

Neogenin, an Avian Cell Surface Protein Expressed during Terminal Neuronal Differentiation, Is Closely Related to the Human Tumor Suppressor Molecule Deleted in Colorectal Cancer

Jost Vielmetter, Jon F. Kayyem, Janet M. Roman, and William J. Dreyer

Division of Biology, 156-29, California Institute of Technology, Pasadena, California 91125

Abstract. Using a monoclonal antibody, we have identified and characterized a previously unknown cell surface protein in chicken that we call neogenin and have determined its primary sequence. The deduced amino acid sequence and structure of neogenin characterize it as a member of the immunoglobulin (Ig) superfamily. Based on amino acid sequence similarities, neogenin is closely related to the human tumor suppressor molecule DCC (deleted in colorectal cancer). Neogenin and DCC define a subgroup of Ig superfamily proteins structurally distinct from other Ig molecules such as N-CAM, Ng-CAM, and Bravo/Nr-CAM. As revealed by antibody staining of tissue sections and Western blots, neogenin expression correlates with the

onset of neuronal differentiation. Neogenin is also found on cells in the lower gastrointestinal tract of embryonic chickens. DCC has been observed in human neural tissues and has been shown to be essential for terminal differentiation of specific cell types in the adult human colon. These parallels suggest that neogenin, like DCC, is functionally involved in the transition from cell proliferation to terminal differentiation of specific cell types. Since neogenin is expressed on growing neurites and downregulated at termination of neurite growth, it may also play an important role in many of the complex functional aspects of neurite extension and intercellular signaling.

DURING embryogenesis, cells undergo a transition from proliferation to differentiation. This transition often involves specific cell surface proteins with receptor and signal transduction functions. The same molecular mechanisms regulating cell growth and differentiation also play a key role in aberrant and uncontrolled growth leading to tumors (Aaronson, 1991). Tumors are in many cases caused by the malfunction of molecular processes mediating the transition from a proliferative to a differentiated state. In this context, proto-oncogenes have been defined as molecules that are actively involved in keeping cells in the proliferative state (Solomon et al., 1991). On the other hand, tumor suppressor genes have been defined as genes whose expression is essential for normal control of growth and cessation of proliferation leading to differentiation, therefore deletion of these genes promotes abnormal growth. This latter category of proteins includes p53, the retinoblastoma gene (RB), the Wilms tumor gene (WT), the erbA gene, the neurofibromatosis tumor suppressor gene (NF-1), and the deleted in colorectal carcinoma gene (DCC),¹ the only cell

surface protein among the known tumor suppressor molecules (Weinberg, 1991, 1993). Because of its role in tumor growth DCC is likely to be associated with receptor and signal transduction functions mediating the transition from proliferation to terminal differentiation in certain cell types (Fearon et al., 1990; Hedrick et al., 1992; Lawlor and Narayanan, 1992).

Neuronal precursor cells, like other cell types, undergo terminal differentiation at an appropriate point in development. In neurons this differentiation is associated with the outgrowth of cell processes, the neurites. Many cell surface molecules that play a role in neurite outgrowth are members of the Ig superfamily (Jessell, 1988; Rathjen et al., 1992; Sonderegger and Rathjen, 1992; Brummendorf and Rathjen, 1993; Goodman and Shatz, 1993). Some of these molecules are known to mediate general adhesion between cells or their processes, some appear to mediate specific recognition between certain cell or fiber populations, and others cause repulsion of approaching neurites (Patterson, 1988; Jessell, 1988; Kaprielian and Patterson, 1994). All of these molecules are believed to contribute to the proper assembly of the nervous system, and similar mechanisms are believed to be

Address all correspondence to J. Vielmetter, Division of Biology, 156-29, California Institute of Technology, Pasadena, CA 91125. Tel.: (818) 395-4961. Fax: (818) 449-0756.

1. *Abbreviations used in this paper:* DCC, deleted in colorectal carcinoma; E5, embryonic day 5; EGL, external granule cell layer; FNIII, fibronectin

type III; Ig, immunoglobulin; IGL, internal granule cell layer; IPL, inner plexiform layer; KLH, keyhole limpet hemocyanin; ML, molecular layer; OFL, optic fiber layer; OPL, outer plexiform layer; ORF, open reading frame; PFA, paraformaldehyde; Pl, posthatching day 1.

involved in the formation of other organs during embryonic development.

Here we describe the cloning, sequencing and certain aspects of the expression of neogenin. We choose the name "neogenin" because we find it expressed mainly in newly generated tissues as their cells undergo terminal differentiation. Throughout the terminal differentiation of neurons, neogenin undergoes dynamic changes in its level of expression, being upregulated on neuronal precursor cells with the initiation of process outgrowth and cessation of cell division, and downregulated as neurite growth is completed. We have found that, unlike other neuronal cell surface molecules studied so far, neogenin shares significant similarities with the tumor suppressor molecule DCC, and therefore we propose that neogenin and DCC constitute a new subfamily within the Ig superfamily of cell surface molecules. Based on its relationship to DCC and on its expression primarily in regions of neuronal differentiation, we suggest that neogenin is involved in terminal differentiation, neurite extension, and cellular signaling.

Materials and Methods

Production of Monoclonal Antibodies

The monoclonal antibody 10-22A8 used in these experiments was generated using a procedure previously described (Kayyem et al., 1992b). In brief: to isolate cell surface molecules from developing neural tissues, intact tecta from 1,000 embryonic day 8 (E8) chicken embryos (obtained from Lake View Farm, Lake View, CA) were dissected. The tissue (20 tecta at a time) was labeled with biotin-X-NHS (Calbiochem-Behring Corp., La Jolla, CA). The biotinylated tissue was washed and homogenized, and the homogenate containing biotinylated cell surface molecules was loaded onto an avidin-(monomeric)-agarose column (Sigma Chemical Co., St. Louis, MO) at 18°C, washed, and eluted with 1 mg/ml free d-biotin (Sigma Chemical Co., St. Louis, MO). The biotinylated cell surface molecules were then fractionated on the basis of their molecular mass using a TSK G-4000 SW size exclusion HPLC column (Pharmacia-LKB, now available through Toso Haas, Montgomeryville, PA) and the eluate was collected in 12 fractions covering the entire molecular mass range. Part of each fraction (50 µg) was coupled to keyhole-limpet hemocyanin (KLH) (Calbiochem-Behring Corp.). Antibodies were generated in Robertsonian 8.12 female mice (Jackson Laboratories, Bar Harbor, ME). Initial immunizations were carried out by injecting emulsions of KLH-coupled antigen (ca. 50 µg) in complete Freund's adjuvant (Sigma Chemical Co.) intraperitoneally and subcutaneously into the foot pads and the tail. The initial immunization was followed by three subsequent booster immunizations given with ~20 µg unmodified antigen emulsified in incomplete Freund's adjuvant (Sigma Chemical Co.) at 3-mo intervals. The final immunization was done by injecting unmodified antigen directly into the spleen. Fusions were performed using the spleen cells of the immunized mice and FOX-NY myeloma cells (Taggart and Samloff, 1983). Supernatants of the resulting hybridoma clones were screened for reactivity on cryostat sections containing retina, optic nerve, and tectum of E8 chicken embryos using the immunohistological techniques described below. Hybridoma clones producing monoclonal antibodies with restricted staining patterns in the retina, tectum and/or optic nerve (such as 10-22A8) were subcloned and ascites fluids were prepared for further studies (Mishell and Shiigi, 1980).

Immunohistochemistry

For immunohistology, chicken embryos were washed in PBS and fixed for 5–8 h in 4% paraformaldehyde, 11% sucrose in 100 mM KH₂PO₄ buffer, pH 7.4. The fixed embryos were rinsed with PBS and soaked in 25% sucrose solution for 12 h. Embryos were transferred to fresh 25% sucrose solution and stored up to 3 wk at 4°C. Embryonic tissue was sectioned on a Tissue Tek II cryostat (Miles Laboratories, Elkhart, IN) after embedding in O.T.C.-embedding medium (Miles Laboratories, Elkhart, IN). 10-µm-thick sections were collected on pretreated microscope slides. Pretreatment of slides was as follows: washing with 1 N HCl, 95% ethanol, 1 N NaOH, and

H₂O, followed by drying, and treatment with 3-aminopropyl-triethoxysilane (APTS). Slides were rinsed with H₂O and fixed with 1% glutaraldehyde for 30 min. Collected sections were dried overnight and stored at –20°C.

Sections were rehydrated for 15 min with blocking solution consisting of PBS pH 7.4 with 10% fetal calf serum (GIBCO-BRL, Gaithersburg, MD) and then stained with a 1/1,000 dilution of ascites from mAb 10-22A8 in blocking solution. After a 1-h incubation in a moisturized chamber sections were washed three times with PBS and incubated with secondary antibody (fluorescein isothiocyanate-conjugated Fab' fragments of goat anti-mouse heavy and light chain IgG antibody, Jackson ImmunoResearch Laboratories, West Grove, PA) for 1 h. Sections were washed three times with PBS, mounted with glycerol, and photographed with a photo microscope (Zeiss, Oberkochen, Germany).

Protein Analysis and Amino-terminal Sequencing

Anti-neogenin mAb 10-22A8 was purified from mouse ascites by chromatography on a ProteinG Sepharose 4 Fast Flow column at 25°C (Pharmacia, Alameda, CA). After loading the column with ascites it was washed with 100 ml of 20 mM sodium phosphate buffer at pH 7.0 and eluted with 20 ml of 100 mM glycine-HCl at pH 2.7. The purified antibody was then dialyzed against 0.1 M NaHCO₃ buffer at pH 8.3 at 25°C. 3 µg of purified mAb was coupled to 1 ml of CNBr Sepharose 4B as recommended by the manufacturer (Pharmacia). A column was prepared and washed with 0.1 M NaHCO₃ at pH 8.5 and 0.1 M sodium acetate buffer with 0.5 M NaCl at pH 4.5. The column was then equilibrated with 0.02 M Tris-HCl, 0.14 M NaCl, 0.5% NP-40 (Sigma Chemical Company), and 0.5% Zwittergent 3-14 (Calbiochem-Behring Corp.).

Neogenin protein was purified from posthatching day 1 (P1) chicken brains following a procedure described earlier (Kayyem et al., 1992b). After ultracentrifugation (100,000 g, 5 h, 4°C) the supernatant from 200 brains was gravity loaded onto a 1 ml affinity column at a flow rate of approximately 50 ml/h at 20°C. To remove weakly bound material, the column was washed with 10 ml of 50 mM Tris-HCl, pH 8.0, containing 0.5 M NaCl and 0.5% NP-40. The antigen was eluted with 50 mM triethanolamine, pH 11.5, containing 150 mM NaCl and 0.1% NP-40. Fractions of 1 ml were collected in tubes containing 0.3 ml of 1 M Tris-HCl, pH 6.7 to neutralize the high pH solution. The fractions were analyzed using SDS-PAGE and peak fractions were concentrated by centrifugation in a Centricon-30 vial (Amicon Corp., Danvers, CO).

In preparation for amino-terminal sequencing, neogenin (50–100 pmol) was purified by affinity chromatography as described above and electroblotted onto polyvinylidene difluoride (PVDF) membrane (Applied Biosystems, Foster City, CA) (Matsudaira, 1987) following SDS-PAGE under reducing conditions. The membrane was stained with 0.1% Coomassie blue in 40% methanol and 1% acetic acid. Visible protein bands with a molecular weight corresponding to neogenin were cut from the membrane and subjected to a 20 h wash in 30% methanol. Amino-terminal sequencing was attempted on a protein sequenator (model 477A; Applied Biosystems, Foster City, CA). Due to amino-terminal blockage, no sequence data could be obtained.

Purified proteins or detergent lysates from tissues were separated by SDS-PAGE (8 or 10% polyacrylamide) using standard methods (Laemmli, 1970) and silver stained using the method of Ansorge (1985). DTT (10 mM) or iodoacetamide (10 mM) was added to the protein samples before analysis by SDS-PAGE to produce reducing or non-reducing conditions, respectively. Molecular weights were determined by comparison with prestained molecular weight standards from Bio-Rad Laboratories: 205 kD, myosin; 116 kD, β-galactosidase; 80 kD, BSA; and 49 kD, ovalbumin (Bio-Rad Laboratories).

For Western blots, proteins were transferred to nitrocellulose membranes (Schleicher & Schuell) using published procedures (Towbin et al., 1979). After blocking sites on the membrane overnight with 5% nonfat dry milk (Carnation Company, Los Angeles, CA) in PBS at pH 7.0, the membranes were incubated with ascites fluid at a 1/1,000 dilution in 5% nonfat dry milk in PBS at pH 7.0, for 3 h at room temperature. After washing the filters three times with PBS containing 0.05% Tween-20 (Sigma Chemical Co.), the filters were incubated in alkaline phosphatase-conjugated goat anti-mouse IgG (GIBCO-BRL) diluted 1/5,000 in the same solution as the primary antibody for 3 h. After washing the filters three times with PBS containing 0.05% Tween-20, proteins were visualized by reacting the filters with nitro blue tetrazolium (NBT, 300 µg/ml) and 5-bromo-4-chloro-3-indolyl-phosphate (BCIP, 150 µg/ml) (Stratagene) in 100 mM Tris-HCl, pH 9.5, containing 100 mM NaCl and 5 mM MgCl₂.

DNA Cloning and Sequencing

To isolate total RNA, embryonic brain tissues were homogenized in cesium trifluoroacetic acid (Carter et al., 1983). Poly(A)⁺ RNA was isolated from total RNA using oligo deoxy thymidine (dT) spin columns from Pharmacia Fine Chemicals (Alameda, CA). cDNA libraries were constructed using random primers for cDNA synthesis and the λ -ZAPII expression-cloning system (Stratagene) for cDNA cloning. One cDNA library was prepared for us by Stratagene from RNA prepared as described above and a second one was kindly provided to us by Peter Sonderegger and colleagues (Zürich, Switzerland). In addition, an amplified λ -FIXII genomic library prepared from male White Leghorn chicken spleen was used (Stratagene; kindly provided by Hirdeyapal Bhathal, San Diego, CA).

The E18-cerebellum cDNA library (prepared by Stratagene) was screened with a 7.5×10^{-4} dilution of ascites fluid from monoclonal antibody 10-22A8 using standard procedures (Sambrook et al., 1989). Clones reactive with mAb 10-22A8 were subcloned by in vivo excision into the plasmid pBluescript and sequenced using primers synthesized in the Caltech Microchemical Facility (Pasadena, CA). Sequencing was done using the dideoxy sequencing technique with the Sequenase version 2 kit (United States Biochemicals, Cleveland, OH) (Sanger et al., 1977).

To identify additional neogenin sequences, fragments from the initial clones were amplified by PCR and labeled with ³²P using a random oligonucleotide priming kit (Boehringer Mannheim Biochemicals) (Feinberg and Vogelstein, 1984). These probes were used to screen cDNA and genomic libraries (Sambrook et al., 1989). Clone purification and subcloning was done following standard procedures (Sambrook et al., 1989). Clones were mapped by PCR using polylinker primers in combination with primers matching the coding sequence.

Sequence analysis of genomic clones identified with 5' terminal cDNA probes revealed a large intron at position 29 presenting an obstacle to the identification of the start codon. To elucidate this region anchored PCR was carried out using a method that implements single strand ligation of an anchor sequence. Here the anchor oligonucleotide is attached to the 3' end of the first strand cDNA representing the 5' end of the RNA (Clontech, Palo Alto, CA) (Dumas et al., 1991). The anchored PCR experiment failed to yield 5' terminal sequence including the start codon, probably due to secondary structure at the GC-rich 5' end (Fig. 2, bases 1-22).

Sequence data were analyzed and compiled using the GCG software package on a VAX/VMS system (Devereux et al., 1984). The translated neogenin sequence was compared pairwise with a protein translation from GenBank using the BLAST program (Altschul et al., 1990). Gene trees were compiled using the program Pileup (Higgins and Sharp, 1989).

Northern Blot and PCR Analysis to Confirm Alternative Splicing

For Northern blot analysis poly(A)⁺ RNA from E8 chicken brain (see above) and "0.24-9.5-kb RNA Ladder" as size standards (GIBCO-BRL) were fractionated on 0.7% denaturing agarose gels (7.4% formaldehyde, 20 mM MOPS, 5 mM sodium acetate, 1 mM EDTA (modified after Sambrook

et al., 1989). The RNA was blotted onto nitrocellulose membranes following standard procedures (Sambrook et al., 1989). The size markers were visualized by staining the blot with methylene blue (Sambrook et al., 1989). Probes were generated by PCR amplification and radiolabeled as described above for the library screening. Prehybridization was carried out at 37°C for 2 h in 4× SSPE (35 mM NaH₂PO₄, 600 mM NaCl, 4 mM EDTA, pH 7.4), 5× Denhardt's solution, 50% formamide, 1% SDS, and 100 µg/ml fragmented and heat-denatured salmon sperm DNA (modified after Sambrook et al., 1989). For probe hybridization 10% dextran sulfate was dissolved in the prehybridization solution and 2 ng/ml α -³²P-deoxy-ATP-labeled heat-denatured probe (1×10^5 – 5×10^5 cpm/ml) was added. Hybridization was carried out at 37°C for 20 h. Washes were as follows: 1) 10 min at 25°C in 2× SSC (300 mM NaCl, 34 mM sodium citrate, pH 7.0) + 0.1% SDS; 2) 10 min at 25°C in 0.5× SSC (75 mM NaCl, 8.5 mM sodium citrate, pH 7.0) + 0.1% SDS, and 3) 10 min at 65°C in 0.1× SSC (15 mM NaCl, 1.7 mM sodium citrate, pH 7.0) + 0.1% SDS. Autoradiography was for 16 h with enhancing screen at –70°C using X-AR 5 x-ray film (Eastman Kodak, Rochester, NY) for 16 h.

PCR experiments to confirm alternative splicing were performed under the following conditions: using standard conditions for AMV-reverse transcriptase (Boehringer Mannheim Biochemicals) two reverse transcriptions were performed, in each of which 5 µg poly(A)⁺ RNA from E8 chicken brain were primed with neogenin specific primer 1 (PCR1): 5'-AGGGCAAT-CACACCACTGAG-3' and primer 2 (PCR2): 5'-CTGGAACCAACACAGG-CAT-3'. 1 µl of reverse-transcribed, single-stranded neogenin template was used in each 20 µl PCR reaction. The primers used in the PCR reactions were primer 3 (PCR3): 5'-TTCACCAAAGGCGGAAGTCC-3', primer 4 (PCR4): 5'-AAAGACTGGAGCTAAAGCCC-3', and primer 5 (PCR5): 5'-TCAGAGCTCGGTGTGTTCCG-3'. In one PCR experiment PCR1 primed template and primer pair PCR3/PCR5 were used (Fig. 4 b, lane 2). In another reaction PCR2 primed template and primer pair PCR4/PCR5 were used (Fig. 4 b, lane 3). The 20 µl reaction mixtures were composed of: 1 µl template per 20 µl reaction, 1 mM each primer, 200 mM each deoxy nucleotide, 10 mM Tris-HCl, pH 8.3, 50 mM KCl, 1.5 mM MgCl₂, 0.001% gelatine, 0.5 U Amplitaq (Perkin-Elmer, Norwalk, CT). To avoid unspecific priming the enzyme was added after the reaction mixture was covered with mineral oil and heated to 93°C (hot start PCR). The thermocycling conditions were: 60°C for 1 min, 70°C for 1 min, and 93°C for 30 s repeated for 35 cycles, followed by 70°C for 5 min. PCR products and 100-bp ladder (Pharmacia, Alameda, CA) as DNA size standards were separated on 1% agarose gels and stained with ethidium bromide following standard protocols (after Sambrook et al., 1989).

Results

Isolation of cDNA Clones and Determination of Sequence Coding for Neogenin

The monoclonal antibody 10-22A8 was used to screen a random-primed λ -ZAPII cDNA library constructed from

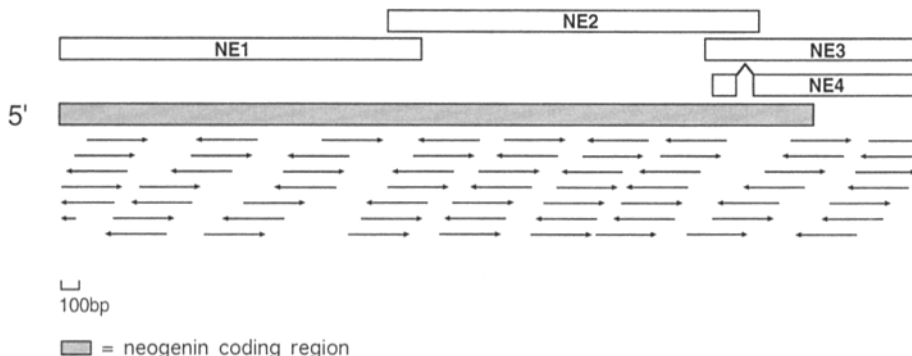


Figure 1. Schematic illustration of neogenin cDNA clones and sequence fragments. Four overlapping clones NE1 to NE4 containing the sequence coding for neogenin are represented by the open bars, with the 5' ends on the left side and the 3' ends on the right side. The gap in clone NE4 bridged by the angled line indicates the sequence missing from this cDNA clone, which corresponds to the intracellular alternative exon. The shaded bar indicates the extent of the coding sequence. The arrows represent sequencing extended from particular primer fragments and indicate their direction.

E18 chicken cerebellar mRNA. From a screening of $\sim 2 \times 10^6$ phage plaques, seven clones were isolated, one of which (NE1) was further analyzed (Fig. 1). Additional overlapping cDNA clones NE2, NE3, and NE4 were obtained by screening the same cDNA library and an additional oligo-dT primed λ -ZAPII cDNA library (constructed from E14 chicken brains) with radio-labeled fragments from clones obtained in preceding screenings. A total of four overlapping cDNA clones were selected to determine the sequence coding for neogenin (Fig. 1). A contiguous 4,781-bp sequence obtained from these clones and the deduced protein sequence are shown in Fig. 2, and a schematic representation of the domain structure in Fig. 3. Only one large open read-

ing frame (ORF) could be identified and it codes for a protein of at least 1,442 amino acids with a calculated molecular mass of 157.9 kD.

The 5' terminal sequence identified is extremely G/C rich (Fig. 2). Secondary structure artifacts in this G/C-rich region might be involved in the premature termination of cDNA synthesis and the resulting absence of an identified start codon. Sequence similarity with DCC (Fig. 6) and the nearly identical apparent molecular weights of the proteins suggest that the missing cDNA sequence most likely encodes only amino acid residues of the processed signal peptide. However, since the amino terminus of neogenin is blocked and the specific amino-terminal end for processed DCC pro-

```

1  R S G P R S P L T G S V V R T F T P T F L V E P M D I L S V R G A S Y T H
2  C T A G C G C C C C C G A G C C C C C C A G C G C C C C T A G C A G C C C C C T A G C A G C C C C C T A G C A G C C C C C
3  S V S C E T P P K I E N K K D G C T C L N L V S D O R R Q L L P D G S L I N S V
4  T A T T A T T T T A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A
5  V H S K N H P D S V V V Q G V A T V E S L G S I V S R T A K L V A C L P R P
6  G R U C L A T T C A G C A T A T A A A C C G A T A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A
7  T Q P E L S S V Y K U N S A I L N C E V N V D L A P F V R W E D R O P L S L
8  A C C A G C A G C A G C A T A T C T C T A T A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A
9  D R V F P L P S G A L L I C N A T D T D G O P Y R V I E S G D T P K Y S E R
10 G A T A C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C
11 A E L K L L P D S S V S L V F V V D P S S L T K V T C G V A V F P Q V I Q D
12 G A G A C A C A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A
13 F P T P V V R M T K N C E E L I T E D S E R F A L R A G O S L L I S D V T E E D
14 T T T C A A T C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C
15 V G T C T T A D N R N E T I P A Q A B L A V Q V P P F E L K R P A N I V A H E
16 C T T C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C
17 S K D I V F S E T T U X P T P T V X N V K N G D V I P F S F K V I V H N
18 T A T T A G C A T C T C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C
19 L Q V L G L V K S D E G F Y O I A E N D V G N A Q A Q L T I L D L V A I
20 C T C A G C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C
21 P T L P P T S L T S A T N D H A P A T T O P L P T A P R D V V A T L V S T R F
22 C T A A K A T C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C
23 I R I T F V S D P Q G D L T V S I F Y T K E G I N R R V E T S E R P G
24 A T C A G C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C
25 E T O V H I O N L M P E T V V V F R V V A Q N K K H G H E S S A P L K V A T Q P
26 G A G A C A C A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A
27 E V O L P G P A P H I R A V A G S P T S V T V T E T P L S G N G E I Q N Y K L
28 G A G T P K A C C T A C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C
29 V Y M E K C O D S E Q C V D V A C L V P T I T O L K V T E S F R V Y A N K
30 T A C T A C T A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A
31 H P Q V S T A Q D V V R T L T V S P A A P O L T L E A R N S K S I N L R
32 C A C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C
33 O P P A G T H S G Q I T G Y K I R Y R P V S R K S D V T E S V G C T O L F Q L
34 C A G C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C
35 I E G L E R C T E Y N F R I A A H T V R O T G P A T D W T S A E T F E S D L R
36 A T H A G A C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C
37 S R V P E V P S S L H V R P L V T S I V V S T P P E N Q N I V V R O Y A I G Y
38 G C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C
39 C T G S P H A O T I X V O Y K Q R Y V T L E N L D P S S H Y V I T K A F N N V
40 G C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C
41 O G G I P L Y S S A V T S P H S D T S E V O L F V I N A P Y T P V P D P S M H
42 C T G A A G A T P T C T A T A G A G C C A G C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C
43 P P V G V Q A S I L S H D T I R I T A D N S L P K N Q K I T D A R Y T T V R W
44 C A C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C
45 A T N I P A R A T V T V R I A T V R I A T V R I A T V R I A T V R I A T V R I A T V R I A T V R I A T V R I A T V R I A T
46 A A A A A A T A T T C T A A T A C A A A T A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A
47 R S S T W S H T A R G T T F P L V P T S P P K D V T V S K E G K P R T I E V
48 Q A A G A C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C
49 N Q P P S E A N G K I T G Y I L Y S T D V N A E I H D W V I S P V V G R L
50 A A C T A G C A C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C
51 T H O I D E L T L D T P F V P R I C A R N S R K N G F M S E A V O R T P K A E
52 A C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C
53 C S D K M P N D Q A S C S A G C G S K P V D V O P D Y K P P L S C S N S P H G S
54 T L C T A G A T A A A A T A A T A A T A A T A A T A A T A A T A A T A A T A A T A A T A A T A A T A A T A A T A A T A A
55 P T S P L O S N H L L V I L V S V G V I T V V V V V A V F C T R R T T S H Q
56 C G A T T C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C
57 H K R A A C K S V D H N H K E R P D V R I E S P A V A V A R G S A Y D P P L
58 A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A
59 K S P D N P I M T D T P I P R N S Q D I P V O N S H D S N T H O R N N S V R
60 A A A T T C A G A T C A T A C A G A T A C C C A A T C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C
61 U H E S R D S N S T L A C R P D R K R K N N N F D S Q P P Q P P L V I A N A P T H
62 G A G A T M P T C G A G A A A A T P T T C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C
63 S L O N P H M H F S G S L A S P T R S Y L N N Q V S P H P V G T E M S E D R
64 A C T C A T A A C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C
65 A N S T E S V R N T P S S D T M P A S S S Q P C A D H Q D P D S S R S C A V L Q S
66 A G E A D A C S U P T A H V S R I E L S F A V A V A R G S A Y D P P L
67 A Q E Z A A C S U P T A H V S R I E L S F A V A V A R G S A Y D P P L
68 C T C A G A G A G A G A G A G A G A G A G A G A G A G A G A G A G A G A G A G A G A G A G A G A G A G A G A G A G A G
69 P S T P L L T Q A P S R P V H S V K T A S I G T L Q R T R P P H P V V P A
70 C C A A C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C
71 F D V O E T R M L D S E S V E P D E L T R E M A H L E G L M K C L N A I T
72 C T T C K C T C G A G A A C A C A G A T C A G A G A G A G A G A G A G A G A G A G A G A G A G A G A G A G A G A G A G A
73 T A
74 A G A T A C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C
75 G A G A G A T C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C
76 C T C C A T T A C A G A T C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C
77 T A C A T T C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C
78 T A C A T T C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C
79 T A C A T T C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C
80

```

Figure 2. Nucleotide sequence and deduced amino acid sequence of neogenin. The ORF codes for 1,442 amino acids. The 23 amino acids (1089 to 1111) spanning the putative transmembrane region are underlined. The nucleotides comprising an alternative exon in the intracellular region are underlined (3695 to 3853). Note the accumulation of histidines with potential metal binding properties within this exon. Eight potential sites for asparagine-linked glycosylation are marked with inverted triangles and 14 potential intracellular phosphorylation sites are indicated by asterisks. The conserved cysteines of the Ig domains are circled and the conserved tryptophans of the FNIII domains are boxed. The sequence data are available from EMBL/GenBank/DBJ under accession number U07644.

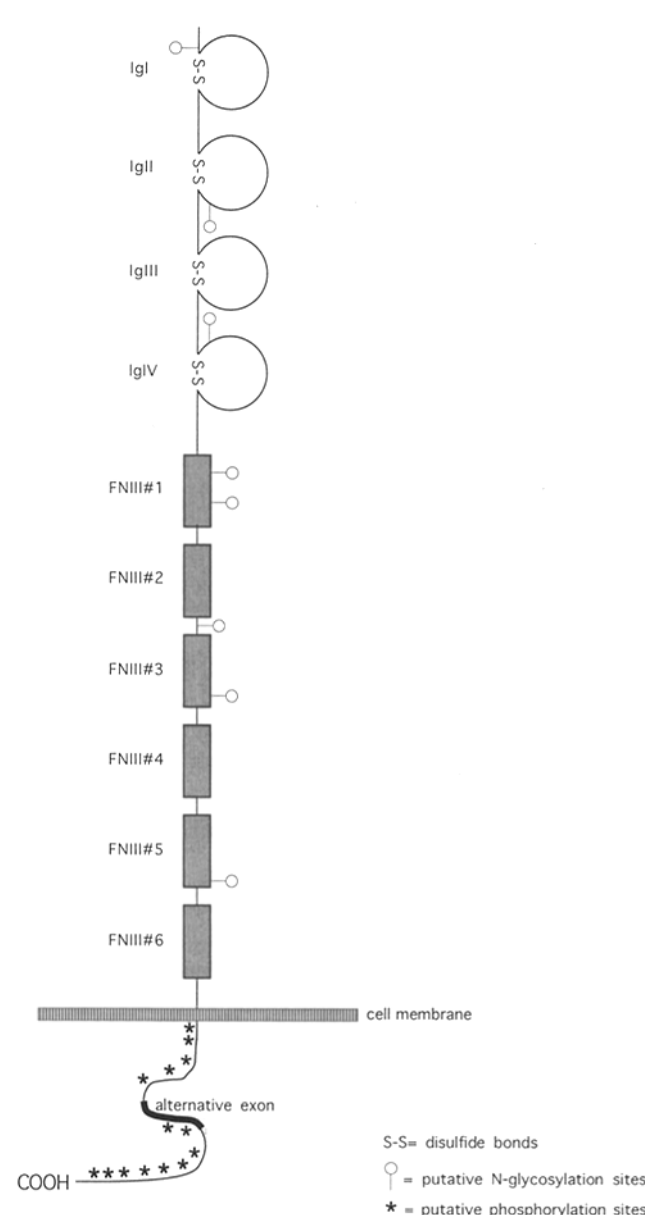


Figure 3. Schematic structural model of neogenin. The four Ig domains numbered Igl through IgIV are illustrated as disulfide-linked (S-S) loops as expected for Ig domains. The six FNIII domains numbered FNIII#1 through FNIII#6 are represented by the filled boxes. The stemmed circles denote potential asparagine-linked glycosylation sites.

tein has not been determined, we cannot conclude which amino acid constitutes the exact amino terminus of either protein in its mature form. For orientation we therefore refer to the first base and amino acid determined from cDNA sequences as base and amino acid 1.

A stop codon terminates the large ORF at position 4326, and a potential polyadenylation signal is encoded at position 4762 (Proudfoot and Brownlee, 1976). No polyadenylation tail was observed in any of the clones found.

Neogenin Features Structural Elements of the Immunoglobulin Superfamily of Proteins

Based on the knowledge about the structure of proteins belonging to the Ig superfamily, the amino acid sequence of neogenin allows us to predict four types of major structural characteristics as indicated in Fig. 3. The extracellular part of the molecule is comprised of four Ig domains or loops, each containing a highly conserved pair of putatively disulfide-bonded cysteines (circled in Fig. 2). All four Ig domains show both the spacing between the conserved cysteine residues of about 50 amino acids as well as the presence of conserved motifs near the cysteines, features that place them into the C2 subcategory of Ig domains (Williams and Barkley, 1988).

The Ig domain region is followed by six fibronectin-type III (FNIII) repeats that are characterized by a tryptophan (boxed in Fig. 2) followed by a tyrosine residue typically separated by about 50-amino acid residues (Engel, 1991).

The extracellular domains are followed by the strongly hydrophobic putative transmembrane region (residues 1089 to 1111, underlined in Fig. 2) as determined by a hydrophobicity analysis (Kyte and Doolittle, 1982). The remaining 331 amino acid residues constitute the cytoplasmatic domain, which contains a site for potential alternative splicing (nucleotides are underlined in Fig. 2). This is revealed by cDNA clone NE3, which contains an additional 159-bp sequence (nucleotides 3694 to 3853) coding for amino acids 1232 to 1285, in contrast to clone NE4, which is missing this sequence (Fig. 1).

The extracellular part of neogenin contains eight putative sites for asparagine-linked glycosylation indicated by the inverted triangles in Fig. 2 (Marshall, 1972). The cytoplasmatic domain contains 14 potential serine and threonine phosphorylation sites marked by the asterisks in Fig. 2 (Woodgett et al., 1986).

Analysis of RNA and Alternative Splicing

In order to determine the size of the mRNA coding for neogenin and to obtain additional evidence of alternative splicing, a Northern analysis was performed. A Northern blot of poly(A)⁺ RNA from E8 chicken embryos was probed with a radiolabeled 1 kb neogenin probe spanning nucleotides 22 to 1040. The autoradiograph (Fig. 4 a) reveals a doublet band between 7.5 and 9 kb, with the top band being fainter than the bottom band. This finding indicates size heterogeneity in the visualized mRNA population and suggests alternative RNA splicing. To further substantiate the alternative splicing of the putative 159-bp alternative exon (nucleotides 3694 to 3853), we performed PCR experiments with primers close to the putative alternative exon borders. As template, single stranded cDNA generated from poly(A)⁺ RNA from E8 chicken brain (source identical to

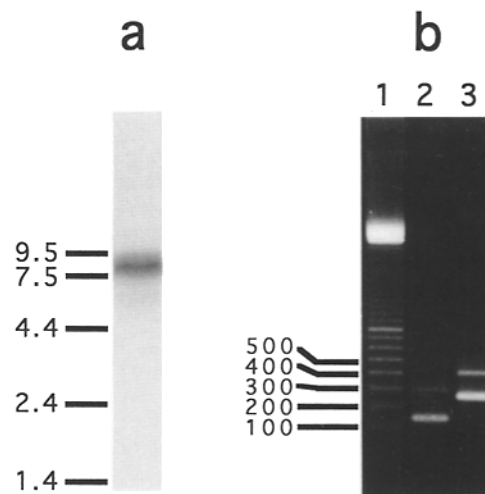


Figure 4. Northern blot and PCR data to demonstrate alternative splicing of neogenin. (a) Autoradiograph of a Northern blot of poly(A)⁺ RNA (10 μ g) from E8 chicken embryos probed with a neogenin probe spanning nucleotides 22 to 1040. The sizes are determined using RNA size standards and are given in kb to the left. Note the doublet band between 7.5- and 9-kb standards, with the larger band of \sim 9 kb being much fainter. (b) PCR products are shown separated on a 1% agarose gel stained with ethidium bromide. Lanes 2 and 3 contain PCR products obtained by using two different sets of primers close to the borders of the putative alternatively spliced exon. The template used in these PCR experiments is single stranded cDNA generated from poly(A)⁺ RNA from E8 chicken brain (source identical to that used in the Northern blot). The expected lengths for the PCR fragments are 307 and 146 bp in lane 2 and, 427 and 266 bp in lane 3. Lane 1 contains molecular size markers ("100-bp ladder") with the sizes indicated in bp to the left.

that used in the Northern blot) was used. For one primer pair the expected fragment lengths are 307 and 146 bp (Fig. 4 b, lane 2) and for the other primer pair the expected lengths are 427 and 266 bp, respectively (Fig. 4 b, lane 3). The actual sizes of these PCR products closely match the expected sizes as can be seen by comparison to the DNA size markers (100-bp DNA ladder, Fig. 4 b, lane 1). These PCR results show that both alternatively spliced RNA forms (the one with the 159-bp exon and the one without this exon) are present in the poly(A)⁺ RNA population.

Apparent Molecular Weight of Neogenin Protein

Neogenin was visualized using immunoblotting with mAb 10-22A8. Whole brain lysates from E12, P1, and adult chickens were resolved on an 8% polyacrylamide gel under non-reducing conditions (Fig. 5): a compact band at 190 kD appears for embryonic (E12, lane 1), newly hatched (P1, lane 2), and adult (lane 3) chicken brain extract. Occasionally very faint, diffuse bands at 100 and 170 kD were detected. These bands were sometimes more prevalent in samples that had been subjected to conditions where proteolysis might have occurred (data not shown). We therefore believe that these bands include mainly proteolytic fragments of the native molecule and that the apparent molecular weight of the neogenin protein is 190 kD. The difference between 190 kD seen in SDS-PAGE and a theoretical molecular mass of

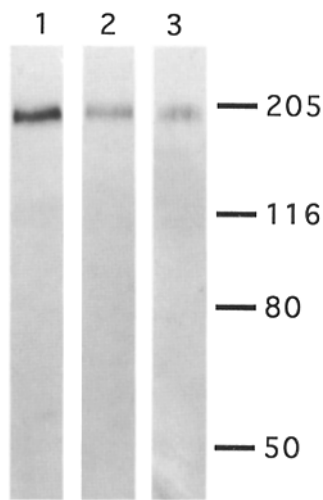


Figure 5. Immunoblots of E13 (lane 1), P1 (lane 2), and adult chicken brain (lane 3) detergent extracts. Each lane contains protein extracted from approximately 2 mg wet weight tissue. The protein was separated on an 8% SDS-polyacrylamide gel under non-reducing conditions. The mAb 10-22A8 was used to detect neogenin and was visualized with alkaline phosphatase-conjugated goat anti-mouse secondary antibody followed by nitro blue tetrazolium (NBT) and 5-bromo-4-chloro-3-indolyl-phosphate (BCIP) color development. Prestained molecular weight standards (molecular weights are shown in kD) were used to determine the apparent molecular weight of the visualized antigen.

150 kD (calculated from the amino acid sequence) is presumed to be due to glycosylation, as is typical for other Ig superfamily molecules.

Sequence Similarities of Neogenin with Known Proteins

Comparison of the deduced amino acid sequence of neogenin with known protein sequences in the GenBank database revealed significant similarity with the coding sequence of the tumor suppressor molecule DCC (Cho et al., 1993; GenBank accession number: X76132). Fig. 6 shows an alignment of the deduced amino acid sequences of neogenin and DCC. The amino acid similarity and identity scores for this alignment are 70 and 53%, respectively (Needleman and Wunsch, 1970).

No significant similarity of the intracellular domain to any known protein other than DCC could be found by searching the database.

The Expression Pattern of Neogenin at Different Developmental Stages Correlates with the Onset of Terminal Differentiation

We studied the expression of neogenin in the chick retina and cerebellum by staining cryostat sections of different embryonic stages with the monoclonal antibody 10-22A8. These results show a significant correlation of neogenin expression with areas of neuronal maturation. At E5, staining is visible on retinal ganglion cell axons that comprise the optic fiber layer (OFL) (Fig. 7a). At this stage more fibers have developed on the temporal half of the retina (the left side) than on the nasal half (Goldberg and Coulombre, 1972), and correlated with this neuronal development, neogenin expression extends farther on the temporal side than on the nasal side (left). By E6 staining is observed in both nasal and temporal retina, though still confined mainly to the OFL (Fig. 7). At E8 the staining of the temporal optic fiber layer (left) is fainter than on the nasal part (right) of the retina (Fig. 7

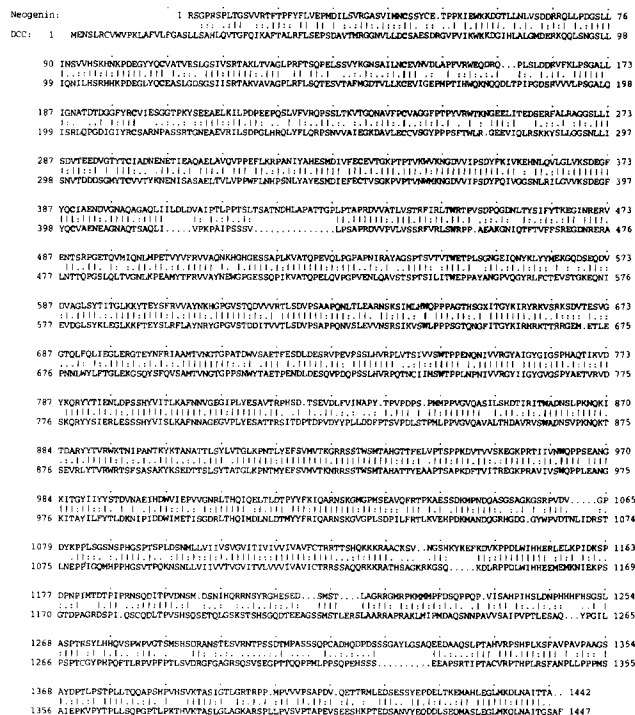


Figure 6. Alignment of the amino acid sequences of neogenin and DCC. To align both sequences the Gap program (GCG) was used. Identity of residues is indicated by a vertical dash, close similarity with a double dot, and more distant similarity with a single dot. The conserved cysteines of the Ig domains and the conserved tryptophans of the FNIII repeats are highlighted by boldface type. Note that domains differ in the extent of conservation of sequence similarity. The alternatively spliced domain is included in the neogenin illustrated here.

c). This again parallels the asymmetric maturation of retinal tissue (Mey and Thanos, 1992). In addition, E8 sections reveal staining in the inner plexiform layer (IPL), which contains ganglion cell dendrites and also neurites from bipolar and amacrine cells. Staining of neogenin in the IPL appears later than in the OFL, which correlates with the fact that the processes that form this layer initiate growth later (Mey and Thanos, 1992). At E8 a very faint staining is also seen in the outer plexiform layer (OPL) that contains neurites from bipolar and horizontal cells. At E12 (Fig. 7d) the staining in the OFL has become faint, but staining is still seen in both the inner and outer plexiform layers. This correlates with the fact that the neurites continue to grow at this stage in the IPL and OPL, but the axons of the OFL have reached the tectum and have ceased growth. In the newly hatched chicken (P1), most of the neogenin staining is diminished in all layers of the retina, thus again showing the parallels between neogenin expression and neuronal terminal maturation including extension of processes. A ring of neogenin staining remains around the optic nerve exit in the IPL (Fig. 7e). No staining could be detected in sections of the adult retina, although sensitive Western blot techniques revealed the presence of some neogenin protein in adult retinal lysates (not shown), comparable to what is detected in whole brain lysates from adult chicken (Fig. 5). Thus, neogenin seems to be expressed at very low levels in the adult retina (Table I).

temporal

nasal

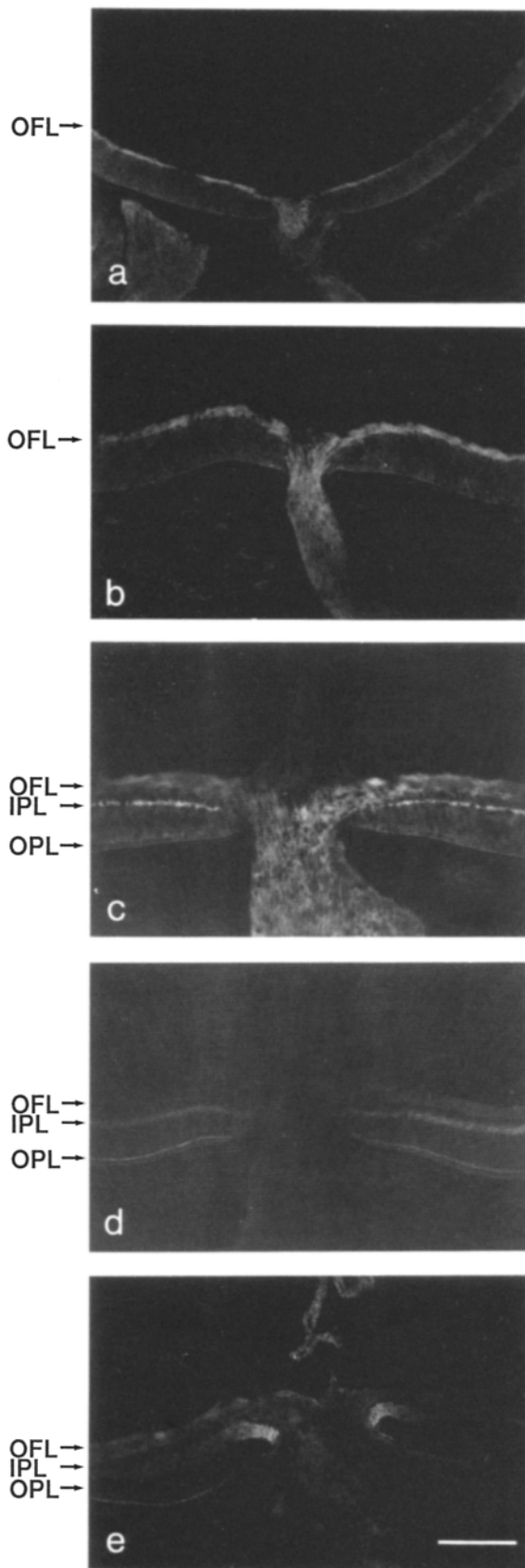


Figure 7. Immunofluorescence staining of chicken retinæ at various embryonic stages with neogenin specific antibody 10-22A8. The

In summary, the temporal appearance of neogenin expression first in the optic fiber layer, followed by the inner plexiform and finally the outer plexiform layer, correlates with the sequential maturation of neuritic layers starting with the innermost (vitreal) and ending with the outermost layers. Furthermore, the downregulation of neogenin expression seen in the optic fiber layer, starting on the temporal side, again correlates with the maturation profile of retinal axons as they extend to and reach their target tissue in the brain, the optic tectum. In each case neogenin expression is greatly reduced in concert with termination of process outgrowth.

Similar parallels between neuronal maturation and neogenin expression can be found in the cerebellum. Granule cells are generated in the posterior pole or in the rhombic lip. They migrate tangentially along the cerebellar surface, establishing the external granule cell layer (EGL), before they migrate radially into deeper cerebellar layers to form the internal granule cell layer (IGL) (LeDouarin, 1993). As a consequence of these developmental events, a maturation gradient extends from posterior to anterior. At E8, very early in cerebellar development, bright neogenin staining is seen in a posterior (left) segment, whereas the rest of the cerebellum is only faintly stained (Fig. 8 *a*). High levels of neogenin staining continue to be limited to the posterior part of the cerebellum at later developmental stages, for example, at E10 (Fig. 8, *b* and *c*) high levels of neogenin expression are detected in the most posterior segment of the cerebellum (left). Anterior to this region, where most neurons have differentiated, bright neogenin staining can still be seen in the molecular layer, consisting of the axons of granule cells. This bright staining is not seen in more anterior regions, where older axons are located. This pattern of elevated posterior staining is also seen at E13, as exemplified in Fig. 8 (*d* and *e*). It becomes less obvious in later embryonic stages but can be detected up to E17. In the cerebellum of the newly hatched chicken, the overall staining is very faint compared to the staining in embryonic tissues. Although Western blots revealed the presence of neogenin in cell lysates of adult cerebella, neogenin expression was not detected on sections stained with the anti-neogenin antibody, indicating very low levels of expression in the adult cerebellum (data not shown).

Since DCC has been found to be expressed not only in human central nervous system tissues but also in the adult colon, we studied the expression of neogenin in the lower gastrointestinal tract of embryonic, newly hatched and adult chickens. As is the case for retinal and cerebellar tissues, the lower gastrointestinal tissues express neogenin during those developmental stages that involve organogenesis (Fig. 9). Subsequent to these developmental stages (i.e., in newly hatched or adult chickens) neogenin was not detected in stained tissue sections of the lower gastrointestinal tract. It

photographs depict a central region of the retina including the optic nerve exit. Left hand corresponds to temporal and right hand to nasal retinal positions. (*a*) E5: Only the OFL is stained. The OFL does not extend over all of the nasal part of the retina at this stage. (*b*) E6: Only the OFL is stained; optic fibers now cover the entire retina. (*c*) E8: The OFL and IPL are stained; the OPL is stained very faintly. (*d*) E12: The IPL and OPL are stained. The OFL is only faintly stained. (*e*) Newly hatched chicken, P1: Staining has mostly disappeared from OFL, IPL, and OPL but staining is left in a small ring on the IPL around the optic nerve exit. Bar, 200 μ m.

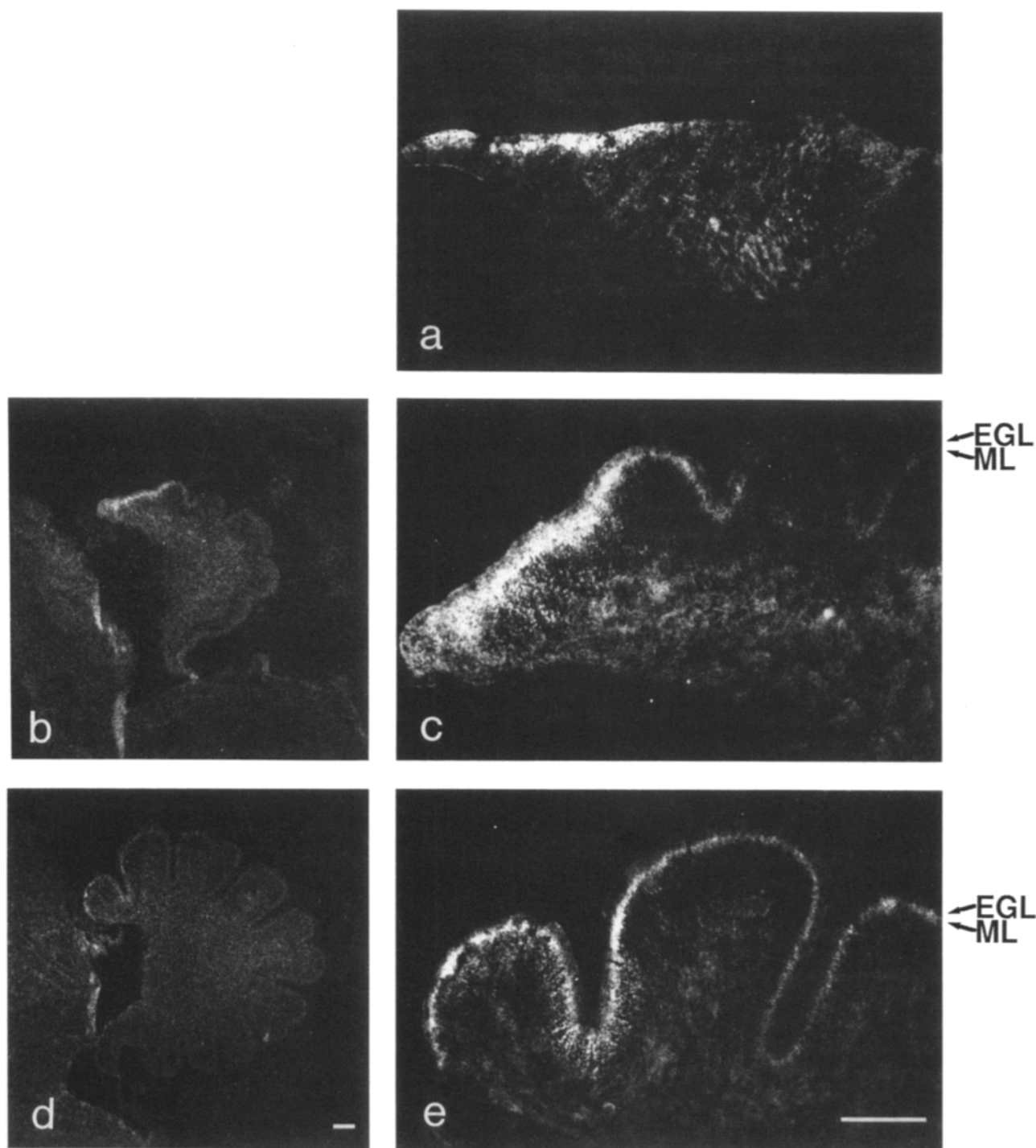


Figure 8. Immunofluorescence staining of chicken cerebella at various embryonic stages with neogenin specific antibody 10-22A8. The pictures show sagittal sections through embryonic cerebella. (a) E8: Entire cerebellum. Staining is concentrated in a posterior patch. (b) E10: Low magnification overview of cerebellum. Staining is again concentrated in a posterior patch. (c) E10: Magnified view of posterior region from b. In the most posterior part staining is spread out, whereas more anterior staining is concentrated in the ML. (d) E13: Low magnification overview of cerebellum. Staining is again concentrated in a posterior region. (e) E13: Magnified view of posterior region from d. Staining is more concentrated in the ML than in earlier stages and diminishes in intensity within the third folium. Bars: (a, c, and e) and (b and d) 200 μ m.

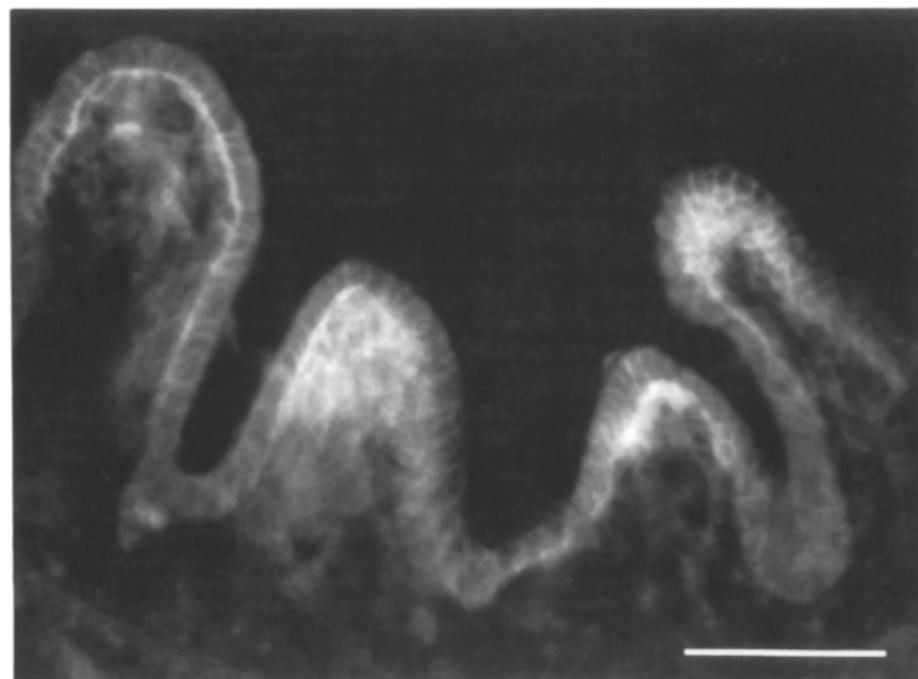


Figure 9. Immunofluorescence staining of lower intestinal tract of E19 chicken with neogenin specific antibody 10-22A8. Seen are four folds of the intestinal wall with the lumen located on the upper side of this picture. Staining is seen on numerous cells of the tissue adjacent to the gut epithelium and on the base of the cylindrical epithelial cells. Fainter staining is also seen at the contact sites between endothelial cells. Bar, 100 μ m.

was, however, detected in the more sensitive Western blots of lysed tissues. These findings are summarized and compared to our data from retina and cerebellum in Table I.

Discussion

Using the monoclonal antibody 10-22A8, we have identified and characterized a previously unknown cell surface protein, neogenin, which is expressed predominantly in the developing nervous system of the chicken.

By screening cDNA expression libraries with this monoclonal antibody, we were able to isolate and sequence cDNA clones encoding neogenin. By comparison of the deduced amino acid sequence with known protein sequences in the GenBank database we have deduced that neogenin is a member of the Ig superfamily, and is comprised, in part, of four Ig domains followed by six FNIII repeats. The sequence further suggests transmembrane and intracellular domains. The database searches also revealed a significant sequence similarity of neogenin to the human tumor suppressor molecule (DCC).

In addition to their sequence homologies there appear to be developmental and functional parallels between neogenin and DCC. First, these two molecules share similar tissue

distribution, being expressed both in neuronal and gastrointestinal tissues (Fearon et al., 1990; Hedrick et al., 1992). Secondly, the expression of both of these molecules is compatible with their involvement in terminal differentiation of specific cell types.

Neogenin is expressed in the chick retina and cerebellum in regions containing mainly cells that have initiated the transition from proliferation to terminal differentiation and cessation of cell division. We also find neogenin expression in the lower gastrointestinal tract of the chick embryo. Expression declines during embryonic development and is extremely low in the newly hatched chick and is only detected on Western blots after hatching. The timing and extent of neogenin expression differs from that of DCC, since DCC can be detected in stained tissue sections of the adult human colon (Hedrick et al., 1992).

Both neogenin in the chick nervous system and the related human DCC protein appear to be involved in providing cells with signals that trigger differentiation and prevent further cell divisions. In the case of DCC this hypothesis is strongly supported by the fact that abolishing its function by mutation in somatic cells promotes uncontrolled cell divisions, resulting in tumor growth (Fearon et al., 1990). In vitro observations further corroborate the notion that expression of DCC

Table I. Neogenin Expression in Specific Tissues during Development

	E8	E12	P1	Adult
Retina	nasal OFL bright	reduced intensity	mostly negative staining	detected only in Western blots
Cerebellum	posterior bright	posterior bright	faint staining	detected only in Western blots
Lower gastrointestinal tract	bright staining	reduced intensity	detected only in Western blots	detected only in Western blots

Summary of the results of staining either tissue sections and/or Western blots made from tissue lysates with the neogenin mAb 10-22A8. The different tissues studied are listed in the first column, and the different developmental stages examined are indicated in the first row. In most tissues of newly hatched and adult chickens, neogenin was downregulated to the extent that it was only detectable in Western blots of lysed tissues. The apparent molecular weight of the protein derived from the lower gastrointestinal tract appeared somewhat lower than normal, perhaps due to proteolytic enzymes in this tissue.

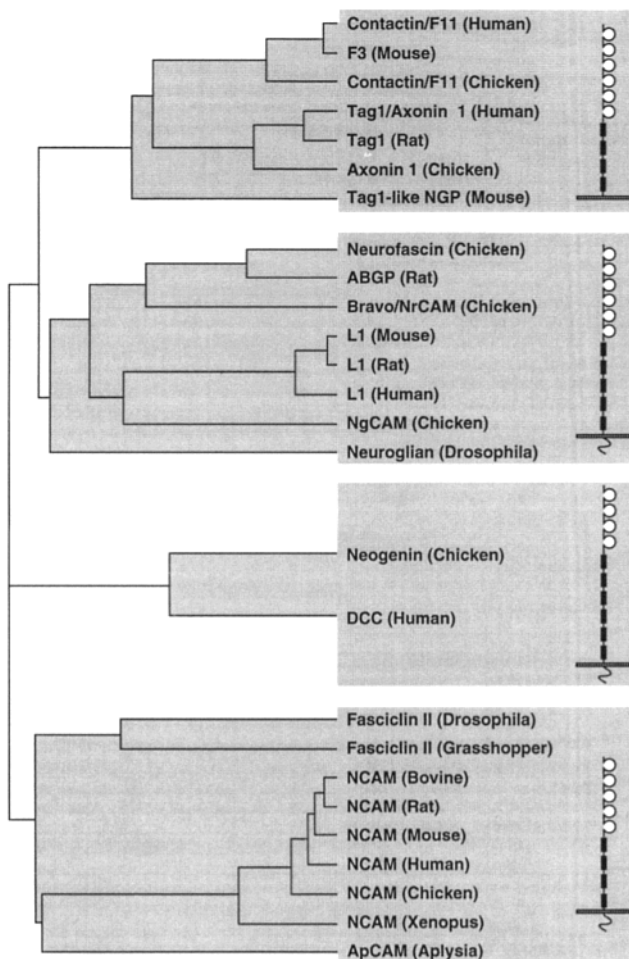


Figure 10. Dendrogram comparing the extracellular portion of neogenin, DCC and other cell surface molecules of the Ig superfamily. The dendrogram was generated using the program Pileup (GCG). The branchpoints represent average scores of pairwise comparisons between clusters attached to each branch. The closer a branch point is to the base of the dendrogram (to the left), the lower the similarity between molecules. Shaded areas indicate subgroups of molecules that share common structural models, as represented by the cartoons. The structural models read the same way as the model in Fig. 3. (Sequence data were derived from GenBank and Swiss Protein Database based on the following publications: Barthels et al., 1988; Berglund and Ranscht, 1993; GenBank accession number U07819; Bieber et al., 1989; Burgoon et al., 1991; Cho, 1993; GenBank accession number: X76132; Connely et al.: GenBank accession number: L01991; Cunningham et al., 1987; Davis et al., 1993; Dickson et al., 1987; Djabali et al., 1990; Fearon et al., 1990; Gennarini et al., 1989; Grenningloh et al., 1991; Grumet et al., 1991; Harrelson and Goodman, 1988; Hasler et al., 1993; Kayyem et al., 1992a; Lipkin and Khramtsov, 1989; Mayford et al., 1992; Miura et al., 1992; Moos et al., 1988; Ranscht and Dours, 1988; Small et al., 1987; Tonissen and Krieg, 1993; Volkmmer et al., 1992; Zuellig et al., 1992).

is required for terminal differentiation in cell lines with neuronal characteristics (Lawlor and Narayanan, 1992).

How do these findings compare to the role of other proteins in growth and differentiation of cells? It appears that no cell surface molecule other than neogenin has been described with such a dynamic range of expression during retinal or cerebellar development. For example, Ng-CAM, an-

other cell surface protein of the Ig superfamily mainly expressed on axons with long projections, is not strongly downregulated as neurite growth terminates (Thiery et al., 1985). Bravo/Nr-CAM, axonin-1, contactin/F11, and neurofascin, some of which are downregulated during late embryonic stages, are more evenly expressed in developing tissues than neogenin (Rathjen et al., 1987a,b; Ranscht and Dours, 1988; Brummendorf et al., 1989; de la Rosa et al., 1990; Grumet et al., 1991; Stoeckli et al., 1991).

In order to examine the structural relationship of neogenin to other Ig family members, pairwise comparisons were made of the extracellular parts of neogenin, DCC and several related members of the Ig superfamily. The resulting gene tree (Fig. 10) reveals that these molecules can be divided into four major subgroups, a grouping that has been described previously for three of them: the N-CAM-, the L1-, and the Tag-1-like subfamilies (Edelman and Crossin, 1991; Sonderegger and Rathjen, 1992; Brummendorf and Rathjen, 1993). These subfamilies are characterized by particular combinations of Ig and FNIII domains. This is indicated schematically by the model presentations to the right in Fig. 10. Interestingly, the gene tree illustrates that neogenin and DCC constitute a separate subfamily with close sequence similarity to each other and unique structural characteristics.

A pairwise comparison of the Ig domains of neogenin and DCC (data not shown) reveals that some domains are more conserved than others. The Ig domain, which is most conserved between neogenin and DCC is domain IV. Also, corresponding domains of neogenin and DCC are more closely related than neighboring domains on the same molecule. For example, domain IgI of neogenin is more closely related to domain IgI of DCC than it is to domain IgII of neogenin. This leads to the conclusion that Ig domain shuffling has not taken place within these molecules subsequent to the evolutionary separation, which gave rise to birds and mammals. These findings suggest functional constraints on both the domains themselves and their sequential arrangement.

The extracellular portion of neogenin contains structural elements typical of cell recognition and adhesion molecules. However, the intracellular part of neogenin does not show similarities to any previously published sequences except DCC. A peculiar feature of the presumptive alternative exon (nucleotides 3695 to 3853 underlined in Fig. 1), however, is the unusually high accumulation of histidines, not found in its DCC counterpart. Histidine-rich regions in some proteins such as DNA binding proteins and metalloproteases are associated with metal ions such as Cu^{2+} and Zn^{2+} (Christianson, 1991). This exon is also unusually rich in proline residues.

The close similarity of neogenin to a tumor suppressor molecule and neogenin's intriguing expression during development suggest a possible receptor role for this molecule in the mediation of cell signaling events leading to cessation of cell division and to terminal differentiation.

We thank Inder Nangiana, Susan Ou, and Jennifer Kim for their excellent technical assistance, Uli Schwarz for his support for initiation of this project during our visit in Tübingen, Peter Sonderegger and Richard Zuellig for kindly providing us with an E14 chicken λ -ZAPII cDNA library, Hirdeypal Bhatthal for kindly providing us with a chicken genomic-FIXII library, and Paul Patterson for critically reading the manuscript.

Jost Vielmetter was supported in part by a fellowship from Deutsche For-

schungs Gemeinschaft (Germany). This work was also supported by the National Institutes of Health grant EY07725.

Received for publication 15 March 1994 and in revised form 19 September 1994.

References

- Aaronson, S. A. 1991. Growth factors and cancer. *Science (Wash. DC)*. 254:1146-1153.
- Altschul, S. F., W. Gish, W. Miller, E. W. Myers, and D. J. Lipman. 1990. Basic local alignment search tool. *J. Mol. Biol.* 215:403-410.
- Ansorge, W. 1985. Fast and sensitive detection of protein and DNA bands by treatment with potassium permanganate. *J. Biochem. Biophys. Methods*. 11:13-20.
- Barthels, D., G. Vopper, and W. Wille. 1988. NCAM-180, the large isoform of the neural cell-adhesion molecule of the mouse, is encoded by an alternatively spliced transcript. *Nucleic Acids Res.* 16:4217-4225.
- Bieber, A. J., P. M. Snow, M. Hortsch, N. H. Patel, J. R. Jacobs, Z. R. Traquina, J. Schilling, and C. S. Goodman. 1989. Drosophila neuroglian: a member of the immunoglobulin superfamily with extensive homology to the vertebrate neural adhesion molecule L1. *Cell*. 59:447-460.
- Brummendorf, T., and F. G. Rathjen. 1993. Axonal glycoproteins with immunoglobulin and fibronectin Type-III-Related domains in vertebrates—structural features, binding activities, and signal transduction. *J. Neurochem.* 61:1207-1219.
- Brummendorf, T., J. M. Wolff, R. Frank, and F. G. Rathjen. 1989. Neural cell recognition molecule F11—homology with fibronectin type-III and immunoglobulin type-C domains. *Neuron*. 2:1351-1361.
- Burgoon, M. P., M. Grumet, V. Mauro, G. M. Edelman, and B. A. Cunningham. 1991. Structure of the chicken neuron-glia cell adhesion molecule, Ng-CAM—origin of the polypeptides and relation to the Ig superfamily. *J. Cell Biol.* 112:1017-1029.
- Christianson, D. W. 1991. Structural biology of zinc. *Adv. Protein Chem.* 42:281-355.
- Cunningham, B. A., J. J. Hemperly, B. A. Murray, E. A. Prediger, R. Brackenbury, and G. M. Edelman. 1987. Neural cell-adhesion molecule—structure, immunoglobulin-like domains, cell-surface modulation, and alternative RNA splicing. *Science (Wash. DC)*. 236:799-806.
- Davis, J. Q., T. McLaughlin, and V. Bennett. 1993. Ankyrin-binding proteins related to nervous system cell adhesion molecules: candidates to provide transmembrane and intercellular connections in adult brain. *J. Cell Biol.* 121:121-133.
- de la Rosa, E. J., J. F. Kayyem, J. M. Roman, Y. D. Stierhof, W. J. Dreyer, and U. Schwarz. 1990. Topologically restricted appearance in the developing chick retinotectal system of Bravo, a neural surface protein: experimental modulation by environmental cues (published erratum appears in 1991. *J. Cell Biol.* 112:1049.) *J. Cell Biol.* 111:3087-3096.
- Devereux, J., P. Haeblerli, and O. Smithies. 1984. A comprehensive set of sequence-analysis programs for the VAX. *Nucleic Acids Res.* 12:87-395.
- Dickson, G., H. J. Gower, C. H. Barton, H. M. Prentice, and V. L. Elsom. 1987. Human-muscle neural cell-adhesion molecule (N-CAM): identification of a muscle-specific sequence in the extracellular domain. *Cell*. 50:1119-1130.
- Djabali, M., M. G. Mattei, C. Ngyen, D. Roux, and J. Demengeot. 1990. The gene encoding L1, a neural adhesion molecule of the immunoglobulin family, is located on the X-chromosome in mouse and man. *Genomics*. 7:587-593.
- Dumas, J. B., M. Edwards, J. Delort, and J. Mallet. 1991. Oligodeoxyribonucleotide ligation to single-stranded cDNAs: a new tool for cloning 5' ends of mRNAs and for constructing cDNA libraries by in vitro amplification. *Nucleic Acids Res.* 19:5227-5232.
- Edelman, G. M., and K. L. Crossin. 1991. Cell adhesion molecules: implications for a molecular histology. *Annu. Rev. Biochem.* 60:155-190.
- Engel, J. 1991. Common structural motifs in proteins of the extracellular matrix. *Curr. Opin. Cell Biol.* 3:779-785.
- Fearon, E. R., K. R. Cho, J. M. Nigro, S. E. Kern, J. W. Simons, J. M. Ruppert, S. R. Hamilton, A. C. Preisinger, G. Thomas, K. W. Kinzler, et al. 1990. Identification of a chromosome 18q gene that is altered in colorectal cancers. *Science (Wash. DC)*. 247:49-56.
- Feinberg, A. P., and B. Vogelstein. 1983. A technique for radiolabeling DNA restriction endonuclease fragments to high specific activity. *Anal. Biochem.* 132:6-13.
- Gennarini, G., G. Cibelli, G. Rougon, M. G. Mattei, and C. Goridis. 1989. The mouse neuronal cell surface protein F3: a phosphatidylinositol-anchored member of the immunoglobulin superfamily related to chicken contactin. *J. Cell Biol.* 109:775-788.
- Goldberg, S., and A. J. Coulombre. 1972. Topographical development of the ganglion cell fiber layer in the chick retina. A whole mount study. *J. Comp. Neurol.* 146:507-518.
- Goodman, C. S., and C. J. Shatz. 1993. Developmental mechanisms that generate precise patterns of neuronal connectivity. *Cell*. 72 (Suppl.):77-98.
- Grenningloh, G., E. J. Rehm, and C. S. Goodman. 1991. Genetic analysis of growth cone guidance in Drosophila—Fasciclin-II functions as a neuronal recognition molecule. *Cell*. 67:45-57.
- Grumet, M., V. Mauro, M. P. Burgoon, G. M. Edelman, and B. A. Cunningham. 1991. Structure of a new nervous system glycoprotein, Nr-CAM, and its relationship to subgroups of neural cell adhesion molecules. *J. Cell Biol.* 113:1399-1412.
- Harrelson, A. L., and C. S. Goodman. 1988. Growth cone guidance in insects: fasciclin-II is a member of the immunoglobulin superfamily. *Science (Wash. DC)*. 242:700-708.
- Hasler, T. H., C. Rader, E. T. Stoeckli, R. A. Zuellig, and P. Sonderegger. 1993. cDNA cloning, structural features, and eucaryotic expression of human TAG-1/Axonin-1. *Eur. J. Biochem.* 211:29-339.
- Hedrick, L., K. R. Cho, J. Boyd, J. Risinger, and B. Vogelstein. 1992. DCC: a tumor suppressor gene expressed on the cell surface. *Cold Spring Harbor Symp. Quant. Biol.* 57:45-51.
- Higgins, D. G., and P. M. Sharp. 1989. Fast and sensitive multiple sequence alignments on a microcomputer. *Comp. Appl. Biosci.* 5:151-153.
- Jessell, T. M. 1988. Adhesion molecules and the hierarchy of neural development. *Neuron*. 1:3-13.
- Kaprielian, Z., and P. H. Patterson. 1994. The molecular basis of retinotectal topography. *Bioessays*. 16:1-11.
- Kayyem, J. F., J. M. Roman, E. J. Delarosa, U. Schwarz, and W. J. Dreyer. 1992a. Bravo/Nr-CAM is closely related to the cell adhesion molecules L1 and Ng-CAM and has a similar heterodimer structure. *J. Cell Biol.* 118:1259-1270.
- Kayyem, J. F., J. M. Roman, Y. Vonboberg, U. Schwarz, and W. J. Dreyer. 1992b. A method for the generation of monoclonal antibodies against rare cell-surface molecules. *Eur. J. Biochem.* 208:1-8.
- Kyte, J., and R. F. Doolittle. 1982. A simple method for displaying the hydrophobic character of a protein. *J. Mol. Biol.* 157:105-132.
- Laemmli, U. K. 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature (Lond.)*. 227:680-685.
- Lawlor, K. G., and R. Narayanan. 1992. Persistent expression of the tumor suppressor gene DCC is essential for neuronal differentiation. *Cell Growth & Differ.* 3:609-616.
- Ledouarin, N. M. 1993. Embryonic neural chimaeras in the study of brain development. *Trends Neurosci.* 16:64-72.
- Lipkin, V. M., and N. V. Khramtsov. 1989. Calmodulin-independent bovine brain adenylate-cyclase—amino-acid sequence and nucleotide-sequence of the corresponding cDNA. *FEBS (Fed. Eur. Biochem. Soc.) Lett.* 254 (1-2):69-73.
- Marshall, R. D. 1972. Glycoproteins. *Annu. Rev. Biochem.* 41:673-702.
- Matsudaira, P. 1987. Sequence from picomole quantities of proteins electroblotted onto polyvinylidene difluoride membranes. *J. Biol. Chem.* 262:10035-10038.
- Mayford, M., A. Barzilai, F. Keller, S. Schacher, and E. R. Kandel. 1992. Modulation of an NCAM-Related adhesion molecule with long-term synaptic plasticity in Aplysia. *Science (Wash. DC)*. 256:638-644.
- Mey, J., and S. Thanos. 1992. Development of the visual system of the chick—a review. *J. Hirnforschung*. 33:673-702.
- Mishell, B. B., and S. M. Shiigi. 1980. *Selected Methods in Cellular Immunology*. Freeman and Co., New York. 486 pages.
- Miura, M., M. Kobayashi, H. Asou, and K. Uyemura. 1991. Molecular cloning of cDNA encoding the rat neural cell adhesion molecule L1: two L1 isoforms in the cytoplasmic region are produced by differential splicing. *FEBS (Fed. Eur. Biochem. Soc.) Letts.* 289:91-95.
- Moos, M., R. Tacke, H. Scherer, D. Teplow, K. Fruh, and M. Schachner. 1988. Neural adhesion molecule L1 as a member of the immunoglobulin superfamily with binding domains similar to fibronectin. *Nature (Lond.)*. 334:701-703.
- Needleman, S. B., and C. D. Wunsch. 1970. A general method applicable to the search for similarities in the amino acid sequence of two proteins. *J. Mol. Biol.* 48:443-453.
- Patterson, P. H. 1988. On the importance of being inhibited, or saying no to growth cones. *Neuron*. 1:263-267.
- Proudfoot, N. J., and G. G. Brownlee. 1976. 3' non-coding region sequences in eukaryotic messenger RNA. *Nature (Lond.)*. 263:211-214.
- Ranscht, B., and M. T. Dours. 1988. Sequence of contactin, a 130-kD glycoprotein concentrated in areas of interneuronal contact, defines a new member of the immunoglobulin supergene family in the nervous system. *J. Cell Biol.* 107:1561-1573.
- Rathjen, F. G., J. M. Wolff, R. Frank, F. Bonhoeffer, and U. Rutishauser. 1987a. Membrane glycoproteins involved in neurite fasciculation. *J. Cell Biol.* 104:43-53.
- Rathjen, F. G., J. M. Wolff, S. Chang, F. Bonhoeffer, and J. A. Raper. 1987b. Neurofascin: a novel chick cell-surface glycoprotein involved in neurite-neurite interactions. *Cell*. 51:841-849.
- Rathjen, F. G., U. Norenberg, and H. Volkmer. 1992. Glycoproteins implicated in neural cell adhesion and axonal growth. *Biochemical Soc. Trans.* 20:405-409.
- Sambrook, J., E. F. Fritsch, and T. Maniatis. 1989. *Molecular Cloning, a Laboratory Manual*. Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.
- Sanger, F., S. Nicklen, and A. R. Coulson. 1977. DNA sequencing with chain-terminating inhibitors. *Proc. Natl. Acad. Sci. USA*. 74:5463-5467.
- Solomon, E., J. Borrow, and A. D. Goddard. 1991. Chromosome aberrations

- and cancer. *Science (Wash. DC)*. 254:1153-1160.
- Sonderegger, P., and F. G. Rathjen. 1992. Regulation of axonal growth in the vertebrate nervous system by interactions between glycoproteins belonging to 2 subgroups of the immunoglobulin superfamily. *J. Cell Biol.* 119:1387-1394.
- Stoeckli, E. T., T. B. Kuhn, C. O. Duc, M. A. Ruegg, and P. Sonderegger. 1991. The axonally secreted protein Axonin-1 is a potent substratum for neurite growth. *J. Cell Biol.* 112:449-455.
- Taggart, R. T., and I. M. Samloff. 1983. Stable antibody-producing murine hybridomas. *Science (Wash. DC)*. 219:1228-1230.
- Thiery, J. P., A. Delouvec, M. Grumet, and G. M. Edelman. 1985. Initial appearance and regional distribution of the neuron-glia cell adhesion molecule in the chick embryo. *J. Cell Biol.* 100:442-456.
- Tonissen, K. F., and P. A. Krieg. 1993. Two neural-cell adhesion molecule(NCAM)-encoding genes in *Xenopus-laevis* are expressed during development and in adult tissues. *Gene*. 127:243-247.
- Towbin, H., T. Staehelin, and J. Gordon. 1979. Electrophoretic transfer of proteins from polyacrylamide gels to nitrocellulose sheets: procedure and some applications. *Proc. Natl. Acad. Sci. USA*. 76:4350-4354.
- Volkmer, H., B. Hassel, J. M. Wolff, R. Frank, and F. G. Rathjen. 1992. Structure of the axonal surface recognition molecule neurofascin and its relationship to a neural subgroup of the immunoglobulin superfamily. *J. Cell Biol.* 118:149-161.
- Weinberg, R. 1993. Tumor suppressor genes. *Neuron*. 11:191-196.
- Weinberg, R. A. 1991. Tumor suppressor genes. *Science (Wash. DC)*. 254:1138-1146.
- Williams, A. F., and A. N. Barclay. 1988. The immunoglobulin superfamily—domains for cell surface recognition. *Annu. Rev. Immunol.* 6:381-405.
- Woodgett, J. R., K. L. Gould, and T. Hunter. 1986. Substrate-specificity of protein kinase C: use of synthetic peptides corresponding to physiological sites as probes for substrate recognition requirements. *Eur. J. Biochem.* 161:177-184.
- Zuellig, R. A., C. Rader, A. Schroeder, M. B. Kalousek, F. Vonbohlen, T. Osterwalder, C. Inan, E. T. Stoeckli, F. Halbach, H. U. Affolter, A. Fritz, E. Hafen, and P. Sonderegger. 1992. The axonally secreted cell adhesion molecule, axonin-1: primary structure, immunoglobulin-like and fibronectin-type-III-like domains and glycosyl-phosphatidylinositol anchorage. *Eur. J. Biochem.* 204:453-463.