MOLECULAR EVOLUTION OF GPCRS

Somatostatin/urotensin II receptors

Hervé Tostivint, Daniel Ocampo Daza1, Christina A Bergqvist1, Feng B Quan, Marion Bougerol, Isabelle Lihrmann2 and Dan Larhammar1

Evolution des Régulations Endocriniennes, UMR 7221 CNRS and Muséum National d’Histoire Naturelle, Paris, France
1Department of Neuroscience, Science for Life Laboratory, Uppsala University, Uppsala, Sweden
2Inserm U982, Neuronal and Neuroendocrine Differentiation and Communication Laboratory, Institute for Research and Innovation (IRIB), Rouen University, Mont-Saint-Aignan, France

Correspondence should be addressed to H Tostivint, I Lihrmann or D Larhammar
Emails hto@rmsn.fr, isabelle.lihrmann@univ-rouen.fr or Dan.Larhammar@neuro.uu.se

Abstract

Somatostatin (SS) and urotensin II (UII) are members of two families of structurally related neuropeptides present in all vertebrates. They exert a large array of biological activities that are mediated by two families of G-protein-coupled receptors called SSTR and UTS2R respectively. It is proposed that the two families of peptides as well as those of their receptors probably derive from a single ancestral ligand–receptor pair. This pair had already been duplicated before the emergence of vertebrates to generate one SS peptide with two receptors and one UII peptide with one receptor. Thereafter, each family expanded in the three whole-genome duplications (1R, 2R, and 3R) that occurred during the evolution of vertebrates, whereupon some local duplications and gene losses occurred. Following the 2R event, the vertebrate ancestor is deduced to have possessed three SS (SS1, SS2, and SS5) and six SSTR (SSTR1–6) genes, on the one hand, and four UII (UII, URP, URP1, and URP2) and five UTS2R (UTS2R1–5) genes, on the other hand. In the teleost lineage, all these have been preserved with the exception of SSTR4. Moreover, several additional genes have been gained through the 3R event, such as SS4 and a second copy of the UII, SSTR2, SSTR3, and SSTR5 genes, and through local duplications, such as SS3. In mammals, all the genes of the SSTR family have been preserved, with the exception of SSTR6. In contrast, for the other families, extensive gene losses occurred, as only the SS1, SS2, UII, and URP genes and one UTS2R gene are still present.

Key Words
- neuropeptides
- somatostatin
- urotensin II
- GPCR
- multigene families
- duplication
- tetraploidization

Introduction

The idea that the genomes of vertebrates are the result of two rounds of genome duplications was first suggested by Ohno (1970) based on chromosome numbers and genome sizes. Information about gene locations on chromosomes in human and mouse genomes provided evidence supporting this notion in the late 1980s and early 1990s (Lundin 1993). Since then, it has been constantly strengthened by a growing body of knowledge, in particular, owing to the availability of a large variety of assembled and annotated vertebrate genomes in comparison with deuterostome invertebrate genomes. It is now well established that vertebrates underwent two major rounds of whole-genome duplications during their evolution (Nakatani et al. 2007, Putnam et al. 2008, Van de Peer et al. 2010). These tetraploidization events, referred to as 1R and 2R, probably occurred before the split of the extant agnathan (jawless vertebrates) and gnathostome (jawed vertebrates) lineages (Smith et al. 2013).
An additional round of tetraploidization, referred to as 3R, occurred in the stem lineage of teleost fish (Jaillon et al. 2004, Kasahara et al. 2007).

It has been suggested that the 2R event yielded many new genes that have contributed to vertebrate features (Shimeld & Holland 2000). Among them, the establishment of the hypothalamic-pituitary axis, leading to the neuroendocrine control of many complex functions, is probably one of the most critical (Sower et al. 2009). The activity of this system is based on the action of multiple peptidic factors belonging to families whose expansion has been shown to mainly coincide with the 2R event. Such is the case, for instance, of the gonadotropin-releasing hormone (GNRH; Kim et al. 2011, Tostivint 2011, Decatur et al. 2013) and corticotropin-releasing hormone (CRH; Hwang et al. 2013) families that have been intensively studied during the last few years. It is particularly worth noting that the 2R event has contributed to the diversification of not only the peptides but also their receptors (Kim et al. 2011, Hwang et al. 2013).

Somatostatin (SS) is another important component of the hypothalamic-pituitary axis, involved in the control of growth hormone (GH) secretion (Gahete et al. 2010). The aim of the present review is to demonstrate how critical the 2R event has been in the evolution of the SS family and its receptors. The history of the urotensin II (UII) family and its receptors is also discussed herein because of the evolutionary relationships between SS-ergic and UII-ergic systems (Tostivint et al. 2006).

Peptides of the SS family

The SS family in vertebrates is composed of at least six paralogous genes named SS1, SS2, SS3, SS4, SS5, and SS6 according to the nomenclature initially proposed by Tostivint et al. (2008) and then expanded by Liu et al. (2010).

SS1

SS1 is a cyclic peptide that was originally isolated from ovine hypothalamic extracts as a result of its ability to inhibit the release of GH by pituitary cells (Brazeau et al. 1973). SS1 has subsequently been demonstrated to exist in two forms, SS-14, composed of 14 amino acids (Brazeau et al. 1973), and SS-28, its N-terminally extended form, composed of 28 residues, first characterized in porcine gut (Esch et al. 1980, Pradayrol et al. 1980). Both peptides are generated from the same precursor molecule (called prepro-SS1 or PSS1) through post-translational processing at two cleavage sites, a dibasic Arg-Lys site and a monobasic Arg site respectively (Patel & Galanopoulou 1995). Note that in this review, the term SS1 will be used to collectively describe all the PSS1 products containing the SS-14 sequence at their N-terminus, whatever their size. In contrast, the term SS-14 will be preferred to specifically designate the PSS1 product of 14-residues.

Products of the SS1 gene have been characterized in representatives of all vertebrate classes (reviewed in Conlon et al. (1997) and Tostivint et al. (2008)). The primary structure of SS-14 is identical in almost all species investigated so far (Fig. 1). In this regard, it probably represents one of the most strongly conserved peptides during the evolution of vertebrates (Conlon et al. 1997). The only exceptions concern the Pacific ratfish (Hydrolagus colliei; Conlon 1990) and elephant shark (Callorhinus milii; the present study), two cartilaginous fish belonging to the group of chimaeras. In both species, Asn at position 5 is substituted by Ser in SS1 (Fig. 1). Cleavage sites generating SS-14 and its N-terminally extended form have also been fully conserved (Fig. 1; Conlon et al. 1997, Tostivint et al. 2008). In mammals, SS-14 and SS-28 peptides are synthesized in variable amounts depending on the tissue, as a result of differential precursor processing: SS-14 is predominantly produced by neurons and secretory cells in the CNS and peripheral nervous system and SS-28 in the gastrointestinal tract (Patel 1999). SS-14 is also predominant in the pancreas of species from all vertebrate classes that have been studied, with the exception of bowfin (Amia calva; Wang et al. 1993) and hagfish (Myxine glutinosa; Conlon et al. 1988a). It is worth noting that the N-terminally extended forms of SS-14 have never been detected in teleost fish, indicating that structural features of the N-terminal domain of the precursor may be important for recognition by the processing enzyme (Conlon et al. 1997).

PSS1 has recently been shown to produce an additional 13-amino acid amidated peptide called neuronostatin, immediately following the signal peptide (Samson et al. 2008). This peptide was initially predicted to exist by bioinformatic analyses of evolutionarily conserved sequences, although its occurrence has only been demonstrated in mammals thus far.

SS2/cortistatin

SS2 was first purified from brain extracts of the marsh frog Pelophylax ridibundus (previously Rana ridibunda) concurrently with SS1 (Vaudry et al. 1992). The cDNA encoding its precursor (PSS2) was characterized soon thereafter (Tostivint et al. 1996). An ortholog of the frog
SS1

Human Homo sapiens -- RANSNPAMAP---GCKNNFVKTPTSC BC032625
Mouse Mus musculus -- RANSNPAMAP---GCKNNFVKTPTSC NM_009215.1
Chicken Gallus gallus -- RANSNPAMAP---GCKNNFVKTPTSC X60191
Clawed frog Xenopus laevis -- RANSNPAMAP---GCKNNFVKTPTSC AI316318.1
Zebrafish Danio rerio -- RAN---G---QPPLAP---GCKNNFVKTPTSC BC076254
Medaka Oryzias latipes -- RAAAGPPLAP---GCKNNFVKTPTSC ENSORLGD00000019152
Fugu Takifugu rubripes -- RAAAGPPLAP---GCKNNFVKTPTSC ENSTRUGGD00000010048
Bowfin Amia calva -- RANPALAP---GCKNNFVKTPTSC AA682439.1
Spotted gar Lepisosteus oculatus -- RANPALAP---GCKNNFVKTPTSC JX964750
Elephant shark Callorhinchus milii -- RANPALAP---GCKNNFVKTPTSC ESGP Sc451; 7423-7343
Ratfish Hydrodolus colliei -- AVERPRDGQVHEPPGRDK GCKNNFVKTPTSC Conlon 1990
Hagfish Myxine glutinosa

SS2 (CST)

Human Homo sapiens RABLIAAPP--QQSARKK---PCIKNNFVKTPTSC AF013252
Mouse Mus musculus RQERFPP--QPFFPLKR---PCIKNNFVKTPTSC AF501516.1
Opossum Monodelphis domestica RALP----HVRPLKR---PCIKNNFVKTPTSC ENSMODGD0000002663
Chicken Gallus gallus RAL---OQLLSQRIK---PCIKNNFVKTPTSC AY221945.1
Marsh frog Pelophylax ridibundus ---L---L---L---PCIKNNFVKTPTSC U68136.1
Zebrafish Danio rerio VA-VNYRLS-QLPQRDK---PCIKNNFVKTPTSC BG307388
Medaka Oryzias latipes VA-VNYRLS-QLPQRDK---PCIKNNFVKTPTSC BJJ85293.1
Fugu Takifugu rubripes VA-VNYRLS-QLPQRDK---PCIKNNFVKTPTSC ENSTRUGGD00000012370
Spotted gar Lepisosteus oculatus VA-VNYRLS-QLPQRDK---PCIKNNFVKTPTSC ENSLOGD00000006494
Dogfish Scyliorhinus canicula VA-VNYRLS-QLPQRDK---PCIKNNFVKTPTSC JX964752
Elephant shark Callorhinchus milii VA-VNYRLS-QLPQRDK---PCIKNNFVKTPTSC ESGP Sc58:1649040-1649075

SS2-like

Little skate Leucoraja erinacea ---YAQAEITPRKQ---PCIKNNFVKTPTSC ESGP Sc93:2822783-2822824

SS3

Zebrafish Danio rerio RENSNHPARDEGCIKNNFVKTPTSC XM_689061
Medaka Oryzias latipes EAUAN--VPFFPLRDGCIKNNFVKTPTSC AU168379.1
Fugu Takifugu rubripes EAUPHANLSPPLRDGCIKNNFVKTPTSC AU168379.1
Anglerfish Lophius americanus EAUTSPNPLPRRDGCIKNNFVKTPTSC P01170.2
Spotted gar Lepisosteus oculatus EAUTSPNPLPRRDGCIKNNFVKTPTSC ENSLOGD00000009445

SS4

Zebrafish Danio rerio V---K---K---K---GCKNNFVKTPTSC NM_131727.1
Medaka Oryzias latipes RF---G---T---K---GCKNNFVKTPTSC ENSLG17:24768900-24769028
Fugu Takifugu rubripes RFDLNVRSQPCRRFMWKSSTYCC ENSC7:374465-5145
Catfish Ictalurus punctatus RF---DN-TVTPPLNCMNYWKSSTAC J00945.1

SS5

Zebrafish Danio rerio P1VPSQREFGCIKNNFVKTPTSC XM_001333046
Medaka Oryzias latipes RM---V---K---GCKNNFVKTPTSC ENSLG18:1406975-1407049
Fugu Takifugu rubripes RMRVFPSQREFGCIKNNFVKTPTSC ENSLG18:1406975-1407049
Spotted gar Lepisosteus oculatus RP---LPLAQRDEGCIKNNFVKTPTSC ENSLOGD00000013292
Dogfish Scyliorhinus canicula RPNNP---DRREGCIKNNFVKTPTSC JX964751

SS6

Zebrafish Danio rerio V---DK---SK---K---GCKNNFVKTPTSC NM_001128784
Spotted gar Lepisosteus oculatus V---MPPVPSQREFGCIKNNFVKTPTSC ENSLG18:12213192-12213236

Other

Sea lamprey Petromyzon marinus

References

---RSANSSPVLAP--RERK
---RAA--GPMLAP--RERK
---RAAASGPLLAP--RDRK
---SAN--PALAP--RERK
---RSAN--PALAP--RERK
---RSAN--PNLAQ--RERK
AGCK
RR--DN-TVTSKPLN
RRDLDNLRVQQRQFG
RR--GVEIATQR--P
RRD----TKEERKTG
---RA--HSRQ--RP
---RAVIRFTQRERK
---AALRAAAVAGSPQQLPRLGQRERK
---RS--VFGPSVSRERTP
---RS-VFGPSVSRERTP
---RS-VFGPSVSRERTP
---RS-A--G---P---STP
---RS--VFGPSVSRERTP
---RS--VFGPSVSRERTP
---RS--VFGPSVSRERTP
---RS--VFGPSVSRERTP
---RS--VFGPSVSRERTP
---RS--VFGPSVSRERTP
---RS--VFGPSVSRERTP
---RS--VFGPSVSRERTP

Figure 1
Alignment of SS sequences including their putative N-terminally extended forms from some representative vertebrate species. The overall structure of the SS precursors is shown. Conserved residues are shown in bold. Putative cleavage sites are enclosed by boxes. References correspond to GenBank or Ensembl accession numbers. For the non-annotated genes, the chromosomal location is indicated according to the Ensembl (ENS) or the Elephant Shark Genome Project (ESGP) data. SP, signal peptide; SS, somatostatin; LG, linkage group; sc, scaffold.
SS2 gene has been subsequently revealed by cDNA cloning in goldfish (Carassius auratus; Lin et al. 1999b), West African lungfish (Protopterus annectens; Trabucchi et al. 1999), white sturgeon (Acipenser transmontanus; Trabucchi et al. 2002), and chicken (Gallus gallus; Trabucchi et al. 2003). It is now well established that the SS2 gene exists in all gnathostome lineages (reviewed in Tostivint et al. 2008, 2013). In most species examined so far, SS2 exhibits only one substitution (Pro for Ala2) compared with SS1. The exceptions are SS2 in the chondrichthyan elephant shark (Callorhinichus milii) and in the marsh frog, where there are additional substitutions (Arg for Val1 and Met for Ser13 respectively) (Fig. 1). It is worth noting that additional SS2-related sequences have been identified in the elephant shark and little skate (Leucoraja erinacea), two chondrichthyan species, albeit without Pro2 in little skate (Fig. 1; Liu et al. 2010).

Cortistatin (CST) was discovered by de Lecea et al. (1996) as a putative product of a novel 112-amino acid polypeptide sharing 11 amino acids with SS1 at its C-terminus. It was named CST in recognition of its strong immunoreactivity with the same chromatographic properties as synthetic CST-14 (Ramírez et al. 2002, de Lecea 2008).

It is not clear whether CST is the mammalian counterpart of SS2, and the question has long been debated (Conlon et al. 1997, Tostivint et al. 2004a). Although both peptides have a Pro2, which is considered a SS2 signature, they are derived from precursors whose sequences are very divergent, except for their C-terminal ends. Moreover, CST possesses a C-terminal Lys that has never been found in SS2 peptides outside of placental mammals (Fig. 1). However, definitive evidence that the CST gene is indeed an ortholog of the SS2 gene has been provided by synteny analysis showing that the neighboring regions of the two genes are composed of the same set of orthologous genes (Tostivint et al. 2004b).

SS3

The SS3 gene (previously called SSII) encodes a precursor (PSS3) that contains the [Tyr7, Gly10]SS-14 sequence at its C-terminus (Fig. 1). PSS3 products, which were the first SS variants to be discovered (Hobart et al. 1980), have been characterized in numerous teleost fish species, including American angelfish (Lophius americanus; Andrews et al. 1984), coho salmon (Oncorhynchus kisutch; Plisetskaya et al. 1986), European flounder (Platichthys flesus; Conlon et al. 1987), shortnose sculpin (Cottus scorpius; Conlon et al. 1987), European eel (Anguilla anguilla; Conlon et al. 1988b), Nile tilapia (Oreochromis niloticus; Nguyen et al. 1995), and goldfish (Carassius auratus; Usaka et al. 1995, Lin et al. 1999b). Most of them are 25–28 amino acids in length and correspond to the N-terminally extended forms of the [Tyr7, Gly10]SS-14 sequence (Fig. 1). In contrast, the [Tyr7, Gly10]SS-14 peptide, though flanked by a putative dibasic Arg–Lys cleavage site at its N-terminus, is generally not produced (Conlon et al. 1997). Notably, SS3 exhibits an additional substitution within the [Tyr7, Gly10]SS-14 sequence in some species, such as eel (Anguilla anguilla), tilapia (Oreochromis niloticus), and medaka (Oryzias latipes) (reviewed in Tostivint et al. 2013).

Although the SS3 gene has long been thought to exist in only teleosts, recent genome database analyses have revealed the presence of its counterpart in the spotted gar (Lepisosteus oculatus), a non-teleost actinopterygian fish (Tostivint et al. 2013). In the spotted gar, SS3 exhibits a Tyr residue at position 7, but no Gly at position 10.

SS4

SS4 was originally identified in the channel catfish (Ictalurus punctatus; Magazin et al. 1982) and subsequently in other closely related species, such as zebrafish (Danio rerio; Devos et al. 2002), leading to the notion that it was restricted to the group of ostariophys (Tostivint et al. 2008). However, it is now apparent that an SS4 gene is present in all teleost species for which whole-genome sequences are currently available (Liu et al. 2010).

SS4 exhibits a very atypical structure compared with SS1 and other SS isoforms, sharing only four residues with them: the two Cys residues involved in disulfide bridge formation and the Phe–Trp motif inside the cyclic domain (Fig. 1). The structure of SS4 is highly divergent between species, in both length and primary sequence, mainly because of the erratic positions of the dibasic cleavage sites (reviewed in Liu et al. 2010).

SS5

The occurrence of the SS5 gene has been established only recently by Liu et al. (2010). The first representative of this gene was reported in the Japanese pufferfish or fugu.
However, the second peptide, a 33- to 37-residue peptide corresponds to SS1 as it exhibits the conserved SS-14 sequence. Variants have been characterized (Andrews et al. 1994), but their position in the current classification remains unclear. In the lampreys Petromyzon marinus, Lampetra fluviatilis, and Geotria australis, two SS variants have been characterized (Andrews et al. 1988, Sower et al. 1994, Conlon et al. 1995a,b, Wang et al. 1999). The first one is a tetradecapeptide that probably corresponds to SS1 as it exhibits the conserved SS-14 sequence. However, the second peptide, a 33- to 37-residue peptide that contains the [Ser^{12}]SS-14 sequence at its C-terminus, is totally unique (Fig. 1). One possibility could be that the [Ser^{12}]SS-14 variant is a product of an SS5 gene because, as has been mentioned above, SS5 and SS1 peptides are structurally very similar. Another possibility could be that it is a duplicate that arose in the recently proposed third lamprey-specific tetraploidization (Mehta et al. 2013). Further information regarding the sequences of their precursors and the location of their genes would be needed to definitively resolve this issue. It is worth noting that, until now, no member of the SS family could be retrieved from the sea lamprey (Petromyzon marinus) genome sequence (Decatur et al. 2013).

Peptides of the UII family

The UII family consists of four paralogous genes called UII, URP, URP1, and URP2 according to the nomenclature introduced by Nobata et al. (2011).

Urotensin II

UII is a cyclic peptide first isolated from the teleost urophysis, a neurosecretory organ located in the caudal portion of the spinal cord of teleost fish, on the basis of its ability to induce contractions of the trout hindgut (Pearson et al. 1980). Because the urophysis is a peculiarity of teleosts, the presence of UII was long thought to be restricted to this group (Conlon 2008). This view was proven to be incorrect when UII was also identified in other groups such as amphibia (Pelophylax ridibundus; Conlon et al. 1992a), chondrostees (Acipenser ruthenus and Polyodon spathula; McMaster et al. 1992, Waugh et al. 1995), elasmobranchs (Leucoraja erinacea, Raja rhina, and Scylliorhinus canicula; Conlon et al. 1992b, Waugh & Conlon 1993, the present study), chimaeras (Callorhinus milii), and lampreys (Petromyzon marinus and Lampetra fluviatilis; Waugh et al. 1995). It is now well established that a UII gene is present in all vertebrate classes (reviewed in Vaudry et al. (2010) and Lihrmann et al. (2013)). Notably in some teleost species, such as white sucker (Catostomus commersonii; McMaster & Lederis 1983) and zebrafish (Parmentier et al. 2008), the occurrence of two UII isoforms encoded by distinct genes has been reported. This number goes up to four in the common carp (Cyprinus carpio; Ichikawa et al. 1984).

All mature UII peptides described so far in vertebrates (reviewed in Conlon et al. (1997), Leprince et al. (2008), Vaudry et al. (2010) and Lihrmann et al. (2013)) share a common cyclic sequence (Cys–Phe–Trp–Lys–Tyr–Cys), while the N-terminus is rather variable in both length and amino acid sequence (Fig. 2). Additionally, proteolytic cleavage of the UII precursor at dibasic (Arg–Lys) or tribasic...
(Lys–Lys–Arg) sites has the potential to generate mature peptides whose sequences may vary from 11 to 17 amino acids in length (Fig. 2). UII peptides composed of 11 and 17 residues have actually been detected in humans (Chartrel et al. 2004) and mice (Dubessy et al. 2008) respectively. In species other than mammals, the mature UII generally possesses 13 residues, except in the marsh frog Pelophylax ridibundus, which has only 12 residues (Conlon et al. 1992a; Fig. 2). In the flounder Platichthys flesus, the occurrence of a nonapeptide UII (4–12) has been detected in addition to that of the dodecapeptide UII itself, but only as a minor product (Conlon et al. 1990).

**UII-related peptide**

UII-related peptide (URP) was the first UII variant to be discovered when it was isolated from rat brain extracts by Sugo et al. (2003). The cDNA encoding its precursor was cloned in rats, mice, and humans (Sugo et al. 2003) and then in chickens (Tostivint et al. 2006) and the African clawed frog (Xenopus laevis; Konno et al. 2013). The URP gene has previously been thought to exist only in tetrapods, but its occurrence in teleosts has been demonstrated recently (Quan et al. 2012).

In tetrapods, the URP sequence is only seven amino acids in length and is preceded by a single dibasic cleavage site.
Its cyclic region exhibits the same structure as that of UII (Fig. 2). However, the nature of the residue flanking this region is significantly different, namely a non-polar residue (Val or Ile) in URP in place of an acidic residue (Asp or Glu) in UII (Fig. 2). In teleost fish, the cyclic region is also conserved, but it is preceded by a three-residue extension whose sequence differs widely from species to species (Quan et al. 2012; Fig. 2).

URP1

The occurrence of the URP1 gene was first revealed in the Japanese eel (Anguilla japonica) by cDNA cloning (Nobata et al. 2011). The URP1 gene was initially considered to be the teleost counterpart of the URP gene previously identified in tetrapods (Nobata et al. 2011), but this view was invalidated when it appeared that several teleost species possess both the URP and the URP1 genes (Quan et al. 2012). An ortholog of the URP1 gene has also been found in a chondrichthyan, the elephant shark (Callorhinchus milii; Parmentier et al. 2011), in a non-teleost actinopterygian fish, the spotted gar (Lepisosteus oculatus; Tostivint et al. 2013), and in coelacanth (Latimeria chalumnae), a representative of an early diverging lineage among the sarcopterygians (Tostivint et al. 2013). In contrast, it could not be found in tetrapod genomes.

URP1 has the same cyclic region as UII and URP, but it specifically possesses an extra three-residue extension at its C-terminus (Fig. 2). The single residue located just upstream of the cyclic region is a non-polar residue, similar to that in URP (Nobata et al. 2011, Parmentier et al. 2011; Fig. 2).

URP2

Similar to the URP1 gene, the URP2 gene is currently known to exist in chimaeras and actinopterygians and not in tetrapods. It was initially identified in teleost species (Nobata et al. 2011, Parmentier et al. 2011) and subsequently in the spotted gar (Tostivint et al. 2013) and elephant shark (the present study). The primary structure of URP2 is very similar to that of URP1, with the presence of a single non-polar residue at its N-terminus and a three-residue extension at its C-terminus (Fig. 2). It is notable that the URP2 precursor possesses two conserved dibasic cleavage sites, making it possible for an N-terminally extended form of the peptide to be produced (Fig. 2).

Molecular evolution of the SS and UII gene families

The 2R model predicts that vertebrate genomes contain four paralogs for each proto-ortholog of the chordate ancestral genome. These four paralogs are theoretically expected to be present on distinct chromosomes, but remain physically linked to the same groups of paralogous genes that together form a tetraparalogon. In reality, the integrity of these so-called syntenic blocks may be disrupted by gene losses and/or chromosome rearrangements, making them sometimes more difficult to identify.

Genome synteny analysis comparing vertebrate SS-containing genome fragments with human chromosomes allowed the identification of four genome fragments that share several paralogous genes. These notably include ATP1B1/2/3, DVL1/2/3, FGF12/13/14, MBLN1/2/3, RAB2A/B/C, SENP1/3/5, and TP53/63/73 (Tostivint et al. 2006, Liu et al. 2010; Fig. 3). This observation indicates that these paralogous genes are probably members of four early regions containing SS1, SS2, and SS5 genes, as well as a fourth putative SS gene apparently lost today. Very interestingly, the four genes of the UII family, namely UII, URP, URP1, and URP2, were also found on the same fragments (Parmentier et al. 2011; Fig. 3), indicating that they emerged at the same time as the four SS genes. Taken together, these data support the view that the 2R event played a major role in the expansion of the SS and UII families (Fig. 4). The occurrence of the URP1, URP2, and SS genes, both in teleosts and in the spotted gar and/or elephant shark (Callorhinchus milii), confirms that these genes, albeit absent in tetrapods, predate the 3R event and thus were already present in the gnathostome ancestor.
Figure 4

A proposed evolutionary scheme for the evolution of the genes of the SS and UII families in gnathostomes. Crossed-out boxes represent lost genes. The ancestral vertebrate had one SS gene and one UII gene that are thought to have arisen through local duplication. The two basal vertebrate tetraploidizations (1R and 2R) quadrupled each of these genes, whereupon the SS quartet lost one gene. Early in tetrapod evolution, the SS5, URP1, and URP2 genes were lost. In the actinopterygian lineage, the SS1 and SS2 genes were duplicated through local duplication. Then, the SS1 and UII genes were duplicated in the basal teleost tetraploidization (3R). No new SS2, SS5, URP, URP1, or URP2 genes emerged through the 3R event. In the chondrichthyan lineage, three genes of the ancestral SS quartet were apparently preserved, while only one gene of the UII family has been identified so far.
Among vertebrates, teleost fish have the highest number of SS and UII variants (Conlon et al. 1997), and it has been shown that a part of this specific diversity is due to the 3R event. Several genes can be clearly identified as 3R products, namely the SS4 gene and the UIIα and UIIβ genes that arose from common ancestors shared with the SS1 and UII genes respectively (Tostivint et al. 2006, Liu et al. 2010; Fig. 4).

Following a whole-genome duplication event, a large proportion of the duplicated genes (about 80%) are usually lost. However, not all genes have the same probability of suffering this fate (Maere et al. 2005). Current knowledge on the composition of the UII and SS families indicates that gene losses occurred at different rates and time points and were lineage-dependent. Thus, in the teleost lineage, only one of the eight genes generated through the 2R event has been lost, while in the tetrapod lineage, half of them have disappeared. In contrast, gene losses following the 3R event were massive, as only the duplicates of SS1 (SS4) and UII (UIIα and UIIβ) have been preserved (Liu et al. 2010, Tostivint et al. 2013; Fig. 4).

The tetraploidizations by no means explain all vertebrate gene duplications. Numerous examples exist of genes that have been duplicated by other mechanisms, usually by local (tandem) duplication. So far, two genes generated through this mechanism have been detected in the SS family, namely the SS3 (Tostivint et al. 2004a,b) and SS6 (Liu et al. 2010, Tostivint et al. 2013) genes that arose from common ancestors shared with the SS1 and SS2 genes respectively (Fig. 4). The fact that the SS3 and SS6 genes are present both in teleosts and in the spotted gar indicates that these duplications occurred before the 3R event, early in the evolution of actinopterygians. As has been mentioned above, the occurrence of a second SS2-like sequence has been reported in the elephant shark (Callorhinchus milii) and little skate (Leucoraja erinacea). The possibility that this peptide is an SS6 ortholog cannot be totally excluded, although the localization of the gene in both species is currently unknown. If true, this would imply that the SS6 gene appeared much earlier, before the chondrichthyan/osteichthyan split.

**Origin of the UII and SS ancestral genes**

UII was first reported as an SS-like peptide by Pearson et al. (1980) on the basis of its structural similarities with SS1. UII and SS1 actually share several features, including a disulfide bridge and a common Phe–Trp–Lys motif, both of which are essential for their biological activity (Fig. 5). Moreover, they derive from precursors whose general organization is also very similar. From all these observations, it is tempting to assume that UII and SS are derived from a common ancestral gene.

As has been mentioned above, the UII and SS2 genes and the URP and SS1 genes exist in proximity on the same chromosomes in all species investigated so far (Tostivint et al. 2006, Liu et al. 2010). In teleosts, the URP1 and SS5 genes are located on the same region of two paralogous chromosomes (Parmentier et al. 2011; Fig. 4). This indicates that the ancestral SS and UII genes were already closely linked before the 1R event and that both genes may have been generated through a local duplication.

Many peptide hormone families have been shown to have members in both protostomes and deuterostomes. Typical examples are the tachykinins, vasopressin, CRH, cholecystokinin gastrin, neuromedin, etc. (De Loof et al. 2012). The question of whether SS exists in invertebrate species has long been debated. Peptides of the allatostatin C (ASTC) family, which are especially well known in insects, have received particular attention because they are cyclic peptides and interact with receptors that are evolutionarily related to SS receptors (Kreienkamp et al. 2002). Interestingly, they have also been shown to act as inhibitors of endocrine secretions such as SS1 (Bendena et al. 1999). However, in terms of sequence conservation, ASTC and SS do not share any notable similarity aside from the two Cys residues that form the disulfide bridge (Fig. 5). Whether the ASTC and SS families are actually evolutionarily related, as Mirabeau & Joly (2013) claim, or whether
their structural similarities are due to convergent evolution remains to be determined. An SS-like gene sequence recently found in the amphioxus (*Branchiostoma floridae*) genome (Mirabeau & Joly 2013) appears to be a much more credible candidate. This gene encodes a putative peptide that contains a disulfide bridge between two Cys residues and that shares several additional residues with vertebrate SS isoforms, namely the Tyr–Phe–Lys residues in the cyclic region (Fig. 5). However, it should be noted that the putative amphioxus peptide differs from all the known SS by the distance between the two Cys residues.

Evidence of a UII-like peptide has recently been reported in the marine mollusk *Aplysia californica* (Romanova et al. 2012). This peptide was called ApUII. It is 20 amino acid residues in length and shares several structural similarities with UII (Fig. 5). However, the sequence of the peptide is not located at the C-terminus of its precursor. This difference is important, and the question is therefore whether ApUII is an authentic homolog of the vertebrate UII. Indeed, it is worth noting that several other protostomian species have been shown to possess this peptide, but are devoid of any UT-like sequence (O Mirabeau, unpublished observations).

**Receptors of the SS family**

**Background and discovery**

In addition to functional responses to SS1 in living tissues, the existence of specific receptor structures was suggested in the 1970s by binding sites in mammalian tissues. However, the physiological and pharmacological identification of distinct receptor subtypes was hampered by the lack of selective ligands as tools. The complete conservation of the sequence and structure of SS1 throughout vertebrates, as has been described above, meant that evolution had resulted in no peptide variants that could be used to distinguish receptor subtypes. Furthermore, we now know that the different receptor subtypes have very similar signal transduction pathways, thereby further complicating the distinction between them. Studies of signal transduction indicated that SS1 acts via members of the large family of G-protein-coupled receptors (GPCRs; Reisine et al. 1985, Lewis et al. 1986). Binding studies with SS-28 revealed tissue differences in binding properties (Srikant & Patel 1981), and the development of SS analogs allowed the identification of at least two distinct receptor subtypes (Tran et al. 1985). However, similar to many other peptide–receptor systems, it was not until the era of molecular cloning of GPCRs that the SS receptor family in mammals was discovered in its full glory. Already the first report on molecular cloning and functional characterization had described two receptor subtypes in both humans and mice, SSTR1 (STT1) and SSTR2 (STT2) (Yamada et al. 1992a). The discovery of SSTR3 (STT3) ensued later the same year (Yamada et al. 1992b) and the following year that of the mammalian SSTR4 (STT4) and SSTR5 (STT5) receptors was reported (Yamada et al. 1993). Thus, SS1 was rapidly found to have one of the largest receptor families for any neuronal or endocrine peptide in mammals, with as many as five subtypes in all placental mammals that have been investigated in detail. Subsequent genome projects have confirmed that no additional SSTR genes exist in mammals. The nomenclature of the SS receptor family is summarized in Table 1. Henceforth, we will use only the gene nomenclature while referring to them.

**SS receptor properties**

SS1 binds to the five mammalian receptor sequences or elicits responses with very similar Kᵢ or IC₅₀ values in the range of 0.04–2.32 nM (Olias et al. 2004, Epelbaum et al. 2013, Martinez 2013). The variation is probably at least partially due to methodological differences between laboratories. SS1 exhibits the highest affinity for SSTR2. Similarly, SS2 (CST) has values in the range of 0.3–7.0 nM (Olias et al. 2004).

All the five SS receptors decrease adenylyl cyclase activity, reducing the concentration of intracellular cAMP via the pertussis-sensitive G protein family Gᵢ/Gₒ, leading, further down the pathway, to the stimulation of a phosphotyrosine phosphatase (PTP) and the modulation of MAPKs, as well as the stimulation of phospholipase C. Furthermore, some SS receptors activate inwardly
rectifying K⁺ channels and/or inhibit voltage-dependent Ca²⁺ channels. They can also act via pertussis-toxin-independent mechanisms, via the activation of G proteins such as G₁₂ and G₁₄. In some instances, a direct interaction between SS receptors and phosphatases or kinases might occur (Møller et al. 2003, Martinez 2013).

All the five mammalian SSTR genes lack introns in the coding region, except SSTR2, which has two major mRNA variants called SSTR2A and SSTR2B, the latter of which uses an alternatively spliced region encoding a shorter carboxy terminal tail of the receptor (Patel et al. 1993, Vanetti et al. 1993). These two SSTR2 variants are sometimes referred to as separate subtypes, but from a genetic point of view they are more accurately described as splice variants of the same subtype.

The five mammalian receptor subtypes display differential anatomical distributions, but there is also overlap in expression patterns, even at the single-cell level. Briefly, all the five subtypes are expressed in the brain in mammals as well as in the pituitary (Epelbaum et al. 2013). The well-characterized inhibitory effect on GH release in the pituitary is primarily mediated by subtypes SSTR1, SSTR2, and SSTR5. In the gastrointestinal tract, SSTR2 predominates and inhibits secretion and modulates motility, but SSTR1 and SSTR5 are also present and contribute to certain physiological responses (Martinez 2013). All the five subtypes are expressed in pancreatic islets.

Functional studies have been carried out using SS family peptides in a few teleost species (Nelson & Sheridan 2005). As for the mammalian SS receptors, signaling via adenylyl cyclase inhibition has been detected for SSTR1 and SSTR2 in goldfish (Carassius auratus; Lin et al. 1999a, 2000) and SSTR3 in the black ghost knifefish (Apterorhynchus albifrons; Siehler et al. 1999).

SS receptor evolution

The five mammalian receptor subtypes have been sorted into two subfamilies based on their sequence relationships: SSTR1 and SSTR4 are more closely related to each other and SSTR2, SSTR3, and SSTR5 form a separate subfamily (see Møller et al. (2003) and Nelson & Sheridan (2005)). However, the origins of these subtypes remains unclear and conflicting scenarios have been proposed in the literature (for a review, see Ocampo Daza et al. (2012)). In species other than mammals, several receptor sequences have been reported and to some extent characterized, primarily in chickens (Meng et al. 2014) and a few teleost fish species (Nelson & Sheridan 2005, Trainor & Hofmann 2006, Haiyan et al. 2010). Some of these teleosts represent lineages that, in addition to the third tetraploidization in the teleost fish ancestor, also have undergone independent fourth tetraploidizations, namely rainbow trout (Oncorhynchus mykiss) and goldfish (Carassius auratus). Indeed, additional receptor duplicates were found in these species (Nelson & Sheridan 2005, and references therein). The highest number of receptors has been characterized in the goldfish, with eight receptor subtypes. On the other hand, SSTR4 could not be identified in teleosts.

Recently, some of us explored the available vertebrate genome databases for additional sequences to obtain as much information as possible about the SSTR family in each species (Ocampo Daza et al. 2012). Also, we wished to use comparisons of gene synteny to firmly identify orthologous genes and distinguish them from any possible paralogs. To our surprise, the combined sequence data obtained from teleost fish and tetrapods resulted in a phylogenetic tree with six major clades rather than five clades. However, none of the previously investigated vertebrate lineages was found to have all six of these receptor clades. Another unexpected finding was that although all the five mammalian receptor genes are located on separate chromosomes, in teleost fish up to three of these were located on the same chromosome. Our comprehensive phylogenetic analyses showed that teleost fish have an additional receptor clade that is absent in mammals, aside from 3R-generated duplicates or SSTR2, SSTR3, and SSTR5 (Fig. 6). This clade was named SSTR6 and belongs to the SSTR1–SSTR4 subfamily (Fig. 6). The SSTR6 subtype has not been identified in any tetrapod and was presumably lost early in that lineage (Ocampo Daza et al. 2012). Conversely, the SSTR4 subtype seems to be missing in all teleost fish genomes, as well as the genome of the spotted gar (Lepisosteus oculatus), a basal actinopterygian that branched off before the teleost ancestor underwent its third tetraploidization. SSTR4 is present not only in mammals but also in all other tetrapods investigated (Ocampo Daza et al. 2012). This scenario with six ancestral SSTR subtypes was confirmed by analyses of the receptor repertoire in the coelacanth (Latimeria chalumnae), which was found to contain all the six subtypes.

The two rainbow trout genes initially identified as SSTR1A and SSTR1B (Slagter et al. 2004, Hagemeister et al. 2010) could be reclassified as duplicates of SSTR6, presumably resulting from the fourth tetraploidization of the salmonid lineage (Ocampo Daza et al. 2012). This is one of several gene families that have undergone expansion in the salmonid lineage. Another family that is functional in the same regulatory network as SS
The synteny comparisons of several vertebrate species clearly showed that the ancestral vertebrate had two SSTR genes located in different parts of the genome. These two genes were the founders of the two distinct SSTR subfamilies (Fig. 7). The two basal vertebrate tetraploidizations quadrupled each of these two proto-orthologs, whereupon each quartet lost one gene. The remaining six receptors form the two SSTR subfamilies comprising SSTR1, SSTR4, and SSTR6 and SSTR2, SSTR3 and SSTR5 respectively (Fig. 7). This scenario is strongly supported by the detection of several neighboring gene families in the same chromosome regions as the SSTR genes. The synteny of some of the SSTR genes in teleost fish is due to chromosome fusions before the teleost tetraploidization and subsequent gene rearrangements (Fig. 7). This chain of events is supported by recent findings regarding the genome of the spotted gar (Lepisosteus oculatus), because this lineage branched off from the lineage leading to teleosts before the chromosomal rearrangements took place.

In conclusion, the vertebrate ancestor had two SSTR genes, each of which resulted in three surviving copies after the two basal vertebrate tetraploidizations, whereupon the actinopterygian and sarcopterygian lineages lost one receptor each, namely SSTR4 and SSTR6 respectively. To what extent SSTR4 and SSTR6 had evolved distinct functions before they were lost is still unknown. Further studies of each of these receptors in the lineages that still have them are required before this question can be addressed.

### UUI receptors

#### Background and discovery

A few years after the discovery of UII (Pearson et al. 1980), specific receptors were reported in rats based on binding studies as well as functional studies with the radiolabeled peptide that was first discovered in teleosts (as has been described above) (Itoh et al. 1987, 1988). However, it was not until the peptide had been found in humans (Coulouarn et al. 1998) that a receptor was described at the molecular level. In 1999, four laboratories independently reported that a previously identified orphan receptor, named GPR14, responded to human UII (Ames et al. 1999, Liu et al. 1999, Mori et al. 1999, Nothacker et al. 1999). The receptor protein was named UT with the corresponding gene name UTS2R (Table 1). The receptor was already known to display high sequence identity to the SSTR receptors and furthermore its location in the human genome on chromosome 17q23, close to the SSTR2 gene on 17q25, further indicated a relationship by duplication from a common ancestral gene (Tostivint et al. 2006).

#### UUI receptor properties

UII and URP were reported to trigger responses with EC50 values around 1 nM (0.1–3.0 nM), i.e. very similar to the concentration range of the responses elicited by SS1 (Douglas et al. 2004). This signal transduction pathway, which is primarily triggered by human and rat UTS2R, leads via Gq to inositol phosphates, but other G proteins have been found to be activated as well (for a review, see Vaudry et al. (2010)). The human UII receptor is widely expressed both in the nervous system and in peripheral tissues (Douglas & Ohlstein 2000), and the discovery of URP further indicated general or multiple roles for the
To our knowledge, UT receptors have been reported in the literature for only three species other than mammals, namely the European flounder (*Platichthys flesus*; Lu et al. 2006), the orange-spotted grouper (*Epinephelus coioides*; Sun et al. 2013), and the African clawed frog (*Xenopus laevis*; Konno et al. 2013). Similar to the SSTR genes, the *UTS2R* gene in mammals lacks introns in the coding region.

**UII receptor evolution**

After we had resolved the evolution of the vertebrate SSTRs (Ocampo Daza et al. 2012), we wished to explore how the single *UTS2R* reported in the literature is related to the two SSTR subfamilies. Our initial searches in available genome databases revealed a surprising complexity with up to five distinct *UTS2R* gene sequences. These genes represent five ancestral gnathostome *UTS2R* genes. Early in the evolution of tetrapods, the *SSTR6* gene was lost and subsequently there have been differential losses of *UTS2R* genes in different tetrapod lineages, notably massive losses in the mammals. Only the non-avian reptiles have conserved all the five ancestral *UTS2R* genes. The loss of *SSTR4* from the Anole lizard may be a secondary loss, as it can be found in the genomes of other non-avian reptiles (D Ocampo Daza, C Bergqvist and D Larhammar, unpublished observations) and birds. In the teleost lineage, *SSTR4* was lost, while *SSTR2*, *SSTR3*, and *SSTR5* were duplicated in the basal teleost tetraploidization (3R). No new *SSTR1*, *SSTR6*, or *UTS2R* genes emerged through the 3R event. The loss of *UTS2R5* in zebrafish is probably a secondary loss as it can be found in other teleosts. Chromosome fusions and gene rearrangements in the teleost lineage moved some of the SSTR and UTS2R genes to the same chromosomes.

**Figure 7**

Proposed evolutionary scheme for the evolution of the SSTR and UTS2R gene families in gnathostomes. Gene losses are indicated by crossed-out boxes. The ancestral vertebrate had two SSTR proto-orthologs located on different chromosomes; one gave rise to the SSTR1, SSTR4, and SSTR6 genes and another gave rise to the SSTR2, SSTR3, and SSTR5 genes. The ancestral *UTS2R* gene was located in the same chromosomal region as the SSTR2/3/5 ancestor. Thus, it is likely that the ancestral SSTR and UTS2R genes arose through a local duplication before the SSTR duplication that resulted in the SSTR2/3/5 and SSTR1/4/6 ancestors, followed by the translocation of the SSTR1/4/6 ancestor. The two basal vertebrate tetraploidizations (1R and 2R) quadrupled each of the receptor genes, whereupon the two SSTR quartets lost one gene each. In contrast, the UTS2R gene family acquired an additional member through a local duplication following the 2R event, resulting in five ancestral gnathostome UTS2R genes. Early in the evolution of tetrapods, the SSTR6 gene was lost and subsequently there have been differential losses of UTS2R genes in different tetrapod lineages, notably massive losses in the mammals. Only the non-avian reptiles have conserved all the five ancestral UTS2R genes. The loss of SSTR4 from the Anole lizard may be a secondary loss, as it can be found in the genomes of other non-avian reptiles (D Ocampo Daza, C Bergqvist and D Larhammar, unpublished observations) and birds. In the teleost lineage, SSTR4 was lost, while SSTR2, SSTR3, and SSTR5 were duplicated in the basal teleost tetraploidization (3R). No new SSTR1, SSTR6, or UTS2R genes emerged through the 3R event. The loss of UTS2R5 in zebrafish is probably a secondary loss as it can be found in other teleosts. Chromosome fusions and gene rearrangements in the teleost lineage moved some of the SSTR and UTS2R genes to the same chromosomes.
UT receptor subtypes, numbered UTS2R1 through UTS2R5; however, phylogenetic analyses indicated complex relationships among them, including many losses (Fig. 8). For instance, only non-avian reptile lineages have preserved all the five subtypes (Fig. 8), namely the Carolina anole lizard (Anolis carolinensis), the American alligator (Alligator mississippiensis), and the Western painted turtle (Chrysemys picta bellii). Furthermore, the single UTS2R sequence found in the opossum clearly belongs to a different subtype (UTS2R2) compared with the UT receptors (UTS2R1) of placental mammals. All family members that were identified lack introns in the coding region.

To resolve orthology–paralogy relationships for the many UTS2R genes and to deduce the time points of the gene duplications, we analyzed the chromosomal regions for all the UTS2R family members in several sequenced vertebrate genomes, finding that four of the UTS2R genes display synteny with the SSTR2, SSTR3, and SSTR5 genes, as shown in Fig. 9. The UTS2R gene that we named UTS2R4 is located on the fourth chromosome in this paralogon, i.e. the chromosome that has lost its SSTR member. A single UTS2R candidate gene was found in the amphioxus (Branchiostoma floridae) genome, an invertebrate chordate. Thus, we conclude that four of the five UTS2R genes arose from a common ancestral gene as a result of the two basal vertebrate tetraploidizations (D Ocampo Daza, C Bergqvist and D Larhammar, unpublished observations; Fig. 7).

The fifth UTS2R gene, of the subtype that we named UTS2R5, is co-localized with UTS2R1 (and SSTR2), presumably as a result of a local gene duplication (Fig. 7). However, it is difficult to say with certainty when this local duplication had taken place in relation to the two tetraploidizations. The most parsimonious scenario would be that the local duplication happened shortly after the 2R event (Fig. 7), but it remains possible that the local duplication happened earlier and that any 2R duplicate(s) of UTS2R5 have been lost.

The synteny between UTS2R and SSTR genes also means that the UTS2R genes went through the same complicated chromosomal rearrangements in actinopterygians as the SSTRs belonging to the SSTR2, SSTR3, and SSTR5 subfamily (see above; Fig. 7). While all the five UTS2R subtypes are represented among teleosts as a group, no single teleost species has yet been found to possess more than four UTS2R subtypes. Thus, differential losses have occurred, similar to that among tetrapods. It should be noted that no single UTS2R duplicate resulting from the 3R event seems to have survived in the teleosts investigated so far (Fig. 7).

**Functional considerations**

**SS-ergic systems**

Most of the current knowledge regarding the peptides of the SS family concerns SS1 and, to a lesser extent, SS2. The other SS-related peptides, which occur mainly in teleosts, have been studied to a much lesser degree. All these peptides exert their actions through specific receptors, which belong to the GPCR superfamily. As has been described above, the SS receptor family in vertebrates encompasses six distinct subtypes, SSTR1–SSTR6.

**SS1 actions** SS1-producing cells are mainly present in the CNS and peripheral nervous system, in the endocrine
pancreas, and in the gut. However, a number of additional expression sites have been described in the placenta, the kidney, and the retina (Reichlin 1983, Patel 1999). In relation to this widespread distribution, SS1 is known to exert a very large spectrum of biological activities that have been investigated to a large extent in mammals, mostly in humans and rats/mice. SS1 is primarily involved in the control of the secretory activity and proliferation of endocrine and exocrine secretions from the pituitary, pancreas, and gut (Patel 1999). The lowering of intracellular cAMP and Ca$^{2+}$ concentrations is the most relevant mechanism for the inhibition of hormone secretion from endocrine cells, whereas the activation of PTP and MAPKs is postulated to play a role in the antiproliferative effect on tumor cells (Olias et al. 2004).

At the level of the CNS, SS is known to be a neuromodulator of many processes including sleep, sensory and motor activities, and cognitive functions (Martel et al. 2012).

Hypothalamic SS1 plays a key role in the control of pituitary hormone release. It has been proven to be the primary inhibitor of GH release in vertebrates, and this function has been well conserved during the evolution of vertebrates. In mammals, birds, and teleosts, SS1 can act on the basal secretion as well as the GHRH-induced secretion of GH (Patel 1999, Gahete et al. 2010). The selective role of each SSTR subtype in SS1 modulation has no apparent effect on GH release by itself, but it enhances the release of GH stimulated by TRH, the main GH release stimulatory factor in these groups (Gahete et al. 2010). Surprisingly, despite the important role of SS1 in the control of GH secretion, Ss1-null mice do not exhibit excessive growth (Zeyda & Hochgeschwender 2008).

The selective role of each SSTR subtype in SS1 inhibition of GH secretion is not yet fully clarified. In mammals, SS1 is thought to act through SSTR2, SSTR5, and to some extent SSTR1 (Ben-Shlomo & Melmed 2010). Studies in goldfish indicate that SSTR2 may also play an important role in the inhibition of GH secretion in teleost fish (Wong et al. 2006). It is worth noting that SS1 has also been shown to inhibit GH secretion indirectly via the modulation of hypothalamic GHRH release (Zheng et al. 1997), a process in which primarily SSTR1 and SSTR2 seem to be involved (Olias et al. 2004).

In addition to that of GH, SS1 also inhibits the secretion of other pituitary hormones such as thyroid-stimulating hormone (TSH) and adrenocorticotropic hormone (ACTH) through SSTR2 and SSTR5 and prolactin through SSTR2 (Ben-Shlomo & Melmed 2010, Gahete et al. 2010). Although they have mainly been studied in mammals, some of these actions have also been reported in other vertebrates (for a review, see Gahete et al. (2010)). For instance, SS1 inhibits prolactin release in the Mozambique tilapia (Oreochromis mossambicus; Grau et al. 1982) and PACAP38-induced TSH secretion in the American bullfrog (Lithobates catesbeiana, previously Rana catesbeiana) (Okada et al. 2006). Surprisingly, in the trout (Salmo trutta), SS1 seems to exert a stimulatory effect on ACTH secretion (Langhorne 1986).

SS1 has long been known to be an important modulator of pancreatic secretions. In mammals, it has been shown that SS-14 and SS-28 differentially target pancreatic α- or β-cells. In accordance with this, SS-14, which is locally produced by δ-cells, has been reported to be associated with the inhibition of glucagon secretion, while circulating SS-28 seems to primarily inhibit insulin secretion (Strowski & Blake 2008). In rodents, glucagon inhibition is mediated by SSTR2, while insulin inhibition involves SSTR5. In humans, SSTR2 appears to be the main receptor mediating the effects of SS on both insulin and glucagon release, although SSTR1 and SSTR5 also participate in the regulation of insulin secretion. In teleosts, SS1 injection has been reported to reduce plasma glucagon concentrations but not insulin concentrations (Eilertson & Sheridan 1993). The specific receptor subtypes mediating these effects are unknown.
SS1 also mediates a broad range of physiological actions in the gastrointestinal tract. In mammals, SS1 is known to inhibit the release of many gastrointestinal peptides, including cholecystokinin, enteroglucagon, gastric inhibitory peptide, ghrelin, secretin, and vasoactive intestinal polypeptide. It also inhibits numerous gastrointestinal functions such as bowel motility, gastric emptying, gastrointestinal tract transit time, small bowel segmentation, gallbladder contractility, and bowel flow, as well as splanchnic and liver blood flow (Martinez 2013). Finally, SS1 prevents the intestinal absorption of carbohydrates, amino acids, calcium, and triacylglycerols (Gahete et al. 2010). Most of these effects seem to be mediated by SSTR2. The digestive functions of SS1 outside of mammals are very poorly documented. In the Atlantic cod (Gadus morhua), SS1 has been shown to reduce gastric acid secretion (Holstein & Cederberg 1988), while in the Asian bullfrog Hoplobatrachus tigerinus (previously Rana tigrina), it seems to inhibit pepsinogen secretion (Hong et al. 1988).

SS1 is widely expressed in the CNS. It is mostly found in non-glutamatergic cells where it generally co-localizes with γ-aminobutyric acid (GABA), the major inhibitory neurotransmitter in the brain (Martel et al. 2012). SS1 functions as a neuromodulator that mainly acts at the presynaptic level. It inhibits glutamate release presynaptically by acting through SSTR1, SSTR2, or SSTR5, depending on the anatomical structure, thus leading to the activation of potassium leak current and inhibition of voltage-dependent calcium currents. SS1 also exhibits long-lasting effects on glutamatergic synapses by modulating long-term potentiation (Viollet et al. 2008). In mammals, i.c.v. administration of SS1 increases locomotor activity through the activation of SSTR2 and SSTR4, whereas its depletion has been reported to induce learning and memory deficits, mediated by SSTR2 (Viollet et al. 2008).

**SS2/CST actions** SS2 (including its mammalian counterpart CST) is expressed in a few discrete areas in the brain (Tostivint et al. 1996, de Lecea et al. 1997b, Trabucchi et al. 2003, Canosa et al. 2004). For instance, in mammals, CST mRNA can be detected in GABAergic neurons of the cortex and hippocampus that partially overlap with SS1 mRNA (de Lecea et al. 1997b). Subsequent studies have also reported the occurrence of CST in immune cells (Dalm et al. 2004). As expected from their structural similarities, CST can reproduce several of the effects of SS1, such as the inhibition of GH and insulin release and depression of neuronal activity (de Lecea et al. 1996, Gahete et al. 2010, de Lecea 2008). These observations are consistent with the fact that CST and SS1 exhibit a similar binding affinity for the five SSTR subtypes (Siehler et al. 1998). However, there is also evidence that CST and SS1 may produce distinct effects, both in the brain, as initially described (de Lecea et al. 1996, de Lecea 2008), and in the periphery, especially in the immune system (Martin van Hagen et al. 2008). For example, intracerebral injection of CST induces slow-wave sleep and leads to a reduction of locomotor activity, both opposite effects to those observed with SS1. Moreover, CST, but not SS1, has been demonstrated to have potent anti-inflammatory activity (Dalm et al. 2003).

Structure–activity relationship studies have indicated that the Pro moiety at position 2 and the C-terminal Lys residue are crucial for the distinctive effects of CST in the brain (Criado et al. 1999). Based on these data, the existence of specific CST receptors able to bind selectively to CST but not to SS1 has been postulated. Two receptors identified during the last decade appear to fulfill these criteria, namely the ghrelin receptor (Deghenghi et al. 2001) and the MrgX2 receptor (Robas et al. 2003). However, the ability of CST to activate them under physiological conditions remains to be demonstrated (Siehler et al. 2008).

In goldfish (Carassius auratus) as in chickens, SS2 has been demonstrated to inhibit both basal and GHRH-induced secretion of GH (Yunker et al. 2003, Meng et al. 2014). In the American bullfrog (Lithobates catesbeianus), SS2, similar to SS1, only has an effect on GHRH-stimulated GH secretion (Jeandel et al. 1998).

**Actions of other SS-related peptides** SS3 is equipotent to SS1 in inhibiting GH release in goldfish, while SS4 has no apparent effect in catfish (Marchant et al. 1987). In rainbow trout (Oncorhynchus mykiss), it has been shown that the injection of SS3, but not of SS1, reduces plasma insulin concentrations (Eilerston & Sheridan 1993). These distinct actions may, in part, result from differential ligand–SSTR interactions (Klein & Sheridan 2008). For instance, the SSTR6 receptors initially identified as SSTR1A and SSTR1B (see the ‘SS receptor evolution’ section) preferentially bind to PPSS1 products rather than to PPSS3-derived peptides (Gong et al. 2004).

**Ull-ergic systems**

**Actions of UII and URP** UII was first characterized in teleost fish on the basis of its spasmogenic properties more than 30 years ago (Pearson et al. 1980, Bern et al. 1985). The research in this field became more intense when the human isoform was identified (Coulouarn et al. 1998); however, the physiological significance of the UII-ergic
The complexity of the UII system is further emphasized by the discoveries of multiple URPs (Nobata et al. 2011, Parmentier et al. 2011, Tostivint et al. 2013) and the five receptor subtypes that arose in early vertebrate evolution (as has been described above). Surprisingly, URP has received little attention although it was discovered soon after UII. This is probably because it shares the same ability as UII to bind to and activate recombinant UTS2R1 (Sugo et al. 2003, Dubessy et al. 2008). Thus, its effects are discussed along with those of UII in this review. Even less is known about the binding properties of URP1 and URP2 (Nobata et al. 2011, Parmentier et al. 2011). All functional data relate to UTS2R1, the single UTS2R receptor present in placental mammals, as the other UTS2R isoforms have only been obtained recently. UTS2R1 was originally described as being coupled to Gq, but other signal transduction pathways involving Gs and Gi have subsequently been identified, associated with multiple signaling cascades (Ross et al. 2010).

The highest levels of expression of the UII gene are found in the CNS, specifically in motor neurons (Chartrel et al. 1998, Coulouarn et al. 1998, 1999, Dun et al. 2001, Pelletier et al. 2002, 2005, Egginger et al. 2006). However, the bulk of all research into the UII-ergic system has mainly focused on the peripheral effects as the receptor is expressed predominantly in the periphery, in the cardiovascular system (Tal et al. 1995, Ames et al. 1999, Matsushima et al. 2001, Elshourbagy et al. 2002). The UII and/or URP and UTS2R genes are expressed in the vasculature and in the heart in humans and other primates as well as in rodents (Ames et al. 1999, Matsushima et al. 2001, Totsume et al. 2001, Elshourbagy et al. 2002). In line with the strong expression in the cardiovascular system, UII is considered to be the most potent vasoconstrictor identified to date (Ames et al. 1999). However, the effects of UII clearly depend on the nature of the vascular bed location and type (Douglas et al. 2000). These variations are attributable to the fact that UTS2R1 mRNA is present on both smooth muscle cells and endothelium. Consequently, it may act as a potent endothelium-independent vasoconstrictor on smooth muscle cells in dogfish (Scylliorhinus canicula; Hazon et al. 1993), rainbow trout (Onchorhynchus mykiss; Le Mevel et al. 1996), American bullfrog (Lithobates catesbeiana; Yano et al. 1995), rat (Itoh et al. 1988, Douglas et al. 2000), and human arteries (Maguire et al. 2000, Hillier et al. 2001). The response is limited to large elastic vessels and does not include small resistance vessels. On the other side, UII can also promote vasodilation via an endothelium-dependent nitric oxide (NO)-mediated response (Bottrill et al. 2000, Stirrat et al. 2001). It is difficult to describe the effects of UII in vivo as they have not been conserved during evolution. For instance, depending on the species studied, peripheral administration of UII and/or URP promotes either hypotensive or hypertensive cardiovascular responses, or it can be totally ineffective, as in humans (Yano et al. 1995, Conlon et al. 1996, Ames et al. 1999, Wilkinson et al. 2002, Sugo et al. 2003, Le Mevel et al. 2008). In this regard, URP was reported to be equipotent to UII in rats, but tenfold less potent than UII in rainbow trout, and the duration of its response was obviously shorter (Sugo et al. 2003, Le Mevel et al. 2008). Central administration of UII has yielded more constant results, because in most species examined it evokes an increase in mean arterial pressure and heart rate through the activation of the peripheral sympathetic nervous system (Watson & May 2004, Le Mevel et al. 2008). UTS2R1 mRNA can be detected in brainstem nuclei controlling cardiovascular homeostasis and in hypothalamic nuclei such as the paraventricular nucleus, which supports its role as a modulator of cardiovascular function (Jegou et al. 2006). The URP mRNA is highly expressed in neurons that are intermingled with tyrosine hydroxylase-immunoreactive A1/C1 cells in the rostral ventrolateral medulla, indicating that URP may directly control adrenal sympathetic outflow. Furthermore, c-Fos-immunoreactivity is induced in the nucleus of the solitary tract and the central nucleus of the amygdala following the central administration of UII (Watson et al. 2008). In vitro, UII has a potent inotropic effect on isolated perfused human heart (Russell et al. 2001, Quail et al. 2009). In addition, UII functions as an autocrine/paracrine factor mediating cardiac and vasculature remodeling (Sauzeau et al. 2001, Tzanidis et al. 2003).

The urophysis is the main source of UII in teleosts (Bern et al. 1985), which indicates the involvement of UII in osmoregulatory and renal functions. The plasma concentrations of UII fluctuate in the flounder (Platichthys flesus) after adaptation from fresh water to sea water and vice versa (Bond et al. 2002, McCrohan et al. 2007, Lu et al. 2008). The activity of Dahlgren cells is modulated by osmotic stress, and urophysectomy has been shown to impair adaptation to water salinity (Chan & Bern 1976), which further supports the role of UII in osmoregulation. Radiolabeled UII-binding sites were detected in goldfish tissues (kidney, gill, and intestine) involved in osmoregulation (Woo et al. 1985), and it is clear that UII alters water and ion fluxes across a range of osmoregulatory
epithelia, resulting in net absorption in a manner dependent on the acclimation to salinity (Lu et al. 2008).

The possible renal actions of UII in mammals have been amply addressed, mainly because UII was first defined in teleosts as a peptide involved in osmoregulation. The human kidney expresses UII in tuberoepithelial cells, collecting ducts, distal convoluted tubules, and endothelial cells, and UII immunostaining has been demonstrated in rat proximal tubules and inner and outer medullary collecting ducts (Song et al. 2006). The plasma concentrations of UII in humans and rats are lower than the urinary levels, indicating that UII is produced from the kidney and that this organ could represent the most likely source of plasma UII (Matsushita et al. 2001, Song et al. 2006). The role of UII in the kidney is not clear at present. It has been shown to promote endothelial NO-mediated vasodilation in rat renal arteries accompanied by a dose-dependent increase in renal blood flow and an increased urinary water and sodium excretion, but opposite effects have also been reported (Zhang et al. 2003, Ashton 2006, Song et al. 2006). It has been suggested that these discrepancies might result from the limited amounts of unoccupied receptors for binding exogenous UII (Qi et al. 2005). An inhibitory action of UII on renal function is supported by experiments using a UII antagonist called urantide, showing a negative tonic influence of UII (Ashton 2006, Song et al. 2006).

It is likely that circulating UII functions as an endocrine/neuroendocrine factor as UII is present in the circulation, albeit at low concentrations (22 pg/ml in humans; Thompson et al. 2003). Consistently, UII and/or URP are expressed in a number of endocrine tissues, including the pituitary, adrenal gland, pancreas, ovary, and testis (Coulouarn et al. 1998, Sugo et al. 2003, Dubessy et al. 2008). The endocrine effects of the UII-ergic system have been poorly documented to date. UII has been shown to promote the inhibition of prolactin release by the isolated pars distalis of Mozambique tilapia (Oreochromis mossambicus; Rivas et al. 1986). UII also inhibits corticosterone production in dispersed rat adrenocortical cells (Albertin et al. 2006), while it stimulates the secretion of cortisol in European flounder (Platichthys flesus; Kelsall & Balment 1998). Both the stimulation of corticosterone production and the absence of effects have been reported in the amphibians marsh frog (Pelophylax ridibundus) and axolotl (Ambystoma mexicanum; Feuilloley et al. 1994, Gupta & Hanke 1994). Metabolic effects have also been described, including the inhibition of insulin’s response to glucose in isolated rat perfused pancreas with no effect on glucagon and SS1 release (Silvestre et al. 2001, Marco et al. 2008). This suggests a direct effect on β-cells. A UII-mediated inhibition of insulin response has also been described in coho salmon (Oncorhynchus kisutch; Sheridan et al. 2007). UII also acts as a neurohormone. It is produced by the caudal neurosecretory system in teleosts, for instance, and we have already discussed its neuroendocrine osmoregulatory function. In rats, UII regulates the pituitary–sympathoadrenal medullary and pituitary adrenal cortex pathways (Watson & May 2004, Watson et al. 2008). The UII and UTS2R1 mRNAs are co-expressed in the rat pheochromocytoma cell line PC12 and UII increases dopamine secretion (Aita et al. 2010). Centrally, UII controls plasma prolactin and TSH concentrations (Gartlon et al. 2001). Double-immunofluorescence labeling in mouse brain revealed a co-localization of URP and GNRH in fibers in the median eminence and vascular organ of the lamina terminalis as well as in cell bodies in the preoptic area, suggesting that URP is co-released with GNRH (Egginger et al. 2011). In the orange-spotted grouper (Epinephelus coioides), UII may promote growth through the activation of GHRH expression and inhibition of SS1 and SS2 concentrations in the hypothalamus, which stimulates GH expression in the pituitary and increases GH receptor expression (Sun et al. 2013).

A number of behavioral effects of UII have been reported (Do Rego et al. 2008), in line with the broad distribution of UTS2R1 mRNA and radiolabeled UII-binding sites in the CNS (Jegou et al. 2006). It is well established that UII is involved in various mood and sleep processes. Thus, i.c.v. administration of UII promotes anxiogenic and depressant-like effects in mice (Matsumoto et al. 2004, Do Rego et al. 2005) and enhances neural activity in rat peri-aqueductal gray matter (Gartlon et al. 2004), a central region involved in anxiety control. In addition, targeted administration of UII into the pedunculopontine nucleus modulates paradoxical sleep through the direct activation of brainstem cholinergic neurons (Huitron-Resendiz et al. 2005, de Lecea & Bourgin 2008). It has been obvious for many years, ever since the discovery of UII in mammals, that UII and URP are involved in the control of motor functions, although their exact roles are not fully understood. A high density of UII- and URP-binding sites has been found in regions of the human and rat brains that are involved in motor control (Maguire et al. 2000, Jegou et al. 2006, Bucharles et al. 2014), and i.c.v. administered UII and/or URP increases locomotion in mice, rats, and rainbow trout (Gartlon et al. 2001, Lancien et al. 2004, Do Rego et al. 2005). The UII and URP mRNAs are abundantly expressed in the CNS, especially in the adult cranial and spinal motoneurons of many species.
(Chartrel et al. 1998, Coulouarn et al. 1998, 1999, Dun et al. 2001, Pelletier et al. 2002, 2005, Egginger et al. 2006). Notably, UII mRNA is detected early in the formation of the rat embryonic ventral neural tube, at a developmental stage coincident with the differentiation of motor neurons (Coulouarn et al. 2001). Recent studies indicate that this motoneuronal expression pattern is conserved for both UII and URP in the African clawed frog (Xenopus laevis; Konno et al. 2013). The presence of UII receptor on cholinergic nerve terminals, in part originating from the recurrent motoneuron axon collaterals, in mouse spinal cord has given rise to the idea that UII may regulate motoneuron activity in an autocrine/paracrine manner (Bruzzone et al. 2010). Supporting this view, functional UII receptors have been found to be coupled to calcium mobilization in dissociated rat motor neurons (Filipeanu et al. 2002). A presynaptic role at the neuromuscular junction is also likely, as UII immunoreactivity is detected in muscle motor end plate regions coexisting with synaptophysin immunolabeling (Dubessy et al. 2008) and because UII promotes acetylcholine release (Brailoiu et al. 2003, Bruzzone et al. 2010). Finally, UTS2R1 mRNA and UII-binding sites are highly expressed in skeletal muscle (Maguire et al. 2000, Elshourbagy et al. 2002, Dubessy et al. 2008) and functional receptors exist in a human skeletal muscle cell line (Douglas et al. 2004), indicating a possible post-synaptic function. Notably, UII/URP expression is likely to be sexually dimorphic as androgen receptors are co-localized with UII and URP mRNAs in motor neurons, while castration downregulates the expression of both genes (Pelletier et al. 2002, 2005).

### Concluding remarks

Taken together, the discoveries of many additional family members for SS peptides and receptors, as well as UII peptides and receptors, indicate an astounding complexity that was completely unexpected only a few years ago. Almost all the gene duplications can be explained by tetraploidizations that took place in the vertebrate ancestor and teleost fish ancestor. The implication of these findings is that the vertebrate ancestor, before the two tetraploidizations, possessed three receptors for SS/UII-like peptides: one SSTR1/4/6 ancestor, one SSTR2/3/5 ancestor, and one UTS2R ancestor. This setup of ancestral pre-vertebrate receptors encourages further studies of invertebrate deuterostomes to explore their peptide and receptor repertoires and their biological functions.

The multiplicity of components leads to expectations of many additional roles for these systems not only within each species, but also across the range of vertebrates. In particular, the unprecedented repertoire of four or five ancient UII receptors in some vertebrate lineages and five UII family peptides in teleost fish provides intriguing areas of research. It will also be very interesting to learn more about the roles of UII, URP, and SS2 in mammals. The evolutionary dynamics of the SS and UII systems make them both exciting and challenging to explore.

### Declaration of interest

The authors declare that there is no conflict of interest that could be perceived as prejudicing the impartiality of the review.

### Funding

This review was supported by the French National Centre for Scientific Research (CNRS UMR 7221) and The Swedish Research Council.

### Acknowledgements

The authors thank Dr Olivier Mirabeau (Institut Curie, France) for helpful discussions.

### References


Dalm VA, van Hagen PM, van Koetsveld PM, Langerak AW, van der Lely AJ, Lambertsw SW & Holland IJ 2003 Cortistatin rather than somatostatin as...


Quan FB, Kenigfest NR, Mazan S & Tostivint H 2013 Molecular cloning of the cDNAs encoding three somatostatin variants in the dogfish (*Scyllorhinus canicula*). *General and Comparative Endocrinology* 180 1–6. (doi:10.1016/j.ygcen.2012.10.007)


Thematic Review

Somatostatin and urotensin II system evolution

52:3 | T85


Van de Peer Y, Maere S & Meyer A 2010 2R or not 2R is not the question anymore. Nature Reviews. Genetics 11 166. (doi:10.1038/nrg2600-c2)

Vanetti M, Vogt G & Höltt V 1993 The two isoforms of the mouse somatostatin receptor (mSSTR2A and mSSTR2B) differ in coupling efficiency to adenylyl cyclase and in agonist-induced receptor desensitization. FBS Letters 331 260–266. (doi:10.1016/0014-5793(93)80349-Y)


Received in final form 14 March 2014
Accepted 10 April 2014
Accepted Preprint published online 16 April 2014