Goat casein peptides and their potential through an *in silico* approach

Joanderson Pereira Cândido da Silva

Program in Biological Sciences (Biochemistry),
Institute of Chemistry of University of São Paulo,
São Paulo, Brazil
Email: joandersonpereiracs@gmail.com

Paula Perazzo de Souza Barbosa* and Joellington Marinho de Almeida

Program in Food Science and Technology,
Federal University of Paraiba,
João Pessoa, Brazil
Email: paulabio05@hotmail.com
Email: joellingtonalmeida@gmail.com
*Corresponding author

Nathália Miranda de Medeiros

Program in Biological Sciences (Biochemistry),
Institute of Chemistry of University of São Paulo,
São Paulo, Brazil
Email: nathaliamiranda@usp.br

Emmanuel Silva Marinho

Theoretical Chemistry and Electrochemistry Group,
State University of Ceará,
Limoeiro do Norte, Brazil
Email: emmanuel.marinho@uece.br

Tatiane Santi Gadelha

Department of Molecular Biology,
Federal University of Paraiba,
João Pessoa, Brazil
Email: santi.tatiane@gmail.com

Abstract: The caprine caseins αs1, αs2, β, and κ from UniProtKB were characterised for amino acid profile, and hydrolysed *in silico* with pepsin or trypsin. The generated peptides were characterised for physical-defined properties, bioactive potential, Boman index, toxicity and allergenicity. The peptides generated are bioactive, predominating dipeptidyl peptidase IV and

ACE inhibitors activities. Only two peptides are toxic, and most have been proposed to be allergenic by *Allergen Online* or SDAP, because they have < 50% identity with other confirmed allergenic proteins, including bovine milk caseins, however, they were considered non-allergenic by the methods MEME/MAST and BLAST in Algpred, requiring more studies for better evaluation. Despite of variations, bioinformatics tools are useful before bench tests, because allows time and cost savings, and provides studies projection for use by pharmaceutical and food industries, besides to identify sequences that may need other tests to evaluate the safety before being introduced in use.

Keywords: proteolysis; bioactive peptides; bioinformatics; prediction; eating disorders; industrial applications; biopotential; allergenicity; toxicity; physical-defined properties.

Reference to this paper should be made as follows: da Silva, J.P.C., de Souza Barbosa, P.P., de Almeida, J.M., de Medeiros, N.M., Marinho, E.S. and Gadelha, T.S. (2023) 'Goat casein peptides and their potential through an *in silico* approach', *Int. J. Bioinformatics Research and Applications*, Vol. 19, No. 1, pp.70–81.

Biographical notes: Joanderson Pereira Cândido da Silva graduated in Biological Sciences from the Federal University of Paraíba and taking a Master's degree in Biochemistry from the Institute of Chemistry of the University of São Paulo, at the Laboratory of Applied Toxinology at the Butantan Institute.

Paula Perazzo de Souza Barbosa graduated in Biological Sciences from the University of the State of Rio Grande do Norte, Master in Cellular and Molecular Biology from Department of Molecular Biology and PhD from the Postgraduate Program in Food Science and Technology, both at the Federal University of Paraíba.

Joellington Marinho de Almeida graduated in Food Engineering and Master in Food Science and Technology from the Federal University of Paraíba, and currently a PhD student in the Postgraduate Program in Biotechnology at the Federal University of Espírito Santo.

Nathália Miranda de Medeiros is a Chemical Engineer from the Federal University of Paraíba in 2020. She is a PhD student in Biological Sciences (Biochemistry) at the Institute of Chemistry of the University of São Paulo.

Emmanuel Silva Marinho graduated in Chemistry and has a Master's and Doctorate in Biochemistry from the Federal University of Ceará. He is currently an Adjunct Professor at the State University of Ceará, a Collaborating Professor at the Posgraduate Programs in Natural Sciences at the State University of Ceará and Chemistry at the Federal University of Ceará (UFC), and coordinates the Theoretical Chemistry and Electrochemistry Group.

Tatiane Santi Gadelha graduated in Biological Sciences from the Federal University of Santa Maria. She has a Master and Doctor in Biochemistry from the Federal University of Ceará. She is an Associate Professor IV at the Federal University of Paraíba.

1 Introduction

Proteins and peptides are milk components with roles in metabolic functions of organisms and, therefore, for health because they have important hormonal or pharmacological properties. Depending on the primary structure, bioactive peptides exercise different activities and can influence physiological processes (Chakrabarti et al., 2018). They are produced naturally *in vivo* through gastrointestinal digestion of proteins from different food matrices, or by *in vitro* enzymatic hydrolysis (Albenzio et al., 2017).

The *in vitro* enzymatic hydrolysis approach allows yield optimisation of peptides for admitting the choice substrate, specific enzymes and reaction conditions (Espejo-Carpio et al., 2013). Instead, the use of bioinformatics (*in silico*), which is based on information available in databases, can be a useful tool for obtaining bioactive peptides before *in vitro* and *in vivo* tests, since it allows time reduction and costs (Daliri et al., 2017), aiding in the selection of appropriate substrate and enzyme, while to implying precise and quick analyses (FitzGerald et al., 2020).

Considering this scenario, this study addresses *in silico* hydrolysis with pepsin (pH 2.0 or pH 1.3), or trypsin of caprine milk caseins, with biological activities prediction of generated fragments, and some of their physicochemical, toxicity and allergenicity, provide and, theoretically, reinforce the properties of proteins as potential precursors of bioactive peptides with low toxicity, allergenicity and with potential antimicrobial to serve food and pharmaceutical industries.

2 Material and methods

The caprine milk proteins selected for *in silico* hydrolysis were the caseins αs1, αs2, β, and κ (Table 1), whose primary sequences are deposited in *UniProtKB*. Their amino acid composition was determined by *ProtParam*. Each of caprine milk caseins was hydrolysed with pepsin (EC 3.4.23.1; pH > 2 and pH 1.3) or trypsin (EC 3.4.21.4; pH not specified), with the *BIOPEP database*, also used to verification of probable biological activities. Molecular weight, isoelectric point (pI), net charge at pH 7 and solubility of peptides were estimated with *INNOVAGEN Peptide Property Calculator*.

Table 1 Entry name, length and molecular weights of caprine caseins obtained from UniProtKB and selected for in silico hydrolysis

Caseins	*Entrance*	*Length*	*Molecular weight (Da)*
αs1	P18626	214	24.290
αs2	P33049	223	26.389
β	P33048	222	24.865
κ	P02670	192	21.441

The *Boman index* of peptides was investigated in *Antimicrobial Peptide Database* (APD3), with sequences starting at five amino acids in length. Toxicity was predicted with *ToxinPred*; peptides with less than 50 amino acids were analysed in the *batch submision* option and, above that, in *protein scanning*. The support vector machine (SVM) based prediction method with a threshold value of 0.0 and e-value cut-off of 10 was chosen for the toxicity prediction.

Allergenicity was verified in *Allergen Online*; *Structural Database of Allergenic Proteins* – SDAP, using the parameters:

1 the complete sequence (*Full FASTA*) alignment, with E-value of 0.1, looking for identities (ID) > 50%

2 the 80 amino acid sliding window alignment, looking for IDs > 35%

3 small exact *wordmatch*: eight, seven and six contiguous amino acids sliding windows alignment, according to the FAO / WHO allergenicity rules.

The peptides were also evaluated for potential allergenicity using *AlgPred* (Saha and Raghava, 2006). Three methods were used: SVM module, based on amino acid composition prediction, MEME/MAST motif prediction approach and BLAST, similarity-based approach.

3 Results

3.1 Characterisation of peptides generated

The sequences of caseins have the 20 amino acids and glutamic acid is the most abundant amino acid in caseins αs1 and αs2 caseins. In β and k, proline is the most frequent amino acid (Table 2). A total of 253 different peptides were generated through the *in silico* hydrolysis of the four caprine caseins (αs1, αs2, β and κ) in BIOPEP with enzymes pepsin (pH 1.3 or pH 2.0) or trypsin, whose pH is not discriminated in the tool, but acts in a basic range of pH.

Table 2 Amino acid content of caprine caseins

Amino acid		*Amino acid percentage in caseins (%)**			
		αS1	*αs2*	*β*	*k*
Ala	A	7	5.4	3.6	9.4
Arg	R	3.3	3.1	1.4	2.6
Asn	N	5.1	5.8	1.8	4.7
Asp	D	3.3	2.2	1.8	3.6
Cys	C	0.5	1.3	0.5	1.6
Gln	Q	6.5	7.2	9.5	7.8
Glu	E	9.3	11.2	8.6	5.7
Gly	G	4.2	0.9	2.3	1
His	H	1.9	2.2	2.3	2.1
Ile	I	4.7	5.8	5.0	5.7
Leu	L	10.3	5.8	11.7	6.8
Lys	K	6.5	11.2	5.9	4.7

Notes: Values obtained from ProtParam.
*Percentage of each amino acid identified in each casein.

Table 2 Amino acid content of caprine caseins (continued)

Amino acid		*Amino acid percentage in caseins (%)**			
		αs1	*αs2*	*β*	*k*
Met	M	2.8	2.2	3.2	1.6
Phe	F	3.3	4.5	4.1	3.6
Pro	P	8.9	5.8	14.9	10.4
Ser	S	8.4	6.3	6.8	7.3
Thr	T	2.8	6.7	5	8.9
Trp	W	0.9	1.3	0.5	0.5
Tyr	Y	5.1	5.4	1.4	4.7
Val	V	5.1	5.4	10.4	7.3

Notes: Values obtained from ProtParam.
*Percentage of each amino acid identified in each casein.

The fragments varied between dipeptides and sequences with 66 amino acids in length. Pepsin at pH 2.0 was able to hydrolyse the caseins, through which was generated more dipeptides and tripeptides. Pepsin at pH 1.3 or trypsin produced higher amounts of peptides with more than six amino acids, characterised by lower degree of hydrolysis values when compared to a hydrolysis performed with pepsin at pH 2.0 (Table 3).

Table 3 Quantity of peptides by size, released through in silico proteolysis of caprine caseins in three different conditions and degree of hydrolysis

Caseins	*Hydrolysis condition*	*DH**	*Size*					
			Di	*Tri*	*Tetra*	*Penta*	*Hexa*	*> of six aa***
αs1	Pepsin pH 1.3	13.2%	4	7	2	2	0	11
	Pepsin pH 2	67.5%	25	10	1	0	0	0
	Trypsin	9.5%	3	1	1	0	2	15
αs2	Pepsin pH 1.3	9%	2	2	3	0	1	9
	Pepsin pH 2	69%	28	8	0	2	1	0
	Trypsin	14%	6	6	1	1	3	14
β	Pepsin pH 1.3	15%	5	3	2	1	8	11
	Pepsin pH 2	63%	21	5	3	0	0	0
	Trypsin	7%	3	1	0	1	2	9
κ	Pepsin pH 1.3	10%	4	0	0	4	0	8
	Pepsin pH 2	64%	20	5	3	1	0	1
	Trypsin	7%	0	3	1	2	0	8

Notes: *DH: Degree of Hydrolysis; **aa: amino acid.

Peptides molecular weight varied between 5,571.84 g/mol and 162.14 g/mol. These variations are related to the types of enzymes used and the factors that influence enzyme activity, such as pH. The pI of peptides ranged between 11.41 and 0.88, with 62 of them having pI < 7, which suggest their acidic nature, and 36 with pI > 7, which indicates their basic nature. Most of the fragments, about 111, have good solubility in water. In dipeptide, tripeptide, and tetrapeptide sequences, in which at least one amino acid

contains basic side chain such as lysine, arginine, and histidine, presented pI > 5.0. Those with acidic side chain amino acids have pI < 5.0.

3.2 Bioactivity of peptides and Boman index of peptides released

In silico analyses of caprine milk caseins show that they are good precursors of bioactive peptides, predominating dipeptidyl peptidase IV (DPP-IV) inhibitory activities and ACE inhibitors. Other properties were registered for peptides, but less often: antioxidant renin inhibitor, opioid antagonist, glucose-stimulating peptide release, vasoactive stimulating substance, alpha-glucosidase inhibitor, anxiolytic peptide, anti-inflammatory, opioid, inhibitor of prolyl endopeptidase, peptide regulating the metabolism of phosphoinositol, antithrombotic, peptide that regulates the activity of the stomach mucous membrane, CaMPDE inhibitor, activating ubiquitin-mediated proteolysis, ileum contraction peptide, chymosin inhibitor, antifungal, and antibacterial.

According to the APD3, the value of the *Boman index* ranged from 4.69 to –3.32 Kcal/mol for all analysed peptides in this work. Of the 93 peptides analysed, 33 have values > 2.48 Kcal/mol.

3.3 Toxicity of peptides released through in silico hydrolysis of caprine caseins in three different conditions

Only two of the peptides generated from internal sequences were identified as toxic, both from αs1:

- SKDIGSESTEDQAMEDAKQMKAGSSSSSEEIVPNSAEQKYIQKEDVPSERYL, produced with pepsin pH 1.3
- AGSSSSSEEIVPNSAEQK, with trypsin.

In the first peptide, *ToxinPred* identified two regions as toxic, DAKQMKAGSS and AKQMKAGSSS (SVM scores: 0.02 and 0.12, respectively). The second peptide of sequence, SVM-score of 0.11, is encrypted in the first, where in this situation it was not indicated as a toxic region, only when released, which is related to hydrolysis conditions. All other peptides obtained had negative SVM scores and were considered non-toxic.

3.4 Allergenicity of peptides released through in silico hydrolysis of caprine caseins

Allergenicity was also evaluated in 65 peptides which have length starting from 8 amino acids with SDAP and *Allergen Online*; and in 46 peptides which have length starting from 10 amino acids with Algpred. Using SDAP and *Allergen Online*, those whose identity above 50% with known allergens were considered allergenic.

The 8 amino acid window, used by *Allergen Online* and SDAP, identified 47 and 48 peptides allergenic, in respective. The 100 and 80 amino acid windows used only by *Allergen Online* identified 65 and 12 peptides allergenic, in respective.

The peptides analysed also had identity to the four caseins of bovine milk, and the sequence IHPFAQAQSLVYPFTGPIPNSLPQNILPLTQTPVVVPPFLQPEIMGVPK with the β-casein isoform X2 of yak (*Bos mutus*). With Algpred, peptides were predicted

to have potential allergens by the SVM module. However, using the MEME/MAST and BLAST modules, were identified like non-allergens.

4 Discussion

The results obtained through *ProtParam* reinforces the nutritional value of caprine caseins and make of this milk rich source of proteinogenic amino acids, which can be part of other physiological processes: proline, for example, participates in the metabolism of arginine, polyamines, and glutamate via pyrroline-5-carboxylate, and beginning of transcription. Some amino acids are substrates for the synthesis of others, as phenylalanine, which participates in the synthesis of tyrosine, and isoleucine and serine.

The amino acid feature in protein determines the profile in the fragments generated by hydrolysis, depending on the enzyme used, and, therefore, on its biological activities. Proteolytic enzymes hydrolyse proteins at specific temperature and pH values. Due to this, they generate peptides with varying sizes and free amino acids depending on the position in the polypeptide chain.

Pepsin cleaves preferentially at the C-terminal position of phenylalanine and leucine and, to a lesser extent, in bonds where there is glutamic acid, but does not act in places with valine, alanine or glycine. Trypsin cleaves peptides on the C-terminal side of the amino acid residues lysine and arginine. If a proline residue is on the carboxylic side of cleavage site, cleavage will not occur (Worthington Enzyme Manual, 1993). The amounts of these amino acids and their positions in each casein used in this work favour the action of pepsin and trypsin in the formation of dipeptides and tripeptides generated.

In *in vitro* approach, although the enzymes are specific for their substrates, the hydrolysis environment can be manipulated. The enzyme/protein interaction is altered due to the changes that are induced in the enzyme spatial conformation, thus affecting its access in bonds, which can alter the enzyme behaviour and the attack pattern, resulting in the release of different peptides (Cheison et al., 2011). The different pH values promote changes in the protein conformation substrate, interfering with the enzyme access to the cleavage region points, which also influences the profile of generated fragments. Furthermore, it cannot be disregarded that in *in silico* parameters such as time, temperature and substrate concentration are not provided in the hydrolysis tools, and that they also contribute to the hydrolysate profile.

Molecular weight and pI and of peptides are important properties for solubility adjustments, especially for isolation and storage purposes, as well as for performing experiments in which the protein solubility must be considered. Because at pH close to pI, the solubilisation of peptides may be reduced to a minimum. Further terminal groups (NH_2 and COOH) charges, pI depends on the dissociation constants (pKa) of charged amino acids ionisable groups and, therefore, on the amino acid profile (Kozlowski, 2017).

The amino acid profile also influences the peptide solubility. In our work, the peptides with good solubility in water were those whose primary sequence predominated with amino acids with ionisable or polar side chains. Water solubility is an important parameter, because it influences bioavailability, indispensable factor for achieving the desired concentration in the systemic circulation and for obtaining an ideal therapeutic response (Coltescu et al., 2020). For pharmaceutical industry, solubility is a relevant because its applications generally require samples with high concentrations to obtain a satisfactory response.

In silico hydrolysis of caseins produced more peptides with DPP-IV inhibitory activities and ACE inhibitors. Despite these being in all groups of peptides of different sizes, the first one predominated in relation to the second in the dipeptides originating from each casein when submitted to the action of pepsin at pH 2.0.

DPP-IV is a multifunctional serine protease in epithelial tissues of the liver, kidney, and small intestine, and in circulating soluble form. It is involved in several processes, including the degradation of the glucose-dependent insulinotropic polypeptide (GIP) and the glucagon-like peptide-1 (GLP-1) (Jao et al., 2015), that act on the pancreatic islets by stimulating the release of insulin, thereby controlling blood glucose levels (Nadkarni et al., 2014), what makes DPP-IV inhibition has been used as a therapeutic strategy for type 2 diabetes.

DPP-IV acts in the immune system and one of the hypotheses reported is that, in the respiratory tract, it can facilitate the installation and entry of SARS-CoV-2, disease causative COVID-19 agent. On the other hand, DPP-IV inhibitors have the capacity to induce a protromic state through its inhibition that acts on the vascular system like an antithrombotic and anticoagulant (Kow and Hasan, 2021).

Dipeptides with proline residue in the C-terminal region have been showed could act as DPP-IV inhibitors (Hatanaka et al., 2012). In our work, only three peptides exhibited such a characteristic: HP, released from the hydrolysis of αs1-casein; IP and RP released from the hydrolysis of k-casein, all through pepsin (pH 2.0). One sequence with 33 amino acids, released through hydrolysis of αs1-casein with pepsin (pH 1.3) also exhibited it. However, peptides with proline in other positions also presented DPP-IV inhibitory activity.

Authors have showed that in DPP-IV inhibiting peptides, the occurrence of N-terminal tryptophan and proline is frequent as the second amino acid residue of the substrate (Nongonierma and FitzGerald, 2013); and that tripeptides with the side chain of N-terminal tryptophan interacts with Phe357 in enzyme (Lan et al., 2015). In this work, only three peptides generated by hydrolysis (pepsin pH 2.0) of αs1, k and αs2-caseins, respectively, are in this situation: WY, WQ, and WPQ.

The angiotensin I-converting enzyme (ACE) promotes the formation of angiotensin II and inactivates the vasodilator bradykinin. Food-derived peptides that act as ACE inhibitors significantly lower blood pressure (Manzanares et al., 2019). The literature related to the bioactivity of caprine milk with a focus on peptides derived from their proteins with frequent ACE inhibitory activity, however scarce it is in relation to other possible biological activities that could be explored.

The effects of aromatic amino acids in the third position of the tripeptides on the ACE inhibitory activity and of charged amino acids in the second position have been investigated and it was concluded that, for a high ACE inhibitory activity, is essential to have a positively charged residue close to an aromatic residue, because the first one participates in the connection by the ACE, while the second one, which is bulky, prevents the substrates from accessing the ACE active site (Kobayashi et al., 2008). In our work, the peptides that were registered with this activity have aromatic chain amino acids at the positions described, which may explain the prediction.

In this work, the BIOPEP-UWM database showed that the peptides KTVDQHQKAMKPWTQPKTNAIPYVRYL, DDKIAKYIPIQYVL, and YIPIQYVLSR have antibacterial potential. According to APD3, these sequences have four, three, and three hydrophobic amino acid residues, respectively, on the surface and their *Boman indexes* are 2.05 Kcal/mol, 0.63 Kcal/mol and 0.53 Kcal/mol, in respective.

Boman index is a measure of the peptide affinity for proteins and its ability to establish biological interactions (Dziuba and Dziuba, 2014). It is based on the sum of the solubility values for all amino acids in a sequence and represents the peptide potential to bind to the membrane or other proteins (Keikha et al., 2019). The low index value may suggest a potential candidate for antibacterial medication without many side effects was reported, property of importance for the pharmaceutical industry (Boman, 2003). Its mechanism of action is based on the interaction with membranes, whose composition is mainly hydrophobic. Thus, antimicrobial peptides do not tend to interact with other proteins, interesting strategy to expand the options for antimicrobial candidates, because microbial resistance.

In silico models have been developed using the support vector machine learning (SVM) technique to discriminate between toxic and non-toxic peptides. *ToxinPred* uses a data set consisting of 1805 toxic peptides (< = 35 residues) that classifies the peptides as toxic or non-toxic based on an SVM algorithm (Gupta et al., 2013) or points to encrypted toxic regions no strings. Monitoring the toxicity of bioactive compounds is necessary to identify possible harmful effects and one of main steps in the design of drugs. Although natural products are widely considered to be less risky compared to synthetic drugs, they are not completely free from the possibility of toxicity or other adverse effects.

Among the allergens in the databases that have identity with the peptides in this work through the 8 amino acid window, parameters used in both *Allergen Online* and SDAP, are the four caseins of bovine milk. In our study, the 80 amino acid Sliding Window FASTA method was used, which shows the possible cross-reactivity of allergens, but the exact match of 6, 7 and 8 amino acids were considered as an additional parameter and, according to Guarneri (2010), it is unlikely that sequences with less than 50% identity are cross-reactive.

In Algpred, the peptides in study were considered for MEME/MAST and BLAST methods as non-allergenic. The MEME/MAST method consists of the MEME module, used to discover motifs in closely related sequences and the MAST, used to search for matches to a set of motifs (from the MEME output). BLAST, on the other hand, identifies allergens based on similarity with allergens and non-allergens, where the query strings were found in the database (Sharma et al., 2020).

Allergenicity through food is one of humanity's main concerns, and milk proteins are among the most common food allergens in babies, as they can trigger immediate IgE-mediated reactions. Caprine milk is often considered less allergenic, what does it a substitute for bovine milk in exceptional cases. However, there are reports of people who triggered an allergy by ingesting caprine milk (Goh et al., 2019).

The allergic potential of a food also depends, among other factors, on the individual susceptibility. Some foods are related to allergic reactions, but are usually seen as unlikely causes of these conditions, which makes them considered hidden allergens (Skypala, 2019). Another factor in the length; B cell epitopes are generally 5 to 17 amino acids in length (Nevagi et al., 2018), and this analysis, most of the available peptides have lengths smaller than 5 amino acids.

Allergenicity prediction is interesting in bioinformatics, as it is potentially useful for clinical research. However, the tools still have limitations for the peptides size, so the data on the potential allergenicity of shorter peptides could not be analysed. The results suggest that allergic reactions to the human body need to be investigated through biochemical and biological experiments to validate the results obtained by

bioinformatics, and sensitised individuals should avoid foods that trigger such reactions because have no treatment for food allergies.

5 Conclusions

The peptides generated by pepsin and trypsin have predominantly DPP-IV and ACE inhibitor activities, being able to represent news perspectives, because can be tested in SARS-COV-2 and patients with hypertension, respectively; most have antigenic sequences with the presence of cysteine, leucine, valine, serine, lysine, threonine, glutamate and alanine, that can be used in vaccine designs; only two of peptides are toxic, generated from αs1 casein when hydrolysed with pepsin pH 1.3 and trypsin, whose SVM scores were negative.

Despite many peptides have been shown to be allergenic by *Allergen Online* or SDAP, because they have < 50% identity with other confirmed allergenic proteins, they were considered non-allergenic by the methods MEME/MAST and BLAST in Algpred, requiring studies *in vitro* and *in vivo* for better evaluation. Furthermore, B cell epitopes are generally 5 to 17 amino acids in length, and our analysis, most of the available peptides have lengths smaller than 5 amino acids.

In silico analyses cannot replace biochemical and biological tests with allergenic individuals to validate the predictions of allergenic compounds, but they can provide valuable information, despite the limitations, for the design of studies on the subject and to guarantee the correct labeling of these hydrolysates in case of use by the food and pharmaceutical industries.

Financing

National Council for Scientific and Technological Development (Conselho Nacional de Desenvolvimento Científico e Tecnológico – CNPq).

Acknowledgements

Acknowledgement to Federal University of Paraiba (Universidade Federal da Paraíba – UFPB).

References

Albenzio, M., Santillo, A., Caroprese, M., Della Malva, A. and Marino, R. (2017) 'Bioactive peptides in animal food products', *Foods*, Vol. 6, pp.1–4, DOI: 10.3390/foods6050035.

Allergen Online [online] http://www.allergenonline.org/databasefasta.shtml (accessed 15 May 2020).

Antimicrobial Peptide Database [online] https://aps.unmc.edu/prediction (accessed 10 August 2020).

Boman, H.G. (2003) 'Antibacterial peptides: basic facts and emerging concepts', *Journal Internal Medicine*, Vol. 254, pp.197–215, DOI: 10.1046/j.1365-2796.2003.01228.x.

Chakrabarti, S., Guha, S. and Majumder, K. (2018) 'Food-derived bioactive peptides in human health: challenges and opportunities', *Nutrient*, Vol. 10, pp.1–17, DOI: 10.3390/nu10111738.

Cheison, S.C., Leeb, E., Letzel, T. and Kulozik, U. (2011) 'Influence of buffer type and concentration on the peptide composition of trypsin hydrolysates of β-lactoglobulin', *Food Chemistry*, Vol. 125, pp.121–127, DOI: https://doi.org/10.1016/j.foodchem.2010.08.047.

Coltescu, A.R., Butnariu, M. and Sarac, I. (2020) 'The importance of solubility for new drug molecules', *Biomedical and Pharmacology Journal*, Vol. 13, pp.577–583, DOI: https://dx.doi.org/10.13005/bpj/1920.

Daliri, E.B.M., Oh, D.H. and Lee, B.H. (2017) 'Bioactive peptides', *Foods*, Vol. 6, pp.1–21, DOI: 10.3390/foods6050032.

Dziuba, B. and Dziuba, M. (2014) 'New milk protein-derived peptides with potential antimicrobial activity: an approach based on bioinformatic studies', *International Journal of Molecular Science*, Vol. 15, pp.14531–14545, DOI: 10.3390/ijms150814531.

Espejo-Carpio, F.J., de Gobba, C., Guadix, A., Guadix, E.M. and Otte, J. (2013) 'Angiotensin I-converting enzyme inhibitory activity of enzymatic hydrolysates of goat milk protein fractions', *International Dairy Journal*, Vol. 32, pp.175–183, DOI: 10.1016 /j.idairyj.2013.04.002.

FitzGerald, R.J., Cermeño, M., Khalesi, M., Kleekayai, T. and Amigo-Benavent, M. (2020) 'Application of in silico approaches for the generation of milk protein-derived bioactive peptides', *Journal of Functional Foods*, Vol. 64, p.103636, DOI:10.1016/j.jff.2019.103636.

Goh, S.H., Chong, K.W., Loh, W. and Goh, A.E.N. (2019) 'Goat's milk anaphylaxis in a cow's milk tolerant child', *Asia Pacific Allergy*, Vol. 9, pp.1–3, DOI: 10.5415/apallergy.2019.9.e34.

Guarneri, F. (2010) 'In silico allergen identification: Proposal for a revision of FAO/WHO guidelines', *Atti della Accademia Peloritana dei Pericolanti. Classe di Scienze Fisiche, Matematiche e Naturali*, Vol. 88, pp.1–9.

Gupta, S., Kapoor, P., Chaudhary, K., Gautam, A., Kumar, R. and Raghava, G.P. (2013) 'In silico approach for predicting toxicity of peptides and proteins', *PloS One*, Vol. 8, pp.1–10, DOI: 10.1371/journal.pone.0073957.

Hatanaka, T., Inoue, Y., Arima, J., Kumagai, Y., Usuki, H., Kawakami, K. and Mukaihara, T. (2012) 'Production of dipeptidyl peptidase IV inhibitory peptides from defatted rice bran', *Food Chemistry*, Vol. 134, pp.797–802, DOI: 10.1016/j.foodchem.2012.02.183.

INNOVAGEN Peptide property calculator [online] https://pepcalc.com/ (accessed 20 March 2020).

Jao, C.L., Hung, C.C., Tung, Y.S., Lin, P.Y., Chen, M.C. and Hsu, K.C. (2015) 'The development of bioactive peptides from dietary proteins as a dipeptidyl peptidase IV inhibitor for the management of type 2 diabetes', *Biomedicine*, Vol. 5, pp.1–7, DOI: 10.7603/s40681-015-0014-9.

Keikha, M., Rahdar, H.A., Karami-Zarandi, M. and Azadi, D. (2019) 'The new insight for novel antimicrobial peptides designing by computational design and improvement of an antimicrobial peptide derivate of LL-37', *Avicenna Journal of Clinical Microbiology and Infection*, Vol. 6, pp.15–20, DOI: 10.34172/ajcmi.2019.04.

Kobayashi, Y., Yamauchi, T., Katsuda, T., Yamaji, H. and Katoh, S. (2008) 'Angiotensin-I converting enzyme (ACE) inhibitory mechanism of tripeptides containing aromatic residues', *Journal of Bioscience and Bioengineering*, Vol. 106, pp.310–312, DOI: https://doi.org /10.1263/jbb.106.310.

Kow, C.S. and Hasan, S.S. (2021) 'Use of DPP-4 inhibitors in patients with COVID-19', *Acta Diabetologica*, Vol. 58, pp.245–246, DOI: https://doi.org/10.1007/s00592-020-01629-y.

Kozlowski, L.P. (2017) 'Proteome-pI: proteome isoelectric point database', *Nucleic Acids Research*, Vol. 45, pp.1112–1116, DOI: 10.1093/nar/gkw978.

Lan, V.T.T., Ito, K., Ohno, M., Motoyama, T., Ito, S. and Kawarasaki, Y. (2015) 'Analyzing a dipeptide library to identify human dipeptidyl peptidase IV inhibitor', *Food Chemistry*, Vol. 175, pp.66–73, DOI: 10.1016/j.foodchem.2014.11.131.

Manzanares, P., Gandía, M., Garrigues, S. and Marcos, J.F. (2019) 'Improving health- promoting effects of food-derived bioactive peptides through rational design and oral delivery strategies', *Nutrients*, Vol. 11, pp.1–22, DOI: 10.3390/nu11102545.

Nadkarni, P., Chepurny, O.G. and Holz, G.G. (2014) 'Regulation of glucose homeostasis by GLP-1', in Teplow, D. (Ed.): *Progress in Molecular Biology and Translational Science*, pp.23–65, Academic Press/Elsevier Inc, London.

Nevagi, R.J., Toth, I. and Skwarczynski, M. (2018) 'Peptide-based vaccines', in Koutsopoulos, S. (Ed.): *Peptide Applications in Biomedicine, Biotechnology and Bioengineering*, Woodhead Publishing, DOI:10.1016/B978-0-08-100736-5.00012-0.

Nongonierma, A.B. and FitzGerald, R.J. (2013) 'Inhibition of dipeptidyl peptidase IV (DPP-IV) by proline containing casein-derived peptides', *Journal of Functional Foods*, Vol. 5, pp.1909–1917, DOI: https://doi.org/10.1016/j.jff.2013.09.012.

ProtParam [online] https://web.expasy.org/protparam/ (accessed 20 March 2020).

Saha, S. and Raghava, G.P.S. (2006) 'AlgPred: prediction of allergenic proteins and mapping of IgE epitopes', *Nucleic Acids Research*, Vol. 34, pp.202–209, DOI: https://doi.org/10.1093/nar/gkl343.

Sharma, N., Patiyal, S., Dhall, A., Pande, A., Arora, C. and Raghava, G.P. (2020) 'AlgPred 2.0: an improved method for predicting allergenic proteins and mapping of IgE epitopes', *Briefings in Bioinformatics*, Vol. 00, pp.1–13, DOI: 10.1093/bib/bbaa294.

Skypala, I.J. (2019) 'Food-induced anaphylaxis: role of hidden allergens and cofactors', *Frontiers in Immunology*, Vol. 10, pp.1–10, DOI: 10.3389/fimmu.2019.00673.

Structural Database of Allergenic Proteins [online] https://fermi.utmb.edu/ (accessed 15 May 2020).

ToxinPred [online] http://crdd.osdd.net/raghava/toxinpred/index.html (accessed 10 April 2020).

UniProtKB [online] https://www.uniprot.org/uniprot/P02670 (accessed 13 March 2020).

UniProtKB [online] https://www.uniprot.org/uniprot/P33048 (accessed 13 March 2020).

UniProtKB [online] https://www.uniprot.org/uniprot/P33049 (accessed 13 March 2020).

UniProtKB [online] https://www.uniprot.org/uniprot/P18626 (accessed 13 March 2020).

Worthington Enzyme Manual [online] http://www.worthington-biochem.com/index/manual.html (accessed 15 June 2020).