Indian Journal of Biochemistry & Biophysics
Vol. 62, July 2025, pp. 700-711
DOI: 10.56042/ijbb.v62i7.15925

Tracing the evolutionary history of novel hormone asprosin: An *in silico* study across vertebrates

Priyanka Sathoria[1], Bhawna Chuphal[2], Vaishnavi Rajagopalan[1], Kajal[1], Arya Singh[1], Khushi Tyagi[1], Jyoti Singh[1], Umesh Rai[3] & Brototi Roy[1]*

[1]Department of Zoology, Maitreyi College, University of Delhi, New Delhi-110 021, Delhi, India

[2]ICMR-National Institute of Child Health & Development Research, Safdarjang Hospital Campus, New Delhi-110 029, Delhi, India

[3]University of Jammu, Jammu-180 006, Jammu and Kashmir, India

Received 23 December 2024; *revised 04 March 2025*

Although, asprosin is implicated in regulation of various physiological functions and metabolic disorders, there are no reports in non-mammalian vertebrates except in fish *Channa punctata*. Hence, in this study we explore the asprosin across the vertebrate group through *in silico* analysis. This novel hormone is by product of enzymatic cleavage of profibrillin protein (encoded by *FBN1* gene) by furin protease. We have focused on the comparative analysis of physicochemical properties, structure and evolutionary relationship of putative asprosin. The physicochemical properties of putative asprosin across the vertebrate groups revealed thermostability, ex vivo stability and its hydrophilic nature. The secondary and tertiary structures of putative asprosin revealed beta strands that provide the stability and help in folding of protein. The sequence homology of putative asprosin primary sequence reveals more than 50% conservation across the vertebrates. The crucial post-translational modifications such as phosphorylation and glycosylation are present in putative asprosin. Asprosin was observed to be subjected to purifying selection, suggesting limited changes in structure and function of asprosin over extensive evolutionary period. Further, phylogenetic analysis of asprosin showed that bony fishes form a separate clade distinct from mammals, birds, reptiles and amphibians. This study for the first time provides an insight into the conservation of *fbn1* encoded profibrillin protein, furin cleavage site in profibrillin protein and its C-terminal cleavage product, asprosin, across the vertebrate groups. The conserved physicochemical properties and strong purifying selection showed that asprosin was under strong evolutionary pressure.

Keywords: Asprosin, Mammalian, Non-vertebrates, Physicochemical properties

Vertebrate physiology is regulated by the interplay of several complex endogenous physiological systems with the environmental stimuli. The endocrine system plays a key role in regulating the metabolic physiology and reproduction. In recent years, the adipose tissue has been recognized as a crucial and active endocrine organ, which synthesizes and secretes several adipokines such as leptin, adiponectin, visfatin, chemerin, *etc*. Dysfunction in these adipokines secretion results in obesity, polycystic ovarian syndrome and diabetes[1,2]. However, few adipokines have been demonstrated to have different physiological roles in other vertebrate groups compared to the mammals[3-7]. Also, the major site of synthesis of adipokine might also vary depending upon the vertebrate groups[3-7]. Asprosin a novel glucogenic hormone, discovered in 2016, has been recently added to the ever-expanding list of adipokines and is the focus of intense research due to its pleiotropic effect in regulation of multiple key physiological processes[8,9]. However, since its discovery, studies on asprosin are confined to the mammals except a single report in teleost *Channa punctata*[10].

Asprosin in mammals is encoded by the C-terminal region of the *FBN1* gene. This gene containing 66 exons encodes profibrillin protein, which is cleaved by furin enzyme into fibrillin 1 and 140 amino acid long asprosin protein (encoded by 65 and 66 exon)[8]. Although *FBN1* is a widespread gene in vertebrates, asprosin has not been specifically described in non-mammals with one exception[10]. In mammals, asprosin is reported to be an important metabolic hormone having glucogenic[8] and orexigenic properties[11]. It regulates glucose homeostasis[8] and insulin secretion[12]. Moreover, it is also implicated in inflammation[13], apoptosis[14] and reproductive functions[15-19]. Not

*Correspondence:
E-mail: broy@maitreyi.du.ac.in
Suppl. data available on respective page of NOPR

surprisingly, dysregulation in asprosin levels has been implicated in several metabolic disorders such as obesity, insulin resistance[20,21], type II diabetes mellitus[22], and polycystic ovarian syndrome[23,24]. In addition to this, asprosin is also involved in the anti-hypersensitivity effect in case of neuropathic pain models[25] and regulation of several signalling pathways in ovarian cancer[26]. Although, unique asprosin specific receptor has not been discovered so far, asprosin is reported to act through interaction with the Olfr734 (olfactory receptor 734)[27], TLR4 (Toll-like receptor 4)[12] and Ptprd (Protein tyrosine receptor kinase δ)[11] in different tissues.

In the present *in silico* study, we investigate the presence of asprosin in different vertebrate groups and elucidate their physicochemical properties. In addition to this, we seek to illuminate the probable secondary and tertiary structure, andthe post-translational modifications. Further, codon-based selection analysis was performed. The evolutionary relationship of putative asprosin across vertebrate groups was explored using phylogenetic analysis.

Materials and Methods

Retrieval of *fbn1* encoded profibrillin protein and prediction of furin cleavage site

The protein sequence of profibrillin encoded by the *fbn1* gene was retrieved from the NCBI (National Centre for Biotechnology Information) GenBank for Mammals: *Homo sapiens* (NP_000129.3), *Mus musculus* (NP_032019.2), *Rattus norvegicus* (NP_114013.2), *Bos taurus* (NP_776478.1), *Felis catus* (XP_023111155.1), *Myotis lucifugus* (XP_006086058.1), *Sarcophilus harrisii* (XP_031811012.1), *Ornithorhynchus anatinus* (XP_028926273.1); Birds : *Gallus gallus* (XP_015147420.1), *Columba livia* (XP_021153532.1), *Taeniopygia guttata* (XP_030137628.3), *Pseudopodoces humilis* (XP_005521568.1), *Sturnus vulgaris* (XP_014729550.1), *Lepidothrix coronata* (XP_017660646.1), *Falco peregrinus* (XP_027648361.1), *Haliaeetus leucocephalus* (XP_010570170.1); Reptiles: *Chelonia mydas* (XP_037766466.1), *Pelodiscus sinensis* (XP_006135123.1), *Gavialis gangeticus* (XP_019359370.1), *Alligator sinensis* (XP_014377923.2), *Anolis carolinensis* (XP_003229148.2), *Podarcis muralis* (XP_028561406.1), *Gekko japonicus* (XP_015267268.1), *Protobothrops mucrosquamatus* (XP_029140016.1); Amphibians: *Xenopus tropicalis* (XP_002936615.3), *Rhinatremabi vittatum* (XP_029430269.1), *Rana temporaria* (XP_040198586.1), *Bufo gargarizans* (XP_044135564.1), *Microcaecilia unicolor* (XP_030045415.1), *Spea bombifrons* (XP_053321481.1), Bombina bombina (XP_053573434.1); Fishes: *Scomber japonicus* (XP_053179269.1), *Oreochromis niloticus* (XP_005461642.1), *Hippocampus comes* (XP_019742290.1), *Oryzias latipes* (XP_023808941.1), *Takifugu rubripes* (XP_003969883.1), *Paralichthys olivaceus* (XP_019955037.1), *Sparus aurata* (XP_030270453.1), *Danio rerio* (XP_017207479.2), *Esox lucius* (XP_012992552.2) and *Amblyraja radiata* (XM_033014876.1).

The multiple sequence alignment was constructed for the predicted profibrillin protein across the vertebrate groups using Clustal omega. In the aligned sequences, furin cleavage site (R-X-K/R-R↓ or R/K-X-X-X-K/R-R↓; where ↓ represents cleavage site, / represents eitherK or R, X for any amino acid, R and K indicates arginine and lysine) was manually marked near the C-terminal region[28]. The amino acid sequence after the furin cleavage site in the profibrillin protein represents the putative asprosin sequence[8].

Structural analysis of putative asprosin

The primary putative sequences of asprosin across the vertebrate groups predicted on the basis of presence of furin cleavage site were subjected to ProtParam tool (http://expasy.org/cgi-in/protparam) for determining the physicochemical properties such as molecular weight, aliphatic index, instability index and grand average hydropathy (GRAVY) index[29]. Asprosin putative sequences were submitted to I-TASSER (Iterative-Threading ASSEmbly Refinement) server to construct and analyse secondary and tertiary structure[30]. After 3D modelling, the tertiary structures were subjected to the PROCHECK module (http://servicesn.mbi.ucla.edu/SAVES/) for geometry analysis using the Ramachandran plot[31].The various physicochemical properties and Ramachandran plot determines the stability of the protein. The value of instability index below 40 indicates the stability of protein under *ex vivo* conditions[32]. The relative volume occupied by the aliphatic side chain residues in a protein determines its aliphatic index, which ranges from 71.13-143.54[33]. Kyte & Doolittle in 1982[34] had elucidated that positive GRAVY index denotes the hydrophobicity and membrane-bound proteins whereas negative GRAVY value indicates

the hydrophilicity and globular nature of the proteins. Further, to determine the post-translational modifications in the asprosin sequences, Motifscan software (https://myhits.sib.swiss/cgi-bin/motif_scan) was employed. Clustal omega software was employed for construction of multiple sequence alignment of putative asprosin sequence using ClustalW method to determine the conservation of amino acid across the vertebrate groups and calculating the percentage similarity of putative asprosin of all the vertebrate groups with the asprosin of *Homo sapiens*.

Evolutionary analysis of putative asprosin

To explore the evolutionary pattern, codon-based selection analysis was conducted on the codons of *fbn1*coding asprosin. The nucleotide sequences of *fbn1* were retrieved from the NCBI and subjected to ORF finder (https://www.ncbi.nlm.nih.gov/orffinder) to determine the complementary determining sequence (CDS). The codons from the C-terminus region of *fbn1* was selected, which codes for asprosin for all the species, excluding the stop codon. The codon alignment of this file was done using ClustalW in MEGA11 (Molecular Evolutionary Genetics Analysis Version 11) and further analysis was conducted in Datamonkey server[35,36]. FUBAR (fast, unconstrained Bayesian approximation) with 0.9 posterior probability was employed for the detection of site-specific selection for the entire codons of gene, whether undergoing positive or negative selection[37]. For recognition of episodic selection (sudden changes) in the codon sequence for specific lineages, MEME (mixed effects model of evolution; $p = 0.05$) was used[38].

In order to understand the evolutionary history of asprosin across vertebrates, MEGA11 was employed to align all amino acid sequences using ClustalWand neighbor-joining method with Jones-Taylor-Thornton substitution model, uniform rate model and pairwise deletion of missing data treatment was used to construct the midpoint rooted phylogenetic tree of putative asprosin. In addition to this, 1000 bootstrap replicates were used to assess the reliability of constructed tree.

Results

Furin cleavage site and putative asprosin

The retrieved sequences of putative profibrillin proteins from the NCBI were employed for multiple sequence alignment. In the multiple sequence alignment, the furin cleavage site, R/K-X-X-X-K/R-R↓ sequence is conserved across the vertebrate groups as shown in (Fig. 1). Following the furin cleavage

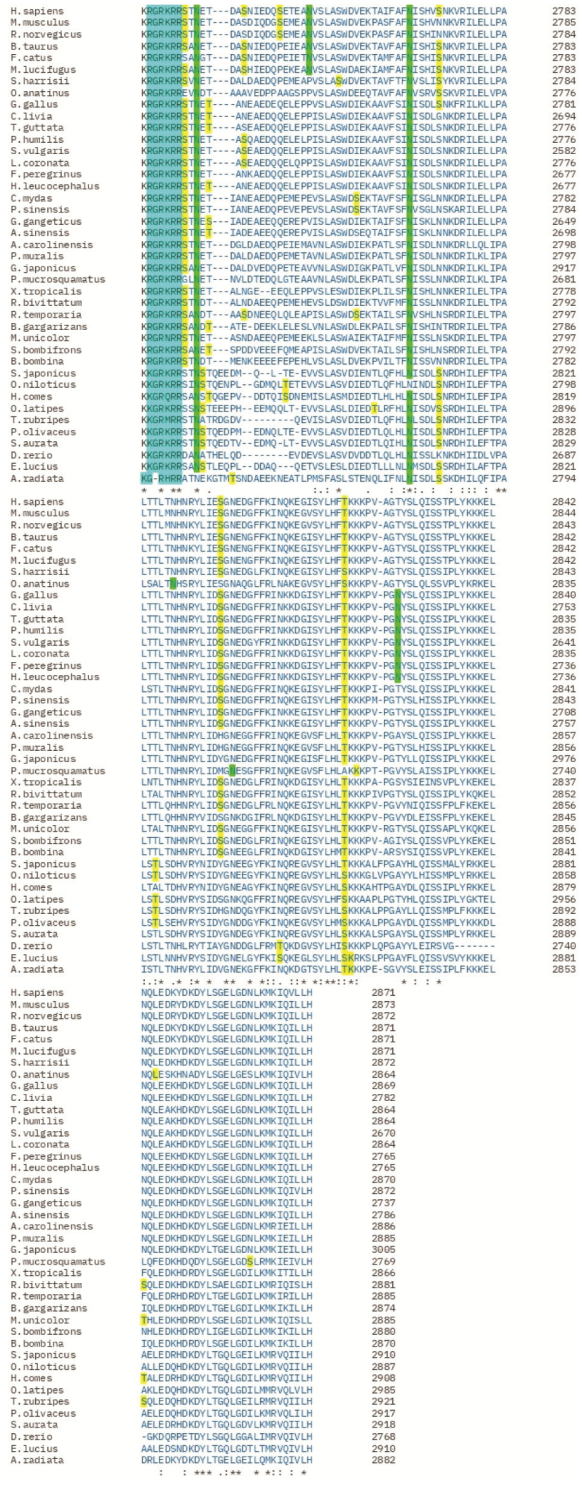

Fig. 1 — Conserved furin cleavage site and post-translational modifications in profibrillin protein. The conserved furin cleavage site, R/K-X-X-X-K/R-R↓ in profibrillin protein is highlighted in turquoise colour. The text in blue colour represents the putative asprosin sequences in different vertebrate groups species. The green and yellow highlighted text indicate the glycosylation and phosphorylation sites in the asprosin

site, the obtained amino acid sequence towards the C-terminal region of profibrillin protein represents the putative asprosin sequences.

Structural analysis of putative asprosin

The analysis of the physicochemical properties of putative asprosin across the vertebrate groups revealed that predicted asprosin in mammalian and reptilian group have 140 amino acids, birds have 139 amino acids, and the amphibians and fish have 139 to 143 amino acids. However, only 128 amino acids and 135 amino acids were present in the predicted asprosin in teleost *D. rerio* and in *T. rubripes*, respectively. The molecular weight of asprosin was found to be approximately 16 kDa across the vertebrate groups. The theoretical pI of all species studied was below 7 except *O. anatinus*. The instability index of all the vertebrate species was found either 40 or below 40 except *B. taurus*, *F. catus*, *O. anatinus*, *X. tropicalis*, *O. latipes* and *P. olivaceous*. The predicted asprosin in all the species across the vertebrate groups were found to have an aliphatic index of approximately 90 or above 90 and had negative GRAVY index value (Table 1).

The secondary and tertiary structural analysis using I-TASSER software predicted presence of multiple beta strands and coils in the putative asprosin across

Table 1 — Physicochemical properties of asprosin protein across the vertebrate groups

Vertebrate groups	Species	Number of amino acids	Molecular weight (kDa)	Theoretical pI	Total number of negatively charged residues (Asp + Glu)	Total number of positively charged residues (Arg + Lys)	Instability index	Aliphatic index	GRAVY index
Mammals	*H. sapiens*	140	15.88	5.55	20	16	37.84	89.86	-0.549
	M. musculus	140	15.89	5.54	20	16	31.00	87.07	-0.558
	R. norvegicus	140	15.88	5.54	20	16	36.37	87.79	-0.551
	B. taurus	140	15.88	5.79	19	16	42.75	94.07	-0.498
	F. catus	140	15.84	6.07	18	16	42.23	91.29	-0.506
	M. lucifugus	140	15.94	6.19	19	17	37.63	90.64	-0.552
	S. harrisii	140	15.88	5.35	21	16	28.18	100.29	-0.396
	O. anatinus	140	15.30	7.19	15	15	42.15	96.14	-0.225
Birds	*G. gallus*	139	15.87	5.80	21	18	39.01	94.75	-0.597
	C. livia	139	15.82	5.36	22	17	37.69	95.47	-0.635
	T. guttata	139	15.77	5.55	21	17	37.34	96.19	-0.580
	P. humilis	139	15.78	5.79	20	17	35.96	98.99	-0.541
	S. vulgaris	139	15.78	5.55	21	17	36.73	96.19	-0.609
	L. coronata	139	15.77	5.79	20	17	39.69	96.19	-0.580
	F. peregrinus	139	15.85	5.80	21	18	40.01	95.47	-0.640
	H. leucocephalus	139	15.85	5.36	22	17	38.30	95.47	-0.637
Reptiles	*C. mydas*	140	15.96	5.20	22	16	30.70	88.50	-0.657
	P. sinensis	140	15.87	5.51	21	16	28.04	87.07	-0.599
	G. gangeticus	140	16.09	5.81	22	19	37.35	94.07	-0.645
	A. sinensis	140	16.08	5.58	22	18	36.97	91.29	-0.679
	A. carolinensis	140	15.90	5.31	22	16	40.35	96.86	-0.583
	P. muralis	140	15.91	5.64	22	17	33.53	94.07	-0.614
	G. japonicus	140	15.94	5.33	22	17	30.84	103.79	-0.479
	P. mucrosquamatus	140	15.75	5.79	19	16	28.99	101.71	-0.352
Amphibians	*X. tropicalis*	138	15.81	5.08	23	15	42.94	102.46	-0.578
	R. bivittatum	141	16.19	5.04	23	15	36.30	94.75	-0.555
	R. temporaria	140	16.04	5.64	21	16	39.35	100.36	-0.497
	B. gargarizans	139	15.99	5.48	24	18	38.43	107.34	-0.491
	M. unicolor	140	15.85	5.81	20	17	33.13	91.36	-0.528
	S. bombifrons	140	16.00	5.32	23	16	35.40	104.50	-0.440
	B. bombina	140	16.15	5.37	24	18	39.26	100.93	-0.547

(*Contd*.)

Table 1 — Physicochemical properties of asprosin protein across the vertebrate groups (*Contd*.)

Vertebrate groups	Species	Number of amino acids	Molecular weight (kDa)	Theoretical pI	Total number of negatively charged residues (Asp + Glu)	Total number of positively charged residues (Arg + Lys)	Instability index	Aliphatic index	GRAVY index
Fishes	*S. japonicus*	138	15.96	5.28	22	14	40.97	96.09	-0.508
	O. niloticus	142	16.26	5.24	21	13	34.94	102.25	-0.406
	H. comes	142	16.14	5.36	22	14	26.67	98.94	-0.490
	O. latipes	141	15.97	5.23	21	12	48.78	95.39	-0.425
	T. rubripes	135	15.37	5.73	19	14	30.17	100.37	-0.450
	P. olivaceous	141	16.06	4.71	25	13	42.74	93.33	-0.593
	S. aurata	140	15.95	5.00	23	14	33.96	98.14	-0.480
	D. rerio	128	14.24	5.36	18	12	35.01	105.08	-0.332
	E. lucius	139	15.56	4.93	19	12	36.51	101.73	-0.334
	A. radiata	143	16.22	5.57	21	17	26.47	93.43	-0.456

Table 2 — Ramachandran plot analysis of putative asprosin of all species across the vertebrate groups

Vertebrate groups	Species	Residues in most favoured region	Residues in additional allowed region	Residues in generously allowed region	Residues in disallowed region
Mammals	*M. musculus*	51.2%	40.2%	3.9%	4.7%
	R. norvegicus	52.0%	37.8%	7.9%	2.4%
	B. taurus	50.0%	39.8%	7.0%	3.1%
	F. catus	46.5%	40.9%	7.9%	4.7%
	M. lucifugus	43.0%	41.4%	10.9%	4.7%
	S. harrisii	48.8%	41.7%	7.1%	2.4%
	O. anatinus	44.4%	39.5%	9.7%	6.5%
Birds	*G. gallus*	48.0%	40.8%	8.0%	3.2%
	C. livia	47.6%	41.9%	9.7%	0.8%
	T. guttata	58.4%	34.4%	4.8%	2.4%
	P. humilis	50.8%	38.1%	8.7%	2.4%
	S. vulgaris	44.8%	44.0%	7.2%	4.0%
	L. coronata	41.6%	48.8%	8.0%	1.6%
	F. peregrinus	52.0%	40.0%	3.2%	4.8%
	H. leucocephalus	37.6%	44.8%	12.8%	4, 8%
Reptiles	*C. mydas*	52.8%	36.8%	8.0%	2.4%
	P. sinensis	48.8%	38.4%	12.0%	0.8%
	G. gangeticus	49.6%	36.2%	10.2%	3.9%
	A. sinensis	48.8%	40.2%	6.3%	4.7%
	A. carolinensis	53.2%	39.5%	4.8%	2.4%
	P. muralis	52.8%	37.6%	7.2%	2.4%
	G. japonicus	48.8%	40.0%	8.0%	3.2%
	P. mucrosquamatus	50.8%	38.1%	7.9%	3.2%
Amphibians	*X. tropicalis*	40.2%	40.2%	13.9%	5.7%
	R. bivittatum	62.0%	31.8%	4.7%	1.6%
	R. temporaria	47.2%	41.1%	5.5%	3.1%
	B. gargarizans	43.7%	42.1%	11.9%	2.4%
	M. unicolor	59.8%	33.9%	4.7%	1.6%
	S. bombifrons	51.2%	38.6%	7.9%	2.4%
	B. bombina	45.3%	41.4%	8.6%	4.7%

(*Contd*.)

the vertebrate groups (S1 Table). In some species, alpha helix was also predicted with low confidence scores. The Ramachandran plot analysis showed that more than 85% residues of asprosin across the vertebrate groups lies in the most favoured region and additional allowed region (Table 2 and Suppl. Table 1).

Further, post-translational modifications such as phosphorylation and glycosylation sites were understood in the putative asprosin across the vertebrate groups. Two N-linked glycosylation sites were conserved across the vertebrate groups as shown in (Fig. 1). However, one extra glycosylation site

Table 2 — Ramachandran plot analysis of putative asprosin of all species across the vertebrate groups (*Contd*.)

Vertebrate groups	Species	Residues in most favoured region	Residues in additional allowed region	Residues in generously allowed region	Residues in disallowed region
Fishes	*S. japonicus*	43.8%	44.5%	7.0%	4.7%
	O. niloticus	48.4%	43.0%	7.0%	1.6%
	H. comes	45.0%	44.2%	6.2%	4.7%
	O. latipes	47.2%	43.3%	7.9%	1.6%
	T. rubripes	49.2%	45.1%	2.5%	3.3%
	P. olivaceous	53.1%	37.5%	7.0%	2.3%
	S. aurata	48.8%	39.5%	4.7%	7.0%
	D. rerio	53.1%	40.7%	5.3%	0.9%
	E. lucius	57.5%	36.2%	5.5%	0.8%
	A. radiata	59.2%	33.8%	4.6%	2.3%

was also found in mammalian group, avian group and one species of reptilian group *P. mucrosquamatus*. Several phosphorylation sites were also predicted in the asprosin protein. Out of several phosphorylation sites, one serine/threonine phosphorylation near C-terminus was well conserved in all the vertebrate groups except *P. mucrosquamatus*. Also, serine phosphorylation at first position of asprosin was conserved in mammals, birds, reptiles and amphibians but not in fishes (Fig. 1 and Table 3).

Using clustal omega, the percent identity matrix of putative asprosin was calculated across the vertebrate groups and compared with the predicted asprosin of *H. sapiens*. The predicted asprosin in other mammals, birds, reptiles and amphibian groups were more than 70% similar with the asprosin of *H. sapiens*. The putative asprosin in teleosts and cartilaginous fish were more than 50% similar to asprosin of *H. sapiens*. However, the predicted asprosin of *D. rerio* showed only 46.83% similarity to *H. sapiens* asprosin (Table 4).

Evolutionary analysis of putative asprosin

The analysis of site-specific selection in codons encoding asprosin revels that no site experienced diversifying selection, while 95 (66.43%) sites experienced purifying selection (FUBAR, 0.9 posterior probability). Further, no site showed episodic positive/diversifying selection (MEME, p = 0.05).In the phylogenetic tree of asprosin protein, the bony fishes formed a separate cluster distinct from other vertebrate groups. The mammals, amphibians, avians and reptiles were clustered together (Fig. 2).

The phylogenetic tree was constructed using the neighbor-joining method with 1000 bootstrap replicates.

Discussion

The present study was conducted to explore novel adipokine asprosin in non-mammalian vertebrates

Table 3 — The sites of post-translational modifications in the putative asprosin across the vertebrate groups

Vertebrate groups	Species	Glycosylation sites	Phosphorylation sites
Mammals	*H. sapiens*	3, 19, 36	1, 8, 14, 41, 66, 87
	M. musculus	3, 19, 36	1, 14, 66, 87
	R. norvegicus	3, 19, 36	1, 14, 66, 87
	B. taurus	3, 19, 36	1, 8, 41, 66, 87
	F. catus	3, 19, 36	8, 41, 66, 87
	M. lucifugus	3, 19, 36	1, 8, 41, 66, 87
	S. harrisii	3, 36	1, 41, 66, 87
	O. anatinus	3, 36, 58	24, 41, 87, 114
Birds	*G. gallus*	3, 35, 94	1, 5, 40, 65, 86
	C. livia	3, 35, 94	1, 5, 65, 86
	T. guttata	3, 35, 94	1, 5, 40, 65, 86
	P. humilis	3, 35, 94	1, 7, 40, 65, 86
	S. vulgaris	3, 35, 94	1, 5, 40 65, 86
	L. coronata	3, 35, 94	1, 5, 40, 65, 86
	F. peregrinus	3, 35, 94	1, 40, 65, 86
	H. leucocephalus	3, 35, 94	1, 5, 40, 65, 86
Reptiles	*C. mydas*	3, 36	1, 27, 66, 87
	P. sinensis	3, 36	1, 27, 66, 87
	G. gangeticus	3, 36	1, 5, 66, 87
	A. sinensis	3, 36	1, 5, 66, 87
	A. carolinensis	3, 36	1, 87
	P. muralis	3, 36	1, 87
	G. japonicus	3, 36	1, 87
	P. mucrosquamatus	3, 36, 68	89, 130
Amphibians	*X. tropicalis*	3, 34	1, 64, 85
	R. bivittatum	3, 36	1, 66, 87, 113
	R. temporaria	3, 36	1, 8, 27, 66, 87
	B. gargarizans	3, 35	1, 5, 86
	M. unicolor	3, 36	1, 66, 87, 112
	S. bombifrons	3, 36	1, 5, 66, 87
	B. bombina	3, 36	1, 66, 87
Fishes	*S. japonicus*	3, 33	4, 38, 52, 84
	O. niloticus	3	4, 16, 42, 56, 88
	H. comes	3, 37	5, 16, 42, 88, 114
	O. latipes	3, 36	4, 30, 41, 55, 87
	T. rubripes	3, 30	35, 49, 81, 107
	P. olivaceous	3, 36	4, 41, 55, 87
	S. aurata	3, 35	4, 9, 40, 54, 86
	D. rerio	3, 31	34, 71, 82
	E. lucius	3, 34	4, 39, 74, 85, 86, 131
	A. radiata	39	9, 44, 90, 91

Table 4 — Percentage similarity of putative asprosin of all the species compared with the asprosin of *Homo sapiens*

Vertebrate groups	Species	Percentage similarity of putative asprosin with asprosin of *Homo sapiens*
Mammals	*M. musculus*	92.14
	R. norvegicus	90.71
	B. taurus	95.00
	F. catus	92.86
	M. lucifugus	92.86
	S. harrisii	87.14
	O. anatinus	67.86
Birds	*G. gallus*	76.98
	C. livia	76.98
	T. guttata	78.42
	P. humilis	78.42
	S. vulgaris	77.70
	L. coronata	77.70
	F. peregrinus	77.70
	H. leucocephalus	77.70
Reptiles	*C. mydas*	83.57
	P. sinensis	81.43
	G. gangeticus	83.57
	A. sinensis	83.57
	A. carolinensis	75.00
	P. muralis	77.14
	G. japonicus	77.86
	P. mucrosquamatus	68.57
Amphibians	*X. tropicalis*	70.29
	R. bivittatum	74.29
	R. temporaria	69.29
	B. gargarizans	67.63
	M. unicolor	75.00
	S. bombifrons	75.00
	B. bombina	72.14
Fishes	*S. japonicus*	60.29
	O. niloticus	56.43
	H. comes	56.43
	O. latipes	58.27
	T. rubripes	56.72
	P. olivaceous	58.99
	S. aurata	58.70
	D. rerio	46.83
	E. lucius	53.28
	A. radiata	60.71

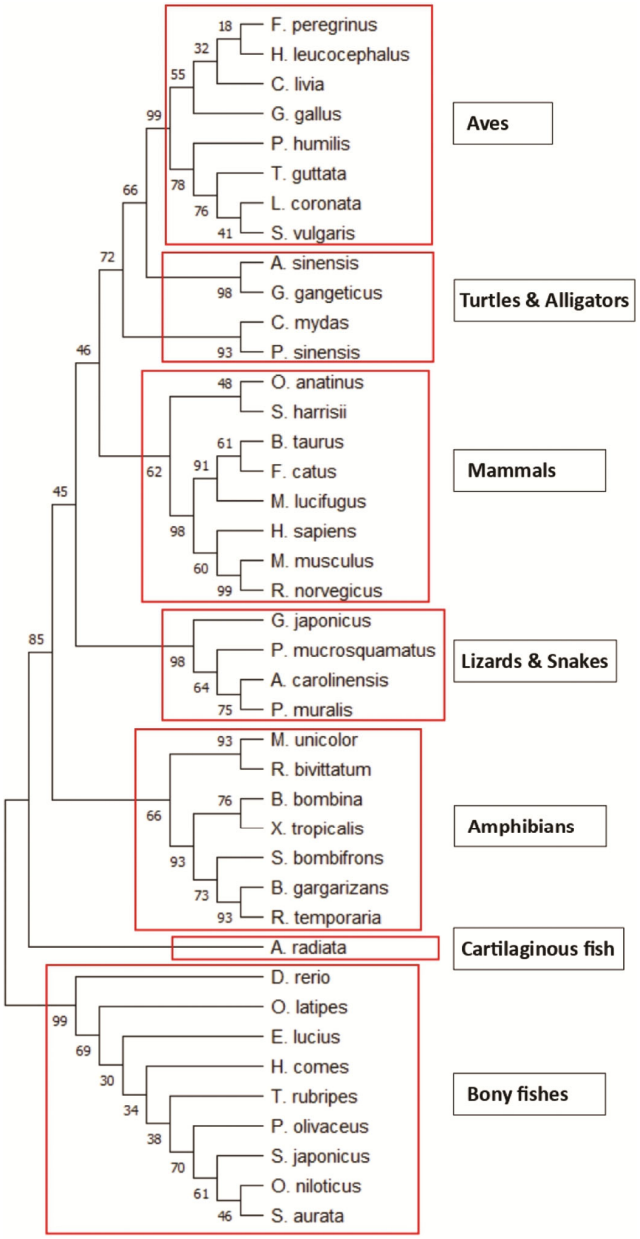

Fig. 2 — Phylogenetic tree showing the evolutionary relationship of putative asprosin across the vertebrate groups

using an *in silico* approach. Multiple sequence analysis of the putative profibrillin protein of all vertebrate groups revealed the presence of well conserved furin cleavage site across the vertebrate groups. In teleost *Channa punctata* also, K-X-X-X-K/R-R↓ has been reported as the furin cleavage site in profibrillin protein[10]. After cleavage of profibrillin protein by furin protease, the cleaved C-terminal region represents the putative asprosin protein. In case of mammalian and reptilian vertebrates, the predicted asprosin comprised 140 amino acids similar to earlier reports by Romere *et al*. in 2016[8]. However, the number of amino acids in predicted asprosin vary in birds, amphibians and bony fishes, ranging from 138 to 142 amino acids. The majority of insertions of amino acids in putative asprosin are found near the N-terminal region of asprosin, as observed in multiple sequence alignment. However, the well-established experimental model *D. rerio* has only 128 residues in putative asprosin protein. The predicted molecular weight of asprosin in all vertebrates is approximately 16 kDa, which is similar to that reported in *H. sapiens* and *C. punctata*[8,10]. The length of the protein, tertiary

structure and sequence of amino acid decides the structural domains and hence the function of proteins. Although under different evolutionary pressure, species tends to evolve differently, however, the length of asprosin protein remains almost consistent across the vertebrates. This might suggest strong universal selective pressure.

The theoretical pI represents the pH value at which proteins carry no net charge, or the sum of positive and negative charges is equal. All the putative asprosin across the vertebrates have pI below 7, suggesting strong acidic nature of the protein across the vertebrate. However, *O. anatinus* has a pI value 7.19, representing its neutral nature. In accordance, the number of negatively charged amino acids (acidic in nature) are higher as compared to the positively charged residues (basic in nature) in putative asprosin of vertebrates except *O. anatinus*, where the number of negatively charged amino acids are equal to the positively charged residues. Further, the instability index of the predicted asprosin in the majority of the vertebrate species studied is either 40 or below, indicating their stability under ex vivo conditions[32]. The asprosin of *H. sapiens* and *C. punctata* are also reported to be stable under ex vivo conditions as their instability index are below 40[10]. All the putative asprosin across vertebrate groups similar to *H. sapiens* and *C. punctata* are thermostable as their aliphatic index values are high. Further, the negative grand average of hydropathicity index (GRAVY) of the predicted asprosin across vertebrate species indicates the hydrophilic nature of protein. Thus, the similarity in the physicochemical properties of asprosin such as acidic nature, thermostability and hydrophilic nature might be because of unique evolutionary constraint on asprosin across the vertebrates. The widespread presence of *fbn1* gene, furin cleavage site in its encoded profibrillin protein and nature of asprosin indicates that the asprosin protein and its gene were subjected to strong evolutionary pressure and asprosin might have conserved functional roles. Similar conserved physicochemical properties of another adipokine, leptin, have been reported in fish and mammals[39].

This *in silico* study for the first time provides a comprehensive account of the secondary and tertiary structure analysis of putative asprosin across vertebrate species. In the absence of NMR/crystal structure of asprosin, this study provides important insights into the structural conformation and stability parameters of the structure. The secondary and tertiary structure of all putative asprosin comprises beta strands, which might be involved in the stabilizing and strengthening of the protein backbone as suggested by Kellis *et al*. in 1988[40]. Similar beta strands are also reported in *H. sapiens* and *C. punctata*[10,41]. Liu *et al*. also predicted the presence of alpha helix in the secondary structure of asprosin in *H. sapiens*[41]. In the current study, although alpha helix is predicted in the secondary structure of asprosin in the species studied, cautious interpretation of the result is needed, as putative asprosin in most of the species has less than 3.6 residues involved in alpha helix formation. Moreover, those species which had more than 3.6 residues in the alpha helix had low confidence score. According to Boyle, 2018, at least 3-4 residues are required for stable alpha helix formation[42]. The Ramachandran plot analysis of the tertiary structure of putative asprosin in different vertebrate species showed that more than 85% residues lie in the most favoured region and additional allowed region. The residues in the disallowed region range from 0.8 to 7%. Ideally, for a protein to be considered stable under physiological conditions, the percentage of residues in the most favoured region should be more than 85%[43]. However, the lesser residues in the most favoured region of predicted asprosin might be due to the absence of reference tertiary NMR/crystal structure of asprosin.

The post-translational modifications, such as glycosylation in asprosin has been reported by Romere *et al*.[8] and Sathoria *et al*.[10] in *H. sapiens* and *C. punctata*, respectively. The asprosin of *H. sapiens* have three glycosylation sites whereas in *C. punctata*, two glycosylation sites have been reported[8,10]. Glycosylation in a protein increases its half-life and enhances stability[44,45]. In addition, glycosylation helps in interaction with their cognate receptor by increasing the affinity and affects the signalling pathway of glycoproteins[46-51]. In all vertebrates studied except *O. niloticus*, we found two conserved N-linked glycosylation sites in the predicted asprosin. In addition, the mammalian group except *S. harrisii*, birds, and *P. mucrosquamatus* belonging to the reptilian group exhibit an extra glycosylation site at variable positions. The presence of glycosylation has been conjectured to be the reason that the molecular weight of the asprosin (~30 kDa) is much higher than the bacterially expressed recombinant asprosin (~16 kDa)[8,10]. Interestingly, conserved

phosphorylation sites are also present across vertebrates. The phosphorylation at the first position residue, serine, is observed in mammals, birds, reptiles and amphibians. In addition to this, one serine/threonine phosphorylation near the C-terminus is found conserved in all the vertebrates except *P. mucrosquamatus*. Although the role of phosphorylation is well illustrated in cell signalling, its functional implication in asprosin is yet to be deciphered.

Further, we have compared the percentage identity matrix of putative asprosin sequence of all vertebrates with the asprosin of *H. sapiens*. The asprosin of mammalian species showed over 87% similarity with the asprosin of *H. sapiens*, indicating the close evolutionary relationship except monotremata with 67% similarity which suggest some divergence. The predicted asprosin in birds showed approximately 77% similarity with *H. sapiens* asprosin, reflecting a more distant common ancestor with mammals. Turtle and alligators asprosin showed high similarity with *H. sapiens* asprosin, indicating greater conservation compared to lizards and snakes with low similarity, reflecting high evolutionary divergence. Amphibians resulted into moderate level of similarity (70%) and conservation with *H. sapiens* asprosin, indicating its intermediate position between reptiles and fishes. The asprosin of fishes have 50-60% similarity with *H. sapiens* asprosin, showing a more distant evolutionary relationship between fishes and humans. We had earlier reported 57% similarity between predicted asprosin of the teleostean species, *C. punctata,* and *H. sapiens*[10].

Until now, no studies have addressed the codon-based selection analysis on asprosin coding region of *fbn1* gene. The pervasive negative/purifying selection on asprosin was detected in the current study. The negative (purifying) selection is prevalent evolutionary force responsible for the conservation of genomic sequence over the extensive evolutionary period[52]. Since, purifying selection is a very strong evolutionary pressure, this might be the reason for the conserved physicochemical properties of asprosin protein. The purifying selection results into elimination of mutation in asprosin sequence and conservation in genome sequence.

To understand the evolutionary history of asprosin across the vertebrate groups, the phylogenetic tree was constructed, In the phylogenetic tree, predicted asprosin of the bony fishes form a separate cluster and are closely related with the amphibians, having a common ancestor. In the teleostean species, asprosin in *D. rerio* belonging to the Cypriniformes class seem to have diverged from other teleostean species at an early stage of evolution. The same divergence of Cyriniformes class with other teleostean species have been reported in another adipokine, the leptin protein[39]. Within the amphibian group, the predicted asprosin form two separate clusters, wherein the limbless caecilians species from Gymnophiona class (*R. bivittatum* and *M. unicolor*) diverge from the tailless amphibians of Anura class (*X. tropicalis*, *B. bombina*, *S. bombifrons*, *R. temporaria* and *B. gargarizans*). The putative asprosin in mammals, aves and reptiles diverge from a common ancestor. In the reptilian group, predicted asprosin of Squamata (*A. carolinensis*, *P. muralis*, *G. japonicus*, and *P. mucrosquamatus*) is within a clade with the Gymnophiona (vermiform or serpentine amphibians), which were originally thought to be related to the snakes. However, asprosin of the semi-aquatic reptiles, Crocodylia (*G. gangeticus* and *A. sinensis*) are sister to the aves clade. The sampled species from the major lineages (amphibian, reptiles, mammals and bony fishes) conform to monophyletic groups, which is similar to reports in leptin[53]. Among mammals, putative asprosin of placental mammals, Eutheria diverge from the egg-laying mammals, Monotremata (*O. anatinus*), and Metatheria (*S. harrisii*) that are characterized by presence of pouch and giving birth to partially developed young ones. Leptin in Monotremata has been also reported to diverge from Eutherian group[53].

The bootstrap proportions provide insights into the reliability of the inferred clades in the phylogeny under various conditions. The bootstrap support greater the 70% indicates high probability (>95%) that clade is accurately presented and strong support for the clade under conditions of low internodal change, symmetric phylogenies and equal rate of change[54]. In the current study, the phylogenetic tree has nodes with low bootstrap supports, which makes it difficult to comment about the conformation of the phylogeny of asprosin with the species tree. Phylogenetic tree in the current study has been constructed using putative as well as sequenced amino acids of asprosin. Although the bootstrap values of many clades were more than 70, others have quite low bootstrap value. It might be possible that the support values vary in case the sequences are characterised. In addition, it can be speculated that difference in the

amino acid length of asprosin protein ranging from 128 in *D. rerio* to 143 in *A. radiata* impacted the number of sites (amino acid position) during alignment and thereafter, has affected the robustness of the phylogenetic tree. Also, the large taxon size and gap treatment during alignment could have led to decrease in bootstrap value. Interestingly, passerine bird clade with moderate bootstrap support (between 70-80%) deviates from the current understanding of passerine evolution[55]. This might be because of the data limitation and taxon sampling for the construction of phylogenetic tree.

Nonetheless, the current study invokes intriguing questions about the significance behind the evolutionary conservation of asprosin. The absence of *fbn1* gene sequence data of the most primitive extant lampreys and cartilaginous fishes limits our ability to understand the evolutionary conservation of asprosin in these groups. Further, the lack of NMR/crystal structure of asprosin impedes our understanding of its tertiary structure and function. Thus, more experimental studies are required to address the functional aspects of asprosin in non-mammalian vertebrates.

Conclusion

The conserved *fbn1* gene and furin cleavage site of profibrillin protein across vertebrates implies that asprosin might be present in all vertebrates. Additionally, the structure of asprosin also seems to be conserved. Further, the current *in silico* study infers that the molecular weight of putative asprosin across vertebrates is consistent, suggesting unique evolutionary constraints acting on asprosin. We also demonstrated the thermostability and hydrophilic nature of putative asprosin across the vertebrate groups and predicted the post-translational modifications. Through critical assessment, we have shown the evolutionary relationship of asprosin in vertebrates. We have also demonstrated that asprosin is subjected to purifying selection, thus conserving the structural and functional aspects. Based on this comprehensive comparative analysis, it can be tentatively suggested that asprosin in non-mammalian vertebrates might be involved in diverse crucial biological functions similar to that observed in mammals. In future, detailed studies across vertebrates could reveal the physiological significance of this important hormone and add a new paradigm to our understanding.

Conflict of interest

All authors declare no conflict of interest.

References

1 Maury E & Brichard SM, Adipokine dysregulation, adipose tissue inflammation and metabolic syndrome. *Mol Cell Endocrinol*, 314 (2010) 1.

2 Ouchi N, Parker JL, Lugus JJ & Walsh K, Adipokines in inflammation and metabolic disease. *Nat Rev Immunol*, 11 (2011) 85.

3 Lin F, Zhou C, Chen H, Wu H, Xin Z, Liu J, Gao Y, Yuan D, Wang T, Wei R & Chen D, Molecular characterization, tissue distribution and feeding related changes of NUCB2A/nesfatin-1 in Ya-fish (*Schizothorax prenanti*). *Gene*, 536 (2014) 238.

4 Sundarrajan L, Blanco AM, Bertucci JI, Ramesh N, Canosa LF & Unniappan S, Nesfatin-1-like peptide encoded in nucleobindin-1 in goldfish is a novel anorexigen modulated by sex steroids, macronutrients and daily rhythm. *Sci Rep*, 6 (2016) 28377.

5 Yang G, Song Q, Sun C, Qin J, Jia J, Yuan X, Zhang Y & Li W, Ctrp9 and adiponectin receptors in Nile tilapia (*Oreochromis niloticus*): molecular cloning, tissue distribution and effects on reproductive genes. *Gen Comp Endocrinol*, 265 (2018) 160.

6 Pi D, Wang J, Zhao M, Liu M, Zhang Y, Qin C, Yang L, Yan X & Nie G, Adiponectin and adiponectin receptors in common carp (*Cyprinus carpio*): Tissue distribution and their expressions in response to high-carbohydrate and high-lipid diets. *Aquacult Rep*, 27 (2022) 101341.

7 Bakshi A, Singh R & Rai U, Trajectory of leptin and leptin receptor in vertebrates: Structure, function and their regulation. *Comp Biochem Physiol B Biochem Mol Biol*, 257 (2022) 110652.

8 Romere C, Duerrschmid C, Bournat J, Constable P, Jain M, Xia F, Saha PK, Del Solar M, Zhu B, York B & Sarkar P, Asprosin, a fasting-induced glucogenic protein hormone. *Cell*, 165 (2016) 566.

9 Farrag M, AitEldjoudi D, González-Rodríguez M, Cordero-Barreal A, Ruiz-Fernández C, Capuozzo M, Gonzalez-Gay MA, Mera A, Lago F, Soffar A & Essawy A, Asprosin in health and disease, a new glucose sensor with central and peripheral metabolic effects. *Front Endocrinol*, 13 (2023) 1101091.

10 Sathoria P, Chuphal B, Rai U & Roy B, Molecular cloning, characterization and 3D modelling of spotted snakehead fbn1 C-terminal region encoding asprosin and expression analysis of fbn1. *Sci Rep*, 13 (2023) 4470.

11 Mishra I, Xie WR, Bournat JC, He Y, Wang C, Silva ES, Liu H, Ku Z, Chen Y, Erokwu BO & Jia P, Protein tyrosine phosphatase receptor δ serves as the orexigenic asprosin receptor. *Cell Metab*, 34 (2022) 549.

12 Lee T, Yun S, Jeong JH & Jung TW, Asprosin impairs insulin secretion in response to glucose and viability through TLR4/JNK-mediated inflammation. *Mol Cell Endocrinol*, 486 (2019) 96.

13 Shabir K, Gharanei S, Orton S, Patel V, Chauhan P, Karteris E, Randeva HS, Brown JE & Kyrou I, Asprosin exerts pro-inflammatory effects in THP-1 macrophages mediated *via* the toll-like receptor 4 (TLR4) pathway. *Int J Mol Sci*, 24 (2022) 227.

14 Zhang Z, Tan Y, Zhu L, Zhang B, Feng P, Gao E, Xu C, Wang X, Yi W & Sun Y, Asprosin improves the survival of

mesenchymal stromal cells in myocardial infarction by inhibiting apoptosis *via* the activated ERK1/2-SOD2 pathway. *Life Sci*, 231 (2019) 116554.

15 Keskin T, Erden Y & Tekin S, Intracerebroventricular asprosin administration strongly stimulates hypothalamic-pituitary-testicular axis in rats. *Mol Cell Endocrinol*, 538 (2021) 111451.

16 Maurya S, Krishna A, Lal B & Singh A, Asprosin promotes steroidogenesis and spermatogenesis with improved glucose metabolism in adult mice testis. *Andrologia*, 54 (2022) e14579.

17 Maylem ER, Spicer LJ, Batalha I & Schutz LF, Discovery of a possible role of asprosin in ovarian follicular function. *J Mol Endocrinol*, 66 (2021) 35.

18 Maylem ER, Spicer LJ, Atabay EP, Atabay EC, Batalha I & Schutz LF, A potential role of fibrillin-1 (FBN1) mRNA and asprosin in follicular development in water buffalo. *Theriogenology*, 178 (2022) 67.

19 Batalha IM, Maylem ER, Spicer LJ, Bello CA, Archilia EC & Schütz LF, Effects of asprosin on estradiol and progesterone secretion and proliferation of bovine granulosa cells. *Mol Cell Endocrinol*, 565 (2023) 111890.

20 Long W, Xie X, Du C, Zhao Y, Zhang C, Zhan D, Li Z, Ning Q & Luo X, Decreased circulating levels of asprosin in obese children. *Horm Respaediatr*, 91 (2019) 271.

21 Wang M, Yin C, Wang L, Liu Y, Li H, Li M, Yi X & Xiao Y, Serum asprosin concentrations are increased and associated with insulin resistance in children with obesity. *Ann Nutr Metab*, 75 (2020) 205.

22 Zhang X, Jiang H, Ma X & Wu H, Increased serum level and impaired response to glucose fluctuation of asprosin is associated with type 2 diabetes mellitus. *J Diabetes Investig*, 11 (2020) 349.

23 Alan M, Gurlek B, Yilmaz A, Aksit M, Aslanipour B, Gulhan I, Mehmet C & Taner CE, Asprosin: a novel peptide hormone related to insulin resistance in women with polycystic ovary syndrome. *Gynecol Endocrinol*, 35 (2019) 220.

24 Jiang Y, Liu Y, Yu Z, Yang P & Zhao S, Serum asprosin level in different subtypes of polycystic ovary syndrome: a cross-sectional study. *Rev Assoc Méd Bras*, 67 (2021) 590.

25 Ozcan S, Kelestemur MM, Hekim MG, Bulmus O, Bulut F, Bilgin B, Canpolat S & Ozcan M, Asprosin, a novel therapeutic candidate for painful neuropathy: an experimental study in mice. *Naunyn-Schmiedeb Arch Pharmacol*, 395 (2022) 325.

26 Kerslake R, Sisu C, Panfilov S, Hall M, Khan N, Jeyaneethi J, Randeva H, Kyrou I & Karteris E, Differential regulation of genes by the glucogenic hormone asprosin in ovarian cancer. *J Clin Med*, 11 (2022) 5942.

27 Li E, Shan H, Chen L, Long A, Zhang Y, Liu Y, Jia L, Wei F, Han J, Li T & Liu X, OLFR734 mediates glucose metabolism as a receptor of asprosin. *Cell metab*, 30 (2019) 319.

28 Molloy SS, Anderson ED, Jean F, Thomas G, Molloy SS, Anderson ED, Jean F & Thomas G, Bi-cycling the furin pathway: from TGN localization to pathogen activation and embryogenesis. *Trends cell biol*, 9 (1999) 28.

29 Gasteiger E, Protein identification and analysis tools on the ExPASy server. The proteomics protocols handbook/*Human Press Inc.* 2005.

30 Zhang Y, I-TASSER server for protein 3D structure prediction. *BMC Bioinform*, 9 (2008) 1.

31 Laskowski RA, MacArthur MW, Moss DS & Thornton JM, PROCHECK: a program to check the stereochemical quality of protein structures. *J Appl Crystallogr*, 26 (1993) 283.

32 Guruprasad K, Reddy BB & Pandit MW, Correlation between stability of a protein and its dipeptide composition: a novel approach for predicting in vivo stability of a protein from its primary sequence. *Protein Eng Des Sel*, 4 (1990) 155.

33 Ikai A, Thermostability and aliphatic index of globular proteins. *J Biochem*, 88 (1980) 1895.

34 Kyte J & Doolittle RF, A simple method for displaying the hydropathic character of a protein. *J Mol Biol*, 157 (1982) 105.

35 Pond SL & Frost SD, Datamonkey: rapid detection of selective pressure on individual sites of codon alignments. *Bioinfo*, 21(2005) 2531.

36 Weaver S, Shank SD, Spielman SJ, Li M, Muse SV, Kosakovsky Pond SL. Datamonkey 2.0: a modern web application for characterizing selective and other evolutionary processes. *Mol Biol Evol*, 35 (2018) 773.

37 Murrell B, Moola S, Mabona A, Weighill T, Sheward D, Kosakovsky Pond SL, Scheffler K. FUBAR: a fast, unconstrained bayesian approximation for inferring selection. *Mol Biol Evol*, 30 (2013) 1196.

38 Murrell B, Wertheim JO, Moola S, Weighill T, Scheffler K & Kosakovsky Pond SL, Detecting individual sites subject to episodic diversifying selection. *PLoS Genet*, 8 (2012) e1002764.

39 Bakshi A & Rai U, *In silico* analyses of leptin and leptin receptor of spotted snakehead *Channa punctata*. *PLoS One*, 17 (2022) e0270881.

40 Kellis Jr JT, Nyberg K, Sali D & Fersht AR, Contribution of hydrophobic interactions to protein stability. *Nature*, 333 (1988) 784.

41 Liu L, Liu Y, Huang M, Zhang M, Zhu C, Chen X, Bennett S, Xu J & Zou J, The effects of asprosin on exercise-intervention in metabolic diseases. *Front Physiol*, 13 (2022) 907358.

42 Boyle AL, Applications of de novo designed peptides. In Peptide applications in biomedicine, Biotechnology and Bioengineering, Woodhead Publishing, (2018) 51.

43 Rao S, Sayeeda M, Prakash T, Padamshree AP, Imran S & Ravi L, Construction of Computational 3d Structures of Protein Drug Targets of *Mycobacterium Tuberculosis*. *Asian J Pharm Clin Res*, 7 (2020) 82.

44 Perlman S, van den Hazel B, Christiansen J, Gram-Nielsen S, Jeppesen CB, Andersen KV, Halkier T, Okkels S & Schambye HT, Glycosylation of an N-terminal extension prolongs the half-life and increases the in vivo activity of follicle stimulating hormone. *J Clin Endocrinol Metab*, 88 (2003) 3227.

45 Sinclair AM & Elliott S, Glycoengineering: the effect of glycosylation on the properties of therapeutic proteins. *J Pharma Sci,* 94 (2005) 1626.

46 Smith PL, Kaetzel D, Nilson J & Baenziger JU, Thesialylated oligosaccharides of recombinant bovine lutropin modulate hormone bioactivity. *J Biol Chem,* 265 (1990) 874.

47 Bishop LA, Robertson DM, Cahir N & Schofield PR, Specific roles for the asparagine-linked carbohydrate residues of recombinant human follicle stimulating hormone in receptor binding and signal transduction. *Mol Endocrinol*, 8 (1994) 722.

48 Grossmann M, Szkudlinski MW, Tropea JE, Bishop LA, Thotakura NR, Schofield PR & Weintraub BD, Expression of human thyrotropin in cell lines with different glycosylation patterns combined with mutagenesis of specific glycosylation sites: Characterization of a novel role for the oligosaccharides in the *in vitro* and *in vivo* bioactivity. *J Biol Chem*, 270 (1995) 29378.

49 Ulloa-Aguirre A, Midgley Jr AR, Beitins IZ & Padmanabhan V, Follicle-stimulating isohormones: characterization and physiological relevance. *Endo Rev*, 16 (1995) 765.

50 Arey BJ, Stevis PE, Deecher DC, Shen ES, Frail DE, Negro-Vilar A & Lopez FJ, Induction of promiscuous G protein coupling of the follicle-stimulating hormone (FSH) receptor: a novel mechanism for transducing pleiotropic actions of FSH isoforms. *Mol Endocrinol*, 11 (1997) 517.

51 Bousfield GR, Butnev VY, Butnev VY, Nguyen VT, Gray CM, Dias JA, MacColl R, Eisele L & Harvey DJ, Differential effects of α subunit asparagine56 oligosaccharide structure on equine lutropin and follitropin hybrid conformation and receptor-binding activity. *Biochemistry*, 43 (2004) 10817.

52 Liu J, Zhang Y, Lei X & Zhang Z, Natural selection of protein structural and functional properties: a single nucleotide polymorphism perspective. *Genome Biol,* 9 (2008) R69.

53 Denver RJ, Bonett RM & Boorse GC, Evolution of leptin structure and function. *Neuroendocrinol*, 94(2011) 21.

54 Hillis DM & Bull JJ, An empirical test of bootstrapping as a method for assessing confidence in phylogenetic analysis. *Syst Biol*, 42 (1993) 182.

55 Prum RO, Berv JS, Dornburg A, Field DJ, Townsend JP, Lemmon EM & Lemmon AR, A comprehensive phylogeny of birds (Aves) using targeted next-generation DNA sequencing. *Nature*, 526 (2015) 569.