Molecular Diversity (2024) 28:2365–2374
https://doi.org/10.1007/s11030-023-10718-3

COMPREHENSIVE REVIEW

Review and perspective on bioinformatics tools using machine learning and deep learning for predicting antiviral peptides

Nicolás Lefin[1] · Lisandra Herrera-Belén[2] · Jorge G. Farias[1] · Jorge F. Beltrán[1]

Received: 2 May 2023 / Accepted: 15 August 2023 / Published online: 26 August 2023

Abstract

Viruses constitute a constant threat to global health and have caused millions of human and animal deaths throughout human history. Despite advances in the discovery of antiviral compounds that help fight these pathogens, finding a solution to this problem continues to be a task that consumes time and financial resources. Currently, artificial intelligence (AI) has revolutionized many areas of the biological sciences, making it possible to decipher patterns in amino acid sequences that encode different functions and activities. Within the field of AI, machine learning, and deep learning algorithms have been used to discover antimicrobial peptides. Due to their effectiveness and specificity, antimicrobial peptides (AMPs) hold excellent promise for treating various infections caused by pathogens. Antiviral peptides (AVPs) are a specific type of AMPs that have activity against certain viruses. Unlike the research focused on the development of tools and methods for the prediction of antimicrobial peptides, those related to the prediction of AVPs are still scarce. Given the significance of AVPs as potential pharmaceutical options for human and animal health and the ongoing AI revolution, we have reviewed and summarized the current machine learning and deep learning-based tools and methods available for predicting these types of peptides.

Keywords Machine learning · Deep learning · Antiviral peptide · Virus · Bioinformatics

Introduction

The discovery of new pathogens significantly impacts infectious diseases and human health worldwide [1]. Evidence of this is the COVID-19 pandemic that began in 2020, which underscores the increasing concern over viral pandemics as human populations become more mobile [2]. Many of the most important clinical and public health outbreaks have been caused by emerging viruses such as the hantavirus [3], the severe acute respiratory syndrome (SARS) coronavirus [4], the 2009 pandemic H1N1 influenza virus [5], the coronavirus EMC [6, 7] the H7N9 avian influenza virus [8], and most recently, the severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) [9]. Most of these originate from animal reservoirs. Despite advances in antiviral development, major problems persist, including antimicrobial resistance and a lack of broad-spectrum antiviral drugs, highlighting the urgent need for additional alternatives to treat infectious diseases [10]. Research and development of new vaccines are often challenging and time-consuming, contributing to the prevalence of viruses as a major cause of human diseases [11, 12]. Therefore, antiviral drug therapy is the most commonly employed approach to manage viral outbreaks [13]. Generally, antivirals target viral and host molecular structures, representing the most prevalent mechanisms of action [14]. Antivirals that target viruses inhibit vital transcriptional and replication enzymes, such as proteases and polymerases [15–17], or directly inactivate viral structural proteins [18, 19]. Conversely, host-directed antivirals inhibit cyclophilins, crucial cellular factors sequestered by some viruses during the replication cycle [20, 21].

Due to their unfavorable side effects and increasing risk of resistance, current antiviral medications have inherent limitations. Consequently, there has been significant scientific interest in finding synthetic [22] and natural chemicals [23–25] with potential antiviral activities. One such promising class of chemicals is antimicrobial peptides (AMPs) with antiviral activity [26]. These antiviral peptides (AVPs)

✉ Jorge F. Beltrán
beltran.lissabet.jf@gmail.com

[1] Department of Chemical Engineering, Faculty of Engineering and Science, University of La Frontera, Ave. Francisco Salazar, 01145 Temuco, Chile

[2] Departamento de Ciencias Básicas,Facultad de Ciencias, Universidad Santo Tomás, Temuco, Chile

can inhibit viral replication through various mechanisms, including the direct inactivation of viral proteins or the inhibition of enzymes involved in the viral replication cycle. The specific action sites and inhibition modes of AVPs can vary depending on the virus type and the replication cycle stage at which they exert their effects [14]. Several AVPs are currently patented and approved by the Food and Drug Administration (FDA) for use against the influenza virus, hepatitis virus, and human immunodeficiency virus (HIV), demonstrating the efficacy of these robust therapeutic alternatives [27].

Bioinformatics tools play a critical role in the discovery of both conventional antivirals and AVPs. Various techniques are available, including the traditional method based on the quantitative structure–activity relationship [28–30], protein-peptide dockings [31–33], protein-small ligand dockings [34–36], and molecular dynamics [36–39], among others. Recently, there has been an increase in the development of bioinformatics tools that use machine-learning techniques for AVPs activity prediction from the primary sequence [40]. By focusing on decoding the information in the primary sequence of peptides, bioinformatics tools using machine learning (ML) and deep learning (DL) approaches for AVPs prediction are powerful strategies for predicting biological activity against certain viral pathogens [41]. These predictive models are trained with datasets of peptides with and without antiviral activity [42, 43].

Given the importance of AVPs to global health and the significant role of ML and DL algorithms in redefining many conventional antiviral drug discovery concepts, this study reviews cutting-edge bioinformatics tools based on both approaches for predicting AVPs. It also discusses the ongoing challenges and future perspectives in this field.

Overview of AVPs: characteristics and mechanisms of action

Peptides have been studied for at least 40 years and have exhibited a range of biological effects. Some are termed "promiscuous" due to their multiple effects. Specifically, AVPs can be an ideal candidate for treating viral infections due to their key features, which include cationic, and amphipathic characteristics, and positive net charges [14].

Several studies have proposed mechanisms of action for AVPs. These peptides may act directly by inhibiting the viral particle or by competing for the membrane binding site (such as heparan sulfate), thereby interfering with the adsorption of the viral particle onto the host cell [14, 44, 45]. Additionally, it has been observed that AVPs can act at other stages of the viral cycle, such as inhibiting the replication, transcription, and translation complex, as depicted in Fig. 1 [45–47]. These interactions are due to the charge, amphipathic nature, and/or hydrophobic properties of AVPs [48].

AVPs are a subset of AMPs, distinguished by their activity against viruses [49]. These types of peptides are produced by organisms and microorganisms as a defense mechanism against viruses, and they have been found in a wide variety of sources, including plants, bacteria, animals, and fungi [14, 49]. Some AVPs are produced constitutively, while others are generated in response to the presence of viruses [50].

Several AVPs are currently under study, including melittin, a peptide found in honeybee venom. This peptide has demonstrated potent antiviral activity against a wide range of viruses, including HIV, herpes simplex virus, dengue virus, Junín virus, coxsackievirus, enterovirus, influenza A viruses, respiratory syncytial virus, vesicular stomatitis virus, tobacco mosaic virus [51] and SARS-CoV-2 [52, 53]. Another AVP, LL-37, is a naturally occurring peptide in the human body that has shown antiviral activity against pandemic influenza A viruses [54], respiratory syncytial virus [55], hepatitis C virus [56], HIV-1 [57], and Kaposi's sarcoma-associated herpesvirus [58]. Plectasin, derived from the fungus *Plectania nigrella*, has also displayed potent antiviral activity against Flaviviruses [59] and the hepatitis C virus [60].

Advantages and disadvantages of using antiviral peptides

Natural AVPs have several advantages over traditional chemical compounds. These include high specificity and effectiveness, low toxicity (since the final breakdown products are amino acids), and biodegradability by peptidases (which prevents accumulation in organisms). Additionally, they can inhibit large surface area interactions (such as binding proteins), have a low molecular weight, and most importantly, can exert broad-spectrum activity against different viruses [61]. However, there are some limitations to using these compounds as antivirals. These include a short half-life (rapid blood clearance), potential immunogenicity, complex modes of action, high production costs, and membrane impermeability (which depends on multiple factors, such as length and amino acid composition). Other drawbacks include solubility issues and low oral absorption [46].

Overcoming these limitations is possible by employing new technologies to modify these molecules and enhance their stability/delivery properties. Efforts have been made to improve the properties of peptides by optimizing physicochemical parameters. This can be done by introducing or fluorinating amino acids, restricting their conformations, and modifying them with polymers. These methodologies can enhance resistance to natural proteases, thereby improving

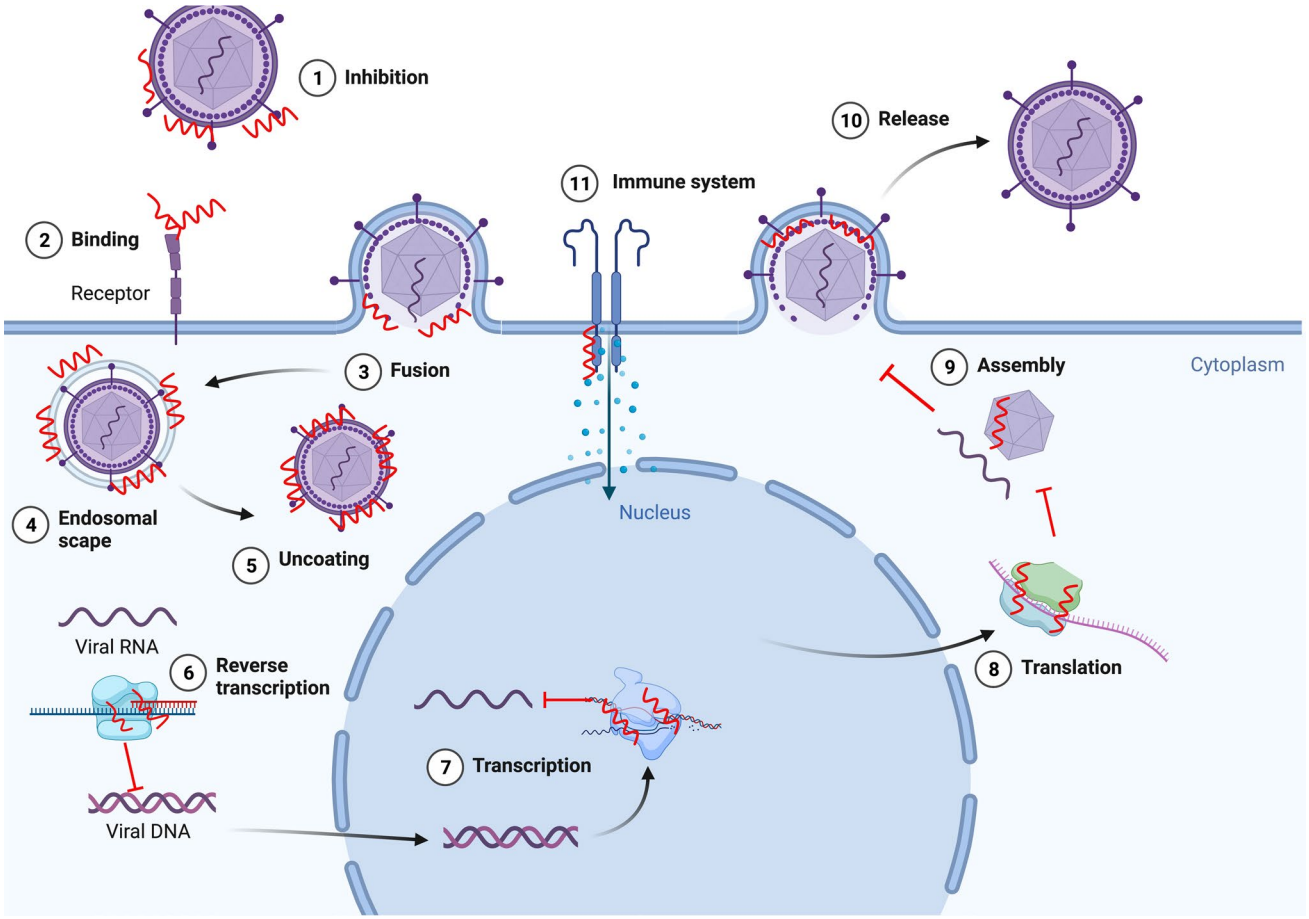

Fig. 1 Antiviral peptides can interfere with various stages of the viral replication cycle by targeting specific sites of inhibition. The peptides that have been studied and whose mechanisms of action are known can be classified into eleven different categories according to the site at which they exert their inhibitory effects: (1) Inhibition of virion formation, (2) Inhibition of adsorption, (3) Blocking of viral fusion, (4) Hindrance of endosomal escape, (5) Inhibition of viral uncoating, (6) Interference with viral genome replication, (7) Inhibition of transcription, (8) Inhibition of translation, (9) Prevention of assembly, (10) Prevention of mature virion release, and (11) Modulation of the host cell's antiviral immune system by up-regulating cytokines

serum half-life and facilitating oral administration [62]. In addition, various drug delivery strategies have been tested, such as lipidation and glycosylation, which can enhance cellular adhesion of peptides and transmembrane transport [63]. Another approach, for example, involves encapsulating enfuvirtide in a nano-liposome [64]. Although resistance to chemical and peptide antivirals is an issue, one alternative to circumvent this resistance involves changing the inhibitory target of the host viruses. The relatively shorter half-life of the peptide could be advantageously utilized by employing small peptides that bind to receptors used by viruses to enter cells. The peptide would degrade faster than chemical molecules, leaving the biological system without the typical byproducts generated by small molecules, which could be harmful and toxic [65]. Peptide cyclization is a common strategy in drug design to increase the stability and biological activity of peptides. Cyclization can limit the conformational flexibility of peptides, enhancing their resistance to enzymatic degradation, increasing their binding affinity to the target, and improving their selectivity [61, 66–69].

Synthetic peptides also hold promise for improving the structural requirements for the antiviral activity of peptides, which often correlate with a high cationic and amphiphilic character, as well as the spatial positioning of charged residues. Moreover, it has been shown that peptide analogs with arginine residues have a higher affinity for heparan sulfate than those substituted with lysines [70]. One approach to reducing production costs is identifying the minimal residues necessary to achieve sufficient antiviral activity. Furthermore, peptide synthesis technologies and amide purification methods are promising approaches for the mass production of these peptides [65].

Rational design of AVPs

In addition to natural peptides, another approach to discovering new antiviral drugs involves predicting and rationally designing new molecules. There are three main methods for rational design: template-based design, de novo design, ensemble docking, and other computational methods, all aimed at creating new peptides or improving existing ones [71]. Template-based design involves modifying a known peptide sequence by adding an amino acid or changing its position to improve selectivity and/or increase activity. This can lead to the creation of a new AVPs, even from inactive peptides. De novo design allows for the creation of new peptides using characteristics and patterns or frequencies of amino acids that correlate with AVP mechanisms and specificity with a target pathogen. These characteristics may include amino acid composition, percentage of hydrophobicity, total charge, peptide size, and disulfide content. The latter methodology, encompassing ensemble docking and other computational methods, can be further divided into two categories: docking-based methods and ML methods. This distinction is important due to the significant potential of machine learning tools in the rational design of viral peptides. Docking-based methods allow for the analysis of the interaction between the designed peptide, whose tertiary structure is designed using molecular dynamics or structural biology techniques such as NMR, and the binding protein [72]. However, this methodology has limitations, such as low precision of structures and conformational changes during binding [73, 74]. On the other hand, the ML method improves and incorporates the aforementioned methods, possessing the ability to integrate models to generate significant improvements in designs [75]. However, it is important to note that the superiority of ML/DL methods over traditional docking methods in terms of precision and accuracy is still being investigated, and results may vary depending on specific contexts and application circumstances. Drug development is a complex multi-objective optimization problem that must meet various criteria, including activity against the biological target, appropriate biophysical and pharmacokinetic properties, and low immunogenicity against the organism. ML deals with algorithms that can learn from samples containing complex patterns, noise, and redundancies [76].

Currently, thousands of AVPs have been reported, and most of them are available in specialized AMP databases that collect this type of information. Examples of these databases are AVPpred [77], HIPdb [78], ACovPepDB [79], APD3 [80], DRAMP [81], and CAMPR [82], among others. These databases present great opportunities for the development of machine learning and deep learning strategies to predict activity.

Overview of machine learning and deep learning for peptide function prediction

In the field of bioinformatics, the use of ML and DL techniques has been shown to be valuable in predicting the function of peptides and proteins from their primary amino acid sequence [83–85]. These approaches have been widely used in various research fields for predicting protein secondary and tertiary structures [86, 87], B and T cell epitopes [88, 89], protein cellular localization [90–92], and AMPs [93], among others. These use cases demonstrate how the use of these techniques in bioinformatics has enabled a better understanding of peptide and protein function, thus opening new possibilities for the development of therapies and treatments.

In recent years, one of the research fields in which ML and DL are being widely used is the prediction of AMP activity [93], which has been made possible due to the growth of databases fed with new reports of AMPs. Both ML and DL are approaches that allow for the detection and learning of complex patterns from biological data where traditional bioinformatics methods are not efficient [94]. It is important to note that ML is a general term that refers to any process by which a system can learn from data, while DL specifically refers to the use of deep artificial neural networks to model complex patterns from large amounts of data. Basically, DL is a subcategory of ML and is based on the use of deep artificial neural networks [95].

Methods and tools based on ML for AVP prediction

AVPpred is the first tool developed for AVP prediction, utilizing a support vector machine (SVM). For the generation of this model, a total of 1056 AVPs was used, divided into 951 for training and 105 as a validation dataset. This process evaluated different features such as motifs and alignment, followed by amino acid composition and physiochemical properties. Based on the physiochemical properties, the best performance measures were reported on the validation dataset with 86% accuracy (Ac) and 0.71 Matthew's Correlation Coefficient (MCC) [77]. In a study by Chang and Yang, the random forest (RF) algorithm was evaluated for the first time. They proposed a model based on physiochemical properties, which improved its performance when assisted with aggregation and secondary structural features, achieving 90% precision (Ps) and 0.79 MCC, surpassing the AVPred strategy based on SVM [96]. AntiVPP 1.0 is a pioneering desktop application that incorporates the RF algorithm for AVP prediction. The model of this tool was generated using various basic protein properties such as

molecular weight, net charge, hydropathy index, number of hydrogen bond donors, and amino acid composition. The model incorporated in AntiVPP 1.0 presented suitable performance measures on an independent dataset with sensitivity (Sn) 0.87, specificity (Sp) 0.97, Ac 0.93, and MCC 0.87 [42]. ClassAMP is the first tool that uses a multiple classification approach for the prediction of peptides with antibacterial, antifungal, and antiviral activity. In this study, the SVM and RF algorithms were proposed for the classification of peptides in the aforementioned categories. Considering the metrics reported during training and testing on two independent datasets, it was observed that the models generated with both algorithms tend to overfit the antiviral category. This is evident from a decrease in performance measures when the trained models are evaluated on the independent dataset [97]. Overfitting refers to a situation in which a machine learning model is trained too well on the training data and as a result, performs poorly on new data [98, 99].

Pseudo amino acid composition (PseAAC) is a widely used method for generating a numerical representation of a peptide sequence. Owing to its usefulness, it has been extensively applied in the field of bioinformatics for the development of predictive models [100]. PseAAC was first utilized in a study carried out by Zare et al., where it was reported that the combined application of the AdaBoost and J48 algorithms can be extremely beneficial for the prediction of AVPs. The results indicated good performance metrics during the cross-validation stage: Ac 93.26%, Sn 92.6%, 93.9% Sp, 98.2% area under the ROC curve (AUC), and 0.86 MCC [101]. iAMP-2L is another multi-classification tool for AMPs that includes AVPs in its model architecture. It uses the following ten categories: antibacterial peptides, anticancer/tumor peptides, antifungal peptides, antiparasitic peptides, anti-protist peptides, AMPs with chemotactic activity, insecticidal peptides, and spermicidal peptides, in addition to anti-HIV peptides and AVPs. This tool primarily relies on PseAAC and the fuzzy K-nearest neighbor algorithm (FKNN), where the PseAAC components incorporate five physiochemical properties. This allows for performance measures of 97.72% Sn, 86.74% Sp, 92.23% Ac, and 0.84 MCC on an independent dataset [102]. Meta-iAVP is an interesting proposal that combines six ML algorithms for the prediction of AVPs [103]. In this study, it was demonstrated that the authors' proposed model shows better performance than AVPred [77], AntiVPP 1.0 [42], and Chang et al.'s method [96] on an independent dataset, exhibiting the following performance measures: 94.90% Ac, 91.70% Sn, 98.30% Sp, and 0.90 MCC. The Meta-iAVP model also includes PseAAC as part of its training [103].

AVPIden currently stands out as one of the most intriguing tools for the prediction of AVPs, due to its superior performance and innovative purpose. This tool not only distinguishes whether a peptide possesses antiviral activity, but it also predicts the specific species and virus family against which an AVP could be active. AVPIden facilitates AVP predictions for six virus families: Herpesviridae, Paramyxoviridae, Coronaviridae, Retroviridae, Flaviviridae, and Orthomyxoviridae, as well as eight species: HPIV3, FIV, HIV, HCV, HSV1, RSV, INFVA, and SARS-CoV. The development of AVPIden's predictive model involved the use of different descriptors such as amino acid composition (AAC), dipeptide composition (DC), the composition of k-spaced amino acid pairs (CKSAAGP), PseAAC, and physicochemical properties (PHYC), along with the RF algorithm. This combination yielded good performance measures for each of the categories included in the classification [104]. The study by Pang et al. is notable for its focus on the specific prediction of AVPs against the coronavirus group. This work evaluated the same descriptors as their subsequent work (AVPIden) [104]. The results of this study reported a 79.42% GMean on an independent dataset using the RF algorithm [105]. FIRM-AVP is another tool for the prediction of AVPs. In its development, three classification algorithms (ANN, RF, SVM) were evaluated to generate predictive models of these peptides. It was found that classification with SVM yields the best results on an independent dataset, with performance measures of 93.3% Sn, 91.1% Sp, 92.4% Ac, and 0.84 MCC [41].

Until now, all the previously mentioned tools and methods have relied on the use of classification algorithms. The only reported tool for predicting AVPs using regression-based methods is AVP-IC_{50}Pred [106] to date. This tool enables the prediction of antiviral peptide activity, expressed in half-maximal inhibitory concentration (IC_{50}) values. Several models for this study were generated using the SVM, RF, IBk, and kStar algorithms. However, the best performance was observed with the SVM and RF algorithms, reporting Pearson correlation coefficients (PCC) of 0.66 and 0.64, respectively. Importantly, this tool can generate AVP predictions not only using the predictive models created with these individual algorithms but also by utilizing a hybrid approach that combines them. All the tools and methods discussed in this work are summarized in Table 1.

Methods and tools based on DL for predicting AVPs

To date, only a few studies have been reported concerning the use of DL for predicting AVPs, a trend that is likely to change in the coming years considering the recent advancements in the development of new artificial neural network (ANN) architectures and the increasing number of new AVPs being reported (Table 1). iAMP-CA2L is a tool for predicting AMPs, but it also enables the prediction of AVPs since it adopts a multiple classification approach,

Table 1 Current machine learning and deep learning tool and methods to predict the AVPs activity

Name	Project	Algorithm	References
AVPpred	http://crdd.osdd.net/servers/avppred/	SVM	[77]
Chang and Yang method	–	RF	[96]
AntiVPP 1.0	Standalone	RF	[42]
ClassAMP	http://www.bicnirrh.res.in/classamp/	SVM and RF	[97]
Zare et al. method	–	AdaBoost (J48)	[101]
iAMP-2L	http://www.jci-bioinfo.cn/iAMP-2L (*not working*)	FKNN	[102]
Meta-iAVP	http://codes.bio/meta-iavp/	hybrid approach (RF, SVM, *k*-NN, RPART, GLM, XGBoost)	[103]
AVPIden	https://awi.cuhk.edu.cn/AVPIden/#/	RF	[104]
Pang et al. method	https://github.com/poncey/PreAntiCoV	RF	[105]
AVP-IC_{50}Pred	http://crdd.osdd.net/servers/ic50avp/	SVM, RF, IBk, kStar	[106]
FIRM-AVP	https://github.com/pmartR/FIRM-AVP https://msc-viz.emsl.pnnl.gov/AVPR	SVM, ANN, RF	[41]
iAMP-CA2L	https://github.com/liujin66/iAMP-CA2L	CNN-BiLSTM-SVM	[107]
ENNAVIA	https://research.timmons.eu/ennavia	ANN	[43]
AI4AVP	https://github.com/LinTzuTang/AI4AVP_predictor	CNN	[108]
Deep-AVPpred	https://deep-avppred.anvil.app/	ANN	[109]

SVM support vector machine, *RF* random forest, *FKNN* fuzzy *k*-nearest neighbor, *k-NN* *k*-nearest neighbor, *RPART* recursive partitioning and regression trees, *GLM* generalized linear model, *XGBoost* extreme gradient boosting, *CNN* convolutional neural network, *BiLSTM* bidirectional long short-term memory recurrent neural network, *ANN* artificial neural network

including categories such as AVPs and anti-HIV peptides. iAMP-CA2L is supported by an intriguing combination of algorithms including convolutional neural network (CNN), bidirectional long short-term memory recurrent neural network (BiLSTM), SVM, and cellular automata image [107]. ENNAVIA is an interesting proposal based on ANN for predicting AVPs using a classification approach. The authors of this study reported performance metrics that surpassed previous works such as AVPred [77], Chang and Yang method [96], AntiVPP 1.0 [42], Meta-iAVP [103], Firm-AVP [41], and Pang et al. method [105] during the cross-validation and test stages on independent datasets.

AI4AVP is another tool where CNN was also used for training the model, but this study uniquely applied a generative adversarial network (GAN) prior to training as a data augmentation technique for AVPs. The authors reported superior performance of their model on an independent dataset compared to AVPpred, AntiVPP 1.0, Meta-iAVP, and FIRM-AVP [108]. Deep-AVPpred is based on the concept of transfer learning with a deep learning algorithm. Using this tool, the authors reported the identification of AVPs from human interferons-α family proteins. Deep-AVPpred achieved Ps of 94% and 93% in the validation and test stages of the model, respectively [109].

General perspectives

The correct selection of the most effective tool or method for predicting AVPs generally depends on the researcher's purpose. It is noteworthy that there has been an increase in these types of tools in recent years, largely due to the lessons learned from the COVID-19 pandemic. However, the performance measures of the most recent tools from the last three years up to the present do not show much difference between them. Therefore, their selection will depend on the specific interests of each researcher.

Nonetheless, we wish to highlight the tool AVPIden, as this tool is specifically designed to predict AVPs at the virus family and species levels, and it has demonstrated solid performance measures with the data available to date [104]. In this article, we encourage future research related to AVP prediction to evaluate the prediction approach proposed by the authors of AVPIden, which could be improved using data augmentation techniques. For example, the authors of AI4AVP [108] propose using Generative Adversarial Networks (GANs) to augment the data for AVPs of those viruses for which information is currently scarce. This tool, unique since its 2015 publication, allows the prediction of the IC_{50} for AVPs. In fact, combining classification and regression methods can be a highly effective strategy. For example, AVPs identified using classification methods could subsequently be re-ranked using regression methods. This process could more accurately identify peptides with a higher probability of exhibiting antiviral activity.

Conclusions

Viruses pose a constant threat to animal and human health, necessitating immediate action to combat the infections they cause. AVPs are regarded as excellent candidates for countering pathogenic viruses, primarily due to their high activity and specificity against the therapeutic targets they engage. The development of ML and DL based tools and methods for predicting AVPs in the coming years could positively influence the discovery of new antiviral drugs. Future research should aim to identify AVPs against specific viruses, explore novel deep learning techniques, and combine classification and regression approaches for the identification and IC_{50} determination of AVPs, respectively.

Author contributions NL and JF wrote the main manuscript. LH and JG reviewed the manuscript. JF supervised the work.

Declarations

Competing interests The authors declare no conflict of interest.

References

1. Chiu CY (2013) Viral pathogen discovery. Curr Opin Microbiol 16:468–478. https://doi.org/10.1016/j.mib.2013.05.001
2. Gan BH, Gaynord J, Rowe SM et al (2021) The multifaceted nature of antimicrobial peptides: current synthetic chemistry approaches and future directions. Chem Soc Rev 50:7820–7880. https://doi.org/10.1039/D0CS00729C
3. Nichol ST, Spiropoulou CF, Morzunov S et al (1993) Genetic Identification of a Hantavirus Associated with an Outbreak of Acute Respiratory Illness. Science 262:914–917. https://doi.org/10.1126/science.8235615
4. Rota PA, Oberste MS, Monroe SS et al (2003) Characterization of a novel coronavirus associated with severe acute respiratory syndrome. Science 300:1394–1399. https://doi.org/10.1126/science.1085952
5. Shinde V, Bridges CB, Uyeki TM et al (2009) Triple-Reassortant Swine Influenza A (H1) in humans in the United States, 2005–2009. N Engl J Med 360:2616–2625. https://doi.org/10.1056/NEJMoa0903812
6. Zaki AM, van Boheemen S, Bestebroer TM et al (2012) Isolation of a novel coronavirus from a man with pneumonia in Saudi Arabia. N Engl J Med 367:1814–1820. https://doi.org/10.1056/NEJMoa1211721
7. van Boheemen S, de Graaf M, Lauber C et al (2012) Genomic characterization of a newly discovered coronavirus associated with acute respiratory distress syndrome in humans. mBio 3:10–1128. https://doi.org/10.1128/mBio.00473-12
8. Gao R, Cao B, Hu Y et al (2013) Human infection with a novel Avian-Origin Influenza A (H7N9) Virus. N Engl J Med 368:1888–1897. https://doi.org/10.1056/NEJMoa1304459
9. Andersen KG, Rambaut A, Lipkin WI et al (2020) The proximal origin of SARS-CoV-2. Nat Med 26:450–452. https://doi.org/10.1038/s41591-020-0820-9
10. Kaufmann SHE, Dorhoi A, Hotchkiss RS, Bartenschlager R (2018) Host-directed therapies for bacterial and viral infections. Nat Rev Drug Discov 17:35–56. https://doi.org/10.1038/nrd.2017.162
11. Mahmoud A (2016) New vaccines: challenges of discovery. Microb Biotechnol 9:549–552. https://doi.org/10.1111/1751-7915.12397
12. Kuri PR, Goswami P (2022) Current update on rotavirus in-silico multiepitope vaccine design. ACS Omega. https://doi.org/10.1021/acsomega.2c07213
13. Martinez MA (2022) Efficacy of repurposed antiviral drugs: Lessons from COVID-19. Drug Discov Today 27:1954–1960. https://doi.org/10.1016/j.drudis.2022.02.012
14. Vilas Boas LCP, Campos ML, Berlanda RLA et al (2019) Antiviral peptides as promising therapeutic drugs. Cell Mol Life Sci 76:3525–3542. https://doi.org/10.1007/s00018-019-03138-w
15. Kumar N, Acharya V (2023) Machine intelligence-guided selection of optimized inhibitor for human immunodeficiency virus (HIV) from natural products. Comput Biol Med 153:106525. https://doi.org/10.1016/j.compbiomed.2022.106525
16. Mótyán JA, Mahdi M, Hoffka G, Tőzsér J (2022) Potential resistance of SARS-CoV-2 Main Protease (Mpro) against protease inhibitors: lessons learned from HIV-1 protease. Int J Mol Sci 23:3507. https://doi.org/10.3390/ijms23073507
17. NunesSantos DAFR, Fonseca STD et al (2022) NS2B-NS3 protease inhibitors as promising compounds in the development of antivirals against Zika virus: a systematic review. J Med Virol 94:442–453. https://doi.org/10.1002/jmv.27386
18. Wang L, Wu Y, Yao S et al (2022) Discovery of potential small molecular SARS-CoV-2 entry blockers targeting the spike protein. Acta Pharmacol Sin 43:788–796. https://doi.org/10.1038/s41401-021-00735-z
19. Jiang S, Tuzikov A, Andrianov A (2022) Small-molecule HIV-1 entry inhibitors targeting the epitopes of broadly neutralizing antibodies. Cell Chem Biol 29:757–773. https://doi.org/10.1016/j.chembiol.2022.03.009
20. Kumar N, Sharma S, Kumar R et al (2020) Host-directed antiviral therapy. Clin Microbiol Rev 33:10–1128. https://doi.org/10.1128/CMR.00168-19
21. Wallis RS, O'Garra A, Sher A, Wack A (2022) Host-directed immunotherapy of viral and bacterial infections: past, present and future. Nat Rev Immunol. https://doi.org/10.1038/s41577-022-00734-z
22. Parra AL, Bezerra LP, Shawar DE et al (2022) Synthetic antiviral peptides: a new way to develop targeted antiviral drugs. Future Virol 17:577–591. https://doi.org/10.2217/fvl-2021-0308
23. Abookleesh FL, Al-Anzi BS, Ullah A (2022) Potential antiviral action of alkaloids. Molecules 27:903. https://doi.org/10.3390/molecules27030903
24. Antonopoulou I, Sapountzaki E, Rova U, Christakopoulos P (2022) Ferulic acid from plant biomass: a phytochemical with promising antiviral properties. Front Nutr 8:777576. https://doi.org/10.3389/fnut.2021.777576
25. Xu X-Y, Wang D-Y, Li Y-P et al (2022) Plant-derived lignans as potential antiviral agents: a systematic review. Phytochem Rev 21:239–289. https://doi.org/10.1007/s11101-021-09758-0
26. Zakaryan H, Chilingaryan G, Arabyan E et al (2021) Natural antimicrobial peptides as a source of new antiviral agents. J Gen Virol 102:001661. https://doi.org/10.1099/jgv.0.001661
27. Mousavi Maleki MS, Sardari S, Ghandehari Alavijeh A, Madanchi H (2022) Recent patents and FDA-approved drugs based on antiviral peptides and other peptide-related antivirals. Int J Pept Res Ther 29:5. https://doi.org/10.1007/s10989-022-10477-z
28. Chayrov R, Parisis NA, Chatziathanasiadou MV et al (2020) Synthetic analogues of aminoadamantane as influenza viral

inhibitors—in vitro, in silico and QSAR studies. Molecules 25:3989. https://doi.org/10.3390/molecules25173989
29. Ding X, Kang D, Sun L et al (2022) Combination of 2D and 3D-QSAR studies on DAPY and DANA derivatives as potent HIV-1 NNRTIs. J Mol Struct 1249:131603. https://doi.org/10.1016/j.molstruc.2021.131603
30. Toropov AA, Toropova AP, Veselinović AM et al (2022) SARS-CoV M ^{pro} inhibitory activity of aromatic disulfide compounds: QSAR model. J Biomol Struct Dyn 40:780–786. https://doi.org/10.1080/07391102.2020.1818627
31. Padhi S, Sanjukta S, Chourasia R et al (2021) A multifunctional peptide from bacillus fermented soybean for effective inhibition of SARS-CoV-2 S1 receptor binding domain and modulation of toll like receptor 4: a molecular docking study. Front Mol Biosci 8:636647. https://doi.org/10.3389/fmolb.2021.636647
32. Pei P, Qin H, Chen J et al (2021) Computational design of ultrashort peptide inhibitors of the receptor-binding domain of the SARS-CoV-2 S protein. Brief Bioinform 22:bbab243. https://doi.org/10.1093/bib/bbab243
33. Singh S, Banavath HN, Godara P et al (2022) Identification of antiviral peptide inhibitors for receptor binding domain of SARS-CoV-2 omicron and its sub-variants: an in-silico approach. 3 Biotech 12:198. https://doi.org/10.1007/s13205-022-03258-4
34. Al-Rabia M, Alhakamy N, Ahmed O et al (2021) Repurposing of sitagliptin- melittin optimized nanoformula against SARS-CoV-2; antiviral screening and molecular docking studies. Pharmaceutics 13:307. https://doi.org/10.3390/pharmaceutics13030307
35. Hagar M, Ahmed HA, Aljohani G, Alhaddad OA (2020) Investigation of some antiviral N-heterocycles as COVID 19 drug: molecular docking and DFT calculations. Int J Mol Sci 21:3922. https://doi.org/10.3390/ijms21113922
36. Kalbhor MS, Bhowmick S, Alanazi AM et al (2021) Multi-step molecular docking and dynamics simulation-based screening of large antiviral specific chemical libraries for identification of Nipah virus glycoprotein inhibitors. Biophys Chem 270:106537. https://doi.org/10.1016/j.bpc.2020.106537
37. Poustforoosh A, Hashemipour H, Tüzün B et al (2022) The impact of D614G mutation of SARS-COV-2 on the efficacy of anti-viral drugs: a comparative molecular docking and molecular dynamics study. Curr Microbiol 79:241. https://doi.org/10.1007/s00284-022-02921-6
38. Jahantigh H, Ahmadi N, Lovreglio P et al (2022) Repurposing antiviral drugs against HTLV-1 protease by molecular docking and molecular dynamics simulation. J Biomol Struct Dyn. https://doi.org/10.1080/07391102.2022.2078411
39. Kumari R, Kumar V, Dhankhar P, Dalal V (2022) Promising antivirals for PLpro of SARS-CoV-2 using virtual screening, molecular docking, dynamics, and MMPBSA. J Biomol Struct Dyn. https://doi.org/10.1080/07391102.2022.2071340
40. Charoenkwan P, Anuwongcharoen N, Nantasenamat C et al (2021) In silico approaches for the prediction and analysis of antiviral peptides: a review. Curr Pharm Des 27:2180–2188. https://doi.org/10.2174/1381612826666201102105827
41. Chowdhury AS, Reehl SM, Kehn-Hall K et al (2020) Better understanding and prediction of antiviral peptides through primary and secondary structure feature importance. Sci Rep 10:19260. https://doi.org/10.1038/s41598-020-76161-8
42. Beltrán Lissabet JF, Belén LH, Farias JG (2019) AntiVPP 1.0: a portable tool for prediction of antiviral peptides. Comput Biol Med 107:127–130. https://doi.org/10.1016/j.compbiomed.2019.02.011
43. Timmons PB, Hewage CM (2021) ENNAVIA is a novel method which employs neural networks for antiviral and anti-coronavirus activity prediction for therapeutic peptides. Brief Bioinform 22:bbab258. https://doi.org/10.1093/bib/bbab258
44. Jenssen H, Hamill P, Hancock REW (2006) Peptide antimicrobial agents. Clin Microbiol Rev 19:491–511. https://doi.org/10.1128/CMR.00056-05
45. Sukmarini L (2022) Antiviral Peptides (AVPs) of marine origin as propitious therapeutic drug candidates for the treatment of human viruses. Molecules 27:2619. https://doi.org/10.3390/molecules27092619
46. Galdiero S, Falanga A, Tarallo R et al (2013) Peptide inhibitors against herpes simplex virus infections. J Pept Sci 19:148–158. https://doi.org/10.1002/psc.2489
47. Vanzolini T, Bruschi M, Rinaldi AC et al (2022) Multitalented synthetic antimicrobial peptides and their antibacterial, antifungal and antiviral mechanisms. Int J Mol Sci 23:545. https://doi.org/10.3390/ijms23010545
48. Badani H, Garry RF, Wimley WC (2014) Peptide entry inhibitors of enveloped viruses: The importance of interfacial hydrophobicity. Biochim Biophys Acta-Biomembranes 1838:2180–2197. https://doi.org/10.1016/j.bbamem.2014.04.015
49. Agarwal G, Gabrani R (2021) Antiviral Peptides: Identification and Validation. Int J Pept Res Ther 27:149–168. https://doi.org/10.1007/s10989-020-10072-0
50. Handfield C, Kwock J, MacLeod AS (2018) Innate antiviral immunity in the skin. Trends Immunol 39:328–340. https://doi.org/10.1016/j.it.2018.02.003
51. Memariani H, Memariani M, Moravvej H, Shahidi-Dadras M (2020) Melittin: a venom-derived peptide with promising antiviral properties. Eur J Clin Microbiol Infect Dis 39:5–17. https://doi.org/10.1007/s10096-019-03674-0
52. Enayathullah MG, Parekh Y, Banu S et al (2022) Gramicidin S and melittin: potential anti-viral therapeutic peptides to treat SARS-CoV-2 infection. Sci Rep 12:3446. https://doi.org/10.1038/s41598-022-07341-x
53. Mousavi Maleki MS, Rostamian M, Madanchi H (2021) Antimicrobial peptides and other peptide-like therapeutics as promising candidates to combat SARS-CoV-2. Expert Rev Anti Infect Ther 19:1205–1217. https://doi.org/10.1080/14787210.2021.1912593
54. Tripathi S, Wang G, White M et al (2015) Antiviral activity of the human cathelicidin, LL-37, and derived peptides on seasonal and pandemic influenza a viruses. PLoS ONE 10:e0124706. https://doi.org/10.1371/journal.pone.0124706
55. Currie SM, Findlay EG, McHugh BJ et al (2013) The human cathelicidin LL-37 has antiviral activity against respiratory syncytial virus. PLoS ONE 8:e73659. https://doi.org/10.1371/journal.pone.0073659
56. Matsumura T, Sugiyama N, Murayama A et al (2016) Antimicrobial peptide LL-37 attenuates infection of hepatitis C virus. Hepatol Res 46:924–932. https://doi.org/10.1111/hepr.12627
57. Bergman P, Walter-Jallow L, Broliden K et al (2007) The antimicrobial peptide LL-37 inhibits HIV-1 replication. Curr HIV Res 5:410–415. https://doi.org/10.2174/157016207781023947
58. Brice DC, Toth Z, Diamond G (2018) LL-37 disrupts the Kaposi's sarcoma-associated herpesvirus envelope and inhibits infection in oral epithelial cells. Antiviral Res 158:25–33. https://doi.org/10.1016/j.antiviral.2018.07.025
59. Ji M, Zhu T, Xing M et al (2019) An antiviral peptide from Alopecosa nagpag spider targets NS2B–NS3 protease of flaviviruses. Toxins 11:584. https://doi.org/10.3390/toxins11100584
60. Abdulrahman AY, Rothan HA, Rashid NN et al (2017) Identification of peptide leads to inhibit hepatitis C virus: inhibitory effect of plectasin peptide against hepatitis C serine protease. Int J Pept Res Ther 23:163–170. https://doi.org/10.1007/s10989-016-9544-6

61. Wang L, Wang N, Zhang W et al (2022) Therapeutic peptides: current applications and future directions. Signal Transduct Target Ther 7:48. https://doi.org/10.1038/s41392-022-00904-4
62. Di L (2015) Strategic approaches to optimizing peptide ADME properties. AAPS J 17:134–143. https://doi.org/10.1208/s12248-014-9687-3
63. Matsuzaki K (2009) Control of cell selectivity of antimicrobial peptides. Biochim Biophys Acta (BBA)-Biomembranes 1788:1687–1692. https://doi.org/10.1016/j.bbamem.2008.09.013
64. Figueira TN, Domingues MM, Illien F et al (2020) Enfuvirtide-Protoporphyrin IX dual-loaded liposomes. in vitro evidence of synergy against HIV-1 entry into cells. ACS Infect Dis 6:224–236. https://doi.org/10.1021/acsinfecdis.9b00285
65. Lalani S, Gew LT, Poh CL (2021) Antiviral peptides against Enterovirus A71 causing hand, foot and mouth disease. Peptides 136:170443. https://doi.org/10.1016/j.peptides.2020.170443
66. Parikesit AA, Tambunan US (2013) Screening of commercial cyclic peptides as inhibitor envelope protein Dengue Virus (DENV) through molecular docking and molecular dynamics. Pak J Biol Sci 16:1836–1848. https://doi.org/10.3923/pjbs.2013.1836.1848
67. Zhang H, Chen S (2022) Cyclic peptide drugs approved in the last two decades (2001–2021). RSC Chem Biol 3:18–31. https://doi.org/10.1039/D1CB00154J
68. Jing X, Jin K (2020) A gold mine for drug discovery: strategies to develop cyclic peptides into therapies. Med Res Rev 40:753–810. https://doi.org/10.1002/med.21639
69. Li P, Roller P (2002) Cyclization strategies in peptide derived drug design. Curr Top Med Chem 2:325–341. https://doi.org/10.2174/1568026023394209
70. Jenssen H, Andersen J, Mantzilas D, Gutteberg T (2004) A wide range of medium-sized, highly cationic, ?-Helical peptides show antiviral activity against herpes simplex virus. Antiviral Res 64:119–126. https://doi.org/10.1016/j.antiviral.2004.08.003
71. Lee Y-CJ, Shirkey JD, Park J et al (2022) An overview of antiviral peptides and rational biodesign considerations. BioDesign Res. https://doi.org/10.34133/2022/9898241
72. Ciemny M, Kurcinski M, Kamel K et al (2018) Protein–peptide docking: opportunities and challenges. Drug Discov Today 23:1530–1537. https://doi.org/10.1016/j.drudis.2018.05.006
73. Oyedele A-QK, Ogunlana AT, Boyenle ID et al (2022) Docking covalent targets for drug discovery: stimulating the computer-aided drug design community of possible pitfalls and erroneous practices. Mol Divers. https://doi.org/10.1007/s11030-022-10523-4
74. Kumalo H, Bhakat S, Soliman M (2015) Theory and applications of covalent docking in drug discovery: merits and pitfalls. Molecules 20:1984–2000. https://doi.org/10.3390/molecules20021984
75. Bai Q, Liu S, Tian Y et al (2022) Application advances of deep learning methods for de novo drug design and molecular dynamics simulation. WIREs Comput Mol Sci 12:e1581. https://doi.org/10.1002/wcms.1581
76. Narayanan H, Dingfelder F, Butté A et al (2021) Machine learning for biologics: opportunities for protein engineering, developability, and formulation. Trends Pharmacol Sci 42:151–165. https://doi.org/10.1016/j.tips.2020.12.004
77. Thakur N, Qureshi A, Kumar M (2012) AVPpred: collection and prediction of highly effective antiviral peptides. Nucleic Acids Res 40:W199–W204. https://doi.org/10.1093/nar/gks450
78. Qureshi A, Thakur N, Kumar M (2013) HIPdb: a database of experimentally validated HIV inhibiting peptides. PLoS ONE 8:e54908. https://doi.org/10.1371/journal.pone.0054908
79. Zhang Q, Chen X, Li B et al (2022) A database of anti-coronavirus peptides. Sci Data 9:294. https://doi.org/10.1038/s41597-022-01394-3
80. Wang G, Li X, Wang Z (2016) APD3: the antimicrobial peptide database as a tool for research and education. Nucleic Acids Res 44:D1087–D1093. https://doi.org/10.1093/nar/gkv1278
81. Shi G, Kang X, Dong F et al (2022) DRAMP 3.0: an enhanced comprehensive data repository of antimicrobial peptides. Nucleic Acids Res 50:D488–D496. https://doi.org/10.1093/nar/gkab651
82. Waghu FH, Barai RS, Gurung P, Idicula-Thomas S (2016) CAMP R3: a database on sequences, structures and signatures of antimicrobial peptides: Table 1. Nucleic Acids Res 44:D1094–D1097. https://doi.org/10.1093/nar/gkv1051
83. Larrañaga P, Calvo B, Santana R et al (2006) Machine learning in bioinformatics. Brief Bioinform 7:86–112. https://doi.org/10.1093/bib/bbk007
84. Li H, Tian S, Li Y et al (2021) Modern deep learning in bioinformatics. J Mol Cell Biol 12:823–827. https://doi.org/10.1093/jmcb/mjaa030
85. Min S, Lee B, Yoon S (2016) Deep learning in bioinformatics. Brief Bioinform. https://doi.org/10.1093/bib/bbw068
86. Wang Y, Mao H, Yi Z (2017) Protein secondary structure prediction by using deep learning method. Knowl Based Syst 118:115–123. https://doi.org/10.1016/j.knosys.2016.11.015
87. AlQuraishi M (2021) Machine learning in protein structure prediction. Curr Opin Chem Biol 65:1–8. https://doi.org/10.1016/j.cbpa.2021.04.005
88. Galanis KA, Nastou KC, Papandreou NC et al (2021) Linear B-cell epitope prediction for in silico vaccine design: a performance review of methods available via command-line interface. Int J Mol Sci 22:3210. https://doi.org/10.3390/ijms22063210
89. Peters B, Nielsen M, Sette A (2020) T cell epitope predictions. Annu Rev Immunol 38:123–145. https://doi.org/10.1146/annurev-immunol-082119-124838
90. Gudenas BL, Wang L (2018) Prediction of LncRNA subcellular localization with deep learning from sequence features. Sci Rep 8:16385. https://doi.org/10.1038/s41598-018-34708-w
91. Almagro Armenteros JJ, Sønderby CK, Sønderby SK et al (2017) DeepLoc: prediction of protein subcellular localization using deep learning. Bioinformatics 33:4049–4049. https://doi.org/10.1093/bioinformatics/btx548
92. Lu Z, Szafron D, Greiner R et al (2004) Predicting subcellular localization of proteins using machine-learned classifiers. Bioinformatics 20:547–556. https://doi.org/10.1093/bioinformatics/btg447
93. Lee EY, Lee MW, Fulan BM et al (2017) What can machine learning do for antimicrobial peptides, and what can antimicrobial peptides do for machine learning? Interface Focus 7:20160153. https://doi.org/10.1098/rsfs.2016.0153
94. Prompramote S, Chen Y, Chen Y-PP Machine learning in bioinformatics. In: Bioinformatics technologies. Springer, Berlin, pp 117–153
95. Shinde PP, Shah S (2018) A review of machine learning and deep learning applications. In: 2018 fourth international conference on computing communication control and automation (ICCUBEA). IEEE, pp 1–6
96. Chang KY, Yang J-R (2013) Analysis and prediction of highly effective antiviral peptides based on random forests. PLoS ONE 8:e70166. https://doi.org/10.1371/journal.pone.0070166
97. Joseph S, Karnik S, Nilawe P et al (2012) ClassAMP: a prediction tool for classification of antimicrobial peptides. IEEE/ACM Trans Comput Biol Bioinform 9:1535–1538. https://doi.org/10.1109/TCBB.2012.89
98. Hawkins DM (2004) The problem of overfitting. J Chem Inf Comput Sci 44:1–12. https://doi.org/10.1021/ci0342472
99. Lever J, Krzywinski M, Altman N (2016) Model selection and overfitting. Nat Methods 13:703–704. https://doi.org/10.1038/nmeth.3968

100. Du P, Gu S, Jiao Y (2014) PseAAC-general: fast building various modes of general form of Chou's pseudo-amino acid composition for large-scale protein datasets. Int J Mol Sci 15:3495–3506. https://doi.org/10.3390/ijms15033495
101. Zare M, Mohabatkar H, Faramarzi FK et al (2015) Using Chou's pseudo amino acid composition and machine learning method to predict the antiviral peptides. Open Bioinform J 9:13–19. https://doi.org/10.2174/1875036201509010013
102. Xiao X, Wang P, Lin W-Z et al (2013) iAMP-2L: a two-level multi-label classifier for identifying antimicrobial peptides and their functional types. Anal Biochem 436:168–177. https://doi.org/10.1016/j.ab.2013.01.019
103. Schaduangrat N, Nantasenamat C, Prachayasittikul V, Shoombuatong W (2019) Meta-iAVP: a sequence-based meta-predictor for improving the prediction of antiviral peptides using effective feature representation. Int J Mol Sci 20:5743. https://doi.org/10.3390/ijms20225743
104. Pang Y, Yao L, Jhong J-H et al (2021) AVPIden: a new scheme for identification and functional prediction of antiviral peptides based on machine learning approaches. Brief Bioinform 22:bbab263. https://doi.org/10.1093/bib/bbab263
105. Pang Y, Wang Z, Jhong J-H, Lee T-Y (2021) Identifying anti-coronavirus peptides by incorporating different negative datasets and imbalanced learning strategies. Brief Bioinform 22:1085–1095. https://doi.org/10.1093/bib/bbaa423
106. Qureshi A, Tandon H, Kumar M (2015) AVP-IC $_{50}$ Pred: multiple machine learning techniques-based prediction of peptide antiviral activity in terms of half maximal inhibitory concentration (IC $_{50}$). Biopolymers 104:753–763. https://doi.org/10.1002/bip.22703
107. Xiao X, Shao Y-T, Cheng X, Stamatovic B (2021) iAMP-CA2L: a new CNN-BiLSTM-SVM classifier based on cellular automata image for identifying antimicrobial peptides and their functional types. Brief Bioinform 22:bbab209. https://doi.org/10.1093/bib/bbab209
108. Lin T-T, Sun Y-Y, Wang C-T et al (2022) AI4AVP: an antiviral peptides predictor in deep learning approach with generative adversarial network data augmentation. Bioinform Adv 2:vbac080. https://doi.org/10.1093/bioadv/vbac080
109. Sharma R, Shrivastava S, Singh SK et al (2022) Deep-AVPpred: artificial intelligence driven discovery of peptide drugs for viral infections. IEEE J Biomed Health Inform 26:5067–5074. https://doi.org/10.1109/JBHI.2021.3130825