Advances and Applications in Bioinformatics and Chemistry Dovepress

Open Access Full Text Article REVIEW

Recent Applications of Artificial Intelligence in Discovery of New Antibacterial Agents

Youcef Bagdad, Maria A Miteva

Université Paris Cité, CNRS UMR 8038 CiTCoM, Inserm U1268 MCTR, Paris, France

Correspondence: Maria A Miteva, Université Paris Cité - CNRS UMR 8038 CiTCoM, Inserm U1268 MCTR, 4 Av. de l'Observatoire, Paris, 75006, France, Email maria.mitev@inserm.fr

Abstract: Antimicrobial resistance (AMR) represents today a major challenge for global public health, compromising the effectiveness of treatments against a multitude of bacterial infections. In recent decades, artificial intelligence (AI) has emerged as a promising technology for the identification and development of new antibacterial agents. This review focuses on AI methodologies applied to discover new antibacterial candidates. Case studies that identified small molecules and peptides showing antimicrobial activity and demonstrating efficiency against pathogenic resistant bacteria by employing AI are summarized. We also discuss the challenges and opportunities offered by AI, highlighting the importance of AI progress for the identification of new promising antibacterial drug candidates to combat the AMR.

Keywords: artificial intelligence, drug discovery, antibacterial agents, antimicrobial resistance

Introduction

Bacteria are unicellular prokaryotic organisms that play a crucial role in many vital biological processes.[1] Their remarkably extensive genetic and metabolic diversity makes them strongly adaptive to various environments, from the most extreme conditions to the most welcoming habitats. In addition to their essential role in biogeochemical cycles, bacteria maintain symbiotic relationships with many living organisms, including humans. A notable example is the human microbiome, a complex assembly of bacterial communities primarily located in the gut,[2] but also in the skin and other parts of the body.[3] These bacteria play a vital role in digestion, vitamin synthesis, protection against pathogens, and modulation of the immune system.[2–4] However, pathogenic bacteria can cause severe infectious diseases, posing significant problems to human health. They have the capacity to provoke a varying range of conditions, from mild infections to serious diseases, and thus represent a major challenge for public health.[5–7]

While the fight against bacterial diseases has always been existing,[8–10] in the 19th century, bacteria were identified as the source of these infections with the foundational work of Louis Pasteur and Robert Koch.[11] Progressing into the early 20th century, Paul Ehrlich pioneered the development of the first synthetic antimicrobial agent, salvarsan, specifically formulated to fight syphilis.[12] However, the discovery of penicillin in 1928 by Alexander Fleming and its further development by Howard Florey and Ernst Chain in 1939 marked a pivotal moment in the fight against infectious diseases.[13,14] This discovery was particularly significant because penicillin has proven to be more effective, less toxic, and with a broader spectrum of activity against various bacterial infections compared to salvarsan. This started the antibiotic era, which significantly improved human survival against infections.[10,14]

In the years following the discovery of penicillin, research in microbiology and pharmacology has shown a remarkable advance, leading to the identification and development of different families of antibiotics. Several class molecules, such as β-lactams[15] (to which penicillin belongs), aminoglycosides,[16,17] tetracyclines,[18] macrolides,[19] and fluoroquinolones,[20] target

Received: 27 June 2024
Accepted: 25 October 2024
Published: 3 December 2024

specific and crucial aspects of bacterial survival. Their diversity in mechanisms of action has led to treatments in which many bacterial diseases could be controlled or eliminated.

Unfortunately, the abundant use of antibiotics triggered the arrival of antimicrobial resistance (AMR), a critical concern for global health. Bacteria have evolved sophisticated mechanisms to counteract the antibiotics action: the expression of β-lactamase enzymes degrading β-lactam antibiotics, thus neutralizing their effect on Penicillin Binding Proteins (PBPs); enhancing the efflux pump activity; reducing the membrane permeability.[21,22] Consequently, this adaptation has led to the emergence of bacterial strains resistant to various antibiotics, complicating treatment efforts and presenting a new public health challenge. Today, the AMR continues to intensify, with around 5 million deaths in 2019 due to treatment resistance.[23] Future projections are even more alarming, the United Nations Environment Program anticipates that by 2050 deaths from bacterial infections could escalate to 10 million annually, with economic losses potentially reaching US$100 trillion.[24,25] Resistance dissemination methods, including horizontal gene transfer, mutation, and conjugation,[26,27] further intensify AMR, decreasing therapeutic options.[28] The growing AMR emphasizes the urgent need for the development of novel therapeutic strategies.

In this context, artificial intelligence (AI) has already proven to be valuable in drug discovery.[29] AI was also proposed to improve diagnostics and the rational use of antibiotics,[30] as demonstrated by Google's Deep Learning system, showing performance comparable to radiologists practice in identifying active pulmonary tuberculosis.[31] Similarly, mobile applications have been developed to quickly and accurately analyze antibiotic susceptibility tests (AST), without expensive equipment.[32,33] AI was also applied to enhance bacteriophage therapy, a promising avenue in combating AMR.[34] Examples include the recognition of the bacteriophage life-cycle using AI,[35,36] prediction of bacteriophage virion proteins,[37,38] phage host prediction[39–41] or the classification of bacteriophages.[42–46]

We focus here on the application of AI in the discovery of new antibacterial agents, given the immense challenges due to AMR,[47] often surpassing the challenges of developing other drugs. We review studies employing AI for identifying new agents targeting bacterial strains, most of which are designated as high-priority threats by the World Health Organization (WHO).[48] We also discuss the methodologies and algorithms used, highlighting the importance of AI progress for the identification of new promising antibacterial drug candidates.

AI Methods Used to Identify New Antibacterial Agents

AI algorithms in drug discovery exploit mathematical models to process large volumes of data, predict molecular interactions, assess the efficacy of various compounds, or more recently generate new compounds with novel properties. To be efficient, these algorithms require databases containing molecules with known activity toward therapeutic targets and pharmacological data to constitute training and test sets. Traditional machine learning (ML) methods like Random Forests (RF), Support Vector Machines (SVM), among others (Figure 1) involve physicochemical molecular descriptors or fingerprints. These descriptors capture structural and chemical features of the molecules of training and test sets. In deep learning (DL), approaches like Convolutional Neural Networks (CNNs), Graph Neural Networks (GNNs), Generative Adversarial Networks (GANs) etc. allow for direct extraction and processing of these features from molecular representations, eg as graphs. Additionally, DL can also blend physicochemical descriptors with autonomously identified features.

Below, we briefly present the ML/DL methods used in the studies discussed in this review.

SVM[49,50] and RF[51–53] are two major ML approaches for classification and regression, widely used to classify molecules based on their biological activity or to quantitatively predict their activity. SVM primarily aims to find an optimal hyperplane that separates data into different classes (eg, active and inactive molecules), while maximizing the distance (or margin) between the closest data points to this hyperplane, known as support vectors, crucial in determining the position of the hyperplane. By maximizing this margin, SVM seeks to enhance the model robustness, making it more stable in case of data variability. The advantage of SVM is its ability to handle nonlinear problems through the use of kernels, which are mathematical functions that transform the original data into a high-dimensional space where they become separable.

RF operates using a collection of decision tree models. It constructs multiple trees using different random subsets of the data and molecular descriptors, then combines their predictions for a more stable and robust outcome. Unlike a single decision tree, RF creates a large number of trees, each tree making an individual decision on classification (or regression), and the final decision is obtained through majority voting in classification, or by averaging predictions in regression.

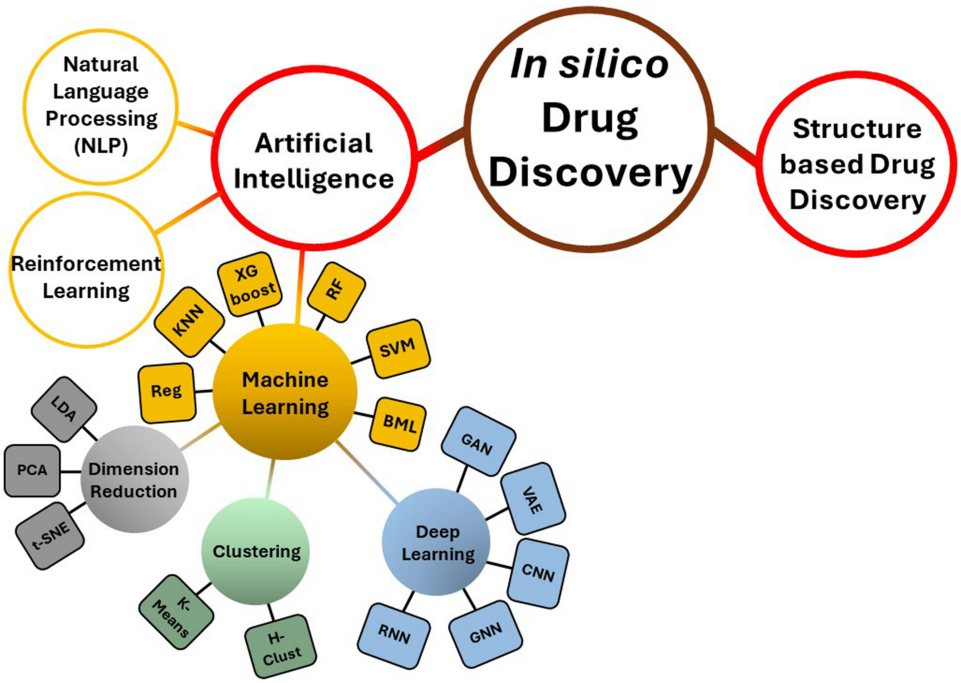

Figure 1 Non-exhaustive mapping of AI techniques for in silico drug discovery including traditional ML methods and cutting-edge DL techniques.
Abbreviations: KNN, K-Nearest Neighbors; SVM, Support Vector Machines; RF, Random Forest; BML, Bayesian Machine Learning; XG Boost, eXtreme Gradient Boosting; Reg, Linear Regression + Logistic Regression; PCA, Principal Component Analysis; LDA, Linear Discriminant Analysis; t-SNE, t-Distributed Stochastic Neighbor Embedding; GAN, Generative Adversarial Network; CNN, Convolutional Neural Network; VAE, Variational Autoencoder; RNN, Recurrent Neural Network.

Bayesian Machine Learning (BML)[54] models start with initial probabilities reflecting an early estimate based on general assumptions or uncertain initial information. These probabilities refine as the model processes new data from the training set. The Bayesian model relies on an iterative process that updates these estimates by combining pre-existing knowledge with newly acquired information, leading to improved and more precise probabilities concerning the properties to be predicted. The final model is based on a series of probabilistic distributions.

Currently, advanced and more complex techniques have emerged, particularly Deep Neural Networks (DNNs) methods and their application in DL.[55,56] DNNs consist of multiple layers of neurons organized in a hierarchy. Each layer processes the outputs from the previous layer, allowing the network to capture increasingly abstract representations of the data through the network. DNNs are particularly effective for handling large, complex and heterogeneous datasets. They are used in various applications, including predicting molecular properties or generating new molecules with specific properties.

CNNs[57] are especially effective in image processing due to their use of convolution operations. This operation involves a repeated application of a filter or kernel that moves across the entire image and performs simple mathematical operations to transform the initial data. A filter is essentially a small matrix compared to the image size, designed to detect specific features like edges, angles, or textures by calculating weighted sums of the pixel values it covers. These principles are well transposable into molecular structure analysis. Molecules can be represented in a format suitable for CNNs, such as distance matrices or connectivity matrices, which describe the spatial relationships between atoms in a molecule. By applying convolutional filters to these matrices, CNNs can extract key information on structural and functional patterns of molecules.

On the other hand, GNNs,[58] representing a significant advancement of AI in drug design, process data in the form of regular matrices as graphs, which naturally corresponds to the representation of molecules where atoms are nodes and chemical bonds are edges. GNNs use a protocol called "message passing", where each node gathers information from its neighbors to adjust its own characteristics. That allows the model to consider both the local interactions between atoms and the overall structure of the molecule. Directed Message Passing Neural Networks (DMPNNs)[59,60] are a specific

variant of GNNs integrating the direction of chemical bonds (eg differences in electronegativity) in the message-passing process.

In DL, generative neural networks provide sophisticated methods for creating compounds with specific properties. Among these techniques, Variational Autoencoders (VAEs)[61] are a class of generative models that learn to compress data into a simplified form, known as the latent space. VAEs apply a probabilistic distribution (usually Gaussian) to this latent space, allowing to generate new instances by randomly sampling from this space. By training a VAE on sets of known molecules, the model learns a mapping of the chemical space into a latent space where similar points exhibit similar chemical properties. By exploring this space, it is possible to generate novel molecules with desired properties in terms of functionality and structure.

The effectiveness of AI algorithms strongly depends on the quality and nature of the data used for their training. Each type of algorithm has specific characteristics that make it particularly well suited to certain types of data, resulting in superior performance in specific domains. For example, CNNs are specially designed to process grid-structured data like digital images. Their architecture, inspired by the functioning of the human visual cortex, allows them to excel in tasks such as image recognition and classification. They outperform other algorithms like RF or SVM, which are not optimized to handle high-dimensional visual data. Conversely, RF and SVM are particularly effective in processing datasets with numerous independent features, where each attribute contributes autonomously to the final prediction. In the context of complex structures like networks or graphs, GNNs stand out for their exceptional performance, surpassing traditional ML methods that are not suited to handle complex relationships and interdependencies within the data. Finally, the optimization of hyperparameters is a fundamental element in the effectiveness of ML/DL models. These parameters, defined before the learning process, directly influence the model's ability to learn and generalize correctly from the provided data. A cautious adjustment of hyperparameters can significantly improve the model's performance on a specific dataset. However, it is crucial to carry out this optimization by considering the compatibility between the data, the chosen model, and the intended objective to avoid overfitting.

Antibacterial Molecules Discovered by AI

Small Molecules

One of the first advances using AI in research on antimicrobial agents is the study by Hu et al[62] aimed at combating the re-emerging diseases caused by *Yersinia spp* species, including *Yersinia pestis*, the agent responsible for the plague.[63] Their approach combined homology modeling, ML, and high-throughput docking to discover inhibitors of *Yersinia* protein kinase A (YpkA), a crucial virulence factor whose activity modulates the pathogenicity of these bacteria. The first phase involved virtual screening through an ML model based on SVM[64] to select potential YpkA inhibitors through a vast chemical library. This model, trained on data of 364 inhibitors and 4220 non-inhibitors and utilizing over 200 3D physicochemical characteristics (descriptors), reached an area under the ROC curve (AUC) of 70% on its test set. With this approach, the authors screened an internal virtual database of over 200 million compounds, and selected 200,000 candidates. In the second phase, the selected compounds underwent docking-based screening against five different YpkA structures, derived from molecular dynamics (MD) simulations of a homology model. This process resulted in 45 compounds chosen for experimental validation, leading to the identification of seven YpkA inhibitors with complete inhibition at concentrations ranging from 225 μM to 450 μM. Notably, compounds labelled as *compound 1_Hu* et al, *compound 3_Hu* et al, and *compound 4_Hu* et al in Figure 2A had the strongest inhibitory effects, with IC50 values of 1.81, 9.72, and 5.87 μM, respectively. Their efficacy was further validated by comparative tests against other kinases, such as Mitogen-Activated Protein Kinase (MAPK) and Protein Kinase C (PKC), showing a preference for inhibiting YpkA, with compounds 3 and 4 exhibiting up to ten-fold greater selectivity.

Another AI applications in antimicrobial research were studies aiming to fight against methicillin-resistant *Staphylococcus aureus* (MRSA),[70] a significant challenge for the antibiotic resistance. In 2014, Wang et al[65] worked on a project using predictive ML models to identify new compounds with potential MRSA inhibitory properties. This investigation was based on a dataset of 5451 compounds from the ChEMBL database[71] and their anti-MRSA activities, including 2066 active molecules with a Minimum Inhibitory Concentration (MIC) of less than 5 μM. To develop these

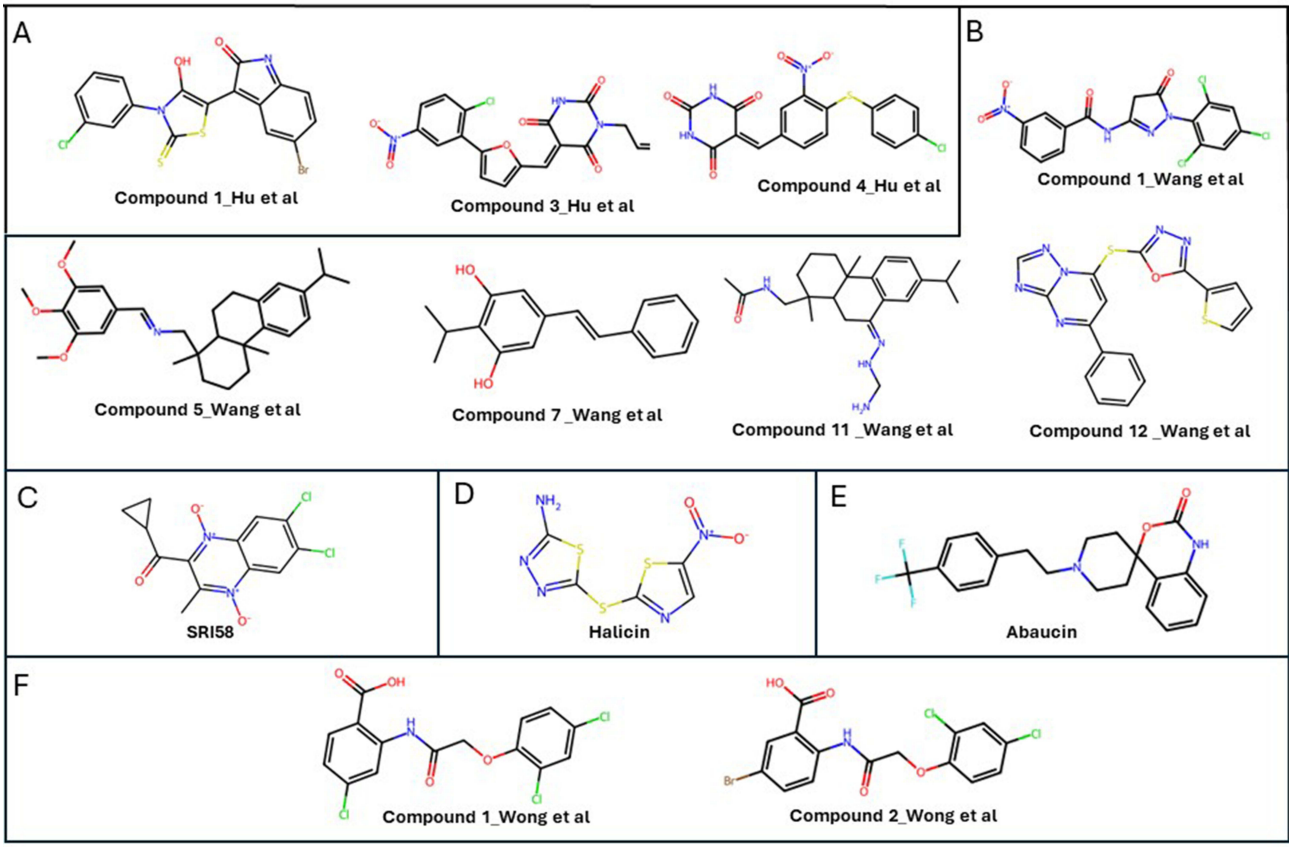

Figure 2 2D structures of different compounds showing antibacterial effects and discovered by AI techniques. (**A**) compounds identified by Hu at al[62] (**B**) compounds identified by Wang et al[65] (**C**) SRI58 - identified by Ekins et al[66] (**D**) halicin - identified by Stokes et al[67] (**E**) abaucin – identified by Liu et al[68] (**F**) Compounds identified by Wong et al[69] The numbers of compounds in the figure correspond to the original numbers of the compounds identified in the cited studies/.

models, researchers utilized various ML algorithms such as Naive Bayes (NB),[72] SVM and k-Nearest Neighbors (kNN).[73] Several models were established using physicochemical descriptors and molecular fingerprints and the most accurate models displayed success rates exceeding 80% on both the training and test datasets. The most effective model, SVM, was applied for the virtual screening of the Guangdong Small Molecule Tangible Library (GSMTL), containing approximately 7500 compounds. Subsequently, 56 compounds were selected and evaluated through micro-dilution assays against three MRSA strains known for their strong resistance (ST239, ST5, and 252), resulting in the validation of 12 new highly promising anti-MRSA agents with MICs ranging from 4 to 64 mg/L (see Figure 2B). Among them, the compound *1_Wang* et al "CID113055", *compound 7_Wang* et al "CID133191", *compound 11_Wang* et al "CID122370897", and *compound 12_Wang* et al "CID11840831" exhibited good activity against the three MRSA strains with MIC values below 32 mg/L, which were superior or comparable to ampicillin sodium, a widely used antibacterial agent (see Table 1 for more details). Wang et al[65] suggested that unlike ampicillin sodium, which acts on penicillin-binding proteins (PBPs) and is less effective against MRSA due to mutations in PBPs, the new compounds might work through new mechanisms or through novel scaffolds targeting known pathways.

In 2015, a work focusing on combating *Mycobacterium tuberculosis* (Mtb)[83] reported by Ekins et al[66] proposed an innovative approach to search for new compounds against this Mtb, by using pharmacophores based on metabolites and substrates of Mtb, combined with *in-house* Bayesian ML models trained to recognize molecules inhibiting the growth of Mtb and having low cytotoxicity to human cells.[84–86] In this study, 20 Mtb enzymes were identified as being of particular interest, and 66 pharmacophores were selected based on substrates and metabolites of these enzymes. These pharmacophores were used to screen 206,000 commercial molecules from the database Asinex Gold,[87] and 14,733 molecules were retained. These molecules were then processed through the Bayesian models, and 110 candidates were selected, and subjected to in vitro testing, resulting in the discovery of two promising hits: *BAS 04912643* and *BAS 00623753*, with

Table 1 Antimicrobial Compounds and AMPs Identified by Use of in silico and AI Approaches and Experimentally Validated

Reference /Figure n°	Compound/Peptide	Targeted Bacteria or Protein	Activity	in silico/AI Methodology	Dataset Screened
[62] Figure 2A	Compound 1 _Hu et al	YpkA	IC50 1.81 μM	SVM	Internal database ≈ 200 million compounds
	Compound 3 _Hu et al		IC50 9.72 μM		
	Compound 4 _Hu et al		IC50 5.87 μM		
[65] Figure 2B	Compound 1_Wang et al	*MRSA*	MIC <32 mg/L	SVM	GSMTL
	Compound 5_Wang et al		MIC 8–16 mg/L		
	Compound 7_ Wang et al		MIC 4–8 mg/L		
	Compound 11_ Wang et al		MIC 8–16 mg/L		
	Compound 12_ Wang et al		MIC <32 mg/L		
[66] Figure 2C	SR158	*Mtb*	MIC 1.25 g/mL	Bayesian ML models	Asinex Gold database
[67] Figure 2D	Halicin	*E. coli, Mtb, CRE*	active	D-MPNN	DRH, ZINC15
[68] Figure 2E	Abaucin	*A. baumannii*	active	D-MPNN	DRH
[69] Figure 2F	Compound 1_Wong et al	*MRSA, VRE*	MIC 4 μg/mL	GNN	Mcule purchasable database and a Broad Institute database
	Compound 2_Wong et al		MIC 3 μg/mL		

[74]	HHC-10	*P. aeruginosa*	MIC 0.8–5.9 μM	ANN	100,000 computer-generated peptides of 9 amino acids.
		MRSA	MIC 1.5 μM		
		E. coli (ESBL)	MIC 1.5 μM		
		*K. pneumoniae (*ESBL)	MIC 6.2–25 μM		
		P.maltophilia	MIC 0.4 μM		
		*E.cloacae (*AmpC β-lactamase)	MIC 3 μM		
		VRE faecalis	MIC 12–99 μM		
		VRE faecium	MIC 1.5–3.1 μM		
	HHC-36	*P. aeruginosa*	MIC 0.7–11 μM		
		MRSA	MIC 1.4 μM		
		E. coli (ESBL)	MIC 2.7–5.4 μM		
		K. pneumoniae (ESBL)	MIC 22–174 μM		
		P. maltophilia	MIC 1.4 μM		
		E. cloacae (AmpC β-lactamase)	MIC 11 μM		
		VRE faecalis	MIC 43–174 μM		
		VRE faecium	MIC 1.3–11 μM		
[75]	Peptide-B4	*E. coli*	MIC 64 μg/mL	ML models (AMPA, ADAM, CAMPR3)	Transcriptome of *sea anemone Cnidopus japonicus*
		B. subtilis	MIC 32 μg/mL		
	Peptide-A3	*B. subtilis*	MIC 128 μg/mL		
	Peptide-B1	*B. subtilis*	MIC 64 μg/mL		
[76]	Peptide-3	*E. coli*	MIC 32 μM	SVM model with Word2vec	Computer-generated peptides
		P. aeruginosa	MIC 64 μM		
	Peptide-4	*E. coli*	MIC 2 μM		
		P. aeruginosa	MIC 16 μM		

(*Continued*)

https://doi.org/10.2147/AABC.S484321

Table 1 (Continued).

Reference /Figure n°	Compound/Peptide	Targeted Bacteria or Protein	Activity	in silico/AI Methodology	Dataset Screened
[77]	Peptide-P3	*B. subtilis*	active	CNN and PseKRAAC	Genome of *Candida glabrata*
		V. parahaemolyticus	active		
[78]	Peptide-P1	*E. coli*	EC50 2.9 μg/mL	VAE	Computer-generated peptides
		A. baumannii	EC50 3.1 μg/mL		
		S. aureus	EC50 0.4 μg/mL		
	Peptide- P2	*E. coli*	EC50 3.1 μg/mL		
		A. baumannii	EC50 5.9 μg/mL		
		S. aureus	EC50 6.6 μg/mL		
	Peptide-P3	*E. coli*	EC50 2.9 μg/mL		
		A. baumannii	EC50 1.8 μg/mL		
		S. aureus	EC50 2.2 μg/mL		
[79]	AMP-12	*C. acnes*	MIC 2 μg/mL	Generative Deep learning pipeline	Computer-generated peptides
	AMP-29		MIC 2 μg/mL		
	AMP-31		MIC 4 μg/mL		
	AMP-33		MIC 4 μg/mL		
	AMP-25		MIC 4 μg/mL		
[80] Figure 3A	Compound 11_Parvaiz et al	*β-lactamase CMY-10 - Drug-resistant Enterobacteriaceae*	MIC 4 mg/mL	RF	In-house database from Chembridge and Maybridge
[81] Figure 3B	Zndm19	*Metallo-β-lactamase NDM-1*	IC50 2.68 μM	DNN	Specs chemical library
[82] Figure 3C	ZINC339204163	*β-lactamase AmpC*	KI 1.3 μM	Molecular docking	170 million compounds
	ZINC549719643	*β-lactamase AmpC*	Ki 77 nM		

MICs of 2.5 and 5 μg/mL, respectively. These two hits led to the discovery of a class of compounds, the quinoxaline di-N-oxides, one of which, *SRI58* "CID156018242" shown in Figure 2C, proved particularly promising results with an MIC of 1.25 μg/mL against Mtb and low cytotoxicity towards Vero cells (mammalian cells). This study highlighted the mechanism of action for *SRI58*, functioning by inducing perturbations in the mRNA levels. This suggested that *SRI58* altered the bacterial membrane potential, a crucial mechanism for bacterial energy maintenance and overall survival. However, despite these encouraging in vitro results, *SRI58* did not show measurable levels in mouse blood, highlighting the challenges related to the in vivo pharmacokinetics of the compounds, and indicating that there is a long way to go from AI-based candidates to in vivo activity for truly effective molecules.

The above discussed studies concern the first steps of AI in the discovery of new antimicrobial compounds. Since 2020, we have observed a strong evolution that marks the growing role of AI in the search for small antimicrobial molecules. This new era is characterized by significant advances in AI techniques, particularly in DL, which have made it possible to identify molecules with remarkable antibacterial properties validated in vitro and in vivo with different mechanisms of action. The study that marked this first turning point is that of Stokes et al[67] using a directed message-passing deep neural network (D-MPNN),[59] a cutting-edge approach compared to the classic ML models utilized in previous works. This technical innovation allowed for a more sophisticated analysis and prediction of compounds with the potential to act as new antibiotics. The study started with the experimental assessment of 2335 molecules, covering FDA-approved drugs[88] and natural products, for their inhibitory effects on *Escherichia coli (E. coli*). That resulted in 120 molecules displaying significant inhibition (over 80%), forming a dataset of 120 inhibitors and 2215 non-inhibitors. This dataset was used to train a D-MPNN model to predict the potential of compounds as effective antibiotics. After the initial training, the model was further applied to screen broader compound libraries, including the Drug Repurposing Hub (DRH), which comprises over 6000 molecules at different stages of development for various diseases. This extensive search led to the discovery of *SU3327*, an inhibitor of c-Jun N-terminal kinase, which was renamed as halicin, a preclinical nitrothiazole candidate (Figure 2D). Originally investigated for diabetes medication, halicin demonstrated potent activity against *E. coli* resistant strains, as well as pathogens like Mtb and carbapenem-resistant *Enterobacteriaceae* (CRE), by disrupting the ΔpH component of the proton-driven force, differently from traditional antibiotics that focus on cell wall, protein synthesis, or DNA replication. This action against the proton- driven force weakens a critical bacterial function: maintaining a proton gradient across the cell membrane, essential for ATP production, nutrient transport, and cell stability. Consequently, halicin has proven to be effective against antibiotic-resistant bacteria by targeting an underexploited bacterial function. Notably, unlike conventional antibiotics, which are primarily effective against actively metabolizing cells, halicin retains its bactericidal activity even against cells in a metabolically repressed state, thanks to its uncommon mechanism of action. This is particularly important for eradicating persistent cells that can survive treatment with other antibiotics and contribute to AMR. Furthermore, halicin was shown to be effective in mouse models against infections caused by pan-resistant *Acinetobacter baumannii (A. baumannii)*.

The AI-driven discovery of small-molecule antibiotics further advanced in 2023 with the significant research conducted by Gary Liu et al[68] leading to the identification of new narrow-spectrum antibiotics. Their focus was on combating *A. baumannii*, a multidrug-resistant Gram-negative pathogen. Utilizing a similar AI model based on a D-MPNN, Liu's team performed a screening of approximately 7500 molecules, which included a mix of off-patent drugs and synthetic chemicals. The authors identified 480 active molecules that demonstrated inhibitory activities against *A. baumannii*. The AI model, adeptly trained with data from these screenings, successfully identified patterns and relationships in molecular structures indicative of antibacterial activity. The screening of the DRH library through the AI model unveiled abaucin (Figure 2E) (previously known as *RS102895*, targeting Chemokine (C-C motif) Receptor 2 (CCR2)), a compound specifically effective against *A. baumannii*. As halicin, abaucin's mechanism of action is different from the traditional antibiotics. It was found to disrupt bacterial lipoprotein trafficking, particularly affecting the function of the Lipoprotein-releasing transmembrane protein (LolE)[89] protein, vital for moving lipoproteins from the inner to the outer membrane in Gram-negative bacteria. This interference in the bacterial cellular process represents a novel strategy in combating antibiotic resistance, especially against hard-to-treat pathogens like *A. baumannii*. Importantly, abaucin was shown to control *A. baumannii* infection in a mouse wound model in vivo, demonstrating its potential for therapeutic application. Moreover, this molecule demonstrates minimal activity against human commensal bacterial species,

underlying its potential for targeted therapy with reduced impact on the host's microbiome. The specificity of abaucin in targeting *A. baumannii* highlights the potential of AI in discovering narrow-spectrum antibiotics, which are crucial in preserving the body's beneficial microbiota and reducing the risk of developing drug resistance.

It is to note that the above-mentioned models based on neural network (NN), such as those contributing to the discovery of halicin[67] and abaucin,[68] were characterized as black box models. Despite their proven success in identifying new potential antibiotics from vast chemical libraries, the understanding of the principles behind these findings remains unclear. However, elucidating new classes of antibiotics goes beyond the simple identification of isolated compounds; it requires a deep understanding of the chemical structures and mechanisms of action that are shared by different compounds. Thus, in continuation of the research conducted by Stokes et al[67] and Liu et al,[68] Wong et al have recently advanced the idea[69] that it would be possible to significantly improve the understanding of models by making their predictions explicit. Specifically, they argued that by decrypting the predictions at the level of chemical substructures using graph search algorithms, it would be possible to identify common patterns or structures among the most effective compounds as antibiotics and to predict structural classes of antibiotics. This was based on the observation that antibiotic classes are generally defined by shared substructures. Wong et al focused on discovering structural classes active against *S. aureus*. They determined experimentally the antibiotic activities and cytotoxicity profiles on human cells of 39,312 compounds containing antibiotics and natural products best known for their growth inhibitory activity against a methicillin-sensitive strain, *S. aureus RN4220*. From these data, they developed two GNN[90] models to predict the antibiotic activity and cytotoxicity of 12,076,365 molecules drawn from the Mcule purchasable database[91] and a Broad Institute database.[92] Using explainable graph algorithms, they identified substructure-based rationales for compounds exhibiting high antibiotic activity and low cytotoxicity. The screening and filtering steps led to a set of 283 compounds, which were tested experimentally. Two compounds (compound 1_Wong et al "CID2942818" and compound 2_Wong et al "CID1314498") from the same structural class, highlighted by the presence of an N-[2-(2-chlorophenoxy)ethyl] aniline core (Figure 2F), demonstrated significant activity against *S. aureus*, including MRSA and *vancomycin-resistant Enterococci* (VRE), showcasing their potential against drug-resistant infections. Specifically, these compounds were tested against 40 US Centers for Disease Control and Prevention (CDC)[93] isolates of different bacterial species containing various resistance factors, exhibiting median MIC values of 4 μg/mL for "CID2942818" and 3 μg/mL for "CID1314498", with a range of MICs from 2 to 16 μg/mL. In vivo tests were also conducted, showing significant antibacterial activity in mouse models, which further showed their therapeutic potential. In vitro toxicity testing showed that the two compounds exhibited low cytotoxicity against human cell lines, including HepG2 liver carcinoma cells, primary human skeletal muscle cells, and IMR-90 lung fibroblasts, at a concentration of 10 μM. The promising in vitro and in vivo antibacterial activity of these compounds, coupled with their non-conventional mechanism of action and their minimized risk to human cells, makes these compounds highly promising for further development.

This approach enabled the DL-guided discovery of structural classes of antibiotics and demonstrates that AI models in drug discovery can be explained, thus providing insight into the chemical substructures that underlie the selective activity of antibiotics.

β-Lactamase Inhibitors

β-lactamases are enzymes produced by bacteria that hydrolyze β-lactam antibiotics, rendering them ineffective in treating bacterial infections. This mechanism of resistance is one of the most significant in AMR. Particularly important among these are: ESBLs,[94,95] which can hydrolyze penicillins, cephalosporins, and monobactams; carbapenemases, such as KPC (*Klebsiella pneumoniae* carbapenemase)[96] and NDM (New Delhi metallo-β-lactamase),[97] which can degrade carbapenems (considered as last-resort treatments);[98] Ampicillinase C (AmpC) β-lactamases that provide resistance against a broad range of β-lactam antibiotics.[99] These enzymes have been identified in various pathogenic bacteria classified as critical threats by the WHO, including *K. pneumoniae, E. coli, A. baumannii* and *P. aeruginosa*, leading to infections increasingly difficult to treat and control. To address β-lactamase-mediated resistance, inhibitors of these enzymes are co-administered with β-lactam antibiotics to restore their efficacy against bacteria with this particular type of resistance.[100–105]

AI has also been integrated to identify inhibitors of different β-lactamases, several promising compounds were identified with AI or in silico modeling and experimental validation (see Figure 3A–C). A recent example is the work of

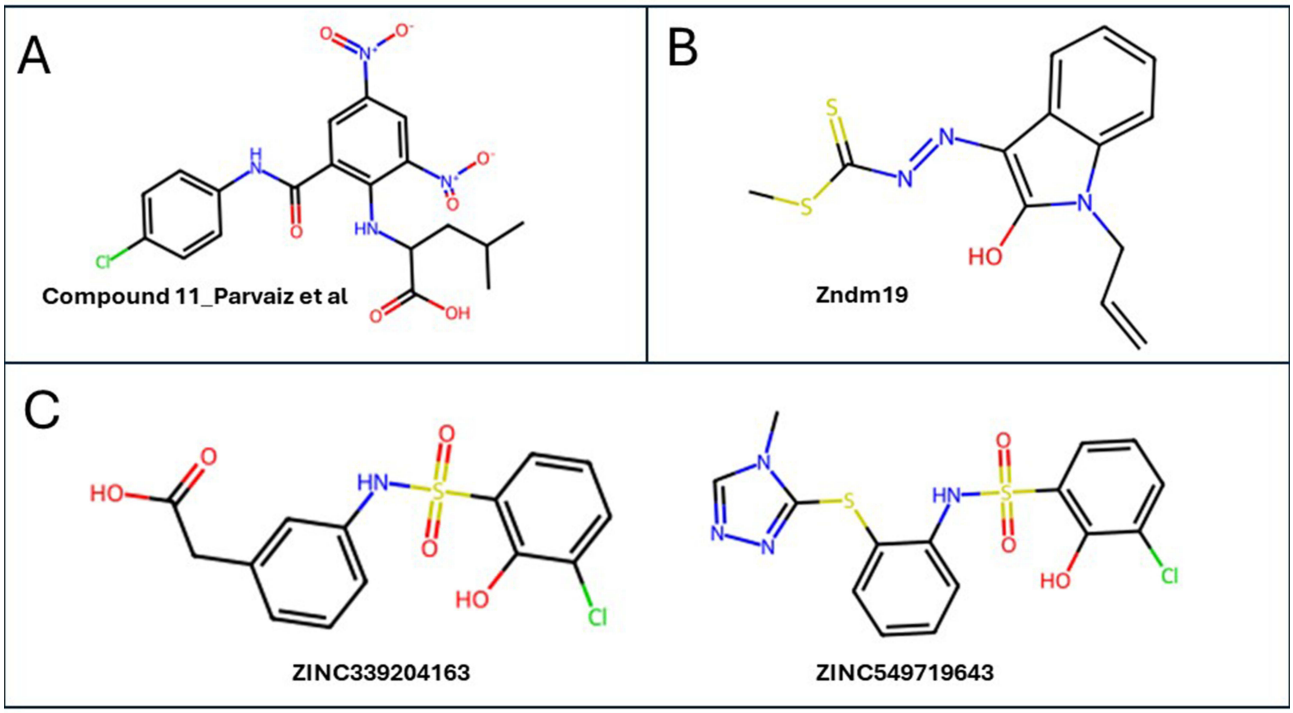

Figure 3 AmpC β -lactamase inhibitors discovered by use of in silico and AI techniques. (**A**) compound 11 - CMY-10 β-lactamase inhibitor identified by Parvaiz et al[80] "11" corresponds to the original compound number in Parvaiz et al (**B**) Zndm19 - metallo-β-lactamase NDM-1 inhibitor identified by Lv et al (**C**) non-covalent AmpC inhibitor ZINC339204163 and its analogue ZINC549719643 (Lyu et al[82]).

Parvaiz[80] in 2021. Their research centered on a class C β-lactamase called CMY-10 (AmpC β-lactamase), a key factor in drug resistance to a variety of *Enterobacteriaceae*. They initially employed the SILCS (Site-Identification by Ligand Competitive Saturation) method[106] to perform a detailed analysis of specific functional groups in the CMY-10 β-lactamase pocket and pharmacophore modeling. The created pharmacophore models were then used as a basis for a large-scale screening of a vast, in-house, chemical library. The library comprised around 700,000 compounds. The authors built a RF[107] model to screen compounds and identify potential inhibitors, considering various structural and chemical parameters such as molecular shape, electrostatic properties, polarity, and solubility. This approach yielded accurate predictions of the compounds' potential as effective inhibitors of the β-lactamase CMY-10. Consequently, 74 candidates were identified and subsequently subjected to experimental tests. Among these, compound 11_Parvaiz et al "CID2849737" (Figure 3A), demonstrated notable efficacy in β-lactamase assays and against multi-drug resistant (MDR) clinical strains, with MIC of 4 mg/mL against all tested MDR strains (*E. coli, Enterobacter cloacae* (*E. cloacae), Enterobacter agglomerans* (*E. agglomerans), Enterobacter alvei* (*E. alvei*) among others. Synergistic tests confirmed the potential of this compound as a β-lactam enhancer and β-lactamase inhibitor. *CID2849737* was particularly noted for its mechanism of action not based on a conventional β-lactam structure. It was proposed as a promising candidate for future optimizations and evaluations in view of clinical trials.

In the context of in silico screening of vast libraries for the identification of β-lactamase inhibitors, molecular docking methods have also proven their efficiency. As a recent example, the research conducted by Lyu et al[82] in 2019 highlighted the application of docking on a set of 170 million compounds, specifically targeting β-lactamase AmpC. The compounds that achieved the best scores were selected for synthesis and experimental evaluation of their interaction with AmpC, resulting in the synthesis of 44 molecules. That led to the identification of a non-covalent AmpC inhibitor, *ZINC339204163* "CID134613160" (Figure 3C), characterized by an inhibitory activity with a Ki of 1.3 µM. This inhibitor was identified as the most effective reversible agent against AmpC discovered by screening. It was subsequently

optimized to produce an analogue, *ZINC549719643* "CID132373998" (Figure 3C), exhibiting a Ki of 77 nM. This optimization placed it as one of the most effective non-covalent AmpC inhibitors to date.

Recently, we have developed a series of ML and DL models to predict non-covalent inhibitors of AmpC β-lactamase. Our work utilized a dataset comprising both inhibitors and non-inhibitors of AmpC to train SVM, RF, and Feed-Forward Neural Networks (FFNN).[108] The models demonstrated cross-validation accuracies between 80% and 82%, largely outperforming others recently reported models.[109] We provided an innovative approach for identifying new potential non-covalent inhibitors, non-covalent inhibitors could represent an alternative strategy against β-lactamases. The developed models are publicly available to foster innovation and collaboration and in this important field.

In 2023, Lv et al reported a study[81] aimed at identifying novel, effective inhibitors against the metallo-β-lactamase NDM-1. To achieve this, a curated set of 547 compounds, consisting of 271 inhibitors and 276 non-inhibitors of NDM-1, was used. These compounds were encoded using molecular fingerprints and served as the training dataset for a deep neural network (DNN). Post-training, the model showcased very good performance, achieving AUC of 0.90. It was then utilized to screen the Specs chemical library of over 210,000 compounds (2019 v. http://www.specs.net), leading to the selection of 6252 molecules. Molecular docking was subsequently applied for the investigation of the interactions between the inhibitor candidates and the NDM-1 enzyme, using Autodock4Zn[110] for the analysis. The candidate *Zndm19* "CID1221715" (illustrated in Figure 3B) emerged as a very efficient inhibitor of NDM-1. Experimental validation of *Zndm19*, via specific NDM-1 enzymatic activity assays, confirmed its potency with an IC50 value of 2.678 μM. Biological evaluations on mouse peritonitis infection models demonstrated its capability to reinstate the bactericidal effect of meropenem against NDM-1-positive *E. coli* strains. Furthermore, this inhibitor exhibited no cytotoxicity at less than 32 μg/mL towards a variety of human cell lines and enhanced survival rates in a mouse model of peritonitis, underlying its potential as a safe and effective therapy for carbapenem-resistant bacterial infections mediated by NDM-1.

Antimicrobial Peptides

The exploration of new antimicrobial agents using AI has also focused on antimicrobial peptides (AMPs).[111] AMPs are short sequences of amino acids essential to the innate immune response in various organisms, effectively combating a wide range of bacteria by disrupting their cell membranes. This mode of action poses a significant hurdle for AMR development, showcasing the AMP's robust antimicrobial ability. AMPs are typically 12−50 amino acids in length with 2−9 excess basic residues (arginine or lysine) and up to 50% hydrophobic amino acids; they fall into four major structural categories based on their amphiphilic conformations that can be performed or adopted after membrane interaction, namely, β-structures with 2−4 β-strands, amphipathic α-helices, loop structures, and extended structures.[112]

One of the first studies utilizing AI for the discovery of effective new AMPs was that of Cherkasov et al[74] in 2009, where the predictive capabilities of an artificial neural network (ANN) combined with chemical descriptors to identify and design effective AMPs were explored. By applying quantitative structure–activity relationship (QSAR) analysis, the team trained a model on existing peptide data that had been experimentally tested. This model was used to predict the activity of 100,000 virtual peptides, each made up of 9 amino acids. These were ranked according to their hypothetical antimicrobial potential, and among the most promising results, two peptides, *HHC-10* (KRWWKWIRW) and *HHC-36* (KRWWKWWRR), were identified as particularly effective in vitro against *Pseudomonas aeruginosa* (*P. aeruginosa*). These peptides were then tested in vitro against a range of resistant pathogens, including multidrug-resistant strains of *P. aeruginosa, MRSA, E. coli, Klebsiella pneumoniae* (*K. pneumoniae*) producers of extended-spectrum β-lactamase (ESBL), and *VRE faecalis* and *faecium*. They showed an MIC between 0.3 and 11 μM (see Table 1 for more detail), surpassing certain major antibiotics like tobramycin, ciprofloxacin, ceftazidime, and imipenem. In vivo tests further demonstrated the peptides' efficacy, with both *HHC-10* and *HHC-36* significantly reducing bacterial counts in a mouse model of invasive *S. aureus* infection, highlighting their efficacy. Additionally, these peptides revealed minimal toxicity, with negligible hemolytic activity at all concentrations tested, thereby confirming their pathogen-specific nature.

In 2018, Grafskaia et al[75] conducted a transcriptomic study on the *sea anemone Cnidopus japonicus* to extract, analyze, and assess its peptides and their antimicrobial activity. A total of 63,484 peptides were identified in the sea anemone and analyzed using various online ML models to detect the most promising peptides. AMPA,[113] ADAM,[114] and

CAMPR3[115] were utilized to estimate antimicrobial properties, along with predictors of eukaryotic cells, while CellPPD,[116] CPPpred,[117] and CPPred-RF[118] were employed to evaluate cytotoxicity. Consequently, 10 peptides were ultimately selected and synthesized. Among them, three peptides, *B4 KVKYFKRWLR, A3 NSVRNRVMLWRTKR*, and *B1 GIWTCRKKRA*, exhibited moderate antimicrobial activity. *B4* showed activity against both *E. coli* (Gram-negative) with a MIC of 64 µg/mL and *Bacillus subtilis* (*B. subtilis*) (Gram-positive) with a MIC of 32 µg/mL, indicating its broader spectrum of antimicrobial activity. On the other hand, *A3* and *B1* inhibited the growth of Gram-positive bacteria, *B. subtilis*, with a MIC of 128 µg/mL and 64 µg/mL, respectively, suggesting their potential use in targeting Gram-positive bacterial infections.

Recently, several studies have focused on the application of generative DL in the research of AMPs. Fields et al[76] employed a combination of ML and biophysical characteristics analysis to identify and develop AMPs derived from bacteriocins. An SVM model was trained on a dataset comprising 346 active bacteriocins and an equivalent number of inactive sequences. The sequences were encoded using the word-embedding algorithm Word2vec to transform each peptide into a high-dimensional vector space. This method encodes the sequences by capturing the relationships between amino acid trimers, thereby enhancing the SVM model's ability to discern structural and functional nuances of peptides, and effectively differentiate between active and inactive sequences. This model succeeded in isolating 676 candidates as potentially active from a library of generated peptides. Then, from the isolated 676 putative bacteriocins, they produced 28,895 peptides of 20 amino acids using a sliding window technique. Consequently, sixteen sequences were chosen based on physicochemical parameters, synthesized and tested, and among them, two peptides, peptide *3 IKKIGKKAAKKVIVKAIQAI* and *peptide 4 KKIGKKAAKKVIVKAIQAIV*, demonstrated the best activities. Peptide 3 demonstrated an MIC of 32 µM against *E. coli* and an MIC of 64 µM against *P. aeruginosa*. Peptide *4* displayed even better MICs, with an MIC of 2 µM against *E. coli* and a MIC of 16 µM against *P. aeruginosa*. Moreover, cytotoxicity assays demonstrated that these peptides specifically target bacterial membranes without adversely affecting mammalian cells.

Another recent study focused on predicting and designing AMPs, especially short-length AMPs.[77] The authors introduced Deep-AmPEP30, a method employing PseKRAAC (pseudo K-tuple reduced amino acid composition) and a CNN. This methodology was used to predict AMPs of about 30 amino acids, with performance comparable to other existing ML-based methods in a balanced benchmark dataset of 188 samples. The performance metrics of the model demonstrated an accuracy of 77% and achieved 85% of AUC. Deep-AmPEP30's utility was demonstrated through its application to the genome of *Candida glabrata*, a gut commensal fungus. The model successfully identified a 20 amino-acid peptide, *P3 FWELWKFLKSLWSIFPRRRP* with robust antibacterial activity against *B. subtilis* and *Vibrio parahaemolyticus*, comparable to ampicillin. This discovery showcases the high model's capability in identifying potent AMPs from genomic data.

Alongside these developments, in 2020 Dean and Walper[78] utilized a Variational Autoencoder (VAE), a type of generative DL algorithm. The VAE was trained on a database containing known active AMP sequences as well as their scrambled counterparts. These latter are variants where the order of amino acids is randomly altered, rendering the peptides biologically inactive. The VAE's ability to interpolate between active AMP sequences and their scrambled versions enabled the generation of novel active peptides. Among these, the peptides *P1 (RKLKKLWRKFR), P2 (RRFVKKVRKLVK*), and *P3 (FRWLRKWFRR)* demonstrated significant antimicrobial activity against various bacterial strains such as *E. coli, A.baumannii*, and *S.aureus*, showing promising EC50 values (see Table 1 for more details). These results validated the VAE's effectiveness in identifying potent AMPs and demonstrated the potential of AI-driven approaches in the rapid generation and optimization of new AMPs for specific therapeutic purposes.

In 2024, the exploration of AMPs as novel therapeutic agents fostered by AI continued to yield more promising results. A study conducted by Dong et al[79] targeted the antibiotic resistance of *Cutibacterium acnes* (*C. acnes*), implicated in acne vulgaris, which affects a vast majority of adolescents and young adults. Their efforts highlighted the use of a DL pipeline to generate peptides with specific inhibitory activity against this bacterium. The study began by training a base generator model on sequences of known active AMPs, exploring public data to capture a diversity of bioactive motifs. At this stage, pre-trained protein embedding (representations of protein sequences as numerical vectors), such as those from ProtTrans[119] and Evolutionary Scale Modeling (ESM),[120] was used to enrich the models' understanding of the complex relationships between amino acid sequences and their functions. To refine the model's

specificity towards *C. acnes*, it was fine-tuned using sequences of AMPs effective against this bacterium. A phylogenetic tree was constructed to select closely related bacterial species for training, improving the relevance of the generated peptides. The model generated 660,000 new peptide sequences, which were filtered by activity and hemolysis classifiers, reducing them to 24,579 candidates. These underwent further selection based on length and clustering, identifying 42 peptides for in vitro testing. Five peptides showed specific activity against *C. acnes* with MIC values of 2–4 μg/mL (see Table 1). Toxicity tests indicated low toxicity for all, suggesting a favorable safety profile.

Challenges and Future Opportunities in Antibacterial Drug Discovery

The discovery of new antibiotics is becoming more and more difficult. Various reasons can be cited. Indeed, the declining development of new antibiotics by the pharmaceutical industry results in part in a lack of sufficient investments.[121] Further, antibiotics were discovered largely through screening microbes for secondary metabolites that inhibited the growth of pathogenic bacteria. This approach has led to the discovery of many of the antibiotic classes used clinically, as β-lactams, aminoglycosides, polymyxins, glycopeptides, etc.[122] New approaches are needed to increase the discovery of new antibiotics while reducing the cost of early drug discovery for next- generation therapeutics. Today, metagenomics and metatranscriptomics are becoming important diagnostic tools for screening and detecting pathogens and antibiotic resistance genes, assessing the effects of antibiotics and characterizing the microbiome[123]. CRISPR-Cas9, evolving as a state-of-the-art genetic engineering tool for altering specific genes across diverse microorganisms, is a promising tool for addressing antibiotic resistance by selectively modifying genes in diverse microorganisms[124].

In the era of big data and next-generation sequencing, it is almost impossible to manually find a drug with low AMR risk. AI is a promising technology successfully used in drug discovery that will speed the development of efficient antibiotic alternatives. Evidently, in silico approaches in early drug discovery enable the exploration of rapidly expanded vast chemical spaces that are beyond the reach of traditional experimental approaches. Indeed, the recent progress of ML and DL offers possibilities to efficiently explore high-dimensional data, to speed up target identification, preclinical and clinical development[125] and design candidate compounds with desired properties, including those with antibacterial activity[126]. More widely, AI and in particular large language modelling have been outlined with possible benefits in personalized medicine, drug discovery, and applications to improve diagnosis and clinical decisions. Yet, limitations have also been noted as risks of bias, lack of transparency and reliability, and ethical consequences[127].

The advancements in ML and DL allow now the application of such algorithms for prediction of molecular properties in order to identify novel classes of antibiotics. Modern neural network-based molecular representations have the potential to increase the true positive rate of identifying structurally novel compounds with the desired bioactivity.[67,108] Traditional QSAR models[128] built on known scaffolds cannot identify novel scaffolds. The power of neural network approaches is in their capability to automatically learn molecular representation.[129] AI is particularly valuable for novel antibiotic discovery, where identifying structurally divergent from conventional antibiotics is crucial. For example, Stokes et al[67] utilized a cutting-edge approach, a directed message-passing deep neural network (D-MPNN), to discover structural classes of antibiotics allowing for the efficient substructure-based exploration of vast chemical spaces.

Although, AI demonstrated its utility in early drug discovery, it still faces important challenges. In order to create high-confidence predictive models, AI approaches need high quality and quantity reproducible data. In the context of antibacterial discovery, the compound libraries should contain molecules with physicochemical properties appropriate for antibacterial drugs, yet sufficiently diverse. Positive and negative training data with important chemical diversity are critical for building models capable of generating new scaffolds. A lack of diversity or balanced representation can lead to significant biases in the models' predictions. Furthermore, DL requires very large volumes of data to be able to extract the most relevant and important information from the inputs and achieve optimal performance. The availability, quantity, and quality of data is thus critical for models' performance. Several databases[125] provide detailed information on resistance genes, molecular mechanisms, chemical structures of antibiotics, among others. The Comprehensive Antibiotic Resistance Database (CARD)[130] accessible at https://card.mcmaster.ca/ compiles for instance detailed data on resistance genes, their mutations, and associated proteins, facilitating an understanding of the molecular mechanisms of resistance. The NCBI AMR (National Center for Biotechnology Information Antimicrobial Resistance) https://www.ncbi.nlm.nih.gov/pathogens/antimicrobial-resistance/ compiles genes, sequences, and phenotypic profiles related to AMR

in different bacterial species. MEGARes 3.0[131–133] includes resistance sequences for antimicrobial drugs and biocides. These databases focus on resistance trends and offer expanded data to develop more effective treatments. The advanced bioinformatics tool ResFinder of the Bacterial Antimicrobial Resistance Reference Gene Database[134] was designed to identify AMR genes in bacterial genomic sequences derived from next-generation sequencing data. Such genomics-based data combined with well-performing ML/DL –based tools can evolve toward accurate prediction of new chemical structures and activities of genomically encoded antibiotics.[135]

Employing AI will be crucial in the near future to efficiently treat heterogeneous and huge information on AMR by compiling genes, sequences, and phenotypic profiles, antibiotic structures and their mechanisms of action and side effects. Yet, AI models are often black boxes in nature, that is why decoding the structural diversity of compounds for antimicrobial and mechanistic investigations is critical[136]. In order to optimally exploit the available AMR, structural, and pharmacological data, integrating AI approaches based on high-dimensional modelling with mechanistic modelling approaches would overcome the current challenges, and would ultimately accelerate antimicrobial drug discovery.

Conclusion

The integration of AI in the fight against AMR has opened new horizons in the development of novel antibacterial agents. Here, we illustrated how AI can facilitate the identification and development of new promising agents with antimicrobial properties. Covering from the discovery of small-molecule inhibitors targeting specific bacterial virulence factors to the innovative design of AMPs and β-lactamase inhibitors, the implication of AI can change the paradigm in combating AMR. These achievements not only highlight the importance of AI in accelerating the discovery and optimization processes but also underline the urgent need for massive investment in AI technologies in drug discovery and in particular in antimicrobial research. Finally, to move forward, it is crucial to bridge the gap between the numerous AI-driven applications and current clinical investigations to foster the IA technological innovations toward new effective treatments related to AMR.

Acknowledgments

We acknowledge the French ANR (project NASPEC), Université Paris Cité and INSERM Institute for their support. We thank Pr. M. Ethève-Quelquejeu, Dr. L. Iannazzo (Université Paris Cité), Dr. M. Arthur (Sorbonne Université-Université Paris Cité) and Dr. K. Cariou (PSL University) for helpful discussions.

Disclosure

The authors report no conflicts of interest in this work.

References

1. Madigan MT, Bender KS, Buckley DH, Brock TD, Sattley WM, Stahl DA. *Brock Biology of Microorganisms*. 15th ed. ed. Pearson; 2018.
2. Hattori M, Taylor TD. The human intestinal microbiome: a new frontier of human biology. *DNA Res Int J Rapid Publ Rep Genes Genomes*. 2009;16(1):1–12. doi:10.1093/dnares/dsn033
3. Ogunrinola GA, Oyewale JO, Oshamika OO, Olasehinde GI. The human microbiome and its impacts on health. *Int J Microbiol*. 2020;2020:8045646. doi:10.1155/2020/8045646
4. Malard F, Dore J, Gaugler B, Mohty M. Introduction to host microbiome symbiosis in health and disease. *Mucosal Immunol*. 2021;14(3):547–554. doi:10.1038/s41385-020-00365-4
5. Fuchs TM. Molecular mechanisms of bacterial pathogenicity. *Naturwissenschaften*. 1998;85(3):99–108. doi:10.1007/s001140050463
6. Jones KE, Patel NG, Levy MA, et al. Global trends in emerging infectious diseases. *Nature*. 2008;451(7181):7181):990–993. doi:10.1038/nature06536
7. Bertuzzo E, Azaele S, Maritan A, Gatto M, Rodriguez-Iturbe I, Rinaldo A. On the space-time evolution of a cholera epidemic. *Water Resour Res*. 2008;44(1). doi:10.1029/2007WR006211
8. Mukhtar O. The miraculous fever tree: malaria, medicine and the cure that changed the world. *BMJ*. 2003;327(7407):169. doi:10.1136/bmj.327.7407.169-a
9. Porter R *The Greatest Benefit to Mankind: A Medical History of Humanity*. New York: W. W. Norton; 1998. Available from http://archive.org/details/greatestbenefitt00port. Accessed October 7, 2024.
10. Gould K. Antibiotics: from prehistory to the present day. *J Antimicrob Chemother*. 2016;71(3):572–575. doi:10.1093/jac/dkv484
11. Pasteur L. *Germ Theory and Its Applications to Medicine and on the Antiseptic Principle of the Practice of Surgery*. Prometheus; 1996.
12. Valent P, Groner B, Schumacher U, et al. Paul Ehrlich (1854-1915) and his contributions to the foundation and birth of translational medicine. *J Innate Immun*. 2016;8(2):111–120. doi:10.1159/000443526

13. Gaynes R. The Discovery of Penicillin—New Insights After More Than 75 Years of Clinical Use. *Emerg Infect Dis*. 2017;23(5):849–853. doi:10.3201/eid2305.161556
14. Mohr KI. History of antibiotics research. *Curr Top Microbiol Immunol*. 2016;398:237–272. doi:10.1007/82_2016_499
15. Kong KF, Schneper L, Mathee K. Beta-lactam antibiotics: from antibiosis to resistance and bacteriology. *APMIS Acta Pathol Microbiol Immunol Scand*. 2010;118(1):1–36. doi:10.1111/j.1600-0463.2009.02563.x
16. Mingeot-Leclercq MP, Glupczynski Y, Tulkens PM. Aminoglycosides: activity and resistance. *Antimicrob Agents Chemother*. 1999;43(4):727–737. doi:10.1128/AAC.43.4.727
17. Revillo Imbernon J, Weibel JM, Ennifar E, Prévost G, Kellenberger E. Structural analysis of neomycin B and kanamycin A binding Aminoglycosides Modifying Enzymes (AME) and bacterial ribosomal RNA. *Mol Inform*. 2024;43(7):e202300339. doi:10.1002/minf.202300339
18. Chopra I, Roberts M. Tetracycline antibiotics: mode of action, applications, molecular biology, and epidemiology of bacterial resistance. *Microbiol Mol Biol Rev*. 2001;65(2):232–260. doi:10.1128/MMBR.65.2.232-260.2001
19. Kanoh S, Rubin BK. Mechanisms of action and clinical application of macrolides as immunomodulatory medications. *Clin Microbiol Rev*. 2010;23(3):590–615. doi:10.1128/CMR.00078-09
20. Chan T, Bunce PE. Fluoroquinolone antimicrobial drugs. *CMAJ Can Med Assoc J*. 2017;189(17):E638. doi:10.1503/cmaj.160938
21. Reygaert WC. An overview of the antimicrobial resistance mechanisms of bacteria. *AIMS Microbiol*. 2018;4(3):482–501. doi:10.3934/microbiol.2018.3.482
22. Blair JMA, Webber MA, Baylay AJ, Ogbolu DO, Piddock LJV. Molecular mechanisms of antibiotic resistance. *Nat Rev Microbiol*. 2015;13(1):42–51. doi:10.1038/nrmicro3380
23. Murray CJL, Ikuta KS, Sharara F, Antimicrobial Resistance Collaborators. Global burden of bacterial antimicrobial resistance in 2019: a systematic analysis. *Lancet Lond Engl*. 2022;399(10325):629–655. doi:10.1016/S0140-6736(21)02724-0
24. O'Neill J Antimicrobial resistance: tackling a crisis for the health and wealth of nations. *Rev Antimicrob Resist*; 2014. Available from: https://cir.nii.ac.jp/crid/1370857593729357568. Accessed January 18, 2024.
25. Environment UN.Bracing for superbugs: strengthening environmental action in the one health response to antimicrobial resistance.*UNEP - UN Environment Programme*.2023;30:Available from:http://www.unep.org/resources/superbugs/environmental-action. Accessed, 2023.
26. Lerminiaux NA, Cameron ADS. Horizontal transfer of antibiotic resistance genes in clinical environments. *Can J Microbiol*. 2019;65(1):34–44. doi:10.1139/cjm-2018-0275
27. Ellabaan MMH, Munck C, Porse A, Imamovic L, Sommer MOA. Forecasting the dissemination of antibiotic resistance genes across bacterial genomes. *Nat Commun*. 2021;12(1):2435. doi:10.1038/s41467-021-22757-1
28. Tenover FC. Mechanisms of antimicrobial resistance in bacteria. *Am J Med*. 2006;119(6, Supplement 1):S3–S10. doi:10.1016/j.amjmed.2006.03.011
29. Lv Q, Zhou F, Liu X, Zhi L. Artificial intelligence in small molecule drug discovery from 2018 to 2023: does it really work? *Bioorg Chem*. 2023;141:106894. doi:10.1016/j.bioorg.2023.106894
30. Amin D, Garzón-Orjuela N, Garcia Pereira A, Parveen S, Vornhagen H, Vellinga A. Artificial intelligence to improve antibiotic prescribing: a systematic review. *Antibiot Basel Switz*. 2023;12(8):1293. doi:10.3390/antibiotics12081293
31. Kazemzadeh S, Yu J, Jamshy S, et al. Deep learning detection of active pulmonary tuberculosis at chest radiography matched the clinical performance of radiologists. *Radiology*. 2023;306(1):124–137. doi:10.1148/radiol.212213
32. Pascucci M, Royer G, Adamek J, et al. AI-based mobile application to fight antibiotic resistance. *Nat Commun*. 2021;12(1):1173. doi:10.1038/s41467-021-21187-3
33. Scherr TF, Gupta S, Wright DW, Haselton FR. Mobile phone imaging and cloud-based analysis for standardized malaria detection and reporting. *Sci Rep*. 2016;6(1):28645. doi:10.1038/srep28645
34. Hatfull GF, Dedrick RM, Schooley RT. Phage therapy for antibiotic-resistant bacterial infections. *Annu Rev Med*. 2022;73(1):197–211. doi:10.1146/annurev-med-080219-122208
35. McNair K, Bailey BA, Edwards RA. PHACTS, a computational approach to classifying the lifestyle of phages. *Bioinforma Oxf Engl*. 2012;28(5):614–618. doi:10.1093/bioinformatics/bts014
36. Tynecki P, Guziński A, Kazimierczak J, Jadczuk M, Dastych J, Onisko A. PhageAI - bacteriophage life cycle recognition with machine learning and natural language processing. *BioRxiv*. 2020;2020:198606. doi:10.1101/2020.07.11.198606
37. Seguritan V, Alves N, Arnoult M, et al. Artificial neural networks trained to detect viral and phage structural proteins. *PLoS Comput Biol*. 2012;8(8):e1002657. doi:10.1371/journal.pcbi.1002657
38. Manavalan B, Shin TH, Lee G. PVP-SVM: sequence-based prediction of phage virion proteins using a support vector machine. *Front Microbiol*. 2018;9:476. doi:10.3389/fmicb.2018.00476
39. Edwards RA, McNair K, Faust K, Raes J, Dutilh BE. Computational approaches to predict bacteriophage-host relationships. *FEMS Microbiol Rev*. 2016;40(2):258–272. doi:10.1093/femsre/fuv048
40. Villarroel J, Kleinheinz KA, Jurtz VI, et al. HostPhinder: a phage host prediction tool. *Viruses*. 2016;8(5):116. doi:10.3390/v8050116
41. Zhang M, Yang L, Ren J, Ahlgren NA, Fuhrman JA, Sun F. Prediction of virus-host infectious association by supervised learning methods. *BMC Bioinf*. 2017;18(Suppl 3):60. doi:10.1186/s12859-017-1473-7
42. Amgarten D, Braga LPP, da Silva AM, Setubal JC. MARVEL, a tool for prediction of bacteriophage sequences in metagenomic bins. *Front Genet*. 2018;9:9. doi:10.3389/fgene.2018.00304
43. Ren J, Ahlgren NA, Lu YY, Fuhrman JA, Sun F. VirFinder: a novel k-mer based tool for identifying viral sequences from assembled metagenomic data. *Microbiome*. 2017;5(1):69. doi:10.1186/s40168-017-0283-5
44. Roux S, Enault F, Hurwitz BL, Sullivan MB. VirSorter: mining viral signal from microbial genomic data. *PeerJ*. 2015;3:e985. doi:10.7717/peerj.985
45. Shang J, Jiang J, Sun Y. Bacteriophage classification for assembled contigs using graph convolutional network. *Bioinforma Oxf Engl*. 2021;37(Suppl_1):i25–i33. doi:10.1093/bioinformatics/btab293
46. Chibani CM, Farr A, Klama S, Dietrich S, Liesegang H. Classifying the unclassified: a phage classification method. *Viruses*. 2019;11(2):195. doi:10.3390/v11020195

47. Silver LL. Challenges of antibacterial discovery. *Clin Microbiol Rev*. 2011;24(1):71–109. doi:10.1128/CMR.00030-10
48. Institute TD. WHO global priority pathogens list of antibiotic-resistant bacteria. December 2, 2021. Available from: https://www.doherty.edu.au/news-events/news/who-global-priority-pathogens-list-of-antibiotic-resistant-bacteria/. Accessed November 18, 2022.
49. Cortes C, Vapnik V. Support-vector networks. *Mach Learn*. 1995;20(3):273–297. doi:10.1023/A:1022627411411
50. Keerthi SS. Efficient tuning of SVM hyperparameters using radius/margin bound and iterative algorithms. *IEEE Trans Neural Netw*. 2002;13(5):1225–1229. doi:10.1109/TNN.2002.1031955
51. Breiman L. Random Forests. *Mach Learn*. 2001;45(1):5–32. doi:10.1023/A:1010933404324
52. Iranzad R, Liu X. A review of random forest-based feature selection methods for data science education and applications. *Int J Data Sci Anal*. 2024. doi:10.1007/s41060-024-00509-w
53. Liaw A, Wiener M. Classification and regression by randomForest. *R News*. 2002;2.
54. Bharadiya JP. A review of bayesian machine learning principles, methods, and applications. *Int J Innov Sci Res Technol*. 2023;8(5):2033–2038. doi:10.5281/zenodo.8002438
55. Vogt M. Using deep neural networks to explore chemical space. *Expert Opin Drug Discov*. 2022;17(3):297–304. doi:10.1080/17460441.2022.2019704
56. He J, You H, Sandström E, et al. Molecular optimization by capturing chemist's intuition using deep neural networks. *J Cheminformatics*. 2021;13(1):26. doi:10.1186/s13321-021-00497-0
57. Kim P. Convolutional Neural Network. In: Kim P, editor. *MATLAB Deep Learning: With Machine Learning, Neural Networks and Artificial Intelligence*. Apress; 2017:121–147. doi:10.1007/978-1-4842-2845-6_6
58. Scarselli F, Gori M, Tsoi AC, Hagenbuchner M, Monfardini G. The graph neural network model. *IEEE Trans Neural Netw*. 2009;20(1):61–80. doi:10.1109/TNN.2008.2005605
59. Kubara K Introduction to message passing neural networks. Medium. October 5, 2020. Available from.: https://towardsdatascience.com/introduction-to-message-passing-neural-networks-e670dc103a87. Accessed April 12, 2024.
60. Swanson KW. *Message Passing Neural Networks for Molecular Property Prediction*. Thesis. Massachusetts Institute of Technology; 2019. Available from:https://dspace.mit.edu/handle/1721.1/123133. Accessed October 1, 2024.
61. Pinheiro Cinelli L, Araújo Marins M, da Silva EA B, Lima Netto S. Variational Methods for Machine Learning with Applications to Deep Networks. In: Cinelli LP, Marins MA, Barros da Silva EA, Netto SL editors. *Variational Autoencoder*. Springer International Publishing; 2021:111–149. doi:10.1007/978-3-030-70679-1_5.
62. Hu X, Prehna G, Stebbins CE. Targeting plague virulence factors: a combined machine learning method and multiple conformational virtual screening for the discovery of Yersinia protein kinase A inhibitors. *J Med Chem*. 2007;50(17):3980–3983. doi:10.1021/jm070645a
63. Zietz BP, Dunkelberg H. The history of the plague and the research on the causative agent *Yersinia pestis*. *Int J Hyg Environ Health*. 2004;207(2):165–178. doi:10.1078/1438-4639-00259
64. Hearst MA, Dumais ST, Osuna E, Platt J, Scholkopf B. Support vector machines. *IEEE Intell Syst Their Appl*. 1998;13(4):18–28. doi:10.1109/5254.708428
65. Wang L, Le X, Li L, et al. Discovering new agents active against methicillin-resistant staphylococcus aureus with ligand-based approaches. *J Chem Inf Model*. 2014;54(11):3186–3197. doi:10.1021/ci500253q
66. Ekins S, Madrid PB, Sarker M, et al. Combining metabolite-based pharmacophores with bayesian machine learning models for mycobacterium tuberculosis drug discovery. *PLoS One*. 2015;10(10):e0141076. doi:10.1371/journal.pone.0141076
67. Stokes JM, Yang K, Swanson K, et al. A deep learning approach to antibiotic discovery. *Cell*. 2020;180(4):688–702.e13. doi:10.1016/j.cell.2020.01.021
68. Liu G, Catacutan DB, Rathod K, et al. Deep learning-guided discovery of an antibiotic targeting Acinetobacter baumannii. *Nat Chem Biol*. 2023;19(11):1342–1350. doi:10.1038/s41589-023-01349-8
69. Wong F, Zheng EJ, Valeri JA, et al. Discovery of a structural class of antibiotics with explainable deep learning. *Nature*. 2024;626(7997):7997):177–185. doi:10.1038/s41586-023-06887-8
70. Lakhundi S, Zhang K. Methicillin-resistant staphylococcus aureus: molecular characterization, evolution, and epidemiology. *Clin Microbiol Rev*. 2018;31(4):e00020–18. doi:10.1128/CMR.00020-18
71. ChEMBL Database.ChEMBL Database. Available from: https://www.ebi.ac.uk/chembl/. Accessed April 12, 2024.
72. Yang FJ An Implementation of Naive Bayes Classifier. In: *2018 International Conference on Computational Science and Computational Intelligence (CSCI)*. (CSCI) 2018:301–306. doi:10.1109/CSCI46756.2018.00065.
73. Kramer O. *K-Nearest Neighbors*. In: Kramer O editor. Dimensionality Reduction with Unsupervised Nearest Neighbors. Springer; 2013:13–23. doi:10.1007/978-3-642-38652-7_2.
74. Cherkasov A, Hilpert K, Jenssen H, et al. Use of artificial intelligence in the design of small peptide antibiotics effective against a broad spectrum of highly antibiotic-resistant superbugs. *ACS Chem Biol*. 2009;4(1):65–74. doi:10.1021/cb800240j
75. Grafskaia EN, Polina NF, Babenko VV, et al. Discovery of novel antimicrobial peptides: a transcriptomic study of the sea anemone Cnidopus japonicus. *J Bioinform Comput Biol*. 2018;16(2):1840006. doi:10.1142/S0219720018400061
76. Fields FR, Freed SD, Carothers KE, et al. Novel antimicrobial peptide discovery using machine learning and biophysical selection of minimal bacteriocin domains. *Drug Dev Res*. 2020;81(1):43–51. doi:10.1002/ddr.21601
77. Yan J, Bhadra P, Li A, et al. Deep-AmPEP30: improve short antimicrobial peptides prediction with deep learning. *Mol Ther Nucleic Acids*. 2020;20:882–894. doi:10.1016/j.omtn.2020.05.006
78. Dean SN, Walper SA. Variational autoencoder for generation of antimicrobial peptides. *ACS Omega*. 2020;5(33):20746. doi:10.1021/acsomega.0c00442
79. Dong Q, Wang S, Miao Y, Luo H, Weng Z, Yu L. Novel antimicrobial peptides against Cutibacterium acnes designed by deep learning. *Sci Rep*. 2024;14(1):14. doi:10.1038/s41598-024-55205-3
80. Parvaiz N, Ahmad F, Yu W, MacKerell AD, Azam SS. Discovery of beta-lactamase CMY-10 inhibitors for combination therapy against multi-drug resistant Enterobacteriaceae. *PLoS One*. 2021;16(1):e0244967. doi:10.1371/journal.pone.0244967
81. Lv H, Zhu Z, Qian C, et al. Discovery of isatin-β-methyldithiocarbazate derivatives as New Delhi metallo- β-lactamase-1 (NDM-1) inhibitors against NDM-1 producing clinical isolates. *Biomed Pharmacother*. 2023;166:115439. doi:10.1016/j.biopha.2023.115439

82. Lyu J, Wang S, Balius TE, et al. Ultra-large library docking for discovering new chemotypes. *Nature*. 2019;566(7743):224–229. doi:10.1038/s41586-019-0917-9
83. Gagneux S. Ecology and evolution of mycobacterium tuberculosis. *Nat Rev Microbiol*. 2018;16(4):202–213. doi:10.1038/nrmicro.2018.8
84. Ekins S, Reynolds RC, Franzblau SG, Wan B, Freundlich JS, Bunin BA. Enhancing hit identification in mycobacterium tuberculosis drug discovery using validated dual-event bayesian models. *PLoS One*. 2013;8(5):e63240. doi:10.1371/journal.pone.0063240
85. Ekins S, Reynolds RC, Kim H, et al. Bayesian models leveraging bioactivity and cytotoxicity information for drug discovery. *Chem Biol*. 2013;20(3):370–378. doi:10.1016/j.chembiol.2013.01.011
86. Ekins S, Freundlich JS, Reynolds RC. Fusing dual-event data sets for Mycobacterium tuberculosis machine learning models and their evaluation. *J Chem Inf Model*. 2013;53(11):3054–3063. doi:10.1021/ci400480s
87. Asinex. Asinex.com. Available from: https://www.asinex.com/. Accessed April 12, 2024.
88. Commissioner O of the. U.S. Food and Drug Administration. FDA. November 4, 2024. Available from: https://www.fda.gov/. Accessed April 12, 2024.
89. Tillery LM, Barrett KF, Dranow DM, et al. Toward a structome of Acinetobacter baumannii drug targets. *Protein Sci Publ Protein Soc*. 2020;29(3):789–802. doi:10.1002/pro.3826
90. Sanchez-Lengeling B, Reif E, Pearce A, Wiltschko AB. A gentle introduction to graph neural networks. *Distill*. 2021;6(9):e33. doi:10.23915/distill.00033
91. Mcule. mcule database. Available from: https://mcule.com/database/. Accessed May 6, 2024.
92. Broad Institute database. Broad Institute. 2017. Available from: https://www.broadinstitute.org/epigenomics/data. Accessed May 6, 2024.
93. CDC. Centers for Disease Control and Prevention. Centers for Disease Control and Prevention. 2024. Available from: https://www.cdc.gov/index.htm. Accessed April 12, 2024.
94. Bradford PA. Extended-spectrum β-lactamases in the 21st century: characterization, epidemiology, and detection of this important resistance threat. *Clin Microbiol Rev*. 2001;14(4):933–951. doi:10.1128/CMR.14.4.933-951.2001
95. Paterson DL, Bonomo RA. Extended-spectrum beta-lactamases: a clinical update. *Clin Microbiol Rev*. 2005;18(4):657–686. doi:10.1128/CMR.18.4.657-686.2005
96. Nordmann P, Cuzon G, Naas T. The real threat of Klebsiella pneumoniae carbapenemase-producing bacteria. *Lancet Infect Dis*. 2009;9(4):228–236. doi:10.1016/S1473-3099(09)70054-4
97. Khan AU, Maryam L, Zarrilli R. Structure, genetics and worldwide spread of New Delhi Metallo-β-lactamase (NDM): a threat to public health. *BMC Microbiol*. 2017;17(1):101. doi:10.1186/s12866-017-1012-8
98. Meletis G. Carbapenem resistance: overview of the problem and future perspectives. *Therapeutic Advances in Infectious Disease*. 2016;3(1):15–21. doi:10.1177/2049936115621709
99. Jacoby GA. AmpC β-Lactamases. *Clin Microbiol Rev*. 2009;22(1):161–182. doi:10.1128/CMR.00036-08
100. Bouchet F, Barnier JP, Sayah I, et al. Ruthenium-catalyzed cycloaddition for introducing chemical diversity in second-generation β-lactamase inhibitors. *ChemMedChem*. 2023;18(9):e202300077. doi:10.1002/cmdc.202300077
101. Lahiri SD, Mangani S, Durand-Reville T, et al. Structural insight into potent broad-spectrum inhibition with reversible recyclization mechanism: avibactam in complex with CTX-M-15 and pseudomonas aeruginosa AmpC β-lactamases. *Antimicrob Agents Chemother*. 2013;57(6):2496–2505. doi:10.1128/AAC.02247-12
102. Lomovskaya O, Sun D, Rubio-Aparicio D, et al. Vaborbactam: spectrum of beta-lactamase inhibition and impact of resistance mechanisms on activity in Enterobacteriaceae. *Antimicrob Agents Chemother*. 2017;61(11):e01443–17. doi:10.1128/AAC.01443-17
103. Langley GW, Cain R, Tyrrell JM, et al. Profiling interactions of vaborbactam with metallo-β-lactamases. *Bioorg Med Chem Lett*. 2019;29(15):1981–1984. doi:10.1016/j.bmcl.2019.05.031
104. Livermore DM, Mushtaq S, Vickers A, Woodford N. Activity of aztreonam/avibactam against metallo-β-lactamase-producing Enterobacterales from the UK: impact of penicillin-binding protein-3 inserts and CMY-42 β-lactamase in Escherichia coli. *Int J Antimicrob Agents*. 2023;61(5):106776. doi:10.1016/j.ijantimicag.2023.106776
105. Drawz SM, Bonomo RA. Three decades of β-lactamase inhibitors. *Clin Microbiol Rev*. 2010;23(1):160–201. doi:10.1128/CMR.00037-09
106. Raman EP, Yu W, Guvench O, Mackerell AD. Reproducing crystal binding modes of ligand functional groups using Site-Identification by Ligand Competitive Saturation (SILCS) simulations. *J Chem Inf Model*. 2011;51(4):877–896. doi:10.1021/ci100462t
107. Wójcikowski M, Ballester PJ, Siedlecki P. Performance of machine-learning scoring functions in structure-based virtual screening. *Sci Rep*. 2017;7(1). doi:10.1038/srep46710
108. Bagdad Y, Sisquellas M, Arthur M, Miteva MA. Machine learning and deep learning models for predicting noncovalent inhibitors of AmpC β-lactamase. *ACS Omega*. 2024;9(40):41334–41342. doi:10.1021/acsomega.4c03834
109. Anant PS, Gupta P. Application of machine learning in understanding bioactivity of beta-lactamase AmpC. *J Phys Conf Ser*. 2022;2273(1):012005. doi:10.1088/1742-6596/2273/1/012005
110. Santos-Martins D, Forli S, Ramos MJ, Olson AJ. AutoDock4Zn: an improved autodock force field for small-molecule docking to zinc metalloproteins. *J Chem Inf Model*. 2014;54(8):2371–2379. doi:10.1021/ci500209e
111. Yan J, Cai J, Zhang B, Wang Y, Wong DF, Siu SWI. Recent progress in the discovery and design of antimicrobial peptides using traditional machine learning and deep learning. *Antibiotics*. 2022;11(10):1451. doi:10.3390/antibiotics11101451
112. Hancock REW, Sahl HG. Antimicrobial and host-defense peptides as new anti-infective therapeutic strategies. *Nat Biotechnol*. 2006;24(12):1551–1557. doi:10.1038/nbt1267
113. Torrent M, Di Tommaso P, Pulido D, et al. AMPA: an automated web server for prediction of protein antimicrobial regions. *Bioinformatics*. 2012;28(1):130–131. doi:10.1093/bioinformatics/btr604
114. Lee HT, Lee CC, Yang JR, Lai JZC, Chang KY. A large-scale structural classification of antimicrobial peptides. *BioMed Res Int*. 2015;2015:475062. doi:10.1155/2015/475062
115. Waghu FH, Barai RS, Gurung P, Idicula-Thomas S. CAMPR3: a database on sequences, structures and signatures of antimicrobial peptides. *Nucleic Acids Res*. 2016;44(Database issue):D1094–D1097. doi:10.1093/nar/gkv1051
116. Osorio D, Rondón-Villarreal P, Torres R. Peptides: a package for data mining of antimicrobial peptides. *R J*. 2015;7(1):4. doi:10.32614/RJ-2015-001

117. Holton TA, Pollastri G, Shields DC, Mooney C. CPPpred: prediction of cell penetrating peptides. *Bioinforma Oxf Engl*. 2013;29(23):3094–3096. doi:10.1093/bioinformatics/btt518
118. Wei L, Xing P, Su R, Shi G, Ma ZS, Zou Q. CPPred-RF: a sequence-based predictor for identifying cell-penetrating peptides and their uptake efficiency. *J Proteome Res*. 2017;16(5):2044–2053. doi:10.1021/acs.jproteome.7b00019
119. Elnaggar A, Heinzinger M, Dallago C, et al. ProtTrans: towards cracking the language of life's code through self-supervised deep learning and high performance computing. *IEEE Transactions on Pattern Analysis and Machine Intelligence*. 2021;44(10):7112–7127. doi:10.48550/arXiv.2007.06225
120. Lin Z, Akin H, Rao R, et al. Language models of protein sequences at the scale of evolution enable accurate structure prediction. *BioRxiv*. 2022;2022:500902. doi:10.1101/2022.07.20.500902
121. Analysis shows continued deficiencies in antibiotic development since 2014. July 22, 2019. Available from: https://pew.org/30FFj0e. Accessed October 7, 2024.
122. Wright GD. Opportunities for natural products in 21st century antibiotic discovery. *Nat Prod Rep*. 2017;34(7):694–701. doi:10.1039/c7np00019g
123. Asante J, Osei Sekyere J. Understanding antimicrobial discovery and resistance from a metagenomic and metatranscriptomic perspective: advances and applications. *Environ Microbiol Rep*. 2019;11(2):62–86. doi:10.1111/1758-2229.12735
124. Pandey P, Vavilala SL. From gene editing to biofilm busting: CRISPR-CAS9 against antibiotic resistance—a review. *Cell Biochem Biophys*. 2024;82(2):549–560. doi:10.1007/s12013-024-01276-y
125. Talat A, Khan AU. Artificial intelligence as a smart approach to develop antimicrobial drug molecules: a paradigm to combat drug-resistant infections. *Drug Discov Today*. 2023;28(4):103491. doi:10.1016/j.drudis.2023.103491
126. Cesaro A, Bagheri M, Torres M, Wan F, de la Fuente-Nunez C. Deep learning tools to accelerate antibiotic discovery. *Expert Opin Drug Discov*. 2023;18(11):1245–1257. doi:10.1080/17460441.2023.2250721
127. Sallam M. ChatGPT utility in healthcare education, research, and practice: systematic review on the promising perspectives and valid concerns. *Healthcare*. 2023;11(6):887. doi:10.3390/healthcare11060887
128. Muratov EN, Bajorath J, Sheridan RP, et al. QSAR without borders. *Chem Soc Rev*. 2020;49(11):3525–3564. doi:10.1039/d0cs00098a
129. Yang K, Swanson K, Jin W, et al. Analyzing Learned Molecular Representations for Property Prediction. *J Chem Inf Model*. 2019;59(8):3370–3388. doi:10.1021/acs.jcim.9b00237
130. McArthur AG, Waglechner N, Nizam F, et al. The comprehensive antibiotic resistance database. *Antimicrob Agents Chemother*. 2013;57(7):3348. doi:10.1128/AAC.00419-13
131. Bonin N, Doster E, Worley H, et al. MEGARes and AMR++, v3.0: an updated comprehensive database of antimicrobial resistance determinants and an improved software pipeline for classification using high-throughput sequencing. *Nucleic Acids Res*. 2023:51(D1):D744–D752. doi:10.1093/nar/gkac1047.
132. Doster E, Lakin SM, Dean CJ, et al. MEGARes 2.0: a database for classification of antimicrobial drug, biocide and metal resistance determinants in metagenomic sequence data. *Nucleic Acids Res*. 2020;48(D1):D561–D569. doi:10.1093/nar/gkz1010
133. Lakin SM, Dean C, Noyes NR, et al. MEGARes: an antimicrobial resistance database for high throughput sequencing. *Nucleic Acids Res*. 2017;45(Database issue):D574–D580. doi:10.1093/nar/gkw1009
134. Florensa AF, Kaas RS, Clausen PTLC, Aytan-Aktug D, Aarestrup FM. ResFinder – an open online resource for identification of antimicrobial resistance genes in next-generation sequencing data and prediction of phenotypes from genotypes. *Microb Genomics*. 2022;8(1):000748. doi:10.1099/mgen.0.000748
135. Skinnider MA, Johnston CW, Gunabalasingam M, et al. Comprehensive prediction of secondary metabolite structure and biological activity from microbial genome sequences. *Nat Commun*. 2020;11(1):6058. doi:10.1038/s41467-020-19986-1
136. Qiu Z, Huang R, Wu Y, Li X, Sun C, Ma Y. Decoding the structural diversity: a new horizon in antimicrobial prospecting and mechanistic investigation. *Microb Drug Resist Larchmt N*. 2024;30(6):254–272. doi:10.1089/mdr.2023.0232