



HAL
open science

MEGF9: a novel transmembrane protein with a strong and developmentally regulated expression in the nervous system

Ulrike Brandt-Bohne, Douglas R Keene, Fletcher A White, Manuel Koch

► **To cite this version:**

Ulrike Brandt-Bohne, Douglas R Keene, Fletcher A White, Manuel Koch. MEGF9: a novel transmembrane protein with a strong and developmentally regulated expression in the nervous system. *Biochemical Journal*, 2006, 401 (2), pp.447-457. 10.1042/BJ20060691 . hal-00478592

HAL Id: hal-00478592

<https://hal.science/hal-00478592v1>

Submitted on 30 Apr 2010

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

EGFL5: a novel transmembrane protein with a strong and developmentally regulated expression in the nervous system

Ulrike BRANDT-BOHNE^{*}, Douglas R. KEENE[#], Fletcher A. WHITE[§] and Manuel KOCH^{*+†1}

^{*}Center for Biochemistry, [†]Center for Molecular Medicine Cologne, and [‡]Department of Dermatology, Medical Faculty, University of Cologne, D-50931 Cologne, Germany, and [#]Shriners Hospital for Children Research Center, Portland Oregon, 97329, USA, and [§]Department of Cell Biology, Neurobiology and Anatomy, Loyola University Medical Center, Maywood, IL 60153, USA

Short title: EGFL5, a novel transmembrane protein

¹Corresponding author:

Manuel Koch,

Center for Biochemistry, Medical Faculty, University of Cologne,

Joseph-Stelzmann-Str. 52, D-50931 Cologne, Germany,

Tel: +49 221 478 6990; Fax: +49 221 478 6977;

E-mail: Manuel.Koch@uni-koeln.de

SYNOPSIS

EGFL5, a novel transmembrane protein with multiple Epidermal Growth Factor (EGF)-like repeats is predominantly expressed in the developing and adult central and peripheral nervous system. The domain structure of EGFL5 consists of an N-terminal region with several potential O-glycosylation sites followed by five EGF-like domains, which are highly homologous to the short arms of laminins. Following one single pass transmembrane domain, a highly conserved short intracellular domain with potential phosphorylation sites is present. The protein was recombinantly expressed and characterized as a tissue component. To further study the expression pattern, immunohistochemistry was performed and staining was detected in Purkinje cells of the cerebellum and in glial cells of the peripheral nervous system. Additional expression was observed in the epidermal layer of skin, papillae of the tongue and the epithelium of the gastrointestinal tract. By immunoelectron microscopy, EGFL5 was detected in glial cells of the sciatic nerve facing the basement membrane. EGFL5 represents a novel putative receptor, expressed in neuronal and non-neuronal tissues, that is regulated during development and could function as a guidance or signaling molecule.

Keywords: EGFL5, transmembrane protein, nervous system, glycosylation, glial cell, skin.

INTRODUCTION

EGF-like domains are major modular components present in many proteins of the extracellular matrix (ECM). These modules are involved in cell adhesion, receptor-ligand interactions and tissue repair [1-3]. The EGF molecule is a small protein of 53 amino acids with three internal disulfide bridges stabilizing its native structure. It consists of a small domain folded to an S-shaped structure and two short stretches of anti-parallel β -sheet. The localization of cysteine and glycine residues and the distance between them define its conformation [2, 3]. Homologous sequences of about 40 amino acids in one or more copies have been described in a large number of proteins [3]. A common feature is that this domain is present in the extracellular region of membrane bound or in secreted proteins, with the exception of the intracellular prostaglandin G/H synthase. In addition to these prototype EGF-like domains, laminins contain laminin type EGF-like (LE) domains, consisting of about 60 amino acids. The tertiary structure is remotely similar at its N-terminal to that of the classical EGF-like module. The two domain types differ in their number of conserved cysteines; six in classical EGF-like and eight in laminin-type EGF-like domains. The sequences between conserved cysteines vary in length [2-4].

Proteins containing EGF-like domains are thought to be involved in the development, maintenance and injury response of the nervous system. Candidates for such processes are the Notch family of proteins, reelin, slit and tenascin-C [5-8] as well as laminins, which play a role in supporting the survival

of neurons [9]. Mutations in these proteins have drastic effects like hypertrophy of the nervous system in the case of Notch or its ligand Delta in *Drosophila* or zebrafish [10, 11] or cerebellar and cerebral cortex disorders found in the *reeler* mice [6]. Accordingly, human disorders like Marfan syndrome (mutations in fibrillin-1; [12]), cerebral autosomal dominant arteriopathy with subcortical infarcts and leukoencephalopathy (Notch-3; [13]), Alagille syndrome (jagged 1; [14]), and familial hypercholesterolemia (LDL receptor; [15]) arise from mutations in EGF-like domains. Furthermore, EGF-like domain containing proteins are also involved in connective tissue function and, for example, severe skin alterations (*cutis laxa*) can be observed if fibulin-5 is miss-expressed [16]. Mutations in laminin chains (e.g. laminin α 3, β 3, γ 2) result in skin diseases such as junctional epidermolysis bullosa [17]. The identification and characterization of further proteins containing multiple EGF-like repeats is likely to help our understanding of cell-cell interactions or ligand-receptor interactions in the nervous system and in other tissues. In this study, we describe the initial identification, characterization and expression profile of a novel transmembrane protein, called EGFL5.

EXPERIMENTAL

Northern blot analysis

A 724-bp PCR product (forward 5'-dgacccaactgccgagtc-3' and reverse 5'-dttttaattgggtccacatggcc-3') was labeled with ^{33}P -dCTP (NEN Life Science Products) using the rediprime DNA labeling system (GE Healthcare). Northern blots (CLONTECH Laboratories, Inc.) were prehybridized in 50% formamide, 5x SSPE, 1x Denhardt's, 1% SDS, 10% dextran sulfate, and 0.1 mg/ml salmon sperm DNA (Invitrogen) at 42°C for 2 h. Without further purification, the probe was denatured in the same buffer plus 1/10 vol/vol human Cot-1 DNA (Boehringer Mannheim), and 1/10 vol/vol sheared salmon testis DNA (Invitrogen) at 94°C for 5 min, placed on ice, added to the blots, and hybridized for 20 h. Blots were washed three times in 2x SSC, 1% SDS at 42°C and two times in 0.1x SSC, 1% SDS at 42°C. Blots were placed on BioMax MR film with a BioMax TranScreen-LE intensifying screen (both from Eastman Kodak Co.) for 20 h at -70°C.

RT-PCR

RNA was isolated from animal tissues using the RNeasy kit (QIAGEN), and 1 μ g of total RNA was reverse-transcribed using an RT-PCR kit (CLONTECH). PCR was performed on these cDNAs using a long expand PCR kit (Boehringer Mannheim) with GADPH primers (forward: 5'-dtgaaggtcggtgtgaacgga-3'; reverse: 5'-dgatggcatggactgtgtgca-3') or actin primers (forward: 5'-gccattcaggcgggtgctgtc; reverse: 3'-gtgctaggtgctagggctgt) and the amount of template was normalized for each tissue. A range of cycle numbers was tested to ensure that the amounts were normalized in the linear range of the reaction. With the gene-specific primers (forward: 5'-dcccgccacccttcggaggacgg-3'; reverse: 5'-dggccatgtggaccaattaaactcc-3'), another PCR was performed on the normalized cDNA templates.

Expression and purification of recombinant EGFL5

The EGFL5 cDNA was generated by PCR on full length cDNA derived from mouse brain tissues using specific primers to amplify a 78 kb fragment coding for the five EGF-like domains of EGFL5. The primers used were: forward 5'-gaggctagccatattgttaacttctgaggttc-3' and reverse 5'-ccaggatccttacttgatgcagtttctctgaagat-3'. The sequenced construct was cloned into a modified pCEP-Pu expression vector carrying an N-terminal 6xHis tag and a thrombin cleavage site and the plasmid was transfected into EBNA 293 cells. Protein was purified from the collected cell culture supernatant using CoCl₂-Sepharose (CLONTECH Laboratories, Inc.) and eluted with increasing imidazole concentrations in 150 mM NaCl, 50mM Tris-HCl, pH 8. Finally, the protein was dialyzed against PBS and stored at -20°C.

Preparation of specific EGFL5 antibodies

The purified recombinant protein was used to immunize a rabbit and a guinea pig. The antisera were affinity purified on CNBr-activated Sepharose (GE Healthcare) with recombinant EGFL5 protein coupled. The specific antibodies were eluted with 150 mM NaCl, 0.1 M triethylamine, pH 11.5, and the eluate neutralized with 1 M Tris-HCl, pH 6.8 and dialyzed against PBS before freezing. The antibodies are called pAbKR18 (rabbit origin) and pAbKG20 (guinea pig origin).

Rotary shadowing and electron microscopy

For rotary shadowing experiments the protein was dialyzed against 150 mM ammonium bicarbonate. Samples were diluted to a final concentration of 70% (v/v) glycerol, sprayed onto freshly cleaved mica, and dried under vacuum (Balzers BAE 250 evaporator). Rotary shadowing was performed as described previously (Sakai et al., 2004). Replicas were examined at 80 kV in a transmission electron microscope (Philips 410LS). The protein images were measured using the imageJ (NIH, freeware) software.

Tissue preparation and extraction

Mice were sacrificed to collect tissues for analysis. Brains were dissected, weighed, extracted by homogenization in chilled buffer (1x complete, 0.2 mM EDTA, 1% NP40), and incubated on ice for 30 min. Homogenates were centrifuged at 4°C at 13000rpm for 15 min. Supernatants were mixed with 5x SDS-PAGE sample buffer supplemented with 4% β-mercaptoethanol and applied to a 10-12% SDS-polyacrylamide gel.

Removal of N-linked glycans

PNGase F (Biolabs) was used to digest purified proteins and tissue extracts according to manufacturers specifications.

SDS-polyacrylamide gel electrophoresis and immunoblotting

SDS-PAGE was performed as described [18]. For immunoblots the proteins were transferred to nitrocellulose, blocked in TBS containing 0.05% Tween and 3% low fat milk powder. The blots were incubated with the respective antibodies diluted in TBS containing 0.05% Tween and 3% low fat milk

powder for 2 h at room temperature or over night at 4°C. Bound antibodies were detected by luminescence using peroxidase-conjugated goat anti-rabbit IgG (Sigma), goat anti-mouse IgG (DAKO) or goat anti-guinea-pig IgG (DAKO).

Tissue preparation, frozen sections and immunohistochemistry

Tissues were embedded in Tissue Tek mounting medium, frozen in nitrogen vapor and 10 µm frozen sections cut. The dried sections were permeabilized in ice cooled methanol for 1 min, washed and blocked for 1 h at room temperature with 5% normal goat serum (NGS) in PBS containing 0.5% Tween 20. Primary antibodies were diluted in blocking solution and incubated on the sections over night at 4°C. For double immunofluorescence labeling, sections were washed and incubated with a second primary antibody for 2 h at room temperature. Primary antibodies were the affinity purified rabbit polyclonal antibody KR18 and guinea pig polyclonal antibody KG20 recognizing EGFL5, rat monoclonal anti-laminin γ2 (Chemicon), monoclonal mouse anti-neurofilament heavy chain (AbCAM) and polyclonal rabbit anti-GFAP (DAKO). Sections were washed with PBS and incubated with the detection antibody carrying a Cy3 (Cy3 goat anti-rabbit IgG or Cy3 goat anti-guinea pig IgG, Sigma) or Alexa 488 (Alexa 488 goat anti-mouse IgG, Alexa 488 goat anti-rat IgG, Alexa 488 goat anti-rabbit IgG, Invitrogen) fluorochrome for 2 h at room temperature. Alexa 488 conjugated lectin from Griffonia simplicifolia (Molecular Probes) was incubated for one hour at room temperature on sections treated with 4% paraformaldehyde. Sections were washed and mounted in Mowiol mounting medium.

Nerve whole mount preparations

Nerve whole mount preparations were performed with adult tissues. Nerves were dissected, treated with clostridial collagenase (3.5 mg/ml, Sigma) and dispase II (Roche Diagnostics) for 20-30 min at 37°C, separated on collagen coated coverslips, fixed with 2% paraformaldehyde for 5-10 min at room temperature and washed before immunohistochemical staining.

RN2 schwannoma cell culture and staining

RN2 rat schwannoma cells were plated and grown on chamber slides for 2 days in 10% fetal calf serum (FCS) containing DMEM/F12 medium (Invitrogen). The primary antibody was incubated with the cells for 30 min at room temperature. Cells were washed, fixed with 2% PFA for 2-5 min and permeabilized with 0.2% Triton for 1-5 min at room temperature. Cells were blocked with NGS for subsequent secondary antibody incubation. Nuclear staining was performed using bisbenzimidine (0.001mg/ml in methanol diluted in PBS, Serva), whereas the F-actin cytoskeleton of RN2 cells was detected by FITC-conjugated phalloidin (Molecular Probes).

Immuno-electronmicroscopy

Mouse sciatic nerve was diced and immersed in primary KG20 antibody diluted in serum free DMEM medium overnight at 4°C. The tissue was washed in DMEM for 4 h and immersed in a suspension of 1-nm gold labeled secondary antibody in medium overnight at 4°C. Following an extensive

wash in medium, the tissue was immersed in gold enhancement solution, rinsed, fixed in 1.5% glutaraldehyde/1.5% paraformaldehyde with 0.05 % tannic acid, osmicated, and prepared for transmission electron microscopy by a standard protocol.

RESULTS

Identification and characterization of the EGFL5 genes

By screening the EST (expressed sequence tags) database for Epidermal Growth Factor like (EGF-like) domains, we identified a murine sequence with high homology to the laminin β 2 chain and cloned the gene. The encoded type I single pass transmembrane protein, named EGFL5, contains five EGF-like domains.

The EGFL5 genes map to syntenic regions in the human and mouse genome on chromosome 9 (9q32-q33.3) and 4 (CA2), respectively. At the 5'-end, the human and murine genes are flanked by the CDK5 regulatory subunit associated protein 2, the DBC1 (deleted in bladder cancer 1) gene, the TLR4 (toll-like receptor 4), and astrotactin 2 genes. The latter is already located in the adjacent human chromosomal region 9q31 and in murine chromosomal region 4CA1. There are no further genes downstream in the same syntenic region.

The human (XM_376905) and mouse (NM_172694) genomic sequences are fully annotated in the public database. The organization of the exon/intron size, length and codon phase is very similar in both species. The human and mouse EGFL5 genes consist of 6 exons and span over 113.66 kb and 106.09 kb, respectively. The corresponding cDNAs are about 6298 bp (human) and 6146 bp (murine) long (Figure 1). Exon 1 encodes the N-terminal signal peptide sequence and a region of unknown function, whereas exons 2-5 code for EGF-like repeats 1-5, and exon 6 a transmembrane region and a cytoplasmic tail. A large portion of exon 6 encodes an unusually long untranslated region of 4326 bp and at the 3' end a conserved polyadenylation site is present.

Analysis of EGFL5 domain structure

EGFL5 is highly conserved between different species in vertebrates; however no homologous gene can be located in the *C. elegans* or *D. melanogaster* genome. The murine sequence is highly homologous (68%) with its human counterpart (Figure 2) and the proteins have an identical domain structure. The protein sequence consists of 600 amino acid residues in both mouse and human, which yields a calculated mass of 63 kDa. An N-terminal signal peptide (amino acids 1-34) is followed by a stretch of 168 amino acids. This region is highly homologous to the EGFL5 sequences in rat, chimpanzee and human (Supplementary Figure 1A).

As indicated by its name, EGFL5 contains five EGF-like domains (EGF1: amino acids 202-251, EGF2: 252-298, EGF3: 299-346, EGF4: 347-397 EGF5: 398-449). They match the EGF motif consensus sequence CxCxxxxGxxC, where x represents any residue, G an often conserved glycine, and C indicates conserved cysteines, providing disulfide bridges. The EGF-like domains of EGFL5 are 47-52 amino acids in length, contain eight cysteines and are similar to the LE domains found in laminins. Multiple alignments of all five EGF-like motifs were performed using the BlastP algorithm with the blosum62 matrix. In an end-weight pileUp run, the consensus sequence of conserved cysteine and glycine residues in all five EGF-like domains is shown (Supplementary figure 1B). In the human orthologue, the EGF-like domains share 81% sequence homology with the ones found in mouse. Pronounced sequence similarities are found to the laminin β 1 and β 2 chains while other members of the laminin group, such as the laminin α 1, -2, -3, and -4 chains as well as laminin γ 1 and γ 3 chains display weaker homologies to EGFL5. The shared regions are restricted to the EGF-like domains of EGFL5, and multiple alignments are shown in supplementary figure 2A for the murine EGF like domains 1-5.

A stretch of 60 amino acids, which displays no significant homology to any other protein, might be a putative linker region and separates the EGF-like domains from the transmembrane region. The short transmembrane region comprises 21 amino acids, followed by a highly conserved cytoplasmic stretch. This intracellular region is almost identical to the sequence of the human EGFL5 (98%) with only one arginine being replaced by a glycine. This unique sequence is highly conserved between the orthologous genes in different species (Supplementary figure 2B).

Prediction of posttranslational modifications

Putative posttranslational modifications can in part be predicted by analyzing the amino acid sequence for motifs similar to those described in other proteins. Eight putative N-glycosylation sites were detected and are indicated by asterisks in figure 1 (NetNGlyc 1.0 server), with six being present in the EGF-like domains. Numerous putative glycosylation sites for O-linked glycans (NetOGlyc 3.1 server) were predicted within the N-terminal region and for the C-terminal stretch between the EGF-like domains and the transmembrane region (dotted lines in figure 1), whereas no such O-GalNAc (mucin-type) glycosylations were predicted for the EGF-like domains or the intracellular region. In contrast, YinOYan 1.2 prediction results for O-(β)-GlcNAc sites, mainly found in intracellular or nuclear proteins, indicate a putative site at position T591, but the prediction level is just above threshold. This same threonine is identified by the motif scan results (CTAIL) with medium stringency, indicating a potential proline dependent serine/threonine kinase group for the cyclin dependant kinase-5 (CDK5). Searching with low stringency results in additional predicted sites possibly recognized by calmodulin dependent kinase-2 (T590 and S559).

The amino acid composition of murine and human EGFL5 shows 42 cysteine residues in the sequence, representing 7% of all amino acids. 40 out of these 42 cysteines are located in the EGF-like domains; none are present in the N-terminal region of unknown function and in the intracellular domain, while two occur in the signal peptide.

The C-terminal region contains frequent serine (12.5%) and threonine (16.7%) residues, which may serve as putative O-glycosylation sites as in mucins. The amino acid composition of this unknown region shows that 75.6% out of the total 186 amino acids are accounted for by six amino acids, serine (12.5%), threonine (16.7%), proline (20.8%), glycine (8.3%), alanine (11.9%) and valine (5.4%). The proportions of positively charged residues (i.e. arginine and lysine) in the murine and human intracellular EGFL5 stretches are 16% and 15% respectively, but these are not clustered. Three putative phosphorylation sites are predicted in the intracellular region (S561, S562, and Y563) (NetPhos2.0, Supplementary figure 2B), which are present in both mouse and human EGFL5. Other programs confirm this prediction indicating the SYHD sequence as a possible phosphorylation site for casein kinase II (CK2). In addition, one predicted phosphorylation site is located in the short stretch separating the EGF-like domains from the transmembrane region. Finally, the hydrophobicity plots of both human and mouse EGFL5 show hydrophobic stretches in the N-terminal region, which corresponds to the signal peptide and the C-terminal transmembrane region (www.bioinformatics.weizmann.ac.il).

EGFL5 gene expression

To determine the length of EGFL5 mRNA we performed northern blot hybridization with mRNA derived from different embryonic mouse tissues and developmental stages (Figure 3A and B). The first expression of EGFL5 mRNA was detected at embryonic day E11. A variety of tissue (heart, brain, spleen, kidney and testis) gave a clear band of 7.0 kb, whereas weaker hybridization signals were found in lung and liver, and no signal in skeletal muscle. RT-PCR analysis confirmed gene expression in several regions of the brain and weaker expression in dorsal root ganglia, skin and keratinocytes (Figure 3C).

EGFL5 does not seem to be alternatively spliced, since only one single band was detected by northern blot analysis and since no alternative EST clones were identified. In silico analysis of different EST clones from expressed single sequence tag gene bank shows that EGFL5 is expressed in embryonic and adult tissues. Several clones originated from nervous tissues, especially from diencephalon, eye and ganglia. In non-neuronal tissues, numerous transcripts can be found in pancreas, lymph node and spleen, whereas few clones are from genital tracts, heart, bone and thymus.

Recombinant EGFL5 expression and antibody production

In order to characterize the EGFL5 protein, a 930 bp insert (EGFL5-EGF), which encodes the five EGF-like domains, was amplified using specific primers (see Material and Methods). The construct was

transfected into EBNA 293 cells and the expressed protein purified. SDS-PAGE of the purified EGFL5-EGF protein showed a single band at ~75 kDa. Rotary shadowing experiments confirmed the homogeneity of the sample and showed 12.56 nm +/- 1.3 nm long rod like particles with some bends (Figure 4A). The recombinant EGFL5-EGF was used for immunization of a rabbit and a guinea pig. The affinity purified polyclonal rabbit KR18 and guinea pig KG20 antibodies were highly specific and recognized the recombinant protein as well as the full-length protein in tissue extracts of brain as a single band (Figure 4B and C).

The mass of the purified EGFL5 ectodomain protein by SDS-PAGE (~75 kDa) was much higher than the 33 kDa predicted from the sequence, while MALDI-TOF mass spectrometry gave a peak at 41 kDa and after PNGase F digestion 32 kDa. Similarly, the full length protein detected by immunoblot in murine brain extracts showed a mass of ~180 kDa, in contrast to the calculated 63 kDa. The presence of N-linked glycans, predicted by sequence comparison, was confirmed by N-glycosidase F digestion, which resulted in a shift of ~30 kDa in both the recombinant EGFL5-EGF and in full length EGFL5 from brain tissue extracts (Figure 4B and C).

No Ca²⁺-binding EGF-like modules are predicted in EGFL5, and recombinant EGFL5-EGF as well as full length EGFL5 detected in brain tissue extracts show no difference in their electrophoretic mobility under either reduced or non-reduced SDS-PAGE (Figure 4D and E).

EGFL5 expression in the central nervous system

Immunohistochemistry using affinity purified antibodies raised against the recombinant EGFL5-EGF fragment showed EGFL5 protein expression throughout the E18.5 embryonic brain and a more restricted expression by postnatal day 1 (Figure 5A and B). Notably, higher protein levels were found in the outer layers of the cortex than in inner layers. High expression was detected in the CA1 region, the stratum radium, and the oriens layers of the hippocampus, whereas the nuclear layers in CA1 and CA2 were negative (Figure 5E). Additional signals were found in the diencephalon, superior and inferior colliculus, cerebellum, facial nucleus and dorsal medulla oblongata of embryonic and newborn brain. Cranial sensory ganglia and nerves also express EGFL5 (Figure 5A and B). The cerebellar expression becomes more limited late in development (Figure 5A-C) and EGFL5 is subsequently expressed only by Purkinje cells of the adult cerebellum (Figure 5D). Staining was also detected in the olfactory bulb and olfactory epithelium (Figure 5F and G).

In eyes from the stage E12.5, positive signals were seen only in the lens (Figure 5H), whereas on postnatal day 1 (P1) EGFL5 was present throughout the lens and retinal ganglion cell layer (Figure 5I). In the adult eye, EGFL5 expression was confined to interneural layers of the retina, including the inner and outer plexiform layers, the outer region of the photoreceptor layer and the Bruchs membrane. No signal was detected in the nuclear layers or in the pigment epithelium (Figure 5J).

In the spinal cord, EGFL5 was present throughout the lamina of the dorsal horn and focally in the ventral commissure (Figure 5K). Little to no expression was detected in the ventral horn of the spinal cord.

EGFL5 expression in the peripheral nervous system

EGFL5 expression is prominent in both sensory ganglia (cranial and somatosensory) and associated peripheral nerves. Sections through embryonic mice showed strong staining in the trigeminal ganglia, dorsal root ganglia and thoracic sympathetic chain ganglia (Figure 5A, B and K-N). The immunopositive cells within these sensory ganglia are likely to represent myelinating and non-myelinating Schwann cells and satellite glia cells. Prominent staining was also present throughout the length of the peripheral nerves, including the trigeminal nerve (see also Figure 5A and B). In co-staining experiments, an overlap was seen with laminin present in basement membranes of Schwann cells ensheating peripheral nerves (not shown). EGFL5 was also detected in whole mount preparations of sciatic nerves and in sections depicting afferent fibers of the tongue and the whiskers (Figure 6J and 7D).

To determine the nature of EGFL5-immunopositive peripheral glial cells, further immunostaining was performed on both longitudinal and transverse sections of sciatic nerves. Co-staining with a glia specific lectin showed that EGFL5 occurs in structures around dorsal root nerve cells, representing associated glia cells (Figure 6A and B). EGFL5 showed a colocalization with laminins containing the $\gamma 1$ chain (Figure 6C) and was present on glia throughout peripheral nerves, as seen with antibodies against neurofilament heavy chain and against GFAP (Figure 6D-F). Accordingly, EGFL5 is present only on cells surrounding nerve fibers and not in the actual nerve fibers. In cell culture, EGFL5 is expressed on rat RN2 schwannoma (Figure 6G and H) and on rat C6 oligodendroglial cells (Figure 6I). Immunogold labeling of sciatic nerves with the KG20 antibodies localized EGFL5 to the outer membrane of myelinating Schwann cells (Figure 6K and L).

EGFL5 expression on keratinocytes of the skin and tongue

EGFL5 was also detected in tongue and skin epithelium. In agreement with the lower EGFL5 mRNA levels in skin compared to brain, the immunofluorescence signal in skin was also weaker than those in the nervous system (Figure 7A). In skin, all epidermal layers express EGFL5 and strong staining was also seen within the external root sheath of hair follicles (Figure 7B and C). In the tongue, the lingual filiform papillae, which cover the entire anterior part of the back of the tongue and consist of cone-shaped structures with a core of connective tissue, are EGFL5 positive (Figure 7D and E). EGFL5 was also detected in the esophageal and intestinal epithelium (Figure 7F and G).

DISCUSSION

EGFL5, a novel transmembrane protein, was identified by screening the EST database for proteins containing EGF-like domains. Related proteins, containing multiple EGF-like domains, often play crucial roles in the development and maintenance of the nervous system. Even though the EGFL5 gene has been detected earlier and both the murine and the human sequences submitted to the database, no detailed characterization was performed up to date. The human cDNA of EGFL5 was identified in a search for high-molecular-weight proteins with multiple EGF-like motifs using motif-trap screening methods [19] and the mouse gene was identified and shown to be located on chromosome 4 [20].

We could demonstrate that the human and mouse EGFL5 genes are clearly orthologous. They are located on syntenic areas of the genome, the exon-intron organization is similar and the amino acid sequences are highly conserved (68%). On the mRNA level a single EGFL5 transcript was detected, indicating the absence of alternative splice variants. Translation of the cDNA sequence reveals a signal peptide sequence. The subsequent region shows no homology to any other sequences found in the database. The middle part of EGFL5 in the extracellular region shows homology to EGF like domains. Laminins also contain numerous EGF-like domains and represent heterotrimeric proteins consisting of three chains: one α , one β , and one γ chain [21]. By database analysis of this region, EGF repeats from the laminin β 2 chain show the highest degree of conservation at the amino acid level, followed by the laminin β 1 subunit, the α 1 and -2 chains, and the laminin γ 1 chain precursor. The characteristics of the laminin-type of EGF-like domains are that they contain eight instead of six conserved cysteines. Additional analysis of the EGFL5 EGF repeat reveals that no putative Ca-binding consensus sequences are present. Following the short transmembrane region, an unique intracellular stretch containing putative phosphorylation sites for signal transduction events are located at the C-terminus. This region is highly conserved between different species (about 98%).

Posttranslational modifications, such as glycosylations, were predicted at the amino acid level and could in part be verified experimentally by glycan digestion. Glycosylations have been described to be involved in folding and stability of proteins, but also in cell – cell and cell – matrix interactions [22]. Especially high amounts of O-linked glycosylations of the mucin type, commonly found on mucins, were predicted for the N-terminal region of EGFL5. A general feature of mucins is the presence of regions rich in threonine and/or serine whose hydroxyl groups are in O-glycosydic linkage with the oligosaccharides involved [23]. As shown for other mucin-type glycosylated proteins [24], the C-terminal region of EGFL5 is also rich in serine, threonine, glycine, alanine, and proline, which together account for more than three-fourths of the amino acid residues.

As large number of proteins that are not classical mucins carry mucin-type glycosylation [25] and are thought to be involved in recognition of other proteins or cells resulting in adhesion [26], EGFL5

might be involved in similar processes. High levels of EGFL5 transcripts and protein were detected in neuronal and glial cells of the central and peripheral nervous system. Classical mucins are thought to be absent in the nervous system, while other proteins with a mucin-type glycosylation, e.g. dystroglycans, are expressed in brain and nerves [27] and have been implicated in myelination and synaptogenesis [28]. Deglycosylation of dystroglycans lead to an impairment of ligand binding and causes muscular dystrophies. In addition to the N-linked glycans predicted in the region of unknown function, six additional putative sites are present in the EGF-like motifs. We could verify the presence of N-linked glycans on both recombinant and native murine full length EGFL5, as deglycosylation of EGFL5 followed by SDS-PAGE/ immunoblot analysis resulted in a shift in mass for both. Further glycosylation and other posttranslational modifications can be expected since the size of the protein after PNGase F digestion is still higher than the calculated mass.

We generated specific antibodies against the EGFL5 mouse protein which specifically reacted with one band in brain extracts. By immunohistochemistry the main expression of EGFL5 is confined to nervous tissues. It cannot only be detected in neuronal cells in the central nervous system, but also in the peripheral nervous system. Neural cell bodies and Purkinje cells in the cerebellum as well as nerve associated cell types express EGFL5 at their surface. The latter could be identified as Schwann cells by immunofluorescent staining of tissue sections and schwannoma cells also express EGFL5. By immunogold labeling and electronmicroscopy EGFL5 was detected on the outer membrane region of myelinated Schwann cells ensheating sciatic nerves.

Similarly to laminins, EGFL5 was shown to be present in a variety of tissues, especially during development. Whereas in some tissue, like the olfactory bulb and the retina, laminin and EGFL5 stainings do not overlap; in others a close localization was observed. Also in the central nervous system, the cellular sources of the two proteins appear to overlap partially, including glia and neurons. Laminins are detected in astrocyte cultures but are absent from oligodendrocytes [28], playing a role in axon outgrowth. Glia cells of the central nervous system also seem to express EGFL5, as staining of brain sections appear diffuse and can not always be assigned to specific neurons. In contrast to laminins, EGFL5 can be detected on cultured C6 oligodendroglia.

Glial cells control the survival of associated neurons in both *Drosophila* and mammals, but this control is dependent on the prior neuronal triggering of glial cell fate commitment and trophic factor expression. Differentiating glial cells and neurons define the course of nervous system development by reciprocal interactions between the two cell types [29]. Therefore, the expression pattern of specific proteins at a critical time point may play a crucial role. EGFL5 expression is regulated during development, with protein being detectable from around embryonic day E12.5 up to the adult stage. In mammals, it has been shown that many proteins containing EGF-like repeats play a role in both events. Similarly, early differentiating neurons and their associated glia rely on reciprocal signaling to establish

the basic axon scaffolds from which neuronal connections evolve. The importance of such interactive signaling events is illustrated by the action of glial transcription factors and of glial axon guidance cues such as netrin-1 [30] and slit [7], which together regulate the commissural crossing of pioneer axons at the neural midline.

In such events, the defining principle is the mutually reinforcement and dependent signaling that occurs in a network of developing neurons and glia [29]. EGFL5 could represent a guidance molecule, as its expression pattern is present during the critical time points in development. The commencing broad expression could represent a migration or differentiation signal, which becomes more confined to specific structures during development. In addition it could also play a role in maintenance and stabilization processes, as it still persists in adulthood. In addition, EGFL5 could be involved in myelination, as it is present in the basement membranes of myelinated Schwann cells in the peripheral nervous system.

ACKNOWLEDGEMENTS

We would like to thank Wilhelm Stoffel for the kind donation of RN2 schwannoma and C6 glial cells, and Markus Plomann and Mats Paulsson for the helpful discussion. This work was supported by the Center for Molecular Medicine Cologne, the Federal Ministry of Education and Research in Germany, an award from the Alexander von Humboldt Foundation, the Federal Ministry of Education and Research in Germany, and by Köln Fortune.

REFERENCES

- 1 Campbell I D , B. P. (1993) Epidermal growth factor-like modules Current Opinion in Structural Biology **3**, 385-392
- 2 Engel, J. (1998) EGF-like domains in extracellular matrix proteins: localized signals for growth and differentiation? FEBS Lett. **231**, 1-7
- 3 Appella, E. W., I.T.; Blasi, F. (1988) Structure and function of epidermal growth factor-like regions in proteins FEBS Lett. **231**, 1-4
- 4 Martin, G. R. (1988) Basement membrane proteins: molecular structure and function Adv. Protein Chemistry **39**, 1-50
- 5 Lasky, J. L. and Wu, H. (2005) Notch signaling, brain development, and human disease Pediatr. Res. **57**, 104R-109R
- 6 Fatemi, S. H. (2005) Reelin glycoprotein: structure, biology and roles in health and disease Mol. Psychiatry **10**, 251-7
- 7 Fan, X., Labrador, J. P., Hing, H. and Bashaw, G. J. (2003) Slit stimulation recruits Dock and Pak to the roundabout receptor and increases Rac activity to regulate axon repulsion at the CNS midline Neuron **40**, 113-27
- 8 Kusakabe, M., Mangiarini, L., Laywell, E. D., Bates, G. P., Yoshiki, A., Hiraiwa, N., Inoue, J. and Steindler, D. A. (2001) Loss of cortical and thalamic neuronal tenascin-C expression in a transgenic mouse expressing exon 1 of the human Huntington disease gene J. Comp. Neurol. **430**, 485-500
- 9 Chen, Z. L., Indyk, J. A. and Strickland, S. (2003) The hippocampal laminin matrix is dynamic and critical for neuronal survival Mol. Biol. Cell **14**, 2665-76
- 10 de Celis, J. F., Barrio, R., del Arco, A. and Garcia-Bellido, A. (1993) Genetic and molecular characterization of a Notch mutation in its Delta- and Serrate-binding domain in Drosophila Proc. Natl. Acad. Sci. U. S. A. **90**, 4037-41
- 11 Bernardos, R. L., Lentz, S. I., Wolfe, M. S. and Raymond, P. A. (2005) Notch-Delta signaling is required for spatial patterning and Muller glia differentiation in the zebrafish retina Dev. Biol. **278**, 381-95
- 12 Dowgiert, J., Sosne, G. and Kurpakus-Wheat, M. (2004) Laminin-2 stimulates the proliferation of epithelial cells in a conjunctival epithelial cell line Cell Prolif. **37**, 161-75
- 13 Haritunians, T., Boulter, J., Hicks, C., Buhrman, J., DiSibio, G., Shawber, C., Weinmaster, G., Nofziger, D. and Schanen, C. (2002) CADASIL Notch3 mutant proteins localize to the cell surface and bind ligand Circ. Res. **90**, 506-8
- 14 Pilia, G., Uda, M., Macis, D., Frau, F., Crisponi, L., Balli, F., Barbera, C., Colombo, C., Frediani, T., Gatti, R., Iorio, R., Marazzi, M. G., Marcellini, M., Musumeci, S., Nebbia, G., Vajro, P., Ruffa, G., Zancan, L., Cao, A. and DeVirgilis, S. (1999) Jagged-1 mutation analysis in Italian Alagille syndrome patients Hum. Mutat. **14**, 394-400
- 15 Webb, J. C., Sun, X. M., McCarthy, S. N., Neuwirth, C., Thompson, G. R., Knight, B. L. and Soutar, A. K. (1996) Characterization of mutations in the low density lipoprotein (LDL)-receptor gene in patients with homozygous familial hypercholesterolemia, and frequency of these mutations in FH patients in the United Kingdom J. Lipid Res. **37**, 368-81
- 16 Loeys, B., Van Maldergem, L., Mortier, G., Coucke, P., Gerniers, S., Naeyaert, J. M. and De Paepe, A. (2002) Homozygosity for a missense mutation in fibulin-5 (FBLN5) results in a severe form of cutis laxa Hum. Mol. Genet. **11**, 2113-8
- 17 Airene, T., Haakana, H., Sainio, K., Kallunki, T., Kallunki, P., Sariola, H. and Tryggvason, K. (1996) Structure of the human laminin gamma 2 chain gene (LAMC2): alternative splicing with different tissue distribution of two transcripts Genomics **32**, 54-64
- 18 Laemmli, U. K. (1970) Nature **227**, 680-685
- 19 Nakayama, M., Nakajima, D., Nagase, T., Nomura, N., Seki, N. and Ohara, O. (1998) Identification of high-molecular-weight proteins with multiple EGF-like motifs by motif-trap screening Genomics **51**, 27-34

- 20 Okazaki, N., Kikuno, R., Ohara, R., Inamoto, S., Koseki, H., Hiraoka, S., Saga, Y., Nagase, T., Ohara, O. and Koga, H. (2003) Prediction of the coding sequences of mouse homologues of KIAA gene: III. the complete nucleotide sequences of 500 mouse KIAA-homologous cDNAs identified by screening of terminal sequences of cDNA clones randomly sampled from size-fractionated libraries *DNA Res.* **10**, 167-80
- 21 Timpl, R. and Brown, J. C. (1994) The laminins *Matrix Biol.* **14**, 275-81
- 22 Haltiwanger, R. S. and Lowe, J. B. (2004) Role of glycosylation in development *Annu. Rev. Biochem.* **73**, 491-537
- 23 Gendler, S. J. and Spicer, A. P. (1995) Epithelial mucin genes *Annu. Rev. Physiol.* **57**, 607-34
- 24 Eckhardt, A. E., Timpte, C. S., Abernethy, J. L., Zhao, Y. and Hill, R. L. (1991) Porcine submaxillary mucin contains a cystine-rich, carboxyl-terminal domain in addition to a highly repetitive, glycosylated domain *J. Biol. Chem.* **266**, 9678-86
- 25 Perez-Vilar, J. and Hill, R. L. (1999) The structure and assembly of secreted mucins *J. Biol. Chem.* **274**, 31751-4
- 26 Melnick, M., Chen, H., Zhou, Y. and Jaskoll, T. (2001) An alternatively spliced Muc10 glycoprotein ligand for putative L-selectin binding during mouse embryonic submandibular gland morphogenesis *Arch. Oral Biol.* **46**, 745-57
- 27 Michele, D. E. and Campbell, K. P. (2003) Dystrophin-glycoprotein complex: post-translational processing and dystroglycan function *J. Biol. Chem.* **278**, 15457-60
- 28 Chiu, A. Y., Espinosa de los Monteros, A., Cole, R. A., Loera, S. and de Vellis, J. (1991) Laminin and s-laminin are produced and released by astrocytes, Schwann cells, and schwannomas in culture *Glia* **4**, 11-24
- 29 Jessen, K. R. (2004) Glial cells *Int. J. Biochem. Cell Biol.* **36**, 1861-7
- 30 Deiner, M. S. and Sretavan, D. W. (1999) Altered midline axon pathways and ectopic neurons in the developing hypothalamus of netrin-1- and DCC-deficient mice *J. Neurosci.* **19**, 9900-12

FIGURE LEGENDS

Figure 1: Exon/intron organization of the human and murine EGFL5 genes and domain structures of the EGFL5 proteins. The genes are about 110 kb long and consist of six exons. Potential N-glycosylation sites are marked by asterisks, whereas predicted O-glycosylation sites are marked by dotted lines. SP, signal peptide; TM, transmembrane; CT, cytoplasmatic tail.

Figure 2: Homology of the mouse and human EGFL5 amino acid sequences. The signal peptide (SP) is underlined, the unknown region is shown in dotted boxes, the EGF-like domains in closed boxes, and the transmembrane region in a dashed box. Identical amino acid residues are marked in grey. m, mouse; h, human.

Figure 3: Northern blot and RT-PCR analysis. A: Northern blot analysis of murine tissues detects EGFL5 mRNA as a single band at ~7.4 kb in brain. Weaker signals at the same size are present in non-neuronal tissues as heart, spleen, liver, kidney, and testis. B: mRNA levels are regulated during development. C: Semi-quantitative RT-PCR analysis depicts different levels of mRNA in neuronal tissues and in skin.

Figure 4: Structural analysis of EGFL5 protein. A: Rod like particles representing the recombinant EGFL5-EGF protein were detected by rotary shadowing electron microscopy. B: The affinity purified rabbit anti-EGFL5 antibody KR18 recognizes the recombinant protein at ~75 kDa in SDS-PAGE immunoblot. PNGase F digestion results in a mass shift of ~30 kDa. C: EGFL5 is detected as a single ~180 kDa band in brain extract that shifts to ~150 kDa upon PNGase F digestion. The electrophoretic mobility is analyzed under reduced or non-reduced condition in E: recombinant EGFL5-EGF and F: a brain tissue extract detected with the KR18 antibody by immunoblotting.

Figure 5: Protein expression in the central nervous system. EGFL5 expression is detected in all structures of (A) embryo (E18.5) and (B) newborn brain (P1). The strongest staining is in (C) cerebellum (E18.5), mesencephalon, hindbrain, and in the hippocampus (E: adult). (D) In the adult cerebellum expression is restricted to Purkinje cells (D; adult). Facial nuclei, dorsal medulla oblongata, trigeminal ganglia and nerves also express EGFL5 (A and B, respectively, see arrows), as do cells of the olfactory bulb (F; adult) and olfactory epithelium (G; P1). Further signals are found in the developing eye. (H) At stage E12.5 protein expression can be detected in the anterior lens and by P1, throughout the lens and inner retinal layers (I). (J) In adult eyes, the staining is restricted to specific cell layers, e.g. as the inner and outer plexiform layer, the photoreceptor layer and Bruchs membrane. (K) EGFL5 is expressed in

cells of the spinal cord, the dorsal horn, and the ventral root (E19.5). Additional staining is detected in dorsal root ganglia (L), thoracic sympathetic chain ganglia of E19.5 embryos (M), and DRGs of adult animals (N). Co, Cortex; Po, posterior thalamic nuclei group; DpMe, deep mesencephalic nuclei; Pn, pontine nuclei; Fn, facial nuclei; TriN, trigeminal nerve; TriG, trigeminal ganglia; hip, hippocampus; LPG, lateral paragigantocellular nuclei; Mo, medulla oblongata; Sp5C, spinal 5 nuclei caudal part; Cpu, caudate putamen striatum; DG, dentate gyrus of hippocampus; SC, superior colliculus; IC, inferior colliculus; C, cerebellum; PC, Purkinje cells of the cerebellum; CA1 and CA2, field CA1 and CA2 of hippocampus; Mi, mitral cell layer of olfactory bulb; Gro, granular cell layer of the olfactory bulb; lo, lateral olfactory tract; EPI, external plexiform layer of olfactory bulb; RGCl, retinal ganglion cell layer; IPL; INL, inner plexiform and nuclear layer; OPL; ONL, outer plexiform and nuclear layer; PRl, photoreceptor layer; BM, Bruchs membrane; PE, pigment epithelium, SC, spinal cord; DHSC, dorsal horn of spinal cord; VHSC, ventral horn of spinal cord; DRG, dorsal root ganglion; VR, ventral root; SCG: sympathetic chain ganglion.

Figure 6: EGFL5 protein expression in the peripheral nervous system. EGFL5 (red) staining is (A) present in satellite glia cells and (B) around dorsal root nerves co-stained with IB4 (green), a marker for non-myelinated fibers. (C) EGFL5 expression overlaps with that for laminins containing the $\gamma 1$ chain present in the basement membranes of Schwann cells. (D, E) Labeling of EGFL5 and neurofilament 200 kDa (green) in sciatic nerves shows that EGFL5 is expressed in myelinating Schwann cells. EGFL5 and (F) GFAP, expressed by subsets of Schwann cells, also show a limited overlap. In cell culture, (G, H) both RN2 schwannoma cells (H: costain for F-actin) and (I) C6 oligodendroglia cell lines are both positive for EGFL5. (J) EGFL5 expression in peripheral nerves is detected in whole mount preparations of trigeminal nerves. (K, L) Electron microscopy images of immunogold labeled sciatic nerve sections. EGFL5 is present close to the basement membrane of myelinating Schwann cells.

Figure 7: EGFL5 protein expression in skin and other keratinized tissues: On cryosections of embryonic and newborn tissues, EGFL5 is detected in (A-C) skin and around hair follicles (P1), (D, E) in keratinocytes of the tongue (P1), (F) oesophagus (E16.5), and (G) intestine (E16.5). TriG, trigeminal ganglia; TriN, trigeminal nerve; Wh, whiskers; Hf, hair follicle; olfE, olfactory epithelium; K, keratinocytes.

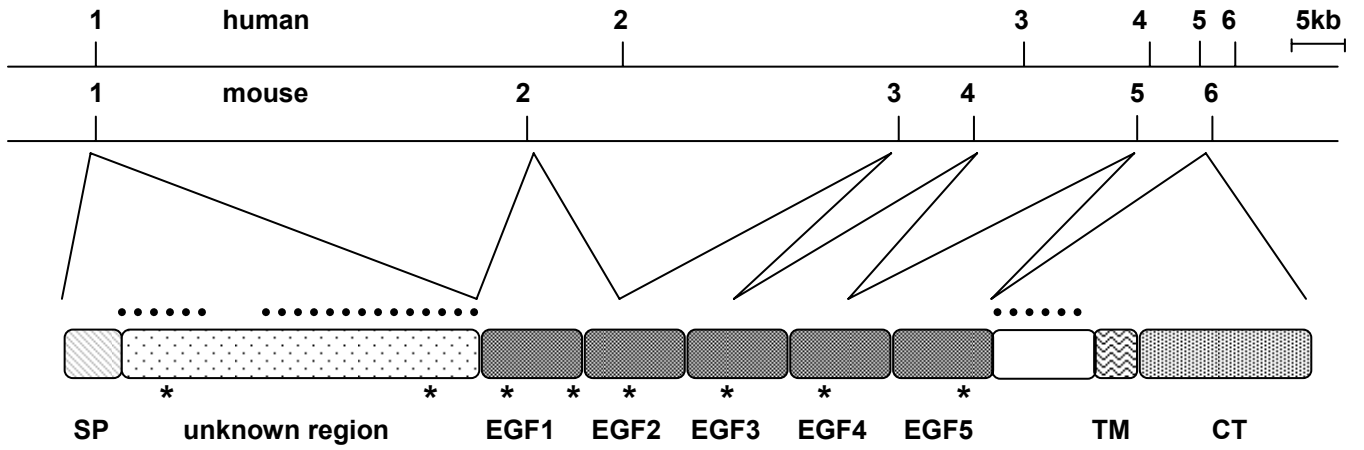


Fig. 1:

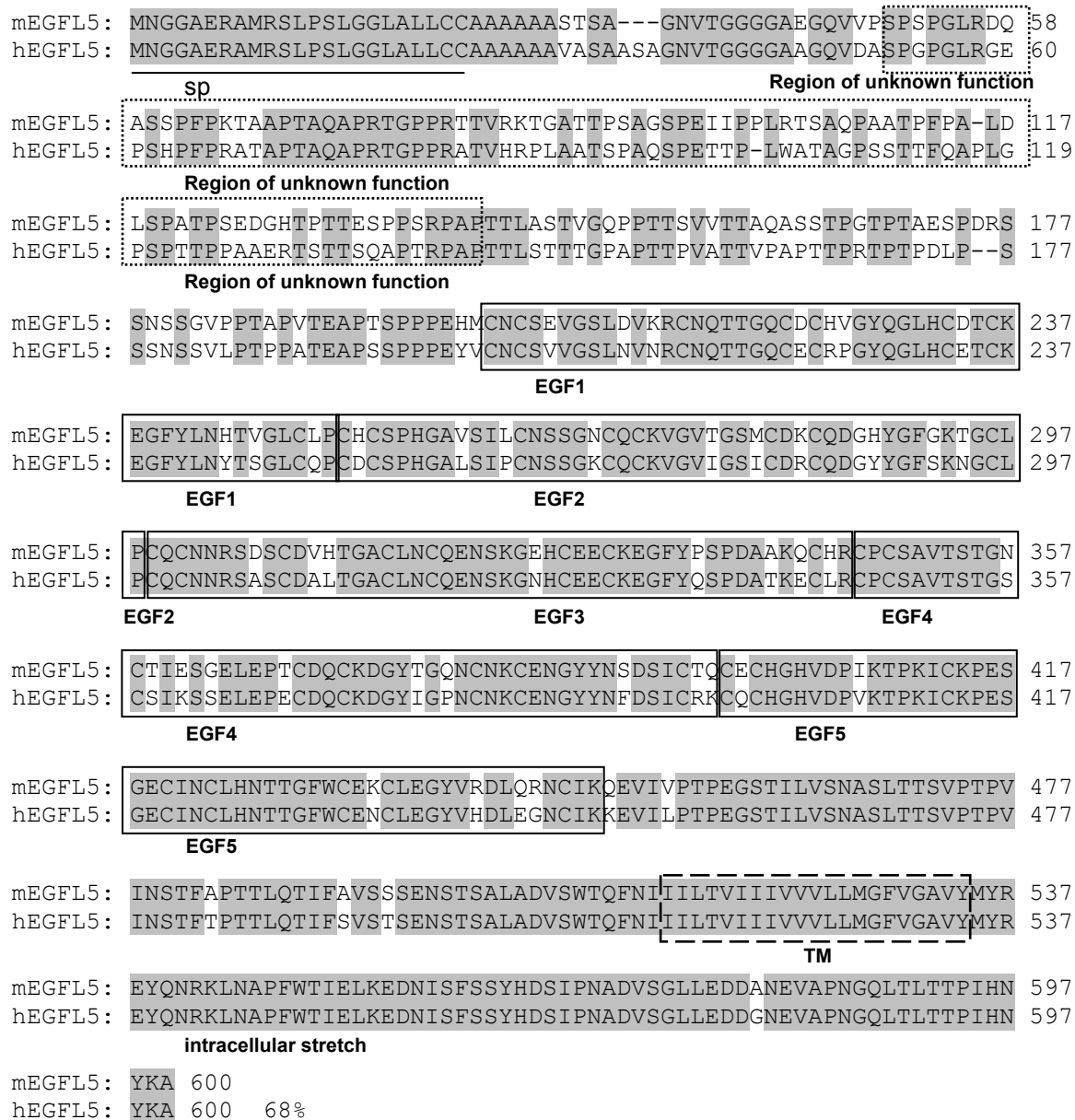


Fig. 2:

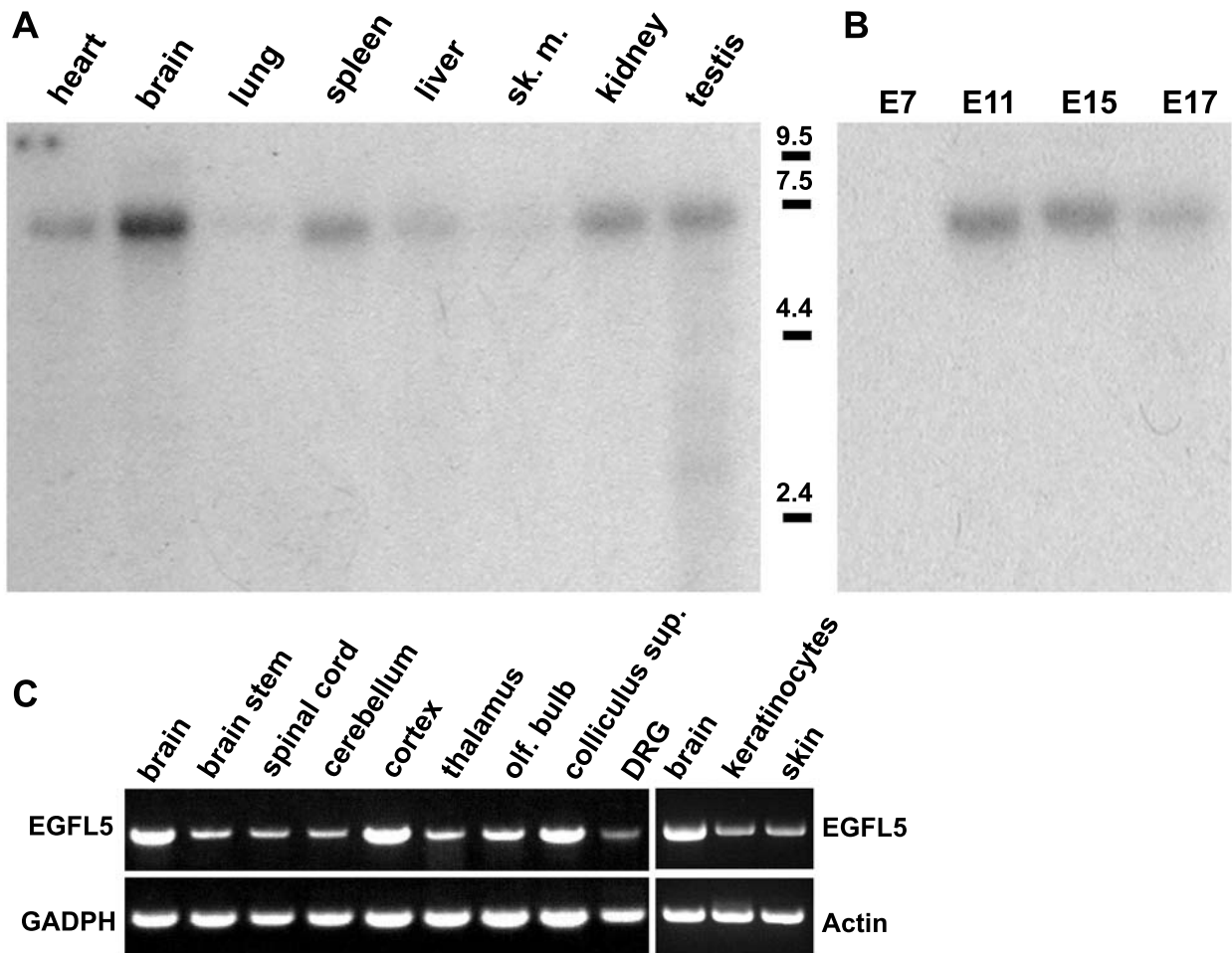


Fig. 3:

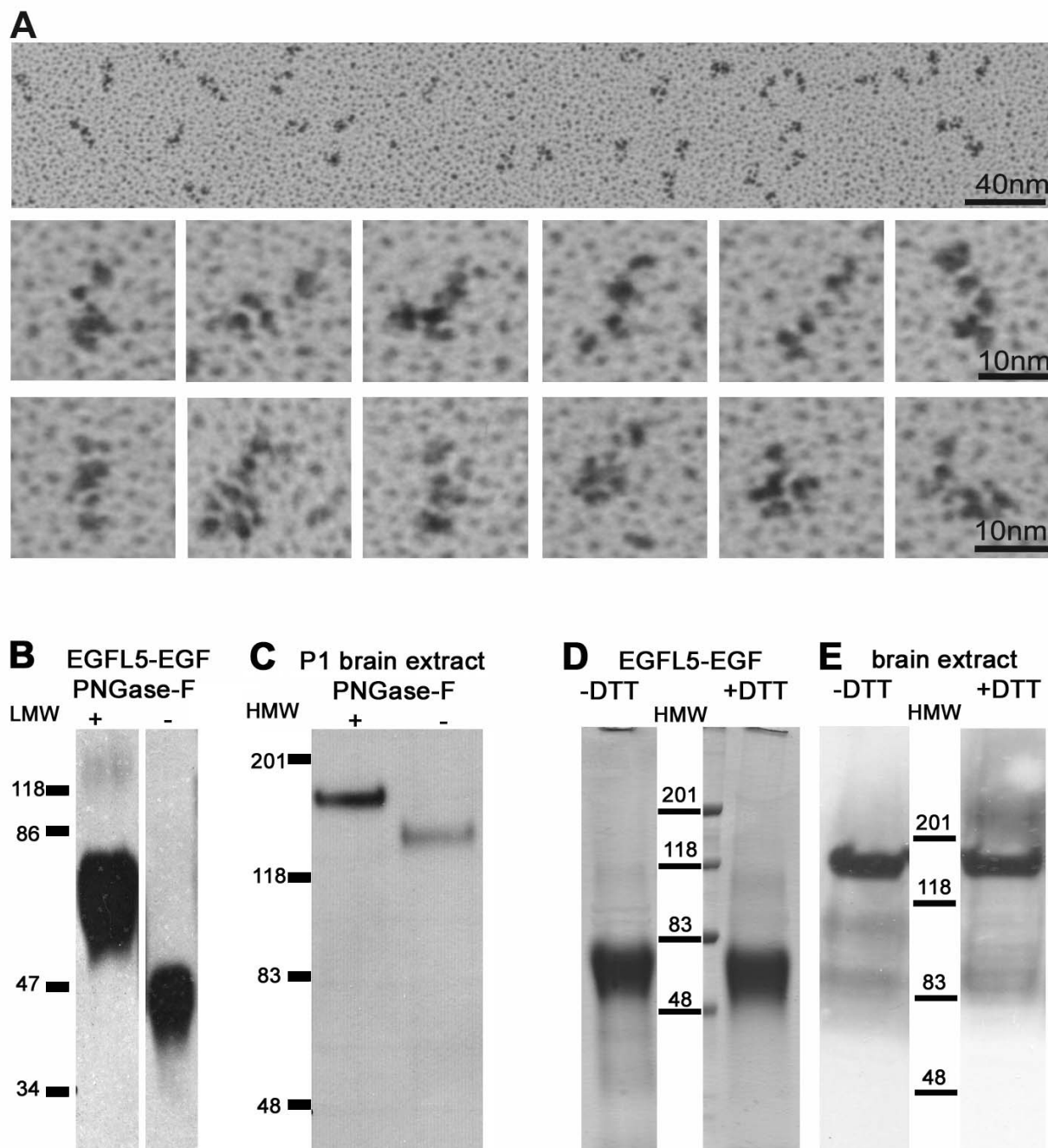


Fig. 4:

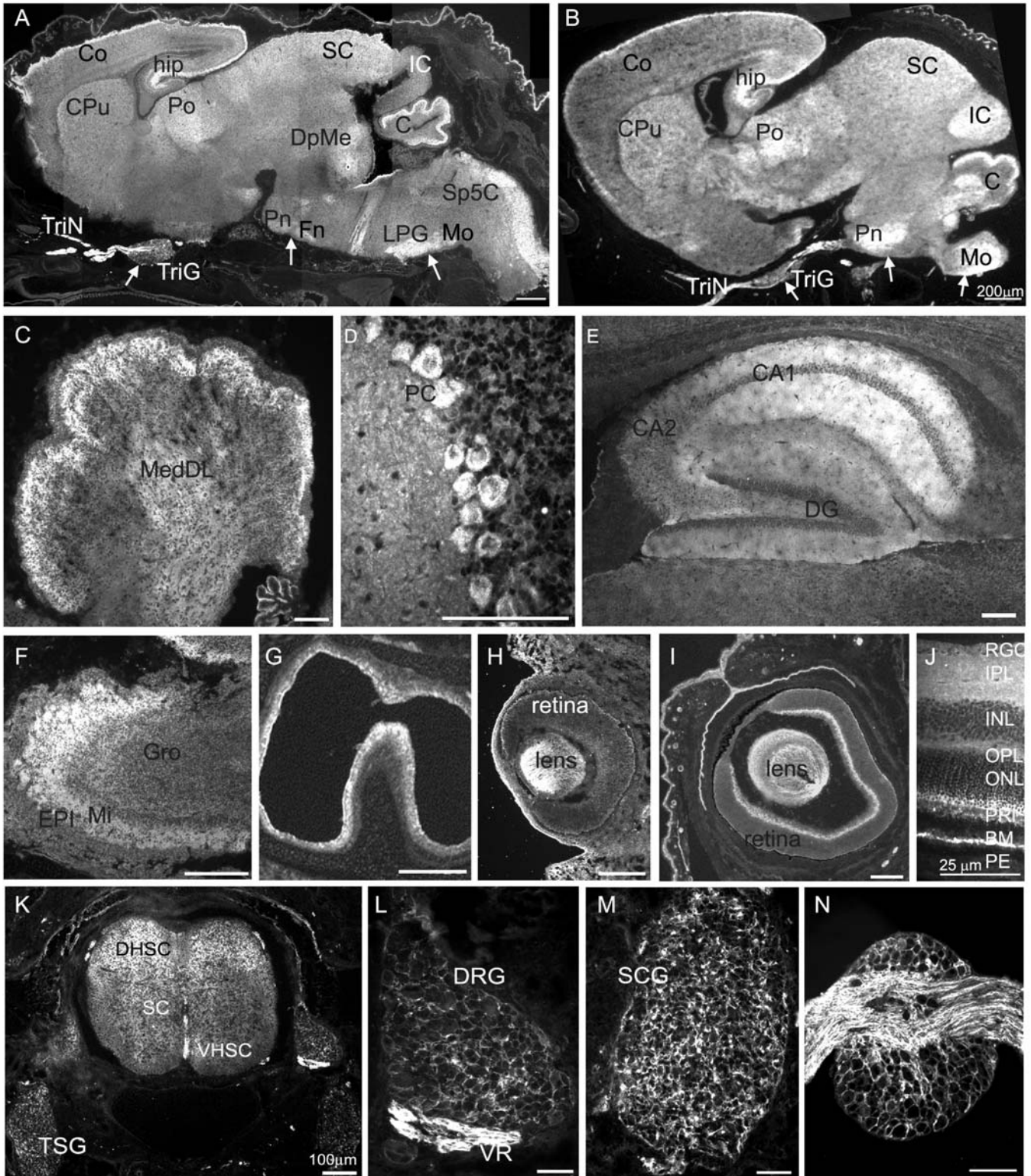


Fig. 5:

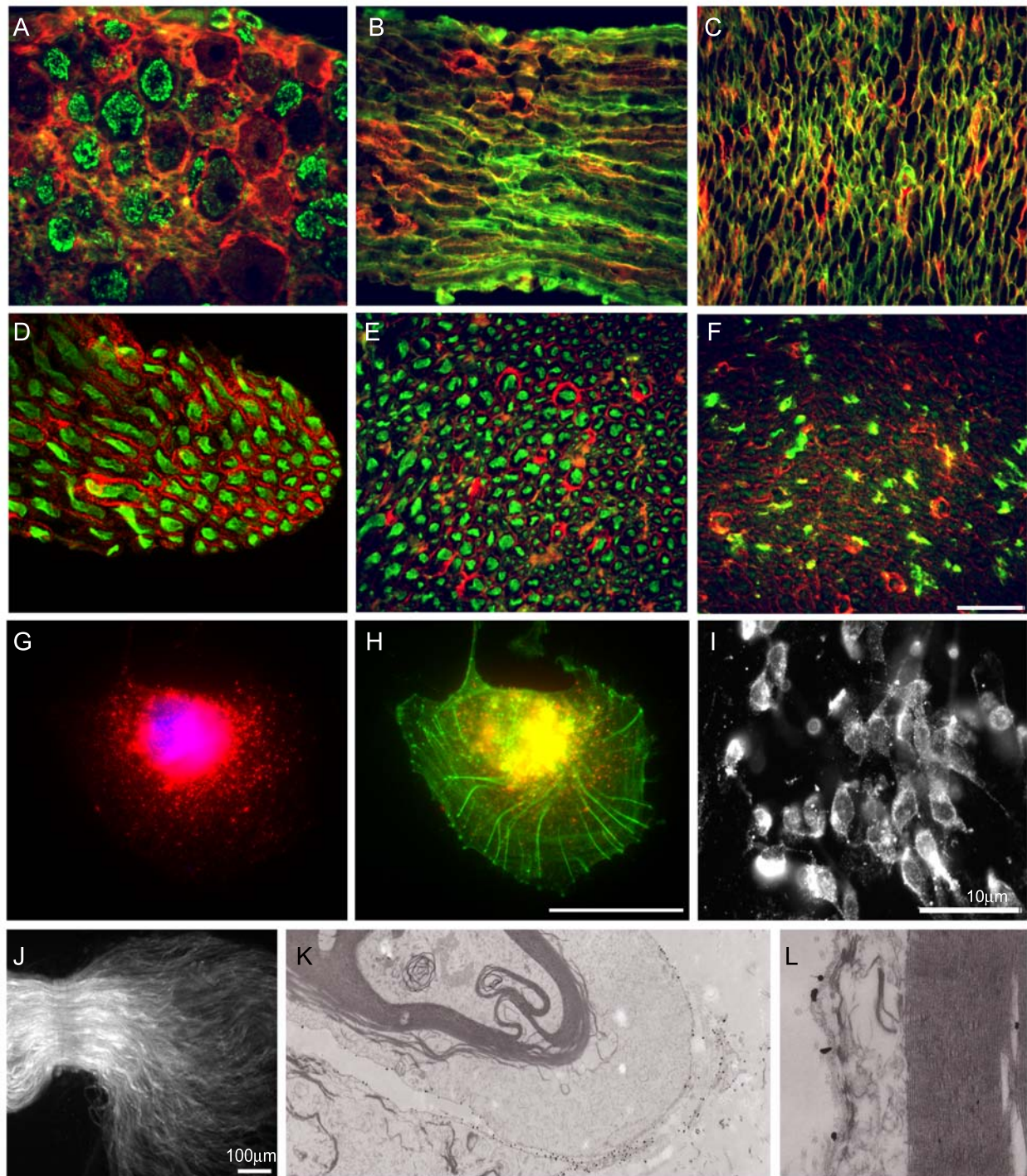


Fig. 6:

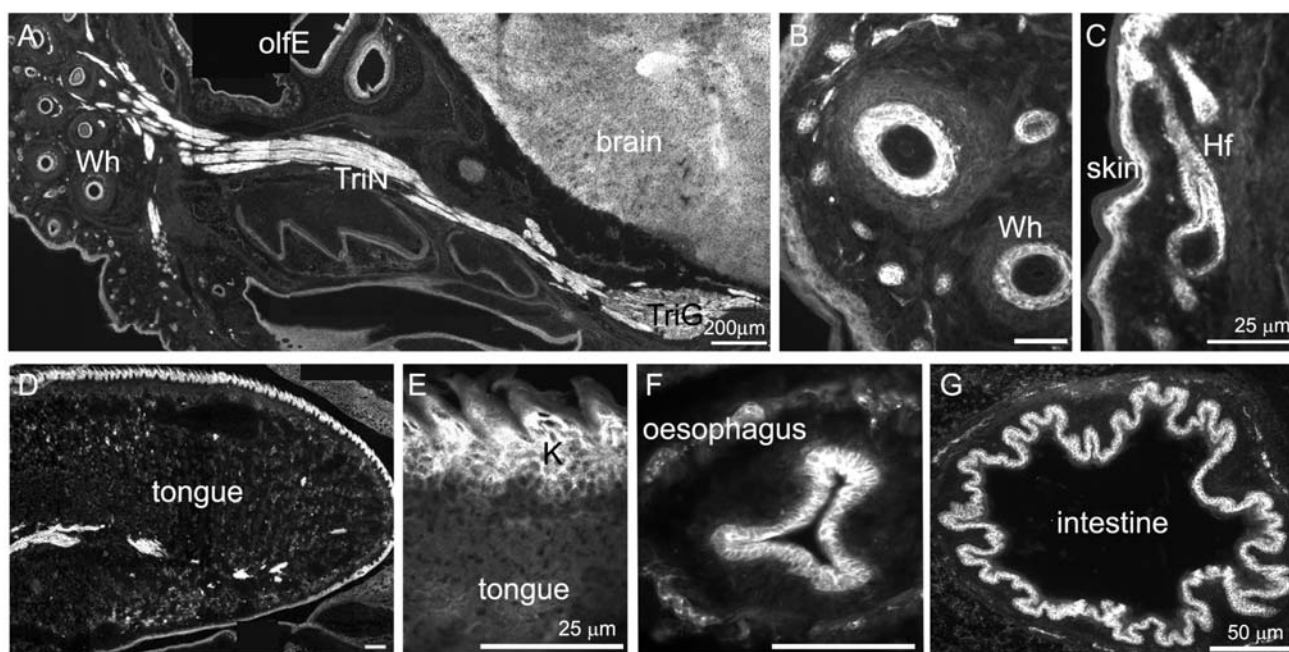


Fig. 7: