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.
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 (Wil- to fibronectin type III repeats. Both structures are known to
nervous system. In combination with the immunoglobulin-
super gene family expressed in the
Protein Isolation and Sequencing
and its localization in areas of the nervous system where neu-
rons 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 [32P]-protein A (400,000 cpm/ml) as previously described (Ranscht et al., 1984). For amino acid sequence analy-

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)-cellu-
lose (type 3; Collaborative Research, Inc., Waltham, MA). CDNA was syn-
thetized according to Gubler and Hoffman (1983). First strand synthesis was for 2-4 h in 50% deionized formamide, 5x SSPE, 1% SDS, and

Isolation of cDNA Clones
Nitrocellulose replica filters of the unamplified λgt.11 library were screened with monoclonal and polyclonal antibodies against GPI30 (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. Horserad-
ish peroxidase conjugated secondary antibodies (Cappel Laboratories, Inc., Cochraneville, PA) and the diaminobenzidine enzyme substrate were used to monitor binding of the antibodies (Graham and Karnovsky, 1966). Of ~105 independent recombinants screened, seven positive clones were selected and plaque purified. The antibody-selected clones were rescreened on nylon replica filters (Biodyne Electronic, 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 GPI30. Mixed oligonucleotides were synthe-
sized by Dr. D. Rigby at the Harvard Microchemistry Facility. The oligo-
nucleotide probes were labeled with 32P-ATP (Amersham Corp., Arlington Heights, IL) using T4-Polymerase 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 nonspecific 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 pre-
hybridized at 37°C for 4-16 hr in 6 x SSPE, 0.1% SDS, 1% NaCl, 0.2
 M NaH2PO4, H2O, 0.02 M EDTA, Na+, pH 7.4, 5 x Denhardt’s, and

Subcloning and Sequence Determination
λgt.11 DNAs containing selected cDNA inserts were purified from liter by-
sates 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 separa-
ted 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 elec-
 troelution (Seakam agarase; 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 oli-
gonucleotide 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 iso-
lated from the insert containing λgt.11 clones.

RNA Analysis
Total cellular RNA was isolated according to the guanidinium isothio-
cyanate 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 20x SSPE. Membranes were cross-linked with UV light and baked for 60 min under vacuum at 80°C. Prehybridiza-
tion was for 2-4 h in 50% deionized formamide, 5x SSPE, 1% SDS, and

Immunohistochemistry
Chickens were perfused with 100 mM Na-peridate, 75 mM lysine, and 3% paraformaldehyde. Nerve tissue was removed and immersion fixed in the above fixative for an additional 2 hr. The tissue was cryoprotected by over-
night 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

The Journal of Cell Biology, Volume 107, 1988 1562
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 125I-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 Agt-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 125I–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 to which mixed oligonucleotides were synthesized.
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 if 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 NH2-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 J). 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.

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. 4b. 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. 4b, 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. 4b) 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. 5A. 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 (DXGXYXCVX/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. 5B; 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) type III segments. The overall similarity of contactin type 2 repeats to FN type III blocks is between 20 and 31%, whereas amino acids of FN type III repeats are more conserved in contactin type 2 repeat I as compared with repeat II. In Fig. 6b, 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; Argurak et al., 1988), or components of the erythrocyte membrane skeleton band 3 (Kopito and Lodish, 1985) and glycophorin (Tomita et al., 1978).

\textbf{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-
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 potential sites for asparagine-linked glycosylation are marked with solid triangles. A potential nucleotide-binding site is marked with an asterisk.

The Journal of Cell Biology, Volume 107, 1988 1566
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 GPI130 protein sequence using the Microgenie sequence analysis program (Queen and Korn, 1984). This comparison included members of the cadherin and immunoglobulin superfamilies (E-cadherin: Naga-

fuchi et al., 1987; N-cadherin: Hatta et al., 1988; L-CAM: Gallin et al., 1987; N-CAM: Cunningham et al., 1987; Hem-

perley 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 GPI130 amino acids 70–335  
and 356–450 share a 31% identity with NCAM residues 62–  
322 and 360–452, respectively (Fig. 6 a). GPI130 appears to  
be less similar to MAG (Salzer et al., 1987; Sutcliffe et al.,  
1983), since the Microgenie sequence analysis program re-

vealed only short stretches of amino acid identity. In addi-
tion, 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,β-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 trans-
forming protein v-fms (Hampe et al., 1984) whose cellular  
counterpart has been proposed to act as a receptor for macro-
phage CSF-1 (Sherr et al., 1985; Coussens et al., 1986). The  
similarity of the 130-kD glycoprotein with v-fms was de-

tected 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 immu-

noglobulin supergene family in the nervous system and sug-

gest 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 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 extracel-

lular region resemble fibronectin type  
III repeats.
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 (GXXGFG, 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. 4b) 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 (Hunter 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.

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 with 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 immuno-histochemical 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 localization 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 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 intrachain disulfide bonds.

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 Δ.
interaction is its localization in the fiber-rich areas of the retina, particular the significant conservation of amino acid residues to members of the immunoglobulin supergene family and in the two molecules indicate similar binding capacities. De-
cell-binding region of NCAM has not been closely mapped,
mains (Cunningham et al., 1983, 1987). Since, however, a that of NgCAM/L1/G4/NILE/8D9 (Daniloff et al., 1986; Beasley and Stallcup, 1987), contactin staining inina. Even though the pattern of contactin expression in the adult retina is distinct from that of NgCAM/L1/G4/NILE (Daniloff 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 ki-
ase activity are known, it is not clear if the nucleotide bind-
ing and the autophosphorylation sequence motifs are of any functional relevance.

Contactin shares significant similarity (31%) with the neu-
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 do-
ains (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. De-
pite the similarities of contactin and NCAM in their globu-
lar domains, amino acid sequences carboxy terminal to these regions are distinct. The structural resemblance of contactin to members of the immunoglobulin superfine 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 cell-
ular communication in the nervous system.

In line with a prospective role of contactin in cell–cell in-
teraction is its localization in the fiber-rich areas of the reti-
a. Even though the pattern of contactin expression in the adult retina is distinct from that of Ng-CAM/NILE (Danillof et al., 1986; Beasley and Stallcup, 1987), contactin staining may overlap during restricted phases of development with that of NgCAM/LI/G4/NILE/8D9 (Danillof et al., 1986; Stallcup et al., 1985; Beasley and Stallcup, 1987; Lemmon and McLoon, 1986) neurofascin (Rathjen et al., 1987b), or F11 (Rathjen et al., 1987a) (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 rec-
ognized by monospecific antibodies (Fig. 1 left and Ranscht et al., 1984). In contrast, the 130–140kd components of NgCAM/LI/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., 1987a, b). (b) Contactin is highly in-
soluble in non-ionic detergents. An isolation procedure based on affinity chromatography as described for LI/G4, F11, and neurofascin (Rathjen and Schachner, 1984; Rathjen et al., 1987a, b) cannot be applied to isolate contactin. (c) The re-
striction 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 LI/G4 (Rathjen et al., 1987a) is distinct from the contactin sequence (Figs. 1 right and 4 b). Despite these molecular differences, contactin may be related structurally and func-
tionally 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/LI/G4/NILE/8D9, neurofascin (Rathjen et al., 1987b), F11 (Rathjen et al., 1987a), N-Cal-
CAM (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 com-
ponents 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 immunoglobu-
lin superfine 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 glycopro-
tins and if contactin interacts with any of those. If a homo-
typic-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 solu-
tions (Moss, 1986). Recent work also demonstrated an in-
teraction 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 con-ceivable.

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 cytoskele-
ton could serve, for example, to immobilize membrane compo-
ents 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 filament 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; Ar-
graves 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 cytoskeletal 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. Kinney 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. Kinney and E. Ruslalati for valuable discussions and comments on the manuscript. I also thank Dr. D. Cleveland for providing the 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 M. G. Edelman. 1987. Molecular localization 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.
Bennett, M. K., and M. Kennedy. 1987. Deduced primary structure of the 13-subunit of brain type II Ca2+/calmodulin-dependent protein kinase determined by molecular cloning. Proc. Natl. Acad. Sci. USA. 84:1794-1798.
Bironstein, 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. 411-652.
Chidam, G. J., F. Savour, B. J. B. West, M. Karin, J. A. Martial, and J. Baxter. 1983. A method for isolation of intact, translationally active ribonucleic acid (RNA) (IV). 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. Kirchmann, and U. K. Laemmli. 1977. Peptide mapping by limited proteolysis in sodium dodecyl sulfate and analysis by gel electrophoresis. J. Biol. Chem. 252:1102-1106.
Cleveland, D. W., M. A. MacDonald, N. J. Cowan, W. J. Rutter, and M. W. Kirchmann. 1980. Number and evolutionary conservation of α- and β-tubulin and cytoplasmic β- and γ-actin genes using specific cloned cDNA probes. Cell. 20:95-107.
Cook, R. S. Pratt, and J. Lillen. 1984. Turnover and orientation of the major neuronal retina cell surface protein protected from trypsin cleavage by calmodulin. Biochemistry. 23:899-904.
Cousins, L., C. Van Beveren, D. Smith, B. C. Sorkin, G. M. Edelman, and B. A. Cunningham. 1987. Isolation of cDNA clones encoding the 20K non-glycosylated NCAM from human brain. Proc. Natl. Acad. Sci. USA. 84:8502-8506.
Dallman, R. J., and 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. 6:179-187.
Dayhoff, M. O., W. C. Barker, and L. T. Hunt. 1983. Establishing homologies in protein sequences. Methods Enzymol. 91:524-545.
Deck, F., I. Kiss, K. Sparks, S. W. Argraves, G. Hampkian, 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. Galin. 1987. Cellular expression of liver and neuronal cell adhesion molecules after transfection with their cDNAs results in specific cell-cell binding. Proc. Natl. Acad. Sci. USA. 84:8502-8506.
Galil, 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 adherens 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.
Goid, D., J. M. Puck, C. L. Pettry, M. Cho, J. Coligan, J. N. Woody, and C. G. Berry. 1986. cDNA cloning of DNA encoding the 20K non-glycosylated polypeptide chain of the human T-cell receptor/3 complex. Nature (Lond.). 321:431-434.
Graham, R. C., and M. J. Karnovsky. 1964. The early stages of adsorption of immunoglobulins to a plastic surface and the role of calcium. J. Histochem. Cytochem. 12:291-302.
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. Sharp, 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 neuronal calcium-dependent cell adhesion molecule: its identity with the cadherin gene family. J. Cell Biol. 106:873-881.
Hampe, A. J., B. A. Murray, G. M. Edelman, and B. A. Cunningham. 1984. Sequence of a cDNA genome encoding the polypeptide acid-rich and cytoplasmic domains of the neuronal cell adhesion molecule NCAM. Proc. Natl. Acad. Sci. USA. 83:3037-3041.
Horwitz, A., K. Duggan, C. Beck, M. C. Beckerle, and K. Burridge. 1986. Interaction of plasma membrane fibronectin receptor with talin: a transmembrane linkage. Nature (Lond.). 320:531-533.
Hunkapiller, M. W., E. Lujan, P. Osterlander, and L. Hood. 1983. Isolation of microgram quantities of protein 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.
Ullrich, A., J. R. Bell, E. Y. Chen, R. Herrera, L. M. Petrezelli, T. J. Dull, A. Gray, L. Coussens, Y. C. Liao, M. Tsukokawa, A. Mason, P. H. Seeburg, C. Grunfeld, 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. 82:1585-1589.

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 Agt.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.