Vol. 188, No. 1, 1992 BIOCHEMICAL AND BIOPHYSICAL RESEARCH COMMUNICATIONS
October 15, 1992 Pages 395-401

CLONING OF A MAJOR HEPARAN SULFATE PROTEOGLYCAN FROM BRAIN AND IDENTIFICATION AS THE RAT FORM OF GLYPICAN+

L. Karthikeyan, P. Maurel, U. Rauch, R.K. Margolis, and R.U. Margolis*

Department of Pharmacology, New York University Medical Center, New York, NY 10016, and Department of Pharmacology, State University of New York, Health Science Center, Brooklyn, NY 11203

Received August 31, 1992

We have obtained the complete coding sequence of a highly conserved heparan sulfate proteoglycan which we previously characterized biochemically after isolation from rat brain. An open reading frame of 558 amino acids encodes a protein with a molecular mass of 62 kDa containing three peptide sequences present in the isolated proteoglycan. The total sequence obtained is 3.5 kb long, including 1.6 kb of 3'-untranslated sequence and 0.2 kb of 5'-untranslated sequence. The deduced amino acid sequence and the 3'- and 5'-untranslated sequences have 89% and 66-80% identity, respectively, with those of a phosphatidylinositol-anchored human lung fibroblast heparan sulfate proteoglycan (glypican) for which mRNA is detectable in a large number of human cell lines. Our data therefore demonstrate that this major heparan sulfate proteoglycan of brain is the rat form of glypican.

We previously described the isolation and characterization of a major membrane-associated heparan sulfate proteoglycan from rat brain, which was very similar in many of its properties (e.g., size, charge, density) to the population of more abundant chondroitin sulfate proteoglycans, from which it was separated by affinity chromatography on lipoprotein lipase-Sepharose (1). This ~220 kDa proteoglycan contains a 55 kDa core glycoprotein, ~15 kDa heparan sulfate chains, and predominantly tri- and tetraantennary *N*-glycosidic as well as *O*-glycosidic oligosaccharides (1,2). We have also described some of the structural properties of its heparan sulfate chains (2). N-terminal amino acid sequences of the core glycoprotein and of tryptic peptides derived from it demonstrated a high degree of identity with sequences originally reported for a glycosyl phosphatidylinositol (GPI)-anchored heparan sulfate proteoglycan (glypican) of human lung fibroblasts (3), and which was recently shown by Northern blot analysis to also be present in a large number of human cell lines including glioma and neuroblastoma cells (4). Using an RNA probe

+Sequence data from this article have been deposited with the EMBL/GenBank Data Libraries under Accession No. L02896.

*To whom correspondence should be addressed.

0006-291X/92 $4.00

to screen a rat brain cDNA library, we have now determined the primary structure of this heparan sulfate proteoglycan from rat brain.

EXPERIMENTAL PROCEDURES

Preparation of peptides and amino acid sequence analysis. The N-terminal sequence of the heparan sulfate proteoglycan core glycoprotein was determined using the 55 kDa bands transferred to a ProBlott membrane after SDS-PAGE of heparitinase-treated proteoglycan. Heparitinase digestions were performed for 5.5 h at 37°C in 0.1 M Tris-HCl buffer, pH 7.2, using a ratio of 2.8 milli international units of heparitinase (EC 4.2.2.8, Seikagaku America, Inc., Rockville, MD) per 100 μg proteoglycan protein.

Electrophoresis was performed on a 10% gel as described by Laemmli (5), with the modifications that the gel was pre-electrophoresed for 30 min at 100 volts (30-40 mA) at 4°C using stacking gel buffer containing 0.1 mM thioglycolic acid in the upper buffer chamber and sample wells (after which both buffers were changed), and 0.1 mM thioglycolic acid was added to the upper buffer for electrophoresis of the samples. Separated proteins were transferred to ProBlott membranes (Applied BioSystems, Foster City, CA) for 30 min at room temperature in 10 mM CAPS buffer (pH 11) containing 10% methanol. Membranes containing transferred proteins were washed several times with water followed by a few seconds in methanol, stained for 1 min in 50% methanol containing 1% acetic acid and 0.1% Coomassie Blue, destained in 50% methanol, washed with water, and air-dried. Transferred protein bands were excised from the membrane sheet and sequenced on an Applied Biosystems Model 475A instrument equipped with the Blott cartridge. To obtain internal amino acid sequence data, the 55 kDa core protein was blotted to nitrocellulose, the excised bands were blocked with PVP-40 and digested with trypsin, and the resulting tryptic peptides were separated by HPLC and used for microsequencing (6).

Amplification of cDNA, cDNA library screening, and DNA sequencing. Based on the N-terminal amino acid sequence and an internal peptide sequence determined for the 55 kDa core glycoprotein of the heparan sulfate proteoglycan, sense and anti-sense oligonucleotide primers were synthesized with *Xba* I and *Eco*R I linkers. Deoxyinosine was substituted in positions where the codon degeneracy is >2, and mixed pairs of bases were used at positions where there are only two triplet codons. We initially used 20 cycles of the PCR to amplify all inserts in a 6-week rat brain λZAPII cDNA library (Stratagene), employing the Bluescript forward and reverse (BSK and BKS) primers to which we added 6 bases in each case. Aliquots of this product and of the original cDNA library were then tested using the heparan sulfate proteoglycan-specific primers for further amplification with *Taq* polymerase, using 40 temperature-step cycles of 94°C (0.5 min), 55°C (1.5 min), and 72°C (with the extension time increasing from 2 to 6 min). Agarose-ethidium bromide gel electrophoresis of the PCR reaction products demonstrated a 1.2 kbp band which was not produced from a control reaction containing only empty λZAPII vector or single primers.

The PCR product was treated sequentially with polynucleotide kinase and the Klenow fragment of DNA polymerase I, purified by agarose gel electrophoresis and Qiaex extraction (Qiagen Inc., Chatsworth, CA), and subcloned into the *Sma* I site of pGEM-7Zf+ (Promega) for further analysis of recombinant plasmids amplified in *E. coli* XL1-Blue cells. The pGEM-7Z/PCR product plasmid was then cut with *Sma* I, and used to generate a 344 base RNA sense probe for screening the original rat brain λZAPII cDNA library. Bluescript clones were tested by the PCR to determine orientation and insert size. DNA sequencing was performed on restriction fragments subcloned into pGEM-7 and with synthetic primers corresponding to Bluescript vector

sequences or to the ends of previously determined sequences, using the Applied Biosystems Model 373A DNA sequencing system with *Taq* polymerase and dye-labeled terminators. Both strands of the coding region of the cloned cDNA were sequenced, with sequence alignment and analysis accomplished with the software package from the Genetics Computer Group (Madison, Wisconsin). The reading frame was verified by our N-terminal amino acid sequence data for the 55 kDa core glycoprotein and for tryptic peptides derived from it.

RESULTS AND DISCUSSION

Degenerate oligonucleotide primers based on the amino acid sequences of the heparan sulfate proteoglycan (Fig. 1A) were used for PCR amplification of rat brain cDNA. A protein database search revealed identities of our peptide sequences from the brain proteoglycan with published sequences present in a human lung fibroblast proteoglycan (3), and indicated that the sequence of our tryptic peptide 31 began immediately after the last sequenced amino acid (lysine) in the N-terminus of the proteoglycan core protein. The N-terminal glycine of tryptic peptide 31 was therefore used together with the last seven sequenced amino acids of the proteoglycan core protein for design of a PCR sense primer (Fig. 1B).

The 1.2 kb PCR product was subcloned into pGEM for sequencing, and after linearization was also used as template for the *in vitro* transcription of an RNA probe which was used to screen a rat brain λZAPII cDNA library. Nineteen out of 37 positive plaques were tested by PCR, and four of these were selected for secondary screening based on their ability to serve as templates for PCR products of the appropriate size. DNA was prepared after *in vivo* excision of Bluescript clones and the largest of these, with an insert size of 3.5 kb, was then sequenced (Fig. 2).

An open reading frame of 558 amino acids encodes a protein with a molecular mass of 61,734 Da and contains the three peptide sequences found in the isolated

A

N-terminal sequence: DPA-K-R---EVRQIYGAK

Tryptic peptide 31: GFSLSDVPQAEIS

Tryptic peptide 45: YLPEVMGDGLANQINNPEVEVD

B

Sense primer (from N-terminal sequence):

5' GTTCTAGAATTC(A/C)GICA(A/G)ATITA(C/T)GGIGCIAA(A/G)GG 3'

Antisense primer (from tryptic peptide 45):

5' AT(G/A)TCGAATTCIAC(C/T)TCIGG(G/A)TT(G/A)TTIAT(C/T)TGG 3'

Fig. 1. **A.** N-terminal amino acid sequence of the heparan sulfate proteoglycan and sequences of two tryptic peptides derived from it. Underlined sections of the peptide sequences indicate regions used for the design of oligonucleotide primers. **B.** Sequences of synthetic oligonucleotide PCR primers prepared on the basis of the peptide sequences shown above.

proteoglycan. The sequenced portion of this clone includes 1.6 kb of 3'-untranslated sequence and 0.2 kb of 5'-untranslated sequence. The deduced amino acid sequence and the 3'- and 5'-untranslated sequences have 89% and 66-80% identity, respectively, to those reported for the GPI-anchored human lung fibroblast heparan sulfate proteoglycan, glypican (3). The sequence of the rat brain proteoglycan contains two potential *N*-glycosylation sites and two potential threonine *O*-glycosylation sites (7). There is no obvious transmembrane domain, and it is therefore likely that this heparan sulfate proteoglycan is also GPI-anchored to the cell surface. Herndon and Lander (8) have previously reported data indicating that two heparan sulfate proteoglycans of rat brain, with core proteins of 50 and 59 kDa, are linked to the membrane by GPI anchors. Removal of the 23 amino acid signal peptide and the 28 C-terminal amino acids at the site of attachment of the phosphatidylinositol tail (3) would yield a core protein with a molecular size of 55,890, which is in good agreement with the apparent molecular size of the proteoglycan core protein (~55 kDa) on SDS-PAGE (2). There are seven serine-glycine sequences which could serve as potential glycosaminoglycan attachment sites, but only serine-55, -486, -488, and -490 fit the "consensus sequence" of a serine-glycine dipeptide preceded by at least one acidic amino acid (9,10). A glycine-serine sequence at serine-315 preceded by a glutamic acid could represent a fifth glycosaminoglycan attachment site. The probable maximum of five ~15 kDa heparan sulfate chains (1) which could be accommodated in this core protein together with some *N*- and *O*-glycosidic oligosaccharides would appear to yield a proteoglycan with a mass somewhat less than the estimate of ~220 kDa determined by gel filtration under dissociative conditions (1). This figure is close to the size of >200 kDa estimated for human lung fibroblast glypican (11), which however contains much larger heparan sulfate chains with a size close to 40 kDa (12).

Glypican is the only major heparan sulfate proteoglycan of adult brain to have been cloned. A partial cDNA sequence has been reported (13) for a transmembrane heparan sulfate proteoglycan (named *N*-syndecan) with 62-72% amino acid identity with the transmembrane and cytoplasmic domains of murine syndecan (14) and of fibroglycan, a human lung fibroblast heparan sulfate proteoglycan (15), but *N*-syndecan has a unique extracellular sequence. High levels of message for *N*-syndecan were found in neonatal rat brain, heart, and Schwann cells, whereas its mRNA was barely detectable in neonatal or adult liver, or in adult brain. Rat brain mRNA has also been

---

Fig. 2. Nucleotide sequence and deduced amino acid sequence from the heparan sulfate proteoglycan core protein cDNA. The cDNA sequence is shown together with the translation of an open reading frame of 558 residues. Potential *N*-glycosylation sites are indicated with filled triangles and potential threonine *O*-glycosylation sites with open triangles. Serine-glycine sequences representing the most likely glycosaminoglycan attachment sites are double underlined. Peptides from which amino acid sequence data were obtained are underlined, and amino acid sequences used for the design of oligonucleotide primers for enzymatic amplification reactions are boxed. The probable site of attachment of the phosphatidylinositol tail is indicated by an arrow, and the AATAAA polyadenylation signal is underlined.

```
                                                    CGCCCGCCTGAGTGGCCGGACCTCGCGCCTCGCGCCTCGCG   41
TCCCGCGCTGCAGCCGCAGTCGGCTTTTGTTGTCTCCGCCTCCTCGTCGGTCCCCGACTCTGGACCACGAGCGGCGCGCGCTGGGACCTT  131
GGCTCTGCCCTTCGCGCAGACTGGGCTGAGCGGGCCGGGCGGCCATCCAGGGAGGCGCGCACGGGCGGTCGGGGAGCCGCGGGCCCCGCC  221

ATGGAGCTCCGGGCCCGAGGCTGGTGGCTGCTGTGCGCGGCCGCCGCGCTAGTCGCCTGCACCCGCGGGGACCCCGCCAGCAAGAGCCGG  311
M  E  L  R  A  R  G  W  W  L  L  C  A  A  A  A  L  V  A  C  T  R  G  D  P  A  S  K  S  R    30

AGCTGCAGCGAAGTCCGCCAGATCTACGGGGCTAAGGGCTTTAGCCTGAGCGACGTGCCCCAGGCAGAGATCTCGGGAGAGCACCTGCGG  401
S  C  S  E  V  R  Q  I  Y  G  A  K  G  F  S  L  S  D  V  P  Q  A  E  I  S  G  E  H  L  R    60

ATCTGCCCCCAGGGCTACACCTGCTGCACCAGTGAGATGGAGGAGAACCTGGCCAACCACAGCCGGATGGAGCTGGAGACCGCACTCCAC  491
I  C  P  Q  G  Y  T  C  C  T  S  E  M  E  E  N  L  A  N  H  S  R  M  E  L  E  T  A  L  H    90
                                                    ▲
GACAGCAGCCGTGCCCTGCAGGCTACACTGGCCACCCAGCTGCATGGTATCGATGACCACTTCCAGCGCCTGCTGAATGACTCGGAGCGT  581
D  S  S  R  A  L  Q  A  T  L  A  T  Q  L  H  G  I  D  D  H  F  Q  R  L  L  N  D  S  E  R   120
                                                          ▲
ACACTGCAGGATGCTTTTCCCGGGGCCTTTGGGGACCTGTACACGCAGAACACTCGGGCCTTCCGGGACCTGTATGCTGAGCTGCGTCTC  671
T  L  Q  D  A  F  P  G  A  F  G  D  L  Y  T  Q  N  T  R  A  F  R  D  L  Y  A  E  L  R  L   150

TACTACCGAGGGGCCAACCTACACCTTGAGGAGACACTGGCCGAGTTCTGGGCACGGCTGCTGGAGCGTCTCTTCAAGCAGCTGCACCCC  761
Y  Y  R  G  A  N  L  H  L  E  E  T  L  A  E  F  W  A  R  L  L  E  R  L  F  K  Q  L  H  P   180

CAGCTGCTGCTGCCCGATGACTATCTGGACTGCCTGGGCAAGCAGGCGGAGGCACTGCGGCCGTTTGGGGATGCCCCTCGAGAACTGCGC  851
Q  L  L  L  P  D  D  Y  L  D  C  L  G  K  Q  A  E  A  L  R  P  F  G  D  A  P  R  E  L  R   210

CTGAGGGCCACCCGTGCTTTTGTGGCGGCACGATCCTTTGTGCAGGGCCTGGGTGTGGCCAGTGACGTAGTCCGAAAGGTGGCCCAGGTT  941
L  R  A  T  R  A  F  V  A  A  R  S  F  V  Q  G  L  G  V  A  S  D  V  V  R  K  V  A  Q  V   240

CCTCTGGCCCCAGAATGTTCTCGGGCTGTCATGAAGTTGGTCTACTGTGCCCCATTGCCGGGGAGTCCCTGGTGCCCGGCCCTGTCCCGAC 1031
P  L  A  P  E  C  S  R  A  V  M  K  L  V  Y  C  A  H  C  R  G  V  P  G  A  R  P  C  P  D   270

TATTGCCGAAATGTGCTCAAAGGCTGCCTTGCCAACCAGGCCGACCTGGATGCCGAGTGGAGGAACCTCCTGGACTCCATGGTGCTCATC 1121
Y  C  R  N  V  L  K  G  C  L  A  N  Q  A  D  L  D  A  E  W  R  N  L  L  D  S  M  V  L  I   300

ACTGACAAGTTCTGGGGCCCGTCGGGTGCGGAGTATGTCATTGGCAGTGTGCATATGTGGCTGGCGGAGGCCATCAACGCCCTCCAGGAC 1211
T  D  K  F  W  G  P  S  G  A  E  Y  V  I  G  S  V  H  M  W  L  A  E  A  I  N  A  L  Q  D   330

AACAAGGACACACTCACAGCTAAGGTCATCCAGGGCTGCGGAAACCCCAAGGTCAATCCCCATGGCTCTGGGCCTGAGGAGAAGCGTCGC 1301
N  K  D  T  L  T  A  K  V  I  Q  G  C  G  N  P  K  V  N  P  H  G  S  G  P  E  E  K  R  R   360

CGTGCCAAACTGGCACTGCAGGAGAAGTCCTCCACAGGTACTCTGGAAAAGCTGGTCTCTGAGGCCAAGGCCCAGCTCCGAGACATTCAG 1391
R  A  K  L  A  L  Q  E  K  S  S  T  G  T  L  E  K  L  V  S  E  A  K  A  Q  L  R  D  I  Q   390

GACTACTGGATCAGCCTCCCAGGGACACTGTGTAGTGAGAAGATGGCCATGAGTCCTGCCAGCGATGACCGCTGCTGGAATGGGATTTCC 1481
D  Y  W  I  S  L  P  G  T  L  C  S  E  K  M  A  M  S  P  A  S  D  D  R  C  W  N  G  I  S   420
                           Δ
AAGGGCCGGTACCTACCTGAGGTGATGGGTGATGGGCTGGCCAACCAGATCAACAACCCTGAAGTGGAGGTGGACATCACCAAGCCGGAT 1571
K  G  R  Y  L  P  E  V  M  G  D  G  L  A  N  Q  I  N  N  P  E  V  E  V  D  I  T  K  P  D   450

ATGACCATCCGGCAGCAGATCATGCAGCTCAAGATCATGACCAACCGTTTACGTGGCGCCTACGGTGGAAATGATGTGGACTTCCAGGAT 1661
M  T  I  R  Q  Q  I  M  Q  L  K  I  M  T  N  R  L  R  G  A  Y  G  G  N  D  V  D  F  Q  D   480

GCCAGTGATGACGGTAGTGGTTCCGGCAGCGGTGGCGGATGCCCAGATGACGCCTGTGGTCGGAGGGTCAGCAAGAAGAGCTCCAGCTCC 1751
A  S  D  D  G  S  G  S  G  S  G  G  G  C  P  D  D  A  C  G  R  R  V  S  K  K  S  S  S  S   510

CGGACCCCCTTGATCCATGCCCTCCCCGGCTTGTCAGAACAGGAGGGACAGAAGACCTCGGCCGCCACTCGCCCAGAGCCTCACTACTTC 1841
R  T  P  L  I  H  A  L  P  G  L  S  E  Q  E  G  Q  K  T  S  A  A  T  R  P  E  P  H  Y  F   540
   Δ                                                        ↑
TTTCTGCTCTTCCTGTTCACCTTGGTCCTTGCTGCAGCCAGGCCCAGGTGGCGGTAACTGCCCCCTAGCCCAAAGGACTGTCTTGGCCAA 1931
F  L  L  F  L  F  T  L  V  L  A  A  A  R  P  R  W  R  *                                    558

AACATGCAACAGACCATATTTACTTCCCTTGGCCTTCGAGGCCCAGGGCAGGATAAGGAGACAGTAGCTCTGAGTGCTGGGGCAGGGCGC 2021
ATGGGGTCCTGGCCTTCCTGGGTCTGGCCACGCCTGTCACCCTAGCTTCTAGTTGTTGTATCAGGTCAGCTGCGAGCCAGTGTCCCCAAA 2111
AGCCATGTTTTTCAGGGACCTCAGGGGCACCTCTGGCTGCACACTCCTCCCCTACCCTCCTGTACCACCCCAGAAGCTCACGAGGTCACC 2201
CAGAAGGGCGGTTATTAGCTACAACCCATCGGAGACCTCAAGTGAGTCTGTGTCTTCCTCTCCTGCCTCTTCCCGGTGGGGACTCCCCAC 2291
CAGACCCCATGGGACACAGATGTCAGAAATTGAGGCCCATCCCGCAGCTCCCCAGGAAGCCTGGAAGGGATGCCAGTATGTCGCTGACCA 2381
GGCTCTGGCAGGGCCTACAAGTTTATGCATGATACCTTCTCCTCAGAAGAAGCTCTGCAGGCAGGATCCCAACACACCAGACCAGGAC 2471
CCTGTGCCACTGTGGAGCTCAGTGAATCTGGTTCTCAAAGAAGACCCACTGTGGGGTTCCTCTAGTGTGACATAGGTGAGGTGGCAGCTG 2561
GCAAGGCCGCACTAGGCTGCCACACTGTTTGCCCCCTCAGATGGGCTCATAAAGACTGGGCTTGAGGGTCCACATGAAGAGCCTCACTTC 2651
AGGGGAATAGCCGGCCACGTCTAGCCCCTATCCCACTTAGAAAGGTCACCTGGCCACGGTGATGCTGGGTGGCTGATGAAACTTAAGCTC 2741
AGGGTCAGTGGGACCTGGCACTGCAGGTCAAGAGGACGCCTGGGCCCCTTCCTGACCCAGGCAGCTCTGCAGGAGCTGGGGGAGCATTGA 2831
CAGTCAAGGGCTTTTATAGACATACACATTCAGACCCTCGGTGTCCTTGTCCACTGAGTGCTGTATCTCATCGTATCTCAATCTTCATTG 2921
ACAGCACTGGAGAGGCTCGGGGCACCACTTGGAGCCTTGTATCCCGCAGGCTTGAGACCTGGGGGCACCTGGCTTACCAAGGTTGGCAAG 3011
GCTCCATCTTCCCTTCCAGGGGCTTGGGGATGCTAAGTTGCTGTATCCAGGAAGACAGTAGTCCTCACATGCAGAAGGCCTGGGAAAAGG 3101
CTGCTTGTCGCTTTTTTTTTTCCTTTCTTTTTGTCACTGGGTTAGAGAGGCTCCCGCCAGCACAGCACCAAGGCCTGCCCAGTAGTCAG 3191
GTCTCCTGGTCCAGGATGGGTGTGCTCTGTAGTTGGTCTGTGGGTTTCTTAGGCCCATGCCCTGAGCACATTCACCCAGCCACATCCCTG 3281
CTAGTGACACTCAGGCAGGGGCACTGGGAACCAGACCTGGGCTCTGGCATCAACGAGTGCCTAGGTGTCAGGACAGCGGCACCCATCCTG 3371
TACAGGGCAGCCAGGATAGTGGCCAAGCTACTGTGTCCTTTCTCCATGAGGCTCCCTGTCACTCGGTGCCAGTGGGTAATGTGTGTTCTT 3461
TGAGTCCTTATATGAATAAAAGGCTGGAGACCTACCAAAAAAAAAAAA                                           3509
```

shown to hybridize with a probe for a recently cloned major cell surface heparan sulfate proteoglycan of rat liver (16), whose deduced amino acid sequence (which codes for a 23 kDa protein) has a high degree of identity with the predicted partial sequence of human fibroglycan (15). The major 49 kDa core protein in the liver heparan sulfate proteoglycan preparation was also found to be reactive with an antibody to the cytoplasmic domain of fibroglycan. Because the available monoclonal antibodies, S1 and 1G11, to glypican (3) do not react with rat tissues or with fixed human brain (A. Flaccus and R. Margolis, unpublished results, and G. David, personal communication), we have raised polyclonal antibodies to a synthetic peptide from rat glypican for immunocytochemical studies of its localization in developing brain.

Although the GPI anchor of glypican could provide a means for its selective release from the cell surface in brain and other tissues by stimulation of a specific phospholipase under appropriate physiological conditions, as has been demonstrated for other GPI-anchored proteins (17,18), there is as yet no definitive evidence that such a process actually occurs for heparan sulfate proteoglycans. In the case of rat ovarian granulosa cells none of the GPI-anchored heparan sulfate proteoglycans are shed into the medium, but it appears that the GPI-anchored and membrane-intercalated heparan sulfate proteoglycans in these cells have distinct secretory, endocytic, and intracellular degradation pathways which may be due to differences in their anchor structures (19).

It is also interesting that the OCI-5 cDNA clone, corresponding to a developmentally regulated transcript in rat intestine, codes a 67 kDa protein having 23% identity and 47% similarity with rat glypican throughout the two sequences. The protein coded by clone OCI-5, which contains a spacing of cysteine residues and potential glycosaminoglycan attachment sites very similar to those in glypican, may be involved in cytoskeletal organization or in cell attachment (20).

The 89% identity of amino acid sequence in the rat and human proteoglycans and 66-80% identity over ~2 kb in the nucleotide sequences of the untranslated regions indicate that the glypican gene has been highly conserved, and the apparently wide cell and tissue distribution of the proteoglycan suggests that it serves a fundamental biological function. Cloning this rat brain heparan sulfate proteoglycan now provides a means for studying its cellular localization in developing nervous tissue, and its possible roles in cell interactions and in modulating the effects of neuronal and glial growth factors.

Acknowledgments: This research was supported by grants NS-09348 and NS-13876 from the National Institutes of Health, and by a Research Scientist Award to R.U.M. from the NIMH. We thank A. Flaccus, M. Flad, and M. Engel for skillful technical assistance.

REFERENCES

1. Klinger, M.M., Margolis, R.U. and Margolis, R.K. (1985) J. Biol. Chem. 260, 4082-4090.
2. Ripellino, J.A., and Margolis, R.U. (1989) J. Neurochem. 52, 807-812.

3. David, G., Lories, V., Decock, B., Marynen, P., Cassiman, J.-J., and Van den Berghe, H. (1990) J. Cell Biol. 111, 3165-3176.
4. Lories, V., Cassiman, J.-J., Van den Berghe, H., and David G. (1992) J. Biol. Chem. 287, 1116-1122.
5. Laemmli, U.K. (1970) Nature (Lond.) 227, 680-685.
6. Aebersold, R.H., Leavitt, J., Saavedra, R.A., Hood, L.E., and Kent, S.B. (1987) Proc. Natl. Acad. Sci. USA 84, 6970-6974.
7. Gooley, A.A., Classon, B.J., Marschalek, R., and Williams, K.L. (1991) Biochem. Biophys. Res. Comm. 178, 1194-1201.
8. Herndon, M.E., and Lander, A.D. (1990) Neuron 4, 949-961.
9. Bourdon, M.A., Krusius, T., Campbell, S., Schwartz, N.B., and Ruoslahti, E. (1987) Proc. Natl. Acad. Sci. U.S.A. 84, 3194-3198.
10. Krueger, R.C., Fields, T.A., Hildreth, J., and Schwartz, N.B. (1990) J. Biol. Chem. 265, 12075-12087.
11. Lories, V., David, G., Cassiman, J.J., and Van den Berghe, H. (1986) Eur. J. Biochem. 158, 351-360.
12. Lories, V., Cassiman, J.J., Van den Berghe, H., and David, G. (1989) J. Biol. Chem. 264, 7009-7016.
13. Carey, D.J., Evans, D.M., Stahl, R.C., Asundi, V.K., Conner, K.J., and Garbes, P., and Cizmeci-Smith, G. (1992) J. Cell Biol. 117, 191-201.
14. Saunders, S., Jalkanen, M., O'Farrell, S., and Bernfield, M. (1989) J. Cell. Biol. 108, 1547-1556.
15. Marynen, P., Zhang, J., Cassiman, J.-J., Van den Berghe, H., and David, G. (1989) J. Biol. Chem. 264, 7017-7024.
16. Pierce, A., Lyon, M., Hampson, I.N., Cowling, G.J., and Gallagher, J.T. (1992) J. Biol. Chem. 267, 3894-3900.
17. Low, M.G., and Saltiel, A.R. (1988) Science 239, 268-275.
18. Lisanti, M.P., Rodriguez-Boulan, E., and Saltiel, A.R. (1990) J. Membrane Biol. 117, 1-10.
19. Yanagishita, M. (1992) J. Biol. Chem. 267, 9505-9511.
20. Filmus, J., Church, J.G., and Buick, R.N. (1988) Mol. Cell Biol. 8, 4243-4249.