

基于人工智能的新型抗菌肽的筛选、设计和优化

王晓宇 雒家玉 于晓琳

中山大学附属口腔医院, 光华口腔医学院, 广东省口腔医学重点实验室, 广东省口腔疾病临床医学研究中心, 广州 510055

通信作者: 于晓琳, Email: yuxlin3@mail.sysu.edu.cn



于晓琳

【摘要】 抗菌肽(AMP)凭借广谱抗菌、低耐药诱导等优势,已成为新型抗菌制剂的研究热点。传统AMP的设计依赖天然产物分离提取和有限的序列修饰,导致新型AMP挖掘范围较窄,人工筛选效率低。随着人工智能(AI)技术的快速发展,AI已在AMP虚拟筛选、抑菌活性预评估、AMP的从头生成及序列定向优化领域取得突破,可

大幅缩短传统肽分子筛选周期、降低实验室合成与验证成本。本文通过梳理主要AMP数据库的特征和相关AI算法的技术原理,阐明AI在AMP筛选、设计和优化中的应用策略,并分析当前研究中存在的数据集标准化不足、体内验证体系不完善及学科交叉衔接薄弱等问题,为新型AMP的研发和临床应用提供参考。

【关键词】 人工智能; 抗菌肽; 机器学习; 深度学习

基金项目: 国家自然科学基金(U22A20316, 81801012); 中山大学大学生创新创业训练计划项目(20260616)

引用著录格式: 王晓宇, 雒家玉, 于晓琳. 基于人工智能的新型抗菌肽的筛选、设计和优化[J/OL]. 中华口腔医学研究杂志(电子版), 2026, 20(4): 279-293.

DOI: 10.3877/cma.j.issn.1674-1366.2026.04.005

Screening, design, and optimization of novel antimicrobial peptides based on artificial intelligence

Wang Xiaoyu, Luo Jiayu, Yu Xiaolin

Hospital of Stomatology, Guanghua School of Stomatology, Sun Yat-sen University, Guangdong Provincial Key Laboratory of Stomatology, Guangdong Provincial Clinical Research Center of Oral Diseases, Guangzhou 510055, China

Corresponding author: Yu Xiaolin, Email: yuxlin3@mail.sysu.edu.cn

【Abstract】 Antimicrobial peptides (AMPs) have become a major focus in the development of novel antimicrobial agents due to their broad-spectrum antimicrobial properties and extremely low risk of inducing resistance. Traditional AMP

design strategies have relied primarily on the isolation and extraction of natural products, along with limited sequence modifications, which has greatly restricted the discovery of new candidate molecules and led to inefficient *in vitro* screening processes. With rapid advancements in artificial intelligence (AI), significant breakthroughs have been achieved in the virtual screening of AMPs, preliminary prediction of antibacterial activity, de novo design of peptide sequences, and targeted optimization of peptide structures and functions, thereby significantly shortening research timelines and reducing overall development costs. This article systematically reviews the core features of mainstream antimicrobial peptide databases and the fundamental mechanisms of commonly used AI algorithms, and further elaborates on feasible strategies for applying AI to antimicrobial peptide screening, design, and structural optimization. At the same time, it identifies key bottlenecks in current research, such as the lack of standardized datasets, immature *in vivo* evaluation systems, and insufficient interdisciplinary collaboration. This review provides valuable insights and theoretical support for the subsequent research, translational development, and clinical application of next-generation antimicrobial peptides.

【Key words】 Artificial intelligence; Antimicrobial peptides; Machine learning; Deep learning

Fund programs: National Natural Science Foundation of China(U22A20316, 81801012); Undergraduate Innovation and Entrepreneurship Training Program of Sun Yat-sen University (20260616)

DOI: 10.3877/cma.j.issn.1674-1366.2026.04.005

传统小分子抗生素存在作用靶点单一、易诱导菌株产生耐药突变和研发迭代速度缓慢等问题,难以应对日益严峻的临床感染^[1-2]。在此背景下,抗菌肽(antimicrobial peptides, AMP)凭借广谱抑菌、多靶点作用和不易诱发耐药等优势,成为极具应用前景

的新型抗菌候选药物,并受到广泛关注^[3]。但传统AMP的研发依赖天然产物分离提取和有限的序列修饰,导致新型AMP挖掘范围较窄,人工筛选效率低,耗时久,难以满足耐药感染新药研发的现实需求^[4]。随着人工智能(artificial intelligence, AI)技术的快速发展,机器学习(machine learning, ML)、深度学习(deep learning, DL)和生成式大模型等AI技术已逐步应用于AMP的研发和优化^[5-6],可通过数据挖掘序列与学习抗菌序列的内在规律,实现对候选肽快速筛查、全新AMP的生成与多肽性能改良,研究表明其可将AMP的研发周期从数年缩短至48 d^[7]。本文将围绕基于AI的新型AMP筛选、设计和优化等方面进行综述(完整的技术流程见图1),并探讨其当前存在的不足和未来发展方向,旨在为新型AMP的开发和临床应用提供支持和参考。

综述文献检索策略为:以“antimicrobial peptides AND artificial intelligence OR machine learning OR deep learning”为检索式,检索Web of Science、PubMed数据库,排除重复研究、综述类文献(除核心综述外)及实验数据不完整的研究,最终纳入有效文献130篇。

一、人工智能驱动抗菌肽研究的相关技术基础

1. 抗菌肽常用训练数据集:AI算法的表现很大程度上取决于训练数据的构成。各类数据库中AMP信息各有侧重,应用于AI算法搭建的模型,可以应对不同需求的AMP训练任务。因此,高质量、多样化的训练数据集是构建可靠预测模型的基石。目前,已形成多个功能互补、层次分明的AMP数据集。以dbAMP3.0^[8]、DRAMP4.0^[9]、CAMP_{R4}^[10]及LAMP2.0^[11]为代表的综合型数据库,广泛整合文献、专利及公共数据库中的AMP条目,为模型训练提供了规模庞大的基础来源;APD6^[12]专注于天然AMP的高质量收集,DBAASpV3^[13]则聚焦肽的结构与活性数据,二者深度揭示了AMP的作用机制;SATPdb^[14]与YADAMP^[15]

凭借清晰的分类体系与细致的活性标注,进一步拓展了数据覆盖的广度与维度;DAMPD^[16]中的条目则经过人工核验校对,可输出低误差的AMP理化数据特征,为AI提供高可靠性的训练样本。这些来源不同、侧重各异的数据库在相互补充的同时,也因序列重叠问题为模型的准确训练带来了挑战,上述数据集的产生和演变推动了AI在该领域的发展。常用AMP数据库见表1。

2. 核心人工智能算法:包括传统机器学习算法和深度学习算法。

(1)传统机器学习算法:传统ML方法在AMP预测的早期阶段占据主导地位。这类方法通常依赖手工提取构建的肽段特征,从肽序列中提取氨基酸组成、二肽组成和理化性质等统计特征,再输入分类器进行判别。尽管近年来DL模型发展迅猛,但传统ML方法因其可解释性强,在AMP筛选任务中仍有不可替代的应用价值^[17]。

在AMP的筛选与预测的ML研究中,不同的ML算法在AMP预测任务中发挥各自的应用优势,常用ML算法如表2所示^[18-22]。

在AMP的筛选与预测任务中,支持向量机(support vector machine, SVM)在多个基准数据集上表现优异,与随机森林(random forest, RF)和极限梯度提升(extreme gradient boosting, XGBoost)共同构成AMP预测工具中最常采用的算法组合^[23]。一项研究构建了EnAMP模型,结合了词嵌入特征与统计特征的两个深度神经网络算法,同时训练RF和SVM两个机器学习算法,通过不同算法架构的平均结果给出最终预测。并基于来自LAMP、APD、DBAASP及CAMPR3的多个数据集,将EnAMP模型与其他模型进行比较,结果显示EnAMP的性能优于大多数计算成本相似的模型^[24],但研究中并未对EnAMP模型进行湿实验验证。

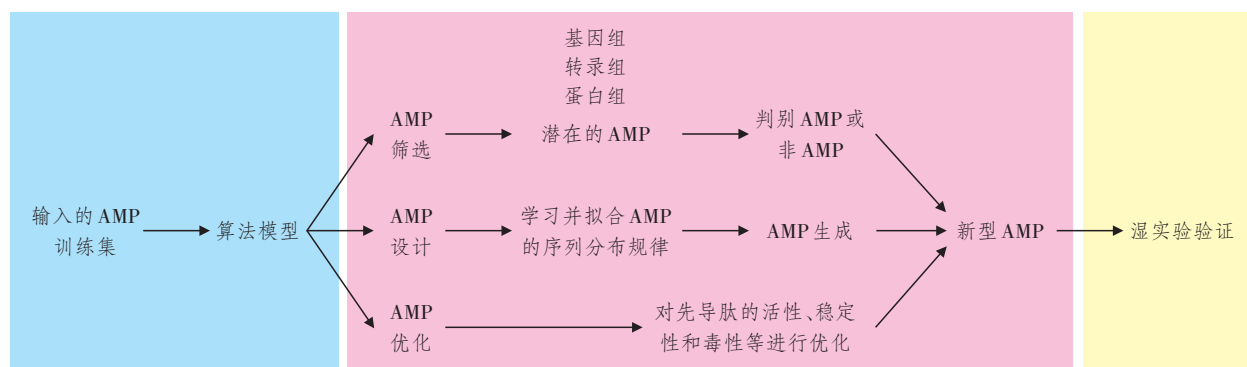


图1 基于人工智能(AI)的抗菌肽(AMP)筛选、设计和优化技术流程图

表1 常用抗菌肽(AMP)数据库

数据库	简称	作用	数据量	数据库链接
Database of Antimicrobial Peptides ^[8]	dbAMP3.0	整合 AMP 数据库和文献的综合型数据库	33 065	https://ycclab.cuhk.edu.cn/dbAMP/
Data Repository of Antimicrobial Peptides ^[9]	DRAMP4.0	综合 AMP 数据库,包含专利 AMP	30 260	https://dramp.cpu-bioinfor.org/
Collection of Anti-Microbial Peptides ^[10]	CAMP _{R4}	综合 AMP 数据库,包含专利和预测的 AMP	24 243	https://camp.bicnirrh.res.in/
Linking Antimicrobial Peptides ^[11]	LAMP2.0	综合 AMP 数据库	23 253	http://biotechlab.fudan.edu.cn/database/lamp/
Antimicrobial Peptide Database ^[12]	APD6	专注于天然 AMP 的数据库	5 188	https://aps.unmc.edu/AP/
Database of Antimicrobial Activity and Structure of Peptides ^[13]	DBAASPv3	专注于 AMP 结构和活性数据	25 069	https://www.dbaasp.org/home/
Structured Annotated Therapeutic Peptides Database ^[14]	SATPdb	整合治疗性肽的综合数据库	19 192	http://crdd.osdd.net/raghava/satpdb/
Yet Another Database of Antimicrobial Peptides ^[15]	YADAMP	专注于 AMP 的综合数据库,详细标注了收录 AMP 的抗菌活性	2 133	http://yadamp.unisa.it/
Dragon Antimicrobial Peptide Database ^[16]	DAMPD	基于 UniProt 蛋白数据库手工整理的 AMP 数据库(10.1093/nar/gkr1063)	1 232	http://apps.sanbi.ac.za/dampd/

表2 常用传统机器学习(ML)算法

算法名称	特点
SVM	在特征空间寻找最优超平面,最大化类别间隔,是 AMP 预测最广泛应用的算法
RF	集成多棵决策树投票分类,有效避免过拟合,可评估特征重要性
kNN	基于相似度度量对新样本分类,适用于小样本 AMP 识别
QDA	通过高维多样性测度提取序列信息,适用于多种生物活性的 AMP 分类
XGBoost	通过迭代添加决策树以降低预测误差,具备优异的计算效率与并行处理能力,分类性能表现突出

注:SVM为支持向量机;AMP为抗菌肽;RF为随机森林;kNN为k-近邻;QDA为二次判别分析;XGBoost为极限梯度提升。

(2)深度学习算法:DL在AMP预测中展现出超越传统ML方法的独特优势,不同于传统的ML算法,DL算法不依赖人工手动提取、筛选、构建氨基酸理化特征、序列统计特征,其能够自主学习氨基酸排列规律、残基间长距离依赖、肽段空间隐含关联等复杂信息,这大大提高了AMP的筛选和设计效率^[25]。DL模型分为判别式深度学习模型和深度生成模型,前者主要用于AMP的特征提取、分类与预测,主要包括卷积神经网络(convolutional neural networks, CNN)、循环神经网络(recurrent neural networks, RNN)及其变体长短期记忆网络(long short-term memory, LSTM)、Transformer编码器模型等。后者则学习数据集的内在分布,实现全新AMP的生成,主要包括变分自编码器(variational autoencoder, VAE)、生成对抗网络(generative adversarial networks, GAN)、Transformer解码器生成模型、LSTM自回归生成模型和扩散模型(diffusion models, DM)等。常用DL算法见表3。

①判别式深度学习模型:CNN擅长提取局部序列模式,在AMP预测中,可通过不同尺寸卷积核捕

捉二肽、三肽等局部氨基酸序列特征用于后续分类。RNN及其变体LSTM能够处理序列的顺序依赖关系,LSTM架构在嵌入层表示肽段后,可通过双向长短期记忆网络(bi-directional long short-term memory, Bi-LSTM)捕获氨基酸序列的上下文信息,其优势在于同时考虑每个残基的双侧序列,对理解AMP特定序列氨基酸的功能意义至关重要^[25]。

注意力机制(attention mechanism, ATT)是提升深度学习模型性能的关键组件。在AMP预测中,ATT存在多种应用形式。Li等^[26]开发的AMPlify模型结合Bi-LSTM与ATT,经平衡数据集训练后,成功从美洲牛蛙基因组中鉴定出4种新型AMP,并进行了体外最小抑菌浓度(minimum inhibitory concentration, MIC)的检测;Zhang等^[27]则提出AMP-EF框架,将XGBoost与Bi-LSTM和ATT相结合,构建了一个双分支集成模型:一方面利用XGBoost作为经典机器学习分支处理one-hot编码和k-mer特征,另一方面利用Bi-LSTM捕获肽序列的上下游氨基酸序列的依赖关系,最终通过ATT折中融合获得预测结果。该方法在8个独立测试数据集(XUAMP、YADAMP、

表3 常用深度学习(DL)算法

算法名称	类型	特点
CNN	判别式深度学习模型	捕捉二肽、三肽等局部序列特征,用于AMP分类预测
RNN及LSTM	判别式深度学习模型	可处理序列的顺序依赖关系,捕获上下文信息,适用于长序列AMP的特征提取
ATT	判别式深度学习模型	提升深度学习模型性能的关键组件
Transformer架构	判别式深度学习模型	以ATT机制为核心,可直接建模远距离残基间的依赖关系,适用于具有非连续保守序列特征的AMP
PLM	判别式深度学习模型	以Transformer为基础,通过自监督预训练提取高阶特征,从而提升AMP的识别性能
AR	深度生成模型	能有效拟合序列依赖关系,结构简单且易于实现,但在生成长肽时易出现序列冗余、结构不合理的问题
VAE	深度生成模型	训练过程稳定,隐空间具备可控性,支持无约束生成,但生成的序列新颖性相对有限
GAN	深度生成模型	生成的序列新颖性高、多样性丰富,但训练过程不稳定,存在较高的模式崩溃风险
Transformer解码器	深度生成模型	可挖掘全局序列关联,生成结构合理、活性潜力高的AMP,但对算力需求较高
DM	深度生成模型	数据拟合能力强,序列完整性优异,但离散序列适配性尚不成熟,目前仍处于探索阶段

注:CNN为卷积神经网络;AMP为抗菌肽;RNN为循环神经网络;LSTM为长短期记忆网络;ATT为注意力机制;Transformer为变换器;PLM为蛋白质语言模型;AR为自回归生成模型;VAE为变分自编码器;GAN为生成对抗网络;DM为扩散模型。

DRAMP、CAMP、LAMP、APD3、dbAMP和DBAASP)上取得了77.9%~99.8%的准确率,且在同一数据集下AMP-EF模型的性能均优于同期主流算法,但研究中并未对AMP-EF模型进行湿实验验证。

Transformer架构以ATT机制为核心,可通过计算序列各位置间的ATT权重,直接建模任意距离残基的依赖关系,尤其适用于处理AMP序列中非连续保守的氨基酸序列^[28]。Devlin等^[29]于2018年提出了基于Transformer架构的BERT预训练模型, Lee等^[30]随后基于BERT设计了AMP-BERT模型,无需手工设计氨基酸统计特征,即可从原始肽序列中自动提取含丰富生物学上下文的嵌入特征,捕捉肽分子潜在结构与功能信号; Gao等^[31]对比现有基于BERT的AMP预测工具差异,提出集成学习改进模型,提升了多项性能指标,但研究中并未进行湿实验验证。蛋白质语言模型(protein language model, PLM)作为深度学习分支,以Transformer为基础,借鉴自然语言处理(natural language processing, NLP)的预训练思想,通过海量无标注序列自监督预训练提取蛋白质高阶特征, UniproLcad便是其代表性工作,该模型集成ESM-2、ProtBert和UniRep蛋白质语言模型获取全面特征,输入Bi-LSTM与CNN构成的网络,经ATT分配特征权重,在AMP识别中表现出优异性能^[32],但研究中并未对UniproLcad模型进行湿实验验证。

除此之外值得注意的是,早期AI模型仅对AMP的一维氨基酸序列进行编码,无法反映肽段三维空间构象、残基拓扑关系等关键信息,而AMP的抗菌活性、细胞毒性高度依赖空间结构。为此,近年来三维拓扑特征感知技术得到了重点发展,借助AlphaFold2、

ESMPair等工具由序列预测三维结构,通过图神经网络(graph neural networks, GNN)编码残基拓扑图、3D-CNN处理体素化空间特征、几何深度学习提取分子表面特征,最终可将三维拓扑特征与传统序列特征进行多模态融合^[33-34],以更准确地识别AMP。GNN是目前应用最成熟的方案,但仍存在结构预测偏差、动态构象表征缺失、算力成本偏高等问题^[17]。

②深度生成模型:自回归生成模型(autoregressive model, AR)依托循环结构拟合氨基酸序列的依赖关系,其算法结构简单、易于实现^[35],但受限于长距离依赖捕捉缺陷,长肽生成易出现序列冗余、残基重复及结构不合理等问题。VAE作为经典概率生成模型,依靠编码器与解码器构建连续高斯隐变量空间,通过变分推断实现序列特征压缩与重构生成,模型训练稳定性强、隐空间可控性好,能够完成AMP的无约束生成^[36]。GAN则通过生成器与判别器的对抗博弈学习天然肽序列分布,生成的AMP结构新颖性突出、序列多样性丰富,但模型易出现训练不稳定等问题^[37]。Transformer解码器的自回归生成模型能够充分挖掘氨基酸残基间全局关联、进化信息及功能约束规律,生成结构合理、生物活性潜力优异的新型肽序列^[38]。DM作为新一代生成架构,依靠正向加噪与逆向去噪的分步生成机制具备极强的数据分布拟合能力,生成序列整体结构完整、细节表现力优异^[39]。不过该模型目前多应用于图像领域,在氨基酸离散序列建模中适配性尚未成熟,现阶段在AMP生成研究中仍处于初步探索阶段。

二、人工智能驱动新型抗菌肽研究现状

1. 基于人工智能的抗菌肽筛选:基于AI的AMP

筛选是指借助 ML、DL 等算法,对多肽序列库、虚拟分子库及生物信息学数据库开展自动化、高通量预测与排序,高效识别并筛选具备潜在抗菌活性的候选 AMP,以此降低传统实验筛选的工作量与成本的筛选策略。目前,AI 在 AMP 领域的研究多集中于此方向,通过宏基因组、转录组及蛋白质/多肽组数据对未知 AMP 进行挖掘筛选,可获取大量新型 AMP。

基因组作为天然 AMP 的重要资源库,为新型 AMP 挖掘提供了潜在的数据支撑。Santos-Júnior 等^[40]构建了基于随机森林(random forest, RF)算法的 Macrel 预测模型,对全球范围内 63 410 个宏基因组与 87 920 个原核基因组序列进行系统筛选,构建了包含 863 498 条非冗余候选序列的 AMPphere 数据库,其中超 90% 为全新未知 AMP。该研究选取并合成了 100 条候选肽,其中 79 条具备体外抗菌活性,部分 AMP 在小鼠感染模型中表现出与多黏菌素 B 相当的抗感染效果。Ma 等^[41]则基于 LSTM、ATT 与 BERT 架构开发了 C_AMP_Predict 模型,从人体肠道微生物的基因组中筛选并合成 216 条候选 AMP,其中 181 条表现出抗菌活性。针对已灭绝物种的蛋白组数据,de la Fuente 团队先后构建了基于 RF 的 panCleave 模型^[42]与基于全连接神经网络(fully connected neural networks, FCNN)的 APEX 模型^[43],前者从尼安德特人及丹尼索瓦人蛋白组中挖掘 AMP,合成的 5 条肽均在体外及小鼠体内实验中表现出抗菌活性;后者进一步对灭绝生物蛋白组进行扫描筛选,经合成验证的 69 条肽多通过破坏细菌细胞膜电位发挥抗菌作用,在小鼠皮肤脓肿及大腿感染模型中有效清除多重耐药鲍曼不动杆菌。

AI 亦可应用于虚拟肽库的高通量筛选。Huang 等^[44]对长度为 6~9 个氨基酸的全虚拟肽库进行挖掘筛选,获得 3 条强效六肽,该类肽对多重耐药临床菌株具有广谱抗菌活性,在小鼠肺炎模型中,经雾化给药的候选肽展现出与青霉素相当的治疗效果。总之,现阶段基于 AI 的 AMP 筛选研究已经积累了大量成果。该领域的 AI 算法框架类型多样,所用序列数据来源广泛,既涵盖基因组、转录组、蛋白质组、多肽组和宏基因组等测序数据,也可采用人工构建的虚拟肽库。相关研究的验证体系也在持续完善,早期仅通过计算完成初步预判,当前验证已覆盖体外抑菌实验、细胞毒性检测、抑菌机制以及动物感染模型体内药效评价等多个环节。为系统梳理该领域的研究成果,本文通过表 4 罗列了相关研究中所用的主要算

法类型、序列来源及实验验证方式,可为了解该领域的研究进展提供更为完整的参考^[10,25,30,33,40-99]。

2. 基于人工智能的抗菌肽设计与生成:基于 AI 的 AMP 设计与生成,是指在不依赖天然 AMP 序列模板或仅利用少量已知活性肽作为先验知识的前提下,通过 ML、DL 和生成式模型等 AI 算法,直接自主生成具备预设结构特征、理化性质与抗菌功能的全新多肽序列。与传统筛选模式不同,该类研究不局限于对已有数据库的检索与预测,而是对多肽序列空间进行从头(de novo)设计,在序列新颖性、功能定向性及结构多样性方面具备显著优势。具体而言,AMP 的虚拟筛选是指从现有序列中挖掘、筛选出具备潜在抗菌活性的候选 AMP,序列本身并非 AI 创造;AI 从头设计与生成是指不依赖现有活性肽模板,仅以已知 AMP 的序列规律、理化特征(两亲性、电荷和疏水性等)为学习依据,通过生成式 AI 模型自主创造全新氨基酸序列。

目前,DM、VAE、GAN、Transformer 和 LSTM 等多种算法已被用于 AMP 的设计和生成。Wang 等^[39]基于 DM、VAE 和 BERT 算法,通过学习数千条活性 AMP 的序列模式、两亲性、电荷分布与疏水性等关键特征,在不依赖模板的条件下生成了大量结构全新的多肽,生成的 AMP 对革兰氏阳性菌、革兰氏阴性菌及多重耐药临床菌株均表现出强效抗菌活性。Liu 等^[100]则基于 VAE 和 CNN 算法,结合体外抗菌实验、生物相容性检测与抗菌机制探索,验证了生成的自组装 AMP 对急性肠道感染小鼠模型的治疗潜力。此外,研究表明,基于 Transformer 算法构建的 TACa 模型^[101]、GAN 与图卷积网络(graph convolutional networks, GCN)算法构建的 AMPredictor 模型^[102]、XGBoost 与 LSTM 算法构建的 AMPGen 模型^[103]、沙普利可加性解释算法(SHapley additive explanations, SHAP)与 BERT 算法构建的 DLFea4AMPGen 模型^[104],均实现了高效的序列生成与活性预测。其中 AMPredictor、DLFea4AMPGen 进行了体外活性检测与小鼠急性感染体内模型的检测,TACa、AMPGen 仅开展体外湿实验,完成了抗菌活性、生物相容性及协同用药测试,未进行体内动物实验。

为提升生成肽的可靠性,深度生成算法还可与分子动力学(molecular dynamics, MD)模拟结合。Das 等^[7]在生成潜在 AMP 序列后,通过分子动力学模拟动态评估多肽与细菌细胞膜的相互作用模式、膜插入能力及结合稳定性。经实验验证,该策略筛选出的

表4 基于人工智能(AI)的抗菌肽(AMP)筛选研究成果汇总

文献	主要算法框架	生物学序列类型	生物学序列来源	湿实验验证
Gawde 等 ^[10]	SVM+RF+ANN	/	/	未进行湿实验验证
Wang 等 ^[25]	CNN+ATT+LSTM+Transformer	基因组	人类口腔微生物组	体外抗菌实验,生物相容性检测,抗菌机制探索,小鼠大腿感染模型
Lee 等 ^[30]	NLP+BERT	/	/	未进行湿实验验证
Hao 等 ^[33]	GAT+PLM	/	/	未进行湿实验验证
Santos-Júnior 等 ^[40]	RF	宏基因组和基因组	全球微生物	体外抗菌实验,生物相容性检测,抗菌机制探索,小鼠皮肤感染模型
Ma 等 ^[41]	ATT+LSTM+BERT	宏基因组	人类肠道微生物	体外抗菌实验,生物相容性检测,小鼠肺部感染模型
Maasch 等 ^[42]	RF	蛋白质组	灭绝亲属尼安德特人和丹尼索瓦人	体外抗菌实验,生物相容性检测,抗菌机制探索,小鼠皮肤及大腿感染模型
Wan 等 ^[43]	FCNN	蛋白质组	灭绝物种	体外抗菌实验,生物相容性检测,抗菌机制探索,小鼠皮肤及大腿感染模型
Huang 等 ^[44]	XGBoost+LSTM+RF+CNN	计算机虚拟的理论上所有可能的短肽序列空间	/	体外抗菌实验,生物相容性检测,抗菌机制探索,小鼠大腿及肺部感染模型
Ruiz Puentes 等 ^[45]	GCN	基因组	大肠杆菌	体外抗菌实验,抗菌机制探索
Torres 等 ^[46]	FCNN	蛋白质组	古细菌	体外抗菌实验,生物相容性检测,抗菌机制探索,小鼠皮肤及大腿感染模型
Sharma 等 ^[47]	CNN+SVM	基因组	水蛭	未进行湿实验验证
Yan 等 ^[48]	CNN+RAAC	基因组	光滑念珠菌	体外抗菌实验
Luo 等 ^[49]	PLM+GAT	多肽库	UniRef90 数据库	体外抗菌实验,生物相容性检测
Li 等 ^[50]	BERT+CNN	基因组	大型原生生物	体外抗菌实验,生物相容性检测
Duque-Salazar 等 ^[51]	RF+SVM+DA+ANN	蛋白质组	菌、藻类和无脊椎动物在内的九种生物	体外抗菌实验,生物相容性检测
Ran 等 ^[52]	SVM+RF+CNN+XGBoost	多肽组	汇总的对禾谷镰孢菌具有抗真菌活性的抗菌肽	体外抗菌实验,抗菌机制探索,分子动力学模拟
Kim 等 ^[53]	SVM+RF+ANN	转录组	美洲蟑螂	体外抗菌实验,生物相容性检测
Yan 等 ^[54]	CNN+ATT	/	/	未进行湿实验验证
Chen 等 ^[55]	SVM+RF+ANN+kNN+CNN+VAE+RNN+LSTM	非抗菌蛋白库	章鱼双斑类	体外抗菌实验,生物相容性检测,抗菌机制探索
Lin 等 ^[56]	NLP+LSTM	/	/	未进行湿实验验证
Gull 等 ^[57]	SVM+XGBoost	/	/	未进行湿实验验证
Fjell 等 ^[58]	HMM	/	/	未进行湿实验验证
Yang 等 ^[59]	RF+LSTM	/	/	未进行湿实验验证
Lin 等 ^[60]	SVM+RF	/	/	未进行湿实验验证
Lata 等 ^[61]	ANN+QM+SVM	/	/	未进行湿实验验证
Cai 等 ^[62]	BERT	/	/	未进行湿实验验证
Hartman 等 ^[63]	CNN	多肽组	串联质谱生成的急性伤口液、未感染伤口及感染金黄色葡萄球菌伤口的肽组数据集	未进行湿实验验证
Shehadeh 等 ^[64]	SVM	/	/	未进行湿实验验证
Guan 等 ^[65]	FCNN	蛋白质组	毒液组学	体外抗菌实验,生物相容性检测,抗菌机制探索,小鼠皮肤感染模型
Jhong 等 ^[66]	RF	/	/	未进行湿实验验证
Veltri 等 ^[67]	DNN+LSTM	/	/	未进行湿实验验证
Zhao 等 ^[68]	PLM+CNN+BiLSTM+CBAM	/	/	未进行湿实验验证

续表4

文献	主要算法框架	生物学序列类型	生物学序列来源	湿实验验证
Chen 等 ^[69]	DT+kNN+LightGBM+NB+RF+SVM+XGBoost+ANN+CNN+LSTM	/	设计的半胱胺酶钉结合肽	体外抗菌实验,生物相容性检测
Yu 等 ^[70]	LightGBM+LSTM+ATT	多肽组	牛奶	体外抗菌实验
Yang 等 ^[71]	ATT+LSTM+BERT	基因组	细菌基因组编码的非核糖体肽	体外抗菌实验,生物相容性检测,抗菌机制探索
Meng 等 ^[72]	CNN+LSTM	/	/	未进行湿实验验证
Wang ^[73]	MLP	/	/	未进行湿实验验证
Fang 等 ^[74]	RF+XGBoost+SVM+ET+LightGBM+CatBoost+GB	/	CPPsite 2.0 数据库	体外抗菌实验,生物相容性检测,小鼠皮下及腹膜感染模型
Xiao 等 ^[75]	CNN+BiLSTM+SVM	/	蛋白数据库训练	未进行湿实验验证
Xu 等 ^[76]	CNN+LSTM	/	/	未进行湿实验验证
Pang 等 ^[77]	BERT	/	/	未进行湿实验验证
Biswas 等 ^[78]	RF+ANN	多肽库	天然宿主防御肽数据库、已知具有抗结核活性的先导肽	体外抗菌实验,生物相容性检测
Zhang 等 ^[79]	GA+ML	多肽	由车虾生产的 LBDMj 肽作为模型肽	体外抗菌实验
Giguere 等 ^[80]	K-means+SVM	多肽库	CAMP 数据集	体外抗菌实验
Song 等 ^[81]	CNN+LSTM	蛋白质组	巨型长牡蛎黏液	体外抗菌实验,生物相容性检测
Han 等 ^[82]	PLM	/	/	未进行湿实验验证
Cordoves-Delgado 等 ^[83]	PLM+GAT	/	/	未进行湿实验验证
Meher 等 ^[84]	SVM	/	/	未进行湿实验验证
Ng 等 ^[85]	SVM	/	/	未进行湿实验验证
Olcay 等 ^[86]	RF+SVM+LR+kNN+XGBoost	/	/	未进行湿实验验证
Xia 等 ^[87]	FCNN	多肽组	蛋白酶体数据库	未进行湿实验验证
Medina-Ortiz 等 ^[88]	SVM+RF+kNN+XGBoost	肽图谱多肽数据库 中活性未知肽 从新合成的多肽筛选抗菌肽	/	未进行湿实验验证
Hussain ^[89]	CNN	/	/	未进行湿实验验证
Wu 等 ^[90]	RF	宏基因组	南美白对虾养殖池塘水体	体外抗菌实验,生物相容性检测
Guo 等 ^[91]	RF+XGBoost	多肽库	实验数据获得的肽和 DBAASP 数据库	体外抗菌实验,生物相容性检测,抗菌机制探索,体内抗龋效果分析,体内安全性分析
Zhou 等 ^[92]	CNN+LSTM	/	/	未进行湿实验验证
Du 等 ^[93]	RF	基因组	乳杆菌科	未进行湿实验验证
Barroso 等 ^[94]	SVM+RF+DNN	基因组、转录组、蛋白质组数据	刺胞动物	体外抗菌实验,生物相容性检测
Shen 等 ^[95]	CNN+LSTM+BERT	宏基因组	反刍动物胃肠道微生物组	体外抗菌实验,生物相容性检测,小鼠皮肤感染模型
Kavousi 等 ^[96]	NB+kNN+SVM+RF+XGBoost	/	/	未进行湿实验验证
Xu 等 ^[97]	LSTM+ATT+BERT	宏基因组	污泥	体外抗菌实验
Grafksaia 等 ^[98]	DNN+HMM	转录组	海葵	体外抗菌实验
Henson 等 ^[99]	ML	多肽库	虚拟肽库	体外抗菌实验,生物相容性检测,抗菌机制探索

注:SVM为支持向量机;RF为随机森林;ANN为人工神经网络;CNN为卷积神经网络;ATT为注意力机制;LSTM为长短期记忆网络;Transformer为变换器;NLP为自然语言处理;BERT为双向编码器表示的变换器;GAT为图注意力网络;PLM为蛋白质语言模型;FCNN为全连接神经网络;XGBoost为极端梯度提升;GCN为图卷积网络;RAAC为简化氨基酸组成;DA为判别分析;kNN为k-近邻;VAE为变分自编码器;RNN为循环神经网络;HMM为隐马尔可夫模型;QM为Q矩阵;DNN为深度神经网络;BiLSTM为双向长短期记忆网络;CBAM为卷积块注意力模型;DT为决策树;LightGBM为轻量级梯度提升;NB为朴素贝叶斯;MLP为多层感知机;ET为极端随机树;CatBoost为高性能梯度提升;GB为梯度提升;GA为遗传算法;ML为机器学习;K-means为K均值聚类;LR为逻辑回归。

生成肽具有明确的膜破坏机制,对多重耐药菌具有高效抑菌活性,显著提升了AMP的生成成功率。

除此之外,通过AI设计引入特异性功能基团,可将AMP与聚合物复合,构建水凝胶缓释递送体系。研究显示,通过AI引导设计平台(AMP-hydrogel-Designer)可生成抗菌活性高和富含半胱氨酸残基的新型AMP,其硫醇基团可与四臂硫醇聚乙二醇(4-arm PEG-SH, PEG-4SH)水凝胶耦合形成复杂网络结构^[105]。借助Cu²⁺和巯基的配位作用,还可进一步实现Cu改性钛酸钡(Cu-BTO)的交联。体外抗菌结果显示,AI-AMP-水凝胶对耐甲氧西林金黄色葡萄球菌(methicillin-resistant *Staphylococcus aureus*,

MRSA)和大肠杆菌(*Escherichia coli*, *E. coli*)具备优良的杀菌效率。现阶段,基于AI的AMP智能递送载体的相关设计研究仍较为有限,后续仍需进一步开展相关探索,推动功能性递送体系的开发与完善。总之,目前基于AI的AMP设计与生成的算法框架种类多样,设计得到的AMP可通过体外、体内实验完成抗菌性能检验,部分研究还结合分子动力学模拟与生物材料递送体系开展了拓展研究。表5汇总了该方向29篇相关文献,系统梳理了各研究采用的计算模型与验证方案,可为了解该领域的研究进展提供参考^[7,37,39,100-125]。

为进一步横向对比近年主流AMP计算模型的

表5 基于人工智能(AI)的抗菌肽(AMP)的设计和生成研究成果汇总

文献	主要算法框架	湿实验验证
Das等 ^[7]	VAE+WAE	体外抗菌实验,生物相容性检测,抗菌机制探索
Tian等 ^[37]	GAN	体外抗菌实验,生物相容性检测,抗菌机制探索
Wang等 ^[39]	DM+VAE+BERT	体外抗菌实验,生物相容性检测,体内小鼠皮肤/肺部感染模型
Liu等 ^[100]	VAE+CNN	体外抗菌实验,生物相容性检测,抗菌机制探索,体内小鼠肠道感染模型
de Albernaz等 ^[101]	Transformer	体外抗菌实验,生物相容性检测,协同作用试验;抗菌肽与抗生素美罗培南
Dong等 ^[102]	GAN+GCN	体外抗菌实验,生物相容性检测,抗菌机制探索,体内小鼠多重耐药鲍曼杆菌急性感染模型
Jin等 ^[103]	XGBoost+LSTM	体外抗菌实验,生物相容性检测,抗菌机制探索
Gao等 ^[104]	SHAP+BERT	体外抗菌实验,生物相容性检测,抗菌机制探索,体内小鼠急性感染模型
Jiang等 ^[105]	PLM+PT+KD+RL	体外抗菌实验,生物相容性检测,抗菌机制探索,体内大鼠颈部感染模型
Zhang等 ^[106]	LSTM+CNN	未进行湿实验验证
Pikalyova等 ^[107]	WAE+GTM	体外抗菌实验,生物相容性检测
Zhao等 ^[108]	VAE+Transformer	未进行湿实验验证
Liu等 ^[109]	RNN+PLM+RTKU	体外抗菌实验,生物相容性检测,抗菌机制探索
Zhao等 ^[110]	CVAE+DM	未进行湿实验验证
Wu等 ^[111]	GGD+RL	体外抗菌实验,生物相容性检测,抗菌机制探索
Wang等 ^[112]	GRU+ATT+MCTS	抗菌活性、溶血分析、细胞毒性分析
Yang等 ^[113]	CVAE+RF	体外抗菌实验,生物相容性检测
Pandi等 ^[114]	VAE+CNN+LSTM	体外抗菌实验,生物相容性检测,抗菌机制探索
Boone等 ^[115]	RST+GA	体外抗菌实验
Wang等 ^[116]	PL+DM	体外抗菌实验,生物相容性检测,抗菌机制探索
Cesaro等 ^[117]	VAE	体外抗菌实验,生物相容性检测,动物模型内抗感染实验
Whelan等 ^[118]	SVM+RF+DA	未进行湿实验验证
Swanson等 ^[119]	DNN+MCTS	体外抗菌实验,耐药概率测试
Yin等 ^[120]	XGBoost	体外抗菌实验
Dong等 ^[121]	RNN+Transformer+GRU+LSTM	体外抗菌实验,生物相容性检测,抗菌机制探索
Zare-Zardini等 ^[122]	LightGBM+GA	体外抗菌实验,生物相容性检测,抗菌机制探索,伤口愈合能力分析
Dean等 ^[123]	VAE	体外抗菌实验
Dean等 ^[124]	VAE	体外抗菌实验
Tucs等 ^[125]	GAN	体外抗菌实验,氨苄西林药效对标分析

注:VAE为变分自编码器;WAE为瓦瑟斯坦自编码器;GAN为生成对抗网络;DM为扩散模型;BERT为双向编码器表示的变换器;CNN为卷积神经网络;Transformer为变换器;GCN为图卷积网络;XGBoost为极端梯度提升;LSTM为长短期记忆网络;SHAP为沙普利可加性解释;PLM为蛋白质语言模型;PT为预训练;KD为知识蒸馏;RL为强化学习;GTM为生成式拓扑映射;RNN为循环神经网络;RTKU为实时知识更新;CVAE为条件变分自编码器;GGD为图语法蒸馏;GRU为门控循环单元;ATT为注意力机制;MCTS为蒙特卡洛树搜索;RF为随机森林;RST为粗糙集理论;GA为遗传算法;PL为提示学习;SVM为支持向量机;DA为判别分析;DNN为深度神经网络;LightGBM为轻量级梯度提升。

设计思路、实验支撑与应用优势,本文汇总了几种代表性的判别及生成类AI模型,从特征编码、靶向菌种、干湿实验验证和核心创新结论等方面进行梳理,详细对比信息见表6。

3. 基于AI的AMP优化:是在已有AMP基础上,通过机器学习、深度学习、分子模拟与进化算法等手段,对多肽序列、氨基酸组成、电荷分布、疏水性及空间构象进行定向改造,以实现抗菌活性提升、溶血毒性降低、体内稳定性增强、膜选择性改善等目标^[126-130]。

天然L残基AMP易被酶降解,体内稳定性较差。通过机器学习模型对氨基酸位点取代效应进

行预测,定向引入D-氨基酸,以优化天然L残基AMP的两亲性与螺旋稳定性,同时保留关键抗菌位点^[126]。体内外实验表明,经优化的多肽对多重耐药菌的抑菌活性显著提升,且抗蛋白酶水解能力增强,在MRSA和*E. coli*感染的小鼠皮肤缺损模型中展现出优异的治疗效果。

天然AMP普遍存在膜靶向特异性不足、细胞溶血不良反应较强、抗菌活性有限等缺陷,限制了其临床转化应用。基于AI的AMP优化可解析抗菌肽关键氨基酸位点的功能差异,定向完成序列改造与结构优化,从而实现AMP性能的定向改良。Randall等^[128]通过对先导肽(protegrin-1, PG-1)进行深度突

表6 代表性的判别及生成类人工智能(AI)模型

核心AI模型	算法框架	特征编码方式	靶标菌株	湿实验验证方式	核心结论
Macrel ^[40]	RF	氨基酸组成、理化特性、二肽及三肽的分布、分子量等一维序列特征	ESKAPEE(屎肠球菌、金黄色葡萄球菌、肺炎克雷伯菌、鲍曼不动杆菌、铜绿假单胞菌、肠杆菌属和大肠埃希菌)等11种临床致病菌	体外抗菌实验 生物相容性检测 抗菌机制探索 小鼠皮肤鲍曼不动杆菌感染模型	对全球范围内63 410个宏基因组与87 920个原核基因组序列进行系统筛选,构建了AMPphere数据库,其中超90%为全新未知AMP。该研究选取并合成了100条候选肽,其中79条具备体外抗菌活性,部分AMP在小鼠感染模型中表现出与多黏菌素B相当的抗感染效果
APEX ^[43]	FCNN	氨基酸组成、理化特性、二肽及三肽的分布、分子量等一维序列特征	ESKAPEE等11种临床致病菌	体外抗菌实验 生物相容性检测 抗菌机制探索 小鼠皮肤和大腿深部鲍曼不动杆菌感染模型	对灭绝生物蛋白组进行扫描筛选,经合成验证的69条肽多通过破坏细菌细胞膜电位发挥抗菌作用,可在小鼠皮肤脓肿及大腿感染模型中有效清除多重耐药鲍曼不动杆菌
AMPlify ^[26]	Bi-LSTM+ATT	仅包括氨基酸序列信息,不包括理化特性等	WHO重点耐药致病菌(耐药性大肠杆菌、金黄色葡萄球菌、化脓性链球菌和铜绿假单胞菌)	体外抗菌实验	擅长从真核基因组中挖掘新型AMP,灵敏度与特异度显著提升
PGAT-ABPp ^[33]	GNN+PLM	PLM全局序列嵌入,氨基酸残基构建拓扑图,叠加理化属性节点特征	广谱抗菌	未进行湿实验验证	PLM可捕获氨基酸全局上下文规律,GNN有效挖掘残基局部相互作用,二者融合能弥补单一模型缺失结构信息的缺陷;PGAT-ABPp在多套测试数据集上识别准确率显著优于现有主流预测工具
PepVAE ^[124]	VAE	氨基酸离散序列嵌入编码,将氨基酸字符映射为低维连续向量;隐空间约束多肽净电荷、疏水性、长度等关键理化参数,实现可控序列生成	大肠杆菌、金黄色葡萄球菌、铜绿假单胞菌	体外抗菌实验	PepVAE可实现多肽生成与活性预测,相比传统生成模型训练过程更稳定,不易出现序列失效问题
PepGAN ^[125]	GAN	氨基酸组成、理化特性、二肽及三肽的分布、分子量等一维序列特征	大肠杆菌	体外抗菌实验 并与氨苄西林药效对标	活性约束的GAN可定向产出高抑菌活性多肽,提升有效候选肽的产出效率

注:RF为随机森林;AMP为抗菌肽;FCNN为全连接神经网络;Bi-LSTM为长短期记忆网络;ATT为注意力机制;WHO为世界卫生组织;GNN为图神经网络;PLM为蛋白质语言模型;VAE为变分自编码器;GAN为生成对抗网络。

变扫描(deep mutational scanning, DMS),构建了基于PG-1的突变体,并筛选出决定细菌膜靶向性与宿主细胞膜毒性的关键氨基酸位点。优化后的AMP在显著提升对细菌膜选择性的同时降低了溶血毒性。Mishra等^[127]则利用ML对具有弱抗MRSA活性的AMP(citropin 1.1)进行氨基酸水平的优化,体内实验表明,优化后的AMP实现了抗菌活性的提升,并在小鼠皮肤缺损感染模型中可有效清除病原菌,治疗效果显著优于未优化的citropin 1.1。

三、现有研究的争议、不足与挑战

当前,AI辅助AMP预测与设计研究发展迅速,但诸多争议与关键问题仍未解决,制约了模型可靠性与临床转化进程。

模型预测结果的可靠性是核心争议点。不同研究构建的数据集在序列冗余度、活性标准和菌种类型上差异明显,导致同一模型在不同任务中的性能波动较大,且多数模型仅能预测体外抗菌活性,与体内真实微环境、药代行为及实际疗效存在偏差,预测结果难以直接用于实际应用。此外,抗菌活性、溶血性和稳定性等多目标优化缺乏统一权重标准,难以实现各功能的协同。

与此同时,公开数据集普遍存在正负样本不平衡、有效样本量偏小、抗菌谱与物种覆盖范围有限等问题,易导致算法模型的过拟合。针对上述问题,对比学习与蛋白质大语言模型的自监督预训练可提供解决方案。对比学习通过构建正负样本对进行特征相似度训练,无需大幅扩充标注样本总量,即可引导模型挖掘AMP与非AMP的核心特征差异。该方法能在现有不平衡数据集基础上,强化模型对关键功能特征的判别能力,弱化样本数量差距的负面影响,一定程度上缓解小样本训练引发的过拟合问题。蛋白质大语言模型则依托无监督学习处理大量无标注天然蛋白与多肽序列,在预训练过程中自主总结通用序列规律,隐式勾勒出非AMP的完整特征空间。这一过程无需人工专门标注大量负样本,仅依靠自然界海量普通序列即可区分AMP与非AMP的本质差异,弥补了负样本稀缺、样本不平衡的缺陷。

过高的算力要求也是AI辅助AMP预测与设计面临的另一大困境,传统ML、DL及大模型在计算效率与算力成本上存在显著差异。传统ML方法(如SVM、RF、XGBoost)算力需求低,训练与推理耗时短,适用于小样本快速筛选场景,研发成本较低。DL模型(如CNN、LSTM、Transformer)训练时间长、

硬件成本高,更适合大规模数据与高精度预测任务。蛋白质大模型与生成式大模型虽性能优异,但算力需求高、训练成本昂贵且推理速度慢,推广与落地难度较大。这种算法的差异导致中小实验室难以开展基于大模型的AMP研发工作。

此外,多数研究仅止步于计算筛选与体外验证,缺乏系统的体内抗菌实验、药代动力学及毒理学评价,加之AMP规模化合成成本较高,给药途径有限,进一步阻碍了从AI预测到临床应用的转化路径。最后,计算研究者与实验生物医学研究者存在明显脱节,AI模型研发多侧重于算法性能提升,未能针对不同临床问题的代表性菌种进行设计,导致模型设计与实验验证脱节,难以研发出真正具备应用价值的新型AMP。为此,应组建融合计算、实验及临床人员的联合团队,围绕耐药菌感染的临床需求开展定向研究,并依据活性与毒性实验数据优化理论预测体系。此外,还应推动跨学科人才培养和标准化共享数据工具平台的建立以降低操作门槛,以促进AI研究者和生物医学研究者的深度协同。

四、总结与展望

AI已全面应用于AMP的高通量筛选、设计生成及优化,实现了对宏基因组、转录组、蛋白组及虚拟肽空间的高效筛选和AMP的新生成,获得了大量的新型AMP。AI算法可与分子模拟结合,进一步提升AMP的活性、膜选择性与体内稳定性,缩短了研发周期并降低了成本。在细菌耐药日趋严重的背景下,AI为AMP研发提供了新型技术手段,对开发新型抗菌制剂、应对临床耐药感染具有重要意义。

尽管当前基于AI的AMP研究已取得相当多的研究成果,但目前研究整体仍集中于计算模拟与体外验证阶段,体内药效与安全性数据相对匮乏,给药递送形式单一,长效缓释与靶向递送系统仍较为缺乏,距离临床转化尚有较大距离。

未来基于AI的AMP的研发应重点关注以下几个方面:首先,应建设高质量的数据集,为模型训练与性能优化提供坚实数据支撑。其次,进一步优化已有算法,在提升抗菌活性的同时,实现对膜选择性、酶解稳定性、免疫原性等关键指标的优化和筛选。其三,应加快推进临床转化研究,系统评价AMP的体内药效、药代动力学及生物相容性,并结合水凝胶、纳米载体、可降解生物材料等开发新型递送系统,为解决临床耐药感染难题提供新一代抗菌策略。最后,可搭建自动化合成与平行生物检测

的一体化实验平台,依托AI的筛选设计能力与高通量自动化设备,批量完成多肽制备、体外抗菌活性及细胞相容性评价,进一步提升AMP的开发效率。

综上所述,AI已广泛应用于AMP的高通量筛选、设计生成及优化,结合分子模拟可显著提升AMP的研发效率、降低成本和缩短周期,为AMP的研发提供了新的参考方向。

利益冲突 所有作者均声明不存在利益冲突

作者贡献声明 王晓宇、雒家玉:资料收集、文献查阅和文章撰写;于晓琳:文章审校、修改指导

参 考 文 献

- [1] Kariuki S. Global burden of antimicrobial resistance and forecasts to 2050 [J]. *Lancet*, 2024, 404 (10459): 1172-1173. DOI: 10.1016/S0140-6736(24)01885-3.
- [2] Brinkac L, Voorhies A, Gomez A, et al. The threat of antimicrobial resistance on the human microbiome [J]. *Microb Ecol*, 2017, 74(4): 1001-1008. DOI: 10.1007/s00248-017-0985-z.
- [3] Luo Y, Song Y. Mechanism of antimicrobial peptides: Antimicrobial, anti-inflammatory and antibiofilm activities [J]. *Int J Mol Sci*, 2021, 22(21): 11401. DOI: 10.3390/ijms222111401.
- [4] Melo M C R, Maasch J R M A, de la Fuente - Nunez C. Accelerating antibiotic discovery through artificial intelligence [J]. *Communications Biology*, 2021, 4(1): 1050. DOI: 10.1038/s42003-021-02586-0.
- [5] Wan F, Wong F, Collins J J, et al. Machine learning for antimicrobial peptide identification and design [J]. *Nat Rev Bioeng*, 2024, 2(5): 392-407. DOI: 10.1038/s44222-024-00152-x.
- [6] Szymczak P, Szczurek E. Artificial intelligence - driven antimicrobial peptide discovery [J]. *Curr Opin Struct Biol*, 2023 (83): 102733. DOI: 10.1016/j.sbi.2023.102733.
- [7] Das P, Sercu T, Wadhawan K, et al. Accelerated antimicrobial discovery via deep generative models and molecular dynamics simulations [J]. *Nat Biomed Eng*, 2021, 5(6): 613-623. DOI: 10.1038/s41551-021-00689-x.
- [8] Yao L, Guan J, Xie P, et al. dbAMP 3.0: Updated resource of antimicrobial activity and structural annotation of peptides in the post-pandemic era [J]. *Nucleic Acids Res*, 2025, 53(D1): D364-D376. DOI: 10.1093/nar/gkae1019.
- [9] Ma T, Liu Y, Yu B, et al. DRAMP 4.0: An open-access data repository dedicated to the clinical translation of antimicrobial peptides [J]. *Nucleic Acids Res*, 2025, 53(D1): D403-D410. DOI: 10.1093/nar/gkae1046.
- [10] Gawde U, Chakraborty S, Waghv F H, et al. CAMPR4: A database of natural and synthetic antimicrobial peptides [J]. *Nucleic Acids Res*, 2023, 51(D1): D377-D383. DOI: 10.1093/nar/gkac933.
- [11] Zhao X, Wu H, Lu H, et al. LAMP: A database linking antimicrobial peptides [J]. *PLoS One*, 2013, 8(6): e66557. DOI: 10.1371/journal.pone.0066557.
- [12] Wang G, Schmidt C, Li X, et al. APD6: The antimicrobial peptide database is expanded to promote research and development by deploying an unprecedented information pipeline [J]. *Nucleic Acids Res*, 2026, 54 (D1): D363 - D374. DOI: 10.1093/nar/gkaf860.
- [13] Pirtskhalava M, Armstrong A A, Grigolava M, et al. DBAASP v3: Database of antimicrobial/cytotoxic activity and structure of peptides as a resource for development of new therapeutics [J]. *Nucleic Acids Res*, 2021, 49 (D1): D288-D297. DOI: 10.1093/nar/gkaa991.
- [14] Singh S, Chaudhary K, Dhanda S K, et al. SATPdb: A database of structurally annotated therapeutic peptides [J]. *Nucleic Acids Res*, 2016, 44(D1): D1119-D1126. DOI: 10.1093/nar/gkv1114.
- [15] Piotto S P, Sessa L, Concilio S, et al. YADAMP: Yet another database of antimicrobial peptides [J]. *Int J Antimicrob Agents*, 2012, 39(4): 346-351. DOI: 10.1016/j.ijantimicag.2011.12.003.
- [16] Seshadri Sundararajan V, Gabere M N, Pretorius A, et al. DAMPD: A manually curated antimicrobial peptide database [J]. *Nucleic Acids Res*, 2012, 40 (D1): D1108 - D1112. DOI: 10.1093/nar/gkr1063.
- [17] Brizuela C A, Liu G, Stokes J M, et al. AI methods for antimicrobial peptides: Progress and challenges [J]. *Microb Biotechnol*, 2025, 18 (1): e70072. DOI: 10.1111/1751 - 7915.70072.
- [18] Cortes C, Vapnik V. Support-vector networks [J]. *Mach Learn*, 1995(20): 273-297. DOI: 10.1007/BF00994018.
- [19] Ho TK. The random subspace method for constructing decision forests [J]. *IEEE Trans Pattern Anal Mach Intell*, 1998, 20(8): 832-844. DOI: 10.1109/34.709601.
- [20] Altman NS. An introduction to kernel and nearest - neighbor nonparametric regression [J]. *Am Stat*, 1992, 46(3): 175-185. DOI: 10.1080/00031305.1992.10475879.
- [21] Wani MA, Garg P, Roy KK. Machine learning-enabled predictive modeling to precisely identify the antimicrobial peptides [J]. *Med Biol Eng Comput*, 2021, 59(11-12): 2397-2408. DOI: 10.1007/s11517-021-02443-6.
- [22] Chen W, Luo L. Classification of antimicrobial peptide using diversity measure with quadratic discriminant analysis [J]. *J Microbiol Methods*, 2009, 78(1): 94-96. DOI: 10.1016/j.mimet.2009.03.013.
- [23] Xu J, Li F, Leier A, et al. Comprehensive assessment of machine learning - based methods for predicting antimicrobial peptides [J]. *Brief Bioinform*, 2021, 22(5): bbab083. DOI: 10.1093/bib/bbab083.
- [24] Zhuang J, Gao W, Su R. EnAMP: A novel deep learning ensemble antibacterial peptide recognition algorithm based on multi - features [J]. *J Bioinform Comput Biol*, 2024, 22(1): 2450001. DOI: 10.1142/S021972002450001X.
- [25] Wang B, Lin P, Zhong Y, et al. Explainable deep learning and virtual evolution identifies antimicrobial peptides with activity against multidrug-resistant human pathogens [J]. *Nat Microbiol*, 2025, 10(2): 332-347. DOI: 10.1038/s41564-024-01907-3.
- [26] Li C, Sutherland D, Hammond S A, et al. AMPlify: Attentive deep learning model for discovery of novel antimicrobial peptides effective against WHO priority pathogens [J]. *BMC Genomics*,

- 2022,23(1):77. DOI:10.1186/s12864-022-08310-4.
- [27] Zhang S, Zhao Y, Liang Y. AMP-EF: An ensemble framework of extreme gradient boosting and bidirectional long short-term memory network for identifying antimicrobial peptides [J]. *Match*, 2024,91(1):109-131. DOI:10.46793/match.91-1.109z.
- [28] Madan S, Lentzen M, Brandt J, et al. Transformer models in biomedicine [J]. *BMC Med Inform Decis Mak*, 2024, 24(1): 214. DOI:10.1186/s12911-024-02600-5.
- [29] Devlin J, Chang M W, Lee K, et al. BERT: Pre-training of deep bidirectional transformers for language understanding [C/OL]// Burstein J, Doran C, Solorio T. *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*. Minneapolis: Assoc Comput Linguist, 2019: 4171-4186. <https://www.semanticscholar.org/reader/df2b0e26d0599ce3e70df8a9da02e51594e0e992>. DOI: 10.18653/v1/N19-1423.
- [30] Lee H, Lee S, Lee I, et al. AMP - BERT: Prediction of antimicrobial peptide function based on a BERT model [J]. *Protein Sci*, 2023,32(1):e4529. DOI:10.1002/pro.4529.
- [31] Gao W, Zhao J, Gui J, et al. Comprehensive assessment of BERT-based methods for predicting antimicrobial peptides [J]. *J Chem Inf Model*, 2024, 64(19): 7772-7785. DOI: 10.1021/acs.jcim.4c00507.
- [32] Wang X, Wu Z, Wang R, et al. UniproLead: Accurate identification of antimicrobial peptide by fusing multiple pre-trained protein language models [J]. *Symmetry*, 2024, 16(4): 464. DOI:10.3390/sym16040464.
- [33] Hao Y, Liu X, Fu H, et al. PGAT-ABPp: Harnessing protein language models and graph attention networks for antibacterial peptide identification with remarkable accuracy [J]. *Bioinformatics*, 2024,40(8):btac497. DOI:10.1093/bioinformatics/btac497.
- [34] Chen B, Xie Z, Qiu J, et al. Improved the heterodimer protein complex prediction with protein language models [J]. *Briefings in Bioinformatics*, 2023,24(4):bbad221. DOI:10.1093/bib/bbad221.
- [35] Nagarajan D, Nagarajan T, Roy N, et al. Computational antimicrobial peptide design and evaluation against multidrug-resistant clinical isolates of bacteria [J]. *J Biol Chem*, 2018, 293(10): 3492-3509. DOI: 10.1074/jbc.M117.805499.
- [36] Szymczak P, Możejko M, Grzegorzec T, et al. Discovering highly potent antimicrobial peptides with deep generative model HydrAMP [J]. *Nat Commun*, 2023, 14(1): 1453. DOI: 10.1038/s41467-023-36994-z.
- [37] Tian C, Hao Y, Fu H, et al. From AI-driven sequence generation to molecular simulation: A comprehensive framework for antimicrobial peptide discovery [J]. *J Chem Inf Model*, 2025, 65(18): 9566-9575. DOI: 10.1021/acs.jcim.5c00892.
- [38] Madani A, Krause B, Greene ER, et al. Large language models generate functional protein sequences across diverse families [J]. *Nat Biotechnol*, 2023,41(8):1099-1106. DOI:10.1038/s41587-022-01618-2.
- [39] Wang Y, Song M, Liu F, et al. Artificial intelligence using a latent diffusion model enables the generation of diverse and potent antimicrobial peptides [J]. *Sci Adv*, 2025, 11(6): eadp7171. DOI:10.1126/sciadv.adp7171.
- [40] Santos-Júnior C D, Torres M D T, Duan Y, et al. Discovery of antimicrobial peptides in the global microbiome with machine learning [J]. *Cell*, 2024, 187(14): 3761-3778. e16. DOI: 10.1016/j.cell.2024.05.013.
- [41] Ma Y, Guo Z, Xia B, et al. Identification of antimicrobial peptides from the human gut microbiome using deep learning [J]. *Nat Biotechnol*, 2022, 40(6): 921-931. DOI: 10.1038/s41587-022-01226-0.
- [42] Maasch J R M A, Torres M D T, Melo M C R, et al. Molecular de-extinction of ancient antimicrobial peptides enabled by machine learning [J]. *Cell Host Microbe*, 2023, 31(8): 1260-1274. e6. DOI:10.1016/j.chom.2023.07.001.
- [43] Wan F, Torres M D T, Peng J, et al. Deep-learning-enabled antibiotic discovery through molecular de-extinction [J]. *Nat Biomed Eng*, 2024, 8(7): 854-871. DOI: 10.1038/s41551-024-01201-x.
- [44] Huang J, Xu Y, Xue Y, et al. Identification of potent antimicrobial peptides via a machine-learning pipeline that mines the entire space of peptide sequences [J]. *Nat Biomed Eng*, 2023, 7(6): 797-810. DOI: 10.1038/s41551-022-00991-2.
- [45] Ruiz Puentes P, Henao M C, Cifuentes J, et al. Rational discovery of antimicrobial peptides by means of artificial intelligence [J]. *Membranes*, 2022, 12(7): 708. DOI: 10.3390/membranes12070708.
- [46] Torres M D T, Wan F, de la Fuente-Nunez C. Deep learning reveals antibiotics in the archaeal proteome [J]. *Nat Microbiol*, 2025, 10(9): 2153-2167. DOI: 10.1038/s41564-025-02061-0.
- [47] Sharma R, Shrivastava S, Kumar Singh S, et al. AniAMPpred: Artificial intelligence guided discovery of novel antimicrobial peptides in animal kingdom [J]. *Brief Bioinform*, 2021, 22(6): bbab242. DOI: 10.1093/bib/bbab242.
- [48] Yan J, Bhadra P, Li A, et al. Deep-AmPEP30: Improve short antimicrobial peptides prediction with deep learning [J]. *Mol Ther Nucleic Acids*, 2020(20): 882-894. DOI: 10.1016/j.omtn.2020.05.006.
- [49] Luo J, Zhao K, Chen J, et al. iMFP-LG: Identify novel multi-functional peptides using protein language models and graph-based deep learning [J]. *Genomics Proteomics Bioinform*, 2025, 22(6): qzae084. DOI: 10.1093/gpbjnl/qzae084.
- [50] Li W, Zhu G, Zubair M, et al. Deep learning-driven discovery of novel antimicrobial peptides from large-scale protist genomes and experimental characterization [J]. *J Chem Inf Model*, 2025, 65(19): 9962-9973. DOI: 10.1021/acs.jcim.5c01196.
- [51] Duque-Salazar G, Mendez-Otalvaro E, Ceballos-Arroyo AM, et al. Design of antimicrobial and cytolytic peptides by computational analysis of bacterial, algal, and invertebrate proteomes [J]. *Amino Acids*, 2020, 52(10): 1403-1412. DOI: 10.1007/s00726-020-02900-w.
- [52] Ran Y, Li S, Wang Y J, et al. AI-accelerated identification of novel antimicrobial peptides for inhibiting *Fusarium graminearum* [J]. *J Agric Food Chem*, 2025, 73(28): 17471-17482. DOI: 10.1021/acs.jafc.5c03429.

- [53] Kim I W, Lee J H, Subramaniyam S, et al. De novo transcriptome analysis and detection of antimicrobial peptides of the American cockroach *Periplaneta americana* (Linnaeus) [J]. PLoS One, 2016, 11(5):e0155304. DOI:10.1371/journal.pone.0155304.
- [54] Yan J, Zhang B, Zhou M, et al. A deep learning method for predicting the minimum inhibitory concentration of antimicrobial peptides against *Escherichia coli* using multi-branch-CNN and attention[J]. mSystems, 2023, 8(4):e00345-23. DOI:10.1128/mSystems.00345-23.
- [55] Chen P, Zhang T, Li C, et al. Aggregation-prone antimicrobial peptides target gram-negative bacterial nucleic acids and protein synthesis [J]. Acta Biomater, 2025 (192): 446-460. DOI: 10.1016/j.actbio.2024.12.002.
- [56] Lin T T, Yang L Y, Lu I H, et al. AI4AMP: An antimicrobial peptide predictor using physicochemical property-based encoding method and deep learning[J]. mSystems, 2021, 6(6):e00299-21. DOI:10.1128/mSystems.00299-21.
- [57] Gull S, Shamim N, Minhas F. AMAP: Hierarchical multi-label prediction of biologically active and antimicrobial peptides [J]. Comput Biol Med, 2019 (107): 172-181. DOI: 10.1016/j.compbio.2019.02.018.
- [58] Fjell C D, Hancock R E W, Cherkasov A. AMPper: A database and an automated discovery tool for antimicrobial peptides [J]. Bioinformatics, 2007, 23 (9): 1148-1155. DOI: 10.1093/bioinformatics/btm068.
- [59] Yang S, Yang Z, Ni X. AMPFinder: A computational model to identify antimicrobial peptides and their functions based on sequence-derived information [J]. Anal Biochem, 2023 (673): 115196. DOI:10.1016/j.ab.2023.115196.
- [60] Lin Y, Cai Y, Liu J, et al. An advanced approach to identify antimicrobial peptides and their function types for penaeus through machine learning strategies [J]. BMC Bioinform, 2019, 20(S8):291. DOI:10.1186/s12859-019-2766-9.
- [61] Lata S, Sharma B, Raghava G. Analysis and prediction of antibacterial peptides [J]. BMC Bioinform, 2007, 8 (1): 263. DOI:10.1186/1471-2105-8-263.
- [62] Cai J, Yan J, Un C, et al. BERT-AmPEP60: A BERT-based transfer learning approach to predict the minimum inhibitory concentrations of antimicrobial peptides for *Escherichia coli* and *Staphylococcus aureus* [J]. J Chem Inf Model, 2025, 65(7):3186-3202. DOI:10.1021/acs.jcim.4c01749.
- [63] Hartman E, Wallblom K, van der Plas M J A, et al. Bioinformatic analysis of the wound peptidome reveals potential biomarkers and antimicrobial peptides [J]. Front Immunol, 2021 (11):620707. DOI:10.3389/fimmu.2020.620707.
- [64] Shehadeh F, Mishra B, Ferrer-Espada R, et al. CAMPER: Mechanistic artificial intelligence for designing peptides that target MRSA persists [J]. Nat Commun, 2026, 17(1):3689. DOI:10.1038/s41467-026-70348-9.
- [65] Guan C, Torres M D T, Li S, et al. Computational exploration of global venoms for antimicrobial discovery with venomomics artificial intelligence [J]. Nat Commun, 2025, 16(1):6446. DOI: 10.1038/s41467-025-60051-6.
- [66] Jhong J H, Chi Y H, Li W C, et al. dbAMP: An integrated resource for exploring antimicrobial peptides with functional activities and physicochemical properties on transcriptome and proteome data [J]. Nucleic Acids Res, 2019, 47 (D1): D285-D297. DOI:10.1093/nar/gky1030.
- [67] Veltri D, Kamath U, Shehu A. Deep learning improves antimicrobial peptide recognition [J]. Bioinformatics, 2018, 34 (16):2740-2747. DOI:10.1093/bioinformatics/bty179.
- [68] Zhao J, Liu H, Kang L, et al. deep-AMPpred: A deep learning method for identifying antimicrobial peptides and their functional activities [J]. J Chem Inf Model, 2025, 65(2):997-1008. DOI: 10.1021/acs.jcim.4c01913.
- [69] Chen R, You Y, Liu Y, et al. Deep-learning-based approaches for rational design of stapled peptides with high antimicrobial activity and stability [J]. Microb Biotechnol, 2025, 18 (3): e70121. DOI:10.1111/1751-7915.70121.
- [70] Yu W, Zhang X, Yu Y, et al. Discovery of milk-derived antimicrobial peptides in human milk by DeepMAMP based on peptidomics technology and deep learning method [J]. Food Chem, 2025 (491): 145174. DOI: 10.1016/j.foodchem.2025.145174.
- [71] Yang C L, Wang P P, Zhou Z Y, et al. Discovery of naturally inspired antimicrobial peptides using deep learning [J]. Bioorg Chem, 2025(160):108444. DOI:10.1016/j.bioorg.2025.108444.
- [72] Meng Q, Chen G, Lin B, et al. DMAMP: A deep-learning model for detecting antimicrobial peptides and their multi-activities [J]. IEEE ACM Trans Comput Biol Bioinform, 2024, 21(6):2025-2034. DOI:10.1109/TCBB.2024.3439541.
- [73] Wang S C. E-CLEAP: An ensemble learning model for efficient and accurate identification of antimicrobial peptides [J]. PLoS One, 2024, 19(5):e0300125. DOI: 10.1371/journal.pone.0300125.
- [74] Fang Y, Fan D, Feng B, et al. Harnessing advanced computational approaches to design novel antimicrobial peptides against intracellular bacterial infections [J]. Bioact Mater, 2025 (50): 510-524. DOI:10.1016/j.bioactmat.2025.04.016.
- [75] Xiao X, Shao Y T, Cheng X, et al. iAMP-CA2L: A new CNN-BiLSTM-SVM classifier based on cellular automata image for identifying antimicrobial peptides and their functional types [J]. Brief Bioinform, 2021, 22 (6): bbab209. DOI: 10.1093/bib/bbab209.
- [76] Xu J, Li F, Li C, et al. iAMPCN: A deep-learning approach for identifying antimicrobial peptides and their functional activities [J]. Brief Bioinform, 2023, 24(4):bbad240. DOI:10.1093/bib/bbad240.
- [77] Pang Y, Yao L, Xu J, et al. Integrating transformer and imbalanced multi-label learning to identify antimicrobial peptides and their functional activities [J]. Bioinformatics, 2022, 38(24):5368-5374. DOI:10.1093/bioinformatics/btac711.
- [78] Biswas D, Benson S, Matunis A, et al. Lead informed artificial intelligence mining of antitubercular host defense peptides [J]. Biomacromolecules, 2025, 26(5):3167-3179. DOI:10.1021/acs.biomac.5c00244.

- [79] Zhang H, Wang Y, Zhu Y, et al. Machine learning and genetic algorithm - guided directed evolution for the development of antimicrobial peptides[J]. J Adv Res, 2025(68):415-428. DOI: 10.1016/j.jare.2024.02.016.
- [80] Giguere S, Laviolette F, Marchand M, et al. Machine learning assisted design of highly active peptides for drug discovery[J]. PLoS Comput Biol, 2015, 11(4): e1004074. DOI: 10.1371/journal.pcbi.1004074.
- [81] Song J, Liu K, Jin X, et al. Machine learning-driven discovery and evaluation of antimicrobial peptides from *Crassostrea gigas* mucus proteome[J]. Mar Drugs, 2024, 22(9): 385. DOI: 10.3390/md22090385.
- [82] Han J, Kong T, Liu J. PepNet: An interpretable neural network for anti-inflammatory and antimicrobial peptides prediction using a pre-trained protein language model[J]. Commun Biol, 2024, 7(1):1198. DOI:10.1038/s42003-024-06911-1.
- [83] Cordoves-Delgado G, Garcia-Jacas C R. Predicting antimicrobial peptides using ESMFold-predicted structures and ESM-2-based amino acid features with graph deep learning[J]. J Chem Inf Model, 2024, 64(10): 4310-4321. DOI: 10.1021/acs.jcim.3c02061.
- [84] Meher P K, Sahu T K, Saini V, et al. Predicting antimicrobial peptides with improved accuracy by incorporating the compositional, physico-chemical and structural features into Chou's general PseAAC[J]. Sci Rep, 2017, 7(1):42362. DOI: 10.1038/srep42362.
- [85] Ng X Y, Rosdi B A, Shahrudin S. Prediction of antimicrobial peptides based on sequence alignment and support vector machine-pairwise algorithm utilizing LZ-complexity[J]. Biomed Res Int, 2015(2015):1-13. DOI:10.1155/2015/212715.
- [86] Olcay B, Ozdemir G D, Ozdemir M A, et al. Prediction of the synergistic effect of antimicrobial peptides and antimicrobial agents via supervised machine learning[J]. BMC Biomed Eng, 2024, 6(1):1-19. DOI:10.1186/s42490-024-00075-z.
- [87] Xia X, Torres M D T, de la Fuente-Nunez C. Proteasome-derived antimicrobial peptides discovered via deep learning [PP/OL]. bioRxiv (2025-06-03) [2026-04-28]. <https://doi.org/10.1101/2025.03.17.643752>.
- [88] Medina - Ortiz D, Contreras S, Fernandez D, et al. Protein language models and machine learning facilitate the identification of antimicrobial peptides[J]. Int J Mol Sci, 2024, 25(16):8851. DOI:10.3390/ijms25168851.
- [89] Hussain W. sAMP - PFPDeep: Improving accuracy of short antimicrobial peptides prediction using three different sequence encodings and deep neural networks[J]. Brief Bioinform, 2022, 23(1):bbab487. DOI:10.1093/bib/bbab487.
- [90] Wu K, Xu G, Tian Y, et al. Synthesis and evaluation of aquatic antimicrobial peptides derived from marine metagenomes using a high-throughput screening approach[J]. Mar Drugs, 2025, 23(4):178. DOI:10.3390/md23040178.
- [91] Guo H, Yu Y, Zhang Z, et al. The discovery of wheat-derived anticaries peptides by a ML-based computational strategy[J]. Bioorg Med Chem, 2025(129): 118303. DOI: 10.1016/j.bmc.2025.118303.
- [92] Zhou W, Liu Y, Li Y, et al. TriNet: A tri-fusion neural network for the prediction of anticancer and antimicrobial peptides[J]. Patterns, 2023, 4(3): 100702. DOI: 10.1016/j.patter.2023.100702.
- [93] Du R, Han F, Li Z, et al. Uncovering encrypted antimicrobial peptides in health-associated *Lactobacillaceae* by large-scale genomics and machine learning[J]. Microbiome, 2025, 13(1): 151. DOI:10.1186/s40168-025-02145-3.
- [94] Barroso RA, Aguero - Chapin G, Sousa R, et al. Unlocking antimicrobial peptides: In silico proteolysis and artificial intelligence-driven discovery from cnidarian omics[J]. Molecules, 2025, 30(3):550. DOI:10.3390/molecules30030550.
- [95] Shen H, Li Y, Pi Q, et al. Unveiling novel antimicrobial peptides from the ruminant gastrointestinal microbiomes: A deep learning-driven approach yields an anti-MRSA candidate[J]. J Adv Res, 2025(77):821-835. DOI:10.1016/j.jare.2025.01.005.
- [96] Kavousi K, Bagheri M, Behrouzi S, et al. IAMPE: NMR-assisted computational prediction of antimicrobial peptides[J]. J Chem Inf Model, 2020, 60(10):4691-4701. DOI: 10.1021/acs.jcim.0c00841.
- [97] Xu J, Xu X, Jiang Y, et al. Waste to resource: Mining antimicrobial peptides in sludge from metagenomes using machine learning[J]. Environ Int, 2024(186): 108574. DOI: 10.1016/j.envint.2024.108574.
- [98] Gafskaia E N, Polina N F, Babenko V V, et al. Discovery of novel antimicrobial peptides: A transcriptomic study of the sea anemone *Cnidopus japonicus* [J]. J Bioinform Comput Biol, 2018, 16(2):1840006. DOI:10.1142/S0219720018400061.
- [99] Henson B A B, Li F, Alvarez-Huerta J A, et al. Novel active Trp- and Arg-rich antimicrobial peptides with high solubility and low red blood cell toxicity designed using machine learning tools[J]. Int J Antimicrob Agents, 2025, 65(1): 107399. DOI: 10.1016/j.ijantimicag.2024.107399.
- [100] Liu H, Song Z, Zhang Y, et al. De novo design of self-assembling peptides with antimicrobial activity guided by deep learning[J]. Nat Mater, 2025, 24(8):1295-1306. DOI:10.1038/s41563-025-02164-3.
- [101] de Albernaz D T F, Allend S O, Neto A C P S, et al. Novel antimicrobial peptides against *Pseudomonas aeruginosa*: In silico design and experimental validation[J]. J Appl Microbiol, 2025, 136(12):lxaf287. DOI:10.1093/jambio/lxaf287.
- [102] Dong R, Liu R, Liu Z, et al. Exploring the repository of de novo-designed bifunctional antimicrobial peptides through deep learning[J]. eLife, 2025(13): RP97330. DOI: 10.7554/eLife.97330.
- [103] Jin S, Zeng Z, Xiong X, et al. AMPGen: An evolutionary information-reserved and diffusion-driven generative model for de novo design of antimicrobial peptides[J]. Commun Biol, 2025, 8(1):839. DOI:10.1038/s42003-025-08282-7.
- [104] Gao H, Guan F, Luo B, et al. DLFea4AMPGen de novo design of antimicrobial peptides by integrating features learned from deep learning models[J]. Nat Commun, 2025, 16(1): 9134.

- DOI:10.1038/s41467-025-64378-y.
- [105] Jiang Z, Feng J, Wang F, et al. AI - guided design of antimicrobial peptide hydrogels for precise treatment of drug - resistant bacterial infections [J]. *Adv Mater*, 2025, 37 (20) : 2500043. DOI:10.1002/adma.202500043.
- [106] Zhang H, Saravanan K M, Wei Y, et al. Deep learning-based bioactive therapeutic peptide generation and screening [J]. *J Chem Inf Model*, 2023, 63(3) : 835-845. DOI:10.1021/acs.jcim.2c01485.
- [107] Pikalyova K, Akhmetshin T, Orlov A, et al. Design of highly potent antibiofilm, antimicrobial peptides using explainable artificial intelligence[J]. *J Chem Inf Model*, 2026, 66(1) : 744-755. DOI:10.1021/acs.jcim.5c01992.
- [108] Zhao W, Hou K, Shen Y, et al. A conditional denoising VAE-based framework for antimicrobial peptides generation with preserving desirable properties [J]. *Bioinformatics*, 2025, 41 (2) : btaf069. DOI:10.1093/bioinformatics/btaf069.
- [109] Liu J, Cui T, Wang T, et al. A multi - property optimizing generative adversarial network for de novo antimicrobial peptide design[J]. *Adv Sci*, 2025, 12(38) : e03443. DOI:10.1002/advs.202503443.
- [110] Zhao W, Hou K, Tang C, et al. A novel generative framework for designing pathogen - targeted antimicrobial peptides with programmable physicochemical properties [J]. *PLoS Comput Biol*, 2025, 21 (12) : e1013833. DOI: 10.1371/journal.pcbi.1013833.
- [111] Wu T, Zhou M, Zou J, et al. AI-guided few-shot inverse design of HDP-mimicking polymers against drug-resistant bacteria[J]. *Nat Commun*, 2024, 15 (1) : 6288. DOI:10.1038/s41467-024-50533-4.
- [112] Wang J, Wu R, Zhang X, et al. AMPGP: Discovering highly effective antimicrobial peptides via deep learning[J]. *J Chem Inf Model*, 2025, 65 (17) : 9303 - 9312. DOI: 10.1021/acs.jcim.5c00647.
- [113] Yang D, Li Y, Li C, et al. Biochemical - knowledge - driven machine learning pipeline for generating potent antimicrobial peptides[J]. *Brief Bioinform*, 2026, 27(2) : bbag115. DOI: 10.1093/bib/bbag115.
- [114] Pandi A, Adam D, Zare A, et al. Cell - free biosynthesis combined with deep learning accelerates de novo-development of antimicrobial peptides [J]. *Nat Commun*, 2023, 14 (1) : 7197. DOI:10.1038/s41467-023-42434-9.
- [115] Boone K, Wisdom C, Camarda K, et al. Combining genetic algorithm with machine learning strategies for designing potent antimicrobial peptides [J]. *BMC Bioinform*, 2021, 22 (1) : 239. DOI:10.1186/s12859-021-04156-x.
- [116] Wang Y, Li M, Huang F, et al. Controllable generation of pathogen - specific antimicrobial peptides through knowledge - aware prompt diffusion model [J]. *Adv Sci*, 2025, 12 (41) : e07457. DOI:10.1002/advs.202507457.
- [117] Cesaro A, Wan F, Torres M D T, et al. Design of multimodal antibiotics using deep learning[PP/OL]. *bioRxiv* (2025-10-31) [2026-04-28]. <https://doi.org/10.1101/2024.12.20.629780>. DOI: 10.1101/2024.12.20.629780.
- [118] Whelan C, Roark B, Sonmez K. Designing antimicrobial peptides with weighted finite-state transducers[J]. *Annu Int Conf IEEE Eng Med Biol Soc*, 2010(2010) : 764-767. DOI:10.1109/IEMBS.2010.5626357.
- [119] Swanson K, Liu G, Catacutan D B. Generative AI for designing and validating easily synthesizable and structurally novel antibiotics [J]. *Nat Mach Intell*, 2024, 6 (3) : 338 - 353. DOI: 10.1038/s42256-024-00809-7.
- [120] Yin K, Xu W, Ren S, et al. Machine learning accelerates de novo design of antimicrobial peptides[J]. *Interdiscip Sci Comput Life Sci*, 2024, 16 (2) : 392 - 403. DOI: 10.1007/s12539-024-00612-3.
- [121] Dong Q, Wang S, Miao Y, et al. Novel antimicrobial peptides against *Cutibacterium acnes* designed by deep learning [J]. *Sci Rep*, 2024, 14(1) : 4529. DOI:10.1038/s41598-024-55205-3.
- [122] Zare - Zardini H, Seyedjavadi S S. Potential application of Healtide-GP1, a novel antibacterial peptide, in wound healing: *In vitro* studies[J]. *Sci Rep*, 2025, 15(1) : 31078. DOI:10.1038/s41598-025-17002-4.
- [123] Dean S N, Walper S A. Variational autoencoder for generation of antimicrobial peptides [J]. *ACS Omega*, 2020, 5 (33) : 20746 - 20754. DOI:10.1021/acsomega.0c00442.
- [124] Dean S N, Alvarez J A E, Zabetakis D, et al. PepVAE: Variational autoencoder framework for antimicrobial peptide generation and activity prediction [J]. *Front Microbiol*, 2021 (12) : 725727. DOI:10.3389/fmicb.2021.725727.
- [125] Tucs A, Tran D P, Yumoto A, et al. Generating ampicillin-level antimicrobial peptides with activity-aware generative adversarial networks [J]. *ACS Omega*, 2020, 5 (36) : 22847 - 22851. DOI: 10.1021/acsomega.0c02088.
- [126] Zhao Y, Kong Q, Gong H, et al. AI - based D - amino acid substitution for optimizing antimicrobial peptides to treat multidrug - resistant bacterial infection [J]. *Adv Sci*, 2026, 13 (10) : e18522. DOI:10.1002/advs.202518522.
- [127] Mishra B, Basu A, Shehadeh F, et al. Antimicrobial peptide developed with machine learning sequence optimization targets drug resistant *Staphylococcus aureus* in mice [J]. *J Clin Invest*, 2025, 135(12) : e185430. DOI:10.1172/JCI185430.
- [128] Randall J R, Vieira L C, Wilke C O, et al. Deep mutational scanning and machine learning for the analysis of antimicrobial-peptide features driving membrane selectivity [J]. *Nat Biomed Eng*, 2024, 8(7) : 842-853. DOI:10.1038/s41551-024-01243-1.
- [129] Cao Q, Ge C, Wang X, et al. Designing antimicrobial peptides using deep learning and molecular dynamic simulations [J]. *Brief Bioinform*, 2023, 24(2) : bbad058. DOI:10.1093/bib/bbad058.
- [130] Fjell C D, Jenssen H, Cheung W A, et al. Optimization of antibacterial peptides by genetic algorithms and cheminformatics [J]. *Chem Biol Drug Des*, 2011, 77 (1) : 48 - 56. DOI:10.1111/j.1747-0285.2010.01044.x.

(收稿日期:2026-05-06)

(本文编辑:王嫚)