Food-Derived Bioactive Peptides And Artificial Intelligence Techniques For Their Prediction: A Brief Review

Margarita Terziyska, Ivelina Desseva, Zhelyazko Terziyski

Abstract: Biologically active peptides (BAPs) have a positive effect on human health, which is why they are used as a basis for drug and functional foods development. They are therefore of economic interest. However, the process of their isolation is too expensive and time-consuming. Hence, it is necessary to develop more effective methods to predict the potential activity of peptides. An appropriate solution could be an in silico approach, in particular the use of computational methods based on artificial intelligence (AI) techniques. The use of AI approaches may facilitate the identification of bioactive peptides. Thus, in this paper, along with some basic information about food-derived BAPs, a brief review of the AI techniques used for their activity prediction is made.

Index Terms: artificial intelligence, activity prediction, food-derived peptides, deep learning, neural networks.

1 INTRODUCTION

Artificial intelligence (AI) can be defined as a computer science of teaching machines to think like humans. AI concepts include different techniques that mimic biological systems or societal behavioral patterns. Such techniques are widely used to solve various problems in science, technology, business, education, medical diagnostics, handwriting recognition in fraud detection, image retrieval, etc. One problem in which AI could also be useful in predicting the biological activity of food-derived peptides. In recent years, there has been a significant increase in interest in food-derived BAPs. This is because BAPs possess a number of properties that allow them to control diseases and support human health. They are used to lower blood pressure, to improve immunity, to lower cholesterol, have anti-cancer [1] and antidiabetic [2] properties. In the literature, there are numerous excellent books and review papers regarding BAPs and their impact on human health [3], [4]. Because of their health-enhancing potential and safety profiles, they may be used as basic components in drugs or functional foods. Therefore, BAPs are of commercial interest as well. To date, most BAPs have been isolated following a so-called “conventional” approach. This includes the experimental in vivo and in vitro protocols, which are expensive and time-consuming. Recently, in silico approaches have been used to predict BAPs in a more effective way. The term in silico stems from the computer component silicium and in silico methods, refer to methods or predictions using computational approaches. In this sense, the algorithms of Machine Learning (ML) and AI, in turn, refer to in silico methods. This paper includes synthesized basic information about food-derived BAPs, as well as about the individual stages in the development of AI-based BAPs predictive model. It also provides a brief overview of the AI techniques used for BAPs prediction, comments on their features, and outlines the main trends in their application. The article can be considered as a kind of guide for choosing a suitable technique for predicting the potential activity of peptides.

2 BIOACTIVE PEPTIDES

Proteins are fundamental components of living organisms. They play a significant role in all cellular processes and functions. Proteins are buildup of amino acids that are ranged in polypeptide chains and thanks to the different bonds that are formed between the amino acid residues, a complex spatial structure of the proteins is formed. Peptides differ from proteins mainly by size. They can be described as proteinous fragments that consist of between 2 to 50 amino acids. Peptides can be naturally produced or be derived from native proteins (cryptides) [5]. The latter are released mainly as a result of one of the following processes: microbial fermentation; gastrointestinal digestion or proteolytic hydrolysis which will be discussed in details below. In general, what determines the potential activity of peptides is their amino acid sequence and composition [6]. These peptides can be defined as specific amino acids sequences that have a beneficial effect on the body and can have a positive effect on human health that can be measurable [6]. At the same time, they should be harmless. Therefore, researchers and scientists are most interested in those cryptides that display some kind of biological activity.

2.1 Production of Food-Derived Bioactive Peptides

In addition to their nutritional value, dietary proteins are considered a promising source of bioactive peptides. So far, animal proteins have been the most widely studied [7], [8], [9] followed by plant, vegetable and marine proteins [10], [11]. These peptides are initially in an inactive form as part of the precursor protein. Generally, they could be liberated via three ways: microbial fermentation; gastrointestinal digestion or proteolytic hydrolysis. During microbial fermentation, food components are used by the microorganisms for their life cycle and as a result, a breakdown of the macromolecules occurs. Peptide bonds in proteins are hydrolyzed by the microbial proteases and release in this way peptides and amino acids. The process can be controlled by using microbial cultures with predictable behavior. Fermented dairy products are excellent examples in this regard [12]. On this basis, many functional drinks and foods were marketed and promoted. Another way of release of bioactive peptides is throughout the gastrointestinal digestion. Then, the proteins taken with food are subjected to sequential hydrolysis by different proteases in the gastric and intestinal compartments of the digestive tract. In addition, the individual microflora further processes the resulting digesta ending up to potential liberation of bioactive peptides [13]. However, the gastrointestinal digestion is a very complex process and the real outcome of food breakdown remains hardly predictable for each individual. Many research groups are focused on the establishment of realistic in vitro

gastrointestinal method that can be used for further study on how the food is transformed. So far, INFOGEST's protocol seems to be the most detailed [14]. The third way to obtain bioactive peptides from the precursor is a direct hydrolysis with analytical grade proteolytic enzymes. Proteases are a very diverse group of enzymes that break down specific peptide bonds in protein's polypeptide chain. Therefore, different proteases will generate different hydrolysates i.e. different peptides from the same substrate. This approach is the most predictable one but at the same time remains the most expensive.

2.2 Types of Bioactive Peptides

So far, many bioactive peptide activities are recorded, among them, several are the most common. In Fig. 1 are shown the most relevant one. Antihypertensive peptides are certainly the most reported food-derived peptides [15]. These peptides inhibit a key enzyme in a renin-angiotensin system named angiotensin-converting enzyme (ACE), resulting in a lowering of blood pressure. Several peptides are even commercialized. Ile-Pro-Pro (IPP) and Val-Pro-Pro (VPP) are milk-derived peptides that possess well-characterized antihypertensive effects in vitro and in vivo [16], [17].

The existence of antimicrobial peptides (AMPs) is known for a very long time ago because they are found in nearly all life forms. They are host defense peptides and are part of the innate immune response of living organisms. AMPs usually represent a sequence of 12 to 50 amino acids, generally with cationic and hydrophobic residues that attack the integrity of the cell membrane and cell wall [18], [19]. The target cells could be Gram-positive bacteria, Gram-negative bacteria, fungi, viruses, or cancer cells [20]. Peptides with antimicrobial action are also encoded in food-grade proteins [21]. In this regard, milk-derived peptides with antimicrobial property have been extensively studied [22]. Antioxidant peptides are another large group of bioactive peptides that attract constantly scientific attention. They are able to diminish reactive oxygen species production [23]. Usually, they are very low molecular weight peptides ranging from 2 to 15 amino acids. As for all type of biologically active peptides, the amino acid sequence and composition is crucial for the manifestation of antioxidant activity. The presence of hydrophobic amino acids seems decisive in this case [24]. Extensive information regarding the food-grade protein sources is summarized in many reviews [23], [25], [26].

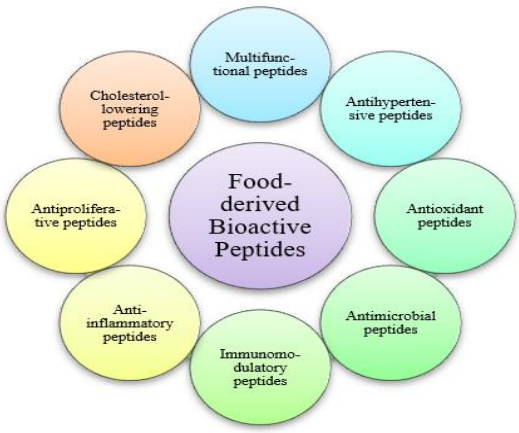

Fig. 1. Different types of bioactive peptides.

DEVELOPING PROCEDURE

Peptides possess specific biological activity, which depends on their structure i.e. on their amino acid sequence. Therefore, the most widely used approach for predicting the BAPs is one based on their sequence. The procedure of building an AI-based BAPs predictive model includes several stages, which are schematically presented in Fig. 2. The first stage is related to data preparation. The data can be collected by their function, source, length, etc. from public databases. Some of the databases with food-derived peptides are summarized in Table. 1. An important aspect in data preparation is the generation of negative samples. According to [27] the negative samples should generally consist of the following peptides: (1) random peptides retrieved from the UniProt; (2) random shuffling of positive samples; or (3) peptides with different functions rather than the desired function. Finally, the data set should be divided into training and testing data sets. The proportions are decided according to the size and type of the available data, but also on the model, you are training. In the AI field the rule 80%/20% is best adopted, i.e. 80% of the original data are used to train the predictive model and 20% are used to assess the developed model.

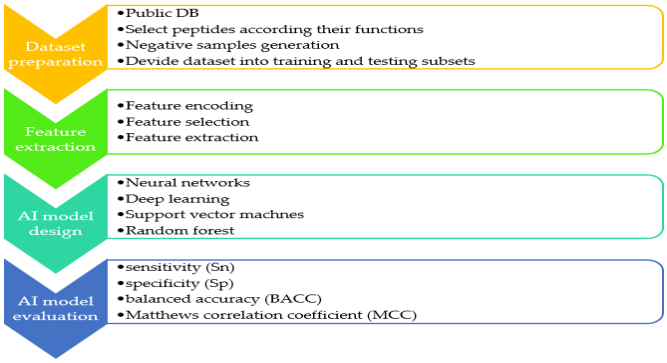

Fig. 2. Four stages BAPs predictive AI-based model developing procedure.

The second stage of the process of building the predictive AI-based model is related to the feature representation of peptides. Firstly, the peptides amino acid sequence is converted into a numerical vector of features X (which are also known as descriptors). This process is known as feature encoding. There is a huge variety of methods and mechanisms for feature encoding. Some of the most frequently used in AI methods are amino acid index (AAI),

TABLE 1

CURRENTLY AVAILABLE DATABASES OF FOOD-DERIVED BIOACTIVE PEPTIDES

Database	[a]Website
FeptideDB [28] Database of food-derived bioactive peptides	http://www4g.biotec.or.th/FeptideDB/
BioPepDB [30] Database of bioactive peptides of food origin	http://bis.zju.edu.cn/biopepdbr/
AHTPDB [31] Database of food-derived AHTPs	http://crdd.osdd.net/raghava/ahtpdb/info2.php
BIOPEP-UWM DB [29] Database of bioactive peptides of food origin	http://www.uwm.edu.pl/biochemia/index.php/en/biopep
MBPDB [32] Milk Bioactive Peptide database	http://mbpdb.nws.oregonstate.edu

[a]All websites indicated in the table were accessed in July, 2021.

amino acid composition (AAC), pseudo amino acid composition (PseAAC), binary profile (BPF), dipeptide composition (DPC), Enhanced amino acid composition (EAAC), Enhanced grouped amino acid composition (EGAAC), Grouped amino acid composition (GAAC), Physicochemical properties (PCP) and etc. Some of them are detailed described in [27]. Developing an appropriate encoding is a major challenge, which has not been entirely solved so far. For this reason, the development of novel amino acid encodings is established as stand-alone research. At the same time, several research groups have proposed useful web servers or standalone programs to compute feature descriptors from the provided peptide sequences. The most popular of them are summarized in Table 2.

Not all of the selected descriptors are relevant to the property of the peptides to be predicted. In addition, not all of them are independent of each other. This requires the exclusion of irrelevant or of mutually correlated features. This can be achieved through feature extraction or feature selection. Both techniques aim for dimensionality reduction and redundancy elimination. Feature selection is a process of choosing a subset of the original pool of features, while feature extraction is a process of getting useful features from existing data. Feature selection can be realized by filter-based, wrapper-based, or embedded methods. The things are not so simple when discussing feature extraction. In practice, there are no limits to the ways of creating features. Extraction of meaningful features often means a great deal of extensive exploration and lots of time. Therefore, it is useful to do extraction in an automatic manner. Here is the biggest advantage of DL architecture. Deep neural networks (DNNs) can extract useful features automatically because they are trained to do so. In the third stage of the process, the predictive model (or classifier) is designed by using some of the AI techniques, presented in the next section and is trained to do the accurate classification of any unseen data. During the training, feature descriptors along with the class (positive or negative), are inputted to an AI model, where it learns the relationship between feature descriptors (X) and response variable (Y). After the training, the model can make subsequent predictions for any newly provided data sets. In such a way, the designed model realized so-called mapping, which involves empirically discovering a function that maps between the feature vectors and the property of interest, denoted by the symbol Y.

TABLE 2

CURRENTLY AVAILABLE WEB SERVER OR STANDALONE PROGRAMS TO COMPUTE FEATURE DESCRIPTORS

Software	[a]*Website*
Propy [38]	code.google.com/archive/p/protpy/
Pse-in-One [39]	bliulab.net/Pse-in-One/
Pse-in-One 2.0 [40]	bliulab.net/Pse-in-One2.0/
iFeature [33]	ifeature.erc.monash.edu
Pfeature [34]	webs.iiitd.edu.in/raghava/pfeature/
iLearn [35]	ilearn.erc.monash.edu
Seq2Feature [41]	www.iitm.ac.in/bioinfo/SBFE/
PyBioMed [42]	github.com/gadsbyfly/PyBioMed
PyFeat [37]	github.com/mrzResearchArena/PyFeat/
VisFeature [36]	github.com/wangjun1996/VisFeature

[a]All websites indicated in the table were accessed in July, 2021.

The last step of the process of building the predictive AI-based model is related to the model assessment. Four sets of metrics are commonly used to evaluate prediction performance. These are sensitivity (Sn), specificity (Sp), balanced accuracy (BACC), and Mattew correlation coefficient (MCC) which can be computed by the following expressions:

$$Sn = \frac{TP}{TP + FN} \quad (1)$$

$$Sp = \frac{TN}{TN + FP} \quad (2)$$

$$BACC = \frac{1}{2}\left(\frac{TP}{TP + FN}\right) + \frac{1}{2}\left(\frac{TN}{TN + FP}\right) \quad (3)$$

$$MCC = \frac{TP * TN - FP * FN}{\sqrt{(TP + FN)(TP + FP)(TN + FP)(TN + FN)}} \quad (4)$$

where TP is the number of positive samples correctly classified in prediction and TN is the number of negative samples correctly classified by predictors. FP and FN represent the number of positive and negative samples misclassified, respectively. The area under the receiver operating characteristic (ROC) curve (AUC) also is calculated to evaluate the performance of models. The range of AUC values is from 0 to 1, and a perfect classifier can be found as AUC=1.0 while the classifier has no discriminative power as AUC=0.5.

4 ARTIFICIAL INTELLIGENCE TECHNIQUES

4.1 Artificial neural networks

Artificial neural networks (ANN) represent powerful machine learning-based techniques. In fact, machine learning (ML), as well as soft computing, is a subfield of AI. Hence, ANNs fall under the umbrella of AI. ANNs can be represented as a computer architecture in which a number of processors (neurons) are interconnected, mimicking the connections between neurons in the human brain, and which is able to learn through trial and error. The most popular type of ANN is Multilayer Perceptron (MLP), also known as feedforward neural network (FNN). ANNs have been applied in life sciences, including in the peptides bioactivity prediction. Li et al. [43] used FNN to predict the bioactivity of fish protein-derived peptides (bighead carp muscles with alcalase) and concluded that ANN can be promising in predicting the biological properties of these peptides. A simple guide on how to apply ANNs to predict the biological activity of AMPs is presented in [44]. It is described how to build and curate datasets, how to train an ANN, and validate the results. The ANN is realized in the software package R. A novel ANN approach, the so-called physics and chemistry-driven artificial neural network (Phys-ChemANN) is proposed by [45]. This approach combines aspects of QSAR (Quantitative structure-activity relationship) and ANN, all its layers, parameters and coefficients have clear physical and chemical implications and provides high accuracy in predicting the bioactivity of proteins and peptides. A standard FNN with a backpropagation learning rule and with a single hidden layer is used in [46] to determine the key parameters in the bioactivity of antifungal peptides. It is proven that α-helical content, hydrophobic interaction, and

charge positivity are the most important parameters in antifungal activity. An ANN approach, based on the AMPs physicochemical characteristics, that is able to identify active peptides but also to assess their antimicrobial potency is presented in [47].

4.2 Deep learning

In recent years, deep learning (DL) architectures such as DNNs, deep belief networks, recurrent neural networks, and convolutional neural networks have become extremely popular. They have been applied successfully to different fields including computer vision, speech recognition, natural language processing, audio recognition, social network filtering, but also in bioinformatics, cheminformatics, drug design, and medical image analysis. The term "deep" in DL means that these structures have multiple hidden layers. There is not a universal definition of when DL begins. It is considered that DL architectures must have at least three nonlinear transformations, i. e. equal or more than two hidden layers and one output layer. DL structures offer the ability to find structure-activity relationships in large sets of heterogeneous data by extracting combinations of linear and nonlinear features [48]. An extensive review that has examined how DL can be applied to make sense of biomedical data is presented in [49]. An adaptive fully-connected DL network that performs cascaded feature extraction from molecular descriptors while finding sparse analytic structure-activity relations is presented in [50]. The adaptability of the network consists of the ability to adjust the number of information processing layers, their size, and their connectivity. In [51] bidirectional recursive neural networks (BRNN) are designed which take as input the protein sequence and a number of predicted structural features of the protein and predict bioactive peptides within the protein sequence. The data for precursor protein sequences on a combination of many classes of bioactive peptides with different functionality, but with shared properties, are retrieved from PeptideDB. A new DNN classifier with convolutional and recurrent layers for AMPs recognition is presented in [52]. This classifier achieved an accuracy of 91.01% on the test dataset, outperforming the state-of-the-art classification models.

4.3 Fuzzy Logic

Fuzzy logic, along with ANNs, is one of the most popular AI techniques. Among the main advantages of FL are its interpretability since it models in a human-like reasoning manner, and its ability to reduce uncertainties. Specifically, for the BAPs prediction, the use of fuzzy logic is quite limited. It is mainly used in combination with other AI techniques. For example, in [53] an adaptive neuro-fuzzy inference system (ANFIS) model is used for AMPs prediction; an improved fuzzy K-nearest neighbor (FKNN) algorithm is employed to establish the two-stage AMPs prediction in [54]. It should be noted that a number of authors consider ANFIS as a deep architecture [53], [55]. As such, it provides high prediction accuracy as reported in [53]. The authors conclude that ANFIS is a more efficient model for pattern recognition in antimicrobial peptide prediction in comparison to ANN, SVM, and RF. Also, the ANFIS methodology is shown to deal effectively with the fuzzy nature of the correlation between physicochemical properties and antimicrobial activity. The ANFIS structure, however, does not have the ability to automatic feature encoding. Traditional fuzzy inference systems (FIS) work with exact membership functions. However, they can be ineffective in a number of real-world problems where insufficient information is available or the model works with incomplete or noisy data. To overcome these problems, there is a need to incorporate additional uncertainty into FISs. Therefore, several extensions of FISs have been introduced such as Interval-valued (Type 2) fuzzy sets and Intuitionistic fuzzy sets (IFSs). They have attracted considerable attention due to their capacity to dealing with imprecise and vagueness concepts. A number of modifications to the ANFIS structure have been developed, including those that use Type 2 fuzzy logic and/or Intuitionistic fuzzy logic [56], [57], [58], [59]. At the time of writing this paper, no sources have been found to report the use of Type 2 or Intuitionistic fuzzy logic in predicting biologically active peptides.

4.4 Support Vector Machine

Support Vector Machine (SVM) is a powerful supervised ML algorithm, used for classification, regression, and prediction analysis. To solve these tasks SVM use two concepts - linear separation with hyperplanes and nonlinear kernel functions. The optimal separation is achieved by converting the inputs into a higher feature order. When the feature set is not distinct or discrete, then the primary non-separable characteristics are changed into a linear separable feature set, by using the kernels. Four classes of kernel functions that often employed for this purpose. These are a linear function, radial function (RBF), polynomial function, and sigmoid function. In the classification of various therapeutic peptides, the most commonly used function is RBF. SVM is widely used in computational biology due to its high accuracy and ability to work with large datasets in higher-dimensional space. Also, the SVM algorithm is computationally simple and effective. An SVM-based meta-predictor for the identification of Anticancer Peptides (ACPs) is presented in [60]. A novel method, a combination of the sequence alignment method, Lempel-Ziv (LZ) complexity, and SVM-pairwise algorithm, for AMPs prediction, is introduced in [61]. Algorithms, using Chou's pseudo amino acid composition (PseAAC) and SVM classifier, for peptides identification are proposed in [62], [63]. An algorithm, which combined orthogonal signal correction with SVM (OSC-SVM), is presented in [64].

4.5 k-Nearest Neighbors

K-Nearest Neighbors algorithm (k-NN) is a classic non-parametric ML method. It is a simple but powerful algorithm that gives very good results. This determines its widespread usage in bioinformatics and cheminformatics. At the same time, for BAPs prediction k-NN has limited application. Current trends show that k-NN is usually used in combination with other ML algorithms [65] and as a basis for comparison with new-designed methods [66].

4.6 Random Forest

Random forest (RF) is an ensemble learning method for classification and regression. Detailed descriptions of the RF algorithm have been provided in [67], [68]. RF model for identifying antiviral peptides (AVPs) is presented in [69]. TargetAntiAngio tool for the prediction and analysis of anti-angiogenic peptides, based on RF classifier is discussed in [70]. In [71] an RF-based method to predict anti-inflammatory peptides is developed. All the above papers reported for high accuracy of the designed models, but the main difficulty is how

to obtain feature vector (descriptor) for inputs of these models.

5 Conclusion

The focus of this article is on food-derived bioactive peptides derived from food. This paper gives some basic information about food-derived BAPs, as well as about the process of the development of AI-based predictive models. It also provides a brief overview of the AI techniques used for BAPs prediction. Comparing the considered methods (ANN, SVM, FL, RF, DL, etc.) it can be concluded, that the DL algorithms outperform all the other AI techniques. Deep learning architectures can make feature extraction automatic and have higher prediction accuracy than the other techniques. At the same time, DL involves much more parameters, which leads to some difficulties during the training process, especially under the circumstances when the samples are not enough or the feature matrix is sparse. In addition, DL algorithms require notable computational resources. Therefore, it is reasonable to prefer such a DL structure in cases where a lot of data is available. In this case, a lot of data should be understood as a huge number of data points consisting of hundreds and even thousands of observations. On the other side, the k-NN and SVM are relatively simple to implement, less powerful than DL and ANN, and require fewer computational resources. Current trends show that k-NN, SVM, and fuzzy logic are usually used in combination with other ML algorithms and as a basis for comparison with new-designed methods.

Acknowledgment

This research presented in this paper is supported by The Bulgarian National Program “Young Scientists and Postdoctoral Students”.

References

[1] Sharma, P., Kaur, H., Kehinde, B., Chhikara, N., Sharma, D., Panghal, A.: Food-Derived Anticancer Peptides: A Review. International Journal of Peptide Research and Therapeutics, 1-16 (2020).

[2] Patil, P., Mandal, S., Tomar, S., Anand, S.: Food protein-derived bioactive peptides in management of type 2 diabetes. European journal of nutrition, 54(6), 863-880 (2015).

[3] Daliri, E. B. M., Oh, D. H., Lee, B. H. Bioactive peptides. Foods, 6(5), 32 (2017).

[4] Imai, K., Ji, D., Nwachukwu, I., Agyei, D., Udenigwe, C.: Bioinformatics and Chemometrics for Discovering Biologically Active Peptides From Food Proteins (2019).

[5] Malaguti, M., Dinelli, G., Leoncini, E., Bregola, V., Bosi, S., Cicero, A., Hrelia, S.: Bioactive peptides in cereals and legumes: agronomical, biochemical and clinical aspects. International journal of molecular sciences, 15(11), 21120-21135 (2014).

[6] Sanchez, A., Vazquez, A.: Bioactive peptides: A review. Food Quality and Safety, 1(1), 29-46 (2017).

[7] Bhat, Z., Kumar, S., Bhat, H.: Bioactive peptides from egg: a review. Nutrition & Food Science (2015).

[8] Anusha, R., Bindhu, O.: Bioactive peptides from milk. MILK PROTEINS, 101 (2016).

[9] Ryan, J., Ross, R., Bolton, D., Fitzgerald, G., Stanton, C.: Bioactive peptides from muscle sources: meat and fish. Nutrients, 3(9), 765-791 (2011).

[10] Kamran, F., Reddy, N.: Bioactive peptides from legumes: Functional and nutraceutical potential. Recent Advances in Food Science, 1(3), 134-149 (2018).

[11] Le Gouic, A., Harnedy, P., FitzGerald, R.: Bioactive peptides from fish protein by-products. Bioactive molecules in food. Cham: Springer International Publishing, 1-35 (2018).

[12] Pihlanto, A.: Lactic fermentation and bioactive peptides. Lactic Acid Bacteria—R & D for Food, Health and Livestock Purposes, 309-332 (2013).

[13] Caron, J., Cudennec, B., Domenger, D., Belguesmia, Y., Flahaut, C., Kouach, M., Lesage, J., Goossens, J., Dhulster, P., Ravallec, R.: Simulated GI digestion of dietary protein: Release of new bioactive peptides involved in gut hormone secretion. Food Research International, 89, 382-390 (2016).

[14] Brodkorb, A., Egger, L., Alminger, M., et al.: INFOGEST static in vitro simulation of gastrointestinal food digestion. Nature protocols, 14(4), 991-1014 (2019).

[15] Balgir, P., Kaur, T., Sharma, M.: Antihypertensive peptides derived from food sources. MOJ Food Processing & Technology, 2(1), 1-6 (2016).

[16] Yamaguchi, N., Kawaguchi, K., Yamamoto, N.: Study of the mechanism of antihypertensive peptides VPP and IPP in spontaneously hypertensive rats by DNA microarray analysis. European journal of pharmacology, 620(1-3), 71-77 (2009).

[17] Sipola, M., Finckenberg, P., Santisteban, J., Korpela, R., Vapaatalo, H., Nurminen, M.: Long-term intake of milk peptides attenuates development of hypertension in spontaneously hypertensive rats. J Physiol Pharmacol, 52(4 Pt 2), 745-754 (2001).

[18] Haney, E., Straus, S., Hancock, R.: Reassessing the host defense peptide landscape. Frontiers in chemistry, 7, 43 (2019).

[19] Pizzo, E., Cafaro, V., Di Donato, A., Notomista, E.: Cryptic Antimicrobial Peptides: identification methods and current knowledge of their immunomodulatory properties. Current pharmaceutical design, 24(10), 1054-1066 (2018).

[20] Udenigwe, C.: Bioinformatics approaches, prospects and challenges of food bioactive peptide research. Trends in Food Science & Technology, 36(2), 137-143 (2014).

[21] Pellegrini, A.: Antimicrobial peptides from food proteins. Current pharmaceutical design, 9(16), 1225-1238 (2003).

[22] Mohanty, D., Jena, R., Choudhury, P., Pattnaik, R., Mohapatra, S., Saini, M.: Milk derived antimicrobial bioactive peptides: a review. International Journal of Food Properties, 19(4), 837-846 (2016).

[23] Wong, F., Xiao, J., Wang, S., Ee, K., Chai, T.: Advances on the antioxidant peptides from edible plant sources. Trends in Food Science & Technology (2020).

[24] Nwachukwu, I., Aluko, R.: Structural and functional properties of food protein-derived antioxidant peptides. Journal of Food Biochemistry, 43(1), e12761 (2019).

[25] Najafian, L., Babji, A.: A review of fish-derived antioxidant and antimicrobial peptides: their production, assessment, and applications. Peptides, 33(1), 178-185 (2012).

[26] Zou, T., He, T., Li, H., Tang, H., Xia, E.: The structure-activity relationship of the antioxidant peptides from natural proteins. Molecules, 21(1), 72 (2016).

[27] Basith, S., Manavalan, B., Hwan Shin, T., Lee, G.: Machine intelligence in peptide therapeutics: A next-generation tool for rapid disease screening. Medicinal research reviews (2020).

[28] Panyayai, T., Ngamphiw, C., Tongsima, S., Mhuantong, W., Limsripraphan, W., Choowongkomon, K.,

Sawatdichaikul, O.: FeptideDB: A web application for new bioactive peptides from food protein. Heliyon, 5(7), e02076 (2019).
[29] Minkiewicz, P., Iwaniak, A., Darewicz, M.: BIOPEP-UWM database of bioactive peptides: Current opportunities. International journal of molecular sciences, 20(23), 5978 (2019).
[30] Li, Q., Zhang, C., Chen, H., Xue, J., Guo, X., Liang, M., Chen, M.: BioPepDB: An integrated data platform for food-derived bioactive peptides. International Journal of Food Sciences and Nutrition, 69(8), 963-968 (2018).
[31] Kumar, R., Chaudhary, K., Sharma, M., Nagpal, G., Chauhan, J., Singh, S., Gautam, A., Raghava, G.: AHTPDB: a comprehensive platform for analysis and presentation of antihypertensive peptides. Nucleic acids research, 43(D1), D956-D962 (2015).
[32] Nielsen, S. D., Beverly, R. L., Qu, Y., & Dallas, D. C.: Milk bioactive peptide database: A comprehensive database of milk protein-derived bioactive peptides and novel visualization. Food Chemistry, 232, 673-682 (2017).
[33] Chen Z, Zhao P, Li F, et al. iFeature: a Python package and web server for features extraction and selection from protein and peptide sequences. Bioinformatics. 2018;34(14):2499-2502.
[34] Pande, A., Patiyal, S., Lathwal, A., et al.: Computing wide range of protein/peptide features from their sequence and structure. bioRxiv, 599126 (2019).
[35] Chen Z, Zhao P, Li F, et al. iLearn: an integrated platform and meta-learner for feature engineering, machine-learning analysis and modeling of DNA, RNA and protein sequence data. Brief Bioinform. 2019.
[36] Wang J, Du PF, Xue XY, et al. VisFeature: a stand-alone program for visualizing and analyzing statistical features of biological sequences. Bioinformatics. 2019.
[37] Muhammod R, Ahmed S, Md Farid D, Shatabda S, Sharma A, Dehzangi A. PyFeat: a Python-based effective feature generation tool for DNA, RNA and protein sequences. Bioinformatics. 2019;35(19):3831-3833.
[38] Cao D-S, Xu Q-S, Liang Y-Z. Propy: a tool to generate various modes of Chou's PseAAC. Bioinformatics. 2013;29(7): 960-962.
[39] Liu B, Liu F, Wang X, Chen J, Fang L, Chou KC. Pse-in-One: a web server for generating various modes of pseudo components of DNA, RNA, and protein sequences. Nucleic Acids Res. 2015;43(W1):W65-W71.
[40] Liu B, Wu H, Chou K-C. Pse-in-One 2.0: an improved package of web servers for generating various modes of pseudo components of DNA, RNA, and protein sequences. Nat Sci. 2017;9(04):67-91.
[41] Nikam R, Gromiha MM. Seq2Feature: a comprehensive web-based feature extraction tool. Bioinformatics. 2019;35: 4797-4799.
[42] Dong J, Yao ZJ, Zhang L, et al. PyBioMed: a python library for various molecular representations of chemicals, proteins and DNAs and their interactions. J Cheminform. 2018;10(1):16.
[43] Li, L., Wang, J., Zhao, M., Cui, C., Jiang, Y.: Artificial Neural Network for Production of Antioxidant Peptides Derived from Bighead Carp Muscles with Alcalase. Food Technology & Biotechnology, 44(3) (2006).
[44] Andreu, D., Torrent, M.: Prediction of bioactive peptides using artificial neural networks. In Artificial Neural Networks (pp. 101-118). Springer, New York, NY (2015).
[45] Huang, R., Du, Q., Wei, Y., Pang, Z., Wei, H., Chou, K.: Physics and chemistry-driven artificial neural network for predicting bioactivity of peptides and proteins and their design. Journal of theoretical biology, 256(3), 428-435 (2009).
[46] Soltani, S., Keymanesh, K.: Evaluation of structural features of membrane acting antifungal peptides by artificial neural networks (2008).
[47] Torrent, M., Andreu, D., Nogues, V., Boix, E.: Connecting peptide physicochemical and antimicrobial properties by a rational prediction model. PloS one, 6(2), e16968 (2011).
[48] Roudi, Y., Taylor, G.: Learning with hidden variables. Current opinion in neurobiology, 35, 110-118 (2015).
[49] Baldi, P.: Deep learning in biomedical data science. Annual review of biomedical data science, 1, 181-205 (2018).
[50] Müller, A., Kaymaz, A., Gabernet, G., Posselt, G., Wessler, S., Hiss, J., Schneider, G.: Sparse Neural Network Models of Antimicrobial Peptide-Activity Relationships. Molecular informatics, 35(11-12), 606-614 (2016).
[51] Mooney, C., Haslam, N., Holton, T., Pollastri, G., Shields, D.: PeptideLocator: prediction of bioactive peptides in protein sequences. Bioinformatics, 29(9), 1120-1126 (2013).
[52] Veltri, D., Kamath, U., Shehu, A.: Deep learning improves antimicrobial peptide recognition. Bioinformatics, 34(16), 2740-2747 (2018).
[53] Fernandes, F., Rigden, D., Franco, O.: Prediction of antimicrobial peptides based on the adaptive neuro-fuzzy inference system application. Peptide Science, 98(4), 280-287 (2012).
[54] Xiao, X., Wang, P., Lin, W., Jia, J., Chou, K.: iAMP-2L: a two-level multi-label classifier for identifying antimicrobial peptides and their functional types. Analytical biochemistry, 436(2), 168-177 (2013).
[55] S.-J. Heo, Z. Chunwei, and E. Yu, ``Response simulation, data cleansing and restoration of dynamic and static measurements based on deep learning algorithms," Int. J. Concrete Struct. Mater., vol. 12, no. 1, p. 82, Dec. 2018.
[56] Terziyska, M., Todorov, Y., Doneva, M., & Metodieva, P. (2019). Distributed Adaptive Neuro Intuitionistic Fuzzy Architecture for prediction of the dose in gamma irradiated milk products. IFAC-PapersOnLine, 52(25), 75-80.
[57] Terziyska, M., Todorov, Y., & Olteanu, M. (2016, June). Input space selective fuzzification in intuitionistic semi fuzzy-neural network. In 2016 8th International Conference on Electronics, Computers and Artificial Intelligence (ECAI) (pp. 1-7). IEEE.
[58] Todorov, Y., & Terziyska, M. (2014, September). Modeling of chaotic time series by interval type-2 neo-fuzzy neural network. In International Conference on Artificial Neural Networks (pp. 643-650). Springer, Cham.
[59] Terziyska, M., & Todorov, Y. (2016, September). Intuitionistic Neo-Fuzzy Network for modeling of nonlinear systems dynamics. In 2016 IEEE 8th International Conference on Intelligent Systems (IS) (pp. 616-621). IEEE.
[60] Boopathi, V., Subramaniyam, S., Malik, A., Lee, G., Manavalan, B., Yang, D.: mACPpred: a support vector machine-based meta-predictor for identification of anticancer peptides. International journal of molecular sciences, 20(8) (2019).
[61] Ng, X., Rosdi, B., Shahrudin, S.: Prediction of

antimicrobial peptides based on sequence alignment and support vector machine-pairwise algorithm utilizing LZ-complexity. BioMed research international, (2015).
[62] Mousavizadegan, M., Mohabatkar, H.: Computational prediction of antifungal peptides via Chou's PseAAC and SVM. Journal of bioinformatics and computational biology, 16(04), 1850016 (2018).
[63] Meher, P., Sahu, T., Saini, V., Rao, A.: Predicting antimicrobial peptides with improved accuracy by incorporating the compositional, physico-chemical and structural features into Chou's general PseAAC. Scientific reports, 7(1), 1-12 (2017).
[64] Guan, X., Liu, J.: QSAR study of angiotensin i-converting enzyme inhibitory peptides using svhehs descriptor and osc-svm. International Journal of Peptide Research and Therapeutics, 25(1), 247-256 (2019).
[65] Li, B., Zhang, Y., Jin, M., Huang, T., Cai, Y.: Prediction of protein-peptide interactions with a nearest neighbor algorithm. Current Bioinformatics, 13(1), 14-24 (2018).
[66] Wang, L., Niu, D., Wang, X., Shen, Q., Xue, Y.: A Novel Machine Learning Strategy for Prediction of Antihypertensive Peptides Derived from Food with High Efficiency. BioRxiv (2020).
[67] Lee, J., Lee, K., Joung, I., Joo, K., Brooks, B., Lee, J.: Sigma-RF: prediction of the variability of spatial restraints in template-based modeling by random forest. BMC bioinformatics, 16(1), 94 (2015).
[68] Manavalan, B., Lee, J., Lee, J.: Random forest-based protein model quality assessment (RFMQA) using structural features and potential energy terms. PloS one, 9(9), e106542 (2014).
[69] Chang, K., Yang, J.: Analysis and prediction of highly effective antiviral peptides based on random forests. PloS one, 8(8), e70166 (2013).
[70] Laengsri, V., Nantasenamat, C., Schaduangrat, N., Nuchnoi, P., Prachayasittikul, V., Shoombuatong, W.: TargetAntiAngio: A sequence-based tool for the prediction and analysis of anti-angiogenic peptides. International journal of molecular sciences, 20(12), 2950 (2019).
[71] Manavalan, B., Shin, T., Kim, M., Lee, G., "AIPpred: sequence-based prediction of anti-inflammatory peptides using random forest." Frontiers in pharmacology, 9, 276 (2018).