

Title: AMP-RNNpro: a two-stage approach for identification of antimicrobials using probabilistic features

Author Names: Md. Shazzad Hossain Shaon, Tasmin Karim, Md. Fahim Sultan, Md. Mamun Ali, Kawsar Ahmed, Md. Zahid Hasan, Ahmed Moustafa, Francis M. Bui & Fahad Ahmed Al-Zahrani

Abstract: Antimicrobials are molecules that prevent the formation of microorganisms such as bacteria, viruses, fungi, and parasites. The necessity to detect antimicrobial peptides (AMPs) using machine learning and deep learning arises from the need for efficiency to accelerate the discovery of AMPs, and contribute to developing effective antimicrobial therapies, especially in the face of increasing antibiotic resistance. This study introduced AMP-RNNpro based on Recurrent Neural Network (RNN), an innovative model for detecting AMPs, which was designed with eight feature encoding methods that are selected according to four criteria: amino acid compositional, grouped amino acid compositional, autocorrelation, and pseudo-amino acid compositional to represent the protein sequences for efficient identification of AMPs. In our framework, two-stage predictions have been conducted. Initially, this study analyzed 33 models on these feature extractions. Then, we selected the best six models from these models using rigorous performance metrics. In the second stage, probabilistic features have been generated from the selected six models in each feature encoding and they are aggregated to be fed into our final meta-model called AMP-RNNpro. This study also introduced 20 features with SHAP, which are crucial in the drug development fields, where we discover AAC, ASDC, and CKSAAGP features are highly impactful for detection and drug discovery. Our proposed framework, AMP-RNNpro excels in the identification of novel Amps with 97.15% accuracy, 96.48% sensitivity, and 97.87% specificity. We built a user-friendly website for demonstrating the accurate prediction of AMPs based on the proposed approach

Keywords: Antimicrobial peptides (AMPs) Recurrent Neural Network (RNN) Deep learning–based peptide classification Feature encoding techniques Explainable AI (SHAP)

<https://www.nature.com/articles/s41598-024-63461-6>