J Mol Evol
DOI 10.1007/s00239-007-9055-3

The Lengthening of a Giant Protein: When, How, and Why?

Olivier Meiniel · Robert Meiniel · Fabrice Lalloué · Robert Didier · Marie-Odile Jauberteau · Annie Meiniel · Daniel Petit

Received: 2 April 2007 / Accepted: 7 November 2007

Abstract Subcommissural organ (SCO)-spondin is a giant glycoprotein of more than 5000 amino acids found in Vertebrata, expressed in the central nervous system and constitutive of Reissner's fiber. For the first time, in situ hybridization performed on zebrafish (*Danio rerio*) embryos shows that the gene encoding this protein is expressed transitionally in the floor plate, the ventral midline of the neural tube, and later in the diencephalic third ventricle roof, the SCO. The modular organization of the protein in Echinodermata (*Strongylocentrotus purpuratus*), Urochordata (*Ciona savignyi* and *C. intestinalis*), and Vertebrata (Teleostei, Amphibia, Aves and Mammalia) is also described. As the thrombospondin type 1 repeat motifs represent an increasingly large part of the protein during Deuterostomia evolution, the duplication mechanisms leading to this complex organization are examined. The functional significance of the particularly well-preserved arrangement of the series of SCO-spondin repeat motifs and thombospondin type 1 repeats is discussed.

Keywords SCO-spondin · Central nervous system · Thrombospondin type 1 repeat · SCO-spondin repeat · Deuterostomia

Electronic Supplementary Material The online version of this article (doi:10.1007/s00239-007-9055-3) contains supplementary material, which is available to authorized users.

O. Meiniel · R. Meiniel · R. Didier · A. Meiniel
Faculté de Médecine, INSERM, UMR 384, 28 place Henri Dunant, 63001 Clermont-Ferrand cedex, France

F. Lalloué · M.-O. Jauberteau
Faculty of Medicine, EA 3842, 2 rue du Docteur Marcland, 87025 Limoges cedex, France

D. Petit (✉)
INRA, UMR 1061, Université de Limoges, 123 avenue Albert Thomas, 87060 Limoges cedex, France
e-mail: daniel.petit@unilim.fr

Introduction

The body plan organization in Chordata is characterized by bilateral muscles, pharyngeal gill slits, and the coexistence of two dorsal structures: a rigid notochord and a hollow neural tube or nerve cord. The ventral midline of the neural tube, called the floor plate (FP), plays a crucial role during ontogenesis in cell fate speciation and axonal pathfinding (Charron and Tessier-Lavigne 2005). Another important permanent structure of this phylum is Reissner's fiber (RF) identified by Reissner (1860) in *Petromyzon fluviatilis*. In Vertebrata, the RF is secreted in the cavity of the neural tube and extends from the edge of the subcommissural organ (SCO), through the Sylvian aqueduct, the fourth ventricle, and along the central canal of the spinal cord to the *ampulla caudalis*. The exact function of the RF remains poorly understood. Immunological approaches using antibodies directed against bovine RF have shown that RF components are synthesized by both the FP and the SCO during ontogenesis (Naumann et al. [1987] in various Vertebrata, Lopez-Avalos et al. [1997] in *Scyliorhinus canicula*, Schoebitz et al. [1986] in *Gallus gallus*, Rodriguez et al. [1996] in *Rattus norvegicus*) and only by the SCO in the adult (see Sterba et al. [1982] and Rodriguez et al. [1984] for comparative evolutionary studies). The RF has been characterized in all chordate species studied but, to our knowledge, not yet in humans (Oksche 1961; Olsson 1961; Rodriguez et al. 1992). The main component of the RF has been identified in the bovine as a giant

glycoprotein of 5146 amino acids, SCO-spondin (Gobron et al. 1996, 2000). Southern blot analyses performed with a bovine probe revealed that the gene encoding SCO-spondin was present in the chordate phylum, i.e., Urochordata, Cephalochordata, and Craniata including humans (Gobron et al. 1999). This raises the question of the role of SCO-spondin in this phylum and whether it occurs, at least partly, in other organisms in Deuterostomia and possibly Protostomia.

This protein displays a unique arrangement of several conserved domains, including von Willebrand factor type D (vWD) domains, SCO-spondin repeat (SCOR) motifs, low-density lipoprotein receptor type A (LDLrA) domains, thrombospondin type 1 repeat (TSR) motifs, and a C-terminal cystin knot (CTCK) domain (for a review see Meiniel and Meiniel 2007). Further, Mammalia show a highly conserved series of 26 TSRs spanning most of the protein sequence, suggesting that these motifs play a crucial role in SCO-spondin function (Meiniel and Meiniel 2007). This is supported by results of in vitro experiments performed on neuronal cell cultures, showing that a highly conserved fragment of a TSR module can alone promote cell differentiation, and mainly neurite outgrowth (Monnerie et al. 1998). From the sequence analysis of TSR motifs in Mammalia, three groups were defined, differing in the number and position of conserved cysteine residues (Meiniel and Meiniel 2007).

The first aim of this work was to describe where and when the gene encoding SCO-spondin is expressed during the development of a vertebrate model, the zebrafish, and whether this pattern fits the sites of RF production in Chordata. We demonstrate here that this gene is transitionally expressed in the FP during ontogenesis. This transitional expression pattern observed during CNS ontogenesis of zebrafish leads us to consider a putative relationship between development and evolution of SCO-spondin in the chordate phylum.

The second aim of this work was to analyze the mosaic organization of SCO-spondin in Deuterostomia. This was to give insight into the variations in the number and composition of TSRs in vertebrate SCO-spondin, by comparison with the situation in Ascidia, and by extension in the sea urchin *Strongylocentrotus purpuratus*. Our purpose was to perform a phylogenetic analysis of all the TSRs of this protein, in order to decipher the duplication and deletion events leading to this complex modular structure. However, given the low information content of these short-length sequences (average length, 55 amino acids; range, 43–91), the arrangement of the different groups of TSRs interspersed with SCORs plays a signature role that can help in assessing the TSR orthology relationships among Vertebrata, Urochordata, and Echinodermata. The probable explanation for the conservation of this arrangement is discussed in the context of central nervous system (CNS) evolution.

Materials and Methods

In Situ Hybridization

We first generated, by reverse transcriptase-polymerase chain reaction (RT-PCR), full-length coding desoxyribonucleic acid (cDNA) from zebrafish 24 h post-fertilization (hpf)-stage total ribonucleic acids (RNA). A 648-base pair (bp) SCO-spondin cDNA fragment was then amplified using the following primers: 5′-TTTGCAGTCTGGAT GGTGAAGACC-3′ (forward) and 5′-cgctccagctcataca ggtggagg-3′ (reverse).

The denaturing was performed at 94°C for 30 s, annealing at 60°C for 30 s, and elongation at 72°C for 45 s; 30 cycles were used. The polymerase chain reaction (PCR) product was purified and ligated directly into a T-tailed pGEM vector (Promega, Madison, WI, USA).

Whole-mount in situ hybridization experiments were performed as described (see http://zfin.org/zf_info/zfbook/chapt9/9.82.html). Digoxigenin-labeled antisense RNA probes were synthesized from a *Spe*I-linearized DNA template using T7 RNA polymerase (Roche Diagnostics, Mannheim, Germany). Sense and antisense probes raised against SCO-spondin messenger RNA (mRNA) were used to perform in situ hybridization on 5 hpf to 5-days-post-fertilization (dpf) zebrafish embryos.

Protein Sequences

A full search of GenBank (http://www.ncbi.nlm.nih.gov/Genbank/) and Ensembl (http://www.ensembl.org/index.html) using the bovine mRNA sequence (Gobron et al. 2000) as a probe revealed the counterpart of the gene in the genomes of *Strongylocentrotus purpuratus*, *Ciona intestinalis*, *Ciona savignyi*, *Tetraodon nigroviridis*, *Danio rerio*, and *Silurana tropicalis*. The proteins were predicted from these genomic sequences using Wise2 software (http://www.sanger.ac.uk/Software/Wise2/), taking advantage of the experimentally established bovine sequence. Sequences in *Gallus gallus*, *Homo sapiens*, *Rattus norvegicus*, *Mus musculus*, and *Canis familiaris* have already been published (Didier et al. 2007; Meiniel and Meiniel 2007). The accession numbers of all the sequences studied are given in Table 1. We note that the gene symbols differ from one species to another: SSPO at HGNC (http://www.gene.ucl.ac.uk/nomenclature/), Sspo at MGI

(http://www.informatics.jax.org/), and scospondin at ZFIN (http://zfin.org/).

Conserved Domains

The search for conserved motifs was performed directly on the SCO-spondin sequences using SMART sequence analysis software (http://smart.embl-heidelberg.de/). The length of each conserved motif is reported in Figs. 2 and 3. Information on conserved protein motifs was found under the description headings of the following databases: SMART, PFAM (http://pfam.cgb.ki.se/index.html), PROSITE (http://www.expasy.org/prosite/), MEROPS (http://merops.sanger.ac.uk/), and InterPro (http://www.ebi.ac.uk/interpro/index.html).

Alignments and Phylogeny

A total of 195 TSR motifs were found in SCO-spondin proteins in the 11 species analyzed and aligned with ClustalX 1.83 (Thompson et al. 1997; Jeanmougin et al. 1998). The small size of the TSRs, average length 55 amino acids, results in a small number of informative sites, a serious disadvantage for obtaining high bootstrap values. We therefore used two phylogenetic and molecular evolutionary analyses: minimum evolution using MEGA version 3.1 (Kumar et al. 2004) and maximum likelihood using PhyML (Guindon and Gascuel 2003). In both cases, the JTT model of amino acid changes was chosen. Robustness of branches was assessed with percentages calculated from 500 bootstrap replicates.

Table 1 Accession numbers of studied SCO-spondin protein sequences, experimentally determined or bioinformatically predicted

Sequence	Accession no.
Bos taurus (experimental)	CAC94914.1
Homo sapiens (predicted)	BN000852
Rattus norvegicus (predicted)	CAF33425.1
Mus musculus (predicted, modified)	CAD42654.1
Canis familiaris (predicted)	BN000732
Gallus gallus (experimental)	CAI29216
Silurana tropicalis (predicted from Ensembl scaffold_1041)	Unedited
Tetraodon nigroviridis (predicted, modified)	BN000853
Danio rerio (predicted; supported by AM159048 and AM158248)	BN000851
Ciona savignyi (predicted)	BN001016
Ciona intestinalis (predicted)	BN001015
Strongylocentrotus purpuratus (predicted, modified)	BN001014

Results

Early Development Expression Pattern of SCO-Spondin in Zebrafish

In situ hybridization experiments were performed on *Danio rerio* embryos using a SCO-spondin specific riboprobe. SCO-spondin expression was observed in some individual cells of the presumptive cephalic FP from the bud tail stage, at 10 hpf (Fig. 1A). The labeling intensity strongly increased in the flexural organ (FO), i.e., the anterior FP, and in the FP at 18 hpf (Figs. 1B and C, respectively). At 24 hpf, the labeling was limited to the FP and the FO, where it was maximal (Fig. 1D). At 28 hpf, the labeling occurred in the SCO and had already decreased in the FP, while the FO labeling moved toward a more rostral location (Fig. 1E). From 36 hpf (Fig. 1F), the labeling intensity progressively decreased in the FP until its complete disappearance at 48 hpf (Fig. 1G). The FO continued to migrate rostrally and the SCO labeling intensity increased. At 5 dpf (Fig. 1H), only the SCO labeling was observed.

Conservation of SCO-Spondin Modular Organization in Deuterostomia

From the published sequence organization in Mammalia (Gobron et al. 2000; Goncalves-Mendes et al. 2004; Meiniel and Meiniel 2007) and Aves (Didier et al. 2007), and the newly defined genomic sequences, we searched for the SCO-spondin gene and predicted the modular structure of the protein in Teleostei, Amphibia, Urochordata, and Echinodermata (Table 1 and Fig. 2). Except for the signal peptide, which was not found in *Ciona* species or *S. purpuratus*, the N-terminal organization is conserved and includes 3 vWDs, 4 SCORs, and 7 to 10 LDLrAs. The vWD domain has an average length of 148 amino acids and is involved in the protein multimerization process (PFAM accession number PF00094); the SCOR domain corresponds to a serine protease inhibitor sequence of 98 amino acid residues (Meiniel and Meiniel 2007); the LDLrA domain ensures the recognition of low-density lipids (PROSITE accession number PS50068). The rest represents more than half of the protein, in which TSRs (PROSITE accession number PS50092) and SCORs irregularly alternate. TSR motifs of SCO-spondin are thought to be relevant to adhesion and are involved in neurite extension and axonal pathfinding processes (Gobron et al. 2000). The number of TSRs ranges from 7 in *S. purpuratus* to 27 in *G. gallus*. There are 26 TSRs in other vertebrate classes and 18 in *Ciona*. Similarly, the number of SCORs also ranges from 7 in *S. purpuratus* to 15–16 in

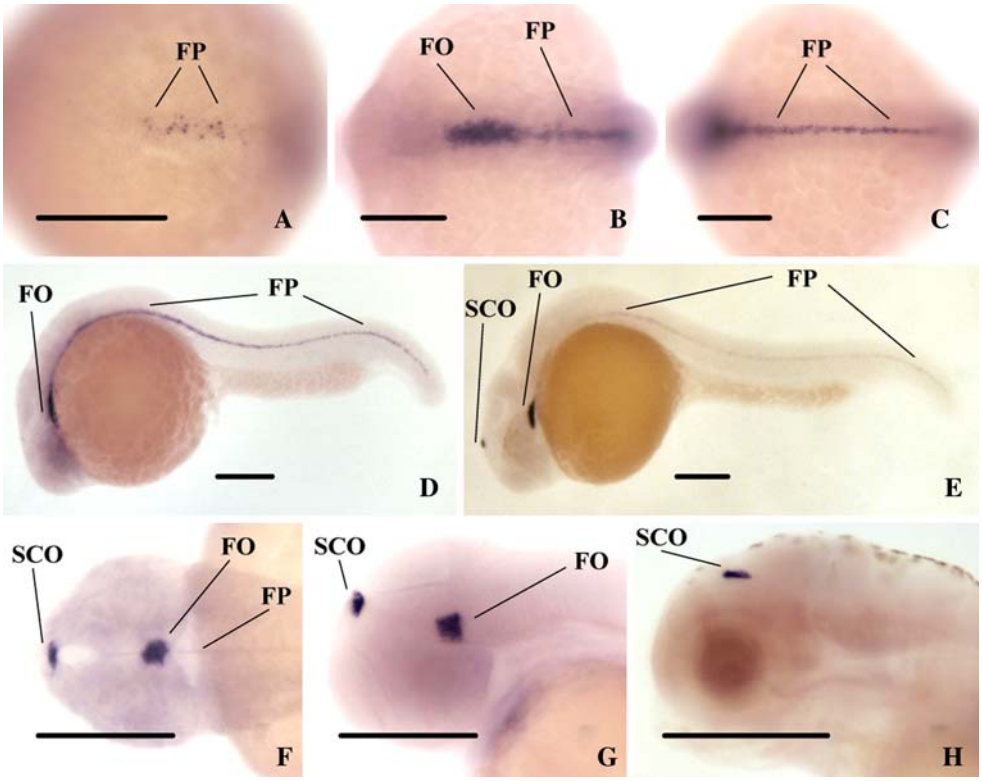

Fig. 1 Scale bar: 0.25 mm. In situ hybridization pattern of *Danio rerio* embryos at different developmental stages performed with a subcommissural organ (SCO)-spondin specific riboprobe: (**A**) dorsal view of a 10-h-postfertilization (hpf) whole embryo; (**B**, **C**) dorsal views of an 18-hpf whole embryo focusing on head and trunk, respectively; (**D**) lateral view of a 24-hpf whole embryo; (**E**) lateral view of a 28-hpf whole embryo; (**F**) dorsal view of a 36-hpf embryo's head; (**G**) lateral view of a 48-hpf (hatching) embryo's head; (**H**) lateral view of a 5-days-postfertilization (dpf) larva's head. The earliest signs of SCO-spondin expression were observed from 10 hpf (tail bud) in the presumptive cephalic floor plate (FP). The labeling intensity then increases in the FP and the cephalic FP, also called the flexural organ (FO), up to 24 hpf. At 24 hpf, the labeling is limited to the FP and the FO. At 28 hpf, transition of the expression to the SCO has started and the FP labeling decreases. At 36 hpf, the FP signal is almost restricted to the FO. Up to 48 hpf, the SCO-spondin transcripts are present in both the ventral FO and the dorsal SCO. After hatching, the labeling intensity decreases in the FO, and in larval stages (5 dpf), only the SCO remains positive

Teleostei, Amphibia, Aves, and Mammalia. We note that an additional series of 1 to 3 LDLrA domains occurs between TSR 3 and TSR 4 in Chordata. The C-terminal part of the protein ends with a highly conserved CTCK module (PROSITE accession number PS01225), which is thought to be involved in the dimerization of SCO-spondin monomers.

Diversity of TSR Motifs in Deuterostomia

TSR motifs are characterized by a stretch of about 55 amino acids and a conserved pattern of cysteine (C), tryptophan (W), and arginine (R) residues. The general features of this motif are wxxWxxWxxCs in the N-terminal of the module sequence, two conserved arginines (Rxr) in the midpart of the module, and the positioning of six to eight cysteine residues (Chen et al. 2000; Adams 2001). Three groups of TSRs were defined in Mammalia (Meiniel and Meiniel 2007) on the basis of the cysteine distribution pattern (Fig. 3). Group 1 TSRs share six cysteines, have no cysteine in position 2 of the previously given NH_2 consensus (xxwxxWxxWxxCs), and show a cysteine in position 25 in the central part of the motif, resulting in the core sequence RxRxCxx. Group 2 TSRs also share six cysteines but show a cysteine in the amino terminus (CxwxxWxxWxxCs) and no cysteine in the central region of the TSR (Rxrxxxx). Group 3 TSRs share eight cysteines, the first one within the canonical N-terminal consensus (xxwCxWxxWxxCs), and have two additional cysteine residues in positions +2 and +4 of the conserved arginine region (RxrxCxC). The three groups of TSRs were found in all the Deuterostomia studied. In Vertebrata, there are 15–16 TSRs of group 1, 8–10 TSRs of group 2, and 2 TSRs of group 3. In Urochordata there are 14 TSRs of group 1, 4 TSRs of group 2 and 1 TSR of group 3. In Echinodermata, there are 2 TSRs of group 1, 4 TSRs of group 2, and 1 TSR of group 3. Therefore, TSRs of group 1 prevail in Chordata, while the ratio of group 2 TSRs is higher in Echinodermata.

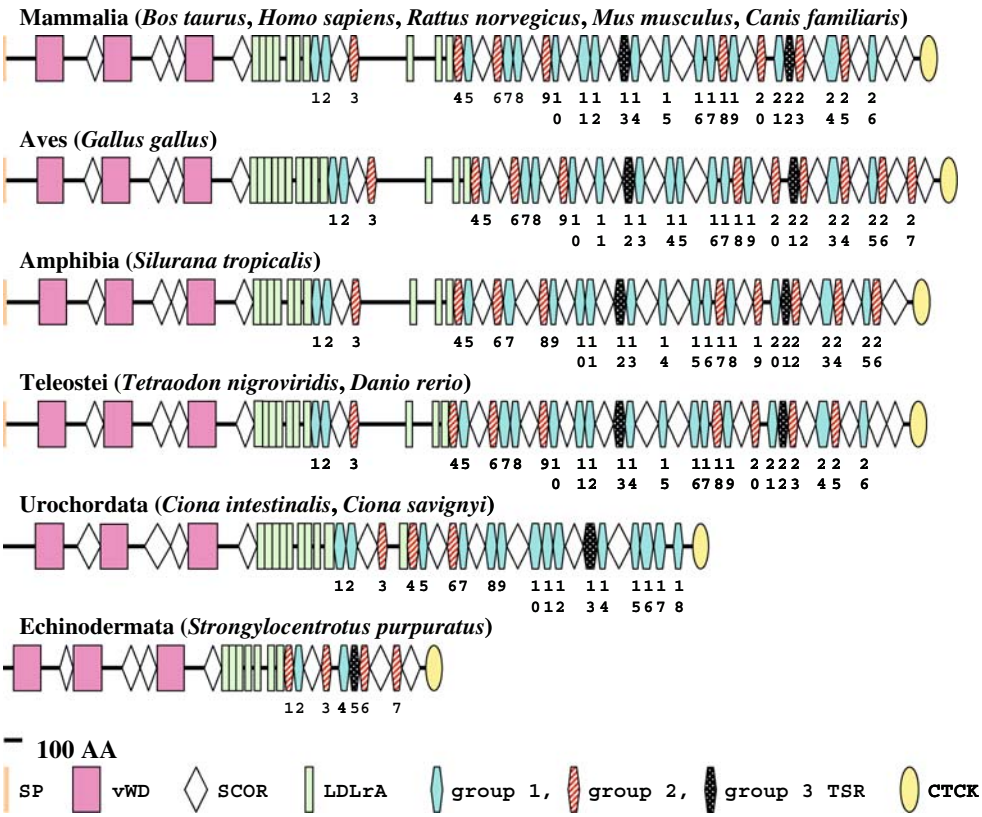

Fig. 2 Conservation of the multidomain organization of SCO-spondin among the superphylum of Deuterostomia. The scheme of the mosaic structure is drawn for four vertebrate classes, Urochordata, and Echinodermata. The various structural domains are represented by different shapes and the three groups of thrombospondin type 1 repeats (TSRs) are highlighted by three kinds of shadings. TSRs are numbered from the N-terminus to the C-terminus of a given protein. The structure, including three von Willebrand factor D (vWD) and SCO-spondin repeat (SCOR) motifs and the C-terminal cystine knot (CTCK) domain, was found to be conserved among Deuterostomia. Note the variations in the number of low-density lipoprotein receptor type A (LDLrA) and TSR/SCOR stretches

Orthology Relationships Between TSRs

The maximum likelihood algorithm implemented in PhyML enables us to assess the orthology of each vertebrate TSR with high bootstrap values (see tree in Supplementary data 1), ranging from 45 to 100% (mean, 81%), despite the short length of the analyzed sequences. A largely concordant tree was obtained with the minimum evolution algorithm implemented in MEGA3.1 (Supplementary Data 2). We summarize the correspondences between the Vertebrata TSRs in Fig. 4A. The order of TSR modules is highly conserved within the four Classes of Vertebrata. However, a few variations were observed. If we retain the numbering of TSRs in *T. nigroviridis* as a reference, the 8th is lacking in *S. tropicalis*, the 12th and the 21st in *G. gallus*. By contrast, a new TSR appears in *G. gallus* between the equivalents of the 14th and 15th, significantly close to the 2nd one (bootstrap value = 94%). Also, after the equivalent of the 26th, two close TSRs (bootstrap value = 98%) are inserted in *S. tropicalis* and in *G. gallus*, related to the equivalent of the 10th (bootstrap value = 39%). The origin of the last TSR of *G. gallus* (g27) is obscure and may be related to the equivalent of the 21st (from minimum evolution) or of the 20th (from maximum likelihood). Oddly, the 22nd of *T. nigroviridis* is associated with equivalents of the 4th, but it may be an artifact, as it is also related to the other 22nd with bootstrap value 42% (from maximum likelihood).

These variations concerning the *T. nigroviridis* sequence of TSRs are all supported by the positions of intervening SCOR motifs. For example, the 11th and 12th of *T. nigroviridis* are flanked by two SCORs, whereas in *G. gallus*, there is only one TSR between these two SCORs. In the same species, the 15th of *T. nigroviridis*, flanked by a pair of SCORs, corresponds to two TSRs.

Interestingly, when the arrangement of TSRs and SCORs of *Ciona* is compared with the one observed in Teleostei, the pattern closest to Urochordata's among the Vertebrata studied, there is a striking identity from the 1st TSR to the 15th of the equivalent of *T. nigroviridis* (N-terminal TSRs) (Fig. 4B). The phylogenetic analyses support this view, the TSRs 1, 3, 6, and 13 of *Ciona* and *T. nigroviridis* are recognized as orthologous (Supplementary Data 1 and 2). However, the 14th in *Ciona* is probably orthologous to the equivalent of the 24th in Vertebrata. The C-terminal TSRs from the 15th onward are more numerous in Vertebrata (about 12) than in *Ciona* (5). The 20th of Teleostei may be orthologous to the 16th, 17th, and 18th of *Ciona*.

For Echinodermata and Vertebrata, several orthology relationships can be assessed: the 2nd, 5th, and 6th in *S. purpuratus* correspond to the 19th, 22nd, and 25th in

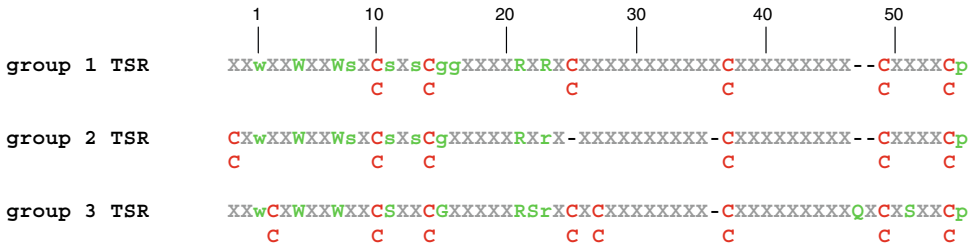

Fig. 3 Consensus sequences and cysteine patterns of group 1, 2, and 3 thrombospondin type 1 repeats (TSRs) in Mammalia. Highly conserved cysteines are shown in red. Other conserved amino acids are shown in green. Any amino acid is denoted by a gray X. Conserved cysteine patterns distinguishing group 1, 2, and 3 TSRs were found in all vertebrate classes, Urochordata, and Echinodermata

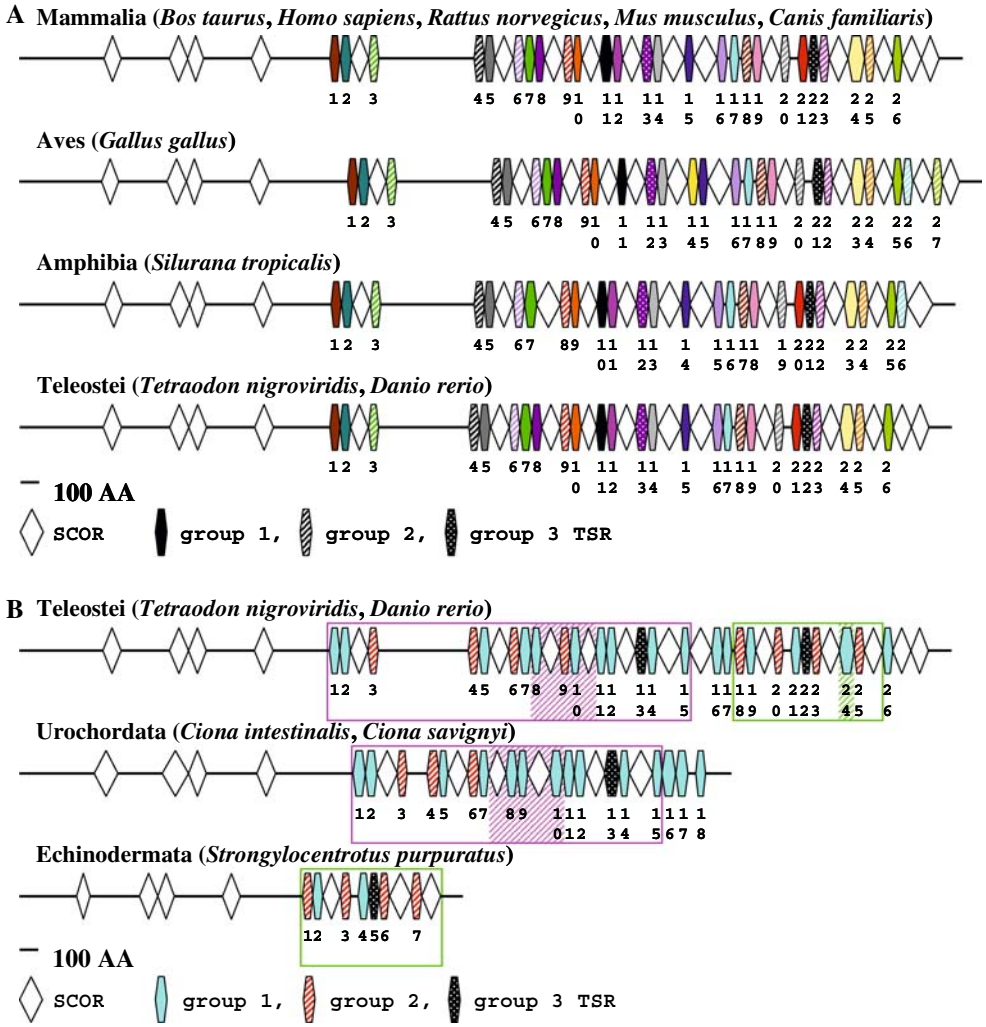

Fig. 4 **(A)** Conservation of thrombospondin type 1 repeat (TSR) orthologues in vertebrate classes. One color is attributed to each orthologous TSR motif among species. **(B)** Comparison of TSR/SCO-spondin repeat (SCOR) patterns in Deuterostomia. The regions sharing common TSR/SCOR arrangements between *Tetraodon* and *Ciona* and between *Tetraodon* and *S. purpuratus* are boxed in pink and green, respectively. Differences in TSR and (or) SCOR composition between common regions are striped

T. nigroviridis, respectively (Supplementary Data 3). Interestingly, the stretches of TSRs and SCORs in both organisms show the same pattern at the C-terminal end of the protein, from the 18th to the 25th TSR in *T. nigroviridis* (Fig. 4B).

Paralogy Relationships Between TSRs

Surprisingly, 14 TSRs (more than half of the structure in Vertebrata) are associated in pairs in the tree generated by PhyML, separated by a distance of about 11 TSRs: 2–13,

5–14, 7–17, 8–21, 9–18, 10–26 (*G. gallus* and *S. tropicalis* 26, corresponding to the equivalent of the 27th), and 11–23. The correlation between the position of the N-terminal and C-terminal TSRs is highly significant ($r = 0.88$, $p < 0.01$). This situation suggests that the duplication events that led to these TSRs are not independent, but result from block duplication (Fig. 5). Other pairs of TSRs are separated by one to four TSRs: 1–4, 3–6, 10–15, 12–16, 22–26, and 24–25. A few TSRs are single: 19 and 20.

Discussion

Evolution of the Sites of SCO-Spondin Production

A great deal of work has been done on the ontogenetic and phylogenetic origins of the RF (for a review see Olsson 1993). Material secreted by the FP has been shown to participate in RF formation in all Chordata, but whether SCO-spondin is the protein participating in the early formation of this threadlike structure has remained controversial. For the first time, we clearly demonstrate a transitional expression of SCO-spondin in the FP cells of a vertebrate model, *Danio rerio*, prior to the expression in the SCO at the larval stage. These *in situ* hybridization results correlate with FP immunoreactions observed with an anti-RF antibody (unpublished results; see Gobron et al. [1999] for the details of the antibody).

Phylogenetically, the SCO is exclusively a vertebrate feature (Oksche 1961, 1969). Protochordata never pass the FP stage of the embryonic Craniata. In *Branchiostoma lanceolata* (Cephalochordata), the fiber source is the infundibular organ (IO), the equivalent in Craniata of the most rostral part of the FP called the flexural organ (FO) (Olsson and Wingstrand 1954). In ascidian tadpole larvae of *Ciona intestinalis* (Urochordata), the site of RF secretion is not known, but the neural tube lined by four rows of ependymal cells contains an RF (Olsson 1972). We hypothesize that the ependymal cells of the embryonic neural tube may be the counterpart of FP cells expressing the SCO-spondin gene in zebrafish, for example, and that they are probably the source of SCO-spondin at least partly constituting the RF. In *Oikopleura dioica* (Appendicularia), a tiny RF structure is secreted by one fibrogen cell (FC) (Holmberg and Olsson 1984). There is a transition of the site of SCO-spondin secretion between Prochordata, where it is ventral (IO, FP, FO, and FC), and Vertebrata, where it becomes dorsal in larvae (SCO).

We show that an analogous evolutionary pattern is conserved during vertebrate ontogenesis as the zebrafish *D. rerio* transitionally expresses SCO-spondin in the FP and FO territories during early developmental stages. This ventral expression then progressively decreases in the FP but remains in the FO, while at the same time SCO dorsal expression increases. SCO-spondin expression then stops in FO and remains only in the SCO of larvae. Thus a ventrodorsal shift in the sites of SCO-spondin synthesis appears during evolution and this shift is observable during ontogenesis. The kinship among SCO, FO, IO, FC, and FP is undeniable but it is still unclear why primary ventral sources are replaced by a dorsal location at more advanced ontogenetical stages of vertebrate development.

In Chordata, the RF is concomitant with the appearance of a hollow neural tube. No hollow neural tube was found in Echinodermata, but the ectoneural system is made of neuronal and glial cells (Holland 1984). These glial cells may secrete SCO-spondin, as they have been shown to synthesize a Reissner's substance-like product recognized by an anti-RF antibody in the sea star *Asterias rubens* (Viehweg et al. 1997). However, this remains to

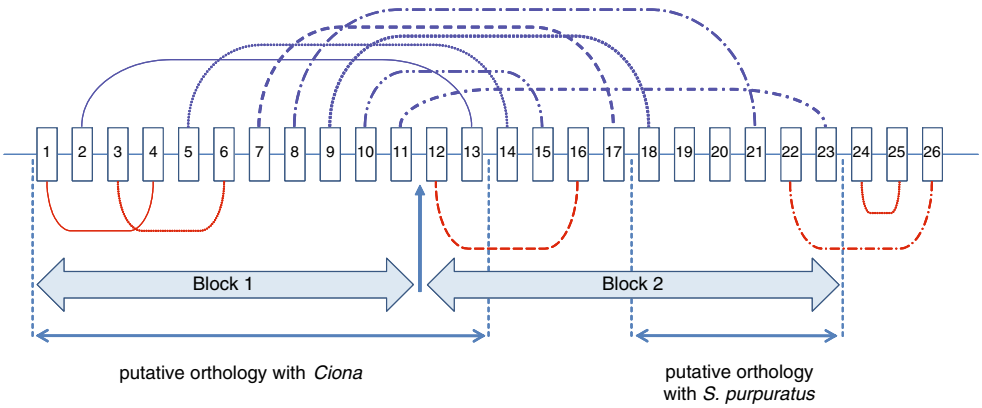

Fig. 5 Paralogy relationships between Teleostei TSRs and hypothetical block duplication model. Each numbered box corresponds to a *Tetraodon* TSR. The relationships between paralogues are given by the phylogenetic tree in Supplementary Data 1. Blue links above the boxes support the block duplication scenario. Red links below correspond to subsequent arrangements within the two blocks. The probable extent of each block is indicated by a large arrow. The putative orthologous regions between *Tetraodon* and *Ciona*, and between *Tetraodon* and *S. purpuratus*, are illustrated by thin arrows

be analyzed further using a SCO-spondin probe and (or) antibody.

The Lengthening of SCO-Spondin During Evolution

We show from both experimental and predicted sequences (see Table 1) that the general mosaic organization of SCO-spondin, combining vWD, LDLrA, TSR, SCOR, and CTCK motifs, is conserved among Deuterostomia. The orthologous gene coding for this protein was identified bioinformatically, not only in Vertebrata and Urochordata, but also in *Strongylocentrotus purpuratus* (Echinodermata). This last finding correlates well with the anti-RF immunoreaction observed in *Asterias rubens* and raises the question of the original role of SCO-spondin and how it evolved.

Although the predicted protein is already a large one in the sea urchin *S. purpuratus* (2730 amino acids), the stretch of alternating SCORs and TSRs is notably lengthened in Chordata. This stretch is more than twice as long in *Ciona* and three times as long in Vertebrata as in Echinodermata. We found that the arrangement of SCORs, TSRs, and the CTCK in sea urchin is faithfully found in the C-terminal part of the vertebrate SCO-spondin, and probably closely resembles to the ancestral stretch in Deuterostomia.

If the Teleostei arrangement is taken as a reference for Vertebrata, most TSRs (7) between the 1st and the 11th have a paralogue in the region between the 12th and the 23rd. This suggests that a block duplication has occurred (Fig. 5), lengthening the SCO-spondin by doubling the number of TSRs. This event was followed by local translocations, as the order of paralogues in the two blocks is not conserved. In addition, subsequent duplication and (or) genic conversion events have taken place within each block.

The observation of the situation in *Ciona* and *S. purpuratus* enables us to deduce the relative dates of these events. The complementary information given by orthology and TSR/SCOR arrangement in *Ciona* and Vertebrata shows a common organization, encompassing block 1 and the beginning of block 2. This would mean that the block duplication occurred in the common ancestor of Urochordata and Vertebrata. Also, the 14th and the 16th/17th/18th TSRs in *Ciona* are orthologous to the 24th and 20th, respectively, in Teleostei. We thus hypothesize that *Ciona* has secondarily lost most of its C-terminal region. Concerning postblock duplication events, some identical traces are found in *Ciona* and Vertebrata, e.g., the 3rd6th TSRs duplication within block 1, as a result of an event earlier than the Urochordata-Vertebrata divergence.

As stated previously, the sea urchin *S. purpuratus* presents a stretch of TSRs/SCORs organized similarly to the C-terminus part of the vertebrate block two (Figs. 4B and 5). Taken together, these elements prompt the conclusion that the block duplication event occurred in the chordate lineage after the split of Echinodermata and Chordata and prior to the divergence of Urochordata and Vertebrata. We propose that block 2 is the ancestral block, which was duplicated in the common ancestor of Urochordata and Vertebrata. This hypothesis has to be tested and a thorough study of SCO-spondin in other Deuterostomia may provide useful clues.

Probable Significance of SCOR-TSR Association

The highly conserved modular pattern of SCO-spondin probably has a profound functional significance. For example, the presence of SCORs may protect the protein from protease cleavages. The protective effect of SCORs between the main domains probably helps the protein to maintain its activity throughout the time of its propagation from the SCO to the *ampulla caudalis* in Chordata, but the question arises of their presence in echinoderms, as these do not have a neural tube. In *Mus musculus*, the RF grows at a rate of about 10% of its length per day (Ermisch 1973) and is degraded in the *ampulla caudalis* (Peruzzo et al. 1987).

TSR motifs represent an increasingly large part of SCO-spondin during Deuterostomia evolution and the role of these modules has been assessed by several authors. The experiments of Monnerie et al. (1998) have shown the ability of SCO-spondin TSRs to promote cell-to-cell or cell-to-matrix adhesion and neurite outgrowth. This is consistent with the effect of various proteins of the TSR superfamily expressed in the CNS (Adams 2000). Thrombospondin 1, which is expressed early in the matrix of the CNS, contains three group 1 TSRs and has been shown to promote adhesion and neurite outgrowth of central and peripheral neurons (Neugebauer et al. 1991; O'Shea et al. 1991). This particular activity has been localized within the TSR region (Osterhout et al. 1992; DeFreitas et al. 1995. Similarly, an FP-expressed protein, F-spondin, which contains six group 2 TSRs (Klar et al. 1992), has been shown to influence pathfinding of commissural axons (Burstyn-Cohen et al. 1999) and repair of injured peripheral sensory axons (Burstyn-Cohen et al. 1998). Group 3 TSRs were first described in SCO-spondin (Meiniel and Meiniel 2007) and to date they have been found only in this protein. Also, mixing of group 1, 2, and 3 TSRs in SCO-spondin is an original feature, which is conserved from Echinodermata to Vertebrata. The significance of the combination of TSRs belonging to different groups and the subtle functional differences among the three groups need to be investigated.

In Vertebrata, SCO-spondin may participate in the formation of commissures during both the development and the evolution of the CNS; TSR motifs are thought to be relevant to the adhesive property and promoting effect on neuronal differentiation of the protein (see Introduction). We hypothesize that the lengthening of TSRs and SCOR arrangements in the SCO-spondin structure is correlated to the increase in the complexity of networks during the evolution of the deuterostomian CNS.

In conclusion, we note that SCO-spondin is a matricial protein involved in neuronal cell differentiation. The issue remains whether it can be secreted in Echinodermata by radial glia cells and then act locally. The presence of an RF in Chordata depends on the polymerization of SCO-spondin proteins; the question arises whether this process is linked to the cavitation of the neural tube during evolution. In Deuterostomia, we have demonstrated the conservation of the multidomain structure of SCO-spondin from Echinodermata to Vertebrata and propose a scenario of the lengthening of this giant protein. The generation of complexity by block duplication is an important feature of genome evolution (Petit et al. 2006). Until now, no gene coding for SCO-spondin has been identified in the genomes of Protostomia, but this issue needs to be revisited regularly.

Acknowledgments We thank Sridhar Ramachandran of the ZFIN for editing the scospondin gene. We acknowledge the contributions of genome sequencing projects that have generated some sequences useful for our analyses. Special thanks go to Olivier Jaillon at the Genoscope for providing *T. nigroviridis* sequencing accession numbers and to Kerstin Howe and Mario Caccamo at Sanger for helping with *D. rerio* genomic data. We are grateful to Bernard and Christine Thisse for their generous collaboration in establishing zebrafish expression patterns. We thank Richard Ryan for proofreading. Work at the UMR 384 is supported by INSERM. O. Meiniel is supported by a DRRT postdoctoral fellowship (FNADT, Région Auvergne-Limousin).

References

Adams JC, Tucker RP (2000) The thrombospondin type 1 repeat (TSR) superfamily: diverse proteins with related roles in neuronal development. Dev Dyn 218:280–299

Adams JC (2001) Thrombospondins: multifunctional regulators of cell interaction. Annu Rev Cell Dev Biol 17:25–51

Burstyn-Cohen T, Frumkin A, Xu YT, Scherer SS, Klar A (1998) Accumulation of F-spondin in injured peripheral nerve promotes the outgrowth of sensory axons. J Neurosci 18(21):8875–8885

Burstyn-Cohen T, Tzarfaty V, Frumkin A, Feinstein Y, Stoeckli E, Klar A (1999) F-Spondin is required for accurate pathfinding of commissural axons at the floor plate. Neuron 23:233–246

Charron F, Tessier-Lavigne M (2005) Novel brain wiring functions for classical morphogens: a role as graded positional cues in axon guidance. Development 132(10):2251–2262

Chen H, Herndon ME, Lawler J (2000) The cell biology of thrombospondin 1. Matrix Biol 19:597–614

DeFreitas MF, Yoshida CK, Frazier WA, Mendrick DL, Kypta RM, Reichardt LF (1995) Identification of integrin $\alpha_3\beta_1$ as a neuronal thrombospondin receptor mediating neurite outgrowth. Neuron 15:333–343

Didier R, Meiniel O, Meiniel A (2007) Molecular cloning and early expression of chick embryo SCO-spondin. Cell Tissue Res 327:111–119

Ermisch A (1973) Zur Charakterisierung des Komplexes Subcommissuralorgan-Reissnerscher Faden und seiner Beziehung zum Liquor unter besonderer Berücksichtigung autoradiographisher Untersuchungen sowie funktioneller Aspekte. Math Naturwiss R (Wiss Z Karl Marx Univ Leipzig) 22:297–336

Gobron S, Monnerie H, Meiniel R, Creveaux I, Lehmann W, Lamalle D, Dastugue B, Meiniel A (1996) SCO-spondin: a new member of the thrombospondin family secreted by the subcommissural organ is a candidate in the modulation of neuronal aggregation. J Cell Sci 109:1053–1061

Gobron S, Creveaux I, Meiniel R, Didier R, Dastugue B, Meiniel A (1999) SCO-spondin is evolutionnarily conserved in the central nervous system of the chordate phylum. Neuroscience 88:655–664

Gobron S, Creveaux I, Meiniel R, Didier R, Herbet A, Bamdad M, El Bitar F, Dastugue B, Meiniel A (2000) Subcommissural organ/Reissner's fiber complex: characterization of SCO-spondin, a glycoprotein with potent activity on neurite outgrowth. Glia 32:177–191

Goncalves-Mendes N, Simon-Chazottes D, Creveaux I, Meiniel A, Guénet J-L, Meiniel R (2003) Mouse SCO-spondin, a gene of the thrombospondin type 1 repeat (TSR) superfamily expressed in the brain. Gene 312:263–270

Guindon S, Gascuel O (2003) PHYML: a simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. Syst Biol 52(5):696–704 (available at http://atgc.lirmm.fr/phyml/)

Holland ND (1984) Epidermal cells. In: Bereiter-Hahn J, Matoltsy AG, Richards KS (eds) Biology of the Integument. 1. Invertebrates. Springer, New York, 756–774

Holmberg K, Olsson R (1984) The origin of Reissner's fibre in an appendicularian, *Oikopleura dioica*. Vidensk. Medd Dansk Naturhist Foren 145:43–52

Jeanmougin F, Thompson JD, Gouy M, Higgins DG, Gibson TJ (1998) Multiple sequence alignment with Clustal X. Trends Biochem Sci 23:403–405

Klar A, Baldassare M, Jessel TM (1992) F-spondin: a gene expressed at high levels in the floor plate encodes a secreted protein that promotes neural cell adhesion and neurite extension. Cell 69:95–110

Kumar S, Tamura K, Nei M (2004) MEGA3: Integrated software for Molecular Evolutionary Genetics Analysis and sequence alignment. Brief Bioinform 5:150–163

Lopez-Avalos MD, Cifuentes M, Grondona JM, Miranda E, Perez J, Fernandez-Llebrez P (1997) Rostral floor plate (flexural organ) secretes glycoproteins immunologically similar to subcommissural organ glycoproteins in dogfish (*Scyliorhinus canicula*) embryos. Dev Brain Res 102:69–75

Meiniel O, Meiniel A (2007) The complex multidomain organization of SCO-spondin protein is highly conserved in mammals. Brain Res Rev 53(2):321–327

Monnerie H, Dastugue B, Meiniel A (1998) Effect of synthetic peptides derived from conserved domain on chick cortical and spinal-cord neurons in cell cultures. Cell Tissue Res 293:407–418

Naumann W, Müller G, Kloss P (1987) Immunoreactive glycoprotein of the subcommissural organ in the embryonic stages of the vertebrate brain. Math Naturwiss R (Wiss Z Karl Marx Univ Leipzig) 36:17–20

Neugebauer KM, Emmett CJ, Venstrom KA, Reichardt LF (1991) Vitronectin and thrombospondin promote retinal neurite outgrowth: developmental regulation and role of integrins. Neuron 6:345–358

Oksche A (1961) Vergleichende Untersuchungen über die sekretorische Aktivität des Subcommissuralorgans und den Gliacharakter seiner Zellen. Z Zellforsch 54:549–612

Oksche A (1969) The subcommissural organ. J Neuro-Visc Relat (Suppl) 9:111–139

Olsson R (1961) Subcommissural ependyma and pineal organ development in the human fetuses. Gen Comp Endocr 1:117–123

Olsson R (1972) Reissner's fiber in ascidian tadpode larvae. Acta Zool (Stockh) 53:17–21

Olsson R (1993) Reissner's fiber mechanisms: some common denominators. In: Oksche A, Rodriguez EM, Fernandez-Llebrez P (eds) The subcommissural organ: an ependymal brain gland. Springer Verlag, Berlin, pp 33–39

Olsson R, Wingstrand KG (1954) Reissner's fibre and the infundibular organ in Amphioxus. Publ Biol Stat (Univ Bergen Årbok) 14:1–14

O'Shea KS, Liu L-HJ, Dixit VM (1991) Thrombospondin and a 140 Kd fragment promote cell adhesion and neurite outgrowth from embryonic central and peripheral neurons and from PC 12 cells. Neuron 7:231–237

Osterhout DJ, Frazier WA, Higgins D (1992) Thrombospondin promotes process outgrowth in neurons from the peripheral and central nervous systems. Dev Biol 150:256–265

Peruzzo M, Rodriguez S, Delannoy L, Hein S, Rodriguez EM, Oksche A (1987) Ultrastructural immunocytochemical study of the massa caudalis of the subcommissural-Reissner's fiber complex in lamprey larvae (*Geotrica australis*). Evidence for a terminal vascular route of secretory material. Cell Tissue Res 247:367–376

Petit D, Maftah A, Julien R, Petit JM (2006) En bloc duplications, mutation rates and densities of amino acid changes clarify the evolution of vertebrate α1,3/4 fucosyltransferases. J Mol Evol 63(3):353–364

Reissner E (1860) Beiträge zur Kenntnis vom Bau des Rückenmarkes von *Petromyzon fluviatilis* L. Arch Anat Physiol Wiss Med (Leipzig) 545–588

Rodriguez EM, Oksche A, Hein S, Rodriguez S, Yulis R (1984) Comparative immunocytochemical study of the subcommissural organ. Cell Tissue Res 237:427–441

Rodriguez EM, Oksche A, Hein S, Yulis R (1992) Cell biology of subcommissural organ. Int Rev Cytol 135:39–121

Rodriguez EM, Jara P, Richter H, Montecinos H, Flandez B, Wiegand R, Oksche A (1993) Evidence for the release of CSF-soluble secretory material from the subcommissural organ, with particular reference to the situation in the human. In: Oksche A, Rodriguez EM, Fernandez-Llebrez P (eds) The subcommissural organ: an ependymal brain gland. Springer Verlag, Berlin, pp 121–131

Rodriguez EM, Del Brio Leon MA, Riera P, Menendez J, Schoebitz K (1996) The floor plate of the hindbrain is a highly specialized gland. Immunocytochemical and ultrastructural characteristics. Dev Brain Res 97:153–168

Schoebitz K, Garrido O, Heinrichs M, Speer L, Rodriguez E.M (1986) Ontogenetical development of the chick and duck subcommissural organ and immunocytochemical study. Histochemistry 84:31–40

Sterba G, Kiessig C, Naumann W, Petter H, Kleim I (1982) The secretion of the subcommissural organ. A comparative immunocytochemical investigation. Cell Tissue Res 226:427–439

Thompson JD, Gibson TJ, Plewniak F, Jeanmougin F, Higgins DG (1997) The ClustalX windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. Nucleic Acids Res 25:4876–4882

Viehweg J, Naumann WW, Olsson R (1997) Secretory Radial Glia in the Ectoneural System of the Sea Star *Asterias rubens* (Echinodermata). Acta Zool (Stockholm) 79(2):119–131