Neuroglycan C, a Novel Membrane-spanning Chondroitin Sulfate Proteoglycan That Is Restricted to the Brain*

(Received for publication, May 22, 1995, and in revised form, August 21, 1995)

Eiji Watanabe‡§, Nobuaki Maeda§, Fumiko Matsui, Yoichi Kushima**, Masaharu Noda§, and Atsuhiko Oohira‡

From the Department of Perinatology and Neuroglycoscience, Institute for Developmental Research, Kasugai, Aichi 480-03, Japan, the Division of Molecular Neurobiology, National Institute for Basic Biology, and Department of Molecular Biomechanics, Ibaraki University for Advanced Studies, Okazaki 444, Japan, and the Tokyo Research Institute, Sekagaku Corporation, Higashi-yamato, Tokyo 207, Japan

Monoclonal antibodies were raised to membrane-bound proteoglycans derived from rat brain, and four monoclonal antibodies that recognized a 150-kDa chondroitin sulfate proteoglycan with a core glycoprotein of 120 kDa were obtained. Immunohistological study revealed that the proteoglycan was associated with developing neurons. We screened rat brain cDNA libraries using the four monoclonal antibodies and isolated overlapping cDNA clones that encoded the entire core protein of 514 amino acids plus a 30-residue signal peptide. The deduced amino acid sequence suggested an integral membrane protein divided into five structurally different domains: an N-terminal domain to which chondroitin sulfate chains might be attached, a basic amino acid cluster consisting of seven arginine and two lysine residues, a cysteine-containing domain, a membrane-spanning segment, and a C-terminal cytoplasmic domain of 95 amino acids. On Northern blots, the cDNA hybridized with a single mRNA of 3.1 kilobases that was detectable in brains of neonatal and adult rats but not in kidney, liver, lung, and muscle of either. The sequence of the proteoglycan did not exhibit significant homology to any other known protein, indicating that the proteoglycan, designated neuroglycan C, is a novel integral membrane proteoglycan.

Proteoglycans are molecules in which glycosaminoglycan chains are covalently linked to core glycoproteins. They are major glycoconjugates, located at the cell surface and in extracellular spaces, and they are considered to have various regulatory effects on cells by binding to chemotrophic factors, cell adhesion molecules, and extracellular matrix glycoproteins (1, 2). Early biochemical studies of the mammalian brain revealed the presence of a varied set of proteoglycans whose expression is regulated precisely during brain development (3, 4). The developmentally regulated expression of such molecules has been implicated in the formation of several complicated structures in the brain, such as the mossy fiber (5), visual (6), and cortical barrel systems (7-9). Studies in vitro have also shown that neuronal proteoglycans are involved in the morphogenesis of neurons, for example, in neuritogenesis. They are capable of exerting both positive (10, 11) and negative (12-15) effects on the outgrowth of neurites in vitro.

Our understanding of the functions of neuronal proteoglycans is expanding as the primary structures of core proteins are elucidated. Such analyses have revealed that a number of proteoglycan core proteins share similar sequence motifs and can be grouped into several families. Neurocan (16), brevican (17), versican (18), and BEHAB (19) are members of a soluble chondroitin sulfate proteoglycan family; their sequences include the tandem repeats of hyaluronan-binding domain. N-Syndecan (20) is a member of the syndecan family of heparan sulfate proteoglycans, which have extensively similar transmembrane and cytoplasmic domains. Moreover, it was recently revealed that a chondroitin sulfate proteoglycan, phosphacan/684 proteoglycan, is an extracellular variant of a receptor-like protein-tyrosine phosphatase, RPTPβ (21, 22). Although primary structures of a number of core proteins have been reported, more proteoglycans in the brain remain to be characterized. In particular, the complete primary structures of only three membrane-bound proteoglycans have been reported, namely NG2 (23), glypican (24), and cerebroglycan (25), although there appear to be at least 16 core proteins in the membrane-bound fraction of the rat brain (4). Therefore, we attempted to raise monoclonal antibodies against membrane-bound proteoglycans and to isolate the cDNA clones that encoded for a proteoglycan-core protein.

Here, we report a novel membrane-spanning chondroitin sulfate proteoglycan that is recognized by a panel of monoclonal antibodies. The predicted protein is an integral membrane protein containing a signal peptide, a chondroitin sulfate-attachment domain, a cluster of basic amino acids, a cysteine-containing domain, a 24-amino acid transmembrane domain, and a cytoplasmic tail. The amino acid sequence of the core protein exhibits little similarity to other known proteins, indicating that the proteoglycan is a novel species of integral membrane proteoglycan. We propose the name neuroglycan C (NGC₁ from neuronal proteoglycan with chondroitin sulfate) for this proteoglycan.

EXPERIMENTAL PROCEDURES

Extraction of Membrane-bound Proteoglycans—Brains from 10-day-old Sprague-Dawley rats (usually 100 rats per experiment; SlcSD strain; SLC Inc., Shizuoka, Japan) were homogenized with a Teflon-
glass homogenizer in 2 ml per brain of ice-cold phosphate-buffered saline (PBS) that contained 20 mM EDTA, 10 mM N-ethylmaleimide, and 2 mM phenylmethylsulfonyl fluoride as protease inhibitors. Homogenization and subsequent steps were carried out at 4 °C unless otherwise indicated. The homogenate was centrifuged at 27,000 × g for 40 min. The pellet was subjected to recentrifugation in the same solution. After centrifugation, the PBS-insoluble material was homogenized in 2 ml per brain of PBS that contained the protease inhibitors plus 1% Nonidet P-40 at 6°C. This homogenate was centrifuged for 60 min on a magnetic stirrer and then centrifuged at 27,000 × g for 40 min. The supernatant was retained, and the pellet was treated again with the detergent-containing buffer. After centrifugation, both detergent-containing supernatant and pellet were combined and lyophilized. Detergent-containing supernatants were combined and lyophilized.

Partial Purification of Membrane-bound Proteoglycans—The lyophilized residue obtained from the pooled supernatants was suspended in 1 ml per brain of 4 M urea, 50 mM Tris-HCl, pH 7.5, that contained 0.2% detergent-containing buffer. After centrifugation, both detergent-containing supernatant and pellet were combined and lyophilized. Partial purification of proteoglycans was performed as described previously (8). The core glycoprotein moiety was then treated sequentially with 1 M guanidine HCl, 50 mM Tris-HCl, pH 7.5 (guanidine HCl buffer), at room temperature. The enzymatic digestion and subsequent steps were carried out at 4°C unless otherwise indicated.

Preparation of tissue sections from the rat cerebral cortex and immunostaining of the tissue sections were performed by previously described standard methods (9).

Cultured Rat Neurons—To purify the 120-kDa core glycoprotein, rat neurons were partially purified by the method used for partial purification of proteoglycan-type protein-tyrosine phosphatases that was described by Maeda et al. (22). To prepare the neuron preparation, 2 ml per brain of PBS that contained 2% Triton X-100 and the protease inhibitors. The proteoglycan fraction (Kw from 0.11 to 0.65) was pooled and concentrated to 3 ml on a Diaflo YM-10 membrane (Amicon Corp., Danvers, MA). The concentrated solution was chromatographed on a column (1.6 cm, inner diameter, × 100 cm) of Sepharose CL-4B (Pharmacia) in 4 M guanidine HCl, 50 mM Tris-HCl, pH 7.5, that contained 0.2% Nonidet P-40 and the protease inhibitors. The proteoglycan fraction (Kw from 0.11 to 0.65) was pooled and concentrated to 3 ml on a Diaflo YM-10 membrane. Proteoglycans were precipitated from the solution by addition of 3 volumes of 95% ethanol that contained 1.3% (w/v) potassium acetate at 0°C, and then it was dried over P2O5 in vacuo.

Screening of cDNA Libraries and Isolation of cDNA Clones—A novel proteoglycan of the brain (NGC3) was cloned by the limiting dilution method. mAbs C1, C3, C5, and C15 reacted with a chondroitin sulfate proteoglycan with a 120-kDa core glycoprotein. All of the four antibodies were determined to be IgG1 with a mouse monoclonal antibody (isotyping kit (Immunotype, Sigma).

Preparation of tissue sections from the rat cerebral cortex and immunostaining of the tissue sections were performed by previously described standard methods (9).

Cultured Rat Neurons—To purify the 120-kDa core glycoprotein, rat neurons were partially purified by the method used for partial purification of proteoglycan-type protein-tyrosine phosphatases that was described by Maeda et al. (22). To prepare the neuron preparation, 2 ml per brain of PBS that contained 2% Triton X-100 and the protease inhibitors. The proteoglycan fraction (Kw from 0.11 to 0.65) was pooled and concentrated to 3 ml on a Diaflo YM-10 membrane. Proteoglycans were precipitated from the solution by addition of 3 volumes of 95% ethanol that contained 1.3% (w/v) potassium acetate at 0°C, and then it was dried over P2O5 in vacuo.
primer-cDNA library) had been isolated by immunoscreening, subsequent clones were obtained by plaque hybridization using 32P-labeled cDNA probes. cDNA probes for plaque hybridization and Northern blotting were labeled with the Megaprime DNA labeling system (Amersham International, Amersham, UK). cDNA inserts from agt11 clones were purified and subcloned into the pBluescript II plasmid vector (Stratagene).

Sequencing of DNA—Subclones for sequencing were generated by the introduction of deletions by restriction enzymes. DNA sequencing was performed with Taq polymerase and dye-labeled primers using an automated DNA Sequencer (Alf II; Alfred combined system; Pharmacia). Editing and analysis of sequences were performed with GENETYX software (Softwate Development Co., Tokyo). The reading frame was verified by reference to our data for the N-terminal amino acid sequence of the 120-kDa core glycoprotein and for the CNBr-generated fragments. Comparisons with sequences in the GenBank (Release 87) and Swiss-Prot (Release 31) data bases were performed with a local alignment search tool (30).

Northern Blot Analysis—Total RNA was extracted from brains of rats on postnatal day 7 and adult rats, and also from kidney, liver, lung, and muscle of adult rats as described by Chomczynski and Sacchi (31).

Poly(A)+ RNA was purified with oligotex-dT30 as described in the instruction manual from the manufacturer (TAKARA Biomedicals, Otsu, Japan). Poly(A)+ RNA (5 μg) was subjected to electrophoresis on a 1% agarose gel that contained 2.2 M formaldehyde and transferred to a PVDF membrane (Immobilon N, Millipore). A 682-bp insert from the NGC1 was labeled with 32P-dCTP and used as a probe at a concentration of 3 x 106 cpm/ml. The filter was washed finally in 0.2 x SSC (SSC = 0.15 M NaCl and 0.015 M sodium citrate) at 68 °C and exposed to x-ray film (Fuji, Tokyo) for autoradiography.

Reverse Transcriptase-mediated Polymerase Chain Reaction (PCR)—Poly(A)+ RNA from brains of P7 rats was used for reverse transcriptase-mediated PCR. The first strand of cDNA was generated using diox-IDT primers and reverse transcriptase (Time saver; Pharmacia). Then, portions or of the entire region coding for the proteoglycan core protein was amplified by PCR using Taq DNA polymerase (Ampli Taq Perkin Elmer Corp., Norwalk, CT) and DNA primers. The DNA primers for PCR were probes with the sequences 5'-ACCGCTGGGGCTGTACCG-GCAC-3' (primer 1, nucleotides 91–112 in Fig. 3), 5'-ATCGAGGCT-GAGAGCTGGTT-3' (primer 2, nucleotides 130–149), 5'-TGGTG-GCAGCATCTCCTCTTA-3' (primer 3, nucleotides 1026–1045), 5'-TAAGAGAGATGCTGCCACCA-3' (primer 4, the complementary sequence corresponding to nucleotides 1026–1045), 5'-CAGTCACTGTTATTTCTGGAG-3' (primer 5, the complementary sequence corresponding to nucleotides 1027–1047), and 5'-GTACAAAGAGCGACAGAC-GGAACC-3' (primer 6, the complementary sequence corresponding to nucleotides 1,839–1,862). Primers 1 and 6, primers 2 and 4, primers 2 and 5, and primers 3 and 5 were paired, respectively. The mixture for PCR consisted of 50 mM Tris-HCl (pH 8.8), 50 mM KCl, 2.5 mM MgCl2, 0.2 mM each deoxyribonucleotide triphosphate, 200 nM each primer, 2 units of Taq DNA polymerase, and a heat-treated solution of the first strand of cDNA (0.5 μg of mRNA). Amplification was carried out for 30 cycles (94 °C, 1 min; 55 °C, 2 min; 72 °C, 3 min) in a programmable temperature-control system (F) 2000; Perkin Elmer).

Fig. 1. Western blot of a partially purified preparation of membrane-bound proteoglycans and a preparation of PBS-soluble proteoglycans from brains of 10-day-old rats. The preparation of partially purified membrane-bound proteoglycans (A1) was digested with chondroitinase ABC (A2), chondroitinase ABC plus heparitinase I (A3 and B1), neuraminidase (B2), O-glycansase (B3), or N-glycansase (B4). In B, samples were digested sequentially with these glycosidases. Samples (2 nmol of uronic acid for the preparation of intact membrane-bound proteoglycans (A1) and 0.2 nmol for others) were subjected to electrophoresis on a 6% polyacrylamide gel, blotted onto a PVDF membrane, and stained with mAb C5. The positions of molecular mass markers are indicated in kDa.

Fig. 2. Spatiotemporal patterns of expression of NGC in the rat cerebral cortex. A and B, a section of coronal tissue from the cerebral cortex of a 0- (A) and a 7- (B) day-old rat was immunostained with mAb C5. Immunopositive material associated with neurons can be seen throughout the cortex. Scale = 100 μm. C, developmental changes in the relative amounts of the core glycoprotein of the novel proteoglycan. The intensities of the immunolabeled bands of protein on a Western blot were quantified by densitometry. The levels are shown as percentages of the level on P20.
RESULTS AND DISCUSSION

Characterization of Monoclonal Antibodies—We established four hybridoma cell lines (designated C1, C3, C5, and C15), each of which produced an antibody that reacted with a preparation of membrane-bound proteoglycans on a Western blot. All of the four mAbs recognized a very diffuse band that corresponded to an average molecular mass of 150 kDa when the preparation of membrane-bound proteoglycans that had been used as immunogen was examined (Fig. 1A, immunoblot with mAb C5). When the preparation of membrane-bound proteoglycans was digested with chondroitinase ABC, the band of a 120-kDa immunopositive protein became detectable. Treatment with heparitinase did not change the mobility of this band. When a large amount of a preparation of PBS-soluble proteoglycans derived from brains of 10-day-old rats was used to prepare a Western blot, a 75-kDa core glycoprotein was weakly detected with mAb C5 (data not shown). mAb C5 did not stain any of the proteoglycan core glycoproteins prepared from extra-neuronal tissues, such as liver, lung, kidney, muscle, and cartilage (data not shown).

To examine the possibility that the proteoglycan might have oligosaccharide side chains in addition to chondroitin sulfate chains, the core glycoprotein was digested sequentially with three glycosidases, as shown in Fig. 1B, and the molecular mass of the deglycosylated core protein was estimated by SDS-PAGE. Each treatment with glycosidase resulted in an increase in the mobility of the band of the core protein on SDS-PAGE. The ability of mAb C5 to recognize the core protein was not affected by digestion with any of the glycosidases. Therefore, it appeared that the epitope recognized by mAb C5 was present on the polypeptide moiety.

The spatial expression of the proteoglycan in the rat cerebral cortex was examined by immunohistochemical staining with mAb C5. Fresh medium containing normal mouse IgG (5 mg/ml, Sigma) was used in place of mAb solutions for negative control of immunostaining. The control sections were virtually free of immunolabelings.

To quantify the temporal expression of the proteoglycan in the cerebral cortex, PBS-insoluble extracts of the cerebral cortex at various developmental stages (from embryonic day 12 (E12) to adulthood) were digested with chondroitinase ABC and processed for immunoblotting, and the intensity of the immunoblabeled band was quantified. A small amount of the proteoglycan was detected on E16 through E18, and the amount of the proteoglycan increased to reach a maximum...
level around P20 (Fig. 2C). After P20, the amount of the proteoglycan component decreased, and in the mature cerebral cortex this proteoglycan was expressed at approximately half of the peak level of expression. This biochemical data confirmed our immunohistochemical data.

Isolation of cDNA Clones—We purified the 120-kDa core glycoprotein, cleaved it into several peptides by treatment with CNBr, and determined the partial N-terminal amino acid sequence of the intact core glycoprotein and of two peptides. The N-terminal sequence of the intact core glycoprotein was VPAREAGSAIEAEEL. Entirely different N-terminal sequences were found in the two products of degradation by CNBr (15 and 24 kDa), (M)VPGGSISLRPRPGDKLA and N-terminal sequence of the intact core glycoprotein was determined by amino acid analysis. This arrangement meets the criteria for a transmembrane domain.

One positive clone (lNGC1) from the λgt11 cDNA library (random primers) derived from a rat brain on P8 and another positive clone (lNGC3) from the λgt11 cDNA library (oligo(dT) primers) derived from a rat brain on P8 were initially isolated by immunoscreening with a mixture of the four mAbs, C1, C3, C5, and C15. lNGC1 and lNGC3 contained inserts of 682 bp (nucleotides 108–790 in Fig. 3) and 1,398 bp (nucleotides 710–2,107), respectively. The fusion protein with β-galactosidase from lNGC1 reacted with only mAb C1 and not with mAbs C3, C5, and C15. By contrast, the fusion protein from lNGC3 reacted with mAbs C3, C5, and C15 but not with mAb C1. The inserts of lNGC1 and lNGC3 were subcloned into the pBluescript II plasmid vector for further analyses (to yield pNGC1 and pNGC3, respectively). The authenticity of these clones was unequivocally established by identification of an amino acid sequence from the 120-kDa core glycoprotein and the 24-kDa CNBr-generated fragment within the amino acid sequence deduced from the insert in pNGC1 and by that from the 15-kDa CNBr fragment within the amino acid sequence deduced from the insert in pNGC3 (see Fig. 3, underlining).

The insert in pNGC1 was labeled with 32P dCTP and used as a probe to identify overlapping cDNA clones. Four overlapping cDNA clones (lNGC7 (nucleotides 339–1, 219 in Fig. 3), lNGC13 (nucleotides 694; 1.6 kb), lNGC15 (nucleotides 205–683), and lNGC19 (nucleotides 752; 2.2 kb)), obtained from the λgt11 cDNA library (random primers), were covered the entire coding region of the proteoglycan, which we designated as NGC.

With the same probe, mRNA for NGC was detected by Northern blot analysis (Fig. 4). A single transcript of 3.1 kb was detected in analyses of brains of 7-day-old and adult rats. This transcript was not detected in analyses of kidney, liver, lung, and muscle.

Predicted Structure of the Core Protein of NGC—The complete cDNA sequence encoding the core protein of NGC and the deduced amino acid sequence are shown in Fig. 3. The clones covered more than 2,107 bp, which included an open reading frame that encoded 544 amino acids. The calculated molecular mass was 58,612 Da. The nucleotide sequence around the ATG triplet at position 13 corresponded to the consensus sequence for an initiation site of translation (32). The 5′-region of 12 bp of the cDNA contributes to the 5′-untranslated region. There is a stretch of hydrophobic amino acids from the initial methionine to the first amino acid residue (valine) of the mature core protein, and it is probably a signal sequence. Hydropathy analysis of the predicted protein by the method of Kyte and Doolittle (33) revealed a second hydrophobic segment near the C terminus, composed of 24 amino acids (residues 426–450), which was followed by two basic residues (Fig. 3, dashed line). This arrangement meets the criteria for a transmembrane domain proposed by Sabatini et al. (34). The 3′-untranslated region consists of 463 nucleotides and included a polyadenylation signal (AAUAAA).

The size of the open reading frame was confirmed by reverse transcriptase-mediated PCR. All of the amplified DNA fragments using multiple pairs of primers (see “Experimental Procedures”) had reasonable sizes, which are predicted from the cDNA sequence. Primers 1 and 2 used for reverse transcriptase-mediated PCR are positioned on the N-terminal sequence of the intact core glycoprotein determined by amino acid sequence analysis. In addition, primer 5 is positioned on the predicted C-terminal sequence. Therefore, the size of the open reading frame shown in Fig. 3 can be considered to be appropriate. After removal of oligosaccharide side chains, however, the estimated molecular mass of the core protein was 100 kDa (Fig. 1), which is still considerably larger than the calculated molecular mass (55.8 kDa) of the mature core protein with 514 amino acid residues encoded by the cloned cDNA. This discrepancy might be attributable to the anomalously slow electrophoretic migration of glycosylated proteins that results from decreased binding of SDS (35), since after digestion by the glycosidases some oligosaccharides might still remain on the core protein. Differences between actual and apparent molecular masses have been reported for the core proteins of other proteoglycans, such as syndecan (36), versican (18, 37), and neurocan (16).

The core protein of NGC was rich in glycine (10.5%), leucine (9.9%), proline (9.0%), glutamic acid (8.1%), serine (7.9%), and threonine (7.7%) residues. The calculated pI is 4.9. We found a total of 10 cysteine residues in the core protein. All of them were localized in the central domain, just outside the transmembrane domain. The predicted extracellular domain of the NGC core protein contained three potential sites of N-glycosylation (38) and three serine-threonine clusters (residues 143-
144, 188–189, and 271–272) that could serve as acceptors for O-linked carbohydrates (39).

The N-terminal region, from amino acid residues 31–281, contained eight serine-glycine (SG) or glycine-serine (GS) dipeptide sequences. These dipeptides have been proposed to be core portions of the consensus sequence of attachment sites for chondroitin sulfate (18, 40). The amino acid sequences around the eight serine residues differed from the consensus sequences for attachment sites for glycosaminoglycans, which are SGXG (40) and (E/D)GSG(E/D) (18). However, these consensus sequences are not the only ones that allow attachment of glycosaminoglycans to core proteins. In neurocan, for example, the sequence around a serine residue to which a chondroitin sulfate chain is attached is EEVASGQED (16). The alignment of amino acids characteristic of attachment sites for chondroitin sulfate in reported proteoglycans, such as aggrecan, versican, syndecan, decorin, and collagen type IX, is SG/GS with preceding and following acidic amino acid residues. The importance of acidic amino acid residues in the attachment of chondroitin sulfate has also been noted elsewhere (18, 40, 41). The eight SG/GS dipeptides of the NGC core protein were arrayed in sequence with this consensus-type arrangement. NGC contained two additional dipeptide sequences (serine residues at positions 341 and 374) in its putative extracellular domain. However, these dipeptides were not associated with this consensus-type organization. Therefore, the N-terminal domain is putatively defined as the chondroitin sulfate attachment domain.

A short stretch of basic amino acids KRRKRRRRIR (residues 282–291) is found (double underline in Fig. 3) adjacent to this putative chondroitin sulfate-attachment domain. The basic amino acid residues could contribute to the proteolytic processing of this proteoglycan since they correspond to sites of cleavage by serine proteases that might cooperate in the regulation of neurite outgrowth (42–47). The soluble type of NGC might be such a proteolytic product since NGC mRNA was detected as a single 3.1-kb transcript on Northern blots of mRNAs from the rat brain (Fig. 4). Moreover, we reported previously the proteolytic processing of another chondroitin sulfate proteoglycan, neurocan (48).

The cytoplasmic domain of NGC contained threonine and serine residues that could be potential sites of phosphorylation by protein kinases. The amino acid sequences around the threonine residue at position 465 (KLRRTNK) and around the serine residue at position 521 (SPK) are similar to the consensus sequence for the sites of phosphorylation by protein kinase C (XRKX(S/T)X(R/K)) proposed by Graff et al. (49). Although it remains to be determined whether NGC is phosphorylated, it is possible that NGC might be involved in signal transduction. The putative cytoplasmic domain contains two tyrosine residues, but the sequences adjacent to these tyrosine residues differ from the reported consensus sequences for tyrosine phosphorylation (50). Fig. 5 is a schematic representation of the proposed domain structure of the NGC core protein.

A search of data base analyses at the amino acid level indicated that small portions of NGC are homologous to regions of the core protein of aggrecan, a large chondroitin sulfate proteoglycan of cartilage (51). All of the homologous fragments are distributed in the putative extracellular domain of NGC. In the putative chondroitin sulfate-attachment domain, we identified eleven homologous clusters: residues 45–94 of NGC and residues 907–956 of aggrecan (32% identity/48% chemical similarity); 26–70 and 967–1011 (29%/56%), 53–103 and 955–1005 (22%/47%), 88–113 and 1585–1610 (42%/54%), 117–151 and 1649–1683 (31%/40%), 43–92 and 1130–1179 (28%/38%), 50–94 and 892–936 (27%/47%), 59–108 and 1233–1282 (22%/44%), 122–163 and 1296–1337 (29%/36%), 109–153 and 1705–1749 (31%/40%), and 90–114 and 1108–1132 (32%/56%). All of the corresponding homologous clusters in aggrecan are localized in the putative chondroitin sulfate-attachment domain. In the cysteine-containing domain of NGC, we found three homologous clusters: 287–322 and 1827–1862 (31% identity/42% chemical similarity), 348–386 and 1898–1936 (23%/38%), and 401–411 and 1968–2008 (55%/55%). All of the corresponding homologous clusters in aggrecan are also localized in one of the cysteine-containing domains of aggrecan. A schematic alignment of homologous sequences in rat aggrecan and NGC is shown in Fig. 6. Aggrecan consists of several structurally different domains, namely an immunoglobulin-like domain and tandem repeats of a hyaluronic acid-binding region in the N-terminal portion, a chondroitin sulfate-attachment domain in the central portion, and epidermal growth factor-like domains, a lectin-like domain, and a complement regulatory protein-like domain in the C-terminal portion (51). Several other chondroitin sulfate proteoglycans, such as versican (18), neurocan (16), brevican (17), and BEHAB (19), also include these domains. Therefore, they are all considered to be members of the aggrecan family. NGC is not a member of this family since the NGC core protein does not have any of these domains except for a chondroitin sulfate-attachment domain.

Acknowledgments—We are very grateful to Drs. S. Hirano, H. Hamanaka, T. Shintani, T. Nishiwaki, and Y. Yamada for helpful advice and to Dr. M. Mizoguchi for expert assistance with DNA sequencing.

REFERENCES
1. Ruoslahti, E. (1989) J. Biol. Chem. 264, 13369–13372
2. Ruoslahti, E., and Yamaguchi, Y. (1991) Cell 64, 867–869
3. Oshira, A., Matsui, F., Matsuura, M., Takida, Y., and Kuboki, Y. (1988) J. Biol.

Fig. 6. Schematic alignment of amino acid sequences in rat aggrecan (52) with those in NGC. Sequences were compared with a local alignment search tool (30). Central positions of homologous clusters in rat aggrecan and in NGC are connected by lines. CS, CS1, CS2, and CS3, chondroitin sulfate-attachment domains; H1 and H2, hyaluronic acid-binding domains; KS, a keratan sulfate-attachment domain; GELC, a globular domain that includes epidermal growth factor-like domains, a lectin-like domain, and a complement regulatory protein-like domain; B, a stretch of basic amino acid residues; G, a globular domain; T, a transmembrane region; CP, a cytoplasmic domain. A scale in amino acid residues is shown.
A Novel Proteoglycan of the Brain

4. Herndon, M. E., and Lander, A. D. (1990) Neuron 4, 949–961
5. Maeda, N., Matsui, F., and Oohira, A. (1992) Dev. Biol. 151, 564–574
6. Sur, M., Frost, D. O., and Hockfield, S. (1988) J. Neurosci. 8, 874–882
7. Crossin, K. L., Hoffman, S., Tan, S.-S., and Edelman, G. M. (1989) Dev. Biol. 136, 381–392
8. Oohira, A., Matsui, F., Watanabe, E., Kushima, Y., and Maeda, N. (1994) Neuroscience 60, 145–157
9. Watanabe, E., Aono, S., Matsui, F., Yamada, Y., Naruse, I., and Oohira, A. (1995) Eur. J. Neurosci. 7, 547–554
10. Lander, A. D., Fujii, D. K., and Reichardt, L. F. (1985) Proc. Natl. Acad. Sci. U. S. A. 82, 2183–2187
11. Riopelle, R. J., and Dow, K. E. (1990) Neuron 7, 1007–1018
12. Cole, G. J., and McCabe, C. F. (1991) J. Biol. Chem. 266, 547–554
13. Oohira, A., Matsui, F., Watanabe, E., Kushima, Y., and Maeda, N. (1994) Neuroscience 60, 145–157
14. Grumet, M., Flaccus, A., and Margolis, R. U. (1993) J. Cell Biol. 120, 813–824
15. Dou, C.-L., and Levine, J. M. (1994) J. Neurosci. 14, 7616–7628
16. Rauch, U., Karshtikayan, L., Maurel, P., Margolis, R. U., and Margolis, R. K. (1992) J. Biol. Chem. 267, 19536–19547
17. Yamada, H., Watanabe, K., Shimonaka, M., and Yamaguchi, Y. (1994) J. Biol. Chem. 269, 10119–10126
18. Zimmermann, D. R., and Ruoslahti, E. (1989) J. Biol. Chem. 264, 13224–13232
19. Hansen, M. E., and Lander, A. D. (1990) Proc. Natl. Acad. Sci. U. S. A. 87, 5466–5470
20. Hawkins, R. L., and Seeds, N. W. (1986) J. Biol. Chem. 261, 1365–1372
21. Maeda, N., Hamanaka, H., Shintani, T., Nishikawa, T., and Noda, M. (1994) FEBS Lett. 354, 67–70
22. Maeda, N., Hamanaka, H., Shintani, T., Nishikawa, T., and Noda, M. (1994) FEBS Lett. 354, 67–70
23. Nishiyama, A., Dahlin, K. J., Prince, J. T., Johnstone, S. R., and Stallcup, W. B. (1991) J. Cell Biol. 114, 359–371
24. Karshtikayan, L., Maurel, P., Rauch, U., Margolis, R. K., and Margolis, R. U. (1994) J. Cell Biol. 124, 149–160
25. Bitter, T., and Muir, H. E. (1962) J. Mol. Biol. 4, 330–334
26. Watanabe, E., Fujita, S. C., Murakami, F., Hayashi, M., and Matsumura, M. (1989) Neuroscience 29, 645–657
27. Scott, M. G., Crimmings, D. L., McCourt, D. W., Tarrand, J. J., Eyerman, M. C., and Nahm, M. H. (1988) Biochem. Biophys. Res. Commun. 155, 1353–1359
28. Sambrook, J., Fritsch, E. F., and Maniatis, T. (1989) Molecular Cloning: A Laboratory Manual 2nd Ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY
29. Altschul, S. F., Gish, W., Miller, W., Myers, E. W., and Lipman, D. J. (1990) J. Mol. Biol. 215, 403–410
30. Chomczynski, P., and Sacchi, N. (1989) Anal. Biochem. 162, 156–159
31. Monard, D. (1988) Trends Neurosci. 11, 541–544
32. Pittman, R. N., Ivins, J. K., and Buetner, H. M. (1989) J. Neurosci. 9, 4269–4280
33. Monard, D. (1988) Trends Neurosci. 11, 541–544
34. Monard, D., Niday, E., Limat, A., and Solomon, F. (1983) Proc. Brain Res. 58, 359–364
35. Monard, D. (1988) Trends Neurosci. 11, 541–544
36. Monard, D. (1988) Trends Neurosci. 11, 541–544
37. Bourdon, M. A., Krusius, T., Campbell, S., Schwartz, N. B., and Ruoslahti, E. (1987) Proc. Natl. Acad. Sci. U. S. A. 84, 3194–3198
38. Krueger, R. C., Jr., Fields, T. A., Hildreth, J., Jr., and Schwartz, N. B. (1990) J. Biol. Chem. 265, 12075–12087
39. Hawkins, R. L., and Seeds, N. W. (1986) Brain Res. 396, 63–70
40. Jalink, K., and Moenkena, W. H. (1992) J. Cell Biol. 118, 411–419
41. Saunders, S., Jalkanen, M., O’Farrell, S., and Bernfield, M. (1989) J. Biol. Chem. 264, 13224–13232
42. Suiden, H. S., Stone, S. R., Hemmings, B. A., and Monard, D. (1992) Neuron 9, 363–375
43. Matsui, F., Watanabe, E., and Oohira, A. (1994) Neurochem. Int. 25, 425–431
44. Graff, J. M., Stumpo, D. J., and Blackshear, P. J. (1989) J. Biol. Chem. 264, 11212–11191
45. Karshtikayan, L., Maurel, P., Rauch, U., Margolis, R. K., and Margolis, R. U. (1992) Biochem. Biophys. Res. Commun. 188, 395–401
46. Karshtikayan, L., Maurel, P., Rauch, U., Margolis, R. K., and Margolis, R. U. (1992) Biochem. Biophys. Res. Commun. 188, 395–401
47. Karshtikayan, L., Maurel, P., Rauch, U., Margolis, R. K., and Margolis, R. U. (1992) Biochem. Biophys. Res. Commun. 188, 395–401
48. Karshtikayan, L., Maurel, P., Rauch, U., Margolis, R. K., and Margolis, R. U. (1992) Biochem. Biophys. Res. Commun. 188, 395–401
49. Karshtikayan, L., Maurel, P., Rauch, U., Margolis, R. K., and Margolis, R. U. (1992) Biochem. Biophys. Res. Commun. 188, 395–401
50. Karshtikayan, L., Maurel, P., Rauch, U., Margolis, R. K., and Margolis, R. U. (1992) Biochem. Biophys. Res. Commun. 188, 395–401
51. Karshtikayan, L., Maurel, P., Rauch, U., Margolis, R. K., and Margolis, R. U. (1992) Biochem. Biophys. Res. Commun. 188, 395–401
52. Karshtikayan, L., Maurel, P., Rauch, U., Margolis, R. K., and Margolis, R. U. (1992) Biochem. Biophys. Res. Commun. 188, 395–401