PART I. TENEURINS - AN EMERGING FAMILY OF TRANSMEMBRANE PROTEINS

Teneurins are a newly discovered glycoprotein family with conserved domain architectures and amino acid sequences from invertebrates to vertebrates. Moreover, teneurin expression patterns are similar across phyla, suggesting that their function is also maintained between species [1].

Since most of data available for teneurins is limited to invertebrate expression profiles and *in vitro* studies, I will first describe the expression pattern in invertebrates.

1. 1 INVERTEBRATE TENEURINS

Teneurins were initially discovered in *Drosophila melanogaster* in the earlier 1990s by two screenings performed by two groups independently, one searching for invertebrate homologs of the tenascin protein family [2][3], and the other one performing a screen for tyrosine phosphorylated proteins [4]. Screening results performed by the groups found the same novel molecule, which was termed "tenascin-like molecule major" (*Ten-m*) by Baumgartner et al. (1994) or coined by Levine et al. (1994) as "*odd Oz*" (or *Odz*) because mutations in the gene resembled the phenotype observed in "odd pair-rule" mutants. This last finding was unexpected for a transmembrane protein, since until then, the segmental patterning in *Drosophila* was only known to be regulated by traditional nuclear transcription factors [4][5].

Tenascin is a large extracellular matrix glycoprotein involved in cellular adhesion during cell motility, proliferation and differentiation [6]; a series of tenascin-type EGF-like repeats present on the teneurin molecule explain its appearance in both screens [3][4][5].

1.1.1 Teneurin Expression in *Drosophila melanogaster*

During *Drosophila* development, segmentation occurs in a stepwise hierarchical manner, thus anterior-posterior axis formation is initiated by expression of maternal genes that then activate zygotic genes. Under the effect of maternal genes, Gap genes activate a subset of pair rule genes, which subsequently activate segment-polarity genes [7]. Ten-m/Odz was found to be expressed in alternating odd-numbered parasegments in the blastoderm stage where its expression overlaped with the *even-skipped* stripes [3]. By their initial description, teneurins were considered as a pair-rule gene. In this regard, Baumgartner et al. (1994) proposed that Ten-m/Odz was a "secondary" pair rule gene whose activity during segmental pattering depended on the "primary" pair-rule genes ftz (*fushi tarazu*) and *eve* (*even-skipped*), with Ten-m/Odz acting upstream of the segment polarity-genes: *prd* (*paired*), *slp1* (*sloppy paired 1*), *gsb* (*gooseberry*), *en* (*engrailed*) and *wg* (*wingless*). Mutations in *odd-paired* (*opa*), another pair-rule gene, cause a phenotype characterized by partial fused denticles belts, deletion of alternative parasegments and naked cuticle embryos [8]. Similarly mutations in the *Ten-m/Odz* gene caused embryonic lethality due to the fusion of adjacent denticle belts resembling the *opa* mutants phenotype [3][8]. Since expression patterns of *prd*, *gsb* and *slp1* were identical in both mutant types, whereas Ten-m/Odz expression pattern remained virtually unaffected in the *opa* mutants, Baumgartner et al. (1994) postulated that Ten-m/Odz modulates Opa protein activity during fly blastoderm and germband extension stages. Because Ten-m/Odz

expression was found in some stripes whereas *opa* may be expressed in all primordium cells, it was suggested that both these genes were epistatically related in the regulation of downstream genes such as *prd*, *slp* and *gsb*, when segmental pattering occurs [3].

Despite the consideration of Ten-m/Odz as "earlier" pair-rule gene, Levine et al. (1994), describe Ten-m/Odz as later acting pair-rule gene that is expressed when the syncytial blastoderm becomes a cellular blastoderm and within which, cell-cell communication can take place [4].

Many genes involved in pattern formation during earlier embryogenesis stages are also important in the morphogenetic processes occurred later, and for teneurin this might also be the case [5][7]. Regardless of when Ten-m/Odz starts to be expressed, these *in vivo* studies highlight that it is fundamental for the regulation of segmental pattern formation in *Drosophila* embryogenesis [2][3][4].

In addition to above, during embryo fly development Ten-m/Odz is also expressed in cardiac cells, in the lymph gland and in the tracheal system, as well as in all segmental furrows that fold during gastrulation. After hatching and during larval development, Ten-m/Odz is present in the ventral nerve cord, in a subset of neurons in the brain, in pioneering commissural axons as well as around muscle attachment sites [3][4]. Ten-m/Odz is particularly strongly expressed in the developing eye during pupal stages indicating a function in eye patterning [9]. In the eye imaginal disc, Ten-m/Odz is present in the morphogenetic furrow, which implies a function in the differentiation of imaginal disc cells [9][10]. As ommatidia mature, Ten-m/Odz is expressed in the photoreceptor R7 and may be involved in the specification of this cell type. Mutations in Ten-m/Odz result in R7 photoreceptor and

visible light photoreceptor loss and other defects such as ommatidial size, shape and rotation [11].

The Ten-m/Odz paralog Tenascin-like molecule accessory (Ten-a) have an expression pattern that overlaps with Ten-m/Odz during fly development, suggesting that both proteins may act cooperatively [2][12]. Although mutants for Ten-a alone do not show an odd pair-rule phenotype, this phenotype is enhanced when mutations of Ten-a are presents in the later Ten-m/Odz mutants embryos [5].

1.1.2 Teneurin Expression in *Caenorhaditis elegans*

In *C. elegans* there is only one gene homolog to *Ten-m/Odz* named Ten-1 from which, two coding isoforms are generated by alternative promoters [13]. The long form, Ten-1L, encodes for the full-length protein whereas the short form, Ten-1S, is missing most of the N-terminal region. Both isoforms have complex and distinct expression patterns as evidenced by GFP-fluorescence microscopy. Transfection experiments of Ten-1L promoter-GFP translational fusion protein indicate that it is already expressed in the germ line and in all cells during early embryogenesis. Later in development it is found among several other locations and in the mesoderm-derived distal tip cell, which plays a crucial function in regulating the morphogenesis of the gonad. Moreover, Ten-1S promoter is mainly found in the dorsal hypodermis and in the anterior leading cell of the ventral hypodermis during embryonic development; later is found in a subset of neurons. Thus, the Ten-1L promoter is mainly active in the mesoderm whereas Ten-1S is more highly expressed in the ectoderm. On the basis of their pattern of expression, it was inferred that the main function of teneurin-1 might

be the promotion of cell–cell recognition, often associated with pattern formation and cell migration process [13].

The *in vivo* function of Ten-1 was analyzed using deletion-mutants and RNAi against both long and short forms of teneurin. The embryonic lethality of worms injected with RNAi against both Ten-1 transcripts most likely resulted from abnormal embryonic elongation due to hypodermal cell migration defects, whereas nerve cord defasciculation and aberrant axonal pathfinding were found in depleted Ten-1S worms. Further, interference with Ten-1L protein function showed a phenotype that included defects in gonad migration and somatic gonad organization, but these abnormalities can be partially rescued with a full-length ten-1-containing cosmid [13][14]. Additionally, homozygous Ten-1 deletion worms that lack part of the C-terminal region are viable, but 15–20% of them are sterile or burst-through-the-vulva due to germ cell leakage in the middle of the gonad [14]. A more detailed analysis of gonad rupture and bursting of the early gonads formation in these mutants suggested that this phenotype must be probably caused by defects in basement membrane (BM) formation or in its maintenance. In fact, the Ten-1 mutant phenotype resembled in many aspects the phenotypes of *laminin-like protein* (*epi-1)* as well as the laminin binding receptors *integrin-alpha* (*ina-1*) and *distroglican-1 (dgn-1)* mutants, and double mutants between *ten-1*, *ina-1,* and *dgn-1* showed a synergistic genetic effect on BM maintenance, implying that these three genes act in similar developmental processes and that they have partly redundant functions on the regulation of BM integrity in early development of the gonad. However, the mechanism of teneurin signaling as BM adhesion molecule remains to be elucidated [14].

In addition, Drabikowski et al. (2005) showed that the intracellular domain of Ten-1L can be detected in cell nuclei when anti-bodies against the teneurin N-terminal domain are

used. This suggests that Ten-1 signaling may function through proteolytic cleavage and the N-terminal domain may be subsequently translocated to the nucleus. Although the processing mechanism for teneurin-1 cleavage is unclear, the authors proposed that Ten-1L constitutes a novel member of the family of membrane-anchored transcription modulators [13]. Similar observations were evidenced in vertebrate counterpart supporting this idea, as over-expression of the chicken teneurin-2 intracellular domain has also shown a nuclear localization in cultured cells [15].

1.2 VERTEBRATE TENEURINS

A few years after their initial description in *Drosophila*, several groups independently described Ten-m/Odz-like proteins in vertebrates. The first report came from Wang et al. (1998) who, using a differential expression screen found the murine Ten-m/Odz ortholog, named DOC4 (for Downstream of CHOP 4 protein), which was found expressed in the brain of mouse embryos [16]. This work was followed by successive reports of teneurin orthologs in zebrafish, chickens, mice, rats and humans. Different names for protein were assigned depending on the species where they were found.

In vertebrates, four teneurin paralogs were found and named teneurin-1 to -4 [5]. Hereafter, we will refer to the *Drosophila* proteins by their original names (i.e., Ten-a and Ten-m/Odz) [2][4], while we will use the "teneurins" term to describe these proteins in vertebrates. The term Teneurin was adopted referring to a fusion name between its original discovery as **ten**ascin-like molecule and the observation that it is prevalently expressed in developing **neur**ons [9]. A summary of the respective citations, names and chromosomal localizations of teneurin homologs is shown in Table 1.1.

Table 1.1. Chromosomal Location and Nomenclature for Teneurins

(Adapted from Tucker et al., 2006)[5]

	Human	Mouse	G. Gallus	Synonyms	D. melanogaster	C. elegans
ODZ1 gene	Xq25	XA3.3	4			R13A5.9
Teneurin-	1	1	1	ten-m1[a] ; odz1[b] ; odz3[c] ; ten-m/odz1[d]; tenascine-M1[e]		Ten-1S; Ten-1L
ODZ2 gene	5q34	11A4	13		3L	
Teneurin-	2	2	2	ten-m2[a]; odz2[b]; odz1[c]; ten-m/odz2[d]; neurestin[f]	Ten-m/Odz; Ten-a	
ODZ3 gene	4q35.1	8B1.1	4			
Teneurin-	3	3	3	ten-m3[a]; odz3[b]; ten-m/odz3[d]		
ODZ4 gene	11q14.1	7E1	1			
Teneurin-	4	4	4	ten-m4[a];odz4[b]; ten-m/odz4[d]; DOC4[g]		

[17] a Oohashi et al.(1999)

[18] b Ben-Zur et al.(2000)

[19] c Ben-Zur & Wides (1999)

[20] d Zhou et al.(2003)

[21][22] e Brandau et al.(1999); Mieda et al.(1999)

[23] f Otaki & Firestein (1999)

[16] g Wang et al.(1998)

1.2.1 Domain Organization of The Teneurin Family

Computer modelling and experimental evidence predict that teneurins are type II transmembrane glycoproteins with a single pass transmembrane-spanning helical domain [24][25]. Regardless of the gene-product or species of origin considered, all teneurins share common basic characteristics (Figure 1.1):

- They are large proteins, with a molecular mass of approximately 300kDa, composed of between 2500 and 2800 amino acids.

- The N-terminal (300-375 aa) domain lies inside the cell. The intracellular domain (ICD) contains two EF-hand like motifs in which calcium can bind; two polyproline rich motifs, typically associated to SH3-protein docking, and several conserved tyrosines, which are predicted to be phosphorylated. The vertebrate ICD is highly conserved and can be aligned along its entire length between different species [5].

- The remaining large C-terminal domain of about 2400 residues is remarkably well conserved between vertebrates and invertebrates. This domain is positioned outside the cell, and 200-300 aa away from the transmembrane segment there are eight phylogenetically well conserved tenascin-type EGF-like repeats that feature the unique and conserved replacement of cysteine in repeat 2 with a tyrosine residue, and the conserved replacement of a cysteine in repeat 5 with a phenylalanine (or another tyrosine in the case of Ten-a). The EGF-like repeats are followed by about 700-800 residues that do not contain any recognizable repeats or homologies to any proteins other than the teneurins themselves. In this region, 17 cysteine residues are present that, after the EGF-like repeats, are the second most conserved regions between invertebrate and vertebrate teneurins. All teneurins present a stretch of approximately 800 residues in the distal half of the extracellular domain (ECD) where 26 YD-repeats are found that do not appear in any other eukaryotic proteins [5][26][27].

The overall identity between vertebrate teneurins is about 60-98%, depending whether orthologs or paralogs are considered, while conservation between vertebrate teneurins and fly and worm teneurins is 30-40% [5][26][27].

There is good evidence that teneurins exist as either homo- or hetero-dimers in the cell membrane. For example, under reducing conditions, both Ten-a and murine teneurins have an apparent molecular mass of 300 kDa, but under non-reducing conditions they run at 600 kDa [12][24]. Electron microscopic analysis also shows purified murine teneurin-1 as two large globules linked by thin stalks with the ends of their sticks tied together. The dimerization most likely takes place between the EGF-like repeats of two teneurins arrayed side-by-side, as evidenced by a recombinant protein assay consisting of the murine teneurin-1 EGF-like repeats fused with alkaline phophatase, which binds to purified teneurin-1 on a far

Western Blot [17]. Additionally, a more conclusive evidence that the dimerization comes from covalent bisulphite binding between the conserved, unpaired cysteine residues in the 2nd and 5th EGF-like repeats came in the detailed study by Feng et al. (2002), who found that both homo -and heterotypic dimerization were possible *in-vitro*.

Three postulated cleavage sites are present in teneurins (Figure 1.1): There is a conserved furin cleavage site in ECD between the transmembrane (TM) domain and the first EGF-like repeat, which was shown to be functional in teneurin-2 [1]. The resulting ectodomain shedding appears to take place *in vivo*, since an antibody recognizing the teneurin-2 EGF-like repeats labels both the cell surface and the extracellular matrix in the chick embryo. This cleavage site is also present in invertebrate teneurins, suggesting that ectodomain shedding is important for the function of teneurins [13][25]. A putative cleavage site within or near the TM domain serves to release the teneurin ICD inside the cell [4][15]. Finally, there is a conserved furin cleavage site near the C-terminus, which was proposed to release the Teneurin C-terminal Associated Peptide (TCAP), with neuromodulatory activity [28][29]. The modular domain composition of the teneurins and the evidence for proteolytic cleavage processing suggest that different parts of the teneurins may act separately to mediate distinct teneurin functions [25][5].

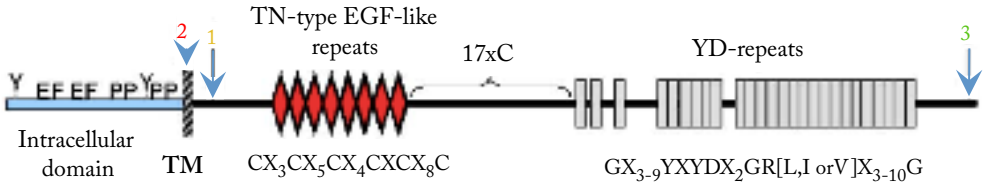

Figure 1.1.Organization of teneurin domains. The N-terminal intracellular domain harbouring two EF hand-like motifs that may bind Ca^{2+} ions (EF), two proline-rich stretches (PP) and conserved tyrosines residues predicted to be phosphorylated (Y). About 200 aa after the transmembrane domain (TM) there are eight tenascin-type (TN-type) EGF-like repeats of the consensus sequence indicated, followed by a trait of 700-800aa containing 17 conserved cysteine residues (17xC). The YD-repeats of the consensus sequence are also indicated. Arrows indicate postulated furin-cleavage sites: (1) indicates a site found in teneurin-2; (2) arrowhead marks a cleavage site in or near the TM resulting in the release of the ICD; (3) Near to the C-terminal, all teneurins have the consensus RXRR cleavage site postulated to be involved in the release of Teneurin C-terminal Associated Peptide (TCAP) *(adapted from Tucker et al., 2006)* [5].

Teneurins may be post-translationally modified by glycosylation, as treatment with N-glycosidase F alters their electrophoretic mobility [24]. The teneurin EGF-like repeats are most closely related to those found in Tenascin-C and Notch where glycosylation is required for proper ligand interaction and for Notch signalling induction [30][31-34]. A database search for the EGF-repeats of all teneurins and Tenascin-C revealed a protein-O-fucosyltransferase- 1 (POFUT1) consensus site, which is therefore predicted to be O-fucosylated [32]. Additionally, the YD–repeats are also predicted to be glycoslyated, and were shown to bind glycosaminoglycan and heparin *in vitro* [9].

1.2.2 Vertebrate Teneurin Expression and Functions

All studies investigating the expression of vertebrate teneurins in mice, chicken and zebrafish agree that teneurins are prominently expressed in the developing central nervous system, but at least some are also expressed in other tissues where they may regulate pattern formation during morphogenesis, as well as at sites of cell migration and muscle attachment points [2-4][13][35].

Teneurin-1, -2 and -4 knockout mice or zebrafish teneurin morpholino knockdowns have not been reported to date, but genetic analyses of their homologs in invertebrates have been helpful in demonstrating their functions [35]. Moreover, the expression and function of vertebrate teneurins have been studied most extensively in chicken and mice.

Therefore, I will first describe the expression pattern and their predicted functions in these species and then I report the *in-vitro* studies and the ongoing model for teneurin function.

1.2.2.1 Avian Teneurins

In the visual chick system teneurin-1, -2 and -4 are prominently expressed in a non-overlapping manner in a subset of developing neurons as evidenced by *in-situ* hybridization and immunohistochemistry experiments [25][36-38]. Briefly, at day 7 of embryonic development (stage E7), which coincides with the onset of synaptogenesis of the chicken visual system, teneurin-2 is detected in the stratum of the optic tectum; a few days later, the protein is present in various mesencephalic and thalamic nuclei that make up the thalamofugal visual system; then, teneurin-2 is prevalently expressed in the hippocampus and in parts of the forebrain, two receiving areas of the thalamofungal visual center inputs [25]. At this stage of development, teneurin-4 is detected in the retina and in the optic nerve, and this expression pattern is maintained throughout the stage E17. In central retina layers, all four teneurins are present with related but distinct patterns [39]. For example, at E17 teneurin-1 distribution is concentrated in the laminae of inner plexiform layer and nuclei of the tectofugal visual system, two interconnected populations of neurons that make up the tectofugal visual pathway in chicken. Both teneurin-2 and -1 are found in the retina inner plexiform layer, but they are concentrated in different laminae [36]. Teneurin-4 is expressed in the outermost lamina of the plexiform layer, whereas teneurin-3 resembles the teneurin-1 expression pattern [39].

In the developing central nervous system of the chicken embryo, the teneurin-1 transcript is widely expressed at embryonic day 17 (E17) and similarly to teneurin-2 and -4, it is strongly expressed in cells of the olfactory bulb and in some neuron populations in the hippocampus. Teneurin-1 transcripts are also presents in brain areas involved in motor sensation and where auditory information is processed, as well as in the cerebellum and spinal

cord. At this stage, teneurin-1 -2 and -3 are prominently expressed in the cerebellum, but teneurin-4 immunostaining is faint [36][39]. These related and almost complementary patterns of expression in the chick nervous system suggest that teneurins may play a role in cell-cell recognition for appropriate synapse formation, and/or in signalling events during the development of specific neural networks, as well as in axon fasciculation in the developing visual system [36][39].

Teneurins are also present in non-neural tissues. For example, teneurin-4 is found in peripheral tissues during development, as evidenced by strong staining around the lung primordium, the gut epithelium and the kidney-surrounding mesenchyme using an antisera raised against the ICD and ECD domains of teneurin-4. Moreover, teneurin-1 is found in some regions of the ectoderm and mesoderm in the limb bud [39][40].

In developing the chicken limb, Teneurin-2 and -4 are also expressed. While teneurin-2 and -4 are present in the apical ectoderm ridge (AER)[36][37], in the zone of polarizing activity (ZPA) only teneurin-4 is expressed but its expression is transient, since it is soon relocated into the anterodistal mesenchyme of the limb bud [39]. Because the expression of teneurin-2 in the AER coincides with the presence of the morphogenetic factor FGF8, it was suggested that teneurin-2 might be a downstream target of FGF8 signalling [38]. AER is induced by fibroblast growth factor (FGF) which determine the proliferation of underlying mesenchymal cells and the distal outgrowth of the limb bud; and ZPA is induce by the Sonic hedgehog (Shh) factor to determine the anterior-posterior axis formation [41][42]. Because teneurin-4 is particularly expressed in both crucial organizer-signaling centers: the AER and ZPA, Kenzelmann (2008) proposed that teneurin-4 expression might also be regulated by FGFs signals. In this context, teneurin-2 and -4 induced by FGFs may function as adhesion

molecules and mediate the maintenance of reciprocal coordination between these both centers during proximo-distal and anterior-posterior pattern specifications in the developing limb [40].

At least three splice variants of chicken teneurin-2 have been identified, whose expression has been confirmed by RT-PCR and *in situ* hybridization. A short splice variant is composed only of the intracellular domain, the transmembrane region and the first seven EGF-like repeats. The two long variants differ from each other only in the presence or absence of eight amino acids between the 7th and 8th EGF-like repeats [38]. Moreover, Northern blot experiments with teneurin-2 probes show multiple messengers in adult mouse tissues [17], but it is unclear if these represent splice variants of the various teneurins or if they are a hybridization artefact [38].

1.2.2.2 Murine Teneurins

Zhou et al. (2003) studied teneurin expression in mice, concentrating on the hippocampus and cerebellum. They found that each teneurin is expressed by distinctive subpopulations of neurons, but that these populations often overlap. For example, in the adult hippocampus teneurin -1 and -3 transcripts are most prominent in CA2 subfield of the pyramidal layer, whereas teneurin-2 is expressed throughout this region and teneurin-4 is barely detectable. However, in the dentate gyrus all teneurins are presents. Furthermore, Teneurins-2, -3 and -4 are expressed in the cerebellum Purkinje cell layer whereas teneurin-1 is present in the granule cell layer. Interestingly, there is also a faint *in situ* hybridization signal with a teneurin-4 probe in cerebellar white matter, indicating that at least some glia may be able to express this teneurin as well [20].

In the adult mouse, teneurin-1 and -3 extra-neuronal expression was detected by Northern blotting, *In situ* hybridization and Immunohistochemistry, particularly in both the kidney and testes. However, teneurin-3 Northern blot showed that it was expressed at a much

lower level than in brain [20].

Recently, two studies of mouse teneurin-3 shed light on the *in vivo* function of vertebrate teneurins. In a microarray study comparing gene expression between the visual and somatosensory cortex at the time of establishment of neuronal connectivity, several teneurins (teneurin-2, -3 and -4) were found to be specific for the visual cortex [43]. Further experiments demonstrated that the cell bodies of neurons transfected *in-uterus* with GFP-teneurin-3 tend to aggregate and exhibit increased neurite outgrowth that persists during development [44]. The same group published the first analysis of a teneurin-3 knockout mouse. The teneurin-3 knockouts homozygous (-/-) are viable and survive into adulthood. Their brain and visual system appear normal on anatomical and histological level and do not exhibit an overt phenotype. However, KO mice have defects in retinal ganglion cells (RGC) axon targeting of the ipsilateral projections, resulting in an impairment of binocular vision. The authors conclude that teneurin-3 is the first molecule that was shown to affect only ipsilateral, but not contralateral RGC axon guidance [44].

Moreover, an early study of neurestin (rat teneurin-2) shows its expression in the developing olfactory bulb and in hippocampal granule cells, areas where cell regeneration is continuous. Since expression in tufted cells is re-induced during olfactory bulb regeneration after injury, the authors speculate that neurestin is involved in neuronal development and the neural regeneration process [23].

Finally, in a microarray study, murine teneurin-4 was shown to be expressed in the regenerating olfactory bulb, suggesting that teneurins might not only be important during mammalian neuronal development, but also during neuronal regeneration [45].

Thus, based on the expression analysis of the different teneurins, it is inferred that they play fundamental roles in the regulation of pattern formation in both the CNS and non-neural tissues at different stages of vertebrate and invertebrate development. Some examples include the teneurin expression in nerve bundles during invertebrate and vertebrate eye development stage as well as during leg imaginal discs in fly and in the limb buds of chicken and mice. A summary of teneurin expression profiles described to date is provided in Table 1.2

Zebrafish	
teneurin-1	ND
teneurin-2	ND
teneurin-3	developing brain[1], somites[1], notochord[1], pharyngeal arches[1]
teneurin-4	developing brain[1], spinal cord[1]
Chicken	
teneurin-1	developing CNS[2, 3, 4], visual system[2, 3, 4]
teneurin-2	developing CNS[2, 3, 4], visual system[2, 3, 4], AER of limb buds [5] tendon primordia [5], pharyngeal arches [5], heart[5], somites [5], neural tube [5], craniofacial mesenchyme [5]
teneurin-3	ND
teneurin-4	developing CNS[6], ZPA of limb buds[6], pharyngeal arches [6]
Mouse	
teneurin-1	developing and adult CNS [7,8,9], visual system [7,8], smooth muscle cells in lung[8], kidney glomeri[8], adult testes[8]
teneurin-2	developing and adult CNS[9], visual system[7]
teneurin-3	developind and adult CNS[10, 2,11,7,9,12], visual system [10,7], spinal cord[10,12], notochord[12], craniofacial mesenchyme[10], tongue[10], dermis[10], saccule[10],developing limb[10],periosteum[10]
teneurin-4	developing and adult CNS[10,7,13,12], visual system[7], somites[12], spinal cord[10],trachea[10], nasal epithelium[10], saccule[10], joints[10], adipose tissue[10], tail bud and limbs[13]
Human	
teneurin-1	fetal and adult brain*, spinal cord*,cerebellum*, heart*, lung*, kidney*, liver*
teneurin-2	brain*, spinal cord*,cerebellum*, heart*, lung*, kidney*, skin*
teneurin-3	brain*, spinal cord*, cortex*, bone marrow*,cerebellum*, heart*, lung*, kidney*, skin*, ovary*, placenta*, testis, cervix*
teneurin-4	brain[14], cortex[14], cerebellum[14], bone marrow*, spinal cord*, thymus*, heart* scheletal and smooth muscles*, kidney*,lung*, pancreas*, prostate*, thyroid*, ovary*, testis*

Table 1.2. Summary of Teneurin Expression in Vertebrates. These expression data are supported by the following studies: (1) Mieda, 1999 [22], (2) Kenzelmann, 2008 [40], (3) Minet, 1999 [9], (4) Rubin, 1999 [1], (5) Tucker, 2001 [38], (6) Tucker, 2000 [37], (7) Li, 2006 [46], (8) Oohashi, 1999 [17], (9) Zhou, 2003 [20], (10) Ben-Zur ,2000 [18], (11) Leamey, 2007 [43], (12) Leamey, 2008 [44], (13) Lossie ,2004 [47], (14) Chiquet-Ehrismann and Kenzelmann, 2011 [48].

(*) Refers to microarray data that do not have additional supportive experimental evidence. (Reported: HG-U95B, GNF1H and/or HGU133 Plus2 gene chips) [49].

1.2.2.3 *In Vitro* studies of Teneurins

A few studies have addressed the function of teneurins using cell culture or biochemical assays. For example, explanted chicken dorsal root ganglia were induced to form neurites when treated with recombinant teneurin-1-or -2 YD-repeats rich-lysate [9][25]. Further, transfection of neuroblastoma cell lines (Nb2a) with an avian full-length teneurin-2 construct results in the formation of numerous teneurin-2 enriched filopodia and enlarged growth cones, suggesting a possible interaction of teneurin-2 with the cytoskeleton [1]. The formation of filopodia may be dependent on the teneurin ICD, since COS-7 cells transfected with full-length or teneurin-2 lacking the YD repeats form numerous filopodia, but cells transfected with the same constructs lacking the ICD do not [25]. Additionally, the overexpression of teneurin-2 construct containing the YD-repeat in primate fibrosarcoma cell lines (HT1080) promotes cell aggregation, but this does not occur in cells lacking the YD-repeat. Similarly, it was noted that in transiently transfected Nb2a cells, teneurin-2 becomes concentrated at points where adjacent cell bodies are touching [25].

On the other hand, the epistatic relationship previously evidenced between Ten-m/Odz and *opa* in *Drosophila* [3], led Begutti et al. (2003) to study a possible relationship between teneurin-2 and zic-1, a mammalian homolog of *opa*. Their study not only showed that the teneurin-2 ICD is translocated to the nucleus where it co-localizes with PML bodies, markers of transcriptional activity [51], but also that the teneurin-2 ICD decreased zic-1 expression. Another important finding of this study was to show that the teneurin-2 ICD can be detected in the nucleus even though a teneurin-2 construct containing only the TM domain was transfected, suggesting that teneurin-2 could be cleaved from the membrane-spanning portion. In addition, they also showed that nuclear translocation was possible when the full-

length teneurin-2 constructs were transfected into cells that constitutively expressed a construct lacking teneurin-2 ICD, suggesting that homophilic interaction between teneurin-2 ECD may trigger the cleavage of ICD for its nuclear translocation [15].

Further, within the nucleus the teneurin-1 ICD was found to interact with both the methyl-CpG-binding protein MBD1, which could mediate teneurin-dependent modulation of gene expression in the nucleus, and CAP/Ponsin, an adaptor protein that through vinculin could represent a possible link of teneurin to the actin cytoskeleton [50][52].

Thus, all these *in-vitro* evidences point to the consideration of teneurins as putative adhesion molecules associated to cytoskeleton, and that interaction between them could result in translocation to the nucleus and to modify the expression of some genes such as those involved in pattern specification during development. However, still very little is known on how these proteins are regulated and what could be their function in the adult organism. A schematic representation of a current working model for teneurin signaling is depicted in Figure 1.2.

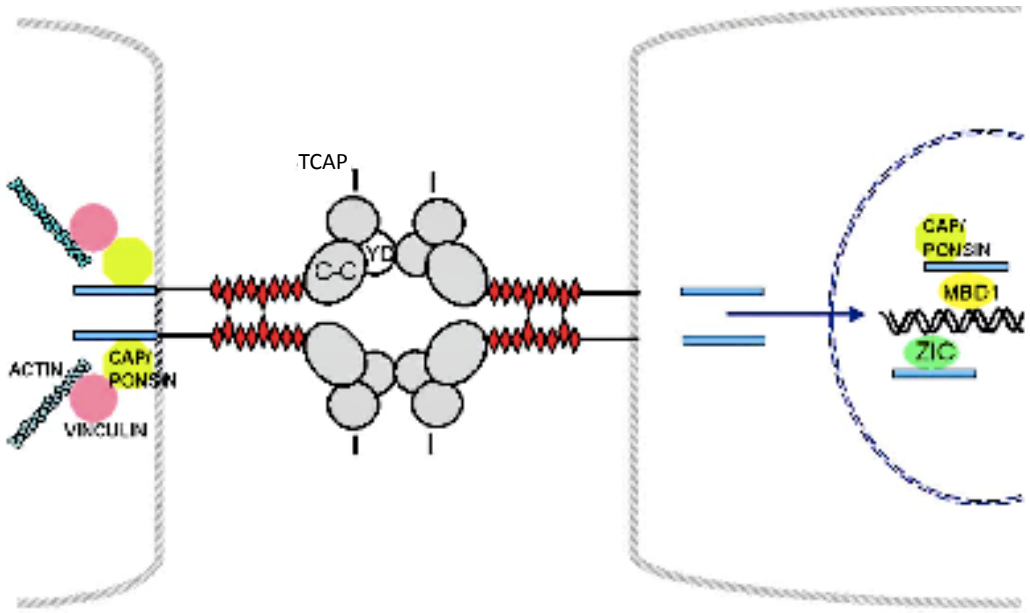

Figure 1.2. Ongoing model proposed for teneurin signaling. In the figure, teneurin is depicted as a dimer linked by cysteine bridges formed among the EGF-like repeats two and five; the homophilic interaction through the extracellular domains distal to the EGF-like repeats triggers receptor dimerization. At left, teneurin is anchored to the cytoskeleton through its interaction with the adaptor protein CAP/ponsin, which in turn is linked to actin through vinculin. Upon homophilic binding with the right cell, where teneurins are also dimerized, the ICD is proteolitycally cleaved and translocated to the nucleus, where it may interact with transcriptional regulators such as MBD1 and zic. The teneurin associated signaling also includes the possible release of an active neuropeptide from the extracellular domain, TCAP. This type of direct signaling from a cell surface receptor to the nucleus is becoming recognized for many other receptor proteins including the case of notch signaling. *[Reproduced from Tucker and Chiquet-Ehrismann, (2006)].*

Recent *in-vivo* evidence supports this functional model: Ten-1 and teneurin-1 -2 and -3 processing and nuclear trafficking were also reported in *c. elegans* and in neurons of chicken embryos, respectively. Thus, not only can the ICD be translocated to the nucleus *in vivo*, but also this event is phylogenetically conserved. A proposed mechanism by which these membrane glycoproteins could act as transcriptional regulators is the Regulated Intramembrane Proteolysis (RIP) mechanism, in which the teneurin ICD could serve as a substrate for the site-2 protease –or the SPP-like proteases-, two members of the intramembrane type II protease family [39][40][53].

1.3 TENEURIN GENES

In vertebrates, genes named ODZs encode the different teneurin proteins [4][5]. Therefore, there are four ODZ genes, one for each teneurin (Table 1.1). As previously described, the teneurin genes were so called because they were identified as odd pair-rule genes in *Drosophila* [4]. However, unlike these genes, ODZs are the only known pair-rule genes that do not seem to encode for traditional transcription factors. Instead, they encode for membrane glycoproteins with tenascin EGF-like repeats that might act as transcriptional regulators [39][40].

In humans, the ODZ genes have different sizes. Nevertheless, the gene-products are around the 300kDa for all members of the family. The result of a database search for human ODZ genes and their characteristics is summarized in Table 1.3.

Table 1.3. ODZs Gene Family and Their Predicted Characteristics in Humans.

Gene	Teneurin	RefSeq	gene size, Kb	exon coding *	mRNA size, Kb	protein size
ODZ1	1	NM_001163279.1	587	32/32	12.9	2731 Aa
ODZ2	2	NM_001122679.1	980	29/29	9.6	2765 Aa
ODZ3	3	NM_001080477.1	479	27/27	10.8	2699 Aa
ODZ4	4	NM_001098816.2	780	30/34	13.6	2769 Aa

* coding exon respect to the total mRNA compoining exons

Accession number of Reference Sequence (RefSeq) for the different ODZ transcripts found on GenBank database [54].

1.3.1 The Human ODZ4 gene

The odd Oz/ten-m homolog 4 (ODZ4) predicted gene structure is composed by 34 exons spanning 787367 bp on human chromosome 11q14.1 [54]. The length of the entire ODZ4 transcript is about 13.6 Kb and four of the thirty-four total exons would not be translated (see Table 1.3, GenBank RefSeq Acc.n NM_001098816.2). The validated Reference Sequence (RefSeq) record on the GenBank database was derived from direct submissions and preliminary reviews; nevertheless, virtually no experimental data regarding its entirely coding-structure is available yet for human. Therefore, the predicted ODZ4 mRNA Reference Sequence was created from genomic sequence data because no single transcript was available for the full length of the gene. Thus, the extension of the transcript is solely supported by transcript alignments and orthologs data [55].

In the mature transcript, a region of about 8.3 Kb composed of exons five to exon thirty-four encode for the protein (CDS), whereas the first four exons form the 5'-untranslated region (5'-UTR). Figure 1.3 depicts the predicted structure of the ODZ4 RefSeq transcript and the relative positions important for translation.

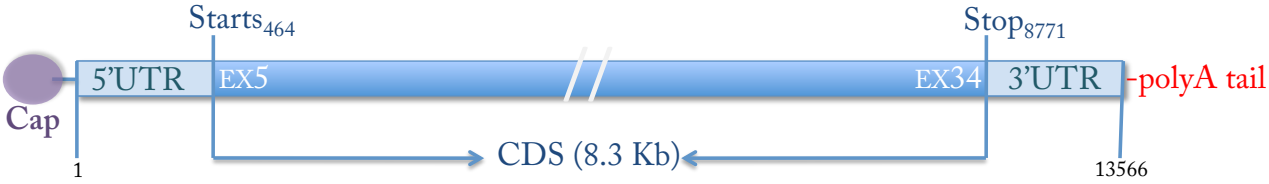

Figure 1.3. Predicted ODZ4 Mature Transcript (13.6 Kb of RefSeq, GenBank Acc.n NM_00108816.2). The figure represents the predicted structure of the full-lenght ODZ4 transcript. The start codon (ATG) localizes within exon five at base position number 464 whereas the stop codon (TGA) resides within exon thirty-four at base position 8771. Exons -1, -2, -3 and 4 form the 5'-UTR region.

1.4 TENEURINS IN HUMAN DISEASES

1.4.1 TCAPs as Neuroprotective Factors

Qian et al. (2004) discovered a novel family of bioactive peptides in teleost named urotensins, which corresponded to the C-terminus of teleost teneurin-3. Urotensin peptides were found during a low stringency screen for additional corticotrophin releasing factor (CRF) paralogues using mammalian urocortin as a probe. Urocortins are stress hormones that bind to CFR receptors; but in fish they are called urotensins [56].

Therefore, TCAPs are the released-form of the teneurin ECD. There are four members from the corresponding processed teneurin paralogues, however, the exact processing mechanism is unknown. Qian et al. (2004) also demonstrated that trout TCAP-3 can upregulate the proliferation of LHRH-neuron derived cell lines (G-11 cells) as well as to activate adenylate cyclase signalling. A more recent study showed similar results for murine TCAP-1 *in vivo* [28].

Over the years, neuropeptide systems have gained greater recognition in their ability to modulate cognitive and mood behaviours, and several studies have pointed to a role for TCAP expression in brain areas involved in psychiatric and mental degenerative disorders. Such an increased expression of TCAP-1 occurred in hippocampal areas CA1, CA2, CA3 and dentate gyrus [57]. Other limbic regions that clearly expressed TCAP include the piriform cortex, the bed nucleus of the stria terminalis, and the amygdala, which in rat brain are known as areas that regulate the emotional and stress responses [57][58].

At present, TCAPs are considered as neuroprotective factors, which may promote cell proliferation, attenuate cell degeneration [58] and upregulate enzymes involved in oxidative-stress responses such as superoxide dismutase (SOD1) and catalase [59]. Preliminary studies

indicate that some of TCAP-1's neuroprotective effects may be mediated by the neurotrophin BDNF [57][60] About this, PT00114 (a TCAP-1 peptide formulation) is currently in preclinical development phase for the treatment of anxiety and depression disorders, but may also find applications in areas of mental neurodegeneration, spinal cord injuries and in oncological approaches [61].

1.4.2 X-linked Mental Retardation

Mental retardation (MR) is a complex phenotype, defined by significant limitations both in intellectual functioning and in adaptive behaviour as expressed in conceptual, social and practical adaptive skills. The MR symptomatology originates before 18 years of age and is not degenerative [62].

Teneurins are predominantly expressed in the developing CNS where they are proposed to have a function in axon guidance and in the establishment of neuronal connectivity. Because human teneurin-1 resides in chromosome Xq25, a locus with low gene density that is compromised in the linkage intervals of several XLMR families, teneurin-1 has been hypothesized as promising candidate gene for XLMR [36][40].

In fact, the *OCRL1* gene located in Xq25 was known to be mutated in Lowe-Syndrome [63] for more than a decade; and since 2006, other Xq25 genes involved in XLMR have been also identified, such as ZDHHC971, which is near OCRL1; and GRIA372, which is very close to the teneurin-1 gene [64]. Additionally, frequent MECP2 mutations are observed in XLMR and in Rett syndrome [65]; for this reason, the entire family of methyl-CpG binding proteins was analyzed and a nonsense mutation in MBD1 in one autistic and mentally retarded patient was revealed [66]. The MBD1 protein was also shown to interact and to co-

localize in the nucleus with the teneurin-1 ICD (76) and Li et al. (2005) suggested that teneurins might be regulated by the homeobox transcriptor factor, Emx2. This observation was recently confirmed *in vivo* as reported by Beckmann et al. (2011), who published that a specific human teneurin-1 splice variant transcript is upregulated by EMX2. Moreover, it has been also reported that mutation of the transcriptional repressor *ZIC-2* causes MR due to abnormal failures to form distinct hemispheres (holoprosencephaly) [68], and that teneurin-2 ICD might repress the transcription repression mediated by another ZIC family member, ZIC-1 [15]. Thus, several proteins that interact with teneurins are also XLMR genes, so it is likely that also teneurins may be implicated in the XLMR manifestation.

1.4.3 Teneurins and Cancer

As mentioned previously, virtually no data are available on the role of teneurins in humans. Nevertheless, recent publications point to the possibility that teneurins might play a role in human cancers. In 2002, a linkage analysis revealed that one or more genes on chromosome 5q33-34 might act as modifiers for breast cancer risk in individuals carrying BRCA1 mutations [69], and the teneurin-2 gene (ODZ2), which maps to chromosome 5q34, was among the candidate genes. Later, it was suggested that ODZ2 could acts as tumor-suppressor gene, since the gene was found to be disrupted in patients with chronic hepatitis caused by HBV infection [70]. Teneurin-2 mRNA was absent in one hepatic tumor and low in four of seven cases. Because the expression of teneurin-2 transcript was higher in normal tissues and decreased in tumors, the authors postulated that ODZ2 may act as a tumor-suppressor gene, and it was suggested that insertional mutagenesis mediated by viral infection, might be an early genetic change leading to hepatocarcinogenesis [70]. The

possibility that teneurin-2 might act as a tumor suppressor is supported by an additional microarray study performed on RNA of microdissected epithelial cells, which revealed that ODZ2 had the most downregulated (14.8-fold) expression in breast hyperplasic enlarged lobular units as compared to normal tissue [71]. These lesions are the earliest histological potential precursors of breast cancer and the loss of ODZ2 expression would thus occur very early in the transformation.

However, a different line of recent evidence proposes that teneurin-2 is being over-expressed in other tumors, suggesting an oncogenic role. For example, Vinatzer et al. (2008) identified a novel translocation involving the IGH and ODZ2 genes in three mucosa-associated lymphoid tissue lymphomas, two of which affected the skin and one the ocular adnexa [72]. These anatomical locations are highly reminiscent of the impairments in the visual system and the hypodermis, areas where also the teneurins are expressed. In this study, sufficient RNA was available only from one tumor for quantitative RT-PCR, which revealed several-fold elevations in teneurin-2 mRNA levels. Although functional data is missing, this evidence supports the role of teneurin-2 as oncogene rather than a tumor-suppressor.

A more recent study of teneurin-2 and lung cancer came from Ziegler et al. (2011), who throughout a combined proteo-transcriptomic strategy identified two new protein-markers, ODZ2 and thy-1/CD90 (THY1), for the discrimination between malignant pleural mesothelioma (MPM) and lung adenocarcinoma (ADCA) [73].

MPM is an aggressive disease mainly caused by asbestosis exposure [74]. The diagnosis requires immunohistochemistry (IHC) with combining sets of antibodies such anti-calretinin, anti-podoplanin and anti-cytokeratins 5/6 or –WT-1, as MPM positive markers, together with carcinoma-markers (negative for MPM). However, the discrimination of MPM from

other malignancies affecting the lung, particularly ADCA, is most challenging for diagnosis [75][76]. In this study, SILAC-based cell surface analysis coupled with mass-spectrometric measures approaches were used for the N-glycosylated cell surface protein profiling quantification in MPM and ADCA cells lines. Validation of result was done at protein level (by Wester blot analysis) as well as assessment of relative mRNA expression by low density array RT-PCR profiling. The provided result supports that teneurin-2, and specially thy-1/CD90, are cell surface accessible markers candidates for the clinical differentiation of MPM from ADCA [73].

On the other hand and although not yet published, in the work conducted by Kenzelmann (2008) an up-regulation of ODZ4 in malignant brain tumors was described [40][48]. The RNA was subject to microarray analysis and consistently elevated levels of the ODZ4 transcript were found in astrocytomas, oligodendriomas, and glioblastomas. Notably, this is the sole work where brain teneurin-4 mRNA was evaluated together with teneurin-4 protein levels (by Western blot and Immunohistochemistry analysis). For protein analysis they used their own antibodies because until then, teneurin-4 antibodies were not commercially available. So, all previously mentioned studies were limited to the analysis of teneurin genes and mRNA expression. In addition, the author showed a prominent teneurin-4 staining of blood vessels, and it was postulated that teneurin-4 might also play a role in tumor angiogenesis [40].

PART II. TENEURINS IN HUMAN OVARIAN CANCER

2.1 OVARIAN CANCER

Ovarian cancer is a highly lethal disease. Despite its rather low incidence, it remains the leading cause of death from gynecologic malignancy, and the fifth most common cause of death from cancer among women. Initial symptoms are usually unspecific, with the consequence that >60% of patients come with advanced stages (III-IV) at time of diagnosis [77].

Therefore, premature detection of ovarian cancer before nodal and extrapelvic dissemination can significantly impact on patient survival, however currently there are not efficient screening procedures for early disease detection. Measurement of serum CA-125 is used for disease diagnosis and prognosis management, but this marker fails to detect ~50% of earlier ovarian tumors [78]. Thus, the search for improved markers of disease aimed to detect ovarian cancer at early stage is ongoing.

The etiology of ovarian cancer is still poorly understood and up to now, no clinically recognizable features can predict future ovarian cancer risk. For the larger group (90%) of sporadic cases, ovarian cancer susceptibility has been explained by recurrent inflammation associated with ovulation [79]. However, the observed association with ovarian carcinogenesis is modest, suggesting that multiple other processes must be involved. For the group with inherited predisposition (10%) ovarian cancer is associated with germline mutations in the BRCA1 -or -2 gene, and in this group of women, often standard CA-125 measure at time of

diagnosis fail to detect the earlier stage of disease [80][81]. In addition, a smaller fraction of familial cases are associated with germline mismatch mutations in some repair genes [82].

2.1.1 Ovarian Cancer Histopathology

Primary tumors of the ovary can arise from three different cellular compartments. In most cases tumors are of the epithelial type (90%), which is thought to be originates from the single-cell layer of peritoneal mesothelium that surrounds the ovary [83] Non-epithelial ovarian cancers are less frequent (10%) and include malignant germ cell tumors (derived from germ cells) and malignant sex cord tumors (derived from ovarian stroma)[84].

The tumor classification is achieved by immunohistochemistry with panels of antibodies directed at specific cell proteins or surface markers, in combination with microscopic tumor evaluation and radiologic assessment whereas the patient prognosis is scored upon the histological subtype evaluation, tumor differentiation grade, and residual disease after surgical treatment [84].

2.2 Ovarian Cancer and Mesothelioma: Common Origin, Common Features

It is thought that epithelial tumors such most of ovarian cancer and malignant mesothelioma may derive from the same cellular compartment, the mesothelium layer. Accordingly with this idea, it was found that among these both tumors, some common features were shared [75][85]. For example, the main risk factor for development of lung

mesothelioma is exposure to asbestos, and although less investigated, asbestos and talc have also been suspected as ovarian carcinogens, even if experimental and epidemiological evidence are not sufficient to establish a direct relationship [86].

Similarities at the level of tumor markers are also evident: 1) Mesothelin, a cell-surface protein identified as a mesothelioma-associated antigen, is highly expressed in mesothelioma, ovarian carcinoma and some other cancers, but only in a restricted number of normal tissues [85]. A soluble mesothelin-derived peptide has been detected in serum of patients with mesothelioma and ovarian carcinoma as well [87]. 2) Osteopontin was shown to detect mesothelioma in high risk individuals exposed to asbestos and in ovarian tumors, the osteopontin expression has been associated with peritoneal metastasis [88][89].

The mesothelioma signalling pathway has been less studied than ovarian cancer, but some of the known common characteristic among both tumors include activation of the PI3-kinase/AKT pathway [90][91] and metalloproteinases MMP2 and MMP9 [92]; overexpression of VEGF100, and IL-8; inactivation or loss of expression of PTEN and p16 ; and involvement of the Wnt- signaling pathway, among others [93-99].

2.3 TENEURINS IN OVARIAN CANCER

As summarized above (section 1.4.3), evidence exists that support a role for some teneurins in the development of human cancer. Furthermore, it has been hypothesized that malignant lung mesothelioma and epithelial ovarian cancer have a common origin, as supported by the molecular similarities found between tumors as well as by the overlapping in the morphologic features. In addition, Ziegler A et al. (2011) demonstrated the differential expression of teneurin-2 in MPM respect to lung ADCA, highlighting the potential of

teneurin-2 as mesothelioma-associated biomarker. On the basis of these findings, a preliminary screen was performed in a series of epithelial human cancer cell lines aimed to evaluate teneurin-2 mRNA expression in this type of cells (Ziegler A. et al., unpublished results). In agreement with publicly available microarray data (sources: Gene Expression Omnibus [100]; Oncomine Research [101]), Figure 2.1 shows the teneurin-2 mRNA expression in one ovarian cancer (SKOV3) and two-breast cancer cell lines (ZR75 and T47D), whereas teneurin-2 transcript was absent in two gastric carcinoma and one cervical carcinoma cell lines. Moreover, in breast and ovarian cell lines, teneurin-2 expression seems to be correlated with the expression of FGF8, which in chicken limb bud induce the teneurin-2 expression and that would be also involve in teneurin tumor expression.

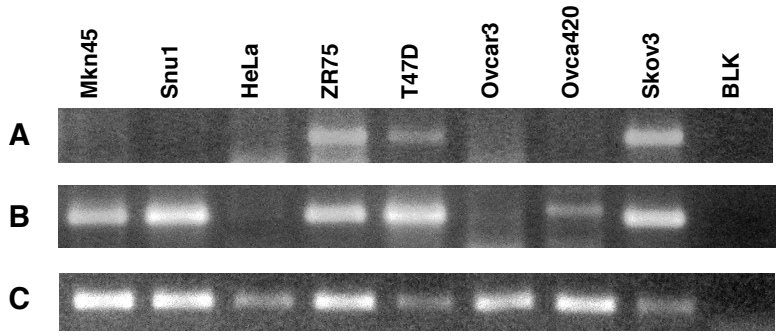

Figure 2.1. Teneurin-2 expression in human cancer cell lines. (A) ODZ2 and (B) FGF8 mRNA detection by RT-PCR amplifications. β2-microglobulin (C) was used as internal control. *SNU1, MKN45*: gastric cancer; *HeLA*: cervical cancer; *ZR75, T47D*: breast cancer; *Ovcar3, Ovcar420, SKOV3*: ovarian cancer cell lines; *BLK*, PCR negative control (Ziegler A. et al., unpublished results).

Thus, in addition to the prominent role of teneurins in development of nervous system, recent evidences suggests that some of these glycoproteins, most notably teneurin-2 and teneurin-4, are expressed and could play a role in some forms of human cancer, nevertheless, the functional role of teneurin expression in human cancers is currently and largely unknown.