

Article

In Silico Analysis of Individual Fractions of Bovine Casein as Precursors of Bioactive Peptides—Influence of Post-Translational Modifications

Piotr Minkiewicz , Małgorzata Darewicz  and Anna Iwaniak 

Department of Food Biochemistry, Faculty of Food Science, University of Warmia and Mazury in Olsztyn, Plac Cieszyński 1, 10-726 Olsztyn, Poland; darewicz@uwm.edu.pl (M.D.); ami@uwm.edu.pl (A.I.)

* Correspondence: minkiew@uwm.edu.pl

Abstract: Bovine casein is one of the most known precursors of bioactive peptides among food proteins. Thus far, in silico investigations addressing casein have taken no account of the impact of modifications of amino acid residues on the feasibility of bioactive peptide release. The present study aimed to determine the effect of such modification on the possibility of release of bioactive peptides from casein during simulated digestion. The α_{s1-} , α_{s2-} , $\beta-$, and κ -casein sequences were deposited in the BIOPEP-UWM protein database considering phosphorylated amino acids, cysteine residues forming disulfide bridges, and pyroglutamic acid residues. The frequency of occurrence of bioactive fragments and the frequency of their release by digestive enzymes were determined for the analyzed modified and unmodified proteins. Peptides found exclusively in the sequences of unmodified proteins were deemed as false-positive results. From 1.74% (β -casein A2) to 4.41% (α_{s2} -casein B and D) of the false-positive results were obtained for the total frequency of occurrence of bioactive fragments (sums of frequencies computed for all activities). In turn, from 1.78% (κ -casein B) to 9.18% (β -casein A2 and A3) of false-positive results were obtained for the predicted total frequency of release of bioactive peptides by the system of digestive enzymes (pepsin, trypsin, and chymotrypsin).

Keywords: bovine casein; bioactive peptides; BIOPEP-UWM; amino acid sequence analysis; post-translational modifications; phosphorylation



Citation: Minkiewicz, P.;

Darewicz, M.; Iwaniak, A. In Silico Analysis of Individual Fractions of Bovine Casein as Precursors of Bioactive Peptides—Influence of Post-Translational Modifications. *Appl. Sci.* **2023**, *13*, 8091. <https://doi.org/10.3390/app13148091>

Academic Editor: Chiara Cavaliere

Received: 19 June 2023

Revised: 7 July 2023

Accepted: 10 July 2023

Published: 11 July 2023



Copyright: © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Given the importance of modifications of amino acid residues, contemporarily-developed computer programs take account of such modifications in protein sequences [1–5]. Databases annotating modified amino acid residues in proteins have recently been established as well [6–8]. Modifications of amino acid residues are also considered in records of peptide sequences in various databases [9–18] and during prediction of both structure [19] and physicochemical properties of peptides such as, for example, the isoelectric point [20]. There are datasets of non-proteinogenic and modified amino acids, such as the SwissSideChain database [21] and Norine list of monomers [14]. New codes allowing the extension of space of annotated amino acids have been recently proposed [22–25]. Recent trends in encoding of biopolymers (e.g., peptides) have been recently discussed by David et al. [26].

The factors determining biological and functional properties of milk proteins (which affect their behavior during technological processes) include post-translational modifications, such as phosphorylation, and also substitutions of amino acid residues and deletions leading to the occurrence of genetic variants [27–30]. Milk proteins have been extensively studied as precursors of bioactive peptides representing activities of various types [31–36].

Though computer-assisted methods have been widely applied in research addressing bioactive peptides derived from food proteins, including milk [36–40], the impact of modifications of amino acid residues on the profiles of their bioactive peptides has not been explored so far.

In view of the above, the present study aimed to quantify the effect of considering modifications of amino acid residues, including mainly phosphorylation, on the feasibility of bioactive peptide release from individual fractions of bovine milk casein.

2. Materials and Methods

Amino acid sequences of individual casein fractions derived from the UniProt database [41,42] were modified via the removal of signal peptides and the addition of the localization of phosphate residues and L-pyroglutamic acid residues based on information provided in the UniProt database. In the case of phosphorylation, only those residues were considered whose localization had been confirmed experimentally, while those included in the database as theoretically predicted based on the similarity of sequences with other proteins (annotation: “By similarity”) were omitted. Account was also taken of the fact that cysteine residues form disulfide bridges in individual casein fractions [43,44]. In the case of κ -casein, analyses included only the sequences without saccharide residues. The list of post-translational modifications considered in the present work is provided in Table 1.

Table 1. Modified amino acids considered in sequences of the analyzed proteins.

Name	Symbol	ID in BIOPEP-UWM ¹ Database	CID in PubChem ² Database
Phosphoserine	<S[3*]>	108	68,841
Phosphothreonine	<T[3*]>	109	3,246,323
L-Pyroglutamic acid	<P[4O]>	110	7405
Cysteine constituting disulfide bridges	<C>	111	5862

¹ Repository of amino acids and modifications in the BIOPEP-UWM database [45–47]. ² PubChem database [48,49].

Amino acid sequences were deposited in the BIOPEP-UWM database of proteins [33]. Substitutions of amino acids or deletions of fragments in individual genetic variants were provided based on information derived from the UniProt database and a review by Caroli et al. [50]. Protein sequences used in this work are presented in Table S1.

Profiles of potential bioactivity of protein fragments and simulation of protein proteolysis by the coupled action of digestive enzymes were performed for the chosen protein sequences: pepsin at pH > 2 (EC 3.4.23.1, BIOPEP-UWM ID 39), trypsin (EC 3.4.21.4, BIOPEP-UWM ID 12), and chymotrypsin A (EC 3.4.21.1, BIOPEP-UWM ID 11), following the INFOGEST experimental protocol [51]. Bioactive fragments were searched for among the predicted products of proteolysis.

The analyzed sequences were also quantified for the frequency of occurrence of bioactive fragments in protein sequences (parameter A) [52] (Equation (1)).

$$A = a/N \quad (1)$$

a—number of fragments with a given activity in a protein sequence.

N—number of amino acid residues in a protein sequence.

The second analyzed quantitative parameter was the frequency of release of fragments with a given bioactivity as a result of the coupled action of the aforementioned proteolytic enzymes (parameter A_E) [53], computed based on Equation (2):

$$A_E = d/N \quad (2)$$

d—number of peptides with a given activity potentially released by the above-mentioned enzyme combination from a given protein.

N—number of amino acid residues in a protein sequence.

Computer simulations of proteolysis and computations of quantitative parameters were performed using applications available in the BIOPEP-UWM database [46,47].

The casein fractions with and without post-translational modifications were compared in terms of the frequency of occurrence of bioactive fragments in sequences of individual proteins.

The percentage of false-positive (FP) results calculated from Equations (3) and (4) was used as the parameter describing the effect of considering the modifications on the computed frequency of occurrence of bioactive fragments in a protein sequence (A) and the frequency of release of fragments with a given activity by digestive enzymes (A_E):

$$FP = 100\%(A_0 - A)/A_0 \quad (3)$$

A_0 —value of parameter A determined for non-modified protein.

A —value of parameter A determined for protein containing post-translational modifications.

$$FP = 100\%(A_{E0} - A_E)/A_{E0} \quad (4)$$

A_{E0} —value of parameter A_E determined for non-modified protein.

A_E —value of parameter A_E determined for protein containing post-translational modifications.

Identity between a protein sequence with post-translation modifications and a sequence of the same protein without modifications was computed based on the following equation:

$$I = 100\%(N - M)/N \quad (5)$$

N —number of amino acid residues in a protein chain.

M —number of modified amino acid residues in a protein chain.

Visualization of the results in the figures was made by means of HeatMapper software [54,55].

Biological activity (interaction with human body proteins) of di- and tri-peptides containing post-translation modifications was predicted using SwissTargetPrediction software [56,57].

3. Results

The values of A and A_E parameters computed for the sequences with and without post-translational modifications are presented in Figures 1 and 2 and in Tables S2–S9 in the Supplementary Materials. The figures present only those activities for which the predicted values of quantitative parameters changed as a result of considering post-translational modifications in calculations. The values of quantitative parameters determined for all peptides occurring in sequences of the analyzed proteins or being predicted products of their digestion are collated in Tables S2–S9. The percentage of false positive results obtained due omission of post-translational modifications is presented in Figure 3 and Tables S10 and S11.

Consideration of the post-translational modifications during analyses led to changes in the numeric values of parameters characterizing proteins as precursors of bioactive peptides only in the case of certain activities. The percentage of these activities determined for individual fractions and genetic variants of casein is presented in Table 2.

One of the aims of protein amino acid sequence analysis is to predict the likelihood of bioactive peptide release during protein proteolysis followed by their detection deploying a targeted peptidomics strategy. An example of such an experiment was provided in a work by Darewicz et al. [58], describing the detection of bioactive peptides among products of enzymatic hydrolysis of oat proteins. Table 3 lists peptides with a known activity, being putative products of digestion of individual casein fractions.

The putative products of casein digestion include 46 peptides with a known bioactivity, described in the BIOPEP-UWM database. Among these, 38 peptides exhibit more than one activity. The highest number of identified peptides include dipeptidyl peptidase IV inhibitors—36 compounds and angiotensin-converting enzyme inhibitors—24 compounds. Peptides exhibiting these activities have been most extensively investigated in recent years and represented in the highest numbers in the BIOPEP-UWM database [45]. Four of the peptides listed in Table 3 inhibit dipeptidyl peptidase III. Three inhibit α -glucosidase,

and another three exhibit antioxidative activity. Nine of the listed peptides exhibit other types of activity. Peptides identified as inhibitors of the angiotensin-converting enzyme, dipeptidylpeptidase IV, and α -glucosidase as well as those exhibiting antioxidative activity may be valuable in terms of preventing the so-called metabolic syndrome [59].

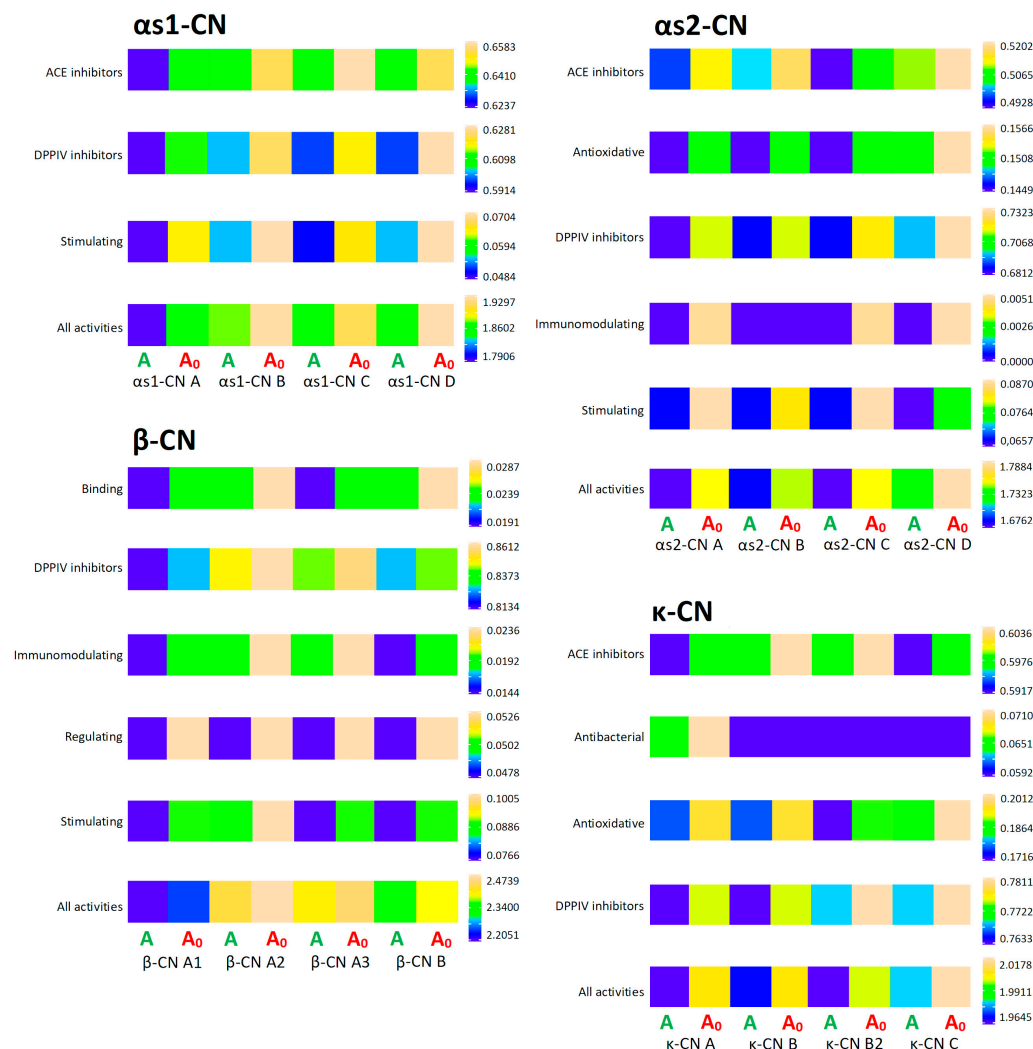


Figure 1. The frequency of occurrence of bioactive fragments (Parameter A; Equation (1)) in sequences of individual casein fractions with (A) and without (A_0) modified amino acids. The figure presents only those activities for which differences were found between the modified and unmodified proteins. Results determined for all activities of casein fragments are provided in Tables S2–S5 in Supplementary Materials.

Simulated *in silico* proteolysis enables predicting the release of peptides containing modified amino acid residues, including mainly those of phosphopeptides. The bioactivity of casein hydrolysates, composed of phosphopeptides, has already been investigated [35]; however, the activities of individual compounds have not been registered in the BIOPEP-UWM database so far. For this reason, theoretical predications of bioactivity were performed by means of the SwissTargetPrediction software. Authors of this software [57] recommend it as a tool for the preliminary analysis of interactions of low-molecular-weight compounds with proteins that enables predicting the bioactivity of di- and tri-peptides.

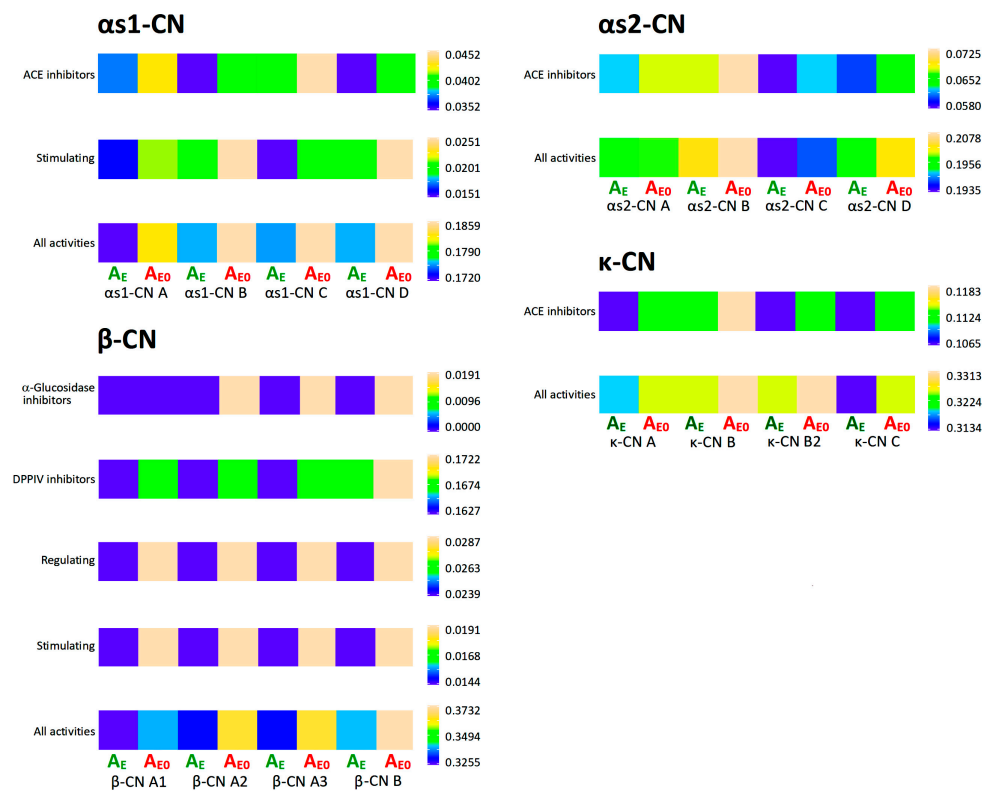


Figure 2. Frequency of predicted release of bioactive fragments (Parameter A_E ; Equation (2)) from individual casein fractions with (A_E) and without (A_{E0}) modified amino acids. The figure presents only those activities for which differences were found between the modified and unmodified proteins. Results determined for all activities of casein fragments are provided in Tables S6–S9 in Supplementary Materials.

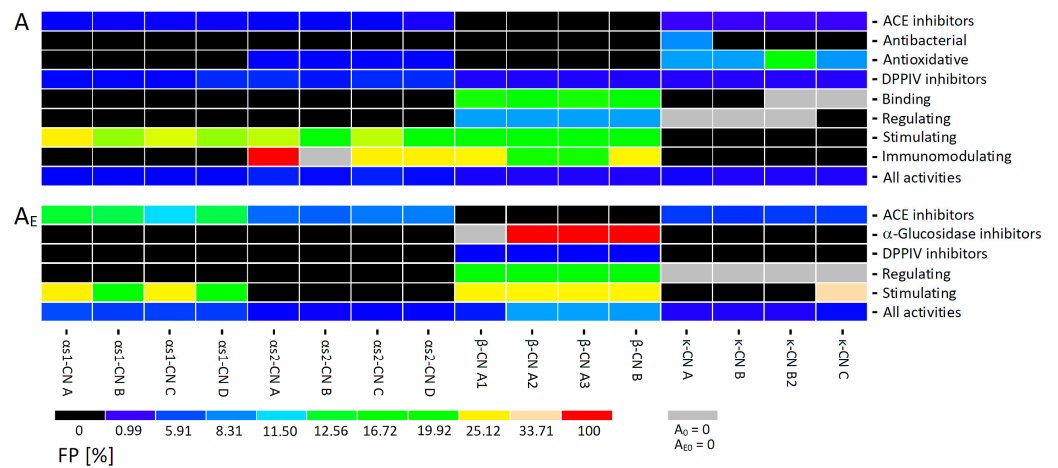


Figure 3. Percentage of false-positive (FP) results for the frequency of occurrence of bioactive fragments in sequences of individual proteins (parameter A) and frequency of release of bioactive fragments during digestion of individual proteins (parameter A_E), computed based on Equations (3) and (4), respectively. Gray color means no possibility of computing FP values based on Equation (3) or (4) (dividing by $A_0 = 0$ or $A_{E0} = 0$). Activities with FP = 0 for all analyzed proteins were omitted. Numeric values of FP for A and A_E parameters are provided in Tables S10 and S11, respectively, in the Supplementary Materials.

Table 2. Percentage of activities in the case of which considering the post-translational modifications affects the values of A and A_E parameters.

Protein	Variant	I [%] ¹	Percentage of Activities in the Case of Which Modifications Affect the Value of Parameter A [%] ^{2,3}	Percentage of Activities in the Case of Which Modifications Affect the Value of Parameter A _E [%] ^{3,4}
α _{s1} -CN	A	95.70	13.04	33.33
	B	95.98	12.50	33.33
	C	95.98	12.50	33.33
	D	95.48	12.50	33.33
α _{s2} -CN	A	94.20	26.32	20.00
	B	94.69	22.22	20.00
	C	94.20	26.32	20.00
	D	95.45	26.32	20.00
β-CN	A1	97.61	17.86	37.50
	A2	97.61	16.13	44.44
	A3	97.61	16.13	44.44
	B	97.61	16.13	44.44
κ-CN	A	97.63	17.39	12.50
	B	97.63	13.04	12.50
	B2	97.63	13.04	12.50
	C	97.63	13.04	12.50

¹ Identity between a protein sequence containing modifications of amino acid residues and a sequence of the same protein devoid of modifications, computed according to Equation (5). ² Value of parameter A computed from Equation (1). ³ Complete list of activities of fragments found in proteins and released upon in silico proteolysis is presented in Tables S2–S9 in the Supplementary Materials. ⁴ Values of parameter A_E computed according to Equation (2).

Table 3. List of bioactive peptides being putative products of digestion of individual casein fractions. Peptide IDs are taken from the BIOPEP-UWM database of bioactive peptides.

Peptide	Protein—Peptide Precursor	ACEi ¹	DPPiV ²	Antioxidative ³	DPPIII ⁴	α-Gluc ⁵	Others
IA	κ-CN A; B; B2; C	7562	8525				
IE	β-CN A1; A2; A3; B κ-CN A; B; B2; C	7827					
IG	α _{s1} -CN A; B; C; D; α _{s2} -CN A; B; C; β-CN A1; κ-CN B	7595					
IH	α _{s1} -CN A; B; C; D; β-CN A1; A2; A3; B		8800		9497		8323 ⁶
IL	κ-CN A; B; B2; C α _{s2} -CN A; B; C; D;	9079	8802				
IN	β-CN A1; A2; A3; B; κ-CN A; B; B2; C		8804				
IP	κ-CN A; B; B2; C	7581	8501				
IPY	α _{s2} -CN A; B; C; D α _{s1} -CN A; B; C; D;	7803					
IQ	α _{s2} -CN A; B; C; D; β-CN A1; A2; A3; B; κ-CN A; B; B2; C		8805				
IR	κ-CN A; B; B2; C	3258	8806	8215			8246 ⁷ ; 8247 ⁸
PA	κ-CN A; B; B2; C		3179				
PE	α _{s1} -CN A; B; C; D; β-CN A1; A2; A3; B				9504	9694	
PF	α _{s1} -CN A; B; C; D; β-CN A1; A2; A3; B		8854			9505	
PG	β-CN A1; A2; A3; B	7625	8855				2754 ⁹ ; 3285 ¹⁰ ; 3460 ¹¹
PH	β-CN A1; A2; A3; B; κ-CN A; B; B2; C	7843	8856				
PK	α _{s1} -CN A; B; C; D; α _{s2} -CN A; B; C; D; β-CN A1; A2; A3; B		8858				

Table 3. Cont.

Peptide	Protein—Peptide Precursor	ACEi ¹	DPPiVi ²	Antioxidative ³	DPPIIIi ⁴	α -Gluc ⁵	Others
PL	α_{s1} -CN A; B; C; D; β -CN A1; A3; B	7513	8638				
PM	α_{s1} -CN A; B; C; D		8859				
PPL	β -CN A1;	7513					
PQ	α_{s1} -CN A; B; C; D; α_{s2} -CN A; B; C; D; β -CN A1; A2; A3; B	7837	8861				
PSY	κ -CN A; B; B2	7559					
PT	α_{s2} -CN A; B; C; D; κ -CN A; B; B2; C	7833	8638				
PW	α_{s2} -CN A; B; C; D		8865	8190			
PY	κ -CN A; B; B2; C		8866				9856 ¹²
SE	α_{s1} -CN A; B; C; D; α_{s2} -CN A; B; D						8330 ¹³
SF	κ -CN A; B; B2; C	7685	8891				9432 ⁷
SG	α_{s1} -CN C	7618					
SK	α_{s1} -CN A; B; C; D; α_{s2} -CN A; B; C; D; β -CN A1; A2; A3; B		8894				
SL	β -CN A1; A2; A3; B		8560				9955 ¹⁴
SM	α_{s1} -CN A; B; C; D				9507		
ST	α_{s2} -CN A; B; D; κ -CN A; B; B2; C	9184					
SW	β -CN A1; A2; A3; B		8896				
VA	α_{s1} -CN B; C; D; α_{s2} -CN A; B; C; D; κ -CN A; B; B2		3172				
VD	α_{s2} -CN A; B; C		8915				
VE	α_{s1} -CN A; B; C; D; β -CN A1; A2; A3; B; κ -CN A; B; B2; C	7829	8916			9693	
VF	α_{s1} -CN A; B; C; D; α_{s2} -CN A; B; C; D	3384	8917				
VK	α_{s2} -CN A; B; D; β -CN A1; A2; B	7558	8921				
VL	α_{s1} -CN B; C; D; α_{s2} -CN A; B; C; D; β -CN A1; A2; A3; B; κ -CN A; B; B2; C		8922				8320 ⁶
VM	β -CN A1; A2; A3	9882	8923				
VN	α_{s1} -CN A; B; C; D		8924				
VP	α_{s2} -CN A; B; C; D	7587	3181				
VPL	α_{s1} -CN A; B; C; D		8347				3166 ¹¹ ; 3350 ¹³
VQ	β -CN A1; κ -CN A; B; B2; C		8925				
VR	α_{s2} -CN A; B; C; D; β -CN A1; A2; A3; B κ -CN B2; C	7628	8594				
VT	κ -CN A; B; B2; C		8927				
VY	α_{s2} -CN A; B; C; D; β -CN A1; A2; B	3492	8929	8224	9509		

¹ Angiotensin-converting enzyme (ACE) (EC 3.4.15.1) inhibitor; ² Dipeptidyl peptidase IV (DPPiV) (EC 3.4.14.5) inhibitor; ³ Antioxidative peptide; ⁴ Dipeptidyl peptidase III (DPPIII) (EC 3.4.14.4) inhibitor; ⁵ α -Glucosidase (EC 3.2.1.20) inhibitor; ⁶ Glucose uptake stimulating peptide (annotated in the BIOPEP-UWM database as stimulating); ⁷ Renin (EC 3.4.23.15) inhibitor; ⁸ Calmodulin-dependent phosphodiesterase 1 (CaMPDE) (EC 3.1.4.17) inhibitor; ⁹ Peptide regulating the stomach mucosal membrane activity (annotated in the BIOPEP-UWM database as regulating); ¹⁰ Antithrombotic peptide; ¹¹ Prolyl oligopeptidase (EC 3.4.21.26) inhibitor (annotated in the BIOPEP-UWM database as anti-amnesic peptide); ¹² Anti-inflammatory peptide; ¹³ Peptide stimulating vasoactive substance secretion (annotated in the BIOPEP-UWM database as stimulating); ¹⁴ Phosphoglycerate kinase (EC 2.7.2.3) activity regulator (annotated in the BIOPEP-UWM database as regulating).

Table 4 lists proteins showing the highest likelihood of interactions with modified peptides.

Table 4. Di- and tri-peptides containing post-translational modifications, potentially released via simulated in silico digestion, and their potential interactions with proteins predicted using the SwissTargetPrediction software. Proteins being receptors of ligands containing phosphate residues or enzymes catalyzing reactions with compounds containing such residues are denoted in italics.

Peptide ¹	Protein—Peptide Precursor	Protein Whose Predicted Ligand Is a Peptide ^{2,3,4}
<P[4O]>E	κ-CN A; B; B2; C	HLA-A (UniProt: P04439; ChEMBL: CHEMBL2632); <i>p</i> = 0.0989 DPPIV (UniProt: P27487; ChEMBL: CHEMBL284); <i>p</i> = 0.0989 XIAP (UniProt: P98170; ChEMBL: CHEMBL4198); <i>p</i> = 0.0989 <i>LPAR3</i> (UniProt: Q9UBY5; ChEMBL: CHEMBL3250); <i>p</i> = 0.1453 <i>PLK1</i> (UniProt: P53350; ChEMBL: CHEMBL3024); <i>p</i> = 0.1133 <i>LPAR1</i> (UniProt: Q92633; ChEMBL: CHEMBL3819); <i>p</i> = 0.1133 BRCA1 (UniProt: P38398; ChEMBL: CHEMBL5990); <i>p</i> = 0.1133 <i>PIN1</i> (UniProt: Q13526; ChEMBL: CHEMBL2288); <i>p</i> = 0.1133
P<S[3*]>K	α _{s2} -CN A; B; C; D	FPR2 (UniProt: P25090; ChEMBL: CHEMBL4227); <i>p</i> = 0.1133 <i>S1PR5</i> (UniProt: Q9H228; ChEMBL: CHEMBL2274); <i>p</i> = 0.1133 <i>S1PR2</i> (UniProt: O95136; ChEMBL: CHEMBL2955); <i>p</i> = 0.1133 HLA-A (UniProt: P04439; ChEMBL: CHEMBL2632); <i>p</i> = 0.1133 <i>PRKCE</i> (UniProt: Q02156; ChEMBL: CHEMBL3582); <i>p</i> = 0.1133 TNFRSF10A (UniProt: O00220; ChEMBL: CHEMBL3551); <i>p</i> = 0.1133 <i>S1PR5</i> (UniProt: Q9H228; ChEMBL: CHEMBL2274); <i>p</i> = 0.0604 <i>S1PR2</i> (UniProt: O95136; ChEMBL: CHEMBL2955); <i>p</i> = 0.0604
<S[3*]>A	α _{s1} -CN A; B; C; α _{s2} -CN A; B; C; D	<i>S1PR3</i> (UniProt: Q99500; ChEMBL: CHEMBL3892); <i>p</i> = 0.0604 <i>S1PR1</i> (UniProt: P21453; ChEMBL: CHEMBL4333); <i>p</i> = 0.0604 <i>LPAR3</i> (UniProt: Q9UBY5; ChEMBL: CHEMBL3250); <i>p</i> = 0.0604 <i>S1PR</i> (UniProt: Q9H228; ChEMBL: CHEMBL2274); <i>p</i> = 0.0429 <i>S1PR2</i> (UniProt: O95136; ChEMBL: CHEMBL2955); <i>p</i> = 0.0429
<S[3*]>G	κ-CN A; B2; C	<i>S1PR3</i> (UniProt: Q99500; ChEMBL: CHEMBL3892); <i>p</i> = 0.0429 <i>S1PR1</i> (UniProt: P21453; ChEMBL: CHEMBL4333); <i>p</i> = 0.0429 <i>LPAR3</i> (UniProt: Q9UBY5; ChEMBL: CHEMBL3250); <i>p</i> = 0.0429 <i>LPAR3</i> (UniProt: Q9UBY5; ChEMBL: CHEMBL3250); <i>p</i> = 0.1006 <i>S1PR5</i> (UniProt: Q9H228; ChEMBL: CHEMBL2274); <i>p</i> = 0.1006 <i>S1PR2</i> (UniProt: O95136; ChEMBL: CHEMBL2955); <i>p</i> = 0.1006 <i>S1PR3</i> (UniProt: Q99500; ChEMBL: CHEMBL3892); <i>p</i> = 0.1006 <i>S1PR1</i> (UniProt: P21453; ChEMBL: CHEMBL4333); <i>p</i> = 0.1006
<S[3*]>E	α _{s1} -CN A; B; C; D; β-CN A1; A2; A3; B	HLA-A (UniProt: P04439; ChEMBL: CHEMBL2632); <i>p</i> = 0.1006 <i>LPAR3</i> (UniProt: Q9UBY5; ChEMBL: CHEMBL3250); <i>p</i> = 0.1120 <i>S1PR5</i> (UniProt: Q9H228; ChEMBL: CHEMBL2274); <i>p</i> = 0.1120 <i>S1PR2</i> (UniProt: O95136; ChEMBL: CHEMBL2955); <i>p</i> = 0.1120 <i>S1PR3</i> (UniProt: Q99500; ChEMBL: CHEMBL3892); <i>p</i> = 0.1120 <i>S1PR1</i> (UniProt: P21453; ChEMBL: CHEMBL4333); <i>p</i> = 0.1120
<S[3*]>L	β-CN A1; A2; A3; B	<i>TK1</i> (UniProt: P04183; ChEMBL: CHEMBL2883); <i>p</i> = 0.1120 <i>LPAR3</i> (UniProt: Q9UBY5; ChEMBL: CHEMBL3250); <i>p</i> = 0.1062 <i>PLK1</i> (UniProt: P53350; ChEMBL: CHEMBL3024); <i>p</i> = 0.1062 BRCA1 (UniProt: P38398; ChEMBL: CHEMBL5990); <i>p</i> = 0.1062
<S[3*]><S[3*]>E	α _{s2} -CN B	HLA-A (UniProt: P04439; ChEMBL: CHEMBL2632); <i>p</i> = 0.1062 <i>LPAR1</i> (UniProt: Q92633; ChEMBL: CHEMBL3819); <i>p</i> = 0.1062 <i>DLG4</i> (UniProt: P78352; ChEMBL: CHEMBL5666); <i>p</i> = 0.1062 <i>PIN1</i> (UniProt: Q13526; ChEMBL: CHEMBL2288); <i>p</i> = 0.1062 <i>LPAR3</i> (UniProt: Q9UBY5; ChEMBL: CHEMBL3250); <i>p</i> = 0.0536 <i>S1PR5</i> (UniProt: Q9H228; ChEMBL: CHEMBL2274); <i>p</i> = 0.0536
<S[3*]>T	α _{s1} -CN A; B; C; D; α _{s2} -CN A; B; C; D	<i>S1PR2</i> (UniProt: O95136; ChEMBL: CHEMBL2955); <i>p</i> = 0.0536 <i>S1PR3</i> (UniProt: Q99500; ChEMBL: CHEMBL3892); <i>p</i> = 0.0536 <i>S1PR1</i> (UniProt: P21453; ChEMBL: CHEMBL4333); <i>p</i> = 0.0536
<T[3*]>M	α _{s1} -CN D	<i>FNTA</i> ; <i>FNTB</i> (UniProt: P49354; P49356; ChEMBL: CHEMBL2094108); <i>p</i> = 0.1006 <i>TK1</i> (UniProt: P04183; ChEMBL: CHEMBL2883); <i>p</i> = 0.1006 <i>FDPS</i> (UniProt: P14324; ChEMBL: CHEMBL1782); <i>p</i> = 0.1006

¹ Symbols of modified amino acids are explained in Table 1; ² Table also presents IDs of proteins available in the UniProt [41,42] and ChEMBL [60,61] databases; ³ *p*—probability that a given peptide will be a ligand of a given protein; ⁴ Table lists proteins for which the probability of interactions with the ligand is the highest among all proteins typed by the SwissTargetPrediction software.

4. Discussion

A comparison of the frequency of occurrence of bioactive fragments (A and A_0) and the frequency of their predicted release during proteolysis (A_E and A_{E0}) indicates that taking account of the post-translation modifications during computations leads to decreased values of these parameters found for many activities. Part of the bioactive fragments containing serine (S symbol) do not appear in the results obtained for the sequences containing post-translation modifications because these sequences contain phosphoserine denoted with the $\langle S[3^*] \rangle$ symbol. A similar observation was made for the cysteine residues involved in the formation of disulfide bridges (denoted with $\langle C \rangle$). Di-peptides and tri-peptides account for major part of bioactive fragments found in both the analyzed protein sequences as well as in sequences of many other proteins [62]. The sites of their occurrence may, accidentally, coincide with the localization of modifications, if the latter are not taken into account during analysis. Prediction of the occurrence of peptides containing non-modified amino acid residues at the localization sites of post-translation modifications or considering such peptides while quantifying parameters that characterize proteins as potential precursors of bioactive fragments should, therefore, be treated as a false-positive result. Such a result means that also other fragments do appear in real protein sequences that are different than those predicted based on sequence analysis taking no account of modifications.

The qualitative analysis results (i.e., “yes” or “no” responses) are divided into four categories: true-positive, false-positive, true-negative, and false-negative results. The percentages of results from these categories are used to assess, e.g., methods of qualitative analysis deployed in analytical chemistry [63]. The present analysis may allow defining a true-positive result (the presence or the probability of release of a bioactive fragment of both a protein containing post-translational modifications or a protein having an identical amino acid sequence but not containing such modifications) or a false-positive result (the presence of a bioactive fragment or the possibility of its release from only a protein devoid of modification, as well as a lack of such a fragment or no possibility of its release during hydrolysis of a protein having an identical amino acid sequence and containing modifications). The value of parameter A or A_E (computed from Equations (3) and (4), respectively) may be interpreted as the number of true-positive results per protein chain length. The $A_0 - A$ or $A_{E0} - A_E$ difference may be interpreted as the number of false-positive results per protein chain length.

The false-negative results cannot be defined while analyzing protein sequences as precursors of bioactive peptides. Such a result would mean a lack of the fragment with a given activity in a protein sequence. In reality, the negative result achieved for a given activity means a lack of information about such fragments. There is no complete information on all possible types of biological activity of all possible fragments of a given protein. The false-negative result would mean that a given fragment would occur in a sequence of the modified protein, whereas the sequence analysis of the unmodified protein would not indicate its presence. In the present analysis, the following inequality has not been met for any activity: $A > A_0$ or $A_E > A_{E0}$. Such an inequality could have been met if information was available about the activity of peptides containing phosphoryl amino acid residues or L-pyroglutamic acid residue. It cannot be excluded, however, that activities of such peptides will be investigated in the future. Antioxidative activity of peptides containing L-pyroglutamic acid has been recently described by Vitale et al. [64].

Identity between a protein sequence containing post-translational modifications and a sequence of the same protein devoid of modifications computed based on Equation (5) (parameter I) has a similar physical sense as the identity computed using specialist software for matching protein sequences such as, for example, BLAST [65], Clustal Omega [66], or HMMER [67]. Equation (5) is very simplified compared to the aforementioned computer programs because it takes no account of insertion and deletion of protein fragments and considers only the equivalent of amino acid substitutions.

Even though the analyzed proteins contain a relatively low number of modified amino acid residues, the changes noted in the predicted values of parameters describing proteins as precursors of bioactive peptides seem to be significant. The appearance of false-positive results for, e.g., four out of nine activities of fragments potentially released from β -casein, is likely (Table 2).

The percentage of false-positive results obtained for particular types of activities varied within a broad range. $FP = 0$ means that the frequency of occurrence of bioactive fragments in protein sequences or the predicted frequency of release of bioactive fragments are identical for the proteins with post-translational modifications considered and for the same sequences analyzed without considering modifications ($A = A_0$ or $A_E = A_{E0}$). In turn, $FP = 100\%$ indicates that the value of the analyzed parameter is 0 for the proteins with post-translational modifications considered and differs from 0 for the same proteins analyzed without considering these modifications ($A = 0$ and $A_0 \neq 0$ or $A_E = 0$ and $A_{E0} \neq 0$). In the case of certain activities and certain genetic variants of individual casein fractions, the bioactive fragments do not occur either in modified or in unmodified sequences ($A = 0$ and $A_0 = 0$ or $A_E = 0$ and $A_{E0} = 0$). In such a case, computing the percentage of false-positive results is impossible. Predictions of frequent activity types (relatively high values of A and A_E), such as inhibition of angiotensin-converting enzyme (EC 3.4.15.1) or inhibition of dipeptidylpeptidase IV (EC 3.4.14.5), and also a sum of all activities yield a relatively low percentage of false-positive results. In the case of the less known activities (relatively low values of A and A_E), the percentage of false-negative results ranged from 0 to 100%.

Among proteins summarized in Table 4, HLA-A is a protein of the immune system responsible for presenting antigens, e.g., those derived from viruses or neoplasms [68–70]. DPPIV (EC 4.4.14.5) is a commonly known proteolytic enzyme whose inhibitors are applied as anti-diabetic drugs [59,71]. Recent studies have focused on DPPIV role in the course of also other diseases, such as Alzheimer's disease [72] or neoplasms [73]. The XIAP protein is ubiquitin ligase (EC 2.3.2.27), involved in apoptotic processes, modulation of inflammatory signals and immune response, and cell spreading, including tumor metastases [74–77]. The LPAR1 and LPAR3 proteins are receptors of lysophosphatidic acid (PubChem CID: 5497152; ChEMBL ID: ChEMBL117021), which together with its receptors are involved in signal transfer in cells of, i.e., nervous, respiratory, and gastrointestinal systems [78]. In turn, the PLK1 (EC 2.7.11.21) and PRKCE (EC 2.7.11.13) proteins are kinases catalyzing the phosphorylation of amino acid residues (serine in particular) in proteins. Phosphorylation is the most common type of post-translational modifications of proteins in live organisms [79,80]. BRCA1 (Breast cancer type 1 susceptibility protein) (EC 2.3.2.27) is a ligase catalyzing protein ubiquitination. Its malfunction due to, among other things, mutations has been implicated in the course of breast cancer [81,82]. The PIN1 protein is an enzyme (EC 5.2.1.8) catalyzing the reaction of cis-trans isomerization of peptide bonds formed by amine groups of proline and carboxyl groups of phosphorylated serine or threonine. Its malfunction is observed, e.g., in the course of the Alzheimer's disease [83], whereas its over-expression and hyper-activity play a significant role in carcinogenesis [84]. The FPR2 acronym (N-formyl peptide receptor 2) denotes a protein being a receptor of chemotactic peptides possessing a formic acid residue at the N-terminus. These receptors play a meaningful role during inflammatory states and angiogenesis [85]. The latter process is essential to wound healing and also to cancer development. A group of proteins referred to as S1PR (Sphingosine 1-phosphate receptor) serves the function of sphingophospholipid receptors. Their hyperactivity is observed in the course of non-specific inflammatory bowel disease. Compounds that modulate the activity of S1PR receptors have been investigated as potential drugs to be used in this disease condition [86]. The use of S1PR receptor modulators is also considered in the treatment of multiple sclerosis [87]. The TNFRSF10A (Tumor necrosis factor receptor superfamily member 10A) protein (UniProt: P50591) is a receptor of a cytotoxic ligand TRAIL that stimulates apoptosis [88]. The protein abbreviated as TK1 (EC 2.7.1.21) is a kinase accelerating the phosphorylation of thymidine nucleoside, being one of the key enzymes to the metabolism of nucleosides and nucleotides and, thereby, to

DNA synthesis [89]. The DLG4 (Disks large homolog 4) compound is a homolog of proteins that play a key role in the formation of synapses [90]. The FNT acronym (FNTA; FNTB) denotes farnesyltransferase (EC 2.5.1.58), i.e., an enzyme involved in the modification of cysteine residues by attaching oligo-isoprene residues. Its malfunction contributes to the improper pattern of post-translational modifications of proteins in certain neoplasms [91]. Also, peptides built of the non-modified protein amino acids may—theoretically—serve as farnesyltransferase inhibitors [92]. The FDPS (Farnesyl diphosphate synthase) or FPPS (Farnesyl pyrophosphate synthase) (EC 2.5.1.1) enzyme is a synthetase of farnesyl diphosphate (pyrophosphate) (PubChem: No 445713), being a substrate of farnesyltransferase.

Summing up data collated in Table 4, it may be concluded that most of the proteins listed therein are related to either reactions (as enzymes) or bioactivities (as receptors) of compounds containing phosphate residues (e.g., of phospholipids). The malfunctions of these proteins are observed in the course of various diseases such as, for example, neoplasms. In the case of peptides built from unmodified protein amino acids, proteolytic enzymes serve as the main group of potentially interacting proteins [93]. The highest probability of peptide–protein interactions, presented in Table 4, is lower than the analogous probability predicted for the peptides described in previous works [92–94]. The difference in the probability of interactions with proteins noted between the peptides built of the non-modified amino acids and phosphopeptides may stem from two reasons. The first entails differences in the structure and physicochemical properties triggered by modifications, whereas the other one may be due to knowledge gaps. The SwissTargetPrediction software compares the structure of analyzed compounds with data from the ChEMBL database [57]. This database provides a load of experimental data related to the interactions of short, unmodified peptides with proteins [95], while it lacks information about phosphopeptides and, therefore, has insufficient data for comparative analyses. Further studies are, hence, needed to resolve the doubts regarding the putative bioactivity of phosphopeptides.

Considering that modifications of amino acid residues would enable more precise description of the bioactivity of protein fragments, the sequence analysis made took no account of such modifications results in the appearance of false-positive results. On the other hand, the analysis of sequences built exclusively from protein amino acids is simpler than taking into account the changes induced by chemical or enzymatic modifications of amino acid residues. This fact may prove essential while working with large sets of sequences. Bovine milk proteins (including casein) have been thoroughly investigated as precursors of bioactive peptides. These investigations addressed even the impact of single substitutions of amino acids, leading to the formation of genetic variants of individual casein fractions of various species [96–100]. Certain cases require a different strategy, wherein large sets of sequences can be reduced. After selecting the first sequence, proteins with an identity higher than a pre-determined threshold, e.g., 90%, are eliminated. Such a strategy has been deployed in our previous research [58,101,102]. The proteins with 90% identity may be expected to belong to the same family defined based on the presence of appropriate domains, described in the InterPro database [103]. Proteins belonging to the same family share similar profiles of potential biological activity [104]. Substitutions of amino acid residues cannot be deemed the drivers of false-positive results but can be the likely reasons of changes in the contents of bioactive fragments. Considering or neglecting differences in the contents of bioactive fragments in analyses is up to the decision of individual authors. Taking account of modifications may be recommended in the analysis of peptide sequences with a relatively high content of untypical amino acids such as, for example, peptides containing hydroxyproline and methionine sulfoxide described by Bougatef et al. [105].

Peptides containing the modified residues may exhibit bioactivity. Even when the methodology of their isolation or synthesis is unavailable, they may be investigated by means of *in silico* methods [106,107]. Due to the established repository of amino acids and modifications, the BIOPEP-UWM database may prove well in such analyses.

Several databases of bioactive peptides, containing components other than proteino-genic amino acids, have been recently launched [9–18]. In comparison with them, BIOPEP-UWM possesses both strong points and limitations. Three main advantages of the BIOPEP-UWM database may be pointed out:

- Possibility of matching peptide and protein sequences, containing modified amino acids annotated using code available in the database;
- Possibility of proteolysis simulation using sequences with modified residues;
- Possibility of easy translation of sequences into SMILES code, used in chemical databases (e.g., PubChem) and programs from the area of cheminformatics (e.g., SwissTargetPrediction). It is possible due to direct input of modified residues into sequences. For instance, proteins in UniProt database are annotated as sequences containing unmodified proteino-genic amino acid residues, with information about modifications added as a text. Only unmodified sequences in UniProt are available for processing (e.g., sequence alignments).

Possibilities of annotation in the BIOPEP-UWM database are restricted to linear pep-tides or proteins. For instance, the current version of BIOPEP-UWM code allows for the provision of information that the given cysteine residue is involved in the formation of disulfide bond, but there is no detailed location of this bond in peptide or protein se-quence (Table S1 in Supplementary Materials). On the other hand, macrocyclic or branched peptides may be annotated using, for example, BILN code [25], but there is no simple algorithm for translation of such peptide structures into SMILES code. Another restriction of BIOPEP-UWM is that it provides only possibility of exact match, i.e., finding in protein sequences fragments with 100% identity and the same length as queries. However, all existing algorithms and programs for sequence alignments (e.g., BLAST, Clustal Omega, and HMMER) accept sequences containing only unmodified, proteinogenic amino acids. Development of algorithms and programs for automated sequence alignments incorporat-ing post-translational modifications may require radically new ideas. The third restriction of BIOPEP-UWM is that this tool is able to annotate and process sequences containing only residues present in repository of amino acids and modifications. Enrichment of space of peptides and proteins annotated in BIOPEP-UWM may require prior submission of new amino acid residues or other monomers. The addition of individual monomers to the repository may depend on the possibility of adapting its SMILES representation to the CHUCKLES algorithm [108] and further validation of the result of rearrangement as recommended previously [109]. Taking the above into account, we would like to en-courage BIOPEP-UWM users to submit new compounds to the repository of amino acids and modifications.

5. Conclusions

Today, the BIOPEP-UWM database serves as a tool enabling the analysis of peptide and protein sequences with amino acid residues that underwent chemical and enzymatic modifications (e.g., of phosphopeptides and phosphoproteins). The likely benefit of imple-menting the new options is the increased precision of predictions of bioactivity of protein fragments. On the one hand, taking into account modifications of amino acid residues makes it possible to avoid false-positive results, while on the other hand, it allows for research on modified peptides. The analysis of individual casein fractions presented in this paper is the first example of a comprehensive sequence analysis of proteins containing modified amino acid residues as precursors of biologically active peptides.

Supplementary Materials: The following supporting information can be downloaded at: <https://www.mdpi.com/article/10.3390/app13148091/s1>, Table S1: Protein sequences used for analyses; Table S2: Frequency of occurrence of bioactive fragments A and A₀ in individual genetic variants of α_{s1} -casein; Table S3: Frequency of occurrence of bioactive fragments A and A₀ in individual genetic variants of α_{s2} -casein; Table S4: Frequency of occurrence of bioactive fragments A and A₀ in individual genetic variants of β -casein; Table S5: Frequency of occurrence of bioactive fragments A and A₀ in individual genetic variants of κ -casein; Table S6: Frequency of predicted release of bioactive fragments A_E and A_{E0} in individual genetic variants of α_{s1} -casein; Table S7: Frequency of predicted release of bioactive fragments A_E and A_{E0} in individual genetic variants of α_{s2} -casein; Table S8: Frequency of predicted release of bioactive fragments A_E and A_{E0} in individual genetic variants of β -casein; Table S9: Frequency of predicted release of bioactive fragments A_E and A_{E0} in individual genetic variants of κ -casein; Table S10: Percentage of false-positive (FP) results for the frequency of occurrence of bioactive fragments in sequences of individual proteins (parameter A); Table S11: Percentage of false-positive (FP) results for the frequency of release of bioactive fragments during digestion of individual proteins (parameter A_E).

Author Contributions: Conceptualization, P.M., M.D. and A.I.; methodology, P.M., M.D. and A.I.; software, P.M., M.D. and A.I.; validation, P.M., M.D. and A.I.; formal analysis, P.M., M.D. and A.I.; investigation, P.M., M.D. and A.I.; resources, P.M., M.D. and A.I.; data curation, P.M. and A.I.; writing—original draft preparation, P.M., M.D. and A.I.; writing—review and editing, P.M., M.D. and A.I.; visualization, P.M.; supervision, M.D.; project administration, M.D.; funding acquisition, M.D. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by the Polish National Science Center (grant number 2021/43/B/NZ9/01111).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: All data are present in the article and Supplementary Materials.

Acknowledgments: We thank Mariusz Falkowski and Krzysztof Sieniawski (Enter Krzysztof Sieniawski, Olsztyn, Poland) for IT support, especially during launching of the repository of amino acids and modifications and also Monika Hryniewicz and Marta Godlewska for their help in inserting new data to the database.

Conflicts of Interest: The authors declare no conflict of interest.

Abbreviations

ACE	Angiotensin-converting enzyme (EC 3.4.15.1)
ACEi	Inhibitor of angiotensin-converting enzyme
AChE	Acetylcholinesterase (EC 3.1.1.7)
BChE	Butyrylcholinesterase (EC 3.1.1.8)
BRCA1	Breast cancer type 1 susceptibility protein (EC 2.3.2.27)
<C>	Cysteine involved in formation of disulfide bonds
CaMPDE	3',5'-cyclic-nucleotide phosphodiesterase (Calmodulin-dependent phosphodiesterase 1) (EC 3.1.4.17)
CID	Compound ID in the PubChem database
α_{s1} -CN	α_{s1} -Casein
α_{s2} -CN	α_{s2} -Casein
β -CN	β -Casein
κ -CN	κ -Casein
DLG4	Disks large homolog 4
DPPIII	Dipeptidyl peptidase III (EC 3.4.14.4)
DPPIIIi	Inhibitor of dipeptidyl peptidase III
DPPIV	Dipeptidyl peptidase IV (EC 3.4.14.5)
DPPIVi	Inhibitor of dipeptidyl peptidase IV
FDPS	Farnesyl pyrophosphate synthase (EC:2.5.1.1)
FNTA, FNTB	Farnesyltransferase/geranylgeranyltransferase type-1 (EC 2.5.1.58) subunit α and β , respectively

FPR2	Formyl peptide receptor 2
α -Glu	α -Glucosidase (EC 3.2.1.20)
α -Glui	Inhibitor of α -Glucosidase
HLA-A	HLA class I histocompatibility antigen, A α chain
HMG-CoA	3-Hydroxy-3-methylglutaryl coenzyme A
LPAR	Lysophosphatidic acid receptor
<P[4O]>	L-pyroglutamic acid
PIN1	Peptidyl-prolyl cis-trans isomerase NIMA-interacting 1 (EC 5.2.1.8)
PLK1	Serine/threonine-protein kinase PLK1 (EC 2.7.11.21)
PRKCE	PRKCE Protein kinase C ϵ type (EC 2.7.11.13)
S1PR	Sphingosine 1-phosphate receptor
<S[3*]>	Phosphoserine
<T[3*]>	Phosphothreonine
TK1	Thymidine kinase, cytosolic (EC 2.7.1.21)
TNFRSF10A	Tumor necrosis factor receptor superfamily member 10A
UWM	University of Warmia and Mazury in Olsztyn
XIAP	E3 ubiquitin-protein ligase (EC 2.3.2.27)

References

- Zhang, J.; Xin, L.; Shan, B.; Chen, W.; Xie, M.; Yuen, D.; Zhang, W.; Zhang, Z.; Lajoie, G.A.; Ma, B. PEAKS DB: De novo sequencing assisted database search for sensitive and accurate peptide identification. *Mol. Cell. Proteomics* **2012**, *11*, M111.010587. [[CrossRef](#)] [[PubMed](#)]
- Paul, P.; Muthu, M.; Chilukuri, Y.; Haga, S.M.; Chun, S.; Oh, J.-W. In silico tools and phosphoproteomic software exclusives. *Processes* **2019**, *7*, 869. [[CrossRef](#)]
- Seymour, R.W.; van der Post, S.; Mooradian, A.D.; Held, J.M. ProteoSushi: A software tool to biologically annotate and quantify modification-specific, peptide-centric proteomics data sets. *J. Proteome Res.* **2021**, *20*, 3621–3628. [[CrossRef](#)] [[PubMed](#)]
- Ergin, E.K.; Uzozie, A.C.; Chen, S.; Su, Y.; Lange, P.F. SQuAPP—Simple quantitative analysis of proteins and PTMs. *Bioinformatics* **2022**, *38*, 4956–4958. [[CrossRef](#)] [[PubMed](#)]
- Barente, A.S.; Villén, J.A. Python package for the localization of protein modifications in mass spectrometry data. *J. Proteome Res.* **2023**, *22*, 501–507. [[CrossRef](#)] [[PubMed](#)]
- Lin, S.; Wang, C.; Zhou, J.; Shi, Y.; Ruan, C.; Tu, Y.; Yao, L.; Peng, D.; Xue, Y. EPSD: A well-annotated data resource of protein phosphorylation sites in eukaryotes. *Brief Bioinform.* **2021**, *22*, 298–307. [[CrossRef](#)]
- Zhang, W.; Tan, X.; Lin, S.; Gou, Y.; Han, C.; Zhang, C.; Ning, W.; Wang, C.; Xue, Y. CPLM 4.0: An updated database with rich annotations for protein lysine modifications. *Nucleic Acids Res.* **2022**, *50*, D451–D459. [[CrossRef](#)] [[PubMed](#)]
- Das, A.; Giri, K.; Behera, R.N.; Maity, S.; Ambatipudi, K. BoMiProt 2.0: An update of the bovine milk protein database. *J. Proteom.* **2022**, *267*, 104696. [[CrossRef](#)]
- Tyagi, A.; Tuknait, A.; Anand, P.; Gupta, S.; Sharma, M.; Mathur, D.; Joshi, A.; Singh, S.; Gautam, A.; Raghava, G.P. CancerPPD: A database of anticancer peptides and proteins. *Nucleic Acids Res.* **2015**, *43*, D837–D843. [[CrossRef](#)]
- Mathur, D.; Prakash, S.; Anand, P.; Kaur, H.; Agrawal, P.; Mehta, A.; Kumar, R.; Singh, S.; Raghava, G.P.S. PEPlife: A repository of the half-life of peptides. *Sci. Rep.* **2016**, *6*, 36617. [[CrossRef](#)]
- Singh, S.; Chaudhary, K.; Dhanda, S.K.; Bhalla, S.; Usmani, S.S.; Gautam, A.; Tuknait, A.; Agrawal, P.; Mathur, D.; Raghava, G.P. SATPdb: A database of structurally annotated therapeutic peptides. *Nucleic Acids Res.* **2016**, *44*, D1119–D1126. [[CrossRef](#)] [[PubMed](#)]
- Mathur, D.; Mehta, A.; Firmal, P.; Bedi, G.; Sood, C.; Gautam, A.; Raghava, G.P. TopicalPdb: A database of topically delivered peptides. *PLoS ONE* **2018**, *13*, e0190134. [[CrossRef](#)] [[PubMed](#)]
- Usmani, S.S.; Agrawal, P.; Sehgal, M.; Patel, P.K.; Raghava, G.P.S. ImmunoSPdb: An archive of immunosuppressive peptides. *Database* **2019**, *2019*, baz012. [[CrossRef](#)] [[PubMed](#)]
- Flissi, A.; Ricart, E.; Campart, C.; Chevalier, M.; Dufresne, Y.; Michalik, J.; Jacques, P.; Flahaut, C.; Lisacek, F.; Leclère, V.; et al. Norine: Update of the nonribosomal peptide resource. *Nucleic Acids Res.* **2020**, *48*, D465–D469. [[CrossRef](#)] [[PubMed](#)]
- Mathur, D.; Kaur, H.; Dhalla, A.; Sharma, N.; Raghava, G.P.S. SAPdb: A database of short peptides and the corresponding nanostructures formed by self-assembly. *Comp. Biol. Med.* **2021**, *133*, 104391. [[CrossRef](#)] [[PubMed](#)]
- Balogh, B.; Ivánczi, M.; Nizami, B.; Beke-Somfai, T.; Mándity, I.M. ConjuPepDB: A database of peptide–drug conjugates. *Nucleic Acids Res.* **2021**, *49*, D1102–D1112. [[CrossRef](#)]
- D’Aloisio, V.; Dognini, P.; Hutcheon, G.A.; Coxon, C.R. PepTherDia: Database and structural composition analysis of approved peptide therapeutics and diagnostics. *Drug Discov. Today* **2021**, *26*, 1409–1419. [[CrossRef](#)]
- Li, J.; Yanagisawa, K.; Sugita, M.; Fujie, T.; Ohue, M.; Akiyama, Y. CycPeptMPDB: A comprehensive database of membrane permeability of cyclic peptides. *J. Chem. Inf. Model.* **2023**, *63*, 2240–2250. [[CrossRef](#)]
- Singh, S.; Singh, H.; Tuknait, A.; Chaudhary, K.; Singh, B.; Kumaran, S.; Raghava, G.P.S. PEPstrMOD: Structure prediction of peptides containing natural, non-natural and modified residues. *Biol. Direct.* **2015**, *10*, 73. [[CrossRef](#)]

20. Frolov, A.I.; Chankeshwara, S.V.; Abdulkarim, Z.; Ghiandoni, G.M. pIChemist—Free tool for the calculation of isoelectric points of modified peptides. *J. Chem. Inf. Model.* **2023**, *63*, 187–196. [CrossRef]
21. Gfeller, D.; Michielin, O.; Zoete, V. SwissSidechain: A molecular and structural database of non-natural sidechains. *Nucleic Acids Res.* **2013**, *41*, D327–D332. [CrossRef] [PubMed]
22. Zhang, T.; Li, H.; Xi, H.; Stanton, R.V.; Rotstein, S.H. HELM: A hierarchical notation language for complex biomolecule structure representation. *J. Chem. Inf. Model.* **2012**, *52*, 2796–2806. [CrossRef] [PubMed]
23. Milton, J.; Zhang, T.; Bellamy, C.; Swayze, E.; Hart, C.; Weisser, M.; Hecht, S.; Rotstein, S. HELM software for biopolymers. *J. Chem. Inf. Model.* **2017**, *57*, 1233–1239. [CrossRef]
24. Minkiewicz, P.; Darewicz, M.; Iwaniak, A.; Turło, M. Proposal of the annotation of phosphorylated amino acids and peptides using biological and chemical codes. *Molecules* **2021**, *26*, 712. [CrossRef] [PubMed]
25. Fox, T.; Bieler, M.; Haebel, P.; Ochoa, R.; Peters, S.; Weber, A. BILN: A human-readable line notation for complex peptides. *J. Chem. Inf. Model.* **2022**, *62*, 3942–3947. [CrossRef] [PubMed]
26. David, L.; Thakkar, A.; Mercado, R.; Engkvist, O. Molecular representations in AI-driven drug discovery: A review and practical guide. *J. Cheminform.* **2020**, *12*, 56. [CrossRef]
27. Barłowska, J.; Sawicka-Zugaj, W.; Ślaska, B.; Król, J.; Brodziak, A.; Teter, A.; Chabuz, W. Genetic Analysis of CSN2 in Local and International Cattle Breeds Raised in Poland. *Anim. Sci. Papers Rep.* **2022**, *40*, 411–422. Available online: <https://www.igbzpan.pl/uploaded/FSiBundleContentBlockBundleModelTranslatableBlockTranslatableFilesElement/filePath/2249/str411-422.pdf> (accessed on 4 May 2023).
28. Cieślińska, A.; Fiedorowicz, E.; Rozmus, D.; Sienkiewicz-Szlapka, E.; Jarmołowska, B.; Kamiński, S. Does a little difference make a big difference? Bovine β -Casein A1 and A2 variants and human health—An update. *Int. J. Mol. Sci.* **2022**, *23*, 1563. [CrossRef]
29. Daniloski, D.; McCarthy, N.A.; Huppertz, T.; Vasiljevic, T. 2022, What is the impact of amino acid mutations in the primary structure of caseins on the composition and functionality of milk and dairy products? *Curr. Res. Food Sci.* **2022**, *5*, 1701–1712. [CrossRef]
30. Nadugala, B.H.; Pagel, C.N.; Raynes, J.K.; Ranadheera, C.S.; Logan, A. The effect of casein genetic variants, glycosylation and phosphorylation on bovine milk protein structure, technological properties, nutrition and product manufacture. *Int. Dairy J.* **2022**, *133*, 105440. [CrossRef]
31. Bielecka, M.; Cichosz, G.; Czeczot, H. Antioxidant, antimicrobial and anticarcinogenic activities of bovine milk proteins and their hydrolysates—A review. *Int. Dairy J.* **2022**, *127*, 105208. [CrossRef]
32. Bu, T.; Zheng, J.; Liu, L.; Li, S.; Wu, J. Milk proteins and their derived peptides on bone health: Biological functions, mechanisms, and prospects. *Compr. Rev. Food Sci. Food Saf.* **2021**, *20*, 2234–2262. [CrossRef] [PubMed]
33. Mirzapour-Kouhdasht, A.; Garcia-Vaquero, M. Cardioprotective peptides from milk processing and dairy products: From bioactivity to final products including commercialization and legislation. *Foods* **2022**, *11*, 1270. [CrossRef] [PubMed]
34. Stobiecka, M.; Król, J.; Brodziak, A. Antioxidant activity of milk and dairy products. *Animals* **2022**, *12*, 245. [CrossRef]
35. Tenenbaum, M.; Deracinois, B.; Dugardin, C.; Matéos, A.; Romelard, A.; Auger, J. Identification, production and bioactivity of casein phosphopeptides—A review. *Food Res. Int.* **2022**, *157*, 111360. [CrossRef]
36. Dhar, H.; Verma, S.; Dogra, S.; Katoch, S.; Vij, R.; Singh, G.; Sharma, M. Functional attributes of bioactive peptides of bovine milk origin and application of in silico approaches for peptide prediction and functional annotations. *Crit. Rev. Food Sci. Nutr.* **2023**. [CrossRef]
37. Iwaniak, A.; Minkiewicz, P.; Darewicz, M.; Protasiewicz, M.; Mogut, D. Chemometrics and cheminformatics in the analysis of biologically active peptides from food sources. *J. Funct. Foods* **2015**, *16*, 334–351. [CrossRef]
38. Iwaniak, A.; Darewicz, M.; Mogut, D.; Minkiewicz, P. Elucidation of the role of in silico methodologies in approaches to studying bioactive peptides derived from foods. *J. Funct. Foods* **2019**, *61*, 103486. [CrossRef]
39. Du, Z.; Comer, J.; Li, Y. Bioinformatics approaches to discovering food-derived bioactive peptides: Reviews and perspectives. *Trends Anal. Chem.* **2023**, *162*, 117051. [CrossRef]
40. Barati, M.; Jabbari, M.; Davoodi, S.H. In-silico methods for milk-derived bioactive peptide prediction. In *Enzymes Beyond Traditional Applications in Dairy Science and Technology*; Rajput, Y.S., Sharma, R., Eds.; Elsevier Inc.: Amsterdam, The Netherlands, 2023; pp. 137–162, ISBN 978-0-323-96010-6. [CrossRef]
41. UniProt Website. Available online: <https://www.uniprot.org/> (accessed on 31 January 2023).
42. The UniProt Consortium UniProt: The Universal Protein Knowledgebase in 2023. *Nucleic Acids Res.* **2023**, *51*, D523–D531. [CrossRef]
43. Rasmussen, L.K.; Højrup, P.; Petersen, T.E. The multimeric structure and disulfide-bonding pattern of bovine κ -casein. *Eur. J. Biochem.* **1992**, *207*, 215–222. [CrossRef] [PubMed]
44. Rasmussen, L.K.; Højrup, P.; Petersen, T.E. Localization of two interchain disulfide bridges of bovine α_{s2} -casein. Parallel and antiparallel alignments of the polypeptide chains. *Eur. J. Biochem.* **1992**, *203*, 381–386. [CrossRef] [PubMed]
45. BIOPEP-UWM Website. Available online: <https://biochemia.uwm.edu.pl/en/biopep-uwm-2/> (accessed on 30 March 2023).
46. Minkiewicz, P.; Iwaniak, A.; Darewicz, M. BIOPEP-UWM database of bioactive peptides: Current opportunities. *Int. J. Mol. Sci.* **2019**, *20*, 5978. [CrossRef] [PubMed]
47. Minkiewicz, P.; Iwaniak, A.; Darewicz, M. BIOPEP-UWM Virtual—A novel database of food-derived peptides with in silico-predicted biological activity. *Appl. Sci.* **2022**, *12*, 7204. [CrossRef]

48. PubChem Website. Available online: <https://pubchem.ncbi.nlm.nih.gov/> (accessed on 31 January 2023).
49. Kim, S.; Chen, J.; Cheng, T.; Gindulyte, A.; He, J.; He, S.; Li, Q.; Shoemaker, B.A.; Thiessen, P.A.; Yu, B.; et al. PubChem 2023 update. *Nucleic Acids Res.* **2023**, *51*, D1373–D1380. [[CrossRef](#)] [[PubMed](#)]
50. Caroli, A.M.; Chessa, S.; Erhardt, G.J. Invited review: Milk protein polymorphisms in cattle: Effect on animal breeding and human nutrition. *J. Dairy Sci.* **2009**, *92*, 5335–5352. [[CrossRef](#)]
51. Brodtkorb, A.; Egger, L.; Alminger, M.; Alvito, P.; Assunção, R.; Ballance, S.; Bohn, T.; Bourlieu-Lacanal, C.; Boutrou, R.; Carrière, F.; et al. INFOGEST static in vitro simulation of gastrointestinal food digestion. *Nature Protoc.* **2019**, *14*, 991–1014. [[CrossRef](#)]
52. Dziuba, J.; Iwaniak, A.; Minkiewicz, P. Computer-Aided Characteristics of Proteins as Potential Precursors of Bioactive Peptides. *Polimery* **2003**, *48*, 50–53. Available online: <https://polimery.ichp.vot.pl/index.php/p/article/view/1836> (accessed on 18 June 2023). [[CrossRef](#)]
53. Minkiewicz, P.; Dziuba, J.; Michalska, J. Bovine meat proteins as potential precursors of biologically active peptides—A computational study based on the BIOPEP database. *Food Sci. Technol. Int.* **2011**, *17*, 39–45. [[CrossRef](#)]
54. HeatMapper Website. Available online: <http://www.heatmapper.ca/> (accessed on 28 April 2023).
55. Babicki, S.; Arndt, D.; Marcu, A.; Liang, Y.; Grant, J.R.; Maciejewski, A.; Wishart, D.S. Heatmapper: Web-enabled heat mapping for all. *Nucleic Acids Res.* **2016**, *44*, W147–W153. [[CrossRef](#)]
56. SwissTargetPrediction Website. Available online: <http://www.swisstargetprediction.ch/> (accessed on 31 March 2023).
57. Daina, A.; Michielin, O.; Zoete, V. SwissTargetPrediction: Updated data and new features for efficient prediction of protein targets of small molecules. *Nucleic Acids Res.* **2019**, *47*, W357–W364. [[CrossRef](#)] [[PubMed](#)]
58. Darewicz, M.; Pliszka, M.; Borawska-Dziadkiewicz, J.; Minkiewicz, P.; Iwaniak, A. Multi-bioactivity of protein digests and peptides from oat (*Avena sativa* L.) kernels in the prevention of the cardiometabolic syndrome. *Molecules* **2022**, *27*, 7907. [[CrossRef](#)]
59. Iwaniak, A.; Darewicz, M.; Minkiewicz, P. Peptides derived from foods as supportive diet components in the prevention of the metabolic syndrome. *Compr. Rev. Food Sci. Food Saf.* **2018**, *17*, 63–81. [[CrossRef](#)] [[PubMed](#)]
60. ChEMBL Website. Available online: <https://www.ebi.ac.uk/chembl/db/> (accessed on 30 April 2023).
61. Mendez, D.; Gaulton, A.; Bento, A.P.; Chambers, J.; De Veij, M.; Félix, E.; Magariños, M.P.; Mosquera, J.F.; Mutowo, P.; Nowotka, M.; et al. ChEMBL: Towards direct deposition of bioassay data. *Nucleic Acids Res.* **2019**, *47*, D930–D940. [[CrossRef](#)] [[PubMed](#)]
62. Minkiewicz, P.; Darewicz, M.; Iwaniak, A.; Sokołowska, J.; Starowicz, P.; Bucholska, J.; Hryniewicz, M. Common amino acid subsequences in a universal proteome—relevance for food science. *Int. J. Mol. Sci.* **2015**, *16*, 20748–20773. [[CrossRef](#)]
63. Cárdenas, S.; Valcárcel, M. Analytical features in qualitative analysis. *Trends Anal. Chem.* **2005**, *24*, 477–487. [[CrossRef](#)]
64. Vitale, G.A.; Scarpato, S.; Mangoni, A.; D’Auria, M.V.; Della Sala, G.; de Pascale, D. Enhanced molecular networking shows *Microbacterium* sp. V1 as a factory of antioxidant proline-rich peptides. *Mar. Drugs* **2023**, *21*, 256. [[CrossRef](#)]
65. Altschul, S.F.; Madden, T.L.; Schaffer, A.A.; Zhang, J.; Zhang, Z.; Miller, W.; Lipman, D.J. Gapped BLAST and PSI-BLAST: A new generation of protein database search programs. *Nucleic Acids Res.* **1997**, *25*, 3389–3402. [[CrossRef](#)]
66. Sievers, F.; Higgins, D.G. Clustal omega, accurate alignment of very large numbers of sequences. *Methods Mol. Biol.* **2014**, *1079*, 105–116. [[CrossRef](#)]
67. Potter, S.C.; Luciani, A.; Eddy, S.R.; Park, Y.; Lopez, R.; Finn, R.D. HMMER web server: 2018 update. *Nucleic Acids Res.* **2018**, *46*, W200–W204. [[CrossRef](#)]
68. Jelachich, M.L.; Cowan, E.P.; Turner, R.V.; Coligan, J.E.; Biddison, W.E. Analysis of the molecular basis of HLA-A3 recognition by cytotoxic T cells using defined mutants of the HLA-A3 molecule. *J. Immunol.* **1988**, *141*, 1108–1113. [[CrossRef](#)] [[PubMed](#)]
69. Madden, D.R.; Garbocci, D.N.; Wiley, D.C. The antigenic identity of peptide-MHC complexes: A comparison of the conformations of five viral peptides presented by HLA-A2. *Cell* **1993**, *75*, 693–708. [[CrossRef](#)]
70. Song, I.; Gil, A.; Mishra, R.; Ghersi, D.; Selin, L.K.; Stern, L.J. Broad TCR repertoire and diverse structural solutions for recognition of an immunodominant CD8+ T cell epitope. *Nature Struct. Mol. Biol.* **2017**, *24*, 395–406. [[CrossRef](#)] [[PubMed](#)]
71. Wiedeman, P.E. DPP4V inhibition: Promising therapy for the treatment of type 2 diabetes. *Progr. Med. Chem.* **2007**, *45*, 63–109. [[CrossRef](#)]
72. Chalichem, N.S.S.; Kiran, P.S.S.; Basavan, D. Possible role of DPP4 inhibitors to promote hippocampal neurogenesis in Alzheimer’s disease. *J. Drug Target.* **2018**, *26*, 670–675. [[CrossRef](#)]
73. Kawakita, E.; Koya, D.; Kanasaki, K. CD26/DPP-4: Type 2 diabetes drug target with potential influence on cancer biology. *Cancers* **2021**, *13*, 2191. [[CrossRef](#)]
74. Schimmer, A.D.; Dalili, S.; Batey, R.A.; Riedl, S.J. Targeting XIAP for the treatment of malignancy. *Cell Death Differ.* **2006**, *13*, 179–188. [[CrossRef](#)]
75. Galbán, S.; Duckett, C.S. XIAP as a ubiquitin ligase in cellular signaling. *Cell Death Differ.* **2010**, *17*, 54–60. [[CrossRef](#)]
76. Abbas, R.; Larisch, S. Killing by degradation: Regulation of apoptosis by the ubiquitin-proteasome-system. *Cells* **2021**, *10*, 3465. [[CrossRef](#)]
77. Beck, D.B.; Werner, A.; Kastner, D.L.; Aksentjevich, I. Disorders of ubiquitylation: Unchained inflammation. *Nature Rev. Rheumatol.* **2022**, *18*, 435–447. [[CrossRef](#)]
78. Yanagida, K.; Shimizu, T. Lysophosphatidic acid, a simple phospholipid with myriad functions. *Pharmacol. Therapeut.* **2023**, *246*, 108421. [[CrossRef](#)] [[PubMed](#)]
79. Hunter, T. Why nature chose phosphate to modify proteins. *Philos. Trans. R. Soc. B Biol. Sci.* **2012**, *367*, 2513–2516. [[CrossRef](#)] [[PubMed](#)]

80. Kamerlin, S.C.L.; Sharma, P.K.; Prasad, R.B.; Warshel, A. Why nature really chose phosphate. *Q. Rev. Biophys.* **2013**, *46*, 1–132. [[CrossRef](#)] [[PubMed](#)]
81. Czegle, I.; Huang, C.; Soria, P.G.; Purkiss, D.W.; Shields, A.; Wappler-Guzzetta, E.A. The role of genetic mutations in mitochondrial-driven cancer growth in selected tumors: Breast and gynecological malignancies. *Life* **2023**, *13*, 996. [[CrossRef](#)]
82. Zhu, J.W.; Charkhchi, P.; Adekunle, S.; Akbari, M.R. What is known about breast cancer in young women? *Cancers* **2023**, *15*, 1917. [[CrossRef](#)]
83. Malter, J.S. Pin1 and Alzheimer's disease. *Transl. Res.* **2023**, *254*, 24–33. [[CrossRef](#)]
84. Caligiuri, I.; Vincenzo, C.; Asano, T.; Kumar, V.; Rizzolio, F. The metabolic crosstalk between PIN1 and the tumour microenvironment. *Sem. Cancer Biol.* **2023**, *91*, 143–157. [[CrossRef](#)]
85. Prevete, N.; Liotti, F.; Marone, G.; Melillo, R.M.; De Paulis, A. Formyl peptide receptors at the interface of inflammation, angiogenesis and tumor growth. *Pharmacol. Res.* **2015**, *102*, 184–191. [[CrossRef](#)]
86. Tourkochristou, E.; Mouzaki, A.; Triantos, C. Unveiling the biological role of sphingosine-1-phosphate receptor modulators in inflammatory bowel diseases. *World J. Gastroenterol.* **2023**, *29*, 110–125. [[CrossRef](#)]
87. Dumitrescu, L.; Papathanasiou, A.; Coclitu, C.; Garjani, A.; Evangelou, N.; Constantinescu, C.S.; Popescu, B.O.; Tanasescu, R. An update on the use of sphingosine 1-phosphate receptor modulators for the treatment of relapsing multiple sclerosis. *Expert Opin. Pharmacother.* **2023**, *24*, 495–509. [[CrossRef](#)]
88. Pan, G.; O'Rourke, K.; Chinnaiyan, A.M.; Gentz, R.; Ebner, R.; Ni, J.; Dixit, V.M. The receptor for the cytotoxic ligand TRAIL. *Science* **1997**, *276*, 111–113. [[CrossRef](#)] [[PubMed](#)]
89. Desler, C.; Lykke, A.; Rasmussen, L.J. The effect of mitochondrial dysfunction on cytosolic nucleotide metabolism. *J. Nucl. Acids* **2010**, *2010*, 701518. [[CrossRef](#)] [[PubMed](#)]
90. Brigidi, G.S.; Santyr, B.; Shimell, J.; Jovellar, B.; Bamji, S.X. Activity-regulated trafficking of the palmitoyl-acyl transferase DHHC5. *Nature Commun.* **2015**, *6*, 8200. [[CrossRef](#)] [[PubMed](#)]
91. Haney, S.L.; Holstein, S.A. Targeting the isoprenoid biosynthetic pathway in multiple myeloma. *Int. J. Mol. Sci.* **2023**, *24*, 111. [[CrossRef](#)]
92. Fatoki, T.H.; Aluko, R.E.; Udenigwe, C.C. In silico investigation of molecular targets, pharmacokinetics, and biological activities of chicken egg ovalbumin protein hydrolysates. *J. Food Bioact.* **2022**, *17*, 34–48. [[CrossRef](#)]
93. Iwaniak, A.; Minkiewicz, P.; Pliszka, M.; Mogut, D.; Darewicz, M. Characteristics of biopeptides released in silico from collagens using quantitative parameters. *Foods* **2020**, *9*, 965. [[CrossRef](#)]
94. Fatoki, T.H.; Chukwuejim, S.; Udenigwe, C.C.; Aluko, R.E. In silico exploration of metabolically active peptides as potential therapeutic agents against amyotrophic lateral sclerosis. *Int. J. Mol. Sci.* **2023**, *24*, 5828. [[CrossRef](#)]
95. Iwaniak, A.; Minkiewicz, P.; Darewicz, M.; Hryniewicz, M. Food protein-originating peptides as tastants—Physiological, technological, sensory, and bioinformatic approaches. *Food Res. Int.* **2016**, *89*, 27–38. [[CrossRef](#)]
96. Weimann, C.; Meisel, H.; Erhardt, G. Bovine kappa-casein variants result in different angiotensin I converting enzyme (ACE) inhibitory peptides. *J. Dairy Sci.* **2009**, *92*, 1885–1888. [[CrossRef](#)]
97. Pauciullo, A.; Giambra, I.J.; Iannuzzi, L.; Erhardt, G. The β -casein in camels: Molecular characterization of the CSN2 gene, promoter analysis and genetic variability. *Gene* **2014**, *547*, 159–168. [[CrossRef](#)]
98. Erhardt, G.; Shuiep, E.T.S.; Lisson, M.; Weimann, C.; Wang, Z.; El Zubeir, I.E.Y.M.; Pauciullo, A. Alpha S1-casein polymorphisms in camel (*Camelus dromedarius*) and descriptions of biological active peptides and allergenic epitopes. *Tropic. Anim. Health Prod.* **2016**, *48*, 879–887. [[CrossRef](#)] [[PubMed](#)]
99. Iwaniak, A.; Mogut, D.; Minkiewicz, P.; Żulewska, J.; Darewicz, M. Gouda cheese with modified content of β -casein as a source of peptides with ACE- and DPP-IV-inhibiting bioactivity: A study based on in silico and in vitro protocol. *Int. J. Mol. Sci.* **2021**, *22*, 2949. [[CrossRef](#)] [[PubMed](#)]
100. Iwaniak, A.; Mogut, D.; Minkiewicz, P.; Żulewska, J.; Darewicz, M. An integrated approach to the analysis of antioxidative peptides derived from Gouda cheese with a modified β -casein content. *Sci. Rep.* **2022**, *12*, 13314. [[CrossRef](#)] [[PubMed](#)]
101. Darewicz, M.; Borawska, J.; Pliszka, M. Carp proteins as a source of bioactive peptides—An in silico approach. *Czech J. Food Sci.* **2016**, *34*, 111–117. [[CrossRef](#)]
102. Borawska-Dziadkiewicz, J.; Darewicz, M.; Tarczyńska, A.S. Properties of peptides released from salmon and carp via simulated human-like gastrointestinal digestion described applying quantitative parameters. *PLoS ONE* **2021**, *16*, e0255969. [[CrossRef](#)]
103. Paysan-Lafosse, T.; Blum, M.; Chuguransky, S.; Grego, T.; Pinto, B.L.; Salazar, G.A.; Bileschi, M.L.; Bork, P.; Bridge, A.; Colwell, L.; et al. InterPro in 2022. *Nucleic Acids Res.* **2023**, *51*, D418–D427. [[CrossRef](#)]
104. Minkiewicz, P.; Mattison, C.P.; Darewicz, M. Quantitative in silico evaluation of allergenic proteins from *Anacardium occidentale*, *Carya illinoensis*, *Juglans regia* and *Pistacia vera* and their epitopes as precursors of bioactive peptides. *Curr. Issues Mol. Biol.* **2022**, *44*, 3100–3117. [[CrossRef](#)]
105. Bougatef, H.; de la Vega-Fernández, C.; Sila, A.; Bougatef, A.; Martínez-Alvarez, O. Identification of ACE I-inhibitory peptides released by the hydrolysis of tub gurnard (*Chelidonichthys lucerna*) skin proteins and the impact of their in silico gastrointestinal digestion. *Mar. Drugs* **2023**, *21*, 131. [[CrossRef](#)]
106. Vidal-Limon, A.; Aguilar-Toalá, J.E.; Liceaga, A.M. Integration of molecular docking analysis and molecular dynamics simulations for studying food proteins and bioactive peptides. *J. Agric. Food Chem.* **2022**, *70*, 934–943. [[CrossRef](#)]

107. Zhang, J.; Li, F.; Shen, S.; Yang, Z.; Ji, X.; Wang, X.; Liao, X.; Zhang, Y. More simple, efficient and accurate food research promoted by intermolecular interaction approaches: A review. *Food Chem.* **2023**, *416*, 135726. [[CrossRef](#)]
108. Siani, M.A.; Weininger, D.; Blaney, J.M. CHUCKLES: A method for representing and searching peptide and peptoid sequences on both monomer and atomic levels. *J. Chem. Inf. Comput. Sci.* **1994**, *34*, 588–593. [[CrossRef](#)] [[PubMed](#)]
109. Minkiewicz, P.; Iwaniak, A.; Darewicz, M. Annotation of peptide structures using SMILES and other chemical codes—practical solutions. *Molecules* **2017**, *22*, 2075. [[CrossRef](#)] [[PubMed](#)]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.