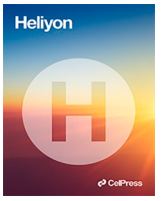
Heliyon 10 (2024) e40265

Contents lists available at ScienceDirect

Heliyon

journal homepage: www.cell.com/heliyon



Review article

Therapeutic peptide development revolutionized: Harnessing the power of artificial intelligence for drug discovery

Samaneh Hashemi [a,b], Parisa Vosough [a,b], Saeed Taghizadeh [b,c,**], Amir Savardashtaki [b,d,*]

[a] Student Research Committee, Shiraz University of Medical Sciences, Shiraz, Iran
[b] Department of Medical Biotechnology, School of Advanced Medical Sciences and Technologies, Shiraz University of Medical Sciences, Shiraz, Iran
[c] Pharmaceutical Science Research Center, Shiraz University of Medical Sciences, Shiraz, Iran
[d] Infertility Research Center, Shiraz University of Medical Sciences, Shiraz, Iran

ARTICLE INFO

Keywords:
Therapeutic peptides
Artificial intelligence
Machine learning
Drug discovery

ABSTRACT

Due to the spread of antibiotic resistance, global attention is focused on its inhibition and the expansion of effective medicinal compounds. The novel functional properties of peptides have opened up new horizons in personalized medicine. With artificial intelligence methods combined with therapeutic peptide products, pharmaceuticals and biotechnology advance drug development rapidly and reduce costs. Short-chain peptides inhibit a wide range of pathogens and have great potential for targeting diseases. To address the challenges of synthesis and sustainability, artificial intelligence methods, namely machine learning, must be integrated into their production. Learning methods can use complicated computations to select the active and toxic compounds of the drug and its metabolic activity. Through this comprehensive review, we investigated the artificial intelligence method as a potential tool for finding peptide-based drugs and providing a more accurate analysis of peptides through the introduction of predictable databases for effective selection and development.

Abbreviations: **AI**, artificial intelligence; **ML**, machine learning; **DL**, deep learning; **ANNs**, artificial neural networks; **QSAR**, quantitative structure-activity relationship; **SGF**, simulated gastric fluid; **SIF**, small intestinal fluid; **SVM**, support vector machine; **SVMC**, support vector machine classifier; **PSMs**, peptide spectrum matches; **XIC**, extracted ion chromatogram; **DNNs**, deep neural networks; **MHC-I**, major histocompatibility complex class I; **PseAAC**, pseudo amino acid composition; **DPC**, dipeptide composition; **AAC**, amino acid composition; **MCC**, Matthews' correlation coefficient; **BiCGAN**, bidirectional conditional generative adversarial network; **FACS**, Fluorescence-activated cell sorting; **CVAE**, conditional variational autoencoder model; **AP**, average precision; **GANs**, generative adversarial networks; **CNNs**, convolutional neural networks; **NMMs**, neural network models; **HSM**, Hierarchical statistical mechanical modeling; **RNNs**, recurrent neural networks; **VAEs**, variational autoencoders; **AUROC**, area under the receiver operating characteristic curve; **KNN**, K-Nearest Neighbor; **RF**, Random Forest; **MD**, molecular dynamics; **LSTM**, long short-term memory; **BiLSTM**, bidirectional long short-term memory; **PTDs**, Protein transduction domains; **CPPs**, cell-penetrating peptides; **AMPs**, antimicrobial peptides; **ACPs**, anti-cancer peptides; **AHTPs**, antihypertensive peptides; **AIPs**, anti-inflammatory peptides; **AFPs**, antifungal peptides; **AVPs**, antiviral peptides; **AOPs**, Antioxidant peptides; **AADs**, Amino acid descriptors; **AUC**, area under the curve; **BPFs**, binary profile features; **GDL**, Graph-based deep learning; **BERT**, bidirectional encoder representations from transformers; **DD**, Deep Docking platform; **CPANNs**, counter-propagation artificial neural networks; **ECM**, extracellular matrix; **NLP**, Natural language processing.

* Corresponding author. Department of Medical Biotechnology, School of Advanced Medical Sciences and Technologies, Shiraz University of Medical Sciences, Shiraz, Iran.

** Corresponding author. Department of Medical Biotechnology, School of Advanced Medical Sciences and Technologies, Shiraz University of Medical Sciences, Shiraz, Iran.

E-mail addresses: Staghizadeh@sums.ac.ir (S. Taghizadeh), dashtaki63@gmail.com (A. Savardashtaki).

https://doi.org/10.1016/j.heliyon.2024.e40265
Received 20 February 2024; Received in revised form 7 October 2024; Accepted 7 November 2024
Available online 8 November 2024

1. Introduction

The discovery of new drugs is essential to healthcare and aims to identify and characterize molecules with the potential to improve patient outcomes and address unmet medical needs [1]. This process can include developing new drugs to treat various diseases, such as chronic conditions and emerging infectious diseases [2]. In recent years, therapeutic peptides, short chains of amino acids with the potential as targeted therapies with high specificity and low toxicity, have attracted considerable attention. They offer advantages such as improved drug delivery, reduced side effects, and increased efficacy compared to traditional small-molecule drugs [3]. While proteins and nucleic acids offer therapeutic potential, peptides possess distinct advantages. Due to their smaller size, peptides have better tissue penetration and faster clearance from the body, reducing potential toxicity compared to larger proteins [4]. Peptides are also easier to design and modify, allowing for improved stability and specificity. In contrast to nucleic acids, which face significant delivery challenges and risk degradation, peptides are more stable and can be tailored to target specific molecular interactions with greater precision [5,6].

More than 80 therapeutic peptides have been approved to treat various diseases, including infectious diseases, cardiovascular diseases, dysmetabolic diseases, and cancers. Hundreds of peptides are currently being studied in preclinical and clinical trials [7]. In recent years, scientific publications and patents on peptide-based therapies have increased exponentially. Therapeutic peptides are expected to continue to attract investment and research efforts due to their therapeutic potential, market prospects, and economic value [6]. In 2017, sales of peptide drugs reached $20 billion, and projections indicate they will surpass $50 billion by 2024. Furthermore, the market size is expected to grow from $25.3 billion in 2024 to $41.7 billion by 2030, with a Compound Annual Growth Rate (CAGR) exceeding 8 % [8]. These market projections reflect the increased interest in peptide-based therapies due to their therapeutic potential and growing demand.

However, discovering new therapeutic peptides faces several challenges. First, the synthesis of peptides and their stability are primary concerns; Peptides are sensitive to enzyme degradation and exhibit limited stability under physiological conditions, necessitating the discovery of peptides that remain active enough in the body to produce effective therapeutic effects [9]. Second, identifying suitable targets for therapeutic peptides is a complex task that involves determining specific proteins or receptors involved in disease pathways and selecting targets that peptides can effectively modulate [10]. Third, peptides often face obstacles related to poor bioavailability and efficient delivery to the target site, such as rapid clearance from the bloodstream and inability to cross the cell membrane, limiting their effectiveness [11]. Overcoming these obstacles is essential to maximize the therapeutic potential of peptides. Fourth, designing peptides with minimal off-target effects and potential toxicity is critical. Peptides should show low immunogenicity and not interact with non-target proteins or receptors [12]. Finally, traditional peptide discovery methods, such as high-throughput screening and combinatorial chemistry, have limitations in exploring the vast chemical space of potential peptide sequences. These methods are time-consuming and expensive and may not efficiently identify promising targets [13].

In recent years, the development of artificial intelligence (AI) has impacted most therapeutic modalities, including targeted protein degradation, antibodies, gene therapy, and oligonucleotide and vaccine design. This advancement has captured the attention of investors, industrial scientists, and legislators, who have shown great interest in applying AI to drug discovery [14]. AI technologies, including techniques such as machine learning (ML), deep learning (DL), and data mining, have offered innovative approaches to overcome the challenges faced in developing and optimizing new therapies [2]. These computational approaches analyze large datasets, identify patterns, and make unprecedented accuracy and efficiency predictions. AI accelerates the discovery and development of therapeutic peptides by analyzing diverse datasets, including genomic information, protein structures, and clinical data [15]. Identifying potential peptide candidates with desirable properties will help solve synthesis, stability, and target selection challenges. Using deep learning and AI methods, novel functional and antimicrobial peptides (AMPs) have been discovered from various sources, including the human proteome and microbiome. As antibiotic resistance increases, there is a pressing need for alternative antibacterial drugs [16]. In addition, AI and deep learning approaches in drug discovery will contribute to identifying functional molecules and developing much-needed medicines [17].

Furthermore, AI helps predict peptide-protein interactions, optimize peptide structures, and even predict pharmacokinetics and toxicity profiles [18]. The integration of AI into the drug discovery process has the potential to significantly reduce the time and cost required to identify and develop novel therapeutic peptides. These capabilities empower researchers to make more informed decisions throughout drug development, enhancing success rates through data-driven insights [19]. It is possible to identify the most effective approaches for specific drug development challenges by compiling comprehensive comparisons of different AI techniques, which have made significant inroads into drug discovery. AI's real-world impact and potential in therapeutic drug development must be demonstrated through more case studies that demonstrate its successful integration. Analyzing these case studies would provide valuable insights into how AI techniques can accelerate drug discovery. By addressing these gaps in knowledge, the field of AI in drug discovery can continue to evolve and optimize its approaches for more effective and efficient therapeutic development [20].

The pharmaceutical industry will benefit from generative AI by developing more advanced algorithms that address the complexity inherent in molecular interactions and human biology. Drug design has been significantly enhanced by AI methodologies that introduce generative algorithms that streamline the identification and refinement of drug candidates. These algorithms utilize vast bioactivity, toxicity, and protein structure data to identify and refine drug candidates. Pharmaceutical companies are integrating these algorithms into their drug design processes, with some AI-designed drugs progressing to clinical trials [21]. However, AI has limitations, including bias in training data and challenges in interpreting complex models [22]. Understanding these limitations is critical to responsible and effective drug discovery integration.

This article presents a comprehensive overview of the challenges in discovering new therapeutic peptides and discusses the

potential of AI in the drug discovery process. Different AI techniques and their applications in introducing therapeutic peptides will be examined, and the challenges and future directions will be discussed.

2. Literature search methodology

To comprehensively review the integration of artificial intelligence in developing therapeutic peptides, we meticulously searched through renowned databases such as Google Scholar, PubMed, Scopus, and Web of Science. These platforms, known for their extensive repository of biomedical literature, provided a broad spectrum of high-quality research articles. The search was tailored using specific keywords, including "therapeutic peptides," "drug discovery," "AI in peptide design," and "biomedical applications," focusing on articles published from 2010 to the present to ensure the inclusion of the most recent and relevant advancements. This period was chosen to reflect the rapid evolution in the field, particularly in integrating AI technologies. Filters were applied to restrict the search to English-language publications to maintain consistency and reliability in understanding and evaluating the studies. Two authors independently reviewed the search results, selecting articles that met predefined inclusion criteria.

- Full-text articles written in English.
- Studies specifically discussing the application of artificial intelligence in peptide discovery.
- Publications detailing therapeutic peptides' synthesis, design, or biomedical applications.

Exclusion criteria were set to streamline the focus of the review, omitting.

- Studies published before 2010 to concentrate on recent advancements.
- Non-English publications to avoid discrepancies in translation and interpretation.
- Papers not centrally focused on peptides or artificial intelligence within a biomedical context.

This methodological approach allowed us to thoroughly explore the state of the field, ensuring a holistic inclusion of significant studies highlighting the latest innovations in peptide discovery and AI-driven methodologies.

3. Traditional vs. AI-driven approaches in peptide discovery

The traditional methods of peptide discovery involve a step-by-step and repetitive screening, synthesis, and optimization process. Initially, researchers screen extensive libraries of compounds to identify potential peptide candidates, which can be time-consuming and expensive [23]. High-throughput screening techniques, where thousands or even millions of compounds are tested for their biological activity tests, have limitations in exploring the vast chemical space of peptides [24]. Once potential peptide candidates are identified, they must be synthesized and tested individually, which can be challenging and costly, especially for longer and more complex sequences [25]. After synthesis, the peptides undergo different assays to evaluate their biological activity, selectivity, stability, and toxicity. This experimental validation process can take significant time and resources [26]. Then, the researchers proceeded to optimize the properties of the peptides. Optimization can include modifying the peptide sequence, backbone, or side chains to improve potency, selectivity, or other desired properties. However, the optimization process often involves trial and error and is highly dependent on the expertise and intuition of researchers [3]. Although traditional approaches have led to the discovery of several successful therapeutic peptides, they have limitations. This process is time-consuming and requires years of research and development. It is also resource-intensive and requires significant financial investments and peptide chemistry and biology expertise. In addition, traditional approaches may not fully explore the vast chemical space of peptides and have limitations in discovering new and effective therapeutic candidates.

AI-driven approaches are a promising alternative to traditional methods in discovering therapeutic peptides. AI for analyzing large datasets uses various computational techniques, including pattern recognition and making predictions [27]. These techniques enable researchers to systematically and comprehensively explore the chemical space of peptides. Virtual screening is a crucial advantage of AI-driven approaches [28]. AI algorithms can display large compound databases virtually instead of physically testing individual compounds. By analyzing structural and functional data, AI can identify potential peptide candidates with desirable properties, such as high potency, selectivity, and low toxicity. This virtual screening process significantly reduces the time and resources required compared to traditional initial screening. Machine learning algorithms play an essential role in AI-based peptide discovery. These algorithms can analyze large datasets of known peptides and their biological activities to predict the activity of new peptides. Machine learning models can make accurate predictions about the biological activity of peptides by learning from the patterns and relationships in the data. This enables researchers to prioritize and select candidates for further experimental validation [29,30].

By using artificial intelligence algorithms, it is possible to analyze the structural features of target peptides and proteins and predict the binding affinity and specificity of peptide-target interactions. Thus, AI can also help predict peptide-protein interactions [31]. This information is invaluable in selecting highly probable peptides binding to their intended targets. Furthermore, AI can assist in designing peptide scaffolds and optimizing peptide synthesis routes. ML models can learn from existing peptide structures and properties to generate novel sequences with desired characteristics [32]. AI algorithms can also optimize the synthesis process by predicting the most efficient and cost-effective routes for peptide synthesis. Table 1 compares the benefits and drawbacks of conventional techniques and those that rely on AI in peptide exploration.

In peptide discovery, AI-powered approaches present numerous benefits compared to conventional methods. These AI-driven

methods not only speed up the discovery phase and cut down on expenses but also delve into a vastly expanded chemical space. AI furnishes valuable guidance by offering insights and predictions, aiding researchers in making informed decisions and identifying the most promising peptide candidates for subsequent experimental validation. By harnessing the power of AI, the anticipation is high for unearthing innovative and potent therapeutic peptides tailored to combat a spectrum of diseases.

4. AI techniques enhancing peptide discovery

AI strategies have revolutionized the design and discovery of peptides [33]. Some AI strategies, such as machine learning, deep learning, and generative models, have been used to improve the synthesis of efficient peptides and their properties [34,35]. AI algorithms can validate chemical compounds, identify targets, monitor drugs, and determine their effectiveness [36]. It is necessary to integrate AI and peptide libraries to develop innovative peptides. This combination offers unique advantages in activity, bioavailability, and safety [37]. Early computational methods in peptide design included quantitative structure-activity relationship (QSAR) modeling and molecular dynamics (MD) simulations. QSAR models considered pharmacophore and molecular shape factors relevant to a peptide's activity through molecular descriptors. Molecular dynamics simulations were used to study the mechanism of antimicrobial peptide action by simulating peptide interactions with membranes [38,39]. The application of AI in drug discovery has progressed significantly, starting with basic machine learning algorithms and evolving to sophisticated deep learning models. Key milestones include the development of AI-driven QSAR models, the advent of AlphaFold for protein structure prediction, and the integrating of AI in designing AMPs [12].

Several studies have demonstrated the potential of AI in therapeutic peptide development. For example, using AI models facilitated the development of selective peptide therapies in peptidomics [40]. The emergence of AlphaFold and other AI techniques has enhanced peptide-based drug discovery, offering new opportunities for innovation [41]. Another significant study applied AI for the accelerated discovery of antimicrobial sequences, showcasing the practical benefits of AI integration in peptide research [5].

Methods of bioinformatics, drug discovery, and protein structure prediction all rely on feature encoding methods for peptide sequences [42,43]. Peptide-level features capture physicochemical or structural properties of peptide sequences. These characteristics can help predict a peptide's biological activity or understand its function by better representing the peptide [44]. Several bioinformatics features can be used to define and forecast the properties of peptides based on their structure and sequence [45]. Molecular docking and molecular simulations provide insight into the functions and interactions of the peptides and validate predictions [46].

4.1. Machine learning applications in peptide discovery

ML algorithms are pivotal in analyzing peptide sequences and prediction-making based on training data. Support Vector Machine (SVM), RF (Random Forests), and K-nearest Neighbors (KNN) are some of the commonly used machine learning methods [47]. SVM is used chiefly to describe the molecular structure and mode of action [48]. KNN is a classifier based on the similarity between data points [49]. Peptide-protein interactions and regression and classification problems have been studied using RF [47,50]. ML substantially enhances the efficiency and effectiveness of identifying novel therapeutic peptides by employing data-driven insights and high-throughput screening processes. The analysis of large datasets of peptide sequences, biological activities, and structural information by ML algorithms reveals patterns that assist in predicting peptides with desired therapeutic properties. This method minimizes the necessity for exhaustive laboratory experiments by prioritizing intriguing peptide candidates [51].

Integrating molecular docking with ML accelerates peptide ligand screening. LightGBM, among eight ML algorithms, demonstrated superior efficiency and accuracy in predicting peptide-protein docking results [52]. In a study with 160,000 tetrapeptides targeting viral envelope proteins, training LightGBM on just 1 % of the dataset accurately predicted the remaining 99 %, significantly reducing computational costs. This approach, independent of docking software, increases screening speed tenfold while maintaining precision, making it an efficient tool for discovering bioactive compounds [53]. In another study, researchers developed a transparent ML framework combined with a genetic algorithm to optimize sequence space and design AMPs specifically targeting the oral pathogen *Streptococcus gordonii*. Using set theory to define decision boundaries, the model classifies peptide sequences based on their

Table 1
Comparing efficiency: conventional vs. AI-driven peptide discovery.

Aspect	Conventional Approaches	AI-Driven Approaches
Time and Resource Consumption	Extensive, requiring years of research and development	Accelerated, reducing time and resource requirements
Screening Process	Time-consuming physical testing of individual compounds	Virtual screening through computational analysis
Chemical Space Exploration	Limited exploration of the vast chemical space of peptides	Systematic and comprehensive exploration of chemical space
Synthesis and Testing	Challenging and costly, especially for longer sequences, using chemical and biochemical methods	Enhanced efficiency and cost-effectiveness in synthesis, utilizes neural networks, Generating innovative peptides without prior knowledge
Biological Activity Prediction	Relies on experimental validation and trial-and-error	Machine learning models predict activity based on data
Peptide Design and Optimization	Guided by trial-and-error and researcher intuition	Machine learning assists in designing and optimizing properties
Sequence Design	The chemical synthesis of peptides	Developing peptide sequences using deep learning algorithms

antibacterial activity. This approach overcomes the limitations of traditional neural networks by providing high specificity in AMP classification. Moreover, the system is adaptive and capable of incorporating new data without retraining, thanks to the integration of Modified Learning, making it a flexible and efficient tool for AMP discovery [54].

4.2. Deep learning innovations in peptide discovery

A subset of ML, deep learning algorithms based on artificial neural networks (ANNs), have emerged as a powerful tool for peptide discovery. These algorithms leverage multiple layers of nonlinear processing to generate highly accurate predictions [55]. Deep generative models have been successful in creating novel drug-like molecules and identifying potential antimicrobial peptides, paving the way for further laboratory studies [56,57].

Advanced deep learning models such as recurrent neural networks (RNNs), convolutional neural networks (CNNs), variational autoencoders (VAEs), bidirectional encoder representations from transformers (BERT), and generative adversarial networks (GANs) play a crucial role in the design and synthesis of synthetic peptides [58–61]. For instance, RNNs, specifically the Long Short-Term Memory (LSTM) model, are instrumental in de novo peptide design, overcoming challenges in sequence generation [62,63]. CNNs are adept at recognizing local patterns and are widely applied in MHC–peptide binding prediction, while deep neural networks (DNNs) represent an enhanced form of ANNs, offering superior prediction capabilities [55].

Recent advancements in deep learning have notably impacted peptide discovery, particularly in predicting multi-functional therapeutic peptides (MFTPs). Multi-label deep learning methodologies enable the simultaneous classification of peptides across various functional categories, advancing our understanding of their therapeutic potential. Ensemble learning techniques, which combine multiple models, further enhance predictive performance by leveraging diverse perspectives [64].

Researchers have also successfully integrated deep learning with cell-free protein synthesis (CFPS) to accelerate the design, production, and screening of over 500 AMPs within 24 h. This process, which used deep generative variational autoencoders to generate novel peptides, resulted in the discovery of 30 functional AMPs, including six with strong antimicrobial activity against multidrug-resistant pathogens [59] (see Fig. 1).

Additionally, UniDL4BioPep employs a pretrained biological language model for peptide sequence embeddings, providing a universal deep-learning architecture for bioactive peptide classification. It supports efficient transfer learning, enabling high-precision predictions without requiring deep expertise. The model has demonstrated exceptional performance in 15 out of 20 peptide bioactivity prediction tasks, significantly improving metrics such as accuracy, Matthews correlation coefficient, and area under the curve (AUC) [65]. Fig. 2 provides an insightful overview of the machine learning and deep learning techniques integral to therapeutic peptide discovery.

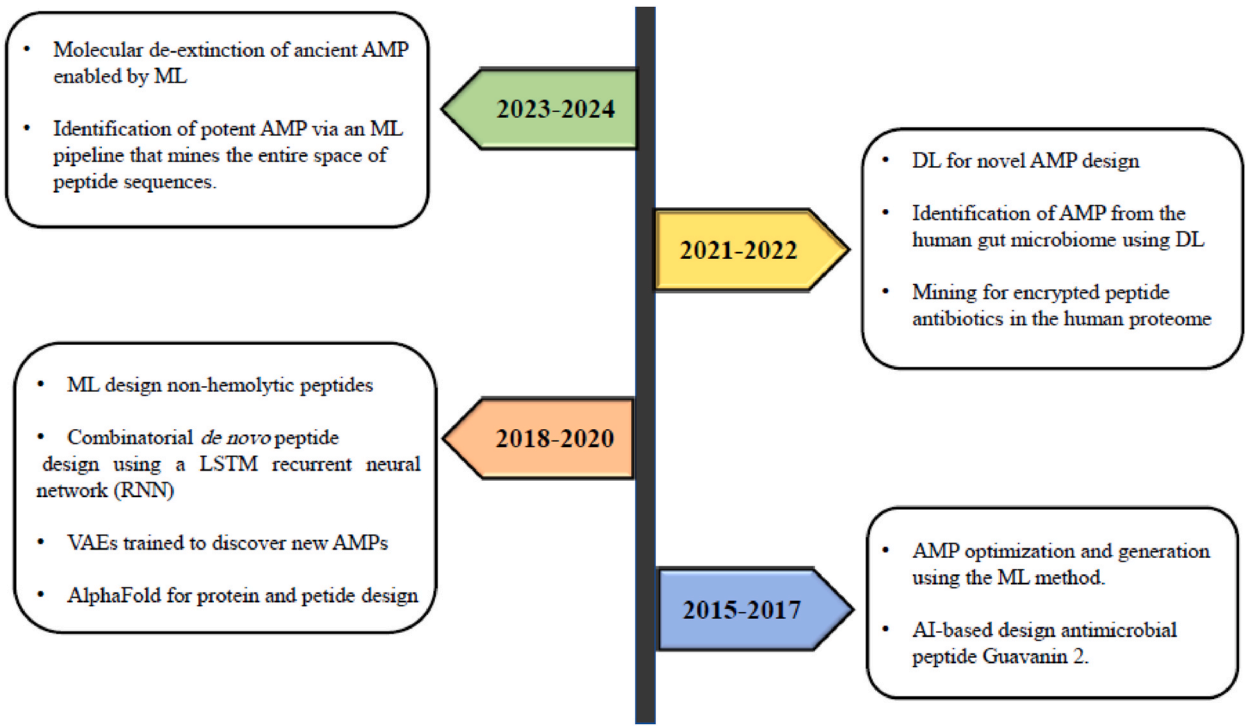


Fig. 1. Timeline of AI advancements in peptide discovery.

5. AI-driven advances in peptide design

Peptide sequence design has become a crucial aspect of drug development, especially in the medical and pharmaceutical industries, given the high value of peptides [66]. ML techniques have proven instrumental in anticipating, creating, and optimizing biological sequences and metabolic pathways [67]. DL algorithms, specifically those based on ANNs, play a significant role in peptide sequence design [55]. Table 2 shows some AI tools in peptide design. Various approaches are commonly employed, each offering a unique perspective. Structural-based design, protein mimicry, and short motif design are key methods for creating therapeutic peptides, emphasizing 3D structures. The integration of AI software enables the training of deep generative models, combined with DL, facilitating the design of novel synthetic peptides with therapeutic properties [68].

Three deep-learning sequence-based prediction models have been developed to design peptide properties, encompassing hemolysis, solubility, and resistance to nonspecific interaction. Each model is meticulously trained on specific datasets, and their performance is rigorously assessed in terms of accuracy and the area under the receiver operating characteristic curve (AUROC). The RNNs model simplifies peptide feature design, facilitating accessibility and reproducibility in predicting the solubility of short peptides. Notably, the web-based MahLooL outperforms current state-of-the-art methods, achieving the highest accuracy [69]. Recent studies showcase the versatile applications of machine learning in predicting peptide drug stability in simulated gastric and small intestinal fluids. Models such as KNN and XGBoost, coupled with feature selection techniques, demonstrate effectiveness in predicting stability based on factors like molecular weight, hydrophobicity/hydrophilicity, and conformational flexibility [70].

5.1. Enhancing peptide vaccine efficacy with AI

Regarding therapeutic potential in vaccines, selecting peptides firmly bound to the major histocompatibility complex (MHC) is crucial. The mouse MHC I allele model (H-2Db) trains support vector machine classifier (SVMC) models to predict binding peptides. Residues 5 and 9 in the peptide sequence emerge as crucial predictors of binding affinity, underscoring their essential role. While models based on physicochemical descriptor sets demonstrate similarity to an advanced tool like NetMHCpan in accuracy and specificity, their sensitivity exhibits a lower rate [71]. In a separate study aimed at improving the accuracy of MHC binding predictions, docked peptide-MHC models are generated using an ANN method. Based on a subset of 1271 nonameric peptides interacting with a murine MHC allele (H-2Db) with known binding affinity, the structure-based approach significantly improves false positives compared to NetMHCpan 4.0. When tested on human alleles, an increase in positive predictive value is observed for tightly binding peptides [72].

ForestMHC, utilizing mass spectrometry, discovers an extensive database of human peptides for training RF classifiers. It exhibits superior precision compared to NetMHC and NetMHCpan, with RF scores correlating with known chemical binding affinities. Notably, positions two and nine, recognized as anchor positions, prove particularly significant in the classification process. Additionally, the study demonstrates gene expression's role in presenting peptides by MHC-I [73].

5.2. Facilitating the development of peptide databases by AI

Introducing the Peptipedia web application, an easily navigable peptide database for researchers, integrating information from 30

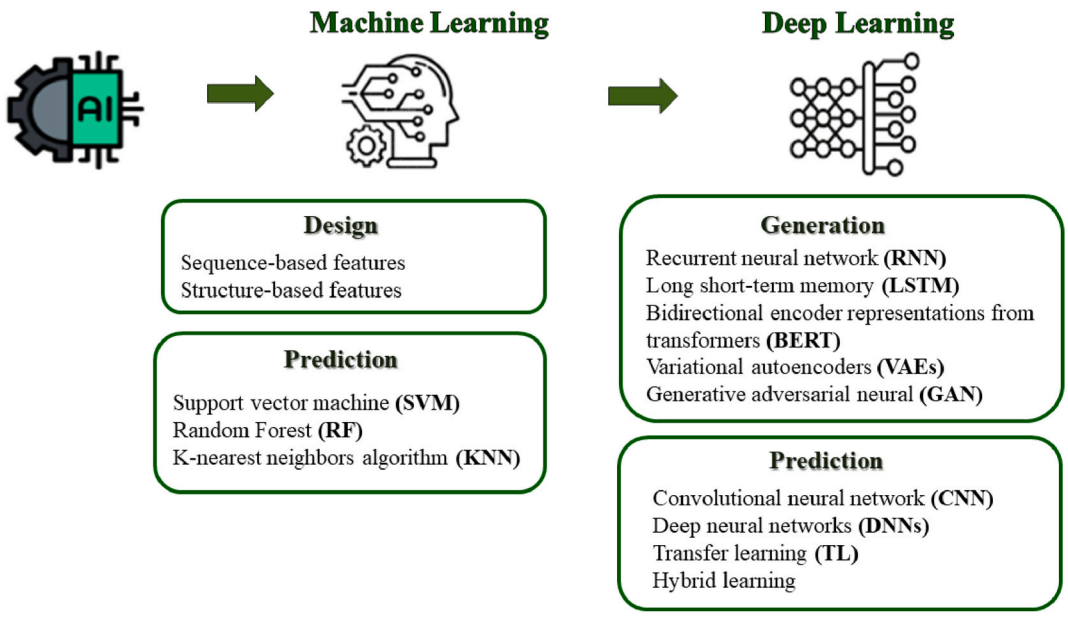


Fig. 2. An overview of the machine learning and deep learning techniques used in therapeutic peptide discovery.

Table 2
Some AI tools in peptide design.

AI tools	Methodology	Applications	Refs
NetMHCpan	Training support vector machine classifier	Predict peptide binding and affinity	[71]
ForestMHC	training random forest classifiers	Predict peptide presentation by MHC with known chemical binding affinities	[73]
HemoPImod	many machine learning methods, including logistic regression, support vector classifiers, decision trees, Random Forest, gradient boosting	Discover and design non-hemolytic peptides	[74]
A.M.P.G.A.N. v2	BiCGAN model	AMP design tool with specific properties	[75]
BChemRF-CPPred	artificial neural networks, support vector machines, and Gaussian process classifiers	Predict cell-penetrating peptides	[76]
Peptipedia	A web application for an easy-to-use peptide database and a binary classification system	Predict putative biological activities	[77]
GM-Pep	A conditional variational autoencoder model and a deep multiclassifier model	Design de novo functional peptide sequences	[78]
MahLooL	recurrent neural networks (RNN)	Predict peptide solubility	[69]

previously published databases using machine learning. With 92,055 registered peptide sequences, the system trains 44 models for different biological activity categories, achieving over 83 % accuracy compared to previously developed classification systems. Utilizing a binary classification system, Peptipedia can predict putative biological activities for peptide sequences [77].

5.3. Novel AI methods for designing peptides

A novel GM-Pep approach involves de novo designing functional peptide sequences with a single bioactivity, utilizing deep learning based on the physicochemical properties of the training data. The generated sequences exhibit statistical consistency (p-value >0.05), with the deep multiclassifier model providing up to 96.41 percent sequence accuracy. Importantly, peptides generated with AI exhibit a shorter length than random peptides, simplifying synthesis and enabling the discovery of potential therapeutic peptides [78]. The HemoPImod ML algorithm is also introduced for discovering and designing non-hemolytic peptides, identifying interesting sequences, and addressing challenges in peptide-based drug design. Utilizing three publicly accessible datasets (HemoPI-1 to HemoPI-3, Antimicrobial Peptide Dataset - APD, and a dataset of known hemolytic antimicrobial peptides - HAMP), various ML algorithms are employed. Gradient-boosting classifiers and extreme gradient-boosting classifiers prove more accurate in predicting the hemolytic nature and activity of peptides, achieving accuracy rates ranging from 95 % to 97 % [74].

5.4. AI-driven peptide target identification

Target identification represents a pivotal phase in the drug development process. Integrating rational peptide design with predictive modeling empowers scientists to systematically explore a vast array of peptide sequences, identifying promising candidates for further investigation [79]. Proteomic methods strive to pinpoint biological targets that exert the desired therapeutic impact while minimizing side effects [80]. Target identification methods span computational, multi-omics, and experimental approaches, widely utilized in drug development [81]. Integrating AI algorithms into the computational identification of potential drug targets is a powerful tool for analyzing extensive biological data, encompassing genetic information, protein structures, and signaling pathways. This approach proves instrumental in curing and preventing various diseases and disorders [82]. Predicting peptides with high response rates for targeted proteomics involves a data-independent acquisition experiment with equimolar synthetic peptides, training an ANN with eleven physiochemical properties. PREGO software significantly enhances the selection of high-response peptides, improving accuracy in predicting peptide responses [83].

5.5. AI-driven personalized therapeutic predictions

Integrating multi-omics data with AI algorithms enhances the understanding of peptide-target interactions and aids in developing personalized peptide therapies [84]. This approach allows for the rapid validation of candidate antimicrobial peptides by integrating diverse protein sequences, including those from the extinct human proteome. AI algorithms can identify patterns and interactions within complex biological systems that are not apparent when analyzing individual omics levels. This leads to a more targeted, effective, and personalized approach to drug discovery [85]. For instance, AI can detect genetic variants associated with disease susceptibility or drug response by analyzing genomic data [86]. A multi-omics approach can reveal molecular interactions and identify new therapeutic targets for various diseases, including cancer. It also helps in understanding how disease mechanisms and outcomes are influenced by proteome complexity. Identifying molecular signatures of diseases and drug responses, prioritizing potential drug targets, and designing validation experiments can guide the development of computational models and experimental design [87].

6. AI techniques for identifying therapeutic peptides

The surge in experimental and computational peptide data has prompted bioinformaticians to harness ML techniques for identifying therapeutic peptides, including AMPs, anti-cancer peptides (ACPs), antiviral peptides (AVPs), antihypertensive peptides

(AHTPs), anti-inflammatory peptides (AIPs), and antifungal peptides (AFPs) [33]. Brief introductions to some methods and algorithms for peptide prediction are presented in Table 3.

In the pursuit of novel anti-diabetic peptides targeting factors contributing to diabetes and blood pressure, a machine-learning approach is employed. The predictive architecture is refined through a predict-test-refine loop using active learning techniques. In vitro experiments, guided by the model's predictions of peptides with specific GLUT4 translocation activity, resulted in a 74 % success rate [88]. To improve the prediction of antihypertensive peptides using DL methods, the model outperformed state-of-the-art techniques on independent and benchmarking datasets. Combining feature extraction with CNN and SVMC, the ensemble method exhibited high accuracy [89]. The machine learning model dedicated to these peptides, Ensemble-AHTPpred, enhances performance by incorporating various computed features and optimizing weighted classifiers [90].

PredAPP is a computational tool for predicting anti-parasitic peptides (APPs). It addresses class imbalance through undersampling and utilizes machine learning to combine nine feature groups and six algorithms, creating 54 classifiers. Logistic regression then integrates the best classifiers, resulting in PredAPP achieving superior accuracy (0.880) and AUC (0.922) compared to existing methods like AMPfun [91]. Similarly, BBPpred is designed to predict blood-brain barrier peptides (BBPs). It employs logistic regression, extracting amino acid sequence features and selecting seven key features to improve accuracy. BBPpred demonstrated strong performance, achieving AUC and AUPR scores of 0.8764 and 0.8757 in 10-fold cross-validation on datasets containing 119 BBPs [92].

LSTM models are employed for de novo peptide generation to generate novel peptides with therapeutic efficacy, refined through training on known bioactive peptides. This method combines docking, DL, and MD simulation to assess potential active peptides targeting specific proteins. Techniques like Rosetta for flexible docking, ZDOCK for attaching peptides to target regions, and Tr-rosetta for generating 3D peptide structures from sequences are utilized. An iterative screening method employing LSTM_pep and DeepPep identifies peptides with enhanced affinity for their targets [56].

6.1. Advancing AMP discovery through AI

One of the most important medical challenges in the world is the issue of antibiotic resistance, which causes deaths and imposes high economic costs on governments [114]. The successful design of antimicrobial peptides is essential in solving this challenge. In a research in 2023, GAN-designed peptides demonstrated antibacterial effects, even against bacteria resistant to antibiotics [115]. Integrating genetic algorithms and machine learning strategies offers a promising avenue for designing potent AMPs. A genetic algorithm uses codon representation for peptides to expedite the AMP design process. Following the initial generation, a crucial step involves modifying the peptides to preserve their antimicrobial activity and simplify synthesis, ensuring their effectiveness and practicality for real-world applications [116]. AMPGAN v2 is an AMP design tool based on ML. Innovative AMP candidates with specific properties are generated using the BiCGAN model. Conditioning variables were used to get the training data for AMPGAN v2. Models made new and different sequences with AMP characteristics. This algorithm used the MIC50 levels for target mechanisms and sequence length to construct the conditioning vectors more effectively than the previous version [75].

Furthermore, Researchers employed DL algorithms and MD simulations to identify novel AMPs with significant medicinal potential. The model achieved remarkable outcomes through training on DL techniques and utilizing an extensive dataset of peptides

Table 3
Some AI tools and algorithms in peptide identification and prediction.

Prediction type	AI Tools	Methodology	Refs
Predict Antimicrobial Peptides (AMPs)	Deep-AmPEP30	Convolutional Neural Network (CNN)	[93]
	HMD-AMP	Deep Forest	[94]
	LMPred	Deep Learning	[95]
	sAMP-VGG16	image-DNN prediction	[96]
	AMP-BERT	Deep Learning	[97]
	AMP-EBiLSTM	Deep Learning	[98]
Predict Anti-cancer Peptides (ACPs)	xDeep-AcPEP	Convolutional Neural Network (CNN) and multitask learning (MTL)	[99]
	ACPred-LAF	Machine Learning	[100]
	ACPNet	Hybrid learning	[101]
	ACP-MCAM.	Convolutional Neural Network (CNN) and attention model	[102]
	ACP_MS	Based on feature extraction	[103]
	C.A.C.P.P.	Convolutional Neural Network (CNN)	[104]
Predict Antiviral Peptides (AVPs)	AVPIden	Based on multiple peptide descriptors and Machine Learning	[105]
	FIRM-AVP	Generative Adversarial Network (GAN)	[106]
	AI4AVP	Generative Adversarial Network (GAN)	[107]
Predict Antioxidant Peptides (AOPs)	AnOxPP	Bidirectional Long Short-Term Memory (BiLSTM)	[108]
	AnOxPePred	Deep Learning	[109]
Predict Anti-Inflammatory Peptides (AIPs)	AIPStack	Machine Learning	[110]
	IF-AIP	Machine Learning	[111]
Predict Antifungal peptides (AFPs)	iAFP-ET	Combining hybrid features	[112]
	DeepAFP	Deep Learning	[113]
Predict antihypertensive peptides (AHTPs)	Ensemble-AHTPpred	Machine Learning	[90]
anti-parasitic peptides (APPs)	PredAPP	Machine learning	[91]
blood-brain barrier peptides (BBPs)	BBPpred	Machine learning	[92]

from 19 public databases. The classification of peptides based on their antimicrobial activities, encompassing properties such as antimicrobial characteristics and Minimum Inhibitory Concentrations (MICs) against various bacterial strains, demonstrated the model's comprehensive understanding. Compared to existing methods, the AMPs-Net algorithm accurately and efficiently predicted the average precision of 95.76 percent in antimicrobial activity and antiviral resistance [117]. For mining AMPs, some NNMs models are combined to form a pipeline using AI-assisted methods, like LSTM (long short-term memory) and Bidirectional Encoder Representations from transformers (BERT). Predicting AMPs with the pipeline was highly precise and had a low false-positive rate. Analyzing association networks and cross-validating using metaproteomic data reduced the number of potential AMPs. The antibacterial activity of novel AMPs was successfully identified against multidrug-resistant bacteria. One of these AMPs did not develop resistance in mouse models infected with *Klebsiella pneumoniae* [118].

In a recent study for antimicrobial activity prediction, some methods based on ML, like the three-dimensional structure of peptides and sequences, are converted into graphs, physicochemical properties are extracted, and the graphs are converted back into peptide sequences. The graph-based deep learning (GDL) method, which utilizes deep neural models in non-Euclidean settings, achieves superior reliability, accuracy, and area under the curve (AUC) compared to alternative methods. GDL has potential applications in predicting and designing innovative targeted molecules based on AMP structure modeling [119]. Deep-AmPEP30, a DL-based method, also indicates short-length AMPs from genomic sequences. Its optimal feature set, derived from reduced amino acid compositions and combined with CNNs, proves effective, with identified peptides exhibiting potency comparable to ampicillin [93]. A recent study proposes improvements to AMP identification, employing a machine-learning pipeline with multiple stages, including empirical selection, classification, and regression analysis. The dataset is enhanced with various preprocessing steps, leading to the identification of C-terminal amidated peptides with robust antimicrobial activity against *S. aureus* [120].

ML techniques are crucial in identifying AMPs from the human gut microbiome. A pipeline integrating multiple Natural Language Processing (NLP) neural network models was developed to predict AMPs and applied to large-scale metagenomic data. The pipeline effectively identified 2349 AMP candidates, and upon chemically synthesizing 216 of these peptides, antimicrobial testing demonstrated a high percentage (83.8 %) of antibacterial activity. Notably, the AMPs exhibited potential against antibiotic-resistant Gram-negative bacteria, such as *Escherichia coli* DH5α and *Pseudomonas aeruginosa*. They effectively reduced bacterial burdens in a murine model of bacterial lung infection [121].

In addition, AI can accelerate and automate lab processes, including generating novel AMPs, peptide synthesis, and multi-assay validation. Advances in optimizing AMP properties, improving evaluation methods, and automating lab processes are necessary to advance AI-driven AMP discovery [122].

6.2. Advancing CPP discovery through AI

Cell-penetrating peptides (CPPs) are being prioritized and understood using molecular dynamics simulations and AI. MD simulations were used to rank peptides based on their permeability. In all simulations, Pep-MD, the top-scoring peptide, showed conformational changes. According to wet-lab experiments, Pep-MD has better permeability and lower toxicity than a clinically used CPP called TAT. Fluorescence-activated cell sorting (FACS) analysis and confocal laser scanning microscopy demonstrate Pep-MD's effective internalization in over 45 % of fluoresced cells [123]. The model generated for this peptide, BChemRF-CPPred, relies on ANN, SVM, and Gaussian process classifiers. Four features significantly impact the algorithm's performance, including compositions (FCs), predicting CPPs with FC-4, and combining structure- and sequence-based descriptors. Notably, FC-4 exhibits the highest accuracy rate at 90.66 %, surpassing other machine-learning-based tools in accuracy, sensitivity, and specificity [76].

DL identified a short nuclear targeting peptide, P6, to efficiently deliver antisense oligomers. Using extended sequences of PMO-CPP in the training process, the DL algorithm optimized short CPPs with minimal arginine content. Flow cytometry evaluated the model's performance by analyzing the fluorescence emitted by PMO after nuclear transport. The model optimizes sequences by increasing predicted activity, minimizing length and arginine content, reducing potential toxicity to cells, and enhancing the efficient delivery of therapeutic macromolecules [124].

6.3. Advancing ACP discovery through AI

Anti-cancer peptides possess potential for anti-cancer treatment due to their selectivity, high penetration, minimal side effects, and ease of chemical modification [33]. AI can expedite the development of ACPs, enhancing efficiency and reducing costs in cancer therapy [125]. The DeepImmuno-CNN method employs beta-binomial distributions to evaluate peptides based on their sequences alone. DeepImmuno-CNN identifies critical residues in T-cell antigen recognition by outperforming traditional ML models and other immunogenicity prediction algorithms. Additionally, the method employs a Generative Adversarial Network (GAN) model in DeepImmuno-GAN to simulate immunogenic peptides with similar physicochemical properties to real antigens [126].

The xDeep-AcPE predicts the biological activity against six tumor cells of ACP. Feature importance weights determine each residue's contribution to the expected activity. The data testing was conducted on the CancerPPD dataset, and the training data were preprocessed to prevent redundancy. In addition, the study examines the effectiveness of the models at different data scales and determines the applicability domain for improving prediction accuracy [99].

ACPred-LAF, an anti-cancer peptide predictor grounded in ML, leverages feature embedding to learn from 29 different features for generating feature descriptors in the training dataset. The integration of LAF and hand-crafted features represents a significant enhancement in prediction performance [100]. In another study, ACP-DA focuses on augmenting prediction performance through data augmentation. Peptide sequences are encoded using binary profile features (BPFs) and amino acid database characteristics, with

additional samples added within the training set feature space during augmentation to improve accuracy in distinguishing between ACPs and non-ACPs [101].

A significant study utilized counter-propagation artificial neural networks (CPANNs) to create ACPs targeting lung and breast cancer cells. The modLAMP de novo design engine generated 1000 peptide sequences with helical amino acid distributions resembling those found in ACPs. By employing an ensemble approach that includes four CPANNs, each designed to recognize different sequence-activity patterns, researchers identified peptides with a predicted activity of at least 50 %. These highly active peptides were then chosen for synthesis and in vitro testing. Fourteen peptides were synthesized and tested on breast cancer (MCF7) and lung cancer (A549) cells, revealing six synthesized peptides with anti-cancer activity, including five showing activity against both cell lines [127].

6.4. Advancing AVP discovery through AI

The AVPIden is an antiviral peptide identification and characterization method based on ML that can predict antiviral activity and identify targeted viruses. This algorithm uses multiple peptide descriptors and imbalanced learning strategies to improve prediction performance. The first step is to use a classifier trained on a comprehensive peptide dataset to distinguish AVPs from other peptides. Then, through cross-validation, metrics such as AUC and GMean are used to evaluate the algorithm's performance [105].

6.5. Advancing AOP discovery through AI

Antioxidant peptides (AOPs) have a therapeutic impact on diseases by reducing reactive oxygen species (ROS) levels. AOPs are known for their ability to scavenge free radicals and reduce oxidative stress, a critical factor in various diseases [128]. These peptides can capture and neutralize ROS, thus preventing the oxidative damage associated with cellular components like proteins, lipids, and DNA. Reducing ROS levels contributes to mitigating the progression of diseases linked to oxidative stress, such as neurodegenerative disorders, cardiovascular diseases, and certain types of cancer [129]. A bidirectional Long Short-Term Memory (BiLSTM) neural network is used to develop the AnOxPP platform. Applied amino acid descriptors (AADs) are also applied to convert the sequence characteristics of peptides into feature codes for input into the model. AnOxPP 's platform identifies AOP with steric, hydrophobic, electronic properties, and hydrogen bond contributions [108].

6.6. Advancing AIP and AFP discovery through AI

For treating inflammation, the AIPStack model for predicting AIPs is based on an ensemble model using two stacking layers and a dataset from the Immune Epitope Database (IEDB). A 10-fold cross-validation was used to evaluate and compare the model with existing methods. Various metrics, including MCC (Matthews' correlation coefficient), AUC, accuracy, and precision, showcase the model's high performance in composition analysis and the selection of ML algorithms and feature encoding for building models [110].

Pretrained protein models prove highly effective as an antifungal target peptide feature extractor. Peptides with negative and positive inhibitory concentration values against fungal species are obtained from the DBAASP database, and six pre-trained protein models are assessed using various ML classifiers [130]. To enhance antifungal prediction accuracy, the iAFP-ET model incorporates

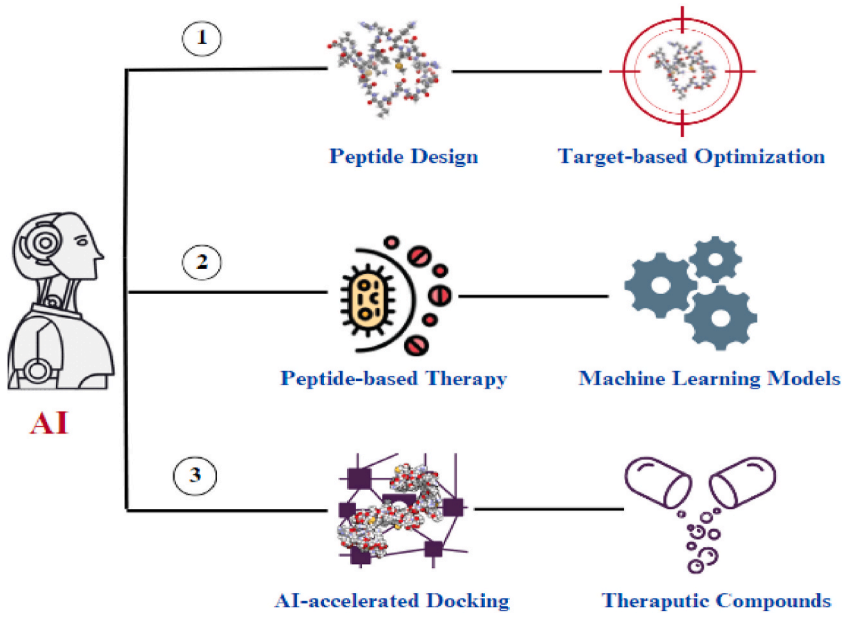


Fig. 3. AI in peptide-base-drug discovery.

feature extractions using the same method [112].

7. AI revolutionizes the drug discovery process

Artificial intelligence significantly contributes to drug development in multiple ways. It provides access to novel biological systems, superior chemistry, higher success rates, and more efficient innovation trials. AI techniques are used in various stages of drug development, including real-time image-based cell sorting, quantum physics calculations, in-silico organic synthesis, assay construction, and forecasting the 3D structures of target proteins [131]. AI also enhances the management and development of pharmaceutical products, clinical trial design and observation, and pharmaceutical manufacturing.

AI streamlines the drug discovery process by identifying predictive biomarkers, validating targets, and analyzing digital pathology data, reducing costs and improving efficiency [132]. In drug discovery, molecular libraries are created to identify new drug candidates with optimal properties to predict protein biological functions, and DL plays a crucial role in these processes [133]. The drug development journey involves various data-driven mechanisms, including target identification, drug screening, assessment, and preclinical and clinical trials [134]. The prediction of target-drug interactions is facilitated by ML methods, which can also assist in modeling drug metabolism and its relationship with biological outcomes and metabolic processes [135]. The Deep Docking platform (DD) expedites drug discovery by enabling rapid and efficient docking of billions of molecular structures [136]. Fig. 3 illustrates the application of AI in peptide-based drug discovery.

AI approaches and docking methods were used to screen potential peptides to treat liver cancer. Toll-like receptor-2 and casein kinase-2 were selected as targets, with the ADME tool predicting pharmacokinetic properties and toxicity. Docking employs tools such as DOCK, CB Dock-2, and AI approaches, including SVM, Gaussian classifiers, and cosine similarity. Docking analysis identified Phe-Tyr as a potent blocker of casein kinase-2, while cosine similarity analysis revealed Pro-Ser as most similar to standard drugs [137]. Moreover, AI-expert, developed as an ML approach, expedites the discovery of self-assembling peptide sequences. The method combines RF, Monte Carlo tree search, and MD simulations. An experimental scoring system, based on sample opacity versus HPLC reaction time, was established to assess the success of human and AI-expert algorithms. AI-experts outperformed human specialists with a success rate of 66.7 % [138].

The peptide Pep_RTE62G, identified by AI, is under investigation for its anti-aging properties and potential therapeutic effects on the extracellular matrix (ECM). This peptide stimulates the proliferation and migration of keratinocytes and the production of ECM proteins like collagen and elastin. In a study involving 94 healthy female faces treated with an emulsion containing pep_RTE62G for 28 days, parameters such as average roughness, average relief, and maximum relief amplitude of cutaneous relief were examined using three-dimensional microtopography. No statistically significant differences were found compared to the placebo group. Quantitative analysis revealed a 26 % increase in elastin expression and a 12 % increase in collagen expression in skin explants treated with pep_RTE62G [139].

AI techniques in peptide design, such as ANN and genetic algorithms, were born out of molecular dynamics simulations and QSARs, which have since been used to automate peptide design and generate AMP sequences [66,140]. Examples of their peptide drugs include SPR206, which is in the completed Phase I development stage and utilizes SAR-based design. It is effective against carbapenem-resistant *Acinetobacter baumannii* (CRAB), carbapenem-resistant *Pseudomonas aeruginosa* (CRPA), and carbapenem-resistant *Klebsiella pneumoniae* (CRKP) [141]. Several AMPs and synthetic antimicrobial peptidomimetics SMAMPs have been tested successfully in clinical trials. Notable examples include ethylamine brilliantin (PMX-30063) and peptidomimetic lytixar (LTX-109). Other synthesized compounds, such as di-cationic exeporfinium chloride (XF-73) and arylamide brilacidin (NP-213), have also undergone Phase II trials [142]As the challenges related to balancing activity and toxicity are addressed, AMPs and SMAMPs are expected to be used systemically in addition to their current topical applications.

8. Challenges in AI peptide research

Ensuring the performance of AI-designed peptides under various conditions and real-world applications requires rigorous robustness validation. This involves extensive testing to confirm stability, efficacy, and safety, which is crucial for successful utilization. Comprehensive assessments are essential to identify potential risks associated with AI-generated peptides, such as toxicity and immunogenicity [143]. Developing specific regulatory guidelines for AI-generated therapeutics is crucial, as current regulations may not fully address the unique aspects of these molecules. Updated standards and protocols will facilitate their approval and integration into clinical practice. Translating AI advances into clinical practice involves ensuring the reproducibility and generalizability of AI models, which requires collaboration between AI developers, clinicians, and regulatory bodies [144]. AI significantly expedites the drug discovery process, reduces costs, and improves the precision of peptide design. AI-designed peptides can effectively target diseases, offering new treatment options and personalized therapies [2]. However, several challenges must be addressed to harness AI's full potential in advancing peptide therapeutics.

Key challenges include the lack of interpretability and reproducibility in ML-based peptide predictions [33]. AI models can introduce bias if not properly managed, and high-quality data is essential for accurate predictions. Privacy and security concerns arise from processing personal health data, necessitating explicit consent and robust protection measures. Integrating AI models into clinical workflows can be challenging, and ethical considerations must be addressed [145,146].

Challenges also arise in implementing ML models for potential therapeutic peptides due to unclear data and irreproducible results. Access restrictions to datasets, ML models, and model codes from established prediction models hinder researchers from developing new experimental models [147]. Predicting accurate binding affinity during drug interactions using DL models poses challenges that

can be addressed by incorporating domain knowledge and utilizing high-quality datasets. While DL excels in areas like de novo protein structure prediction and chemical reaction modeling, its practical applications in drug discovery and medicinal chemistry are still limited. Demonstrating measurable impacts of DL will require time and greater acceptance in interdisciplinary settings [148].

Standardized guidelines and best practices in applying AI to therapeutic peptide development are needed to improve the reproducibility and comparability of studies. Selecting appropriate molecular descriptors for peptide features is challenging due to a limited understanding of the underlying biological or molecular mechanisms. Using descriptors that describe ligand-protein interactions could enhance the molecular understanding of predictions. Additionally, ML models often face the risk of overfitting, which can be mitigated by using more relevant features and validating them with external data. The use of deep learning techniques and combined classifiers shows great promise in rational drug discovery, particularly in predicting drug-target interactions and drug-induced toxicities [149].

9. Future directions in therapeutic peptides and artificial intelligence

Developing novel computational approaches with superior performance and usability, alongside continuous innovation and improvement of existing ML methods for peptide predictions, is essential for future advancements. Recent advancements in deep learning hold significant potential to reduce costs and increase efficiency, fostering collaborations between pharmaceutical companies and AI businesses [150]. DL has emerged as a powerful tool in peptide research, especially in evaluating protein and DNA sequences, with high-throughput techniques positioning DL as a standard option for large-scale data processing and analysis [55].

Designing therapeutic peptides requires collaboration among experts in peptide chemistry, artificial intelligence, and drug discovery. The interpretability and transparency of AI models are crucial for ensuring the safety and efficacy of proposed treatments. Scientists and regulators must understand how AI influences drug design and optimization decisions. As AI becomes a standard tool in pharmaceutical research, it is essential to develop regulatory and ethical frameworks to ensure its appropriate and safe use. These guidelines must support innovation without compromising patient safety, necessitating collaboration between researchers, industry, and regulatory bodies [151].

AI in therapeutic development shows great promise, but ethical and practical considerations are necessary. Regulatory bodies, academia, and industry can work together to create collaborative ecosystems that facilitate faster development of innovative therapies. When scientific collaboration is encouraged, data is shared, and best practices are adopted, the potential for accelerated breakthroughs increases. AI and ML can streamline drug development and clinical trials, optimize dosage forms, and enhance decision-making, resulting in higher-quality products [152].

In the United States, the Food and Drug Administration (FDA) has approved AI-based applications for healthcare and implemented the Digital Healthcare Innovation Action Plan, which includes the Software Precertification (PreCert) Pilot Program for AI-based digital health projects [153]. However, the regulatory landscape for healthcare AI is still evolving. In Europe, the European Commission has established a High-Level Expert Group to develop AI ethics guidelines, prioritizing data protection, digital rights, and ethical standards in AI and robotics [154].

The intersection of therapeutic peptides and AI represents a transformative frontier in drug discovery. By leveraging AI's predictive capabilities, researchers can significantly enhance the identification of novel therapeutic peptides through advanced algorithms that analyze vast biological datasets, uncovering patterns that may elude traditional methods [155]. Additionally, AI can deepen our understanding of peptide mechanisms of action by modeling molecular interactions, enabling predictions of how peptides interact with specific biological targets.

Moreover, AI can optimize the delivery and efficacy of therapeutic peptides. Machine learning algorithms can design peptide formulations that improve stability and bioavailability while minimizing potential side effects [156]. Researchers can rapidly iterate on peptide designs by integrating AI-driven simulations with high-throughput screening techniques, refining their properties for specific therapeutic applications.

Ultimately, fostering collaboration between computational scientists and peptide chemists will be crucial in realizing the full potential of AI in therapeutic peptide development. This interdisciplinary approach can lead to innovative therapies that are safer, more effective, and tailored to patient needs, marking a significant advancement in personalized medicine.

10. Limitations of the review

The primary limitations of this review include:

Scope and Breadth: While we aimed to cover a comprehensive range of AI applications in peptide drug discovery, the rapidly evolving nature of the field means that some recent advancements may not be included.

Data availability: The review relies on published studies and available data, which may introduce a publication bias. Unpublished or proprietary data from industry sources are not considered, potentially limiting the completeness of the review.

Heterogeneity of Studies: The methodologies, datasets, and evaluation metrics used across different studies vary significantly, making direct comparisons challenging and potentially affecting the consistency of our conclusions.

Technical Complexity: The implementation of AI techniques can be highly specialized, and detailed technical nuances may not be fully explored in this review, potentially oversimplifying complex methodologies.

11. Conclusion

In conclusion, the significance of peptide medicines not only in their specificity reaction and high ability to eliminate a wide range

of pathogens but also in effectively treating cancer is growing fast and has contributed to the drug development process. Enhance therapeutic peptides using AI-driven techniques, reducing the costs associated with research and development (R&D) and improving quality control in all industrial fields. Companies are forced to employ AI-based techniques to accelerate the analysis of large and complex data ranging from natural to synthetic compounds to the classification of active pharmaceutical ingredients in drug manufacturing. The value of ML and DL methods in prediction and operations will increase the success rate of multitarget treatment; thus, researchers have greatly benefited from databases based on AI algorithms that predict the performance of different peptides to find solutions in cancer therapy and antibiotic resistance as a major global concern. It is important to remember that using AI requires full compliance with various ethical principles.

CRediT authorship contribution statement

Samaneh Hashemi: Writing – review & editing, Writing – original draft, Methodology, Investigation. **Parisa Vosough:** Writing – review & editing, Writing – original draft, Methodology, Investigation. **Saeed Taghizadeh:** Writing – review & editing, Writing – original draft, Project administration, Methodology, Investigation. **Amir Savardashtaki:** Writing – review & editing, Writing – original draft, Supervision, Project administration, Methodology, Investigation.

Ethics approval and consent to participate

Not applicable. This article contains no studies performed by authors with human participants or animals. It is a comprehensive review, synthesizing insights from previously published articles.

Consent for publication

This consent is not applicable as this manuscript does not contain any person's data in any form (including personal details, images, or videos).

Data availability

The data used to support the findings of this study are included in the article.

Use of AI and AI-assisted technologies

AI technologies were not employed when creating this manuscript. The content is derived from the authors' interpretations of existing literature and informed discussions based on their expertise.

Funding

Not applicable. This research did not receive a specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

[1] S. D'Souza, K. Prema, S. Balaji, Machine learning models for drug–target interactions: current knowledge and future directions, Drug Discov. Today 25 (2020) 748–756.
[2] L.K. Vora, A.D. Gholap, K. Jetha, R.R.S. Thakur, H.K. Solanki, V.P. Chavda, Artificial intelligence in pharmaceutical technology and drug delivery design, Pharmaceutics 15 (2023).
[3] L. Wang, N. Wang, W. Zhang, X. Cheng, Z. Yan, G. Shao, X. Wang, R. Wang, C. Fu, Therapeutic peptides: current applications and future directions, Signal Transduct. Targeted Ther. 7 (2022) 48.
[4] N. Tsomaia, Peptide therapeutics: targeting the undruggable space, Eur. J. Med. Chem. 94 (2015) 459–470.
[5] L. Wang, N. Wang, W. Zhang, X. Cheng, Z. Yan, G. Shao, X. Wang, R. Wang, C. Fu, Therapeutic peptides: current applications and future directions, Signal Transduct. Targeted Ther. 7 (2022) 48.
[6] G. Rossino, E. Marchese, G. Galli, F. Verde, M. Finizio, M. Serra, P. Linciano, S. Collina, Peptides as therapeutic agents: challenges and opportunities in the green transition era, Molecules 28 (2023) 7165.
[7] P. Barman, S. Joshi, S. Sharma, S. Preet, S. Sharma, A. Saini, Strategic approaches to improvise peptide drugs as next generation therapeutics, Int. J. Pept. Res. Therapeut. 29 (2023) 61.
[8] Y. Zhang, H. Zhang, D. Ghosh, R.O. Williams III, Just how prevalent are peptide therapeutic products? A critical review, Int. J. Pharm. 587 (2020) 119491.
[9] G. Rossino, E. Marchese, G. Galli, F. Verde, M. Finizio, M. Serra, P. Linciano, S. Collina, Peptides as therapeutic agents: challenges and opportunities in the green transition era, Molecules 28 (2023).
[10] M.S.R. Abid, S. Mousavi, J.W. Checco, Identifying receptors for neuropeptides and peptide hormones: challenges and recent progress, ACS Chem. Biol. 16 (2021) 251–263.

[11] C. Adessi, C. Soto, Converting a peptide into a drug: strategies to improve stability and bioavailability, Curr. Med. Chem. 9 (2002) 963–978.
[12] T.D.D. Kazmirchuk, C. Bradbury-Jost, T.A. Withey, T. Gessese, T. Azad, B. Samanfar, F. Dehne, A. Golshani, Peptides of a feather: how computation is taking peptide therapeutics under its wing, Genes 14 (2023) 1194.
[13] J.M. Mata, E. van der Nol, S.J. Pomplun, Advances in ultrahigh throughput hit discovery with tandem mass spectrometry encoded libraries, J. Am. Chem. Soc. 145 (2023) 19129–19139.
[14] C. Hasselgren, T.I. Oprea, Artificial intelligence for drug discovery: are we there yet? Annu. Rev. Pharmacol. Toxicol. 64 (2024) 527–550.
[15] M. Craig, J. Caicedo, P. Das, J. Collins, F. Grisoni, C. de la Fuente-Núñez, Y.-S. Lin, J. Tang, AI and drug discovery, Cell Reports Physical Science 3 (2022) 101142.
[16] J. Xu, F. Li, C. Li, X. Guo, C. Landersdorfer, H.H. Shen, A.Y. Peleg, J. Li, S. Imoto, J. Yao, T. Akutsu, J. Song, iAMPCN: a deep-learning approach for identifying antimicrobial peptides and their functional activities, Briefings Bioinf. 24 (2023).
[17] X. Zhang, C. de la Fuente-Nunez, J. Wang, Artificial intelligence accelerates efficient mining of functional peptides, Life Medicine 2 (2023) lnad005.
[18] S. Bechelli, J. Delhommelle, AI's role in pharmaceuticals: assisting drug design from protein interactions to drug development, Artificial Intelligence Chemistry (2023) 100038.
[19] M.B.M.A. Rashid, Artificial intelligence effecting a paradigm shift in drug development, SLAS Technology 26 (2021) 3–15.
[20] S.K. Niazi, Z. Mariam, Computer-aided drug design and drug discovery: a prospective analysis, Pharmaceuticals 17 (2023) 22.
[21] D. Crucitti, C. Pérez Míguez, J.Á. Díaz Arias, D.B. Fernandez Prada, A. Mosquera Orgueira, De novo drug design through artificial intelligence: an introduction, Frontiers in Hematology 3 (2024) 1305741.
[22] N. Norori, Q. Hu, F.M. Aellen, F.D. Faraci, A. Tzovara, Addressing bias in big data and AI for health care: a call for open science, Patterns (N Y) 2 (2021) 100347.
[23] E. Pirogova, T. Istivan, E. Gan, I. Cosic, Advances in methods for therapeutic peptide discovery, design and development, Curr. Pharmaceut. Biotechnol. 12 (2011) 1117–1127.
[24] P. Szymański, M. Markowicz, E. Mikiciuk-Olasik, Adaptation of high-throughput screening in drug discovery-toxicological screening tests, Int. J. Mol. Sci. 13 (2012) 427–452.
[25] A. Isidro-Llobet, M.N. Kenworthy, S. Mukherjee, M.E. Kopach, K. Wegner, F. Gallou, A.G. Smith, F. Roschangar, Sustainability challenges in peptide synthesis and purification: from R&D to production, J. Org. Chem. 84 (2019) 4615–4628.
[26] A.A. Robles-Loaiza, E.A. Pinos-Tamayo, B. Mendes, J.A. Ortega-Pila, C. Proaño-Bolaños, F. Plisson, C. Teixeira, P. Gomes, J.R. Almeida, Traditional and computational screening of non-toxic peptides and approaches to improving selectivity, Pharmaceuticals 15 (2022) 323.
[27] H. Aldewachi, R.N. Al-Zidan, M.T. Conner, M.M. Salman, High-throughput screening platforms in the discovery of novel drugs for neurodegenerative diseases, Bioengineering 8 (2021) 30.
[28] T.A.d. Oliveira, M.P.d. Silva, E.H.B. Maia, A.M.d. Silva, A.G. Taranto, Virtual screening algorithms in drug discovery: a review focused on machine and deep learning methods, Drugs and Drug Candidates 2 (2023) 311–334.
[29] C. Sarkar, B. Das, V.S. Rawat, J.B. Wahlang, A. Nongpiur, I. Tiewsoh, N.M. Lyngdoh, D. Das, M. Bidarolli, H.T. Sony, Artificial intelligence and machine learning technology driven modern drug discovery and development, Int. J. Mol. Sci. 24 (2023) 2026.
[30] Q. Wu, H. Ke, D. Li, Q. Wang, J. Fang, J. Zhou, Recent progress in machine learning-based prediction of peptide activity for drug discovery, Curr. Top. Med. Chem. 19 (2019) 4–16.
[31] S. Romero-Molina, Y.B. Ruiz-Blanco, J. Mieres-Perez, M. Harms, J. Münch, M. Ehrmann, E. Sanchez-Garcia, PPI-affinity: a web tool for the prediction and optimization of protein–peptide and protein–protein binding affinity, J. Proteome Res. 21 (2022) 1829–1841.
[32] S. Giguère, F. Laviolette, M. Marchand, D. Tremblay, S. Moineau, X. Liang, É. Biron, J. Corbeil, Machine learning assisted design of highly active peptides for drug discovery, PLoS Comput. Biol. 11 (2015) e1004074.
[33] S. Basith, B. Manavalan, T. Hwan Shin, G. Lee, Machine intelligence in peptide therapeutics: a next-generation tool for rapid disease screening, Med. Res. Rev. 40 (2020) 1276–1314.
[34] M. Yoshida, T. Hinkley, S. Tsuda, Y.M. Abul-Haija, R.T. McBurney, V. Kulikov, J.S. Mathieson, S.G. Reyes, M.D. Castro, L. Cronin, Using evolutionary algorithms and machine learning to explore sequence space for the discovery of antimicrobial peptides, Chem 4 (2018) 533–543.
[35] S. Mohapatra, N. Hartrampf, M. Poskus, A. Loas, R. Gomez-Bombarelli, B.L. Pentelute, Deep learning for prediction and optimization of fast-flow peptide synthesis, ACS Cent. Sci. 6 (2020) 2277–2286.
[36] D. Paul, G. Sanap, S. Shenoy, D. Kalyane, K. Kalia, R.K. Tekade, Artificial intelligence in drug discovery and development, Drug Discov. Today 26 (2021) 80.
[37] A.T. Müller, G. Gabernet, J.A. Hiss, G. Schneider, modlAMP: Python for antimicrobial peptides, Bioinformatics 33 (2017) 2753–2755.
[38] L.D. Saravolatz, J. Pawlak, L. Johnson, H. Bonilla, L.D. Saravolatz 2nd, M.G. Fakih, A. Fugelli, W.M. Olsen, In vitro activities of LTX-109, a synthetic antimicrobial peptide, against methicillin-resistant, vancomycin-intermediate, vancomycin-resistant, daptomycin-nonsusceptible, and linezolid-nonsusceptible Staphylococcus aureus, Antimicrob. Agents Chemother. 56 (2012) 4478–4482.
[39] G.S. Dijksteel, M.M.W. Ulrich, E. Middelkoop, B. Boekema, Review: lessons learned from clinical trials using antimicrobial peptides (AMPs), Front. Microbiol. 12 (2021) 616979.
[40] L. Perpetuo, J. Klein, R. Ferreira, S. Guedes, F. Amado, A. Leite-Moreira, A.M.S. Silva, V. Thongboonkerd, R. Vitorino, How can artificial intelligence be used for peptidomics? Expert Rev. Proteomics 18 (2021) 527–556.
[41] L. Chang, A. Mondal, B. Singh, Y. Martínez-Noa, A. Perez, Revolutionizing peptide-based drug discovery: advances in the post-AlphaFold era, Wiley Interdiscip. Rev. Comput. Mol. Sci. 14 (2024) e1693.
[42] C. Fang, Y. Moriwaki, C. Li, K. Shimizu, Prediction of antifungal peptides by deep learning with character embedding, IPSJ Transactions on Bioinformatics 12 (2019) 21–29.
[43] D. Agyei, A. Tsopmo, C.C. Udenigwe, Bioinformatics and peptidomics approaches to the discovery and analysis of food-derived bioactive peptides, Anal. Bioanal. Chem. 410 (2018) 3463–3472.
[44] L. Nanni, A. Lumini, A new encoding technique for peptide classification, Expert Syst. Appl. 38 (2011) 3185–3191.
[45] F. Cui, Z. Zhang, Q. Zou, Sequence representation approaches for sequence-based protein prediction tasks that use deep learning, Briefings in Functional Genomics 20 (2021) 61–73.
[46] P. Agrawal, H. Singh, H.K. Srivastava, S. Singh, G. Kishore, G.P. Raghava, Benchmarking of different molecular docking methods for protein-peptide docking, BMC Bioinf. 19 (2019) 105–124.
[47] M. Attique, M.S. Farooq, A. Khelifi, A. Abid, Prediction of therapeutic peptides using machine learning: computational models, datasets, and feature encodings, IEEE Access 8 (2020) 148570–148594.
[48] F. Khan, S. Akbar, A. Basit, I. Khan, H. Akhlaq, Identification of anticancer peptides using optimal feature space of Chou's split amino acid composition and support vector machine. Proceedings of the 2017 4th International Conference on Biomedical and Bioinformatics Engineering, 2017, pp. 91–96.
[49] A.B. Hassanat, Two-point-based binary search trees for accelerating big data classification using KNN, PLoS One 13 (2018) e0207772.
[50] H. Feng, F. Wang, N. Li, Q. Xu, G. Zheng, X. Sun, M. Hu, G. Xing, G. Zhang, A random forest model for peptide classification based on virtual docking data, Int. J. Mol. Sci. 24 (2023) 11409.
[51] M.D. Aguilera-Puga, F. Plisson, Structure-aware machine learning strategies for antimicrobial peptide discovery, Sci. Rep. 14 (2024) 11995.
[52] J. Ye, A. Li, H. Zheng, B. Yang, Y. Lu, Machine learning advances in predicting peptide/protein-protein interactions based on sequence information for lead peptides discovery, Advanced Biology 7 (2023) 2200232.
[53] J.-R. Codina, M. Mascini, E. Dikici, S.K. Deo, S. Daunert, Accelerating the screening of small peptide ligands by combining peptide-protein docking and machine learning, Int. J. Mol. Sci. 24 (2023) 12144.
[54] K. Boone, N. Tjokro, K.N. Chu, C. Chen, M.L. Snead, C. Tamerler, Machine learning enabled design features of antimicrobial peptides selectively targeting peri-implant disease progression, Frontiers in dental medicine 5 (2024).

[55] X. Chen, C. Li, M.T. Bernards, Y. Shi, Q. Shao, Y. He, Sequence-based peptide identification, generation, and property prediction with deep learning: a review, Molecular Systems Design & Engineering 6 (2021) 406–428.
[56] H. Zhang, K.M. Saravanan, Y. Wei, Y. Jiao, Y. Yang, Y. Pan, X. Wu, J.Z. Zhang, Deep learning-based bioactive therapeutic peptide generation and screening, J. Chem. Inf. Model. 63 (2023) 835–845.
[57] C. Li, D. Sutherland, S.A. Hammond, C. Yang, F. Taho, L. Bergman, S. Houston, R.L. Warren, T. Wong, L.M. Hoang, AMPlify: attentive deep learning model for discovery of novel antimicrobial peptides effective against WHO priority pathogens, BMC Genom. 23 (2022) 77.
[58] Q. Cao, C. Ge, X. Wang, P.J. Harvey, Z. Zhang, Y. Ma, X. Wang, X. Jia, M. Mobli, D.J. Craik, Designing antimicrobial peptides using deep learning and molecular dynamic simulations, Briefings Bioinf. 24 (2023) bbad058.
[59] A. Pandi, D. Adam, A. Zare, V.T. Trinh, S.L. Schaefer, M. Burt, B. Klabunde, E. Bobkova, M. Kushwaha, Y. Foroughijabbari, Cell-free biosynthesis combined with deep learning accelerates de novo-development of antimicrobial peptides, Nat. Commun. 14 (2023) 7197.
[60] A. Capecchi, X. Cai, H. Personne, T. Köhler, C. van Delden, J.-L. Reymond, Machine learning designs non-hemolytic antimicrobial peptides, Chem. Sci. 12 (2021) 9221–9232.
[61] Y. LeCun, Y. Bengio, G. Hinton, Deep learning, Nature 521 (2015) 436–444.
[62] D. Veltri, U. Kamath, A. Shehu, Deep learning improves antimicrobial peptide recognition, Bioinformatics 34 (2018) 2740–2747.
[63] Y. Yu, X. Si, C. Hu, J. Zhang, A review of recurrent neural networks: LSTM cells and network architectures, Neural Comput. 31 (2019) 1235–1270.
[64] H. Fan, W. Yan, L. Wang, J. Liu, Y. Bin, J. Xia, Deep learning-based multi-functional therapeutic peptides prediction with a multi-label focal dice loss function, Bioinformatics 39 (2023) btad334.
[65] Z. Du, X. Ding, Y. Xu, Y. Li, UniDL4BioPep: a universal deep learning architecture for binary classification in peptide bioactivity, Briefings Bioinf. 24 (2023) bbad135.
[66] M.H. Cardoso, R.Q. Orozco, S.B. Rezende, G. Rodrigues, K.G. Oshiro, E.S. Cândido, O.L. Franco, Computer-aided design of antimicrobial peptides: are we generating effective drug candidates? Front. Microbiol. 10 (2020) 3097.
[67] J.-L. Faulon, L. Faure, In silico, in vitro, and in vivo machine learning in synthetic biology and metabolic engineering, Curr. Opin. Chem. Biol. 65 (2021) 85–92.
[68] F. Wan, D. Kontogiorgos-Heintz, C. de la Fuente-Nunez, Deep generative models for peptide design, Digital Discovery 1 (2022) 195–208.
[69] M. Ansari, A.D. White, Serverless prediction of peptide properties with recurrent neural networks, J. Chem. Inf. Model. 63 (2023) 2546–2553.
[70] F. Wang, N. Sangfuang, L.E. McCoubrey, V. Yadav, M. Elbadawi, M. Orlu, S. Gaisford, A.W. Basit, Advancing oral delivery of biologics: machine learning predicts peptide stability in the gastrointestinal tract, Int. J. Pharm. 634 (2023) 122643.
[71] M.P. Aranha, C. Spooner, O. Demerdash, B. Czejdo, J.C. Smith, J.C. Mitchell, Prediction of peptide binding to MHC using machine learning with sequence and structure-based feature sets, Biochim. Biophys. Acta Gen. Subj. 1864 (2020) 129535.
[72] M.P. Aranha, Y.S. Jewel, R.A. Beckman, L.M. Weiner, J.C. Mitchell, J.M. Parks, J.C. Smith, Combining three-dimensional modeling with artificial intelligence to increase specificity and precision in peptide–MHC binding predictions, J. Immunol. 205 (2020) 1962–1977.
[73] K.M. Boehm, B. Bhinder, V.J. Raja, N. Dephoure, O. Elemento, Predicting peptide presentation by major histocompatibility complex class I: an improved machine learning approach to the immunopeptidome, BMC Bioinf. 20 (2019) 1–11.
[74] F. Plisson, O. Ramírez-Sánchez, C. Martínez-Hernández, Machine learning-guided discovery and design of non-hemolytic peptides, Sci. Rep. 10 (2020) 16581.
[75] C.M. Van Oort, J.B. Ferrell, J.M. Remington, S. Wshah, J. Li, AMPGAN v2: machine learning-guided design of antimicrobial peptides, J. Chem. Inf. Model. 61 (2021) 2198–2207.
[76] E.C.L. de Oliveira, K. Santana, L. Josino, A.H. Lima e Lima, C. de Souza de Sales Júnior, Predicting cell-penetrating peptides using machine learning algorithms and navigating in their chemical space, Sci. Rep. 11 (2021) 7628.
[77] C. Quiroz, Y.B. Saavedra, B. Armijo-Galdames, J. Amado-Hinojosa, Á. Olivera-Nappa, A. Sanchez-Daza, D. Medina-Ortiz, Peptipedia: a user-friendly web application and a comprehensive database for peptide research supported by machine learning approach, Database 2021 (2021) baab055.
[78] Q. Chen, C. Yang, Y. Xie, Y. Wang, X. Li, K. Wang, J. Huang, W. Yan, GM-Pep: A high efficiency strategy to de novo design functional peptide sequences, J. Chem. Inf. Model. 62 (2022) 2617–2629.
[79] P.G. Aronica, L.M. Reid, N. Desai, J. Li, S.J. Fox, S. Yadahalli, J.W. Essex, C.S. Verma, Computational methods and tools in antimicrobial peptide research, J. Chem. Inf. Model. 61 (2021) 3172–3196.
[80] L. Sleno, A. Emili, Proteomic methods for drug target discovery, Curr. Opin. Chem. Biol. 12 (2008) 46–54.
[81] F.W. Pun, I.V. Ozerov, A. Zhavoronkov, AI-powered therapeutic target discovery, Trends Pharmacol. Sci. (2023).
[82] A.L. Oliveira, Biotechnology, big data and artificial intelligence, Biotechnol. J. 14 (2019) 1800613.
[83] B.C. Searle, J.D. Egertson, J.G. Bollinger, A.B. Stergachis, M.J. MacCoss, Using data independent acquisition (DIA) to model high-responding peptides for targeted proteomics experiments, Mol. Cell. Proteomics 14 (2015) 2331–2340.
[84] Z. Ahmed, S. Wan, F. Zhang, W. Zhong, Artificial intelligence for omics data analysis, BMC Methods 1 (2024) 4.
[85] M. Pammi, N. Aghaeepour, J. Neu, Multiomics, artificial intelligence, and precision medicine in perinatology, Pediatr. Res. 93 (2023) 308–315.
[86] A.C. Pushkaran, A.A. Arabi, From understanding diseases to drug design: can artificial intelligence bridge the gap? Artif. Intell. Rev. 57 (2024) 86.
[87] T. Ivanisevic, R.N. Sewduth, Multi-omics integration for the design of novel therapies and the identification of novel biomarkers, Proteomes 11 (2023) 34.
[88] R. Casey, A. Adelfio, M. Connolly, A. Wall, I. Holyer, N. Khaldi, Discovery through machine learning and preclinical validation of novel anti-diabetic peptides, Biomedicines 9 (2021) 276.
[89] A. Rauf, A. Kiran, M.T. Hassan, S. Mahmood, G. Mustafa, M. Jeon, Boosted prediction of antihypertensive peptides using deep learning, Appl. Sci. 11 (2021) 2316.
[90] S. Lertampaiporn, A. Hongsthong, W. Wattanapornprom, C. Thammarongtham, Ensemble-AHTPpred: a robust ensemble machine learning model integrated with a new composite feature for identifying antihypertensive peptides, Front. Genet. 13 (2022) 883766.
[91] W. Zhang, E. Xia, R. Dai, W. Tang, Y. Bin, J. Xia, PredAPP: predicting anti-parasitic peptides with undersampling and ensemble approaches, Interdiscipl. Sci. Comput. Life Sci. (2022) 1–11.
[92] R. Dai, W. Zhang, W. Tang, E. Wynendaele, Q. Zhu, Y. Bin, B. De Spiegeleer, J. Xia, BBPpred: sequence-based prediction of blood-brain barrier peptides with feature representation learning and logistic regression, J. Chem. Inf. Model. 61 (2021) 525–534.
[93] J. Yan, P. Bhadra, A. Li, P. Sethiya, L. Qin, H.K. Tai, K.H. Wong, S.W. Siu, Deep-AmPEP30: improve short antimicrobial peptides prediction with deep learning, Mol. Ther. Nucleic Acids 20 (2020) 882–894.
[94] Q. Yu, Z. Dong, X. Fan, L. Zong, Y. Li, HMD-AMP: protein language-powered hierarchical multi-label deep forest for annotating antimicrobial peptides, arXiv preprint arXiv:2111.06023 (2021).
[95] W. Dee, LMPred: predicting antimicrobial peptides using pre-trained language models and deep learning, Bioinformatics Advances 2 (2022) vbac021.
[96] P. Pandey, A. Srivastava, sAMP-VGG16: drude polarizable force-field assisted image-based deep neural network prediction model for short antimicrobial peptides, bioRxiv (2023), 2023.2006. 2004.543607.
[97] H. Lee, S. Lee, I. Lee, H. Nam, AMP-BERT: prediction of antimicrobial peptide function based on a BERT model, Protein Sci. 32 (2023) e4529.
[98] Y. Wang, L. Wang, C. Li, Y. Pei, X. Liu, Y. Tian, AMP-EBiLSTM: employing novel deep learning strategies for the accurate prediction of antimicrobial peptides, Front. Genet. 14 (2023).
[99] J. Chen, H.H. Cheong, S.W. Siu, xDeep-AcPEP: deep learning method for anticancer peptide activity prediction based on convolutional neural network and multitask learning, J. Chem. Inf. Model. 61 (2021) 3789–3803.
[100] W. He, Y. Wang, L. Cui, R. Su, L. Wei, Learning embedding features based on multisense-scaled attention architecture to improve the predictive performance of anticancer peptides, Bioinformatics 37 (2021) 4684–4693.
[101] X.-g. Chen, W. Zhang, X. Yang, C. Li, H. Chen, Acp-da: improving the prediction of anticancer peptides using data augmentation, Front. Genet. 12 (2021) 698477.
[102] X. Wu, W. Zeng, F. Lin, P. Xu, X. Li, Anticancer peptide prediction via multi-kernel CNN and attention model, Front. Genet. 13 (2022) 887894.

[103] C. Zhou, D. Peng, B. Liao, R. Jia, F. Wu, ACP_MS: prediction of anticancer peptides based on feature extraction, Briefings Bioinf. 23 (2022) bbac462.
[104] X. Yang, J. Jin, R. Wang, Z. Li, Y. Wang, L. Wei, CACPP: a contrastive learning-based siamese network to identify anticancer peptides based on sequence only, J. Chem. Inf. Model. (2023).
[105] Y. Pang, L. Yao, J.-H. Jhong, Z. Wang, T.-Y. Lee, AVPIden: a new scheme for identification and functional prediction of antiviral peptides based on machine learning approaches, Briefings Bioinf. 22 (2021) bbab263.
[106] A.S. Chowdhury, S.M. Reehl, K. Kehn-Hall, B. Bishop, B.-J.M. Webb-Robertson, Better understanding and prediction of antiviral peptides through primary and secondary structure feature importance, Sci. Rep. 10 (2020) 19260.
[107] T.-T. Lin, Y.-Y. Sun, C.-T. Wang, W.-C. Cheng, I.-H. Lu, C.-Y. Lin, S.-H. Chen, AI4AVP: an antiviral peptides predictor in deep learning approach with generative adversarial network data augmentation, Bioinformatics Advances 2 (2022) vbac080.
[108] D. Qin, L. Jiao, R. Wang, Y. Zhao, Y. Hao, G. Liang, Prediction of antioxidant peptides using a quantitative structure– activity relationship predictor (AnOxPP) based on bidirectional long short-term memory neural network and interpretable amino acid descriptors, Comput. Biol. Med. 154 (2023) 106591.
[109] T.H. Olsen, B. Yesiltas, F.I. Marin, M. Pertseva, P.J. García-Moreno, S. Gregersen, M.T. Overgaard, C. Jacobsen, O. Lund, E.B. Hansen, AnOxPePred: using deep learning for the prediction of antioxidative properties of peptides, Sci. Rep. 10 (2020) 21471.
[110] H. Deng, C. Lou, Z. Wu, W. Li, G. Liu, Y. Tang, Prediction of anti-inflammatory peptides by a sequence-based stacking ensemble model named AIPStack, iScience 25 (2022).
[111] S. Gaffar, M.T. Hassan, H. Tayara, K.T. Chong, IF-AIP: a machine learning method for the identification of anti-inflammatory peptides using multi-feature fusion strategy, Comput. Biol. Med. 168 (2024) 107724.
[112] A. Razaa, A. Ahmada, Z. Iqbala, Q. Yasina, H. Javeda, A. Shaha, S. Chaudharya, iAFP-ET: A Robust Approach for Accurate Identification of Antifungal Peptides Using Extra Tree Classifier and Multi-View Fusion.
[113] L. Yao, Y. Zhang, W. Li, C.R. Chung, J. Guan, W. Zhang, Y.C. Chiang, T.Y. Lee, D. eepAFP, An effective computational framework for identifying antifungal peptides based on deep learning, Protein Sci. 32 (2023) e4758.
[114] F. Prestinaci, P. Pezzotti, A. Pantosti, Antimicrobial resistance: a global multifaceted phenomenon, Pathog. Glob. Health 109 (2015) 309–318.
[115] T.-T. Lin, L.-Y. Yang, C.-Y. Lin, C.-T. Wang, C.-W. Lai, C.-F. Ko, Y.-H. Shih, S.-H. Chen, Intelligent de novo design of novel antimicrobial peptides against antibiotic-resistant bacteria strains, Int. J. Mol. Sci. 24 (2023) 6788.
[116] K. Boone, C. Wisdom, K. Camarda, P. Spencer, C. Tamerler, Combining genetic algorithm with machine learning strategies for designing potent antimicrobial peptides, BMC Bioinf. 22 (2021) 239.
[117] P. Ruiz Puentes, M.C. Henao, J. Cifuentes, C. Muñoz-Camargo, L.H. Reyes, J.C. Cruz, P. Arbeláez, Rational discovery of antimicrobial peptides by means of artificial intelligence, Membranes 12 (2022) 708.
[118] Y. Zhang, P. Li, Y. Ma, J. Wang, Y. Chen, Artificial intelligence accelerates the mining of bioactive small molecules from human microbiome, Clin. Transl. Med. 12 (2022).
[119] F.C. Fernandes, M.H. Cardoso, A. Gil-Ley, L.V. Luchi, M.G. da Silva, M.L. Macedo, C. de la Fuente-Nunez, O.L. Franco, Geometric deep learning as a potential tool for antimicrobial peptide prediction, Frontiers in Bioinformatics 3 (2023).
[120] J. Huang, Y. Xu, Y. Xue, Y. Huang, X. Li, X. Chen, Y. Xu, D. Zhang, P. Zhang, J. Zhao, Identification of potent antimicrobial peptides via a machine-learning pipeline that mines the entire space of peptide sequences, Nat. Biomed. Eng. (2023) 1–14.
[121] Y. Ma, Z. Guo, B. Xia, Y. Zhang, X. Liu, Y. Yu, N. Tang, X. Tong, M. Wang, X. Ye, Identification of antimicrobial peptides from the human gut microbiome using deep learning, Nat. Biotechnol. 40 (2022) 921–931.
[122] P. Szymczak, E. Szczurek, Artificial intelligence-driven antimicrobial peptide discovery, Curr. Opin. Struct. Biol. 83 (2023) 102733.
[123] D.P. Tran, S. Tada, A. Yumoto, A. Kitao, Y. Ito, T. Uzawa, K. Tsuda, Using molecular dynamics simulations to prioritize and understand AI-generated cell penetrating peptides, Sci. Rep. 11 (2021) 10630.
[124] E.M. López-Vidal, C.K. Schissel, S. Mohapatra, K. Bellovoda, C.-L. Wu, J.A. Wood, A.B. Malmberg, A. Loas, R. Gómez-Bombarelli, B.L. Pentelute, Deep learning enables discovery of a short nuclear targeting peptide for efficient delivery of antisense oligomers, JACS Au 1 (2021) 2009–2020.
[125] J.S. Hwang, S.G. Kim, T.H. Shin, Y.E. Jang, D.H. Kwon, G. Lee, Development of anticancer peptides using artificial intelligence and combinational therapy for cancer therapeutics, Pharmaceutics 14 (2022) 997.
[126] G. Li, B. Iyer, V.S. Prasath, Y. Ni, N. Salomonis, DeepImmuno: deep learning-empowered prediction and generation of immunogenic peptides for T-cell immunity, Briefings Bioinf. 22 (2021) bbab160.
[127] F. Grisoni, C.S. Neuhaus, M. Hishinuma, G. Gabernet, J.A. Hiss, M. Kotera, G. Schneider, De novo design of anticancer peptides by ensemble artificial neural networks, J. Mol. Model. 25 (2019) 1–10.
[128] Y. Zhang, Y. Li, Z. Quan, P. Xiao, J.-A. Duan, New insights into antioxidant peptides: an overview of efficient screening, evaluation models, molecular mechanisms, and applications, Antioxidants 13 (2024) 203.
[129] J. Liu, X. Han, T. Zhang, K. Tian, Z. Li, F. Luo, Reactive oxygen species (ROS) scavenging biomaterials for anti-inflammatory diseases: from mechanism to therapy, J. Hematol. Oncol. 16 (2023) 116.
[130] F. Lobo, M.S. González, A. Boto, J.M. Pérez de la Lastra, Prediction of antifungal activity of antimicrobial peptides by transfer learning from protein pretrained models, Int. J. Mol. Sci. 24 (2023) 10270.
[131] A. Blanco-González, A. Cabezón, A. Seco-González, D. Conde-Torres, P. Antelo-Riveiro, Á. Piñeiro, R. Garcia-Fandino, The role of AI in drug discovery: challenges, opportunities, and strategies, Pharmaceuticals 16 (2023).
[132] A. Kate, E. Seth, A. Singh, C.M. Chakole, M.K. Chauhan, R.K. Singh, S. Maddalwar, M. Mishra, Artificial Intelligence for Computer-Aided Drug Discovery, Drug research, 2023.
[133] J. Jiménez-Luna, F. Grisoni, G. Schneider, Drug discovery with explainable artificial intelligence, Nat. Mach. Intell. 2 (2020) 573–584.
[134] S. Ekins, A.C. Puhl, K.M. Zorn, T.R. Lane, D.P. Russo, J.J. Klein, A.J. Hickey, A.M. Clark, Exploiting machine learning for end-to-end drug discovery and development, Nat. Mater. 18 (2019) 435–441.
[135] H. Shi, S. Liu, J. Chen, X. Li, Q. Ma, B. Yu, Predicting drug-target interactions using Lasso with random forest based on evolutionary information and chemical structure, Genomics 111 (2019) 1839–1852.
[136] F. Gentile, V. Agrawal, M. Hsing, A.-T. Ton, F. Ban, U. Norinder, M.E. Gleave, A. Cherkasov, Deep docking: a deep learning platform for augmentation of structure based drug discovery, ACS Cent. Sci. 6 (2020) 939–949.
[137] A.G. Abhishek, A. Saini, K. Nagarajan, P. Bhatt, G. Kapoor, S. Mehta, S. Mishra, S. Shrivastava, A. Ahlawat, A. Shrivastava, Comparative in-silico screening of potent peptide leads using docking strategy & AI approaches for the treatment of liver cancer, J. Pharm. Negat. Results (2023) 3144–3152.
[138] R. Batra, T.D. Loeffler, H. Chan, S. Srinivasan, H. Cui, I.V. Korendovych, V. Nanda, L.C. Palmer, L.A. Solomon, H.C. Fry, Machine learning overcomes human bias in the discovery of self-assembling peptides, Nat. Chem. 14 (2022) 1427–1435.
[139] K. Kennedy, R. Cal, R. Casey, C. Lopez, A. Adelfio, B. Molloy, A. Wall, T. Holton, N. Khaldi, The anti-ageing effects of a natural peptide discovered by artificial intelligence, Int. J. Cosmet. Sci. 42 (2020) 388–398.
[140] C.D. Fjell, J.A. Hiss, R.E. Hancock, G. Schneider, Designing antimicrobial peptides: form follows function, Nat. Rev. Drug Discov. 11 (2012) 37–51.
[141] G.-Y. Liu, D. Yu, M.-M. Fan, X. Zhang, Z.-Y. Jin, C. Tang, X.-F. Liu, Antimicrobial resistance crisis: could artificial intelligence be the solution? Military Medical Research 11 (2024) 7.
[142] Y. Jiang, Y. Chen, Z. Song, Z. Tan, J. Cheng, Recent advances in design of antimicrobial peptides and polypeptides toward clinical translation, Adv. Drug Deliv. Rev. 170 (2021) 261–280.
[143] E.V. Bernstam, P.K. Shireman, F. Meric-Bernstam, M. N Zozus, X. Jiang, B.B. Brimhall, A.K. Windham, S. Schmidt, S. Visweswaran, Y. Ye, Artificial intelligence in clinical and translational science: successes, challenges and opportunities, Clinical and translational science 15 (2022) 309–321.
[144] C. Carini, A.A. Seyhan, Tribulations and future opportunities for artificial intelligence in precision medicine, J. Transl. Med. 22 (2024) 411.
[145] B. Jones, M. Archer, S. Germain, Reality check: the limitations of artificial intelligence in clinical medicine, Int. J. Dent. Hyg. 1 (2021).

[146] M. Feehan, L.A. Owen, I.M. McKinnon, M.M. DeAngelis, Artificial intelligence, heuristic biases, and the optimization of health outcomes: cautionary optimism, J. Clin. Med. 10 (2021) 5284.
[147] B. Peters, S.E. Brenner, E. Wang, D. Slonim, M.G. Kann, Putting Benchmarks in Their Rightful Place: the Heart of Computational Biology, Public Library of Science San Francisco, CA USA, 2018 e1006494.
[148] J. Bajorath, Artificial intelligence in interdisciplinary life science and drug discovery research, Future Science OA 8 (2022) FSO792.
[149] Q. Wu, H. Ke, D. Li, Q. Wang, J. Fang, J. Zhou, Recent progress in machine learning-based prediction of peptide activity for drug discovery, Curr. Top. Med. Chem. 19 (2019) 4–16.
[150] S. Vatansever, A. Schlessinger, D. Wacker, H.Ü. Kaniskan, J. Jin, M.M. Zhou, B. Zhang, Artificial intelligence and machine learning-aided drug discovery in central nervous system diseases: state-of-the-arts and future directions, Med. Res. Rev. 41 (2021) 1427–1473.
[151] Á. Canal-Alonso, P. Jiménez, N. Egido, J. Prieto Tejedor, J.M. Corchado Rodríguez, Revolutionizing Pharmaceuticals: Applications and Potential of Generative Artificial Intelligence in Drug Discovery, 2022.
[152] C. Selvaraj, I. Chandra, S.K. Singh, Artificial intelligence and machine learning approaches for drug design: challenges and opportunities for the pharmaceutical industries, Mol. Divers. (2021) 1–21.
[153] U. Food, D. Administration, The Software Precertification (Pre-cert) Pilot Program: Tailored Total Product Lifecycle Approaches and Key Findings, 2023.
[154] K. Sharma, P. Manchikanti, Regulation of artificial intelligence in drug discovery and health care, Biotechnol. Law Rep. 39 (2020) 371–380.
[155] J. Fetse, S. Kandel, U.-F. Mamani, K. Cheng, Recent advances in the development of therapeutic peptides, Trends Pharmacol. Sci. 44 (2023) 425–441.
[156] M. Goles, A. Daza, G. Cabas-Mora, L. Sarmiento-Varón, J. Sepúlveda-Yañez, H. Anvari-Kazemabad, M.D. Davari, R. Uribe-Paredes, Á. Olivera-Nappa, M. A. Navarrete, Peptide-based drug discovery through artificial intelligence: towards an autonomous design of therapeutic peptides, Briefings Bioinf. 25 (2024).