Research Article

In Silico Characterization of Bioactive Peptides Derived from Sheep Milk Caseins for Sustainable Animal Health

Anila Hoda[1]*, Sulltane Ajçe[2], Ilia Mikerezi[1], Xhelil Koleci[3]

[1]Academy of Sciences of Albania, Tirane, Albania; [2]University Fan S. Noli, 7001 Korçë, Albania; [3]Agricltural University of Tirana, Tiranë, Albania.

Abstract | Sheep milk caseins are promising sources of bioactive peptides with health-enhancing properties. In this study, αS1-, αS2-, β-, and κ-caseins were analyzed using *in silico* tools assess their physicochemical properties including molecular weight, isoelectric point (pI), instability index, aliphatic index, and GRAVY (Grand Average of Hydropathicity) score as well as their secondary structure composition, post-translational modifications (phosphorylation and glycosylation sites), and conserved sequence motifs. β-casein showed the highest instability index (100.37) and aliphatic index (89.76), suggesting high flexibility and thermostability. αS2-casein was the most hydrophilic, with the lowest GRAVY score (–1.060) and a balanced charge profile. Secondary structure analysis revealed that β- and κ-caseins are largely disordered, while αS1- and αS2-caseins showed higher alpha-helix content. Simulated gastrointestinal digestion using pepsin, trypsin, and chymotrypsin generated peptide fragments, which were evaluated for bioactivity potential, fragment length, and digestion susceptibility. Antimicrobial activity was predicted using four machine learning classifiers: Random Forest (RF), Support Vector Machine (SVM), Artificial Neural Network (ANN), and Discriminant Analysis (DA). Peptides were further assessed for toxicity, solubility, charge, hydrophobicity, and pI. Machine learning models predicted strong antimicrobial potential for peptides such as PL, AW, PW, and IPIQY. These peptides also exhibited multifunctional activities, including ACE (Angiotensin Converting Enzyme) inhibition and antioxidant effects, and were non-toxic with favorable physicochemical profiles. The findings support the use of sheep milk-derived peptides as natural, multifunctional compounds, contributing to sustainable livestock health management under the One Health framework.

Keywords | Sheep milk, Casein-derived peptides, Antimicrobial peptides (AMPs), Bioactive peptides, Gastrointestinal digestion, One Health, Sustainable livestock

Received | July 17, 2025; **Accepted** | October 01, 2025; **Published** | November 17, 2025
***Correspondence** | Anila Hoda, Academy of Sciences of Albania, Tirane, Albania; **Email:** anila.hoda@akad.gov.al
Citation | Hoda A, Ajçe S, Mikerezi I, Koleci X (2025). *In silico* characterization of bioactive peptides derived from sheep milk caseins for sustainable animal health. J. Anim. Health Prod. 13(4): 1216-1227.
DOI | https://dx.doi.org/10.17582/journal.jahp/2025/13.4.1216.1227
ISSN (Online) | 2308-2801

INTRODUCTION

Milk is not only a source of nutrition but also a functional food. It contains many helpful compounds, especially bioactive peptides (BAPs). These peptides are short chains of 2 to 20 amino acids. They are major compound of milk proteins and are released during digestion, fermentation, or food processing (El-Salam and El-Shibiny, 2013; Hati *et al*., 2024). Once released, BAPs can have many health effects. They may fight microbes, reduce oxidation, lower blood pressure, support the immune system, and help control blood sugar (Dhar *et al*., 2024; Quintieri *et al*., 2024).

Recent studies have focused more on BAPs from non-bovine milk, such as sheep, goat, camel, and buffalo that are widely used in Mediterranean and Asian countries. They also contain unique peptides that may offer better or different

health benefits than cow milk (de la Fuente *et al*., 2013; Hati *et al*., 2024). Sheep milk has gained special interest. It is rich in protein and fat and is a good source of bioactive peptides with useful health effects (Sathika *et al*., 2025).

In silico tools have made it easier to study bioactive peptides (BAPs). These computational tools help simulate protein breakdown and predict the activity of peptides. They are used to infer the allergenicity, toxicity or otherwise of the peptide. In addition, they show the chemical properties of the peptides. This saves time and money by reducing lab work and helps find the best peptides for food or medicine use (Dhar *et al*., 2024; Du *et al*., 2023).

There is growing interest in milk from non-bovine species, as studies increasingly highlight the health benefits of milk-derived peptides. In this context, sheep milk has emerged as a particularly rich source of bioactive peptides with promising functional properties. The main goal of this study was to explore the potential of sheep milk caseins as a source of bioactive peptides. Their digestion was stimulated to find short peptides with possible health benefits. The study aimed to identify and also characterize the peptides for their predicted antimicrobial, antioxidant, and antihypertensive effects using in silico tools. This characterization included machine learning-based activity prediction, physicochemical profiling, and toxicity assessment, providing a comprehensive evaluation of their multifunctional potential. These findings support the use of natural peptides in One Health approach, sustainable farming and the circular bioeconomy approach.

MATERIALS AND METHODS

This study was conducted by the use of publicly available protein sequence data for sheep (*Ovis aries*) caseins retrieved from the UniProt database. All analyses were performed using online bioinformatics platforms and locally installed computational tools.

Study design

The study employed a computational in silico design involving:

Data acquisition: Retrieval of αS1-, αS2-, β-, and κ-casein sequences.

- Physicochemical analysis: Calculation of molecular weight, isoelectric point, instability index, aliphatic index, and GRAVY score.
- Structural analysis; Prediction of secondary structure composition using SOPMA and identification of conserved motifs.
- Post-translational modification prediction: Identification of phosphorylation (NetPhos 3.1) and N-linked glycosylation sites (NetNGlyc 1.0).
- Phylogenetic analysis: Construction of trees to assess evolutionary relationships.
- Simulated gastrointestinal digestion: Prediction of bioactive peptides released by pepsin, trypsin, and chymotrypsin hydrolysis (BIOPEP-UWM).
- Bioactivity prediction: Evaluation of antimicrobial, antioxidant, ACE-inhibitory, and DPP-IV-inhibitory potential using multiple machine learning models.
- Physicochemical and toxicity assessment: Characterisation of identified peptides for hydrophobicity, charge, isoelectric point, molecular weight, and predicted toxicity.

Sequence characterization and chemical analysis

The amino acid sequences of the main proteins found in sheep milk in FASTA format were retrieved from the UniProt Knowledgebase (UniProtKB) to be used in bioinformatic analysis. The specific proteins studied were α-s1 casein (ID: P04653), α-s2 casein (P04654), β-casein (P11839), and κ-casein (P02669).

The ExPASy ProtParam tool (https://web.expasy.org/protparam/) (Gasteiger *et al*., 2005) was used to determine detailed physicochemical parameters of the proteins, including molecular weight, theoretical isoelectric point (pI), instability index, aliphatic index, extinction coefficient, amino acid composition**,** and GRAVY score (Grand Average of Hydropathicity)**.** These parameters provide insights into protein solubility, thermostability, and hydrophobicity. To obtain the final molecular weight of the mature protein, the signal peptide mass was subtracted from the total protein mass using the ExPASy PeptideMass tool (https://web.expasy.org/peptide_mass/).

In silico prediction of structural features and post-translational modifications of casein proteins

The structural and functional features of sheep αs1-casein, αs2-casein, β-casein, and κ-casein proteins (from the species *Ovis aries*) were studied using different bioinformatic tools. One of these tools, called SOPMA (Self-Optimized Prediction Method with Alignment) (Geourjon and Deléage, 1995), was used to predict the protein's secondary structure.

To study post-translational modifications in sheep casein proteins, the NetPhos 3.1 Server (Blom *et al*., 1999) was used to predict potential phosphorylation sites specifically on serine, threonine, and tyrosine residues while the NetNGlyc 1.0 Server (Blom *et al*., 1999) was applied to identify possible N-linked glycosylation sites based on conserved sequence motifs. Both tools rely on established biochemical patterns and machine learning models to assess the likelihood of these modifications, which can influence protein structure, stability, and function.

In addition, the MEME (Multiple EM for Motif Elicitation) suite (Bailey *et al*., 2009) was used to find conserved motifs short, important patterns in the β-casein and κ-casein proteins. These motifs may help the proteins stay stable, perform their biological functions, or interact with other molecules.

Phylogenetic analysis and multiple sequence alignment

The amino acid sequences of casein proteins were collected and used to search for similar proteins using the BLAST tool (http://blast.ncbi.nlm.nih.gov/Blast.cgi), which compares sequences to a large protein database. The matching protein sequences were then aligned using a program called Clustal Omega (version 1.2.4), and the alignment was improved with CLUSTAL-W, which is part of the MEGA 11 software (Tamura *et al*., 2021). Phylogenetic analysis was carried out using the Neighbor-Joining (NJ) method based on distance matrices, applying the default P-distance model for evolutionary distance calculations. A consensus tree was generated and evaluated using ClustalX, while the final rooted tree was visualized using the NJPlot application. To assess the robustness of the phylogenetic groupings, bootstrapping was performed with 1,000 replicates.

In silico digestion and peptide functional prediction

The protein sequences of major sheep milk proteins, with signal peptides excluded, were individually subjected to *in silico* enzymatic digestion. The theoretical generation of peptide fragments were carried out using the BIOPEP-UWM database and its enzymatic simulation tool (Minkiewicz *et al*., 2008). To mimic gastrointestinal digestion, a combination of three proteolytic enzymes were applied: chymotrypsin A (EC 3.4.21.1), trypsin (EC 3.4.21.4), and pepsin at pH 1.3 (EC 3.4.23.1). The resulting theoretical peptides were subsequently screened using the "Search for active fragments" function within BIOPEP to identify sequences with documented bioactivities. Digestion was performed separately for each milk protein, and three parameters were calculated: (1) the total number of peptides generated, (2) the number of peptides with known biological activities, (3) the distribution of peptides by length (di- to heptapeptides and longer).

In silico evaluation of antimicrobial and physicochemical properties of casein-derived peptides

The CAMPR3 database (http://www.camp.bicnirrh.res.in) was utilized to assess the antibacterial potential of peptides generated via *in silico* digestion. For this reason a combination of multivariate machine learning models, including Random Forest (RF), Support Vector Machines (SVM), Artificial Neural Networks (ANN), and Discriminant Analysis (DA) (Waghu *et al*., 2014) was applied. These models predict antimicrobial activity by analyzing peptide sequences based on structural and physicochemical features. Peptides with prediction scores above 0.49 in the RF, SVM, and DA models were classified as antimicrobial peptides (AMPs).

The molecular weight, hydrophobicity, and isoelectric point (pI) of sheep milk-derived peptides were analyzed using the ToxinPred tool (https://webs.iiitd.edu.in/raghava/toxinpred/index.html). Additionally, the net charge and water solubility of the peptides was estimated with PepCalc (https://pepcalc.com/), providing further insight into their physicochemical properties.

RESULTS

Sequence analysis

The key physicochemical parameters of four major casein proteins from *Ovis aries* (sheep) αS1-casein, αS2-casein, β-casein, and κ-casein are shown in Table 1. The proteins vary in sequence length from 211 amino acids for κ-casein to 223 for αS2-casein. αS2-casein has the highest pre-protein mass (24,661.78 Da) and extinction coefficient (34,505). β-casein has the highest instability index (100.37). The isoelectric points (pI) range from 5.13 for β-casein to 7.14 for αS2-casein. The GRAVY (Grand Average of Hydropathicity) index values range from –1.060 for αS2-casein to –0.330 for β-casein. The aliphatic index is highest for β-casein (89.76). Signal peptide masses differ among the proteins, and all mature proteins retain values close to their full pre-protein mass.

Table 1: Physicochemical properties of sheep casein proteins based on *in silico* analysis.

Accession no.	Protein name	No. of AA	Pre-protein mass (Da)	Signal peptide mass (Da)	Mature protein mass (Da)	Theoretical pI	Extinction coefficient	Instability index	Aliphatic index	negatively charged residues (Asp + Glu	positively charged residues (Arg + Lys)	GRAVY index
P04653	Alpha-S1-casein	214	22,751.65	14.24	22,737.41	5.19	27390	60.17	75.98	27	20	-0.674
P04654	Alpha-S2-casein	223	24661.78	15.39	24646.39	7.14	34505	56,42	62.31	31	31	-1.060
P11839	Beta casein	222	23351.09	14.83	23336.26	5.13	9970	100.37	89.76	23	15	-0.330
P02669	Kappa casein	211	19143.46	11.96	19131.50	5.53	18910	50.94	70.23	18	14	-0.589

Computational prediction of structural and post-translational modification of sheep casein proteins

The secondary structure analysis of sheep casein proteins αS1-, αS2-, β-, and κ-casein predicted by the SOPMA server (Table 2) showed that αS1- and αS2-caseins had alpha-helix contents of 47.20% and 55.16%, respectively. β-casein and κ-casein had alpha-helix contents of 22.07% and 13.02%, respectively, and random coil regions accounted for 73.42% and 82.81%, respectively. The content of extended strands ranged from 3.74% to 5.38% across all proteins. No beta bridges were detected in any of the proteins.

Table 2: Secondary structure composition (%) of casein fractions predicted by SOPMA.

Secondary structure	αS1-casein	αS2-casein	β-casein	κ-casein
Alpha helix	47.20	55.16	22.07	13.02
Extended strand	3.74	5.38	4.50	4.17
Random coil	49.07	39.46	73.42	82.81

Phosphorylation site predictions (Supplementary Figure S1) using NetPhos 3.1 showed potential post-translational modifications in all caseins except αS1-casein, which had no high-confidence phosphorylation sites. In αS2-casein, high-confidence predictions were observed at S32, T138, and T198. β-casein contained clusters of predicted phosphorylatable residues within S30–S34 and S137–T143. κ-casein exhibited predicted phosphorylation sites in its C-terminal region, including S108, S172, and T188. Glycosylation analysis using NetNGlyc 1.0 (Supplementary Figure S2) revealed one high-confidence N-linked glycosylation site in αS2-casein at position 30, with the consensus sequence NISQ. No N-glycosylation sites were predicted in αS1-, β-, or κ-caseins. O-linked glycosylation was not evaluated in this analysis.

The MEME Suite (v5.5.7) was used to identify conserved sequence motifs in the four main sheep milk casein proteins. Using the "any number of repetitions" (anr) model, three statistically significant motifs were discovered in each protein sequence. The analysis revealed meaningful patterns associated with the structural and functional roles of these proteins (Supplementary Figure S3).

In αS1-casein, the identified motifs were LRCLVA, AWFYPPLGRQFY, and KQMKEG. In αS2-casein, the most prominent motif was WPQYLKYLDQGPKVLKPWDQPKRNAGP, spanning 27 residues, along with FCIFSC and KKTIDD. In β-casein, motif 1 was VLILAC, motif 2 was GPFRGPFPN, and motif 3 was QRWMPQ. In κ-casein, motif 1 was RFCCEK, motif 2 was RHPHPGLNFM, and motif 3 was KKCQDK. The combined p-value for all identified motifs was 5.98e-19.

Phylogenetic analysis of sheep casein proteins

The phylogenetic trees of αS1-, αS2-, β-, and κ-casein proteins show the evolutionary relationships among *Ovis aries* (sheep), *Capra hircus* (goat), *Bos taurus* (cow), and *Bubalus bubalis* (buffalo) (Supplementary Figure S4). In the αS2-casein tree, sheep and goat formed a clade with 100% bootstrap support, while cow and buffalo formed a separate clade with the same level of support. The β-casein tree showed the same grouping, with 100% support for the sheep–goat clade and 99% support for the cow–buffalo clade. The κ-casein and αS1-casein trees also showed this structure, with sheep and goat in one clade and buffalo and cow in another, supported by bootstrap values of 99–100%.

Analysis of *in silico*-derived peptide fragments and their bioactive potential

The *in silico* hydrolysis of sheep milk caseins using pepsin, trypsin, and chymotrypsin, analyzed through the BIOPEP-UWM database, produced 46 peptides from αS1-casein, 42 from αS2-casein, 38 from β-casein, and 34 from κ-casein (Figure 1). Dipeptide counts were 16 for αS1-casein, 11 for β-casein, 11 for αS2-casein, and 5 for κ-casein. Tripeptide counts were 13 for αS1-casein, 10 for β-casein, 8 for αS2-casein, and 5 for κ-casein. κ-casein yielded 8 peptides longer than 7 amino acids (Table 3).

Table 3: Distribution of peptides by length generated from *in silico* digestion of milk proteins.

Peptide length	α-s1 casein	α-s2 casein	β casein	κ casein
Dipeptides (2 AA)	16	11	11	5
Tripeptides (3 AA)	13	6	6	6
Tetrapeptides (4 AA)	5	3	4	4
Pentapeptides (5 AA)	6	4	5	7
Hexapeptides (6 AA)	3	5	5	1
Heptapeptides (7AA)	1	7	2	3
Longer (>7 AA)	2	6	5	8

Comparative analysis of AMP predictions from casein-derived peptides using four classifiers

A comparative antimicrobial peptide (AMP) prediction for peptides derived from αS1-, αS2-, β-, and κ-caseins was performed using four machine learning classifiers: Support Vector Machine (SVM), Random Forest (RF), Artificial Neural Network (ANN), and Discriminant Analysis (DA) (Figure 2). From αS1-casein, three peptides EL, PL, and SM were identified by ANN; PL was also predicted as AMP by SVM and DA, EL by SVM, and SM had no additional confirmation. From αS2-casein, four peptides AL, AW, EH, and PW were identified by

ANN; AW and PW were also predicted as AMP by SVM and DA, EH by SVM and RF, and AL had no additional confirmation. From β-casein, five peptides DM, GPF, IH, PF, and VPK were identified by ANN; DM and PF were also predicted as AMP by all other classifiers, GPF by SVM, IH by RF, and VPK by no other classifier. From κ-casein, ANN identified GL, IAK, and IPIQY; IPIQY was also predicted as AMP by RF and DA, and GL and IAK had no additional confirmation.

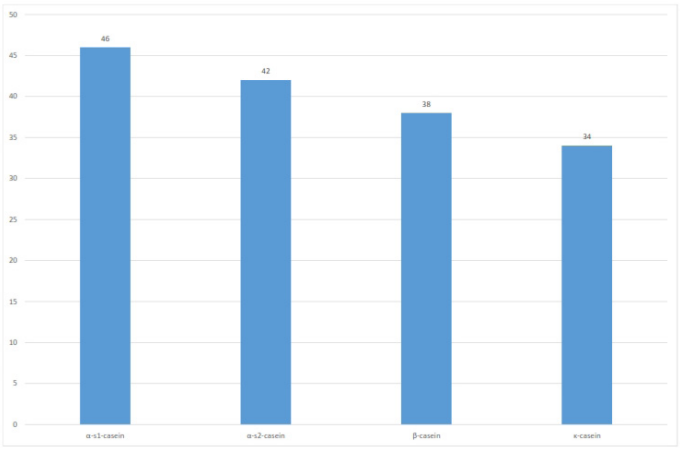


Figure 1: Total number of peptides derived from α-s1-casein, α-s2-casein, β-casein, κ-casein proteins of sheep milk by *in silico* gastrointestinal digestion.

In silico PREDICTION OF BIOACTIVE PEPTIDES

The BIOPEP-UWM database was used to identify biological activities associated with each peptide sequence. Table 4 presents the profile of bioactive peptides derived from αS1-, αS2-, β-, and κ-caseins, including their sequences, locations within the parent proteins, biological activities, and physicochemical properties. Across all casein types, peptides were predicted as non-toxic. Hydrophobicity values ranged from −0.92 to +0.54. The isoelectric points (pI) were observed in two ranges: near-neutral (~5.88) and basic (≥9.11), with charges varying from −2 to +1.

DISCUSSION

Food derived peptides, particularly those from dairy proteins, are gaining increasing attention as bioactive compounds with diverse health benefits. These include antioxidant, antihypertensive, and antimicrobial effects, making them valuable for functional food development (Capriotti *et al.*, 2016). Advanced analytical and computational techniques are now widely used to explore their structure–function relationships, mechanisms of action, and potential applications in nutrition and health.

In this study, a comprehensive *in silico* approach was applied to the main casein proteins of sheep milk (αS1-, αS2-, β-, and κ-casein) to predict their structural characteristics, enzymatic cleavage patterns, and potential bioactivities.

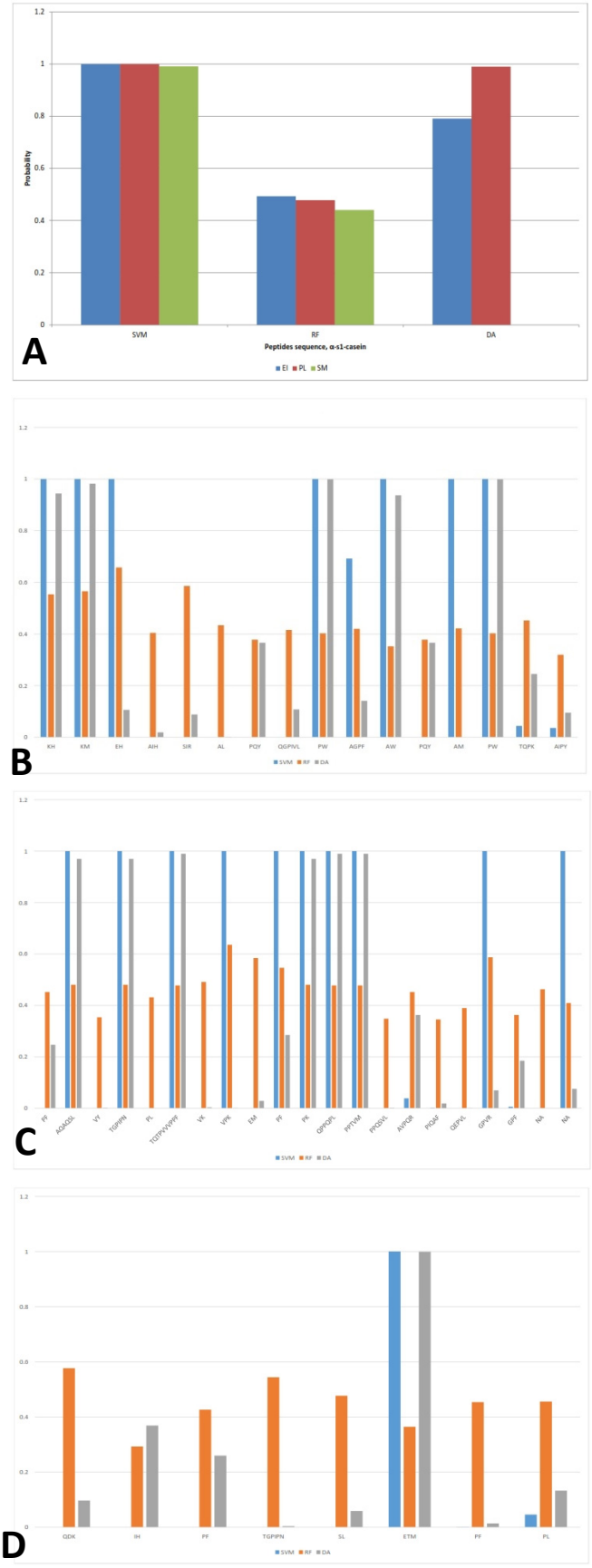


Figure 2: Prediction of antimicrobial peptides using statistical models from CAMPR3 database including random-Forest (RF), support-vector machines (SVM), artificial neural network (ANN), and discriminant-analysis (DA); (a) α-s1-casein, (b) α-s2-casein, (c) β-casein, (d) κ-casein.

Table 4: Physicochemical properties and predicted water solubility of predicted antimicrobial peptides from sheep casein proteins.

Sequence	Location	Activity	Hydrophobicity	Hydropathicity	Hydrophilicity	Charge	pI	Mol wt	Prediction
αS1-casein									
AY	[143-144]	ACE inhibitor; antioxidative; dipeptidyl peptidase IV inhibitor; tubulin-tyrosine ligase inhibitor	0.14	0.25	-1.4	0	5.88	252.28	Non-Toxin
EDVPSER	[84-90]	osteoanabolic; antioxidative	-0.5	-1.89	1.54	-2	4.14	830.93	Non-Toxin
EL	[39-40]	Antioxidative	-0.04	0.15	0.6	-1	4	260.31	Non-Toxin
GTQY	[170-173]	ACE inhibitor; antioxidative	-0.17	-1.47	-0.62	0	5.88	467.53	Non-Toxin
GY	[93-94]	ACE inhibitor; dipeptidyl peptidase IV inhibitor; inhibitor of tripeptidyl peptidase II	0.09	-0.85	-1.15	0	5.88	238.26	Non-Toxin
IN	[37-38]	dipeptidyl peptidase IV inhibitor	0.04	0.5	-0.8	0	5.88	245.3	Non-Toxin
PK	[2-3]	dipeptidyl peptidase IV inhibitor; antioxidative	-0.59	-2.75	1.5	1	9.11	243.32	Non-Toxin
PL	[168-169]	ACE inhibitor; dipeptidyl peptidase IV inhibitor; xaa-pro inhibitor; lactocepin inhibitor	0.23	1.1	-0.9	0	5.88	228.31	Non-Toxin
QF	[152-153]	dipeptidyl peptidase IV inhibitor; renin inhibitor	-0.04	-0.35	-1.15	0	5.88	293.34	Non-Toxin
QK	[131-132]	ACE inhibitor	-0.9	-3.7	1.6	1	9.11	274.34	Non-Toxin
QL	[155-156]	dipeptidyl peptidase IV inhibitor	-0.08	0.15	-0.8	0	5.88	259.33	Non-Toxin
SK	[41-42]	dipeptidyl peptidase IV inhibitor	-0.68	-2.35	1.65	1	9.11	233.28	Non-Toxin
SM	[122-123]	dipeptidyl peptidase III inhibitor	0	0.55	-0.5	0	5.88	236.3	Non-Toxin
αS2-casein									
AL	[82-83]	dipeptidyl peptidase IV inhibitor	0.39	2.8	-1.15	0	5.88	202.27	Non-Toxin
AW	[176-177]	ACE inhibitor; antioxidative; dipeptidyl peptidase IV inhibitor	0.31	0.45	-1.95	0	5.88	275.32	Non-Toxin
EH	[5-6]	dipeptidyl peptidase IV inhibitor	-0.51	-3.35	1.25	-0.5	5.25	284.29	Non-Toxin
EK	[34-35]	ACE inhibitor; dipeptidyl peptidase IV inhibitor	-0.86	-3.7	3	0	6.35	275.32	Non-Toxin
PR	[31-32]	ACE inhibitor; dipeptidyl peptidase III inhibitor	-0.92	-3.05	1.5	1	10.11	271.33	Non-Toxin
PW	[109-110]	antioxidative; antioxidative; dipeptidyl peptidase IV inhibitor; dipeptidyl peptidase IV inhibitor	0.15	-1.25	-1.7	0	5.88	301.36	Non-Toxin
QF	[88-89]	dipeptidyl peptidase IV inhibitor; renin inhibitor	-0.04	-0.35	-1.15	0	5.88	293.34	Non-Toxin
QK	[80-81]	ACE inhibitor	-0.9	-3.7	1.6	1	9.11	274.34	Non-Toxin
QY	[98-99]	dipeptidyl peptidase IV inhibitor	-0.33	-2.4	-1.05	0	5.88	309.34	Non-Toxin
SK	[136-137]	dipeptidyl peptidase IV inhibitor	-0.68	-2.35	1.65	1	9.11	233.28	Non-Toxin
TK	[149-150]	dipeptidyl peptidase IV inhibitor	-0.64	-2.3	1.3	1	9.11	247.31	Non-Toxin
TN	[199-200]	dipeptidyl peptidase IV inhibitor	-0.41	-2.1	-0.1	0	5.88	233.24	Non-Toxin
VR	[205-206]	ACE inhibitor; dipeptidyl peptidase IV inhibitor	-0.61	-0.15	0.75	1	10.11	273.35	Non-Toxin
β-casein									
DM	[182-183]	ACE inhibitor	-0.23	-0.8	0.85	-1	3.8	264.31	Non-Toxin
GPF	[201-203]	dipeptidyl peptidase IV inhibitor	0.23	0.27	-0.83	0	5.88	319.39	Non-Toxin

Table continues on next page.....

Sequence	Location	Activity	Hydrophobicity	Hydropathicity	Hydrophilicity	Charge	pI	Mol wt	Prediction
IH	[49-50]	dipeptidyl peptidase IV inhibitor; stimulating; dipeptidyl peptidase IV inhibitor; ACE inhibitor; neuropeptide	0.16	0.65	-1.15	0.5	7.1	268.34	Non-Toxin
IN	[26-27]	dipeptidyl peptidase IV inhibitor	0.04	0.5	-0.8	0	5.88	245.3	Non-Toxin
PF	[51-52]	dipeptidyl peptidase IV inhibitor; ACE2 inhibitor	0.27	0.6	-1.25	0	5.88	262.32	Non-Toxin
PK	[112-113]	dipeptidyl peptidase IV inhibitor antioxidative	-0.59	-2.75	1.5	1	9.11	243.32	Non-Toxin
PL	[76-77]	ACE inhibitor; dipeptidyl peptidase IV inhibitor; xaa-pro inhibitor; lactocepin inhibitor	0.23	1.1	-0.9	0	5.88	228.31	Non-Toxin
SL	[69-70]	dipeptidyl peptidase IV inhibitor; regulating	0.14	1.5	-0.75	0	5.88	218.27	Non-Toxin
TL	[126-127]	dipeptidyl peptidase IV inhibitor; hypouricemic	0.18	1.55	-1.1	0	5.88	232.3	Non-Toxin
TQTPV-VVPPF	[78-87]	hypouricemic	0.1	0.57	-0.76	0	5.88	1084.42	Non-Toxin
VK	[98-99]	ACE inhibitor; dipeptidyl peptidase IV inhibitor	-0.28	0.15	0.75	1	9.11	245.34	Non-Toxin
VL	[170-171]	stimulating; dipeptidyl peptidase IV inhibitor	0.54	4	-1.65	0	5.88	230.33	Non-Toxin
VPK	[103-105]	ACE inhibitor	-0.21	-0.43	0.5	1	9.11	342.47	Non-Toxin
VY	[59-60]	ACE inhibitor; antioxidative; dipeptidyl peptidase IV inhibitor	0.28	1.45	-1.9	0	5.88	280.34	Non-Toxin
κ-casein									
AK	[62-63]	neprilysin inhibitor; ACE inhibitor	-0.43	-1.05	1.25	1	9.11	217.28	Non-Toxin
AR	[96-97]	ACE inhibitor; neprilysin inhibitor	-0.76	-1.35	1.25	1	10.11	245.29	Non-Toxin
GL	[39-40]	ACE inhibitor; dipeptidyl peptidase IV inhibitor	0.35	1.7	-0.9	0	5.88	188.25	Non-Toxin
IAK	[22-24]	ACE inhibitor; antibacterial; hypotensive	-0.04	0.8	0.23	1	9.11	330.46	Non-Toxin
IN	[51-52]	dipeptidyl peptidase IV inhibitor	0.04	0.5	-0.8	0	5.88	245.3	Non-Toxin
IPIQY	[26-30]	dipeptidyl peptidase IV inhibitor	0.14	0.52	-1.14	0	5.88	632.83	Non-Toxin
PH	[103-104]	ACE inhibitor; dipeptidyl peptidase IV inhibitor; antioxidative	-0.24	-2.4	-0.25	0.5	7.1	252.29	Non-Toxin
PN	[80-81]	dipeptidyl peptidase IV inhibitor	-0.35	-2.55	0.1	0	5.88	229.25	Non-Toxin
PSY	[36-38]	ACE inhibitor	-0.1	-1.23	-0.67	0	5.88	365.41	Non-Toxin
PY	[57-58]	dipeptidyl peptidase IV inhibitor; phospholipase A2 inhibitor; neuropeptide; anti inflammatory; antioxidative	-0.03	-1.45	-1.15	0	5.88	278.32	Non-Toxin
QF	[54-55]	dipeptidyl peptidase IV inhibitor; renin inhibitor	-0.04	-0.35	-1.15	0	5.88	293.34	Non-Toxin
QW	[75-76]	dipeptidyl peptidase IV inhibitor	-0.16	-2.2	-1.6	0	5.88	332.38	Non-Toxin
SF	[106-107]	ACE inhibitor; dipeptidyl peptidase IV inhibitor; renin inhibitor	0.17	1	-1.1	0	5.88	252.28	Non-Toxin
VL	[31-32]	stimulating; dipeptidyl peptidase IV inhibitor	0.54	4	-1.65	0	5.88	230.33	Non-Toxin

Physicochemical analysis revealed notable differences among the four caseins. β-casein showed the highest instability and aliphatic index, indicating lower stability but greater adaptability within micellar structures. In contrast, αS2-casein was the most hydrophilic and had the highest isoelectric point, linked to its balanced composition of acidic and basic residues. These differences likely affect solubility, digestion behavior, and bioactive peptide release.

Secondary structure predictions showed that β- and κ-caseins are largely disordered, with high proportions of random coil regions. This flexibility supports their roles in micelle formation. Conversely, αS1- and αS2-caseins had a higher alpha-helix content, suggesting more stable structures. These features influence their enzymatic breakdown and the types of peptides they release.

Post-translational modification predictions identified multiple phosphorylation sites in β- and κ-caseins, particularly in regions associated with calcium binding and micelle stability. αS2-casein was the only protein predicted to have an N-glycosylation site, which may enhance its structural stability and resistance to digestion. Conserved motif analysis revealed short sequence patterns involved in disulfide bonding and the generation of functional peptide domains.

Simulated enzymatic digestion showed that αS1-casein generated the highest number of short peptides, such as di- and tripeptides, indicating high susceptibility to enzymatic cleavage. In contrast, κ-casein produced more longer peptides, suggesting greater resistance to digestion. These patterns align with the proteins' structural characteristics.

Partially hydrolyzed caseins are known to release peptides with beneficial biological effects. Many of these peptides are catalogued in the updated Milk Bioactive Peptide Database (MBPDB), which supports their potential use from laboratory studies to clinical applications (Nielsen *et al*., 2024).

Machine learning-based predictions revealed antimicrobial activity in several peptides across all casein types. Peptides such as PL (from αS1- and β-casein), AW and PW (from αS2-casein), and IPIQY (from κ-casein) demonstrated strong antimicrobial potential. These candidates also exhibited favorable properties, including low molecular weight, neutral charge, and non-toxicity.

Motif analysis further supported the functional potential of these peptides. κ-casein motifs rich in cysteine and proline suggest roles in structural stabilization and enzyme interactions. Similarly, β-casein motifs containing proline and glycine are typical of bioactive regions involved in blood pressure regulation and immune modulation.

Milk-derived peptides were assessed for potential bioactivity using the PeptideRanker tool, (Mooney *et al*., 2012). Peptides with scores exceeding 0.5 were considered bioactive, following the threshold suggested by (Imai *et al*., 2021).

AMP prediction confirmed several strong candidates, especially from αS1- and β-caseins. Many of these peptides contain basic or hydrophobic residues, which may help them disrupt bacterial membranes. The oligopeptides are particularly noteworthy as they are more likely to retain specific structural motifs linked to biological functions such as antihypertensive, antioxidant, or antimicrobial activities (Korhonen and Pihlanto, 2006; Udenigwe and Aluko, 2012).

Some peptides also showed multiple biological functions. These include ACE inhibition, antioxidant activity, and DPP-IV inhibition. The presence of similar motifs in different caseins and their fragments suggests that these functional sequences are evolutionarily conserved. This underlines the value of caseins as rich sources of multifunctional, health-promoting peptides.

Milk-derived peptides, particularly those from lactoferrins, have shown strong antioxidant, anti-inflammatory, and antihypertensive effects. These properties support their use as sustainable nutraceuticals for both human and animal health, in line with the One Health approach (Borges *et al*., 2025).

These peptides can be identified using classical enzymatic digestion and advanced bioinformatics tools. They offer several health benefits, such as reducing blood pressure, combating diabetes, lowering oxidative stress, and fighting harmful microbes (Kashung and Karuthapandian, 2025). This makes them valuable ingredients for functional foods, supplements, and health products.

The current results are in line with recent studies that emphasize the multifunctional potential of milk-derived peptides. For example, Singh and Gaur (2024) reviewed their antioxidant, antihypertensive, and antimicrobial effects. Advances in peptidomics and bioinformatics have made it easier to discover and analyze these peptides.

Despite their benefits, milk protein-derived peptides face challenges as functional ingredients. Many have low activity, are unstable in the digestive system, and are poorly absorbed. *in silico* tools such as QSAR modelling, molecular docking, and design of experiments (DoE) offer solutions. These methods help identify, optimize, and design peptides with improved potency and stability (FitzGerald *et al*., 2020).

Barati *et al.* (2020) used *in silico* approaches to assess digestion resistance and bioactive peptide content in milk proteins. While informative, their findings need experimental validation. Future research should include lab-based digestion studies, absorption assessments, and long-term clinical trials. These are essential to confirm the true health impacts of food-derived peptides. Today, bioinformatics plays a central role in peptide research. Tools like proteolysis simulation, QSAR analysis, and molecular docking accelerate the screening and design of effective peptides. This supports their development for applications in food, pharmaceuticals, and cosmetics (Du *et al.*, 2023). This study was conducted entirely through *in silico* analyses, which, while efficient and comprehensive, cannot fully replicate the complexity of *in vitro* or *in vivo* conditions (Hoda *et al.*, 2023). The predicted bioactivities, structural features, and digestion profiles of casein-derived peptides require experimental validation to confirm their functional relevance. Additionally, bioactivity predictions were based on specific machine learning algorithms (SVM, RF, ANN, DA) and reference databases (BIOPEP-UWM, PeptideRanker), which may be limited by dataset size, species coverage, and algorithm-specific biases.

The findings of this study reveal the strong potential of sheep milk-derived peptides such as PL, AW, PW, and IPIQY as natural bioactive compounds with antimicrobial, antioxidant, and antihypertensive properties. Their multifunctional activities, combined with favorable safety and physicochemical characteristics, make them promising candidates for enhancing animal health and resilience under farm conditions. In light of these results, we plan to apply them at the Center for Agricultural Technology Transfer in Korça through pilot trials and on-farm testing. This effort aims to evaluate their practical benefits as functional feed additives, reduce reliance on antibiotics, and promote more sustainable and health-oriented livestock practices. By introducing these peptides in a controlled and locally adapted setting, the Center can take a leading role in translating research into practice supporting innovation in animal nutrition aligned with the principles of One Health and the circular bioeconomy.

CONCLUSION

This *in silico* study comprehensively characterized αS1, αS2-, β-, and κ-caseins from sheep milk, analyzing their physicochemical properties, secondary structures, post-translational modifications, conserved motifs, and evolutionary relationships. Simulated gastrointestinal digestion revealed several promising bioactive peptides notably PL, AW, PW, DM, PF, and IPIQY with predicted antimicrobial, antioxidant, ACE-inhibitory, and DPP-IV-inhibitory activities. All candidate peptides were predicted to be non-toxic and exhibited favorable physicochemical characteristics, supporting their potential for safe use. These findings provide a solid foundation for experimental validation and future application-oriented research, with the potential to benefit both human and animal health while promoting sustainable production systems.

ACKNOWLEDGMENTS

This research was supported by the National Agency for Scientific Research and Innovation (NASRI), Albania, under the Technology and Innovation Projects (TIP) framework, as part of the project titled "Integration of genetic and environmental factors to enhance productivity and health in indigenous sheep."

NOVELTY STATEMENT

To our knowledge, this study provides the first comprehensive in silico exploration of bioactive peptides derived from sheep milk caseins, integrating physicochemical profiling, secondary structure prediction, post-translational modification analysis, motif discovery, phylogenetic assessment, simulated gastrointestinal digestion, and multi-model machine learning screening for antimicrobial, antioxidant, and antihypertensive activity. The study identifies novel multifunctional peptide candidates (such as PL, AW, PW, DM, PF, and IPIQY) characterized by predicted safety, stability, and therapeutic potential. These findings highlight sheep milk–derived peptides as promising natural bioactives that can enhance sustainable livestock health management within the frameworks of *One Health* and the circular bioeconomy.

AUTHOR'S CONTRIBUTION

Anila Hoda: Conceptualization, supervision, project administration, methodology design, data interpretation, and original draft preparation.
Sulltane Ajçe: Data collection, visualization, and review the manuscript..
Ilia Mikerezi: Sequence analysis, computational analysis, phylogenetic evaluation and review the manuscript.
Xhelil Koleci: Functional annotation, antimicrobial peptide screening, and validation of physicochemical predictions and review the manuscript.

Generative AI and AI-assisted technology statement

Generative AI tools (ChatGPT, OpenAI, USA) were used to assist in the refinement of the manuscript's language, structure, and clarity. All scientific content, analysis, data interpretation, and conclusions were developed entirely by the authors. AI tools were not used to generate or alter

data, figures, tables, or scientific results. The authors have carefully reviewed and validated all text edited with AI assistance to ensure accuracy and integrity in accordance with the Journal's publication ethics.

Conflict of interest

The authors have declared no conflict of interest.

REFERENCES

Bailey TL, Boden M, Buske FA, Frith M, Grant CE, Clementi L, Ren J, Li WW, Noble WS (2009). MEME SUITE: Tools for motif discovery and searching. Nucl. Acids Res., 37: W202-208. https://doi.org/10.1093/nar/gkp335

Barati M, Javanmardi F, Jabbari M, Mokari-Yamchi A, Farahmand F, Eş I, Farhadnejad H, Davoodi SH, Mousavi Khaneghah A (2020). An *in silico* model to predict and estimate digestion-resistant and bioactive peptide content of dairy products: A primarily study of a time-saving and affordable method for practical research purposes. LWT 130: 109616. https://doi.org/10.1016/j.lwt.2020.109616

Blom N, Gammeltoft S, Brunak S (1999). Sequence and structure-based prediction of eukaryotic protein phosphorylation sites. J. Mol. Biol., 294: 1351–1362. https://doi.org/10.1006/jmbi.1999.3310

Borges T, Coelho P, Prudêncio C, Gomes A, Gomes P, Ferraz R (2025). Bioactive peptides from milk proteins with antioxidant, anti-inflammatory, and antihypertensive activities. Foods, 14: 535. https://doi.org/10.3390/foods14030535

Capriotti AL, Cavaliere C, Piovesana S, Samperi R, Laganà A (2016). Recent trends in the analysis of bioactive peptides in milk and dairy products. Anal. Bioanal. Chem., 408: 2677–2685. https://doi.org/10.1007/s00216-016-9303-8

Dhar H, Verma S, Dogra S, Katoch S, Vij R, Singh G, Sharma M (2024). 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., 64: 9432–9454. https://doi.org/10.1080/10408398.2023.2212803

Du Z, Comer J, Li Y (2023). Bioinformatics approaches to discovering food-derived bioactive peptides: Reviews and perspectives. TrAC Trends Anal. Chem., 162: 117051. https://doi.org/10.1016/j.trac.2023.117051

El-Salam MHA, El-Shibiny S (2013). Bioactive peptides of buffalo, camel, goat, sheep, mare, and yak milks and milk products. Food Rev. Int., 29: 1–23. https://doi.org/10.1080/87559129.2012.692137

FitzGerald RJ, Cermeño M, Khalesi M, Kleekayai T, Amigo-Benavent M (2020). Application of *in silico* approaches for the generation of milk protein-derived bioactive peptides. J. Funct. Foods 64: 103636. https://doi.org/10.1016/j.jff.2019.103636

de la Fuente MA, Mercedes R, Isidra R, Manuela J (2013). Sheep milk. In: Milk and dairy products in human nutrition. John Wiley and Sons, Ltd, 554–577. https://doi.org/10.1002/9781118534168.ch25

Gasteiger E, Hoogland C, Gattiker A, Wilkins MR, Appel RD, Bairoch A, others (2005). Protein identification and analysis tools on the ExPASy server. The proteomics protocols handbook: 571–607. https://doi.org/10.1385/1-59259-890-0:571

Geourjon C, Deléage G (1995). SOPMA: Significant improvements in protein secondary structure prediction by consensus prediction from multiple alignments. Bioinformatics, 11: 681–684. https://doi.org/10.1093/bioinformatics/11.6.681

Hati S, Singh BP, Gawai K, Sarkar P (2024). Production, characterization and bio functionalities of bioactive peptides from non-bovine species of milk: A review. Indian J. Anim. Health,

Hoda A, Bixheku X, Lika (Çekani) M (2023). Computational analysis of non-synonymous single nucleotide polymorphism in the bovine PKLR gene. J. Biomol. Struct. Dyn., 0: 1–14. https://doi.org/10.1080/07391102.2023.2219315

Imai K, Shimizu K, Honda H (2021). Machine learning screening of bile acid-binding peptides in a peptide database derived from food proteins. Sci. Rep., 11: 16123. https://doi.org/10.1038/s41598-021-95461-1

Kashung P, Karuthapandian D (2025). Milk-derived bioactive peptides. Food Prod. Process. Nutr., 7: 6. https://doi.org/10.1186/s43014-024-00280-2

Korhonen H, Pihlanto A (2006). Bioactive peptides: Production and functionality. Int. Dairy J., 16: 945–960. https://doi.org/10.1016/j.idairyj.2005.10.012

Minkiewicz P, Dziuba J, Iwaniak A, Dziuba M, Darewicz M (2008). BIOPEP and other programs for processing bioactive peptide sequences. J. AOAC Int., 91: 965–980. https://doi.org/10.1093/jaoac/91.4.965

Mooney C, Haslam NJ, Pollastri G, Shields DC (2012). Towards the improved discovery and design of functional peptides: Common features of diverse classes permit generalized prediction of bioactivity. PLoS One, 7: e45012. https://doi.org/10.1371/journal.pone.0045012

Nielsen SD-H, Liang, Ningjian, Rathish, Harith, Kim, Bum J, Lueangsakulthai, Jiraporn, Koh, Jeewon, Qu, Yunyao, Schulz, Hans-Jörg, Dallas DC (2024). Bioactive milk peptides: An updated comprehensive overview and database. Crit. Rev. Food Sci. Nutr., 64: 11510–11529. https://doi.org/10.1080/10408398.2023.2240396

Quintieri L, Fanelli F, Monaci L, Fusco V (2024). Milk and its derivatives as sources of components and microorganisms with health-promoting properties: Probiotics and bioactive peptides. Foods, 13: 601. https://doi.org/10.3390/foods13040601

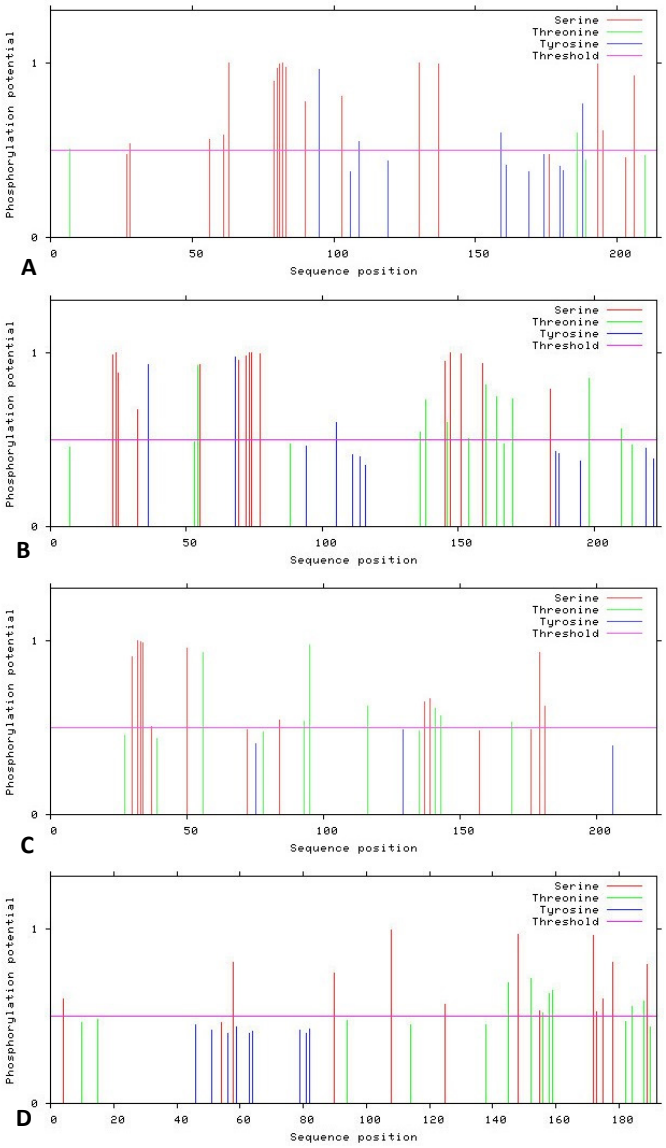
Sathika, K, Priyadharshini M, Harini B., (2025). Revealing the potential of bioactive peptides from sheep milk casein and whey: An *in-silico* exploration. BIO Web Conf., 172: 02008. https://doi.org/10.1051/bioconf/202517202008

Singh N, Gaur S (2024). New insights into multifunctional aspects of milk derived bioactive peptides: A review. Food Chem. Adv., 4: 100628. https://doi.org/10.1016/j.focha.2024.100628

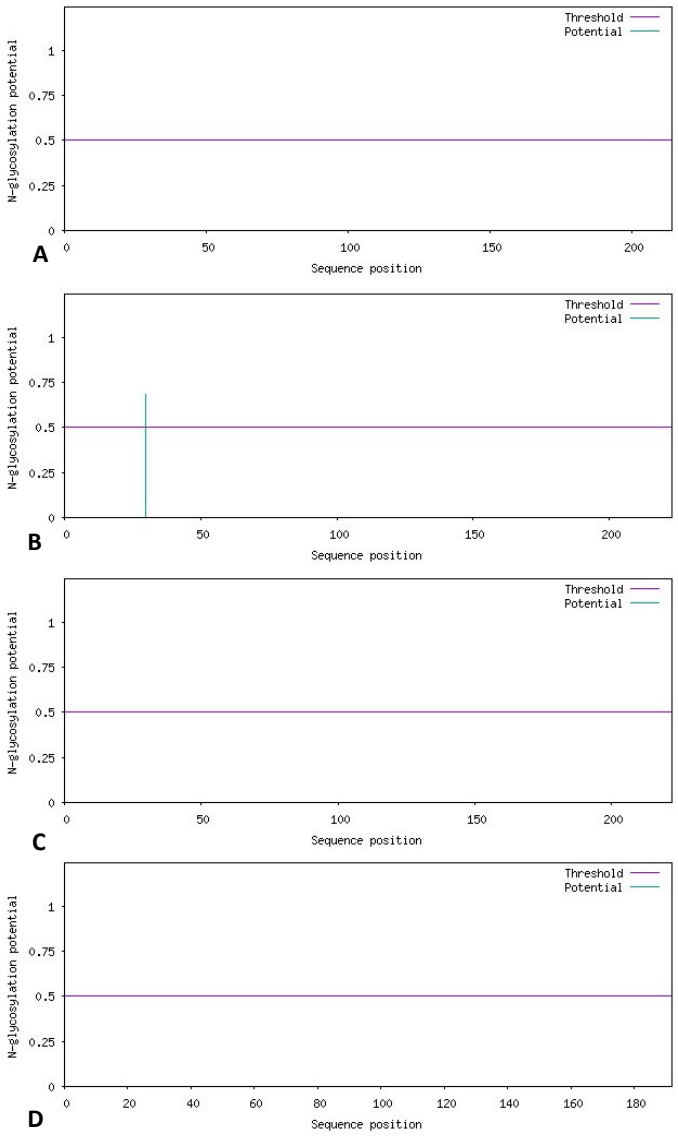
Tamura K, Stecher G, Kumar S (2021). MEGA11: molecular evolutionary genetics analysis version 11. Mol. Boil. Evolut., 38: 3022–3027. https://doi.org/10.1093/molbev/msab120

Udenigwe CC, Aluko RE (2012). Food protein-derived bioactive peptides: Production, processing, and potential health benefits. J. Food Sci., 77: R11–R24. https://doi.org/10.1111/j.1750-3841.2011.02455.x

Waghu FH, Gopi L, Barai RS, Ramteke P, Nizami B, Idicula-Thomas S (2014). CAMP: Collection of sequences and structures of antimicrobial peptides. Nucl. Acids Res., 42: D1154-1158. https://doi.org/10.1093/nar/gkt1157

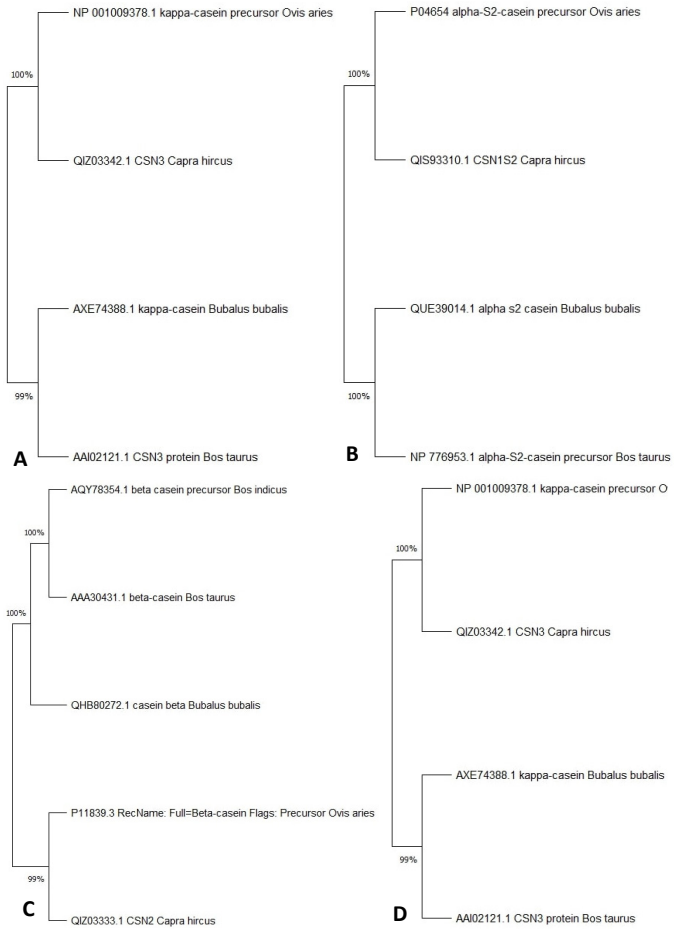


Supplementary Figure S1: Graphical representation of predicted phosphorylation sites (a) αS1-casein; (b) αS2-casein, (c) β-casein and (d) κ-casein, highlighting the specific positions of predicted phosphorylation residues along each protein sequence.



Supplementary Figure S2: Graphical representation of N-glycosylation sites (a) αS1-casein; (b) αS2-casein, (c) β-casein and (d) κ-casein.

Supplementary Figure S3: Conserved motif sequences identified by MEME suite tool in (a) αS1-casein; (b) αS2-casein, (c) β-casein and (d) κ-casein protein.



Supplementary Figure S4: Phylogenetic trees of (a) αS1-casein, (b) αS2-casein, (c) β-casein, and (d) κ-casein proteins were constructed using the Neighbor-Joining (NJ) method. Bootstrap values from 1000 replicates are indicated at the nodes to assess the robustness of the inferred relationships.