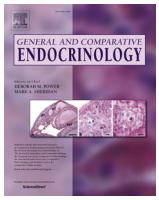
General and Comparative Endocrinology 320 (2022) 114011

Contents lists available at ScienceDirect

General and Comparative Endocrinology

journal homepage: www.elsevier.com/locate/ygcen



Research paper

A comparative phylogenetic analysis of prolactin cleavage sites for the generation of vasoinhibin in vertebrates

Andreas Leuchs [a,1], Nils Davies [a,1], Christin Friedrich [a], Sabrina Trier [a], Carmen Clapp [b], Thomas Bertsch [a], Jakob Triebel [a,*]

[a] Institute for Clinical Chemistry, Laboratory Medicine and Transfusion Medicine, Nuremberg General Hospital & Paracelsus Medical University, Nuremberg, Germany
[b] Instituto de Neurobiología, Universidad Nacional Autónoma de México (UNAM), Campus UNAM-Juriquilla, Querétaro, Mexico

ARTICLE INFO

Keywords:
Prolactin
Vasoinhibin
Cleavage sites
Cathepsin D

ABSTRACT

Vasoinhibin is a pleiotropic protein hormone with endocrine, autocrine, and paracrine effects on blood vessel growth, permeability, and dilation, and a role in several human diseases. It is generated by proteolytic cleavage of the pituitary hormone prolactin by cathepsin D. Several isoforms with a variation in the number of amino acids and corresponding molecular mass exist. This in silico study investigated the cathepsin D cleavage sites in prolactin responsible for the generation of vasoinhibin in vertebrate species. Ninety-one prolactin protein sequences from species of the taxa primates, rodents, laurasiatheria, mammals, sauropsida, and fish were retrieved, and a multiple sequence alignment was performed. Each sequence was investigated for the presence of a vasoinhibin-generating cathepsin D cleavage site and its corresponding substrate affinity using a scoring system. Primates demonstrated the highest substrate affinity for the generation of the 15 kDa vasoinhibin isoform, and fish the highest affinity for the 16.8 kDa isoform. In both cases, this associates to the presence of leucine in the cleavage site, which is not present in species of the other taxa. In primate evolution, the presence of leucine in the cleavage site occurs with the emergence of simiiformes 42 million years ago and is conserved in higher primates across all subsequent speciation nodes. The 17.2 kDa vasoinhibin isoform has a constant substrate affinity in all taxa. The presence of leucine in vasoinhibin generating cleavage sites appears as an important feature of the molecular evolution of vasoinhibin.

1. Introduction

The pituitary hormone prolactin (PRL), best known for its role in mammary growth and lactation (UniProt ID P01236), evolved ~400 million years ago in fish and regulates multiple functions in the various vertebrate groups, for example osmoregulation, reproduction, metabolism, behavior, and vascular processes (Clapp et al., 2012; Hiroshi and Sower, 2006; Rand-Weaver and Kawauchi, 1993). The regulation of blood vessel growth, permeability and dilation is essential for all vertebrate species and is, in part, under control of PRL and vasoinhibin (Clapp et al., 2012). Vasoinhibin is a protein hormone with a diverse array of endocrine, paracrine, and autocrine effects, ranging from the regulation of blood vessel growth, permeability and dilatation (Clapp et al., 2015) to non-vascular effects, which include the stimulation of vasopressin release (Mejía et al., 2003), thrombolytic effects (Bajou et al., 2014), and the stimulation of anxiety- and depression-related behavior (Zamorano et al., 2014). Vasoinhibin has been investigated in the context of several human diseases, and pathophysiological roles emerged in vasoproliferative retinopathies (Arnold et al., 2010; Triebel et al., 2009) and pregnancy-associated diseases (González et al., 2007; Haghikia et al., 2015; Hilfiker-Kleiner et al., 2007). Vasoinhibin is generated by the proteolytic cleavage of its precursor molecule PRL. The proteolysis of PRL generating vasoinhibin takes place in various anatomical compartments and tissues, such as the pituitary gland (Cruz-Soto et al., 2009), the retina (Arnold et al., 2010), the placenta (Perimenis et al., 2014), the heart (Hilfiker-Kleiner et al., 2007), and the cartilage (Macotela et al., 2006). PRL cleaving enzymes, such as cathepsin D and matrix metalloproteinases, utilize multiple cleavage sites, and the resulting molecular mass of the vasoinhibin isoforms varies from 5 to 18 kDa (Ge et al., 2007; Macotela et al., 2006; Zamora et al., 2021). The functional motif responsible for the vascular effects of vasoinhibin is the His-Gly-Arg motif corresponding to amino acids

* Corresponding author.
E-mail address: Jakob.Triebel@gmx.de (J. Triebel).
[1] Shared first authorship and equal contributions.

https://doi.org/10.1016/j.ygcen.2022.114011
Received 19 October 2021; Received in revised form 16 February 2022; Accepted 24 February 2022
Available online 26 February 2022

Table 1
Representative cleavage sites in full-length PRL for the generation of the 15 kDa vasoinhibin isoform by cathepsin D.

Taxon/Species Molecular Mass	Amino-acid Position/Amino Acid/Score								8P score
	P4	P3	P2	P1	P1'	P2'	P3'	P4'	
Primates									
Human 15 kDa	157	158	159	160	161	162	163	164	
	Gly	Met	Glu	Leu	Ile	Val	Ser	Gln	
	42	26	94	416	119	131	55	52	935
Gorilla 15 kDa	158	159	160	161	162	163	164	165	
	Gly	Met	Glu	Leu	Ile	Val	Ser	Gln	
	42	26	94	416	119	131	55	52	935
Rodents									
Rat 14.69 kDa	156	157	158	159	160	161	162	163	
	Gly	Ile	Glu	Lys	Ile	Ile	Ser	Gln	
	42	76	94	6	119	50	55	52	494
Rabbit 15.17 kDa	158	159	160	161	162	163	164	165	
	Gly	Met	Glu	Lys	Ile	Val	Gly	Gln	
	42	26	94	6	119	131	64	52	534
Laurasiatheria									
Pig 15.14 kDa	159	160	161	162	163	164	165	166	
	Gly	Met	Glu	Lys	Ile	Val	Gly	Gln	
	42	26	94	6	119	131	64	52	534
Cow 14.93 kDa	159	160	161	162	163	164	165	166	
	Gly	Met	Glu	Met	Ile	Phe	Gly	Gln	
	42	26	94	35	119	16	64	52	448
Mammals									
Armadillo 15.12 kDa	159	160	161	162	163	164	165	166	
	Gly	Met	Lys	Lys	Ile	Ile	Gly	Gln	
	42	26	31	6	119	50	64	52	390
Elephant 14.85 kDa	158	159	160	161	162	163	164	165	
	Gly	Ile	Glu	Lys	Ile	Val	Asp	Gln	
	42	76	94	6	119	131	57	52	577
Sauropsida									
Duck 15.48 kDa	163	164	165	166	167	168	169	170	
	Gly	Met	Glu	Lys	Ile	Val	Gly	Arg	
	42	26	94	6	119	131	64	2	484
American Alligator 15.87 kDa	166	167	168	169	170	171	172	173	
	Gly	Met	Glu	Lys	Ile	Val	Gly	Arg	
	42	26	94	6	119	131	64	2	484
Fish									
Zebrafish 12.61 kDa	141	142	143	144	145	146	147	148	
	Gly	Leu	Glu	His	Val	Val	His	Lys	
	42	99	94	0	98	131	3	81	548
Amazon molly 14.43 kDa	168	169	170	171	172	173	174	175	
	Gly	Leu	Asp	Ile	Leu	Ser	Gly	Lys	
	42	99	63	1	110	62	64	81	522

Multiple sequence alignment of PRL sequences across a several taxa at the cleavage site at which the 15 kDa vasoinhibin isoform is being generated by cathepsin D with their associated 8P scores. Blue coloring indicates amino acid leucine in P1, grey coloring indicates amino acids other than leucine in P1. Cleavage occurs between P1 and P1′, indicated by the red line. Primates demonstrate a significantly higher 8P score compared to the other species ($p = <0.000001$).

74–76 (or 46–48 without the signal peptide, respectively), that is conserved in mammals, birds, reptiles, amphibians, and fish (Robles et al., 2022; Robles et al., 2018).

PRL shows an episodic pattern of molecular evolution in which long periods of slow evolution contrast with sustained bursts of rapid changes, with one of such bursts occurring during the evolution of primate PRL that leads to marked sequence differences between PRL in humans and non-primate mammals (Wallis, 2000). Two explanations for the episodes of accelerated evolution of PRL in primates propose: 1) an increased acceptance of neutral mutations associated with the relaxation from purifying selection following loss of function; and 2), positive selection associated with an adaptive change in function (Wallis et al., 2005). Since there is no evidence suggesting loss of function, positive selection associated with gain-of-function is more likely and it has been

Table 2
Representative cleavage sites in full-length PRL for the generation of the 15 kDa vasoinhibin isoform of primates by cathepsin D.

Taxon/Species Molecular Mass	Amino-acid Position/Amino Acid/Score								8P score
	P4	P3	P2	P1	P1'	P2'	P3'	P4'	
Primates									
Haplorrhini									
Tarsier 15.19 kDa	158	159	160	161	162	163	164	165	
	Gly	Met	Glu	Lys	Ile	Val	Gly	Gln	
	42	26	94	6	119	131	64	52	534
Simiiformes									
Marmoset 14.95 kDa	158	159	160	161	162	163	164	165	
	Gly	Met	Glu	Leu	Ile	Leu	Ser	Gln	
	42	26	94	416	119	80	55	52	884
Cercopithecinae									
Macaque 15.06 kDa	158	159	160	161	162	163	164	165	
	Gly	Met	Glu	Leu	Ile	Val	Ser	Gln	
	42	26	94	416	119	131	55	52	935
Olive baboon 15.03 kDa	158	159	160	161	162	163	164	165	
	Gly	Met	Glu	Leu	Ile	Val	Ser	Gln	
	42	26	94	416	119	131	55	52	935
Vervet-AGM 15 kDa	157	158	159	160	161	162	163	164	
	Gly	Met	Glu	Leu	Ile	Val	Ser	Gln	
	42	26	94	416	119	131	55	52	935
Hominoidea									
Gibbon 14.99 kDa	158	159	160	161	162	163	164	165	
	Gly	Met	Glu	Leu	Ile	Val	Ser	Gln	
	42	26	94	416	119	131	55	52	935
Great apes									
Orangutan 14.99 kDa	157	158	159	160	161	162	163	164	
	Gly	Met	Glu	Leu	Ile	Val	Ser	Gln	
	42	26	94	416	119	131	55	52	935
Hominines									
Gorilla 14.92 kDa	158	159	160	161	162	163	164	165	
	Gly	Met	Glu	Leu	Ile	Val	Ser	Gln	
	42	26	94	416	119	131	55	52	935
Chimpanzee 14.89 kDa	157	158	159	160	161	162	163	164	
	Gly	Met	Glu	Leu	Ile	Val	Ser	Gln	
	42	26	94	416	119	131	55	52	935
Human 15 kDa	157	158	159	160	161	162	163	164	
	Gly	Met	Glu	Leu	Ile	Val	Ser	Gln	
	42	26	94	416	119	131	55	52	935
Strepsirrhini									
Mouse Lemur 15.10 kDa	161	162	163	164	165	166	167	168	
	Gly	Met	Glu	Lys	Ile	Val	Gly	Gln	
	42	26	94	6	119	131	64	52	534
Bushbaby 15.17 kDa	161	162	163	164	165	166	167	168	
	Gly	Met	Glu	Lys	Ile	Val	Gly	Gln	
	42	26	94	6	119	131	64	52	534

Multiple sequence alignment of PRL sequences in primates at the cleavage site at which the 15 kDa vasoinhibin isoform is being generated by cathepsin D with their associated 8P scores. Blue coloring indicates amino acid leucine in P1, grey coloring indicates amino acids other than leucine in P1. Cleavage occurs between P1 and P1′, indicated by the red line. Simians, or higher primates, demonstrate significantly higher 8P scores compared to haplorrhini and strepsirrhini ($p = <0.000001$).

hypothesized that amino acid substitutions in PRL cleavage sites generating vasoinhibin constitute such gain-of-function event (Triebel et al., 2015).

The variation of the PRL sequence among vertebrate species likely translates into a different levels and total composition of vasoinhibin, since amino acid changes may affect cleavage sites, making them either more or less susceptible to specific proteases. Small scale sequence alignments provided initial confirmation of this assumption and demonstrated that particular vasoinhibin-generating cleavage sites are better conserved than others (Triebel et al., 2015). Understanding the evolution of the prolactin/vasoinhibin axis should help explain the generation and physiology of vasoinhibin. Further, the reconstruction of vasoinhibin emergence throughout speciation may help uncover human diseases attributed to specific vasoinhibin isoforms from the evolutionary perspective. Here, we report an in-silico study comprising multiple sequence alignments of the PRL-sequences across 91 species focussed on cathepsin-D-vasoinhibin-generating cleavage sites, along with a projection of how changes in the PRL sequence impact the

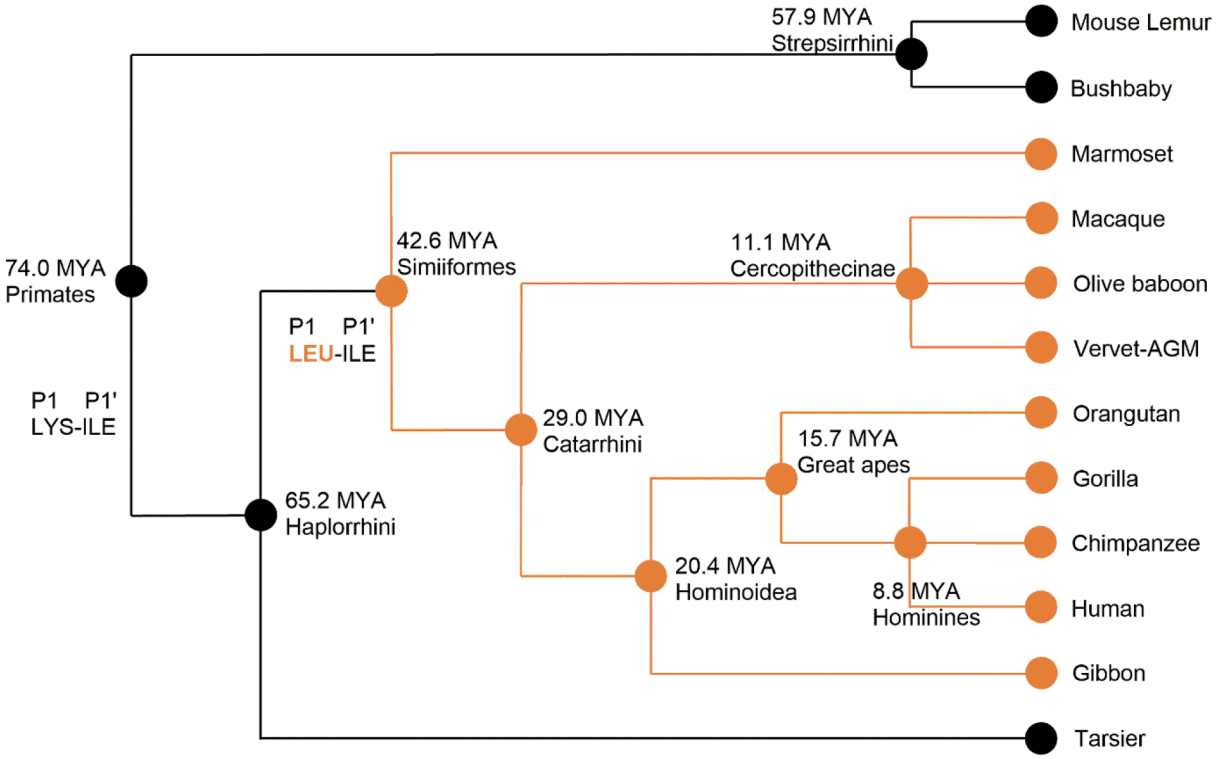


Fig. 1. Phylogenetic tree of primate evolution with P1-P1′ residues of the 15 kDa vasoinhibin cleavage site. A phylogenetic PRL-gene tree of primate evolution, as derived from ENSEMBL, together with an analysis of the cathepsin D clevage site in the PRL sequence responsible for the generation of the 15 kDa vasoinhibin isoform demonstrates the emergence of leucine in P1 between the haplorrhine and simian speciation node. Leucine in P1 is the preferred substrate of cathepsin D, and is thus associated with a high substrate affinity.

generation of vasoinhibin isoforms.

2. Methods

Selection of vasoinhibin-isoforms and cleavage sites. Each cleavage site was defined as 8 amino acids neighboring the cleavage site (P1-P4, and P1′-P4′), that means four residues towards the N-terminus and four residues towards the C-terminus of the uncleaved PRL sequence according to the nomenclature of Schechter and Berger (Schechter and Berger, 1967). The following three vasoinhibin isoforms, for which experimentally established cathepsin D-cleavage site information are available (Piwnica et al., 2004), were selected for this analysis: the 15, 16.8 and 17.2 kDa vasoinhibin isoform. The cleavage sites were located by multiple sequence alignment. The respective cleavage sites, including their amino acid position numbered with and without (inside parenthesis) the signal peptide, are as follows: 15 kDa vasoinhibin: GMELIVSQ, 160–161 (132–133), designated as the LEU160-ILE161 site; 16.8 kDa vasoinhibin: NEIY-PVWS, 175–176 (147–148), designated as the TYR175-PRO176 site; 17.2 kDa vasoinhibin: YPVW-SGLP, 178–179 (150–151), designated as the TRP178-SER179 site.

Selection of taxa and species. Species from the following taxa were selected: primates (n = 17), rodents (n = 11), laurasiatheria (n = 30), mammals (n = 9), sauropsidas (n = 13) and fish (n = 11). The full list of species is shown in the Supplementary Data Table S1.

Retrieval and alignment of sequences. PRL amino acid sequences across 91 species were retrieved from ENSEMBL, UniProt, Genbank, or NCBI (Cunningham et al., 2015). The sequences identifiers/accession numbers of these PRL-sequences are presented in the Supplementary Data Table S1. The sequences were aligned using MEGA, version 7 (clustal W alignment) (Kumar et al., 2016). Likewise, cathepsin D sequences from 48 species were retrieved (primates, n = 10, rodents n = 9, laurasiatheria n = 9, mammals n = 3, sauropsidas n = 7, and fish n = 10). Cathepsin D sequences identifiers/accession numbers are presented in Supplementary Data Table S2.

Evaluation of substrate affinity to cathepsin D. The MEROPS database of proteolytic enzymes, substrates, and inhibitors was interrogated to obtain information about the specificity of the cleavage sites in PRL (Rawlings et al., 2014). The MEROPS database contains information on the number of experimentally observed cleavages corresponding to amino acids in the 8 positions defining a cleavage site. The specificity matrix for cathepsin D was retrieved and entries of numbers of cleavages for each amino acid at P1-P4 and P1′-P4′locations of cleavage sites were added to obtain an '8P-score', - i.e., a calculated parameter estimating the suitability of a given sequence for cleavage by cathepsin D, or, more precisely, the affinity of a specific cleavage site in the PRL sequence to cathepsin D. The scores were statistically analyzed with a *t*-test using GraphPad Prism Version 9, GraphPad, San Diego, CA as described in the results section.

3. Results

3.1. PRL and cathepsin D are conserved in similar species

The present study focused on the PRL sequence of 91 species from within the classes Mammalia (Primates, Rodentia, Laurasiatheria), Sauropsida (Reptilia and Aves), and Fish (Supplemental Data, Table S1). The wide distribution of PRL in vertebrates is well known, as is the fact that PRL exists before the evolution of mammals. The cathepsin D gene was found in 48 of the 91 species in which the PRL sequence was analyzed. The 48 cathepsin D sequences were distributed over the six analyzed taxa (Supplemental Data, Table S2).

3.2. Cathepsin D cleavage sites for the generation of vasoinhibin in vertebrates

The 8P-score for the LEU160-ILE161 site of the 15 kDa vasoinhibin

Table 3
Representative cleavage sites in full-length PRL for the generation of the 16.8 kDa vasoinhibin isoform by cathepsin D.

Taxon/Species Molecular Mass	Position/Amino acid/ single position score								8P score CD-cleavage
	P4	P3	P2	P1	P1'	P2'	P3'	P4'	
Primates									
Human 16.8 kDa	172	173	174	175	176	177	178	179	
	Asn	Glu	Ile	Tyr	Pro	Val	Trp	Ser	
	29	104	70	37	3	131	3	65	442
Gorilla 16.69 kDa	173	174	175	176	177	178	179	180	
	Asn	Glu	Ile	Tyr	Pro	Val	Trp	Ser	
	29	104	70	37	3	131	3	65	442
Rodents									
Rat 16.37 kDa	171	172	173	174	175	176	177	178	
	Asn	Glu	Ile	Tyr	Leu	Val	Trp	Ser	
	29	104	70	37	110	131	3	65	549
Rabbit 16.84 kDa	173	174	175	176	177	178	179	180	
	Asn	Glu	Ile	Tyr	Ser	Val	Trp	Ser	
	29	104	70	37	41	131	3	65	480
Laurasiatheria									
Pig 16.8 kDa	174	175	176	177	178	179	180	181	
	Asn	Glu	Val	Tyr	Ser	Val	Trp	Ser	
	29	104	94	37	41	131	3	65	504
Cow 16.56 kDa	174	175	176	177	178	179	180	181	
	Thr	Glu	Pro	Tyr	Pro	Val	Trp	Ser	
	41	104	2	37	41	131	3	65	424
Mammals									
Armadillo 16.73 kDa	174	175	176	177	178	179	180	181	
	Ser	Glu	Ala	Tyr	Ser	Val	Trp	Ser	
	35	104	82	37	41	131	3	65	498
Elephant 16.50 kDa	173	174	175	176	177	178	179	180	
	Asn	Lys	Ala	Tyr	Ser	Val	Trp	Ser	
	29	34	82	37	41	131	3	65	422
Sauropsida									
Duck 17.10 kDa	178	179	180	181	182	183	184	185	
	Asn	Glu	Ile	Tyr	Ser	Gln	Trp	Glu	
	29	104	70	37	41	50	3	78	412
American Alligator 17.46 kDa	181	182	183	184	185	186	187	188	
	Asn	Glu	Val	Tyr	Ser	Arg	Trp	Ser	
	29	104	137	37	41	20	3	65	436
Fish									
Zebrafish 14.06 kDa	155	156	157	158	159	160	161	162	
	Asn	Leu	Ser	Thr	Leu	Pro	Phe	Asn	
	29	99	52	12	110	2	21	37	362
Amazon molly 15.80 kDa	182	183	184	185	186	187	188	189	
	Ala	Ile	Ser	Phe	Leu	Pro	Tyr	Thr	
	67	76	52	175	110	2	6	55	543

Multiple sequence alignment of PRL sequences across taxa at the cleavage site at which the 16.8 kDa vasoinhibin isoform is being generated by cathepsin D with their associated 8P scores. Cleavage occurs between P1 and P1′, indicated by the red line. Grey coloring indicates amino acids different than those in the human sequence. No difference is present across the taxa (primates vs. all other taxa, $p = 0.65$).

isoform was calculated for each of the species or PRL-sequences, respectively, and scores for representative species from each taxon are presented in Table 1. Primates had the highest scores with a mean value of approximately 935. All other sequences had a score between 390 and 577 (primates vs. all other taxa, $p = <0.000001$). The factor responsible for the difference in the score between rodents (and other taxa) and primates was the residue leucine at position P1 in the cleavage site. Leucine at P1 is a preferred residue in substrates of cathepsin D and its presence at P1 alone increased the score by 416 points. Sequences from all other taxa feature lysine, methionine, histidine or isoleucine at this site, which was associated with markedly lower numbers of observed cleavages (Table 1). Among the twelve primate sequences investigated, there are three outliers: the Tarsier, the Mouse Lemur, and the Bushbaby (Table 2). The PRL-sequences of these species feature lysine instead of

Table 4
Representative cleavage sites in full-length PRL of Primates aligned to Fish for the generation of the 16.8 kDa.

Taxon/Species Molecular Mass	Amino-acid/Position/Score								8P score
	P4	P3	P2	P1	P1‘	P2‘	P3‘	P4‘	
16.8 kDa Vasoinhibin Isoform									
Primates									
Human 16.8 kDa	172	173	174	175	176	177	178	179	
	Asn	Glu	Ile	Tyr	Pro	Val	Trp	Ser	
	29	104	70	37	3	131	3	65	442
Gorilla 16.69 kDa	173	174	175	176	177	178	179	180	
	Asn	Glu	Ile	Tyr	Pro	Val	Trp	Ser	
	29	104	70	37	3	131	3	65	442
Fish									
Zebrafish 14.17 kDa	156	157	158	159	160	161	162	163	
	Leu	Ser	Thr	Leu	Pro	Phe	Asn	Gly	
	136	50	67	416	3	16	43	89	820
Amazon molly 15.92 kDa	183	184	185	186	187	188	189	190	
	Ile	Ser	Phe	Leu	Pro	Tyr	Thr	Gly	
	44	50	19	416	3	12	55	89	688

Multiple sequence alignment of PRL sequences in primates and fish at the cleavage site at which the 16.8 kDa vasoinhibin isoform is being generated by cathepsin D with their associated 8P scores. Cleavage occurs between P1 and P1′, indicated by the red line. Blue coloring indicates the amino acids leucine in position P1. Fish demonstrated a significantly higher 8P score than primates ($p = 0.042$).

leucine at P1 of the 15 kDa vasoinhibin cleavage site and, consequently, demonstrate only a score of 534 (primates vs. outliers, $p = <0.000001$). Transferring this information to a phylogenetic PRL-gene tree of primate evolution (ENSEMBL) demonstrated that the high 8P-score due to leucine at P1 is a unique feature of simians. The simian speciation node dates to ~ 43 MYA and emerged from the lineage of the haplorrhine speciation node at ~ 65 MYA. This located the emergence of leucine at P1 at some time in between the two nodes (Fig. 1). The 8P-score for the TYR175-PRO176 site of the 16.8 kDa vasoinhibin isoform demonstrates no significant variation when analyzed in an alignment set containing all sequences (primates vs. all other taxa, $p = 0.65$) (Table 3). However, it was noted that there is a leucine in the cleavage site of fish sequences, not located at P1, but on P1′, whereas P1 is occupied by threonine or phenylalanine, respectively (Table 3). Since the presence of leucine on P1 or on P1′ has an impact on the 8P-score, an alternative alignment using only primate and fish sequences was performed (Table 4). In this alignment, leucine is aligned to P1, hence to the position at which it is the most preferred residue by cathepsin D, and this was associated with a significantly higher score for fish (Fish vs. primates, $p = 0.042$). The 8P-score for the TRP178-SER179 site of the 17.2 kDa vasoinhibin isoform has an even distribution of cleavage scores between primates, rodents, laurasiatheria, and mammals, sauropsida, and fish (primates vs. all other taxa, $p = 0.72$) (Table 5). An alternative alignment using only primate and fish sequences also showed no difference (primates vs. fish, $p = 0.12$) (Table 6).

4. Discussion

Cathepsin D is an aspartic-type endopeptidase and typically present in lysosomes, phagosomes, and endosomes. Cathepsin D is the primary protease responsible for the generation of vasoinhibin in adenohypophyseal PRL secretory granules and homozygous mice in which the cathepsin D gene was disrupted (cathepsin D-null mice) are devoid of hypophyseal vasoinhibin (Cruz-Soto et al., 2009). Several studies reported that cathepsin D cleaves PRL to generate vasoinhibin at an acidic but not at a neutral pH (Baldocchi et al., 1992, 1993; Clapp et al., 1993) and that this cleavage may therefore be confined to the acidic milieu of intracellular compartments or to tumor environment. Indeed, cathepsin D in human serum does not cleave PRL to generate vasoinhibin (Triebel et al., 2020), although studies reported an association of cathepsin D activity and abundance of vasoinhibin in the human circulation (Hilfiker-Kleiner et al., 2007; Nakajima et al., 2015). These studies, however, did not provide evidence that the cleavage occurs in the circulation and not elsewhere. Noteworthy, secreted cathepsin D cleaves PRL to generate vasoinhibin at neutral/physiological pH in kidney explant cultures (Piwnica et al., 2006). A physiological role for cathepsin D-mediated generation of vasoinhibin has been observed in mammary gland involution (Ishida et al., 2014), and pathological alterations of cathepsin D-mediated vasoinhibin generation were reported in peripartum cardiomyopathy and preeclampsia (González et al., 2007; Hilfiker-Kleiner et al., 2007). Altogether, cathepsin D can be regarded as the most studied enzyme responsible for vasoinhibin generation.

The present analysis investigated the phylogenesis of experimentally established cleavage sites utilized by cathepsin D to generate vasoinhibin by proteolysis of human PRL. Amino acid sequences of vertebrate species were retrieved and aligned, and the change in substrate affinity that occurs by changes in the sequences was projected using a calculated score. It was found that the cathepsin D cleavage sites are conserved throughout vertebrate evolution as they can be found in the PRL sequences of all six investigated taxa, ranging from primates to fish. The substrate affinity however varies significantly and some of the variations appear relevant. The human LEU160-ILE161 site (15 kDa vasoinhibin) demonstrated the highest 8P score in primates emerging from the speciation node of Simiiformes 42.6 MYA (Table 1 & Fig. 1). This node coincides with the substitution of lysine by leucine in P1 of the cleavage site, which is the responsible event in the cleavage site leading to high substrate affinity, and this substitution is conserved across all subsequent speciation nodes in higher primate species. Remarkably, the PRL gene of the Tarsier, the Mouse Lemur and the Bushbaby, which diverged prior to Simiiformes, did not undergo this substitution (Fig. 1). The human TYR175-PRO176 site (16.8 kDa vasoinhibin), the site for the generation of a well-studied vasoinhibin isoform, demonstrated the highest score in fish, for which, again, leucine in P1 accounts responsible. Leucine in P1 is substituted by tyrosine in Sauropsidas and all other vertebrate taxa, accounting for a relatively stable cleavage site. However, the high score in fish is only reached when aligning primate against fish PRL sequences only, as the score is lower when the sequences of all taxa are aligned altogether. It is unclear which of the two scenarios is true, and this finding is therefore to be interpreted with caution. The human TRP178-SER179 site (17.2 kDa vasoinhibin), maintains a stable score over Primates, Rodents, Laurasiatheria, Mammals, Sauropsidas and Fish.

As speculated before, the emergence of higher affinity cleavage sites in PRL for the generation of vasoinhibin might constitute gain-of-

Table 5
Representative cleavage sites in full-length PRL for the generation of the 17.2 kDa vasoinhibin isoform by cathepsin D.

Taxon/Species Molecular Mass	Amino-acid/Position/Score								8P score
	P4	P3	P2	P1	P1'	P2'	P3'	P4'	
Primates									
Human 17.2 kDa	175	176	177	178	179	180	181	182	
	Tyr	Pro	Val	Trp	Ser	Gly	Leu	Pro	
	33	30	137	29	41	56	112	55	493
Gorilla 17.07 kDa	176	177	178	179	180	181	182	183	
	Tyr	Pro	Val	Trp	Ser	Gly	Leu	Pro	
	33	30	137	29	41	56	112	55	493
Rodents									
Rat 16.77 kDa	174	175	176	177	178	179	180	181	
	Tyr	Leu	Val	Trp	Ser	Gln	Leu	Pro	
	33	99	137	29	41	38	112	55	544
Rabbit 17.22 kDa	176	177	178	179	180	181	182	183	
	Tyr	Ser	Val	Trp	Ser	Gly	Leu	Pro	
	33	50	137	29	41	56	112	55	513
Laurasiatheria									
Pig 17.18 kDa	177	178	179	180	181	182	183	184	
	Tyr	Ser	Val	Trp	Ser	Gly	Leu	Pro	
	33	50	137	29	41	56	112	55	513
Cow 16.94 kDa	177	178	179	180	181	182	183	184	
	Tyr	Pro	Val	Trp	Ser	Gly	Leu	Pro	
	33	30	137	29	41	56	112	55	493
Mammals									
Armadillo 17.10 kDa	177	178	179	180	181	182	183	184	
	Tyr	Ser	Val	Trp	Ser	Asp	Leu	Pro	
	33	50	137	29	41	39	112	55	496
Elephant 16.88 kDa	176	177	178	179	180	181	182	183	
	Tyr	Ser	Val	Trp	Ser	Gly	Leu	Pro	
	33	50	137	29	41	56	112	55	513
Sauropsida									
Duck 17.50 kDa	181	182	183	184	185	186	187	188	
	Tyr	Ser	Gln	Trp	Glu	Gly	Leu	Pro	
	33	50	25	29	71	56	112	55	431
American Alligator 17.89 kDa	184	185	186	187	188	189	190	191	
	Tyr	Ser	Arg	Trp	Ser	Gly	Leu	Pro	
	33	50	32	29	41	56	112	55	408
Fish									
Zebrafish 14.42 kDa	158	159	160	161	162	163	164	165	
	Thr	Leu	Pro	Phe	Asn	Gly	Asn	Asn	
	41	99	2	175	26	56	43	37	479
Amazon molly 16.18 kDa	184	185	186	187	188	189	190	191	
	Phe	Leu	Pro	Tyr	Thr	Gly	Gly	Thr	
	67	99	2	37	34	56	64	55	414

Multiple sequence alignment of PRL sequences across taxa at the cleavage site at which the 17.2 kDa vasoinhibin isoform is being generated by cathepsin D with their associated 8P scores. Cleavage occurs between P1 and P1′, indicated by the red line. Grey coloring indicates amino acids different than those in the human sequence. No difference is present across the taxa (primates vs. all other taxa, $p = 0.72$).

function events and might have been conserved under positive selection (Triebel et al., 2015). This view is supported by the present analysis, particularly regarding LEU161-ILE162, which demonstrates the highest substrate affinity score in higher primates. The present study also provides interesting perspective on the role of point mutations in the PRL gene that affect the amino acid sequence in the vasoinhibin cleavage sites. For example, the substitution of leucine in P1 with a non-hydrophobic residue would significantly reduce cleavage by cathepsin D at this site. A major question relates to whether vasoinhibin isoforms differ in the significance of their vascular role. This possibility is challenged by the fact that all cathepsin D generated vasoinhibin isoforms contain the functional motif (His46-Gly47-Arg48) and, thereby, should be equally active. The fact that only the cleavage site responsible for the 15 kDa vasoinhibin evolved in primates, implies that this isoform

Table 6
Cleavage site in full-length PRL of Primates aligned to Fish for the generation of the 17.2 kDa vasoinhibin isoform by cathepsin D.

17.2 kDa Vasoinhibin Isoform									
Primates									
Human 17.2 kDa	175	176	177	178	179	180	181	182	
	Tyr	Pro	Val	Trp	Ser	Gly	Leu	Pro	
	33	30	137	29	41	56	112	55	493
Gorilla 17.07 kDa	176	177	178	179	180	181	182	183	
	Tyr	Pro	Val	Trp	Ser	Gly	Leu	Pro	
	33	30	137	29	41	56	112	55	493
Fish									
Zebrafish 14.53 kDa	159	160	161	162	163	164	165	166	
	Leu	Pro	Phe	Asn	Gly	Asn	Asn	Leu	
	136	30	19	5	8	19	43	63	217
Amazon molly 16.28 kDa	186	187	188	189	190	191	192	193	
	Leu	Pro	Tyr	Thr	Gly	Gly	Thr	Asp	
	136	30	26	12	8	56	55	50	373

Multiple sequence alignment of PRL sequences in primates and fish at the cleavage site at which the 17.2 kDa vasoinhibin isoform is being generated by cathepsin D with their associated 8P scores. Cleavage occurs between P1 and P1′, indicated by the red line. No difference was present between primates and fish ($p = 0.12$).

strengthens vasoinhibin action in primates, whereas other isoforms compensate its limited effect in non-primate vertebrates. Along this line, point mutations in cleavage sites of PRL in patients in whom a pathological role of vasoinhibin is suspected, such as in preeclampsia or peripartum cardiomyopathy, may impact vasoinhibin generation and should be investigated.

5. Conclusion

Aminoacid substitutions in the PRL sequence occurring during vertebrate evolution are indeed associated to gain-of-function events, namely the generation of vasoinhibin. Leucine in P1 favoring the generation of the 15 kDa vasoinhibin isoform is a unique feature of simians, occurred between 42.6 and 65.2 MYA, is conserved in all higher primates, and may have emerged under positive selection. This agrees with earlier assumptions that PRL-encoded vasoinhibin is generated by specific, targeted proteolysis and contributes to the conservation of PRL throughout vertebrate phylogeny (Triebel et al., 2015). Moreover, this observation agrees with the nested information systems theory recently proposed by Campbell et al., that protein hormones do not follow the dogmatic pattern of synthesis, primary action, and rapid, complete proteolysis, but instead are specifically cleaved at conserved sites to generate fragments that alter extra- or intracellular functions in the primary target tissue of their precursor or elsewhere (Campbell et al., 2021), comparable to the principles underlying the prolactin/vasoinhibin axis which controls vasoinhibin function.

6. Author

A poster entitled: “The Number of Prolactin Cleavage Sites Generating Vasoinhibins Varies in Primates” was presented by JT at the Society for Molecular Biology and Evolution Conference 2016, Broadbeach, Australia, July 3–7, 2016 (SMBE, 2016).

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.ygcen.2022.114011.

References

Arnold, E., Rivera, J.C., Thebault, S., Moreno-Páramo, D., Quiroz-Mercado, H., Quintanar-Stéphano, A., Binart, N., Martínez de la Escalera, G., Clapp, C., 2010. High levels of serum prolactin protect against diabetic retinopathy by increasing ocular vasoinhibins. Diabetes 59 (12), 3192–3197. https://doi.org/10.2337/DB10-0873.

Bajou, K., Herkenne, S., Thijssen, V.L., D'Amico, S., Nguyen, N., Bouché, A., Tabruyn, S., Srahna, M., Carabin, J., Nivelles, O., Paques, C., Cornelissen, I., Lion, M., Noel, A., Gils, A., Vinckier, S., Declerck, P., Griffioen, A., Dewerchin, M., Martial, J.A., Carmeliet, P., Struman, I., 2014. PAI-1 mediates the antiangiogenic and profibrinolytic effects of 16K prolactin. Nat. Med. 20 (7), 741–747. https://doi.org/10.1038/nm.3552.

Baldocchi, R.A., Lily, T., Nicoll, C.S., 1992. Processing of rat prolactin by rat tissue explants and serum in vitro. Endocrinology 130 (3), 1653–1659.

Baldocchi, R.A., Tan, L., King, D.S., Nicoll, C.S., 1993. Mass spectrometric analysis of the fragments produced by cleavage and reduction of rat prolactin: evidence that the cleaving enzyme is cathepsin D. Endocrinology 133 (2), 935–938. https://doi.org/10.1210/ENDO.133.2.8344226.

Campbell, K.L., Haspel, N., Gath, C., Kurniatash, N., Nouduri Akkiraju, I., Stuffers, N., Vadher, U., 2021. Protein hormone fragmentation in intercellular signaling: hormones as nested information systems. Biol. Reprod. 104 (4), 887–901. https://doi.org/10.1093/BIOLRE/IOAA234.

Clapp, C., Martial, J.A., Guzman, R.C., Rentier-Delure, F., Weiner, R.I., 1993. The 16-kilodalton N-terminal fragment of human prolactin is a potent inhibitor of angiogenesis. Endocrinology 133 (3), 1292–1299. https://doi.org/10.1210/ENDO.133.3.7689950.

Clapp, C., Martínez de la Escalera, L., Martínez de la Escalera, G., 2012. Prolactin and blood vessels: A comparative endocrinology perspective. Gen. Comp. Endocrinol. 176 (3), 336–340. https://doi.org/10.1016/J.YGCEN.2011.12.033.

Clapp, C., Thebault, S., Macotela, Y., Moreno-Carranza, B., Triebel, J., Martínez de la Escalera, G., 2015. Regulation of blood vessels by prolactin and vasoinhibins. Adv. Experiment. Med. Biol. 846, 83–95. https://doi.org/10.1007/978-3-319-12114-7_4.

Cruz-Soto, M.E., Cosío, G., Jeziorski, M.C., Vargas-Barroso, V., Aguilar, M.B., Cárabez, A., Berger, P., Saftig, P., Arnold, E., Thebault, S., Martínez de la Escalera, G., Clapp, C., 2009. Cathepsin D is the primary protease for the generation of adenohypophyseal vasoinhibins: Cleavage occurs within the prolactin secretory granules. Endocrinology 150 (12), 5446–5454.

Cunningham, F., Amode, M.R., Barrell, D., Beal, K., Billis, K., Brent, S., Carvalho-Silva, D., Clapham, P., Coates, G., Fitzgerald, S., Gil, L., Girón, C.G., Gordon, L., Hourlier, T., Hunt, S.E., Janacek, S.H., Johnson, N., Juettemann, T., Kähäri, A.K., et al., 2015. Ensembl 2015. Nucl. Acids Res. 43 (Database issue), D662–D669. https://doi.org/10.1093/NAR/GKU1010.

Ge, G., Fernández, C.A., Moses, M.A., Greenspan, D.S., 2007. Bone morphogenetic protein 1 processes prolactin to a 17-kDa antiangiogenic factor. PNAS 104 (24), 10010–10015. https://doi.org/10.1073/PNAS.0704179104.

González, C., Parra, A., Ramírez-Peredo, J., García, C., Rivera, J.C., Macotela, Y., Aranda, J., Lemini, M., Arias, J., Ibargüengoitia, F., Martínez de la Escalera, G., Clapp, C., 2007. Elevated vasoinhibins may contribute to endothelial cell dysfunction and low birth weight in preeclampsia. Lab. Investig. 87 (10), 1009–1017. https://doi.org/10.1038/labinvest.3700662.

Haghikia, A., Podewski, E., Berliner, D., Sonnenschein, K., Fischer, D., Angermann, C.E., Böhm, M., Röntgen, P., Bauersachs, J., Hilfiker-Kleiner, D., 2015. Rationale and design of a randomized, controlled multicentre clinical trial to evaluate the effect of bromocriptine on left ventricular function in women with peripartum cardiomyopathy. Clin. Res. Cardiol. 104 (11), 911–917. https://doi.org/10.1007/S00392-015-0869-5.

Hilfiker-Kleiner, D., Kaminski, K., Podewski, E., Bonda, T., Schaefer, A., Sliwa, K., Forster, O., Quint, A., Landmesser, U., Doerries, C., Luchtefeld, M., Poli, V.,

Schneider, M.D., Balligand, J., Desjardins, F., Ansari, A., Struman, I., Nguyen, N., Zschemisch, N., Klein, G., Heusch, G., Schulz, R., Hilfiker, A., Drexler, H., 2007. A cathepsin D-cleaved 16 kDa form of prolactin mediates postpartum cardiomyopathy. Cell 128 (3), 589–600. https://doi.org/10.1016/j.cell.2006.12.036.

Hiroshi, K., Sower, S.A., 2006. The dawn and evolution of hormones in the adenohypophysis. Gen. Comp. Endocrinol. 148 (1), 3–14. https://doi.org/10.1016/J.YGCEN.2005.10.011.

Ishida, M., Maehara, M., Watanabe, T., Yanagisawa, Y., Takata, Y., Nakajima, R., Suzuki, M., Harigaya, T., 2014. Vasoinhibins, N-terminal mouse prolactin fragments, participate in mammary gland involution. J. Mol. Endocrinol. 52 (3), 279–287. https://doi.org/10.1530/JME-13-0189.

Kumar, S., Stecher, G., Tamura, K., 2016. MEGA7: molecular evolutionary genetics analysis version 7.0 for bigger datasets. Mol. Biol. Evol. 33 (7), 1870–1874. https://doi.org/10.1093/MOLBEV/MSW054.

Macotela, Y., Aguilar, M.B., Guzmán-Morales, J., Rivera, J.C., Zermeño, C., López-Barrera, F., Nava, G., Lavalle, C., Martínez de la Escalera, G., Clapp, C., 2006. Matrix metalloproteases from chondrocytes generate an antiangiogenic 16 kDa prolactin. J. Cell Sci. 119 (Pt 9), 1790–1800. https://doi.org/10.1242/JCS.02887.

Mejía, S., Torner, L.M., Jeziorski, M.C., Gonzalez, C., Morales, M.A., de la Escalera, G.M., Clapp, C., 2003. Prolactin and 16K prolactin stimulate release of vasopressin by a direct effect on hypothalamo-neurohypophyseal system. Endocrine 20 (1–2), 155–161. https://doi.org/10.1385/ENDO:20:1-2:155.

Nakajima, R., Ishida, M., Kamiya, C.A., Yoshimatsu, J., Suzuki, M., Hirota, A., Ikeda, T., Harigaya, T., 2015. Elevated vasoinhibin derived from prolactin and cathepsin D activities in sera of patients with preeclampsia. Hypertension Res. 38 (12), 899–901. https://doi.org/10.1038/HR.2015.99.

Perimenis, P., Bouckenooghe, T., Delplanque, J., Moitrot, E., Eury, E., Lobbens, S., Gosset, P., Devisme, L., Duvillie, B., Abderrahmani, A., Storme, L., Fontaine, P., Froguel, P., Vambergue, A., 2014. Placental antiangiogenic prolactin fragments are increased in human and rat maternal diabetes. BBA 1842 (9), 1783–1793. https://doi.org/10.1016/J.BBADIS.2014.06.026.

Piwnica, D., Fernandez, I., Binart, N., Touraine, P., Kelly, P.A., Goffin, V., 2006. A new mechanism for prolactin processing into 16K PRL by secreted cathepsin D. Mol. Endocrinol. (Baltimore, Md). 20 (12), 3263–3278. https://doi.org/10.1210/ME.2006-0044.

Piwnica, D., Touraine, P., Struman, I., Tabruyn, S., Bolbach, G., Clapp, C., Martial, J.A., Kelly, P.A., Goffin, V., 2004. Cathepsin D processes human prolactin into multiple 16K-like N-terminal fragments: Study of their antiangiogenic properties and physiological relevance. Mol. Endocrinol. 18 (10), 2522–2542.

Rand-Weaver, M., Kawauchi, H., 1993. Growth Hormone, prolactin, and somatolactin: a structural overview. In: Hochaka, P.W., Mammsen, T.P. (Eds.), Biochemistry and Molecular Biology of Fishes. Elsevier Science Publishers, pp. 39–56.

Robles, J.P., Zamora, M., Velasco-Bolom, J.L., Tovar, M., Garduño-Juárez, R., Bertsch, T., Martínez, G., de la Escalera, J., Triebel, C Clapp, 2018. Vasoinhibin comprises a three-helix bundle and its antiangiogenic domain is located within the first 79 residues. Sci. Rep. 8 (1) https://doi.org/10.1038/S41598-018-35383-7.

Rawlings, N.D., Waller, M., Barrett, A.J., Bateman, A., 2014. MEROPS: the database of proteolytic enzymes, their substrates and inhibitors. Nucleic Acids Res. 42 (D1), D503–D509. https://doi.org/10.1093/NAR/GKT953.

Robles, J.P., Zamora, M., Siqueiros-Marquez, L., Adan-Castro, E., Ramirez-Hernandez, G., Nuñez, F., Lopez-Casillas, F., Millar, R.P., Bertsch, T., Martínez de la Escalera, G., Triebel, J., Clapp, C., 2022. The HGR motif is the antiangiogenic determinant of vasoinhibin: implications for a therapeutic orally active oligopeptide. Angiogenesis 25 (1), 57–70. https://doi.org/10.1007/s10456-021-09800-x.

Schechter, I., Berger, A., 1967. On the size of the active site in proteases. I. Papain. Biochem. Biophys. Res. Commun. 27 (2), 157–162. https://doi.org/10.1016/S0006-291X(67)80055-X.

Triebel, J., Huefner, M., Ramadori, G., 2009. Investigation of prolactin-related vasoinhibin in sera from patients with diabetic retinopathy. Eur. J. Endocrinol. 161 (2), 345–353. https://doi.org/10.1530/EJE-09-0130.

Triebel, J., Bertsch, T., Bollheimer, C., Rios-Barrera, D., Pearce, C.F., Hüfner, M., Martínez de la Escalera, G., Clapp, C., 2015. Principles of the prolactin/vasoinhibin axis. Am. J. Physiol. Regul. Integr. Comp. Physiol. 309 (10), R1193–R1203. https://doi.org/10.1152/AJPREGU.00256.2015.

Triebel, J., Schauer, N., Zamora, M., Moreno-Vega, A.I., Escalera, G.M., Clapp, C., Bertsch, T., 2020. Matrix metalloproteases and cathepsin D in human serum do not cleave prolactin to generate vasoinhibin. Clinical laboratory. 66 (5), 877–886. https://doi.org/10.7754/CLIN.LAB.2019.191017.

Wallis, M., 2000. Episodic evolution of protein hormones: molecular evolution of pituitary prolactin. J. Mol. Evol. 50 (5), 465–473. https://doi.org/10.1007/S002390010049.

Wallis, O.C., Mac-Kwashie, A.O., Makri, G., Wallis, M., 2005. Molecular evolution of prolactin in primates. J. Mol. Evol. 60 (5), 606–614. https://doi.org/10.1007/S00239-004-0239-9.

Zamora, Magdalena, Robles, Juan Pablo, Aguilar, Manuel B., Romero-Gómez, Sergio de Jesús, Bertsch, Thomas, Martínez de la Escalera, Gonzalo, Triebel, Jakob, Clapp, Carmen, 2021. Thrombin cleaves prolactin into a potent 5.6-kDa vasoinhibin: implication for tissue repair. Endocrinology 162 (12). https://doi.org/10.1210/endocr/bqab177.

Zamorano, M., Ledesma-Colunga, M.G., Adán, N., Vera-Massieu, C., Lemini, M., Méndez, I., Moreno-Carranza, B., Neumann, I.D., Thebault, S., Martínez, G., de la Escalera, L., Torner, C Clapp, 2014. Prolactin-derived vasoinhibins increase anxiety- and depression-related behaviors. Psychoneuroendocrinology. 44, 123–132. https://doi.org/10.1016/J.PSYNEUEN.2014.03.006.

Society for Molecular Biology and Evolution Conference 2016, Broadbeach, Australia, July 3–7, 2016.