ELSEVIER

Gene Expression Patterns 6 (2006) 952–963

GEP

www.elsevier.com/locate/modgep

Comparison of neurolin (ALCAM) and neurolin-like cell adhesion molecule (NLCAM) expression in zebrafish

Christopher J. Mann, Yaniv Hinits, Simon M. Hughes *

MRC Centre for Developmental Neurobiology and Randall Division of Cell and Molecular Biophysics, New Hunt's House, Guy's Campus, King's College London, London SE1 1UL, UK

Received 25 January 2006; received in revised form 19 March 2006; accepted 24 March 2006
Available online 30 March 2006

Abstract

Many immunoglobulin (Ig)-superfamily cell adhesion molecules influence skeletal muscle formation. In *Drosophila*, *dumbfounded* (*duf*/*kirre*), *irreC*, *sticks and stones* and *hibris* encode related Ig-family proteins expressed in subsets of neurons and muscle precursor cells. The family mediates cell migration, axon guidance and fusion of myoblasts. Despite the importance of these genes in invertebrate myogenesis, no obvious functional parallels are known in vertebrate myogenesis. Here we investigate the gene expression pattern and phylogenetic and protein–structural relationships of the *duf*-related molecules neurolin and neurolin-like cell adhesion molecule (NLCAM), members of the activated leukocyte cell adhesion molecule (ALCAM) sub-family of Ig-molecules. These proteins are among the closest to Duf/Kirre by sequence. During zebrafish development, *neurolin* is expressed in subsets of somite and muscle cells, heart and numerous sites of neuronal maturation. The new *ALCAM*-family member, *NLCAM*, appears to have arisen by duplication of *neurolin*/*ALCAM*. *NLCAM* is expressed widely during gastrulation, particularly in the nascent neural plate, but later becomes predominantly expressed in sites of muscle and nerve maturation and in the fin fold. The expression of each gene is often in groups of cells in similar parts of the embryo; for example, in the region of Rohon Beard neurons, trigeminal ganglion and fusing fast and migrating slow muscle fibres. However, expression can also be distinct and dynamic; for example, muscle pioneer fibres express *neurolin* but not *NLCAM* at high level. Both molecules are expressed in subsets of muscle precursors at times prior to fusion.

Keywords: Muscle; Zebrafish; Neurolin; NLCAM; ALCAM; CD166; BEN; JC7; DM-GRASP; SC1; zn-5; zn-8; MEMD; SB-10; MCAM; BCAM; MUC18; CD146; Kirre; Dumbfounded; Nephrin; CD6; NgCAM; Immunoglobulin superfamily; Motorneuron; Muscle pioneers; Brain; Placode; Rohon Beard; Fin fold; Hindbrain; Telencephalon; Retina; External cells; Gastrula; Gene expression

1. Results

1.1. Sequence and phylogeny analysis

The *Drosophila melanogaster* gene *dumbfounded* (*duf*) (Ruiz-Gomez et al., 2000), also known as *kin-of-IrreC* (*kirre*) (Strunkelnberg et al., 2001), encodes a 959 amino acid (aa) member of the Ig-superfamily with five extracellular Ig-like domains and a cytoplasmic tail. By sequence, *duf* is closely related to *irreC*/*rst*, located nearby on the *Drosophila* X-chromosome, and both function redundantly in myogenesis (Boschert et al., 1990; Ramos et al., 1993; Strunkelnberg et al., 2001). *Duf* is expressed early during gastrulation in the invaginating mesoderm and later exclusively in muscle founder cells, but not in fusion competent myoblasts (FCMs), during *Drosophila* myogenesis (Ruiz-Gomez et al., 2000). Based upon sequence data and the characteristic $VVC_2C_2C_2$ Ig-domain structure, early reports suggested that Duf protein may be a homologue of vertebrate nephrin (Ruiz-Gomez et al., 2000), or a member of the neurolin/ALCAM family (Dworak et al., 2001). It is now known that the Duf protein is not a close homologue of nephrin, as the latter has eight Ig-domains. *Nephrin* is, however, similar to other genes involved in myoblast

* Corresponding author. Tel.: +44 20 7848 6445; fax: +44 20 7848 6550.
E-mail address: simon.hughes@kcl.ac.uk (S.M. Hughes).

1567-133X/$ - see front matter © 2006 Elsevier B.V. All rights reserved.
doi:10.1016/j.modgep.2006.03.010

fusion, namely *Hibris* (*Hbs*) (Dworak et al., 2001) and *Sticks and stones* (*Sns*) (Fig. 1A; Bour et al., 2000). Members of the *ALCAM*-family have five extracellular Ig-like domains in the $VVC_2C_2C_2$ arrangement, similar to Duf (Swart, 2002; Swart et al., 2005). In addition to ALCAM itself, the family also includes *MCAM* (also known as *MUC18* or *CD146*) (Lehmann et al., 1987; Lehmann et al., 1989), and *BCAM* (also known as the Lutheran blood group antigen; Campbell et al., 1994; Udani et al., 1998). Moreover, recent studies have identified several members of a family of Kirre-like (KIRREL; also, confusingly, known as NEPH) proteins in vertebrate species which also have $VVC_2C_2C_2$ structure (Donoviel et al., 2001; Sellin et al., 2003; Sun et al., 2003). To determine the relationship of *duf* to *neurolin*/*ALCAM* and other family members, we performed phylogenetic sequence comparisons. Duf/irreC are roughly equally divergent from the vertebrate nephrin, neurolin/ALCAM sub-families and only marginally more related to the KIRREL sub-family, in which we found several other predicted zebrafish genes (Fig. 1A). Due to reported *neurolin* expression in fish muscle, we focussed on this family to ascertain any possible role in myogenesis and/or the existence of founder-like myoblasts.

The literature on neurolin/ALCAM and family members is complicated; so a brief outline is necessary. Neurolin was the name given to a 86 kDa protein cloned from goldfish (*Carassius auratus*) using monoclonal antibody E21 (Paschke et al., 1992). Subsequently, two groups reported the cloning of the zebrafish homologue (Kanki et al., 1994; Laessing et al., 1994). Neurolin was identified as the fish homologue of a cell surface glycoprotein reported by many groups under a myriad of names in different species, including SC1 (Tanaka and Obata, 1984; Tanaka et al., 1989; Tanaka et al., 1991), BEN (Pourquie et al., 1990; Pourquie et al., 1992), DM-GRASP (Burns et al., 1991), JC7 (el Deeb et al., 1992), ALCAM/CD166 (Bowen et al., 1995; Bowen et al., 1997), SB-10 (Bruder et al., 1998) and MEMD (Degen et al., 1998). For the sake of clarity, we henceforth refer to the zebrafish gene as *neurolin* and use the name ALCAM (for activated leukocyte cell adhesion molecule) to refer to both the gene family, and to close homologues of neurolin in other vertebrate species.

The most characterised roles of *ALCAM*-family molecules are in T- and B-cell activation, tumour metastasis and axon guidance and fasciculation, which appear to be mediated through either homophilic adhesions or heterophilic interactions with the ligands CD6 or NgCAM (Corbel et al., 1992; Bowen et al., 1995; Skonier et al., 1997; Fashena and Westerfield, 1999; Ott et al., 2001; Swart et al., 2005). Mice lacking ALCAM show no morphological defects, but subtle misdirection and poor fasciculation of axons suggest some parallels with Duf/Kirre function (Weiner et al., 2004). A detailed mRNA expression profile of *neurolin* in embryonic zebrafish is not available, despite excellent description at some stages and locations (Laessing and Stuermer, 1996; Fashena and Westerfield, 1999; Thisse and Thisse, 2004). Neurolin investigations in zebrafish have often relied upon the monoclonal antibody zn-5/zn-8; which are duplicate isolates of the same hybridoma (Trevarrow et al., 1990; Fashena and Westerfield, 1999; Westerfield, 2000; Kawahara et al., 2002).

We searched the Sanger zebrafish genome database (Zv3-Zv5; www.ensembl.org) with the Duf protein sequence and confirmed neurolin as a closely related match. Additionally, we identified another homologue with a high degree of similarity to neurolin that was annotated in the database as *neurolin-like cell adhesion molecule* (*NLCAM*). To our knowledge, the only other data on *NLCAM* is a partial expression pattern (Rauch et al., 2003). No other significant Neurolin/NLCAM homologues were identified and BLAST searching mammalian genomes identified only a single gene, ALCAM (data not shown). We identified a predicted zebrafish MCAM homologue, and this gene clusters together with BCAM into the ALCAM sub-family (Fig. 1A).

Neurolin is located on chromosome 10 (Unigene Dr.20912) whereas *NLCAM* is located on chromosome 15 (Unigene Dr.36577; Dr. Mario Caccamo, Wellcome Trust Sanger Institute, pers. commun.). The syntenic relationship of the three adjacent known genes either side of zebrafish *neurolin* and *NLCAM* and human and mouse *ALCAM* is shown (Fig. 1B). The evolutionary retention of the *Cblb* gene next to *neurolin*/*ALCAM* emphasises that *neurolin* and *ALCAM* are homologues, whereas *NLCAM* is a paralogue, as predicted by the Inparanoid program (inparanoid.cgb.ki.se). We searched for possible Ig-family molecules in the human genome in the vicinity of *DYRK1A* and *FBXL16*, the nearest known neighbours of zebrafish *NLCAM*, but found none (data not shown). The human and mouse *ALCAM* loci also show strong conservation with one another. We cloned and sequenced *neurolin*/*NLCAM* cDNAs (see Section 2; Fig. 1C). RT-PCR analysis of total RNA extracted from 18 som and 28 som stage embryos revealed a single band for both *neurolin* and *NLCAM* at each stage, suggesting no alternative splicing in the region amplified and stages examined (data not shown). Comparing the available and acquired exonic organisation of *neurolin* and *NLCAM*, we observed a high degree of conservation in exon size (Fig. 1C). This finding suggests that the genes probably arose from a common *ALCAM* ancestor by duplication, although the lack of synteny makes it unclear whether this was a genome-wide duplication in zebrafish/teleosts (Taylor et al., 2001).

Neurolin encodes a 564aa protein (Kanki et al., 1994; Laessing et al., 1994), whereas *NLCAM* is predicted to have 562 residues (Fig. 1D). The predicted protein has four functional domains: an N-terminal signal peptide, the extracellular $VVC_2C_2C_2$ region with several potential glycosylation sites, a transmembrane domain and a C-terminal cytoplasmic tail (Paschke et al., 1992; el Deeb et al., 1992; Laessing et al., 1994). Neurolin and NLCAM are highly similar (44.6% identical at the protein level), but diverge considerably from Duf and IrreC (Fig. 1A). As, to our knowledge, the zn-5/zn-8 epitope on neurolin is not known (Trevarrow et al., 1990), it is possible that studies with this monoclonal antibody have