2023 年臺灣國際科學展覽會
優勝作品專輯

作品編號 **080010**

參展科別 生物化學

作品名稱 **In silico identification and physicochemical analysis of potential novel antimicrobial peptides from Momordica charantia L.**

得獎獎項

國　　家 **Philippines**

就讀學校 **Philippine Science High School - Main Campus**

指導教師 **Elliard Roswell Yanza**

作者姓名 **Janssen Reign B. Pena**

Rysa Antonette H. Ang

Vivien Leila A. Ercia

關鍵詞 **antibiotic resistance、antimicrobial peptides、bioinformatics**

作者照片

ABSTRACT

The emergence of antibacterial resistance has necessitated the development of alternative treatments, such as antimicrobial peptides (AMPs). AMPs are part of the innate immune systems of various organisms such as *Momordica charantia* L., a known medicinal plant in Southeast Asia. In this study, potential novel AMPs from *M. charantia* were derived *in silico* to provide prospective antibiotic alternatives using promising plant-based peptides. *M. charantia* protein sequences that were 500 amino acids long were digested using proteolytic enzymes, resulting in 3,621 peptides. Each resulting sequence was characterized as either AMP or Non-AMP using four statistical analysis tools, and those identified as AMPs were analyzed. This led to 102 AMPs, 53 of which were unregistered on the Data Repository for Antimicrobial Peptides, indicating that they have yet to be derived from other species. Six of the eight studied physicochemical properties show strong correlations with each other, suggesting that subsequent AMP design studies may focus on these six properties. As such, *M. charantia* may be a rich source of potential AMPs and, thereby, alternative antibiotics. The *in vitro* examination of these novel AMPs is also recommended to further understand their potential as alternative antibiotics sourced from locally available plants.

INTRODUCTION

Antibiotic resistance is a growing global threat to human health, causing around 700,000 deaths worldwide every year. Since available antibiotics have become less effective against resistant infections (Edwards et al., 2021; HM Government, 2019), new anti-infectives such as antimicrobial peptides (AMPs) have been recognized as potential drug components for their capacity to eliminate pathogenic microbes (Mahlapuu et al., 2020). However, the Inter-Agency Committee on Antimicrobial Resistance (2019) reports that there is no research in the Philippines on novel anti-infectives.

Momordica charantia L. is a common vegetable in Southeast Asia that has a significant potential antimicrobial activity due to the presence of compounds such as charantin and plumericin (Villarreal-La Torre et al., 2020).

This study focused on identifying potential novel antimicrobial peptides *in silico* from *M. charantia* and analyzing their physicochemical properties as potential components for alternative antibiotics. Specifically, the study determined the AMP prediction score of the peptides as a metric for identifying the potential AMPs derived from *M. charantia*. Furthermore, the study quantified the physicochemical properties of the identified AMPs *in silico* and determined the correlation of these properties. Finally, the study ascertained their novelty.

M. charantia-derived AMPs identified through the study could be considered by drug developers as candidates for alternatives to antibiotics, benefiting the pharmaceutical industries of not only the Philippines but also Southeast Asia. The current study only used *in silico* processes to predict antimicrobial potential and physicochemical properties.

METHODOLOGY

Preparation of M. charantia Proteins

Momordica charantia L. proteins were first identified through the National Center for Biotechnology (NCBI) database (https://www.ncbi.nlm.nih.gov/). All proteins that were 500 amino acids in length were then stored.

Determination of Enzymes

Based on the methods conducted by Suryawanshi & Chouhan (2016), this study used the following enzymes to cut the selected proteins into peptides: pepsin, trypsin, and proteinase K. Exactly 10 of the 500-amino acid long *M. charantia* proteins were selected via simple random sampling.

These 10 proteins were then run through ExPASy's PeptideCutter tool (https://web.expasy.org/peptide_cutter/) to cut them into peptides. As such, because pepsin, trypsin, and proteinase K were able to cut all of the sampled proteins into peptides, they were the enzymes used for the study.

Derivation of M. charantia Potential AMPs

After the proteins and enzymes were selected, all the identified proteins were run through the PeptideCutter tool on ExPASy. The resulting peptides that were at least three amino acids long were then recorded, with recurring amino acid sequences omitted from the data tables (Suryawanshi & Chouhan, 2016).

Each *M. charantia*-derived peptide was processed using the AMP Prediction Tool of the Collection of Antimicrobial Peptides (CAMP) online database (http://www.camp3.bicnirrh.res.in/predict/). Four algorithms were run: (1) Support Vector Machine (SVM), (2) Random Forest (RF), (3) Artificial Neural Network (ANN), and (4) Discriminant Analysis (DA) (Suryawanshi & Chouhan, 2016). The results produced from the four models each included a peptide class of either AMP or NAMP (Non-AMP); a numerical AMP probability score was also included in all aforementioned algorithms except for ANN (Suryawanshi & Chouhan, 2016). The resulting AMP prediction scores and peptide classes for each *M. charantia*-derived peptide were processed in batches and then recorded.

Collected AMP probability scores from the statistical analyses were examined to find sequences that obtained a predictability score of 0.435 at minimum, in the SVM, RF, and DA statistical tests, and an AMP rating in the ANN test (Lei et al., 2021; Suryawanshi & Chouhan, 2016). Peptide sequences that satisfied these conditions were identified as potential antimicrobial peptides. The list of identified potential *M. charantia*-derived AMPs was then prepared.

Data Gathering

The physicochemical properties of the identified potential AMPs were computed through the "Multiple Peptides" page of CellPPD (https://webs.iiitd.edu.in/raghava/cellppd/multi_pep.php). AMPs were processed in batches of 10, each beginning by inputting the sequences of each peptide in FASTA format. The SVM-based prediction method was selected and the SVM threshold was set to -0.1. The following physicochemical properties were calculated: length, molecular weight, amphipathicity, hydrophobicity,

hydropathicity, side bulk, net hydrogen charge, and isoelectric point (Tomazou et al., 2019). To verify the physicochemical values, each *M. charantia*-derived potential AMP was processed on CellPPD three times. The resulting value for each physicochemical property of each potential AMP was recorded afterward through data tables.

The sequence of each identified *M. charantia*-derived potential AMP was inputted in the "Sequence Search" page of CAMP, which compared it to the sequences of all the patented peptides found in the database. If no matches were found, each sequence was then inputted in the "Quick Search" section of another database called Data Repository of Antimicrobial Peptides (DRAMP) (http://dramp.cpu-bioinfor.org/). It compared the same inputted sequence to the sequences of all patented peptides found in the database. If no matches were found in both the databases of CAMP and DRAMP, the potential AMP was considered novel.

Data Analysis

The physicochemical properties were analyzed through a Pearson's correlation coefficient test, via the R software (Version 4.1.3). The correlation between the physicochemical properties' combinations, such as length and molecular weight, hydrophobicity and hydropathicity, and every other possible combination, was measured. The Pearson product-moment correlation type and pairwise-complete observations were used. The pairwise p-values of each physicochemical property combination were also included in the matrix.

Each resulting correlation from the correlation matrix was visualized through a scatterplot generated via the R software. The data was presented using histograms and the following options were also shown: least-squares lines, smooth lines, and spread. The resulting high correlation coefficient values (values between ±0.50 and ±1) from any of the physicochemical property combinations were then noted.

RESULTS

Potential M. charantia-derived AMPs

Obtaining all 500-amino acid long *M. charantia* proteins in CAMP resulted in a total of 31 protein sequences. When ran through ExPASy with enzymes pepsin, trypsin, and proteinase K, 3,621 unique peptide sequences were derived. Among these, CAMP predicted that only 102 may be considered potential AMPs. This meant that they received a prediction score of 0.435 on the SVM, RF, and DA models, and a classification of "AMP" on the ANN test.

Further, of these 102, 53 were found to be unregistered on both the databases of CAMP and DRAMP. Thus, 53 *M. charantia*-derived potential AMPs were considered to be novel in this study.

Physicochemical Properties

To determine the trend of AMPs and NAMPs, eight of their physicochemical properties were investigated. Table 1 below shows the mean values of the physicochemical properties recorded for each *M. charantia*-derived AMPs and NAMPs.

Table 1. Mean values of physicochemical properties of AMPs and NAMPs

Physicochemical property	Mean of AMPs	Mean of NAMPs
peptide length	15.87	8.68
molecular weight	1,733.89	938.33
amphipathicity	0.91	0.59
hydrophobicity	-0.24	-0.24
hydropathicity	-0.58	-0.73
side bulk	0.60	0.61
net hydrogen charge	0.98	0.93
isoelectric point	9.80	19.64

Table 2, on the other hand, shows the Pearson's correlation coefficients of the significant physicochemical property pairs. It was determined that 17 physicochemical property combinations possessed a high correlation coefficient value out of the 28 distinct combinations. Moreover, six of the eight physicochemical properties have high correlation coefficient values or strong relationships with at least five other properties. For several property pairs, their Pearson's correlation coefficient values are similar to those of Suryawanshi & Chouhan (2016).

Table 2. Physicochemical properties of *Momordica charantia* L. potential AMPs with strong Pearson correlation

Properties with strong correlation	Pearson's correlation coefficient
peptide length - molecular weight	0.9935
peptide length - amphipathicity	-0.7142
peptide length - hydrophobicity	0.6221
peptide length - hydropathicity	0.5528
peptide length - net hydrogen charge	-0.5904
molecular weight - amphipathicity	-0.6942
molecular weight - hydrophobicity	0.5881

Table 2 (continued).

molecular weight - hydropathicity	0.5184
molecular weight - net hydrogen charge	-0.5448
amphipathicity - hydrophobicity	-0.8779
amphipathicity - hydropathicity	-0.8222
amphipathicity - net hydrogen charge	0.7972
hydrophobicity - hydropathicity	0.9625
hydrophobicity - net hydrogen charge	-0.9599
hydrophobicity - isoelectric point	-0.5043
hydropathicity - net hydrogen charge	-0.8964
net hydrogen charge - isoelectric point	0.5318

DISCUSSION

M. charantia-derived Potential AMPs

Out of the 102 identified potential AMPs, 76 were categorized as AMPs by all four predictions. However, for the rest of the potential AMPs, the RF algorithm predicted 23 to be non-antimicrobial, which was more than all the other models. Thomas et al. (2010) drew similar results, noting that DA and SVM accurately classified the antimicrobial potential of peptides that RF could not. Thus, the observed trend supports the suggestion that all three tests should be conducted together in analysis.

In comparison to the study's 102 *M. charantia*-sourced AMPs, Suryawanshi and Chouan (2016) note that out of the 1,813 peptides they derived from *Curcuma longa* L., 28 were deemed potential AMPs. This study found more potential AMPs, but both studies are considered significant in aiding to find potential antibiotic alternatives. Before the investigations, the plant used in each study were already reported to have antimicrobial effects due to them being traditionally used as medicine. Such productivity in these *in silico* screenings supports the reported traditional medical use of both plants as antibiotics were inaccessible. (Suryawanshi & Chouhan, 2016; Villarreal-La Torre et al., 2020).

Physicochemical Properties

As seen in Table 1, it was found that there is a large difference between the mean peptide lengths ($\Delta = 7.19$) and molecular weights ($\Delta = 795.56$) of AMPs and NAMPs. The trend observed may be attributed to how the protein's molecular weight is the total molecular weight of all its amino acids (Guan et al., 2015). Since novel AMPs are seen to have longer peptide lengths, their mean molecular weight would also be generally

higher. In addition, longer AMPs are more commonly found to possess antimicrobial activity, which explains how novel AMPs have greater peptide lengths compared to NAMPs (Clark et al., 2021).

In Table 2, the peptide length-molecular weight (0.9935) property pair was seen to possess a strong correlation. Peptide length and molecular weight being strongly correlated support statements from a previous study by Suryawanshi & Chouhan (2016), as well as the similar trends observed between the means of the *M. charantia*-derived AMPs and NAMPs. According to Clark et al. (2021), antimicrobial activity is usually common in longer peptide lengths. This could imply that antimicrobial activity may be seen more in peptides with greater molecular weight because of how a longer peptide length includes peptides having more amino acids. With this, the molecular weight of each amino acid contributes to the peptide's total molecular weight. It could further be concluded that peptide length is strongly correlated with molecular weight, making these two properties important when considering the antimicrobial activity of peptides.

As such, all of the strong correlations between physicochemical properties are deemed important in investigating the antimicrobial activity. The significant difference in the properties' means compared to NAMPS is also important, so the correlations and means may suggest the effects of physicochemical properties on antimicrobial activity.

SUMMARY AND CONCLUSION

The study identified 102 potential antimicrobial peptides (AMPs) from *Momordica charantia* L., 53 of which are novel as they have yet to be registered in the used AMP databases. Six of eight physicochemical properties were found to have strong correlations with each other. Along with the significant differences in the mean prediction scores of AMPs and NAMPs, these suggest that physicochemical properties may have an effect on the antimicrobial activity of peptide sequences. Future studies may examine the antimicrobial activity of AMPs from *M. charantia in vitro* and further analyze the properties with strong correlations.

REFERENCES

Chang, D. K., & Liang, C. C. (1994). Influence of bulky side chains of amino acids on the solution conformation of peptide fragment (81–92) derivatives of CD4, TYICEVEDQKEE, as studied by

NMR spectroscopy and molecular modeling. *Biochimica et Biophysica Acta (BBA)-Protein Structure and Molecular Enzymology*, *1205*(2), 262-267. https://doi.org/10.1016/0167-4838(94)90243-7

Clark, S., Jowitt, T. A., Harris, L. K., Knight, C. G., & Dobson, C. B. (2021). The lexicon of antimicrobial peptides: A complete set of arginine and tryptophan sequences. *Communications Biology*, *4*, 605. https://doi.org/10.1038/s42003-021-02137-7

Edwards, F., MacGowan, A., & Macnaughton, E. (2021). Antibiotic resistance. *Medicine*, *49*(10), 632-637. https://doi.org/10.1016/j.mpmed.2021.07.006

Guan, Y., Zhu, Q., Huang, D., Zhao, S., Jan Lo, L., & Peng, J. (2015). An equation to estimate the difference between theoretically predicted and SDS PAGE-displayed molecular weights for an acidic peptide. *Scientific Reports*, *5*(1), 1-11. https://doi.org/10.1038/srep13370

HM Government (2019, January 24). Tackling antimicrobial resistance 2019–2024: The UK's five-year national action plan. https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/784894/UK_AMR_5_year_national_action_plan.pdf

Inter-Agency Committee on Antimicrobial Resistance. (2019). *The Philippine action plan to combat antimicrobial resistance: A one health approach.* World Health Organization. https://www.who.int/publications/m/item/philippines-the-philippine-action-plan-to-combat-antimicrobial-resistance-one-health-approach

Lei, Y., Li, S., Liu, Z., Wan, F., Tian, T., Li, S., Zhao, D., & Zeng, J. (2021). A deep-learning framework for multi-level peptide–protein interaction prediction. *Nature communications, 12*(1), 1-10. https://doi.org/10.1038/s41467-021-25772-4

Mahlapuu, M., Björn, C., & Ekblom, J. (2020). Antimicrobial peptides as therapeutic agents: Opportunities and challenges. *Critical Reviews in Biotechnology*, *40*(7), 978-992. https://doi.org/10.1080/07388551.2020.1796576

Suryawanshi, S. K., & Chouhan, U. (2016). Application of bioinformatics in the prediction and identification of potential antimicrobial peptides from *Curcuma longa*. *Bioscience Biotechnology Research Communications, 9*(2), 216-224. Retrieved October 1, 2021, from http://bbrc.in/bbrc/papers/pdf%20files/Volume%209%20-%20No%202%20Apr-June%202016/BBRC7_008.pdf

Thomas, S., Karnik, S., Barai, R. S., Jayaraman, V. K., & Idicula-Thomas, S. (2010). CAMP: A useful resource for research on antimicrobial peptides. *Nucleic Acids Research*, *38*(suppl_1), D774-D780. https://doi.org/10.1093/nar/gkp1021

Tomazou, M., Oulas, A., Anagnostopoulos, A. K., Tsangaris, G. T., & Spyrou,

G. M. (2019). *In silico* identification of antimicrobial peptides in the proteomes of goat and sheep milk and feta cheese. *Proteomes, 7*(4), 32. https://doi.org/10.3390/proteomes7040032

Villarreal-La Torre, V. E., Guarniz, W. S., Silva-Correa, C., Cruzado-Razco, L., & Siche, R. (2020). Antimicrobial activity and chemical composition of *Momordica charantia*: A review. *Pharmacognosy Journal, 12*(1). https://doi.org/10.5530/pj.2020.12.32

【評語】080010

The author applied computer simulation to identify potential antimicrobial peptides (AMP) from Momordica charantia L.，a common vegetable in Southeast Asia with a significant potential antimicrobial activity. Identification of novel AMPs is vital in treating multiple drug-resistant bacteria. However，the authors did not explain what these peptides derived from or how to define non-AMP and AMP. Whether the selected AMP indeed has antimicrobial activities are required for further confirmation. What is the significance of the strong correlations of physicochemical properties ? How to purify the AMP or make the identified peptides for future biological assays may need to be considered for pharmaceutical use.

Suggestions:

1. Only in silico analysis was used，but not the real experiments to confirm the peptides are really AMPs.
2. Previously，the compounds such as charantin and plumericin were identified to have antimicrobial activity in M. charantia (Villarreal-La Torre et al.，2020). Why were they so sure there will be AMPs in this plant ? How were they sure the digested peptides by pepsin，

trypsin，and proteinase K could be potent enough AMPs？There are different types of AMPs，and each type has their characteristic sequence requirement. How did they set the criteria for in silico analysis to select different types of AMPs？I know they were using the AMP Prediction Tool of the Collection of Antimicrobial Peptides (CAMP) online database (http://www.camp3.bicnirrh.res.in/predict/)，but how accurate the prediction is and how potent the identified peptide could be？Did they use known AMPs to test the tool for scoring.

3. For the Physicochemical properties of the potential AMPs listed in Table 1，are these properties important measurements for AMPs？For example，what is the unit for peptide length？Is the mean molecular weight of 1733.89 too large for AMPs？

4. In Table 2，the peptide length-molecular weight (0.9935) property pair to possess a strong correlation is not surprising. Even though antimicrobial activity is usually common in longer peptide lengths，it does not mean the longer peptides could be antibiotics because they do not have favorable drug-like properties.