ISSN 2074-9414 (Print)
ISSN 2313-1748 (Online)

https://doi.org/10.21603/2074-9414-2022-1-46-57

Review article
https://fptt.ru/en

Hybrid Strategy of Bioinformatics Modeling (*in silico*): Biologically Active Peptides of Milk Protein

Alexandr G. Kruchinin*, Ekaterina I. Bolshakova

All-Russian Dairy Research Institute, Moscow, Russia

Received: 10.01.2022
Revised: 31.01.2022
Accepted: 14.02.2022

**e-mail: a_kruchinin@vnimi.org*

Abstract.
Bioinformatic analysis methods are an auxiliary tool in the preliminary stage of research into biocatalytic conversion of proteins with predicted release of biologically active peptides. However, there are a number of factors ignored in current strategies for designing biologically active peptides, which prevents the complete prediction of their biological properties. This determines the relevance of the research objective, i.e. developing a hybrid strategy for bioinformatic modeling to study biologically active peptides of milk protein. The new strategy ranks key criteria based on high-performance algorithms of proteomic database.
The research featured the scientific publications on *in silico* methods applied to biologically active peptides. Modern taxonometric methods of information retrieval were applied using the RSCI, Scopus and Web of Science databases.
The article introduces and describes step by step the optimal in silico hybrid strategy algorithm for studying biologically active milk protein peptides. The algorithm takes into account the safety assessment of all hydrolysis products, their physicochemical and technological properties. The strategy algorithm relies on analytical data on the protein profile, the amino acid sequence of proteins that make up the raw material, taking into account their polymorphism, and the subsequent identification of bioactive amino acid sites in the protein structure. The algorithm selects optimal enzyme preparations, as well as models the hydrolysis and assesses the peptide bioactivity using proteomic databases.
At the preliminary stage of protein hydrolysis, the new in silico strategy scientifically predicts the targeted release of stable peptide complexes of biologically active peptides with proven bioactivity, safety and sensory characteristics. The hybrid algorithm contributes to accumulation of the necessary primary data so as to reduce the time and cost of laboratory experiments.

Keywords. Milk proteins, peptides, database, bioinformatics, *in silico*

Funding. The research was supported by the Russian Science Foundation (RSF), grant No. 21-76-00044.

For citation: Kruchinin AG., Bolshakova EI. Hybrid Strategy of Bioinformatics Modeling (*in silico*): Biologically Active Peptides of Milk Protein. Food Processing: Techniques and Technology. 2022;52(1):46–57. https://doi.org/10.21603/2074-9414-2022-1-46-57

https://doi.org/10.21603/2074-9414-2022-1-46-57

Обзорная статья
https://fptts.ru

Гибридная стратегия биоинформатического моделирования (*in silico*) для изучения биологически активных пептидов молочного белка

А. Г. Кручинин*, Е. И. Большакова

Всероссийский научно-исследовательский институт молочной промышленности, Москва, Россия

Поступила в редакцию: 10.01.2022
Поступила после рецензирования: 31.01.2022
Принята в печать: 14.02.2022

**e-mail: a_kruchinin@vnimi.org*

Аннотация.
Методы биоинформатического анализа – вспомогательный инструмент в проведении предварительного этапа исследований процесса биокаталитической конверсии белков с прогнозируемым высвобождением биологически активных пептидов. Однако существует ряд факторов, не учитывающихся в современных стратегиях при проектировании биологически активных пептидов, что препятствует полномасштабному прогнозированию их биологических свойств. Это обуславливает актуальность выбранной цели исследования – разработку гибридной стратегии биоинформатического моделирования для изучения биологически активных пептидов молочного белка с учетом ранжирования ключевых критериев на основе высокопроизводительных алгоритмов протеомных баз данных.
Объектом исследования является научная литература, касающаяся методов *in silico* биологически активных пептидов. Применялись современные таксонометрические методы поиска информации с использованием баз данных РИНЦ, Scopus и Web of Science.
Сформирован и поэтапно описан оптимальный алгоритм гибридной стратегии *in silico* изучения биологически активных пептидов молочного белка с учетом оценки безопасности всех продуктов гидролиза, их физико-химических и технологических свойств. Алгоритм стратегии сформирован исходя из аналитических данных о белковом профиле, аминокислотной последовательности белков, входящих в состав сырья с учетом их полиморфизма, и последующей идентификации биоактивных аминокислотных сайтов в структуре белка. В алгоритм включен подбор оптимальных ферментных препаратов и моделирование гидролиза с оценкой биоактивности пептидов по протеомным базам данных. Предложенная стратегия *in silico* позволит на предварительном этапе проведения гидролиза белка научно прогнозировать направленное высвобождение стабильных пептидных комплексов биологически активных пептидов с доказанными биоактивностью, безопасностью и сенсорными характеристиками. Гибридный алгоритм будет способствовать аккумулированию необходимых первичных данных для сокращения временных и финансовых затрат на проведение реальных экспериментов.

Ключевые слова. Молочные белки, пептиды, база данных, биоинформатика, *in silico*

Финансирование. Исследование выполнено за счет гранта Российского научного фонда (РНФ) № 21-76-00044.

Для цитирования: Кручинин А. Г., Большакова Е. И. Гибридная стратегия биоинформатического моделирования (*in silico*) для изучения биологически активных пептидов молочного белка // Техника и технология пищевых производств. 2022. Т. 52. № 1. С. 46–57. (На англ.). https://doi.org/10.21603/2074-9414-2022-1-46-57

Introduction

Recent years have seen an increase in the number of biotechnological studies aimed at assessing the role of biologically active peptides derived from food raw materials for regulating body functions, maintaining immunological status, and reducing the risk of chronic disease development [1, 2]. Scientists proved that biologically active peptides demonstrate antimicrobial, hypocholestermic, antihypertensive, antioxidant, antidiabetic, immunomodulatory and otherproperties [3–9]. PeptidOme of dairy raw materials is considered one of the most valuable sources for isolating bioactive peptides encoded in its structure [10]. Most biologically active peptides identified in dairy products range from 2 to 20 amino acids in length. This corresponds to a molecular weight range of 0.24–2.50 kDa. As the length of the peptide increases, the probability of forming secondary structure elements rises, which results in steric hindrances to the manifestation of various biological activities. Exposure to proteases brings about the release of bioactive peptides from the amino acid sequence of a protein. This exposure takes place during gastrointestinal digestion, fermentation of milk proteins using proteolytic systems of lactic acid bacteria in the process of ripening, technological treatment of raw materials (homogenization, high temperature treatment, ultrasound, etc.) and bio-conversion of protein raw materials with purified preparations of proteolytic enzymes [11–13].

The classical strategy for research of biologically active peptides relies on the unpredictable cleavage of peptide bonds in the protein structure by proteases *in vitro*, followed by the purification of hydrolysis products and evaluation of their bioactivity *in vivo*. However, this strategy suffers from a number of shortcomings, including a high labor intensity and a long process, as well as high financial costs [14]. With computer technology and in-depth analytical research methods developing rapidly, integrated proteomic data banks, such as NCBI, BIOPEP, UniProt, PepBank, SwePep, etc. were created. Implementing bioinformatic analysis algorithms on these platforms allows the detection of peptide bonds in the protein structure sensitive to proteolytic cleavage, amino acid sequences of proteins and derived peptides, their functionality, allergenicity, chelating ability, etc. [15–17].

Methods of bioinformatic analysis (*in silico*) are an auxiliary tool in preliminary studying the biocatalytic conversion of proteins (using "digital twin" models) by different proteases with predicted release of biologically active peptides. Since peptides, like proteins, exhibit a high degree of structure-activity relationship, the presence and location of certain amino acid residues (biomarkers) can indicate the properties and potential bioactivity of peptides [18]. For example, E.Yu. Agarkova and A.G. Kruchinin showed in their article that redox-active amino acid residues (C, H, Y, W and M) are an important structural descriptor of antioxidant peptides [19]. Residues of hydrophobic amino acid enhance the antioxidant properties of peptides in systems containing the lipid phase. Amino acids with ionogenic groups in side radicals are responsible for binding metal ions of variable valence. Thus, predictive modeling of biological activities in peptides based on biomarkers reduces the number and duration of experiments to obtain representative data [18]. Bioinformatic analysis integrated into research developed new strategies for discovering bioactive peptides and proving their role at the organismic level. Most *in silico* working strategies are based on a paradigm that selects protein substrate and enzymes to generate bioactive peptides (taking into account the frequency and release efficiency criteria), carry out molecular docking, and screen virtually peptide sequences for further optimization of biopeptide release from food protein substrates [20, 21].

However, the design and generation of biologically active peptides neglect a number of factors. For example, the genetic polymorphism of milk proteins associated with amino acid mutations in its structure can affect the type and biological activity of the released peptides [22]. Diversity of the protein matrix of food raw materials should be considered another important factor, as well as their bioavailability for enzymatic cleavage, taking into account the conformational and intermolecular changes during technological processing. Considering peptidomics as an integral part of fudomics, one should pay special attention to predicting the sensory characteristics of hydrolysis products, aim to minimize the formation of free amino acids at the *in silico* stage, as well as level out the formation of bitterness and non-specific flavor as much as possible. A key criterion in the development and identification of biologically active peptides is food safety. That is why a bioinformatic approach to modeling biologically active peptides should predict such factors as toxicity and allergenicity of the peptides released from the protein structure. In terms of technological properties, an important factor is predicting the stability of biologically active peptides during *in silico* modeling. Bioinformatics can predict the average molecular weight, thermal stability (aliphatic index), solubility (hydropathy index), etc. This enables assessment of stability for hydrolysis products during further technological processing and storage. Since bioactive peptides can be completely or partially degraded by digestive proteases in the human gastrointestinal tract and subsequently lose biological activity,

bioinformatic modeling of the resistance of bioactive peptides to hydrolysis by digestive enzymes is considered an important part of the final stage. For example, proline in biologically active peptides increases their resistance to GI peptidases [19].

The foregoing determines the relevance of the study objective, i.e. developing a hybrid strategy for bioinformatic modeling so as to study biologically active peptides of milk protein, taking into account the ranking of key criteria based on high-performance proteomic database algorithms.

Study objects and methods

Analysis embraced Russian and foreign scientific publications dealing with the use of bioinformatic data banks in studying proteins or peptides of food biosystems. It was carried out on the main scientometric databases RSCI, Scopus and Web of Science. The search query excluded teaching materials, as well as conference materials and proceedings. Search descriptors in article titles, keywords, and abstracts included the following words and phrases: food proteins, bioactive peptides, database, bioinformatics, *in silico*. The depth of analysis for scientific publications was limited to a 20-year period. This approach allowed us to identify key actualizable databases and form the fundamental criteria for bioinformatic modeling of targeted hydrolysis of food proteins in order to predict the release of biopeptides from their structures.

Results and discussion

Resultant from the development of principles for the bioinformatic approach in peptidomics, numerous databases were created, including data banks of proteins, as well as enzymes, sensory, allergenic, bioactive and hypothetically bioactive peptides. In addition to listing members of each group, the databases contain associated analytical bioinformatics tools. Thanks to them, one can extract information about the dis-/similarity of given protein structures, their amino acid sequence, theoretical enzymatic cleavage, physicochemical properties, chelating ability, proven or predicted functionality, allergenicity, toxicity, etc.

In a number of studies, scientists used various bioinformatic resources successfully to create algorithms and strategies for predicting the isolation of biologically active peptide from food raw materials [23–25]. Taking into account the characteristics of raw materials or the process of generating biologically active peptides, the authors point out that each individual food object requires appropriate *in silico* modeling tools.

Analysis and systematization of international experience resulted in development and thorough description of an optimal algorithm for a hybrid strategy of bioinformatic modeling so as to study biologically active peptides of milk protein. The strategy takes into account the most significant criteria that increase the probability of obtaining peptides with predictable bioactivity, safety, and acceptable sensory characteristics (Fig. 1).

Protein profile analysis of milk raw materials. The fractional composition of raw milk is not constant and depends on paratypical (period of the year, feeding ration, lactation period, animal health, etc.), genotypical (heredity, breed, individual genotype, etc.) and technological (heat treatment, homogenization, membrane processing, etc.) factors [26]. In this regard, the preliminary proteomic studies require qualitative and quantitative determination of protein fractions for dairy raw materials due to their instability. To determine the total content of casein and serum proteins and to identify protein fractions, one needs to use a set of multi-directional techniques, such as the Kjeldahl method, one- or two-dimensional gel electrophoresis with isoelectric focusing, high-performance liquid chromatography, etc. In addition, high-performance liquid chromatography with time-of-flight mass spectrometry will assess changes in the peptide profile in dairy raw materials depending on various technological factors.

Thus, complete systematic mapping of proteins in dairy raw materials, taking into account the conformational and proteomic changes associated with the technological features of modern production, seems to be a powerful tool at the initial stage of the bioinformatic modeling strategy.

Protein amino acid sequence analysis taking into account genetic polymorphism. The next stage of the strategy involves obtaining data on the amino acid sequences of all protein fractions identified in the composition of raw milk. Data on the amino acid sequence, including the protein gene polymorphism (if necessary), its codifiers, molecular weight, and source, can be retrieved from bioinformatic databases and associated tools: NCBI, Uniprot and BIOPEP [27]. These resources are often used to identify the amino acid sequences of proteins while studying *in silico* new bioactive peptides from animal raw materials and creating databases of sensory peptides [25, 28, 29]. However, *in silico* studies do not take into account information about the genetic variability of protein structures.

The polymorphism of the gene, encoding the amino acid sequence in the protein structure, plays an essential role in the strategy for bioinformatic modeling of enzymatic bioconversion of milk proteins. Amino acid mutations result in the random

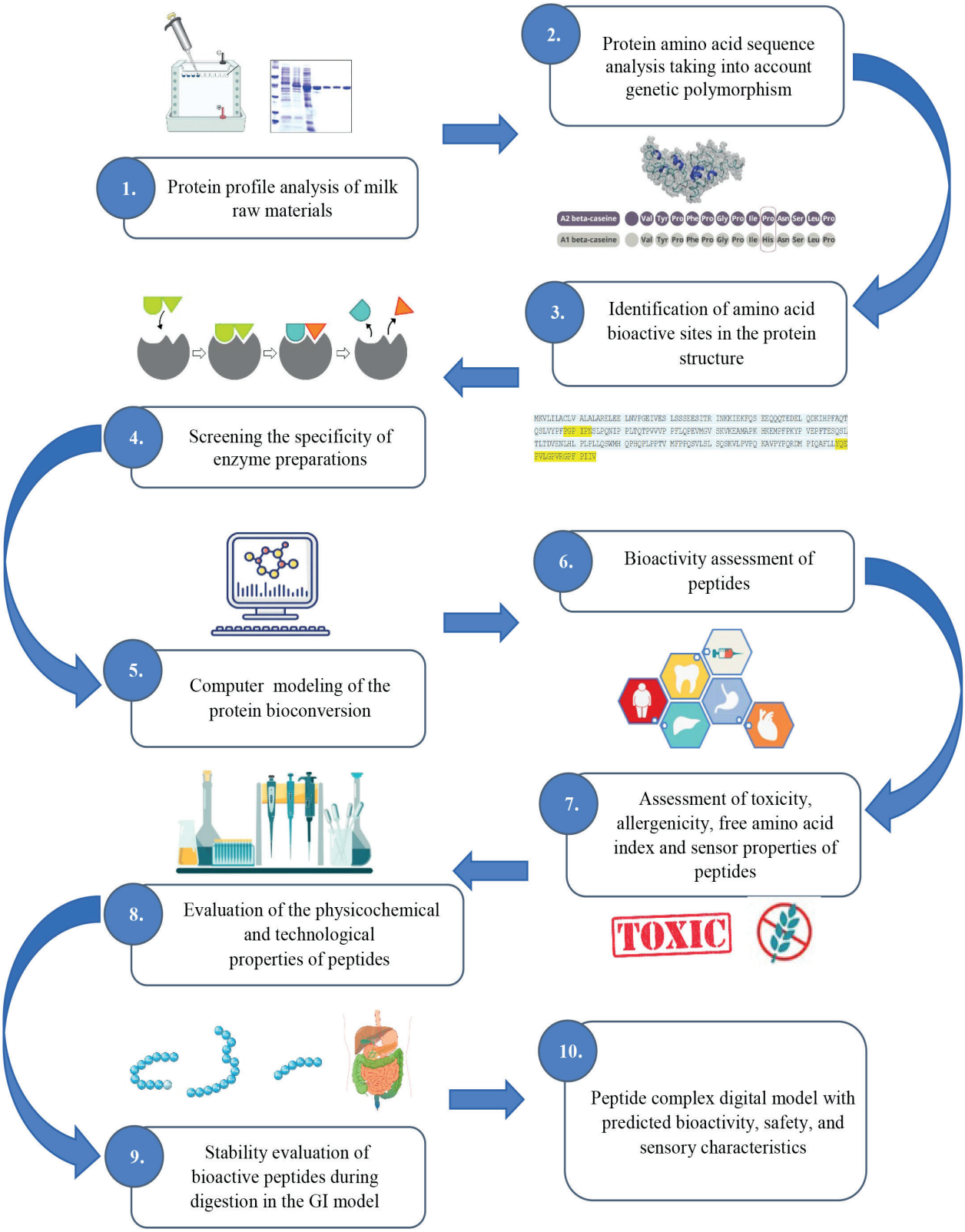

Figure 1. Hybrid strategy algorithm of bioinformatic *in silico* modeling to be used in research on biologically active peptides of milk protein

replacement of single amino acids in the protein structure, which affects its properties as well as the bioactivity and degree of peptide release. The effects of gene polymorphism on the amino acid sequence have been noted in a number of studies and constitute a proven fact [30, 31]. Researchers at the University of Limerick stated that the genetic polymorphism of dairy proteins in raw milk obtained from producing animals of the same breed affects the types of bioactive peptides it contains [24]. The direction of hydrolysis can also depend on the genetic variation of the protein. This effect has been mentioned in the study of polymorphic variants of *β*-casein and their effect on digestion in the GI tract *ex vivo* [32]. Consequently, when modeling the targeted hydrolysis of milk protein raw materials, it is necessary to take into account their genotypic traits because they can determine the direction of hydrolysis and the composition of bioactive sites within the protein structure.

The fact that dairy plants receive milk from farms in a bulk milk tank (mixed) poses the main problem for genetic identification of expressed protein fractions in raw milk. Milk collected from different cows is characterized by heterogeneity of genetic variants of a certain protein, which complicates its controlled bioconversion. The laboratory of canned milk at the All-Russian Dairy Research Institute has developed a modern technique for molecular genetic evaluation of the ratio of relative shares of the *CSN3* gene alleles in mixed dairy products [33]. Based on the proposed technique, the authors developed a bioinformatic analysis program *Calculating the ratio of the relative proportions of κ-casein alleles in collected milk*, available at www.tinyurl.com/allelesprog. Improving this technique and projecting it onto other biotechnologically relevant protein fractions will allow integration of this tool into the strategy of bioinformatic modeling (*in silico*) from the position of rational processing milk raw materials for the predicted release of biologically active peptides.

Identification of amino acid bioactive sites in the protein structure. A key step in *in silico* modeling of hydrolysis is identifing locations of bioactive sites encoded in the amino acid sequences of protein substrates, taking into account genetic polymorphism with the aim of their further targeted release. The evaluation criterion is the frequency of bioactive sites occurrence in the protein structure. Bioactive peptides within the amino acid structure of a protein may be searched by its identifier using the bioinformatic database tools MBPDB and BIOPEP [34]. Bioinformatic algorithms of these databases are able to perform a search query in the following variations: searching for bioactive peptides in the structure of a particular protein; searching for a specific amino acid sequence and assessing homology of biofunctional properties, as well as identifying precursor proteins [35, 36]. The resultant set contains data of bioactive peptides with annotated amino acid sequences included in the studied protein (peptide mapping), their functions, level of bioactivity, and references to primary sources of research data. The data set allows one to simplify the process and reduce labor costs of releasing bioactive peptides from complex protein matrices [37, 38]. The targeted hydrolysis will result in the release of not only the maximum possible number of functional peptides, but also those whose bioactivity is not annotated. The bioinformatic tools BLAST NCBI, Expasy SIM Alignment Tool and Uniprot (ALIGN) are used to compare amino acid sequences (alignment) in order to identify protein structures similar in motifs and functionality [39]. It is worth noting that working with these resources requires care in formulating conclusions. R.A. González-Pech *et al.* have drawn attention to cases of incorrect interpretation of the data obtained through these algorithms [40].

Most other tools used for identifying bioactive peptides, such as APD, PeptideDB, BioPepDB, etc., operate on the basis of an inverse algorithm [41, 42]. This algorithm focuses on the amino acid sequences of peptides whose isolation from the protein requires prior use of resources modelling enzymatic cleavage. This approach forms many options for directing the hydrolysis, since enzyme complexes or individual enzyme preparations will have an individual bioinformatic scheme of cleavage. Processing such a data set implies a time cost, provided that there are no limitations in the number of enzyme systems. A number of publications on *in silico* studies of protein microstructures of collagens, tomato seeds, mung beans, etc. also used this classical algorithm – from enzymatic cleavage to evaluation of peptide properties [43–45].

Screening the specificity of enzyme preparations. The task of the next stage of the bioinformatic modeling strategy is to screen the specificity of enzyme preparations taking into account the hydrolysable peptide bonds at the sites of bioactive peptides. The bioinformatic tool Expasy Peptide Cutter extracts information about the enzymes appropriate for selected protein substrates and indicates the hydrolysable peptide bond between amino acids. Using this information, BIOPEP's "Batch Processing" provides a list of selected amino acid sequences and a list of bioactive peptides included in it.

Enzymatic screening can also be performed with another BIOPEP tool, "Find the enzyme for peptide

release", where the raw data are bioactive peptides and the amino acid sequence of the protein from which they are to be extracted. It is important to enter peptides in FASTA format as follows: "> peptide 1 IPP (amino acid sequence of bioactive peptide)". There can be several peptides, and each must be specified with a new line and a new number. The result of the data processing is a list of enzymes suitable for targeted hydrolysis.

Computer modeling of the protein bioconversion. After suitable enzymes are selected in this way, all enzymatic cleavage products can be analyzed in BIOPEP's section "Enzymes action" by selecting the option "Enzymes action for your sequence". This tool features the complete picture of protein hydrolysis into peptides. Even taking into account the poly-enzyme system. Computer modeling of bioconversion should be performed on a "digital twin" model of the substrate. A digital twin is formed from the analytical data on the protein profile of the raw milk used. Bioconversion modeling is carried out for each protein fraction, after which the hydrolysis products are combined and analyzed. The only drawback of this scheme is that this tool does not take into account the hydrolysis conditions, namely temperature, duration, substrate-enzyme ratio and pH. This offers the basis for studies to optimize the conditions of enzymatic hydrolysis, taking into account technological factors *in vitro*.

Bioactivity assessment of peptides. After targeted hydrolysis on the "digital twin" model of the complex protein matrix of dairy raw materials with enzymes selected after screening, all reaction products should be evaluated for biofunctionality by means of tools. They are listed in "Identification of Bioactive Amino Acid Sites in Protein Structure". In addition to the described bioinformatic resources used to determine the bioactivity of peptides, another tool, Peptide Ranker, is worth mentioning. In the study by S. Nebbia *et al.*, it helped select only 10 out of 30 000 prognostically formed peptides for further study [35]. This resource identifies the biological activity of peptides according to certain structural characteristics on a scale from 0 to 1, in which any peptide scoring above 0.5 is considered biologically active [44, 46]. Using this tool Y. Gu *et al.* evaluated the effect of different types of cultures on the peptide profile of yogurts [47]. M. Tu *et al.* studied biologically active peptides derived from casein hydrolysis [48]. In addition, there are a number of narrowly focused databases that will help in the targeted search for bioactivity. Among such databases, MilkAMP (antimicrobial peptide database), AHTPDB (antihypertensive peptide database), etc. stand out.

Assessment of toxicity, allergenicity, free amino acid index and sensor properties of peptides. Since one of the main objectives of biotechnology is to ensure the safety of isolated substances, a necessary step consists in testing peptides obtained by targeted hydrolysis for adverse effects.

According to the publications, there are approximately 170 food allergens that cause IgE-mediated allergic reactions. 90% of these reactions are caused by food allergens representing 8 groups, including milk and dairy products [49, 50]. Almost all milk proteins are immunoreactive due to a large number of antigenic determinants (epitopes) in their amino acid sequences [51, 52]. On this basis, a prerequisite for *in silico* analysis is to predict the residual antigenicity of all hydrolysis products. It is possible by means of IUIS and BIOPEP databases containing up-to-date information on allergenic protein epitopes. In addition to the search systems of these two bases, there are bioinformatic tools such as Allergenic Protein Sequence Searches and AlgPred2 [53]. They help predict the allergenicity of isolated peptides and the protein as a whole by amino acid sequence. To perform alignment, AlgPred2 is paired with IEDB, which is a database of experimental data on antibody epitopes studied in the context of infectious diseases, allergy, autoimmunity and transplantation, as well as with the NCBI BLAST tool. It is also coupled with the MERCI software to identify allergenic sites in the protein structure [54].

Bioinformatic data on the allergenicity of protein microstructures will allow correcting the hydrolysis process by changing the proteolytic system or adding a second hydrolysis step to break down allergenic sites, which is used in practice to reduce food allergenicity [55].

Apart from allergenicity, toxicity of substances should be taken into account. It is evaluated using ToxinPred. It is a web server based on a peptide dataset consisting of 1805 toxic peptides obtained from various databases (ATDB, Arachno-Server, Conoserver, DBETH, BTXpred, NTXpred and SwissProt) [56]. There is evidence that certain amino acid residues, such as *Cys*, *His*, *Asn*, *Pro*, or the *Phe-Lys-Lys*, *Leu-Lys-Leu*, *Lys-Lys-Leu-Leu*, *Lys-Trp-Lys*, *Cys-Tyr-Cys-Arg* sites, are frequently found in toxic peptides, whereas *Arg*, *Leu*, *Lys*, and *Ile* are the least common [56, 57]. Bioinformatic tools for predicting toxicity *in silico* work on the principle of analyzing amino acid sequence for specific amino acid sites [58]. Current computational approaches used in toxicology are thoroughly described in studies of antidiabetic, antihy-

pertensive, antioxidant peptides and other biological objects for bioinformatic safety assessments [59–63].

For the food industry or pharmaceuticals to continue using bioactive peptides, it is necessary to predict their flavor profile and sensory characteristics in combination. Sensory characteristics of biologically active peptides are another significant descriptor that bioinformatics tools provide for analysis. The taste profile can be predicted due to the BIOPEP, which contains a database of sensory peptides, as well as the BitterDB, which contains peptides with bitter taste [64]. In addition to sensory peptides with bitter, sweet and umami tastes, the abnormal taste profile for hydrolysates can be formed due to a high index of free amino acids (FAA) [65]. This indicator can be evaluated and corrected during computer modeling of the targeted protein bioconversion *in silico*.

Evaluation of the physicochemical and technological properties of peptides. The amino acid sequence in the structure of peptides obtained as a result of hydrolysis affects the stability of the system, physicochemical and technological properties. They will affect the application scope for the obtained components. The bioinformatic tool PepCalc was successfully used in a number of studies to predict physicochemical properties. It can be used to predict peptide solubility in water, theoretical molecular weight, isoelectric point, total charge as a function of pH, extinction coefficient, and instability index [66–68]. The importance of predicting the instability index, characterizing intramolecular stability, lies in the correlation of this index with the thermostability of peptides. This is a significant factor in the technological process (heat treatment) and in the microbiological safety of hydrolysis products [69]. Therefore, the instability index can be viewed as one of the criteria for evaluating the targeted hydrolysis model or a basis for its possible adjustment.

The Expasy ProtParam and ProtPi tools can also be used to predict the instability index, half-life, extinction coefficient, hydropathicity (GRAVY) and some other characteristics.

Stability evaluation of bioactive peptides during digestion in the gastrointestinal model. The structure of biologically active peptides can be destroyed in the gastrointestinal tract by the action of digestive enzymes with complete or partial loss of biofunctional properties. Therefore, it is pointless to extract biologically active peptides blindly, without taking into account degradation in the GI tract. Evaluating peptide stability during simulated digestion is an important final step in a hybrid strategy of bioinformatic modeling (*in silico*) for targeted hydrolysis. *In silico* modeling of digestion can be accomplished via the bionformatic resources described earlier in “Screening the Specificity of Enzyme Preparations”. To simulate digestion in the gastrointestinal tract, three main digestive enzymes, produced in the human body, are used: trypsin, chymotrypsin and pancreatic elastase [70].

Digital model of a peptide complex. Based on the sequentially generated algorithm *in silico*, it seems objectively possible to create a digital model of the peptide complex. The peptide complex with predicted bioactivity, safety, and sensory characteristics may be an object of subsequent scaling studies in real experimental conditions.

Conclusion

By evaluating the capabilities of multi-directional bioinformatic analysis methods combined with high-performance algorithms of proteomic database, it is possible to combine and integrate them into a hybrid strategy for the bioinformatic modeling (*in silico*) of hydrolysis for targeted release of stable peptide complexes with predictable bioactivity, stability, safety and sensory characteristics from complex protein matrices of dairy raw materials. In the generated hybrid strategy algorithm for a bioinformatic modeling, the mainemphasis is placed on safety due to excluding the formationof peptide forms that have a negative impact on the functioning of human organs and human health in general.

The data obtained by bioinformatic modeling (*in silico*) do not always fully correlate with the experimental data obtained *in vitro* and *in vivo* during targeted hydrolysis of milk protein and yet the hybrid algorithm presented in this article facilitate s the accumulation of the necessary primary data to reduce the time and financial costs of real experiments.

However, despite all the advantages of bioinformatics and various strategies, *in silico* remains only a preliminary step in a cascade of studies for biologically active milk protein peptides due to the impossibility of predicting the theoretical enzymatic cleavage under various technological conditions (temperature, duration, active acidity, substrate-enzyme ratio). This offers the basis for studies to optimize the conditions of enzymatic hydrolysis, taking into account technological factors *in vitro*.

Contribution

All the authors contributed equally to the study and bear equal responsibility for information published in this article.

Conflict of interest

The authors declare that there is no conflict of interest regarding the publication of this article.

References

1. Karami Z, Akbari-Adergani B. Bioactive food derived peptides: a review on correlation between structure of bioactive peptides and their functional properties. Journal of Food Science and Technology. 2019;56(2):535–547. https://doi.org/10.1007/s13197-018-3549-4

2. Hafeez Z, Cakir-Kiefer C, Roux E, Perrin C, Miclo L, Dary-Mourot A. Strategies of producing bioactive peptides from milk proteins to functionalize fermented milk products. Food Research International. 2014;63:71–80. https://doi.org/10.1016/j.foodres.2014.06.002

3. Kamali Alamdari E, Ehsani MR. Antimicrobial peptides derived from milk: A review. Journal of Food Biosciences and Technology. 2017;7(1):49–56.

4. Ryazantseva KA, Agarkova EYu, Fedotova OB. Continuous hydrolysis of milk proteins in membrane reactors of various configurations. Foods and Raw Materials. 2021;9(2):271–281. https://doi.org/10.21603/2308-4057-2021-2-271-281

5. Kruchinin AG, Savinova OS, Glazunova OA, Moiseenko KV, Agarkova EYu, Fedorova TV. Hypotensive and hepatoprotective properties of the polysaccharide-stabilized foaming composition containing hydrolysate of whey proteins. Nutrients. 2021;13(3). https://doi.org/10.3390/nu13031031

6. Peighambardoust SH, Karami Z, Pateiro M, Lorenzo JM. A review on health-promoting, biological, and functional aspects of bioactive peptides in food applications. Biomolecules. 2021;11(5). https://doi.org/10.3390/biom11050631

7. Bielecka M, Cichosz G, Czeczot H. Antioxidant, antimicrobial and anticarcinogenic activities of bovine milk proteins and their hydrolysates – A review. International Dairy Journal. 2021;127. https://doi.org/10.1016/j.idairyj.2021.105208

8. Huang S, Gong Y, Li Y, Ruan S, Roknul Azam SM, Duan Y, *et al.* Preparation of ACE-inhibitory peptides from milk protein in continuous enzyme membrane reactor with gradient dilution feeding substrate. Process Biochemistry. 2020;92:130–137. https://doi.org/10.1016/j.procbio.2020.02.030

9. Chalamaiah M, Yu W, Wu J. Immunomodulatory and anticancer protein hydrolysates (peptides) from food proteins: A review. Food Chemistry. 2018;245:205–222. https://doi.org/10.1016/j.foodchem.2017.10.087

10. Giacometti J, Buretić-Tomljanović A. Peptidomics as a tool for characterizing bioactive milk peptides. Food Chemistry. 2017;230:91–98. https://doi.org/10.1016/j.foodchem.2017.03.016

11. Hayes M, Stanton C, Fitzgerald GF, Ross RP. Putting microbes to work: Dairy fermentation, cell factories and bioactive peptides. Part II: Bioactive peptide functions. Biotechnology Journal. 2007;2(4):435–449. https://doi.org/10.1002/biot.200700045

12. Etemadian Y, Ghaemi V, Shaviklo AR, Pourashouri P, Sadeghi Mahoonak AR, Rafipour F. Development of animal/plant-based protein hydrolysate and its application in food, feed and nutraceutical industries: State of the art. Journal of Cleaner Production. 2021;278. https://doi.org/10.1016/j.jclepro.2020.123219

13. Sánchez A, Vázquez A. Bioactive peptides: A review. Food Quality and Safety. 2017;1(1):29–46.

14. Nongonierma AB, FitzGerald RJ. Strategies for the discovery and identification of food protein-derived biologically active peptides. Trends in Food Science and Technology. 2017;69:289–305. https://doi.org/10.1016/j.tifs.2017.03.003

15. Yu Z, Chen Y, Zhao W, Zheng F, Ding L, Liu J. Novel ACE inhibitory tripeptides from ovotransferrin using bioinformatics and peptidomics approaches. Scientific Reports. 2019;9(1). https://doi.org/10.1038/s41598-019-53964-y

16. Tu M, Cheng S, Lu W, Du M. Advancement and prospects of bioinformatics analysis for studying bioactive peptides from food-derived protein: Sequence, structure, and functions. TrAC – Trends in Analytical Chemistry. 2018;105:7–17. https://doi.org/10.1016/j.trac.2018.04.005

17. Barati M, Javanmardi F, Jabbari M, Mokari-Yamchi A, Farahmand F, Eş I, *et al.* An *in silico* model to predict and estimate digestion-resistant and bioactive peptide content of dairy products: A primarily study of a time-saving and affordable method for practical research purposes. LWT. 2020;130. https://doi.org/10.1016/j.lwt.2020.109616

18. Panyayai T, Ngamphiw C, Tongsima S, Mhuantong W, Limsripraphan W, Choowongkomon K, *et al.* FeptideDB: A web application for new bioactive peptides from food protein. Heliyon. 2019;5(7). https://doi.org/10.1016/j.heliyon.2019.e02076

19. Agarkova EYu, Kruchinin AG. Enzymatic conversion as a method of producing biologically active peptides. Vestnik of MSTU. 2018;21(3):412–419. (In Russ.). https://doi.org/10.21443/1560-9278-2018-21-3-412-419

20. FitzGerald RJ, Cermeño M, Khalesi M, Kleekayai T, Amigo-Benavent M. Application of *in silico* approaches for the generation of milk protein-derived bioactive peptides. Journal of Functional Foods. 2020;64. https://doi.org/10.1016/j.jff.2019.103636

21. Agyei D, Tsopmo A, Udenigwe C. Bioinformatics and peptidomics approaches to the discovery and analysis of food-derived bioactive peptides. Analytical and Bioanalytical Chemistry. 2018;410(15):3463–3472. https://doi.org/10.1007/s00216-018-0974-1

22. Ryskaliyeva A, Henry C, Miranda G, Faye B, Konuspayeva G, Martin P. Alternative splicing events expand molecular diversity of camel CSN1S2 increasing its ability to generate potentially bioactive peptides. Scientific Reports. 2019;9(1). https://doi.org/10.1038/s41598-019-41649-5

23. Ryazantzeva KA, Agarkova EYu. Using in silico methods to obtain bioactive peptides of whey. Food Industry. 2021;(5):32–35. (In Russ.). https://doi.org/10.52653/PPI.2021.5.5.007

24. Peredo-Lovillo A, Hernández-Mendoza A, Vallejo-Cordoba B, Romero-Luna HE. Conventional and *in silico* approaches to select promising food-derived bioactive peptides: A review. Food Chemistry: X. 2021;13. https://doi.org/10.1016/j.fochx.2021.100183

25. Lafarga T, O'Connor P, Hayes M. Identification of novel dipeptidyl peptidase-IV and angiotensin-I-converting enzyme inhibitory peptides from meat proteins using *in silico* analysis. Peptides. 2014;59:53–62. https://doi.org/10.1016/j.peptides.2014.07.005

26. Kruchinin AG, Turovskaya SN, Illarionova EE, Bigaeva AV. Evaluation of the effect of *κ*-casein gene polymorphism in milk powder on the technological properties of acid-induced milk gels. Food Processing: Techniques and Technology. 2021;51(1):53–66. (In Russ.). https://doi.org/10.21603/2074-9414-2021-1-53-66

27. Bateman A, Martin MJ, O'Donovan C, Magrane M, Apweiler R, Alpi E, *et al.* UniProt: a hub for protein information. Nucleic Acids Research. 2015;43(D1):D204–D212. https://doi.org/10.1093/nar/gku989

28. Nongonierma AB, FitzGerald RJ. Enhancing bioactive peptide release and identification using targeted enzymatic hydrolysis of milk proteins. Analytical and Bioanalytical Chemistry. 2018;410(15):3407–3423. https://doi.org/10.1007/s00216-017-0793-9

29. Iwaniak A, Minkiewicz P, Darewicz M, Sieniawski K, Starowicz P. BIOPEP database of sensory peptides and amino acids. Food Research International. 2016;85:155–161. https://doi.org/10.1016/j.foodres.2016.04.031

30. Dyachenko EA, Slugina MA. Intraspecific variability of the *Sus1* sucrose synthase gene in *Pisum sativui* accessions. Vavilov Journal of Genetics and Breeding. 2018;22(1):108–114. (In Russ.). https://doi.org/10.18699/VJ18.338

31. Kruchinin AG, Turovskaya SN, Illarionova EE, Bigaeva AV. Molecular genetic modifications of *κ*-casein. News of Institutes of Higher Education. Food Technology. 2020;376(4):12–16. (In Russ.). https://doi.org/10.26297/0579-3009.2020.4.3

32. Asledottir T, Le TT, Petrat-Melin B, Devold TG, Larsen LB, Vegarud GE. Identification of bioactive peptides and quantification of *β*-casomorphin-7 from bovine *β*-casein A1, A2 and I after ex vivo gastrointestinal digestion. International Dairy Journal. 2017;71:98–106. https://doi.org/10.1016/J.IDAIRYJ.2017.03.008

33. Gilmanov KhKh, Semipyatnyi VK, Bigaeva AV, Vafin RR, Turovskaya SN. New determination method for the ratio of the relative proportions of ϰ-casein alleles in milk powder. Food Processing: Techniques and Technology. 2020;50(3):525–535. (In Russ.). https://doi.org/10.21603/2074-9414-2020-3-525-535

34. Nielsen SD, Beverly RL, Qu Y, Dallas DC. Milk bioactive peptide database: A comprehensive database of milk protein-derived bioactive peptides and novel visualization. Food Chemistry. 2017;232:673–682. https://doi.org/10.1016/j.foodchem.2017.04.056

35. Nebbia S, Lamberti C, Lo Bianco G, Cirrincione S, Laroute V, Cocaign-Bousquet M, *et al.* Antimicrobial potential of food lactic acid bacteria: Bioactive peptide decrypting from caseins and bacteriocin production. Microorganisms. 2021;9(1). https://doi.org/10.3390/microorganisms9010065

36. Nielsen SD, Beverly RL, Underwood MA, Dallas DC. Differences and similarities in the peptide profile of preterm and term mother's milk, and preterm and term infant gastric samples. Nutrients. 2020;12(9). https://doi.org/10.3390/nu12092825

37. Pa'ee KF, Razali N, Sarbini SR, Ramonaran Nair SN, Yong Tau Len K, Abd-Talib N. The production of collagen type I hydrolyzate derived from tilapia (*Oreochromis sp*.) skin using thermoase PC10F and its in silico analysis. Food Biotechnology. 2021;35(1):1–21. https://doi.org/10.1080/08905436.2020.1869040

38. Minkiewicz P, Iwaniak A, Darewicz M. BIOPEP-UWM database of bioactive peptides: Current opportunities. International Journal of Molecular Sciences. 2019;20(23). https://doi.org/10.3390/ijms20235978

39. Camacho C, Coulouris G, Avagyan V, Ma N, Papadopoulos J, Bealer K, *et al.* BLAST+: architecture and applications. BMC Bioinformatics. 2009;10. https://doi.org/10.1186/1471-2105-10-421

40. González-Pech RA, Stephens TG, Chan CX. Commonly misunderstood parameters of NCBI BLAST and important considerations for users. Bioinformatics. 2019;35(15):2697–2698. https://doi.org/10.1093/bioinformatics/bty1018

41. Wang Z, Wang G. APD: The antimicrobial peptide database. Nucleic Acids Research. 2004;32:D590–D592.

42. Li Q, Zhang C, Chen H, Xue J, Guo X, Liang M, *et al.* BioPepDB: an integrated data platform for food-derived bioactive peptides. International Journal of Food Sciences and Nutrition. 2018;69(8):963–968. https://doi.org/10.1080/09637486.2018.1446916

43. Iwaniak A, Minkiewicz P, Pliszka M, Mogut D, Darewicz M. Characteristics of biopeptides released in silico from collagens using quantitative parameters. Foods. 2020;9(7). https://doi.org/10.3390/foods9070965

44. Kartal C, Türköz BK, Otles S. Prediction, identification and evaluation of bioactive peptides from tomato seed proteins using in silico approach. Journal of Food Measurement and Characterization. 2020;14(4):1865–1883. https://doi.org/10.1007/s11694-020-00434-z

45. Kusumah J, Real Hernandez LM, de Mejia EG. Antioxidant potential of mung bean (*Vigna radiata*) albumin peptides produced by enzymatic hydrolysis analyzed by biochemical and in silico methods. Foods. 2020;9(9). https://doi.org/10.3390/foods9091241

46. Baghban R, Ghasemali S, Farajnia S, Hoseinpoor R, Andarzi S, Zakariazadeh M, *et al.* Design and in silico evaluation of a novel cyclic disulfide-rich anti-VEGF peptide as a potential antiangiogenic drug. International Journal of Peptide Research and Therapeutics. 2021;27(4):2245–2256. https://doi.org/10.1007/s10989-021-10250-8

47. Gu Y, Li X, Liu H, Li Q, Xiao R, Dudu OE, *et al.* The impact of multiple-species starters on the peptide profiles of yoghurts. International Dairy Journal. 2020;106. https://doi.org/10.1016/j.idairyj.2020.104684

48. Tu M, Liu H, Zhang R, Chen H, Fan F, Shi P, *et al.* Bioactive hydrolysates from casein: generation, identification, and *in silico* toxicity and allergenicity prediction of peptides. Journal of the Science of Food and Agriculture. 2018;98(9):3416–3426. https://doi.org/10.1002/jsfa.8854

49. Nutten S, Schuh S, Dutter T, Heine RG, Kuslys M. Design, quality, safety and efficacy of extensively hydrolyzed formula for management of cow's milk protein allergy: What are the challenges? Advances in Food and Nutrition Research. 2020;93:147–204. https://doi.org/10.1016/bs.afnr.2020.04.004

50. Gromov DA, Borisova AV, Bakharev VV. Food allergens and methods for producing hypoallergenic foods. Food Processing: Techniques and Technology. 2021;51(2):232–247. (In Russ.). https://doi.org/10.21603/2074-9414-2021-2-232-247

51. Loh W, Tang MLK. The epidemiology of food allergy in the global context. International Journal of Environmental Research and Public Health. 2018;15(9). https://doi.org/10.3390/ijerph15092043

52. Flom JD, Sicherer SH. Epidemiology of cow's milk allergy. Nutrients. 2019;11(5). https://doi.org/10.3390/nu11051051

53. Sharma N, Patiyal S, Dhall A, Pande A, Arora C, Raghava GPS. AlgPred 2.0: An improved method for predicting allergenic proteins and mapping of IgE epitopes. Briefings in Bioinformatics. 2021;22(4). https://doi.org/10.1093/bib/bbaa294

54. Vens C, Rosso MN, Danchin EG. Identifying discriminative classification-based motifs in biological sequences. Bioinformatics. 2011;27(9):1231–1238. https://doi.org/10.1093/bioinformatics/btr110

55. Kharitonov VD, Agarkova EYu, Kruchinin AG, Ryazantseva KA, Korolyeva OV, Fedorova TV, *et al.* Impact of new fermented dairy product with whey protein hydrolysate on tolerance and dynamics of atopic dermatitis manifestation in children suffering from cow's milk protein allergy. Problems of Nutrition. 2015;84(5):56–63. (In Russ.). https://doi.org/10.24411/0042-8833-2015-00048

56. Gupta S, Kapoor P, Chaudhary K, Gautam A, Kumar R, Raghava GPS. *In silico* approach for predicting toxicity of peptides and proteins. PloS ONE. 2013;8(9). https://doi.org/10.1371/journal.pone.0073957

57. Chaudhary K, Kumar R, Singh S, Tuknait A, Gautam A, Mathur D, *et al.* A web server and mobile app for computing hemolytic potency of peptides. Scientific Reports. 2016;6. https://doi.org/10.1038/srep22843

58. Parthasarathi R, Dhawan A. In silico approaches for predictive toxicology. In: Dhawan A, Kwon S, editors. In vitro toxicology. Academic Press; 2018. pp. 91–109. https://doi.org/10.1016/B978-0-12-804667-8.00005-5

59. Yap PG, Gan CY. In vivo challenges of anti-diabetic peptide therapeutics: Gastrointestinal stability, toxicity and allergenicity. Trends in Food Science and Technology. 2020;105:161–175. https://doi.org/10.1016/j.tifs.2020.09.005

60. Guo H, Hao Y, Richel A, Everaert N, Chen Y, Liu M, *et al.* Antihypertensive effect of quinoa protein under simulated gastrointestinal digestion and peptide characterization. Journal of the Science of Food and Agriculture. 2020;100(15):5569–5576. https://doi.org/10.1002/jsfa.10609

61. Ji D, Udenigwe C, Agyei D. Antioxidant peptides encrypted in flaxseed proteome: An *in silico* assessment. Food Science and Human Wellness. 2019;8(3):306–314. https://doi.org/10.1016/j.fshw.2019.08.002

62. Lin K, Zhang L-W, Han X, Xin L, Meng Z-X, Gong P-M, *et al.* Yak milk casein as potential precursor of angiotensin I-converting enzyme inhibitory peptides based on *in silico* proteolysis. Food Chemistry. 2018;254:340–347. https://doi.org/10.1016/j.foodchem.2018.02.051

63. Tu M, Qiao X, Wang C, Liu H, Cheng S, Xu Z, *et al. In vitro* and *in silico* analysis of dual-function peptides derived from casein hydrolysate. Food Science and Human Wellness. 2021;10(1):32–37. https://doi.org/10.1016/j.fshw.2020.08.014

64. Dagan-Wiener A, Di Pizio A, Nissim I, Bahia MS, Dubovski N, Margulis E, *et al.* BitterDB: taste ligands and receptors database in 2019. Nucleic Acids Research. 2019;47(D1):D1179–D1185. https://doi.org/10.1093/nar/gky974

65. Wang W, Zhang L, Wang Z, Wang X, Liu Y. Physicochemical and sensory variables of Maillard reaction products obtained from *Takifugu obscurus* muscle hydrolysates. Food Chemistry. 2019;290:40–46. https://doi.org/10.1016/j.foodchem.2019.03.065

66. Stan GM, Constantinescu-Aruxandei D, Oancea F. In silico analysis of the formation of bioactive peptides from silver carp (*Hypophthalmichthys molitrix*) collagen. Proceedings. 2020;57(1). https://doi.org/10.3390/proceedings2020057024

67. Pooja K, Rani S, Prakash B. *In silico* approaches towards the exploration of rice bran proteins-derived angiotensin-I-converting enzyme inhibitory peptides. International Journal of Food Properties. 2017;20. https://doi.org/10.1080/10942912.2017.1368552

68. Shi P, Fan F, Chen H, Xu Z, Cheng S, Lu W, *et al.* A bovine lactoferrin–derived peptide induced osteogenesis via regulation of osteoblast proliferation and differentiation. Journal of Dairy Science. 2020;103(5):3950–3960. https://doi.org/10.3168/jds.2019-17425

69. Tu M, Liu H, Zhang R, Chen H, Mao F, Cheng S, *et al.* Analysis and evaluation of the inhibitory mechanism of a novel angiotensin-I-converting enzyme inhibitory peptide derived from casein hydrolysate. Journal of Agricultural and Food Chemistry. 2018;66(16):4139–4144. https://doi.org/10.1021/acs.jafc.8b00732

70. Sayd T, Dufour C, Chambon C, Buffière C, Remond D, Santé-Lhoutellier V. Combined *in vivo* and *in silico* approaches for predicting the release of bioactive peptides from meat digestion. Food Chemistry. 2018;249:111–118. https://doi.org/10.1016/j.foodchem.2018.01.013