Antibacterial peptides, now discovered with artificial intelligence!

- Up to 13% improvement in antibacterial effect prediction accuracy, contributing to the reduction of antibiotic development time

- Professor Hojung Nam's team published a thesis in 「Protein Science」, a renowned academic journal in the field of protein

▲ (From left) Professor Hojung Nam and integrated student Hansol Lee

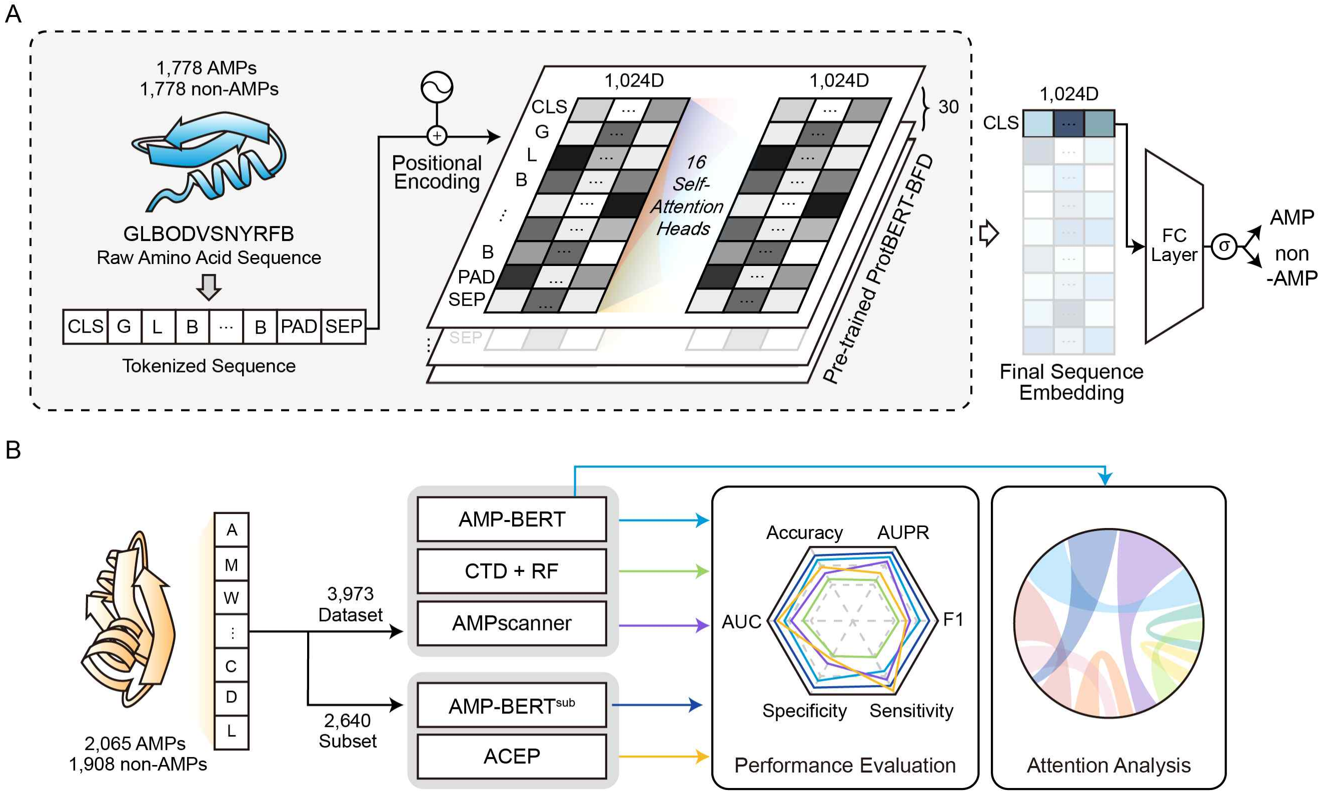
A research team at GIST (Gwangju Institute of Science and Technology, President Kiseon Kim) developed an artificial intelligence model that quickly discovers peptide molecules with antibacterial functions.

As an artificial intelligence platform for drug development to prepare for the upcoming era of antibiotic-resistant bacteria pandemic and is expected to contribute to shortening the development time of new drugs by being used in the early stage of developing peptide molecule-based antibiotics.

Penicillin, the first antibiotic, opened a new chapter in treating various diseases in humans and animals, but indiscriminate misuse of antibiotics led to the emergence of antibiotic-resistant bacteria. Accordingly, research on antimicrobial peptide (AMP), a new antibiotic platform, is drawing attention.

Antimicrobial peptides (AMPs) are small proteins that occur naturally in all areas of life. It has excellent antibacterial activity even against bacteria resistant to existing antibiotics and hardly induces the emergence of new resistant bacteria, so it is in the limelight as a next-generation antibiotic. In addition, it is attracting attention as an effective alternative treatment for patients infected with bacteria, viruses, and fungi or cancer patients.

In general, the development of new antibacterial drugs takes a lot of time and money because the activity is verified by directly screening the antibacterial agent against microorganisms cultured in the laboratory. In order to solve this problem, research on computational models using computers has been actively conducted recently.



▲ Overview of the AMP-BERT study. It comprehensively shows the data used in the study, the detailed structure of the model, and the evaluation and analysis methods.

School of Electrical Engineering and Computer Science Professor Hojung Nam's research team (integrated student Hansol Lee, Ph.D. student Songyeon Lee, and Dr. Ingoo Lee) has developed an artificial intelligence prediction technology that can more accurately determine the antibacterial effect of peptide molecules to develop antibacterial peptide treatments (AMP-BERT: Prediction of Antimicrobial Peptide Function Based on a BERT Model).

This study showed a minimum of 2% and a maximum of 13% higher antimicrobial prediction accuracy depending on the experimental conditions compared to various prediction models published so far. In addition, the interpretability of the prediction results was added by providing information on the partial sequence of major peptides associated with antimicrobial activity, which was not possible in previous studies.

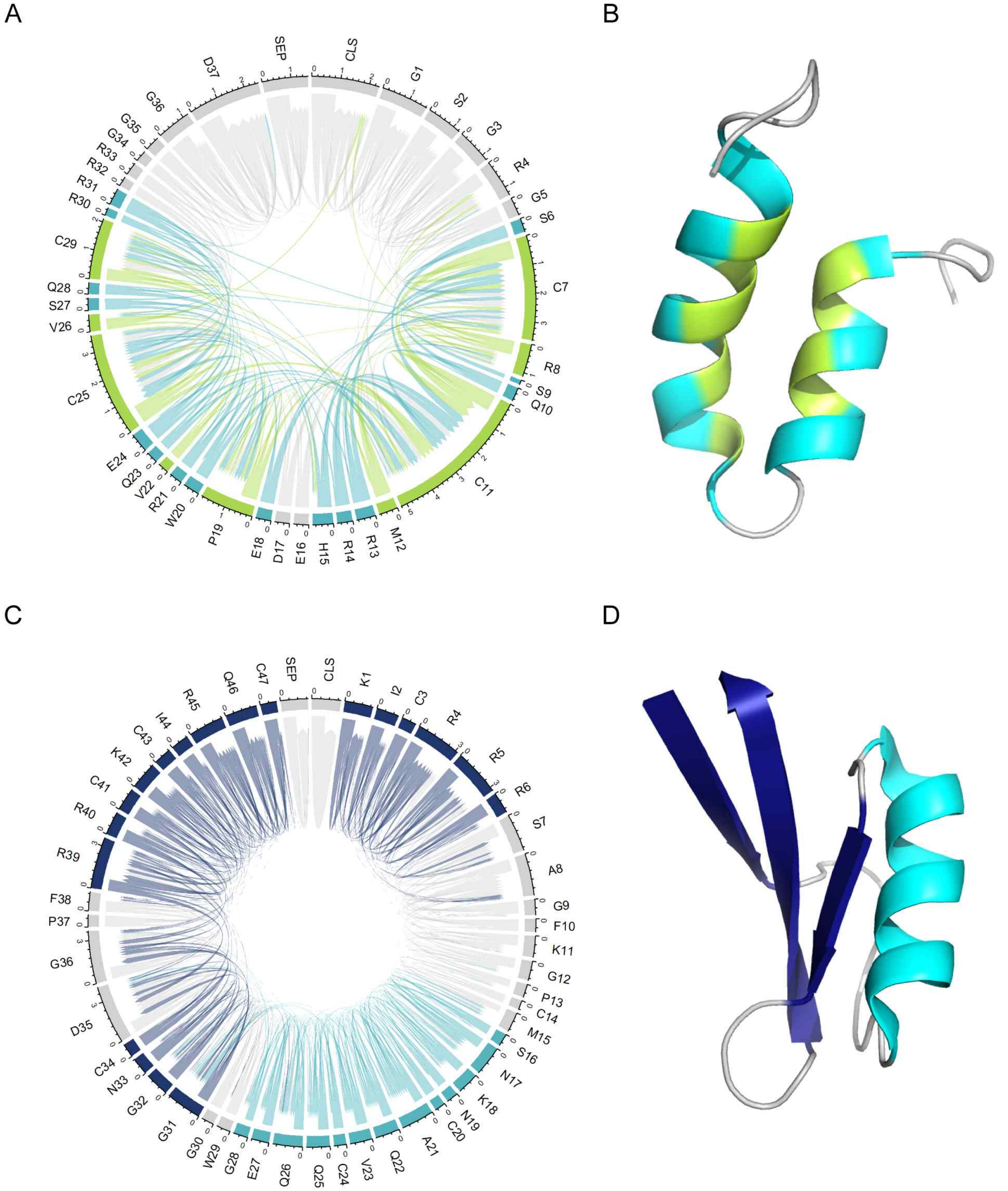
The research team developed an antibacterial peptide activity prediction model that complements the limitations of previous studies by applying self-attention technology* to a BERT neural network model* pretrained with large-scale protein sequences.

* BERT (Bidirectional Encoder Representation from Transformers) neural network model: A natural language processing model that pre-trains with a large amount of data and fine-tunes it with data for a specific task by adding a neural network

* Self-attention technology: This is a process of learning each word in an input sentence and learning the relative properties of the current word in the sentence by referring to the context of words in other positions.

In particular, as a result of fine-tuning a model pre-trained with large-capacity protein data, which has not been used in previous antimicrobial peptide activity prediction studies, with antimicrobial peptide-related sequence data, it showed better prediction accuracy than previous related studies.

In addition, attention technology, which detects key features of protein sequences, was used to find trends in amino acid sequences associated with antibacterial effects. Through this, the reliability and interpretability of the model were verified by confirming that the proposed model correctly detects the major substructures related to the antimicrobial peptide activity.



▲ Prediction results and interpretation results of the model. It can be confirmed that the model not only accurately predicts but also emphasizes and predicts amino acid partial structures that actually cause antibacterial activity.

Professor Hojung Nam said, "This achievement is a technology that trains the general characteristics of protein sequences and specific information of antimicrobial peptides together to increase the accuracy of predicting the antibacterial effect of peptide molecules and provides key peptide partial sequences involved in antibacterial effects. We hope that artificial intelligence models will contribute to discovering various antibiotic candidates in a short period of time and consequently increasing the possibility of success in developing new antibiotic drugs."

This study was carried out with the support of the National Research Foundation of Korea's 'mid-level researcher support project' and was published online on December 3, 2022, in 「Protein Science」, a renowned academic journal in the field of protein.

GIST
Since 1993