Published March 1, 1991

Structure of the Chicken Neuron–Glia Cell Adhesion Molecule, Ng-CAM: Origin of the Polypeptides and Relation to the Ig Superfamily

Mark P. Burgoon, Martin Grumet, Vincent Mauro, Gerald M. Edelman, and Bruce A. Cunningham

The Rockefeller University, New York, 10021

Abstract. The neuron–glia cell adhesion molecule (Ng-CAM) mediates both neuron–neuron and neuron–glia adhesion; it is detected on SDS-PAGE as a predominant 135-kD glycoprotein, with minor components of 80, 190, and 210 kD. We have isolated cDNA clones encoding the entire sequence of chicken Ng-CAM. The predicted extracellular region includes six immunoglobulin-like domains followed by five fibronectin-type III repeats, structural features that are characteristic of several neural CAMs of the N-CAM superfamily. The amino acid sequence of chicken Ng-CAM is most similar to that of mouse L1 but the overall identity is only 40% and Ng-CAM contains a short fibronectin-like segment with an RGD sequence that has no counterpart in L1. These findings suggest that Ng-CAM and L1 may not be equivalent molecules in chicken and mouse.

The amino-terminal sequences of the 210-, 190-, and 135-kD components of Ng-CAM are all the same as the predicted amino terminus of the molecule, whereas the 80-kD component begins within the third fibronectin repeat. The cDNA sequence is continuous across the junction between the 135- and 80-kD components, and a single 170-kD Ng-CAM polypeptide was isolated from tunicamycin-treated cells. In addition, all cDNA probes hybridized on Northern blots to a 6-kb RNA, and most hybridized to single bands on Southern blots. These results indicate that the Ng-CAM components are derived from a single polypeptide encoded by a single gene, and that the 135- and 80-kD components are generated from the 210/190-kD species by proteolytic cleavage. The 135-kD component contains most of the extracellular region including all of the immunoglobulin-like domains. It has no transmembrane segment, but it is tightly associated with the membrane. The 80-kD component contains two and a half type III repeats plus the RGD-containing segment, as well as the single transmembrane and cytoplasmic domains. These structural features of Ng-CAM provide a framework for understanding its multiple functions in neuron–neuron interactions, neurite fasciculation, and neuron–glia interactions.

CELL adhesion molecules (CAMs)[1] mediate a variety of cell–cell interactions that are important both during embryogenesis and in the maintenance of adult tissues (18). These cell surface glycoproteins are of special interest in the nervous system because they are critical in establishing the basic neuroanatomy of an organism, and because their action, coupled with activity-driven events, influences the organization of functional neural networks. CAMs such as the neuron-glia CAM (Ng-CAM; 32, 34) are excellent candidates for such functions because they are expressed early in the development of the nervous system in spatially and temporally restricted patterns.

Two major families of CAMs are known and both are different from cell surface molecules that serve as receptors (e.g., integrins; 47) for cell substrate adhesion molecules. One is represented by the neural-CAM (N-CAM) (13) and mediates calcium-independent adhesion. The other is a group of calcium-dependent CAMs or cadherins (88) exemplified by the liver CAM (L-CAM) (28). Members of both families are found in the nervous system, although to date the N-CAM family is much larger. The distinguishing structural feature of N-CAM (13, 42) is the presence of five consecutive segments similar to each other and to the individual domains of Ig. This initial demonstration of Ig-like segments in a CAM prompted the suggestion (17) that the N-CAM precursor, consisting of a single or even half an Ig domain, was the evolutionary precursor of the entire Ig superfamily; this notion is supported by the identification of N-CAM-like molecules in insects (31, 83). Subsequently other CAMs as well as other cell surface proteins have been shown to have the C2-type of Ig domains (95) that resemble those in N-CAM. In addition, N-CAM and many of the related molecules contain segments that resemble the type III repeating unit of fibronectin, an extracellular matrix adhesion molecule (50, 82).

Recently, several members of the N-CAM family have been identified in the nervous system and characterized in detail. These include the myelin-associated glycoprotein (1,

1. *Abbreviations used in this paper*: CAM, cell adhesion molecule; L-CAM, liver CAM; N-CAM, neural CAM; Ng-CAM, neuron-glia CAM; NILE, NGF-inducible large external; PCR, polymerase chain reaction.

56, 79), and P_0 glycoprotein (57) which contain only Ig-like segments, and chicken contactin/F11 (and its murine homologue F3) (7, 30, 69), mouse L1 (63), rat TAG-1 (27), and the insect molecules fasciclin II (38) and neuroglian (5), all of which contain multiple Ig-like domains and fibronectin type III repeats. Earlier data (33, 35) suggested that Ng-CAM might also be a member of this family, and this notion is confirmed by the results presented here.

Ng-CAM and N-CAM appear throughout the development of the nervous system, but they are distinct molecules as shown by a variety of structural, immunological, and functional criteria. N-CAM is seen on some of the earliest embryonic cells and in a variety of nonneural tissues, including derivatives of all three germ layers (11, 89). In contrast, Ng-CAM first appears during neural development and its expression is restricted to postmitotic neurons and Schwann cells (15, 35, 90). It is involved in both neuron–neuron and neuron–glia adhesion, as well as the fasciculation of neurites, and the migration of neurons along Bergmann glial fibers during cerebellar development (25, 34, 35, 43). Ng-CAM on one neuron binds homophilically to Ng-CAM on another neuron; the molecule can also bind heterophilically to an as yet unidentified ligand(s) on glia (33). The distribution of Ng-CAM on neuronal cell surfaces appears polarized as development progresses, becoming more prevalent on outgrowing axons than on cell bodies and dendrites (90).

Mouse L1 and the NGF-inducible large external glycoprotein (NILE) (78) in rat, which are very similar to each other (6, 77), are also similar to Ng-CAM in that all three molecules have very similar anatomical distributions during development (4, 70, 90) and some antibodies cross-react with each. For example, antibodies raised against chicken Ng-CAM immunoprecipitated the NILE glycoprotein from extracts of rat PC12 pheochromocytoma cells and were used to show that expression of this molecule is enhanced by NGF (25).

Despite their similarities, there are differences in the structure and activities of Ng-CAM and the rodent molecules. The most obvious structural difference is that L1 and NILE are seen on SDS-PAGE primarily as glycoproteins of ~200 kD (22, 78). Ng-CAM is detected as a predominant 135-kD glycoprotein with less prevalent components of 80, 190, and 210 kD (35). Components comparable to the 135- and 80-kD Ng-CAM species are seen in L1 preparations in small amounts and can be generated by proteolysis (22). The 135- and 80-kD components of Ng-CAM are antigenically distinct but both are related to the larger forms suggesting that they are derived from them by proteolysis. The 135-kD form of Ng-CAM, however, is significantly more prevalent than the other forms both on the surface of chicken cells and after isolation, leaving open the possibility that the various forms of Ng-CAM are derived from different mRNAs.

To analyze the chemical structure of Ng-CAM, to define the origin of its components, and to determine its relationship to other CAMs in the nervous system, we have isolated and characterized cDNA clones representing the entire coding sequence of Ng-CAM. The deduced amino acid sequence indicates that the 135- and 80-kD components are derived from the same mRNA, and are apparently generated from larger species by proteolytic cleavage. The predicted protein sequence is most similar to that of L1 but there are a surprisingly large number of differences in the two sequences, raising the possibility that they are functionally as well as structurally distinct molecules.

Materials and Methods

Protein

Ng-CAM was purified from 14-d chicken embryo brain membranes by affinity chromatography using monoclonal anti–Ng-CAM antibodies (32, 34). Intact Ng-CAM and V8-protease digests of the molecule were resolved on SDS-PAGE, and either stained with Coomassie blue (9, 55) or transferred to Immobilon (Millipore Continental Water Systems, Bedford, MA), stained with Coomassie blue and sequenced (62). Ng-CAM containing all four components (Fig. 1) was treated with CNBr (94) and the fragments were initially separated by gel filtration on Sephadex G-25. After further separation on a C_3 column by HPLC, fragments that were homogeneous as assessed by SDS-PAGE and silver staining (64) were sequenced. A 40-kD CNBr fragment was further digested with trypsin and the peptides were fractionated by HPLC; two of these peptides were sequenced. In addition, the 80-kD component was isolated by SDS-PAGE on 6% acrylamide gels, the band was cut from the gel and treated with V8 protease (9). The resulting peptides were separated by SDS-PAGE on 15% acrylamide gels, transferred to Immobilon and sequenced. Amino-terminal sequencing of various intact and fragmented forms of Ng-CAM was performed at the Rockefeller University Protein Sequencing Facility by automated Edman degradation.

Immunoblots of protein samples were performed after resolution on SDS-PAGE and transferred to nitrocellulose (91). Anti–Ng-CAM antibodies were isolated from rabbits after immunization with boiled Ng-CAM (34).

Biosynthetic labeling of Ng-CAM was performed for 16 h in suspension cultures of 9-d chicken embryo brains. Twenty brains were incubated in ^{3}H-fucose (0.15 mCi) (New England Nuclear, Boston, MA) at 37°C in 50 ml DME (Gibco Laboratories, Grand Island, NY). The tissue was harvested by sedimentation at 4°C, washed twice with PBS, extracted with PBS/0.5% NP40/200 U/ml Trasylol (Mobay Chemical Corp., Pittsburgh, PA) and insoluble material removed by centrifugation at 100,000 *g* for 30 min. Labeled protein from the extract was immunoprecipitated with anti–Ng-CAM mAb 16F5 (32), resolved on 6% SDS-PAGE and detected by autoradiography. Tissue labeling with ^{3}H-leucine (2 mCi) (New England Nuclear) was performed in leucine-free DME and treatments with 1 μg/ml of tunicamycin were begun 6 h before addition of the label to inhibit N-linked carbohydrate processing (45).

DNA

cDNA libraries were constructed in λgt11 from total RNA or poly $(A)^+$ RNA isolated from 9- to 14-d embryonic chicken brains. cDNA was synthesized by the RNAse H method (37) using oligo (dT) or synthetic oligonucleotides as primers. After methylation with Eco RI methylase and S-adenosyl methionine, the cDNA was ligated to Eco RI linkers, and then ligated to Eco RI-digested λgt11 DNA (61, 97) and packaged using Gigapack Plus (Stratagene, La Jolla, CA). 16 cDNA libraries were constructed in λgt11 from embryonic chick brain RNA using oligo dT as primers or synthetic oligonucleotides generated from Ng-CAM sequence. Several libraries were prepared from RNA that had been extensively denatured by incubation at 65°C and treated with Actinomycin D (51) and methyl mercuric hydroxide (61). Three genomic libraries were also constructed. Chicken genomic DNA (Clontech, Palo Alto, CA) was partially digested with Sau3A1 or Bam HI, fragments of 10–25 kb were size selected on 10–40% sucrose gradients, ligated into the Bam HI sites of EMBL3 DNA (26, 61), and packaged using Gigapack Plus (Stratagene). Three chicken λgt11 cDNA libraries prepared from adult brain, embryonic cerebellum, and total embryo RNA were purchased from Clontech. A λgt10 cDNA library constructed from chicken embryo brain RNA was kindly provided by Drs. Joan Levi and Hidesabura Hanafusa (Rockefeller University). Antibody screening of the libraries was performed as described (96), using polyclonal antibodies against denatured Ng-CAM protein which recognize the 210-, 190-, 135-, and 80-kD components of the molecule. Positive clones were isolated to homogeneity, and the inserts were excised from λgt11 arms by restriction with Eco RI endonuclease.

cDNA inserts were labeled with ^{32}P-dCTP using the oligolabeling protocol of Feinberg and Vogelstein (23), and used to screen λgt11 libraries to obtain overlapping cDNA clones.

For sequence analysis cDNA inserts were subcloned into M13mp18 and

Published March 1, 1991

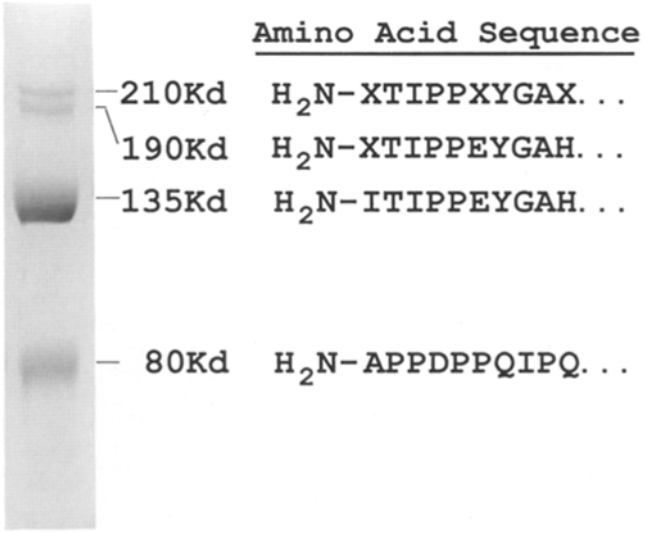


Figure 1. Components of Ng-CAM resolved on SDS-PAGE. Ng-CAM from embryonic chick brain was fractionated on 6% acrylamide gels and stained with Coomassie blue. The amino-terminal sequence of each component (210, 190, 135, 80 kD) was determined after transfer to Immobilon.

M13mp19 vectors (Bethesda Research Laboratories, Gaithersburg, MD) and sequenced by the dideoxynucleotide chain-termination method using Sequenase (United States Biochemical Corp., Cleveland, OH) (80). The sequence of larger inserts was obtained after deletion cloning in M13 vectors, prepared by one of two methods: digestion with restriction endonucleases which recognized sites in the M13 polylinker and within the insert; or with the M13 cyclone cloning system (International Biotechnologies, New Haven, CT), which is based on the rapid deletion cloning method of Dale et al. (14).

The polymerase chain reaction (PCR) was used to synthesize cDNA clones between pairs of sense and antisense oligonucleotides synthesized using protein or DNA sequences in Ng-CAM. To isolate the initial Ng-CAM clone, two degenerate oligonucleotides synthesized on the basis of the amino acid sequence of a 35-residue segment of a CNBr fragment were used to generate a PCR product of 105 nucleotides (Figs. 2 and 3). The subcloned product encoded the appropriate amino acid sequence and it was therefore used to synthesize a 57-base oligonucleotide that was used to screen a λgt10 cDNA library. PCR techniques used Taq Polymerase and Perkins Elmer Corp. (Norwalk, CT) reagent kits in a programmable thermal cycler (USA/Scientific Plastics, Inc., Ocala, FL) (76). For example, cDNA templates were denatured at 94°C for 1 min, annealed with synthetic oligonucleotide primers at 55°C for 1 min, and then DNA was synthesized with Taq Polymerase at 72°C for 1 min. The cycle was repeated 20 times to amplify the DNA segment between the oligonucleotide primers. Oligonucleotides used for PCR, priming cDNA synthesis, and priming sequencing reactions, were synthesized at the Rockefeller University Protein Sequencing Facility.

RNA, Northern Blots, and Southern Blots

Total RNA was prepared from embryonic chicken brains by rapid sonication of freshly dissected tissue in a Polytron homogenizer either in LiCl/Urea (29) or in 4 M guanadinium-thiocyanate (60) followed by centrifugation through a 5.7 M CsCl cushion (8). Poly $(A)^+$ RNA was isolated on oligo (dT) columns (Pharmacia Fine Chemicals, Piscataway, NJ). RNA transfer blots and DNA transfer blots were performed using standard techniques (61). All restriction enzymes were purchased from Bethesda Research Laboratories. For genomic Southern analyses, chicken liver genomic DNA (Clontech), was digested to completion with appropriate restriction endonucleases and resolved on individual lanes of 0.7% agarose gels (61). Nucleic acids were transferred to Hybond (Amersham Corp., Arlington Heights, IL) or Genescreen (New England Nuclear), the membranes were fixed, and hybridized to ^{32}P-labeled cDNA probes (61).

Analysis of Protein Sequences

Sequence data were compiled using the Staden ANALSEQ programs (85). The Dayhoff protein sequence database (National Biomedical Research Foundation, Washington, DC) and the translated Genbank database (release 63) were searched using the rapid homology search program LFASTA (Ktup = 2; 65). The LFASTA program was also used in the pairwise alignment of the internal tandem repeats in Ng-CAM and L1 and in comparisons to other neural CAMs. Hydrophobicity analyses of the sequence were performed using an automated hydropathy program (21) using parameters given by Kyte and Doolittle with a sliding window of 19 residues (54).

Results

The amino terminal sequence of each of the Ng-CAM polypeptides (Fig. 1) was determined after resolving the components on SDS-PAGE and transferring them to Immobilon. The sequences of the 210-, 190-, and 135-kD components were identical to each other, but differed significantly from that of the 80-kD component. To obtain additional sequences, Ng-CAM and the 80-kD component were treated with CNBr, *Staphylococcus aureas* V8 protease, or trypsin, and the peptides purified by gel filtration and HPLC or by SDS-PAGE. The amino terminal sequences of these peptides were all found in the complete sequence derived from cDNA clones (Fig. 2) and are shown in Fig. 3.

Isolation and Sequencing of cDNA Clones

Two strategies were used to obtain initial Ng-CAM cDNA clones. In the first, the PCR, using primers (underlining arrows in Fig. 3) based on the amino acid sequence of a peptide, was used to obtain an authentic Ng-CAM nucleic acid probe; this approach gave cDNA clone λN903. In the second, polyclonal antibodies specific for Ng-CAM were used to isolate directly from expression libraries clones λN902 and λN925 (Fig. 2). The insert in λN902 encodes 81 amino acids near the 3′ end of the coding sequence (1061–1141; see Fig. 3). The insert in clone λN925 encodes 113 amino acids (85–197) and overlaps with the sequence of λN903 (Fig. 2). The amino acid sequence encoded by λN925 included segments that were identical to those of two different CNBr fragments obtained from Ng-CAM (see Fig. 3), verifying that it and λN903 encode segments of Ng-CAM.

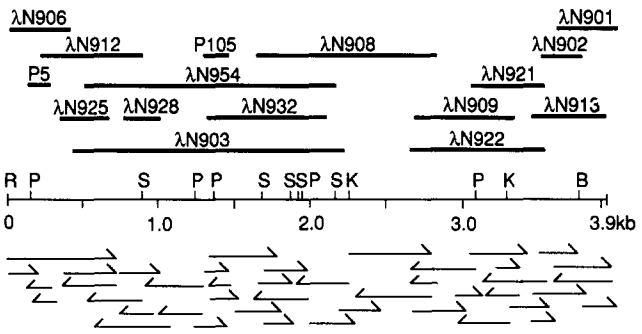


Figure 2. Schematic representation of Ng-CAM cDNA clones and the sequencing strategy. The 14 cDNA clones and two PCR products used to determine the sequence of Ng-CAM are indicated (to scale) in a 5′ to 3′ direction. Restriction sites are: (*R*) Eco RI; (*P*) Pst I; (*S*) Sma I; (*K*) Kpn I; (*B*) Bam HI. Direction and extent of sequencing are indicated by arrows below the scale bar.

Published March 1, 1991

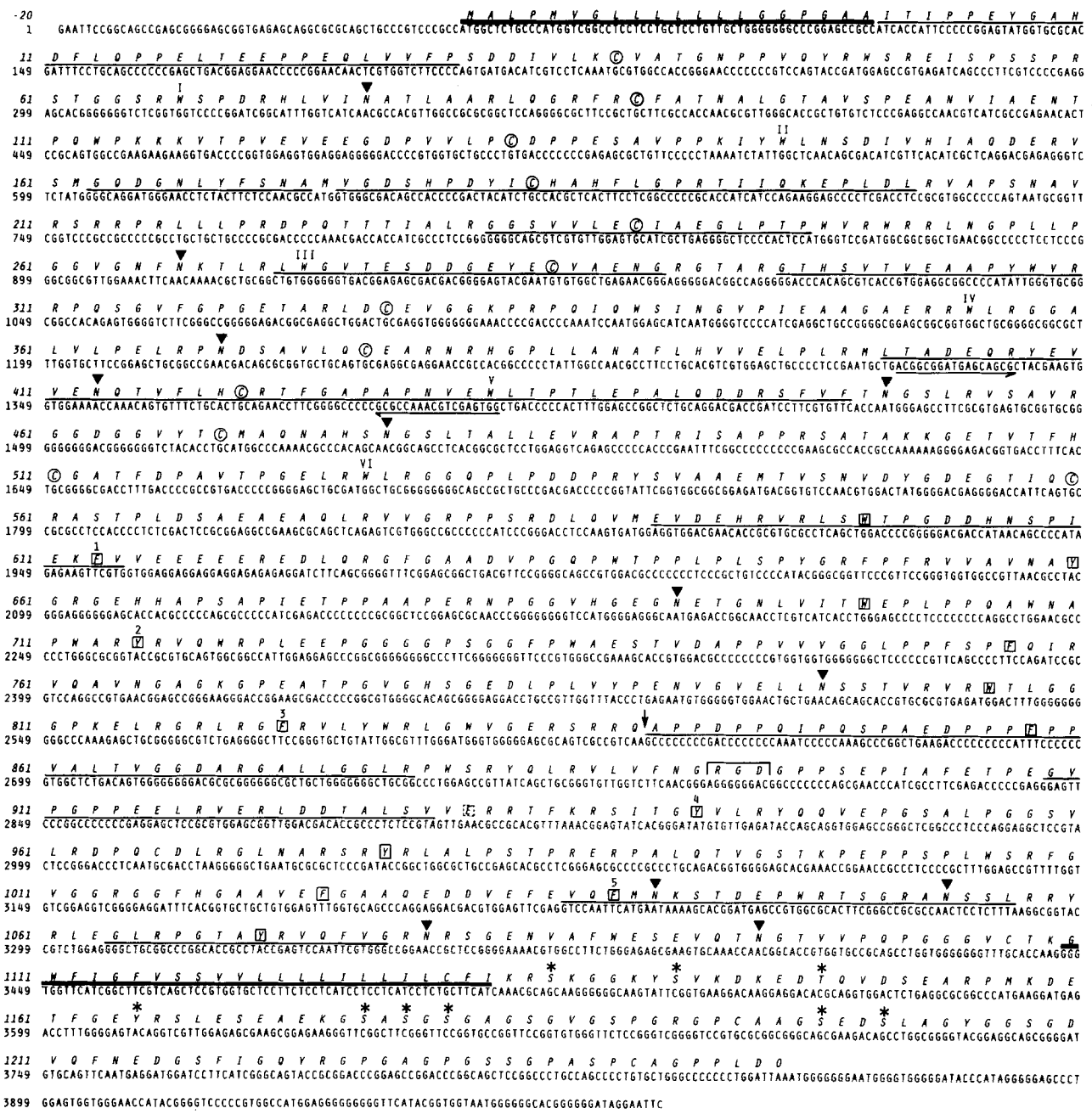


Figure 3. Nucleotide sequence and deduced amino acid sequence of Ng-CAM. The longest open reading frame contains 1,265 amino acids terminating at an ocher termination codon (*O*). The two hydrophobic regions representing the signal peptide (−20–−1) and the transmembrane region (1,110–1,132) are underlined by a thick bar. Thinner underlining denotes amino acid sequences determined by protein chemical analyses of the 210-, 190-, 135-, and 80-kD components of Ng-CAM and peptides obtained after treatment with CNBr, V8 protease, or trypsin. Underlining arrows represent oligonucleotide primers used to create PCR product P105. Potential sites of asparagine-linked glycosylation (45) are marked with inverted triangles. Potential casein-type III kinase phosphorylation sites are indicated by asterisks (*). The predicted proteolytic cleavage site between the 135- and 80-kD components is indicated by a vertical arrow. The immunoglobulin-like domains are numbered from I to VI over the conserved tryptophan with the characteristic cysteines indicated by circles. The fibronectin-type III repeating units are numbered 1 to 5 with the characteristic tryptophan and tyrosine (or phenylalanine) residues boxed. A bracket highlights the RGD sequence. These sequence data are available from EMBL/Genbank/DDBJ under accession number X56969.

To obtain clones that encoded the entire Ng-CAM molecule, restriction fragments of λN902, 903, and 925 were used to screen different libraries, including new libraries generated with oligonucleotides corresponding to specific Ng-CAM sequences. Both strands of the inserts from 14 different cDNA clones (Fig. 2) were sequenced to provide the DNA sequence that encodes the entire protein. The sequence between clones λN903 and λN922 is especially rich in poly-G and -C and more than 30 × 10^6 pfu from 23 cDNA and genomic libraries were screened without detecting clones coding for this region. The one clone, λN908 (Fig. 2), that extended across this region (nucleotide

Published March 1, 1991

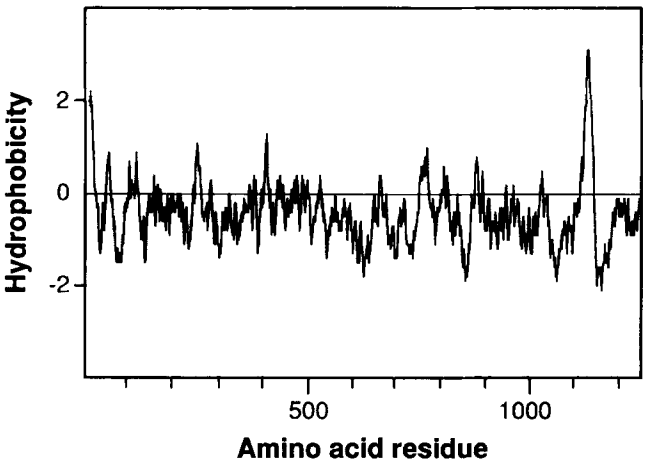


Figure 4. Hydrophobicity plot of the predicted amino acid sequence of Ng-CAM. Amino acid residues are numbered from the translation start site. Positive values indicate hydrophobicity.

1701–2766, Fig. 3), was isolated from a library prepared from poly $(A)^+$ RNA that had been extensively denatured and primed with an oligonucleotide specific for Ng-CAM.

Amino Acid Sequence

The cDNA sequence and the deduced protein sequence are summarized in Fig. 3. Specific structural features are highlighted in comparison to mouse L1 in Fig. 7 and a proposed model is shown in Fig. 8.

The combined data gave a continuous sequence of 3989 nucleotides, with an open reading frame of 1265 amino acids; other reading frames contained multiple stop codons throughout the sequence. The predicted sequence includes all sequences obtained by direct analysis of each of the polypeptide components and of peptides derived from Ng-CAM and the 80-kD component. These sequences (Fig. 3) comprise >10% of the deduced protein sequence and are distributed throughout the extracellular portion of the molecule, confirming the reading frame.

Hydrophobicity analysis (Fig. 4) indicated that there are two nonpolar domains in Ng-CAM. The first (amino acids −20 to −1) has all the features of a signal sequence and the most likely cleavage site for a signal peptide (67, 93) is at the end of this segment. The second hydrophobic domain (amino acid 1110–1132) has features that are characteristic of transmembrane segments. Its location predicts that Ng-CAM has an extracellular region of 1109 amino acids containing twelve potential asparagine-linked glycosylation sites and a cytoplasmic domain of 113 amino acids that includes a number of potential phosphorylation sites. Threonine 1148 and serine 1201 are located in regions which are consensus recognition sites for casein kinase I (39, 92). In addition, there are six serines (1135, 1141, 1176–1180, and 1198) that fit the more general description of recognition sites for several kinases (19), and tyrosine 1165 is a potential site for tyrosine kinase (46).

The extracellular region of Ng-CAM includes two motifs of repeated domains which have been found in several neural CAMs. Beginning at the amino terminus are six similar domains that resemble those found in immunoglobulins, N-CAM, and related CAMs (17). All have paired cysteines and surrounding amino acids characteristic of the C-2 subtype of Ig domains (95), the most common subtype found in the neural CAMs. These Ig-like regions are followed by five domains that resemble each other and the type III repeats found in fibronectin (50). Each contains ~100 residues, and the first three are distinguished by their conserved with the first tryptophan and tyrosine residues. The fourth is less well conserved tryptophan replaced by a glutamate; the fifth repeat is the least similar to other fibronectin-like domains with the tryptophan and tyrosine residues both replaced by phenylalanine.

Ng-CAM contains one RGD sequence (amino acids 893–895) at the end of the third type III repeating unit. It is included in a short segment (29 amino acids) that resembles

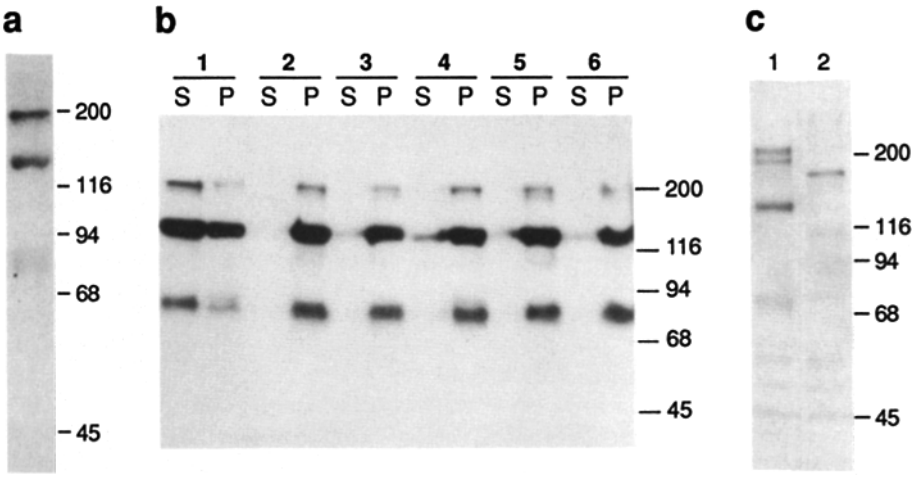


Figure 5. Ng-CAM biosynthesis, glycosylation, and membrane association. Immunoprecipitation with anti-Ng-CAM antibodies of detergent extracts of 10-d chick embryo brains biosynthetically labeled with (*a*) ^{3}H-fucose and (*c*) ^{3}H-leucine in the absence (lane *1*) or presence (lane *2*) of tunicamycin. Proteins were resolved on 6% SDS-PAGE, the gels were dried, and the label was detected by autoradiography. (*b*) 14-d chick embryo brain membranes (150 μl) were treated with 0.5 ml of PBS/0.5% NP40 (lane *1*), PBS/5 mM EDTA (lane *2*), PBS/1 M NaCl (lane *3*), 50 mM diethylamine, pH 11.5 (lane *4*), PBS diluted 10-fold with H_2O (lane *5*), and PBS (lane *6*), and centrifuged in a microfuge for 30 min. The supernatant (*S*) and pellet (*P*) fractions were separated and 2% of each fraction was resolved on SDS-PAGE, transferred to nitrocellulose, and immunoblotted with 100 μg of anti-Ng-CAM antibodies. Molecular weight standards are indicated at the right as $M_r \times 10^{-3}$.

Published March 1, 1991

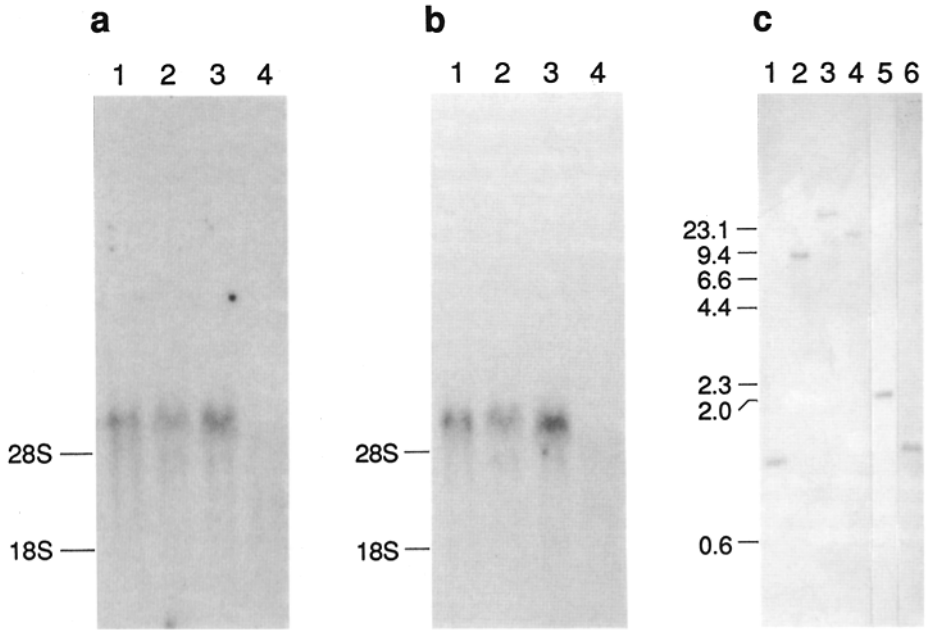


Figure 6. Northern and Southern blots using Ng-CAM cDNA probes. Poly-(A)$^+$ RNA (1 μg) from 12-d chick embryo brain (lane *1*), total brain RNA (15 μg) from 9- (lane *2*) and 12-d (lane *3*) embryos, and 14-d chick embryo liver (lane *4*) were electrophoresed on 0.8% formaldehyde gels transferred to Hybond and hybridized with Ng-CAM cDNA probes N903 (*a*) and N922 (*b*). Positions of 28S and 18S ribosomal RNA are indicated. (*c*) Recognition of genomic DNA by Ng-CAM cDNA clones. 8 μg of adult chicken genomic DNA was restricted with Hga I (lane *1*), Bam HI (lane *2*), Dra I (lane *3*), Ban I (lane *4*), Sma I (lane *5*), and Sst II (lane *6*), electrophoresed on 0.7% agarose gels, transferred to Hybond, and probed with Ng-CAM cDNA N906. Relative migration of molecular weight standards is indicated at the left in kilobases.

the carboxyl terminal portion of a type III repeat (see Fig. 7 *b*) suggesting that DNA specifying this portion of the third repeat may at one time have undergone a duplication.

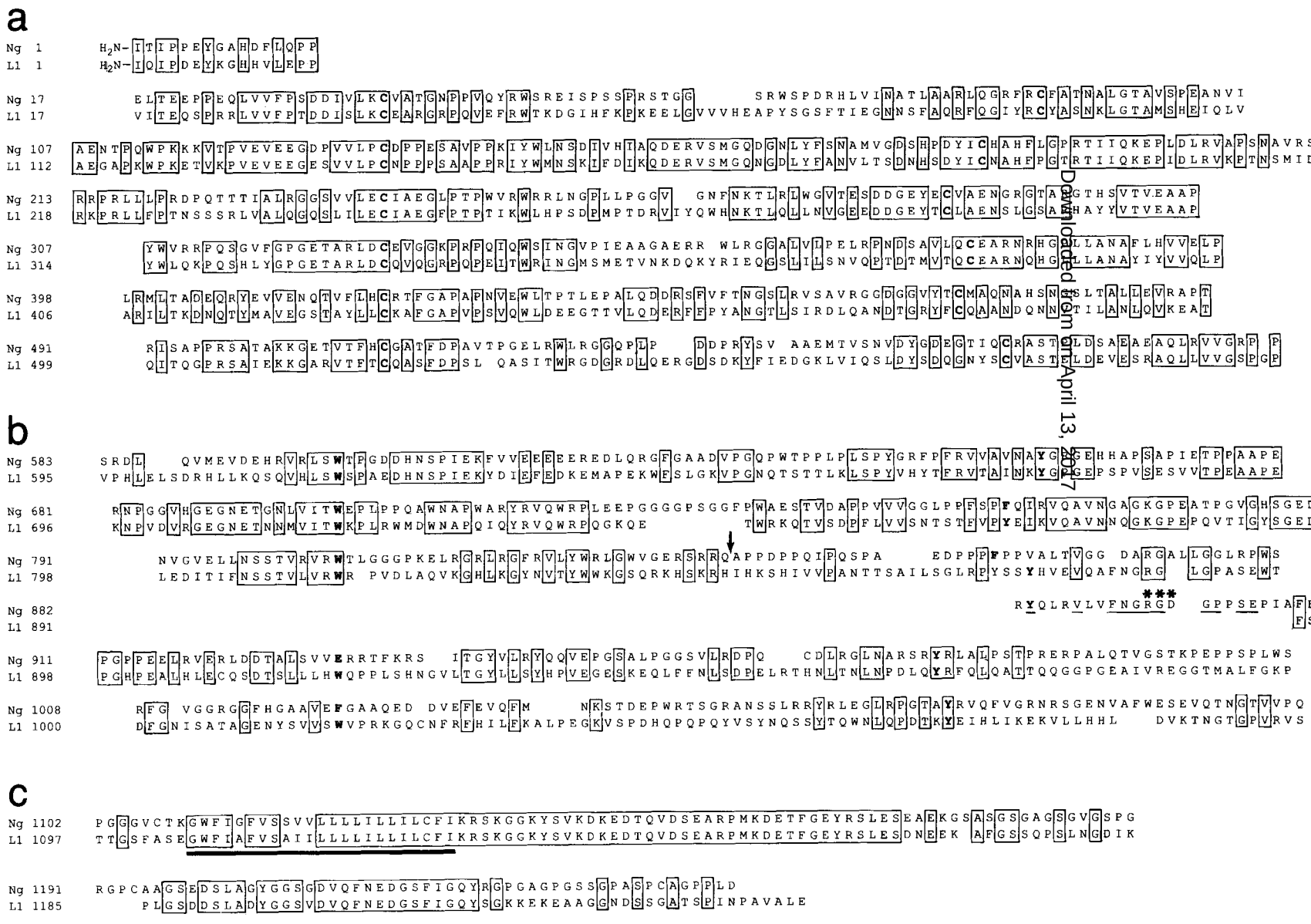
Organization of Ng-CAM Components

The amino-terminal sequences of the 210-, 190-, and 135-kD components (Fig. 1) all correspond to the presumed amino terminus of the protein (Fig. 3). The amino terminus of the 80-kD component is at amino acid 840 in the middle of the third fibronectin type III repeat (Figs. 1 and 3). The carboxyl side of this region is rich in proline residues and the amino side is rich in arginine, suggesting that this could be a highly exposed segment sensitive to proteases such as trypsin. The precise distance, if any, between the carboxyl terminus of the 135-kD component and the amino terminus of the 80-kD component, is as yet unknown. They probably do not overlap; all sequences of CNBr fragments from total Ng-CAM (which is predominantly the 135-kD component) were found in the region amino terminal to where the 80-kD component begins and all peptides isolated from the purified 80-kD component were detected only in the region on the carboxy-terminal side of this site.

The sequence data suggest that the entire 135-kD component is extracellular and includes all six Ig domains plus the first two fibronectin repeats. The 80-kD component contains the last two fibronectin repeats, the transmembrane segment, and the cytoplasmic domain. Both species have extracellular regions with several potential glycosylation sites, eight in the 135-kD component and four in the 80-kD component. In accord with these observations, the 210-, 190-, 135-, and 80-kD forms incorporated ^{3}H-fucose (Fig. 5 *a*). Moreover, endoglycosidase F treatment suggested there are three asparagine-linked oligosaccharides in each of the 135- and 80-kD components (Grumet, M., and G.M. Edelman, unpublished observations). In accord with the notion that the 80 kD has a cytoplasmic domain but the 135-kD segment does not, the 80- but not the 135-kD form can be phosphorylated in cell culture (35). Preliminary experiments indicate that the 200- and 80-kD components of purified Ng-CAM can also be phosphorylated in solution either by itself, or by a kinase that copurifies with it. In these and the previous experiments (35) the band at 200 kD was too broad to discern whether the 210, the 190 kD or both components were phosphorylated. Mouse L1 is similarly phosphorylated, presumably by a kinase which copurifies with it (75).

These data indicate that the 135-kD component does not extend to the transmembrane segment or contain an alternative such region. However, it is tightly associated with brain plasma membranes (Fig. 5 *b*) in that detergents, such as deoxycholate, CHAPS (data not shown), and NP-40 are required (Fig. 5 *b*, lane *1*) to release it. Detergents release all

Figure 7. Comparison of chicken Ng-CAM domains to corresponding domains in mouse L1. The position of the first amino acid residue of each domain is given. Sequences were initially aligned pairwise using the LFASTA program (ktup = 2; 65), and then by visual inspection, to achieve the best overall match. Gaps were introduced to maximize the identities between the sequences. Identical residues between the sequences are boxed. (*a*) Alignment of six consecutive immunoglobulin-like domains in Ng-CAM to corresponding domains in L1. Characteristic cysteine residues likely to be involved in intrachain disulfide bonds are highlighted in bold type, and the six comparisons are aligned through the first cysteines in each pair. (*b*) Alignment of five consecutive fibronectin-type III repeating units in Ng-CAM to corresponding repeats in L1. Characteristic tryptophan and tyrosine residues (or substituted phenylalanine or glutamate) are highlighted in bold type. The five pairs of domains are aligned through the first tryptophans in each. The 29-amino acid segment between the third and fourth fibronectin repeats of Ng-CAM is aligned to show its similarity to the preceding domain. (*c*) Alignment of the transmembrane and cytoplasmic segments in Ng-CAM to corresponding segments in L1.



Published March 1, 1991

Table I. Relationships between Ng-CAM and Other Neural CAMs

	Ng-CAM	L1	Neuroglian	Contactin/F11	F3	TAG-1	Chick N-CAM	Mouse N-CAM	Fasciclin II
Ng-CAM (C)	–								
L1 (M)	**40** (1284)	–							
Neuroglian (I)	23 (1269)	28 (1224)	–						
Contactin/F11 (C)	27 (815)	28 (882)	28 (778)	–					
F3 (M)	27 (873)	29 (886)	27 (781)	**78** (979)	–				
TAG-1 (R)	29 (865)	29 (993)	24 (972)	**49** (1033)	**50** (1009)	–			
N-CAM (C)	22 (600)	21 (706)	21 (710)	19 (676)	23 (680)	19 (730)	–		
N-CAM (M)	21 (584)	20 (712)	22 (713)	22 (551)	23 (670)	21 (707)	**81** (1109)	–	
Fasciclin II (I)	21 (365)	21 (376)	22 (411)	19 (572)	24 (670)	19 (613)	26 (562)	27 (395)	–

Protein sequences for each of the neural CAMs listed were compared in a pairwise fashion using the LFASTA program of Pearson and Lipman (1988; ktup = 1). Scores indicate the percentage identity between the protein sequences across the number of amino acid residues indicated in parentheses. Numbers in boldface highlight the highest similarities between more closely related proteins. Parentheses at left indicate origin of protein. (*C*) chicken; (*M*) mouse; (*I*) insect; (*R*) rat.

forms of Ng-CAM from the membrane, but it was difficult to resolve the 210- and 190-kD components. In contrast, treatments with EDTA (lane *2*), high salt (lane *3*), high pH (lane *4*), and low salt (lane *5*) buffers, which extract peripheral membrane proteins, did not release significant amounts of the 135-kD component or any other Ng-CAM species from the membranes. Although there are several cysteines in the 135- and 80-kD components, treatment of brain plasma membranes with buffers containing dithiothreitol (10 mM) did not release the 135-kD species, or other forms of Ng-CAM. These results suggest that the 135-kD component may be tightly associated noncovalently with the other Ng-CAM components or has another mode of attachment to the membrane.

The fact that the cDNA sequence is continuous and includes sequences from both the 135- and 80-kD components indicates that they are both derived from the larger (210/190 kD) components. In accord with this notion, a single Ng-CAM polypeptide is seen in tunicamycin-treated cells. When brain tissue was incubated in ^{3}H-leucine, Ng-CAM components of 210, 190, 135, and 80 kD were immunoprecipitated with specific anti–Ng-CAM antibodies (Fig. 5 *c*). In the presence of tunicamycin which inhibits N-linked carbohydrate processing (53), however, one predominant component of 170 kD was immunoprecipitated with anti–Ng-CAM antibodies (Fig. 5 *c*). Although the cDNA predicts a single polypeptide of 139 kD, the difference (170 vs. 139 kD) could be because of other posttranslational modifications. Such differences between predicted and observed sizes have been seen for other proteins, including the various forms of N-CAM (13) and probably reflect a variety of factors including the limits of SDS-PAGE in determining molecular weight. In any case, these results suggest that Ng-CAM is synthesized as one polypeptide, which is then glycosylated and further processed.

Ng-CAM mRNA and the Ng-CAM Gene

The conclusion that Ng-CAM is initially synthesized as a single polypeptide is substantiated by Northern and Southern blots. Inserts from cDNA clones λN903 and λN902 were hybridized to RNA isolated from different chicken embryonic tissues. These two probes were chosen because λN902 should encode sequences only in the 80-kD region whereas λN903 encodes sequences only in the 135-kD region. Both λN902 and λN903 hybridized selectively to one band at 6 kb present in poly A$^+$ brain RNA (Fig. 6), suggesting that there is a single mRNA. Total embryonic brain RNA included an Ng-CAM mRNA of the same size, but no RNA from 14-d embryonic chicken liver (Fig. 6), heart, skin, or gizzard reacted with the probes (data not shown). These data are consistent with the model that the 135- and 80-kD components of Ng-CAM are encoded by a single neural-specific mRNA which is large enough to encode the 210/190-kD forms.

Southern blot analyses (Fig. 6 *c*) with the insert from λN903 yielded single bands in Bam HI, Dra I, Ban I, Sma I, and Sst II digests. These data indicate that Ng-CAM is encoded by a single gene.

Similarity of Ng-CAM to Other Neural CAMs

A comparison of the amino acid sequence of Ng-CAM to the Dayhoff protein sequence database, the translated Genbank database, and specifically to the sequences of mouse L1 (63), chicken contactin (69) and F11 (7), mouse F3 (30), and other neural CAMs, showed that it was most similar to that of mouse L1 (Table I). However, the two proteins are less similar than expected for equivalent molecules in different species. While they are identical in some segments, particularly in the cytoplasmic domains (Fig. 7), the overall sequence

Published March 1, 1991

identity is only 40% (Table I). Ng-CAM is significantly less similar (21–29% identity) to other neural CAMs.

The FASTA rapid homology search program (65) showed that the six Ig domains in Ng-CAM are similar to each other, and that the individual Ig domains of chicken Ng-CAM are most similar to the corresponding domains in mouse L1 with a one-to-one correspondence of highest FASTA scores (ranging from 162–383; ktup = 1) between each domain in Ng-CAM and the corresponding domain in L1. These segments are compared in Fig. 7 *a*. The identities between the corresponding domains range from 40% in the fifth and sixth domains to 66 and 54% in the second and third domains, respectively. Many of the identities occur in strings of five to ten consecutive residues.

The five fibronectin-type III repeats of Ng-CAM are similar to each other and to five corresponding segments in L1 (Fig. 7 *b*), although these similarities are weaker than those between Ig domains, ranging from 41% in the first domain to 13% in the fifth domain. The amino terminus of the 80-kD component is in the third repeat. The 18 amino acids surrounding this site differ significantly from the corresponding region of L1 with only two amino acid identities.

Between the third and fourth type III repeats in Ng-CAM is an extended segment of 29 amino acids which includes an RGD sequence. The first 21 amino acids of this segment align most favorably with the last part of the preceding type III repeat (Fig. 7 *b*). However, there appears to be no corresponding segment in L1. This segment is 32% identical to the type III repeat in human fibronectin that contains the RGD-containing sequence that binds to integrins (47, 50).

The predicted transmembrane domain and portions of the cytoplasmic region of Ng-CAM are more similar to L1 than any other region (Fig. 7 *c*). The 23-residue transmembrane segment (Ng-CAM residues 1110–1132) is 81% identical, and a continuous segment of 50 amino acids beginning in the center of the transmembrane region and extending into the cytoplasmic domain (amino acids 1121–1170) is identical in the two proteins. This region is followed by a segment of 26 residues that is only 31% identical to L1 and then 27 amino acids that are 89% identical. The remaining carboxyl-terminal segments of the two proteins are relatively dissimilar (23% identity in 22 amino acids).

Discussion

We have isolated and characterized cDNA clones encoding chicken Ng-CAM and formulated a model of the molecule (Fig. 8). The molecule contains six Ig domains and five fibronectin-type III repeats, a transmembrane segment, and a cytoplasmic domain. The proposed linear arrangement of the components is supported by the following observations: (*a*) the amino-terminal sequences of the 210-, 190-, and 135-kD components are the same and differ from the 80-kD species (Fig. 1); (*b*) the 135- and 80-kD components are immunologically and structurally distinct but both are related to the 200-kD forms (35); (*c*) the predicted protein sequence is continuous and all peptide sequences from the 80-kD component are contained within a different part of the sequence than those from the 135-kD component (Fig. 3); (*d*) all components are glycosylated, but only the ~200- and 80-kD components incorporate phosphate (35) and palmitate (Sorkin, B. C., M. Grumet, B. A. Cunningham, and G. M. Edelman. 1985. *J. Neurosci.* 11:1138); (*e*) all cDNA probes detect a 6-kb Ng-CAM mRNA; and (*f*) Southern blots indicate that there is a single Ng-CAM gene.

The data indicate that a single mRNA encodes a large translation product; by posttranslational modification this gives rise to the 190/210-kD forms of the molecule, which is proteolytically cleaved into an extracellular 135-kD component and a transmembrane 80-kD component. The region around the presumed site of cleavage is unusually abundant in prolines and basic residues and could thus be particularly susceptible to proteases. The sequence in this region is very different from that of L1 (21% identical) which may explain why in L1 and NILE the 200-kD forms predominate. The difference between the 190- and 210-kD components of Ng-CAM is unclear, but the detection of a single size mRNA and of a single protein component synthesized in tunicamycin-treated cells suggests that it may be posttranslational. Because extensive alternative splicing is seen in N-CAM (13) and other CAMs (44, 56, 79), however, it is possible that there are alternatively spliced forms of Ng-CAM.

The extracellular domain of Ng-CAM contains all of the Ig-like domains and fibronectin type III repeats. In N-CAM (12), and presumably most other CAMs of this type, the Ig-

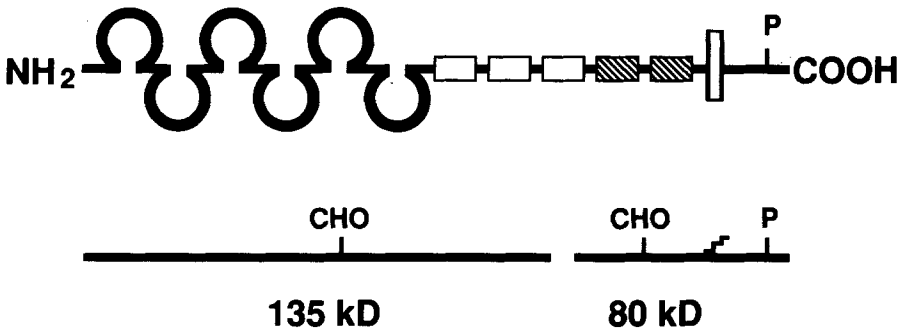


Figure 8. Model of the domain structure of Ng-CAM. Immunoglobulin-like domains are shown as loops in the amino-terminal portion of the polypeptide and fibronectin-type III repeats are shown as boxes. The open boxes represent repeats that have more of the distinguishing features than the others (see text). The predicted transmembrane region is indicated by a vertical bar. Linear representations suggest that the 135- and 80-kD components are generated proteolytically from the ~200-kD component. General locations of carbohydrates (*CHO*), phosphate (*P*), and palmitate (*ladder*) are indicated.

like domains include the regions that mediate cell–cell binding. Of the six Ig domains in Ng-CAM, the second one is the most highly conserved when compared to the second Ig-like domains of L1, N-CAM and neuroglian; in N-CAM this segment has been implicated in binding (10, 12). It is likely that this region of Ng-CAM also plays a role in its homophilic-binding activity.

Although the function of all of the type III repeats in fibronectin is not well understood, certain repeats contain sequences, including an RGD sequence, that are involved in specific cell binding (66). The function of type III repeats in other proteins is unknown. Ng-CAM contains a 29-amino acid segment with one RGD sequence. This segment follows the third type III repeat (amino acid 893), resembles part of a type III repeat, and shows significant similarity to the RGD region in fibronectin. L1 contains two RGD sequences that are close to each other in the sixth immunoglobulin domain (63) and TAG-1 contains an RGD sequence in its second fibronectin type III repeat (27). It remains to be determined whether such sequences in any CAM are involved in cell binding. In Ng-CAM, the RGD-containing segment and each of the fibronectin-type III repeats ends with a sequence containing the nucleotides AG, present at most exon/intron borders. The RGD-containing segment may be a discrete exon; it is sufficiently small that if it were differentially spliced in either Ng-CAM or L1, the alternative mRNAs would be difficult to detect on RNA transfer blots.

The RGD sequence in Ng-CAM is in the 80-kD region of the molecule, raising the possibility that there could be two kinds of Ng-CAM binding: one mediated via the Ig domains and one via the RGD sequence. The cleavage between the 135- and 80-kD regions could allow separation of these domains so that the activities could be separated from each other and regulated independently.

The 135-kD component is tightly associated with the plasma membranes and requires detergent to release it, but it apparently contains no transmembrane segment (Fig. 5). In addition, the 210-, 190-, and 80-, as well as the 135-kD components of Ng-CAM can all be efficiently reconstituted into liposomes, indicating that the 135-kD form binds directly or indirectly to lipid membranes (33). Moreover, all three components moved together in sucrose gradients, suggesting the presence of a supramolecular complex (Becker, J., M. Grumet, G. M. Edelman, and B. A. Cunningham, unpublished observations). These results imply that the 135-kD component is tightly associated with the other Ng-CAM components noncovalently. Such a complex could account for the association of the 135-kD form with the cell surface, but the association must be quite strong because neither high salt nor 0.1 N NaOH released the 135-kD component from membranes (Fig. 5). An alternate possibility is that it may contain lipids which directly anchor it in the plasma membrane. Preliminary experiments, however, indicated that it is not released by phosphatidylinositol-specific phospholipase C as are some forms of N-CAM (40, 41).

The extracellular portion of Ng-CAM includes 12 potential N-linked glycosylation sites and at least six of them appear to be used. Five of the potential sites are in corresponding positions in L1 and are therefore likely to be among those used. Conservation of these sites may reflect the importance of particular carbohydrate moieties to the conformation, distribution, or binding properties of each protein. Such roles have been suggested for the carbohydrates in L1, particularly with regard to their ability to influence apparent L1/N-CAM interactions (48). Both Ng-CAM and L1 contain at least one sulfated oligosaccharide epitope recognized by the mAbs HNK-1, NC-1, and L2 (35, 52). This carbohydrate is present on several other proteins, including N-CAM (35), cytotactin (36) and the myelin-associated glycoprotein (52), although no functional role for it has been demonstrated.

The fibronectin repeats in Ng-CAM are followed by a transmembrane segment and a relatively small (113 amino acids) cytoplasmic domain. The cytoplasmic region includes several serines, a threonine, and a tyrosine that are potential phosphorylation sites. Previous studies (35) showed that components of ~200- and 80-, but not the 135-kD component can be phosphorylated. Sequences including tyrosine 1165 resemble the autophosphorylation sites in viral tyrosine protein kinases (46) and preliminary experiments indicated that Ng-CAM can be autophosphorylated or that a kinase copurifies with the molecule. Such modification may serve to modulate Ng-CAM's binding to cell–surface ligands or intracellular molecules and may have important consequences in transformed cells.

It is becoming increasingly clear that Ng-CAM and other immunologically related chicken CAMs which have been studied in other laboratories are the same molecule. G4 (72), which is involved in interactions between neuronal fibers in culture, has an amino–terminal sequence that is identical to that of Ng-CAM. The 8D9 antigen (58), which acts as an excellent substrate for neurite extension, is identical to Ng-CAM in terms of its polypeptide profile and tissue distribution. In contrast, other molecules that are involved in neurite fasciculation, including neurofascin and axonin-1 are immunologically and structurally distinct from Ng-CAM (71, 73).

The relationship between chicken Ng-CAM and the rodent proteins L1 and NILE is unclear and a detailed comparison to NILE must await the completion of its structure (68). Nevertheless, Ng-CAM, L1, and NILE share many properties and antibodies raised against these molecules cross react (6, 25). The amino acid sequences of Ng-CAM and L1 are nearly identical in their transmembrane segments and in portions of their cytoplasmic domains. Other similarities between Ng-CAM, L1, and NILE include: (*a*) involvement in neuronal aggregation, neurite fasciculation and migration of granule cells in cerebellar slices (24, 35, 43, 59, 70, 87); (*b*) similar tissue distribution in most locations (70, 86, 90); (*c*) polypeptide components that include ~200-kD glycoproteins which can give rise to 135–140- and 80-kD forms, and these different forms probably interact with each other (22, 68, 74); and (*d*) organization of Ig-like domains, fibronectin repeats and cytoplasmic domains (Fig. 8; 63).

Ng-CAM and L1 (and presumably NILE) also show significant differences. The amino acid sequences have segments that are unrelated and overall they are only 40% identical. Other differences include: (*a*) involvement of Ng-CAM but not L1 in neuron-glia adhesion (33, 49); (*b*) expression of Ng-CAM but not L1 on commissural neurons early during spinal cord development (16, 84); (*c*) prevalence of the 135-kD component in Ng-CAM (33, 35); and (*d*) differences in the location of RGD sequences (Fig. 7; 63).

The fact that the sequences of the two proteins are only 40% identical is the most difficult to reconcile with the idea that they are closely related CAMs. Chicken (13) and mouse

N-CAM (2, 3, 81) sequences are >81% identical (Table I). Moreover, chicken contactin/F11 is 78% identical to its presumed murine homologue F3. In contrast, rat TAG-1, which clearly is different from F3 and contactin, is ~50% identical both to mouse F3 and to chicken contactin/F11. These observations suggest that the same CAMs in different species are in general highly conserved (70–90% identical) and that different CAMs in different species (TAG-1, contactin) can show a higher degree of similarity to each other than is seen for chicken Ng-CAM and mouse L1. Ng-CAM and L1 may thus be different, but closely related members of a family of neural CAMs. Preliminary experiments designed to identify an L1 homologue in chickens and an Ng-CAM homologue in mice have not yet revealed such molecules.

The notion that there are subfamilies of closely related CAMs is supported by our recent characterization of a new neural CAM, Nr-CAM (Grumet, M., M. P. Burgoon, V. Mauro, G. M. Edelman, and B. A. Cunningham. 1989. *J. Neurosci.* 15:568). This molecule contains six Ig domains and several fibronectin-type III repeats. Its sequence resembles Ng-CAM and L1 more closely than it does N-CAM or other neural CAMs such as contactin. It is no more similar to mouse L1 than it is to chicken Ng-CAM, so it is not the chicken equivalent of L1. The similarity between Nr-CAM and Ng-CAM (40% identical) and the similarity of TAG-1 to contactin (50%) indicate that there are structural subfamilies of neural CAMs. Molecules within such subfamilies may serve closely related specialized functions during neural development, particularly in such processes as neurite fasciculation and outgrowth. In insects, molecules resembling N-CAM (i.e., fasciclin II) and Ng-CAM (i.e., neuroglian) can act as CAMs and have partially overlapping distributions on growing neurite fascicles (5, 31, 38). Ng-CAM and other closely related CAMs may have similar patterns in vertebrates. The key question, however, is whether these molecules can act to direct neural patterning or whether they serve other functions.

All of the various neural CAMs are large cell surface glycoproteins comprised of different types of domains (20). Such domains may subserve distinct functions including cell–cell adhesion, cell binding to extracellular material, anchorage in the plasma membrane, and release from membranes. The results presented here suggest that different regions of Ng-CAM may mediate neuron–neuron and neuron–glia adhesion, and molecular-binding studies (33) have indicated that there are different ligands for Ng-CAM on neurons and glia. Moreover, the different binding functions of Ng-CAM could be modulated by cleavage into the 135- and 80-kD components. Preliminary experiments indicated that both neurons and glia bound selectively to Ng-CAM–coated substrates; however, it remains to be determined which of the components contribute to the two binding activities. This problem can now be addressed directly by using a variety of Ng-CAM cDNA constructs to express different regions of the Ng-CAM molecule in transfected cells.

We thank Caroline A. Albanese, Alice Huang, Nancy Torres, Michele Mohamadi, Lisa Rothschild, Meg Mueller, and Won Chuong for excellent technical assistance, Drs. Hidesabura Hanafusa and Joan Levi for a chick brain library, and Dr. Jan Geliebter for advice.

This work was supported by United States Public Health Service Grants NS-21629 and HD-16550, and a Senator Jacob Javits Center for Excellence in Neuroscience Grant (NS-22789). M. Grumet is a recipient of an Irma T. Hirschl Career Scientist Award.

Received for publication 5 October 1990 and in revised form 19 November 1990.

References

1. Arquint, M., J. Roder, L. S. Chia, J. Down, D. Wilkinson, H. Bayley, P. Braun, and R. Dunn. 1987. Molecular cloning and primary structure of myelin-associated glycoprotein. *Proc. Natl. Acad. Sci. USA.* 84:600–604.
2. Barbas, J. A., J.-C. Chaix, M. Steinmetz, and C. Goridis. 1988. Differential splicing and alternative polyadenylation generates distinct NCAM transcripts and proteins in the mouse. *EMBO (Eur. Mol. Biol. Organ.) J.* 7:625–632.
3. Barthels, D., M.-J. Santoni, W. Wille, C. Ruppert, J.-C. Chaix, C. Hirsch, J. C. Fontecilla-Camps, and C. Goridis. 1987. Isolation and nucleotide sequence of mouse NCAM cDNA that codes for a Mr 79,000 polypeptide without a membrane-spanning region. *EMBO (Eur. Mol. Biol. Organ.) J.* 6:907–914.
4. Beasley, L., and W. B. Stallcup. 1987. The nerve growth factor-inducible large external (NILE) glycoprotein and neural cell adhesion molecule (N-CAM) have distinct patterns of expression in the developing rat central nervous system. *J. Neurosci.* 7:708–715.
5. 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.
6. Bock, E., C. Richter-Landsberg, A. Faissner, and M. Schachner. 1985. Demonstration of immunochemical identity between the nerve growth factor-inducible large external (NILE) glycoprotein and the cell adhesion molecule-L1. *EMBO (Eur. Mol. Biol. Organ.) J.* 4:2765–2768.
7. 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.
8. Chirgwin, J. M., A. E. Przybyla, R. J. MacDonald, and W. J. Rutter. 1979. Isolation of biologically active ribonucleic acid from sources enriched in ribonuclease. *Biochemistry.* 18:5294–5299.
9. Cleveland, D. W., S. G. Fischer, M. W. Kirschner, and U. K. Laemmli. 1977. Peptide mapping by limited proteolysis in sodium dodecyl sulfate and analysis by gel electrophoresis. *J. Biol. Chem.* 252:1102–1106.
10. Cole, G. J., and R. Akeson. 1989. Identification of a heparin binding domain of the neural cell adhesion molecule N-CAM using synthetic peptides. *Neuron.* 2:1157–1165.
11. Crossin, K. L., C.-M. Chuong, and G. M. Edelman. 1985. Expression sequences of cell adhesion molecules. *Proc. Natl. Acad. Sci. USA.* 82: 6942–6946.
12. Cunningham, B. A., S. Hoffman, U. Rutishauser, J. J. Hemperly, and G. M. Edelman. 1983. Molecular topography of N-CAM: surface orientation and location of sialic acid-rich and binding regions. *Proc. Natl. Acad. Sci. USA.* 80:3116–3120.
13. 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.
14. Dale, R. M. K., B. A. McClure, and J. P. Houchins. 1985. A rapid single-stranded cloning strategy for producing a sequential series of overlapping clones for use in DNA sequencing: application to sequencing the corn mitochondrial 18 S rDNA. *Plasmid.* 13:31–40.
15. Daniloff, J., C.-M. Chuong, G. Levi, and G. M. Edelman. 1986. Differential distribution of cell adhesion molecules during histogenesis of the chicken nervous system. *J. Neurosci.* 6:739–758.
16. Dodd, J., S. B. Morton, D. Karagogeos, M. Yamamoto, and T. M. Jessell. 1988. Spatial regulation of axonal glycoprotein expression on subsets of embryonic spinal neurons. *Neuron.* 1:105–116.
17. Edelman, G. M. 1987. CAMs and Igs: cell adhesion and the evolutionary origins of immunity. *Immunol. Rev.* 100:11–45.
18. Edelman, G. M. 1988. Morphoregulatory molecules. *Biochemistry.* 27: 3533–3543.
19. Edelman, A. M., and D. K. Blumenthal. 1987. Protein Serine/Threonine Kinases. *Annu. Rev. Biochem.* 56:567–613.
20. Edelman, G. M., and B. A. Cunningham. 1990. Place-dependent cell adhesion, process retraction, and spatial signalling in neural morphogenesis. *Cold Spring Harbor Symp. Quant. Biol.* In press.
21. Eisenberg, D. 1984. Three-dimensional structures of membrane and surface proteins. *Annu. Rev. Biochem.* 53:595–623.
22. Faissner, A., D. B. Teplow, D. Kubler, G. Keilhauer, V. Kinzel, and M. Schachner. 1985. Biosynthesis and membrane topography of the neural cell adhesion molecule L1. *EMBO (Eur. Mol. Biol. Organ.) J.* 4:3105–3113.
23. Feinberg, A. P., and B. Vogelstein. 1984. A technique for radiolabeling DNA restriction endonuclease fragments to high specific activity. *Anal. Biochem.* 137:266–267.

24. Fischer, G., J. Künemund, and M. Schachner. 1986. Neurite outgrowth patterns in cerebellar microexplant cultures are affected by antibodies to the cell surface glycoprotein L1. *J. Neurosci.* 6:605–612.
25. Friedlander, D. R., M. Grumet, and G. M. Edelman. 1986. Nerve growth factor enhances expression of neuron-glia cell adhesion molecule in PC12 cells. *J. Cell Biol.* 102:413–419.
26. Frischauf, A. -M., H. Lehrach, A. Polstka, and N. M. Murray. 1983. Lambda replacement vectors carrying-polylinker sequences. *J. Mol. Biol.* 170:827–842.
27. Furley, A. J., S. B. Morton, D. Manalo, D. Karagogeos, J. Dodd, and T. M. Jessel. 1990. The axonal glycoprotein TAG-1 is an immunoglobulin superfamily member with neurite-promoting activity. *Cell.* 61: 157–170.
28. Gallin, W. J., B. C. Sorkin, G. M. Edelman, and B. A. Cunningham. 1987. Sequence analysis of a cDNA clone encoding the liver cell adhesion molecule, L-CAM. *Proc. Natl. Acad. Sci. USA.* 84:2808–2812.
29. Geliebter, J. 1987. Dideoxynucleotide sequencing of RNA and uncloned cDNA. *BRL. Focus.* 9:1:5–8.
30. 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.
31. Grenningloh, G., A. Bieber, J. Rehm, P. Snow, Z. Traquina, M. Hortsch, N. Patel, and C. S. Goodman. 1990. Molecular genetics of neuronal recognition in Drosophila: evolution and function of immunoglobulin superfamily cell adhesion molecules. *Cold Spring Harbor Symp. Quant. Biol.* In press.
32. Grumet, M., and G. M. Edelman. 1984. Heterotypic binding between neuronal membrane vesicles and glial cells is mediated by a specific cell adhesion molecule. *J. Cell Biol.* 98:1746–1756.
33. Grumet, M., and G. M. Edelman. 1988. Neuron-glia cell adhesion molecule interacts with neurons and astroglia via different binding mechanisms. *J. Cell Biol.* 106:487–503.
34. Grumet, M., S. Hoffman, and G. M. Edelman. 1984. Two antigenically related neuronal CAMs of different specificities mediate neuron-neuron and neuron-glia adhesion. *Proc. Natl. Acad. Sci. USA.* 81:267–271.
35. Grumet, M., S. Hoffman, C.-M. Chuong, and G. M. Edelman. 1984. Polypeptide components and binding functions of neuron-glia cell adhesion molecules. *Proc. Natl. Acad. Sci. USA.* 81:7989–7993.
36. Grumet, M., S. Hoffman, K. L. Crossin, and G. M. Edelman. 1985. Cytotactin, an extracellular matrix protein of neural and non-neural tissues that mediates glia-neuron interaction. *Proc. Natl. Acad. Sci. USA.* 82:8075–8079.
37. Gubler, U., and B. J. Hoffman. 1983. A simple and very efficient method for generating cDNA libraries. *Gene.* 25:263–269.
38. 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).* 242:700–707.
39. Hathaway, G. M., and J. A. Traugh. 1982. Casein kinases—multipotential protein kinases. *Curr. Top. Cell. Regul.* 21:101–127.
40. He, H. T., J. Barbet, J. C. Chaix, and C. Goridis. 1986. Phosphatidylinositol is involved in the membrane attachment of NCAM-120, the smallest component of the neural cell adhesion molecule. *EMBO (Eur. Mol. Biol. Organ.) J.* 5:2489–2494.
41. Hemperly, J. J., G. M. Edelman, and B. A. Cunningham. 1986. cDNA clones of N-CAM lacking a membrane-spanning region consistent with evidence for membrane attachment via a phosphatidylinositol intermediate. *Proc. Natl. Acad. Sci. USA.* 83:9822–9826.
42. Hemperly, J. J., B. A. Murray, G. M. Edelman, and B. A. Cunningham. 1986. Sequence of a cDNA clone encoding the polysialic acid-rich and cytoplasmic domains of the neural cell adhesion molecule N-CAM. *Proc. Natl. Acad. Sci. USA.* 83:3037–3041.
43. Hoffman, S., D. R. Friedlander, C.-M. Chuong, M. Grumet, and G. M. Edelman. 1986. Differential contributions of Ng-CAM and N-CAM to cell adhesion in different neural regions. *J. Cell Biol.* 103:145–158.
44. Hortsch, M., A. J. Bieber, N. H. Patel, and C. S. Goodman. 1990. Differential splicing generates a nervous system-specific form of drosophila neuroglian. *Neuron.* 4:697–709.
45. Hubbard, S. C., and R. J. Ivatt. 1981. Synthesis and processing of asparagine-linked oligosaccharides. *Annu. Rev. Biochem.* 50:555–583.
46. Hunter, T., and J. A. Cooper. 1985. Protein-tyrosine kinases. *Annu. Rev. Biochem.* 54:897–930.
47. Hynes, R. O. 1987. Integrins: a family of cell surface receptors. *Cell.* 48:549–554.
48. Kadmon, G., A. Kowitz, P. Altevogt, and M. Schachner. 1990. Functional cooperation between the neural adhesion molecules L1 and N-CAM is carbohydrate dependent. *J. Cell Biol.* 110:209–218.
49. Keilhauer, G., A. Faissner, and M. Schachner. 1985. Differential inhibition of neurone-neurone, neurone-astrocyte and astrocyte-astrocyte adhesion by L1, L2, and N-CAM antibodies. *Nature (Lond.).* 316:728–730.
50. Kornblihtt, A. R., K. Umezawa, K. Vibe-Pedersen, and F. E. Baralle. 1985. Primary structure of human fibronectin: differential splicing may generate at least 10 polypeptides from a single gene. *EMBO (Eur. Mol. Biol. Organ.) J.* 4:1755–1759.
51. Krug, M. S., and S. L. Berger. 1987. First strand cDNA synthesis primed with Oligo (dT). *Methods Enzymol.* 152:316–325.
52. Kruse, J., R. Mailhammer, H. Wernecke, A. Faissner, I. Sommer, C. Goridis, and M. Schachner. 1984. Neural cell adhesion molecules and myelin-associated glycoprotein share a common carbohydrate moiety recognized by monoclonal antibodies L2 and HNK-1. *Nature (Lond.).* 311:151–153.
53. Kuo, S.-C., and J. O. Lampen. 1974. Tunicamycin—An inhibitor of yeast glycoprotein synthesis. *Biochem. Biophys. Res. Commun.* 58:287–295.
54. Kyte, J., and R. F. Doolittle. 1982. A simple method for displaying the hydropathic character of a protein. *J. Mol. Biol.* 157:105–132.
55. Laemmli, U. K. 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature (Lond.).* 227:680–685.
56. Lai, C., M. A. Brow, K.-A. Nave, A. B. Noronha, R. H. Quarles, F. E. Bloom, R. J. Milner, and J. G. Sutcliffe. 1987. Two forms of 1B236/myelin-associated glycoprotein, a cell adhesion molecule for postnatal neural development, are produced by alternative splicing. *Proc. Natl. Acad. Sci. USA.* 84:4337–4341.
57. Lemke, G., and R. Axel. 1985. Isolation and sequence of a cDNA encoding the major structural protein of peripheral myelin. *Cell.* 40:501–508.
58. Lemmon, V., and S. C. McLoon. 1986. The appearance of an L1-like molecule in the chick primary visual pathway. *J. Neurosci.* 6:2987–2994.
59. Lindner, J., F. G. Rathjen, and M. Schachner. 1983. L1 Mono- and polyclonal antibodies modify cell-migration in early postnatal mouse cerebellum. *Nature (Lond.).* 305:427–430.
60. MacDonald, R. J., G. H. Swift, A. E. Przybyla, and J. M. Chirgwin. 1987. Isolation of RNA using guanidinium salts. *Methods Enzymol.* 152:219–227.
61. Maniatis, T., E. F. Fritsch, and J. Sambrook. 1982. Molecular Cloning: A Laboratory Manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, NY. 545 pp.
62. Matsudaira, P. 1987. Sequence from picomole quantities of proteins electroblotted onto polyvinylidene difluoride membranes. *J. Biol. Chem.* 262:10035–10038.
63. 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.
64. Morrissey, J. H. 1981. Silver stain for proteins in polyacrylamide gels—A modified procedure with enhanced uniform sensitivity. *Anal. Biochem.* 117:307–310.
65. Pearson, W. R., and D. J. Lipman. 1988. Improved tools for biological sequence comparison. *Proc. Natl. Acad. Sci. USA.* 85:2444–2448.
66. Pierschbacher, M. D., and E. Ruoslahti. 1984. Cell attachment activity of fibronectin can be duplicated by small synthetic fragments of the molecule. *Nature (Lond.).* 309:30–33.
67. Popowicz, A. M., and P. F. Dash. 1988. SIGSEQ: a computer program for predicting signal sequence cleavage sites. *Computer Appl. Biosci.* 4:405–406.
68. Prince, J. T., N. Milona, and W. B. Stallcup. 1989. Characterization of a partial cDNA clone for the NILE glycoprotein and identification of the encoded polypeptide domain. *J. Neurosci.* 9:876–883.
69. Ranscht, B. 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.
70. Rathjen, F. G., and M. Schachner. 1984. Immunocytological and biochemical characterization of a new neuronal cell surface component (L1 antigen), which is involved in cell adhesion. *EMBO (Eur. Mol. Biol. Organ.) J.* 3:1–10.
71. Rathjen, F. G., J. M. Wolff, S. Chang, F. Bonhoeffer, and J. A. Raper. 1987. Neurofascin: a novel chick cell-surface glycoprotein involved in neurite-neurite interactions. *Cell.* 51:841–849.
72. Rathjen, F. G., J. M. Wolff, R. Frank, F. Bonhoeffer, and U. Rutishauser. 1987. Membrane glycoproteins involved in neurite fasciculation. *J. Cell Biol.* 104:343–353.
73. Ruegg, M. A., E. T. Stoeckli, R. B. Lanz, P. Streit, and P. Sonderegger. 1989. A homologue of the axonally secreted protein axonin-1 is an integral membrane protein of nerve fiber tracts involved in neurite fasciculation. *J. Cell Biol.* 109:2363–2378.
74. Sadoul, K., R. Sadoul, A. Faissner, and M. Schachner. 1988. Biochemical characterization of different molecular forms of the neural cell adhesion molecule L1. *J. Neurochem.* 50:510–521.
75. Sadoul, R., F. Kirchhoff, and M. Schachner. 1989. A protein kinase activity is associated with and specifically phosphorylates the neural cell adhesion molecule L1. *J. Neurochem.* 53:1471–1478.
76. Saiki, R. K., D. H. Gelfand, S. Stoffel, S. J. Scharf, R. Higuchi, G. T. Horn, K. B. Mullis, and H. A. Erlich. 1988. Primer-directed enzymatic amplification of DNA with a thermostable DNA polymerase. *Science (Wash. DC).* 239:487–491.
77. Sajovic, P., E. Kouvelas, and E. Trenkner. 1986. Probable identity of NILE glycoprotein and the high-molecular-weight component of L1 antigen. *J. Neurochem.* 47:541–546.
78. Salton, S. R. J., M. L. Shelanski, and L. A. Greene. 1983. Biochemical

properties of the nerve growth factor-inducible large external (NILE), glycoprotein. *J. Neurosci.* 3:2420–2430.
79. Salzer, J. G., W. P. Holmes, and D. R. Colman. 1987. The amino acid sequences of the myelin-associated glycoproteins—Homology to the immunoglobulin gene superfamily. *J. Cell Biol.* 104:957–965.
80. Sanger, F., S. Nicklen, and A. R. Coulson. 1977. DNA sequencing with chain-terminating inhibitors. *Proc. Natl. Acad. Sci. USA.* 74:5463–5467.
81. Santoni, M.-J., D. Barthels, J. A. Barbas, R. Hirsh, C. Goridis, and W. Wille. 1987. Analysis of cDNA clones that code for the transmembrane forms of the mouse neural cell adhesion molecule (NCAM) and are generated by alternative RNA splicing. *Nucleic Acids Res.* 15:8621–8641.
82. Schwarzbauer, J. E., R. S. Patel, D. Fonda, and R. O. Hynes. 1987. Multiple sites of alternative splicing of the rat fibronectin gene transcript. *EMBO (Eur. Mol. Biol. Organ.) J.* 6:2573–2580.
83. Seeger, M. A., L. Haffley, and T. C. Kaufman. 1988. Characterization of amalgam: a member of the immunoglobulin superfamily from Drosophila. *Cell.* 55:589–600.
84. Shiga, T., R. W. Oppenheim, M. Grumet, and G. M. Edelman. 1990. Neuron-glia cell adhesion molecule (Ng-CAM) expression in the chick spinal cord: observations on the earliest developing intersegmental interneurons. *Dev. Brain Res.* 55:209–217.
85. Staden, R. 1984. Graphic methods to determine the function of nucleic acid sequences. *Nucleic Acids. Res.* 12:521–538.
86. Stallcup, W. B., and L. L. Beasley. 1985. Involvement of the nerve growth factor-inducible large external glycoprotein (NILE) in neurite fasciculation in primary cultures of rat brain. *Proc. Natl. Acad. Sci. USA.* 82:1276–1280.
87. Stallcup, W. B., L. L. Beasley, and J. M. Levine. 1985. Antibody against nerve growth factor-inducible large external (NILE) glycoprotein labels nerve fiber tracts in the developing rat nervous system. *J. Neurosci.* 5:1090–1101.
88. Takeichi, M. 1988. The cadherins: cell-cell adhesion molecules controlling animal morphogenesis. *Development.* 102:639–655.
89. Thiery, J.-P., J.-L. Duband, U. Rutishauser, and G. M. Edelman. 1982. Cell adhesion molecules in early chick embryogenesis. *Proc. Natl. Acad. Sci. USA.* 79:6737–6741.
90. Thiery, J.-P., A. Delouvée, 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.
91. 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.
92. Tuazon, P. T., E. W. Bingham, and J. A. Traugh. 1979. Cyclic nucleotide-independent protein kinases from rabbit reticulocytes. *Eur. J. Biochem.* 94:497–504.
93. vonHeijne, G. 1986. A new method for predicting signal sequence cleavage sites. *Nucleic Acids Res.* 14:4683–4690.
94. Waxdal, M. J., W. H. Konigsberg, W. L. Henley, and G. M. Edelman. 1968. The covalent structure of a human G-immunoglobulin. II. Isolation and characterization of the cyanogen bromide fragments. *Biochemistry.* 7:1959–1966.
95. Williams, A. F., and A. N. Barclay. 1988. The immunoglobulin superfamily—Domains for cell surface recognition. *Annu. Rev. Immunol.* 6:381–405.
96. Young, R. A., and R. W. Davis. 1983. Yeast RNA polymerase III genes: isolation with antibody probes. *Science (Wash. DC).* 222:778–782.
97. Young, R. A., and R. W. Davis. 1983. Efficient isolation of genes by using antibody probes. *Proc. Natl. Acad. Sci. USA.* 80:1194–1198.