Trends in Peptide and Protein Sciences
Volume 11 (2026): e1

DOI: https://doi.org/10.22037/tpps.v11i1.53063

Review Article

Next-Generation Antimicrobial Peptides: The Emerging Role of Artificial Intelligence in Addressing Antimicrobial Resistance

Maryam Tabarzad

Protein Technology Research Center, Shahid Beheshti University of Medical Sciences, Tehran, Iran.

Article history:
Received: 22 June 2026
Accepted: 26 August 2026

Keywords:
Antimicrobial peptides
Artificial intelligence
Antimicrobial resistance
Generative AI
Drug discovery

Maryam Tabarzad: https://orcid.org/0000-0001-9923-0639

**Corresponding Author:*
Email: m_tabarzad@sbmu.ac.ir (M. Tabarzad)

HIGHLIGHTS

- Artificial intelligence (AI) accelerates discovery of antimicrobial peptides (AMPs).
- Three key AI approaches in this field are data mining, functional prediction, and generative models.
- The fusion of AI and peptide bioengineering represented a significant shift in AMPs discovery.

ABSTRACT

The rising issue of antimicrobial resistance (AMR) is prompting an urgent need for new treatments that go beyond conventional antibiotics. Antimicrobial peptides (AMPs) are emerging as viable options because of their varied actions and lower tendency to develop resistance. This review highlights how artificial intelligence (AI) can play a transformative role in speeding up the discovery, enhancement, and design of these peptides. The current AI methodologies can be classified into three main areas: (i) analyzing genomic and metagenomic data with machine learning tools like the Metagenomic Antimicrobial Peptide Predictor (MACREL) and AMPSphere; (ii) employing protein language models and advanced deep learning techniques, particularly Convolutional Neural Network-Long Short-Term Memory (CNN-LSTM) networks and Transformers, for large-scale functional predictions; and (iii) using generative models such as Generative Adversarial Networks (GANs), Variational Autoencoders (VAEs), and diffusion frameworks to design new peptide candidates in a controlled manner. Challenges like data heterogeneity, inconsistent benchmarking that can lead to exaggerated performance claims, and the gap between computational predictions and their practical applicability in clinical settings were also discussed. Additionally, the emerging strategies that aim to balance enhancing antimicrobial effectiveness with reducing toxicity to the host were discussed. The review tried to draw a forward-looking roadmap proposing a closed-loop, active-learning system that integrates standardized datasets and rigorous laboratory validation. Overall, the fusion of AI and peptide bioengineering represented a significant shift, offering exciting possibilities for developing next-generation AMPs aimed at combating multidrug-resistant pathogens.

Cite this article as: Tabarzad, M. (2026). Next-Generation Antimicrobial Peptides: The Emerging Role of Artificial Intelligence in Addressing Antimicrobial Resistance. *Trends Pept. Protein Sci.*, **11**: e1.

This open-access article is distributed under the terms of the Creative Commons Attribution Non Commercial 4.0 License (CC BY-NC 4.0).

Introduction

Antimicrobial resistance (AMR) poses a serious threat to global health, causing around 700,000 deaths each year. If we don't take action soon, that number could rise to ten million by 2050. In the U.S., infections that are resistant to multiple drugs account for over 20% of hospital admissions annually, with more than 80% of these cases starting in community settings. This issue also has serious economic implications, especially for low- and middle-income countries (Fong, 2023). AMR is a serious global health challenge, leading the World Health Organization (WHO) to identify priority pathogens that require new antibiotics. The development of these antibiotics is a lengthy and costly process, with only 18 novel antibiotics approved since 2014. Recent advancements in artificial intelligence (AI) have significantly enhanced the efficiency of antibiotic discovery (Liu et al., 2024).

Between 2014 and 2021, a number of antibiotics were approved for various resistant pathogens. The rise of antibiotic resistance, particularly among the ESKAPE pathogens (*Enterococcus faecium*, *Staphylococcus aureus*, *Klebsiella pneumoniae*, *Acinetobacter baumannii*, *Pseudomonas aeruginosa*, and *Enterobacter* spp.), poses a significant threat to public health as they remain critical multidrug-resistant bacteria (Miller and Arias, 2024). On May 17, 2024, the WHO updated its Bacterial Priority Pathogens List, categorizing 24 drug-resistant bacteria into three priority groups: critical, high, and medium. The critical priority pathogens include *A. baumannii* and drug-resistant *Mycobacterium tuberculosis*. High-priority pathogens encompass *Salmonella* spp, *Shigella* spp, *P. aeruginosa*, *S. aureus*, *Neisseria gonorrhoeae*, and *Enterococcus faecium*. Medium-priority pathogens include Group A and B *streptococci*, *Streptococcus pneumoniae*, and *Haemophilus influenzae*. This categorization aims to guide research, development, and public health strategies against AMR (Jesudason, 2024).

In 2014, dalbavancin, tedizolid, and oritavancin, all targeting Gram-positive pathogens, were approved, along with ceftolozane/tazobactam, aimed at β-lactamase enzyme-producing bacteria. The following year, 2015, saw the approval of cefazidime/avibactam for carbapenem-resistant Enterobacteriaceae (CRE) and isavuconazonium as an antifungal triazole. In 2017, delafloxacin, a fluoroquinolone for Gram-positive pathogens, and vaborbactam/meropenem for CRE received approval. The year 2018 introduced plazomicin, eravacycline, and omadacycline, all targeting CRE and methicillin-resistant *Staphylococcus aureus* (MRSA). In 2019, a novel combination of relebactam with imipenem/cilastatin targeting CRE and showing potential against carbapenem-resistant *P. aeruginosa* (CRPA) was approved, alongside lefamulin for MRSA, pretomanid for extensively drug-resistant tuberculosis (XDR-TB), and cefiderocol, which targets carbapenem-resistant *A. baumannii* (CRAB), CRPA, and CRE. The year 2020 saw the approval of levonadifloxacin, a fluoroquinolone for Gram-positive pathogens, while contezolid, an oxazolidinone for MRSA, was approved in 2021 (Liu et al., 2024).

Antimicrobial peptides (AMPs) are a promising alternative to traditional antibiotics in addressing antibiotic resistance. These small peptides are part of the innate immune system and are effective against a wide range of pathogens, including drug-resistant bacteria like the ESKAPE pathogens. AMPs target multiple sites in microbes, reducing the likelihood of resistance and maintaining efficacy at low doses. Overall, AMPs offer an innovative approach to combating antibiotic resistance (Girdhar et al., 2024).

In fact, AMPs are evolutionarily significant chemicals found in both prokaryotes and humans. The first human antimicrobial protein, lysozyme, was discovered by Alexander Fleming in nasal mucus in 1922. However, this finding was overshadowed by his 1928 discovery of penicillin, which led to its medical applications and earned him the 1945 Nobel Prize in Medicine. As a result, the therapeutic potential of AMPs like lysozyme was largely overlooked during the "Golden Age of Antibiotics" in the 1940s. The emergence of multidrug-resistant infections in the 1960s sparked interest in AMPs as potential defense agents. AMPs are produced by all known living organisms, with early discoveries dating back to the 1960s, such as lactoferrin from milk and brombinin from frogs. A groundbreaking study in 1981 showed that injecting bacteria into Cecropia moths led to the production of strong AMPs, and by 1987, scientists had isolated

This open-access article is distributed under the terms of the Creative Commons Attribution Non Commercial 4.0 License (CC BY-NC 4.0).

cationic AMPs from the African clawed frog, *Xenopus laevis*. The vital role of AMPs in insect immunity was highlighted in the mid-1990s when research indicated that fruit flies lacking AMP synthesis were highly susceptible to severe fungal infections. Presently, the AMP Database catalogs over 3,569 AMPs from six different kingdoms of life (https://aps.unmc.edu/home) (Hetta et al., 2024).

AMPs combat multidrug-resistant bacteria through diverse mechanisms, though bacteria have evolved counter-strategies. Understanding these resistance pathways has guided rational design approaches, including amino acid modifications, nanoparticle-based delivery, and cell-penetrating peptide conjugates. Synergistic combinations with conventional antibiotics and antibiotic-peptide conjugates further enhance efficacy while mitigating toxicity (Hetta et al., 2024). Representative examples demonstrating this potential are Cbf-K16 synergized with β-lactams against MRSA (Li et al., 2016) and Ib-AMP4 that also synergized with bacitracin E50-52 against MRSA (Satei et al., 2021).

Antimicrobial peptides: Opportunities and remaining challenges

AMPs are typically small and positively charged molecules. Their effectiveness on infectious pathogens is influenced not only by their charge but also by their structure, with most being cationic and some anionic. They are predominantly categorized into either α-helical or β-sheet structural families. Notable examples of the α-helical group include cryptin-4, human alpha-defensins (HD-5 and 6), magainins 1 and 2, melittin, and muricin; conversely, the β-sheet category encompasses peptides such as human beta-defensins (hBD-1–6), lactoferricin B, protegrin-1, and tachyplesin I (Fathi et al., 2024). A subset of AMPs adopts a hybrid structure containing both α-helical and β-sheet elements, typically stabilized by three or four disulfide bridges. This cysteine-stabilized α/β (CSαβ) structural motif was originally characterized in insect defensins and scorpion neurotoxins (Alzain et al., 2025). Unlike αβ peptides, non-αβ peptides lack defined secondary structures, characterized instead by high aqueous flexibility and enrichment in residues like tryptophan and proline. For instance, indolicidin utilized its tryptophan residues to anchor into lipid bilayers, facilitating membrane penetration and subsequent intracellular targets, such as DNA, that finally inhibited bacterial growth (Mishra et al., 2018).

However, limited numbers of AMPs have been approved by the U.S. Food and Drug Administration (FDA) for clinical use. Gramicidin D (Neocidin), Vancomycin (Vancocin), Colistin (Coly-Mycin M), Daptomycin (Cubicin), Telavancin (Vibativ), Dalbavancin (Dalvance), and Oritavancin (Kimyrsa) are those that have been approved before 2015. Gramicidin D (Neocidin) has been used topically for infected surface wounds and infections in the eye, nose, and throat, targeting bacterial membranes. Moreover, a glycopeptide like Vancomycin (Vancocin) has been administered intravenously or orally and could treat serious infections such as septicemia and Clostridioides difficile-associated diarrhea; it consists of D-alanyl-D-alanine moieties. In the group of lipopeptides, Colistin (Coly-Mycin M) and Daptomycin (Cubicin) are effective against multidrug-resistant Gram-negative bacteria and Gram-positive infections, respectively. Among lipoglycopeptides, Telavancin, Dalbavancin, and Oritavancin, have been specifically indicated for complicated skin and skin structure infections, and they contain D-alanyl-D-alanine moieties (Al Musaimi, 2025).

In developing novel AMPs, the potential strategies based on structure, such as modifying amino acids and linking them with various molecular components like small functional groups, synthetic polymers, peptides, and antibiotics, could help overcome existing barriers in the treatment of resistant infections and promote the development of effective AMP therapies (Tabarzad et al., 2026).

While AMPs represent a promising therapeutic frontier, with numerous candidates currently undergoing clinical and preclinical evaluation and several having already secured regulatory approval from the FDA and European Medicines Agency (EMA), the current AMP market remains modest. This is significantly smaller than the conventional antibiotic market. Even though over 60 therapeutic peptides have received FDA approval and more than 400 are currently in trials, only seven have reached clinical application, primarily as topical or systemic injectable treatments. Furthermore, the clinical translation of AMPs is significantly hindered by

This open-access article is distributed under the terms of the Creative Commons Attribution Non Commercial 4.0 License (CC BY-NC 4.0).

pharmacokinetic limitations, including poor gastrointestinal stability, unfavorable absorption and distribution profiles, and rapid metabolic degradation, all of which contribute to compromised bioavailability and potential cytotoxicity (Botelho Sampaio de Oliveira et al., 2023).

To address these inherent challenges, bioengineering has introduced several innovative strategies. Among the most promising approaches are the development of chimeric recombinants, which fuse AMPs with stable scaffolds like human serum albumin or antibody Fc domains, and chemical modifications, such as lipidation. A review explored the fundamental aspects of AMPs, including their origins, structures, and antimicrobial mechanisms, while also examining recent breakthroughs in bioengineering platforms aimed at enhancing their therapeutic potential. Ultimately, they concluded that by merging natural AMPs with advanced bioengineering and nanotechnology, the way for a new generation of antibiotics could be realized (Kim et al., 2025). These limitations also highlight the need for computational approaches capable of simultaneously optimizing multiple peptide properties. AI is rapidly transforming the discovery, optimization, and translational development of antimicrobial peptides, offering unprecedented opportunities to shorten development timelines while improving therapeutic performance (Brizuela et al., 2025). In general, while AMPs hold significant potential in addressing multidrug resistance, the prohibitive costs associated with traditional wet-lab screening have necessitated the integration of AI in this field. Although the field has transitioned from classical machine learning to sophisticated deep learning architectures, which successfully accelerate peptide discovery in preclinical models, comprehensive evaluations of emerging technologies remain scarce. Specifically, the transformative potential of Large Language Models (LLMs), Graph Neural Networks (GNNs), and structure-guided design methodologies has not been fully synthesized in existing literature (Brizuela et al., 2025).

Artificial intelligence in antimicrobial peptide discovery

In fact, AI has emerged as a transformative force in the landscape of AMP discovery, offering an efficient means to navigate the vast and complex sequence space. Several studies have delineated the growing role of AI, categorizing its applications into two fundamental paradigms: AMP mining and AMP generation, supplemented by the indispensable utility of discriminative modeling. While AMP mining leverages large-scale sequence scanning followed by discriminative prediction of bioactivity and toxicity to identify validated candidates, AMP generation represents a shift toward *de novo* design. By employing generative modeling, researchers can move beyond natural templates to engineer synthetic peptides optimized for superior potency and minimal toxicity. Although the challenge of maintaining sequence biological plausibility remains, generative AI holds the potential to unlock unprecedented levels of novelty and diversity in antimicrobial therapeutics (Meng, 2026, Szymczak et al., 2025).

Recent advancements demonstrate that by leveraging iterative peptide array technology and accumulating chemical biology data, it is possible to design small-molecule, broad-spectrum peptides with minimal prior assumptions. As one of the primary studies, a group of researchers in 2008 developed robust *in silico* models capable of predicting the efficacy of vast virtual peptide sets by applying Artificial Neural Networks (ANNs) to data derived from iteratively optimized 9-amino-acid libraries. Notably, these models identified high-potency peptides that outperform both standard antibiotics and leading clinical AMP candidates, demonstrating significant protective efficacy against Staphylococcus aureus in animal models (Cherkasov et al., 2008). The following investigations advanced the milestones in AI-augmented AMP discovery. Various advanced generative strategies, including discriminator-guided filtering, positive-only learning, latent space sampling, and conditional optimization, enabled the precise engineering of AMPs with tailored functional properties (Szymczak and Szczurek, 2023).

The Antimicrobial Peptide Database (APD) played a foundational role in this evolution, providing the first empirical prediction frameworks and enabling the application of early machine learning (ML) algorithms. Current ML-based predictors have evolved from simple single-label

This open-access article is distributed under the terms of the Creative Commons Attribution Non Commercial 4.0 License (CC BY-NC 4.0).

classifications of antimicrobial activity (e.g., AntiBP (Lata et al., 2010), CAMP (Thomas et al., 2010), and iAMPpred (Meher et al., 2017)) to sophisticated multi-label models (e.g., iAMP-2L (Xiao et al., 2013), MLAMP (Lin and Xu, 2016), and AMAP (Gull and Shamim, 2019). The state-of-the-art advanced models can now characterize a diverse array of functional roles, including antifungal, antiviral, antibiofilm, and anti-inflammatory activities, among others. Furthermore, modern predictive methodologies have begun to integrate complex biological parameters such as post-translational modifications, three-dimensional structural data, and microbial strain-specific information. By comparing ML-derived essential amino acids with the residue patterns of natural peptide classes, a review evaluated the current state of the field, identified existing limitations, and outlined a roadmap toward integrated predictive pipelines designed to accelerate the discovery of novel AMP-based therapeutics (Wang et al., 2022).

To expedite the discovery process, computational approaches aim to predict novel AMPs directly from genomic or proteomic sequences. As AI and ML continue to reshape various industries, they are now playing a transformative role in drug discovery. This shift is particularly critical given the rising threat of antimicrobial resistance and the inherent difficulties in developing new antibiotics. While AMPs hold great promise as therapeutic agents, their clinical adoption has been stalled by several hurdles, including toxicity, poor stability, and high production costs. Current ML methodologies aimed at overcoming existing development barriers and the limitations that must be addressed for clinical translation and highlighted the emerging opportunities within data-driven peptide design (Wan et al., 2024).

Recent advancements in AMP discovery have been significantly driven by the integration of hybrid deep learning architectures. To overcome the limitations of single-model approaches, researchers have increasingly turned to hybridizing Convolutional Neural Networks (CNNs) with Recurrent Neural Networks (RNNs), particularly Long Short-Term Memory (LSTM) layers. In these frameworks, CNNs are typically employed for automated local feature extraction and dimensionality reduction, while LSTM modules capture long-range sequential dependencies within peptide sequences, effectively mitigating the vanishing gradient problem (Yan et al. 2022). Notable examples include the ACEP model, which utilizes heterogeneous feature fusion (Fu et al. 2020), and DeepAVP, which employs a parallel CNN and Bi-LSTM architecture to achieve state-of-the-art predictive accuracy for antiviral peptides (AVPs) (Li et al., 2020).

In the following sections, these approaches will be classified and discussed in more detail. Table 1 summarizes the computational tools, databases, and artificial intelligence frameworks for antimicrobial peptide identification, functional classification, *de novo* design, and structural analysis.

AMP identification from genomes

AMP genes are difficult to identify because they evolve quickly, undergo frequent gain and loss, and encode very short peptides that homology-based methods often miss (Hanson et al., 2019). As a result, supervised machine-learning approaches that use physicochemical features have become an attractive alternative, and related tools may help improve AMP discovery from genome or transcriptome data.

The inherent challenges of detecting small, rapidly evolving AMP genes have driven the development of specialized computational pipelines. In 2020, two complementary frameworks emerged, MACREL and ampir. MACREL (https://github.com/BigDataBiology/macrel) established an end-to-end solution for mining genomes and metagenomes by engineering 22 peptide features that match state-of-the-art prediction accuracy while improving precision in identifying antimicrobial and hemolytic activities (Santos-Júnior et al., 2020).

This open-access article is distributed under the terms of the Creative Commons Attribution Non Commercial 4.0 License (CC BY-NC 4.0).

Table 1: Comprehensive overview of computational tools, databases, and artificial intelligence frameworks for antimicrobial peptide identification, functional classification, *de novo* design, and structural analysis.

Tool / Model	Primary prediction type	Functional granularity	Key strength / Distinction	Core methodology	Ref(s)
BINARY PREDICTORS (General AMP vs. Non-AMP)					
AntiBP (AntiBP2)	Binary (AMP)	Low (General)	Widely used for rapid, general-purpose antibacterial peptide screening.	Machine learning	(Lata et al., 2010)
CAMP	Binary (AMP)	Low (General)	One of the earliest foundational resources.	Machine learning	(Thomas et al., 2010)
iAMPpred	Binary (AMP)	Low (General)	Incorporates compositional and physicochemical features.	Machine learning	(Meher et al., 2017)
$CAMP_{R3}$ (RF)	Binary (AMP)	Low (General)	Most accurate general-purpose binary tool in benchmarks (highest AUC).	Random Forest (RF)	(Waghu et al., 2016)
ADAM	Binary (AMP)	Low (General)	Web-based predictor from early benchmarking studies.	Machine learning	(Lee et al., 2015)
ampir	Binary (AMP)	Low (General)	Fast, genome-wide AMP classification with high precision.	Machine learning	(Fingerhut et al., 2020)
MACREL	Binary (AMP)	Low (General)	End-to-end pipeline for genomes and metagenomes.	Machine learning	(Santos-Júnior et al., 2020)
AMPlify	Binary (AMP)	Low (General)	Validated against WHO priority pathogens (e.g., carbapenemase-producing *E. coli*).	Deep learning (Attention)	(Li et al., 2022)
Target-AMP	Binary (AMP)	Low (General)	Integrates PSSM; achieves >97% accuracy.	SVM (Hybrid)	(Jan et al., 2022)
AMP-BERT	Binary (AMP)	Low (General)	Self-attention maps specific residues to antimicrobial function.	BERT (Transformer)	(Lee et al., 2023)
Deep-ABPpred	Binary (ABP)	Low (General)	Validated via synthesis from bacteriophage tail proteins.	biLSTM + word2vec	(Sharma et al., 2021)
StaBle-ABPpred	Binary (ABP)	Low (General)	Meta-level ensemble boosts accuracy and precision.	Stacked Ensemble (biLSTM+Attention+ RF/GB)	(Singh et al., 2022)
iAMP-DL	Binary (short AMPs)	Low (Short peptides)	Optimized for rapid screening of short, high-potency peptides.	CNN + LSTM	(Nguyen et al., 2024)
AMP-Detector	Binary / Generative	Low (General)	Bridges predictive screening and therapeutic synthesis.	PLM + ML	(Medina-Ortiz et al., 2024)
CalcAMP	Binary (AMP)	Low (Short peptides)	High-accuracy prediction specifically for peptides <35 aa.	Machine learning	(Bournez et al., 2023)
ENNAACT	Binary (Anticancer)	Low (Specific class)	98.3% accuracy for anticancer peptides.	DNN	(Timmons and Hewage, 2021a)
ENNAVIA	Binary (Antiviral)	Low (Specific class)	95.7% accuracy for antiviral/coronavirus peptides.	DNN	(Timmons and Hewage, 2021b)
Deep-AntiFP	Binary (Antifungal)	Low (Specific class)	High efficacy for antifungal peptide prediction.	DNN	(Ahmad et al., 2021)
DeepAVP	Binary (Antiviral)	Low (Specific class)	State-of-the-art for antiviral peptides.	CNN + Bi-LSTM	(Li et al., 2020)
ACEP	Binary (AMP)	Low (General)	Automatic feature fusion and amino acid embedding.	CNN + LSTM	(Fu et al., 2020)
iAMP-CRA	Binary (AMP)	Low (General)	0.919 accuracy with high interpretability.	CNN + RNN + Self-Attention	(Lu et al., 2025)
MULTI-LABEL & MULTI-OBJECTIVE					
iAMP-2L	Multi-label	Moderate (2 levels)	Two-level classifier for functional types.	Machine learning	(Xiao et al., 2013)
MLAMP	Multi-label	Moderate (Multiple)	Handles imbalanced, multi-label datasets.	ML-SMOTE + Grey PseAAC	(Lin and Xu, 2016)
AMAP	Multi-label (14 functions)	Moderate (14 activities)	High generalization for novel, non-canonical sequences.	Hierarchical ML	(Gull and Shamim, 2019)
iAMPCN	Multi-label (22 functions)	**High (22 activities)**	Highest functional resolution; uses 4 sequence data types.	Deep learning (CNN)	(Xu et al., 2023)
PepNet	Multi-label (AIPs & AMPs)	Moderate	Interpretable; public web server available.	ResNet + Transformer (PLM)	(Han et al., 2024)
MAPLE	Multi-label (14 activities)	**High (Selectivity)**	Identifies selective motifs; analyzes potency-hemolysis coupling.	PLM + Physicochemical	(Liu et al., 2026)

This open-access article is distributed under the terms of the Creative Commons Attribution Non Commercial 4.0 License (CC BY-NC 4.0).

MIC & STRAIN-SPECIFIC					
DBAASP (Prediction)	Strain-Specific (MSS)	**High (Specific strains)**	Predicts efficacy against specific strains using genomic features. DBAASP v2 offers target-specific ranking, >300 MD trajectory analyses with movies, PDB links, and browser-based structural viewing for drug design. DBAASP v3 provides automated MD trajectories for >3,200 peptides and PDB links for >400, enabling structure-based rational peptide design.	RF + AdaBoost	(Gogoladze et al., 2014, Pirtskhalava et al., 2021, Pirtskhalava et al., 2016)
LLAMP	MIC Regression	**High (Species-aware)**	Screens ~5.5M peptides for MIC prediction.	Large Language Model (LLM)	(Bae et al., 2025)
GENERATIVE & *DE NOVO*					
AMPGAN v1	Generative (De novo)	Low (Design)	Initial proof-of-concept for AMP design.	GAN	(Ferrell et al., 2021)
AMPGAN v2	Generative (*De novo*)	Low (Design)	Enables iterative refinement for customized generation.	BiCGAN	(Van Oort et al., 2021)
AMPGAN v3 / PepCraft	Generative (Multi-objective)	Low (Design + Mods)	Designs D-amino acids/modifications; automated pipeline.	Conditional GAN + Multi-agent	(Jung et al., 2026)
WGAN-GP framework	Generative (*De novo*)	Low (Design)	7/8 synthesized active; 2 broad-spectrum against MRSA/CRPA.	Wasserstein GAN	(Lin et al., 2023)
dsAMP / dsAMPGAN	Binary / Generative	Low (Class + Design)	Universal platform for recognition and *de novo* generation.	CNN-Attention-BiLSTM / GAN	(Zhao et al., 2024)
ProteoGPT	Generative / Screening	Low (Design)	Matches clinical antibiotics in mouse models; no organ toxicity.	Pre-trained protein LLM	(Wang et al., 2025)
DATA RESOURCES & DATABASES					
AMPer	Database / HMM	Low (Discovery)	Uncovered 229 novel AMPs from Swiss-Prot.	HMM	(Fjell et al., 2007)
LAMP	Database / Query-based retrieval	Moderate (Antimicrobial activity & Cytotoxicity)	Contains 5,547 entries (3,904 natural + 1,643 synthetic); uniquely integrates both antimicrobial activity AND cytotoxicity data to aid clinical translation	Keyword / Combinatorial condition search with cross-linking and "top similar AMPs" matching	(Zhao et al., 2013)
dbAMP 2.0	Database & Prediction	Low (Genomic/Proteomic)	26,447 AMPs; integrates structure prediction.	I-TASSER + ML	(Jhong et al., 2022)
AMPSphere	Database (Catalog)	Low (Global Microbiome)	863,498 peptides; 91% novel.	Prodigal + Macrel	(Santos-Júnior et al., 2024)
AMPDiscover	Web server (Multi-label)	Moderate (4 classes)	First to implement Applicability Domain analysis.	RF (AF-QSAMs)	(Pinacho-Castellanos et al., 2021)

Multi-label (Biological Activity), Binary (Yes/No)

Abbreviations: AMP, Antimicrobial Peptide; ABP, Antibacterial Peptide; AIP, Anti-Inflammatory Peptide; MSS, Microbial Strain Specificity; MIC, Minimum Inhibitory Concentration; ML, Machine Learning; DL, Deep Learning; RF, Random Forest; SVM, Support Vector Machine; GB, Gradient Boosting; AdaBoost, Adaptive Boosting; CNN, Convolutional Neural Network; RNN, Recurrent Neural Network; LSTM, Long Short-Term Memory; biLSTM, Bidirectional Long Short-Term Memory; BERT, Bidirectional Encoder Representations from Transformers; PLM, Protein Language Model; LLM, Large Language Model; GAN, Generative Adversarial Network; BiCGAN, Bidirectional Conditional Generative Adversarial Network; WGAN-GP, Wasserstein Generative Adversarial Network with Gradient Penalty; HMM, Hidden Markov Model; I-TASSER, Iterative Threading ASSEmbly Refinement; AF-QSAMs, Alignment-Free Quantitative Sequence–activity models; TAPE, Tasks assessing protein embeddings; GDL, Geometric deep learning; PSSM, Position-specific scoring matrix; AUC, Area under the receiver operating characteristic curve.

This open-access article is distributed under the terms of the Creative Commons Attribution Non Commercial 4.0 License (CC BY-NC 4.0).

Meanwhile, ampir provided a fast, scalable machine-learning classifier tailored for whole-genome scans, integrating smoothly into standard bioinformatics pipelines (at https://github.com/legana/ampir) (Fingerhut et al., 2020). Building on these foundational tools, the deep learning-based AMPlify model leveraged an attentive architecture to screen the bullfrog (*Rana* [*Lithobates*] *catesbeiana*) genome, successfully identifying four novel peptides with confirmed activity against WHO priority pathogens, including a multidrug-resistant carbapenemase-producing *Escherichia coli* isolate (Li et al., 2022).

To support such mining efforts, the dbAMP 2.0 platform was expanded into a comprehensive repository, curating 26,447 AMPs and 2,262 antimicrobial proteins across more than 3,000 organisms. Beyond its extensive dataset, dbAMP 2.0 incorporates I-TASSER-based structural prediction and provides an online tool for rapid identification from genomic, metagenomic, and proteomic inputs, effectively bridging the gap between sequence discovery and translational research (Jhong et al., 2022).

Applying these tools at scale, two major metagenome-wide studies demonstrated the power of data-driven prospecting. Mining the human gut microbiome with a hybrid pipeline combining LSTM, attention mechanisms, and bidirectional encoder representations from transformers (BERT) models, researchers predicted over 2,300 candidate sequences. The synthesis of 216 of these yielded an impressive 181 active peptides, with the top 11 candidates showing not only low similarity to known AMPs but also achieving a tenfold reduction in bacterial burden in a mouse lung infection model (Ma et al., 2022). Expanding this approach globally, the AMPSphere catalog was constructed from tens of thousands of metagenomes and microbial genomes, housing 863,498 non-redundant peptides, over 90% of which were absent from existing databases. As depicted in Fig. 1, the discovery pipeline recovered 118,051 non-singleton clusters organized into 8,788 families, with habitat sharing remaining remarkably limited. Experimental validation of a subset confirmed broad activity, predominantly through membrane disruption (Santos-Júnior et al., 2024).

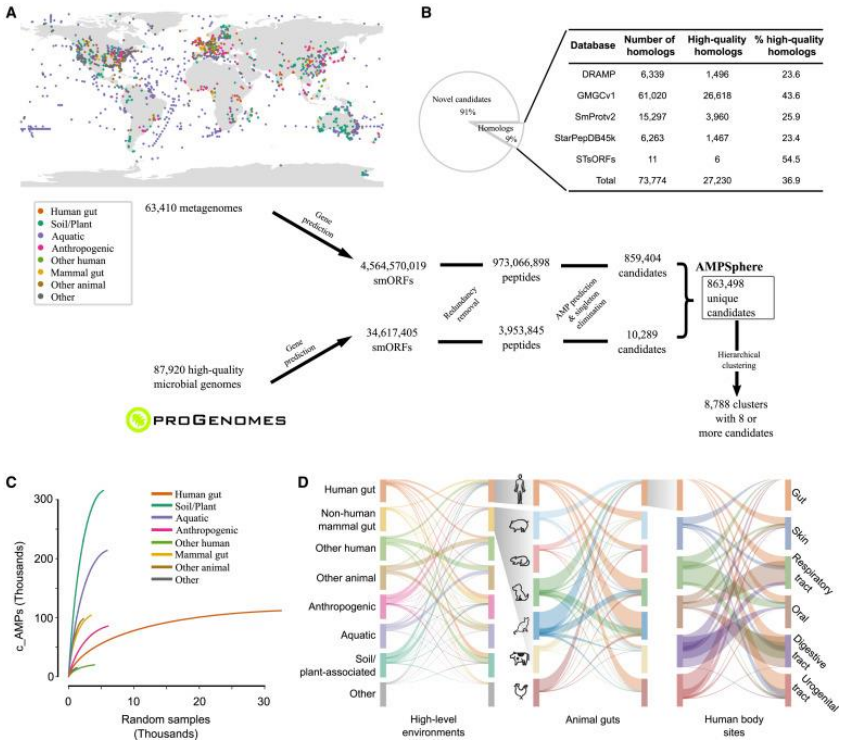

Figure 1: AMPSphere contains 836,498 non-redundant c_AMPs identified from 63,410 metagenomes and 87,920 microbial genomes using Prodigal and Macrel. (A) Pipeline for c_AMP discovery and clustering; 118,051 non-singleton clusters were found at 75% identity, including 8,788 families. (B) Only 9% of c_AMPs had homologs in other peptide or protein databases. (C) Rarefaction curves show that more sampling yields more AMP discovery. (D) c_AMP sharing across habitats is limited (Santos-Júnior et al., 2024). [This is an open access article distributed under the terms of the Creative Commons CC-BY license, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.]

This open-access article is distributed under the terms of the Creative Commons Attribution Non Commercial 4.0 License (CC BY-NC 4.0).

Narrowing the focus to a specific microbial family, a recent genome-mining and machine-learning workflow explored the Lactobacillaceae family, analyzing over 10,000 genomes to uncover 9,601 AMPs organized into 2,092 gene cluster families (GCFs). These clusters exhibited pronounced species specificity and habitat uniqueness, with machine-learning predictions ranking 664 as highly promising (predicted MIC below 100 μM). From this pool, 16 synthesized candidates were tested, of which 10 demonstrated efficacy against 11 common pathogens. Notably, the *Lactobacillus delbrueckii*-derived peptide delbruin_1 exhibited potent broad-spectrum activity, reinforcing the underappreciated potential of health-associated bacteria as reservoirs for new antimicrobials (Du et al., 2025).

AMP identification from proteome

The journey toward computational AMP discovery began in 2007 with the development of an HMM-based framework that achieved up to 99% recognition accuracy by building separate models for mature peptides and propeptides across major AMP classes. Iterative scanning of Swiss-Prot uncovered 229 previously unidentified AMPs, which were compiled into the open-access AMPer database (Fjell et al., 2007). This early success established the foundation for sequence-based prediction, but the field soon evolved toward more sophisticated approaches.

A major paradigm shift occurred with the adaptation of protein language models (PLMs) from large language model architectures, treating amino acid sequences as linguistic structures. These models accelerate protein structure prediction, functional annotation, and *de novo* design, though their substantial computational demands remain an active area of optimization (Leclercq and Droit 2026). During this transition, traditional machine learning continued to prove its value. AmPEP, for instance, demonstrated that a Random Forest (RF) classifier trained on the distribution of physicochemical properties across 3,268 AMPs and nearly 167,000 non-AMP sequences could achieve high predictive performance, with the notable finding that a reduced feature set retained much of its effectiveness (Bhadra et al., 2018).

Addressing the persistent challenge of limited labeled data, researchers benchmarked several PLMs, including ProtBert, ProtAlbert, ProteinBERT, and TAPE, against neural-based classifiers through transfer learning. A 2025 study confirmed that increasing model scale and employing PLM-derived representations with shallow classifiers yielded state-of-the-art performance, with parameter fine-tuning further optimizing accuracy and providing a robust framework for peptide identification despite data scarcity (Georgoulis et al., 2025). Building on this concept, a dual-objective ML framework leveraging ProtBERT was trained on approximately 25,000 sequences to simultaneously screen for antimicrobial activity and hemolytic activity. Applied to a library of 20,000 Melectin-derived variants (6–12 amino acids), this approach yielded 16 hit candidates, two of which contained at least four mutations relative to their parent sequences. Experimental validation confirmed potent activity against MRSA and *P. aeruginosa*, with the D-enantiomer retaining efficacy in serum-rich media and 3D-spheroid models confirming strong antimicrobial properties with minimal host cytotoxicity (Das et al., 2026).

To further maximize the utility of peptide sequences, PepNet emerged as an interpretable neural network integrating pre-trained PLMs with a two-stage architecture: residual dilated convolution blocks capture physicochemical and structural arrangements, while residual Transformer blocks extract complex functional information. This design delivered superior predictive performance for both anti-inflammatory peptides and AMPs, with robust interpretability and a public web server available at http://liulab.top/PepNet/server (Han et al., 2024). Around the same time, Target-AMP demonstrated that combining sequential information with evolutionary profiles, which used PSSM (Position-Specific Scoring Matrix), pseudo amino acid composition, and dipeptide composition, could achieve 97.07% accuracy on independent datasets and 95.71% on training sets when paired with an SVM-based hybrid model (Jan et al., 2022).

Another transformer revolution also produced AMP-BERT, a fine-tuned BERT architecture that extracted high-dimensional structural and functional motifs through self-attention mechanisms. Rigorous

This open-access article is distributed under the terms of the Creative Commons Attribution Non Commercial 4.0 License (CC BY-NC 4.0).

validation against state-of-the-art benchmarks confirmed its superior classification performance, while attention weight analysis enabled interpretable feature mapping directly linking residues to antimicrobial function (Lee et al., 2023). Regarding antibacterial peptides (ABPs) specifically, Deep-ABPpred employed a biLSTM architecture with feature extraction, achieving 97% precision on test datasets and 94% on independent sets. Its practical utility was demonstrated by identifying ABPs within *Streptococcus* bacteriophage tail proteins, which were then synthesized and validated *in vitro* against Gram-positive and Gram-negative bacteria (Sharma et al., 2021). To refine this concept further, StaBle-ABPpred introduced a stacked ensemble approach combining biLSTM with attention mechanisms and a meta-level ensemble of RF, Gradient Boosting, and Logistic Regression, significantly outperforming existing classifiers. Its web application (https://stable-abppred.anvil.app) successfully identified novel ABPs within the *Streptococcus* phage T12 genome (Singh et al., 2022).

The trend toward hybridization continued with iAMP-DL, which integrated CNNs for motif detection and LSTMs for sequential pattern recognition, providing a stable and efficient platform for rapid screening of short, high-potency AMPs with superior predictive accuracy and robustness (Nguyen et al., 2024). Building on this notion, AMP-Detector emerged as a classification pipeline (Fig. 2) synergizing PLMs with machine learning algorithms to train multiple specialized models for diverse biological activities, achieving high precision (avg. >83%).

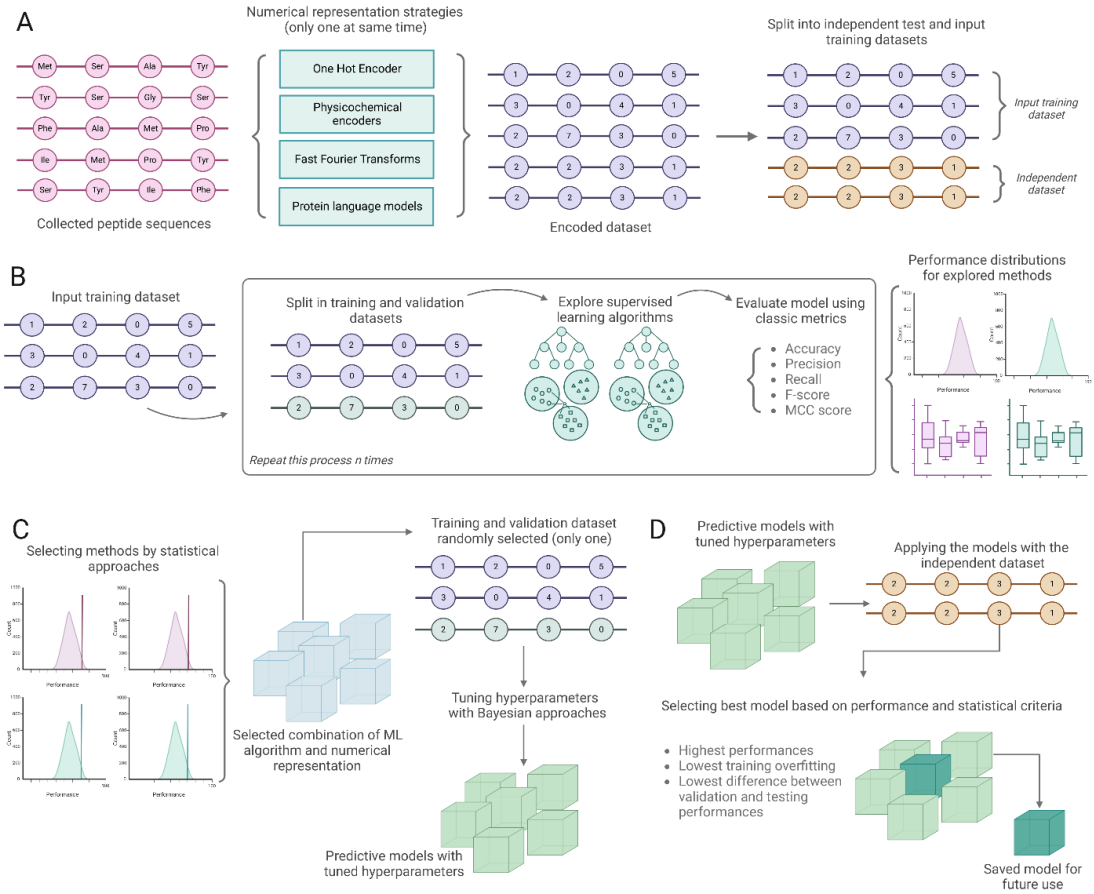

Figure 2: Overview of the model generation and evaluation workflow: (A) Encoding of sequence datasets using one-hot, physicochemical, or pre-trained embedding strategies, followed by a 90:10 split into development and benchmark sets. (B) Iterative model training and validation using 80:20 partitions to derive performance distributions across different architectures and hyperparameters. (C) Statistical selection of optimal algorithm-encoding pairs followed by Bayesian hyperparameter tuning. (D) Final performance validation of the optimized models against the independent benchmark set and comparative analysis with existing methods (Medina-Ortiz et al., 2024). [Published under an open access Creative Commons CC BY license, any part of the article may be reused without permission provided that it is clearly cited.]

This open-access article is distributed under the terms of the Creative Commons Attribution Non Commercial 4.0 License (CC BY-NC 4.0).

Uniquely, the framework extended beyond classification by incorporating generative learning capabilities for *de novo* design, bridging predictive screening and therapeutic synthesis. Its multi-stage pipeline employed one-hot, physicochemical, or embedding-based encodings, iterative 80:20 training/validation partitions, and Bayesian hyperparameter optimization to establish optimal model configurations (Medina-Ortiz et al., 2024).

Shaping on the generative paradigm, a recent landmark study introduced ProteoGPT, a pre-trained protein large language model adapted into specialized submodels for AMP discovery and design. The researchers constructed a sequential pipeline that leveraged transfer learning to perform three interconnected tasks: screening hundreds of millions of peptide sequences, identifying the most promising candidates, and generating entirely novel peptides optimized for both antimicrobial potency and minimal cytotoxicity. This dual-objective approach directly addresses one of the field's persistent bottlenecks—the need to balance efficacy against host cell safety. Reaching an unprecedented scale, ProteoGPT screened hundreds of millions of sequences to generate peptides that matched clinical antibiotics in mouse models without toxicity or microbiota disruption (Wang et al., 2025).

De novo design of novel AMPs

The generative design of antimicrobial peptides has evolved rapidly through successive innovations in deep learning architectures. The pioneering AMPGAN series, initiated by Ferrell et al., established the feasibility of adversarial networks for AMP generation, training an initial model on a dataset of 500,000 non-AMP peptides and nearly 8,000 labeled AMP sequences while incorporating membrane binding tendency as a design criterion (Ferrell et al., 2021). This foundational work provided promising proof-of-concept for AI-driven peptide design. Building upon this foundation, AMPGAN v2 introduced a bidirectional conditional generative adversarial network (BiCGAN) that leverages a generator-discriminator framework conditioned on specific variables to guide peptide generation. A trained encoder projecting data samples into the generator's latent space enabled iterative refinement, allowing the model to produce novel, diverse peptides customized for particular applications (Van Oort et al., 2021). Most recently, AMPGAN v3 advanced the field further as a multi-objective conditional GAN capable of designing peptides incorporating D-amino acids and N- or C-terminal modifications such as amidation. By decoupling the adversarial objective from activity prediction through two distinct discriminators, the model achieved improved training stability and superior performance on external classifiers. Concurrently, the PepCraft multi-agent system was introduced to accelerate discovery through an automated pipeline coordinated by a planning agent managing design, filtering, and validation stages (Jung et al., 2026).

In parallel with these generative advances, substantial efforts have focused on enhancing predictive specificity. A critical limitation of most computational tools has been their inability to predict efficacy against specific microbial strains, a deficiency largely attributable to the scarcity of strain-specific datasets. To address this gap, the MSSPM (Microbial Strain Specificity Predictive Model) approach was developed, integrating AMP sequence properties with genomic characteristics of target strains to enable accurate predictions even for organisms with no prior experimental data. Through rigorous evaluation of feature engineering techniques and machine learning algorithms, RF and AdaBoost emerged as the most effective models, with genomic features significantly enhancing predictive performance compared to sequence-only approaches. This tool is publicly accessible through the DBAASP database (Vishnepolsky et al., 2022).

Recognizing the therapeutic promise of short peptides, Bournez et al. introduced CalcAMP, a machine learning framework engineered specifically for high-accuracy prediction of short-sequence AMPs (fewer than 35 amino acids). Leveraging a specialized dataset from diverse public repositories and empirical results, CalcAMP systematically analyzes both physicochemical descriptors and sequence-based features to achieve robust predictive performance against both Gram-positive and Gram-negative strains, circumventing the high costs of conventional *in vitro* screening (Bournez et al., 2023).

The broader methodological landscape has witnessed a paradigm shift from traditional machine

This open-access article is distributed under the terms of the Creative Commons Attribution Non Commercial 4.0 License (CC BY-NC 4.0).

learning, which relies on intensive domain-specific feature engineering, toward deep learning architectures capable of automated feature extraction. While classical models like SVM and RF remain robust, deep neural networks have demonstrated superior performance in complex bioinformatics tasks. Specialized DNN architectures have achieved remarkable accuracy across functional classes: ENNAACT attained 98.3% for anticancer peptides (Timmons and Hewage, 2021a), ENNAVIA reached 95.7% for antiviral peptides (Timmons and Hewage, 2021b), and Deep-AntiFP demonstrated high efficacy for antifungal prediction (Ahmad et al., 2021). However, the literature advises judicious adoption of deep learning, as increased computational complexity must be weighed against performance gains over shallower, more efficient models (Yan et al., 2022).

Exemplifying this deep learning approach, a study employing a Wasserstein Generative Adversarial Network with gradient penalty (WGAN-GP) generated entirely novel peptide sequences from existing AMP datasets, filtering candidates through an AI-based classifier. From this pipeline, eight top-tier peptides were synthesized and rigorously tested via disc diffusion assays and MIC determinations. Remarkably, seven of the eight AI-designed peptides demonstrated clear antibacterial activity, with two candidates showing significant broad-spectrum efficacy against notoriously difficult pathogens including MRSA and carbapenem-resistant *P. aeruginosa*. This framework exemplified the power of bridging generative AI with wet-bench validation, demonstrating that such integrated approaches can significantly accelerate the identification of effective peptides against both common and drug-resistant pathogens, with potential adaptability to other functional peptide classes (Lin et al., 2023).

Addressing the pervasive challenge of dataset bias and redundancy, another group developed a novel modeling framework using the largest experimentally validated, non-redundant peptide dataset reported to date. Their alignment-free quantitative sequence–activity models (AF-QSAMs), built on RF algorithms, were constructed to identify both general AMPs and specific functional classifications—antibacterial, antifungal, antiparasitic, and antiviral. A pioneering contribution was the implementation of applicability domain analysis, performed for the first time in the context of AMP recognition to ensure prediction reliability. Benchmarking against thirteen established programs, including ClassAMP, iAMP-2L, and Meta-iAVP, revealed that the proposed models achieved superior performance across all endpoints, while most existing methods exhibited weak-to-random predictive agreement. Rigorous validation through Y-scrambling and repeated k-fold cross-validation confirmed the robustness of these predictions, with three distinct chemometric analyses confirming descriptor relevance. These models were made accessible through the AMPDiscover web tool (Pinacho-Castellanos et al., 2021).

Recent breakthroughs have yielded increasingly comprehensive platforms. The discoverAMP (dsAMP) framework combined CNN-Attention-BiLSTM modules with transfer learning to create a powerful classifier, complemented by dsAMPGAN, a generative adversarial network for *de novo* design. This unified approach provides an efficient *in silico* platform for both AMP discovery and design applicable to therapeutic development against antimicrobial resistance (Zhao et al., 2024). Concurrently, addressing the significant imbalance among different functional AMP families, researchers developed ML-SMOTE, a synthetic minority over-sampling technique specifically designed for imbalanced multi-label datasets. Integrated with grey pseudo amino acid composition features, this approach yielded the MLAMP classifier, achieving a subset accuracy of 0.4846 and a Hamming loss of 0.16. A publicly accessible web server now enables identification of AMP functional families (Lin and Xu, 2016).

The trajectory toward more interpretable and complex architectures is exemplified by iAMP-CRA, which integrates Convolutional Recurrent Neural Networks with Self-Attention mechanisms to adaptively incorporate heterogeneous features, including evolutionary information and traditional protein descriptors, achieving competitive accuracy of 0.919 on independent benchmarks (Lu et al., 2025).

Fundamental to all these advances is the recognition that AMP design is constrained by the immense complexity of protein sequence space.

 This open-access article is distributed under the terms of the Creative Commons Attribution Non Commercial 4.0 License (CC BY-NC 4.0).

With 20 standard amino acids, the combinatorial possibilities for a peptide of length L scale as 20^L, creating an astronomically vast landscape that renders exhaustive search computationally intractable. The primary challenge lies in controllable sampling and multi-objective optimization, simultaneously fine-tuning for antimicrobial potency, minimal hemolytic activity, and metabolic stability. Deep generative models, Variational Autoencoders (VAEs), Generative Adversarial Networks (GANs), and Diffusion Models, have emerged to address this challenge. VAEs employ probabilistic mapping of sequences into a latent Gaussian space, enabling continuous interpolation and decoding; GANs utilize adversarial training between generator and discriminator; and diffusion models follow an iterative denoising process within a latent manifold structured via encoder–decoder frameworks to transform noise into functional peptide sequences (Sun et al., 2026).

On the prediction front, LLAMP (Large Language Model for AMP activity prediction) leveraged pre-trained LLMs for species-aware MIC prediction. Through large-scale virtual screening of approximately 5.5 million peptide sequences, the model identified Trp, Lys, and Phe as critical activity determinants via attention mechanism analysis. These interpretability insights guided rational design, where the highly selective peptide 13 was engineered via targeted modifications to enhance amphipathicity, yielding peptide 13-5. When benchmarked against the clinical AMP pexiganan, both peptide 13-5 and peptide 16 demonstrated comparable or superior therapeutic potential (Bae et al., 2025).

Finally, addressing the persistent challenge of selectivity, MAPLE (Multifunctional AMP Learning Engine) emerged as a dual-stream, interpretable machine learning framework integrating PLM embeddings with explicit physicochemical descriptors for systematic functional profiling across 14 activity categories. To ensure robustness against severe label imbalance, the model combined high-dimensional sequence representations with hand-crafted physical features. Systematic κ-mer enrichment analysis mapped motif-level selectivity within the sequence space, investigating whether potency–hemolysis coupling represented a universal property or was regulated by specific sequence motifs. A prioritization workflow leveraging identified antibacterial-selective motifs, followed by structural modeling, confirmed that the predicted conformations were consistent with amphipathic α-helices (Liu et al., 2026).

Multi-objective optimization

While early computational tools effectively distinguished AMPs from non-AMPs, they largely failed to predict specific biological functions. Multi-objective optimization has emerged to address this limitation, enabling simultaneous classification of multiple therapeutic activities. This paradigm shift toward comprehensive functional profiling has culminated in advanced deep learning architectures with unprecedented predictive resolution. Based on a study provided in bioinformatics, researchers identified a significant gap in the field of AI-based AMP identification: while many existing tools were good at identifying whether a sequence was an AMP, very few could actually predict the specific functional activities of those peptides. To address this, the study introduced a new deep learning model called iAMPCN (identification of AMPs based on CNNs). This study introduced a much smarter tool that can actually identify 22 different specific functions within those peptides. While previous tools were pretty limited in their scope, this new deep-learning model uses four different types of sequence data to paint a much clearer picture of what a peptide actually did. When put to the test, it outperformed existing top-tier methods, proving that we could now move past broad, general predictions and start pinpointing specific therapeutic activities with much higher precision (Xu et al., 2023).

Recognizing the absence of standardized evaluation criteria for AMP prediction tools, Gabere and Noble conducted a comprehensive comparative study in 2017, benchmarking six widely used platforms: ADAM, $CAMP_{R3}$(RF), $CAMP_{R3}$(SVM), MLAMP, and DBAASP. Using the Area Under the Receiver Operating Characteristic Curve (AUC-ROC) as the primary performance metric, their analysis revealed that $CAMP_{R3}$(RF) achieved a statistically significant lead over all five competitors, establishing it as the most reliable general-purpose binary classifier at that time. This study underscored a persistent challenge in the field,

This open-access article is distributed under the terms of the Creative Commons Attribution Non Commercial 4.0 License (CC BY-NC 4.0).

including the difficulty of objectively assessing tool performance due to heterogeneous training datasets and evaluation protocols, and emphasized the necessity for standardized benchmarking frameworks to guide researchers in selecting appropriate computational platforms for AMP discovery (Gabere and Noble, 2017). This benchmarking effort was later revisited, which identified a substantial discrepancy in the original DBAASP performance metrics. Rather than the 121 correctly identified peptides reported by Gabere and Noble, the corrected analysis revealed that DBAASP actually achieved 306 correct identifications, a 2.5-fold difference that significantly altered the comparative performance landscape. To address the underlying issue of methodological inconsistency, the authors developed "compatible" datasets specifically tailored to respect the algorithmic constraints of each prediction tool, concluding that valid comparative assessments require test datasets that align with each tool's technical limitations. Concurrently, a publicly accessible web server for the MLAMP classifier was established at http://www.jci-bioinfo.cn/MLAMP, facilitating broader adoption and functional characterization of AMPs (Vishnepolsky and Pirtskhalava, 2019).

To address the gap in functional specificity, the AMAP model was then developed using an ML approach focused on multi-label classification. Unlike traditional predictors, AMAP is engineered to simultaneously predict 14 distinct biological functions, providing a multidimensional profile of peptide activity. A key strength of the AMAP framework lies in its demonstrated generalization capability. Beyond standard cross-validation, the model was rigorously benchmarked against recently published, experimentally verified peptide datasets that were absent from its training set. This benchmarking confirms that AMAP does not merely "memorize" known motifs but successfully captures the underlying physicochemical principles governing biological activity, making it a robust tool for the discovery of novel, non-canonical peptides (Gull and Shamim, 2019).

Geometric deep learning (GDL), which uses deep neural networks built for non-Euclidean data such as graphs and manifolds, was also used in AMP design. Its main point is that antimicrobial peptide sequences and structures are better modeled in this geometric framework than with standard flat representations, because their folding and function depend on complex 3D structure (Fernandes et al., 2023). As the field advances toward more sophisticated multi-label predictors, the establishment of reproducible benchmarking frameworks with carefully tailored test sets remains essential for accurate tool comparison and informed selection by researchers.

Challenges and future perspectives

Recent reviews converge on a similar picture: AI and machine learning have made AMP identification much faster, but the field is still limited by sparse, heterogeneous, and inconsistently evaluated data. One 2024 review notes that these methods may help overcome barriers such as poor stability, toxicity, and limited cellular penetration, while also creating "new opportunities in data-driven peptide design" (Wan et al., 2024). A 2025 review argues that the next step is not a single model, but multiobjective and multimodal frameworks that combine active learning, protein language models, graph neural networks, mechanistic interpretability, curated data resources, and uncertainty estimation (Saleem et al., 2025). The benchmarking problem is especially important, as warned by a recent article, because heterogeneous datasets and inconsistent protocols can produce data leakage and inflated performance estimates, which means future progress will depend on reproducible benchmarks and stronger wet-lab validation (Cisterna-García et al., 2026).

To sum up, it seems useful that AI-driven AMP discovery moves beyond single-model prediction toward closed-loop, experimentally validated pipelines. The next step is likely to combine active-learning lab cycles, protein language models, graph neural networks, curated data resources, and uncertainty estimation, while also tackling the core benchmark problem caused by heterogeneous datasets and inconsistent evaluation protocols. In practice, the field would need standardized, reproducible benchmarks and stronger wet-lab validation if it were to translate these models into clinically useful peptide design.

This open-access article is distributed under the terms of the Creative Commons Attribution Non Commercial 4.0 License (CC BY-NC 4.0).

References

Ahmad, A., Akbar, S., Khan, S., Hayat, M., Ali, F., Ahmed, A. and M. Tahir, (2021). "Deep-AntiFP: Prediction of antifungal peptides using distant multi-informative features incorporating deep neural networks." *Chemometrics and Intelligent Laboratory Systems,* **208**: 104214. DOI: https://doi.org/10.1016/j.chemolab.2020.104214.

Al Musaimi, O. (2025). "FDA-approved antibacterials and echinocandins." *Antibiotics*, **14**(2):166. DOI: https://doi.org/10.3390/antibiotics14020166.

Alzain, M., Daghistani, H., Shamrani, T., Almoghrabi, Y., Daghistani, Y., Alharbi, O. S., Sait, A. M., Mufrrih, M., Alhazmi, W. and M. A. Alqarni, (2025). "Antimicrobial peptides: mechanisms, applications, and therapeutic potential." *Infection and Drug Resistance,* **18**:4385-4426. DOI: https://doi.org/10.2147/IDR.S514825.

Bae, D., Kim, M., Seo, J. and H. Nam, (2025). "AI-guided discovery and optimization of antimicrobial peptides through species-aware language model." *Briefings in Bioinformatics,* **26**(4): bbaf343. DOI: https://doi.org/10.1093/bib/bbaf343.

Bhadra, P., Yan, J., Li, J., Fong, S. and S. W. I. Siu, (2018). "AmPEP: Sequence-based prediction of antimicrobial peptides using distribution patterns of amino acid properties and random forest." *Scientific Reports,* **8**(1): 1697. DOI: https://doi.org/10.1038/s41598-018-19752-w.

Botelho Sampaio de Oliveira, K., Lopes Leite, M., Albuquerque Cunha, V., Brito da Cunha, N. and O. Luiz Franco, (2023). "Challenges and advances in antimicrobial peptide development." *Drug Discovery Today,* **28**(8): 103629. DOI: https://doi.org/10.1016/j.drudis.2023.103629.

Bournez, C., Riool, M., de Boer, L., Cordfunke, R. A., de Best, L., van Leeuwen, R., Drijfhout, J. W., Zaat, S. A. J. and G. J. P. van Westen, (2023). "CalcAMP: A new machine learning model for the accurate prediction of antimicrobial activity of peptides." *Antibiotics*, **12**(4): 725. DOI: https://doi.org/10.3390/antibiotics12040725.

Brizuela, C. A., Liu, G., Stokes, J. M. and C. de la Fuente-Nunez, (2025). "AI methods for antimicrobial peptides: progress and challenges." *Microbial Biotechnology,* **18**(1): e70072. DOI: https://doi.org/10.1111/1751-7915.70072.

Cherkasov, A., Hilpert, K., Jenssen, H., Fjell, C. D., Waldbrook, M., Mullaly, S. C., Volkmer, R. and R. E. W. Hancock, (2008). "Use of artificial intelligence in the design of small peptide antibiotics effective against a broad spectrum of highly antibiotic-resistant superbugs." *ACS Chemical Biology,* **4**(1): 65-74. DOI: https://doi.org/10.1021/cb800240j.

Cisterna-García, A., González-López, A. M., Vozi, A., Esteban, M. A., Egli, A., Jutzeler, C., Palma, J., Sánchez-Ferrer, A. and J. A. J. b. Botía, (2026). "Antimicrobial peptide databases and prediction tools: Toward a standard evaluation framework." *bioRxiv*, **2026**.05: 19.726290. DOI: https://doi.org/10.64898/2026.05.19.726290.

Das, A., Sen, S., Ahmadi, Z. and M. Debnath, (2026). "Dual-objective protein language model for discovery of non-hemolytic antimicrobial peptides from the bee-melectin sequence space." *Chemical Engineering Journal,* **543**: 178296. DOI: https://doi.org/10.1016/j.cej.2026.178296

Du, R., Han, F., Li, Z., Yu, J., Xu, Y., Huang, Y. and Q. Wu, (2025). "Uncovering encrypted antimicrobial peptides in health-associated Lactobacillaceae by large-scale genomics and machine learning." *Microbiome,* **13**(1): 151. DOI: https://doi.org/10.1186/s40168-025-02145-3.

Fathi, F., Alizadeh, B., Tabarzad, M. V. and M. Tabarzad, (2024). "Important structural features of antimicrobial peptides towards specific activity: Trends in the development of efficient therapeutics." *Bioorganic Chemistry,* 149: 107524. DOI: https://doi.org/10.1016/j.bioorg.2024.107524.

Fernandes, F. C., Cardoso, M. H., Gil-Ley, A., Luchi, L. V., da Silva, M. G. L., Macedo, M. L. R., de la Fuente-Nunez, C. and O. L. Franco, (2023). "Geometric deep learning as a potential tool for antimicrobial peptide prediction." *Frontiers in Bioinformatics,* **3**: 1216362. DOI: https://doi.org/10.3389/fbinf.2023.1216362.

Ferrell, J. B., Remington, J. M., Van Oort, C. M., Sharafi, M., Aboushousha, R., Janssen-Heininger, Y., Schneebeli, S. T., Wargo, M. J., Wshah, S. and J. Li, (2021). "A Generative approach toward precision antimicrobial peptide design." *bioRxiv*, **2020**.10: 02.324087. DOI: https://doi.org/10.1101/2020.10.02.324087.

Fingerhut, L. C. H. W., Miller, D. J., Strugnell, J. M., Daly, N. L. and I. R. Cooke, (2020). "ampir: an R package for fast genome-wide prediction of antimicrobial peptides." *Bioinformatics,* 36(21): 5262-5263. DOI: https://doi.org/10.1093/bioinformatics/btaa653.

Fjell, C. D., Hancock, R. E. W. and A. Cherkasov, (2007). "AMPer: a database and an automated discovery tool for antimicrobial peptides." *Bioinformatics,* **23**(9): 1148-1155. DOI: https://doi.org/10.1093/bioinformatics/btm068.

Fong, I. W. (2023). "Antimicrobial resistance: a crisis in the making." In: Fong, I. W., (ed.) *New Antimicrobials: For the Present and the Future*, Cham: Springer International Publishing. pp. 1-21. DOI: https://doi.org/10.1007/978-3-031-26078-0_1.

Fu, H., Cao, Z., Li, M. and S. Wang, (2020). "ACEP: improving antimicrobial peptides recognition through automatic feature fusion and amino acid embedding." *BMC Genomics,* **21**(1): 597. DOI: https://doi.org/10.1186/s12864-020-06978-0.

Gabere, M. N. and W. S. Noble, (2017). "Empirical comparison of web-based antimicrobial peptide prediction tools." *Bioinformatics,* **33**(13): 1921-1929. DOI: https://doi.org/10.1093/bioinformatics/btx081.

Georgoulis, E., Zervou, M. A. and Y. Pantazis, (2025). "Transfer learning on protein language models improves antimicrobial peptide classification." *Scientific Reports,* **15**(1): 37456. DOI: https://doi.org/10.1038/s41598-025-21223-y.

Girdhar, M., Sen, A., Nigam, A., Oswalia, J., Kumar, S. and R. Gupta, (2024). "Antimicrobial peptide-based strategies to overcome antimicrobial resistance." *Archives of Microbiology,* **206**(10): 411. DOI: https://doi.org/10.1007/s00203-024-04133-x.

Gogoladze, G., Grigolava, M., Vishnepolsky, B., Chubinidze, M., Duroux, P., Lefranc, M.-P. and M. Pirtskhalava, (2014). "DBAASP: database of antimicrobial activity and structure of peptides." *FEMS Microbiology Letters,* **357**(1): 63-68. DOI: https://doi.org/10.1111/1574-6968.12489.

Gull, S. and N. Shamim, (2019). "AMAP: Hierarchical multi-label prediction of biologically active and antimicrobial peptides." *Computers in Biology and Medicine,* **107**: 172-181. DOI: https://doi.org/10.1016/j.compbiomed.2019.02.018.

Han, J., Kong, T. and J. Liu, (2024). "PepNet: an interpretable neural network for anti-inflammatory and antimicrobial peptides prediction using a pre-trained protein language model." *Communications Biology,* **7**(1): 1198. DOI: https://doi.org/10.1038/s42003-024-06911-1.

Hanson, M. A., Lemaitre, B. and R. L. Unckless, (2019). "Dynamic evolution of antimicrobial peptides underscores trade-offs between immunity and ecological fitness." *Frontiers in Immunology,* **10**: 2620. DOI: https://doi.org/10.3389/fimmu.2019.02620.

Hetta, H. F., Sirag, N., Alsharif, S. M., Alharbi, A. A., Alkindy, T. T., Alkhamali, A., Albalawi, A. S., Ramadan, Y. N., Rashed, Z. I. and F. E. Alanazi, (2024). "Antimicrobial peptides: the game-changer in the epic battle against multidrug-resistant bacteria." *Pharmaceuticals*, **17**(11): 1555. DOI: https://doi.org/10.3390/ph17111555.

Jan, A., Hayat, M., Wedyan, M., Alturki, R., Gazzawe, F., Ali, H. and F. K. Alarfaj, (2022). "Target-AMP: Computational prediction of antimicrobial peptides by coupling sequential information with evolutionary profile." *Computers in Biology and Medicine,* **151**: 106311. DOI: https://doi.org/10.1016/j.compbiomed.2022.106311.

Jesudason, T. (2024). "WHO publishes updated list of bacterial priority pathogens." *The Lancet Microbe,* **5**(9). DOI: https://doi.org/10.1016/j.lanmic.2024.07.00.

This open-access article is distributed under the terms of the Creative Commons Attribution Non Commercial 4.0 License (CC BY-NC 4.0).

Jhong, J.-H., Yao, L., Pang, Y., Li, Z., Chung, C.-R., Wang, R., Li, S., Li, W., Luo, M., Ma, R., Huang, Y., Zhu, X., Zhang, J., Feng, H., Cheng, Q., Wang, C., Xi, K., Wu, L.-C., Chang, T.-H., Horng, J.-T., Zhu, L., Chiang, Y.-C., Wang, Z. and T.-Y. Lee, (2022). "dbAMP 2.0: updated resource for antimicrobial peptides with an enhanced scanning method for genomic and proteomic data." *Nucleic Acids Research,* **50**(D1): D460-D470. DOI: https://doi.org/10.1093/nar/gkab1080.

Jung, J. H., Zhang, X., Song, S., Sayedahmed, M., Xiang, C., Xu, Y., AbdelKhalek, A., Schneebeli, S. T., Wargo, M. J., Li, J. and S. Wshah, (2026). "Agentic discovery of non-canonical antimicrobial peptides with AMPGAN v3." *ArXiv*. DOI: https://doi.org/10.48550/arXiv.2606.17127.

Kim, S. H., Min, Y.-H. and M. C. Park, (2025). "Antimicrobial peptides: current status, mechanisms of action, and strategies to overcome therapeutic limitations." *Microorganisms*, **13**(11): 2574. DOI: https://doi.org/10.3390/microorganisms13112574.

Lata, S., Mishra, N. K. and G. P. S. Raghava, (2010). "AntiBP2: improved version of antibacterial peptide prediction." *BMC Bioinformatics,* **11**(1): S19. DOI: https://doi.org/10.1186/1471-2105-11-S1-S19.

Leclercq, M. and A. Droit, (2026). "Protein language models: applications and perspectives." *Journal of Proteome Research,* **25**(2): 507-524. DOI: https://doi.org/10.1021/acs.jproteome.5c00506.

Lee, H.-T., Lee, C.-C., Yang, J.-R., Lai, J. Z. C. and K. Y. Chang, (2015). "A Large-scale structural classification of antimicrobial peptides." *BioMed Research International,* **2015**(1): 475062. DOI: https://doi.org/10.1155/2015/475062.

Lee, H., Lee, S., Lee, I. and H. Nam, (2023). "AMP-BERT: Prediction of antimicrobial peptide function based on a BERT model." *Protein Science,* **32**(1): e4529. DOI: https://doi.org/10.1002/pro.4529.

Li, C., Sutherland, D., Hammond, S. A., Yang, C., Taho, F., Bergman, L., Houston, S., Warren, R. L., Wong, T., Hoang, L. M. N., Cameron, C. E., Helbing, C. C. and I. Birol, (2022). "AMPlify: attentive deep learning model for discovery of novel antimicrobial peptides effective against WHO priority pathogens." *BMC Genomics,* **23**(1): 77. DOI: https://doi.org/10.1186/s12864-022-08310-4.

Li, B., Kang, W., Liu, H., Wang, Y., Yu, C., Zhu, X., Dou, J., Cai, H. and C. Zhou, (2016). "The antimicrobial activity of cbf-k16 against mrsa was enhanced by β-lactamantibiotics through cell wall non-integrity." *Archives of Pharmacal Research*, **39**(7): 978-988. DOI: https://doi.org/10.1007/s12272-016-0769-x.

Li, J., Pu, Y., Tang, J., Zou, Q. and F. Guo, (2020). "DeepAVP: a dual-channel deep neural network for identifying variable-length antiviral peptides." *IEEE Journal of Biomedical and Health Informatics (JBHI),* **24**(10): 3012-3019. DOI: https://doi.org/10.1109/JBHI.2020.2977091.

Lin, T.-T., Yang, L.-Y., Lin, C.-Y., Wang, C.-T., Lai, C.-W., Ko, C.-F., Shih, Y.-H. and S.-H. Chen, (2023). "Intelligent de novo design of novel antimicrobial peptides against antibiotic-resistant bacteria strains." *International Journal of Molecular Sciences*, **24**(7): 6788. DOI: https://doi.org/10.3390/ijms24076788.

Lin, W. and D. Xu, (2016). "Imbalanced multi-label learning for identifying antimicrobial peptides and their functional types." *Bioinformatics,* **32**(24): 3745-3752. DOI: https://doi.org/10.1093/bioinformatics/btw560.

Liu, G.-Y., Yu, D., Fan, M.-M., Zhang, X., Jin, Z.-Y., Tang, C. and X.-F. Liu, (2024). "Antimicrobial resistance crisis: could artificial intelligence be the solution?" *Military Medical Research,* **11**(1): 7. DOI: https://doi.org/10.1186/s40779-024-00510-1.

Liu, H., Shi, Y., Guo, F., Wang, J., Li, J., Wang, G., Zhan, D.-C., Hao, H. and G. Yu, (2026). "MAPLE: interpretable deep learning identifies selective antimicrobial peptides using joint evolutionary–physicochemical analysis." *Briefings in Bioinformatics,* **27**(3): bbag318. DOI: https://doi.org/10.1093/bib/bbag318.

Lu, J., He, Y., Han, G. and L. Zeng, (2025). "iAMP-CRA: Identifying antimicrobial peptides using convolutional recurrent neural network with self-attention." *Health Information Science and Systems,* **13**(1): 25. DOI: https://doi.org/10.1007/s13755-025-00342-w.

Ma, Y., Guo, Z., Xia, B., Zhang, Y., Liu, X., Yu, Y., Tang, N., Tong, X., Wang, M., Ye, X., Feng, J., Chen, Y. and J. Wang, (2022). "Identification of antimicrobial peptides from the human gut microbiome using deep learning." *Nature Biotechnology,* **40**(6): 921-931. DOI: https://doi.org/10.1038/s41587-022-01226-0.

Medina-Ortiz, D., Contreras, S., Fernández, D., Soto-García, N., Moya, I., Cabas-Mora, G. and Á. Olivera-Nappa, (2024). "Protein language models and machine learning facilitate the identification of antimicrobial peptides." *International Journal of Molecular Sciences*, **25**(16): 8851. DOI: https://doi.org/10.3390/ijms25168851.

Meher, P. K., Sahu, T. K., Saini, V. and A. R. Rao, (2017). "Predicting antimicrobial peptides with improved accuracy by incorporating the compositional, physico-chemical and structural features into Chou's general PseAAC." *Scientific Reports,* **7**(1): 42362. DOI: https://doi.org/10.1038/srep42362.

Meng, H. (2026) "AI-driven discovery and design of antimicrobial peptides: progress, challenges, and opportunities." *Probiotics and Antimicrobial Proteins,* **18**(4): 5122-5144. DOI: https://doi.org/10.1007/s12602-025-10856-0.

Miller, W. R. and C. A. Arias, (2024). "ESKAPE pathogens: antimicrobial resistance, epidemiology, clinical impact and therapeutics." *Nature Reviews Microbiology,* **22**(10): 598-616. DOI: https://doi.org/10.1038/s41579-024-01054-w.

Mishra, A. K., Choi, J., Moon, E. and K. H. Baek, (2018). "Tryptophan-rich and proline-rich antimicrobial peptides." *Molecules,* **23**(4): 815. DOI: https://doi.org/10.3390/molecules23040815.

Nguyen, Q. H., Nguyen-Vo, T.-H., Do, T. T. T. and B. P. Nguyen, (2024). "An efficient hybrid deep learning architecture for predicting short antimicrobial peptides." *Proteomics,* **24**(14): 2300382. DOI: https://doi.org/10.1002/pmic.202300382.

Pinacho-Castellanos, S. A., García-Jacas, C. s. R., Gilson, M. K. and C. A. Brizuela, (2021). "Alignment-free antimicrobial peptide predictors: improving performance by a thorough analysis of the largest available data set." *Journal of Chemical Information and Modeling,* **61**(6): 3141-3157. DOI: https://doi.org/10.1021/acs.jcim.1c00251.

Pirtskhalava, M., Amstrong, A. A., Grigolava, M., Chubinidze, M., Alimbarashvili, E., Vishnepolsky, B., Gabrielian, A., Rosenthal, A., Hurt, D. E. and M. Tartakovsky, (2021). "DBAASP v3: database of antimicrobial/cytotoxic activity and structure of peptides as a resource for development of new therapeutics." *Nucleic Acids Research,* **49**(D1): D288-D297. DOI: https://doi.org/10.1093/nar/gkaa991.

Pirtskhalava, M., Gabrielian, A., Cruz, P., Griggs, H. L., Squires, R. B., Hurt, D. E., Grigolava, M., Chubinidze, M., Gogoladze, G., Vishnepolsky, B., Alekseev, V., Rosenthal, A. and M. Tartakovsky, (2016). "DBAASP v.2: an enhanced database of structure and antimicrobial/cytotoxic activity of natural and synthetic peptides." *Nucleic Acids Research,* **44**(D1): D1104-D1112. DOI: https://doi.org/10.1093/nar/gkv1174.

Saleem, N., Kumar, N., El-Omar, E., Willcox, M. and X.-T. Jiang, (2025). "Harnessing machine learning approaches for the identification, characterization, and optimization of novel antimicrobial peptides." *Antibiotics*, **14**(12): 1263. DOI: https://doi.org/10.3390/antibiotics14121263.

Santos-Júnior, C. D., Pan, S., Zhao, X. M. and L. P. Coelho, (2020). "Macrel: antimicrobial peptide screening in genomes and metagenomes." *PeerJ: Bioinformatics and Genomics,* **8**: e10555. DOI: http://dx.doi.org/10.7717/peerj.10555.

Santos-Júnior, C. D., Torres, M. D. T., Duan, Y., Rodríguez del Río, Á., Schmidt, T. S. B., Chong, H., Fullam, A., Kuhn, M., Zhu, C., Houseman, A., Somborski, J., Vines, A., Zhao, X.-M., Bork, P., Huerta-Cepas, J., de la Fuente-Nunez, C. and L. P. Coelho, (2024). "Discovery of antimicrobial peptides in the global microbiome with machine learning." *Cell,* **187**(14): 3761-3778.e16. DOI: https://doi.org/10.1016/j.cell.2024.05.013.

 This open-access article is distributed under the terms of the Creative Commons Attribution Non Commercial 4.0 License (CC BY-NC 4.0).

Satei, P., Ghaznavi-Rad, E., Fahimirad, S. and H. Abtahi, (2021). ″Recombinant production of trx-ib-amp4 and trx-e50-52 antimicrobial peptides and antimicrobial synergistic assessment on the treatment of methicillin-resistant *staphylococcus aureus* under *in vitro* and *in vivo* situations.″ *Protein Expression and Purification*, **188**: 105949. DOI: https://doi.org/10.1016/j.pep.2021.105949.

Sharma, R., Shrivastava, S., Kumar Singh, S., Kumar, A., Saxena, S. and Kumar Singh, R. (2021). ″Deep-ABPpred: identifying antibacterial peptides in protein sequences using bidirectional LSTM with word2vec.″ *Briefings in Bioinformatics,* **22**(5): bbab065. DOI: https://doi.org/10.1093/bib/bbab065.

Singh, V., Shrivastava, S., Kumar Singh, S., Kumar, A. and S. Saxena, (2022). ″StaBle-ABPpred: a stacked ensemble predictor based on biLSTM and attention mechanism for accelerated discovery of antibacterial peptides.″ *Briefings in Bioinformatics,* **23**(1): bbab439. DOI: https://doi.org/10.1093/bib/bbab439.

Sun, Z., Wang, X., Kuang, S., Song, J. and C. de la Fuente-Nunez, (2026). ″From generation to validation: Deep generative models for antimicrobial peptide discovery.″ *Current Opinion in Chemical Biology,* **93**: 102685. DOI: https://doi.org/10.1016/j.cbpa.2026.102685.

Szymczak, P. and E. Szczurek, (2023). ″Artificial intelligence-driven antimicrobial peptide discovery.″ *Current Opinion in Structural Biology,* **83**: 102733. DOI: https://doi.org/10.1016/j.sbi.2023.102733.

Szymczak, P., Zarzecki, W., Wang, J., Duan, Y., Wang, J., Coelho, L. P., de la Fuente-Nunez, C. and E. Szczurek, (2025). ″AI-driven antimicrobial peptide discovery: mining and generation.″ *Accounts of Chemical Research,* **58**(12): 1831-1846. DOI: https://doi.org/10.1021/acs.accounts.0c00594.

Tabarzad, M., Torshabi, M., Haeri, A., Fathi, F. and S. M. Mortazavi, (2026) ″Peptide-based antibiotics: structure-driven strategies to tackle toxicity and resistance of antimicrobial peptides.″ *Bioorganic & Medicinal Chemistry,* **133**: 118486. DOI: https://doi.org/10.1016/j.bmc.2025.118486.

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

Timmons, P. B. and C. M. Hewage, (2021a). ″ENNAACT is a novel tool which employs neural networks for anticancer activity classification for therapeutic peptides.″ *Biomedicine and Pharmacotherapy,* **133**: 111051. DOI: https://doi.org/10.1016/j.biopha.2020.111051.

Timmons, P. B. and C. M. Hewage, (2021b). ″ENNAVIA is a novel method which employs neural networks for antiviral and anti-coronavirus activity prediction for therapeutic peptides.″ *Briefings in Bioinformatics,* **22**(6): bbab258. DOI: https://doi.org/10.1093/bib/bbab258.

Van Oort, C. M., Ferrell, J. B., Remington, J. M., Wshah, S. and J. Li, (2021). ″AMPGAN v2: Machine learning-guided design of antimicrobial peptides.″ *Journal of Chemical Information and Modeling,* **61**(5): 2198-2207. DOI: https://doi.org/10.1021/acs.jcim.0c01441.

Vishnepolsky, B., Grigolava, M., Managadze, G., Gabrielian, A., Rosenthal, A., Hurt, D. E., Tartakovsky, M. and M. Pirtskhalava, (2022) ″Comparative analysis of machine learning algorithms on the microbial strain-specific AMP prediction.″ *Briefings in Bioinformatics,* 23(4): bbac233. DOI: https://doi.org/10.1093/bib/bbac233.

Vishnepolsky, B. and M. Pirtskhalava, (2019). ″Comment on:‘Empirical comparison of web-based antimicrobial peptide prediction tools’.″ *Bioinformatics,* **35**(15): 2692-2694. DOI: https://doi.org/10.1093/bioinformatics/bty1023.

Waghu, F. H., Barai, R. S., Gurung, P. and S. Idicula-Thomas, (2016). ″$CAMP_{R3}$: a database on sequences, structures and signatures of antimicrobial peptides.″ *Nucleic Acids Research,* **44**(D1): D1094-D1097. DOI: https://doi.org/10.1093/nar/gkv1051.

Wan, F., Wong, F., Collins, J. J. and C. de la Fuente-Nunez, (2024). ″Machine learning for antimicrobial peptide identification and design.″ *Nature Reviews Bioengineering,* **2**(5): 392-407. DOI: https://doi.org/10.1038/s44222-024-00152-x.

Wang, G., Vaisman, I. I. and M. L. van Hoek, (2022). ″Machine learning prediction of antimicrobial peptides.″ In Simonson, T., (ed.) *Computational Peptide Science: Methods and Protocols*, New York, NY: Springer US. pp. 1-37. DOI: https://doi.org/10.1007/978-1-0716-1855-4_1.

Wang, Y., Zhao, L., Li, Z., Xi, Y., Pan, Y., Zhao, G. and L. Zhang, (2025). ″A generative artificial intelligence approach for the discovery of antimicrobial peptides against multidrug-resistant bacteria.″ *Nature Microbiology,* **10**(11): 2997-3012. DOI: https://doi.org/10.1038/s41564-025-02114-4.

Xiao, X., Wang, P., Lin, W.-Z., Jia, J.-H. and K. C. Chou, (2013) ″iAMP-2L: A two-level multi-label classifier for identifying antimicrobial peptides and their functional types.″ *Analytical Biochemistry,* **436**(2): 168-177. DOI: https://doi.org/10.1016/j.ab.2013.01.019.

Xu, J., Li, F., Li, C., Guo, X., Landersdorfer, C., Shen, H.-H., Peleg, A. Y., Li, J., Imoto, S., Yao, J., Akutsu, T. and Song, J. (2023). ″iAMPCN: a deep-learning approach for identifying antimicrobial peptides and their functional activities.″ *Briefings in Bioinformatics,* **24**(4): bbad240. DOI: https://doi.org/10.1093/bib/bbad240.

Yan, J., Cai, J., Zhang, B., Wang, Y., Wong, D. F. and S. W. I. Siu, (2022). ″Recent progress in the discovery and design of antimicrobial peptides using traditional machine learning and deep learning.″ *Antibiotics*, **11**(10): 1451. DOI: https://doi.org/10.3390/antibiotics11101451.

Zhao, M., Zhang, Y., Wang, M. and L. Z. Ma, (2024). ″dsAMP and dsAMPGAN: Deep learning networks for antimicrobial peptides recognition and generation.″ *Antibiotics*, **13**(10): 948. DOI: https://doi.org/10.3390/antibiotics13100948.

Zhao, X., Wu, H., Lu, H., Li, G. and Q. Huang, (2013). ″LAMP: A .″ *PLOS ONE,* **8**(6): e66557. DOI: https://doi.org/10.1371/journal.pone.0066557.

This open-access article is distributed under the terms of the Creative Commons Attribution Non Commercial 4.0 License (CC BY-NC 4.0).