scientific reports

OPEN

Machine learning assisted rational design of antimicrobial peptides based on human endogenous proteins and their applications for cosmetic preservative system optimization

Lizhi Yue[1,2], Liya Song[1], Siyu Zhu[3,4], Xiaolei Fu[3,4], Xuhui Li[3,5], Congfen He[1]✉ & Junxiang Li[3,4]✉

Preservatives are essential components in cosmetic products, but their safety issues have attracted widespread attention. There is an urgent need for safe and effective alternatives. Antimicrobial peptides (AMPs) are part of the innate immune system and have potent antimicrobial properties. Using machine learning-assisted rational design, we obtained a novel antibacterial peptide, IK-16-1, with significant antibacterial activity and maintaining safety based on β-defensins. IK-16-1 has broad-spectrum antimicrobial properties against *Escherichia coli*, *Staphylococcus aureus*, *Pseudomonas aeruginosa*, and *Candida albicans*, and has no haemolytic activity. The use of IK-16-1 holds promise in the cosmetics industry, since it can serve as a preservative synergist to reduce the amount of other preservatives in cosmetics. This study verified the feasibility of combining computational design with artificial intelligence prediction to design AMPs, achieving rapid screening and reducing development costs.

The cosmetics industry is one of the most rapidly developing sectors worldwide[1]. Preservatives are essential ingredients in cosmetic products that inhibit microbial growth and prevent spoilage[2]. Common preservatives include parabens, formaldehyde releasers, isothiazolinones, phenoxyethanol, and organic acids[3]. However, preservatives are among the most common allergens in cosmetics[4]. Some preservatives can cause cosmetic contact dermatitis[5], such as methylisothiazolinone (MI)[6], methylchloroisothiazolinone/methylisothiazolinone (MCI/MI)[7], parabens[8], and formaldehyde releasers[9]. Furthermore, certain preservatives in leave-on cosmetics may disrupt skin microbiota balance[10]. Cosmetic products with low or no preservatives have gained popularity. While combinations of various preservatives represents a current method of improvement[11], it cannot fundamentally solve this problem. Therefore, there is an urgent need for safe, potent, broad-spectrum antimicrobial agents.

Other substances present in cosmetic products that also have antibacterial effects are known as 'preservative synergists'[8]. Mixing preservatives and preservative synergists effectively decreases the concentration of allergenic preservatives. Certain antibacterial substances, such as glycolipids[12] and new antimicrobial nanomaterials[13–15], may be used as preservative synergists. However, there are few reports on the development of biomimetic antimicrobial peptide derived from human endogenous proteins for use as cosmetic preservatives.

Antimicrobial peptides (AMPs) are oligopeptides with varying numbers of amino acids (from 5 units to > 100 units)[16]. AMPs are important components of the innate immune system in many organisms and provide the first line of defense against various pathogens[17]. AMPs exhibit various activities against Gram-negative and

[1]Key Laboratory of Cosmetic of China National Light Industry, School of Light Industry Science and Engineering, Beijing Technology and Business University, Beijing, China. [2]School of Chemistry and Chemical Engineering, Qilu Normal University, Shandong, China. [3]AGECODE R&D Center, Yangtze Delta Region Institute of Tsinghua University, Zhejiang, China. [4]Harvest Biotech (Zhejiang) Co., Ltd., Zhejiang, China. [5]Zhejiang Provincial Key Laboratory of Applied Enzymology, Yangtze Delta Region Institute of Tsinghua University, Zhejiang, China. ✉email: congfenhe@126.com; lijunxiang@acrdc.cn

Gram-positive bacteria, fungi, mycobacteria, and some enveloped viruses[18,19]. Most AMPs can be classified into four groups according to their secondary structure: α-helical, β-sheet, loop, and extended peptides[16]. Among these structural groups, α-helix and β-sheet structures are more common[20]. α-helical peptides are the most studied AMPs to date. In α-helix structures, the distance between two adjacent amino acids is around 0.15 nm, and the angle between them with respect to the centre is around 100° from the top view[21]. Human β defensin type 3 (hBD-3) is a cysteine-rich small antibacterial peptide with broad-spectrum antimicrobial activity and may play an important role in innate epithelial defense. Because hBD-3 kills bacteria by disrupting the cell membrane, which is relatively stable in chemical composition, its antibacterial activity will not cause drug resistance[22]. Thus, hBD-3 could serve as a potential AMP for designing novel preservatives. However, natural AMPs are large protein molecules that are unstable in various in vitro environments, which limits their applications. Therefore, developing biomimetic AMPs with improved efficacy, greater stability, and lower costs is challenging. Machine learning (ML) is described as the capacity of a system to learn from problem-specific training data to automate the process of analytical model building and solving associated tasks[23]. ML greatly improves work efficiency and accuracy. Driven by increasing computer power and algorithmic advances, ML has become a powerful tool for finding data patterns[24], and it is widely applied in various fields, such as biology[25,26], medicine[27,28], healthcare[29], environment[30], fuel cells[31] and energy[32]. ML methods can also be used to develop new AMPs. As more insights are gained into the antimicrobial mechanism, sequence, and structure, ML can gather information to expedite AMP screening and design[33–35].

The primary objective of this study was to develop safe and potent broad-spectrum antimicrobial agents to reduce the use of preservatives in cosmetics. In this study, the artificial antibacterial peptide was designed based on hBDs through computational simulation and artificial intelligence (AI) based ML tools, and their antibacterial and haemolytic properties were predicted.

Methods
Designing antimicrobial peptides
AI deep learning based on prediction algorithms and rational molecular design has been used to develop new biomimetic AMPs based on endogenous human proteins[36–38]. Specifically, short analogues of β-defensin were used as template peptides and helical wheel projection was conducted. Negatively charged amino acid residues, including Asp (D) and Glu (E), were substituted with positively charged amino acid residues, including Lys (K) and Arg (R), to increase the net positive charge. Several residues were adjusted such that the hydrophobic and hydrophilic residues were located on opposite sides of the helical surface to improve the pore-forming ability of the microbial membrane[33]. The *ab-initio* three-dimensional structures of peptides were studied to support the α-helical conformations. Peptides were predicted for their antimicrobial ability by some AI programs[39–42]. For safety reasons, the haemolytic activity was also predicted using an available AI algorithm[43]. Antimicrobial activity was predicted using the Antimicrobial Peptide Scanner version 2 (https://www.dveltri.com/ascan/v2/ascan.html), and haemolytic activity was predicted using HAPPENN (http://research.timmons.eu/happenn). All the new top-scoring biomimetic peptides were synthesised and screened for their microbial inhibitory activity.

Disk diffusion test
Four types of bacteria and fungi, namely, *Staphylococcus aureus*, *Escherichia coli*, *Pseudomonas aeruginosa*, and *Candida albicans* (ATCC, Manassas, VA, USA), were cultured under standard conditions. *Candida albicans* was cultured with Sabouraud glucose broth medium (Solarbio Life Science, Beijing, China) at 30 °C and 200 rpm; *Staphylococcus aureus*, *Escherichia coli*, and *Pseudomonas aeruginosa* were cultured with Luria Broth (LB) liquid medium (Solarbio Life Science, Beijing, China) at 37 °C and 200 rpm. The paper disks containing 30 μg Kanamycin (Beyotime Biotechnology, Shanghai, China) and 30 μg Amphotericin (Beyotime Biotechnology) were assigned as the positive control, and the medium was assigned as the negative control. The paper disks incorporated with 30 μg of AMP were deposited on the surface of the agar plates pre-inoculated with the bacterial strain to be tested. Specifically, 60 μg AMPs were used to test against *Candida albicans*. The agar plates were incubated for 24 h under appropriate conditions. *Candida albicans* was cultured in Sabouraud glucose agar medium in an inverted incubator at 30 °C for 12 h, and the remainder was cultured in LB solid medium in an inverted incubator at 37 °C for 12 h. The diameter of the inhibition zone was measured for statistical analysis. Each test was performed in triplicates.

Minimum inhibitory concentration
Four types of bacteria and fungi–*Staphylococcus aureus*, *Escherichia coli*, *Pseudomonas aeruginosa*, and *Candida albicans*–were cultured under standard conditions. *Candida albicans* was cultured in Sabouraud glucose broth medium (Solarbio Life Science, Beijing, China) in a static incubator at 30 °C for 12 h, and the remainder were cultured with LB liquid medium (Solarbio Life Science, Beijing, China) in a static incubator at 37 °C for 12 h. The bacterial order of magnitude was 10^4/mL–10^5/mL. A series of two-fold dilutions were performed from the stock solution to obtain different concentrations of peptides from 250 to 1.95 μg/mL, with a total volume of 200 μL per well in a 96-well plate, and then assayed for absorbance at 540 nm. This medium was used as the solvent and was designated as the negative control. Approximately 1×10^4 cells were incubated in a medium containing peptides. After incubating the cells at the appropriate temperature for 24 h, the minimum inhibitory concentration was the lowest peptide concentration that limited visible microbial growth.

Hemolytic activity
For the haemolysis studies, rabbit erythrocytes (Fenghui Biology, Hunan, China) were obtained from healthy blood donors. Blood samples mixed with 0.25% potassium oxalate were immediately separated by centrifugation

at 1500 rpm for 10 min. Subsequently, the cells were washed three times with phosphate buffered saline (PBS, Yuanpei Biology, Shanghai, China) and diluted to a 5% suspension with PBS. 250 μL of erythrocyte stock dispersion was mixed with 250 μL of PBS containing different concentrations of AMPs and incubated at 37 °C for 1 h. Intact erythrocytes were removed by centrifugation at 10,000 rpm for 5 min. The absorbance of the resulting supernatant was measured at 540 nm. PBS was used as a negative control (0% lysis), and 0.1% SDS (Solarbio Life Sciences) was used as a positive control (100% lysis). The haemolysis rate was determined using Eq. (1).

$$Haemolysis\ Rate\ (\%) = \frac{D_{test} - D_{nc}}{D_{pc} - D_{nc}} \times 100\%$$ (1)

where D_{test}, D_{nc}, and D_{pc} are the 540 nm absorbance of the tested sample, negative control, and positive control, respectively.

Cytotoxicity

HaCaT cells (Fenghui Biology, Hunan, China), a long-lived, spontaneously immortalised human keratinocyte line, was grown in Dulbecco's Modified Eagle Medium (Wisent Bioproducts, Quebec, Canada) supplemented with 10% heat-inactivated foetal bovine serum (FBS, Invitrogen, Carlsbad, CA, USA) and 1% penicillin–streptomycin (Langke Biology, Suzhou, China) at 37 °C and 5% CO_2 until 80–90% of fusion was measured. The experiments were performed in 96-well plates with a final volume of 100 μL of medium per well. HaCaT cell proliferation was tested using the cell counting kit-8 assay (Dojindo Laboratories, Tokyo, Japan). The cells were seeded at a density of 0.5×10^5–1.0×10^5 cells/mL and grown at 37 °C until 80–90% fusion was measured. Subsequently, samples at specific concentrations were added to the medium. After incubating the cells for 24 h, all media were removed, and the cells were washed three times with PBS. Next, the cells were incubated for 4 h with fresh medium (100 μL/well) containing 10 μL of the probe. The absorbance was read at 450 nm using a microplate reader.

Antimicrobial assay in products

The basic preservation system contained 1.8% 1,3-propanediol, 0.09% 1,2-Hexanediol, and 0.036% capryl hydroxamic acid. In the synergy test, the 5 μg/mL antimicrobial peptide was mixed with the preservative system and compared with the preservative system without antimicrobial peptide. Four types of bacteria and fungi–*Staphylococcus aureus*, *Escherichia coli*, *Pseudomonas aeruginosa*, and *Candida albicans*–were cultured under standard conditions. *Candida albicans* was cultured with Sabouraud glucose broth medium (Solarbio Life Science, Beijing, China) at 30 °C and 200 rpm; *Staphylococcus aureus*, *Escherichia coli*, and *Pseudomonas aeruginosa* were cultured with LB liquid medium (Solarbio Life Science, Beijing, China) at 37 °C and 200 rpm. Following culturing, 1.0×10^6 CFU/mL was added to the cosmetic products and mixed thoroughly. At hour 2 (day 0), day 3, day 7, day 14, and day 22, 1.0 mL of the product was collected and neutralised for 30 min using a 9 mL neutraliser. The 1.0 mL of the neutralised product was centrifuged at 12,000 rpm for 5 min to remove the suspension. The product was resuspended in 0.2 mL of fresh medium, and the CFUs were calculated using the spreading plate method. CFU refers to the number of individual colonies of any microorganism that can grow on a plate of media. The standard unit of measure for CFU is the number of culturable microorganisms present per 1 mL of culture (CFU/mL). The colony forming units per mL (CFU/mL) was determined using Eq. (2).

$$Colony\ Forming\ Units\ per\ mL\ (CFU/mL) = \frac{N}{d} \times V$$ (2)

where N is the number of colonies on the counting plate; d is the sampling volume (mL); V is the dilution multiple.
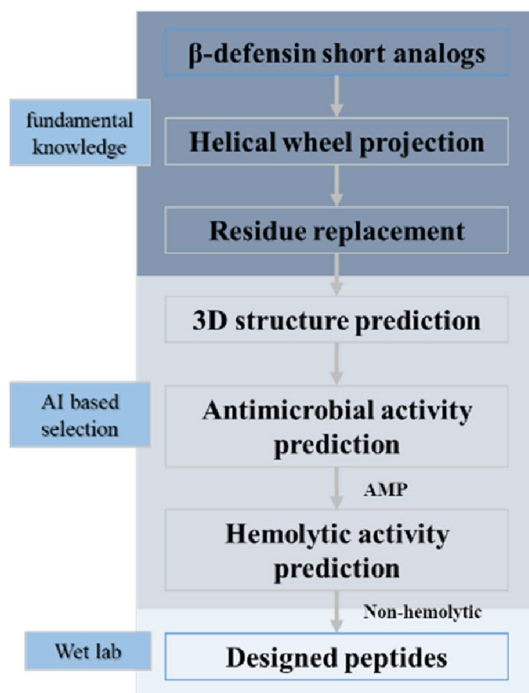
Results

### The design of antimicrobial peptides

The AMP was selected from the mesenchymal stem cell secretome and used as a template. Figure 1 illustrates the design process of the antibacterial peptides. We selected the amino acid sequences from the 30th to 45th position of beta-defensin 103 as the template and optimised it to determine the amino acid sequence of the target AMP IGKVLTRVKLLRRIK, named IK-16-1. As depicted in Fig. 2, the predicted structure was an α-helix. The probability score for AMPs was 1, indicating that IK-16-1 is highly likely to become an AMP. The probability of haemolysis was 0.006, indicating that IK-16-1 was less likely to be haemolytic.

The activities of peptides

To test the inhibitory activity of the peptides, a disk diffusion test and the minimum inhibitory concentration assay was performed. As shown in Table 1, The results showed that IK-16-1 influenced the growth of the four types of microbes. The minimum inhibitory concentration was determined to further test the effects of IK-16-1. As shown in Table 2, IK-16-1 exhibited a broad-spectrum inhibitory effect, although its antifungal activity was not as high as that of antibacterial. Based on the zone of inhibition and minimum inhibitory concentration values, IK-16-1 was found to have broad-spectrum antimicrobial properties against *Escherichia coli*, *Staphylococcus aureus*, *Pseudomonas aeruginosa*, and *Candida albicans*, which are common microbes that cause cosmetic spoilage.

Figure 1. Workflow of antimicrobial peptide design.
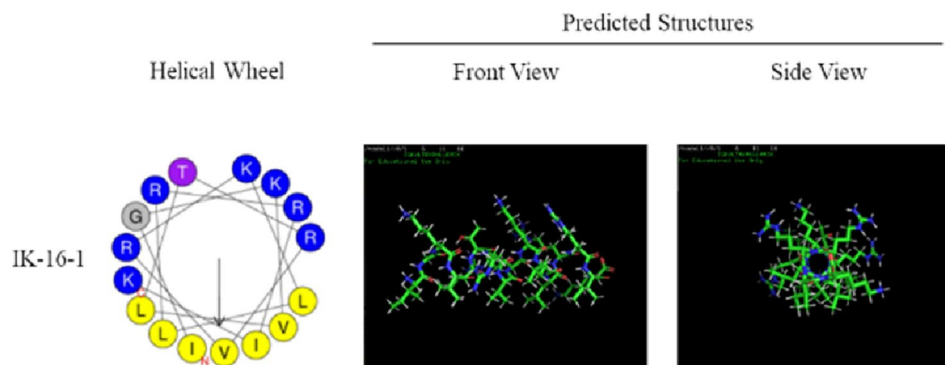
Figure 2. Results of Helical Wheel Projection and predicted structures of beta-defensin analogs.

Species	Samples	Diameter (mm)			Mean ± SD (mm)
Staphylococcus aureus (18 h, 30 μg)	Kanamycin	26.4	25.5	26.1	26.0 ± 0.4
	IK-16-1	12.3	11.2	12.0	11.8 ± 0.5
Escherichia coli (18 h, 30 μg)	Kanamycin	14.6	15.6	15.0	15.1 ± 0.4
	IK-16-1	11.2	11.3	11.5	11.3 ± 0.1
Pseudomonas aeruginosa (18 h, 30 μg)	Kanamycin	16.3	15.1	16.0	15.8 ± 0.5
	IK-16-1	11.5	10.7	10.6	10.9 ± 0.4
Candida albicans (24 h)	Amphotericin (30 μg)	15.0	15.4	15.2	15.2 ± 0.2
	IK-16-1 (60 μg)	9.6	10.3	8.7	9.5 ± 0.7

Table 1. Zone of inhibition of antimicrobial peptides (AMPs) against different species. *SD* standard deviation.

Species	MIC (µg/mL, 18 h)			Mean ± SD (µg/mL, 18 h)
Candida albicans	125.0	125.0	125.0	125.0 ± 0.0
Escherichia coli	125.0	125.0	62.5	104.2 ± 36.1
Pseudomonas aeruginosa	125.0	62.5	62.5	83.3 ± 36.1
Staphylococcus aureus	125.0	62.5	62.5	83.3 ± 36.1

Table 2. Minimum inhibitory concentration (MIC) of IK-16-1 against different species.

The safety of peptides

The cytotoxicity and haemolytic activity of IK-16-1 were evaluated to determine its safety. Compared to the blank group, 10 ppm IK-16-1, a common concentration used in cell models, was found to be non-cytotoxic to keratinocytes. As depicted in Fig. 3, no haemolytic activity was observed even at the highest concentration of the peptide solution used.

The application of peptides in cosmetic products

To validate the use of IK-16-1 as a synergistic preservative in cosmetics, a common preservative formulation was used in combination with IK-16-1. As shown in Table 3, IK-16-1 significantly inhibited microbial growth at an early stage.

Discussion

In the present study, we obtained a novel antibacterial peptide IK-16-1 designed based on β-defensin 103 with significant antibacterial activity and maintaining safety by using machine learning. These results confirmed our hypothesis and verified that the workflow for designing AMPs is feasible. Based on the zone of inhibition and minimum inhibitory concentration values, IK-16-1 was found to have broad-spectrum antimicrobial properties against *Escherichia coli*, *Staphylococcus aureus*, *Pseudomonas aeruginosa*, and *Candida albicans*, which are common microbes that cause cosmetic spoilage. IK-16-1 has a higher antibacterial activity than that of antifungal. IK-16-1 shows good safety, has no cytotoxic to keratinocytes and no haemolytic activity. Notably, the number of preservatives added to cosmetics in the market may exceed the actual requirements. As shown in Table 3, only 30% of the preservatives were used in this study; however, the anticorrosion effect was comparable to that of the
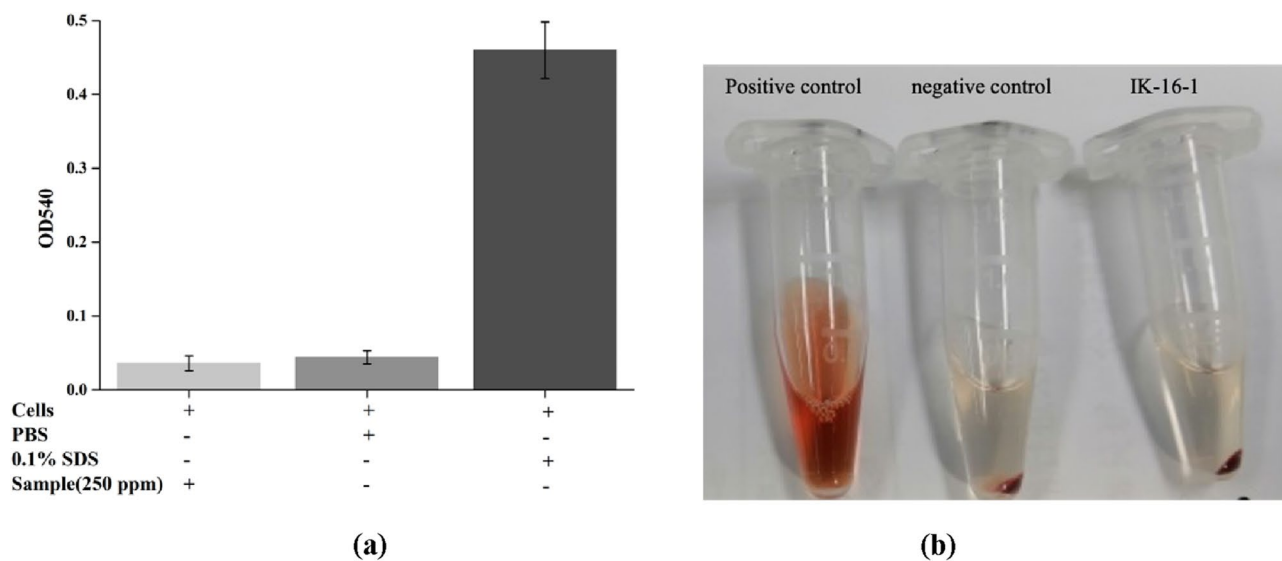
Figure 3. The hemolytic results of IK-16-1 at a concentration of 250 µg/mL (positive control: 0.1% SDS, negative control: PBS).

Species	Day 0 (CFU/mL)	Day3 (CFU/mL)	Day 7 (CFU/mL)	Day 14 (CFU/mL)	Day 22 (CFU/mL)
Escherichia coli	< 10	< 10	< 10	< 10	< 10
Staphylococcus aureus	< 10	< 10	< 10	< 10	< 10
Pseudomonas aeruginosa	< 10	< 10	< 10	< 10	< 10
Candida albicans	1.95 × 10	< 10	< 10	< 10	< 10

Table 3. The colony forming unit (CFU) calculation in cosmetic products at different time.

normal formulation. The results showed that IK-16-1 effectively killed the microbes in cosmetic products at low concentrations. Therefore, IK-16-1 can be used as a synergistic preservative to reduce the use of preservatives in cosmetics.

AMPs are characterised by both hydrophobic and hydrophilic domains. Most are cationic, and their positive net charge allows them to interact with negatively charged bacterial membranes[44]. Gram-negative and Gram-positive bacteria have molecules on their outer membranes that confer a negative net charge, allowing electrostatic interactions with cationic peptides[45]. AMPs can cause bacterial cell death through both membranolytic and non-membranolytic mechanisms by interacting with intracellular targets such as DNA, RNA, and proteins[19,46–52]. Most AMPs have broad-spectrum antimicrobial activity and are less prone to trigger resistance, which has a good prospect. In this study, peptides with helical structures were adjusted to separate hydrophobic and hydrophilic residues on the surface. Theoretically, the ability of AMPs to form pores could be improved.

ML uses mathematical, statistical, and computational processes to learn from data[53]. Developing new AMPs through ML is an emerging topic. Lee et al. developed a support vector machine (SVM) based classifier to investigate α-helical AMPs and the interrelated nature of their functional commonality and sequence homology[54]. Capecchi et al. trained recurrent neural networks (RNN) using data from the Database of Antimicrobial Activity and Structure of Peptides to design short non-haemolytic AMPs[55]. A recent study used in silico methods to investigate potential AMPs against predatory myxobacteria[56]. The sources of AMPs are diverse. In this study, we designed AMPs based on endogenous human proteins. Leon et al. used computer-aided tools to identify multi-functional peptides with antimicrobial, antibiofilm, and antioxidant potential from plant peptides[57]. Tachapuri-punya et al. used biochemical, mass spectrometric, and bioinformatics approaches to comprehensively identify putative bioactive peptides from the mucus proteomes of gastropods[58]. Besides,Ng et al. reported biosurfactants also show antimicrobial properties[59], which can be considered for cosmetic applications.

Due to the lack of fundamental knowledge of antibacterial mechanisms, the characteristics generated by different ML programs cannot fully characterise the properties of AMPs. However, with the rapid development of screening methods[60] and the continuous updating of calculation algorithms[61], challenges in discovering, classifying, and designing AMPs will gradually be solved to provide more alternatives for preventing accidental overgrowth and drug resistance of microorganisms. Notably, wet experiments are the only way to determine the antimicrobial activity of peptides. Integrating ML with targeted experimentation can guide both AMP discovery and design and provide a new understanding of the properties and mechanisms underlying their modes of action[62].

This study has some limitations. The exact mechanism by which IK-16-1 combines with preservatives to inhibit bacterial growth was not elucidated, and the mechanism through which IK-16-1 interacts with bacteria and other compounds in emulsions remains unclear. Furthermore, whether IK-16-1 has a synergistic effect with other preservatives, such as phenoxyethanol, requires further investigation. Because the performance of IK-16-1 was tested only under emulsion conditions to form its theoretical helical structure, the effects of other dosage forms should be considered in future applications. In future research, we intend to investigate the synergistic effects of IK-16-1 in relation to other preservatives.

Notwithstanding these limitations, this study verified the feasibility of combining computational design with artificial intelligence prediction to design AMPs derived from human endogenous proteins. IK-16-1 inhibits the growth of common spoilage-causing bacteria. Therefore, it can be used synergistically to effectively preserve cosmetics while containing lower concentrations of allergenic preservatives.

Conclusion

Preservatives, essential ingredients in cosmetic products, are one of the most common allergens. The irritation caused by preservatives has underscore an urgent need for the development of safe and effective broad-spectrum antimicrobial agents. Using machine learning-assisted rational design, we obtained a novel antibacterial peptide IK-16-1 with significant antibacterial activity and maintaining safety based on β-defensins, which could be a promising preservative synergist to reduce the use of allergenic preservatives in cosmetics. This study demonstrated that ML can significantly improve antibacterial activity and safety, shorten development cycles, and reduce development costs.

Data availability

References

1. Ratajczak, P. et al. The growing market for natural cosmetics in Poland: Consumer preferences and industry trends. *Clin. Cosmet. Investig. Dermatol.* **16**, 1877–1892 (2023).
2. Glaz, P. et al. Effect of commonly used cosmetic preservatives on healthy human skin cells. *Cells* **12**, 1076 (2023).
3. Halla, N. et al. Cosmetics preservation: A review on present strategies. *Molecules* **23**, 1571 (2018).
4. Bruusgaard-Mouritsen, M. A., Garvey, L. H. & Johansen, J. D. Facial contact dermatitis caused by cosmetic-relevant allergens. *Contact Dermat.* **85**, 650–659 (2021).
5. Uter, W., Yazar, K., Kratz, E. M., Mildau, G. & Lidén, C. Coupled exposure to ingredients of cosmetic products: II. Preservatives. *Contact Dermat.* **70**, 219–226 (2014).
6. Burnett, C. L. et al. Amended safety assessment of methylisothiazolinone as used in cosmetics. *Int. J. Toxicol.* **40**, 5S-19S (2021).
7. Qin, O. et al. Patch test in Chinese in Shanghai with cosmetic allergy to cosmetic series and products. *J. Cosmet. Dermatol.* **19**, 2086–2092 (2020).

8. Nowak, K., Jabłońska, E. & Ratajczak-Wrona, W. Controversy around parabens: Alternative strategies for preservative use in cosmetics and personal care products. *Environ. Res.* **198**, 110488 (2021).

9. Goossens, A. & Aerts, O. Contact allergy to and allergic contact dermatitis from formaldehyde and formaldehyde releasers: A clinical review and update. *Contact Dermat.* **87**, 20–27 (2022).

10. Zhang, W. *et al.* Effect of leave-on cosmetic antimicrobial preservatives on healthy skin resident Staphylococcus epidermidis. *J. Cosmet. Dermatol.* **22**, 2115–2121 (2023).

11. Lundov, M. D., Johansen, J. D., Zachariae, C. & Moesby, L. Low-level efficacy of cosmetic preservatives. *Int. J. Cosmet. Sci.* **33**, 190–196 (2011).

12. Ng, Y. J. *et al.* Recent advances and discoveries of microbial-based glycolipids: Prospective alternative for remediation activities. *Biotechnol. Adv.* **68**, 108198 (2023).

13. Meftahi, A. *et al.* Nanocelluloses as skin biocompatible materials for skincare, cosmetics, and healthcare: Formulations, regulations, and emerging applications. *Carbohydr. Polym.* **278**, 118956 (2022).

14. Moradi, N. *et al.* Synthesis of mesoporous antimicrobial herbal nanomaterial-carrier for silver nanoparticles and antimicrobial sensing. *Food Chem. Toxicol.* **165**, 113077 (2022).

15. Bindhu, M. R., Umadevi, M., Esmail, G. A., Al-Dhabi, N. A. & Arasu, M. V. Green synthesis and characterization of silver nano-particles from *Moringa oleifera* flower and assessment of antimicrobial and sensing properties. *J. Photochem. Photobiol. B* **205**, 111836 (2020).

16. Bahar, A. A. & Ren, D. Antimicrobial peptides. *Pharmaceuticals (Basel)* **6**, 1543–1575 (2013).

17. Wu, Q. H., Patočka, J. & Kuča, K. Insect antimicrobial peptides, a mini review. *Toxins* **10**, 461 (2018).

18. Zasloff, M. Antimicrobial peptides of multicellular organisms. *Nature* **415**, 389–395 (2002).

19. Di Somma, A., Moretta, A., Canè, C., Cirillo, A. & Duilio, A. Antimicrobial and antibiofilm peptides. *Biomolecules* **10**, 652 (2020).

20. Powers, J. P. & Hancock, R. E. The relationship between peptide structure and antibacterial activity. *Peptides* **24**, 1681–1691 (2003).

21. Bulet, P., Stöcklin, R. & Menin, L. Anti-microbial peptides: From invertebrates to vertebrates. *Immunol. Rev.* **198**, 169–184 (2004).

22. Zhang, L. Interaction of human beta defensin type 3 (hBD-3) with different PIP2-containing membranes, a molecular dynamics simulation study. *J. Chem. Inf. Model.* **61**, 4670–4686 (2021).

23. Janiesch, C., Zschech, P. & Heinrich, K. Machine learning and deep learning. *Electron. Mark.* **31**, 685–695 (2021).

24. Biamonte, J. *et al.* Quantum machine learning. *Nature* **549**, 195–202 (2017).

25. Crampon, K., Giorkallos, A., Deldossi, M., Baud, S. & Steffenel, L. A. Machine-learning methods for ligand-protein molecular docking. *Drug Discov. Today* **27**, 151–164 (2022).

26. Greener, J. G., Kandathil, S. M., Moffat, L. & Jones, D. T. A guide to machine learning for biologists. *Nat. Rev. Mol. Cell Biol.* **23**, 40–55 (2022).

27. Kang, J., Ullah, Z., Gwak, J. & Brain, M.-B. Tumor classification using ensemble of deep features and machine learning classifiers. *Sensors* **21**, 2222 (2021).

28. Dehkharghanian, T., Mu, Y. Q., Tizhoosh, H. R. & Campbell, C. J. V. Applied machine learning in hematopathology. *Int. J. Lab. Hematol.* **45**(Supplement 2), 87–94 (2023).

29. An, Q., Rahman, S., Zhou, J. W. & Kang, J. J. A comprehensive review on machine learning in healthcare industry: Classification, restrictions, opportunities and challenges. *Sensors (Basel)* **23**, 4178 (2023).

30. Ahmad Sobri, M. Z. *et al.* Kinetic model derived from machine learning for accurate prediction of microalgal hydrogen production via conversion from low thermally pre-treated palm kernel expeller waste. *Chemosphere* **338**, 139526 (2023).

31. Su, D. Q. *et al.* Application of machine learning in fuel cell research. *Energies* **16**, 4390 (2023).

32. Wang, Z. *et al.* The role of machine learning to boost the bioenergy and biofuels conversion. *Bioresour. Technol.* **343**, 126099 (2022).

33. Torres, M. D. T., Sothiselvam, S., Lu, T. K. & de la Fuente-Nunez, C. Peptide design principles for antimicrobial applications. *J. Mol. Biol.* **431**, 3547–3567 (2019).

34. Baindara, P., Korpole, S. & Grover, V. Bacteriocins: Perspective for the development of novel anticancer drugs. *Appl. Microbiol. Biotechnol.* **102**, 10393–10408 (2018).

35. Wang, G. The antimicrobial peptide database provides a platform for decoding the design principles of naturally occurring antimicrobial peptides. *Protein Sci.* **29**, 8–18 (2020).

36. Yang, M., Zhang, C., Zhang, M. Z. & Zhang, S. Beta-defensin derived cationic antimicrobial peptides with potent killing activity against gram negative and gram positive bacteria. *BMC Microbiol.* **18**, 54 (2018).

37. Gunasekera, S., Muhammad, T., Strömstedt, A. A., Rosengren, K. J. & Göransson, U. Alanine and lysine scans of the LL-37-Derived peptide fragment KR-12 reveal key residues for antimicrobial activity. *ChemBioChem* **19**, 931–939 (2018).

38. Esfandiyari, R. *et al.* Performance evaluation of antimicrobial peptide ll-37 and hepcidin and beta-defensin-2 secreted by mesenchymal stem cells. *Heliyon* **5**, e02652 (2019).

39. Veltri, D., Kamath, U. & Shehu, A. Deep learning improves antimicrobial peptide recognition. *Bioinformatics* **34**, 2740–2747 (2018).

40. Yan, J. *et al.* Deep-AmPEP30: Improve short antimicrobial peptides prediction with deep learning. *Mol. Ther. Nucleic Acids* **20**, 882–894 (2020).

41. Xiao, X., Wang, P., Lin, W. Z., Jia, J. H. & Chou, K. C. iAMP-2L: A two-level multi-label classifier for identifying antimicrobial peptides and their functional types. *Anal. Biochem.* **436**, 168–177 (2013).

42. Thomas, S., Karnik, S., Barai, R. S., Jayaraman, V. K. & Idicula-Thomas, S. CAMP: A useful resource for research on antimicrobial peptides. *Nucleic Acids Res.* **38**, D774–D780 (2010).

43. Plisson, F., Ramírez-Sánchez, O. & Martínez-Hernández, C. Machine learning-guided discovery and design of non-hemolytic peptides. *Sci. Rep.* **10**, 16581 (2020).

44. Wieprecht, T., Apostolov, O. & Seelig, J. Binding of the antibacterial peptide magainin 2 amide to small and large unilamellar vesicles. *Biophys. Chem.* **85**, 187–198 (2000).

45. Pasupuleti, M., Schmidtchen, A. & Malmsten, M. Antimicrobial peptides: Key components of the innate immune system. *Crit. Rev. Biotechnol.* **32**, 143–171 (2012).

46. Haney, E. F., Straus, S. K. & Hancock, R. E. W. Reassessing the host defense peptide landscape. *Front. Chem.* **7**, 43 (2019).

47. Mankoci, S. *et al.* Bacterial membrane selective antimicrobial peptide-mimetic polyurethanes: Structure-property correlations and mechanisms of action. *Biomacromolecules* **20**, 4096–4106 (2019).

48. Alghalayini, A., Garcia, A., Berry, T. & Cranfield, C. G. The use of tethered bilayer lipid membranes to identify the mechanisms of antimicrobial peptide interactions with lipid bilayers. *Antibiotics (Basel)* **8**, 12 (2019).

49. Brogden, K. A. Antimicrobial peptides: Pore formers or metabolic inhibitors in bacteria?. *Nat. Rev. Microbiol.* **3**, 238–250 (2005).

50. Cassone, M., Frith, N., Vogiatzi, P., Wade, J. D. & Otvos, L. Induced resistance to the designer proline-rich antimicrobial peptide A3-APO does not involve changes in the intracellular target DnaK. *Int. J. Pept. Res. Ther.* **15**, 121–128 (2009).

51. Shah, P., Hsiao, F. S., Ho, Y. H. & Chen, C. S. The proteome targets of intracellular targeting antimicrobial peptides. *Proteomics* **16**, 1225–1237 (2016).

52. Graf, M. & Wilson, D. N. Intracellular antimicrobial peptides targeting the protein synthesis machinery. *Adv. Exp. Med. Biol.* **1117**, 73–89 (2019).

53. Lee, E. Y., Lee, M. W., Fulan, B. M., Ferguson, A. L. & Wong, G. C. L. What can machine learning do for antimicrobial peptides, and what can antimicrobial peptides do for machine learning?. *Interface Focus* **7**, 20160153 (2017).

54. Lee, E. Y., Fulan, B. M., Wong, G. C. & Ferguson, A. L. Mapping membrane activity in undiscovered peptide sequence space using machine learning. *Proc. Natl Acad. Sci. USA* **113**, 13588–13593 (2016).
55. Capecchi, A. *et al.* Machine learning designs non-hemolytic antimicrobial peptides. *Chem. Sci.* **12**, 9221–9232 (2021).
56. Arakal, B. S. *et al.* In silico and in vitro analyses reveal promising antimicrobial peptides from myxobacteria. *Probiotics Antimicrob. Proteins* **15**, 202–214 (2023).
57. León Madrazo, A. & Segura Campos, M. R. In silico prediction of peptide variants from chia (*S. hispanica* L.) with antimicrobial, antibiofilm, and antioxidant potential. *Comput. Biol. Chem.* **98**, 107695 (2022).
58. Tachapuripunya, V., Roytrakul, S., Chumnanpuen, P. & E-Kobon, T. Unveiling putative functions of mucus proteins and their tryptic peptides in seven gastropod species using comparative proteomics and machine learning-based bioinformatics predictions. *Molecules* **26**, 3475 (2021).
59. Ng, Y. J. *et al.* Recent advances of biosurfactant for waste and pollution bioremediation: Substitutions of petroleum-based surfactants. *Environ. Res.* **212**, 113126 (2022).
60. Tucker, A. T. *et al.* Discovery of next-generation antimicrobials through bacterial self-screening of surface-displayed peptide libraries. *Cell* **172**, 618-628.e13 e613 (2018).
61. Das, P. *et al.* Accelerated antimicrobial discovery via deep generative models and molecular dynamics simulations. *Nat. Biomed. Eng.* **5**, 613–623 (2021).
62. Lee, E. Y., Wong, G. C. L. & Ferguson, A. L. Machine learning-enabled discovery and design of membrane-active peptides. *Bioorg. Med. Chem.* **26**, 2708–2718 (2018).

Acknowledgements

Author contributions

L.Y. wrote the main manuscript text. L.S. and X.F. conducted antibacterial experiments. S.Z. designed artificial AMPs. X.L. reviewed the manuscript. C.H. proposed the framework for this study. J.L. supervised and guided the study. All the authors contributed to the manuscript and approved the submitted version.

Competing interests

The authors declare no competing interests.

Additional information

Correspondence and requests for materials should be addressed to C.H. or J.L.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.