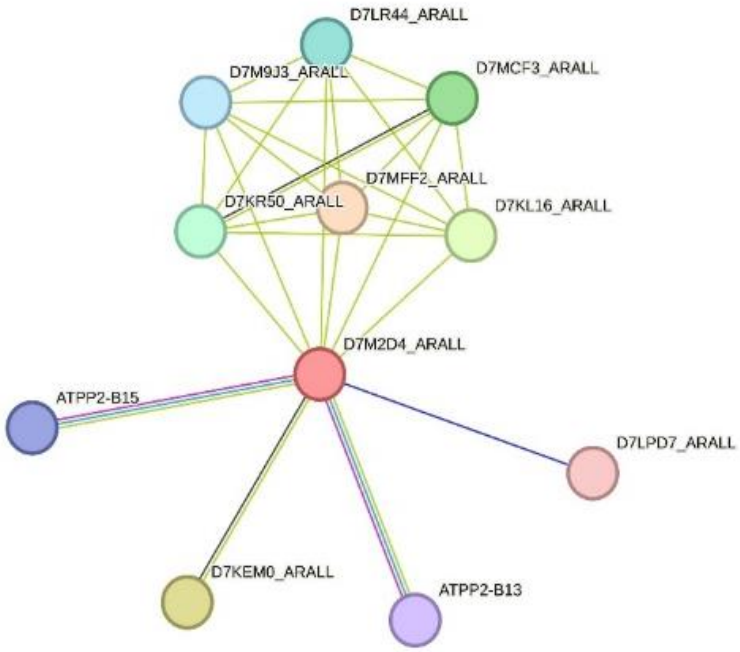

Fig. 6. Ah_AAP interaction network with it's partners

3.5 Tertiary Structure Prediction

Homology modeling was performed with the aid of SWISS-MODEL's automated homology protein modeling server, utilizes ProMod3, an advanced comparative modeling engine (Fig. 4). After that, stereochemistry of the Ah_AAP protein structure was verified through the PROCHECK server [29], This program generated Ramachandran plots by plotting the phi-psi torsion angles of amino acid residues in relation to one another (Fig. 5). The findings indicated that most amino acids align with a phi-psi distribution typical of a right-handed α-helix. Therefore, the protein takes on a structure that is both flexible and stable. Further, MolProbity is also used to assess the accuracy of the Ramachandran plot. In MolProbity, 99.09% of the residues in the Ramachandran plot of the Ah_AAP protein are found in the most favored region, indicating that this protein model is high quality and 0.51% in outliers regions [30].

3.6 Protein Interaction Network

For further investigation about with which type of proteins, our investigated protein is interacting, we used STRING database.Protein-protein

interaction was done using STRING 12.0 to comprehend the interaction network of the Ah_AAP protein. The interaction is depicted in Fig. 6. The confidence score was used to analyze the functional interaction in STRING. Interactions with scores below 0.3 are seen as having low confidence, while scores between 0.3 and 0.7 are considered medium confidence, and scores above 0.7 are deemed high confidence. It showed that our target Ah_AAP interacts with various proteins, some with confirmed functions and others without experimental annotations. Our focused protein is expected to have a significant association with the protein containing the AA_permease_C domain, as well as a lesser interaction with Phloem protein 2-B13 and Phloem protein 2-B15. protein containing the AA_permease_C domain is a membrane protein, which is involved in transport of amino acids across the cell membranes [31]. Phloem protein 2-B13 and Phloem protein 2-B15 proteins are members of phloem protein 2 and present in phloem sap of vascular plants. These proteins are involved in long distance signalling as well as other functions in phloem sap [32,33]. Additionally, the protein interacts with multiple proteins whose functions have not been annotated yet.

4. CONCLUSION

Amino acids are crucial sources of organic nitrogen for plants, carried from source parts to growing sink organs through the phloem and xylem. Various amino acid transporters aid in this distribution. The present study is first report concerning identification and functional characterization of amino acid transporter of *A. hypochondriacus* which has involvement in the transporation and accumulation of amino acids in grain Amaranth. A computational analysis set out to create and confirm the 3D structure of the amino acid transporter Ah_AAP, identifying it as a stable, alkaline protein belonging to the amaranth family. Bioinformatics tools were utilized in the research to create models and examine the primary, secondary, and tertiary structures of the amino acid permease proteins in *Amaranthus hypochondriacus*. Analysis of the primary structure showed that the proteins are both acidic and stable, with the secondary structure mainly consisting of alpha helices. The accuracy of the model was confirmed by the 3D structure prediction, which was validated through PROCHECK's Ramachandran plot. Although more experimentation is necessary, these results offer valuable knowledge on the Ah_AAP transporter's structure and function, showing the potential for computational methods to improve our comprehension of protein mechanisms and interactions in amino acid absorption.

DISCLAIMER (ARTIFICIAL INTELLIGENCE)

Author(s) hereby declare that NO generative AI technologies such as Large Language Models (ChatGPT, COPILOT, etc) and text-to-image generators have been used during writing or editing of manuscripts.

ACKNOWLEDGEMENTS

We expressed our gratitude to the university Govind Bhallabh Pant University of Agriculture & Technology to provide technical assistance.

COMPETING INTERESTS

Authors have declared that no competing interests exist.

REFERENCES

1. Nsholm T, Kielland K, Ganeteg U. Uptake of organic nitrogen by plants. New Phytol. 2009;182:31–48.
2. Rentsch D. Uptake and partitioning of amino acids and peptides. Mol Plant. 2010;3:997–1011.
3. Bloom AJ. The increasing importance of distinguishing among plant nitrogen sources. Curr Opin Plant Biol. 2015;25:10–16.
4. Paungfoo-Lonhienne C, Lonhienne TGA, Rentsch D, Robinson N, Christie M, Webb RI et al. Plants can use protein as a nitrogen source without assistance from other organisms. Proc Natl Acad Sci USA. 2008;105:4524–4529.
5. Yao X, Nie J, Bai R, Sui X. Amino acid transporters in plants: Identification and function. Plants. 2020;9(8):972.
6. Dinkeloo K, Boyd S, Pilot G. Update on amino acid transporter functions and on possible amino acid sensing mechanisms in plants. Semin Cell Dev Biol. 2018;74:105–113.
7. Tegeder M. Transporters for amino acids in plant cells: Some functions and many unknowns. Curr Opin Plant Biol. 2012;15:315–321.
8. Soriano-García M, Aguirre-Díaz IS. Nutritional functional value and therapeutic utilization of Amaranth. InNutritional value of amaranth 2019. IntechOpen.

9. Chauhan R, Prabhakaran S, Singh Aswal K, Pandey D. Molecular docking studies for evaluation of amaranth grain squalene for treatment of melanoma skin cancer in human beings. Lett Org Chem. 2023;20(9):829-36.
10. Pandey D, Bharti M, Rana A, Prabhakaran S, Chauhan R. Molecular docking of phytomolecules of grain amaranth (*Amaranthus hypochondriacus*) with AKR1C3 protein involved in prostate cancer in human beings. Lett Org Chem, 2024;21(8):677-686.
11. Lu Y, Song Z, Lü K, Lian X, Cai H. Molecular characterization, expression and functional analysis of the amino acid transporter gene family (OsAATs) in rice. Acta physiologiae plantarum. 2012 Sep;34:1943-62.
12. Fang ZT, Kapoor R, Datta A, Okumoto S. Tissue specific expression of UMAMIT amino acid transporters in wheat. Scientific Reports. 2022 Jan 10;12(1):348.
13. Siciliano AJ, Zhao C, Liu T, Wang Z. EGG: Accuracy estimation of individual multimeric protein models using deep energy-based models and graph neural networks. Int J Mol Sci. 2024;25(11): 6250.
14. Johnson MS, Srinivasan N, Sowdhamini R, Blundell TL. Knowledge-based protein modeling. Crit Rev Biochem mol biol. 1994;29(1):1-68.
15. Hasani HJ, Barakat K. Homology modeling: An overview of fundamentals and tools. Int. Rev. Model. Simul. 2017 Apr 30;10(2):1-4.
16. Nugent T, Jones DT. Transmembrane protein topology prediction using support vector machines. BMC bioinformatics.2009;10:1-11.
17. Gasteiger E, Hoogland C, Gattiker A, Duvaud S, Wilkins M, Appel RD, Bairoch A, Protein Identification and analysis tools on the expasy server. In: John M. Walker, editor. The Proteomics Protocols Handbook, Humana Press; 2005.
18. Marchler-Bauer A, Derbyshire M K, Gonzales NR, Lu S, Chitsaz F, Geer LY, Geer RC, et al. CDD: NCBI's conserved domain database. Nucleic acids research. 2015;43(D1):D222-D226.
19. Arab M, Najafi Zarrini H, Nematzadeh G, Heidari P, Hashemipetroudi SH, Kuhlmann M. Comprehensive analysis of calcium sensor families, CBL and CIPK, in *Aeluropus littoralis* and their expression profile in response to salinity. Genes. 2023;14(3):753.
20. Nanjareddy K, Guerrero-Carrillo MF, Lara M, Arthikala MK. Genome-wide identification and comparative analysis of the Amino Acid Transporter (AAT) gene family and their roles during Phaseolus vulgaris symbioses. FIG. 2024;24(2):47.
21. Oladipo EK, Obisanya OI, Owoeye VO, Shittu OG, Adamitonde MG, Ikwuka EC, et al. Immunoinformatics assisted design of a multi-epitope kit for detecting *Cronobacter sakazakii* in powdered infant formula (PIF). Food Qual Saf. 2024;fyae005.
22. Hajjari MM, Sharif N. Prolamins' 3D structure: A new insight into protein modeling using the language of numbers and shapes. Food Hydrocolloids. 2024;154:110154.
23. Prasasty VD, Tambunan US, Siahaan TJ. Homology modeling and molecular dynamics studies of EC1 domain of VE-cadherin to elucidate docking interaction with cadherin-derived peptide. OnLine J Biol Sci. 2014;14(2):155.
24. Villoutreix BO, Bastard K, Sperandio O, Fahraeus R, Poyet JL, Calvo F, et al. *In silico In vitro* screening of protein-protein interactions: towards the next generation of therapeutics. Curr Pharm Biotechnol. 2008;9(2):103-22.
25. Yang S, Liu D, Song Y, Liang Y, Yu H, Zuo Y. Designing a structure-function alphabet of helix based on reduced amino acid clusters. Arch Biochem Biophys. 2024; 109942.
26. Firmansyah RP, Alimah SN, Artika IM, Kurniatin PA. *In silico* phylogenetic, physicochemical, and structural characteristics of phytase enzyme from ten Aspergillus species. Menara Perkebunan. 2024;92(1).
27. Thaker K, Patoliya J, Rabadiya K, Patel D, Ponnuchamy M, Rama Reddy NR, Joshi R. An *in silico* approach to unravel the structure of 3-deoxy-D-arabino-heptulosonate 7-phosphate synthase (DAHPS): A critical enzyme for sennoside biosynthesis in Cassia angustifolia Vahl. J Biomol Struct Dyn. 2024;42(8): 3848-61.
28. Bouqellah NA, Farag PF. *In silico* evaluation, phylogenetic analysis, and structural modeling of the class II hydrophobin family from different fungal phytopathogens. Microorganisms. 2023; 11(11):2632.

29. George R, Georrge JJ, Singh SP, Vaidya A. Comparative analysis of thermophilic alpha-amylase using *In silico* approach. In Proceedings of the National Conference on Innovations in Biological Sciences (NCIBS) 2020 Apr 10.
30. Laskowski RA, MacArthur MW, Moss DS, Thornton JM. PROCHECK: A program to check the stereochemical quality of protein structures. J appl crystallogr. 1993;26(2): 283-91.
31. Binbay FA, Rathod DC, George AA, Imhof D. Quality assessment of selected protein structures derived from homology modeling and AlphaFold. Pharmaceuticals. 2023 Nov 29;16(12):1662.
32. Vastermark A, Wollwage S, Houle ME, Rio R, Saier Jr MH. Expansion of the APC superfamily of secondary carriers. Proteins: Struct, Funct, Bioinf. 2014;82 (10):2797-811.
33. Dinant S, Clark AM, Zhu Y, Vilaine F, Palauqui JC, Kusiak C, Thompson GA. Diversity of the superfamily of phloem lectins (phloem protein 2) in angiosperms. Plant physiol. 2003;131(1):114-28.

Peer-review history:
The peer review history for this paper can be accessed here:
https://www.sdiarticle5.com/review-history/119678