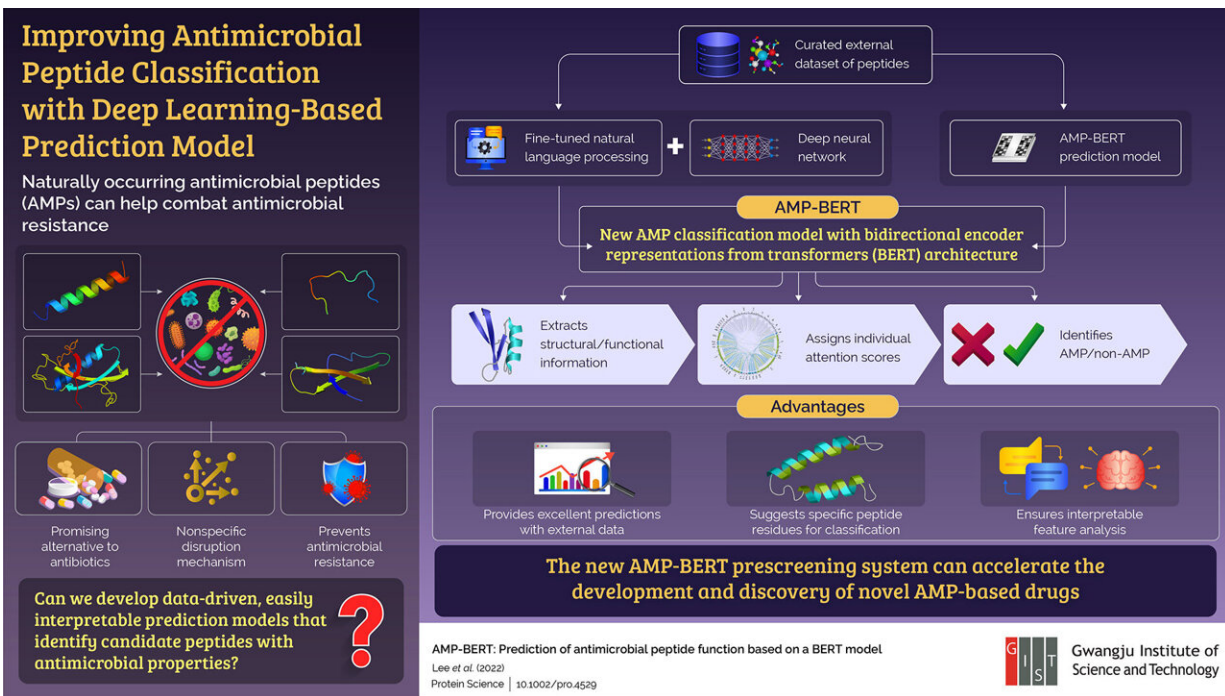# Researchers develop a new AI-based 'finder' of antimicrobial peptides

February 23 2023

Researchers from GIST developed natural language processing (NLP) and machine learning (ML)-based antimicrobial peptide (AMP) prediction model that can identify functional/structural properties of peptides and classify them as AMPs and non-AMPs. Credit: Hojung Nam from Gwangju Institute of Science and Technology

Over the last few decades, antimicrobial resistance has become a major public health concern globally. This has led to a search for alternative

methods of treating microbial infections.

One such innovation is the discovery of antimicrobial properties of certain peptides. Antimicrobial peptides (AMPs) are short peptides found in most animals, plants, and microorganisms as a natural defense against infections. AMPs combat harmful bacteria via a nonspecific mechanism that prevents them from developing antimicrobial resistance.

Despite these exceptional abilities, research on AMPs is being hindered because the existing systems for identifying candidate AMPs are like a black box, where the outputs are not easily interpretable for further analysis.

Now, in a recent breakthrough published in *Protein Science*, a team of researchers from Gwangju Institute of Science and Technology, including Prof. Hojung Nam and Mr. Hansol Lee, proposed an AMP-BERT classification system that uses AI-based bidirectional encoder representation from transformers (BERT) architecture to improve upon the existing AMP classification models.

When asked about the motivation behind developing the classification system, Prof. Nam explains, "The misuse and overuse of antibiotics have resulted in the development of bacteria that cannot be effectively treated with these antibiotics. This has resulted in an increased health risk not only in humans but also agriculture. So, we wanted to develop an AMP pre-screening platform that isn't a black box of algorithms but can be easily interpreted for further research."

The team incorporated a natural language processing (NLP)-based deep neural network that was pre-trained with billions of protein sequences then fine-tuned with thousands of peptide sequences from a benchmark AMP database. This enabled the AMP-BERT model to not only extract the structural and functional information from the input peptide

sequences but also differentiate AMPs from non-AMPs. This enhanced the prediction power allowed the model to make better classifications even with external data.

The team also designed the model to assign individual attention scores to each amino acid from the input peptide sequence. The attention feature then revealed the important subregions of AMPs that play a major role in deciding whether a peptide has antimicrobial properties or not. Furthermore, the prediction results indicated the AMP-BERT model's applicability extends even to unseen peptide data and that it can learn meaningful functional and structural information from those peptides.

The novel AMP-BERT peptide pre-screening model can open new doors for the discovery and development of AMP-based drug candidates for treating antimicrobial-resistant illnesses. The important peptide subregion information provided by this prediction platform can also be used to optimize the antibiotic efficiency of peptides.

"As more AMPs are experimentally validated and new structural information is uncovered using computational methods, we will be able to make more effective antibiotic drugs and potentially stop a new pandemic from spreading across the world in near future," concludes Prof Nam.

Citation: Researchers develop a new AI-based 'finder' of antimicrobial peptides (2023, February