online data

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‡ and Manuel KOCH*§‖[1]

*Center for Biochemistry, Medical Faculty, University of Cologne, Joseph-Stelzmann-Str. 52, D-50931 Cologne, Germany, †Shriners Hospital for Children Research Center, Portland, OR 97329, U.S.A., ‡Department of Cell Biology, Neurobiology and Anatomy, Loyola University Medical Center, Maywood, IL 60153, U.S.A., §Center for Molecular Medicine Cologne, Medical Faculty, University of Cologne, Joseph-Stelzmann-Str. 52, D-50931 Cologne, Germany, and ‖Department of Dermatology, Medical Faculty, University of Cologne, Joseph-Stelzmann-Str. 52, D-50931 Cologne, Germany

MEGF9 [multiple EGF (epidermal growth factor)-like-domains 9], a novel transmembrane protein with multiple EGF-like repeats, is predominantly expressed in the developing and adult CNS (central nervous system) and PNS (peripheral nervous system). The domain structure of MEGF9 consists of an N-terminal region with several potential O-glycosylation sites followed by five EGF-like domains, which are highly homologous with 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 study the expression pattern further, immunohistochemistry was performed and staining was detected in Purkinje cells of the cerebellum and in glial cells of the PNS. Additional expression was observed in the epidermal layer of skin, papillae of the tongue and the epithelium of the gastrointestinal tract. By immunoelectron microscopy, MEGF9 was detected in glial cells of the sciatic nerve facing the basement membrane. MEGF9 represents a novel putative receptor, expressed in neuronal and non-neuronal tissues, that is regulated during development and could function as a guidance or signalling molecule.

Key words: glial cell, glycosylation, MEGF9, nervous system, skin, transmembrane protein.

INTRODUCTION

EGF (epidermal growth factor)-like domains are major modular components present in many proteins of the ECM (extracellular matrix). 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 approx. 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 LE (laminin-type EGF-like) domains, consisting of 60 amino acids. The tertiary structure is remotely similar at its N-terminus to that of the classical EGF-like module. The two domain types differ in their number of conserved cysteine residues: six in classical EGF-like and eight in LE 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 such as 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 such as 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 hypercholesterolaemia (low-density lipoprotein 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 mis-expressed [16]. Mutations in laminin chains (e.g. laminin $\alpha 3$, $\beta 3$ and $\gamma 2$) result in skin diseases such as junctional epidermolysis bullosa [17]. The identification and characterization of further proteins containing multiple EGF-like repeats are likely to help our understanding of cell–cell interactions or ligand–receptor interactions in the nervous system and in other tissues. In the present study, we describe the initial identification, characterization and expression profile of a novel transmembrane protein, called MEGF9 (multiple EGF-like-domains 9).

EXPERIMENTAL

Northern-blot analysis

A 724-bp PCR product (forward 5′-dgaccccaactgccgagtcc-3′ and reverse 5′-dttttaattgggtccacatggcc-3′) was labelled with [^{33}P]dCTP (NEN Life Science Products) using the Rediprime DNA labelling system (GE Healthcare). Northern blots (ClonTech Laboratories) were prehybridized in 50 % (v/v) formamide, 5 × SSPE [0.15M NaCl/10 mM sodium phosphate (pH 7.4)/1 mM EDTA],

Abbreviations used: CDK5, cyclin-dependent kinase-5; P1, postnatal day 1; CNS, central nervous system; DMEM, Dulbecco's modified Eagle's medium; EBNA, Epstein–Barr virus nuclear antigen; EGF, epidermal growth factor; multiple 5; EST, expressed sequence tag; GFAP, glial fibrillary acidic protein; LE, laminin-type EGF-like; MEGF9, multiple EGF-like-domains 9; NGS, normal goat serum; PFA, paraformaldehyde; PNGase F, peptide N-glycosidase F; PNS, peripheral nervous system; RT, reverse transcriptase.

[1] To whom correspondence should be addressed (email Manuel.Koch@uni-koeln.de).

1 × Denhardt's (0.02 % Ficoll 400, 0.02 % polyvinylpyrrolidone and 0.02 % BSA), 1 % SDS, 10 % (w/v) 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 (v/v) human Cot-1 DNA (Boehringer Mannheim), and 1:10 (v/v) 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 2 × SSC (1 × SSC is 0.15 M NaCl and 0.015 M sodium citrate) and 1 % SDS at 42 °C and twice in 0.1 × SSC and 1 % SDS at 42 °C. Blots were placed on BioMax MR film with a BioMax TranScreen-LE intensifying screen (both from Eastman Kodak) for 20 h at − 70 °C.

RT (reverse transcriptase)–PCR

RNA was isolated from animal tissues using the RNeasy kit (Qiagen), and 1 μg 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 GAPDH (glyceraldehyde-3-phosphate dehydrogenase) primers (forward: 5′-dtgaaggtcggtgtgaacgga-3′; reverse: 5′-dgatggcatggactgtggtca-3′) or actin primers (forward: 5′-gccattcaggcggtgctgtc; 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′-dccgccacccctttcggaggacgg-3′; reverse: 5′-dggccatgtggacccaattaaaactcc-3′), another PCR was performed on the normalized cDNA templates.

Expression and purification of recombinant MEGF9

The MEGF9 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 MEGF9. The primers used were: forward 5′-gaggctagccatatgtgtaacttctgaggttc-3′ and reverse 5′-ccaggatccttacttgatgcagtttctctgaagat-3′. The sequenced construct was cloned into a modified pCEP-Pu expression vector carrying an N-terminal 6 × His tag and a thrombin cleavage site and the plasmid was transfected into EBNA (Epstein–Barr virus nuclear antigen) 293 cells. Protein was purified from the collected cell culture supernatant using CoCl–Sepharose (ClonTech Laboratories) and eluted with increasing imidazole concentrations in 150 mM NaCl and 50 mM Tris/HCl (pH 8). Finally, the protein was dialysed against PBS and stored at − 20 °C.

Preparation of specific MEGF9 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 MEGF9 protein coupled. The specific antibodies were eluted with 150 mM NaCl and 0.1 M triethylamine (pH 11.5), and the eluate was neutralized with 1 M Tris/HCl (pH 6.8) and dialysed 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 dialysed against 150 mM ammonium bicarbonate. Samples were diluted to a final concentration of 70 % (v/v) glycerol, sprayed on to freshly cleaved mica, and dried under vacuum (Balzers BAE 250 evaporator). Rotary shadowing was performed as described previously [17a]. Replicas were examined at 80 kV in a transmission electron microscope (Philips 410LS). The protein images were measured using the imageJ [NIH (National Institutes of Health) freeware] software.

Tissue preparation and extraction

Mice were killed to collect tissues for analysis. Brains were dissected, weighed, extracted by homogenization in chilled buffer [1 × Complete™ Protease Inhibitor Cocktail (Roche Diagnostics), 0.2 mM EDTA and 1 % Nonidet P40], and incubated on ice for 30 min. Homogenates were centrifuged at 4 °C at 10000 ***g*** for 15 min. Supernatants were mixed with 5 × SDS/PAGE sample buffer supplemented with 4 % 2-mercaptoethanol and applied to an SDS/10–12 % polyacrylamide gel.

Removal of N-linked glycans

PNGase F (peptide N-glycosidase F) (New England Biolabs) was used to digest purified proteins and tissue extracts according to manufacturer's specifications.

SDS/PAGE and immunoblotting

SDS/PAGE was performed as described in [18]. For immunoblots the proteins were transferred to nitrocellulose, blocked in TBS [Tris-buffered saline (0.15 M NaCl/20 mM Tris/HCl, pH 7.4)] 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 (24 °C) or overnight 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 vapour and frozen sections (10 μm thick) cut. The dried sections were permeabilized in ice-cooled methanol for 1 min, washed and blocked for 1 h at room temperature with 5 % (v/v) NGS (normal goat serum) in PBS containing 0.5 % Tween 20. Primary antibodies were diluted in blocking solution and incubated on the sections overnight at 4 °C. For double immunofluorescence labelling, 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 MEGF9, rat monoclonal anti-laminin γ2 (Chemicon), monoclonal mouse anti-neurofilament heavy chain (Abcam) and polyclonal rabbit anti-GFAP [anti-(glial fibrillary acidic protein)] (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 Fluor® (Alexa Fluor® 488 goat anti-mouse IgG, Alexa Fluor® 488 goat anti-rat IgG and Alexa Fluor® 488 goat anti-rabbit IgG, Invitrogen) fluorochrome for 2 h at room temperature. Alexa Fluor® 488-conjugated lectin from *Griffonia simlicifolia* (Molecular Probes) was incubated for 1 h at room temperature on sections treated with 4 % (w/v) PFA (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

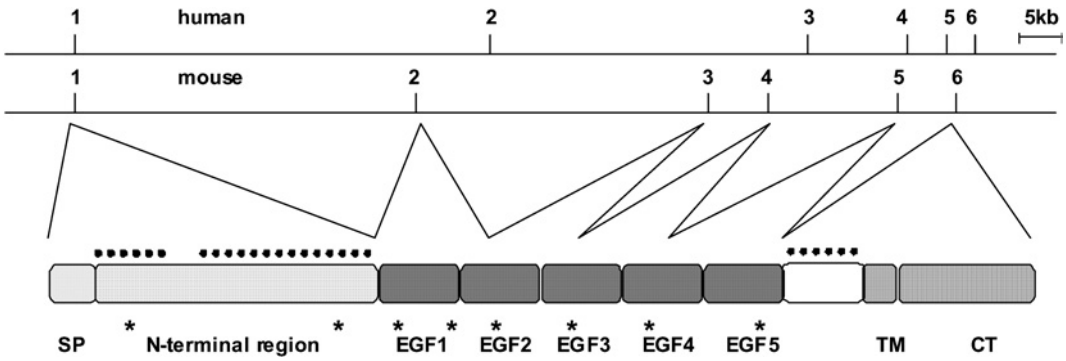

Figure 1 Exon/intron organization of the human and murine MEGF9 genes and domain structures of the MEGF9 proteins

The genes are approx. 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.

with 2 % PFA 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 % FCS (fetal calf serum) containing DMEM (Dulbecco's modified Eagle's medium)/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 (0.001 mg/ml in methanol diluted in PBS, Serva), whereas the F-actin cytoskeleton of RN2 cells was detected by FITC-conjugated phalloidin (Molecular Probes).

Immunoelectron microscopy

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 labelled 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 % (v/v) glutaraldehyde/1.5 % PFA with 0.05 % tannic acid, osmicated, and prepared for transmission electron microscopy by a standard protocol.

RESULTS

Identification and characterization of the MEGF9 genes

By screening the EST (expressed sequence tag) database for EGF-like domains, we identified a murine sequence with high homology to the laminin $\beta 2$ chain and cloned the gene. The encoded type I single-pass transmembrane protein, named MEGF9, contains five EGF-like domains.

The *MEGF9* genes map to syntenic regions in the human and mouse genome on chromosome 9 (9q32–9q33.3) and 4 (CA2) respectively. At the 5′-end, the human and murine genes are flanked by the CDK5 (cyclin-dependent kinase-5) 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 [RefSeq (NCBI Reference Sequence) accession number XM_376905] and mouse (RefSeq accession number 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 MEGF9 genes consist of six exons and span over 113.66 and 106.09 kb respectively. The corresponding cDNAs are approx. 6298 bp (human) and 6146 bp (murine) long (Figure 1). Exon 1 encodes the signal peptide sequence and an N-terminal domain, whereas exons 2–5 code for EGF-like repeats 1–5, and exon 6 codes for 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 MEGF9 domain structure

MEGF9 is highly conserved between different species in vertebrates; however, no homologous gene can be located in the *Caenorhabditis elegans* or *Drosophila 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 with the MEGF9 sequences in rat, chimpanzee and human (Supplementary Figure 1A at http://www.BiochemJ.org/bj/401/bj4010447add.htm).

As indicated by its name, MEGF9 contains five EGF-like domains (EGF1: amino acids 202–251, EGF2: 252–298, EGF3: 299–346, EGF4: 347–397 and EGF5: 398–449). They match the EGF motif consensus sequence CXCXXXXXGXXC, where X represents any residue, G an often conserved glycine residue, and C indicates conserved cysteine residues, providing disulfide bridges. The EGF-like domains of MEGF9 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 at http://www.BiochemJ.org/bj/401/bj4010447add.htm). 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 $\beta 1$ and $\beta 2$ chains, while other members of the laminin group, such as the laminin $\alpha 1$, $\alpha 2$, $\alpha 3$ and $\alpha 4$

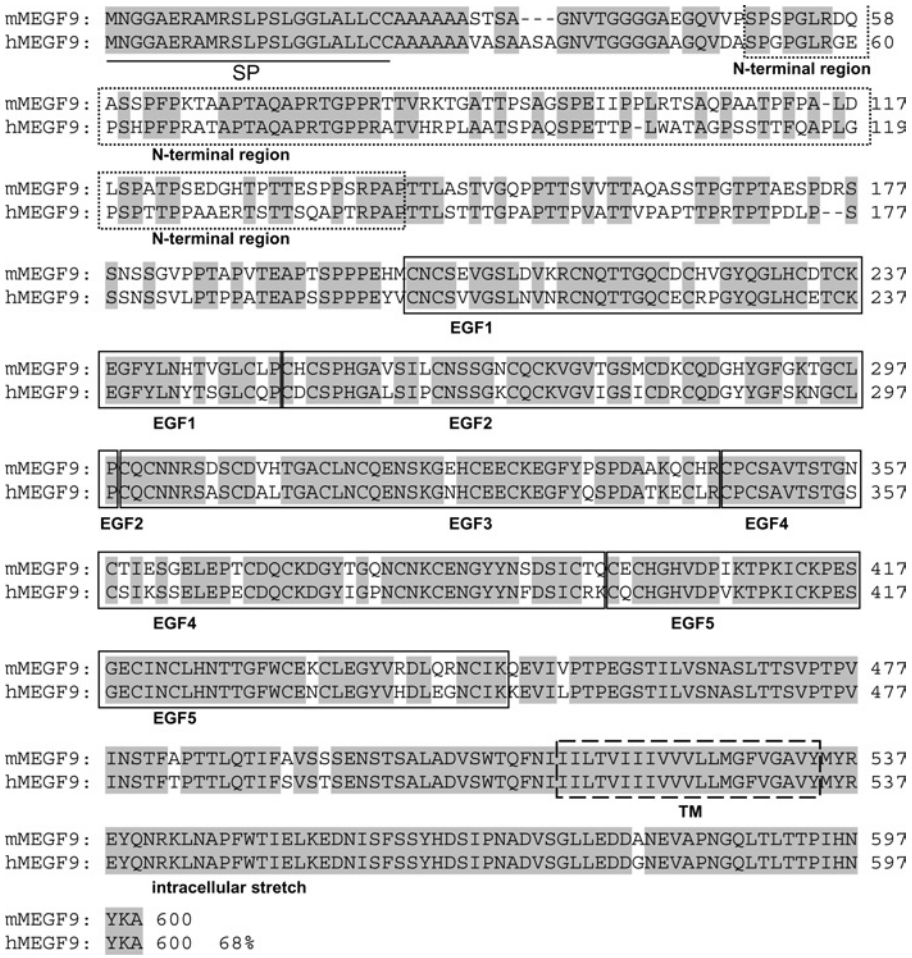

Figure 2 Homology of the mouse and human MEGF9 amino acid sequences

The signal peptide (SP) is underlined, the N-terminal region is shown in dotted boxes, the EGF-like domains are in closed boxes, and the transmembrane region is in a dashed box. Identical amino acid residues are marked in grey. m, mouse; h, human.

chains as well as laminin $\gamma 1$ and $\gamma 3$ chains display weaker homologies to MEGF9. The shared regions are restricted to the EGF-like domains of MEGF9, and multiple alignments are shown in Supplementary Figure 2(A) (http://www.BiochemJ.org/bj/401/bj4010447add.htm) 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 with the sequence of the human MEGF9 (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 at http://www.BiochemJ.org/bj/401/bj4010447add.htm).

Prediction of post-translational modifications

Putative post-translational modifications can in part be predicted by analysing 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 Thr^{591}, 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 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 MEGF9 shows 42 cysteine residues in the sequence, representing 7 % of all amino acids. Forty out of these 42 cysteines are located in the EGF-like domains; none are present in the N-terminal region 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 N-terminal 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 %),

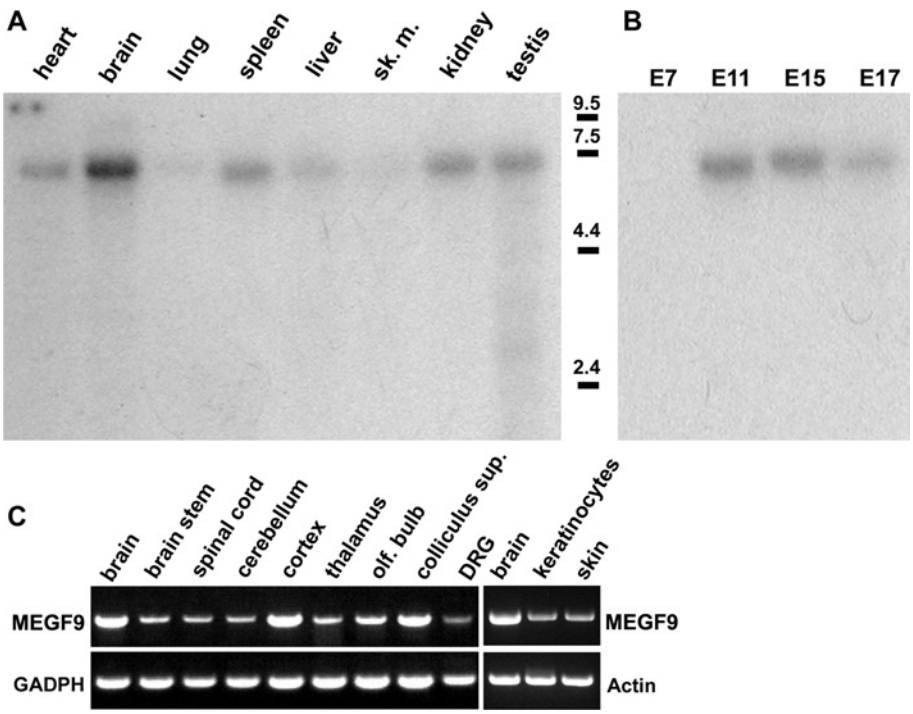

Figure 3 Northern-blot and RT–PCR analyses

(**A**) Northern-blot analysis of murine tissues detects MEGF9 mRNA as a single band at approx. 7.4 kb in brain. Weaker signals at the same size are present in non-neuronal tissues such 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.

alanine (11.9 %) and valine (5.4 %). The proportions of positively charged residues (i.e. arginine and lysine) in the murine and human intracellular MEGF9 stretches are 16 and 15 % respectively, but these are not clustered. Three putative phosphorylation sites are predicted in the intracellular region (Ser^{561}, Ser^{562} and Tyr^{563}) (NetPhos2.0; Supplementary Figure 2B at http://www.BiochemJ.org/bj/401/bj4010447add.htm), which are present in both mouse and human MEGF9. Other programs confirm this prediction, indicating the SYHD sequence as a possible phosphorylation site for protein kinase 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 MEGF9 show hydrophobic stretches in the N-terminal region, which corresponds to the signal peptide and the C-terminal transmembrane region (http://www.bioinformatics.weizmann.ac.il).

MEGF9 gene expression

To determine the length of MEGF9 mRNA we performed Northern blot hybridization with mRNA derived from different embryonic mouse tissues and developmental stages (Figures 3A and 3B). The first expression of MEGF9 mRNA was detected at embryonic day 11. A variety of tissues (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).

MEGF9 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 MEGF9 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 MEGF9 expression and antibody production

In order to characterize the MEGF9 protein, a 930 bp insert (MEGF9–EGF), which encodes the five EGF-like domains, was amplified using specific primers (see the Experimental section). The construct was transfected into EBNA 293 cells and the expressed protein purified. SDS/PAGE of the purified MEGF9–EGF protein showed a single band at approx. 75 kDa. Rotary shadowing experiments confirmed the homogeneity of the sample and showed 12.56 nm $\pm$ 1.3 nm long rod-like particles with some bends (Figure 4A). The recombinant MEGF9–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 (Figures 4B and 4C).

The mass of the purified MEGF9 ectodomain protein by SDS/PAGE ($\sim$75 kDa) was much higher than the 33 kDa predicted from the sequence, while MALDI–TOF MS (matrix-assisted laser-desorption ionization–time-of-flight MS) 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 approx. 180 kDa, in contrast with 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 approx. 30 kDa in both the recombinant MEGF9–EGF and in full-length MEGF9 from brain tissue extracts (Figures 4B and 4C).

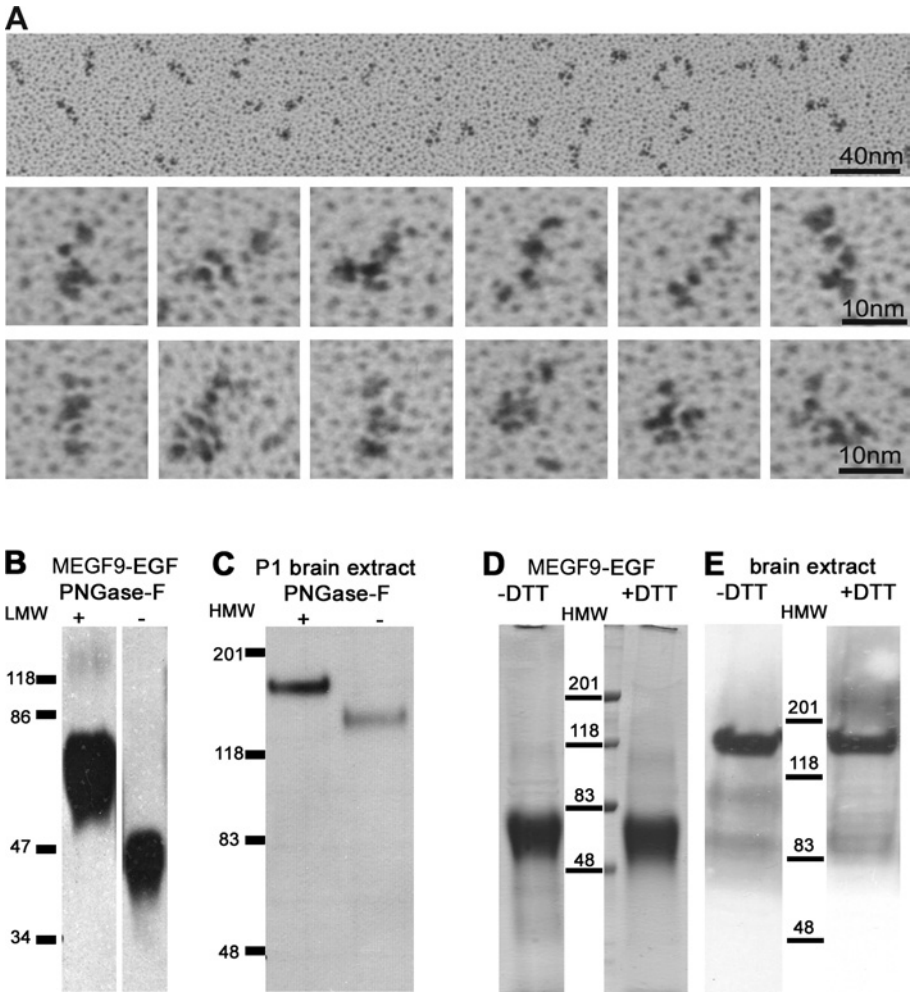

Figure 4 Structural analysis of MEGF9 protein

(**A**) Rod-like particles representing the recombinant MEGF9–EGF protein were detected by rotary shadowing electron microscopy. (**B**) The affinity-purified rabbit anti-MEGF9 antibody KR18 recognizes the recombinant protein at approx. 75 kDa in SDS/PAGE immunoblot. PNGase F digestion results in a mass shift of approx. 30 kDa. (**C**) MEGF9 is detected as a single approx. 180 kDa band in brain extract that shifts to approx. 150 kDa upon PNGase F digestion. The electrophoretic mobility is analysed under reduced or non-reduced condition in (**D**) recombinant MEGF9–EGF and (**E**) a brain tissue extract detected with the KR18 antibody by immunoblotting.

No Ca^{2+}-binding EGF-like modules are predicted in MEGF9, and recombinant MEGF9–EGF as well as full-length MEGF9 detected in brain tissue extracts show no difference in their electrophoretic mobility under either reduced or non-reduced SDS/PAGE (Figures 4D and 4E).

MEGF9 expression in the CNS (central nervous system)

Immunohistochemistry using affinity-purified antibodies raised against the recombinant MEGF9–EGF fragment showed MEGF9 protein expression throughout the E18.5 embryonic brain and a more restricted expression by P1 (postnatal day 1) (Figures 5A and 5B). 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 stratium 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 MEGF9 (Figures 5A and 5B). The cerebellar expression becomes more limited late in development (Figures 5A–5C) and MEGF9 is subsequently expressed only by Purkinje cells of the adult cerebellum (Figure 5D). Staining was also detected in the olfactory bulb and olfactory epithelium (Figures 5F and 5G).

In eyes from the stage E12.5, positive signals were seen only in the lens (Figure 5H), whereas on P1, MEGF9 was present throughout the lens and retinal ganglion cell layer (Figure 5I). In the adult eye, MEGF9 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, MEGF9 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.

MEGF9 expression in the PNS (peripheral nervous system)

MEGF9 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

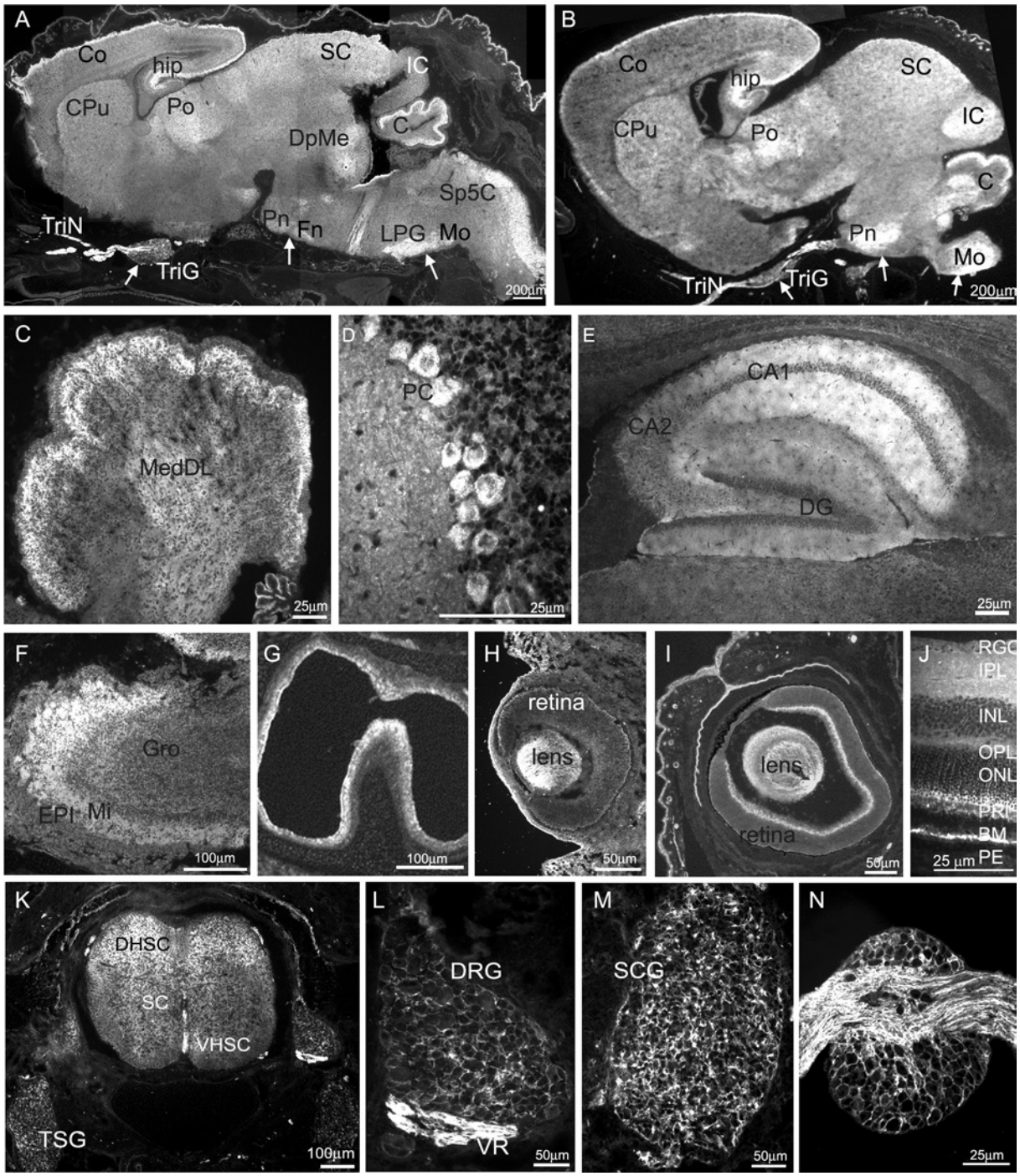

Figure 5 Protein expression in the CNS

MEGF9 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 MEGF9 (**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. the inner and outer plexiform layer, the photoreceptor layer and Bruchs membrane. (**K**) MEGF9 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 dorsal root ganglia of adult animals (**N**). Co, cortex; Po, posteriour 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; EPI, external plexiform layer of olfactory bulb; RGC, retinal ganglion cell layer; IPL, INL, inner plexiform and nuclear layers; OPL, ONL, outer plexiform and nuclear layers; PRI, 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; TSG, thoracal sympathetic ganglia.

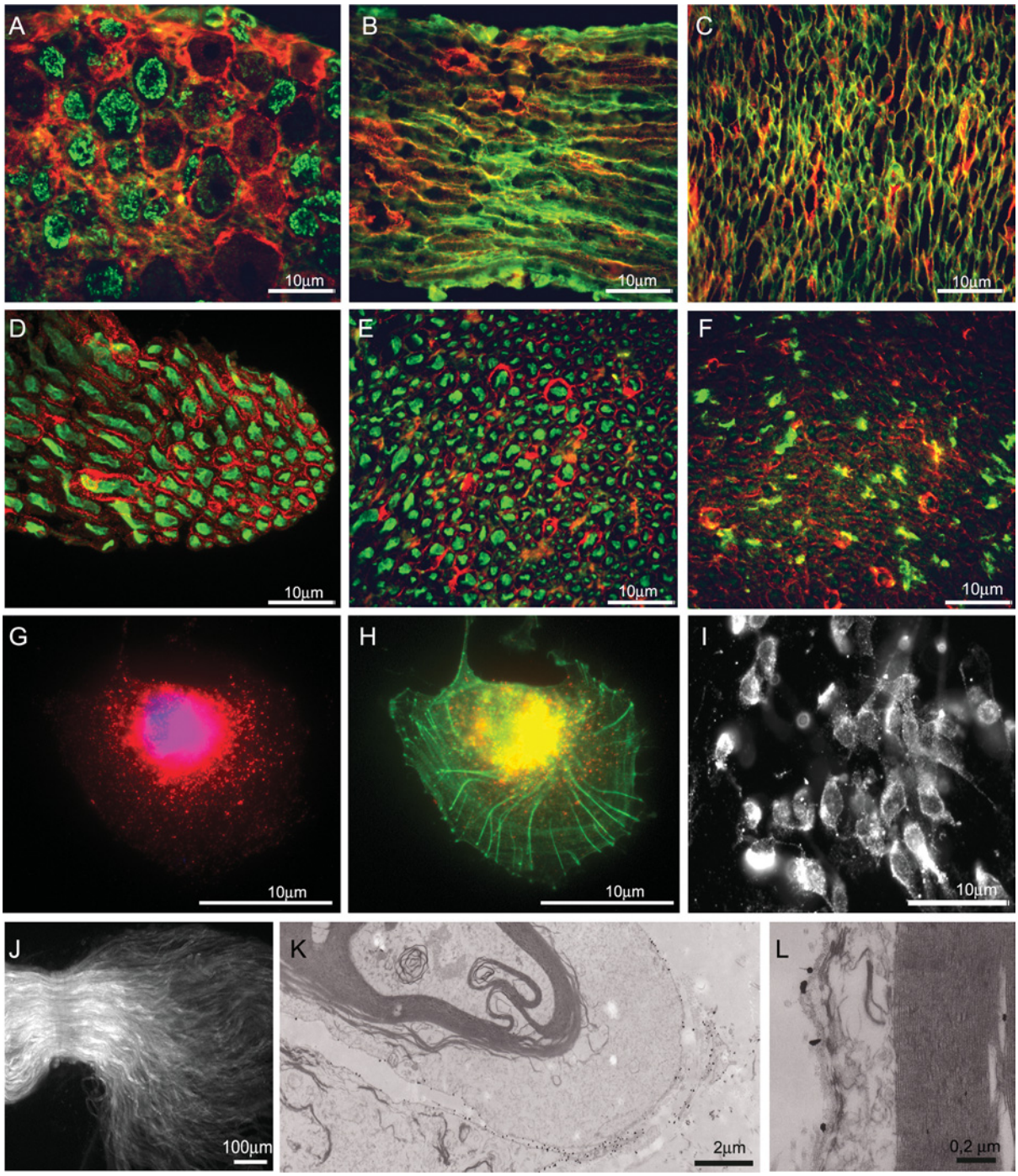

Figure 6 MEGF9 protein expression in the PNS

MEGF9 (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 fibres. (**C**) MEGF9 expression overlaps with that for laminins containing the $\gamma 1$ chain present in the basement membranes of Schwann cells. (**D**, **E**) Labelling of MEGF9 and neurofilament 200 kDa (green) in sciatic nerves shows that MEGF9 is expressed in myelinating Schwann cells. MEGF9 and (**F**) GFAP, expressed by subsets of Schwann cells, also show a limited overlap. In cell culture, (**G**, **H**) both RN2 schwannoma cells (**H**: co-stain for F-actin) and (**I**) C6 oligodendroglia cell lines are both positive for MEGF9. (**J**) MEGF9 expression in peripheral nerves is detected in whole mount preparations of trigeminal nerves. **(K**, **L**) Electron microscopy images of immuno-gold-labelled sciatic nerve sections. MEGF9 is present close to the basement membrane of myelinating Schwann cells.

ganglia (Figures 5A, 5B and 5K–5N). 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 Figures 5A and 5B). In co-staining experiments, an overlap was seen with laminin present in basement membranes of Schwann cells ensheathing peripheral nerves (results not shown). MEGF9 was also detected in whole mount preparations of sciatic nerves and in sections depicting afferent fibres of the tongue and the whiskers (Figures 6J and 7D).

To determine the nature of MEGF9-immunopositive peripheral glial-cells, further immunostaining was performed on both longitudinal and transverse sections of sciatic nerves. Co-staining

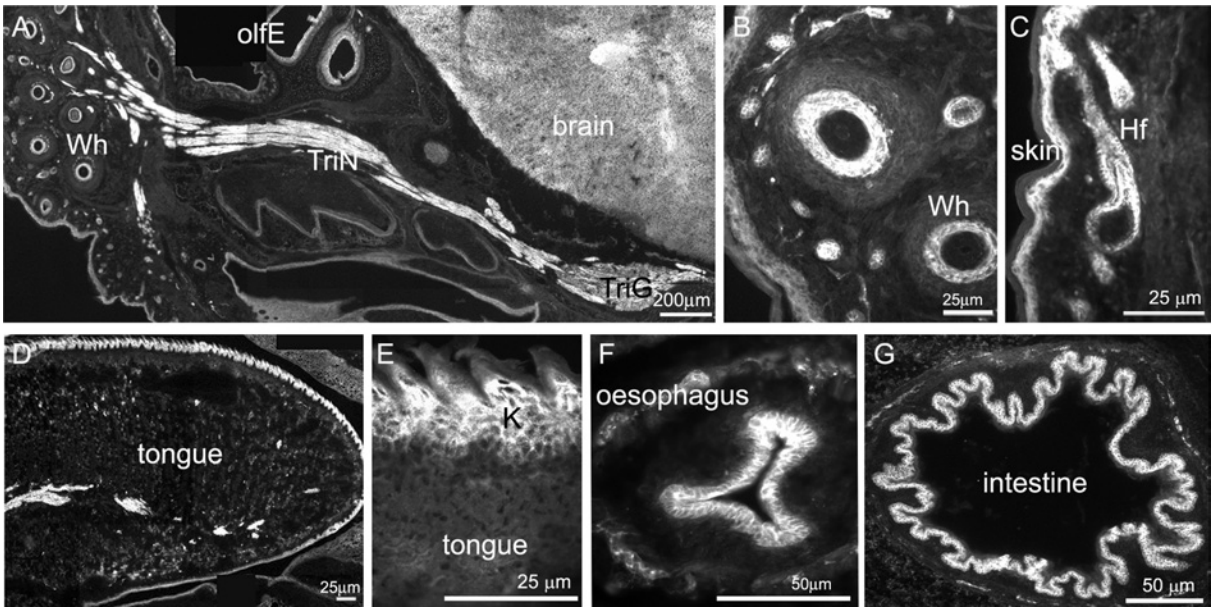

Figure 7 MEGF9 protein expression in skin and other keratinized tissues

On cryosections of embryonic and newborn tissues, MEGF9 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.

with a glia-specific lectin showed that MEGF9 occurs in structures around dorsal root nerve cells, representing associated glia cells (Figures 6A and 6B). MEGF9 showed a co-localization 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 (Figures 6D–6F). Accordingly, MEGF9 is present only on cells surrounding nerve fibres and not in the actual nerve fibres. In cell culture, MEGF9 is expressed on rat RN2 schwannoma (Figures 6G and 6H) and on rat C6 oligodendroglial cells (Figure 6I). Immunogold labelling of sciatic nerves with the KG20 antibodies localized MEGF9 to the outer membrane of myelinating Schwann cells (Figures 6K and 6L).

MEGF9 expression on keratinocytes of the skin and tongue

MEGF9 was also detected in tongue and skin epithelium. In agreement with the lower MEGF9 mRNA levels in skin compared with brain, the immunofluorescence signal in skin was also weaker than those in the nervous system (Figure 7A). In skin, all epidermal layers express MEGF9, and strong staining was also seen within the external root sheath of hair follicles (Figures 7B and 7C). 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 MEGF9-positive (Figures 7D and 7E). MEGF9 was also detected in the oesophageal and intestinal epithelium (Figures 7F and 7G).

DISCUSSION

MEGF9, 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 MEGF9 gene has been detected earlier and both the murine and the human sequences were submitted to the database, no detailed characterization was performed up to date. The human cDNA of MEGF9 was identified in a search for high-molecular-mass 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 MEGF9 genes are 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 MEGF9 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 MEGF9 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 $\beta 2$ chain show the highest degree of conservation at the amino acid level, followed by the laminin $\beta 1$ subunit, the $\alpha 1$ and $\alpha 2$ chains, and the laminin $\gamma 1$ 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 MEGF9 EGF repeat reveals that no putative Ca-binding consensus sequences are present. Following the short transmembrane region, a unique intracellular stretch containing putative phosphorylation sites for signal transduction events is located at the C-terminus. This region is highly conserved between different species ($\sim$98 %).

Post-translational 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 MEGF9. A general feature of mucins is the presence of regions rich in threonine and/or serine whose hydroxy groups are in O-glycosidic linkage with the oligosaccharides involved [23]. As shown for other mucin-type glycosylated proteins [24], the C-terminal region of MEGF9 is also rich in serine, threonine,

glycine, alanine and proline, which together account for more than three-fourths of the amino acid residues.

As a 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], MEGF9 might be involved in similar processes. High levels of MEGF9 transcripts and protein were detected in neuronal and glial cells of the CNS and PNS. 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 leads to an impairment of ligand binding and causes muscular dystrophies. In addition to the N-linked glycans predicted in the N-terminal region, 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 MEGF9, as deglycosylation of MEGF9 followed by SDS/PAGE/immunoblot analysis resulted in a shift in mass for both. Further glycosylation and other post-translational 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 MEGF9 mouse protein which specifically reacted with one band in brain extracts. By immunohistochemistry the main expression of MEGF9 is confined to nervous tissues. It can not only be detected in neuronal cells in the CNS, but also in the PNS. Neural cell bodies and Purkinje cells in the cerebellum as well as nerve-associated cell types express MEGF9 at their surface. The latter could be identified as Schwann cells by immunofluorescent staining of tissue sections and schwannoma cells also express MEGF9. By immuno-gold labelling and electron microscopy MEGF9 was detected on the outer membrane region of myelinated Schwann cells ensheathing sciatic nerves.

Similarly to laminins, MEGF9 was shown to be present in a variety of tissues, especially during development. Whereas in some tissues, like the olfactory bulb and the retina, laminin and MEGF9 stainings do not overlap, in others a close localization was observed. Also in the CNS, 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 CNS also seem to express MEGF9, as staining of brain sections appears diffuse and can not always be assigned to specific neurons. In contrast with laminins, MEGF9 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. MEGF9 expression is regulated during development, with protein being detectable from approximately 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 signalling to establish the basic axon scaffolds from which neuronal connections evolve. The importance of such interactive signalling 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 mutual reinforcement and dependent signalling that occurs in a network of developing neurons and glia [29]. MEGF9 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, MEGF9 could be involved in myelination, as it is present in the basement membranes of myelinated Schwann cells in the PNS.

We thank Dr Wilhelm Stoffel (Laboratory of Molecular Neurosciences, Medical Faculty, University of Cologne, Cologne, Germany) for the donation of RN2 schwannoma and C6 glial cells, and Dr Markus Plomann (Center for Biochemistry, Medical Faculty, University of Cologne, Cologne, Germany) and Professor Mats Paulsson (Center for Biochemistry, Medical Faculty, University of Cologne, Cologne, Germany) 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. and Bork, P. (1993) Epidermal growth factor-like modules. Curr. Opin. Struct. Biol. **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., Weber, I. T. and 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 Chem. **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**, 251–257
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**, 113–127
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**, 2665–2676
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**, 4037–4041
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**, 381–395
12 Dowgiert, J., Sosne, G. and Kurpakus-Wheater, M. (2004) Laminin-2 stimulates the proliferation of epithelial cells in a conjunctival epithelial cell line. Cell Prolif. **37**, 161–175
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**, 506–508
14 Pilia, G., Uda, M., Macis, D., Frau, F., Crisponi, L., Balli, F., Barbera, C., Colombo, C., Frediani, T., Gatti, R. et al. (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**, 368–381
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**, 2113–2118

17 Airenne, 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

17a Koch, M., Murrell, J. R., Hunter, D. D., Olson, P. F., Jin, W., Keene, D. R., Brunken, W. J. and Burgeson, R. E. (2000) A novel member of the netrin family, β-netrin, shares homology with the β-chain of laminin: identification, expression and functional characterization. J. Cell Biol. **151**, 221–234

18 Laemmli, U. K. (1970) Cleavage of structural proteins during the assembly of the head of bacteriophage T4. 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**, 167–180

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**, 607–634

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**, 9678–9686

25 Perez-Vilar, J. and Hill, R. L. (1999) The structure and assembly of secreted mucins. J. Biol. Chem. **274**, 31751–31754

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**, 745–757

27 Michele, D. E. and Campbell, K. P. (2003) Dystrophin–glycoprotein complex: post-translational processing and dystroglycan function. J. Biol. Chem. **278**, 15457–15460

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**, 1861–1867

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**, 9900–9912

Received 10 May 2006/18 August 2006; accepted 18 September 2006
Published as BJ Immediate Publication 18 September 2006, doi:10.1042/BJ20060691