Sequence of Contactin, a 130-kD Glycoprotein Concentrated in Areas of Interneuronal Contact, Defines a New Member of the Immunoglobulin Supergene Family in the Nervous System

Barbara Ranscht, with technical assistance from Maria Teresa Dours*

Massachusetts Institute of Technology, Cambridge, Massachusetts 02139; and
*La Jolla Cancer Research Foundation, Cancer Research Center, La Jolla, California 92037

Abstract. The primary amino acid sequence of contactin, a neuronal cell surface glycoprotein of 130 kD that is isolated in association with components of the cytoskeleton (Ranscht, B., D. J. Moss, and C. Thomas. 1984. *J. Cell Biol.* 99:1803–1813), was deduced from the nucleotide sequence of cDNA clones and is reported here. The cDNA sequence contains an open reading frame for a 1,071-amino acid transmembrane protein with 962 extracellular and 89 cytoplasmic amino acids. In its extracellular portion, the polypeptide features six type 1 and two type 2 repeats. The six amino-terminal type 1 repeats (I–VI) each consist of 81–99 amino acids and contain two cysteine residues that are in the right context to form globular domains as described for molecules with immunoglobulin structure. Within the proposed globular region, contactin shares 31% identical amino acids with the neural cell adhesion molecule NCAM. The two type 2 repeats (I–II) are each composed of 100 amino acids and lack cysteine residues. They are 20–31% identical to fibronectin type III repeats. Both the structural similarity of contactin to molecules of the immunoglobulin supergene family, in particular the amino acid sequence resemblance to NCAM, and its relationship to fibronectin indicate that contactin could be involved in some aspect of cellular adhesion. This suggestion is further strengthened by its localization in neuropil containing axon fascicles and synapses.

CELLULAR adhesion plays a crucial role in guiding development and maintaining the integrity of the nervous system. During their maturation, neurons come in contact with a number of different substrates and cell surfaces that guide their axons to the appropriate targets and determine selectivity of axon fascicles and synapses. In particular during late phases of neuronal development, localized areas of the neuronal cell surface form stable contacts with adjacent cells or the surrounding extracellular matrix. The rationale for our work is the hypothesis that in analogy to adherens junctions (Volk and Geiger, 1984; Geiger et al., 1980, 1983; Singer and Paradiso, 1981) and desmosomes (Kartenbeck et al., 1983), the formation and maintenance of stable contacts of neurons with other cells or substrates involves transmembrane proteins that are tightly linked with the neuronal cytoskeleton.

Components of the membrane skeleton, as the linkage of membrane molecules with the cytoskeleton is referred to (Marchesi, 1985), can be isolated by virtue of their insolubility in nonionic detergents (Sheetz, 1979). We have successfully used this biochemical characteristic to identify and characterize elements of the membrane skeleton in neurons (Moss, 1983; Ranscht et al., 1984). The neuronal membrane skeleton was isolated from embryonic chicken neurons as a detergent-insoluble complex consisting of several polypeptides including actin. The most prominent component of the complex is a glycoprotein of 130 kD, so far referred to as GP130, that can be separated from the other members of the complex under denaturing conditions. Antibodies generated against GP130 demonstrated the presence of this molecule exclusively on neuronal cell surfaces and suggested that its developmental expression may correlate with fasciculation or synapse formation. GP130 is most abundant in the adult chicken brain, where it comprises ~0.1% of the total protein. The late developmental appearance and the association with the cytoskeleton may place this molecule into the group of membrane components found in areas of stable cell–cell contacts.

In recent years, the molecular analysis of polypeptides involved in cellular adhesion has revealed common structural features and defined new gene families (Suzuki et al., 1986; Ruoslahti and Pierschbacher, 1987; Hynes, 1987; Williams, 1987; Shirayoshi et al., 1986). I have undertaken the molecular cloning of the 130-kD protein and describe here the sequence and structure as derived from the cloned cDNAs. The nucleotide sequence encodes a polypeptide of 1,071 amino acids that can be divided into a large extracellular domain, a single membrane-spanning segment, and an 89-amino acid cytoplasmic tail. The structure of the polypeptide

B. Ranscht's present address is La Jolla Cancer Research Foundation, Cancer Research Center, La Jolla, CA 92037.

suggests that this molecule is a new addition to the members of the immunoglobulin supergene family expressed in the nervous system. In combination with the immunoglobulin-like domains, the 130-kD protein contains segments similar to fibronectin type III repeats. Both structures are known to mediate cellular interactions in a variety of systems (Williams, 1987; Hynes, 1985; Mosher, 1984; Yamada, 1983). A role of the 130-kD glycoprotein in some aspect of cellular communication is suggested not only from its structure, but also from its localization in neuropil-containing axon fascicles and synapses. To acknowledge the structural relationship of the 130-kD protein to known cell adhesion molecules and its localization in areas of the nervous system where neurons are in contact with each other, I propose to name this molecule contactin.

Materials and Methods

Protein Isolation and Sequencing

The 130-kD glycoprotein was isolated from 16–18-d-old chicken embryo brains as previously described (Ranscht et al., 1984). For two-dimensional gel analysis, a membrane-skeleton complex was enriched for the 130-kD protein by ion exchange chromatography on DE52 and concentrated 100-fold. The proteins from this fraction were separated sequentially by isoelectric focusing (O'Farrell, 1975) and SDS-PAGE (Laemmli, 1970). The gel was blotted onto nitrocellulose paper (Towbin et al., 1979) and probed with rabbit anti-GP130 antibody (1:100) and ^{125}I-protein A (400,000 cpm/ml) as previously described (Ranscht et al., 1984). For amino acid sequence analysis, preparative polyacrylamide gels were prerun with 0.1 mM sodium thioglycolate in the cathode buffer (Hunkapiller et al., 1983). The polypeptide of 130 kD was excised as a single band from lightly Coomassie Blue-stained gels. This material was electroeluted and used for amino-terminal amino acid sequencing.

A 21-kD peptide was generated by digestion with *Staphylococcus aureus* V8 protease (Miles Laboratories Inc., Naperville, IL). The 130-kD protein was excised from a preparative SDS-PAGE gel as before and placed into the slot of a second preparative gel. Digestion was carried out in the gel slot in the presence of 1 μg/ml V8-protease for 45 min (Cleveland et al., 1977). The tryptic peptides were separated by SDS-PAGE (12% acrylamide). After light staining with Coomassie Blue, the 21-kD peptide was excised and electroeluted for amino acid sequence analysis. Sequencing of the polypeptides was performed by Drs. D. Andrews and W. Lane at the Harvard Microchemistry Facility (Cambridge, MA) using a gas phase microsequenator (Applied Biosystems, Inc., Foster City, CA).

Construction of cDNA Libraries

RNA was extracted from 13-d-old chicken embryo brains according to Cathala et al. (1983) and poly A$^+$ RNA was selected on oligo(dT)-cellulose (type 3; Collaborative Research, Inc., Waltham, MA). CDNA was synthesized according to Gubler and Hoffman (1983). First strand synthesis was with poly A$^+$ RNA primed with oligo(dT) using reverse transcriptase in the presence of actinomycin D and second strand used RNase H and *Escherichia coli* λDNA polymerase I. Internal Eco RI sites were methylated with Eco RI methylase and ends were made blunt with T4 DNA polymerase (Maniatis et al., 1982). After ligation with Eco RI linkers and digestion with Eco RI, large cDNA inserts were selected by filtration on Sephacryl-200. Size-selected cDNA was ligated into the arms of λgt.10 or the dephosphorylated Eco RI sites of λgt.11 and packaged into phage heads using λ packaging extract (prepared and donated by Dr. C. Kintner). The λgt.10 and gt.11 libraries contained 2×10^6 and 1×10^6 independent recombinants, respectively, with an average insert size of 2.2 kb.

Isolation of cDNA Clones

Nitrocellulose replica filters of the unamplified λgt.11 library were screened with monoclonal and polyclonal antibodies against GP130 (Ranscht et al., 1984) following the procedure of Young and Davis (1985). The anti-GP130 rabbit serum was adsorbed on *E. coli* bacterial lysate before use. Horseradish peroxidase conjugated secondary antibodies (Cappel Laboratories, Inc., Cochranville, PA) and the diaminobenzidine enzyme substrate were used to monitor binding of the antibodies (Graham and Karnovsky, 1966). Of $\sim 10^7$ independent recombinants screened, seven positive clones were selected and plaque purified. The antibody-selected clones were rescreened on nylon replica filters (Biodyne Electronics, Santa Monica, CA) with a mixed 23-mer oligonucleotide probe representing amino-terminal sequence of the 130-kD protein or with a mixed 17-mer oligonucleotide representing internal peptide sequence of GP130. Mixed oligonucleotides were synthesized by Dr. D. Rigby at the Harvard Microchemistry Facility. The oligonucleotide probes were labeled with γ-^{32}P-ATP (Amersham Corp., Arlington Heights, IL) using T4-Polynucleotide Kinase (Boehringer Mannheim Biochemicals, Indianapolis, IN) according to standard protocols (Maniatis et al., 1982). For hybridization the procedure of Wood et al. (1985) was adopted, in which an initial nonstringent hybridization with radiolabeled probe is followed by washes in 3.0 M tetramethyl ammonium chloride to control the stringency of the hybridization. In brief, replica filters were prehybridized at 37°C for 4–16 h in 6× SSPE (20× SSPE = 3 M NaCl, 0.2 M $NaH_2PO_4 \cdot H_2O$, 0.02 M EDTA · Na_2, pH 7.4), 5× Denhardt's, and 100 μg/ml heat-denatured salmon sperm DNA. Hybridization was under identical conditions for 12–16 h with the 23-mer or the 17-mer oligonucleotide probe at 0.5–1 × 10^6 cpm/ml. Filters were rinsed three times for 5 min with 6× SSPE at 4°C and then washed in 6× SSPE three times for 30 min at 4°C. The blots were equilibrated two times for 30 min each at 37°C with 3 M tetramethylammonium chloride, 50 mM Tris/HCl, pH 8.0, 2 mM EDTA, and 1% SDS. Washes were in identical buffer at 4°C below the DNA dissociation temperature for 23-mer or 17-mer oligonucleotides in 3.0 M Me_4NCl (Wood et al., 1985), i.e., at 62° and 52°C, respectively. Filters were exposed over night to Kodak XAR 5 film at −70°C.

Subcloning and Sequence Determination

λgt.11 DNAs containing selected cDNA inserts were purified from liter lysates on CsCl gradients (Maniatis et al., 1982). The longest cDNA clone 183 was chosen for nucleotide sequencing by the dideoxy chain termination method of Sanger et al. (1977). Eco RI/Kpn I restriction fragments were separated on agarose gels and isolated either by successive extractions with phenol and phenol/chloroform after melting (low melting temperature agarose [International Biotechnologies Inc., New Haven, CT]) or by electroelution (Seakam agarose; FMC Bioproducts, Rockland, ME). Subcloning was into the polylinker of the transcription vector SP 72 (Melton et al., 1984). Templates for directed sequencing in M13mp18 and M13mp19 (Messing, 1983) were generated by Bal31 exonuclease digestion (Poncz et al., 1982) and restriction site cloning (Maniatis et al., 1982). Synthetic oligonucleotide primers (synthesized by J. Knight, DNA Facility, La Jolla Cancer Research Foundation) were used where sequences from different templates did not overlap, were ambiguous or had been obtained only in one direction. For partial sequencing of cDNA clones 17, 181 and 184, templates were generated from cDNA restriction fragments that were isolated from the insert containing λgt.11 clones.

RNA Analysis

Total cellular RNA was isolated according to the guanidinium isothiocyanate method of Chirgwin et al. (1979). For some experiments, total RNA was separated into poly A$^+$ and poly A$^-$ RNA on oligo(dT) cellulose (Maniatis et al., 1982). RNA was separated on 1% agarose-formaldehyde gels and transferred to nylon membranes (GeneScreen Plus; New England Nuclear, Cambridge, MA) in 20× SSPE. Membranes were cross-linked with UV light and baked for 60 min under vacuum at 80°C. Prehybridization was for 2–4 h in 50% deionized formamide, 5× SSPE, 1% SDS, and 100 μg/ml heat-denatured salmon sperm DNA at 46°C. For hybridization, restriction fragments of the cDNAs were purified from agarose gels as above, labeled by nick translation, and hybridized to the membranes at 0.5–1.5 × 10^7 cpm/ml under prehybridization conditions. Washes were in several changes of 0.2× SSPE, 0.2% SDS at 68°C. RNA sizes were determined by comparison to RNA standards (Bethesda Research Laboratories, Gaithersberg, MD). Chicken β-actin cDNA (a donation of Dr. D. Cleveland) was used as a control.

Immunohistochemistry

Chickens were perfused with 100 mM Na-periodate, 75 mM lysine, and 3% paraformaldehyde. Nerve tissue was removed and immersion fixed in the above fixative for an additional 2 h. The tissue was kryoprotected by overnight incubation in 1.5 M sucrose in PBS, pH 7.2, embedded in Tissue-Tek (Miles Laboratories, Inc.) and frozen at −70°C. Cryostat sections were cut at 15 μm thickness and quenched with 20 mM lysine in PBS. Sections were

stained overnight with rabbit anti-GP130 antibody (1:100) or monoclonal antibody GP130-5 (1:400) (Ranscht et al., 1984). Binding of the monoclonal antibody was monitored after sequential incubation with affinity-purified rabbit anti-mouse Ig (1:200) and rhodamine-conjugated goat anti-rabbit Ig (1:150; Cappel Laboratories Inc.); binding of rabbit anti-GP130 antibodies was monitored by staining with the rhodamine conjugate alone. All antibody dilutions were in PBS containing 10% normal goat serum and 0.02% Triton-X100. The sections were washed after each incubation step by immersion in five changes of PBS and mounted with immunomount (Shandon Southern Instruments, Inc., Sewickley, PA) for visualization.

Results

Characterization of Antibodies and Protein Sequencing

Antibodies and oligonucleotide probes synthesized on the basis of partial protein sequences were used to select cDNA clones. Antibodies were generated against the 130-kD polypeptide excised from one-dimensional preparative SDS-PAGE gels (Ranscht et al., 1984). The specificity of these antibodies was examined by immunoblotting of a crude 130-kD preparation. A membrane-skeleton complex was enriched for the 130-kD glycoprotein by ion exchange chromatography (Ranscht et al., 1984) and separated in two dimensions by isoelectric focusing and SDS-PAGE. Staining with Coomassie Blue revealed only one protein spot of 130-kD which had a pI at pH 6.0. A similar gel was transferred to nitrocellulose and then reacted with the rabbit anti-GP130 antibody and ^{125}I-protein A. The immunoblot analysis detected only one polypeptide at 130 kD (Fig. 1 *left*) indicating that the antiserum was monospecific. Since the antiserum was prepared against the 130-kD protein excised after one-dimensional SDS-PAGE, this result also indicated lack of significant levels of protein contaminants in this preparation. A single amino acid sequence was obtained when the 130-kD protein was excised and electroeluted from one-dimensional preparative SDS-PAGE gels and subjected to amino acid sequencing (Fig. 1 *right*). Peptides were generated by mild digestion of the 130 kD protein with V8-protease (Cleveland et al., 1977) in order to have available amino acid sequence information of an internal part of the molecule. After separation by SDS-PAGE, one peptide of 21 kD was chosen for amino acid sequencing (Fig. 1 *right*) since it could be well separated from other peptides on 12% acrylamide gels.

Isolation and Characterization of cDNA Clones

An unamplified chick brain λgt.11 library was screened with the rabbit anti-GP130, or with a mixture of six monoclonal anti-GP130 antibodies (Ranscht et al., 1984). Seven clones were isolated that gave above background signals with the rabbit antibodies; three of these clones also bound the monoclonal antibodies. All seven clones were plaque purified and further characterized.

To investigate the identity of the isolated cDNAs, the antibody-selected clones were rescreened with mixed oligonucleotide probes. The probes consisted of a 23-mer representing amino acids 5–9 of the amino-terminal sequence and a 17-mer representing amino acids 4–9 of the 21-kD peptide (bars above sequence in Fig. 1 *right*). Four clones (17, 181, 183, and 184) screened positive with the 23-mer oligonucleotide probe. The hybridization signal was mapped by Southern analysis to a 200-bp Eco RI/Xho I restriction fragment at the 5′-ends of the four cDNAs. No hybridization signal, however, was obtained with the shorter 17-mer oligonucleotide. Digestion with restriction enzymes and Southern blotting showed that the four cDNA clones hybridizing to the 23-mer oligonucleotide were related. The restriction maps of 17, 181, 183, and 184 cDNAs are shown in Fig. 2. These four cDNAs share restriction sites over most of their internal sequence and crosshybridize to each other. They were found to differ in the size of their 3′ and 5′ ends. Subsequent nucleotide sequencing revealed that the 3′-end

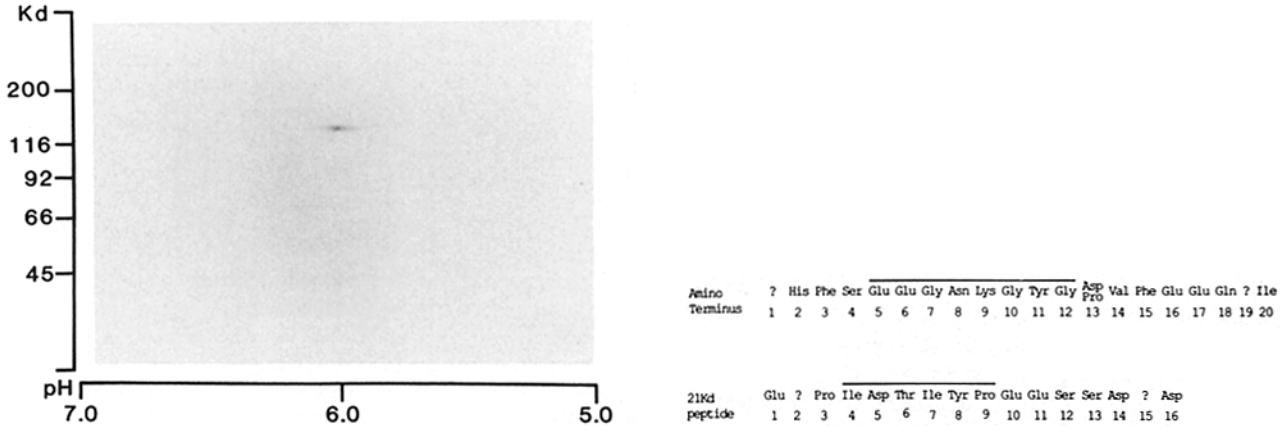


Figure 1. (*Left*) Characterization of antibodies and (*right*) protein sequencing. (*Left*) The 130-kD–containing membrane skeleton complex was resolved by IEF (O'Farrell, 1975) and SDS-PAGE (Laemmli, 1970). The gel was transferred to nitrocellulose and the blot probed with anti-GP130 antiserum and ^{125}I-protein A. A single polypeptide with a pI at pH 6.0 is stained in the 130-kD molecular mass region indicating that the antiserum is monospecific. The pH gradient is indicated on the horizontal axis, the migration of molecular mass standards is shown on the vertical axis. (*Right*) Amino acid sequence of the amino terminus of the 130-kD protein and a 21-kD protein generated by tryptic cleavage of the 130-kD protein with V8-protease (1 μg/ml). Polypeptides were separated by SDS-PAGE, stained lightly with Coomassie Blue, and then excised and electroeluted. Amino acid sequencing was performed at the Harvard Microchemistry Facility by Drs. W. Lane and D. Andrews. Bars above the sequences indicate regions according to which mixed oligonucleotides were synthesized.

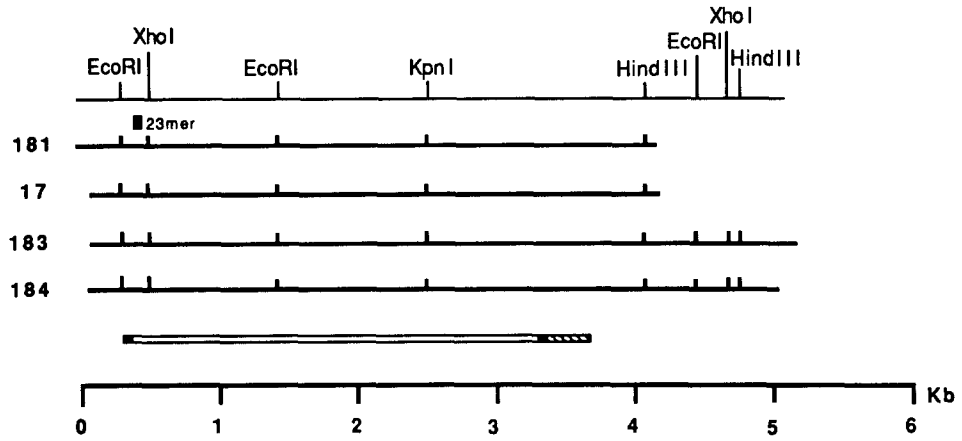


Figure 2. Restriction maps of cDNAs. Four cDNA clones, 17, 181, 183, and 184 were positively identified with antibodies and oligonucleotide probes to represent cDNA encoding the 130-kD protein. Hybridization of the oligonucleotide probe representing the amino-terminal amino acid sequence of the 130-kD glycoprotein is marked with a bar labeled 23-mer. The four cDNA clones are 4.5–5.0 kb long and share restriction sites over their internal sequence. Differences occur in the length of 3′ and 5′ ends. All four cDNA clones have identical 3′-end nucleotide sequences; clones 17, 181, and 184, however, lack part of the 3′-untranslated region present in 183 cDNA. The location of the protein-coding region is indicated below the four cDNAs. Hydrophobic regions are in black, the extracellular portion is in white, and the cytoplasmic region is shaded.

sequences of the shorter cDNAs (not shown) are identical to the longest cDNA 183 (Fig. 4 *b*) but lack part of the 3′-untranslated region. Nucleotide sequences of the cDNA clones differ, however, within 0.1–0.3 kb Eco RI restriction fragments that mapped by Southern analysis to the extreme 5′ ends of the cDNAs. The differences at the 5′ ends could have arisen as an artifact during library construction, or indeed represent genuine differences in transcripts of the 130-kD protein.

Northern Analysis Detects Brain-specific Transcript

Since the expression of the 130-kD protein is restricted to nervous tissue (Ranscht et al., 1984), a genuine RNA transcript would be expected to be distributed accordingly. Northern analysis of RNA from different chicken tissues with the nick-translated cDNA probes detected a major transcript of approximately 5.5 kb in brain and retina (Fig. 3). No hybridization signal was obtained with RNA from sciatic nerve, that contains mainly Schwann cells and fibroblasts, or with RNA from the nonneural tissues heart, liver, lung, kidney, skin, and muscle. The blots were rehybridized with chicken β actin (Cleveland et al., 1980; Kost et al., 1983) to demonstrate that intact RNA was present in each lane. The detection of a large RNA restricted to nervous tissue therefore met the requirements for a genuine transcript of the 130-kD protein.

When the Northern analysis was performed with probes of higher specific activity, an additional brain-specific RNA of ~3.5 kb was detected. The 3.5-kb RNA was substantially less abundant than the major 5.5-kb RNA in both total and poly A^+ RNA from nervous tissue. The presence of a second transcript raised the possibility that different forms of the 130-kD protein exist and arise by alternatively spliced RNA. Hybridization with cDNA probes from the different clones and their restriction fragments, however, did not distinguish between the two transcripts (data not shown). If different molecule forms exist, they are therefore not represented by the isolated cDNAs. It is thus not known, if the smaller transcript corresponds to a different RNA transcript encoding the 130-kD protein or to a transcript from a different, but related gene.

Sequence of the cDNA Clones

The longest cDNA, clone 183, was chosen for sequence analysis and a partial nucleotide sequence of clones 17, 181, and 184 (thick arrows in Fig. 4 *a*) was obtained from templates generated by digestion with restriction enzymes of the cDNA inserts. The nucleotide sequence of 183 cDNA is shown in Fig. 4 *b*.[1] The sequence of the additional three cDNA clones completely matched the sequence of 183 cDNA with the exception of the 5′-Eco RI restriction fragments as discussed above.

In the sequence, the ATG encoding the methionine in the −20 position is flanked by nucleotides resembling the consensus sequence CCACCAUG(G) for eukaryotic initiation sites (Kozak, 1984). The putative AUG initiation codon is followed by a 3,273-nucleotide open reading frame potentially encoding for a 1,091-amino acid polypeptide. The open reading frame ends with a series of three termination codons (Fig. 4 *b*, *boxed*). The untranslated 3′ region extends over a total of 1,676 nucleotides. Even though two potential poly(A) acceptor sites are found (Fig. 4 *b*, *dashed boxes*), the sequence does not reveal poly(A) sequences. This result indicates that parts of the 3′-end sequences present in the messenger RNA are missing in the cDNA clones.

cDNAs Encode a Transmembrane Protein with Immunoglobulin-like Domains and Segments Resembling Fibronectin Type III Repeats

Conceptual translation of the nucleotide sequence predicts a polypeptide of 1,091 amino acids. The amino-terminal region contains a typical signal sequence cleavage site (Von Heijne, 1982). Cleavage of the signal peptide from this site leaves a threonine residue as the NH_2-terminal amino acid, in agreement with results obtained from polypeptide sequencing of the 130-kD protein (Fig. 1 *right* and underlined in Fig. 4 *b* with short dashes). The cDNA thus encodes for a hydrophobic 20 amino acid signal peptide (residues −20 to −1; Fig. 4 *b*, *underlined*), which is not present in the processed protein. The amino acid sequence of the 21-kD peptide ob-

1. These sequence data have been submitted to the EMBL/GenBank Data Libraries under the accession number 400813.

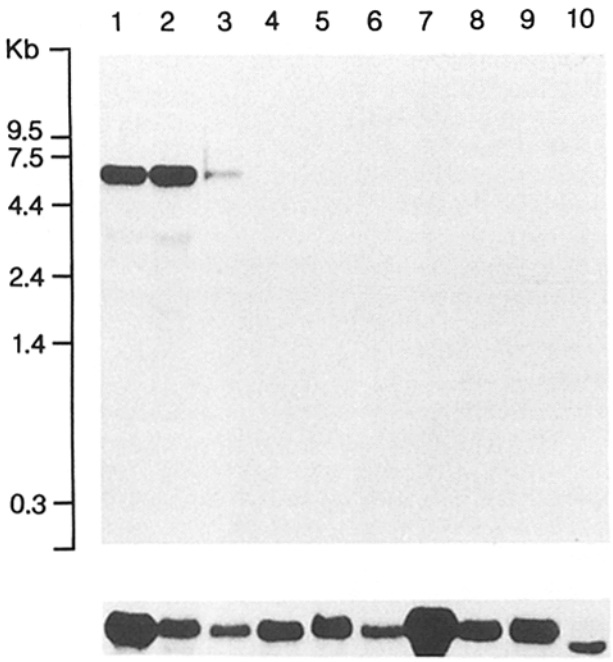


Figure 3. Northern analysis. Total cellular RNA was isolated from 1-2-d-old chicks (lanes *2–10*), and poly A^+ RNA was isolated from brains of 13-d-old chicken embryos (lane *1*). RNA was separated on 1% agarose-formaldehyde gels and blotted to nylon membranes. The blot was probed with nick-translated 183 cDNA representing 3.3-kb coding and 1.0-kb noncoding region. A major nerve tissue-specific transcript of 5.5–6.0 kb is detected (lanes *1–3*). The identity of the smaller and less abundant transcripts is not known. The blot was reprobed with nick-translated β actin to demonstrate that intact RNA was present in each lane (lanes below figure). (*1*) Poly A^+ RNA from chicken brain (embryonic day 13); (*2*) optic lobe; (*3*) retina; (*4*) sciatic nerve; (*5*) skin; (*6*) liver; (*7*) lung; (*8*) kidney; (*9*) heart; (*10*) muscle. Samples *2–10* represent total RNA.

tained by mild tryptic digestion with V8 protease was also confirmed by the nucleotide sequence (Fig. 4 *b*, underlined with long dashes), with the exception of residue 16 of the peptide which was identified as Gln instead of Asp (Fig. 1 *right*).

The mature protein contains a total of 1,071 amino acids with a predicted molecular mass of 119,956 D. The sequence indicates nine potential asparagine-linked glycosylation sites (NXT/S) which are marked by solid triangles in Fig. 4 *b*. Attachment of carbohydrates at these sites thus could explain the molecular mass of 130 kD determined by SDS-PAGE. The hydrophobicity profile of the translated cDNA sequence (performed according to Kyte and Doolittle, 1982; not shown) indicates only one hydrophobic region (amino acids 963–983; Fig. 4 *b*, *underlined*) large enough to span the membrane. Amino acids flanking this region meet the criteria for a membrane-spanning domain (Sabatini et al., 1982). Therefore, the deduced amino acid sequence of the cDNAs predicts a transmembrane protein that traverses the membrane once and extends 962 amino acids into the extracellular space and 89 amino acids into the cytoplasm.

The extracellular part of the deduced polypeptide chain can be subdivided into two domains: the cysteine-containing domain (amino acids 1–561) and the cysteine-free domain (amino acids 562–962). The two domains are separated by a collagen-like amino acid stretch (underlined with a thin line in Fig. 4 *b*) that could serve as a hinge region to provide structural flexibility to the extracellular part of the molecule (Bornstein and Traub, 1979; Miller and Gay, 1987).

The most striking feature of the polypeptide sequence are two types of repeating units in the extracellular portion. Six repeats of type 1 (I–VI) are located in the cysteine-containing amino-terminal half and two type 2 repeats (I and II) are in the cysteine-free carboxy-terminal half of the extracellular domain. Type 1 and type 2 repeats are unrelated to each other.

Type 1 repeats I–VI each contain 81–99 amino acids, which are shown in their best alignment in Fig. 5 *A*. 20–30% of the amino acids within the repeats are identical with the exception of only 14% identity between repeats II and IV. Each repeat contains two cysteine residues spaced 38–56 amino acids apart, of which each second one is flanked by highly conserved amino acids (DXGXYXCXV/A). The cysteine residues and the amino acids flanking them align by visual inspection and using the align function of the Microgenie program (Beckman Instruments, Inc., Palo Alto, CA) with the corresponding residues of immunoglobulins and molecules with immunoglobulin-like structure (Saul et al., 1978; Amzel and Poljak, 1979; Williams and Gagnon, 1982; Cunningham et al., 1987; Salzer et al., 1987; Zimmermann et al., 1987; Yarden et al., 1986; Lewis et al., 1986; Williams, 1987). The cysteine residues of immunoglobulins and immunoglobulin-like molecules are suggested to form globular domains through intrachain disulfide bonding. This is also likely to be the case for this molecule since the 130-kD protein migrates as a single band on SDS-PAGE gels under both reducing and nonreducing conditions, whereby its mobility is increased when reducing agents are omitted (Moss, 1986).

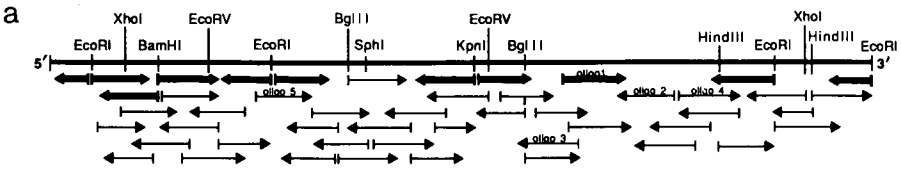
The two type 2 repeats I and II consist each of 100 amino acids that exclude cysteine residues. Their best alignment is shown in Fig. 5 *b*; the similarity between the two repeats is 36%. Amino acids of type 2 repeats align with the characteristic tryptophan and tyrosine residues of fibronectin (FN)[2] type III segments. The overall similarity of contactin type 2 repeats to FN type III blocks is between 20 and 31%, whereby amino acids of FN type III repeats are more conserved in contactin type 2 repeat I as compared with repeat II. In Fig. 6 *b*, the most related regions of contactin type 2 repeat I and bovine FN type III sequences (Peterson et al., 1983) are shown.

The putative cytoplasmic domain of the deduced polypeptide consists of 89 amino acids that could provide a link with the cytoskeleton. There is no apparent similarity to the cytoplasmic fragments of other cytoskeleton-associated membrane proteins such as the receptor for FN (Tamkun et al., 1986; Argraves et al., 1987), or components of the erythrocyte membrane skeleton band 3 (Kopito and Lodish, 1985) and glycophorin (Tomita et al., 1978).

Sequence Similarities of Contactin with Known Proteins

Search of the National Biomedical Research Foundation (Washington, DC) and the Genetic Sequence Data Bank (GenBank, Mountain View, CA) as compiled in the Micro-

2. *Abbreviations used in this paper*: FN, fibronectin; MAG, myelin-associated glycoprotein.



b

Figure 4. (*a*) Sequencing strategy. Indicated restriction sites and deletion with Bal31 exonuclease of 183 cDNA were used to generate sequencing templates (*arrows*). Thick arrows represent templates of cDNAs 17, 181, and 184. (*b*) Nucleotide sequence and deduced amino acid sequence of the 130-kD protein. The longest cDNA 183 contains 4,998 nucleotides with an open reading frame for 1,091 amino acids. Termination codons are boxed. The sequence contains two potential poly(A)-acceptor sites (*dashed boxes*) but lacks polyadenylation. Hydrophobic regions representing the signal peptide (−20 to −1) and the transmembrane region (963–982) are underlined (*thick bar*). The cleavage point of the signal peptide is marked with an arrow. Sequences confirming amino acid sequencing are underlined with dashed lines (*short dashes*, amino-terminal amino acids; *long dashes*, 21-kD peptide sequence). The extracellular part of the polypeptide is divided into a cysteine-containing region (1–561, cysteines are *circled*) and a cysteine-free region (562–962). The two domains are separated by a collagen-like amino acid stretch (amino acids 573–580, *thin line*). Potential sites for asparagine-linked glycosylation are marked with solid triangles. A potential nucleotide-binding site is marked with an asterisk.

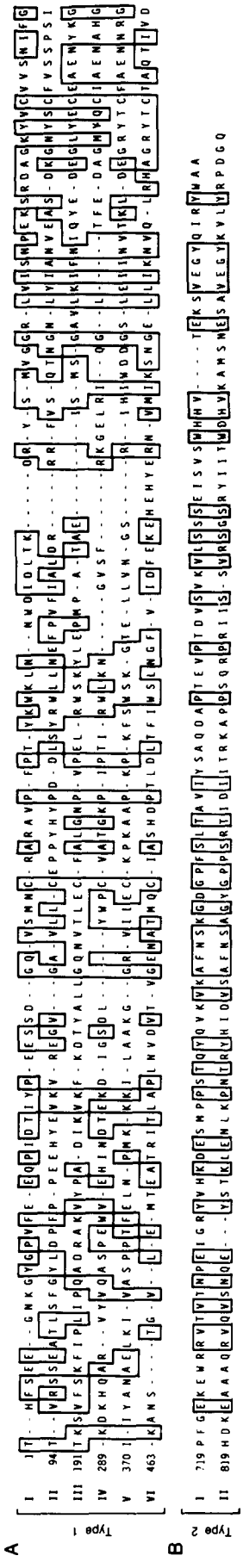


Figure 5. Alignment of internal repeats. (*A*) The six type 1 repeats (I–VI) of 81–99 amino acids each contain two cysteine residues which are flanked by amino acids conserved in molecules with immunoglobulin structure. (*B*) The two type 2 repeats located in the cysteine-free extracellular region resemble fibronectin type III repeats.

genie sequence analysis program revealed no significant homologies of the deduced protein sequence with other proteins. Sequences of cell surface proteins that have recently been published, were placed into the laboratory data base and compared with the GP130 protein sequence using the Microgenie sequence analysis program (Queen and Korn, 1984). This comparison included members of the cadherin and immunoglobulin supergene families (E-cadherin: Nagafuchi et al., 1987; N-cadherin: Hatta et al., 1988; L-CAM: Gallin et al., 1987; N-CAM: Cunningham et al., 1987; Hemperley et al., 1986; Barthels et al., 1987; myelin-associated glycoprotein (MAG): Salzer et al., 1987; I-CAM: Simmons et al., 1988; Staunton et al., 1988; carcinoembryonic antigen: Zimmermann et al., 1987; Thompson et al., 1987; Oikawa et al., 1987). The analysis detected a significant similarity of the 130-kd protein with the neural cell adhesion molecule NCAM (Cunningham et al., 1987). The GP130 amino acids 70–335 and 356–450 share a 31% identity with NCAM residues 62–322 and 360–452, respectively (Fig. 6 *a*). GP130 appears to be less similar to MAG (Salzer et al., 1987; Sutcliffe et al., 1983), since the Microgenie sequence analysis program revealed only short stretches of amino acid identity. In addition, alignment of the type 1 repeats I–VI with each of the five individual immunoglobulin-like domains of NCAM and MAG using the Dayhoff align program (Dayhoff et al., 1983) revealed higher align scores for the 130-kD glycoprotein with NCAM than with MAG (not shown).

Less striking than the sequence similarities of the 130-kD protein with NCAM were similarities with other members of the immunoglobulin supergene family. These include V- and C-regions of immunoglobulins (Saul et al., 1978), T cell surface receptors CD3 and CD8 (Gold et al., 1986; Johnson and Williams, 1986), Fc-receptor (Lewis et al., 1986), carcinoembryonic antigen (Oikawa et al., 1987), α_1B-glycoprotein (Ishioka et al., 1986), peripheral myelin protein P0 (Lemke and Axel, 1985), link-protein (Neame et al., 1986; Deák et al., 1986), platelet-derived growth factor receptor (Yarden et al., 1986), and the kinase-related transforming protein v-fms (Hampe et al., 1984) whose cellular counterpart has been proposed to act as a receptor for macrophage CSF-1 (Sherr et al., 1985; Coussens et al., 1986). The similarity of the 130-kD glycoprotein with v-fms was detected as the most significant match to polypeptides in the Doolittle data base (Doolittle, R., personal communication). The matches of 130-kD sequences to the above polypeptides always include amino acids surrounding one or several of the extracellular cysteines and thus are likely to define part of the immunoglobulin-like structure. The structural resemblance of the 130-kD protein to molecules of the immunoglobulin supergene family and the similarity of its globular domains with NCAM define this molecule as a member of the immunoglobulin supergene family in the nervous system and suggest that it could have some role in cellular communication.

A second group of molecules with similarity to the 130-kD protein are oncogenes and growth factor receptors including the kinase-related transforming proteins src (Takeya and Hanafusa, 1983), yes (Kitamura et al., 1982), raf-1 (Mark and Rapp, 1984), mos (Van Beveren et al., 1981), and insulin growth factor receptor (Ullrich et al., 1985), as well as K-ras (Barbacid, 1987). These sequence similarities were detected in a search of the GenPro data bank (release 52.0; Riverside Scientific Enterprises, Seattle, WA) with the search parame-

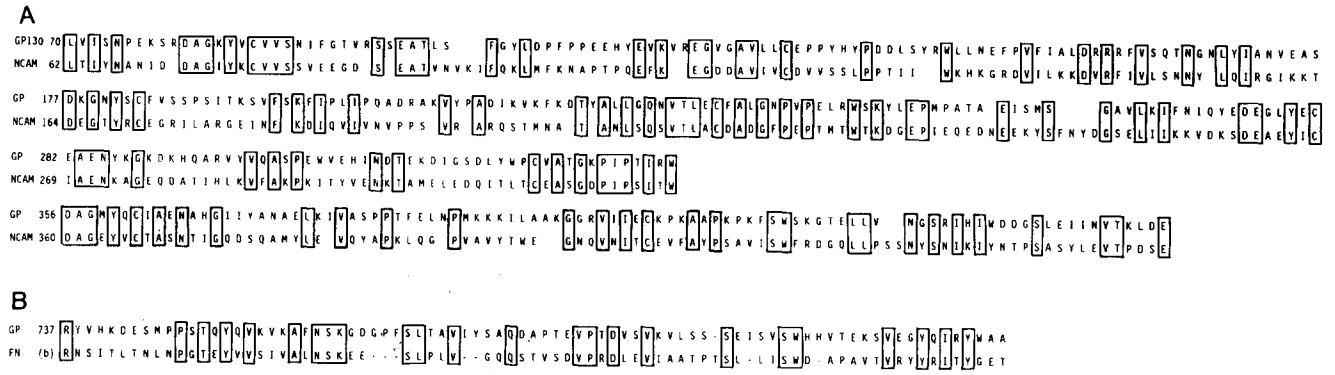


Figure 6. Similarities of contactin with (*A*) NCAM and (*B*) FN. Regions of 31% identity between contactin and NCAM (Cunningham et al., 1987) are shown. Similarities are found between amino acids 70–335 of contactin and 62–322 of NCAM (lines 1–3) and between 356–450 of contactin with 360–452 of NCAM (line 4). (*B*) Regions of similarity between contactin type 2 repeat I and bovine fibronectin type III sequences (Peterson et al., 1983) are shown. 29% of the amino acids are conserved.

ter 12 matches out of 40 and cover extracellular sequence stretches of the 130-kD protein that lack cysteine residues. Many of the oncogenes, growth factor receptors, and tissue-specific kinases contain a characteristic nucleotide-binding lysine (GXGXFG, 16–28 residues amino terminal of VAVK) that has been suggested to mediate kinase activity (Hunter and Cooper, 1985; Van Beveren and Verma, 1986; Ullrich et al., 1984, 1985; Yarden et al., 1986; Shoji et al., 1981; Bennett and Kennedy, 1987). In the polypeptide sequence of the 130-kD protein, lysine 789 (marked with an asterisk in Fig. 4 *b*) was noted to follow closely the above motif. In addition, 10 of the serine and threonine residues within the cysteine-free region are in the right context of amino acids (K/RXXT/S) to serve as potential autophosphorylation sites (Pearson et al., 1985). The significance of this sequence coincidence is not clear, since thus far no molecules with extracellular kinase activity have been reported.

The 130-kD Protein Is Concentrated in Areas of Interneuronal Contact

In view of the suggestion that the 130-kD glycoprotein is related to cell surface molecules that are involved in cell–cell interaction, it is important to know where this molecule is localized in the nervous system. To gain a first understanding about the cells or substrates, the 130-kD protein may be interacting with in vivo, frozen sections of the adult chicken retina were stained in indirect immunofluorescence with anti-GP130 antibodies. The adult retina was chosen for two reasons: (*a*) since 130-kD protein is most abundant in the adult (Ranscht et al., 1984), adult tissue should display best where this molecule accumulates to interact with adjacent cells or substrates; and (*b*) the retina is organized in a way that allows the distinction between neuronal subcompartments.

The localization of the 130-kD protein in the adult chicken retina by staining with a monoclonal antibody is shown in Fig. 7. An identical staining pattern was obtained with the rabbit anti-GP130 antibodies. The immunoreactivity showed a striking concentration in the synaptic inner and outer plexiform layers and the optic fibers that connect the retina with the tectum. Staining of the inner plexiform layer was not uniform, but showed a distinct stratified pattern. Little, if any staining was seen on the neuronal cell bodies. This result provided a first indication that the 130-kD glycoprotein could be associated with axon fascicles and synapses. To further strengthen this observation, sections of the developing spinal cord and the cerebellum were stained with anti-GP130-antibodies. In both systems, a concentration of the 130-kD protein in the fiber-rich areas was observed (not shown).

Discussion

In neurons the linkage of membrane components with the cytoskeleton could play a crucial role in establishing cellular subcompartments by concentrating cell surface receptors to regions of axons and dendrites where they mediate selective fasciculation and participate in the formation and maintenance of synapses. We have isolated a neuronal cell surface molecule of 130 kD, thus far referred to as GP130, by virtue of its association with cytoskeletal elements (Ranscht et al., 1984). The current study defines the primary amino acid sequence of this molecule by molecular cloning and nucleotide sequencing of corresponding cDNAs and suggests that GP130 is a new member of the immunoglobulin supergene family in the nervous system. Molecules with immunoglobulin-like structure are known to mediate various aspects of cellular adhesion or intercellular communication (Williams et al., 1987). A role of the 130-kD protein in cellular communication is suggested by both its structure and its immunohistochemical localization in neuronal cell surface areas that are in contact with adjacent neuronal cells. The association of this molecule with nerve fibers is not only observed in the developing and adult retina (Fig. 7), but also in the developing spinal cord and in the cerebellum (unpublished observations). While the name "GP130" was satisfactory for the initial characterization of this molecule, it now seems rather nondescriptive. To indicate its structural relationship to molecules that are involved in cellular communication and its localisation in areas of cellular contact, I propose the name contactin.

Four cDNAs were isolated with specific antibodies and oligonucleotide probes representing amino-terminal amino acid sequence of the 130-kD glycoprotein. The four cDNAs, 17, 181, 183, and 184 crosshybridize to each other and detect

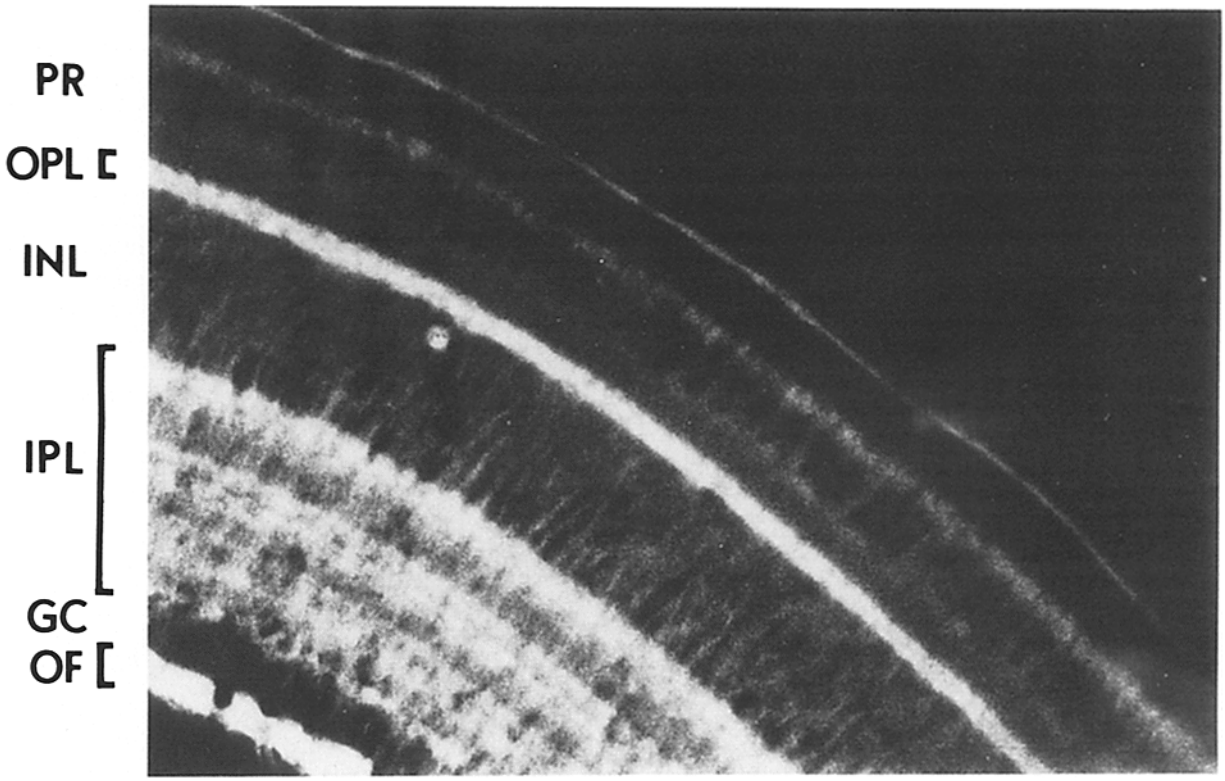


Figure 7. Immunohistochemical localization of the 130-kD protein in the adult chicken retina. Immunoreactivity is concentrated in the synaptic inner and outer plexiform layers and the optic fibers. Little, if any, staining is seen on the neuronal cell bodies. *PR*, photoreceptors; *OPL*, outer plexiform layer; *INL*, inner nuclear layer; *IPL*, inner plexiform layer; *GC*, ganglional bodies; *OF*, optic fibers.

a major brain specific transcript of ~5.5 kb. Conceptual translation of the nucleotide sequence of the cDNAs predicts a 1,071–amino acid polypeptide that traverses the membrane once. The majority of the protein (90%) extends into the extracellular space, leaving only 89 amino acids in the cytoplasm. Experimentally this model is supported by the tryptic cleavage of a 120-kd polypeptide fragment from the surface of cultured neurons (Moss, 1986). The overall structure of contactin is that of a hybrid molecule containing six immunoglobulin-like domains and two segments that resemble fibronectin type III repeats (Fig. 8). The evidence for the inclusion of contactin in the immunoglobulin supergene family is as follows. (*a*) The polypeptide sequence contains repeating units of 81–99 amino acids that each contain two cysteine residues spaced 38–56 amino acids apart. (*b*) Amino acids flanking the cysteines are characteristic for molecules with immunoglobulin structure. For example, each second cysteine of the repeats is preceded by the amino acids DXGXYXC (Williams, 1987). (*c*) The cysteines of contactin and amino acids typically conserved in molecules with immunoglobulin structure align with the corresponding residues of other members of the family. Amongst those are the neural cell adhesion molecule NCAM (Cunningham et al., 1987), MAG (Salzer et al., 1987; Lai et al., 1987), platelet-derived growth factor receptor (Yarden et al., 1986), v-fms (Coussens et al., 1986), Fc-receptor (Lewis et al., 1986), immunoglobulin v- and c- regions (Saul et al., 1978), and others (not shown). The two cysteines within each repeat of these molecules have been implied to form globular domains through intrachain disulfide bonds. The amino acid sequence reported here suggests that contactin contains six globular domains of variable size (38 and 56 amino acids) through in-

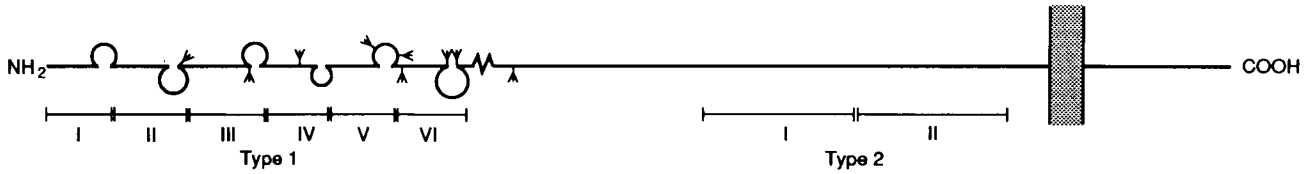


Figure 8. Proposed structure of contactin. The polypeptide contains a single membrane-spanning segment and extends 962 amino acids into the extracellular space and 89 into the cytoplasm. The extracellular domain contains six cysteine-containing repeats of type 1 that resemble those of molecules with immunoglobulin structure. The cysteines are likely to be linked through intrachain disulfide bonds. The two type 2 repeats lack cysteines and resemble the type III repeats of fibronectin. The model is drawn to scale. The plasma membrane is indicated by shaded segments. Carbohydrate attachment sites are marked with ψ and a collagen-like region is indicated with ∧∧.

trachain disulfide bonding of the cysteines (Fig. 8). It is known that the 130-kD glycoprotein does not form interchain disulfide bonds, since the molecule migrates as a single band on SDS gels under both reducing and nonreducing conditions (Moss, 1986).

Contactin is unusual with respect to the size of its extracellular domain. Between the putative globular and transmembrane domains, the protein contains an additional 400 amino acids. This region accommodates two type 2 repeats of 100 amino acids each. Interestingly, 20–31% of the amino acids within these repeats align with fibronectin type III segments, that include the cell-, heparin-, and DNA-binding regions of fibronectin (Hynes, 1985). The arg-gly-asp cell-binding sequence (Pierschbacher and Ruoslahti, 1984*a*; Yamada and Kennedy, 1984) is not contained in contactin. Instead, a conservative amino acid substitution leaves a lys-gly-asp sequence (amino acids 761–763) in contactin. Pierschbacher and Ruoslahti (1984*b*) reported this sequence to be inactive in cell adhesion. This may also be the case for this sequence in contactin. An alternative hypothesis is that neurons use a different, but related cell binding sequence that requires a yet unknown receptor.

In type 2 repeat I, lysine 789 was noted to describe a potential nucleotide-binding site similar to those that mediate kinase activity of oncogenes, growth factor receptors, and tissue specific kinases (Hunter and Cooper, 1985; Van Beveren and Verma, 1986; Ullrich et al., 1984, 1985; Yarden et al., 1986; Shoji et al., 1981; Bennett and Kennedy, 1987). In addition, 10 potential autophosphorylation sites K/RXXT/S (Pearson et al., 1985) were identified within the type 2 repeats. Since no examples of molecules with extracellular kinase activity are known, it is not clear if the nucleotide binding and the autophosphorylation sequence motifs are of any functional relevance.

Contactin shares significant similarity (31%) with the neural cell adhesion molecule NCAM (Cunningham et al., 1987), a molecule that mediates cell–cell interactions in the nervous system via homotypic binding (Edelman, 1987). The region of similarity is restricted to the globular regions and includes the homotypic cell-binding region of NCAM, which is located within the four amino-terminal globular domains (Cunningham et al., 1983, 1987). Since, however, a cell-binding region of NCAM has not been closely mapped, it is not clear whether some of the amino acids conserved in the two molecules indicate similar binding capacities. Despite the similarities of contactin and NCAM in their globular domains, amino acid sequences carboxy terminal to these regions are distinct. The structural resemblance of contactin to members of the immunoglobulin supergene family and in particular the significant conservation of amino acid residues within the globular regions of both contactin and NCAM, support the hypothesis that contactin may be involved in cellular communication in the nervous system.

In line with a prospective role of contactin in cell–cell interaction is its localization in the fiber-rich areas of the retina. Even though the pattern of contactin expression in the adult retina is distinct from that of Ng-CAM/NILE (Daniloff et al., 1986; Beasley and Stallcup, 1987), contactin staining may overlap during restricted phases of development with that of NgCAM/L1/G4/NILE/8D9 (Daniloff et al., 1986; Stallcup et al., 1985; Beasley and Stallcup, 1987; Lemmon and McLoon, 1986) neurofascin (Rathjen et al., 1987*b*), or F11 (Rathjen et al., 1987*a*) (our unpublished observations). Several lines of evidence, however, suggest that contactin is molecularly distinct from the above molecules. (*a*) Contactin is isolated as a single molecular species of 130kd that is recognized by monospecific antibodies (Fig. 1 *left* and Ranscht et al., 1984). In contrast, the 130–140kd components of NgCAM/L1/G4/NILE/8D9, F11, and neurofascin constitute fragments of molecules with higher molecular weights (170–230 kD) (Edelman, 1987; Rathjen and Schachner, 1984; Rathjen et al., 1987*a*, *b*). (*b*) Contactin is highly insoluble in non-ionic detergents. An isolation procedure based on affinity chromatography as described for L1/G4, F11, and neurofascin (Rathjen and Schachner, 1984; Rathjen et al., 1987*a*, *b*) cannot be applied to isolate contactin. (*c*) The restriction map of a cDNA encoding L1 (Tacke et al., 1987) is distinct from that of cDNAs encoding contactin (Figs. 2 and 4 *a*). (*d*) The published amino-terminal sequence for L1/G4 (Rathjen et al., 1987*a*) is distinct from the contactin sequence (Figs. 1 *right* and 4 *b*). Despite these molecular differences, contactin may be related structurally and functionally to these known axonal cell surface glycoproteins.

There is now increasing evidence that several cell surface glycoproteins are simultaneously expressed by the same axons. For example, the retinal ganglion cell axons expose on their cell surface NgCAM/L1/G4/NILE/8D9, neurofascin (Rathjen et al., 1987*b*), F11 (Rathjen et al., 1987*a*), N-CalCAM (Cook et al., 1984), and contactin (Fig. 7). Others may yet be discovered. One hypothesis for selective fasciculation is that the expression level of the individual cell surface components is modulated on distinct fibers or during different phases of development and thus influences the adhesive preferences of individual growth cones. Moreover, the colocalization of these glycoproteins raises the possibility that fasciculation may be mediated by the interaction of these molecules with each other. Molecules of the immunoglobulin supergene family are known to interact with members of their kind (Edelman et al., 1987; Lewis et al., 1986). It will be interesting to know the structure of other axonal glycoproteins and if contactin interacts with any of those. If a homotypic-binding mechanism (contactin–contactin) mediates adhesion between neuronal cell surfaces, then it would be of low affinity, since most, if not all of the extracellular part of the 130-kD glycoprotein exists as a monomer in dilute solutions (Moss, 1986). Recent work also demonstrated an interaction between a member of the immunoglobulin and a member of the integrin families (Marlin and Springer, 1987); with respect to the resemblance of contactin segments to fibronectin type III repeats, a similar mechanism could exist for this molecule. However, other possibilities are conceivable.

The accumulation of the 130-kD protein at discrete areas of the neuronal cell surface could explain its tight association with the membrane skeleton. The linkage with the cytoskeleton could serve, for example, to immobilize membrane components at some, but not other regions of the neuronal cell surface and thus establish neuronal subcompartments. The short length of the cytoplasmic segment indicates that the putative interaction of the 130-kD polypeptide with actin filaments is most likely not a direct one. In analogy with the β chain of the fibronectin receptor, a transmembrane protein with 47 cytoplasmic amino acids (Tamkun et al., 1986; Argraves et al., 1987) that connects the extracellular matrix

with actin via talin and vinculin (Horwitz et al., 1986; Singer and Paradiso, 1981), it seems more appropriate to suggest that the putative interaction of the 130-kD protein with the cytoskeleton is mediated through one or more additional polypeptides. Since GP130 is isolated as part of an actin-containing complex consisting of ~20 polypeptides, it is reasonable to assume that one of these polypeptides serves as an intracellular ligand for the 130-kD transmembrane protein. We are now testing this hypothesis with peptides representing intracellular contactin sequence. The proline-cysteine motif adjacent to the transmembrane domain on the cytoplasmic site could serve as a kink to bend the cytoplasmic region into the correct orientation for its interaction with the cytoskeletal elements.

Contactin is one of the first selectively expressed neuronal glycoproteins with a known structure. The immunoglobulin-like repeats and the sequence similarities to NCAM and fibronectin strongly suggest that contactin is involved in some aspect of cellular adhesion in the nervous system. Because of its putative association with the cytoskeleton and its concentration in axon fascicles and synapses, it appears a reasonable hypothesis that contactin could act in the formation or maintenance of stable contacts between neuronal cells. The probes described here will greatly facilitate the testing of this hypothesis and help to determine the role of contactin in the molecular puzzle of neuronal communication.

I wish to thank Dr. Ron McKay, in whose laboratory at the Massachusetts Institute of Technology (MIT) a large part of this work was carried out for his encouragement and support, Dr. C. Kintner for his help in the construction of the cDNA libraries and Drs. D. Melton and D. Weeks for introducing me to DNA sequencing. The help of Drs. R. F. Doolittle and C. Van Beveren in searching data bank sequences and C. Lai for running the Dayhoff ALIGN Program is gratefully acknowledged. I am thankful to Drs. T. Jessel and C. Goodman for their suggestion to investigate the relationship of contactin with fibronectin and to Drs. C. Kintner and E. Ruoslahti for valuable discussions and comments on the manuscript. I also thank Dr. D. Cleveland for providing the chicken β actin cDNA, Philip Hsu for assistance with the immunohistochemistry, and Candace Farmer and Cindy Clatterbuck for secretarial work.

My work at MIT was supported through the German Gene Technology Program (fellowship and research award 312 402 609/5 obtained through the German Academic Exchange Service [DAAD]) and at the La Jolla Cancer Research Foundation through Public Health Service grant NS 25194 (to B. Ranscht).

Received for publication 18 May 1988, and in revised form 13 July 1988.

References

Amzel, L. M., and R. J. Poljak. 1979. Three-dimensional structure of immunoglobulins. *Annu. Rev. Biochem.* 48:961–997.

Argraves, S., S. Suzuki, H. Arai, K. Thompson, M. Pierschbacher, and E. Ruoslahti. 1987. Amino acid sequence of the human fibronectin receptor. *J. Cell Biol.* 105:1183–1190.

Barbacid, M. 1987. ras-Genes. *Annu. Rev. Biochem.* 56:779–827.

Barthels, D., M. J. Santoni, W. Wille, C. Ruppert, J. C. Chaix, M. R. Hirsch, J. C. Fontecilla-Camps, and C. Gorides. 1987. Isolation and nucleotide sequence of mouse NCAM cDNA the codes for a Mr 79,000 polypeptide without a membrane-spanning region. *EMBO (Eur. Mol. Biol. Organ.) J.* 6: 907–914.

Beasley, L., and W. B. Stallcup. 1987. The nerve growth factor-inducible large external (NILE) glycoprotein and neural cell adhesion molecule (N-CAM) have distinct patterns of expression in the developing rat central nervous system. *J. Neurosci.* 7:708–715.

Bennett, M. K., and M. Kennedy. 1987. Deduced primary structure of the β-subunit of brain type II Ca^{2+}/calmodulin-dependent protein kinase determined by molecular cloning. *Proc. Natl. Acad. Sci. USA.* 84:1794–1798.

Bornstein, P., and W. Traub. 1979. The chemistry and biology of collagen. *In* The Proteins. Neurath Hill, editor. Academic Press, Inc., New York/San Francisco/London. 411–632.

Cathala, G., J. F. Savouret, B. Mendez, B. West, M. Karin, J. A. Martial, and J. Baxter. 1983. A method for isolation of intact, translationally active ribonucleic acid. *DNA (NY).* 2:329–335.

Chirgwin, J. M., A. E. Przybyla, R. J. MacDonald, and W. J. Rutter. 1979. Isolation of biologically active ribonucleic acid from sources enriched in ribonuclease. *Biochemistry.* 18:5294–5299.

Cleveland, D. W., S. G. Fisher, M. W. Kirschner, and U. K. Laemmli. 1977. Peptide mapping by limited proteolysis in sodium dodecylSulfate and analysis by gel electrophoresis. *J. Biol. Chem.* 252:1102–1106.

Cleveland, D. W., M. A. Lopata, R. J. MacDonald, N. J. Cowan, W. J. Rutter, and M. W. Kirschner. 1980. Number and evolutionary conservation of α- and β-tubulin and cytoplasmic β- and γ-actin genes using specific cloned cDNA probes. *Cell.* 20:95–105.

Cook, J., R. S. Pratt, and J. Lilien. 1984. Turnover and orientation of the major neural retina cell surface protein protected from tryptic cleavage by calcium. *Biochemistry.* 23:899–904.

Coussens, L., C. Van Beveren, D. Smith, E. Chen, R. L. Mitchell, C. M. Isacke, I. Verma, and A. Ullrich. 1986. Structural alteration of viral homologue of receptor oncogene fms at carboxyl terminus. *Nature (Lond.).* 320:277–280.

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

Cunningham, B. A., S. Hoffman, U. Rutishauser, J. J. Hemperley, and J. M. Edelman. 1983. Molecular topography of the neural cell adhesion molecule N-CAM: surface orientation and location of sialic acid-rich and binding regions. *Proc. Natl. Acad. Sci. USA.* 80:3116–3120.

Daniloff, J. K., C. M. Chuong, G. Levi, and G. M. Edelman. 1986. Differential distribution of cell adhesion molecules during histogenesis of the chick nervous system. *J. Neurosci.* 6:739–758.

Dayhoff, M. O., W. C. Barker, and L. T. Hunt. 1983. Establishing homologies in protein sequences. *Methods Enzymol.* 91:524–545.

Deák, F., I. Kiss, K. Sparks, S. W. Argraves, G. Hampikian, and P. Goetinck. 1986. Complete amino acid sequence of chicken cartilage link protein deduced from cDNA clones. *Proc. Natl. Acad. Sci. USA.* 83:3766–3770.

Edelman, G. M. 1987. Cell adhesion molecules in the regulation of animal form and tissue pattern. *Annu. Rev. Cell Biol.* 2:81–116.

Edelman, G. M., B. A. Murray, R. M. Mege, B. A. Cunningham, and W. J. Gallin. 1987. 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.

Gallin, W. J., B. C. Sorkin, G. M. Edelman, and B. A. Cunningham. 1987. Sequence analysis of a cDNA clone encoding the liver cell adhesion molecule L-CAM. *Proc. Natl. Acad. Sci. USA.* 84:2808–2812.

Geiger, B., E. Schmid, and W. Franke. 1983. Spatial distribution of proteins specific for desmosomes and adhaerens junctions in epithelial cells demonstrated by double immunofluorescence microscopy. *Differentiation.* 23: 189–205.

Geiger, B., K. T. Tokuyasu, A. H. Dutton, and S. J. Singer. 1980. Vinculin, an intracellular protein localized at specialized sites where microfilament bundles terminate at cell membranes. *Proc. Natl. Acad. Sci. USA.* 77: 4127–4131.

Gold, D. P., J. M. Puck, C. L. Pettey, M. Cho, J. Coligan, J. N. Woody, and C. Terhorst. 1986. Isolation of cDNA clones encoding the 20K non-glycosylated polypeptide chain of the human T-cell receptor/T3 complex. *Nature (Lond.).* 321:431–434.

Graham, R. C., and M. J. Karnovsky. 1966. The early stages of adsorption of injection horseradish peroxidase in the proximal tubes of mouse kidney: ultrastructural cytochemistry by a new technique. *J. Histochem. Cytochem.* 14:291–302.

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

Hampe, A., M. Gobet, C. J. Sherr, and F. Galibert. 1984. Nucleotide sequence of the feline retroviral oncogene v-fms shows unexpected homology with oncogenes encoding tyrosine-specific protein kinases. *Proc. Natl. Acad. Sci. USA.* 81:85–89.

Hatta, K., A. Nose, A. Nagafuchi, and M. Takeichi. 1988. Cloning and expression of cDNA encoding a neural calcium-dependent cell adhesion molecule: its identity with the cadherin gene family. *J. Cell Biol.* 106:873–881.

Hemperley, J. J., B. A. Murray, G. M. Edelman, and B. A. Cunningham. 1986. Sequence of a cDNA clone encoding the polysialic acid-rich and cytoplasmic domains of the neural cell adhesion molecule NCAM. *Proc. Natl. Acad. Sci. USA.* 83:3037–3041.

Horwitz, A., K. Duggan, C. Buck, M. C. Beckerle, and K. Burridge. 1986. Interaction of plasma membrane fibronectin receptor with talin: a transmembrane linkage. *Nature (Lond.).* 320:531–533.

Hunkapillar, M. W., E. Lujan, F. Osterander, and L. Hood. 1983. Isolation of microgram quantities of proteins from polyacrylamide gels for amino acid sequence analysis. *Methods Enzymol.* 91:227–236.

Hunter, T., and J. A. Cooper. 1985. Protein-tyrosine kinases. *Annu. Rev. Biochem.* 54:897–930.

Hynes, R. O. 1985. Molecular biology of fibronectin. *Annu. Rev. Cell Biol.* 1:67–90.

Hynes, R. O. 1987. Integrins: a family of cell surface receptors. *Cell.* 48: 549–554.

Ishioka, N., N. Takahashi, F. W. Putnam. 1986. Amino acid sequence of human plasma α_1B-glycoprotein: homology to the imunoglobulin supergene family. *Proc. Natl. Acad. Sci. USA.* 83:2363–2367.

Johnson, P., and A. F. Williams. 1986. Striking similarities between antigen receptor J pieces and sequence in the second chain of the murine CD8 antigen. *Nature (Lond.).* 323:74–76.

Kartenbeck, J., W. W. Franke, J. G. Moser, and U. Stoffels. 1983. Specific attachment of desmin filaments to desmosomal plaques in cardiac myocytes. *EMBO (Eur. Mol. Biol. Organ.) J.* 2:735–742.

Kitamura, N., A. Kitamura, K. Toyoshima, Y. Hirayama, and M. Yoshida. 1982. Avian sarcoma virus Y73 genome sequence and structural similarity of its transforming gene product to that of rous sarcoma virus. *Nature (Lond.).* 297:205–208.

Kopito, R. R., and H. F. Lodish. 1985. Primary structure and transmembrane orientation of the murine anion exchange protein. *Nature (Lond.).* 316: 234–238.

Kost, T. A., N. Theodorakis, and S. H. Hughes. 1983. The nucleotide sequence of the chick cytoplasmic β-actin gene. *Nucleic Acids Res.* 11:8287–8301.

Kozak, M. 1984. Compilation and analysis of sequences upstream of the translational start site in eukaryotic mRNAs. *Nucleic Acids Res.* 12:857–872.

Kyte, J., and R. F. Doolittle. 1982. A simple method for displaying the hydropathic character of a protein. *J. Mol. Biol.* 157:105–132.

Laemmli, U. K. 1970. Cleavage of the structural proteins during the assembly of the head of bacteriophage T4. *Nature (Lond.).* 227:680–685.

Lai, C., M. D. Brow, K. A. Nave, A. B. Noronah, R. H. Quarles, F. E. Bloom, R. J. Milner, and J. G. Sutcliff. 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.

Lemke, G., and R. Axel. 1985. Isolation and sequence of a cDNA encoding the major structural protein from peripheral myelin. *Cell.* 40:501–508.

Lemmon, V., and S. C. McLoon. 1986. The appearance of an L1-like molecule in the chick primary visual pathway. *J. Neurosci.* 6:2987–2994.

Lewis, V. A., T. Koch, H. Plutner, and I. Mellmann. 1986. A complementary cDNA clone for a macrophage-lymphocyte Fc receptor. *Nature (Lond).* 324:372–375.

Maniatis, T., E. Fritsch, and J. Sambrook. 1982. Molecular cloning. A Laboratory Manual. Cold Spring Harbor Laboratories, Cold Spring Harbor, NY.

Marchesi, V. T. 1985. Stabilizing infrastructure of cell membranes. *Annu. Rev. Cell Biol.* 1:531–561.

Mark, G. E., and U. R. Rapp. 1984. Primary structure of v-raf: relatedness to the src family of oncogenes. *Science (Wash. DC).* 224:285–288.

Marlin, S. D., and T. A. Springer. 1987. Purified intercellular adhesion molecule-1 (ICAM-1) is a ligand for lymphocyte function-associated antigen-1 (LFA-1). *Cell.* 51:813–819.

Melton, D. A., P. A. Krieg, M. R. Rebagliati, T. Maniatis, K. Zinn, and M. R. Green. 1984. Efficient in vitro synthesis of biologically active RNA and RNA hybridization probes from plasmids containing bacterial SP6 promoter. *Nucleic Acids Res.* 12:7035–7056.

Messing, J. 1983. New vectors for M13 cloning. *Methods Enzymol.* 101:20–78.

Miller, E. J., and S. Gay. 1987. The collagens: an overview and update. *Methods Enzymol.* 144:3–41.

Mosher, D. F. 1984. Physiology of fibronectin. *Annu. Rev. Med.* 35:561–575.

Moss, D. J. 1983. Cytoskeleton-associated glycoproteins from chicken sympathetic neurons and chicken embryo brain. *Eur. J. Biochem.* 135:291–297.

Moss, D. J. 1986. Characterisation of a soluble trypsin fragment of GP130: a neuronal glycoprotein associated with the cytoskeleton. *J. Cell Biochem.* 32:97–112.

Nagafuchi, A., Y. Shirayoshi, K. Okazaki, K. Yasuda, and M. Takeichi. 1987. Transformation of cell adhesion properties by exogenously introduced E-cadherin cDNA. *Nature (Lond.).* 329:341–343.

Neame, P. J., J. E. Christner, and J. R. Baker. 1986. The primary structure of link protein from rat chondrosarcoma proteoglycan aggregate. *J. Biol. Chem.* 261:3519–3535.

O'Farrell, P. H. 1975. High resolution two-dimensional electrophoresis of proteins. *J. Biol. Chem.* 250:4007–4021.

Oikawa, S., H. Nakazato, and G. Kosaki. 1987. Primary structure of human carcinoembryonic antigen (CEA) deduced from cDNA sequence. *Biochim. Biophys. Res. Commun.* 142:511–518.

Pearson, R. B., J. R. Woodgett, P. Cohen, and B. E. Kemp. 1985. Substrate specificity of a multifunctional calmodulin-dependent protein kinase. *J. Biol. Chem.* 260:14471–14476.

Peterson, T. E., H. C. Thorgerson, K. Skorstengaard, K. Vibe-Pederson, P. Sahl, L. Sottrup-Jensen, and S. Magnusson. 1983. Partial primary structure of bovine plasma fibronectin: three types of internal homology. *Proc. Natl. Acad. Sci. USA.* 80:137–141.

Pierschbacher, M. D., and E. Ruoslahti. 1984*a*. Cell attachment activity of fibronectin can be duplicated by small synthetic fragments of the molecule. *Nature (Lond.).* 309:30–33.

Pierschbacher, M. D., and E. Ruoslahti. 1984*b*. Variants of the cell recognition site of fibronectin that retain attachment-promoting activity. *Proc. Natl. Acad. Sci. USA.* 81:5985–5988.

Poncz, M., D. Solowiejczyk, M. Ballantine, E. Schwartz, and S. Surrey. 1982. Nonrandom DNA sequence analysis in bacteriophage M13 by the dideoxy chain termination method. *Proc. Natl. Acad. Sci. USA.* 79:4298–4302.

Queen, C., and L. J. Korn. 1984. A comprehensive sequence analysis program for the IBM personal computer. *Nucleic Acids Res.* 12:581–599.

Ranscht, B., D. J. Moss, and C. Thomas. 1984. A neuronal surface glycoprotein associated with the cytoskeleton. *J. Cell Biol.* 99:1803–1813.

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

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

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

Ruoslahti, E., and M. D. Pierschbacher. 1987. New perspectives in cell adhesion: RGD and integrins. *Science (Wash. DC).* 238:491–497.

Sabatini, D. D., G. Kreibich, T. Morimoto, and M. Adesnik. 1982. Mechanisms for the incorporation of proteins in membranes and organelles. *J. Cell Biol.* 92:1–22.

Salzer, J. L., W. P. Holmes, and D. R. Colman. 1987. The amino acid sequences of the myelin-associated glycoproteins: homology with the immunoglobulin gene family. *J. Cell Biol.* 104:957–965.

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

Saul, F. A., L. M. Amzel, and R. J. Poljak. 1978. Preliminary refinement and structural analysis of the Fab fragment from human immunoglobulin new at 2.0Å resolution. *J. Biol. Chem.* 253:585–597.

Sheetz, M. P. 1979. Integral membrane protein interaction with triton cytoskeletons of erythrocytes. *Biochim. Biophys. Acta.* 557:122–134.

Sherr, C. J., C. W. Rettenmier, R. Sacca, M. F. Roussel, A. T. Look, and E. R. Stanley. The c-fms proto-oncogene product is related to the receptor for the mononuclear phagocyte growth receptor, CSF-1. *Cell.* 41:665–676.

Shirayoshi, Y., K. Hatta, M. Hosoda, S. Tsunasawa, F. Sakiyama, and M. Takeichi. 1986. Cadherin cell adhesion molecules with distinct binding specificities share a common structure. *EMBO (Eur. Mol. Biol. Organ.) J.* 5:2485–2488.

Shoji, S., D. C. Parmelee, R. D. Wade, S. Kumar, L. H. Ericsson, K. A. Walsh, H. Neurath, G. L. Long, J. G. Demaille, E. H. Fischer, and K. Titani. 1981. Complete amino acid sequence of the catalytic subunit of bovine cardiac muscle cyclic AMP-dependent protein kinase. *Proc. Natl. Acad. Sci. USA.* 78:848–851.

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

Singer, I. I., and P. R. Paradiso. 1981. A transmembrane relationship between fibronectin and vinculin (130kd protein): serum modulation in normal and transformed hamster fibroblasts. *Cell.* 24:481–492.

Stallcup, W. B., L. Beasley, and J. Levine. 1985. Antibody against nerve growth factor-inducible large external (NILE) glycoprotein labels nerve fiber tracts in the developing rat nervous system. *J. Neurosci.* 5:1090–1101.

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

Sutcliffe, G., R. J. Milner, T. M. Shinnick, and F. E. Bloom. 1983. Identifying the protein products of brain-specific genes with antibodies to chemically synthesized peptides. *Cell.* 33:671–682.

Suzuki, S., S. W. Argraves, R. Pytela, H. Arai, T. Krusius, M. D. Pierschbacher, and E. Ruoslahti. 1986. cDNA and amino acid sequences of the cell adhesion protein receptor recognizing vitronectin reveal a transmembrane domain and homologies with other adhesion protein receptors. *Proc. Natl. Acad. Sci. USA.* 83:8614–8618.

Tacke, R., M. Moos, D. B. Teplow, K. Fruh, H. Scherer, A. Bach, and M. Schacher. 1987. Identification of cDNA clones of the mouse neural cell adhesion molecule L1. *Neurosci. Lett.* 82:89–94.

Takeya, T., and H. Hanafusa. 1983. Structure and sequence of the cellular gene homologous to the RSV src gene and the mechanism for generating the transforming virus. *Cell.* 32:881–890.

Tamkun, J. W., D. W. DeSimone, D. Fonda, R. S. Patel, C. Buck, A. F. Horwitz, and R. O. Hynes. 1986. Structure of integrin, a glycoprotein involved in the transmembrane linkage between fibronectin and actin. *Cell.* 46: 271–282.

Thompson, J. A., H. Pande, R. J. Paxton, L. Shively, A. Padma, R. L. Simmer, C. W. Todd, A. D. Riggs, and J. E. Shively. 1987. Molecular cloning of a gene belonging to the carcinoembryonic antigen gene family and discussion of a domain model. *Proc. Natl. Acad. Sci. USA.* 84:2965–2969.

Tomita, M., H. Furthmayr, and V. T. Marchesi. 1978. Primary structure of human erythrocyte glycophorin A. Isolation and characterization of peptides and complete amino acid sequence. *Biochemistry.* 17:4756–4770.

Towbin, H., T. Staehlin, and J. Gordon. 1979. Electrophoretic transfer of protein from polyacrylamide gels to nitrocellulose sheets: procedure and some applications. *Proc. Natl. Acad. Sci. USA.* 76:4350–4354.

Ullrich, A., J. R. Bell, E. Y. Chen, R. Herrera, L. M. Petrzzelli, T. J. Dull, A. Gray, L. Coussens, Y. C. Liao, M. Tsubokawa, A. Mason, P. H. Seeburg, C. Grunfield, O. M. Rosen, and J. Ramachandran. 1985. Human insulin receptor and its relationship to the tyrosine kinase family of oncogenes. *Nature (Lond.).* 313:756–761.

Ullrich, A., L. Coussens, J. S. Hayflick, T. J. Dull, A. Gray, A. W. Tam, J. Lee, Y. Yarden, T. A. Libermann, J. Schlessinger, J. Downward, E. L. Mayes, N. Whittle, M. D. Waterfield, and P. H. Seeburg. 1984. Human epidermal growth factor receptor cDNA sequence and aberrant expression of the amplified gene in A431 epidermoid carcinoma cells. *Nature (Lond.).* 309:418–425.

Van Beveren, C., and I. Verma. 1986. Homology among oncogenes. *Curr. Top. Microbiol. Immunol.* 123:73–98.

Van Beveren, C., F. Van Straaten, J. A. Galleshaw, and I. M. Verma. 1981. Nucleotide sequence of the genome of a murine sarcoma virus. *Cell.* 27: 97–108.

Volk, T., and B. Geiger. 1984. A 135kd membrane protein of intercellular adherens junctions. *EMBO (Eur. Mol. Biol. Organ.) J.* 3:2249–2260.

Von Heijne, G. 1982. Patterns of amino acids near signal-sequence cleavage sites. *Eur. J. Biochem.* 133:17–21.

Williams, A. F. 1987. A year in the life of the immunoglobulin superfamily. *Immunol. Today.* 8:298–305.

Williams, A. F., and J. Gagnon. 1982. Neuronal cell Thy-1 glycoprotein: homology with immunoglobulin. *Science (Wash. DC).* 216:696–703.

Wood, W. I., J. Gitschier, L. A. Laskey, and R. M. Lawn. 1985. Base composition independent hybridization in tetramethylammonium chloride: a method for oligonucleotide screening of highly complex gene libraries. *Proc. Natl. Acad. Sci. USA.* 82:1585–1588.

Yamada, K., and D. W. Kennedy. 1984. Dualistic natures of adhesive protein function: fibronectin and its biologically active peptide fragments can autoinhibit fibronectin function. *J. Cell Biol.* 99:29–36.

Yamada, K. M. 1983. Cell surface interactions with extracellular materials. *Annu. Rev. Biochem.* 52:761–799.

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

Young, R. A., and R. W. Davis. 1985. Immunoscreening λgt.11 recombinant DNA expression libraries. *In* Genetic Engineering Principles and Methods. Vol. 7. J. Setlow and A. Hollaender, editors. Plenum Publishing Corp., New York. 29–41.

Zimmerman, W., B. Ortlieb, R. Friedrich, and S. Von Kleist. 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.