Artificial Intelligence Evolution
http://ojs.wiserpub.com/index.php/AIE/

UNIVERSAL WISER PUBLISHER

Review

Artificial Intelligence Tools Addressing Challenges of Cancer Progression Due to Antimicrobial Resistance in Pathogenic Biofilm Systems

Vinod K. Mishra[1], Abhijit G. Banerjee[2*]

[1]JBIT College of Applied Sciences, Dehradun, Uttarakhand, India
[2]Genomic Bio-Medicine Research and Incubation, Chhattisgarh, Durg, 491001, India
E-mail: abhijitb@cgbmri.co.in

Received: 22 August 2024; **Revised:** 17 December 2024; **Accepted:** 27 December 2024

Abstract: Infections, inflammation, and progression of multifactorial diseases are found to be integratively linked, including most Cancers. Dysfunctional microbiomes are also associated with several cancers in their tumor microenvironments. Antimicrobial peptides (AMPs) are short, positively charged peptides found in a diverse range of species, including bacteria and humans. As host defense peptides, they can destroy pathogenic infections, particularly those that are multidrug resistant. AMPs have raised hopes in the biomedical and pharmaceutical industries as fresh non-antibiotic strategies for combating infectious diseases. However, *in vitro* and *in vivo* verification of AMPs is problematic and may miss new antimicrobial drugs. Creating computational methods for quick and precise identification of AMPs and their functional forms is critical for developing new and more effective antimicrobial drugs. Machine learning techniques were recently discovered effective at mining, predicting, and producing efficient antimicrobial peptides from a large AMP database. We reviewed 76 articles, after following literature search rubrics to come to the following conclusions. Distance metric-constant K-based nearest neighbor algorithms (KNN), hidden Markov models (HMMs), support vector machine models (SVMs), random forest models (RFs), decision tree models, and deep neural network (DNN)-based models are some of the most popular AI tools for detecting antimicrobial activity in peptide sequence-derived structure and function. Knowledge graphs can further assist in identifying hub genes and antimicrobial peptides that target and block quorum sensing (QS) signals within the microbial networks. In conclusion, we state that currently no single AI method has been found appropriate for AMP discovery and accurately capable of predicting high-efficacy AMPs. Our current literature review and analysis identify cutting-edge algorithms or innovations that might be included in hybrid machine-learning approaches for the most effective AMP identification, creation, and prediction. Non-peptide, natural molecule-based approaches to AMR reduction are also being studied for development, with natural peptide scaffolds serving as the foundation.

Keywords: antimicrobial peptide, biofilms, cancer progression, machine learning, quorum sensing inhibitors, systems biology

1. Introduction

In nature, bacteria predominantly alternate between two forms: unicellular, planktonic cells free-floating in a liquid environment and sessile, multicellular communities called biofilms that form on solid surfaces [1]. Antimicrobial

DOI: https://doi.org/10.37256/aie.6120255553

resistance (AMR) has emerged as a global health concern as antibiotic resistance genes (ARGs) and antibiotic-resistant bacteria (ARBs) is spreading throughout human, animals, and the environment. Despite the upsurge of several new antibacterial agents, their effectiveness is diminishing due to rising antibiotic resistance, rendering these treatments increasingly ineffective. Bacteria that form biofilms exhibit significantly greater resistance to antibiotics. Biofilms are protected by self-produced extracellular matrix polysaccharides (EPS), proteins, extracellular DNA, and lipids. This sticky scaffold enhances cell-cell communication through quorum sensing (QS), binds the community to various surfaces, stores extracellular enzymes and metabolites, and provides a physical barrier that shields cells from environmental hazards [2-5]. The EPS matrix not only provides structural integrity but also plays a crucial role in defending bacteria against antibiotic action. Biofilm-forming bacteria have developed several mechanisms to protect themselves, including restricting antibiotic penetration at the biofilm surface, where the EPS acts as a protective shield.

Additionally, the altered microenvironment within the biofilm, characterized by metabolic changes, reduced oxygen levels, and pH variations, further contributes to antibiotic resistance [6, 7]. Biofilm infections are usually chronic in nature because biofilm-residing bacteria can withstand the immune system, antibiotics, and other treatments [8, 9]. However, bacteria that live in biofilms are more resistant to antibiotics than free-floating (planktonic) bacteria of the same species [10]. The majority of new drugs have been mere modifications, lacking the diversity crucially needed to combat the alarming rise of AMR [11]. The acute dearth of new antibiotics and the frightening growth in resistance to even the most potent existing ones have prompted an urgent need to research new possible antimicrobial agents and/or develop novel biochemical entities to tackle the challenges of antimicrobial resistance. For instance, antibiotics like Tobramycin and Ciprofloxacin rely on oxygen-dependent mechanisms to exert their bactericidal effects, making them less effective in the low-oxygen conditions often found within biofilms. Low oxygen levels can reduce the production of reactive oxygen species (ROS), which are crucial for the efficacy of these antibiotics. Additionally, the activity of aminoglycosides, such as Tobramycin, is highly pH-dependent. In acidic environments, the uptake of aminoglycosides by bacterial cells is reduced, leading to decreased antibiotic efficacy.

Inside biofilms, persister cells represent a small subpopulation of bacteria that can survive antibiotic treatment, further complicating eradication efforts. The biofilm environment promotes horizontal gene transfer, spreading resistance genes among the bacterial community. These mechanisms, along with the unique microenvironment of biofilms, contribute to the heightened antibiotic resistance observed in biofilm-associated infections. Quorum sensing (QS) is a cell-to-cell communication process. The autoinducers produced in QS regulate gene expression based on population density. They play a crucial role in biofilm. This biofilm environment enhances bacterial survival and contributes to antibiotic resistance.

Biofilms and multidrug resistance (MDR) are closely linked, creating significant challenges in treating bacterial infections. Biofilm-associated infections are often chronic and difficult to treat, leading to prolonged illness and increased healthcare costs. Common biofilm-related infections are those associated with medical devices like catheters and implants. Several novel approaches, including AMR and quorum-sensing inhibitors (QI), are being explored to combat biofilm infections. Nowadays, artificial intelligence tools are gaining popularity to address the challenges of antimicrobial resistance in pathogenic biofilm systems. This review article presents the progress made in utilizing artificial intelligence to tackle the challenges of antibiotic resistance in pathogenic biofilms. In our quest for the design of antimicrobial or antibiofilm peptides literature review, a total of 182 full-text articles were initially retrieved through a literature search in medical databases using boolean operators of keywords mentioned. Further nine studies were included through a targeted search of literature. After assessing the eligibility meeting inclusion and exclusion criteria, 76 articles were reviewed for this report, finally. Salient features of this review are presented in the sections below.

1.1 *Antimicrobial peptides: promising antibiotic alternatives*

AMPs are tiny peptides that play an important role in the host's innate defense against a wide range of pathogens, including bacteria, fungi, parasites, and viruses [1-4]. The rise of MDR bacterial pathogens, collectively known as 'ESKAPE' pathogens after their pioneering members such as *Enterococcus faecium*, *Staphylococcus aureus*, *Klebsiella pneumoniae*, *Pseudomonas aeruginosa*, *Acinetobacter baumannii*, and various *Enterobacter spp*., have posed a significant challenge to human health. There is an urgent need for novel antimicrobial strategies to combat these recalcitrant infections [12-14]. AMPs thus have found new hope, as a novel non-antibiotic method for fighting MDR pathogens because of their broad-spectrum activity, multimodal capabilities, and seldom resistance development

[15]. These are produced naturally in almost all forms of life for defense and communication. These biological compounds are polypeptide sequences, typically 12-50 residues in length. They are considered interesting substitutes for antibiotics because they can function in various of ways, such as modulating the host immune system, acting against pathogenic species such as viruses, inhibiting bacterial growth, and causing physical or metabolic disruption of cells. Most antibiotics lack targeting specificity and can kill host-associated commensal bacteria as well as disease-causing pathogens [16-19]. Bacteria and fungi also contain AMPs, which function as defense mechanisms against other microbes or host immunological responses. They have appealing properties as prospective therapeutic agents, including a decreased risk of generating bacterial resistance than standard antibiotics.

1.2 *Biofilms as driver of microbial pathogenesis*

Biofilms present a formidable barrier to antimicrobial approaches towards pathogenesis control measures and concomitant diseases. Biofilms generally build upon commensal bacteria slowly taken over by dominant pathogens and barrier functions represented by extracellular capsular material built up with EPS secretion; by biofilm conglomeration of microbes [20]. Once biofilms develop, it's very difficult to eradicate them completely, as they keep growing back as soon as favorable conditions return within the ecological niche.

1.3 *Anti-biofilm peptides (ABPs)*

AMPs have been utilized to control the formation and removal of mature biofilms. AMPs, both naturally occurring, and synthetic ones have been proven to prevent microbial colonization of surfaces, mitigate target bacterial populations in biofilms, and alter biofilm structures [21, 22]. Therefore, ABPs are considered a subgroup of AMPs that inhibit biofilm formation and eliminate previously established biofilms.

1.4 *Destabilizing outer membrane permeability of prokaryotes*

Most antibiotics work inside the cell and must pass through the bacterial cell membrane. Antibiotics can diffuse through the outer membrane via a lipid-mediated channel for hydrophobic antibiotics and through general diffusion pores for hydrophilic antibiotics. Gram-negative bacteria are protected by a formidable barrier owing to their outer membrane. The asymmetric outer membrane, composed of lipopolysaccharide on the outer leaflet, poses a significant challenge for the permeation of antibiotics and other therapeutics into the cell, and phospholipid on the inner leaflet resists penetration of organic antimicrobials [23, 24]. As a result of their complicated architecture, most antibacterial medicines have lower efficacy in treating gram-negative infections. Antimicrobial drugs with higher membrane permeability are an effective technique for enhancing antibiotic penetration into bacterial cells. Permeabilizers, such as AMPs, are cationic and amphiphilic compounds that interact with polyanionic lipopolysaccharides to break down membrane walls [25, 26].

1.5 *Non-peptide, biofilm, and quorum sensing disruptors in mitigation of AMR*

The disulfide and trisulfide metabolites derived from garlic can suppress LuxR-based QS, in *Pseudomonas aeruginosa* [27]. Rosmarinic acid has been reported to activate quorum sensing-dependent gene expression in *P. aeruginosa*, increase biofilm formation, and the synthesis of the virulence proteins pyocyanin and elastase. Conversely, *Mangifera indica L*. leaf extracts have been shown to prevent the quorum-sensing-regulated synthesis of virulence factors and thereby inhibit biofilm formation in test bacteria [28, 29]. Thus, quorum-quenching (Qq) happens when medicinal plant-derived substances disrupt pathways by interfering with signal molecule formation, inactivating signals, and interfering with signal receptors in bacterial cells, thus blocking target genes under QS regulation.

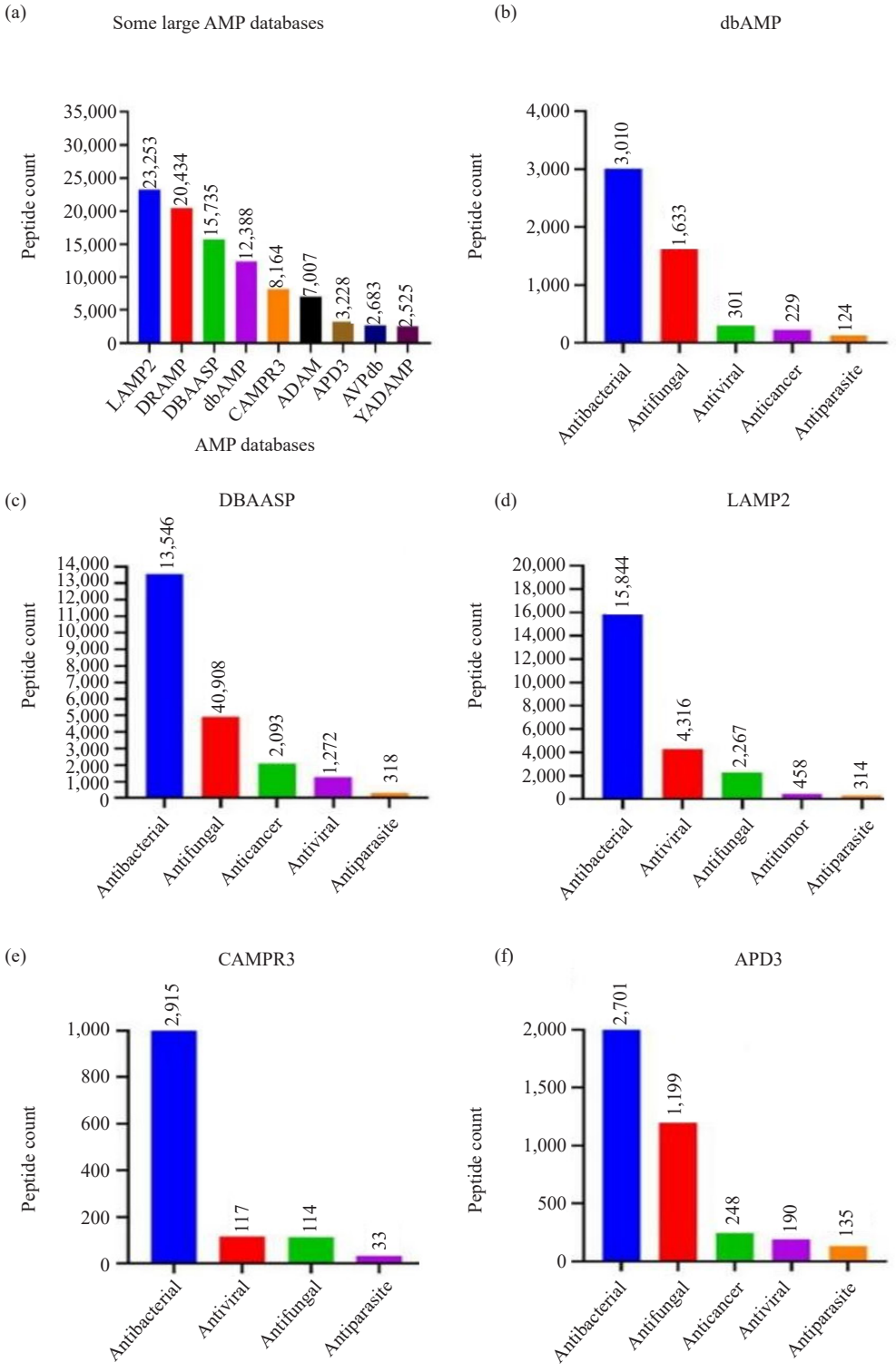

Figure 1. Graphical abstract: Major AMP databases and AI-based prediction functionality
Several AMP-related databases have been developed over the past decade. E.g., LAMP: A database linking major antimicrobial peptides (depicted within squared panel 'a')

2. Computational strategies to elucidate anti-biofilm peptide characteristics

2.1 *Discovery of novel AMPs*

Artificial intelligence (AI) is the study and creation of intelligent machines that can learn, solve problems, and replicate various forms of reasoning similar to natural intelligence. Machine learning (ML) is a subset of artificial intelligence (AI) algorithms that trains mathematical models to predict solutions when supplied with previously unknown data [29, 30]. Notably, AI and ML have the potential to open up new opportunities in healthcare and drive applications for faster, cheaper, and more effective drug discovery and development. Studies have shown potential in employing deep learning, a type of machine learning based on neural networks, to model amino acid sequences and predict antimicrobial activity. These computational methods enable high-throughput virtual screening to uncover novel AMP candidates for experimental validation. By utilizing huge datasets and predictive algorithms, AI and ML can accelerate the development of novel antimicrobial medicines to address the growing public health concern of antibiotic resistance. AI and machine learning can thus accelerate the development of novel antimicrobial medicines to address the growing public health concern of antibiotic resistance [31].

2.2 *AMP databases*

Many databases focusing on AMPs have been built to collect both basic and pharmacological information, as the discovery of new AMPs has increased. The nine AMP databases in terms of peptide count are as follows: LAMP2 [32, 33], DBAASP [34], dbAMP [35], CAMPR3 [35, 36], ADAM [37], APD3 [38], AVPdb [39], and YADAMP [40], six of such databases are illustrated in detail (Figure 1). Antibacterial peptides make up most of the AMPs discovered to date. Several major databases and their important prediction functionality tools are listed in Table 1.

Table 1. Various types of features employed for AMP prediction

Types	Feature extraction parameters
Composition feature	• AAC; Normalized amino acid composition (NAAC); N-gram composition found by t-test (NTC), and pseudo-amino acid composition. • AAPC; (types-AAPC, DPC, TPC), Peptide length and composition. • N-gram composition found by counting (NCC), motif composition (MC).
Position features	• Depends on the location of the bioactive feature (amino acid) within the sequence.
Structural features	• Depends on the amino acid types adjacent to the featured amino acids.
Physicochemical properties	• Isoelectric point and charge.

3. Artificial intelligence and machine learning-based AMP prediction algorithms

A new approach to the design of AMPs was developed by combining an evolutionary algorithm and machine learning-based prediction, followed by *in vitro* bacterial assays. Integrating genetic algorithms, machine learning, and high-throughput screening presents a promising approach for artificial intelligence and machine learning prediction algorithms, effectively searching sequence space and optimizing AMP candidates [41]. This method rapidly improved the antimicrobial activity, achieving a 162-fold increase in activity compared to the original peptide within three generative rounds. In addition to the best peptide, 44 new peptides were highly potent, with a 20-fold decrease in IC50 values compared with the seed WT peptide. During these experiments, the conformation of the selected peptides was observed to change from a random coil to an α-helical form through the optimization process, and this is thought to contribute to the improvement of antimicrobial activity significantly.

3.1 *Conventional machine learning-based AMP predictors*

Identifying antimicrobial peptides (AMPs) from sequence data using machine learning techniques involving the main steps:

(1) Amino acid sequence: The process starts with deriving the amino acid sequence of the peptide.

(2) Feature extraction: The amino acid sequence is converted into a numerical descriptor through various feature extraction methods, resulting in an initial high-dimensional feature set.

(3) Dimensionality reduction: To handle the high-dimensional initial feature set obtained from the peptide sequence, the following approaches are employed: Feature Clustering, where techniques like K-means clustering, Hierarchical clustering, Mean shift, Density-based spatial clustering of applications with noise (DBSCAN), and affinity propagation are used to group similar features together; Feature Selection, where methods such as Chi-square test, Information gain, Mutual information, and Pearson's correlation coefficient are used to identify and select the most relevant features; and then Dimensionality Reduction is implemented, where algorithms like Principal Component Analysis (PCA), Latent Dirichlet Allocation (LDA), and t-distributed Stochastic Neighbor Embedding (t-SNE) are used to project the high-dimensional feature space onto a lower-dimensional subspace. The dimensionality reduction techniques produce a reduced feature vector, which captures the most relevant information from the initial high-dimensional feature set. The reduced feature vector is fed into a machine-learning model, such as support vector machines (SVMs) or random forests (RFs). The machine learning model is trained to predict whether a given peptide sequence is an antimicrobial peptide (AMP) or not, based on the reduced feature vector.

(A) Sequence-based approaches: Feature engineering techniques for representing peptide sequences have garnered significant attention in the field of antibiofilm peptide prediction. A commonly employed approach involves representing the sequence using the frequency or composition of individual amino acids [42]. While this simple representation provides a basic characterization of the sequence, extensions to capture more complex patterns have been explored. These include considering the frequency of dipeptides (adjacent amino acid pairs) or constructing and selecting complex sequence-based features [42, 43]. In addition to sequence-based features, the physicochemical properties of amino acids have been widely used to derive informative features for machine learning models. Properties such as hydrophobicity, charge, and polarity can provide valuable insights into the peptide's behavior and potential interactions [44]. Various scales and indices have been used to convert these qualities into numerical features that can be included in the feature vectors. Aside from these basic representations, academics have investigated more advanced feature engineering strategies to harness additional sources of information. Furthermore, knowledge-based features derived from existing databases or literature, such as known motifs or sequence patterns associated with specific activities, have been integrated into feature vectors to enhance the predictive power of machine learning models [45-52] and presented in Table 2. For instance, structural information such as predicted secondary structure elements or solvent accessibility profiles are incorporated as features in some studies [53]. The choice of feature representation plays a crucial role in the performance of machine learning models for antibiofilm peptide prediction. As a result, careful assessment of the characteristics and their ability to gather essential information is required for the development of accurate and reliable models. Researchers have experimented with various combinations of sequence-based, physicochemical, structural, and knowledge-based characteristics, hoping to use complementary sources of information and improve their models' prediction powers.

(B) Structure-based machine learning models: Specialized machine learning models have been developed to leverage structural information and spatial relationships within peptides. (e.g., graph neural networks, 3D convolutional neural networks, etc.

▪ Graph neural networks (GNNs): GNNs can represent peptides as graphs, where nodes represent amino acids and edges represent their spatial or sequential relationships. These models can capture complex structural patterns and have been applied to antibiofilm peptide prediction tasks [54].

▪ 3D convolutional neural networks (3D CNNs): 3D CNNs can directly operate on 3D structural representations of peptides, capturing spatial and conformational information [55]. These models can be used for predicting antibiofilm activity.

3.2 *Machine learning models for sequence-based prediction*

Various machine learning models have been employed for sequence-based prediction of antibiofilm peptides, as detailed below in Table 2.

Table 2. Discriminative ML algorithms used in AMP identification and prediction

Model	Description	Reference
Support vector machines (SVMs)	An integrated algorithm to predict AMPs by integrating sequence alignment and support vector machine.	[56-58]
Logistic regression	A statistical model that predicts the probability of a binary outcome.	[59, 60]
Hidden markov models (HMMs)	A statistical model that represents the probabilities of sequences of observed events.	[37]
Random forests	An ensemble learning method that uses multiple decision trees to improve accuracy.	[37, 42, 49]
K-nearest neighbors (KNNs)	A classification based on the majority vote of neighbors.	[61]
Neural networks	Computational models inspired by the human brain, capable of learning complex patterns.	[62, 63]
Decision trees	A model that uses a tree-like graph of decisions and their possible consequences.	[58, 64]
Naive bayes	A probabilistic classifier based on Bayes' theorem with strong independence assumptions.	[65]

3.3 *Elaborating on structure-based approaches*

Techniques for predicting and incorporating structural features (e.g., secondary structure, solvent accessibility, physicochemical properties). In addition to sequence-based features, structural information can provide valuable insights into the function and activity of peptides. Several techniques have been employed to predict and incorporate structural features, as detailed below:

▪ Secondary structure prediction: Various computational methods, such as PSIPRED [52] and SPIDER3 [45], have been used to predict the secondary structure elements (e.g., α-helices, β-sheets) of peptides, which can be used as input features for machine learning models.

▪ Solvent accessibility prediction: Tools like ACCpro [53] and SPIDER3 [45] can predict the solvent accessibility of amino acid residues, providing information about the exposed or buried regions of the peptide.

▪ Physicochemical property calculations: Various software packages, such as EMBOSS [66] and GROMACS [67] can be used to calculate physicochemical properties like charge distribution, hydrophobicity, and molecular surface area, which can be used as input features.

3.4 *Integration of molecular dynamics simulations and machine learning*

▪ Molecular dynamics (MD) simulations can provide valuable insights into the dynamic behavior and conformational changes of peptides. Several studies have integrated MD simulations with machine learning techniques to improve the prediction of antibiofilm activity and other peptide properties [67-69].

▪ Integrated MD-ML pipelines: Iterative pipelines that combine MD simulations with machine learning models have been developed, allowing for the refinement of predicted structures and the improvement of activity predictions [70].

▪ Hybrid and ensemble approaches: These approaches combine sequence-based and structure-based features. To leverage the complementary information provided by sequence-based and structure-based features, several studies have explored hybrid approaches that combine these feature types.

▪ Concatenation of feature vectors: feature vectors representing sequence-based and structure-based information can be concatenated and used as input for machine learning models [62].

3.5 *Ensemble learning techniques for integrating multiple models*

Ensemble learning techniques have been employed to combine the predictions of multiple models, potentially improving overall performance and robustness:

▪ Stacking: In stacking, the predictions of individual models (e.g., SVMs, random forests, neural networks) are used as input features for a meta-learner, which combines them to produce the final prediction.

▪ Voting: In voting ensembles, the predictions of multiple models are combined using majority voting (for classification tasks) or averaging (for regression tasks).

▪ Boosting: Techniques like AdaBoost and Gradient Boosting can iteratively combine weak models to create a strong ensemble predictor.

▪ Hybrid architectures (e.g., deep neural networks (DNN) with machine learning models): Hybrid architectures that integrate deep neural networks with traditional machine learning models have been explored for antibiofilm peptide prediction.

▪ Multi-task learning: Neural networks can be trained to simultaneously predict multiple properties in a multi-task learning framework, leveraging shared representations [71].

3.6 *Generative models and De Novo design*

Generative adversarial networks (GANs) have been explored for the de novo design and generation of antibiofilm peptides.

• Sequence generation: GANs can be trained to generate new peptide sequences with desired properties, such as increased antibiofilm activity or improved stability [72].

• Conditional generation: Conditional GANs can generate peptide sequences conditioned on specific properties or constraints, allowing for targeted design of antibiofilm peptides [73].

• Reinforcement learning for peptide optimization: Reinforcement learning techniques have been applied to optimize peptide sequences for improved antibiofilm activity or other desired properties.

• Sequence optimization: Reinforcement learning agents can iteratively modify peptide sequences, guided by a reward function that evaluates the desired properties.

• Constraint-based optimization: Reinforcement learning can be combined with constraint-based optimization techniques to generate peptide sequences that satisfy specific design constraints.

4. Artificial intelligence and machine learning-based predictions

Biofilms' clinical relevance and resistance to standard antibiotics necessitate the aggressive pursuit of alternative therapeutics. A web service called dPABBs was developed to make it easier to anticipate and generate anti-biofilm peptides [74]. Based on the residues' positional preference, selected residue features, and overall amino acid composition, the six SVM and WEKA models applied to dPABBs were discovered to be capable of identifying anti-biofilm peptides, with maximum accuracy (95.24%), sensitivity (92.50%), specificity (97.73%), and MCC of 0.91, on the training datasets. In the case of the antibiofilm peptides, it was observed that the cationic residue R or K is present at all five places on the N-terminus, while in the case of the QS peptides, uncharged polar residue S is in the 1st, 3rd, or 5th position, and anionically charged residues D and E are in the first position. Indeed, in 2014, Dziuba and Dziuba successfully implemented bioinformatics strategies to design effective AMPs from milk proteins [75]. A collection of biological sequences is placed in the database- CAMPR4 and the features used for extraction are presented in Table 3 [76].

Table 3. Database of biological sequences converted into suitable features for model building, followed by prediction of novel sequences [76]

Databases	Characteristics	ML Algorithms
Collection of anti-microbial peptides (CAMPR4)	Sequence-activity and specificity relationships of AMPs	Artificial neural network (ANN), support vector machines and N random forest models

5. Conclusions

The rise in multidrug-resistant bacteria, particularly those that form protective biofilms, poses a severe danger to global health. Conventional antibiotics are losing their potency against these illnesses, thus innovative therapeutic techniques are badly needed. Antimicrobial peptides and antibiofilm peptides have broad-spectrum activity against a number of infections in vitro, making them both interesting replacements. Nonetheless, classic AMP or ABP discovery strategies have limitations, such as a reliance on time-consuming activity assays, that make it difficult to identify AMPs with different processes or specific target sites.

The advent of ML and AI in recent decades has transformed drug discovery. These powerful computational tools offer a breakthrough approach to developing new antibiotics, such as AMPs and ABPs. To improve AMP and ABP predictions, researchers are actively attempting to solve the limits of current AI and ML algorithms utilized in AMP and ABP discoveries. They also hope to close the gap between in vitro and in vivo data, which is necessary for practice specificities. Finally, they hope to overcome the limits of present approaches and fully realize the potential of AMPs and ABPs by augmenting datasets with a broader range of sequences, structures, and target specificities.

Future research should include in vivo data and investigate relevant biological situations for AMPs and ABPs. Furthermore, AI and machine learning models can be trained to predict and minimize potential toxicity during the design phase, resulting in safe and effective AMPs or ABPs. Furthermore, AI and machine learning models can be trained to predict and minimize potential toxicity during the design phase, resulting in safe and effective AMPs or ABPs.

Author contributions

AGB: Conceptualization, strategic supervision, Writing: review and edit the final manuscript. VKM: Writing: original draft, carried out preliminary investigations and writing: first draft, and edited the manuscript. Both authors agree to the final form of the manuscript being submitted for publication.

Acknowledgements

The editorial assistance of Ms. Anushka, Scientific Development and Communications Officer, from NIMHANS (National Institute of Mental Health and Neurosciences), is gratefully acknowledged. The authors are indebted to Dr. Dinesh Kumar (CABin, Indian Agricultural Research Institute, New Delhi) and Dr. V. K. Gopalakrishnan, (CGPH, Saveetha Medical Institute) for critical insights into the bioinformatics tools and biological resources for antimicrobial compounds discovery.

Funding

No funding was received for the study which may construe bias in the reports.

Conflict of interest

No conflicts of interest exist as declared by the authors of this study.

References

[1] Grooters KE, Ku JC, Richter DM, Krinock MJ, Minor A, Li P, et al. Strategies for combating antibiotic resistance in bacterial biofilms. *Frontiers in Cellular and Infection Microbiology*. 2024; 14: 1352273. Available from: https://doi.org/10.3389/fcimb.2024.1352273.

[2] Berlanga M, Guerrero R. Living together in biofilms: the microbial cell factory and its biotechnological implications. *Microbial Cell Factories*. 2016; 15: 1-11. Available from: https://doi.org/10.1186/s12934-016-0569-5.

[3] Limoli DH, Jones CJ, Wozniak DJ. Bacterial extracellular polysaccharides in biofilm formation and function. *Microbiology Spectrum*. 2015; 3(3). Available from: https://doi.org/10.1128/microbiolspec.mb-0011-2014.

[4] Billings N, Ramirez Millan M, Caldara M, Rusconi R, Tarasova Y, Stocker R, et al. The extracellular matrix component psl provides fast-acting antibiotic defense in pseudomonas aeruginosa biofilms. *PLoS Pathogens*. 2013; 9(8): e1003526. Available from: https://doi.org/10.1371/journal.ppat.1003526.

[5] Colvin KM, Irie Y, Tart CS, Urbano R, Whitney JC, Ryder C, et al. The Pel and Psl polysaccharides provide Pseudomonas aeruginosa structural redundancy within the biofilm matrix. Environmental Microbiology. 2012; 14(8): 1913-1928. Available from: https://doi.org/10.1111/j.1462-2920.2011.02657.x.

[6] Olsen I. Biofilm-specific antibiotic tolerance and resistance. *European Journal of Clinical Microbiology & Infectious Diseases*. 2015; 34: 877-886. Available from: https://doi.org/10.1007/s10096-015-2323-z.

[7] Lebeaux D, Ghigo JM, Beloin C. Biofilm-related infections: bridging the gap between clinical management and fundamental aspects of recalcitrance toward antibiotics. *Microbiology and Molecular Biology Reviews*. 2014; 78(3): 510-543. Available from: https://doi.org/10.1128/MMBR.00013-14.

[8] Vestby LK, Grønseth T, Simm R, Nesse LL. Bacterial biofilm and its role in the pathogenesis of disease. *Antibiotics*. 2020; 9(2): 59. Available from: https://doi.org/10.3390/antibiotics9020059.

[9] Bjarnsholt T. The role of bacterial biofilms in chronic infections. *APMIS*. 2013; 121(s136): 1-58. Available from: https://doi.org/10.1111/apm.12099.

[10] Roy R, Tiwari M, Donelli G, Tiwari V. Strategies for combating bacterial biofilms: A focus on antibiofilm agents and their mechanisms of action. *Virulence*. 2018; 9(1): 522-554. Available from: https://doi.org/10.1080/21505594.2017.1313372.

[11] Levy S, Marshall B. Antibacterial resistance worldwide: causes, challenges and responses. *Nature Medicine*. 2004; 10(Suppl 12): S122-S129. Available from: https://doi.org/10.1038/nm1145.

[12] De Oliveira DMP, Forde BM, Kidd TJ, Harris PNA, Schembri MA, Beatson SA, et al. Antimicrobial resistance in ESKAPE pathogens. *Clinical Microbiology Reviews*. 2020; 33(3): e00181-19. Available from: https://doi.org/10.1128/cmr.00181-19.

[13] Wang J, Dou X, Song J, Lyu Y, Zhu X, Xu L, et al. Antimicrobial peptides: Promising alternatives in the post feeding antibiotic era. *Medicinal Research Reviews*. 2019; 39(3): 831-859. Available from: http://dx.doi.org/10.1002/med.21542.

[14] Huan Y, Kong Q, Mou H, Yi H. Antimicrobial peptides: classification, design, application and research progress in multiple fields. *Frontiers in Microbiology*. 2020; 11: 582779. Available from: https://doi.org/10.3389/fmicb.2020.582779.

[15] Drayton M, Kizhakkedathu JN, Straus SK. Towards robust delivery of antimicrobial peptides to combat bacterial resistance. *Molecules*. 2020; 25(13): 3048. Available from: https://doi.org/10.3390/molecules25133048.

[16] Van der Does AM, Hiemstra PS, Mookherjee N. Antimicrobial host defence peptides: Immunomodulatory functions and translational prospects. *Antimicrobial Peptides: Basics for Clinical Application*. 2019; 1117: 149-171. Available from: https://doi.org/10.1007/978-981-13-3588-4_10.

[17] Elnagdy S, AlKhazindar M. The potential of antimicrobial peptides as an antiviral therapy against COVID-19. *ACS Pharmacology & Translational Science*. 2020; 3(4): 780-782. Available from: https://doi.org/10.1021/acsptsci.0c00059.

[18] Cotter PD, Ross RP, Hill C. Bacteriocins-a viable alternative to antibiotics? *Nature Reviews Microbiology*. 2013; 11(2): 95-105. Available from: https://doi.org/10.1038/nrmicro2937.

[19] Lei M, Jayaraman A, Van Deventer JA, Lee K. Engineering selectively targeting antimicrobial peptides. *Annual Review of Biomedical Engineering*. 2021; 23(1): 339-357. Available from: https://doi.org/10.1146/annurev-bioeng-010220-095711.
[20] Di Somma A, Moretta A, Canè C, Cirillo A, Duilio A. Antimicrobial and antibiofilm peptides. *Biomolecules*. 2020; 10(4): 652. Available from: https://doi.org/10.3390/biom10040652.
[21] Bose B, Downey T, Ramasubramanian AK, Anastasiu DC. Identification of distinct characteristics of antibiofilm peptides and prospection of diverse sources for efficacious sequences. *Frontiers in Microbiology*. 2022; 12: 783284. Available from: https://doi.org/10.3389/fmicb.2021.783284.
[22] Yasir M, Willcox MDP, Dutta D. Action of antimicrobial peptides against bacterial biofilms. *Materials*. 2018; 11(12): 2468. Available from: https://doi.org/10.3390/ma11122468.
[23] Delcour AH. Outer membrane permeability and antibiotic resistance. *Biochimica et Biophysica Acta (BBA)-Proteins and Proteomics*. 2009; 1794(5): 808-816. Available from: https://doi.org/10.1016/j.bbapap.2008.11.005.
[24] Nikaido H. Molecular basis of bacterial outer membrane permeability revisited. *Microbiology and Molecular Biology Reviews*. 2003; 67(4): 593-656. Available from: https://doi.org/10.1128/MMBR.67.4.593-656.2003.
[25] Nuri R, Shprung T, Shai Y. Defensive remodeling: How bacterial surface properties and biofilm formation promote resistance to antimicrobial peptides. *Biochimica et Biophysica Acta (BBA)-Biomembranes*. 2015; 1848(11): 3089-3100. Available from: https://doi.org/10.1016/j.bbamem.2015.05.022.
[26] Brogden NK, Brogden KA. Will new generations of modified antimicrobial peptides improve their potential as pharmaceuticals? *International Journal of Antimicrobial Agents*. 2011; 38(3): 217-225. Available from: https://doi.org/10.1016/j.ijantimicag.2011.05.004.
[27] Khan MA, Celik I, Khan HM, Shahid M, Shahzad A, Kumar S, et al. Antibiofilm and anti-quorum sensing activity of *Psidium guajava* L. leaf extract: In vitro and in silico approach. *PLoS One*. 2023; 18(12): e0295524. Available from: https://doi.org/10.1371/journal.pone.0295524.
[28] Corral-Lugo A, Daddaoua A, Ortega A, Espinosa-Urgel M, Krell T. So different and still so similar: The plant compound rosmarinic acid mimics bacterial homoserine lactone quorum sensing signals. *Communicative & Integrative Biology*. 2016; 9(2): e1156832. Available from: https://doi.org/10.1080/19420889.2016.1156832.
[29] Husain FM, Ahmad I, Al-Thubiani AS, Abulreesh HH, AlHazza IM, Aqil F. Leaf extracts of *Mangifera indica* L. inhibit quorum sensing-regulated production of virulence factors and biofilm in test bacteria. *Frontiers in Microbiology*. 2017; 8: 727. Available from: https://doi.org/10.3389/fmicb.2017.00727.
[30] Melo MCR, Maasch JRMA, de la Fuente-Nunez C. Accelerating antibiotic discovery through artificial intelligence. *Communications Biology*. 2021; 4(1): 1050. Available from: https://doi.org/10.1038/s42003-021-02586-0.
[31] Cherkasov A, Hilpert K, Jenssen H, Fjell CD, Waldbrook M, Mullaly SC, et al. Use of artificial intelligence in the design of small peptide antibiotics effective against a broad spectrum of highly antibiotic-resistant superbugs. *ACS Chemical Biology*. 2009; 4(1): 65-74. Available from: https://doi.org/10.1021/cb800240j.
[32] Ye G, Wu H, Huang J, Wang W, Ge K, Li G, et al. LAMP2: a major update of the database linking antimicrobial peptides. *Database*. 2020; 2020: baaa061. Available from: https://doi.org/10.1093/database/baaa061.
[33] Kang X, Dong F, Shi C, Liu S, Sun J, Chen J, et al. DRAMP 2.0, an updated data repository of antimicrobial peptides. *Scientific Data*. 2019; 6(1): 148. Available from: https://doi.org/10.1038/s41597-019-0154-y.
[34] Pirtskhalava M, Gabrielian A, Cruz P, Griggs HL, Squires RB, Hurt DE, et al. DBAASP v.2: an enhanced database of structure and antimicrobial/cytotoxic activity of natural and synthetic peptides. *Nucleic Acids Research*. 2016; 44(D1): D1104-D1112. Available from: https://doi.org/10.1093/nar/gkv1174.
[35] Jhong JH, Yao L, Pang Y, Li Z, Chung CR, Wang R, et al. dbAMP 2.0: updated resource for antimicrobial peptides with an enhanced scanning method for genomic and proteomic data. *Nucleic Acids Research*. 2022; 50(D1): D460-D470. Available from: https://doi.org/10.1093/nar/gkab1080.
[36] Waghu FH, Barai RS, Gurung P, Idicula-Thomas S. CAMPR3: a database on sequences, structures and signatures of antimicrobial peptides. *Nucleic Acids Research*. 2016; 44(D1): D1094-D1097. Available from: https://doi.org/10.1093/nar/gkv1051.
[37] Chang KY, Lin TP, Shih LY, Wang CK. Analysis and prediction of the critical regions of antimicrobial peptides based on conditional random fields. *PLoS One*. 2015; 10(3): e0119490. Available from: https://doi.org/10.1371/journal.pone.0119490.
[38] Wang G, Li X, Wang Z. APD3: the antimicrobial peptide database as a tool for research and education. *Nucleic Acids Research*. 2016; 44(D1): D1087-D1093. Available from: https://doi.org/10.1093/nar/gkv1278.
[39] Qureshi A, Thakur N, Tandon H, Kumar M. AVPdb: a database of experimentally validated antiviral peptides targeting medically important viruses. *Nucleic Acids Research*. 2014; 42(D1): D1147-D1153. Available from:

https://doi.org/10.1093/nar/gkt1191.
[40] Piotto SP, Sessa L, Concilio S, Iannelli P. YADAMP: yet another database of antimicrobial peptides. *International Journal of Antimicrobial Agents*. 2012; 39(4): 346-351. Available from: https://doi.org/10.1016/j.ijantimicag.2011.12.003.
[41] Yoshida M, Hinkley T, Tsuda S, Abul-Haija YM, McBurney RT, Kulikov V, et al. Using evolutionary algorithms and machine learning to explore sequence space for the discovery of antimicrobial peptides. *Chem*. 2018; 4(3): 533-543. Available from: https://doi.org/10.1016/j.chempr.2018.01.005.
[42] Bhadra P, Yan J, Li J, Fong S, Siu SWI. AmPEP: Sequence-based prediction of antimicrobial peptides using distribution patterns of amino acid properties and random forest. *Scientific Report*. 2018; 8(1): 1697. Available from: https://doi.org/10.1038/s41598-018-19752-w.
[43] Veltri D, Kamath U, Shehu A. Improving recognition of antimicrobial peptides and target selectivity through machine learning and genetic programming. *IEEE/ACM Transactions on Computational Biology and Bioinformatics*. 2017; 14(2): 300-313. Available from: https://doi.org/10.1109/TCBB.2015.2462364.
[44] Meher PK, Sahu TK, Saini V, Rao AR. Predicting antimicrobial peptides with improved accuracy by incorporating the compositional, physico-chemical and structural features into Chou's general PseAAC. *Scientific Report*. 2017; 7(1): 42362. Available from: https://doi.org/10.1038/srep42362.
[45] Heffernan R, Paliwal K, Lyons J, Dehzangi A, Sharma A, Wang J, et al. Improving prediction of secondary structure, local backbone angles and solvent accessible surface area of proteins by deep iterative learning. *Scientific Report*. 2015; 5(1): 11476. Available from: https://doi.org/10.1038/srep11476.
[46] Muller AT, Hiss JA, Schneider G. Recurrent neural network model for constructive peptide design. *Journal of Chemical Information and Modeling*. 2018; 58(2): 472-479. Available from: https://doi.org/10.1021/acs.jcim.7b00414.
[47] Han R, Yoon H, Kim G, Lee H, Lee Y. Revolutionizing medicinal chemistry: the application of artificial intelligence (AI) in early drug discovery. *Pharmaceuticals*. 2023; 16(9): 1259. Available from: https://doi.org/10.3390/ph16091259.
[48] Lertampaiporn S, Vorapreeda T, Hongsthong A, Thammarongtham C. Ensemble-AMPPred: robust AMP prediction and recognition using the ensemble learning method with a new hybrid feature for differentiating AMPs. *Genes*. 2021; 12(2): 137. Available from: https://doi.org/10.3390/genes12020137.
[49] Tripathi A, Goswami T, Trivedi SK, Sharma RD. A multi class random forest (MCRF) model for classification of small plant peptides. *International Journal of Information Management Data Insights*. 2021; 1(2): 100029. Available from: https://doi.org/10.1016/j.jjimei.2021.100029.
[50] Sarker IH. Deep learning: a comprehensive overview on techniques, taxonomy, applications and research directions. *SN Computer Science*. 2021; 2(6): 420. Available from: https://doi.org/10.1007/s42979021-00815-1.
[51] Söylemez ÜG, Yousef M, Bakir-Gungor B. Prediction of antimicrobial peptides using deep neural networks. In: *Proceedings of the 16th International Joint Conference on Biomedical Engineering Systems and Technologies (BIOSTEC 2023)*. SciTePress; 2023. p.188-194. Available from: https://doi.org/10.5220/0011690000003414.
[52] Jones DT. Protein secondary structure prediction based on position-specific scoring matrices. *Journal of Molecular Biology*. 1999; 292(2): 195-202. Available from: https://doi.org/10.1006/jmbi.1999.3091.
[53] Pollastri G, Przybylski D, Rost B, Baldi P. Improving the prediction of protein secondary structure in three and eight classes using recurrent neural networks and profiles. *Proteins: Structure, Function, and Bioinformatics*. 2002; 47(2): 142-235. Available from: https://doi.org/10.1002/prot.10082.
[54] Fernandes FC, Cardoso MH, Gil-Ley A, Luchi LV, da Silva MG, Macedo ML, et al. Geometric deep learning as a potential tool for antimicrobial peptide prediction. *Frontiers in Bioinformatics*. 2023; 3: 1216362. Available from: https://doi.org/10.3389/fbinf.2023.1216362.
[55] Wang Y, Gong H, Li X, Li L, Zhao Y, Bao P, et al. De novo multi-mechanism antimicrobial peptide design via multimodal deep learning. *bioRxiv*. 2024. Available from: https://doi.org/10.1101/2024.01.02.573846.
[56] Boopathi V, Subramaniyam S, Malik A, Lee G, Manavalan B, Yang DC. MACppred: A support vector machine-based meta-predictor for identification of anticancer peptides. *International Journal of Molecular Sciences*. 2019; 20(8): 1964. Available from: https://doi.org/10.3390/ijms20081964.
[57] Ng XY, Rosdi BA, Shahrudin S. Prediction of antimicrobial peptides based on sequence alignment and support vector machine-pairwise algorithm utilizing LZ-complexity. *BioMed Research International*. 2015; 2015(1): 212715. Available from: https://doi.org/10.1155/2015/212715.
[58] Hamid MN, Friedberg I. Identifying antimicrobial peptides using word embedding with deep recurrent neural networks. *Bioinformatics*. 2019; 35(12): 2009-2016.

[59] Söylemez ÜG, Yousef M, Bakir-Gungor B. AMP-GSM: Prediction of antimicrobial peptides via a grouping-scoring-modeling approach. *Applied Sciences*. 2023; 13(8): 5106. Available from: https://doi.org/10.3390/app13085106.
[60] Lazzaro BP, Zasloff M, Rolff J. Antimicrobial peptides: Application informed by evolution. *Science*. 2020; 368(6490): eaau5480. Available from: https://doi.org/10.1126/science.aau5480.
[61] Joseph S, Karnik S, Nilawe P, Jayaraman VK, Idicula-Thomas S. ClassAMP: a prediction tool for classification of antimicrobial peptides. *IEEE/ACM Transactions on Computational Biology and Bioinformatics*. 2012; 9(5): 1535-1538.
[62] Yan J, Bhadra P, Li A, Sethiya P, Qin L, Tai HK, et al. Deep-AmPEP30: improve short antimicrobial peptides prediction with deep learning. *Molecular Therapy-Nucleic Acids*. 2020; 20: 882-894. Available from: https://doi.org/10.1016%2Fj.omtn.2020.05.006.
[63] Gombart AF, Bhan I, Borregaard N, Tamez H, Camargo CA Jr, Koeffler HP, et al. Low plasma level of cathelicidin antimicrobial peptide (hCAP18) predicts increased infectious disease mortality in patients undergoing hemodialysis. *Clinical Infectious Diseases*. 2009; 48(4): 418-424. Available from: https://doi.org/10.1086/596314.
[64] Lin Y, Cai Y, Liu J, Lin C, Liu X. An advanced approach to identify antimicrobial peptides and their function types for penaeus through machine learning strategies. *BMC Bioinformatics*. 2019; 20: 1-10. Available from: https://doi.org/10.1186/s12859-019-2766-9.
[65] Wang G, Vaisman II, Van Hoek ML. Machine learning prediction of antimicrobial peptides. *Computational Peptide Science: Methods and Protocols*. 2022; 2405: 1-37.
[66] Rice P, Longden I, Bleasby A. EMBOSS: the European molecular biology open software suite. *Trends in Genetics*. 2000; 16(6): 276-277. Available from: https://doi.org/10.1016/s0168-9525(00)02024-2.
[67] Berendsen HJ, Van der Spoel D, Van Drunen R. GROMACS: A message-passing parallel molecular dynamics implementation. *Computer Physics Communications*. 1995; 91(1-3): 43-56. Available from: https://doi.org/10.1016/0010-4655(95)00042-E.
[68] Chakraborty A, Kobzev E, Chan J, de Zoysa GH, Sarojini V, Piggot TJ, et al. Molecular dynamics simulation of the interaction of two linear battacin analogs with model gram-positive and gram-negative bacterial cell membranes. *ACS Omega*. 2020; 6(1): 388-400. Available from: https://doi.org/10.1021/acsomega.0c04752.
[69] Wang Y, Zhao T, Wei D, Strandberg E, Ulrich AS, Ulmschneider JP. How reliable are molecular dynamics simulations of membrane active antimicrobial peptides? *Biochimica et Biophysica Acta (BBA)-Biomembranes*. 2014; 1838(9): 2280-2288. Available from: https://doi.org/10.1016/j.bbamem.2014.04.009.
[70] Ragoza M, Hochuli J, Idrobo E, Sunseri J, Koes DR. Protein-ligand scoring with convolutional neural networks. *Journal of Chemical Information and Modeling*. 2017; 57(4): 942-957. Available from: https://doi.org/10.1021/acs.jcim.6b00740.
[71] Meng QZ, Tang JJ, Guo F. Multi-AMP: detecting the antimicrobial peptides and their activities using the multi-task learning. In: *2021 IEEE International Conference on Bioinformatics and Biomedicine (BIBM)*. Houston, TX, USA: IEEE; 2021. p.710-713. Available from: https://doi.org/10.1109/BIBM52615.2021.9669452.
[72] Gupta A, Zou J. Feedback GAN for DNA optimizes protein functions. *Nature Machine Intelligence*. 2019; 1(2): 105-111. Available from: https://doi.org/10.1038/s42256-019-0017-4.
[73] Wan F, Kontogiorgos-Heintz D, de la Fuente-Nunez C. Deep generative models for peptide design. *Digital Discovery*. 2022; 1(3): 195-208. Available from: https://doi.org/10.1039/D1DD00024A.
[74] Sharma A, Gupta P, Kumar R, Bhardwaj A. dPABBs: a novel in silico approach for predicting and designing anti-biofilm peptides. *Scientific Reports*. 2016; 6(1): 21839. Available from: https://www.nature.com/articles/srep21839.
[75] Dziuba B, Dziuba M. New milk protein-derived peptides with potential antimicrobial activity: an approach based on bioinformatic studies. *International Journal of Molecular Sciences*. 2014; 15(8): 14531-14545. Available from: https://doi.org/10.3390/ijms150814531.
[76] Gawde U, Chakraborty S, Waghu FH, Barai RS, Khanderkar A, Indraguru R, et al. CAMPR4: a database of natural and synthetic antimicrobial peptides. *Nucleic Acids Research*. 2023; 51(D1): D377-D383. Available from: https://doi.org/10.1093/nar/gkac933.