

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

Follow this and additional works at: [http://digitalcommons.rockefeller.edu/student\\_theses\\_and\\_dissertations](http://digitalcommons.rockefeller.edu/student_theses_and_dissertations)

 Part of the [Life Sciences Commons](#)

---

## Recommended Citation

Burgoon, Mark P., "Structure of the Chicken Neuron-Glia Cell Adhesion Molecule, Ng-CAM: Origin of the Polypeptides and Relation to the Ig Superfamily" (1991). *Student Theses and Dissertations*. 362.  
[http://digitalcommons.rockefeller.edu/student\\_theses\\_and\\_dissertations/362](http://digitalcommons.rockefeller.edu/student_theses_and_dissertations/362)

This Thesis is brought to you for free and open access by Digital Commons @ RU. It has been accepted for inclusion in Student Theses and Dissertations by an authorized administrator of Digital Commons @ RU. For more information, please contact [mcsweej@mail.rockefeller.edu](mailto:mcsweej@mail.rockefeller.edu).





STRUCTURE OF THE CHICKEN NEURON-GLIA CELL ADHESION MOLECULE, Ng-CAM:  
ORIGIN OF THE POLYPEPTIDES AND  
RELATION TO THE Ig SUPERFAMILY

A thesis submitted to the faculty of the Rockefeller University  
in partial fulfillment of the requirements  
for the degree of Doctor of Philosophy  
by

Mark P. Burgoon, B.A.

April 1, 1991

The Rockefeller University  
New York







To my parents who made all this possible



## ACKNOWLEDGEMENTS

I am particularly grateful to Professors Gerald Edelman and Bruce Cunningham for guiding my development here at Rockefeller, and for providing an exceptionally stimulating intellectual environment in which to grow. Dr. Cunningham has been my advisor throughout this work, and I deeply appreciate his expert guidance and constant generosity.

I am indebted to many past and present members of the Edelman/Cunningham laboratory for providing an immensely enjoyable and invigorating environment in which to study and live. I began these studies with Robert Brackenbury and learned many fundamentals of experimental design from his enthusiastic support. Martin Grumet and Vincent Mauro contributed directly to this work. Their skilled collaboration has strengthened the endeavor. I am also grateful to my family and friends who encouraged me through these anxious years.

A  $\lambda$ gt10 cDNA library was donated by Drs. Hidesaburo Hanafusa and Joan Levi of The Rockefeller University. Peptide sequencing and oligonucleotide synthesis were performed at the Rockefeller University Protein Sequencing facility.





## TABLE OF CONTENTS

Abstract.....	1
Chapter 1 BACKGROUND AND INTRODUCTION.....	4
History of Cell Adhesion.....	5
Cell Adhesion Assays.....	11
Identification of Vertebrate CAMs.....	13
Structural Families of CAMs.....	16
Calcium-dependent Neural CAMs.....	20
Neural CAMs of the N-CAM Family.....	21
Other Neural CAMs.....	29
Ng-CAM.....	32
Chapter 2 MATERIALS AND METHODS.....	45
Animals.....	45
Ng-CAM Purification.....	45
Ng-CAM Peptides.....	46
SDS-PAGE.....	47
RNA Purification.....	50
RNA and DNA Transfer Blots.....	51
Library Construction.....	51
Custom Libraries.....	57
Library Screening.....	60
cDNA Probe Labelling.....	62
cDNA Isolation and Purification.....	62
DNA Sequencing.....	63



	Polymerase Chain Reaction.....	66
	Phosphorylation of Ng-CAM.....	66
	General Solutions.....	68
Chapter 3	CHARACTERIZATION OF Ng-CAM.....	69
	Protein Sequence.....	69
	Cloning Strategy.....	75
	cDNA Clones.....	79
	Model of Ng-CAM.....	88
	Domain Structure.....	93
	Organization of the Ng-CAM Components.....	102
	Similarity of Ng-CAM to Other Neural CAMs.....	113
	Search for Other Similar CAMs.....	129
	Discussion.....	135
Chapter 4	Nr-CAM AND SUBGROUPS OF NEURAL CAMs.....	144
	Isolation and Characterization of cDNA Clones...	144
	Potential Alternative Splicing.....	147
	Nr-CAM Protein.....	151
	Localization of Nr-CAM Protein.....	153
	Sequence and Model of Nr-CAM.....	154
	Comparison of Nr-CAM, Ng-CAM, and Ll.....	157
	Similarity to Other CAMs.....	164
	Discussion.....	168
	Families of CAMs.....	169
	Subgroups of Ig-like Neural CAMs.....	172
	Future Prospectives.....	180
	Bibliography.....	182



## TABLE OF ILLUSTRATIONS

Table 1	Peptide sequences.....	49
Table 2	cDNA libraries.....	56
Table 3	Oligonucleotides.....	59
Table 4	FASTA comparison of Ig-like domains.....	116
Table 5	Comparison of domains of Nr-CAM, Ng-CAM, and L1.....	159
Figure 1	Neural cell adhesion molecules.....	18
Figure 2	Monolayer assay for neuron-glia binding.....	34
Figure 3	Ng-CAM components and amino terminal sequences.....	71
Figure 4	Linear model of Ng-CAM.....	73
Figure 5	Isolation of Ng-CAM cDNA clone with PCR.....	77
Figure 6	Isolation of Ng-CAM cDNA clones with antibodies.....	81
Figure 7	Ng-CAM cDNA clones.....	83
Figure 8	Nucleotide sequence and deduced protein sequence of Ng-CAM.....	87
Figure 9	Prediction of signal peptide cleavage sites.....	90
Figure 10	Hydrophobicity plot of Ng-CAM.....	92
Figure 11	Alignment of Ng-CAM Ig-like domains.....	95
Figure 12	Consensus sequences of Ig subtypes.....	98
Figure 13	Alignment of Fn-III repeats of Ng-CAM.....	101
Figure 14	Northern analyses of Ng-CAM.....	105
Figure 15	Southern analyses of Ng-CAM.....	108
Figure 16	Membrane association of Ng-CAM.....	112



Figure 17 Comparison of Ng-CAM and L1 protein sequences.....	119
Figure 18 Phosphorylation of Ng-CAM.....	128
Figure 19 Relationship of Ng-CAM to other neural CAMs.....	131
Figure 20 PCR for similar molecules.....	133
Figure 21 Model for the domain structure of Ng-CAM.....	137
Figure 22 Comparison of NILE protein sequence to Ng-CAM and L1..	143
Figure 23 First Nr-CAM cDNA clone.....	146
Figure 24 Nr-CAM cDNA clones.....	149
Figure 25 Model for the domain structure of Nr-CAM.....	156
Figure 26 Comparison of Ng-CAM, Nr-CAM, and L1.....	161
Figure 27 Relationship of Nr-CAM to other neural CAMs.....	166
Figure 28 Subgroups of the N-CAM-like neural CAMs.....	175





## ABBREVIATIONS

CAM, cell adhesion molecule  
Ng-CAM, neuron-glia cell adhesion molecule  
Nr-CAM, Ng-CAM related cell adhesion molecule  
N-CAM, neural cell adhesion molecule  
L-CAM, liver cell adhesion molecule  
cadherin, calcium-dependent cell adhesion molecule  
NILE glycoprotein, NGF-inducible large external glycoprotein  
MAG, myelin associated glycoprotein  
Ig, immunoglobulin  
Fn-III, fibronectin-type III  
ECM, extracellular matrix  
csA glycoprotein, contact site A glycoprotein  
GalTase, galactosyltransferase  
ICAM-1, intercellular adhesion molecule 1  
ICAM-2, intercellular adhesion molecule 2  
OBCAM, opioid binding protein-cell adhesion molecule  
LEC-CAM, lectin-like cell adhesion molecule  
PCR, polymerase chain reaction  
bp, basepair(s)  
kb, kilobase(s)  
kD, kilodaltons  
d, day  
BSA, bovine serum albumin  
DEPC, diethylpyrocarbonate  
DTT, dithiothreitol  
EDTA, (ethylenedinitrilo)tetraacetic acid (disodium salt)  
EtBr, ethidium bromide  
IPTG, isopropylthio- $\beta$ -D-galactopyranoside  
MOPS, 4-morpholinepropanesulfonic acid  
NP-40, nonidet P-40  
PAGE, polyacrylamide gel electrophoresis  
PBS, phosphate-buffered saline  
PMSF, phenylmethyl-sulfonyl fluoride  
Poly (A)<sup>+</sup> RNA, polyadenylated ribonucleic acid  
SDS, sodium dodecyl sulfate  
SSC, sodium chloride/sodium citrate buffer  
STE, sodium chloride/Tris/EDTA buffer  
TBE, Tris/boric acid/EDTA buffer  
X-Gal, 5-bromo-4-chloro-3-indoyl- $\beta$ -D-galactopyranoside



## ABSTRACT

The neuron-glia cell adhesion molecule, Ng-CAM, is a cell-surface glycoprotein expressed on neurons and Schwann cells in the chicken nervous system. It is expressed as a group of related components of ~200 kD, 135 kD, and 80 kD, which are detected at particular sites during development. Ng-CAM mediates a variety of functions that are known to be important in specific cell-cell interactions and the histogenesis of the nervous system, including neuron-neuron adhesion, neuron-glia adhesion, neurite fasciculation, and neuronal migration on glial cells.

This thesis describes the characterization of cDNA that encode the entire Ng-CAM protein and the determination of the origin of the polypeptide components. An analysis of the deduced amino acid sequence reveals the presence of immunoglobulin-like domains and fibronectin-type III repeats, and establishes the relationship of Ng-CAM to other neural CAMs containing similar structures. The thesis also describes the identification of a similar protein in the embryonic chicken nervous system, the Ng-CAM-related cell adhesion molecule (Nr-CAM). Its close structural and sequence similarities to Ng-CAM indicate its probable involvement in cell adhesion, and defines a closely related subfamily of neural CAMs.

cDNA libraries prepared from embryonic nervous tissue were used to isolate clones encoding Ng-CAM. These clones encode a polypeptide that includes all three known components of Ng-CAM, and hybridize to a single



mRNA species (6 kb) that is restricted to nervous tissue. Single bands were detected in genomic DNA blots with Ng-CAM cDNA probes, indicating that there is a single Ng-CAM gene in chicken. Together these data indicate that the polypeptide components of Ng-CAM all arise from a single translation product, and that the two smaller polypeptide components (135 kD and 80 kD) are products of posttranslational cleavage.

The deduced amino acid sequence of Ng-CAM predicts the presence of six immunoglobulin-like domains, five fibronectin-type III repeats, a single transmembrane domain, and a short cytoplasmic region. Ng-CAM also contains a single RGD sequence in its third fibronectin-type III repeat, which in an equivalent position in several other proteins is known to mediate a cell-binding activity. Ng-CAM is similar to other neural CAMs containing immunoglobulin-like domains and is most similar to its previously presumed equivalent in mouse, L1. However, overall the amino acid sequence of Ng-CAM is only 40% identical to L1, and within the individual domains ranges from only 15% to 66% identity. Known equivalent CAMs (e.g., N-CAM) in the two species are much more closely related (80% identity), indicating that in contrast to previous assumptions Ng-CAM and L1 may be functionally different molecules.

cDNA clones encoding another protein were isolated from embryonic chicken nervous system libraries using antibodies against Ng-CAM. The new protein shares at least one polypeptide epitope with Ng-CAM, and was designated the Ng-CAM-related cell adhesion molecule, or Nr-CAM. Like Ng-CAM, it is predicted to contain six immunoglobulin-like domains, five fibronectin-type III repeats, one transmembrane segment and a small



cytoplasmic domain, each of which is most similar to corresponding segments in Ng-CAM and L1. However, overall Nr-CAM shares only 40% identity to L1, and is thus not the chicken equivalent of L1. In addition, comparison of several similar Nr-CAM cDNA clones indicated that the fifth fibronectin-type III repeat (93 amino acids) and a segment between the second and third immunoglobulin-like domains (20 amino acids) may be differentially spliced in Nr-CAM mRNA; such alternative forms have not been seen for Ng-CAM and L1.

The combined results indicate that Ng-CAM, Nr-CAM and L1 define a structurally related subfamily of neural CAMs. Their similarities suggest that they may perform similar binding functions in the nervous system, possibly at different times and locations during neural development.





The nervous system in vertebrates contains up to one hundred billion cells and thousands of cell types interacting in complex arrays, often over great distances. Understanding the organization and function of this system is one of the largest unsolved problems in biology, primarily because the diversification of cells and the elaborate organization of the nervous system are more complex than in other systems. One hope has been that the fundamental principles of the organization can be discerned by examining the development of the system, before its complexity and rigidity have been firmly established.

Specific cell-cell interactions must play a key role in establishing and maintaining the complex arrays of cells, both during development and in the adult animal. On this premise the molecules that mediate such interactions are of key importance and it is becoming increasingly clear that specific cell adhesion molecules (CAMs), coupled with neuronal activity-driven events, influence and stabilize such cell-cell interactions. The ability of cells to interact is influenced by at least three types of adhesion molecules: (1) CAMs which link cells together and are usually synthesized by the cells that they ligate; (2) substrate adhesion molecules (SAMs) which are usually found in the extracellular matrix but can bind to cells through specific receptors (e.g., integrins); and (3) cell junctional molecules (CJMs) which are found at specialized junctional complexes between cells, such as gap junctions, tight junctions, desmosomes and adherens junctions. Together



these three types of molecules have distinct and overlapping expression patterns and may interact with each other to influence morphogenesis and histogenesis. They have been termed morphoregulatory molecules (Edelman, 1988).

The neuron-glia cell adhesion molecule, Ng-CAM, mediates a variety of neuron-glia and neuron-neuron interactions in the developing chicken nervous system. My studies have focused on the structural characterization of Ng-CAM, and have described the sequencing of cDNA clones encoding the protein. These and related studies establish the origin of the various Ng-CAM polypeptide components and describe Ng-CAM's relationship to other CAMs. The thesis also includes a description of cDNA clones that encode another protein which is very similar to Ng-CAM and therefore may also be involved in cell adhesion and neural development. I begin with a brief review of the history and features of cell adhesion, followed by a more detailed description of neural CAMs and CAM-mediated functions.

### History of Cell Adhesion

One of the earliest *in vitro* demonstrations of adhesive interactions between cells was performed by H.V. Wilson (Wilson, 1907), who dissociated sponges into single cells by pressing them through fine cloth, and observed that the cells reclustered into aggregates, and then differentiated into functional sponges. Subsequent studies showed that the reaggregation of sponge cells was species-specific (Galtsoff, 1925); when sponges from different species were dissociated into single cells and mixed, the cells sorted out and reformed sponges typical for each



species. These early observations were extended in other species (DeMorgan and Drew, 1914), where it was demonstrated that selective adhesion allowed dissociated embryonic cells to sort out into viable differentiating tissues. It was later observed that in certain fish the disassembly and reaggregation of early embryonic cells may be a natural phenomenon (Wourms, 1972); the early blastomeres of these embryos appeared to disaggregate and then reaggregate spontaneously, so that the fish developed from an apparently random rearrangement of cells.

The study of tissue-specific cell adhesion was extended to vertebrates by Holtfreter (Holtfreter, 1948a; Holtfreter, 1948b; Townes and Holtfreter, 1955). Amphibian neurulae were dissociated into single cell suspensions by exposure to high pH. When the pH was returned to normal, the suspensions of mixed cells reaggregated into compact balls and then sorted out into distinct germ layers. Similarly treated gastrula cells reaggregated to form identifiable neural and mesodermal structures. These findings led Holtfreter to propose that embryonic cells acquire different adhesive properties which enable them to separate and sort out according to their tissue specificity.

Selective sorting out was also demonstrated by Moscona with chicken embryonic cells dissociated by trypsin (Moscona, 1962). Mixtures of embryonic scleral and neural retinal cells *in vitro* initially aggregated into mixed balls of cells. The cells then sorted out to form homogeneous aggregates which continued to develop into cartilaginous sheets and identifiable retinal structures, respectively, resembling the original morphology of the tissues. These studies further established that selective cell adhesion is an important force in development. They



also suggested that cell adhesion is involved in inductive events, if only to hold together cells that undergo induction by other means. But the mechanisms of cell adhesion were still unknown.

Several types of mechanisms have been proposed for the sorting out of cells. Two early models treated cell-cell interactions as the establishment of thermodynamic equilibria between adjacent cell membranes. One model proposed that there was a balance of attractive and repulsive physical forces between the two membranes that could account for the observation that there was a 100 Angstrom space between adhering cells (Curtis, 1967). A second model treated cells essentially as droplets of immiscible liquids with quantitatively different "surface free energies" on their membranes, depending on their cell type (Steinberg, 1970). In this model cell sorting resulted from the segregation of different cell types into aggregates of similar surface free energies, establishing an equilibrium condition between the different types. These two models focused on cell interactions over large surface areas, and did not explain such morphogenetic movements as the extension of thin processes of one cell type into a mass of another cell type. Moreover, the models dealt with quantitative adhesive differences, and were indifferent to the chemical or physical means by which the adhesion events must be generated.

A different type of model for cell-sorting proposed that qualitative differences between cell surface components were involved in the recognition and binding of cells (Moscona, 1962). In these models specific cell-surface molecules were primarily responsible for the selective adhesion and histogenetic affinities of cells. The notion of





ligand specificity was also useful for describing cell recognition and the formation of complex patterns of neuronal connections in the nervous system. However, despite intensive efforts, no specific ligands were identified with the assays in use at that time.

In the absence of identified ligands, several models for neuronal specificity were developed and led to predictions about the nature of cell adhesion. A system intensively studied for its precise neuronal connections is the retinotectal projection in the visual system of amphibian embryos. This system has provided the battlefield for several theories of neuronal specificity.

The neural ganglion cells of the differentiating retina project axons to the tectum, creating a highly ordered topographic array of synapses which accurately represents the visual field. Sperry proposed a model for cell recognition in which the presynaptic and postsynaptic elements of this projection contained matching biochemical labels (identification tags) (Sperry, 1963). Individual retinal neurites would synapse with appropriate tectal cells in an exclusively "hard-wired" fashion. In its strictest form this chemoaffinity model predicted a large number of specific biochemical labels (cell addresses). However, the model had difficulty explaining the constantly shifting retinotectal connections which occur in many species (Chung et al., 1974; Gaze et al., 1979; Meyer, 1978; Reh and Constantine-Paton, 1984; Fujisawa et al., 1984). For example, in the goldfish the retinotectal projection is constantly expanding. Ganglion cells are added circumferentially to the retina throughout life (Johns, 1977; Easter et al., 1981). The tectum also continues to grow, but only with the addition of cells on the



caudal side (Meyer, 1978; Raymond and Easter, 1983). Therefore the retinotectal projection must shift along the caudally expanding tectum to accommodate newly arriving axons from the radially expanding retina. These shifting connections implied a fair degree of plasticity between connections and a lack of a "hard-wired" matching of labels between individual cells.

A modification of the chemoaffinity model proposed that the retinotectal specificity was provided by a gradient of ligand on the tectal cells and a corresponding gradient of ligand across the growth cones of the incoming retinal fibers (Bonhoeffer and Gierer, 1984; Gottlieb et al., 1976). By differential adhesion, the incoming axons would form a correct map along the corresponding gradient expressed in the tectum. Such a model would take into account a shifting tectal or retinal field by maintaining topographic order among connections locally.

Several specific molecules had been suggested to operate in this fashion. Roseman, Roth and colleagues proposed that interactions of retinal fibers with specific complex carbohydrates on tectal surfaces may influence the retinotectal projection. They found  $\beta$ 1-4 galactosyltransferases (GalTase) on retinal cell surfaces, and predicted that these enzymes provided the retinal ligand for recognition (Roseman, 1970; Roth et al., 1971). The subsequent discovery of a rostrocaudal gradient of ganglioside binding activity across the tectal surface supported this idea (Marchase, 1977; Balsamo and Lilien, 1975). More recently a binding function has been demonstrated for a galactosyltransferase (GalTase). Antibodies to GalTase inhibited



carcinoma cell aggregation in culture and disrupted late morula compaction (Bayna et al., 1988). The ligand or the direct contribution of GalTase to adhesion has not yet been established, however. The antibodies may have generally coated the GalTase-containing cells to inhibit aggregation, or blocked the interaction of GalTase with another protein which then serves as the actual ligand.

Other molecules with putative adhesive functions have been detected by their distributions. Several have been found in gradients in the embryo (Gottlieb et al., 1976; Trisler et al., 1981; Moskal et al., 1986; also see McCafferty et al., 1990). A monoclonal antibody, JONES, recognized a glycolipid antigen in a dorsoventral gradient on the surface of the embryonic rat retina (Constantine-Paton et al., 1986). However, the functional role of the antigen has yet to be established. Distinct monoclonal antibodies against chicken retinal cell membranes detected an antigen (TOP) in a 35-fold dorsoventral gradient across the developing retina (Trisler et al., 1981). Yet the distribution of TOP also correlates with the density of neurons across the retina at that stage, and may simply reflect the density of neuronal membranes. Such molecules must be assayed for function beyond their correlation with functional gradient models.

A more flexible model for cell adhesion relied on the premise that development is a dynamic process (Edelman, 1985a). In this model, a complex developmental pattern can be achieved through the action of a small number of specific cell adhesion molecules (CAMs) which are expressed transiently at multiple sites and times during development. Their dynamic expression provides the timed signalling for initial



interaction with incoming fibers bearing complementary ligands. This view, together with the development of functional and immunological assays, enabled the identification and characterization of specific CAMs. This model has in turn been strongly supported by the identification of specific CAMs and analyses of their distribution and function during development.

### Cell Adhesion Assays

Evaluation of the various models described above depends ultimately on the identification of the molecules involved and the characterization of their roles in cell adhesion. To identify such molecules it was important to develop assays that dissociated cell adhesion from other cellular processes.

*In vitro* aggregation assays were developed to measure the binding of dissociated cells in suspension. Moscona pioneered the use of a rotating culture of single cells for aggregation studies, but the assay required 24-48 hours before reproducible aggregates were obtained, and in some cases the cells in the aggregates sorted out and began to differentiate during the assay (Moscona, 1962). Clearly a faster assay was needed to separate the initial adhesion events from subsequent cell movements and inductive processes. Roth and Weston modified the assay by measuring the binding of single cells to preformed aggregates over several hours (Roth and Weston, 1967), but even this length of time allowed cell movement after aggregation. However modifications of both of these assays have used controlled conditions to enable aggregation to





be measured in less than an hour (Brackenbury et al., 1977; Bertolotti et al., 1980; Ocklind and Öbrink, 1982).

A short-term immunologically-based assay was developed by Gerisch and his colleagues to identify cell aggregating factors in the slime mold *Dictyostelium* (Beug et al., 1973; Gerisch, 1980). This approach depended on the ability of monovalent Fab' antibody fragments to a putative adhesion molecule to inhibit the binding of cells. The assay eventually led to the identification of a glycoprotein component called contact site A (csA) glycoprotein, which is developmentally regulated and detected when cells differentiate to an aggregation-competent stage (Huesgen and Gerisch, 1975). However, although csA glycoprotein is reported to demonstrate a distant similarity to the hinge regions of immunoglobulins, its characteristics are generally unlike the cell adhesion molecules later identified in vertebrates (Noegel et al., 1986), and therefore may not serve as a comparable system for understanding cell adhesion in the vertebrate nervous system.

Direct assays have also been used to isolate putative adhesion molecules. Such assays depend on the ability of a component to directly enhance or inhibit the aggregation of cells that normally express the molecule on their surfaces. For example, an aggregation factor is released from sponge cells with the depletion of calcium from the growth medium. Sponge cells do not aggregate when this glycosaminoglycan is absent from the medium, but the cells will aggregate when the factor is added (Humphreys, 1975). Similarly cognin is a 50 kD glycoprotein released into the medium from primary cultures of embryonic chick retinal cells. In solution it promotes retina-specific cell aggregation



in rotating cultures (Hausman and Moscona, 1975), and antibodies raised against cognin have been reported to block the reaggregation of retinal cells and cause extensive malformations upon application to the retina (Ophir et al., 1984; Hausman and Moscona, 1979). Other molecules may be present on cell surfaces which, when solubilized, inhibit the aggregation of cells. This presumably occurs by competing for the binding sites of the component's ligand. Ligatin, a filamentous 10 kD protein isolated from rat ileum (Jakoi et al., 1976) or retina (Jakoe and Marchase, 1979), was assumed to be such a molecule because it inhibited the reaggregation of neural retinal cells in rotating cultures (Marchase et al., 1981). It was later shown, however, to serve as a baseplate for the attachment of other cell-surface molecules, rather than mediating adhesion directly (Jakoi and Corley, 1979).

#### Identification of Vertebrate Cell Adhesion Molecules (CAMs)

A major leap forward was made when a modification of Gerisch's immunological assay was used by Edelman and his colleagues to identify and purify a neural cell adhesion molecule (CAM) from the chicken (Brackenbury et al., 1977; Thiery et al., 1977). Fragments of embryonic chicken retina were dissociated with trypsin into single cell suspensions. The cells were allowed to recover in medium to regain cell surface components, and were then assayed for aggregation in a rotating culture assay. Addition of antibodies raised to embryonic brain cells inhibited cell aggregation. To characterize the cell adhesion activity, media conditioned by embryonic brain cells (or a detergent extract from the cells) was preincubated with the antiserum and neutralized the



antibody's inhibition of aggregation. The ability of material to neutralize the antibody was used as an assay to fractionate the active component.

A glycoprotein was partially purified from extracts of brain membranes based on its ability to neutralize the antibody inhibition in the assay. In this manner the chicken neural cell adhesion molecule (N-CAM) was identified. In the nervous system it is detected on the cell bodies and processes of all neurons, and mediates neuron-neuron adhesion (Thiery et al., 1984; Hoffman and Edelman, 1983). N-CAM has also been reported to be present on many classes of macroglia (Noble et al., 1985). N-CAM was subsequently discovered in other species, and was detected in a variety of non-neural tissues, including smooth and cardiac muscle, tendons, and various epithelial structures (Crossin et al., 1985; Chuong and Edelman, 1985; Walsh and Dickson, 1989).

Adhesion mediated by N-CAM is homophilic, i.e., N-CAM on one cell binds to N-CAM on another cell. This was demonstrated directly with N-CAM reconstituted into lipid vesicles (Hoffman and Edelman, 1983) and by expression of N-CAM in cells transfected with N-CAM cDNA (Edelman et al., 1987a; Mege et al., 1988). Several *in vitro* and *in vivo* studies indicated that it mediates a variety of functions. For example, antibodies to N-CAM alter layer formation in the retina *in vitro* (Buskirk et al., 1980), perturb accurate mapping of the retina to the optic tectum *in vivo* (Fraser et al., 1984; Thanos et al., 1984), and inhibit nerve-muscle interactions both *in vitro* (Rutishauser et al., 1983; Bixby et al., 1987) and *in vivo* (Rieger et al., 1988). In



addition, transfected cells expressing N-CAM support neurite outgrowth *in vitro* (Doherty et al., 1989).

The same kind of aggregation assay was used with embryonic chicken hepatocytes to identify the liver cell adhesion molecule (L-CAM) (Bertolotti et al., 1980; Gallin et al., 1983). L-CAM also mediates homophilic binding, and is expressed on a variety of cell types (Cunningham, 1988; Edelman and Gallin, 1987; Thiery et al., 1984). Similar molecules were also found in the mouse (Nagafuchi et al., 1987; Hyafil et al., 1980; Ringwald et al., 1987), human (Damsky et al., 1983) and dog (Imhof et al., 1983), and were assumed to be the mammalian equivalents of L-CAM.

N-CAM and L-CAM appear early in embryogenesis in the chicken blastoderm. They are later expressed on derivatives of all three germ layers (Thiery et al., 1984; Thiery et al., 1985). Detailed analyses of the structures and expression of N-CAM and L-CAM revealed several important new concepts about CAMs and their functions: (1) CAMs appear on both embryonic and adult tissues, and the same CAM may appear in a variety of sites and tissues; (2) The levels of expression, distribution, and chemical states of CAMs may be modulated to influence the activity of the proteins; (3) Multiple CAMs may be expressed at the same sites, and on individual cells; (4) Most of the vertebrate CAMs fall into related classes of molecules. These features have dramatically altered earlier notions about cell adhesion and supported the model that relatively few CAMs acting in dynamic patterns can provide the necessary adhesive specificity during embryogenesis. Many





of these same considerations undoubtedly apply to the role of CAMs in the organization and function of the nervous system.

### Structural Families of CAMS

Since the identification of N-CAM and L-CAM, a number of other CAMs or putative CAMs have been isolated and characterized. Surprisingly most of these can be sorted into one of two structural families - one resembling N-CAM and the other resembling L-CAM (Figure 1). This thesis focuses on neural CAMs in the N-CAM family, which are discussed in detail in the following section. Here I present a brief review of the two families and a short description of the neural CAMs in the L-CAM family.

#### *CAMs Containing Ig-like Domains*

N-CAM-like molecules are distinguished by the presence of segments resembling the structural domains of immunoglobulins and usually mediate calcium-independent adhesion. The constant and variable domains of immunoglobulins (Ig) (Edelman, 1970) each contain about 100 amino acids and include two cysteines that form an internal disulfide loop. In variable regions the loop contains about 65-75 amino acids whereas in constant regions it contains 50-60 residues. Specific amino acids around each cysteine are highly conserved in patterns that are distinct for either variable or constant regions. At the three-dimensional level, each domain is folded into two  $\beta$ -pleated sheets with the disulfide bond linking the two sheets (Amzel and Poljak, 1979). In V domains the two sheets contain four and five antiparallel strands, while

Figure 1      Neural Cell Adhesion Molecules.

The cell surface molecules that are known or predicted to mediate cell adhesion, neurite extension, or neurite fasciculation in the nervous system may be divided into at least three families based on shared structural features. References for proteins are found in the text. Immunoglobulin superfamily members listed mediate calcium-independent adhesion, and each contains domains similar to immunoglobulins. Many also contain segments similar to the type III repeating units of fibronectin. Parentheses indicate Ig-like CAMs found in insects. Calcium-dependent CAMs belong to the cadherin family of CAMs. Each member contains four homologous segments and a cytoplasmic region similar in sequence to the other members. A third group contains CAMs whose structural features are different from the first two groups or have not been characterized sufficiently for comparison.

# Neural Cell Adhesion Molecules

- Immunoglobulin Superfamily (Calcium-independent)

N-CAM	Ng-CAM/G4/8D9	Contactin/Fil	MAG
	L1/NILE	F3	P <sub>0</sub>
	Nr-CAM	TAG-I	
(fasciclin II)	(neuroglial)		OB-CAM

- Calcium-dependent

N-cadherin  
L-CAM/E-cadherin/uvomorulin  
T-cadherin  
R-cadherin

- Others

neurofascin	axonin-1
cognin	Astrotactin
fasciclin III	AMOG



in the smaller C-type domains the two sheets contain three and four antiparallel strands (Amzel and Poljak, 1979).

In N-CAM and related molecules, the Ig-like segments have smaller disulfide loops resembling Ig constant domains, but other conserved residues more closely resemble the pattern seen in Ig variable domains. This distinguishes them from other Ig-like molecules that resemble constant regions or variable regions. The Ig-like segments in N-CAM and presumably in other such CAMs are directly involved in cell adhesion events (Cunningham et al., 1987), although the mechanism of such interactions is unknown.

The similarity of segments of N-CAM to Ig domains prompted the hypothesis that the gene for an N-CAM-like precursor gave rise both to a larger family of cell adhesion molecules and to the key molecules of adaptive immunity in vertebrates (Edelman, 1987). The identification of N-CAM-related proteins in insects (Seeger et al., 1988; Harrelson and Goodman, 1988; Bieber et al., 1989) support this hypothesis, because insects do not appear to contain such basic entities of the adaptive immune system as lymphocytes and immunoglobulins. An inducible protein with Ig-like domains is induced by bacterial infection in the silk moth (Sun et al., 1990); such a molecule could represent the beginnings of an adaptive immune system.

Other molecules have been characterized which also contain Ig-like domains similar to N-CAM. These include other neural CAMs, ICAM-1 and ICAM-2 in the immune system and various viral receptors in the vascular system (Staunton et al., 1988; Staunton et al., 1989; Simmons et al., 1988; Mendelsohn et al., 1989), the carcinoembryonic antigens (CEA)



(Oikawa et al., 1987; Zimmerman et al., 1987) and the non-cross-reacting antigens (NCA) (Zimmerman et al., 1988; Neumaier et al., 1988). These molecules form a family in a larger group of proteins collectively called the Ig superfamily (Edelman, 1987; Williams, 1987). The Ig superfamily also includes immunoglobulins and the molecules in the adaptive immune system, the type III kinase growth factor receptors (Yarden et al., 1986), the macrophage Fc receptors (Lewis and Cowan, 1986; Stengelin et al., 1988), and the core protein of a heparan sulfate proteoglycan (Noonan et al., 1988).

#### *Calcium-dependent Neural CAMs*

L-CAM-like molecules are calcium-dependent CAMs or cadherins. Different cadherins are 45-58% identical in amino acid sequence, but differ in their binding specificities and tissue distributions. They all include internal repeats of about 100 amino acids, but do not resemble immunoglobulins (Gallin et al., 1986; Nose et al., 1987; Nagafuchi et al., 1987; Hatta et al., 1988). Moreover, their expression overlaps the expression of the calcium-independent CAMs and may be present on the same cells, but the cadherins do not appear to directly interact with the N-CAM-like molecules (Takeichi, 1990).

The expression of L-CAM (E-cadherin/uvomorulin) in the developing nervous system is very limited. The early neural plate initially expresses L-CAM, but gradually loses it during invagination and acquires another CAM, N-cadherin (Takeichi, 1988). L-CAM/E-cadherin has also been detected at specific sites in the nervous system later in development (Takeichi, 1991).





N-cadherin (also known as adherens-junction specific CAM, A-CAM; Volk and Geiger, 1984) is the predominant cadherin in the nervous system (Takeichi, 1988). It is first detected in the gastrula, and later it is restricted to particular subpopulations of cells derived from all three germ layers. Transfected fibroblasts expressing N-cadherin support retinal neurite outgrowth in culture, suggesting a role in neurite extension (Matsunaga et al., 1988).

Two distinct neural cadherins have recently been identified: T-cadherin and R-cadherin. T-cadherin (truncated cadherin) was described in a variety of neural and non-neural tissues. It is similar to the other cadherins, but lacks cytoplasmic sequences that are highly conserved in the other cadherins (Ranscht and Bronner-Fraser, 1991). T-cadherin was detected in the caudal half of somitic sclerotomes as neural crest cells exited the neural tube through the rostral half, suggesting a role for T-cadherin as a less permissive or inhibitory substrate for neural crest cell migration (Ranscht and Bronner-Fraser, 1991). R-cadherin is distinct from the other cadherins and has recently been detected in the developing retina. It is reported to be expressed in a complementary manner to N-cadherin (Takeichi, 1991).

#### Neural CAMs of the N-CAM Family

A variety of vertebrate CAMs related to N-CAM have been identified in the nervous system and have several distinctive features. The properties of these neural CAMs provide the framework for my studies on Ng-CAM and a related protein Nr-CAM, and are reviewed here.



N-CAM was the first cell adhesion molecule to be identified as such and the first to be characterized in detail. Variants of N-CAM include nearly all of the structural features found in other CAMs in this family and it is therefore valuable to review its features here. N-CAM is expressed in the nervous system as three predominant forms of 180 kD, 140 kD, and 120 kD, resulting from alternatively spliced primary transcripts of a single gene. The two larger forms are transmembrane glycoproteins and differ in the extent of their cytoplasmic domains. The smaller 120 kD form is associated with the cell membrane by a phosphatidylinositol moiety (Hemperly et al., 1986; He et al., 1986). Although N-CAM is detected in both neural and non-neural tissues, the 180 kD component appears to be restricted to the nervous system (Murray et al., 1986). The phospholipid-anchored form first appears in the brain at the time of glial maturation (Hemperly et al., 1986), and variants of this form containing additional short inserts are found in the nervous system in muscle and heart (Prediger et al., 1988; Covault et al., 1986; Dickson et al., 1987). A number of other variant forms of N-CAM have also been described (Goring et al., 1987; Santoni et al., 1989).

The extracellular region of N-CAM contains five Ig-like domains. It also contains two segments similar to the type III repeat units of fibronectin, an extracellular matrix molecule. Many of the other N-CAM-like CAMs also contain these fibronectin-type III (Fn-III) repeats.

N-CAM has one element that has not been detected on any other protein — large amounts of oligomeric N-acetylneuraminic acid (or polysialic acid) in an  $\alpha$ 2-8 linkage (Hoffman et al., 1982; Ocklind and



Öbrink, 1982). The amount of polysialic acid on N-CAM decreases during development with the adult form of N-CAM containing one third as much. The conversion to an adult form, containing less polysialic acid, increases N-CAM's binding affinity (Hoffman and Edelman, 1983; Cunningham et al., 1987; Rutishauser et al., 1985), although the carbohydrate is not directly involved in the binding (Hoffman and Edelman, 1983; Crossin et al., 1984; Cunningham et al., 1987). Changes in the polysialic acid content of N-CAM are also proposed to regulate intramuscular nerve branching during development (Landmesser et al., 1988).

The myelin-associated glycoprotein (MAG) is a cell-surface glycoprotein of the myelin-forming oligodendrocytes in the CNS and Schwann cells in the PNS, where it is associated with the attachment of the myelin sheath to the axon (Sternberger et al., 1979; Martini and Schachner, 1986; Trapp et al., 1984; Quarles, 1984). MAG contains five Ig-related domains but no fibronectin type III repeats (Salzer et al., 1987; Arquint et al., 1987; Lai et al., 1987); nevertheless, it more closely resembles the proteins in the N-CAM family than other Ig-related molecules. Two polypeptide forms of MAG are expressed in the nervous system ( $M_r=67,000$  and  $72,000$ ), generated from alternatively spliced mRNAs. Like N-CAM, the two forms of MAG differ in their cytoplasmic domains (Salzer et al., 1987; Lai et al., 1987) and are differentially expressed during development. While both forms are expressed in the CNS, the smaller form predominates in the PNS. In the CNS, the appearance of the large form peaks shortly after birth and then declines, while the smaller form is first expressed later and becomes



the predominant form in adults (Lai et al., 1987; Frail and Braun, 1984). Liposomes containing MAG bind to both mouse neurons and oligodendrocytes in culture, further suggesting that it mediates neuron-oligodendrocyte and also oligodendrocyte-oligodendrocyte adhesion (Poltorak et al., 1987). Liposomes containing MAG also bind to neurites in dorsal root ganglia or spinal cord cultures (Johnson and McKnight, 1989).

P<sub>0</sub> is the major glycoprotein in peripheral myelin; it is a 28 kD transmembrane protein containing a single extracellular Ig-like domain. P<sub>0</sub> is thought to function in stabilizing the apposition of the extracellular and cytoplasmic surfaces of compact myelin (Kirschner and Ganser, 1980; Lemke and Axel, 1985; Lemke et al., 1988). It mediates homophilic adhesion between transfected cells in suspension (Filbin et al., 1990; Schneider-Schaulies et al., 1990), and concentrates at apposed surfaces with the increased production of desmosome-like structures between transfected cells in monolayer cultures (D'Urso et al., 1990). As a substrate it promotes neurite outgrowth from dorsal root ganglia (Schneider-Schaulies et al., 1990). P<sub>0</sub> also appears to interact with MAG in cells *in vitro*; when expressed in transfected HeLa cells it colocalizes at the focal contacts with adjacent cells transfected with MAG (Doyle et al., 1990).

The mouse glycoprotein L1 is expressed on subsets of postmitotic neurons in the CNS, and on neurons and Schwann cells in the periphery, as well as transiently on crypt cells in the developing intestine (Rathjen and Schachner, 1984; Thor et al., 1987). It is detected primarily as a 200 kD glycoprotein with minor components at 140 and 80





kD (Faissner et al., 1985; Sadoul et al., 1988). The sequence of L1 contains 6 Ig-like domains and 5 Fn-III repeats (Moos et al., 1988). On neurons, L1 is detected predominantly on axons and can mediate neuronal aggregation and axon fasciculation by a homophilic mechanism (Kadmon et al., 1990a; Rathjen and Schachner, 1984; Fischer et al., 1986). It is also involved in the migration of external granule cells in the cerebellum and can mediate the extension of neurites from dorsal root ganglia explants on substrates of other neurites or Schwann cells in culture (Lindner et al., 1983; Seilheimer and Schachner, 1988; Bixby et al., 1988; Kleitman et al., 1988). In addition, L1 has been reported to interact with N-CAM, resulting in enhanced L1-mediated adhesion of neurons *in vitro* (Kadmon et al., 1990a; Kadmon et al., 1990b).

L1 is immunochemically related to the NGF-inducible large external (NILE) glycoprotein in rat (Bock et al., 1985; Sajovic et al., 1986). NILE was originally detected as a 230 kD glycoprotein induced by NGF in the rat PC12 pheochromacytoma cell line (McGuire et al., 1978). It was also isolated from the rat peripheral nervous system as a protein of this size, but is detected in the developing rat CNS as slightly smaller forms of 200-210 kD (Salton et al., 1983a; Salton et al., 1983b). A partial amino acid sequence for NILE has been reported, and is 97% identical to L1 over a large part of the cytoplasmic domain, suggesting that they are in fact equivalent in the two species (Prince et al., 1989). The distribution of NILE in the rat is similarly restricted to postmitotic neurons of the nervous system, where it is localized to developing neuronal fiber tracts (Salton et al., 1983a;



Stallcup et al., 1985). It is also detected on Schwann cell surfaces in the periphery (Salton et al., 1983a). Antibodies to NILE also inhibited the fasciculation of neurites in cultures of embryonic rat brain cells, but not in cultures of postnatal cerebellum (Stallcup et al., 1985).

Several N-CAM-like molecules have been identified that are not transmembrane proteins, but are probably anchored to the membrane via a phosphatidylinositol anchor. These include contactin, F11, F3, and TAG-1. F11 was identified in embryonic chick brain as an antigen involved in retinal axon fasciculation (Rathjen et al., 1987a). The predicted sequence from cDNA analysis contains 6 Ig-like domains, 4 Fn-III repeats, and a carboxyl segment resembling the consensus sequence for attachment of a glycopospholipid (Brunet et al., 1989). It is probably identical to the chicken protein contactin, which was isolated on the basis of its limited extractability from neuronal cell membranes. Contactin's predicted amino acid sequence contains six Ig-related domains, two fibronectin-type III repeats, a transmembrane segment and a small cytoplasmic domain (Ranscht, 1988). The sequence of F11 is essentially identical, but contains two frame shifts that alter the reading frame from that of contactin; as a result F11 appears to have a carboxyl terminal segment resembling the attachment site for glycopospholipids (Ferguson, 1988; Low and Saltiel, 1988). In accord with this notion, 40% of F11 expressed by cells is released from membranes by phospholipase C suggesting that it is linked to the membrane via phosphatidylinositol (Wolff et al., 1989; Gennarini et al., 1989). The observed difference between the two proteins may reflect two forms of the same protein, or be may be attributable to



artifacts in the cDNA or sequencing. For the purpose of sequence comparisons, these two proteins are jointly referred to in this thesis as contactin/F11. The murine equivalent of contactin/F11, called F3, has been characterized and sequenced; F3 is 81% identical to contactin/F11 in the extracellular region and at least 40% of it can be released from membranes by phospholipase C (Gennarini et al., 1989).

Another phospholipid-anchored protein, TAG-1, is transiently detected on subsets of motor and commissural axons in the developing rat spinal cord (Dodd et al., 1988). It is closely related to the contactin/F11 sequence (49% identical), and contains six Ig-like domains and five Fn-III repeats (Furley et al., 1990). However by comparison to the chicken and mouse sequences it clearly appears to be a different protein. TAG-1 is released from axonal surfaces and may play a role in neurite extension or guidance, but does not appear to be involved in fasciculation (Dodd et al., 1988). Interestingly, TAG-1 and L1 both appear to be present but segregated onto different segments of the same embryonic spinal cord axons. In contrast, both CAMs are coincidentally expressed in developing dorsal root ganglia (Dodd et al., 1988).

A 58 kD opioid-binding protein from bovine brain tissue has been characterized by cDNA cloning (Schofield et al., 1989). Its deduced sequence contains three Ig-like domains similar to N-CAM, and a hydrophobic carboxyl terminal sequence that may indicate attachment to the cell surface by a phosphatidylinositol anchor. Antibodies raised to a peptide from the deduced sequence were used to immunoaffinity purify a protein fraction. Materials in this protein fraction bound opioids, and the opioid binding was inhibited by antibodies raised to the purified



protein. Based on its sequence similarities to the N-CAM family, the protein was predicted to have an adhesive function and has been called an opioid-binding protein-cell adhesion molecule (OBCAM) (Schofield et al., 1989).

#### *Neural CAMs in Invertebrates*

Immunoglobulin-like CAMs have also been detected in the nervous systems of invertebrates. In grasshopper and in *Drosophila*, three proteins (fasciclins I, II, and III) have been described on subsets of commissural and longitudinal axon fascicles in the embryo. They appear in a polarized manner on individual fascicles as the neurites extend toward their target tissues, and are implicated in the selective adhesion of neuronal membranes expressing them, as well as in neurite guidance (Harrelson and Goodman, 1988; Zinn et al., 1988; Patel et al., 1987). The sequence of fasciclin II most closely resembles N-CAM, and contains 5 Ig-like domains and 2 Fn-III repeats (Harrelson and Goodman, 1988). Fasciclin III is a more distantly related member of the Ig superfamily, containing fewer conserved features of the Ig-like domain and lacking the characteristic disulfide bonds (Grenningloh et al., 1990).

Neuroglian in *Drosophila* is most similar to L1 in the mouse. It is expressed as two alternatively spliced forms, both of which contain six Ig-like domains and five fibronectin-type III repeats (Bieber et al., 1989). The two forms differ in the size of their cytoplasmic domains. Although neuroglian is widely expressed in the developing embryo (including the nervous system, salivary glands, trachea, hindgut,





muscle, and ectodermal tissues) (Bieber et al., 1989), the larger form is restricted to subsets of neurons in the nervous system (Hortsch et al., 1990). Neuroglian expressed by transfected cells mediates homophilic binding (Grenningloh et al., 1990).

The sequence of a gene near the antennapodia complex in *Drosophila* encodes a protein with similarities to N-CAM (Seeger et al., 1988). The protein was named amalgam, and was predicted to contain a signal sequence, three Ig-like domains, and a short hydrophobic carboxy-terminal region. Expression of the RNA transcript for amalgam was developmentally regulated, reaching peaks at mid-embryogenesis and during early pupation, and was absent in adult animals. Antisera were raised to a fusion protein of amalgam, created by a  $\beta$ -galactosidase-amalgam gene fusion. The antibodies stained all cell bodies and axons of CNS neurons, as well as subpopulations of mesodermal cells. Its similarity to neural CAMs suggests that amalgam has a cell binding function, although such a function has not yet been demonstrated. In deletion mutants no gross changes in the development of the CNS can be specifically attributed to the absence of the gene for amalgam or for fasciclin I and III (Seeger et al., 1988; Elkins et al., 1990).

#### Other Neural CAMs

Several known or putative CAMs have been identified in the nervous system whose structures have not been sufficiently well characterized to relate them to the N-CAM or L-CAM families of CAMs. These include astrotactin, neurofascin, axonin-1, and Bravo. In addition, an adhesion molecule on glia (AMOG) has been found on mouse cerebellar astrocytes



and Bergmann glial fibers. Monoclonal antibodies to AMOG blocked neuron-astrocyte adhesion by 25% (Antonicek et al., 1987), but AMOG's deduced protein sequence is 40% identical to the  $\beta$  subunit of Na,K-ATPase (Gloor et al., 1990), suggesting that it may perform a function other than cell adhesion.

Astrotactin is a ~100 kD protein expressed transiently during chicken cerebellar development on migrating granule cells (Edmondson et al., 1988). Antibodies to astrotactin inhibited granule neuron membrane binding to astrocytes *in vitro* by 70%, suggesting that this neuronal protein is involved in granule neuron-glia binding (Stitt and Hatten, 1990).

Neurofascin is the name for a group of apparently related neural antigens detected in the chicken (185 kD, 160 kD, 135 - 110 kD, and 92 kD). It is found mainly on fasciculating axons in the embryonic spinal cord, cerebellum and retina. It has also been detected in the cerebrospinal fluid and vitreous humor as a mixture of components (Stoeckli et al., 1991). Antibodies to neurofascin perturbed retinal axon bundling *in vitro* and the extension of sympathetic neuron growth cones on sympathetic axonal surfaces *in vitro* (Rathjen et al., 1987b). The sequence of neurofascin has not yet been determined. Antibodies to the protein did not cross-react with Ng-CAM or F11, but its distribution and function suggest that it may be a related protein (Rathjen et al., 1987b).

Axonin-1 may also play a role in neurite fasciculation in the developing chick embryo (Ruegg et al., 1989a; Stoeckli et al., 1989).



It occurs both as a membrane-bound and a secreted protein, each of which is apparently derived from a different intracellular pool. It was detected as a number of variant forms of  $M_r$  = 132-140 kD (Ruegg et al., 1989b; Ruegg et al., 1989c; Stoeckli et al., 1989), and was localized by immunostaining to developing nerve fiber tracts. When solubilized and added to the media the membrane-bound form of axonin-1 inhibited the fasciculation of neurites in dorsal root ganglia cultures (Ruegg et al., 1989c). The released form also inhibited fasciculation of neurites in dorsal root ganglia explants in culture, possibly by competing with the membrane-bound form on neurites (Stoeckli et al., 1991). The released form was reported to be more abundant and provided a suitable substrate for neurite outgrowth (Stoeckli et al., 1991). Its size, localization and release from axons suggest that axonin-1 may be similar or identical to TAG-1 or contactin/F11.

Recently a novel glycoprotein called Bravo was found in a topographically restricted pattern in the developing retinotectal system of the chicken (de la Rosa et al., 1990). Identified as two components of 130 kD and 140 kD, Bravo was first detected simultaneously on all the fiber layers of the retina and on the optic tectum. While Bravo was found on the optic fibers in the retina, it was *not* found on retinal fibers in the tectum, suggesting that it is involved in axonal outgrowth or guidance.

Intriguingly, Bravo expression on outgrowing retinal fibers *in vitro* was regulated by the substrate on which the fibers grew. Fibers growing out from retinal explants on retinal basal lamina expressed the protein, but similar fibers growing onto collagen gels did not. Bravo



contains the HNK-1 carbohydrate epitope which appears on several other CAMs, but the amino terminal protein sequence reported for Bravo was distinct from known CAMs. This sequence, however, is 74% identical to the amino terminal sequence of a new Ig-like chicken protein (Nr-CAM) described in this thesis. A comparison of the similarities or potential identity between these two proteins is addressed in the Nr-CAM chapter (Chapter 4).

### Ng-CAM

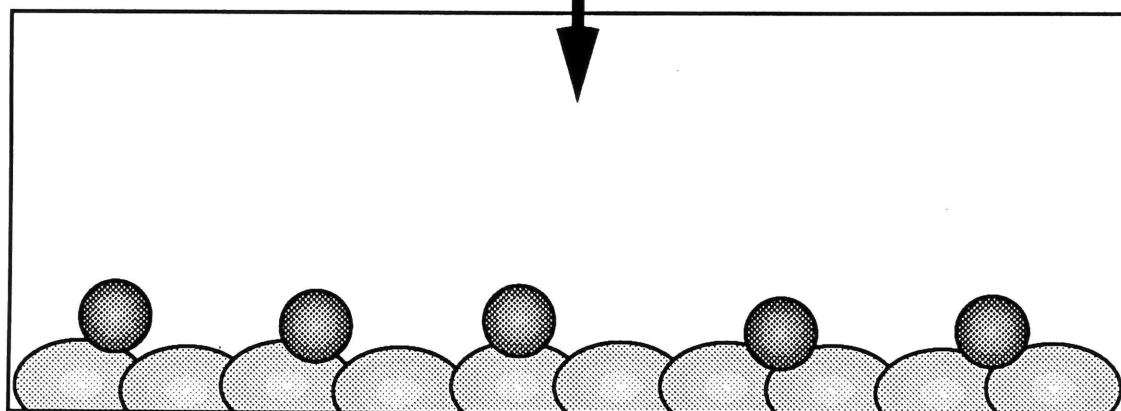
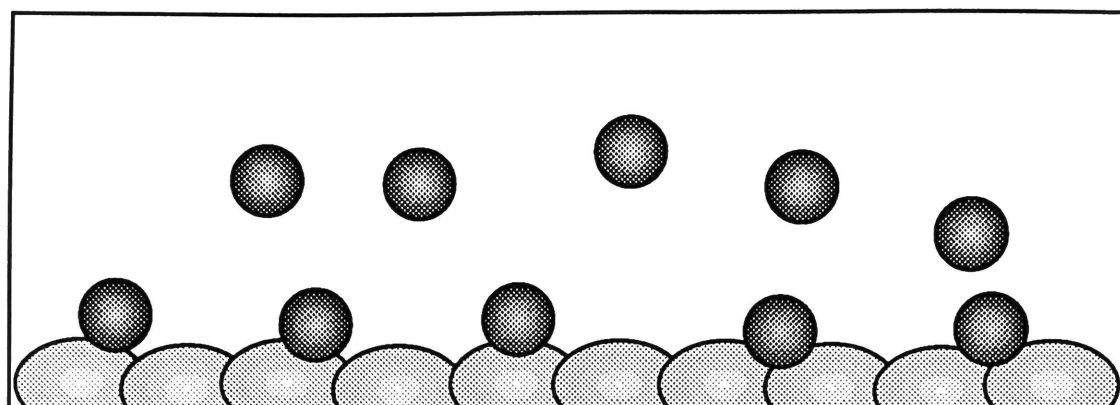
Ng-CAM is important in neural development in that it participates in both neuron-neuron and neuron-glia interactions. It is dynamically expressed on subsets of cells and concentrates on extending neurites, suggesting that Ng-CAM may have a special role in cell recognition. It also shares several features with other known neural CAMs. The main goal of my thesis research was to characterize Ng-CAM in greater structural detail, both to understand the relationship between its protein components and its function and to relate it to other neural CAMs.

Ng-CAM was identified in the early 1980's while searching for a CAM that mediated neuron-glia adhesion. Because two cell types were involved, the assay used to identify N-CAM was adapted to the binding of dissociated neurons or neuronal membrane vesicles to glial monolayers (Grumet et al., 1983; Grumet and Edelman, 1984) (Figure 2). Suspensions of radioactively-labeled or fluorescently-labeled embryonic neuronal membrane vesicles were incubated with monolayers of glia. After rinsing away nonadhering vesicles the amount of vesicle binding

Figure 2      Monolayer Adhesion Assay for Neuron-Glia Binding.

Ng-CAM was initially characterized in an in vitro assay measuring the binding of brain membrane vesicles (darkly shaded circles) to monolayers of glia (light ovals). After incubation of labeled vesicles in suspension with confluent monolayers of glia (upper panel) nonadhering vesicles were rinsed from the dishes (middle panel). Parallel assays were performed in the presence of monovalent Fab' fragments isolated from preimmune antisera, anti-brain membrane antisera or specific anti-N-CAM antisera. Adhering vesicles were scored, and the percentage of inhibition attributed to antibodies relative to maximal binding (preimmune Fab') was calculated.





Fab'	Number of Cells Bound/mm <sup>2</sup>	Inhibition by Fab' (%)
Preimmune	110	—
Anti-Brain	52	53
Anti-N-CAM	118	0



was measured as remaining radioactivity (or fluorescence). Polyclonal antibodies were raised against neuronal cell membranes. When monovalent Fab' fragments of the antibodies were added to the assay mixture, they inhibited the neuronal vesicle-glia binding by 50%. In contrast Fab' fragments of antibodies specifically raised against N-CAM had no effect. This indicated that the neuron-glia ligand recognized by the antibody was present only on the neurons or their membrane vesicles (Grumet and Edelman, 1984).

The released material from the trypsinized membrane vesicles was fractionated, and a protein was partially purified which neutralized the antibody's inhibition of binding. This protein was injected into mice to produce monoclonal antibodies, one of which was able to remove the neutralizing activity from the protein fraction, and by itself was able to inhibit the neuron-glia adhesion (Grumet et al., 1984a). The protein recognized by the monoclonal antibody was purified on an immunoaffinity column. The purified material was designated the neuron-glia cell adhesion molecule (Ng-CAM).

Ng-CAM and N-CAM are distinct molecules by a variety of structural, functional, and immunological criteria (Grumet et al., 1984a; Grumet et al., 1984b; Thiery et al., 1985; Hoffman et al., 1986; Chuong and Edelman, 1984). While several antibodies cross-react between the two proteins, they apparently react to a shared carbohydrate epitope (Grumet et al., 1984b; Grumet and Edelman, 1988). Ng-CAM does not contain the polysialic acid detected on N-CAM. While N-CAM is expressed outside the nervous system, Ng-CAM is detected in the chicken only in the nervous system, where it is found on postmitotic neurons in



the CNS and on neurons and Schwann cells in the periphery (Thiery et al., 1985; Rieger et al., 1986). Within the nervous system the distributions of Ng-CAM and N-CAM differ substantially (Hoffman et al., 1986; Daniloff et al., 1986a) , suggesting that the two molecules are involved in different functions.

N-CAM first appears early in development and on cells derived from all three primary germ layers. In the nervous system it is expressed generally on all neurons and their processes, as well as oligodendrocytes and cerebral glial cells (Bhat and Silberberg, 1986; Nybroe et al., 1985). N-CAM has also been detected on several glial cell lines (Williams et al., 1985). In contrast, Ng-CAM first appears during neural development and its expression is restricted to subsets of neurons and to Schwann cells at particular times (Thiery et al., 1985; Rieger et al., 1986). Ng-CAM and N-CAM may be expressed on the same neurons, but their relative contributions to function appear distinct in at least several sites. In the retina antibodies to N-CAM inhibit cell layering, while antibodies to Ng-CAM have no effect. Conversely, in the cerebellum, antibodies to Ng-CAM inhibit granule cell migration, but antibodies to N-CAM had no effect on this behavior (Hoffman et al., 1986).

The distribution of Ng-CAM on neuronal cell surfaces appears polarized as development progresses, becoming more prevalent on outgrowing axons than on the soma and dendrites (Thiery et al., 1985). The levels of Ng-CAM expression are also modulated in the chicken. It is most prevalent in the chick at approximately embryonic day 14, a time at which the nervous system is rapidly developing. Its expression



diminishes as development progresses, although Ng-CAM is still detected in regions of the adult nervous tissue at low levels (Daniloff et al., 1986a).

#### *Distribution and Function of Ng-CAM in Development*

Antibodies raised to Ng-CAM have been used to determine its localization and expression at various stages of the developing chick nervous system. Immunofluorescent staining with Ng-CAM specific antibodies indicate that Ng-CAM is expressed dynamically and transiently on particular subsets of neurons, correlating with periods of cell migration and neurite extension (Thiery et al., 1985).

Neural development may be divided into two phases: in the first, the neural plate becomes segregated and undergoes induction to form the neural tube, becoming the beginning of the neural axis; in the second phase (histogenesis) specialized structures form upon which future functional activity will depend. The specific expression and modulation of Ng-CAM during this second phase highlight its importance in neural development. However, it is clear that a single neuron may express a variety of CAMs and the expression and activity of any one CAM must be considered in the context of other molecules expressed on that neuron, their relative amounts and distributions, and the patterns of CAMs expressed by neighboring cells.

Prior to neurulation, both N-CAM and L-CAM are expressed in blastoderm cells and then in cells of the ectoderm. As neurulation begins N-CAM is concentrated in the region of the neural plate and disappears in the lateral ectodermal cells. L-CAM conversely is lost in





the neural plate region and remains in the lateral ectoderm. Another calcium-dependent CAM, N-cadherin, is expressed in the neural tube. The neural tube then invaginates to form a cylinder of neuroepithelial cells, the neural tube. These cells are N-CAM positive but L-CAM negative (Edelman, 1985b; Thiery et al., 1984; Takeichi, 1988). During the formation of the nervous system N-CAM is ubiquitously expressed on all neurons and neurites, although at specific locations its expression varies quantitatively and different forms are detected.

Ng-CAM is first seen in the ventral cells of the neural tube at the 31 somite stage, on the developing neurites that will become the white matter of the spinal cord. It is specifically localized and dynamically changes on neurites and migrating neurons in different areas. As neuronal layers and aggregates begin to form and send out neurites to form connections, Ng-CAM is expressed transiently on the neuronal somata and then is detected on the extending neurites (Thiery et al., 1985; Daniloff et al., 1986a). As the connections stabilize and myelination occurs, the levels of Ng-CAM decrease dramatically and become localized to the nodes of Ranvier. At the time of myelination, two other CAMs are expressed on the glial cells, P<sub>0</sub> and MAG (Quarles, 1984). On nonmyelinated axons, Ng-CAM remains uniformly present along the entire surface (Rieger et al., 1986).

Ng-CAM is expressed on neurites and mediates fasciculation. For example, antibodies to Ng-CAM inhibit the fasciculation of neurites extending from dorsal root ganglia explants in culture. Fasciculation is not inhibited by antibodies to N-CAM (Hoffman et al., 1986). Ng-CAM mediates neuron-neuron adhesion, in an N-CAM independent manner.



Antibodies to Ng-CAM inhibit the aggregation of dissociated neurons *in vitro* and antibodies against Ng-CAM also inhibit the binding of Ng-CAM immobilized on Covasphere beads (or liposomes reconstituted with Ng-CAM) to neurons *in vitro*. Preincubation of the Ng-CAM-coated beads or the neurons with the Ng-CAM antibodies inhibit the binding as well, indicating that Ng-CAM-mediated neuron-neuron binding is probably homophilic (Ng-CAM on one surface binds to Ng-CAM on the other surface) (Hoffman et al., 1986; Grumet and Edelman, 1988).

Ng-CAM is detected primarily on neuronal fibers during development, but it is also expressed transiently on migrating neuronal cell bodies. This is evident in the gray matter of the spinal cord and in the external granule cells of the cerebellum (Grumet et al., 1984b; Thiery et al., 1985) as they migrate to deeper layers along radial glial fibers (Rakic, 1971). Ng-CAM expressed on these migrating neurons is functional; in explant cultures of developing cerebellum, anti-Ng-CAM antibodies blocked the migration of the granule cells into the underlying molecular layer (Friedlander et al., 1986). But Ng-CAM mediated binding of neurons to glia is likely to be a heterophilic interaction of Ng-CAM on the neuronal cells binding to an as yet unknown ligand on the glial fiber similar to the neuron-glia interactions by which Ng-CAM was initially discovered. It cannot be ruled out, however, that in the cerebellar explants the Ng-CAM antibodies blocked homophilic Ng-CAM interactions on the migrating cell bodies. This could have effects on other CAM-mediated events which might be blocked in this migration.



In the adult, Ng-CAM is expressed at lower levels in specific areas, such as the local circuit fibers of the molecular layer in the cerebellum, the olfactory bulb, unmyelinated fibers in the CNS, and in peripheral nerves (Thiery et al., 1985). The staining for N-CAM is also reduced in the adult but continues to be more uniform on all neurons. N-CAM also undergoes a conversion from the embryonic highly sialylated form to a less sialylated adult form. This occurs at different rates in the different regions (Chuong and Edelman, 1984). Ng-CAM has no detectable amounts of polysialic acid (Grumet et al., 1984a).

The expression of both Ng-CAM (and N-CAM) is modulated during nerve injury and repair in chickens and in mice. When the sciatic nerve is cut or crushed, the level of Ng-CAM (and N-CAM) increases around the site of the lesion, and N-CAM returns to its embryonic form. Similar increases in the expression of both CAMs are seen after cutting nerves from the dorsal root ganglia or in the spinal cord. Of the two CAMs only N-CAM is present on muscle, and is localized in the adult muscle to the sites of neuromuscular junctions (Daniloff et al., 1986b).

#### *Structure of Ng-CAM*

Ng-CAM is detected in the chicken as a predominant 135 kD component, with minor components of 80 kD and 200 kD (Grumet et al., 1984b). The relative amounts of each are the same in all regions of the nervous system that have been examined. The component at 200 kD is often resolved as a closely spaced doublet of 190 kD and 210 kD and the two components are usually detected in equal amounts.



A variety of data suggested that the 135 kD and 80 kD components are derived from the 200 kD species by proteolysis. Antibodies raised against the 135 kD component recognized both the 135 kD and 200 kD species in immunoblots, but not the 80 kD component. Conversely, antibodies against the 80 kD component recognized both the 80 kD and the 200 kD components, but not the 135 kD component (Grumet et al., 1984b). In addition, phosphate and palmitate were incorporated into the 80 kD and 200 kD components, but not the 135 kD component (Sorkin et al., 1985; Grumet et al., 1984b). A goal of my thesis was to establish whether these components were derived from a single translation product or were synthesized from separate mRNAs.

#### *Relationship of Ng-CAM to Other CAMs*

A key goal of my studies was to establish the relationship of Ng-CAM to the other neural CAMs. There are several other chicken CAMs described in the literature that are likely to be identical to Ng-CAM. The 8D9 antigen has polypeptide components identical to Ng-CAM and it is distributed in the visual pathway in the same manner (Lemmon and McLoon, 1986). The 8D9 antigen is also involved in neuron-neuron binding, neuron-glia binding, and is a potent substrate for neurite outgrowth (Lemmon et al., 1989; Drazba and Lemmon, 1990; Lagenaur and Lemmon, 1987). G4 also has components of the same size as Ng-CAM, and antibodies raised to G4 inhibit neurite fasciculation (Wolff et al., 1987; Chang et al., 1987). Neurofascin is also involved in neurite fasciculation in the chick, but it is immunologically distinct from Ng-CAM (Rathjen et al., 1987b).





Ng-CAM shares many features with two rodent proteins, L1 in mouse and the NGF-inducible large external (NILE) glycoprotein in the rat. As indicated above L1 and NILE appear to be equivalent to each other (Rathjen and Schachner, 1984; Salton et al., 1983a; McGuire et al., 1978; Bock et al., 1985), but their relationship to Ng-CAM is less clear.

Ng-CAM, L1, and NILE have similar anatomical distributions and some antibodies to Ng-CAM and L1 cross react with each other and to NILE. For example, antibodies raised against Ng-CAM immunoprecipitated NILE from extracts of rat PC12 pheochromacytoma cells and were used to demonstrate that expression of that molecule is enhanced by NGF (Friedlander et al., 1986; McGuire et al., 1978). Nevertheless, the detailed distribution of Ng-CAM and L1 differ at several well-characterized sites. L1 is present in the crypt cells of the developing intestine (Thor et al., 1987), whereas Ng-CAM is not detected there (Thiery et al., 1985). In the developing spinal cord a group of commissural circumferential interneurons have been carefully examined in the chick and in the rat for expression of Ng-CAM and L1 respectively, and show clear differences. In the chicken these interneurons are among the first primitive neurons to express Ng-CAM at approximately 2.5 days of development. Ng-CAM is expressed early on the surface of these interneurons as they first project ventrally toward the floorplate at the midline, cross the midline, and course rostrocaudally in the contralateral longitudinal fascicle (Shiga et al., 1990). This is in marked contrast to the same interneurons in the developing rat spinal cord. In the rat, these cells first express a different CAM, TAG-1, as



they project ventrally in the spinal cord. The cells then express L1 only after their neurites cross the midline, and the L1 expression is restricted to the contralateral segments of the neurites in the longitudinal contralateral fascicle (Dodd et al., 1988).

Chicken Ng-CAM and mouse L1 are both detected as components that include a 200 kD doublet, a 135-140 kD component, and a 80 kD component (Grumet et al., 1984b; Rathjen and Schachner, 1984; Faissner et al., 1985). However, L1 is detected as a predominant 200 kD doublet, while the predominant form of Ng-CAM is the 135 kD component. NILE is less well characterized, but is detected in the CNS as a component of ~200 kD.

Ng-CAM and L1 are involved in several similar functions, including neuron-neuron adhesion, neuronal migration in the cerebellum, and neurite fasciculation (Grumet et al., 1984a; Hoffman et al., 1986; Lindner et al., 1983; Rathjen and Schachner, 1984). But Ng-CAM mediates heterophilic neuron-glia adhesion, while L1 does not (Grumet and Edelman, 1984; Keilhauer et al., 1985). L1 has been reported to interact with N-CAM on cell surfaces, resulting in enhanced L1-mediated binding (Kadmon et al., 1990a; Kadmon et al., 1990b). Such interactions have not been demonstrated for Ng-CAM.

These comparisons raised the possibility that Ng-CAM and L1 may be functionally and structurally distinct molecules in the two species. This notion is supported by the more detailed comparison of the structure of Ng-CAM to L1 and NILE in my studies.



## Agenda of the Thesis

In order to analyze the chemical structure of Ng-CAM, discern the relationship of the components, and determine its relationship to other neural CAMs, I isolated and characterized cDNA clones which represent the entire coding sequence of Ng-CAM. In addition more than 10% of the amino acid sequence of the polypeptides were determined. The accumulated sequence data as well as RNA and DNA analyses support the notion that the various components of Ng-CAM are all derived from a single mRNA. The deduced protein sequence indicates that Ng-CAM and Ll may not be equivalent molecules in the chicken and mouse, although they are closely related.

Another protein has been discovered in the embryonic chick brain, Nr-CAM. It is equally similar to Ng-CAM and Ll and suggests a distinction between structural subgroups of neural CAMs. The increasing numbers of N-CAM-like adhesion molecules in the nervous system and their related or overlapping functions suggest that multiple CAMs might be involved in selective neuronal recognition, neurite fasciculation and adhesion, or neurite guidance during development. The identification of Nr-CAM as a new member of this family of neural proteins makes it a logical candidate for such roles. The thesis concludes with a discussion of the subgroups of N-CAM-like molecules defined by these proteins and their possible roles in neural development.



Animals

White Leghorn chickens and embryos were used in experiments.

Ng-CAM Purification

Ng-CAM was purified from membrane preparations of 14-day chicken embryo brains by affinity chromatography using monoclonal anti-Ng-CAM antibodies (Grumet et al., 1984a). Brains were freshly dissected from embryonic chickens and homogenized in ice-cold extraction buffer [0.5% NP40 and 1mM EDTA (pH 7.0) in 1X phosphate buffered saline (PBS), containing 250U per ml of Trasylol (Mobay Chemical Corporation) and 35 $\mu$ g per ml PMSF (Sigma)] at a ratio of 2ml buffer per 1ml tissue and the homogenate cleared by centrifugation in a Sorvall SS34 rotor for 10 min. at 15krpm and 4°C. Anti-Ng-CAM 10F6 and 16F5 monoclonal antibodies (Grumet et al., 1984a) previously coupled to Sepharose CL-2B beads by the cyanogen bromide activation method (Cuatrecasas and Anfinsen, 1971) were mixed with the supernatant and shaken at 4°C for 4 hr. The immunoadsorbent was collected in a column and washed with 10-15 column volumes of extraction buffer. N-CAM was eluted with 0.5% NP-40, 0.05M diethylamine, 1mM EDTA (pH 11.5) and the eluate was neutralized by the addition of 0.1 volumes of 1M potassium phosphate (pH 6.8).

The column was washed to neutrality with 5-10 volumes of extraction buffer. To remove trace contaminants, the immunoadsorbent was removed from the column, added back to the affinity-purified N-CAM and the column procedure repeated. Detergent was removed by incubation





with one fourth volume of BioBeads SM-2 for 1 hr. at 4°C (Holloway, 1973) followed by dialysis against two changes of 12 liters of water for 4 hr. each at 4°C. Intact Ng-CAM and V8-protease digests of the molecule were resolved on 6% acrylamide gel, and stained with Coomassie Blue (Laemmli, 1970; Cleveland et al., 1977).

For sequence analysis Ng-CAM or the digests were resolved on 6% polyacrylamide gels and electroeluted onto Immobilon polyvinylidene difluoride membranes (Millipore Corp., Bedford, MA). The membrane was washed in deionized H<sub>2</sub>O for 5 minutes, stained with Coomassie Blue R-250 in 50% methanol for 5 minutes, and then destained in 50% methanol/10% acetic acid for 5-10 minutes at room temperature (Matsudaira, 1987). The membrane was finally rinsed in deionized H<sub>2</sub>O for 5-10 minutes. The individual components of Ng-CAM (see Figure 3) were excised from the membrane and sequenced.

#### Ng-CAM Peptides

Ng-CAM containing all four components was treated with CNBr (Waxdal et al., 1968) and the fragments were initially separated by gel filtration on Sephadex G-25. After further separation on a C<sub>3</sub> column by HPLC (0-60% acetonitrile gradient in 0.1% trifluoroacetic acid), fragments that were homogeneous as assessed by SDS-PAGE and silver staining (Morrissey, 1981) were sequenced. A 40 kD CNBr fragment was further digested with trypsin and the peptides were fractionated by HPLC (C<sub>3</sub> column, 0-60% gradient of acetonitrile in 0.1% trifluoroacetic acid); 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 (Cleveland et al., 1977). The resulting peptides were separated by SDS-PAGE on 15% acrylamide gels, transferred to Immobilon and sequenced.

The peptide sequences obtained from Ng-CAM are shown in Table 1. 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 in an Applied Biosystem Gas Phase Sequencer.

#### SDS-Polyacrylamide Gel Electrophoresis (SDS-PAGE) and Transfer

SDS-PAGE was performed as described by Laemmli (Laemmli, 1970). 5X Laemmli sample buffer was added directly to the 25-40  $\mu$ l containing 5-50  $\mu$ g protein samples, and electrophoresed in 6% polyacrylamide gels at 40mA (Ausubel et al., 1989).

If gels contained  $^{32}$ P-radiolabelled samples, they were strained in 0.15% Coomassie Blue (Sigma) for 5 min. and destained in 10% methanol and 10% acetic acid for several hours. After drying onto filter paper, the gels were exposed to XAR-5 autoradiographic film (Kodak). Gels containing non-radiographic samples were transferred to nitrocellulose (Schleicher and Schuell) in Laemmli SDS-PAGE buffer containing 20% methanol. This was done by placing the gel-nitrocellulose sandwich between two electrode plates submerged in buffer and applying current using a 12 volt/6 amp DC battery (Schumacher Electric Co., Chicago, IL) for 4 hr. The nitrocellulose was subsequently stained with amido black to reveal the molecular weight standards. Molecular weight standards

Table 1      Ng-CAM Peptide Sequences.

Amino terminal sequences of purified Ng-CAM protein components or peptide fragments obtained after cleavage were sequenced by sequential Edman degradation. The component or cleavage method for obtaining the peptide fragment is indicated. The sequence is listed in single letter amino acid code, beginning at the amino terminus. X represents an ambiguous position which could not be clearly identified from the sequence data.

TABLE 1 - PEPTIDE SEQUENCES

<u>Number</u>	<u>Origin of Fragment</u>	<u>Sequence</u>
1	135kD NH <sub>2</sub>	ITIPPEYGAHDFLQPPELTEEPPEQLVVFP
2	190kD NH <sub>2</sub>	XTIPPEYGAHDF
3	210kD NH <sub>2</sub>	XTIPPHYGAXDFL
4	CNBr	GQDGNLYFSNA
5	CNBr	VQDSHPDYIGHAFLGPRTIIQKEPLDL
6	tryptic	GGSVVLECIAGLPTP
7	tryptic	LWGVTESDDGEYECVAENG
8	tryptic	GTHSVTSEAAPYWVRA
9	CNBr	LTAADQRYEVTEXQTVFLHCRTFGAPA PNVEWLTPTLEGA
10	CNBr	EVDEHRVRLSWTPGDHXNSPIEFV
11	80kD NH <sub>2</sub>	XPPDPPQIPXSPAEDPPPPFFPSA
12	V8- 80 kD	GVPGPPEELRVERLDDTALSV
13	V8- 80 kD	VQFMNKSTDEPWRTSGRANSSL
14	V8- 80 kD	GLRPGTAYRVQFVG



included myosin, Mr=205,000; b-galactosidase, Mr=116,000; phosphorylase b, Mr=97,000; bovine serum albumin, Mr=66,000; and ovalbumin, Mr=45,000.

### RNA Purification

Total RNA was prepared from embryonic chicken brains by rapid sonication of freshly dissected tissue in a Polytron homogenizer either in: (a) 4M guanidinium-thiocyanate (MacDonald et al., 1987); or (b) 3M LiCl/6M urea/ 0.2% SDS (Geliebter, 1987).

(a) Freshly dissected brains or livers were homogenized in a Polytron homogenizer in 4 guanidinium thiocyanate (Fluka)/0.5% Sarkosyl/25 mM sodium citrate (pH 7.0) for 45 seconds at maximum speed (2g tissue/20ml buffer). Large debris was removed by centrifuging in a clinical centrifuge at 200xg for 10 minutes. The supernatants were loaded onto 2.5 ml cushions of 5.7 M CsCl/60 mM EDTA (pH 7) in Beckman SW40 ultracentrifuge tubes [pretreated with 0.01% diethylpyrocarboate (DEPC)]. After centrifugation at 35 k rpm for 16 hours at 20°C, the RNA pellet was rinsed with -20°C ethanol and partially dried with a lyophilizer. The RNA was then redissolved in 1 ml DEPC-treated H<sub>2</sub>O. Its concentration was determined by absorption at 260 nm and 280 nm.

(b) Freshly dissected tissues were immediately placed in liquid nitrogen, and pulverized with mortar and pestle (pretreated with 0.01% DEPC) while frozen. Frozen tissue powder was homogenized upon the addition of ice cold 3M LiCl/6M urea/0.2% SDS with a Polytron homogenizer for 60 seconds at maximum speed (2 ml tissue/25 ml buffer. RNA was precipitated overnight with 0.5 volumes 7.5 M NH<sub>4</sub>OAc/2.5 volumes ethanol at -20°C. After centrifugation in a Sorvall SS-34 rotor at 12k





rpm for 30 minutes at 0°C, the RNA pellet was resuspended in 1.0 ml DEPC-treated H<sub>2</sub>O.

Poly (A)<sup>+</sup> RNA was selected on oligo (dT)-cellulose (Collaborative Research Inc., Bedford MA) in 10-ml RNase-free columns (Ausubel et al., 1989).

#### RNA and DNA Transfer Blots

RNA transfer blots and DNA transfer blots were performed using standard techniques (Maniatis et al., 1982). For Northern analyses, 2 µg Poly(A)<sup>+</sup> RNA from 12 day embryonic brain or liver, or 10 µg total RNA from 9 day and 12 day embryonic brain, liver, heart, gizzard, and skin, were resolved on 0.7% agarose gels containing 1% formaldehyde. All restriction enzymes were purchased from BRL (Gaithersburg, MD). For genomic Southern analyses, chicken liver genomic DNA (Clontech, Palo Alto, CA), was digested to completion with appropriate restriction endonucleases and resolved on individual lanes of 0.7% agarose gels (Maniatis et al., 1982). Nucleic acids were transferred to Hybond (Amersham, Arlington Heights, IL) or Genescreen (New England Nuclear, Cambridge, MA) the membranes were fixed, and hybridized to <sup>32</sup>P-labeled cDNA probes (Maniatis et al., 1982).

#### Library Construction

cDNA libraries were constructed in λgt11 from total RNA or poly (A)<sup>+</sup> RNA isolated from 9-day to 14-day embryonic chicken brains. cDNA was synthesized by the RNase H method (Gubler and Hoffman, 1983) using oligo (dT) or synthetic oligonucleotides as primers. cDNA synthesis



kits from BRL were used for first and second strand synthesis (Gaithersburg, MD).  $1\mu\text{Ci } \alpha\text{-}^{32}\text{P}\text{-dCTP}$  ( $3000\text{Ci}/\text{mmol}$ ) (New England Nuclear, Boston, MA) was added to first strand reactions to monitor the yield of cDNA throughout construction of the libraries (isotope incorporated into the libraries was routinely estimated by counting the entire reactions on the  $^3\text{H}$  channel of a Packard liquid scintillation spectrophotometer (Cerenkov counts)).

For synthesis of the first cDNA strand,  $25\mu\text{l}$  of water containing  $2\text{-}10\mu\text{g}$  of poly(A)+ RNA or containing  $15\mu\text{g}$  of total RNA was added to  $1\text{-}5\mu\text{g}$  oligonucleotide primers, then to  $1\mu\text{Ci}$  dry  $\alpha\text{-}^{32}\text{P}\text{-dCTP}$  ( $3000\text{Ci}/\text{mmol}$ ).  $1\mu\text{l}$  of RNasin (Promega Biotec) was added and the mixture incubated at  $50^\circ\text{C}$  for 10 minutes, then  $37^\circ\text{C}$  for 10 minutes, and finally room temperature for 5 minutes.  $25\mu\text{l}$  containing  $100\text{mM}$  Tris-HCl ( $\text{pH}8.3$ ),  $150\text{mM}$  potassium chloride,  $6\text{mM}$  magnesium chloride and  $20\text{ mM}$  dithiothreitol,  $2.5\mu\text{l}$  of  $10\text{mM}$  dNTPs, and  $2.5\mu\text{l}$  of  $10\text{U}/\mu\text{l}$  MMLV reverse transcriptase were then added and the reaction incubated for 1 hr. at  $37^\circ\text{C}$ .

For second strand synthesis the first strand reaction was placed on ice and the following components added, in order:  $289.5\mu\text{l}$  DEPC-treated  $\text{H}_2\text{O}$ ;  $7.5\mu\text{l}$   $10\text{mM}$  dNTPs;  $40\mu\text{l}$   $10\text{X}$  2nd strand buffer ( $250\text{mM}$  Tris,  $\text{pH } 8.3$ ;  $1\text{M}$  KCl;  $50\text{mM}$   $\text{MgCl}_2$ ,  $50\text{mM}$  dithiothreitol);  $10\mu\text{l}$  of  $10\text{ U}/\mu\text{l}$  E.coli DNA polymerase I;  $1.75\mu\text{l}$  of  $2\text{ U}/\mu\text{l}$  E.coli RNase H. The second strand was then incubated at  $16^\circ\text{C}$  for 2 hours, and stopped with the addition of  $10\mu\text{l}$   $0.5\text{ M}$  EDTA ( $\text{pH } 7.5$ ). The reaction was extracted with phenol/chloroform (1:1) and precipitated with  $3\mu\text{g}$  of yeast tRNA as carrier in  $0.3\text{ M}$  sodium acetate/ $70\%$  ethanol at  $-20^\circ\text{C}$ .



The double-stranded cDNA was then methylated to prevent internal cleavage by EcoRI during subsequent digestions. The cDNA was resuspended in 20  $\mu$ l water and added to 6  $\mu$ l of 5X Methylation Buffer [500mM Tris-HCl (pH8), 5mM EDTA (pH8), 50  $\mu$ M s-adenosylmethionine (Sigma) and 2  $\mu$ g per  $\mu$ l BSA] and 1  $\mu$ l of 80 U per  $\mu$ l EcoRI methylase (BRL) and the reaction incubated at 37°C for 30 min. The efficiency of the methylation was assayed by methylating 1  $\mu$ g of M13 DNA containing a 500 bp EcoRI insert in a parallel tube, and subsequently digesting the M13 DNA with EcoRI restriction enzyme.

The cDNA ends were filled in with Klenow prior to the linker ligation step. The methylation reaction (30  $\mu$ l) was heated at 65° for 15 minutes, then cooled on ice and added to 70  $\mu$ l containing: 36  $\mu$ l H<sub>2</sub>O; 10  $\mu$ l 10X buffer (0.5 M Tris, pH 8/100mM MgCl<sub>2</sub>/1 M NaCl); 20  $\mu$ l of 200  $\mu$ M dNTPs; 5 units Klenow. The reaction was incubated at 37° for 1 hour, then extracted with phenol/chloroform (1:1) and precipitated in 0.3 M sodium acetate/70% ethanol.

The amount of cDNA was determined by detecting <sup>32</sup>P radioactivity on the <sup>3</sup>H channel. The entire cDNA preparation, being as much as 2  $\mu$ g, was resuspended into 13  $\mu$ l of water and added to 1.5  $\mu$ l containing 0.17  $\mu$ g of kinased EcoRI linkers. It was then combined with 4  $\mu$ l of 5X ligase buffer (300mM Tris-HCl (pH7.5), 40mM magnesium chloride, 5mM ATP, 50  $\mu$ M  $\beta$ -mercaptoethanol and 250  $\mu$ g per ml BSA) and 2  $\mu$ l of 400 U per  $\mu$ l T4 DNA Ligase (BRL), and incubated at 15°C overnight.

The cDNA was then digested with EcoRI for subsequent ligation into the EcoRI site of the  $\lambda$ gt11 vector. Only the linkers should be sensitive to the digest since any internal EcoRI sites in the cDNA were



protected by methylation. The ligation reaction was added directly to 80  $\mu$ l containing: 10  $\mu$ l 10X EcoRI digestion buffer (1 M Tris-HCl (pH 7.5)/0.5 M NaCl/50mM MgCl<sub>2</sub>/1  $\mu$ g per  $\mu$ l BSA); 65  $\mu$ l H<sub>2</sub>O; 5  $\mu$ l of 20 U per  $\mu$ l EcoRI (BRL). The reaction was incubated for 6 hrs. at 37°C.

The cDNA was separated from the unligated linkers by passing the digestion reaction through Sephadex G-50 prepared in 1X STE buffer (100 mM NaCl, 10 mM Tris (pH 8.0), 1 mM EDTA) in a 1.8 ml glass pasteur pipette. The column was allowed to flow continuously with buffer after loading the sample, and 3 drop fractions (~80  $\mu$ l) were collected. Each fraction was counted directly on the <sup>3</sup>H channel of a Packard Liquid Scintillation Spectrophotometer. The peak fractions of radioactivity were pooled, extracted with phenol/chloroform, and precipitated from 0.3 M sodium acetate.

The cDNA was ligated to  $\lambda$ gt11 arms (predigested with EcoRI and alkaline phosphatased; BRL) in a 1:50 mass ratio. 100 ng of cDNA and 5  $\mu$ g of  $\lambda$ gt11 arms were brought to 10  $\mu$ l with water containing 1.2  $\mu$ l of 10X ligation buffer. 1  $\mu$ l of 10mM ATP and 1  $\mu$ l of 400 U per  $\mu$ l T4 DNA Ligase were added, and the reaction mixture was incubated at 16°C for 48 hours.

The  $\lambda$ gt11 now containing cDNA was packaged using Gigapack Plus (Stratagene, La Jolla, CA).

Nineteen cDNA libraries were constructed from embryonic 9-day to 14-day chicken brain RNA, and are shown in Table 2. In addition, two embryonic chicken cDNA libraries were purchased from Clontech and one  $\lambda$ gt10 embryonic chicken library was donated by Drs. Joan Levi and

Table 2      cDNA Libraries.

cDNA libraries were constructed or obtained for screening and isolation of Ng-CAM cDNA clones. RNA used to prime synthesis of cDNA was either total RNA or poly A<sup>+</sup> selected on oligo(dT) columns, isolated from embryonic 14 d chick brains unless indicated. Oligonucleotide primers used to prime cDNA synthesis were oligo(dT), random mixtures of hexanucleotides (dN<sub>6</sub> ; Bethesda Research Labs, Gaithersburg, Md.), or specific oligonucleotides generated from Ng-CAM cDNA sequence and listed in Table 3. The number of independent cDNA clones in each library (complexity) was determined by plating out unamplified libraries at serially diluted concentrations, counting the plaque forming units (pfu), and extrapolating to library density. Notations at right indicate the sources of libraries obtained from outside the laboratory.



TABLE 2 - LIBRARIES

	<u>RNA</u>	<u>Primers</u>	<u>Complexity</u>	
1	A <sup>+</sup>	(dT) <sub>n</sub>	2 x 10 <sup>6</sup>	
2	A <sup>+</sup>	(dN) <sub>6</sub>	2.5 x 10 <sup>6</sup>	
3	A <sup>+</sup>	#6	2 x 10 <sup>6</sup>	
4	A <sup>+</sup> ,*	#5	2 x 10 <sup>6</sup>	
5	A <sup>+</sup> ,*	(dT), #14,21	3 x 10 <sup>6</sup>	
6	A <sup>+</sup> ,*	(dN) <sub>6</sub>	3 x 10 <sup>6</sup>	
7	A <sup>+</sup>	(dT) <sub>n</sub>	1 x 10 <sup>6</sup>	Levi/Hanafusa
8	A <sup>+</sup>	#5	5.4 x 10 <sup>6</sup>	
9	A <sup>+</sup>	(dN) <sub>6</sub>	1.4 x 10 <sup>6</sup>	
10	total	#5	1 x 10 <sup>6</sup>	
11	A <sup>+</sup>	#4	4.8 x 10 <sup>6</sup>	
12	total	#4	5.4 x 10 <sup>6</sup>	
13	A <sup>+</sup>	#18	1.5 x 10 <sup>6</sup>	
14	total	#17	2.7 x 10 <sup>6</sup>	
15	A <sup>+</sup>	#18	4.7 x 10 <sup>6</sup>	
16	A <sup>+</sup>	(dT) <sub>n</sub>	3 x 10 <sup>6</sup>	
17	A <sup>+</sup>	#16	7 x 10 <sup>6</sup>	
18	A <sup>+</sup> ,@	#16	3 x 10 <sup>6</sup>	
19	A <sup>+</sup> ,@	#18	3 x 10 <sup>6</sup>	
20	A <sup>+</sup>	(dN) <sub>6</sub>	4 x 10 <sup>5</sup>	
21	embryo A <sup>+</sup> RNA	(dT) <sub>n</sub>	1 x 10 <sup>6</sup>	Clontech
22	emb. cere- bellum A <sup>+</sup>	(dT) <sub>n</sub>	1 x 10 <sup>6</sup>	Clontech

\* RNA selected by binding to cDNA clone  $\lambda$ N932

@ RNA treated with Actinomycin D and methylmercuric hydroxide



Hidesaburo Hanafusa (The Rockefeller University). Each of the cDNA libraries was constructed in the  $\lambda$ gt11 expression vector except library #7, which was constructed in  $\lambda$ gt10. Oligonucleotides used to prime the synthesis of several of these libraries, as well as to sequence the clones obtained, were generated from Ng-CAM protein sequence or from existing cDNA clones, and are shown in Table 3.

### Custom Libraries

The RNA used for the synthesis of libraries # 4-6 was hybrid-selected with the cDNA clone  $\lambda$ N932. Nitrocellulose filters containing  $\lambda$ N932 DNA were incubated with total RNA in hybridization buffer at 37°C overnight. The filters were then washed in hybridization buffer with 0.5x Pipes (50% formamide/ 1x Pipes (pH 6.4)/ 0.1% SDS/ 1mM EDTA) at 50°C and then several times in Pipes/0.1% SDS and finally 0.03x Pipes alone at 50°C, followed by a room temperature rinse in 20 ml Tris-HCl, pH 7.6. The RNA was eluted in 300  $\mu$ l H<sub>2</sub>O by heating to 100°C for 90 seconds, then freezing in a dry ice/ethanol bath. The RNA was then precipitated with 0.3 M NaOAc/ 70% ethanol and centrifugation.

Libraries 18 and 19 were prepared from RNA that had first been extensively denatured by incubation at 65°C and treatment with Actinomycin D (Rhyner et al., 1986) and methyl mercuric hydroxide (Maniatis et al., 1982). For each library 2-10 micrograms of RNA was treated in 29 microliters with 3 mM methylmercuric hydroxide for 10 minutes at room temperature, followed by the addition of 2 microliters of 700 mM 2-mercaptoethanol. 10 microliters of actinomycin D (1mg/ml) was added prior to first strand cDNA synthesis with a BRL kit. After

Table 3      Synthetic Oligonucleotides of Ng-CAM.

Synthetic oligodeoxyribonucleotides were synthesized based on Ng-CAM peptide sequence or on known cDNA sequence. Orientation of oligonucleotide with respect to the coding strand of the cDNA (sense or antisense) is indicated. In sequences the 5' to 3' direction is from left to right. Columns of nucleotides extending below a position indicate mixtures of nucleotides incorporated at that position to create degenerate oligonucleotides, and represent all possible codons (or anticodons).

TABLE 3 - OLIGONUCLEOTIDES

<u>Number</u>	<u>Sense/<math>\alpha</math>-sense</u>	<u>Sequence</u>
1	s	ATCACCATCCCCCGAGTACGGCGCTCACGATTTCTGCAGCCCGA
2	s	CACGATTTCTGCAGCCCCCGAGCTGACCGAGGAGCCCCCGAGCAGCTGGT
3	s	TTAAAATCATGAGCACCATACTC G G G C C G T G G T T
4	$\alpha$	AAGACCACGAGTTGTTCC
5	$\alpha$	CGGCTCCATCGGTACTGGACG
6	$\alpha$	GGACGAAGGGCTGATCTCA
7	s	CGTCCAGTACCGATGGAGCCGTGAGATC
8	s	GATCTGAGCGTTAGAGAAGTACAGGTTGCC
9	$\alpha$	GATGCACTCCAACACGACGC
10	s	ACCCAATAIGGIGCIGCCTC G T
11	$\alpha$	CCAATATGGGGCCGCCT
12	s	GAATTCACAGCAGACGAACAAAG C C T G GC G G T T
13	$\alpha$	CCAAAAGTACGACAATG G C C TG G G G T T
14	$\alpha$	AAGCTTCCACTCAACATTAGGAGC T C G C C G G G T T T
15	$\alpha$	CCGAAGGTTCTGCAGTG
16	s	CACTGCGGGGCGACCTT
17	s	TCCACCCCTCTCGACTC
18	$\alpha$	CACGAACTTCTCTATGG
19	s	AGGCCTGGAACGCCCTG



20	s	GGCGGTTCCCGTTCCGG
21	s	TGGCCGTTAAGCCTACGGGA
22	s	CACCGTGCGCGTGAGATGG
23	s	CCTGCCGTTGGTTTACCCTG
24	s	GGTCCATGGGGAGGGCAATGAGATCAG
25	$\alpha$	GGCGCGGCACTGAATGGTCC
26	$\alpha$	CTTGACGGCGACTGCGCTC
27	$\alpha$	GTCCCCATAGTCCACGTTG
28	$\alpha$	GGGTCTTCAGCCGG
29	$\alpha$	GGTCTTCAGCCGGGCTTTG
30	$\alpha$	CCCACTGTCAGAGCCACGG
31	$\alpha$	CCAACACCCGCAGCTGATAACGGC
32	$\alpha$	GCTCCAAAGCGGGGAGGGCGGTT
33	$\alpha$	GGTTTCGTGCTCCCCACCGTCTGC
34	$\alpha$	ACTCCCCAAAGGTCTCAT
M1	$\alpha$	TTCTCTTTCTTGCCACTGTACTGGCCGATGAAAGAGCCATCCTCATTGAACTGG





the first strand synthesis the reaction mixture was extracted with phenol/chloroform (1:1), and precipitated with 0.5 volumes 7.5 M  $\text{NH}_4\text{OAc}$  and 3 volumes ethanol at  $-20^\circ\text{C}$ . The pellet was resuspended in  $100\ \mu\text{l}$   $\text{H}_2\text{O}$  and used for second strand synthesis in the BRL cDNA synthesis protocol.

Three chicken genomic DNA libraries were also constructed. One hundred micrograms of chicken genomic DNA (Clontech, Palo Alto, CA) was partially digested with Sau3A1 or BamHI as determined in pilot experiments using samples of  $5\ \mu\text{g}$  genomic DNA with serially diluted Sau3A1 or BamHI enzyme concentrations. The digested DNA was separated in 10-40% continuous sucrose gradients by centrifugation in a Beckman SW-40 rotor for 20 hrs at 22k rpm and  $15^\circ\text{C}$ . Using an 18-gauge hypodermic needle, 0.5 ml fractions were collected from the gradients. Five percent of each fraction was electrophoresed on 0.4% agarose gels with ethidium bromide to estimate the size of the DNA in each fraction. Fractions containing fragments of 10-25 kb were pooled and precipitated with 0.3 M  $\text{NaOAc}$ / 70% ethanol, and then ligated into the BamHI sites of EMBL3 DNA (Frischauf et al., 1983; Maniatis et al., 1982), and packaged using Gigapack Plus (Stratagene, La Jolla, CA).

### Library Screening

cDNA libraries were screened according to the protocol of Young and Davis (Young and Davis, 1983) for antibody screening or the protocol of Maniatis (Maniatis et al., 1982) for screening with cDNA probes. For antibody screening  $\lambda\text{gt}11$  libraries were grown with Y1090 cells (Young and Davis, 1983) on 150mm plates in LB-agarose with ampicillin, at a



density of 60,000-80,000 pfu/plate. IPTG (0.33mM) was added during growth to induce production of fusion protein. After 12-16 hours growth at 37°C, the plates were chilled to 4°C and nitrocellulose replicas were lifted. The replicas were treated with PBS/0.5% NP-40/3% skim milk and incubated with polyclonal  $\alpha$ -Ng-CAM antibodies. The antibodies were raised against denatured Ng-CAM protein, and recognize the 210, 190, 135 and 80 kD components of the molecule (Grumet et al., 1984a; Rieger et al., 1986). After washing in PBS/0.5% NP-40 (2 x 10 minutes, 2 x 60 minutes) the nitrocellulose filters were incubated in with [<sup>125</sup>I]-Protein A, and washed again. The filters were autoradiographed on Kodak XAR-5 film with intensifying screens. Positive clones were picked, rescreened until homogeneous and then picked and grown for further analysis.

For nucleic acid screening,  $\lambda$ gt10 or  $\lambda$ gt11 phage libraries were grown with Y1090 cells as above (without IPTG) and duplicate nitrocellulose replicas were lifted, denatured in 0.5 N NaOH/1.5 M NaCl for 1 minute, and neutralized in 1.5 M NaCl/1.0 M Tris (pH 8.0) for 5 minutes. After rinsing in 2x SSC for 5 minutes the filters were baked at 80°C under vacuum for 2 hours, and prehybridized in hybridization buffer (5x Denhardt's/ 6x SSC/ 90mM Tris, pH 8/ 0.05mg per ml salmon sperm DNA/20 mM sodium phosphate, pH 7/ 6 mM EDTA/ 0.1% SDS). Radiolabeled cDNA probes were denatured by heating to 100°C for 10 minutes, and added to the filters in hybridization solution. After overnight incubation at 60-65°C on a rocking platform, the filters were washed with 0.5x SSC/ 0.5% SDS at 65°C (2x 10 minutes, 2x 60 minutes). The filters were autoradiographed overnight on Kodak XAR-5 film with



intensifying screens. The films from duplicate filters were aligned and doubly positive plaques were picked, replated and rescreened to homogeneity.

Colony screening of the EMBL3 plasmid genomic libraries was performed on nitrocellulose filters (Hanahan and Meselson, 1980). The libraries were grown on LB +maltose plates in LE392 cells at a density of  $\sim 8 \times 10^3$  colonies/150mm dish. Replica filters were lifted from the plates, floated on Whatman 3MM paper soaked in 2 % SDS for 2 minutes, and then denatured, neutralized and screened with cDNA probes as indicated above.

#### cDNA Probes

cDNA inserts were labeled with  $\alpha$ -[ $^{32}\text{P}$ ]-dCTP (New England Nuclear) using the oligolabeling protocol of Feinberg and Vogelstein (Feinberg and Vogelstein, 1984) with random-primed labeling kits from Boehringer-Mannheim (Indianapolis,IN).

#### cDNA Isolation and Purification

DNA was purified from lytic cultures of the  $\lambda$ gt11 phage as described (Maniatis et al., 1982). The inserts were excised from  $\lambda$ gt11 arms by restriction with EcoRI endonuclease, separated by electrophoresis in agarose gels containing EtBr (0.5ug/ml), and excised for further analysis. DNA was obtained from the gel slices by one of two methods. When normal melting temperature agarose was used the excised bands were frozen at  $-70^\circ\text{C}$  for 20 minutes and quickly centrifuged at room temperature for 5 minutes in micro-spin filter tubes (Ultrafree-MC,



Millipore Corp., Bedford, MA). When low-melting temperature agarose was used the gel band was melted at 65°C, extracted sequentially with phenol, phenol/sevag (sevag=chloroform:isoamyl alcohol, 24:1), and sevag, and precipitated in 0.3 M sodium acetate/70% ethanol (Maniatis et al., 1982).

### DNA Sequencing

cDNA inserts were sequenced on each strand by the dideoxynucleotide chain-termination method (Sanger et al., 1977) after subcloning into M13mpl8 or M13mpl9 vectors (BRL, Gaithersburg, MD).  $\alpha$ -[<sup>35</sup>S]-dATP and Sequenase and Sequenase II kits (U.S. Biochemical, Cleveland, OH) were used for the sequencing reactions, which use a modified T7 DNA Polymerase enzyme. M13 template DNA was prepared as described (Ausubel et al., 1989) and 1 $\mu$ g was used for each sequencing reaction. The sequencing reactions were electrophoresed on 0.4- 0.8 mm thick (wedge) polyacrylamide gels, (7% polyacrylamide / 7M urea / 1xTBE) at 60 watts in 1X TBE. Gels were fixed in 10% methanol/10% acetic acid for 30 minutes, dried onto Whatman 3MM paper under vacuum, and autoradiographed overnight on Kodak XAR-5 film.

Difficulties in obtaining particular sequences were surmounted by one of several methods: The template DNA was pretreated with single-stranded binding protein (US Biochemical); the concentration of the dideoxynucleotide mixture in the first labeling step was lowered or raised to obtain readings closer or farther from the primer; deoxyinosine triphosphate was substituted for deoxyguanosine triphosphate in several reactions.





The sequence of larger inserts was obtained after deletion cloning in M13 vectors, prepared by one of two methods. The clone was block-deleted by digestion with restriction endonucleases which recognized sites in the M13 polylinker and within the insert, followed by religation of the clone and sequencing across the newly revealed end of the insert.

A second method for obtaining deletion clones utilized a M13 Cyclone cloning system (International Biotechnologies, New Haven, CT), which is based on the rapid deletion cloning method of Dale et al. (Dale et al., 1985). Single-stranded M13mp18 template were hybridized to complementary oligonucleotides along their polylinker and digested at the HindIII site. The clone was then treated with Exonuclease III. 10% aliquots were removed at 5 minute intervals and stopped with 25mM EDTA and placed on ice. Aliquots were then blunt-ended with Mung Bean nuclease, tailed with T4 polynucleotide kinase and 5 mM dCTP, followed by annealing to a 29-base linker and religation. The reaction aliquots were transformed into transformation-competent JM101 cells (Ausubel et al., 1989) and plated on B plates in the presence of 0.33 mM IPTG and 0.03 % X-Gal. White colonies were picked with sterile toothpicks and grown in 2 ml YT media at 37°C overnight. Purified phage were obtained by centrifugation of the overnight cultures at 8000xg for 10 minutes and collection of the supernatants. The phage were sized by incubating 20  $\mu$ L of the supernatant with 4  $\mu$ l c-test dye (0.5% SDS/20% glycerol/0.04% bromphenol blue) at 65°C for 10 minutes and electrophoresis in 0.7% agarose gels with 1X TBE and ethidium bromide (0.05  $\mu$ g/ml), using undeleted M13 clones and M13 vector DNA as markers. From appropriately



sized cDNA clones, M13 single stranded template was prepared and sequenced with M13 universal primer (US Biochemical, Cleveland, OH).

A third method for sequencing the interior of larger cDNA clones utilized specific oligonucleotide primers synthesized from existing sequences in the insert. These were annealed to the M13 template DNA instead of the universal primer, and extended the sequence further into the insert.

### Analysis of Protein Sequence

Sequence data were compiled and analyzed using the Staden ANALYSEQ programs (Staden, 1984) and the Staden shotgun sequencing handling programs (Staden, 1982). Both the FASTA and the LFASTA programs were used to compare protein sequences (Pearson and Lipman, 1988). The FASTA program compares a protein sequence to a protein sequence database or to another protein sequence, and aligns similar sequences along their entire lengths in the single most favorable fashion to maximize identities, introducing gaps in either sequence to extend the regions of similarity. LFASTA compares two sequences to identify regions of sequence similarity, reporting several sequence alignments if there are several similar regions (Pearson, 1990). The Dayhoff protein sequence database (National Biomedical Research Foundation, Washington, DC) and the translated Genbank database (release 12) were searched using the program FASTA (Ktup=2; Pearson and Lipman, 1988). The FASTA program was also used to compare the various neural CAMs to obtain overall percentage identities. The LFASTA program was used in the initial pairwise alignment of Ng-CAM, L1, and Nr-CAM, followed by visual



inspection. Hydrophobicity analyses of the sequence were performed using an automated hydropathy program (Eisenberg, 1984) using the parameters of Kyte and Doolittle with a sliding window of 19 residues (Kyte and Doolittle, 1982).

### PCR

The polymerase chain reaction (PCR) was used to synthesize cDNA clones between pairs of sense and antisense oligonucleotides synthesized from protein or DNA sequences in Ng-CAM. PCR techniques used Taq Polymerase and Perkin Elmer-Cetus (Norwalk, CT) reagent kits in a programmable thermal cycler (USA/Scientific Plastics, Inc., Ocala, FL) (Saiki et al., 1988). General reaction conditions included in a 100  $\mu$ l volume: 0.05-4 $\mu$ g DNA, 100ng of each oligonucleotide primer (15-30 bases long), 1.5mM MgCl<sub>2</sub>, 0.2mM of each dNTP, 10mM Tris (pH 8.3), 2.0 units Taq DNA Polymerase (the MgCl<sub>2</sub> concentration was optimized for individual applications). For example, cDNA templates were denatured at 94°C for 1 minute, annealed with synthetic oligonucleotide primers at 55°C for 1 minute, and then DNA was synthesized with Taq Polymerase at 72°C for 1 minute. 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.

### Phosphorylation

Ng-CAM protein was purified by affinity chromatography as indicated above. Detergent was absent from the elution buffer, and the



protein was concentrated by pressure dialysis in Amicon chambers, using phosphorylation buffer 20 mM Tris, pH 7.4/150 mM NaCl/10 mM MgCl<sub>2</sub>/2 mM MnCl<sub>2</sub>. The protein was used immediately or within 5 days in phosphorylation assays.

Five micrograms of Ng-CAM was incubated in 50  $\mu$ l phosphorylation buffer (Sadoul et al., 1989) with 0.1  $\mu$ Ci  $\gamma$ -<sup>32</sup>P-ATP, at 25°C or 35°C for 30 min. Parallel samples were preheated at 100°C for 5 minutes, or pretreated with 0.1 M vanadate for 30 min at room temperature. After incubation at the appropriate assay temperature, 5  $\mu$ l of 5X SDS sample loading buffer was added to each sample. The samples were heated at 100°C for 5 min and resolved by SDS-PAGE on 6% polyacrylamide gels (Laemmli, 1970) in parallel with appropriate molecular weight standards. The gels were dried onto Whatman 3MM paper under vacuum and exposed to Kodak XAR-5 autoradiography film overnight.

To examine the presence of phosphotyrosine, 5  $\mu$ g Ng-CAM was phosphorylated with nonradioactive ATP under the same conditions as above and resolved on 6% polyacrylamide gels, in parallel with 5  $\mu$ g of extracts from A431 cells containing EGF receptors (ATCC, Rockville, MD). The gels were transferred to Hybond membrane (Amersham, Arlington Heights, IL) and immunoblotted with monoclonal  $\alpha$ -phosphotyrosine antibody G20 (ICN Biomedicals, Inc., Costa Mesa, CA), followed by goat  $\alpha$ -mouse IgG and <sup>125</sup>I-Protein A, and autoradiographed.





## GENERAL SOLUTIONS

### Phosphate-buffered saline (PBS)

#### *10X stock solution, 1 liter:*

80 g NaCl  
2 g KCL  
11.5 g  $\text{Na}_2\text{HPO}_4 \cdot 7\text{H}_2\text{O}$   
2 g  $\text{KH}_2\text{PO}_4$

#### *Working solution, pH 7.0:*

137 mM NaCl  
2.7 mM KCL  
4.3 mM  $\text{Na}_2\text{HPO}_4 \cdot 7\text{H}_2\text{O}$   
1.4 mM  $\text{KH}_2\text{PO}_4$

### 20X SSC

#### *1 liter:*

175.3 g NaCl (3 M)  
88.2 g  $\text{Na}_3\text{citrate} \cdot 2\text{H}_2\text{O}$  (0.3 M)

### 10X TBE electrophoresis buffer

#### *1 liter:*

108 g Tris base  
55 g boric acid  
40 ml 0.5 M EDTA, pH 8.0

#### *Working solution, pH 8.0:*

0.089 M Tris base  
0.089 M boric acid

### 10X MOPS running buffer

[3-(N-morpholino)-propanesulfonic acid]

#### *1 liter, pH 7.0:*

41.8 g MOPS  
16.6 ml 3M Na acetate  
20. ml 0.5 M EDTA, pH 8.0

#### *Working solution:*

20 mM MOPS  
8 mM Na acetate  
1 mM EDTA

### 50X Denhardt's solution

#### *1 liter:*

10 g Ficoll 400  
10 g bovine serum albumin (fraction V; Sigma)  
10 g polyvinylpyrrolidone

### LB growth medium

#### *1 liter, pH 7.5:*

10 g bacto-tryptone  
5 g bacto-yeast extract  
10 g NaCl

### YT growth medium

#### *1 liter, pH 7.5:*

16 g bacto-tryptone  
10 g bacto-yeast extract  
5 g NaCl



A major goal of this thesis was to characterize in detail the structure of Ng-CAM in order to determine the relationship of the polypeptide components and to relate the structure of Ng-CAM to its known functions. In addition, it was important to compare Ng-CAM to other neural CAMs, particularly L1 and NILE.

### Protein Sequence

Ng-CAM is expressed in the chicken as four components, a predominant glycoprotein of 135 kD, with minor species of 80 kD and a doublet of 190 kD and 210 kD. In order to directly compare them, the various components were isolated and their amino terminal sequences determined. Ng-CAM was immunoaffinity purified using monoclonal antibodies and resolved by SDS-PAGE (Figure 3). The components were transferred to Immobilon and individually sequenced. The amino terminal sequences of the 210 kD, 190 kD and 135 kD components were identical to each other, but differed from that of the 80 kD component. These data supported the previous linear model for Ng-CAM shown in Figure 4.

The difference between the 190 kD and 200 kD components is not clear, but probably reflects different posttranslational modifications of a common translation product. In support of this view, Ng-CAM is detected in tunicamycin-treated cells as a single 170 kD species (Burgoon et al., 1991). For convenience, and because they do not always separate on SDS-PAGE, I refer to the 190 kD and 210 kD collectively as the 200 kD component. The model in Figure 4 suggests that one or both

Figure 3 Components of Ng-CAM and amino terminal sequences.

Ng-CAM from embryonic chick brain was fractionated on 6% polyacrylamide gels and stained with Coomassie Blue. The amino terminal sequence of each component (210, 190, 135, 80 kD) was determined after transfer to Immobilon. X represents a position at which the sequence data was ambiguous and precise determination of the residue could not be established.

Amino Acid Sequence

-210Kd H<sub>2</sub>N-XTIPPXYGAX...

-190Kd H<sub>2</sub>N-XTIPPEYGAH...

-135Kd H<sub>2</sub>N-ITIPPEYGAH...

- 80Kd H<sub>2</sub>N-APPDPPQIPQ...

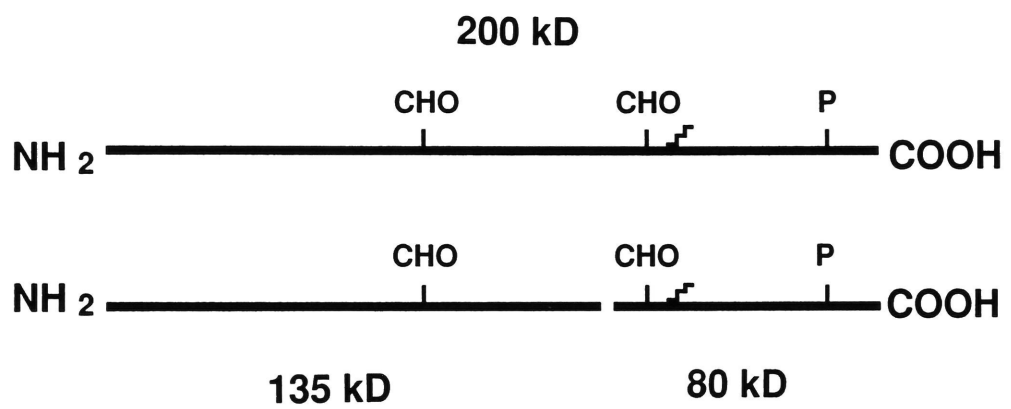




Figure 4      Linear Model of Ng-CAM.

The protein components of Ng-CAM are represented in positions which reflect their similar features. Immunological and biochemical data suggested that the 135 kD and 80 kD components are derived from the larger 200 kD components. General locations of carbohydrate (CHO), phosphate (P), and palmitate (ladder) are indicated.







of these species is proteolytically cleaved, so that the amino terminal portion gives rise to the 135 kD component and the carboxyl terminal segment gives rise to the 80 kD component. Previous immunological evidence agrees with this notion. Antibodies raised to the 135 kD component recognized it and the 200 kD components, but not the 80 kD component. Conversely, antibodies raised to the 80 kD component recognized it and the 200 kD component but not the 135 kD component (Grumet et al., 1984b). The model is also in accord with biochemical studies that showed that the 80 kD and 200 kD components of Ng-CAM were phosphorylated, but the 135 kD component was not (Grumet et al., 1984b; Sorkin et al., 1985).

To obtain additional amino acid sequences, immunoaffinity-purified Ng-CAM and the isolated 80 kD component were treated with CNBr, *Staphylococcus aureus* V8 protease, or trypsin, and the peptides purified by gel filtration and HPLC or by SDS-PAGE. The amino terminal sequences of fourteen peptides were determined (see Table 1). The four sequences obtained from the 80 kD component (peptides 9-12) differed from the sequences derived from proteolytic fragments of total Ng-CAM (predominantly 135 kD), further supporting the model that the 135 kD and 80 kD components are distinct. These sequences together comprised 278 amino acids, and represented about 10% of the predicted size of the 200 kD component. In order to complete the sequence of Ng-CAM, cDNA clones encoding the entire protein were isolated and sequenced.



### Cloning Strategy

The entire sequence of Ng-CAM was obtained from overlapping cDNA clones isolated from embryonic chick brain cDNA libraries. Two strategies were used to obtain initial Ng-CAM cDNA clones. In the first, the polymerase chain reaction (PCR) was used to obtain an Ng-CAM nucleic acid probe, using primers based on the amino acid sequence of a peptide. In the second, polyclonal antibodies specific for Ng-CAM were used to isolate clones from cDNA expression libraries.

The PCR approach used the longest Ng-CAM peptide sequence (# 9 in Table 1). Two degenerate oligonucleotide primers were synthesized based on this sequence, one in the sense orientation and the other in the antisense orientation (# 12, 14 in Table 3; Figure 5a). Each primer was designed to contain all possible combinations of codons across the specified region of the peptide. cDNA from 14-day embryonic chick brain RNA was primed with the antisense oligonucleotide and then amplified with both oligonucleotides. A 105 basepair band (the expected size) was the predominant species after electrophoresis of the reaction products on agarose gels. This band was subcloned into M13mp18 and sequenced (Figure 5b). The sequence contained the two oligonucleotide primers at either end separated by a sequence which encoded the intervening region of the original CNBr peptide of Ng-CAM. Near the 3' end of the PCR product the peptide was encoded by a different reading frame than the 5' region. This reading frame shift probably resulted from a copying error during the amplification of the cDNA by PCR, resulting in a single base deletion. The Taq polymerase used for the PCR is known to have a higher

Figure 5 Isolation of Ng-CAM cDNA Clone Using PCR.

(a) Design of oligonucleotide primers. From the sequence of a CNBr-cleaved peptide of Ng-CAM, represented by single letter amino acid code, two oligonucleotides (one sense, the other antisense ; overlining arrows) were synthesized containing all possible combinations of codons for the regions specified. Restriction endonuclease recognition sequences for EcoRI and HindIII were placed at the 5' end of the sense and antisense oligonucleotides, respectively. Possible codons translating the peptide are represented below.

(b) PCR product and sequence. The oligonucleotides (arrows) were used in PCR with embryonic chick brain RNA to generate a DNA product of 105 basepairs, indicated by a thick bar. The DNA sequence of the PCR product is listed below in a 5' to 3' direction, with regions matching the oligonucleotide primers underlined. Partial translated sequence from one reading frame is listed above the DNA, and the corresponding peptide sequence from Ng-CAM aligned above that. Bracket indicates the DNA sequence from which a 57 base oligonucleotide was synthesized.

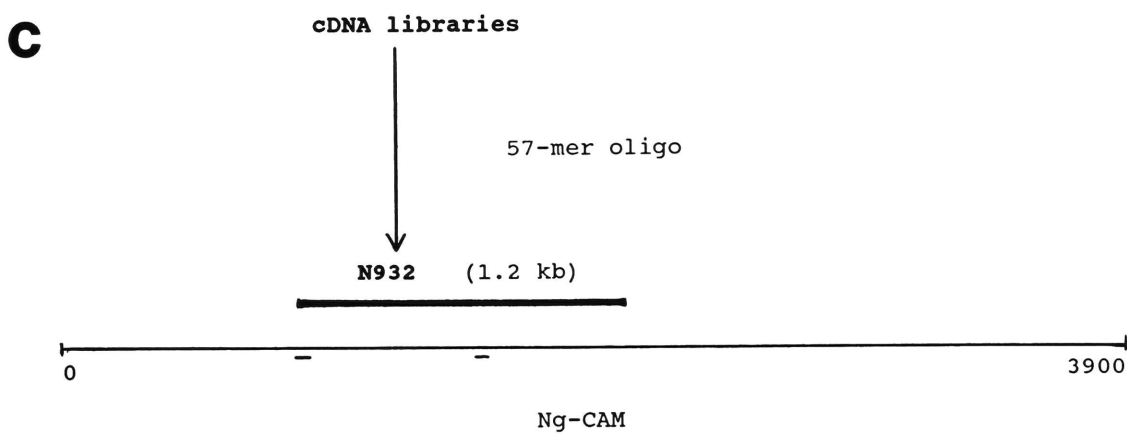
(c) Scheme for isolation of cDNA clone. The 57 base oligonucleotide was used to screen cDNA libraries and isolate clone N932 (thick bar, with size indicated in parentheses). Bar is drawn in scale to corresponding segment of the entire Ng-CAM cDNA sequence, with size delineated in basepairs. Underlining segments show locations of known peptide sequences found in clone N932.

**a**

L	T	A	D	E	Q	R	Y	E	V	T	E	X	Q	T	V	F	L	H	C	R	T	F	G	A	P	A	P	N	V	E	W	L	T	P	T	L	E	G	A
ctaacagcagacgaacaacgatacgaagtaacagaa	caaacagttattcctacactgccgaacattcggagcaccagcaccacaaacgtagaatggc	taacaccaacactagaaggagca																																					
t c c c t g g a c t g c c g	g c c t t c t t a c c t c c c c c t c g	t c c c c t c g c c																																					
g g g	g g g	g g g g g g g g g g																																					
t t t	t t t	t t t t t t t t t t t t t t t t t t																																					

**b**

<div style="border-top: 2px solid black; width: 100%; height: 5px; position: relative;"> <span style="position: absolute; left: 0; top: -10px;">→</span> <span style="position: absolute; right: 0; top: -10px;">←</span> </div>																													
T A D E Q R Y E V V E - Q T V F L H C																													
T A D E Q R Y E V V E N Q T V F L H C																													
<u>GAATTCACAGCGGATGAGCAGCGCTACGAAGTGGTGGAAACCAACAGTGTCTTCTGCACTGCACCTGCAGAACCTTGGGGCCCCCGCCCAACGTCGAATGGAAGCTC</u>																													
<div style="border-top: 1px solid black; width: 100%; height: 5px; position: relative;"> <span style="position: absolute; left: 0; top: -10px;">└─┘</span> </div>																													
57-mer																													







error rate of replication than eukaryotic DNA polymerases (Tindall and Kunkel, 1988).

A 57-base oligonucleotide was synthesized matching the accurately amplified sequence at the 5' end of the PCR product (Figure 5b), and was used as a probe to screen embryonic chick brain cDNA expression libraries. Clone  $\lambda$ N932 was isolated using this probe (Figure 5c).  $\lambda$ N932 encoded the original CNBr peptide at its 5' end. It also encoded another CNBr peptide from Ng-CAM (#10 in Table 1), further verifying its authenticity.

A second clone ( $\lambda$ N903) isolated using the PCR-derived oligonucleotide was 2.8 kb in length and contained  $\lambda$ N932 (see Figure 7). Clone  $\lambda$ N903 contained two EcoRI fragments, both of which were sequenced. The 2.0 kb EcoRI fragment contained sequences overlapping  $\lambda$ N932. However, the 0.7 kb fragment contained multiple stop codons in every reading frame, and did not encode any sequences related to Ng-CAM. The ends of both EcoRI fragments matched the artificial EcoRI linkers used to synthesize the cDNA libraries. This observation suggested that two unrelated cDNAs had been ligated together during the construction of the library. This was later confirmed when other clones were obtained that overlapped the 2.0 kb EcoRI fragment and extended further 5'. These clones differed from the 0.7kb EcoRI fragment and encoded sequences that matched Ng-CAM peptide sequences. The 2.0 kb fragment alone was considered to be clone  $\lambda$ N903.

The second method for obtaining initial clones used polyclonal antibodies raised against Ng-CAM to screen cDNA expression libraries in the vector  $\lambda$ gt11. An antibody which had a high titer to all components



of denatured Ng-CAM was chosen to optimize the detection of fusion proteins. In a primary screen fifteen clones reacted with the antibody with varying intensities (Figure 6). Upon subsequent screening only two of these clones,  $\lambda$ N902 and  $\lambda$ N925, reacted strongly with the  $\alpha$ -Ng-CAM antibody. These two clones were sequenced and contained Ng-CAM sequences. Clone  $\lambda$ N902 encoded 81 amino acids near the 3' end of the coding sequence (1061-1141, see Fig. 8). Overlapping clones were obtained with  $\lambda$ N902 as a probe, and extended in the 5' direction to encode known Ng-CAM protein sequences. The insert in clone  $\lambda$ N925 encoded 113 amino acids (85-197 in Fig. 8) and overlapped with the sequence of  $\lambda$ N903 (Figure 7).  $\lambda$ N925 also encoded the sequence of a tryptic peptide from Ng-CAM (peptide #6 in Table 1), verifying its authenticity.

#### cDNA Clones

To obtain cDNA clones encoding the entire Ng-CAM protein, restriction fragments of  $\lambda$ N902,  $\lambda$ N903 and  $\lambda$ N925 were used to screen a variety of libraries, including thirteen libraries in which cDNA synthesis was primed with Ng-CAM-specific oligonucleotides. In all, fourteen overlapping clones were required to span the entire coding sequence of Ng-CAM (Figure 7).

The cDNA clones were isolated from twenty-two cDNA libraries (Table 2) by screening approximately  $3 \times 10^7$  pfu. Most of the libraries were primed with Ng-CAM-specific oligonucleotides generated from the previous DNA sequences to increase the representation of Ng-CAM sequences in the libraries. Despite the large number of clones

Figure 6      Isolation of Antibody Positive Clones for Ng-CAM.

$\lambda$ gt11 cDNA libraries were screened with antibody #885 raised to denatured Ng-CAM. The 15 clones initially selected were rescreened with the antibody, until two clones were purified to homogeneity based on their reactivity to the antibody. Autoradiographs of the reactivity of the antibody to the two clones ( $\lambda$ N925 and  $\lambda$ N902a) is shown in comparison to a different clone ( $\lambda$ b2b1) which did not react to the antibody.

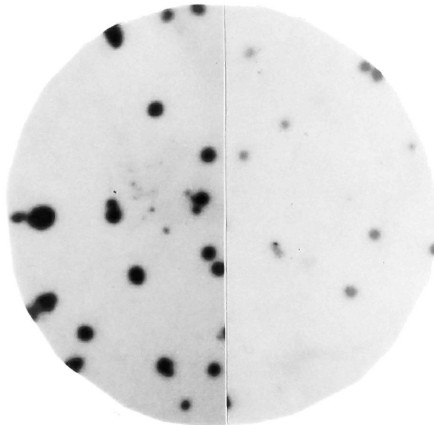
**$\lambda$ gt11 Expression  
Libraries**

Polyclonal  
antibody #885

15 clones

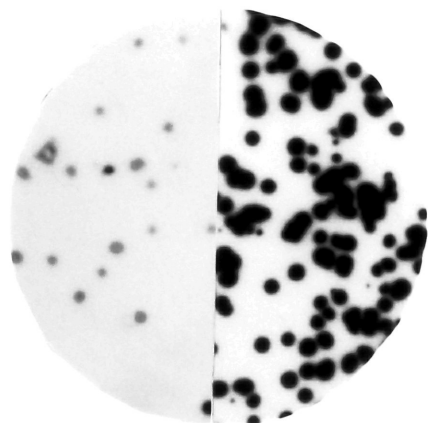
Polyclonal  
antibody #885

2 clones



**$\lambda$ N925**

$\lambda$ b2b1  
(-)



**$\lambda$ N902**

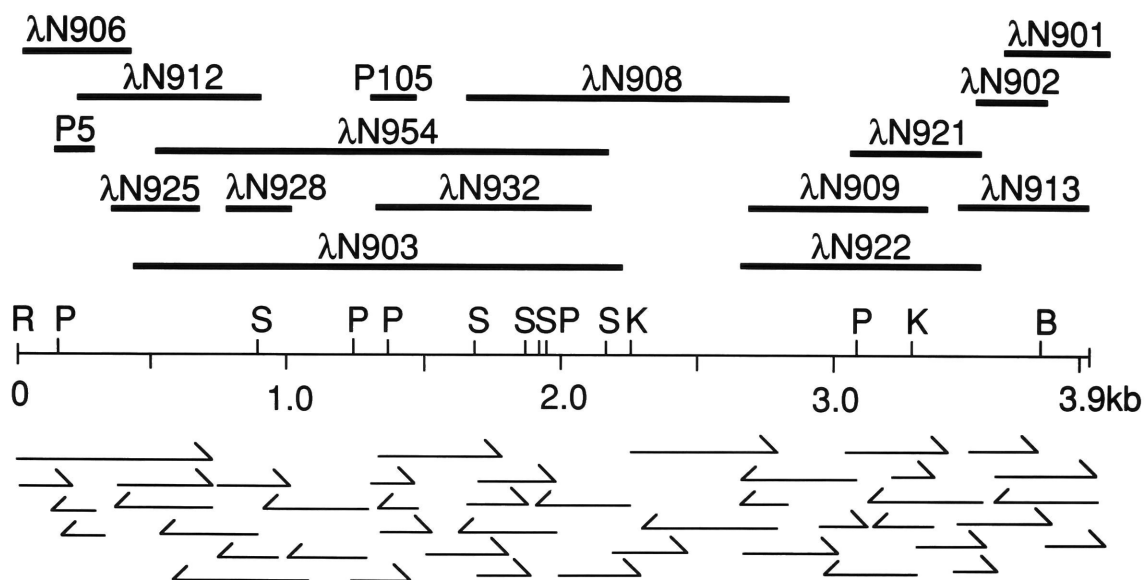




Figure 7 Ng-CAM cDNA clones and the sequencing strategy.

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 endonuclease sites are: R, EcoRI; P, PstI; S, SmaI; K, KpnI; B, BamHI. Direction and extent of sequencing are indicated by arrows below the scale bar.







screened, most in unamplified libraries to maximize the numbers of independent clones screened, there was a serious difficulty in finding Ng-CAM cDNA clones. The largest clone isolated ( $\lambda$ N903) was 2.0 kb; most of the clones ranged in size from 500-1000 basepairs.

The difficulty in finding Ng-CAM clones may have been due to at least two factors: the relative instability of the Ng-CAM mRNA, and the high G-C content of the sequence. The relative instability of the Ng-CAM mRNA was reflected by the difficulty in obtaining sharp RNA hybridizations for Ng-CAM on Northern blots (this same RNA yielded sharp signals for N-CAM and other proteins). Preferential degradation of the Ng-CAM mRNA during RNA isolation would result in difficulties in synthesizing large Ng-CAM cDNA clones. Very rapid isolation methods were required for obtaining full-length Ng-CAM mRNA, including immediate immersion of dissected brains in liquid nitrogen, grinding the tissue under liquid nitrogen and quickly homogenizing the frozen tissue powder in strong denaturing agents.

The relative abundance of deoxycytidines and deoxyguanidines in the entire sequence was particularly evident in the region surrounding nucleotides 2600-2800 (see Figure 8). This region contained extended segments of polyG and polyC which may create secondary structure conformations that inhibit cDNA synthesis and indeed this was the most difficult region to clone. Eight cDNA libraries were constructed and screened for the sole purpose of finding a cDNA which extended from clone  $\lambda$ N922 to  $\lambda$ N903 (Figure 7). Three genomic libraries were also constructed to identify a bridging DNA sequence. Only after screening  $\sim 10 \times 10^6$  pfu was the clone  $\lambda$ N908 found which spanned this region (nucleotide 1701 to 2766).  $\lambda$ N908 was isolated from a cDNA library in



which the poly(A)<sup>+</sup> RNA had been extensively denatured with methylmercuric hydroxide and Actinomycin D, and primed with an oligonucleotide specific for Ng-CAM in that region. It thus seems likely that the polyG's and polyC's abundant in this region of the mRNA can form secondary structures that inhibit either cDNA synthesis or replication of the DNA in bacteria.

The fourteen overlapping clones spanned 3989 basepairs, and are shown in Figure 7. The cDNA sequence encoded a single long open reading frame containing 1265 amino acids; other reading frames on both strands contain multiple stop codons throughout the sequence (Figure 8). The deduced protein sequence represents the entire Ng-CAM protein.

More recently, another cDNA clone was isolated and partially sequenced, which overlaps this sequence at the 3' end for 321 nucleotides and extends the sequence approximately 1.3 kilobases in the 3' direction. At the 3' end of this new clone, a long sequence of poly-A was detected, suggesting that it represents the 3' end of the Ng-CAM mRNA. The sequence obtained upstream of the poly-A segment was ambiguous. The detection of a polyadenylation signal must await deletion cloning and sequencing of the clone.

The validity of the predicted protein sequence is indicated by the fact that it includes all sequences obtained from Ng-CAM peptides. These peptide sequences (Table 1) comprise 10% of the deduced protein sequence, are distributed throughout the extracellular portion of the molecule, and are all encoded by the same reading frame.

**Figure 8** Nucleotide sequence and deduced amino acid sequence of Ng-CAM.

The overlapping cDNA clones extend 3989 basepairs. The longest open reading frame includes 1265 amino acids beginning at the predicted translation start site and terminating at an other termination codon (0). The two hydrophobic regions representing the signal peptide (-20 to -1) and the transmembrane region (1110 - 1132) are underlined by a thick bar. Thinner underlining denotes amino acid sequences determined by protein chemical analyses of the 210 kD, 190 kD, 135 kD, 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 (Hubbard and Ivatt, 1981) are marked with inverted triangles and potential phosphorylation sites are indicated by asterisks (\*). The predicted proteolytic cleavage site between the 135 kD 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/DBJ under accession number X56969.

[illegible]





### Model of Ng-CAM

A typical signal peptide (vonHeijne, 1985) is encoded by a sequence near the 5' end of the cDNA sequence, beginning with a methionine at nucleotide 59 (amino acid residue -20) and extending to the amino terminal sequences of the 135 kD and 200 kD components. The twenty amino acids include a stretch of hydrophobic residues followed by a helix-breaking proline and several small residues (glycine-alanine-alanine), after which signal peptidase is expected to cleave (Perlman and Halvorson, 1983; vonHeijne, 1986). As a further test of this conclusion, the computer program SIGSEQ (Popowicz and Dash, 1988) was used to predict the most likely cleavage site for signal peptidase (Fig. 9). The program uses a weight-matrix approach to identify regions of protein sequence which conform to characteristic signal peptides (Watson, 1984; vonHeijne, 1985). In the entire Ng-CAM protein sequence, the bond between Ala(-1) and Ile(1) was predicted to be the primary cleavage site.

Hydrophobicity analysis was performed using a window size of 19 amino acids to optimize the localization of any membrane-spanning segments (Figure 10). Two extended hydrophobic regions were detected in Ng-CAM. The first occurs at the beginning of the sequence and corresponds to the predicted signal peptide (amino acids -20 to -1 in Fig. 8). The second hydrophobic domain (amino acids 1110 to 1132) has features that are characteristic of transmembrane segments (Davis and Model, 1985; vonHeijne, 1981). The segment contains 23 uninterrupted hydrophobic or nonpolar residues, including a long stretch of leucines

Figure 9      Prediction of signal peptide cleavage sites in Ng-CAM.

The SIGSEQ program of Popowicz (Popowicz and Dash, 1988) was used to predict the optimal signal peptidase cleavage site in the translated amino acid sequence of Ng-CAM. Using a sliding window of 15 amino acid residues and the parameters of von Heijne (von Heijne, 1986), the optimal site in the entire sequence of Ng-CAM for cleavage by signal peptidase occurs between Ala(-1) and Ile(1). Cleavage at this site would expose the amino terminal sequence obtained from the mature protein.

PREDICTION OF SIGNAL SEQUENCE CLEAVAGE SITES BASED ON THE METHOD OF  
G. VON HEIJNE, N.A.R., 14, 4683(1986)

SEQUENCE NAME: ng.aa  
LENGTH: 1265  
SEARCH RANGE: 50  
SEQUENCE TYPE: Eukaryotic

WINDOW	SUBSEQUENCE	NORMALIZED PROBABILITY P[i]/Pmax
1 - 15	MALPMVGLLLLLLL~LG	0.000000
2 - 16	ALPMVGLLLLLLL~GG	0.000054
3 - 17	LPMVGLLLLLLL~GP	0.010673
4 - 18	PMVGLLLLLLL~GG~PG	0.000187
5 - 19	MVGLLLLLLL~GGP~GA	0.000430
6 - 20	VGLLLLLLL~GGPG~AA	0.030419
7 - 21	GLLLLLL~GGPGA~AI	0.045391
8 - 22	LLLLLL~GGPGA~IT	1.000000
9 - 23	LLLLLL~GGPGA~TI	0.000278
10 - 24	LLLL~GGPGA~IT~IP	0.200555
11 - 25	LLL~GGPGA~ITI~PP	0.000000
12 - 26	LL~GGPGA~ITIP~PE	0.000001
13 - 27	LL~GGPGA~ITIPP~EY	0.000001
14 - 28	L~GGPGA~ITIPPE~YG	0.000000
15 - 29	G~GPGA~ITIPPEY~GA	0.000000
16 - 30	G~GPGA~ITIPPEY~AH	0.000000
17 - 31	P~GAA~ITIPPEYGA~HD	0.000000
18 - 32	G~AA~ITIPPEYGAH~DF	0.000000
19 - 33	A~A~ITIPPEYGAHD~FL	0.000000
20 - 34	A~ITIPPEYGAHDF~LQ	0.000000
21 - 35	I~TIPPEYGAHDFL~QP	0.000000
22 - 36	TIPPEYGAHDFLQ~PP	0.000000
23 - 37	IPPEYGAHDFLQP~PE	0.000000
24 - 38	PPEYGAHDFLQPP~EL	0.000000
25 - 39	PEYGAHDFLQPPE~LT	0.000000
26 - 40	EYGAHDFLQPPEL~TE	0.000000
27 - 41	YGAHDFLQPPELT~EE	0.000000
28 - 42	GAHDFLQPPELTE~EP	0.000000
29 - 43	AHDFLQPPELTEE~PP	0.000000
30 - 44	HDFLQPPELTEEP~PE	0.000000
31 - 45	DFLQPPELTEEPP~EQ	0.000000
32 - 46	FLQPPELTEEPPE~QL	0.000000
33 - 47	LQPPELTEEPPEQ~LV	0.000000
34 - 48	QPPELTEEPPEQL~VV	0.000000
35 - 49	PPELTEEPPEQLV~VF	0.000000
36 - 50	PELTEEPPEQLVV~FP	0.000000

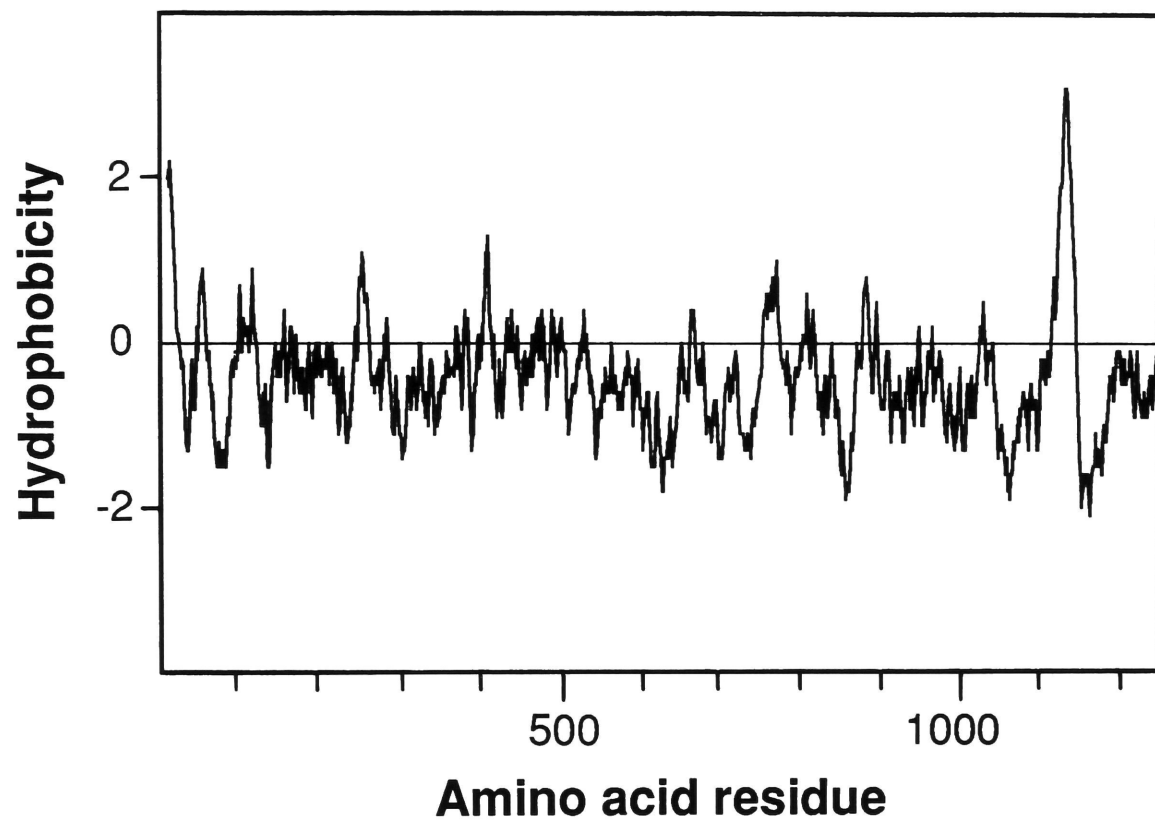
Pmax = 8.386906e+04  
ln(Pmax) = 11.337012 (cf. N.A.R., 14, 4683 (1986))





Figure 10    Hydrophobicity plot of the predicted amino acid sequence of Ng-CAM.

Hydrophobicity was calculated across a sliding window of 19 residues using the parameters of Kyte and Doolittle (Kyte and Doolittle, 1982). Amino acid residues are numbered from the translation start site. Positive values indicate hydrophobicity. Two regions of marked hydrophobicity are apparent, the first denoting the predicted signal peptide at residues 1-20, and a second segment representing the predicted transmembrane region at residues 1109-1132.







and isoleucines, and is bounded on both sides by (charged) lysine residues. The location of the putative membrane-spanning domain predicts that the Ng-CAM protein has an extracellular amino terminal region of 1109 amino acids and a cytoplasmic domain of 113 amino acids.

Palmitate has been incorporated into Ng-CAM (Sorkin et al., 1985), a moiety usually associated with the plasma membrane or adjacent cytoplasmic domain (McIlhinney, 1990). A cysteine is located within the putative transmembrane region and is a likely site for the attachment of palmitate. In several proteins, including the murine histocompatibility antigens, palmitate is associated with the transmembrane segments via a covalent thioester linkage to cysteine (Kaufman et al., 1984).

### Domain Structure

The extracellular region of Ng-CAM includes two types of repeated domains. Beginning near the amino terminus are six repeats that resemble those found in immunoglobulins, N-CAM and other related CAMs (Edelman, 1987; Williams and Barclay, 1988). These are followed by five domains that resemble each other and the type III repeats found in fibronectin (Kornblihtt et al., 1985). Each of the two types of domains have been found in several other neural CAMs.

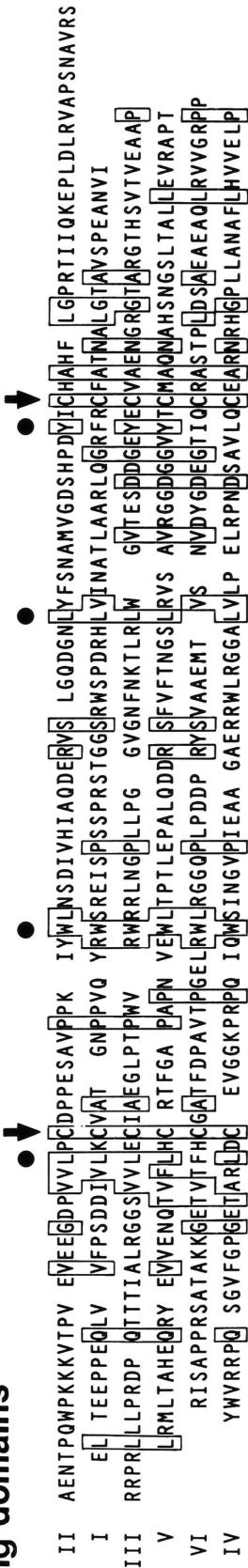
### *Ig-like Domains*

The sequences of the six immunoglobulin-like domains of Ng-CAM are aligned with each other in Figure 11. The domains were first compared to each other using the LFASTA alignment program to establish degrees of

Figure 11 Alignment of the Ig-like domains in Ng-CAM.

The Ig-like region of the Ng-CAM sequence was divided into individual domains based on similarity to immunoglobulins and L1 sequence, and aligned in an order which maximizes identity between adjacent lines (see text). The Ig domain number (I-VI) is indicated to the left of each sequence. Boxed amino acid residues indicate residues which are identical in three or more domains at that position. Positions of cysteines (arrows) and other residues (filled circles) characteristic of the C-2 type of Ig domain are indicated.

Ig domains





highest similarity, and were then displayed in the order that maximizes the identity between adjacent lines of sequence. Thus, the third Ig-like domain is more similar to the first domain than to the second, and more similar to the fifth domain than to the fourth domain; this similarity is reflected in the order in which they are displayed.

Ig-like domains within a single neural CAM are generally about 25% identical to each other. As shown in Figure 11, those in Ng-CAM domains are 25-31% identical to each other. Each of the domains contains approximately 100-110 residues and shares features with both the constant domains and with the variable domains of immunoglobulins. This "hybrid" or C2 domain is also found in other cell adhesion molecules, and has been distinguished from the V and C1 types of Ig domains (Williams and Barclay, 1988). The classifications are based on the lengths of the domains between the conserved cysteine residues, the predicted secondary structure of the domains, and the presence of particular conserved amino acids within the domains. The conserved residues in each type of Ig domain are aligned for comparison in Figure 12.

The Ig-like domains in Ng-CAM are most similar to the C2-type domains, in that the first invariant cysteine in each domain is preceded by a conserved leucine, and is followed 10-15 residues downstream by a completely conserved tryptophan. The second invariant cysteine is preceded by a highly conserved leucine and a less well conserved tyrosine, and is followed two residues downstream by an alanine. Although these residues are similar to both the C2 and V subtypes the

Figure 12 Consensus sequences of Ig domain subtypes.

Ig-like domains are grouped into three subtypes based on the distance between their characteristic cysteines, the predicted folding of their amino acid sequences, and the presence of other conserved amino acid residues. The variable (V), constant-type 1 (C1), and constant-type 2 (C2) sequences are denoted at the left. Only residues characteristic of the subtype are included in the sequence, leaving spaces where intervening nonconserved residues are located. The characteristic cysteines in each subtype are boldfaced, and the more highly conserved residues are capitalized. Lower case amino acid residues are less well conserved in that domain. The dashes indicate that the constant-type domains have fewer amino acid residues located between the cysteines than the variable domains. Underlining indicates the approximate positions of four of the predicted  $\beta$ -strands which are expected to align in antiparallel sheets in each domain. Information in this figure was derived from Williams and Barclay (Williams and Barclay, 1988).

# Ig SUBTYPE CONSENSUS SEQUENCES

V cons	V L C		f w r q	R	L i	D G Y Ca
C1 cons	v L Clv FYP		i v W g	v	s L v	Y C V H
C2 cons	v L C p		l W	r l	L L	d G Y C a n
	<u>β-strand B</u>		<u>β-strand C</u>		<u>β-strand E</u>	<u>β-strand F</u>





distance between the two cysteines in the Ng-CAM domains is 46-51 amino acid residues and therefore closer to the size of the C2 domains.

#### *Fibronectin-like Repeats*

A second set of five repeats of (90-110 amino acids each) are found in Ng-CAM immediately after the Ig-like domains. These segments resemble the type III repeats of fibronectin, and are aligned in Figure 13 to show maximal identity between adjacent domains. In addition, the tenth type III repeat of human fibronectin (Kornblihtt et al., 1985) is included for comparison.

Fibronectin-type III (Fn-III) repeats are characterized by conserved tyrosine and tryptophan residues that are separated by 45-55 amino acids. Other amino acids within the repeats are also conserved, but the homology between Fn-III domains is generally less extensive than between the Ig-like domains. In the amino terminal region of each Fn-III repeat, a tryptophan is followed by a tyrosine 10-15 residues away. In the carboxyl terminal region there is a conserved tyrosine preceded by a leucine six residues away. The first three Fn-III repeats of Ng-CAM have the conserved tryptophan and tyrosine residues, but in the fourth repeat the tryptophan is replaced by a glutamate. The fifth repeat is the least similar to the other domains, with the tryptophan and tyrosine residues both conservatively substituted by a phenylalanine.

Ng-CAM contains a single RGD sequence (amino acids 893-895) at the carboxyl terminal end of the third Fn-III repeat. The tripeptide is contained within a 29 amino acid segment that resembles the carboxyl

Figure 13 Alignment of the fibronectin-type III repeats in Ng-CAM.

The fibronectin-like region of the Ng-CAM sequence was divided into individual domains based on comparison to fibronectin and aligned in an order which maximizes identities between adjacent lines (see text). The number of each Fn-III repeat from its location in Ng-CAM (1-5) is indicated. Below the Ng-CAM sequences is listed the sequence of the tenth type-III repeat of human fibronectin as reported by Kornblihtt (Kornblihtt et al., 1985) for comparison. Asterisks denote characteristic amino acid residues in Fn-III repeats, which are also boxed in the fibronectin sequence. Fn-III repeats typically contain a tryptophan in the amino terminal region, followed 10-15 amino acid residues downstream by a tyrosine. In the carboxyl terminal region a tyrosine is preceded six amino acid residues upstream by a leucine. Other boxing indicates amino acids which are identical in a majority of the residues at that position. In several domains of Ng-CAM conservative amino acid substitutions replace the characteristic residues of Fn-III repeats. The third Fn-III repeat of Ng-CAM contains an additional segment in its center which enlarges the predicted domain compared to the others. This configuration is discussed in the text.

Fn - III domains

1	SRDLQVMEVDEHRVRLSWTPGDDHNSPIEK	FWV	EEEEEREDLQRGFGADVPQDPMTPPLPLSPY	TPPLPLSPYGRFPFR	VVAVNAYGRGEHHASAPI	ETTPAAPE
2	RNPGGVHGEGETGNLVITMEPL	PPQAWNAPWARYVQWR	PIEEPGGGPGSGGFPMAESTV	DAPPVVVGLPIPFSPFQIR	VQAVNNGAGKGPETPGVGHSGEDL	PLVYPE
3	NVGVELLNSSTVRVRMTLGGGPKELRGRLRGFRVLYMR	LGMVGRSRRQAPPDPQISPAE	DPFPFPPVAVLTVGGDARGAL	LGGLRPMRSYQIR	VLVFNGRGDGPPSEPIAFETPEGV	
4	PGPPEELRVERLLDDTALSNNERRTFKRSTGYVLR	YQQV	EPGSALPGGSVLRDPQCDLRGLNARSRYRLAL	PSTPRER	PALOTIVGSTKPEPPSP	LMS
5	RFGVGGRGGFHGAAVEFGAAQEDDVEFEV	QFMNKSTDEPWR	TSGRANSSLR	RYRLIEGLRPGTAYRVQFVGRNRSGENVAFWESEVD	INGTVVVPD	
FN	VSDVPRDLEVVAATPTSLLSISWDAPA	VTVRVRITYGETGGNSPVQ	EFTVPGSKSTATISGKPGVDV	TITVYAVTGRGDS	PASSKPI	SINVRT



terminal region of a Fn-III repeat (it also contains a tyrosine preceded six residues upstream by a leucine), and is aligned in Figure 13 to represent such a position. This alignment consequently indicates an unusually large central region in the third domain. The segment immediately preceding this 29 residue RGD-containing segment also resembles the end of Fn-III repeats, suggesting that either of the two regions may be considered the end of the repeat in Ng-CAM. It is not clear if there are size constraints in the Fn-III repeats, since the three-dimensional structure is not known. However, if this longer configuration were the structure of the third domain, then the RGD sequence is in a position equivalent to the RGD in fibronectin's tenth type III repeat (Kornblihtt et al., 1985). In fibronectin this sequence is involved in the protein's binding to a cell surface receptor (Hynes, 1987; Ruoslahti, 1988). It is possible that this sequence in Ng-CAM performs a similar binding function.

#### Organization of Ng-CAM components

The cDNA sequence of Ng-CAM encodes a single large translation product with a predicted molecular weight of 136,663 Daltons. However, Ng-CAM when isolated from cells is predominantly a 135 kD component, with lesser amounts of an 80 kD component and a 200 kD doublet. Several lines of evidence presented here provide additional support for the hypothesis that a single large translation product is processed to yield the 200 kD components, which are then cleaved into the 135 kD and 80 kD components.



The sequences of all of the Ng-CAM components are contained within the single cDNA sequence. All of the peptide sequences, including those obtained from the purified 80 kD component, are represented, and the amino terminal sequences of the 210 kD, 190 kD and 135 kD components (Figure 3) all correspond to the predicted amino terminus of the mature translocated protein (Figure 8). The amino terminus of the 80 kD component is also contained in this sequence beginning at amino acid 840, and is continuous with the 135 kD sequence preceding it. No cDNA clones were isolated which began encoding the 80 kD component precisely at its amino terminus, as would have been expected if it were encoded by a separate mRNA. ( $\lambda$ N922 ends 29 basepairs 3' from the site of the amino terminus; Figure 7).

The relationship of the cDNA clones and the cDNA sequence indicate that clone  $\lambda$ N922 and clones extending 3' from it (Figure 7) encode all but the first nine amino acids of the 80 kD component, and clone  $\lambda$ N903 and clones 5' of it are predicted to encode most of the 135 kD component (Figure 7). Probes from each of these regions could thus be used to test whether each component is encoded by the same mRNA or from different mRNAs.

#### *Ng-CAM mRNA and the Ng-CAM Gene*

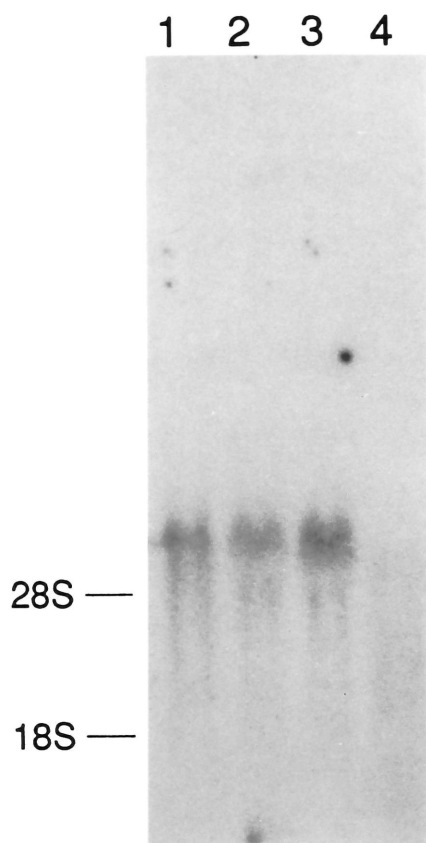
Northern analyses of the Ng-CAM mRNA also support the model that all the Ng-CAM polypeptides are produced by processing of a single translation product. Inserts from cDNA clones  $\lambda$ N903 and  $\lambda$ N902 were hybridized to RNA isolated from chicken embryonic tissues (Figure 14). These two probes were chosen because  $\lambda$ N903 encodes only sequences in the

Figure 14 Northern blot analysis of Ng-CAM.

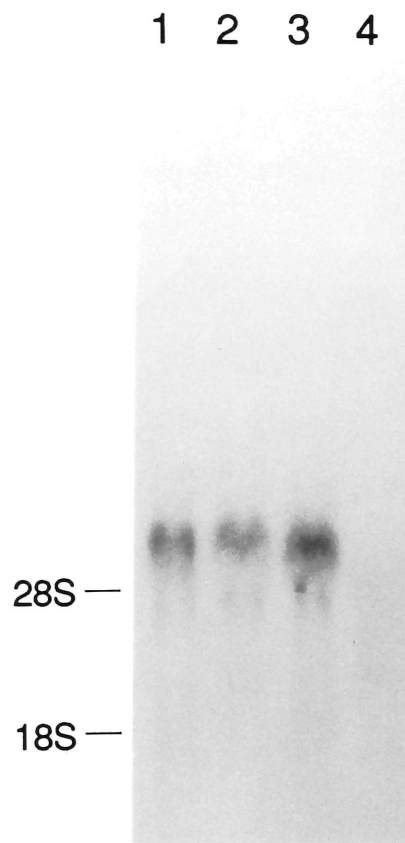
Poly (A)<sup>+</sup> RNA (1  $\mu$ g) from 12d chick embryo brain (lane 1), total brain RNA (15  $\mu$ g) from 9d (lane 2) and 12d (lane 3) embryos, and 14d chick embryo liver (lane 4) were electrophoresed on 0.8% formaldehyde gels and transferred to Hybond. RNA was hybridized with Ng-CAM cDNA probes N903 (a) and N922 (b). Positions of 28S and 18S ribosomal RNA are indicated.



**a**



**b**





135 kD region whereas  $\lambda$ N902 should encode sequences only in the 80 kD region (Figure 7,8). Both  $\lambda$ N902 and  $\lambda$ N903 hybridized to a single band at 6 kb present in total and poly A<sup>+</sup> brain RNA (Figure 14, lanes 1, 2, 3). No hybridizing RNA was detected with either probe in 14-day embryonic chicken liver (lanes 4), heart, skin, or gizzard (data not shown), tissues where Ng-CAM is not detected. The hybridizing RNA was present in 9-day and 12-day embryonic brain, in agreement with those times at which Ng-CAM protein is detected (Thiery et al., 1985). These data support the model that the 135 kD 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 of the molecule.

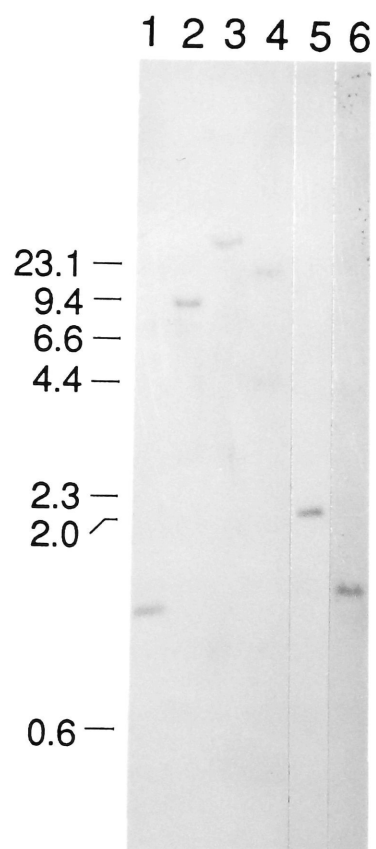
The gene for Ng-CAM was identified in Southern analyses of chicken genomic DNA (Fig. 15). The insert from  $\lambda$ N903 hybridized to single bands in BamHI, DraI, BanI, SmaI, and SstII digests of genomic DNA. These data indicate that Ng-CAM is probably encoded by a single gene, in further accord with the notion that a single Ng-CAM mRNA is transcribed. This result does not rule out the possibility that alternatively-spliced RNA transcripts exist, but suggests that the single-sized mRNA detected by Ng-CAM cDNA probes is the product of a single gene.

### *Protein Sequence*

The amino acid sequence surrounding the amino terminus of the 80 kD Ng-CAM component appears to represent a suitable environment for the proteolytic cleavage of a larger polypeptide into the two smaller chains. Eleven of the first twenty amino acids in the 80 kD component are prolines which could lead to an unusual conformation in this region.

Figure 15 Southern blot analysis of Ng-CAM.

8  $\mu$ g of adult chicken genomic DNA was restriction with HgaI (lane 1), BamHI (lane 2), DraI (lane 3), BanI (lane 4), SmaI (lane 5), and SstII (lane 6), electrophoresed on 0.7% agarose gels, transferred to Hybond and probed with Ng-CAM cDNA  $\lambda$ N906. Relative migration of  $\lambda$ -HindIII and  $\phi$ X174-HaeIII DNA molecular weight markers are indicated at the left in kilobases.





Moreover this segment is preceded by a region that contains three arginines that could be sites of proteolysis by enzymes such as trypsin. If the cleavage attacks multiple sites, the prolines would stabilize attack by aminopeptidases, so that the 80 kD component has a single amino terminus. The carboxyl terminus of the 135 kD component could be frayed; in accord with this notion, we have been unable to detect a single carboxyl terminal residue for this component using carboxypeptidases.

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. The two components probably do not overlap; all sequences of peptides isolated 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 (peptides 11-14 in Table 1) were detected only in the region on the carboxyl terminal side of this site. None of these data rule out the simpler alternative that the two components are cleaved by a relatively specific protease at the 80 kD amino terminus.

The Ng-CAM cDNA sequence does not immediately suggest the nature of the difference between the 210 kD and 190 kD forms of the mature protein (Figure 3). The detection of only one Ng-CAM mRNA suggests that it is posttranslational. In agreement with this notion, Ng-CAM is synthesized in tunicamycin-treated cells as a single 170 kD protein (Burgoon et al., 1991). Tunicamycin blocks the processing of N-linked glycosylation (Kuo and Lampen, 1974), suggesting that the difference between the 200 kD components might arise from differential





glycosylation of the 170 kD polypeptide. However, extensive alternative splicing is seen in N-CAM (Cunningham et al., 1987; Goridis and Wille, 1988) and other CAMs (Hortsch et al., 1990; Salzer et al., 1987), so it is possible that the two largest components of Ng-CAM are derived from a single transcript by alternative splicing.

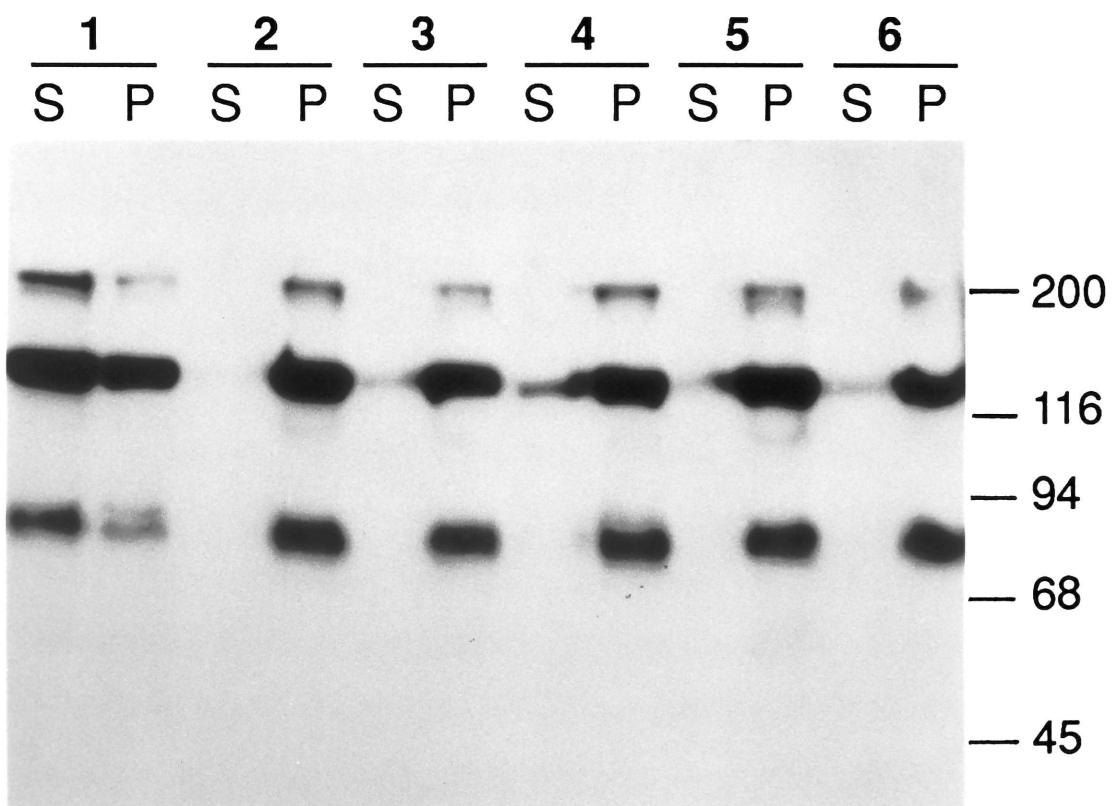
#### *Association with the Membrane*

A variety of data indicate that the 135 kD component is tightly associated with brain plasma membranes, although the sequence data indicate that it does not contain a hydrophobic segment. Detergents such as NP-40 are required to release all the components of Ng-CAM from the membrane (Figure 16, lane 1). In contrast, treatments with EDTA (lane 2), high salt (lane 3), high pH (lane 4), and low salt buffers (lane 5), which extract peripheral membrane proteins, did not release significant amounts of the 135 kD component or any other Ng-CAM species from the membranes (Burgoon et al., 1991).

Several mechanisms may account for the association of the 135 kD component with the cell membrane. It may be anchored to the membrane by phospholipid attachment (Ferguson, 1988), or it may be covalently bound to a membrane protein through a cystine disulfide bond. None of the Ng-CAM polypeptides, however, are released by phosphatidylinositol-specific phospholipase C, and treatment of brain plasma membranes with buffers containing dithiothreitol did not release the 135 kD species, or other forms of Ng-CAM (Burgoon et al., 1991). In addition, the 210 kD, 190 kD and 80 kD as well as the 135 kD components of Ng-CAM can all be efficiently reconstituted into liposomes, indicating that the 135 kD

Figure 16    Membrane association of Ng-CAM.

Freshly dissected 14day chick embryo brains were homogenized and centrifuged at 30,000x g for 30 minutes. The membrane pellet was resuspended in an equal volume of PBS. Aliquots of the membranes (150  $\mu$ l) 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<sub>2</sub>O (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 6% polyacrylamide gels, transferred to nitrocellulose and immunoblotted with 100  $\mu$ g of anti-Ng-CAM antibodies. Molecular weight standards are indicated at the right as  $M_r \times 10^{-3}$ . Figure is provided courtesy of M. Grumet and modified from (Burgoon et al., 1991).





form binds directly or indirectly to lipid membranes (Grumet and Edelman, 1988). These results suggest that the 135 kD component may be tightly associated non-covalently with the other Ng-CAM components or has another mode of attachment to the membrane. Such a complex could account for the association of the 135 kD component 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 (Burgoon et al., 1991).

#### Similarity of Ng-CAM to Other Neural CAMs

Previous studies indicated that other chicken molecules such as G4 (Rathjen et al., 1987a) and 8D9 (Lemmon and McLoon, 1986) are equivalent to Ng-CAM. From its reported amino terminal sequence, it appears that this is true for at least G4. The relationship of Ng-CAM to proteins in other species is less clear. Ng-CAM shares a number of functional and structural characteristics with L1 in mouse and the NILE glycoprotein in the rat, molecules that have previously been thought to represent the equivalents of Ng-CAM in the two mammalian species (Grumet et al., 1984a; Lindner et al., 1983; Keilhauer et al., 1985; Friedlander et al., 1986; McGuire et al., 1978; Sajovic et al., 1986; Bock et al., 1985; Fischer et al., 1986; Stallcup et al., 1985).

A comparison of the amino acid sequence of Ng-CAM to the Dayhoff protein sequence database (release 25), the translated Genbank database (release 63), and specifically to the sequences of mouse L1 (Moos et al., 1988), chicken contactin (Ranscht and Dours, 1989) and F11 (Brummendorf et al., 1989), mouse F3 (Gennarini et al., 1989) and other



neural CAMs, revealed that it was significantly similar to several members of the Ig superfamily as well as to other proteins containing the fibronectin-type III repeats. Not surprisingly, it is most similar to the neural CAMs in the Ig superfamily, and in particular to L1.

Besides sharing the greatest sequence similarity to Ng-CAM, L1 also has an equivalent domain structure, with 6 Ig-like domains, 5 Fn-III repeats, a transmembrane segment, and a short cytoplasmic region similar in size to that in Ng-CAM. Other CAMs in the Ig superfamily contain different numbers of Ig domains and Fn-III repeats; their similarities to Ng-CAM are greatest in the Ig-like and Fn-III regions, but they do not share as many residues with Ng-CAM as does L1. Moreover the cytoplasmic regions of the other neural CAMs were clearly distinct from Ng-CAM.

In order to search for other proteins that may resemble Ng-CAM to a greater extent than L1 over a smaller region of the sequence, 100 amino acid segments of Ng-CAM were also compared individually to the protein databases. Such a localized similarity might be statistically overwhelmed in other comparisons by the overall similarities of L1 or other neural CAMs across multiple domains. However no other obvious similarities were detected.

The individual Ig-like domains of Ng-CAM were more similar to the corresponding domains in L1 than to any other sequences, including other Ng-CAM sequences (Table 4). Comparison of the individual Ig-like domains of Ng-CAM to each other using the FASTA program of Pearson (Pearson and Lipman, 1988) gave modest scores (24 to 90) reflecting the general similarities among all Ig-like domains (Table 4, top). In

Table 4      FASTA comparison of Ng-CAM Ig-like domains to each other and to Ig-like domains of L1.

Top: Values indicate highest FASTA alignment scores (ktup=2 ; Pearson et al., 1988) of pairwise comparisons between the Ig-like domains of Ng-CAM.

Bottom: Values indicate highest FASTA alignment scores (ktup=2 ; Pearson et al., 1988) of pairwise comparisons between individual Ng-CAM Ig-like domains and individual L1 Ig-like domains.



TABLE 4 - FASTA COMPARISON OF Ig DOMAINS

		Ng-CAM					
		I	II	III	IV	V	VI
Ng-CAM							
	I		55	90	-	24	48
	II			29	29	86	50
	III				35	29	64
	IV					27	49
	V						32
	VI						

FASTA

		L1					
		I	II	III	IV	V	VI
Ng-CAM							
	I	244	22	104	24	27	23
	II		408	69	38	82	53
	III			240	57	82	53
	IV				299	80	23
	V					230	24
	VI						172



contrast each of the domains in Ng-CAM was most similar to the corresponding domain in mouse L1, with significantly higher scores (172 to 408) (Table 4, bottom). Because Ng-CAM was most similar to L1, their entire sequences were aligned using the LFASTA program (Pearson et al., 1988) to compare them in more detail (Figure 17). The two CAMS shared many identical residues throughout their sequences but the extent of the similarities varied dramatically from region to region and were surprisingly few for two molecules believed to be equivalent in chickens and mice.

### *Ig-like Domains*

Similarities in the Ig-like domains of Ng-CAM and L1 are between 40% identical in the fifth domain to 66% identical in the second domain (Figure 17a). Many of the identities are clustered around the pairs of cysteines in each domain, and occur as alternate identical residues (most clearly seen in the fourth domain), probably reflecting the alternating polar/apolar residues of  $\beta$ -pleated sheets in the Ig-like domains. At other sites in this region Ng-CAM and L1 are identical over 5-10 consecutive residues. These segments probably serve a common structural or functional role in the two proteins. Although Ng-CAM and L1 are less similar in the central region of each Ig domain, the second and third Ig-like domains contain segments of striking identity in their central areas. Several Ig-like CAMs have been shown to mediate adhesion through their Ig domains (Cunningham et al., 1987; Doyle et al., 1990; Cole and Akeson, 1989). In N-CAM this region has been implicated in homophilic binding (Cunningham et al., 1987) and the binding of heparin

Figure 17 Comparison of chicken Ng-CAM domains to corresponding domains in mouse L1.

Deduced amino acid sequences of Ng-CAM and L1 were initially aligned pairwise using the LFASTA program (ktup=2, Pearson et al., 1988) to achieve the best overall match. A minimal number of gaps were introduced into either sequence to maximize the identities between the sequences. The position of the first amino acid residue of each domain is given. 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, with residues which are identical to L1 underlined. (c) Alignment of the transmembrane and cytoplasmic segments in Ng-CAM to corresponding segments in L1.





(Cole and Akeson, 1989; Reyes et al., 1990). This similarity in Ng-CAM and L1 may reflect a common adhesion site.

### *Fn-III Repeats*

The five fibronectin-type III repeats of Ng-CAM are also similar to the five corresponding segments in L1 (Fig. 17b), although these similarities are weaker than those between Ig domains. The identities range from 50% in the second domain to only 15% in the fifth domain. The amino terminal sequence of the 80 kD component of Ng-CAM is in the middle of the third repeat (indicated by arrow); the amino acids surrounding this site differ significantly from the corresponding region of L1. Although L1 contains several similar basic residues at positions predicted to provide proteolytic cleavage sites in Ng-CAM, there are only two identical amino acid residues. Furthermore L1 lacks the prolines abundant in the amino terminal region of Ng-CAM's 80 kD component. The marked difference between the two CAMs in this region is consistent with the observation that Ng-CAM appears predominantly as the 135 kD component, while in L1 the 200 kD species predominates.

At the site in L1 corresponding to Ng-CAM's 80 kD amino terminus, the L1 sequence contains the segment H1HKSHIVV, which is preceded by an arginine. Trypsin can cleave L1 into two components of the same approximate size as Ng-CAM (135 kD/80 kD) (Faissner et al., 1985; Rathjen and Rutishauser, 1984) and this could be the site of such cleavage. However the sequence in this region of L1 does not contain the abundance of prolines found in Ng-CAM, so if L1 is cleaved at this site, its 80 kD component could be frayed by aminopeptidases; this





could in turn cause it to be more diffuse on SDS-PAGE and more difficult to detect, in accord with biochemical studies of the protein (Rathjen and Rutishauser, 1984).

An additional twenty-nine amino acids are located in the third Fn-III repeat of Ng-CAM that has no counterpart in L1. In the comparison with L1, most of this segment is unique to Ng-CAM and aligns most favorably with the carboxyl terminal region of the preceding Fn-III repeat, as indicated in Figure 17b. The segment contains a tyrosine preceded six residues by a leucine which are characteristic of Fn-III repeats. In comparing the Fn-III repeats within Ng-CAM (Figure 14) this segment aligned most appropriately with the end of the other Fn-III repeats and a segment of amino acids in the center of the repeat appeared to be the least similar to the other repeats. From these comparisons it is not possible to discern which 29 amino acids were added to Ng-CAM or what their placement might be in the domain. Those residues at the end of the repeat which distinguish Ng-CAM from L1 could account for some of the differences in the activity of the two proteins, because this segment contains the RGD sequence.

#### *Cytoplasmic Region*

Ng-CAM is more similar to L1 in its predicted transmembrane and cytoplasmic regions than in any other region (69% identical, Figure 17c). The 23-residue transmembrane segments are 81% identical, differing at only four positions. Beginning in the middle of the transmembrane segment and continuing into the cytoplasmic domain there is a stretch of 50 amino acids (amino acids 1121 to 1170) that are 100%



identical in the two proteins. This region is followed by a segment of 26 residues that is only 31% identical to L1 and then another 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). A search of the translated Genbank or Dayhoff databases or the PROSITE database for consensus motifs did not reveal any similarities between the cytoplasmic region in Ng-CAM and any other proteins other than L1 and neuroglian. The function of this highly conserved region of Ng-CAM and L1 is unclear, but it may be involved in interactions with a common intracellular element.

Several CAMs are known to associate with intracellular components. The large cytoplasmic domain of N-CAM has been reported to react with spectrin (Pollerberg et al., 1987). L-CAM associates with a group of cytoplasmic molecules called catenins (Ozawa et al., 1989), and may form complexes with fodrin and ankyrin (Nelson et al., 1990). In addition, L-CAM colocalizes with actin filaments near the boundaries of cells transfected with L-CAM or N-cadherin (Mege et al., 1988; Matsuzaki et al., 1990). These preliminary results indicate that intracellular functions of the neural CAMs are likely to be important for linking extracellular interactions with the intracellular activities resulting in, for example, changes of cell shape through cytoskeletal interactions.

### N-linked Glycosylation

Previous studies showed that Ng-CAM contained asparagine-linked oligosaccharides on both the 135 kD and 80 kD components (Grumet and



Edelman, 1984). The cDNA sequence reveals that there are 12 potential sites for N-linked glycosylation (Asn-Xaa-Ser/Thr; Hubbard and Ivatt, 1981) distributed throughout the extracellular region, with eight in the 135 kD region and four in the 80 kD region (Figure 8). There are no potential sites for N-linked glycosylation in the predicted cytoplasmic region.

Six of the potential N-linked glycosylation sites in Ng-CAM (Asn 76, 267, 451, 477, 691, and 798) are in corresponding positions in L1, suggesting that they may be glycosylated in both proteins. 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 reported interactions between L1 and N-CAM (Kadmon et al., 1990a).

Both Ng-CAM and L1 contain the HNK-1 carbohydrate epitope (Abo and Balch, 1981; Grumet et al., 1984b; Kruse et al., 1984), which contains sulfated glucuronic acid (Shashua et al., 1986). A number of other cell adhesion related molecules have been reported to contain this epitope, including N-CAM (Grumet et al., 1984b), MAG (McGarry et al., 1983), and the extracellular matrix molecule cytotactin/J1 (Grumet et al., 1985; Kruse et al., 1985) as well as several glycolipids (Ilyas et al., 1984; Chou et al., 1986). The HNK-1 epitope has been reported to be involved in neuron-astrocyte and astrocyte-astrocyte adhesion (Keilhauer et al., 1985), and in interactions between neurons and the extracellular matrix (Riopelle et al., 1986). However, other studies have suggested that specific N-linked carbohydrates are not involved in neuron-Schwann cell



interactions or neurite outgrowth (Ratner et al., 1986). It is not clear what the function of any of the carbohydrate epitopes on Ng-CAM may be, but perhaps the HNK-1 epitope influences or modulates particular binding functions in the CAMs possessing it, as has been shown for the polysialic acid found on N-CAM (Cunningham et al., 1983; Hoffman et al., 1986).

### Phosphorylation

Previous studies showed that the 200 kD and 80 kD components of Ng-CAM, but not the 135 kD component, can be phosphorylated *in vitro* (Grumet et al., 1984b), but it was not clear on which residues. The cytoplasmic region of the sequence includes several residues that are potential phosphorylation sites. Thr-1148 and Ser-1201 each form part of a casein kinase I consensus recognition site (Glu-Xaa-Ser/Thr; Hathaway and Traugh, 1982; Tuazon et al., 1979). The cluster of serines at 1176-1180 correspond to the general description for casein kinase I sites (multiple serines in close proximity; Kuret et al., 1985). Ser-1135 and Ser-1141 each conform to the general description for potential phosphorylation by phosphorylase kinase (clusters of basic residues near a Ser; Graves, 1983) and myosin light chain kinase (2-3 basic residues followed by Ser/Thr; Stull et al., 1986). Ser-1198 is followed by a cluster of acidic residues and could thus be phosphorylated by casein kinase II (Edelman et al., 1987b). Finally, Tyr-1165 resembles the sites autophosphorylated by oncogenic tyrosine kinases (Tyr followed downstream by Lys; Hunter and Cooper, 1985),





although Ng-CAM lacks the characteristic downstream Ala-Pro-Glu sequence present in this group of tyrosine kinases (Hanks et al., 1988).

Several of the potential phosphorylation sites in Ng-CAM are also present at corresponding positions in L1 (serines 1135, 1141, 1180, 1188, threonine 1148, and tyrosine 1165) and also correspond to consensus sites for phosphorylation. Threonine 1148 and its correlate in L1 both correspond to the consensus sequence for phosphorylation by casein kinase I, which is also capable of phosphorylating N-CAM (Mackie et al., 1988). Phosphorylation of hydroxyamino acids may serve to modulate Ng-CAM's binding to cell-surface ligands either on the same cell or on apposing cells, as well as to intracellular elements. Although several CAMs have been reported to be phosphorylated (Edwards et al., 1988; Sorkin et al., 1984; Grumet et al., 1984a; Ehrlich et al., 1986; Faissner et al., 1984; Salton et al., 1983b), there is no evidence as yet for an effect of phosphorylation on the binding activity of CAMs.

L1 is phosphorylated in solution even when purified to apparent homogeneity on affinity columns; a kinase was therefore predicted to copurify with the protein (Sadoul et al., 1989). Because of its other similarities to L1, it seemed probable that Ng-CAM might also be associated with such a kinase activity. Indeed it is phosphorylated in solution either by itself or by a kinase that copurifies with it. When Ng-CAM was immunoaffinity purified using monoclonal antibodies, the isolated protein appeared homogeneous on polyacrylamide gels when detected by either Coomassie Brilliant blue (Figure 3) or by silver



staining (Grumet et al., 1984a). Yet when freshly purified Ng-CAM was incubated with  $\gamma$ -[ $^{32}\text{P}$ ]-ATP in appropriate buffers, the 200 kD and 80 kD components of the protein were labeled (Figure 18a, lane 3). As in previous experiments the 210 kD and 190 kD bands were too broad to discern whether one or both components were phosphorylated. This intrinsic phosphorylation activity was abolished when the preparation was preheated to 100° C, and could be inhibited by the kinase inhibitor AMP-PNP (Figure 18a, lanes 1,2).

In a preliminary attempt to identify the amino acids in Ng-CAM that are phosphorylated, the protein was phosphorylated with nonradioactive ATP to maximize phosphorylation, and probed with a monoclonal antibody specific for phosphotyrosine (Figure 18b). The antibody did not recognize phosphorylated Ng-CAM protein (lane 2), although in a parallel lane it recognized the phosphotyrosine-containing EGF receptor in an extract from A431 cells (lane 1). This result suggests that phosphorylation of Ng-CAM does not occur on tyrosine; it is probably on serine or threonine, although we cannot exclude the possibility that carbohydrates are phosphorylated.

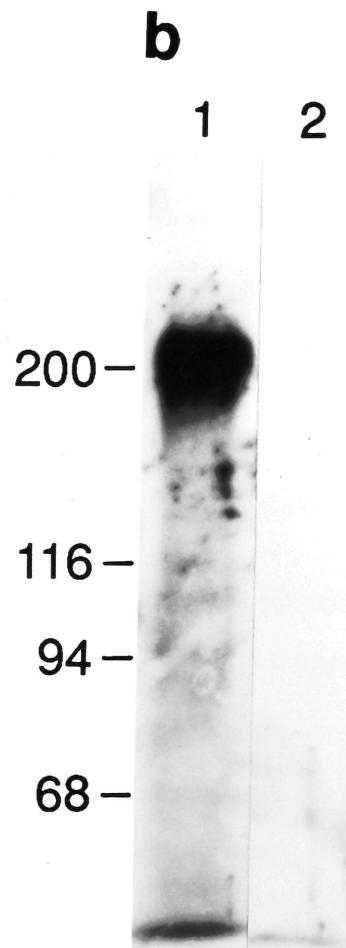
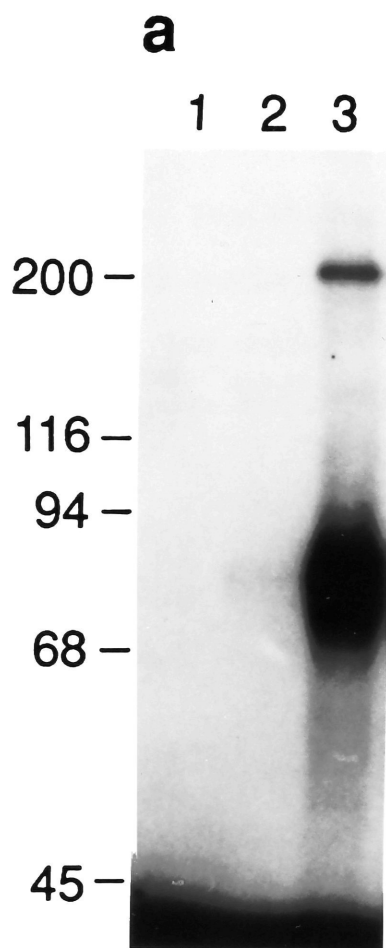
#### Relationship of Ng-CAM and L1

Ng-CAM and L1 display many similarities in their function, distribution, and structure. Despite these similarities the overall sequence identity of the two proteins is surprisingly low: only 40%. Compared to other CAMs in the two species, this is the most difficult observation to reconcile with the idea that Ng-CAM and L1 are equivalent CAMs in chicken and mouse, because other CAMs considered to be

Figure 18 Immunoaffinity purified Ng-CAM is associated with a kinase activity.

The kinase activity associated with purified Ng-CAM and the amino acid residues phosphorylated were characterized. (a) Ng-CAM purified on an affinity column of  $\alpha$ -Ng-CAM monoclonal antibodies was incubated with  $\gamma$ - $^{32}\text{P}$ -ATP, resolved on 6% acrylamide gels and autoradiographed (lane 3). Parallel samples were preincubated at 100 °C for five min (lane 1) or preincubated with the kinase inhibitor AMP-PNP before addition of  $\gamma$ - $^{32}\text{P}$ -ATP (lane 2). Relative migration of molecular weight standards are indicated at the left in  $M_r \times 10^{-3}$ .

(b) An extract of A431 cell membranes containing EGF receptor (lane 1) was resolved on 6% acrylamide gel in parallel with 5  $\mu\text{g}$  of immunopurified Ng-CAM (lane 2), transferred to nitrocellulose and immunoblotted with monoclonal antibodies that specifically recognize phosphotyrosine. Relative migration of molecular weight standards are indicated at left as  $M_r \times 10^{-3}$ .





equivalent in the two species are much more similar as shown in Figure 19. For example chicken and mouse N-CAM sequences are 81% identical, and chicken contactin/F11 is 78% identical to its presumed murine homolog F3. Moreover, rat TAG-1, which is clearly distinct from F3 and contactin, is ~50% identical to both F3 and contactin/F11 sequences. These data indicate 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 chicken Ng-CAM and mouse L1. Based on these comparisons Ng-CAM and L1 appear to be different, but closely related CAMs, which in turn suggests that there may be another molecule in the chicken that is more similar to L1.

#### Search for Other Similar Molecules

A preliminary search for a "chicken L1" was conducted using the polymerase chain reaction. An antisense oligonucleotide primer was synthesized based on the L1 sequence in the cytoplasmic region and a sense oligonucleotide was chosen in the region of the fifth Fn-III repeat, a region over which Ng-CAM and L1 are very dissimilar. The primers were used in PCR reactions with first-strand cDNA from either chicken brain RNA or mouse brain RNA from embryonic ages which express Ng-CAM or L1, respectively. After one 20-cycle round of amplification a mixture of products was obtained. The mixture included a predominant component in both mouse and chicken (Figure 20, lanes 2,3) that was the size (260 basepairs) of the segment expected from the L1 sequence (Moos et al., 1988). Aliquots of each reaction mixture were individually

Figure 19 Relationships between Ng-CAM and other neural CAMs.

Protein sequences for each of the neural CAMs listed were obtained from the translated Genbank database or from the present studies and compared in a pairwise fashion using LFASTA program of Pearson (Pearson et al., 1988) (ktup=2). Scores indicate the percentage identity between the protein sequences across the number of amino acid residues indicated in parentheses. Parentheses at left indicate origin of protein; C = chicken, M = mouse, I = insect, R = rat.



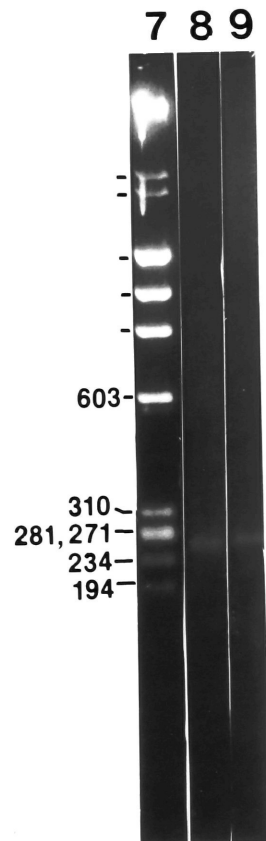
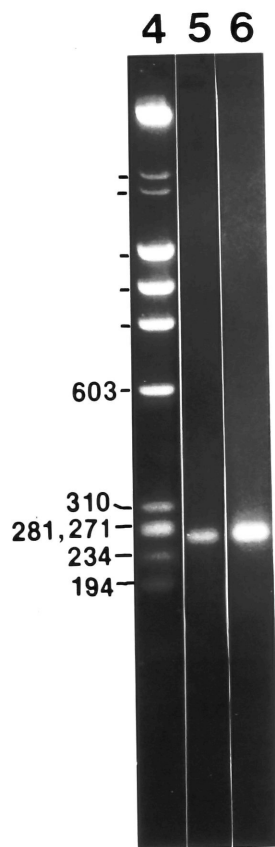
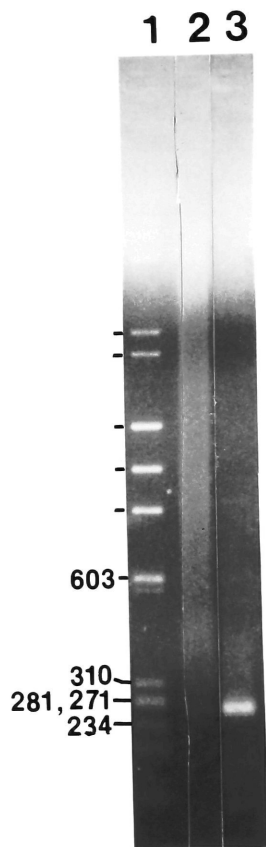
	Ng-CAM	L1	Contactin/ F11	F3	TAG-1	Chick N-CAM	Mouse N-CAM	Neuro- glian
<b>Ng-CAM (C)</b>	—							
<b>L1 (M)</b>	<b>40</b> (1284)							
<b>Contactin/F11 (C)</b>	<b>27</b> (815)	<b>28</b> (882)						
<b>F3 (M)</b>	<b>27</b> (873)	<b>29</b> (886)	<b>78</b> (979)					
<b>TAG-1 (R)</b>	<b>29</b> (865)	<b>29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





Figure 20 Polymerase Chain Reaction (PCR) for L1 equivalent in chicken.

PCR reactions were primed with oligonucleotides derived from mouse L1 sequence and electrophoresed on agarose gels. The starting template used in the PCR reactions for primer extension were as follows: first strand cDNA synthesized from 14d embryonic chick brain RNA (lane 2) or first strand cDNA synthesized from postnatal mouse RNA (lane 3); 5% aliquots of the first PCR of chicken (lane 5) and mouse (lane 6); aliquots of the reamplified 260 basepair fragments from the previous chicken PCR (lane 8) and previous mouse PCR (lane 9).  $\lambda$ -HindIII and  $\phi$ X174-HaeIII DNA standards were electrophoresed in parallel (lanes 1, 4, 7), and the sizes of the smaller standards are indicated at the left in basepairs.





reamplified through another 20 cycles of PCR. Upon reamplification a single 260 bp cDNA was the only product detected in each species (Figure 20, lane 5,6). These products were excised from the gel and successfully reamplified in subsequent PCR reactions (lanes 8,9).

The PCR product derived from the mouse RNA was subcloned and sequenced, and matched the sequence expected from mouse L1 exactly. The chicken-derived product was also subcloned and sequenced. Although it was similar in size to the product from mouse L1, and contained the two (mouse) oligonucleotide primer sequences at its ends, the intervening sequence was completely unrelated to L1 (9% identical). When this sequence was translated in each of the six reading frames, it also differed from the L1 sequence completely, except at the ends which corresponded to the mouse L1 oligonucleotides used for the PCR.

One or more of the other products from the original chicken PCR may be more similar to the mouse L1 sequences. They may have failed to reamplify because they require different optimal PCR conditions. However, these products differed from the predicted size suggesting that they are not closely related to L1. This experiment was not an exhaustive attempt to find a "chicken L1". The other PCR products should be sequenced, and other L1-based oligonucleotides should be synthesized and used in similar PCR reactions. Moreover the primer sequences we used may not have been conserved between L1 and its presumed equivalent in chicken, and it is possible that a different pair of oligonucleotides would allow us to detect an L1 equivalent in chicken.





By itself, 40% sequence identity does not rule out that Ng-CAM and L1 are equivalent molecules, because there are mechanisms such as gene duplication and genetic drift that could rapidly introduce such differences (Futuyma, 1986). Many extended segments in the two proteins are highly similar, particularly in the transmembrane and cytoplasmic segments. Less similar segments of the proteins may represent regions that are not involved in functional interactions, or have evolved differently to serve different functions in the two species.

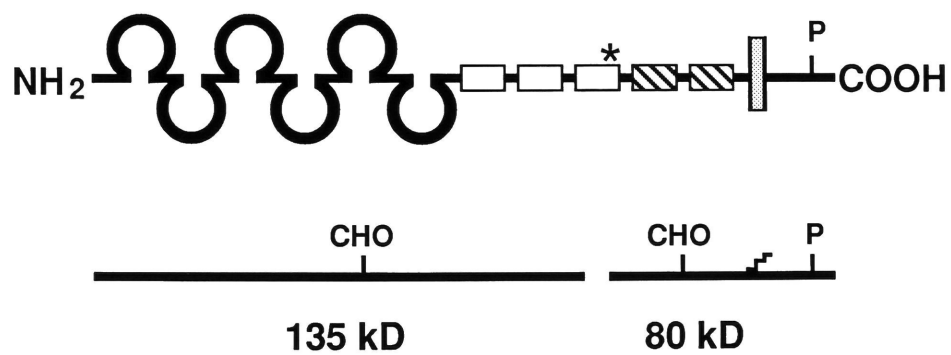
### DISCUSSION

We have isolated and sequenced cDNA clones encoding chicken Ng-CAM. From the data I propose a structural model for the molecule (Figure 21), which suggests that Ng-CAM is synthesized as a single large translation product from a single mRNA, and processed to the 190 kD and 210 kD forms. The predominant 135 kD form of the protein is derived by cleavage from the larger components, and contains the six Ig-like domains and two and a half of the Fn-III repeats. The 80 kD component contains the last two and a half Fn-III repeats, the transmembrane region and cytoplasmic domain.

Both the Ig-like domains and the Fn-III repeats could be involved in Ng-CAM's cell binding activities. In many CAMs in the N-CAM family, adhesion is mediated homophilically between Ig-like domains of the same CAM (e.g., N-CAM, P<sub>0</sub>) on apposing cells or heterophilically between the Ig-like domain of a CAM on one cell (e.g., I-CAM-1) and another ligand (e.g., an integrin) on another cell (Staunton et al., 1988). Ng-CAM,

Figure 21 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 of Fn-III repeats, and indicate that the fourth and fifth Fn-III are less conserved (see text). Asterisk denotes the location of a single RGD tripeptide in the Ng-CAM sequence. The predicted transmembrane region is indicated by a stippled vertical bar. Linear representations suggest that the 135 kD and 80 kD components are generated proteolytically from the 200 kD components. General locations of carbohydrates (CHO), phosphate (P), and palmitate (ladder) are indicated.





however, also has an RGD sequence in a Fn-III repeat and at a location comparable to that of the RGD in fibronectin that allows fibronectin to bind to cells. In the largest form of Ng-CAM this sequence is in the middle of the molecule, where it might not be effective in binding. Cleavage of the 80 kD component, however, leaves a membrane-anchored segment of the molecule with the RGD sequence near the distal end of the extracellular portion. This component of Ng-CAM may contain functions distinct from those predicted in the Ig-like region.

Ng-CAM mediates neuron-neuron adhesion apparently by a homophilic mechanism; it is most likely that this mechanism involves interactions between the Ig-like domains. Studies showed that reconstituted lipid vesicles containing Ng-CAM aggregate. In addition, when neurons and Ng-CAM-coated beads were differentially labeled and either the cells or the beads were preincubated with antibodies against Ng-CAM, coaggregation of the cells and coated beads was inhibited (Grumet and Edelman, 1988). This finding indicated that the ligand involved on neurons as well as the beads was Ng-CAM, and that neurons may be bound by Ng-CAM homophilically. The high levels of expression of Ng-CAM on neurites (relative to Ng-CAM expression on neuronal cell bodies) makes it likely that fasciculation of neurites involves homophilic interactions of Ng-CAM as well.

Ng-CAM's ability to mediate neuron-glia binding must be heterophilic because it does not appear on glia. The receptor on glia is unknown, but it could be an integrin. Furthermore, the integrin could bind the Ig-like domains (comparable to the interactions between ICAM-1 and LFA-1; Staunton et al., 1988) or might involve the RGD



sequence (comparable to the interaction between fibronectin and integrin; Hynes, 1987). The latter possibility is attractive because it could account for the ability of Ng-CAM to bind neurons to glia and explain why L1, which lacks an RGD sequence in its Fn-III repeats, does not. Moreover if the RGD of Ng-CAM has any cell-binding activity, it suggests that there may be a functional significance to the cleavage of the 80 kD component, either to activate this site or to separate its activity from that of the Ig-like domains.

In fibronectin the amino acid residue following the RGD sequence is also related to the efficacy of binding. While RGD is absolutely necessary, the next residue may be substituted by other amino acids with varying effects on binding. Studies have tested the effects of RGD<sub>X</sub>-containing peptides on inhibiting fibronectin adhesion to cells, where X was one of fifteen different amino acid residues. Substitutions at the fourth position by many amino acids resulted in activity (Pierschbacher and Ruoslahti, 1987). However a glycine which is present in Ng-CAM's sequence (RGDG) was not tested. Nevertheless the neutral effect on adhesion from substitution by the similar residue alanine suggests that an RGDG sequence would be functional, particularly in the context of a fibronectin-type III repeat. Ng-CAM and TAG-1 (Furley et al., 1990) are the only CAMs known to contain RGD sequences in Fn-III repeats. The function of this RGD sequence in either protein is as yet unknown, but the availability of Ng-CAM cDNA clones for transfection into cells provides a means for testing such sequences for activity.

The RGD sequence in Ng-CAM may be located on a distinct segment within or near the third Fn-III repeat, raising the possibility that it





may be alternatively spliced in different Ng-CAM mRNAs. Most type III repeats in rat fibronectin are encoded by two discrete exons each (Schwarzbauer et al., 1987); several are known to be alternatively spliced as pairs of exons encoding structural units (Jones et al., 1989; Kornblihtt et al., 1985; Schwarzbauer et al., 1987). Although the structure of the gene for Ng-CAM has yet to be determined, it is reasonable to suppose that the Fn-III repeats of Ng-CAM are also encoded by two exons each. In Ng-CAM, the RGD-containing segment and each of the fibronectin-type III repeats (as shown in Figure 17) ends with a codon containing the nucleotides AG, present at most exon/intron borders. Each of these segments may be predicted to end as discrete exons in the gene. If the RGD-containing segment were a discrete exon, it would be sufficiently small that its differential splicing into Ng-CAM mRNA would be difficult to detect on RNA transfer blots.

Only two Ng-CAM cDNA clones encoding the RGD region were isolated ( $\lambda$ N922,  $\lambda$ N909); both clones encoded the RGD sequence. Nevertheless, it is intriguing to speculate that this segment may be alternatively spliced and confer functional differences to Ng-CAM. Such splicing may result in site-specific expression of particular forms of the protein. For example, in fibronectin the pattern of Fn-III alternative splicing is altered during wound healing and after transformation by oncogenes (French-Constant and Hynes, 1989; Schwarzbauer et al., 1987), suggesting that changes in the type III repeats may influence the functions of proteins expressing them.

The detailed sequence comparisons of Ng-CAM and L1 suggest that they may not be equivalent molecules in the two species. There are



regions of great similarity between the two proteins (cytoplasmic region), but also regions of striking differences (Fn-III repeats). Moreover, compared to the similarities of other Ig-like neural CAMs across species (Figure 19), Ng-CAM and L1 are less similar than expected for two molecules previously considered equivalent in the two species.

A partial sequence of rat NILE has been reported (Prince et al., 1989), and was found to encode a protein sequence 97% identical to a portion of the cytoplasmic region in L1. This sequence is compared to the corresponding sequence of L1 and Ng-CAM in Figure 22. Ng-CAM in this region is 67% identical to each of the two rodent sequences and has a short segment (four residues) not present in the mammalian sequences. Recently a partial sequence of what is presumed to be human L1 has been reported in this region (Harper et al., 1991). It is 96% identical to L1 over the entire cytoplasmic region, but the available extracellular sequence (an adjacent fibronectin-type III repeat) is only 55% identical. Over the same Fn-III repeat Ng-CAM is only 15% identical to either the human or the mouse sequence. Such differences in specific segments make it difficult to discern which of these are equivalent molecules in the various species, and quantitative criteria will eventually have to be established.

The simplest interpretation of the data to date is that Ng-CAM and L1 are distinct but closely-related molecules. In support of this notion we have isolated cDNA clones encoding a different chicken protein that is as similar to both Ng-CAM and to L1 as they are to each other; Nr-CAM does not appear to be the chicken equivalent of L1. This molecule is described in the next chapter.

Figure 22    Comparison of available NILE protein sequence to Ng-CAM and L1.

The deduced amino sequence of NILE (Prince et al., 1989) were aligned by visual inspection to the similar regions of the published L1 sequence (Moos et al., 1988) and the Ng-CAM sequence. Symbols between the lines indicate amino acid residues which are identical between the two proteins. Numbers at the left indicate the positions of the amino terminal residues in each line in the original sequences.

Ng-CAM 1146	EDTQVDSEARPMKDETFGEYRSLESEAEKGSASGSGAGSGVGSPGRGPCAAGSED	
	:::::::::::::::::::::::::::: :	:
L1 1141	EDTQVDSEARPMKDETFGEYRSLESDNEEK AFGSSQPSLNGD	IKPLGSDD
	::: :::::::::::::::::::::::::::::: :	:::::
NILE (1)	EDTEVDSEARPMKDETFGEYRSLESDNEEK AFGSSQPSLNGG	IKPLGSDD

Ng-CAM 1201	SLAGYGGSGDVQFNEDGSFIGQYRGPGAG
	::: :::: :::::::::::::: :
L1 1191	SLADYGGSDVQFNEDGSFIGQYSGKKEK
	:::::::::::::::::::::::::::: :
NILE (51)	SLADYGGSDVQFNEDGSFIGQYSGKKKK



Nr-CAM AND SUBGROUPS OF NEURAL CAMs

While using antibodies against Ng-CAM protein to screen an embryonic chick brain cDNA expression library, an antibody-positive cDNA clone was isolated and found to encode a protein sequence that was similar to Ng-CAM and to Ll. Clones encoding the entire sequence of the new molecule were subsequently isolated. Because of its similarities to Ng-CAM, the new molecule was called the Ng-CAM related cell adhesion molecule or Nr-CAM. Its structure, pattern of expression, and close relationship to other neural CAMs are all consistent with the hypothesis that Nr-CAM functions in cell adhesion during neural development. Furthermore, its similarity to both Ng-CAM and Ll suggest that all three CAMs comprise a subgroup of closely related neural CAMs.

Isolation and Characterization of cDNA Clones

A cDNA clone ( $\lambda$ 701) was isolated from an embryonic chick brain cDNA library using polyclonal antibodies (#810) against denatured Ng-CAM protein (Figure 23a). In subsequent screens clone 701 also reacted strongly with a different antibody (#830) raised against native Ng-CAM (Figure 23a). Analysis of the sequence of clone 701 indicated that the it contained a single open reading frame encoding a 105 amino acid sequence (Figure 23b) that was 40% identical to a portion of the cytoplasmic region of Ng-CAM (Fig. 23c), with one twelve-residue segment identical to Ng-CAM. The sequence was also very similar to the sequence of Ll and NILE in rodents (Moos et al., 1988; Prince et al., 1989).

Figure 23 Isolation of the first Nr-CAM cDNA clone.

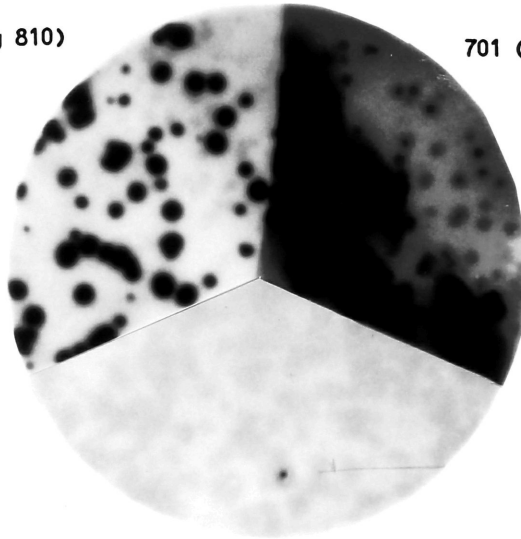
Polyclonal antibodies raised against denatured Ng-CAM (# 810) and native Ng-CAM (# 830) each recognized the fusion protein from a cDNA clone (701) distinct from Ng-CAM cDNA clones. (a) Autoradiograph showing reactivity of antibody #810 to clone 701 and nonreactivity to a different clone (18A), and reactivity of antibody #830 to clone 701. (b) Sequence of clone 701 is represented in the 5' to 3' direction. The first seven nucleotides at the 5' end are identical to the EcoR1 linkers ligated to cDNA during library construction and are shown in lower case. The single open reading frame encoded by the entire cDNA sequence is shown in single letter amino acid code. (c) Alignment of the amino acid sequence encoded by clone 701 with the corresponding region in Ng-CAM to maximize identities. Symbols indicate identical residues in the two sequences, revealing that Ng-CAM and this new peptide sequence are 40 % identical.



**a**

701 ( $\alpha$ Ng 810)

701 ( $\alpha$ Ng 830)



18A ( $\alpha$ Ng 810)

**b**

*K Y P V K E K E D A H A D P E I Q*  
 gaattccGCAAAATATCCAGTGAAGGAAAAGGAGGATGCACATGCTGATCCAGAAATACAG  
 20 40 60  
  
*P M K E D D G T F G E Y S D A E D H K P*  
 CCTATGAAGGAAGATGATGGAACATTTGGTGAATACAGTGATGCAGAGGACCATAAACCT  
 80 100 120  
  
*L K K G S R T P S D R T V K K E D S D D*  
 CTAAAAAAGGAAGTGGGACACCGTCAGACAGAACTGTGAAAAAGAAGACAGTGATGAT  
 140 160 180  
  
*S L V D Y G E G V N G Q F N E D G S F I*  
 AGTTTAGTTGACTATGGAGAAGGTGTAATGGCCAGTTCAATAGGATGGCTCCTTTATT  
 200 220 240  
  
*G Q Y S G K K E K E P A E G N E S S E A*  
 GGACAATACAGCGGTAAAAAAGAGAAAGAACCTGCAGAGGAAATGAAAGTTCTGAGGCT  
 260 280 300  
  
*P S P V N A M N*  
 CCTTCTCCTGTAAATGCCATGAATC  
 320

**c**

	10	20	30	40	50
Nr 701	KYPVKEKEDAHADPEIQPMKEDDGTGFEYS	DAEDHKPLKKGS	RTPSDRTV		
	:: ::::: : : ::	::::: : : ::	:		
Ng-CAM	KYSVKEKEDTQVDSEARPMKDE	TFGEYRSLESEAEGKSASGSGAGSGVGSPPGRGP			
	1150	1160	1170	1180	1190
	60	70	80	90	100
Nr 701	KKEDSDDSLVDYGEVNGQFNEDGSFIGQYSGKKEKEPAEGNESSEAPSPVNMN				
	: :: : :	::::: : : :	:	:	:
Ng-CAM	CAAGSEDSLAGYGGSGDVQFNEDGSFIGQYRGPGAGPGSSGPASPCAGPPLD				
	1200	1210	1220	1230	1240



The immunological cross-reactivity and the sequence similarities indicated that the cDNA encoded a portion of a previously unknown protein in the chicken resembling Ng-CAM that may represent the Ll equivalent in the chicken or be a new neural cell adhesion molecule. It was therefore characterized more fully in collaboration with M. Grumet and V. Mauro in the laboratory (Grumet et al., 1991). This section of the thesis provides an overview of the sequence and characterization of the protein which we have designated Nr-CAM, and its similarity to Ng-CAM and Ll.

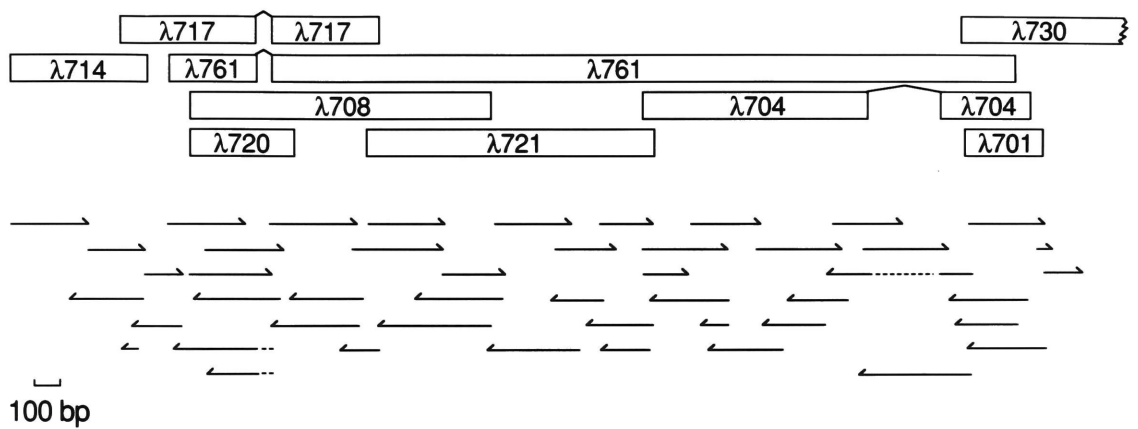
Clone 701 was used as a probe to isolate larger cDNA clones from embryonic chicken brain cDNA libraries. Clone 730 (1.6 kb) extended in the 3' direction, and contained a translation termination codon and clone 704 (1.4 kb) extended in the 5' direction (Figure 24). Additional cDNA clones extending 5' from clone 704 were obtained from different cDNA libraries using cDNA fragments as probes. Nine overlapping cDNA clones were isolated and sequenced in both directions, resulting in a DNA sequence with a single open reading frame containing 1283 amino acid residues.

#### Potential Alternative Splicing

In several areas of the sequence, two sets of cDNA clones were obtained. In one region clone  $\lambda$ 761 contained 279 bp not present in another clone ( $\lambda$ 704), but was otherwise identical in the same region (Figure 24). In addition, in the 5' region two clones were isolated,  $\lambda$ 708 and  $\lambda$ 720, which contained a 57 basepair segment not present in the same region of clones  $\lambda$ 717 and  $\lambda$ 761; outside these segments the

Figure 24 Schematic representation of Nr-CAM cDNA clones and the sequencing strategy.

The nine cDNA clones used to determine the sequence of Nr-CAM are indicated (to scale) in a 5' to 3' direction (left to right). The cDNA clones  $\lambda$ 704,  $\lambda$ 717, and  $\lambda$ 761 do not have sequences found in other clones and the gaps in these clones are indicated by single angled lines. Direction and extent of sequencing are indicated by arrows and gaps are indicated by dashed lines. Size bar is located at lower left in basepairs.





corresponding sequences of the clones were identical. The DNA sequences bordering these two segments did not resemble intron/exon junctions, indicating that the additional segments did not represent unspliced introns from heterogeneous nuclear RNA. Moreover, both segments encoded continuous peptide sequences in frame with the surrounding sequence of the Nr-CAM polypeptide. These data suggested that Nr-CAM RNA transcripts may be alternatively spliced to produce multiple polypeptides. In support of this notion, a single cDNA clone was obtained ( $\lambda$ 761; Figure 24) which extended across both regions of difference and contained the additional 3' segment (found in  $\lambda$ 761) but lacked the 5' insert (found in  $\lambda$ 708 and  $\lambda$ 720). If both segments can be involved in alternative RNA splicing, at least four Nr-CAM polypeptides could be produced.

In Northern blots more than one mRNA species for Nr-CAM was detected, consistent with the notion that there are alternatively spliced forms of the mRNAs. cDNA probes from both clones 721 and 704 hybridized to a major species in poly(A)<sup>+</sup> RNA of ~7 kb and to at least one larger but less prevalent species. The RNA was detected in neural tissues, but it was not detected in chick liver, gizzard, heart, and skin. Levels of Nr-CAM mRNA detected in brain increased until approximately embryonic day 12, and diminished slightly by later embryonic ages. The restricted and dynamic expression of Nr-CAM mRNA in the nervous system suggested that it may play a role in neural development.





### Nr-CAM Protein

In order to characterize the Nr-CAM protein, two synthetic peptides were generated from different regions of the deduced Nr-CAM protein sequence, coupled to keyhole limpet hemocyanin, and injected into rabbits for antibody production. Two polyclonal antibodies were obtained and both specifically bound to a 145 kD species in NP-40 detergent extracts of embryonic chick brain, which was not present in extracts from other embryonic tissues. These antibodies failed to immunoprecipitate the molecule or detect its distribution in immunostaining of tissue sections. The antibodies were therefore used to detect the protein during fractionation and purification procedures.

Nr-CAM was purified from NP-40 detergent extracts of chick brain membranes by a series of chromatographic procedures. Ng-CAM, N-CAM, and cytotactin were first removed from the extracts by affinity absorption and Nr-CAM was then isolated on lentil lectin columns which bind all these molecules. The eluate was fractionated by FPLC to yield fractions containing one major component of 145 kD which was recognized specifically by the anti-peptide antibodies. Polyclonal antibodies were then generated against the purified protein, and they specifically recognized the same 145 kD protein detected by the anti-peptide antibodies and an additional minor species of 170 kD.

To verify that the 145 kD species represents the protein specified by the Nr-CAM cDNA clones, it was resolved on SDS/PAGE, transferred to Immobilon, and the amino terminal sequence was determined. This



sequence was identical to that encoded by sequence near the 5' end of the cDNA clones. From this site the cDNA clones predicted a translated polypeptide of 127-140 kD depending on the presence or absence of the putative alternatively spliced segments.

The sizes of the predicted polypeptides are smaller than the 145 kD (or 170 kD) components detected in brain, indicating that the components expressed in brain are likely to be posttranslationally modified. The fact that Nr-CAM bound to lentil lectin columns suggested that this difference could be due at least in part to glycosylation. In accord with this notion, treatment of the 145 kD species with endoglycosidase F decreased its apparent size by approximately 15 kD. In addition, it was detected on immunoblots by monoclonal antibody HNK-1 which recognizes certain N-linked carbohydrates found in several CAMs and SAMs (Grumet et al., 1984b; Grumet et al., 1985; Kruse et al., 1984; Kruse et al., 1985; Hoffman and Edelman, 1987; Rathjen et al., 1987a; Ruegg et al., 1989a; Gennarini et al., 1989), particularly those of nervous system origin.

The nature of the 170 kD component is less clear, but may represent an alternatively spliced form of Nr-CAM. It is immunologically related to the 145 kD component but it was only detected in immunoblots so we could not determine whether it had the same amino terminal sequence as the 145 kD component. RNA hybridizations with probes specific to the two variant segments of cDNA may answer this question.



### Localization of Nr-CAM Protein

The expression of Nr-CAM in different tissues during development was examined by immunoblotting with antibodies against Nr-CAM protein fragments and against the purified 145 kD species. Nr-CAM was found in neural tissues but not in non-neural tissues, and no reactivity was found in extracts of 3 day chick embryos, a time prior to differentiation of neurons. The levels of Nr-CAM detected in the brain and retina increased progressively until approximately embryonic day 10-14. In addition to the predominant 145 kD species, the 170 kD species was found in lesser amounts in neural tissues both during development and after hatching. These tissues included brain, retina, cerebellum and spinal cord (Grumet et al., 1991).

In *in vitro* cultures of brain from 9-d chicken embryos the staining pattern of antibodies against Nr-CAM were compared to the staining patterns of antibodies against N-CAM, Ng-CAM and other antigens that distinguish neurons from glia in culture (Grumet and Edelman, 1988). In such cultures, the antibodies against Nr-CAM specifically labeled the surface of some, but not all, neurons and neurites; they bound very weakly or not at all to the flat non-neuronal astroblasts. Staining of cultures prepared from retina showed the presence of Nr-CAM on most neuronal cell bodies and neurites but not on the underlying glial cells (Grumet, 1991).

It was consistently observed that while most retinal neurons were Nr-CAM positive, only some were Ng-CAM positive. Conversely, most brain neurons were Ng-CAM positive but only some were Nr-CAM positive. These



results suggest that Nr-CAM and Ng-CAM are differentially expressed on different types of neurons at different stages and locations.

#### Sequence and Model of Nr-CAM

From the cDNA sequence Nr-CAM appeared to be a transmembrane protein similar to Ng-CAM and Ll, with a single transmembrane segment and a short cytoplasmic region. The predicted model of the protein is indicated in Figure 25.

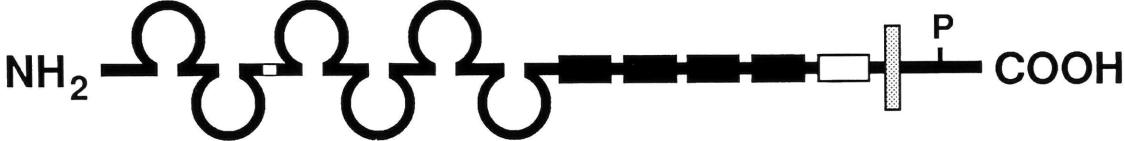
Hydrophobicity analyses revealed a single region of extended hydrophobicity in Nr-CAM. This region is located at amino acid residues 1108-1139 and resembles a membrane-spanning segment (vonHeijne, 1981; Davis and Model, 1985). Another segment was detected immediately preceding the amino terminal sequence of the mature protein; it has all the features of a signal peptide (Watson, 1984; vonHeijne, 1985). These results predict that Nr-CAM contains an extracellular amino terminal region of 1107 amino acids and a smaller cytoplasmic segment of 114 amino acids. In accord with this model, detergents were required to release the 145 kD protein from membranes whereas treatments that usually release peripherally-associated proteins including high pH and salt did not release it (Grumet et al., 1991).

The sequence of Nr-CAM indicated that the extracellular region includes two major motifs of repeated domains also found in Ng-CAM. Beginning at the amino terminus there are six C2-type Ig domains, similar to those found in Ig-like neural CAMs. These domains are followed by five Fn-III repeats. The location of the potentially spliced segments are indicated in the model, one as a segment between Ig

Figure 25 Model of the domain structure of Nr-CAM.

Ig-like domains are shown as loops in the amino terminal portion of the polypeptide and Fn-III repeats are shown as boxes. The two open boxes represent potential RNA splice variants (see text). The predicted transmembrane region is indicated by a vertical bar and the general location of potential phosphorylation sites (P) is indicated.







domains 2 and 3, and the other as the entire fifth Fn-III repeat. In these positions insertion or deletion should not significantly disrupt the structural organization of the molecule.

#### Comparison of Ng-CAM, Nr-CAM, and Ll

The six Ig-like domains in Nr-CAM are most similar to the corresponding domains in Ng-CAM and Ll. A comparison of these domains is shown in Table 5. In the Ig-like domains, Nr-CAM is as similar to Ng-CAM (32-46%) as it is to similar to Ll (26-45% identical). Ng-CAM is generally more similar to Ll (40-66% identical) than to Nr-CAM (31-46% identical). Nevertheless, the similarities of Nr-CAM to Ng-CAM are significantly greater than to other Ig-like molecules. In the Fn-III repeat region the domains of Nr-CAM and Ng-CAM range from 24% identical in the fourth domain to 41% identical in the second domain. These Nr-CAM domains are only slightly less similar to the corresponding domains of Ll (20-39% identical). This range is not as broad as that seen between the same Fn-III region of Ng-CAM and Ll (15-50% identical). In the transmembrane region and the cytoplasmic domain, both Nr-CAM and Ng-CAM are more similar to Ll (54% and 69% identical, respectively) than to each other (48% identical). These data suggest that Ng-CAM and Nr-CAM are very similar to each other and in some regions are more similar to Ll.

In order to compare the sequences of the three proteins in more detail, all three sequences are aligned in Figure 26 and positions of identical amino acids are marked. For the purposes of this thesis, I have focused on the relationship of Ng-CAM to the other two proteins.

Table 5      Comparison of structural domains to corresponding domains in other CAMs.

Top: The individual domains of Nr-CAM were aligned with the corresponding domains Ng-CAM or Ll, using the LFASTA program of Pearson (ktup=2 ; Pearson et al., 1988) and then by visual inspection to maximize identities between the sequences. Values represent the percentage identities between each Ng-CAM domain and the corresponding domain of Ll or Nr-CAM, calculated over the entire length of the domain.

Bottom: The individual domains of Ng-CAM were aligned with the corresponding domains of Nr-CAM or Ll, as above. Values represent the percentage identities between each Ng-CAM domain and the corresponding domain of Nr-CAM or Ll, calculated over the entire length of the domain.

TABLE 5 -            PERCENT IDENTITIES  
                         BETWEEN STRUCTURAL DOMAINS

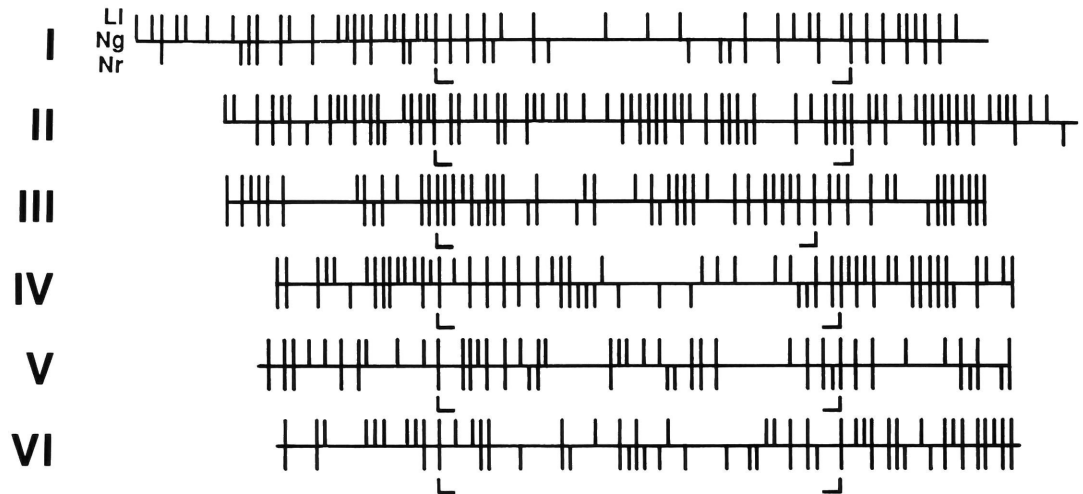
	Nr - CAM	Ng - CAM	L1
Ig	I	32	32
	II	42	43
	III	46	45
	IV	41	35
	V	31	26
	VI	36	26
Fn	1	38	39
	2	41	38
	3	38	31
	4	24	23
	5	27	20
	cyto	48	54

	Ng - CAM	Nr - CAM	L1
Ig	I	32	41
	II	42	66
	III	46	52
	IV	41	52
	V	31	40
	VI	36	43
Fn	1	38	41
	2	41	50
	3	38	26
	4	24	28
	5	27	15
	cyto	48	69

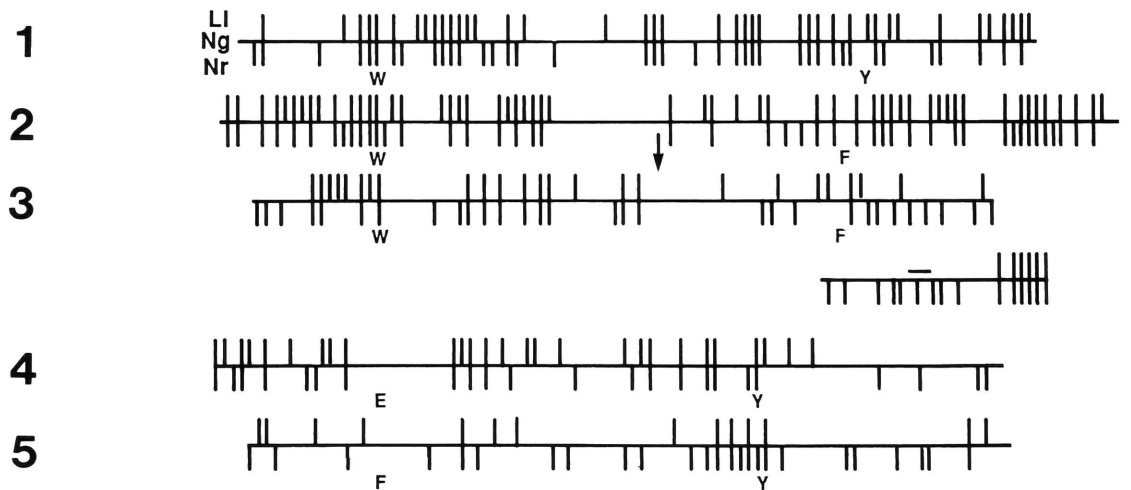
Figure 26 Ng-CAM protein sequence compared to L1 and Nr-CAM.

The protein sequence of Ng-CAM is divided into its various domains (as in Figure 18) and represented by the horizontal baseline. Brackets indicate positions of cysteines predicted to be disulfide linked within each Ig-like domain. Single letter amino acids indicate the positions of Ng-CAM residues characteristic of the Fn-III repeats (or substituted residues at the same positions). The horizontal bar in the Fn-III region indicates the position of the RGD tripeptide sequence in Ng-CAM. Each vertical bar extending above the baseline indicates an amino acid residue identical in Ng-CAM and L1. Each vertical bar extending below the baseline indicates an amino acid residue identical in Ng-CAM and Nr-CAM.

### Ig domains



### Fn-III domains



### Cytoplasmic region







Therefore the Ng-CAM protein sequence (Figure 17) was divided into its domains, and represented by the horizontal baseline. It was compared with the entire sequence of L1 (as also indicated in Figure 17), and with that form of Nr-CAM that shows the most similarity with Ng-CAM and L1 (i.e., it included the fifth Fn-III repeat but excluded the short segment between the second and third Ig-like domains). In this graphic representation identical residues between Ng-CAM and the two other sequences are indicated by vertical bars, so that the patterns of similarities among the three proteins can be discerned.

In the Ig domains, the characteristic cysteines are present in all three proteins (brackets in Figure 26a), as are most of the other residues typical of the C2 Ig domains. In this region, Ng-CAM resembles L1 slightly more than Nr-CAM, as reflected in the extended blocks of identity in domains II and IV, as well as in the numerical comparisons in Table 5. Among all of the Ig domains of Nr-CAM, Ng-CAM, and L1, the most highly conserved regions appear as patterns of alternating amino acids surrounding the highly-conserved cysteines, which probably reflect conserved residues that are important for folding or maintaining the  $\beta$ -pleated sheets found in Ig domains.

An additional highly-conserved segment is found in the middle of the second Ig domain which also shows some similarity to the second Ig domain in N-CAM. In N-CAM, this region is believed to be important for homophilic binding (Cunningham et al., 1987) and for the binding of heparin (Cole and Akeson, 1989; Reyes et al., 1990). This region of Ng-CAM, Nr-CAM and L1 does not contain the putative consensus sequence



for heparin binding found in N-CAM (Reyes et al., 1990), but may play a role in cell binding mediated by these molecules.

It is intriguing that one of the putative spliced segments in Nr-CAM occurs between the second and third Ig-like domains. This spliced variant would place an additional nineteen amino acid segment between the two Ig-like domains without apparently disrupting the structure of either and could provide a means for altering Nr-CAM's binding function. Ng-CAM and L1 are not known to contain a corresponding segment in their sequences.

In the region of the Fn-III repeats, Ng-CAM is approximately equally similar to L1 and Nr-CAM. As observed in the Ng-CAM/L1 comparison (Figure 17b), the Fn-III repeats are again the region of least identity between these proteins. Within this region, the fourth and fifth Fn-III repeats are the least identical domains in the entire sequence. Given this region's close proximity to the plasma membrane in the extracellular portion of these molecules, it may be weakly conserved because it may function only as a spacer.

There are also significant differences between Ng-CAM, Nr-CAM, and L1 in the third Fn-III repeat. Both L1 and Nr-CAM differ significantly from Ng-CAM in the 18 amino acids around the amino terminus of the 80 kD component of Ng-CAM (indicated by the arrow in Figure 26). This suggests that Nr-CAM may not be readily proteolyzed at that site, but like L1, is expressed primarily intact as the largest translation product. If so the predominant 145 kD component detected in brain seems unusually small relative to L1 because both have a similar total number of amino acids. On the other hand, if the 170 kD minor band detected in



immunoblots of Nr-CAM represents the entire molecule, the 145 kD component appears too large to be a component cleaved at a site equivalent to that in Ng-CAM.

The proposed cleavage site in Ng-CAM is in the same Fn-III repeat where Ng-CAM has an RGD sequence in a position comparable to a cell binding site in fibronectin (Hynes, 1987). Nr-CAM, like L1, does not contain any RGD sequences in its Fn-III repeats. This may reflect a functional difference between Ng-CAM and Nr-CAM, as has been noted for Ng-CAM and L1.

The amino acid sequences of Ng-CAM, Nr-CAM, and L1 are more similar to each other in the predicted transmembrane domains and portions of the cytoplasmic regions than in any other region (~50% identical). In these regions, several continuous segments of as many as 12 amino acids are identical in the three proteins. The carboxyl terminal half of the cytoplasmic domain of Ng-CAM is much more similar to L1 (66% identity) than to Nr-CAM (43% identity). However, 90% of the amino acids identical between the cytoplasmic region of Ng-CAM and Nr-CAM are also identical in L1. This provides further evidence that the cytoplasmic regions of Ng-CAM, L1 (and Nr-CAM) serve a common role for all three proteins.

#### Similarity to other CAMs

Computer comparisons between the amino acid sequences of Nr-CAM and other molecules indicated that it is similar to several members of the Ig superfamily, particularly to the neural CAMs (Figure 27). Analyses of many pairwise comparisons also confirm that Nr-CAM, Ng-CAM

Figure 27 Relationships between Ig-like neural CAMs showing similar features.

Protein sequences for each of the neural CAMs listed were obtained from the translated Genbank database (release 63) or from the present studies and compared in a pairwise fashion using the LFASTA program of Pearson (Pearson et al., 1988) (ktup=2). Scores indicate the percentage identity between the protein sequences across the number of amino acid residues indicated in parentheses. Parentheses at left indicate origin of protein; C = chicken, M = mouse, I = insect, R = rat.

	Ng-CAM	L1	Nr-CAM	Contactin/ F11	F3	TAG-1	Chick N-CAM	Mouse N-CAM	Neuro- glian
Ng-CAM (C)									
L1 (M)	40 (1284)								
Nr-CAM (C)	36 (1194)	40 (1263)							
Contactin/F11 (C)	27 (815)	28 (882)	28 (778)						
F3 (M)	27 (873)	29 (886)	25 (908)	78 (979)					
TAG-1 (R)	29 (865)	29 (993)	29 (934)	49 (1033)	50 (1009)				
N-CAM (C)	22 (600)	21 (706)	20 (685)	19 (676)	23 (680)	19 (730)			
N-CAM (M)	21 (584)	20 (712)	20 (720)	22 (551)	23 (670)	21 (707)	81 (1109)		
Neuroglian (I)	23 (1269)	28 (1224)	32 (1227)	28 (778)	27 (781)	24 (972)	21 (710)	22 (713)	
Fasciclin II (I)	21 (365)	21 (376)	24 (460)	19 (572)	24 (670)	19 (613)	26 (562)	27 (395)	22 (411)





and L1 are more closely related to each other than to N-CAM and other CAMs.

*Drosophila neuroglian* (Bieber et al., 1989) also shares a similar organization of domains and shows the highest similarity when compared to Nr-CAM (Figure 27). This is due in part to the carboxyl terminal 111 amino acids in neuroglian which are 38% identical to Nr-CAM; this stretch includes a segment of nine amino acids that is also identical in Nr-CAM, Ng-CAM, L1 and NILE (Burgoon et al., 1991; Moos et al., 1988; Prince et al., 1989).

Other molecules whose structures have not been determined in detail could not be included in these comparisons, but they may be related to Nr-CAM. It has recently been reported that chicken neural proteins of 140 and 130 kDa, called Bravo (de la Rosa et al., 1990), which are detected on retinal neurites in the retina but not in the tectum, have amino terminal sequences similar to mouse L1. This portion of Bravo is even more similar to the amino terminal sequence of Nr-CAM (74% identity). However, ambiguities reported in the amino terminal sequence of Bravo make it difficult to determine whether it is very closely related or identical to Nr-CAM. Neurofascin (Rathjen et al., 1987b) also shares certain properties with Nr-CAM including the general size of some of its components and its localization in the nervous system, but it is known to be immunologically distinct from Bravo (de la Rosa et al., 1990). Better comparisons among these molecules will only be possible when the structures of Bravo and neurofascin have been more completely determined.



## DISCUSSION

Nr-CAM was discovered with reagents designed for the characterization of Ng-CAM. Its overall structure and amino acid sequence resemble both Ng-CAM and L1. In some regions Nr-CAM is more similar to Ng-CAM, and in other regions it is more similar to L1. Overall chicken Nr-CAM is most similar to mouse L1 (40%) but this is only marginally better than the identity between L1 and Ng-CAM (35%). Neither Nr-CAM nor Ng-CAM appears to be the chicken equivalent of mouse L1 given that they are at best only 40% identical and that molecules such as N-CAM in the two species are much more closely related (~80% identity) (Figure 27). A striking feature extracted from the comparisons in Figure 26 is the large number of amino acid residues identical in all three proteins. Ng-CAM and L1 share 511 identical residues and of these 75% are also identical in Nr-CAM. Such conservation may be critical for the structure or the function of the molecule. Ng-CAM and L1 have similar activities and some of these may be shared by Nr-CAM. The three proteins also have many unrelated segments, which could provide different functions or different binding specificities.

The distribution of Nr-CAM and its similarity to other neural CAMs suggests that it functions as a cell adhesion molecule. In the group of molecules that are structurally related to Nr-CAM, it is known that Ng-CAM, L1 and NILE are important for axonal fasciculation (Friedlander et al., 1986; Hoffman et al., 1986; Fischer et al., 1986; Stallcup and Beasley, 1985) and neuroglial has been demonstrated to mediate cell-cell adhesion after transfection into non-neuronal cell lines



(Grenningloh et al., 1990). All of these molecules are found on neurons and are particularly prevalent on axons (Grumet et al., 1984b; Thiery et al., 1985; Rathjen and Schachner, 1984; Stallcup et al., 1985). Initial studies indicated that Nr-CAM is expressed on the surface of neurons during development and is localized at cell-cell junctions (Grumet, 1991). It is restricted to certain neuronal cell bodies and fibers and is dynamically expressed during development, consistent with the hypothesis that it may be important for neural cell adhesion and axonal fasciculation.

The discovery of a neural protein (Nr-CAM) very similar in structure and sequence to both Ng-CAM and L1 prompts the hypothesis that they may comprise a closely related subgroup in the N-CAM family of neural CAMs. Their sequences are more closely related to one another than to the other N-CAM-like molecules and each contains an equivalent number of Ig-like and Fn-III domains. In addition, the transmembrane and cytoplasmic domains of Ng-CAM, Nr-CAM, and L1 do not resemble other CAMs, but are very similar to one another.

The recent discovery of a number of Ig-like neural CAMs in the N-CAM family has supported the idea that this family may be comprised of other closely related subgroups. The next section discusses the classification of these adhesion molecules.

### Families of CAMs

The ability to organize cell adhesion molecules into families and subgroups promises to be a valuable approach in understanding the relationship between CAM structure and function, in identifying the



functions of new CAMs, and in defining the ability of CAMs to influence each other. The remainder of this discussion will focus on our current ability to classify the various CAMs, concluding with a discussion of the specific subgroup of neural CAMs identified by this thesis.

Historically two types of cell adhesive events have been characterized--those that required calcium and those that were calcium independent (Brackenbury et al., 1981; Takeichi et al., 1981; McClain and Edelman, 1982). Both types of adhesion are present in the nervous system as well as in other tissues, and on most cells. The identification of N-CAM and related calcium-independent CAMs and of L-CAM and the other calcium-dependent cadherins has provided a molecular basis for these two mechanisms. This view, however, must be expanded. Integrins are a group of related calcium-dependent cell-surface receptors that influence cell-cell interactions (Kishimoto and , 1989). Integrins are heterodimeric molecules defined by their particular combinations of  $\alpha$  and  $\beta$  subunits (Hynes, 1987; Ruoslahti and Pierschbacher, 1987). They may interact with specific N-CAM-like molecules as in the vascular system with ICAM-1 and ICAM-2 (Staunton et al., 1988; Marlin and Springer, 1987; Simmons et al., 1988; Staunton et al., 1989), or they may interact with substrate adhesion molecules such as fibronectin (Hynes, 1987) or cytotactin/tenascin (Bourdon and Ruoslahti, 1989). N-CAM itself has been reported to mediate calcium-dependent adhesion on oligodendrocytes and thus may be involved in heterophilic binding to other ligands as well (Bhat and Silberberg, 1986). At least one N-CAM-like molecule on endothelial cells, PECAM-1





(Newman et al., 1990), mediates calcium-dependent adhesion and may not involve integrin receptors.

Moreover, lectin-like molecules (LEC-CAMs or selectins; Lasky, 1991; Springer and Lasky, 1991) are involved in lymphocyte and neutrophil adhesion to endothelia, and have also been identified in the nervous system (Zanetta et al., 1985; Zanetta et al., 1987). Antibodies to a cerebellum-specific lectin were shown to block granule cell migration in explant culture (Lehmann et al., 1990), so the role of carbohydrates in cell interactions must be considered seriously. Finally, it is not known whether the proteoglycan-like molecules found in invertebrates, some of which require calcium (e.g. sponge aggregating factor), have counterparts in vertebrates.

Among the various CAMs, the cadherins are a group of closely related molecules that have nearly identical structures and are highly conserved in amino acid sequence (48-68% identity; Takeichi, 1988). They have the same number of internal repeats, highly conserved cytoplasmic regions, and are all synthesized as larger precursors; the greatest difference between the cadherins is the size of the cleared precursor segment. The repeats include the regions involved in cell binding and the binding specificity for the cadherins is likely to reside in the first 113 residues (Nose et al., 1990). The cytoplasmic domains are especially conserved and appear to mediate an indirect interaction with actin filaments through linking proteins (e.g. catenins) (Nelson et al., 1990; Ozawa et al., 1989). This interaction appears obligatory for proper binding functions, and may direct some cadherins to specific cell junctional complexes (Mege et al., 1988;



Ozawa et al., 1990; Gumbiner et al., 1988; Geiger, 1989; Jaffe et al., 1990).

Despite their similarities cadherins have different tissue distributions and different binding specificities. There is also emerging evidence that there may be subgroups of cadherins. For example, T-cadherin has a truncated cytoplasmic domain that does not correspond to other cadherins (Ranscht and Bronner-Fraser, 1991). Some desmosomal proteins have cadherin-like extracellular segments but differ in their cytoplasmic regions (Koch et al., 1990).

Comparisons among the cadherins also reveal a relationship similar to that found between Ng-CAM and Ll. From their binding properties and tissue distribution E-cadherin/uvomorulin was presumed to be the mammalian equivalent of L-CAM in chicken. But in the two species they are only 65% identical in amino acid sequence, while other cadherins in different species are much more closely related (Takeichi, 1990). Moreover, liver membranes from chicken and mouse do not interact with each other, suggesting that L-CAM and uvomorulin/E-cadherin may not be equivalent molecules in the two species (Obrink et al., 1977; Grady and McGuire, 1976).

#### Subgroups of Ig-like Neural CAMs

In contrast to the cadherins, the N-CAM-like molecules have generally appeared much less closely related. As shown in Figure 27, the CAMs in this family are 25-81% identical to each other. However, the members of subgroups that we have begun to identify are about as closely related to each other as the cadherins and similar correlations



of structure and function among members of subgroups may rapidly emerge. The basis for defining such subgroups are still forming, but we can loosely designate four such groups based on their similarities in overall structure and amino acid sequence (Figure 28).

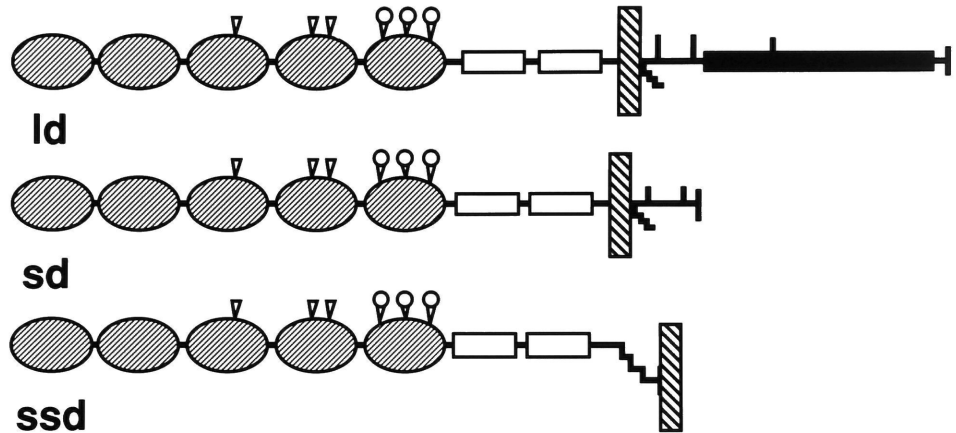
There are so many variants of N-CAM that it could be considered a subgroup of its own. All of the variants or isoforms contain five Ig-like domains and two Fn-III domains, but they differ in their modes of attachment to the membrane or in the presence of alternatively spliced segments in their sequence (Cunningham et al., 1987; Goridis and Wille, 1988; Small et al., 1988). Although N-CAM is ubiquitously expressed in the nervous system, the different isoforms are not equally distributed throughout the nervous system (Schlosshauer et al., 1984; Chuong and Edelman, 1984; Murray et al., 1986; Prieto et al., 1989). *In situ* hybridization studies have shown that this probably reflects the widespread distribution of the shorter transmembrane form of N-CAM. The largest isoform was localized in several areas to particular subsets of neurons, such as Purkinje and internal granule cells of the cerebellum and ependymal cells in the spinal cord but it was absent from several other distinct groups of cells expressing smaller forms of N-CAM (Prieto et al., 1989). The insect protein fasciclin II may also be included in this subgroup. It contains an equivalent domain organization to N-CAM and its sequence is most similar to that of N-CAM (Figure 27).

A less clearly apparent subgroup contains MAG and P<sub>0</sub>, which contain differing numbers of Ig-like domains (five for MAG, one for P<sub>0</sub>), but are distinguished from the other N-CAM-like neural CAMs because they lack Fn-III repeats. They are not closely related to each other in

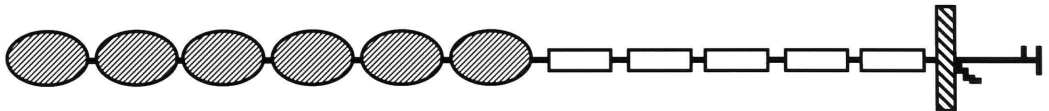
Figure 28 Subfamilies of the N-CAM-like neural CAMs.

Several of the neural CAMs in the Ig superfamily may be grouped into subfamilies based on structural and sequence similarities. Shaded ovals represent Ig-like domains, horizontal open boxes represent Fn-III repeats. Carrots above the Ig-like domains in N-CAM indicate sites of glycosylation. Carrots with filled circles indicate polysialic acid which may undergo an embryonic to adult conversion during development. Stippled vertical bars represent the cell membrane (extracellular regions are to the left), and ladders extending to the right from the membrane indicate the general location of palmitate. Ladders to the right of the membrane region indicate a phospholipid anchor of the protein to the membrane. Filled horizontal boxes indicate alternatively spliced cytoplasmic segments, and lines extending from the cytoplasmic segments indicate general locations of phosphorylation sites.

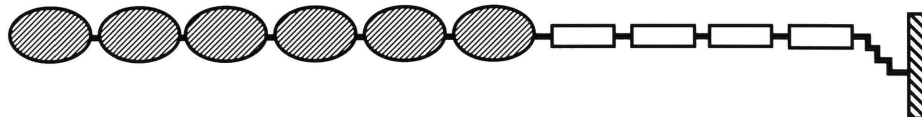
## N-CAM



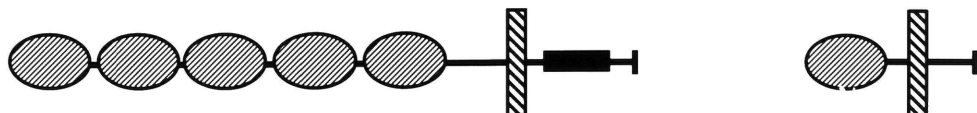
## Ng-CAM, L1, Nr-CAM



## Contactin, F11, F3, Tag-1



## MAG







amino acid sequence, but both are expressed only on oligodendrocytes and Schwann cells. Moreover, in two populations of transfected cells expressing either CAM, MAG and P<sub>0</sub> colocalize to focal contacts between the cells (Doyle et al., 1990), suggesting an interaction between the two molecules.

It has been suggested that the structures of MAG and P<sub>0</sub> are ideally suited for their predicted roles in myelination (Trapp, 1988; Lemke and Axel, 1985). P<sub>0</sub> is involved in maintaining the tight apposition of adjacent Schwann cell membranes in the compact myelin which ensheathes axons. The lack of Fn-III repeats and a single Ig-like domain would make P<sub>0</sub> appear ideally designed for such a role. The single extracellular Ig-like domain is predicted to bind homophilically with P<sub>0</sub> on the apposing membrane (D'Urso et al., 1990; Doyle et al., 1990), while the cytoplasmic domain is hypothesized to mediate intracellular membrane apposition through electrostatic protein-lipid interactions with the opposite membrane (Lemke and Axel, 1985). MAG is a larger protein with five Ig-like domains, and is removed from Schwann cell mesaxon membranes as compact myelin is formed. MAG is localized to the periaxonal regions of Schwann cells near the axon. Near the axonal surfaces MAG has been postulated to maintain the precise periaxonal space and periaxonal cytoplasm of the Schwann cells (Trapp, 1988; Salzer and Colman, 1989).

Two of the subgroups, in addition to containing equivalent numbers of Ig-like domains and Fn-III repeats, have equivalent modes of attachment to the cell membrane. One subgroup includes chicken contactin/F11, its mouse equivalent F3, and the rat TAG-1. Each protein



contains six Ig-like domains and four Fn-III repeats, and appears to be anchored to the membrane via a phosphatidylinositol moiety.

By their apparent lack of association with cytoskeletal elements, members of this subgroup may be expected to exhibit more mobility on the cell surface and may mediate an adhesion independent of cytoskeletal involvement. This may provide a different strength of cell adhesion from the type of adhesion mediated by transmembrane CAMs. Their different mode of attachment to the membrane may also influence their distribution on the cell surface. In addition to their structural similarities, the members of this subgroup of neural CAMs are more similar to each other in amino acid sequence than to any other proteins. These CAM are all similarly concentrated on neurites, and there is evidence that several of these CAMs may be released from the membrane into the fiber tracts (Furley et al., 1990). Such a release could serve several functions. These soluble forms may alter potential neurite binding to the tract, decrease neurite fasciculation, or affect the adhesion of late-arriving neurites and growth cones in the same fiber tract. For example, TAG-1 expression in the rat spinal cord is correlated with the extension of unfasciculated commissural axons from sensory neurons (Dodd et al., 1988; Holley and Silver, 1987), suggesting a role for a released form of TAG-1 in defasciculation.

A subgroup of CAMs is identified in this thesis and contains Ng-CAM, L1 and the new protein Nr-CAM. They have equivalent structures and closely related amino acid sequences. Ng-CAM and L1 are concentrated on neurites like the previously described subgroup, but their transmembrane nature suggest that the role of Ng-CAM, L1, and Nr-CAM may be different



from that of the phospholipid-anchored CAMs. Their activities on the cell surface are probably directly coupled to intracellular events as well, perhaps by activation of a cytoplasmic element or interaction with the cytoskeleton. Portions of their cytoplasmic regions are highly conserved and reflect this view. Moreover, studies showed that antibody binding to L1 at the cell surface increased the intracellular pH and levels of calcium and reduced the levels of the second messengers inositol phosphate 2 and 3 (Schuch et al., 1989).

The extracellular regions in these CAMs are critical for their activities. Besides containing the regions likely to mediate the binding functions, these domains are also expected to provide the specificity of CAM binding. By comparison to the amino acid similarities among the cadherins and their differences in binding specificity, the members of the Ng-CAM/Nr-CAM/L1 subgroup may be expected to exhibit different binding specificities. Any such conclusion, however, will require more knowledge of the binding regions themselves.

Another feature which suggests that these three molecules may have different binding specificities is the presence or absence of the tripeptide RGD in their sequences, which by comparison to fibronectin may have a functional role. A single RGD is present in the third Fn-III repeat of Ng-CAM. There are two such sequences in L1 but they are both in the last Ig-like domain. Nr-CAM does not contain an RGD sequence. Experiments are in progress to determine whether their sequence has any functional significance in Ng-CAM.



While subgroups have been emphasized here the possibility must be raised that such subgroups are simply the reflection of minimal sampling. This raises an old question: how many CAMs are there? If the number is large those that we have identified may be part of a continuum of structures and thus may not be distinct subgroups. There is some evidence for this notion. For example, N-CAM exists as both transmembrane and phospholipid-anchored variants that are differentially expressed and may perform roles similar to members of several subgroups. Moreover Nr-CAM contains a Fn-III repeat which may be alternatively spliced in primary transcripts; it may therefore represent a crossover between two structurally distinct groups based on equivalence of domain structure, the Ng-CAM/Nr-CAM/L1 subgroup with 6 Ig/5Fn-III domains and the contactin/F11/F3/TAG-1 subgroup with 6 Ig/4 Fn-III domains.

Within the N-CAM family the functional significance of the number of Ig-like and Fn-III domains in the CAMs is not clear. In most of the CAMs it is not known how many of the extracellular domains are involved in the binding properties. The specific numbers of domains present in each subfamily may reflect spacing necessary for specific functions in specific locations or add additional functions comparable to the effector functions of immunoglobulins (Davies and Metzger, 1983).

Several reports indicate that CAMs in these different subgroups may interact with each other. N-CAM has been reported to interact with L1 on neuroblastoma cells and by cis-interactions enhance L1-mediated binding (Kadmon et al., 1990a; Kadmon et al., 1990b). In the developing rat spinal cord the spatially segregated expression of TAG-1 and L1 may





be coordinated on the same neurites to influence the guidance of neurite outgrowth (Dodd et al., 1988). The possible interactions of CAMs with each other in the nervous system may provide an additional level of complexity in CAM-mediated adhesion. Such interactions coupled with the number of known CAMs, their complex expression patterns, and the various forms in which several CAMs may be expressed enhances the possibility that relatively few CAMs can provide a wide range of specificities in cell adhesion.

### Future Prospectives

The ability to sort CAMs into structurally related subgroups should facilitate attempts to correlate CAM structure with functional properties and can provide experimental approaches to identify new CAMs.

From the sequence similarities of Ng-CAM, Nr-CAM and Ll, particularly in the cytoplasmic regions, it is now possible to define PCR primers which may detect other CAMs in this subgroup.

Functional studies of Ng-CAM and Nr-CAM will be advanced by the use of several molecular biological approaches. The availability of cDNAs encoding Ng-CAM and Nr-CAM for transfection into cells will now enable the study of their individual functions in *in vitro* systems. The cDNA of Ng-CAM and Nr-CAM may be inserted into the pKSV-10 vector and expressed in L cells which do not normally aggregate and do not appear to express any known CAMs. Expression of these CAMs on transfected cells will isolate their activity from that of other CAMs which may normally be present on the same neurons. Ng-CAM may be expressed as



the individual 135 kD and 80 kD components to assess their particular contributions to adhesion. In addition Nr-CAM contains several potential splice variants which may affect its function. The analysis of various forms of Nr-CAM expressed on transfected cells may indicate their roles in the activity of the molecule.

The cDNA clones identified here may be used to characterize the genes for Ng-CAM and Nr-CAM and will lead to at least two important experimental paths. With the identification of promoter elements in the genes the regulated expression of the CAMs can be analyzed. The promoter regions can be coupled to the reporter gene choline acetyltransferase (CAT) and transfected into neural cells to assay cellular influences on their expression. The genes may also be expressed in transgenic animals or used in gene knockout experiments to investigate the functional role of the CAMs in vivo— particularly during development.



## BIBLIOGRAPHY

- Abo, T., and Balch, C.M. (1981). A differentiation antigen of human NK and K cells identified by a monoclonal antibody (HNK-1). *J. Immunol.* 127, 1024-1031.
- Amzel, L.M., and Poljak, R.J. (1979). Three-dimensional structure of immunoglobulins. *Ann. Rev. Biochem.* 48, 961-997.
- Antonicek, H., Persohn, E., and Schachner, M. (1987). Biochemical and functional characterization of a novel neuron-glia adhesion molecule that is involved in neuronal migration. *J. Cell Biol.* 104, 1587-1595.
- Arquint, M., Roder, J., Chia, L.S., Down, J., Wilkinson, D., Bayley, H., Braun, P., and Dunn, R. (1987). Molecular cloning and primary structure of myelin-associated glycoprotein. *Proc. Natl. Acad. Sci. USA.* 84, 600-604.
- Ausubel, F.M., Brent, R., Kingston, R.E., Moore, D.D., Seidman, J.G., Smith, J.A., and Struhl, K. (1989). *Current Protocols in Molecular Biology*. F. M. Ausubel, R. Brent, R. E. Kingston, D. D. Moore, and J. G. Seidman, eds. (: John Wiley and Sons, New York).
- Balsamo, J., and Lilien, J. (1975). The binding of tissue-specific adhesive molecules to the cell surface. A molecular basis for specificity. *Biochemistry.* 14, 167-171.
- Bayna, E.M., Shaper, J.H., and Shur, B.D. (1988). Temporally specific involvement of cell surface B-1,4 Galactosyltransferase during mouse embryo morula compaction. *Cell.* 53, 145-157.
- Bertolotti, R., Rutishauser, U., and Edelman, G.M. (1980). A cell surface molecule involved in aggregation of embryonic liver cells. *Proc. Natl. Acad. Sci. USA.* 77, 4831-4835.
- Beug, H., Katz, F., and Gerisch, G. (1973). Dynamics of antigenic membrane sites relating to cell aggregation in Dictyostelium discoideum. *J. Cell Biol.* 56, 647-658.
- Bhat, S., and Silberberg, D.H. (1986). Oligodendrocyte cell adhesion molecules are related to neural cell adhesion molecule (N-CAM). *J. Neurosci.* 6, 3348-3354.
- Bieber, A.J., Snow, P.M., Hortsch, M., Patel, N.H., Jacobs, J.R., Traquina, Z.R., Schilling, J., and Goodman, C.S. (1989). Drosophila neuroglian: a member of the immunoglobulin superfamily with extensive homology to the vertebrate neural adhesion molecule L1. *Cell.* 59, 447-460.
- Bixby, J., Lilien, J., and Reichardt, R.F. (1988). Identification of the major proteins that promote neuronal process outgrowth on Schwann cells.



J. Cell Biol. 107, 353-362.

Bixby, J.L., Pratt, R.S., Lilien, J., and Reichardt, L.F. (1987). Neurite outgrowth on muscle cell surfaces involves extracellular matrix receptors as well as Ca<sup>2+</sup>-dependent and -independent cell adhesion molecules. Proc. Natl. Acad. Sci. USA. 84, 2555-2559.

Bock, E., Richter-Landsberg, C., Faissner, A., and Schachner, M. (1985). Demonstration of immunochemical identity between the nerve growth factor-inducible large external (NILE) glycoprotein and the cell adhesion molecule-L1. EMBO J. 4, 2765-2768.

Bonhoeffer, F., and Gierer, A. (1984). How do retinal axons find their targets on the tectum?. TINS. 378-381.

Bourdon, M.A., and Ruoslahti, E. (1989). Tenascin mediates cell attachment through an RGD-dependent receptor. J. Cell Biol. 108, 1149-1155.

Brackenbury, R., Rutishauser, U., and Edelman, G.M. (1981). Distinct calcium-independent and calcium-dependent adhesion systems of chicken embryo cells. Proc. Natl. Acad. Sci. USA. 78, 387-391.

Brackenbury, R., Thiery, J.-P., Rutishauser, U., and Edelman, G.M. (1977). Adhesion among neural cells of the chick embryo. I. An immunological assay for molecules involved in cell-cell binding. J. Biol. Chem. 252, 6835-6840.

Brummendorf, T., Wolff, J.M., Frank, R., and Rathjen, F.G. (1989). Neural cell recognition molecule F11: homology with fibronectin type III and immunoglobulin type C domains. Neuron. 2, 1351-1361.

Brunet, J.-F., Hirsch, M.-R., Naquet, P., Überla, K., Diamantstein, T., Lipinski, M., and Goridis, C. (1989). Developmentally regulated expression of the neural cell adhesion molecule (NCAM) by mouse thymocytes. Eur. J. Immunol. 19, 837-841.

Burgoon, M.P., Grumet, M., Mauro, V., Edelman, G.M., and Cunningham, B.A. (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.

Buskirk, D.R., Thiery, J.-P., Rutishauser, U., and Edelman, G.M. (1980). Antibodies to a neural cell adhesion molecule disrupt histogenesis in cultured chick retinae. Nature (London). 285, 488-489.

Chang, S., Rathjen, F.G., and Raper, J.A. (1987). Extension of neurites on axons is impaired by antibodies against specific cell surface glycoproteins. J. Cell Biol. 104, 355-362.

Chou, D.K.H., Ilyas, A.A., Evans, J.E., Costello, C., Quarles, R.H., and Jungalwala, F.B. (1986). Structure of sulfated glucuronyl glycolipids in the nervous system reacting with HNK-1 antibody and some IgM paraproteins in neuropathy. J. Biol. Chem. 261, 11717-11725.





Chung, S.H., Keating, M.J., and Bliss, T.V.P. (1974). Functional synaptic relations during the development of the retino-tectal projections in amphibians. *Proc. R. Soc. Lond. [Biol.]* 187, 449-459.

Chuong, C.-M., and Edelman, G.M. (1984). Alteration in neural cell adhesion molecules during development of different regions of the nervous system. *J. Neurosci.* 4, 2354-2368.

Chuong, C.-M., and Edelman, G.M. (1985). Expression of cell adhesion molecules in embryonic induction. I. Morphogenesis of nestling feathers. *J. Cell Biol.* 101, 1009-1026.

Cleveland, D.W., Fischer, S.G., Kirschner, M.W., and Laemmli, U.K. (1977). Peptide mapping by limited proteolysis in sodium dodecyl sulfate and analysis by gel electrophoresis. *J. Biol. Chem.* 252, 1102-1106.

Cole, G.J., and Akeson, R. (1989). Identification of a heparin binding domain of the neural cell adhesion molecule N-CAM using synthetic peptides. *Neuron.* 2, 1157-1165.

Constantine-Paton, M., Blum, A.S., Mendez-Otero, R., and Barnstable, C.J. (1986). A cell surface molecule distributed in a dorsoventral gradient in the perinatal rat retina. *Nature.* 324, 459-462.

Covault, J., Merlie, J.P., Goridis, C., and Sanes, J.R. (1986). Molecular forms of N-CAM and its RNA in developing and denervated skeletal muscle. *J. Cell Biol.* 102, 731-739.

Crossin, K.L., Chuong, C.-M., and Edelman, G.M. (1985). Expression sequences of cell adhesion molecules. *Proc. Natl. Acad. Sci. USA.* 82, 6942-6946.

Crossin, K.L., Edelman, G.M., and Cunningham, B.A. (1984). Mapping of three carbohydrate attachment sites in embryonic and adult forms of the neural cell adhesion molecule. *J. Cell Biol.* 99, 1848-1855.

Cuatrecasas, P., and Anfinsen, C.B. (1971). Affinity chromatography. *Methods Enzymol.* 22, 345-378.

Cunningham, B.A. (1988). The structure and function of cell adhesion molecules. *Adv. in Cell Biol.* 2, 25-54.

Cunningham, B.A., Hemperly, J.J., Murray, B.A., Prediger, E.A., Brackenbury, R., and Edelman, G.M. (1987). Neural cell adhesion molecule: Structure, immunoglobulin-like domains, cell surface modulation, and alternative RNA splicing. *Science.* 236, 799-806.

Cunningham, B.A., Hoffman, S., Rutishauser, U., Hemperly, J.J., and Edelman, G.M. (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.

Curtis, A.S.G. (1967). The Cell Surface: Its Molecular Role in Morphogenesis. In *The Cell Surface: Its Molecular Role in Morphogenesis*,



(: Academic Press, New York).

D'Urso, D., Brophy, P.J., Staugaitis, S.M., Gillespie, C.S., Frey, A.B., Stempak, J.G., and Colman, D.R. (1990). Protein Zero of Peripheral Nerve Myelin: Biosynthesis, Membrane Insertion, and Evidence for Homotypic Interaction. *Neuron*. 2, 449-460.

Dale, R.M.K., McClure, B.A., and Houchins, J.P. (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.

Damsky, C.H., Richa, J., Solter, D., Knudsen, K., and Buck, C.A. (1983). Identification and purification of a cell surface glycoprotein involved in cell-cell interactions. *Cell*. 34, 455-466.

Daniloff, J., Chuong, C.-M., Levi, G., and Edelman, G.M. (1986a). Differential distribution of cell adhesion molecules during histogenesis of the chicken nervous system. *J. Neurosci*. 6, 739-758.

Daniloff, J.K., Levi, G., Grumet, M., Rieger, F., and Edelman, G.M. (1986b). Altered expression of neuronal cell adhesion molecules induced by nerve injury and repair. *J. Cell Biol.* 103, 929-945.

Davies, D.R., and Metzger, H. (1983). Structural basis of antibody function. *Ann. Rev. Immunol.* 1, 87-117.

Davis, N.G., and Model, P. (1985). An artificial anchor domain:hydrophobicity suffices to stop transfer. *Cell*. 41, 607-614.

de la Rosa, E.J., Kayyem, J.F., Roman, J.M., Stierhof, Y.D., Dreyer, W.J., and Schwarz, U. (1990). Topologically restricted appearance in the developing chick retinotectal system of Bravo, a neural surface protein: Experimental modulation by environmental cues. *J. Cell Biol.* 111, 3087-3096.

DeMorgan, W., and Drew, H. (1914). A study of the restitution masses formed by the dissociated cells of the hybrids *Antennularia ramosa* and *A. antennine*. *J. Mar. Biol. Assoc. UK*. 10, 440-463.

Dickson, G., Gower, H.J., Barton, C.H., Prentice, H.M., Elsom, V.L., Moore, S.E., Cox, R.D., Quinn, C., Putt, W., and Walsh, F.S. (1987). Human muscle neural cell adhesion molecule (N-CAM)- Identification of a muscle specific sequence in the extracellular domain. *Cell*. 50, 1119-1130.

Dodd, J., Morton, S.B., Karagogeos, D., Yamamoto, M., and Jessell, T.M. (1988). Spatial regulation of axonal glycoprotein expression on subsets of embryonic spinal neurons. *Neuron*. 1, 105-116.

Doherty, P.C., Barton, H., Dickson, G., Seaton, P., Rowett, L.H., Moore, S.E., J., G.H., and Walsh, F.S. (1989). Neuronal process outgrowth of human sensory neurons on monolayers of cells transfected with cDNAs for five human N-CAM isoforms. *J. Cell Biol.* 189, 789-798.



Doyle, J.P., D'Urso, D., Yang, C.W., Gallant, E.M., Cowin, P., and Colman, D.R. (1990). Inhibition of Po mediated cell-cell adhesion by mutagenesis of the extracellular domain. *J. Cell Biol.* 107, 405a (Abstract).

Drazba, J., and Lemmon, V. (1990). The role of cell adhesion molecules in neurite outgrowth on Muller Cells. *Devel. Biol.* 138, 82-93.

Easter, S.S., Rusoff, A.C., and Kish, P.E. (1981). The growth and organization of the optic nerve and tract in juvenile and adult goldfish. *J. Neurosci.* 1, 793-811.

Edelman, A.M., Blumenthal, D.K., and Krebs, E.G. (1987b). Protein Serine/Threonine Kinases. *Ann. Rev. Biochem.* 56, 567-613.

Edelman, G.M. (1970). The covalent structure of a human gamma G-immunoglobulin. XI. Functional implications. *Biochem.* 9(16), 3197-3205.

Edelman, G.M. (1985a). Cell adhesion and the molecular processes of morphogenesis. *Ann. Rev. Biochem.* 54, 135-169.

Edelman, G.M. (1985b). Expression of cell adhesion molecules during embryogenesis and regeneration. *Exp. Cell Res.* 161, 1-16.

Edelman, G.M. (1987). CAMs and Igs: Cell adhesion and the evolutionary origins of immunity. *Immun. Rev.* 100, 11-45.

Edelman, G.M. (1988). *Topobiology: An Introduction to Molecular Embryology.* (: Basic Books, New York).

Edelman, G.M., and Gallin, W.J. (1987). Cell adhesion as a basis of pattern in embryonic development. *Am. Zool.* 27, 645-656.

Edelman, G.M., Murray, B.A., Mege, R.-M., Cunningham, B.A., and Gallin, W.J. (1987a). Cellular expression of liver and neural cell adhesion molecules after transfection with their cDNAs results in specific cell-cell binding. *Proc. Natl. Acad. Sci. USA.* 84, 8502-8506.

Edmondson, J.C., Liem, R.K.H., Kuster, J.E., and Hatten, M.E. (1988). Astrotactin: A novel neuronal cell surface antigen that mediates neuron-astroglial interactions in cerebellar microcultures. *J. Cell Biol.* 106, 505-517.

Edwards, A.M., Arquint, M., Braun, P.E., Roder, J.C., Dunn, R.J., Pawson, T., and Bell, J.C. (1988). Myelin-associated glycoprotein, a cell adhesion molecule of oligodendrocytes, is phosphorylated in brain. *Mol. Cell. Biol.* 8, 2655-2658.

Ehrlich, Y.H., Davis, T.B., Bock, E., Kornecki, E., and Lenox, R.H. (1986). Ecto-protein kinase activity on the external surface of neural cells. *Nature.* 320, 67-70.

Eisenberg, D. (1984). Three-dimensional structures of membrane and surface proteins. *Ann. Rev. Biochem.* 53, 595-623.



Elkins, T., Zinn, K., McAllister, L., Hoffman, F.M., and Goodman, C.S. (1990). Genetic analysis of a drosophila neural cell adhesion molecule: Interaction of fasciclin I and Abelson tyrosine kinase mutations. *Cell*. 60, 565-575.

Faissner, A., Krude, J., Goridis, C., Bock, E., and Schachner, M. (1984). The neural cell adhesion molecule L1 is distinct from the N-CAM related group of surface antigens BSP-2 and D2. *EMBO J.* 3, 733-737.

Faissner, A., Teplow, D.B., Kubler, D., Keilhauer, G., Kinzel, V., and Schachner, M. (1985). Biosynthesis and membrane topography of the neural cell adhesion molecule L1. *EMBO J.* 4, 3105-3113.

Feinberg, A.P., and Vogelstein, B. (1984). A technique for radiolabeling DNA restriction endonuclease fragments to high specific activity. *Anal. Biochem.* 137, 266-267.

Ferguson, M.A.J. (1988). Cell-surface anchoring of proteins via glycosyl-phosphatidylinositol structures. *Ann. Rev. Biochem.* 57, 285-320.

French-Constant, C., and Hynes, R. (1989). Alternative splicing of fibronectin is temporally and spatially regulated in the chicken embryo. *Development*. 106, 375-388.

Filbin, M.T., Walsh, F.S., Trapp, B.D., Pizzey, J.A., and Tennekoon, G.I. (1990). Role of myelin Po protein as a homophilic adhesion molecule. *Nature (Lond)*. 344, 871.

Fischer, G., Kunemund, J., and Schachner, M. (1986). Neurite outgrowth patterns in cerebellar microexplant cultures are affected by antibodies to the cell surface glycoprotein L1. *J. Neurosci.* 6, 605-612.

Frail, D.E., and Braun, P.E. (1984). Two developmentally regulated messenger RNAs differing in their coding region may exist for the myelin-associated glycoprotein. *J. Biol. Chem.* 259, 14857-14866.

Fraser, S.E., Murray, B.A., Chuong, C.-M., and Edelman, G.M. (1984). Alteration of the retinotectal map in Xenopus by antibodies to neural cell adhesion molecules. *Proc. Natl. Acad. Sci. USA*. 81, 4222-4226.

Friedlander, D.R., Grumet, M., and Edelman, G.M. (1986). Nerve growth factor enhances expression of neuron-glia cell adhesion molecule in PC12 cells. *J. Cell Biol.* 102, 413-419.

Frischauf, A.-M., Lehrach, H., Polstka, A., and Murray, N.M. (1983). Lambda replacement vectors carrying polylinker sequences. *J. Mol. Biol.* 170, 827-842.

Fujisawa, H., Thanos, S., and Schwarz, U. (1984). Mechanisms in the development of retinotectal projections in the chick embryo studied by surgical deflection of the retinal pathway. *Dev. Biol.* 102, 356-367.

Furley, A.J., Morton, S.B., Manalo, D., Karagogeos, D., Dodd, J., and Jessel, T.M. (1990). The axonal glycoprotein TAG-1 is an immunoglobulin





superfamily member with neurite-promoting activity. *Cell*. 61, 157-170.

Futuyma, D.J. (1986). *Evolutionary Biology*, Second Edition. (: Sinauer Press, Sunderland, MA).

Gallin, W.J., Chuong, C.-M., Finkel, L.H., and Edelman, G.M. (1986). Antibodies to liver cell adhesion molecules perturb inductive interactions and alter feather pattern and structure. *Proc. Natl. Acad. Sci. USA*. 83, 8235-8239.

Gallin, W.J., Edelman, G.M., and Cunningham, B.A. (1983). Characterization of L-CAM, a major cell adhesion molecule from embryonic liver cells. *Proc. Natl. Acad. Sci. USA*. 80, 1038-1042.

Galtsoff, P.S. (1925). Regeneration after dissociation (an experimental study on sponges) II. Histogenesis of *Microciona prolifera*. *Verr. J. Exp. Zool.* 42, 223-251.

Gaze, R.M., Keating, M.J., Ostberg, A., and Chung, S.-H. (1979). The relationship between retinal and tectal growth in larval *Xenopus*: Implications for the development of retino-tectal projection. *J. Embryol. Exp. Morphol.* 53, 103-143.

Geiger, B. (1989). Cytoskeleton-associated cell contacts. *Curr. Opin. Cell Biol.* 1, 103-109.

Geliebter, J. (1987). Dideoxynucleotide sequencing of RNA and uncloned cDNA. *BRL Focus*. 9:1, 5-8.

Gennarini, G., Cibelli, G., Rougon, G., Mattei, M.-G., and Goridis, C. (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.

Gerisch, G. (1980). Univalent antibody fragments as tools for the analysis of cell interactions in *Dictyostelium*. *Curr. Topics Dev. Biol.* 14, 243-270.

Gloor, S., Antonicek, H., Sweadner, K.J., Pagliusi, S., Frank, R., Moos, M., and Schachner, M. (1990). The adhesion molecule on glia (AMOG) is a homologue of the B subunit of the Na,K-ATPase. *J. Cell Biol.* 110, 165-174.

Goridis, C., and Wille, W. (1988). The three size classes of mouse NCAM proteins arise from a single gene by a combination of alternative splicing and use of different polyadenylation sites. *Neurochem Int.* 12, 269-272.

Goring, D.R., Rossant, J., Clapoff, S., Breitman, M.L., and Tsui, L.-C. (1987). In Situ Detection of  $\beta$ -Galactosidase in Lenses of Transgenic Mice with a  $\beta$ -Crystallin/*lacZ* Gene. *Science*. 235, 456-458.

Gottlieb, D.I., Rock, K., and Glaser, L. (1976). A gradient of adhesive specificity in developing avian retina. *Proc. Natl. Acad. Sci. USA*. 73, 410-414.



Grady, S.R., and McGuire, E.G. (1976). Intercellular adhesive selectivity III. Species selectivity of embryonic liver intercellular adhesion. *J. Cell Biol.* 71, 96-106.

Graves, D.J. (1983). Use of peptide substrates to study the specificity of phosphorylase kinase phosphorylation. *Meth. Enzymol.* 99, 268-278.

Grenningloh, G., Bieber, A., Rehm, J., Snow, P., Traquina, Z., Hortsch, M., Patel, N., and Goodman, C.S. (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).

Grumet, M. (1991). Structure, expression and function of Ng-CAM, a member of the immunoglobulin superfamily involved in neuron-neuron and neuron-glia adhesion, *J. Neurosci. Res.* (in press).

Grumet, M., and Edelman, G.M. (1984). Heterotypic binding between neuronal membrane vesicles and glial cells is mediated by a specific cell adhesion molecule. *J. Cell Biol.* 98, 1746-1756.

Grumet, M., and Edelman, G.M. (1988). Neuron-glia cell adhesion molecule interacts with neurons and astroglia via different binding mechanisms. *J. Cell Biol.* 106, 487-503.

Grumet, M., Hoffman, S., Chuong, C.-M., and Edelman, G.M. (1984b). Polypeptide components and binding functions of neuron-glia cell adhesion molecules. *Proc. Natl. Acad. Sci. USA.* 81, 7989-7993.

Grumet, M., Hoffman, S., Crossin, K.L., and Edelman, G.M. (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.

Grumet, M., Hoffman, S., and Edelman, G.M. (1984a). Two antigenically related neuronal CAMs of different specificities mediate neuron-neuron and neuron-glia adhesion. *Proc. Natl. Acad. Sci. USA.* 81, 267-271.

Grumet, M., Mauro, V., Burgoon, M.P., Edelman, G.M., and Cunningham, B.A. (1991). Structure of a new nervous system glycoprotein, Nr-CAM, and its relationship to subgroups of neural cell adhesion molecules, *J. Cell Biol.* (in press).

Grumet, M., Rutishauser, U., and Edelman, G.M. (1983). Neuron-glia adhesion is inhibited by antibodies to neural determinants. *Science.* 222, 60-62.

Gubler, U., and Hoffman, B.J. (1983). A simple and very efficient method for generating cDNA libraries. *Gene.* 25, 263-269.

Gumbiner, B., Stevenson, B., and Grimaldi, A. (1988). The role of the cell adhesion molecule uvomorulin in the formation and maintenance of the epithelial junctional complex. *J. Cell Biol.* 107, 1575-1587.



- Hanahan, D., and Meselson, M. (1980). Plasmid screening at high colony density. *Gene*. 10, 63-67.
- Hanks, S.K., Quinn, A.M., and Hunter, T. (1988). The protein kinase family: conserved features and deduced phylogeny of the catalytic domains. *Science*. 241, 42-52.
- Harper, J.R., Prince, J.T., Healy, P.A., Stuart, J.K., Nauman, S.J., and Stallcup, W.B. (1991). Isolation and Sequence of Partial cDNA Clones of Human Ll: Homology of Human and Rodent Ll in the cytoplasmic Region. *Journal of Neurochemistry*. 56, 797-804.
- Harrelson, A.L., and Goodman, C.S. (1988). Growth cone guidance in insects: Fasciclin II is a member of the immunoglobulin superfamily. *Science*. 242, 700-707.
- Hathaway, G.M., and Traugh, J.A. (1982). Casein kinases - multipotential protein kinases. *Current Top. Cell. Reg.* 21, 101-127.
- Hatta, K., Nose, A., Nagafuchi, M., and Takeichi, M. (1988). Cloning and expression of cDNA encoding a neural calcium-dependent cell adhesion molecule-Its identity in the cadherin gene family. *J. Cell Biol.* 106, 873-881.
- Hausman, R.E., and Moscona, A.A. (1975). Purification and characterization of the retina-specific cell-aggregating factor. *Proc. Natl. Acad. Sci. USA*. 72, 916-920.
- Hausman, R.E., and Moscona, A.A. (1979). Immunologic detection of retina cognin on the surface of embryonic cells. *Exp. Cell Res.* 119, 191-204.
- He, H.T., Barbet, J., Chaix, J.C., and Goridis, C. (1986). Phosphatidylinositol is involved in the membrane attachment of NCAM-120, the smallest component of the neural cell adhesion molecule. *EMBO J.* 5, 2489-2494.
- Hemperly, J.J., Edelman, G.M., and Cunningham, B.A. (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.
- Hoffman, S., and Edelman, G.M. (1983). Kinetics of homophilic binding by E and A forms of the neural cell adhesion molecule. *Proc. Natl. Acad. Sci. USA*. 80, 5762-5766.
- Hoffman, S., and Edelman, G.M. (1987). A proteoglycan with HNK-1 antigenic determinants is a neuron-associated ligand for cytostatin. *Proc. Natl. Acad. Sci. USA*. 84, 2523-2527.
- Hoffman, S., Friedlander, D.R., Chuong, C.-M., Grumet, M., and Edelman, G.M. (1986). Differential contributions of Ng-CAM and N-CAM to cell adhesion in different neural regions. *J. Cell Biol.* 103, 145-158.
- Hoffman, S., Sorkin, B.C., White, P.C., Brackenbury, R., Mailhammer, R.,



Rutishauser, U., Cunningham, B.A., and Edelman, G.M. (1982). Chemical characterization of a neural cell adhesion molecule purified from embryonic brain membranes. *J. Biol. Chem.* 257, 7720-7729.

Holley, J.A., and Silver, J. (1987). Growth pattern of pioneer chick spinal cord axons. *Dev. Biol.* 123, 375-388.

Holloway, P.W. (1973). A simple procedure for removal of Triton X-100 from protein samples. *Anal. Biochem.* 53, 304-308.

Holtfreter, J. (1948a). The mechanism of embryonic induction and its relation to parthenogenesis and malignancy. *Symp. Soc. Exp. Biol.* 11, 17.

Holtfreter, J. (1948b). Significance of the cell membrane in embryonic processes. *Ann. NY Acad. Sci.* 49, 709-760.

Hortsch, M., Bieber, A.J., Patel, N.H., and Goodman, C.S. (1990). Differential splicing generates a nervous system-specific form of drosophila neuroglian. *Neuron.* 4, 697-709.

Hubbard, S.C., and Ivatt, R.J. (1981). Synthesis and processing of asparagine-linked oligosaccharides. *Ann. Rev. Biochem.* 50, 555-583.

Huesgen, A., and Gerisch, G. (1975). Solubilized contact sites A from cell membranes of *Dictyostelium discoideum*. *FEBS Lett.* 56, 46-49.

Humphreys, T. (1975). Biochemical studies on cell association and its relevance to neoplasia. in: *Cellular Membrane and Tumor Cell Behavior.* 28, 173-192.

Hunter, T., and Cooper, J.A. (1985). Protein-tyrosine kinases. *Ann. Rev. Biochem.* 54, 897-930.

Hyafil, F., Morello, D., Babinet, C., and Jacob, F. (1980). A cell surface glycoprotein involved in the compaction of embryonal carcinoma cells and cleavage stage embryos. *Cell.* 21, 927-934.

Hynes, R.O. (1987). Integrins: A family of cell surface receptors. *Cell.* 48, 549-554.

Ilyas, A.A., Quarles, R.H., and Brady, R.O. (1984). The monoclonal antibody HNK-1 reacts with a peripheral nerve ganglioside. *Biochem. Biophys. Res. Commun.* 122, 1206-1211.

Imhof, B.A., Vollmers, H.P., Goodman, S.L., and Birchmeier, W. (1983). Cell-cell interaction and polarity of epithelial cells- Specific perturbation using a monoclonal antibody. *Cell.* 35, 667-675.

Jaffe, S.H., Friedlander, D.R., Matsuzaki, F., Crossin, K.L., Cunningham, B.A., and Edelman, G.M. (1990). Differential effects of the cytoplasmic domains of cell adhesion molecules on cell aggregation and cell sorting, *Proc. Natl. Acad. Sci. USA* (in press).

Jakoi, E.R., and Marchase, R.B. (1979). Ligatin from embryonic chick





neural retina. *J. Cell Biol.* 80, 642-650.

Jakoi, E.R., and Corley, R.B. (1979). Ligatin: An external cell surface receptor specific for hexosaminidase A isolated from mouse peritoneal macrophages. *J. Cell Biol.* 83, 65a.

Jakoi, E.R., Zampighi, G., and Robertson, J.D. (1976). Regular structures in unit membranes. II. Morphological and biochemical characterization of two water-soluble membrane proteins isolated from the suckling rat ileum. *J. Cell Biol.* 70, 97-111.

Johns, P.R. (1977). Growth of the adult goldfish eye. III. Source of the new retinal cells. *J. Comp. Neurol.* 176, 343-358.

Johnson, P.F., and McKnight, S.L. (1989). Eukaryotic Transcriptional Regulatory Proteins. *Ann. Rev. Biochem.* 58, 799-839.

Jones, F.S., Hoffman, S., Cunningham, B.A., and Edelman, G.M. (1989). A detailed structural model of cytactin: Protein homologies, alternative RNA splicing, and binding regions. *Proc. Natl. Acad. Sci. USA.* 86, 1905-1909.

Kadmon, G., Kowitz, A., Altevogt, P., and Schachner, M. (1990a). The neural cell adhesion molecule N-CAM enhances L1-dependent cell-cell interactions. *J. Cell Biol.* 110, 193-208.

Kadmon, G., Kowitz, A., Altevogt, P., and Schachner, M. (1990b). Functional Cooperation between the Neural Adhesion Molecules L1 and N-CAM Is Carbohydrate Dependent. *J. Cell Biol.* 110, 209-218.

Kaufman, J.F., Auffray, C., Korman, A.J., Shackelford, D.A., and Strominger, J. (1984). The class II molecules of the human and murine major histocompatibility complex. *Cell.* 36, 1.

Keilhauer, G., Faissner, A., and Schachner, M. (1985). Differential inhibition of neurone-neurone, neurone-astrocyte and astrocyte-astrocyte adhesion by L1, L2, and N-CAM antibodies. *Nature (London).* 316, 728-730.

Kirschner, D.A., and Ganster, A.L. (1980). Compact myelin exists in the absence of basic protein in the shiverer mutant mouse. *Nature.* 283, 207-210.

Kishimoto, T.K., and , e.a. (1989). *Adv. Immun.* 46, 149-182.

Kleitman, N., Wood, P.M., Johnson, M.I., and Bunge, R.P. (1988). Schwann cell surfaces but not extracellular matrix organized by Schwann cells support neurite outgrowth from embryonic rat retina. *J. Neurosci.* 8, 653-663.

Koch, P.J., Walsh, M.J., Schmelz, M., Goldschmidt, M.D., Simbelmann, R., and Franke, W.W. (1990). Identification of desmoglian, a constitutive desmosomal glycoprotein, as a member of the cadherin family of cell adhesion molecules. *Eur. J. Cell. Biol.* 53, 1-12.



- Kornblihtt, A.R., Umezawa, K., Vibe-Pedersen, K., and Baralle, F.E. (1985). Primary structure of human fibronectin: differential splicing may generate at least 10 polypeptides from a single gene. *EMBO J.* 4, 1755-1759.
- Kruse, J., Keilhauer, G., Faissner, A., Timpl, R., and Schachner, M. (1985). The J1 glycoprotein- A novel nervous system cell adhesion molecule of the L2/HNK-1 family. *Nature.* 316, 146-148.
- Kruse, J., Mailhammer, R., Wernecke, H., Faissner, A., Sommer, I., Goridis, C., and Schachner, M. (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.
- Kuo, S.-C., and Lampen, J.O. (1974). Tunicamycin - An inhibitor of yeast glycoprotein synthesis. *Biochem. and Biophys. Res. Comm.* 58, 287-295.
- Kuret, J., Woodgett, J.R., and Cohen, P. (1985). Multisite phosphorylation of glycogen synthase from rabbit skeletal muscle. Identification of the sites phosphorylated by casein kinase I. *Eur. J. Biochem.* 151, 39-48.
- Kyte, J., and Doolittle, R.F. (1982). A simple method for displaying the hydropathic 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.* 227, 680-685.
- Lagenaur, C., and Lemmon, V. (1987). An L1-like molecule, the 8D9 antigen, is a potent substrate for neurite extension. *Proc. Natl. Acad. Sci. USA.* 84, 7753-7757.
- Lai, C., Brow, M.A., Nave, K.-A., Noronha, A.B., Quarles, R.H., Bloom, F.E., Milner, R.J., and Sutcliffe, J.G. (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.
- Landmesser, L., Dahm, L., Schultz, K., and Rutishauser, U. (1988). Distinct roles for adhesion molecules during innervation of embryonic chick muscle. *Devel. Biol.* 130, 645-670.
- Lasky, L.A. (1991). Lectin cell adhesion molecules (LEC-CAMs): A new family of cell adhesion proteins involved with inflammation. *J. Cell. Biochem.* 45, 139-146.
- Lehmann, S., Kuchler, S., Theveniau, M., Vincendon, G., and Zanetta, J.-P. (1990). An endogenous lectin and one of its neuronal glycoprotein ligands are involved in contact guidance of neuron migration. *Proc. Natl. Acad. Sci. USA.* 87, 6455-6459.
- Lemke, G., and Axel, R. (1985). Isolation and sequence of a cDNA encoding the major structural protein of peripheral myelin. *Cell.* 40, 501-508.



- Lemke, G., Lamar, E., and Patterson, J. (1988). Isolation and analysis of the gene encoding peripheral myelin protein zero. *Neuron*. 1, 73-83.
- Lemmon, V., Farr, K.L., and Lagenaur, C. (1989). Ll-mediated axon outgrowth occurs via a homophilic binding mechanism. *Neuron*. 2, 1597-1603.
- Lemmon, V., and McLoon, S.C. (1986). The appearance of an Ll-like molecule in the chick primary visual pathway. *J. Neurosci*. 6, 2987-2994.
- Lewis, S.A., and Cowan, N.J. (1986). Anomalous placement of introns in a member of the intermediate filament multigene family: an evolutionary conundrum. *Mol. Cell Biol.* May, 1529-1534.
- Lindner, J., Rathjen, F.G., and Schachner, M. (1983). Ll Mono- and polyclonal antibodies modify cell-migration in early postnatal mouse cerebellum. *Nature (Lond)*. 305, 427-430.
- Low, M.G., and Saltiel, A.R. (1988). Structural and functional roles of glycosyl-phosphatidylinositol in membranes. *Science*. 239, 268-275.
- MacDonald, R.J., Swift, G.H., Przybyla, A.E., and Chirgwin, J.M. (1987). Isolation of RNA Using Guanidinium Salts. *Methods in Enzymology*. 152, 219-227.
- Mackie, E.J., Tucker, R.P., Halfter, W., Chiquet-Ehrismann, R., and Epperlein, H.H. (1988). The distribution of tenascin coincides with pathways of neural crest cell migration. *Development*. 102, 237-250.
- Maniatis, T., Fritsch, E.F., and Sambrook, J. (1982). *Molecular Cloning: A Laboratory Manual*. Cold Spring Harbor Laboratory. .
- Marchase, R.B. (1977). Biochemical investigations of retinotectal adhesive specificity. *J. Cell Biol.* 75, 237-257.
- Marchase, R.B., Harges, P., and Jakoi, E.R. (1981). Ligatin from embryonic chick neural retina inhibits retinal cell adhesion. *Develop. Biol.* 86, 250-255.
- Marlin, S.D., and Springer, T.A. (1987). Purified Intercellular Adhesion Molecule-1 (ICAM-1) is a ligand for Lymphocyte Function-Associated Antigen 1 (LFA-1). *Cell*. 51, 813-819.
- Martini, R., and Schachner, M. (1986). Immunoelectron microscopic localization of neural cell adhesion molecules (Ll, N-CAM, and MAG) and their shared carbohydrate epitope and myelin basic protein in developing sciatic nerve. *J. Cell Biol.* 103, 2439-2448.
- Matsudaira, P. (1987). Sequence from picomole quantities of proteins electroblotted onto polyvinylidene difluoride membranes. *J. Biol. Chem.* 262, 10035-10038.
- Matsunaga, M., Hatta, K., and Takeichi, M. (1988). Role of n-cadherin cell adhesion molecules in the histogenesis of neural retina. *Neuron*. 4, 289-295.



- Matsuzaki, F., Mege, R.-M., Jaffe, S.H., Friedlander, D.R., Gallin, W.J., Goldberg, J.I., Cunningham, B.A., and Edelman, G.M. (1990). cDNAs of cell adhesion molecules of different specificity induce changes in cell shape and border formation in cultured S180 cells. *J. Cell Biol.* 110, 1239-1252.
- McCafferty, P., Neve, R.L., and Drager, U.C. (1990). A dorso-ventral asymmetry in the embryonic retina defined by protein conformation. *Proc. Natl. Acad. Sci. USA.* 87, 8570-8574.
- McClain, D.A., and Edelman, G.M. (1982). A neural cell adhesion molecule from human brain. *Proc. Natl. Acad. Sci. USA.* 79, 6380-6384.
- McGarry, R.C., Helfand, S.L., Quarles, R.H., and Roder, J.C. (1983). Recognition of myelin-associated glycoprotein by the monoclonal antibody HNK-1. *Nature.* 306, 376-378.
- McGuire, J.C., Greene, L.A., and Furano, A.V. (1978). NGF stimulates incorporation of fucose or glucosamine into an external glycoprotein in cultured rat PC12 pheochromocytoma cells. *Cell.* 115, 357-365.
- McIlhinney, R.A. (1990). The fats of life: the importance and function of protein acylation. *TIBS.* 15, 382-391.
- Mege, R.-M., Matsuzaki, F., Gallin, W.J., Goldberg, J.I., Cunningham, B.A., and Edelman, G.M. (1988). Construction of epithelioid sheets by transfection of mouse sarcoma cells with cDNAs for chicken cell adhesion molecules. *Proc. Natl. Acad. Sci. USA.* 85, 7274-7278.
- Mendelsohn, C.L., Wimmer, E., and Racaniello, V.R. (1989). Cellular receptor for poliovirus: molecular cloning, nucleotide sequence, and expression of a new member of the immunoglobulin superfamily. *Cell.* 56, 855-865.
- Meyer, R.L. (1978). Evidence from thymidine labeling for continuing growth of retina and tectum in juvenile goldfish. *Exp. Neurol.* 59, 99-111.
- Moos, M., Tacke, R., Scherer, H., Teplow, D., Fruh, K., and Schachner, M. (1988). Neural adhesion molecule L1 as a member of the immunoglobulin superfamily with binding domains similar to fibronectin. *Nature.* 334, 701-703.
- Morrissey, J.H. (1981). Silver stain for proteins in polyacrylamide gels-A modified procedure with enhanced uniform sensitivity. *Anal. Biochem.* 117, 307-310.
- Moscona, A.A. (1962). Analysis of cell recombinations in experimental synthesis of tissues *in vitro*. *J. Cell Comp. Physiol.* 60, 65-80.
- Moskal, J.R., Trisler, D., Schneider, M.D., and Nirenberg, M. (1986). Purification of a membrane protein distributed in a topographic gradient in chicken retina. *Proc. Natl. Acad. Sci. USA.* 83, 4730-4733.
- Murray, B.A., Owens, G.C., Prediger, E.A., Crossin, K.L., Cunningham,





- B.A., and Edelman, G.M. (1986). Cell surface modulation of the neural cell adhesion molecule resulting from alternative mRNA splicing in a tissue-specific developmental sequence. *J. Cell Biol.* 103, 1431-1439.
- Nagafuchi, A., Shirayoshi, Y., Okazaki, K., Yasuda, K., and Takeichi, M. (1987). Transformation of cell adhesion properties by exogenously introduced E- cadherin cDNA. *Nature.* 329, 341-343.
- Nelson, W.J., Shore, E.M., Wang, A.Z., and Hammerton, R.W. (1990). Identification of a membrane-cytoskeletal complex containing the cell adhesion molecule uvomorulin (E-cadherin), ankyrin, and fodrin in Madin-Darby Canine Kidney epithelial cells. *J. Cell Biol.* 110, 349-357.
- Neumaier, M., Zimmerman, W., Shively, L., Hinoda, Y., Riggs, A.D., and Shively, J.D. (1988). Characterization of a cDNA clone for the nonspecific cross-reacting antigen (NCA) and a comparison of NCA and carcinoembryonic antigen. *J. Biol.Chem.* 263, 3202-3207.
- Newman, P.J., Berndt, M.C., Gorski, J., White, G.C., II, Lyman, S., Paddock, C., and Muller, W.A. (1990). PECAM-1 (CD31) cloning and relation to adhesion molecules of the immunoglobulin gene superfamily. *Science.* 247, 1219-1222.
- Noble, M., Albrechtsen, M., Moller, C., Lyles, J., Bock, E., Goridis, C., Watanabe, M., and Rutishauser, U. (1985). Glial cells express N-CAM/D2-CAM-like polypeptides in vitro. *Nature.* 316, 725-728.
- Noegel, A., Gerisch, G., Stadler, J., and Westphal, M. (1986). Complete sequence and transcript regulation of a cell adhesion protein from aggregating dictyostelium cells. *EMBO J.* 5, 1473-1476.
- Noonan, D.M., Horigan, E.A., Ledbetter, S.R., Vogeli, G., Sasaki, M., Yamada, Y., and Hassel, J.R. (1988). Identification of cDNA clones encoding different domains of the basement membrane heparan sulfate proteoglycan. *J. Biol. Chem.* 263(31), 16379-16387.
- Nose, A., Nagafuchi, A., and Takeichi, M. (1987). Isolation of placental cadherin cDNA- Identification of a novel gene family of cell-cell adhesion molecules. *EMBO J.* 6, 3655-3661.
- Nose, A., Tsuji, K., and Takeichi, M. (1990). Localization of specificity determining sites in cadherin cell adhesion molecules. *Cell.* 61, 147-155.
- Nybroe, O., Albrechtsen, M., Dahlin, J., Linneman, D., Lyles, J.M., McEller, C.J., and Bock, E. (1985). Biosynthesis of the neural cell adhesion molecule: Characterization of polypeptide C. *J. Cell. Biol.* 101, 2310-2315.
- Obrink, B., Kuhlenschmidt, M.S., and Roseman, S. (1977). Adhesive specificity of juvenile rat and chicken liver cells and membranes. *Proc. Natl. Acad. Sci. USA.* 74, 1077-1081.
- Ocklind, C., and Öbrink, B. (1982). Intercellular adhesion of rat



- hepatocytes. Identification of a cell surface glycoprotein involved in the initial adhesion process. *J. Cell Biol.* 257, 6788-6795.
- Oikawa, S., Nakazato, H., and Kozaki, G. (1987). Primary structure of human carcinoembryonic antigen (CEA) deduced from cDNA sequence. *Biochem. Biophys. Res. Commun.* 142, 511-518.
- Ophir, I., Moscona, A.A., and Ben-Shaul, Y. (1984). Cell disorganization and malformation in neural retina caused by antibodies to R-cognin: ultrastructural study. *Cell Differentiation.* 15, 53-60.
- Ozawa, M., Baribault, H., and Kemler, R. (1989). The cytoplasmic domain of the cell adhesion molecule uvomorulin associates with three independent proteins structurally related in different species. *EMBO J.* 8, 1711-1717.
- Ozawa, M., Ringwald, M., and Kemler, R. (1990). Uvomorulin-catenin complex formation is regulated by a specific domain in the cytoplasmic region of the cell adhesion molecule. *Proc. Natl. Acad. Sci. USA.* 87, 4246-4250.
- Patel, N.H., Snow, P.M., and Goodman, C.S. (1987). Characterization and cloning of fasciclin III: a glycoprotein expressed on a subset of neurons and axon pathways in *Drosophila*. *Cell.* 48, 975-988.
- Pearson, C.A., Pearson, D., Shibahara, S., Hofsteenge, J., and Chiquet-Ehrismann, R. (1988). Tenascin: cDNA cloning and induction by TGF- $\beta$ . *EMBO J.* 7, 2977-2982.
- Pearson, W.R. (1990). Rapid and sensitive sequence comparison with FASTP and FASTA. *Meth. Enzymol.* 183, 63-98.
- Pearson, W.R., and Lipman, D.J. (1988). Improved tools for biological sequence comparison. *Proc. Natl. Acad. Sci. USA.* 85, 2444-2448.
- Perlman, K., and Halvorson, H.O. (1983). A putative signal peptidase recognition site and sequence in eukaryotic and prokaryotic signal peptides. *J. Mol. Biol.* 167, 391-409.
- Pierschbacher, M.D., and Ruoslahti, E. (1987). Influence of stereochemistry of the sequence Arg-Gly-Asp-Xaa on binding specificity in cell adhesion. *J. Biol. Chem.* 262(36), 17294-17298.
- Pollerberg, E.G., BurrIDGE, K., Krebs, K.E., Goodman, S.R., and Schachner, M. (1987). The 180-kD component of the neural cell adhesion molecule N-CAM is involved in cell-cell contacts and cytoskeleton-membrane interactions. *Cell Tiss. Res.* 250, 227-236.
- Poltorak, M., Sadoul, R., Keilhauer, G., Landa, C., Fahrig, T., and Schachner, M. (1987). Myelin-associated glycoprotein, a member of the L2/HNK-1 family of neural cell adhesion molecules, is involved in neuron-oligodendrocyte and oligodendrocyte-oligodendrocyte interaction. *J. Cell Biol.* 105, 1893-1899.
- Popowicz, A.M., and Dash, P.F. (1988). SIGSEQ: a computer program for predicting signal sequence cleavage sites. *Computer Appl. Biosci.* 4 (3),



Prediger, E.A., Hoffman, S., Edelman, G.M., and Cunningham, B.A. (1988). Four exons encode a 93-base-pair insert in three neural cell adhesion molecule mRNAs specific for chicken heart and skeletal muscle. *Proc. Natl. Acad. Sci. USA.* 85, 9616-9620.

Prieto, A.L., Crossin, K.L., Cunningham, B.A., and Edelman, G.M. (1989). Localization of mRNA for neural cell adhesion molecule (N-CAM) polypeptides in neural and nonneural tissues by *in situ* hybridization. *Proc. Natl. Acad. Sci. USA.* 86, 9579-9583.

Prince, J.T., Milona, N., and Stallcup, W.B. (1989). Characterization of a Partial cDNA Clone for the NILE Glycoprotein and Identification of the Encoded Polypeptide Domain. *J. Neurosci.* 9, 876-883.

Quarles, R.H. (1984). Myelin-associated glycoprotein in development and disease. *Dev. Neurosci.* 6, 285-303.

Rakic, P. (1971). Neuron-glia relationship during granule cell migration in developing cerebellar cortex. A Golgi and electron microscopic study in *Macacus rhesus*. *J. Comp. Neurol.* 141, 283-312.

Ranscht, B. (1988). Sequence of contactin, 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.

Ranscht, B., and Bronner-Fraser, M. (1991). T-cadherin expression alternates with migrating neural crest cells in the trunk of the avian embryo. *Development.* 111, 15-22.

Ranscht, B., and Dours, M.T. (1989). Selective expression of a novel cadherin in the pathways of developing motor- and commissural axons. *Society for Neuroscience Abstracts* 15, 959 (Abstract).

Rathjen, F.G., and Rutishauser, U. (1984). Comparison of two cell surface molecules involved in neural cell adhesion. *EMBO J.* 3, 461-465.

Rathjen, F.G., and Schachner, M. (1984). Immunocytological and biochemical characterization of a new neuronal cell surface component (L1 antigen), which is involved in cell adhesion. *EMBO J.* 3, 1-10.

Rathjen, F.G., Wolff, J.M., Chang, S., Bonhoeffer, F., and Raper, J.A. (1987b). Neurofascin: A novel chick cell-surface glycoprotein involved in neurite-neurite interactions. *Cell.* 51, 841-849.

Rathjen, F.G., Wolff, J.M., Frank, R., Bonhoeffer, F., and Rutishauser, U. (1987a). Membrane glycoproteins involved in neurite fasciculation. *J. Cell Biol.* 104, 343-353.

Ratner, N., Elbein, A., Bunge, M.B., Porter, S., Bunge, R.P., and Glaser, L. (1986). Specific asparagine-linked oligosaccharides are not required for certain neuron-neuron and neuron-Schwann cell interactions. *J. Cell*



Biol. 103, 159-170.

Raymond, P.A., and Easter, S.S. (1983). Postembryonic growth of the optic tectum in goldfish: I. Location of germinal cells and numbers of neurons produced. *J. Neuroscience*. 3 (5), 1077-1091.

Reh, T.A., and Constantine-Paton, M. (1984). Retinal ganglion cell terminals change their projection sites during larval development of *Rana pipiens*. *J. Neurosci*. 4, 442-457.

Reyes, A.A., Akeson, R., Brezina, L., and Cole, G.J. (1990). Structural requirements for neural cell adhesion molecule-heparin interaction. *Cell Regulation*. 1, 567-576.

Rhyner, T.A., Biguet, N.F., Berard, S., Borbely, A.A., and Maller, J. (1986). An efficient approach for the selective isolation of specific transcripts from complex brain populations. *J. Neurosci. Res*. 16, 167-181.

Rieger, F., Daniloff, J.K., Pinçon-Raymond, M., Crossin, K.L., Grumet, M., and Edelman, G.M. (1986). Neuronal cell adhesion molecules and cytotactin are colocalized at the node of Ranvier. *J. Cell Biol*. 103, 379-391.

Rieger, F., Nicolet, M., Pincon-Raymond, M., Levi, G., and Edelman, G.M. (1988). Distribution and role in regeneration of N-CAM in basal laminae of muscle and Schwann cells. *J. Cell Biol*. 107, 707-719.

Ringwald, M., Schuh, R., Vestweber, D., Eistetter, H., Lohspeich, F., Engel, J., Dotz, R., Jahnig, F., Epplen, J., Mayer, S., Muller, C., and Kemler, R. (1987). The structure of cell adhesion molecule uvomorulin. Insights into the molecular mechanism of Ca-2+-dependent cell adhesion. *EMBO J*. 6, 3647-3653.

Riopelle, R.J., McGarry, R.C., and Roder, J.C. (1986). Adhesion properties of a neuronal epitope recognized by the monoclonal antibody HNK-1. *Brain Research*. 367, 20-25.

Roseman, S. (1970). The synthesis of complex carbohydrates by multiglycosyltransferase systems and their potential function in intercellular adhesion. *Chem. Phys. Lipids*. 5, 270-297.

Roth, S., MacGuire, E.J., and Roseman, S. (1971). Evidence for cell-surface glycosyltransferases. Their potential role in cellular recognition. *J. Cell Biol*. 51, 536-547.

Roth, S.A., and Weston, J.A. (1967). The measurement of intercellular adhesion. *Zool*. 58, 974-980.

Ruegg, M.A., Stoeckli, E.T., Lanz, R.B., Streit, P., and Sonderegger, P. (1989a). 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.

Ruegg, M.A., Stoeckli, E.T., Lanz, R.B., Streit, P., and Sonderegger, P. (1989c). 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.
- Ruegg, M.A., Stoeckli, E.T., Kuhn, T.B., Heller, M., Züellig, R., and Sonderegger, P. (1989b). Purification of axonin-1, a protein that is secreted from axons during neurogenesis. *EMBO J.* 8, 55-63.
- Ruoslahti, E. (1988). Structure and biology of proteoglycans. *Annu. Rev. Cell. Biol.* 4, 229-255.
- Ruoslahti, E., and Pierschbacher, M.D. (1987). New perspectives in cell adhesion- RGD and integrins. *Science.* 238, 491-497.
- Rutishauser, U., Grumet, M., and Edelman, G.M. (1983). N-CAM mediates initial interactions between spinal cord neurons and muscle cells in culture. *J. Cell Biol.* 97, 145-152.
- Rutishauser, U., Watanabe, M., Silver, J., Troy, F.A., and Vimr, E.R. (1985). Specific alteration of NCAM-mediated cell adhesion by an endoneuraminidase. *J. Cell Biol.* 101, 1842-1849.
- Sadoul, K., Sadoul, R., Faissner, A., and Schachner, M. (1988). Biochemical characterization of different molecular forms of the neural Cell Adhesion Molecule L1. *J. Neurochem.* 50, 510-521.
- Sadoul, R., Kirchhoff, F., and Schachner, M. (1989). A protein kinase activity is associated with and specifically phosphorylates the neural Cell Adhesion Molecule L1. *J. Neurochem.* 53, 1471-1478.
- Saiki, R.K., Gelfand, D.H., Stoffel, S., Scharf, S.J., Higuchi, R., Horn, G.T., Mullis, K.B., and Erlich, H.A. (1988). Primer-directed enzymatic amplification of DNA with a thermostable DNA polymerase. *Science.* 239, 487-491.
- Sajovic, P., Kouvelas, E., and Trenkner, E. (1986). Probable identity of NILE glycoprotein and the high-molecular-weight component of L1 antigen. *J. Neurochem.* 47, 541-546.
- Salton, S.R.J., Richter-Landsberg, C., Greene, L.A., and Shelanski, M.L. (1983a). Nerve growth factor-inducible large external (NILE, glycoprotein-Studies of a central and peripheral neuronal marker. *J. Neurosci.* 3, 441-454.
- Salton, S.R.J., Shelanski, M.L., and Greene, L.A. (1983b). Biochemical properties of the nerve growth factor-inducible large external (NILE), glycoprotein. *J. Neurosci.* 3, 2420-2430.
- Salzer, J.G., Holmes, W.P., and Colman, D.R. (1987). The amino acid sequences of the myelin-associated glycoproteins- Homology to the immunoglobulin gene superfamily. *J. Cell Biol.* 104, 957-965.
- Salzer, J.L., and Colman, D.R. (1989). Mechanisms of cell adhesion in the nervous system: role of the immunoglobulin gene superfamily. *Dev. Neurosci.* 11, 377-390.



Sanger, F., Nicklen, S., and Coulson, A.R. (1977). DNA sequencing with chain-terminating inhibitors. *Proc. Natl. Acad. Sci. USA.* 74, 5463-5467.

Santoni, M.J., Barthels, D., Vopper, G., Boned, A., Goridis, C., and Wille, W. (1989). Differential exon usage involving an unusual splicing mechanism generates at least eight types of NCAM cDNA in mouse brain. *EMBO J.* 8, 385-392.

Schlosshauer, B., Schwarz, U., and Rutishauser, U. (1984). Topological distribution of different forms of neural cell adhesion molecule in the developing chick visual system. *Nature.* 310, 141-143.

Schneider-Schaulies, J., Brunn, v., A., S., and , M. (1990). Recombinant peripheral myelin protein Po confers both adhesion and neurite outgrowth-promoting properties. *J. Neurosci. Res.* 27, 286-297.

Schofield, P.R., McFarland, K.C., Hayflick, J.S., Wilcox, J.N., Cho, T.M., Roy, S., Lee, N.M., Loh, H.H., and Seeburg, P.H. (1989). Molecular characterization of a new immunoglobulin superfamily protein with potential roles in opioid binding and cell contact. *EMBO J.* 8, 489-495.

Schuch, U., Lohse, M.J., and Schachner, M. (1989). Neural cell adhesion molecules influence second messenger systems. *Neuron.* 3, 13-20.

Schwarzbauer, J.E., Patel, R.S., Fonda, D., and Hynes, R.O. (1987). Multiple sites of alternative splicing of the rat fibronectin gene transcript. *EMBO J.* 6, 2573-2580.

Seeger, M.A., Haffley, L., and Kaufman, T.C. (1988). Characterization of amalgam: a member of the immunoglobulin superfamily from *Drosophila*. *Cell.* 55, 589-600.

Seilheimer, B., and Schachner, M. (1988). Studies of adhesion molecules mediating interactions between cells of peripheral nervous system indicate a major role for L1 in mediating sensory neuron growth on Schwann cells in culture. *J. Cell Biol.* 107, 341-351.

Shashua, V.E., Daniel, P.F., Moore, H.E., and Jungawala, F.B. (1986). Demonstration of glucuronic acid on brain glycoproteins which react with HNK-1 antibody. *Biochem. Biophys. Res. Comm.* 138, 902.

Shiga, T., Oppenheim, R.W., Grumet, M., and Edelman, G.M. (1990). Neuron-glia cell adhesion molecule (Ng-CAM) expression in the chick spinal cord: Observations on the earliest developing intersegmental interneurons. *Devel. Brain Res.* 55, 209-217.

Simmons, D., Makgoba, M.W., and Seed, B. (1988). ICAM, an adhesion ligand of LFA-1, is homologous to the neural cell adhesion molecule NCAM. *Nature.* 331, 624-627.

Small, S.J., Haines, S.L., and Akeson, R.A. (1988). Polypeptide Variation in an N-CAM Extracellular Immunoglobulin-like Fold Is Developmentally Regulated through Alternative Splicing. *Neuron.* 1, 1007-1017.



Sorkin, B.C., Grumet, M., Cunningham, B.A., and Edelman, G.M. (1985). Structures of two neuronal cell adhesion molecules. *J. Neurosci.* 11, 1138a.

Sorkin, B.C., Hoffman, S., Edelman, G.M., and Cunningham, B.A. (1984). Sulfation and phosphorylation of the neural cell adhesion molecule N-CAM. *Science.* 225, 1476-1478.

Sperry, R.W. (1963). Chemoaffinity in the orderly growth of nerve fiber patterns and connections. *Proc. Natl. Acad. Sci. USA.* 50, 703-710.

Springer, T.A., and Lasky, L.a. (1991). Sticky sugars for selectins. *Nature.* 349, 196-197.

Staden, R. (1982). Automation of the computer handling of gel reading data produced by the shotgun method of DNA sequencing. *Nucleic Acids Res.* 10, 4731-4751.

Staden, R. (1984). Graphic methods to determine the function of nucleic acid sequences. *Nucleic Acids Research.* 12, 521-538.

Stallcup, W.B., and Beasley, L.L. (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.

Stallcup, W.B., Beasley, L.L., and Levine, J.M. (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.

Staunton, D.E., Dustin, M.L., and Springer, T.A. (1989). Functional cloning of ICAM-2, a cell adhesion ligand for LFA-1 homologous to ICAM-1. *Nature.* 339, 61-64.

Staunton, D.E., Marlin, S.D., Stratowa, C., Dustin, M.L., and Springer, T.A. (1988). Primary structure of ICAM-1 demonstrates interaction between members of the immunoglobulin and integrin supergene families. *Cell.* 52, 925-933.

Steinberg, M.S. (1970). Does differential adhesion govern self-assembly processes in histogenesis? Equilibrium configurations and the emergence of a hierarchy among populations of embryonic cells. *J. Exp. Zool.* 173, 395-434.

Stengelin, S., Stamenkovic, I., and Seed, B. (1988). Isolation of cDNAs for two distinct Fc receptors by ligand affinity cloning. *EMBO J.* 7, 1053-1059.

Sternberger, N.H., Quarles, R.H., Itoyama, Y., and Webster, H.d. (1979). Myelin-associated glycoprotein demonstrated immunocytochemically in myelin and myelin-forming cells of developing rat. *Proc. Natl. Acad. Sci. USA.* 76, 1510-1514.



- Stitt, T.N., and Hatten, M.E. (1990). Antibodies that recognize astrotactin block granule neuron binding to astroglia. *Neuron*. 5, 639-649.
- Stoeckli, E.T., Kuhn, T.B., Duc, C.O., Ruegg, M.A., and Sonderegger, P. (1991). The axonally secreted protein Axonin-1 is a potent substratum for neurite growth. *J. Cell Biol.* 112, 449-455.
- Stoeckli, E.T., Lemkin, P.F., Kuhn, T.B., Ruegg, M.A., Heller, M., and Sonderegger, P. (1989). Identification of proteins secreted from axons of embryonic dorsal root ganglia neurons. *Eur. J. Biochem.* 180, 249-258.
- Stull, J.T., Nunnally, M.H., and Michnoff, C.H. (1986). Calmodulin-dependent protein kinases. *The Enzymes*. 17, 114-159.
- Sun, S.-C., Lindstrom, I., Boman, H.G., Faye, I., and Schmidt, O. (1990). Hemolin: An insect-immune protein belonging to the immunoglobulin superfamily. *Sci.* 250, 1729-1732.
- Takeichi, M. (1988). The cadherins: cell-cell adhesion molecules controlling animal morphogenesis. *Development*. 102, 639-655.
- Takeichi, M. (1990). Cadherins: a molecular family important in selective cell-cell adhesion. *Annu. Rev. Biochem.* 59, 237-252.
- Takeichi, M. (1991). Cadherin cell adhesion receptors as a morphogenetic regulator. *Science*. 251, 1451-1455.
- Takeichi, M., Atsumi, T., Yoshida, C., Uno, K., and Okada, T.S. (1981). Selective adhesion of embryonal carcinoma cells and differentiated cells by Ca-2+ dependent sites. *Devl. Biol.* 87, 340-350.
- Thanos, S., Bonhoeffer, F., Rutishauser, a., and , U. (1984). Fiber-fiber interaction and tectal cues influence the development of the chicken retinotectal projection. *Proc. Natl.Acad. Acad. Sci. USA*. 81, 1906-1910.
- Thiery, J.-P., Brackenbury, R., Rutishauser, U., and Edelman, G.M. (1977). Adhesion among neural cells of the chick embryo. II. Purification and characterization of a cell adhesion molecule from neural retina. *J. Biol. Chem.* 252, 6841-6845.
- Thiery, J.-P., Delouvé, A., Gallin, W.J., Cunningham, B.A., and Edelman, G.M. (1984). Ontogenetic expression of cell adhesion molecules: L-CAM is found in epithelia derived from the three primary germ layers. *Dev. Biol.* 102, 61-78.
- Thiery, J.-P., Delouvé, A., Grumet, M., and Edelman, G.M. (1985). Initial appearance and regional distribution of the neuron-glia cell adhesion molecule, in the chick embryo. *J. Cell Biol.* 100, 442-456.
- Thor, G., Probstmeier, R., and Schachner, M. (1987). Characterization of the cell adhesion molecules, Ll, N-CAM and J1 in the mouse intestine. *EMBO*. 6(9), 2581-2586.





- Tindall, K.R., and Kunkel, T.A. (1988). Fidelity of DNA synthesis by the *Thermis aquaticus* DNA polymerase. *Biochemistry*. 27, 6008-6013.
- Townes, P.L., and Holtfreter, J. (1955). Directed movements and selective adhesion of embryonic amphibian cells. *J. Exp. Zool.* 128, 53-120.
- Trapp, B.D. (1988). Distribution of the myelin-associated glycoprotein and P0 protein during myelin compaction in Quaking mouse peripheral nerve. *J. Cell Biol.* 107, 675-685.
- Trapp, B.D., Quarles, R.H., and Suzuki, K. (1984). Immunocytochemical studies of quaking mice support a role for the myelin-associated glycoprotein in forming and maintaining the periaxonal space and periaxonal cytoplasmic collar in myelating Schwann. *J. Cell Biol.* 99, 595-601.
- Trisler, G.D., Schneider, M.D., and Nirenberg, M. (1981). A topographic gradient of molecules in retina can be used to identify neuron position. *Proc. Natl. Acad. Sci. USA.* 78, 2145-2149.
- Tuazon, P.T., Bingham, E.W., and Traugh, J.A. (1979). Cyclic nucleotide-independent protein kinases from rabbit reticulocytes. *Eur. J. Biochem.* 94, 497-504.
- Volk, T., and Geiger, B. (1984). A 135-kd membrane protein of intercellular adherens junctions. *EMBO J.* 3, 2249-2260.
- von Heijne, G. (1981). Membrane proteins. The amino acid composition of membrane-penetrating segments. *Eur. J. Biochem.* 120, 275-278.
- von Heijne, G. (1985). Signal Sequences: The limits of variation. *J. Mol. Biol.* 184, 99-105.
- von Heijne, G. (1986). A new method for predicting signal sequence cleavage sites. *Nucleic Acids Research.* 14, 4683-4690.
- Walsh, F.S., and Dickson, G. (1989). Generation of multiple N-CAM polypeptides from a single gene. *BioEssays.* 11, 83-87.
- Watson, M.E.E. (1984). Compilation of published signal sequences. *Nucl. Acids Res.* 12, 5145-5164.
- Waxdal, M.J., Konigsberg, W.H., Henley, W.L., and Edelman, G.M. (1968). The covalent structure of a human G-immunoglobulin. II. Isolation and characterization of the cyanogen bromide fragments. *Biochem.* 7, 1959-1966.
- Williams, A.F. (1987). A year in the life of the immunoglobulin superfamily. *Immunology Today.* 8, 298-303.
- Williams, A.F., and Barclay, A.N. (1988). The immunoglobulin superfamily-Domains for cell surface recognition. *Ann. Rev. Immunol.* 6, 381-405.
- Williams, R.K., Goridis, C., and Akeson, R. (1985). Individual neural cell types express immunologically distinct N-CAM forms. *J. Cell Biol.* 101,



Wilson, H.V. (1907). On some phenomena of coalescence and regeneration in sponges. *J. Exp. Zool.* 5, 245-258.

Wolff, J.M., Brummendorf, T., and Rathjen, F.G. (1989). Neural cell recognition molecule F11: Membrane interaction by covalently attached phosphatidylinositol. *Biochem. and Biophys. Res. Comm.* 161, 931-938.

Wolff, J.M., Rathjen, F.G., Frank, R., and Roth, S. (1987). Biochemical characterization of polypeptide components involved in neurite fasciculation and elongation. *Eur. J. Biochem.* 168, 551-561.

Wourms, J.P. (1972). The developmental biology of annual fishes III. Pre-embryonic and embryonic diapause of variable duration in the eggs of annual fishes. *J. Exp. Zool.* 182, 389-414.

Yarden, Y., Escobedo, J.A., Kuang, W.J., Yang-Feng, T.L., Daniel, T.O., Tremble, P.M., Chen, E.Y., Ando, M.E., Harkins, R.N., Francke, U., Fried, V.A., Ullrich, A., and Williams\* (1986). Structure of the receptor for platelet-derived growth factor helps define a family of closely related growth factor receptors. *Nature.* 323, 226-232.

Young, R.A., and Davis, R.W. (1983). Yeast RNA Polymerase III Genes: Isolation with Antibody Probes. *Science.* 222, 778-782.

Zanetta, J.P., Dontenwill, M., Meyer, A., and Roussel, G. (1985). Isolation and immunohistochemical localization of a lectin-like molecule from the rat cerebellum. *Dev. Brain Res.* 17, 233-243.

Zanetta, J.P., Meyer, A., Kuchler, S., and Vincendon, G. (1987). Isolation and immunochemical study of a soluble cerebellar lectin delineating its structure and function. *J. Neurochem.* 49, 1250-1257.

Zimmerman, W., Ortlieb, R., Fredrich, R., and Kleist, S. vo (1987). Isolation and characterization of cDNA clones encoding the human carcinoembryonic antigen reveal a highly conserved repeating structure. *Proc. Natl. Acad. Sci. USA.* 84, 2960-2964.

Zimmerman, W., Weber, B., Ortlieb, B., Rudert, F., Schempp, W., Fiebig, H.H., Shively, J.E., Kleist, v., S., T., and , J.A. (1988). Chromosomal localization of the carcinoembryonic antigen gene family and differential expression in various tumors. *Cancer Res.* 148, 2550-2554.

Zinn, K., McAllister, L., and Goodman, C.S. (1988). Sequence analysis and neuronal expression of fasciclin I in grasshopper and *Drosophila*. *Cell.* 53, 577-587.





End