

# BEHAB, a New Member of the Proteoglycan Tandem Repeat Family of Hyaluronan-binding Proteins That Is Restricted to the Brain

Diane M. Jaworski, Gail M. Kelly, and Susan Hockfield

Yale University School of Medicine, Section of Neurobiology, New Haven, Connecticut 06510

**Abstract.** Hyaluronan (HA) is a ubiquitous component of the extracellular matrix of all tissues. In the mammalian central nervous system (CNS) HA is present throughout development and into adulthood. While the functions of HA are likely to be mediated by HA-binding proteins, no cell or tissue specific HA-binding proteins have been reported. In an effort to characterize the composition of the extracellular matrix of the CNS, we sought to identify neural HA-binding proteins.

We report here the isolation and characterization of a cDNA with a high degree of sequence homology to members of the proteoglycan tandem repeat (PTR) family of HA-binding proteins. Unlike other HA-binding proteins, the expression of this cDNA is re-

stricted to the CNS. We propose the name BEHAB, Brain Enriched HyAluronan Binding protein, for this gene. The expression of BEHAB mRNA is developmentally regulated; expression is first detected in the late embryonic period and peaks during the first two postnatal weeks. In the embryo, BEHAB is expressed at highest levels in mitotically active cells. The sequence of BEHAB has long stretches of identity between rat and cat, suggesting that the encoded protein is functionally important. The size and sequence of BEHAB are consistent with the possibility that it could serve a function like link protein, stabilizing interactions between HA and brain proteoglycans. These observations suggest the existence of other tissue specific HA-binding proteins.

**W**HILE the presence of an extracellular matrix in non-neural tissues has been universally accepted, only recently has evidence been adduced for an organized extracellular matrix in the mammalian central nervous system (CNS)<sup>1</sup> (Reichardt and Tomaselli, 1991; Hockfield et al., 1990; Sanes, 1989). Because molecules normally associated with the extracellular matrix, such as laminin and collagen, are not present (or present at very low abundance) in the CNS, the extracellular space in neural tissue has sometimes been considered a preparative artifact. However, the application of a variety of staining techniques demonstrated that the neural extracellular space is filled with a matrix-like material (Nakanishi, 1983; Castejon, 1970; Bomdareff, 1967). With the advent of monoclonal antibody technology, it became clear that many components of the extracellular matrix identified in other tissues are indeed present in the CNS (Margolis and Margolis, 1993; Rauch et al., 1991; Zaremba et al., 1990; Hockfield, 1990; Fujita et al., 1989).

The CNS extracellular matrix consists of a heterogeneous mixture of glycoconjugates, many of which are proteogly-

cans. Proteoglycans are complex macromolecules that consist of a core protein modified with one or more types of glycosaminoglycan chains (Hardingham and Fosang, 1992; Wight et al., 1991). Many functional properties of proteoglycans have been ascribed to glycosaminoglycans. Glycosaminoglycans have been reported to exhibit both adhesive and repulsive properties and, as such, have been suggested to mediate neuronal migration and axon guidance (Prieto et al., 1992; Cole and McCabe, 1991; Snow et al., 1990). Glycosaminoglycans are believed to regulate the local cellular environment primarily by serving as selective filters, facilitating permeability and retention of low molecular weight solutes, including growth factors, while excluding other macromolecules. Hyaluronan (HA) is particularly suited to this function because of its charge density and hydroscopic nature. HA is a ubiquitous component of extracellular matrices of all tissues, including brain, and is believed to organize water and extracellular proteins (Laurent and Fraser, 1992). During development HA plays a role in the regulation of morphogenesis and differentiation of neural tissues (Toole, 1991; Perris and Johansson, 1990).

Because HA is ubiquitously present in extracellular space, cell type specific functions attributed to HA may be mediated through its interaction with HA-binding proteins. Several HA-binding proteins have been reported in the brain, a subset of which have a high degree of sequence similarity to one another, including versican (Zimmermann and Ruos-

Address all correspondence to Dr. Diane M. Jaworski, Yale University School of Medicine, Section of Neurobiology, 333 Cedar St. SHM C-405, New Haven, CT 06510.

1. *Abbreviations used in this paper:* CNS, central nervous system; DRG, dorsal root ganglia; GHAP, glial hyaluronate-binding protein; HA, Hyaluronan; NF, neurofilament-middle; PTR, proteoglycan tandem repeat.

lahti, 1989), link protein (Doege et al., 1986), neurocan (Rauch et al., 1992), glial hyaluronate binding protein (GHAP) (Perides et al., 1989) and CD44 (Culty et al., 1990), and have been called the proteoglycan tandem repeat (PTR) family of HA-binding proteins (Perkins et al., 1989). Additional HA-binding proteins have been isolated from brain, for which sequence is not yet available, including the T1 proteoglycan (Iwata and Carlson, 1993), and the antigen recognized by monoclonal antibody Cat-301 (Fryer et al., 1992).

The spatial distribution and temporal expression of neural extracellular matrix proteoglycans, and HA-binding proteins, in particular, indicate that they may be involved in many events in the development and function of the mammalian CNS. While some HA-binding proteins represent general components of the extracellular matrix (e.g., T1 [Iwata and Carlson, 1993]), others have a restricted pattern of expression on subsets of neurons (e.g., the Cat-301 antigen [Hockfield and McKay, 1983; Hockfield et al., 1983; Hockfield and Sur, 1990; DeYoe et al., 1990]). In addition, while some extracellular matrix molecules are transiently expressed during embryogenesis (e.g., laminin [Letourneau et al., 1989; McLoon et al., 1988] and fibronectin [Stewart and Pearlman, 1987]), others are first expressed late in the postnatal period (e.g., chondroitin sulfate proteoglycans [Watanabe et al., 1989; Aquino et al., 1984; Hockfield et al., 1983; Maeda et al., 1992]), coincident with the decline in developmental synaptic plasticity.

Our interest in the composition and function of the CNS extracellular matrix, along with the ubiquitous expression of HA in the brain, motivated an investigation of neural HA-binding proteins. Here we describe the identification and characterization of a neural specific cDNA that shows a high degree of sequence similarity to the PTR family of HA-binding proteins. We propose to name this gene BEHAB, for Brain Enriched HyAluronan Binding protein.

## Materials and Methods

### Isolation of Rat cDNA Clones

An unamplified postnatal day 12 (P12) rat brain  $\lambda$ gt10 cDNA library was screened with rat aggrecan clone pRCP 4 (Doege et al., 1987).  $4 \times 10^6$  phage (per 150-mm plate) were plated with C600 bacteria, immobilized onto nitrocellulose filters, and prepared for hybridization according to standard techniques (Sambrook et al., 1989). Filters were prewashed for 1 h in 1 M NaCl, 0.1% SDS, 20 mM Tris-HCl (pH 8.0) and 1 mM EDTA at 65°C. Filters were then prehybridized for an additional 4–6 h in 50% formamide, 5 $\times$  SSC (1 $\times$  SSC = 0.15 M sodium chloride, 0.015 M sodium citrate), 1% SDS, 1 $\times$  Denhardt's (0.02% Ficoll, 0.02% BSA [Fraction V], 0.02% polyvinylpyrrolidone), 50 mM sodium phosphate (pH 6.7) and 100  $\mu$ g/ml salmon sperm DNA at 37°C. Hybridization was carried out in the identical solution with the inclusion of  $10^6$  cpm pRCP 4 probe/ml for 24 h at 37°C. For all experiments, radiolabeled probes ( $^{32}$ P-dCTP, Amersham) were prepared by random priming (Feinberg and Vogelstein, 1983) (Boehringer Mannheim Corp., Indianapolis, IN) gel purified cDNA inserts, followed by the removal of unincorporated radionucleotides (NICK column, Pharmacia). One post-hybridization wash in 2 $\times$  SSC, 0.1% SDS and one in 0.2 $\times$  SSC for 1 h each were performed at room temperature. Phage DNA was isolated using DE52 (Whatman) and the cDNA insert excised by EcoRI digestion. The cDNA was gel purified (Gene-Clean, Bio 101), subcloned into pBluescript KS<sup>+</sup> (Stratagene, LaJolla, CA) and transformed into DH5 $\alpha$  (Mandel and Higa, 1970) (GIBCO BRL, Gaithersburg, MD).

### Isolation of Cat cDNA Clones

Random nonamers (1.4 mg) were used to synthesize first strand cDNA from

5  $\mu$ g poly A<sup>+</sup> RNA isolated from P39 cat cortex. cDNA synthesis was performed according to manufacturer's instruction for the production of non-directional libraries (Stratagene) and size-fractionated by column chromatography (GIBCO BRL). 50 ng of cDNA was ligated to 1  $\mu$ g EcoR I cut, phosphatized Lambda Zap II vector and packaged into phage (Gigapack II Gold, Stratagene). This yielded  $0.5 \times 10^6$  recombinants when transfected into XLI-Blue (Stratagene). The unamplified library was screened with rat clone H1. Hybridization was performed in 6 $\times$  SSC, 0.1% SDS, 1 $\times$  Denhardt's and 100  $\mu$ g/ml salmon sperm DNA at 65°C. Filters were washed twice in 2 $\times$  SSC, 0.1% SDS and twice in 0.2 $\times$  SSC at 65°C for 20 min. cDNA inserts of plaque purified positive clones were isolated in pBluescript SK<sup>-</sup> by *in vivo* excision.

### DNA Sequencing and Analysis

DNA sequencing was performed by the dideoxy chain termination method (Sanger et al., 1977) using Sequenase (U.S. Biochemical, Cleveland, OH). Bluescript SK/KS primers or cDNA specific 20-mers were used. Sequence was verified from overlapping clones or by sequencing both strands of DNA. Sequence compressions were resolved using dITP nucleotides. After labeling, the reactions were incubated at 37°C for 30 min in the presence of 1 $\times$  reaction buffer, 1 mM dNTPs (pH 7.0) and 0.5 U terminal deoxynucleotidyl transferase to prevent premature termination caused by the use of dITP. Sequence analyses were performed using the University of Wisconsin Genetics Computer Group (GCG) programs (Devereux et al., 1984).

### Northern Hybridization

For Northern analysis, 25  $\mu$ g total RNA was denatured in 2.2 M formaldehyde, 50% formamide, 1 $\times$  MOPS (3-(N-morpholino) propanesulfonic acid) buffer at 65°C for 15 min. The RNA was electrophoresed on a 1.0% agarose-formaldehyde gel with 1 $\times$  MOPS buffer at 50V with buffer recirculation. The gel was briefly neutralized in transfer buffer (20 $\times$  SSC) and RNA blotted to Zetaprobe (BioRad Labs., Hercules, CA) by capillary transfer. Filters were rinsed briefly in 2 $\times$  SSC, and RNA was immobilized both by UV cross-linking (Church and Gilbert, 1984) and baking *in vacuo* (80°C for 1 h). Hybridization in 7% SDS, 1% BSA, 0.5 M phosphate buffer (pH 6.8) (PB), 1 mM EDTA and  $0.5\text{--}2.5 \times 10^6$  cpm rat H1 probe/ml was carried out for at least 8 h at 65°C. Filters were washed twice in 5% SDS, 0.5% BSA, 40 mM PB, 1 mM EDTA and twice in 1% SDS, 40 mM PB, 1 mM EDTA at 65°C for 20 min (Church and Gilbert, 1984), and exposed to film (Hyperfilm, Amersham) at  $-70^\circ\text{C}$ . Molecular sizes were determined relative to RNA molecular weight standards (GIBCO BRL) and 28S and 18S ribosomal RNA observed during UV illumination. The ubiquitously expressed, non-developmentally regulated gene cyclophilin (Danielson et al., 1988; Lenoir et al., 1986) was used to determine equal loading of lanes. Densitometry was performed using the NIH Image program.

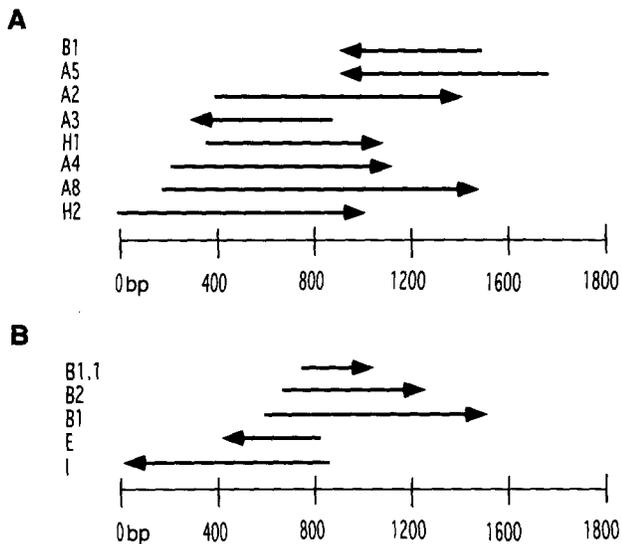
### In Situ Hybridization

Twelve to fourteen micron thick frozen sections were thaw-mounted onto gelatin-coated slides and postfixed in 0.1 M sodium phosphate buffered 4% paraformaldehyde (pH 7.4). Sections were rinsed in 1 $\times$  PBS (137 mM NaCl, 2.7 mM KCl, 10 mM Na<sub>2</sub>HPO<sub>4</sub>, 1.8 mM KH<sub>2</sub>PO<sub>4</sub>), 2 $\times$  SSC and acetylated with 0.5% acetic anhydride in 0.1 M triethanolamine (pH 8.0). Sections were then rinsed in 2 $\times$  SSC, 1 $\times$  PBS, dehydrated in ethanols and delipidated in chloroform. Sections were prehybridized in 2 $\times$  SSC, 50% formamide at 50°C for 1 h, and then hybridized in 0.75 M NaCl, 50% formamide, 1 $\times$  Denhardt's, 10% dextran sulfate, 30 mM DTT, 10 mM Tris-HCl (pH 7.5), 1 mM EDTA, 100  $\mu$ g/ml salmon sperm DNA, 0.5 mg/ml yeast tRNA and  $10^6$  cpm probe per slide at 50°C for 12–15 h. [<sup>35</sup>S]CTP (New England Nuclear, Boston, MA) labeled cRNA probes were synthesized using T3 (GIBCO BRL), SP6, and T7 RNA polymerases (New England Biolabs Inc., Beverly, MA). After hybridization, sections were washed in 2 $\times$  SSC, 50% formamide, 0.1% BME ( $\beta$ -mercaptoethanol) at 50°C for 1 h and treated with 20  $\mu$ g/ml RNase A in 0.5 M NaCl, 10 mM Tris-HCl (pH 8.0) at 37°C for 30 min. Sections were then washed in 2 $\times$  SSC, 50% formamide, 0.1% BME at 58°C for 30 min and 0.1 $\times$  SSC, 0.1% BME at 63°C for 30 min and dehydrated. For initial localization of probe the slides were exposed to film (Hyperfilm, Amersham) for 4 d. Autoradiograms were used as negatives for prints. For higher resolution, the slides were dipped in NTB-2 emulsion (Kodak), developed after 5 d and counter stained with cresyl violet. Neurofilament-middle (NF) antisense (Martin et al., 1992) and rat clone BI sense probes were used as positive and negative controls, respectively.

## Results

### BEHAB Encodes a Protein with Homology to the PTR Family of HA-binding Proteins

To identify HA-binding proteins involved in neural development, rat aggrecan clone pRCP 4, which encodes the HA-binding region (Doege et al., 1987), was used to screen a postnatal day 12 (P12) rat brain  $\lambda$ gt10 cDNA library. A total of  $3.2 \times 10^5$  recombinants were screened resulting in two positives. On Northern blots the two clones recognize the same size mRNA transcript (data not shown). Preliminary sequence analysis showed that both clones have homology to several HA-binding proteins. To obtain full length cDNA sequence, the library was rescreened with one of these clones, resulting in 15 additional clones. The insert size of these clones was determined and partial restriction maps were prepared to eliminate redundant clones. Eight clones were subcloned into pBluescript KS<sup>+</sup> and sequenced (Fig. 1 A). The composite sequence of BEHAB was obtained from these overlapping clones (Fig. 2). The complete BEHAB coding sequence is 1,113 bp. The nucleotide sequence preceding the first AUG contains a consensus sequence for translation initiation (Kozak, 1991, 1984). In the 3' untranslated region only that sequence verified from clones B1, A2, and A8 is presented. The deduced amino acid composition of the BEHAB protein is comprised of 371 amino acids and includes a putative signal peptide cleavage site at Ala-22 (von Heijne, 1986). The resulting mature protein has a predicted molecular mass of 38,447 kD. Analysis of the deduced amino acid sequence indicates the presence of two NX(S/T) consensus sequences for potential N-glycosylation (Marshall, 1974).



**Figure 1.** Overlapping cDNA clones encoding BEHAB. (A) cDNA clones for rat BEHAB were isolated from a P12 rat brain  $\lambda$ gt10 library and subcloned into pBluescript KS<sup>+</sup>. (B) cDNA clones for cat BEHAB were isolated from a P39 cat cortex Lambda Zap II library and excised in pBluescript SK<sup>-</sup>. cDNA alignment was produced using the GCG GELOVERLAP function. The lines indicate the size of the clones in base pairs (bp) and arrows indicate cloning orientation in relation to the coding sequence.

In our previous characterizations of extracellular matrix proteins in brain we focused on expression in the cat visual system. Therefore, we investigated whether the BEHAB gene is also expressed in the cat cortex. Northern hybridization using rat H1 as the probe revealed a single mRNA species in cat visual cortex identical in size to rat BEHAB (see Fig. 6). Rat H1 was used to clone the cat homologue from a P39 cat cortex Lambda Zap II cDNA library. A total of  $3.2 \times 10^5$  recombinants were screened resulting in five positives which were excised in pBluescript SK<sup>-</sup> (Fig. 1 B). The composite sequence for cat BEHAB was obtained from these overlapping clones (Fig. 3). The complete coding sequence for cat BEHAB is 1,134 bp. The first AUG is preceded by both an in-frame termination codon and the translation initiation consensus sequence. The cat BEHAB sequence encodes 378 amino acids which, like the rat, contains a 22 residue signal peptide. However, cat BEHAB contains six additional amino acids at the carboxy terminus, resulting in a predicted molecular mass of 38,955 kD. In the cat, Trp-373 is encoded by TGG, while the corresponding rat sequence of TAG results in termination. This termination sequence was verified in rat clones B1, A5, and A8 and by sequencing both strands of cat clone B1. Cat BEHAB also contains one additional site for potential N-glycosylation not present in the rat.

Database analyses at both the nucleic acid and amino acid levels indicated that BEHAB is a previously unreported member of the PTR family of HA-binding proteins. BEHAB has a substantial degree of amino acid identity to the other members of the PTR family, which includes rat aggrecan (48%) (Doege et al., 1987), rat neurocan (48%) (Rauch et al., 1992), human versican (46%) (Zimmermann and Ruoslahti, 1989), and rat link protein (42%) (Doege et al., 1986). The NH<sub>2</sub>-terminal domain of this family is defined by two structural motifs, (a) an immunoglobulin (Ig) fold and (b) two PTR folds (PTR 1 and PTR 2). The PTR folds have been suggested to mediate binding to HA (Doege et al., 1991; Fosang and Hardingham, 1989; Goetinck et al., 1987). The Ig domain contains two clusters of conserved amino acids around the cysteine residues which generate the disulfide bond of the loop. The consensus sequence YxCxVxH in the COOH-terminal cluster is present in all immunoglobulin and major histocompatibility complex proteins, and is also present in BEHAB (Fig. 4). The most conserved region of the PTR family's HA-binding domain is the sequence CDAGWL(A/S)D(Q/G)(T/S)VRYPYI found in PTR 1 and PTR 2. Two copies of this sequence are also found in BEHAB. The degree of identity of BEHAB between rat and cat is high (84% overall), with the greatest conservation in PTR 1. The identity in PTR 1 is 95% over the entire domain and 100% over 44 amino acids of the domain. PTR 2 shows the next highest homology (86%), followed by the Ig domain (84%) (Fig. 5). The relative degree of homology between the PTR 1, PTR 2, and Ig domains observed in rat and cat is also observed between BEHAB and the other members of the PTR family (Table I and Fig. 4).

### BEHAB mRNA Is Restricted to the Nervous System

The tissue distribution of rat BEHAB mRNA was determined by Northern blot analysis (Fig. 6). A single 3.9-kb mRNA transcript is detected in adult rat cortex, spinal cord, and cerebellum. This transcript is not detected in liver, kid-

	CGAGACCCGCGCAGAGAAGGGAGCGGGTCCCGTGACCGCGCAGAGCCCCCA	52
CGCGGCCAAAGGCGGGGACCGGGGGAAGGCGGGCGCGTGGGAAGAAACCCCTTTTGTGCGGCTCCCGGCGAGCTGGCGCCCCCGTCTCGTCCCGC	151	
GCGCCCGCCTGCTCGGCGCCGCGCATTGCGCGAGTCTCGGCTGCGTGGGACGCGGTGTGGAGGGGACCTCACAAAGTTCTTCCAAGTTTGCAGC	250	
ATG ATC CCA TTG CTT CTG TCC CTG CTG GCA GCT CTG GTC CTG ACC CAA GCC CCT GCA GCC CTC GCT GAT GAC CTG	325	
<u>M I P L L L S L L A A L V L T Q A P A A L A D D L</u>	25	
AAA GAA GAC AGC TCA GAG GAT CGA GCC TTT CGG GTG CGC ATC GGT GCC GCG CAG CTG CGG GGT GTG CTG GGC GGT	400	
K E D S S E D R A F R V R I G A A Q L R G V L G G	50	
TGG GTG GCC ATC CCA TGC CAC GTC CAC CAC CTG AGG CCG CCG CCC AGC CGC CGG GCC GCG CCG GGC TTT CCC CGA	475	
W V A I P C H V H H L R P P P S R R A A P G F P R	75	
GTC AAA TGG ACC TTC CTG TCC GGG GAC CGG GAG GTG GAG GTG CTG GTG GCG CGC GGG CTG CGC GTC AAG GTA AAC	550	
V K W T F L S G D R E V E V L V A R G L R V K V N	100	
GAA GCC TAT CGG TTC CGC GTG GCG CTG CCT GCC TAC CCC GCA TCG CTC ACA GAT GTG TCT TTA GTA TTG AGC GAA	625	
E A Y R F R V A L P A Y P A S L T D V S L V L S E	125	
CTG CGG CCC <u>AAT GAT TCC</u> GGG GTC TAT CGC TGC GAG GTC CAG CAC GGT ATC GAC GAC AGC AGT GAT GCT GTG GAA	700	
L R P <u>N D S</u> G V Y R C E V Q H G I D D S S D A V E	150	
GTC AAG GTC AAA GGG GTC GTC TTC CTC TAC CGA GAG GGC TCT GCC CGC TAT GCT TTC TCC TTC GCT GGA GCC CAG	775	
V K V K G V V F L Y R E G S A R Y A F S F A G A Q	175	
GAA GCC TGT GCT CGC ATC GGA GCC CGA ATT GCC ACC CCT GAG CAG CTG TAT GCT GCC TAC CTC GGC GGC TAT GAA	850	
E A C A R I G A R I A T P E Q L Y A A Y L G G Y E	200	
CAG TGT GAT GCT GGC TGG CTG TCC GAC CAA ACC GTG AGG TAC CCC ATC CAG AAC CCA CGA GAA GCC TGT TAT GGA	925	
Q C D A G W L S D Q T V R Y P I Q N P R E A C Y G	225	
GAC ATG GAT GGC TAC CCT GGA GTG CGG AAT TAC GGA GTG GTG GGT CCT GAT GAT CTC TAC GAT GTC TAC TGT TAT	1000	
D M D G Y P G V R N Y G V V G P D D L Y D V Y C Y	250	
GCC GAA GAC CTA AAT GGA GAA CTG TTC CTA GGT GCC CCT CCC GGC AAG CTG ACG TGG GAG GAG GCT CGG GAC TAC	1075	
A E D L N G E L F L G A P P G K L T W E E A R D Y	275	
TGT CTG GAA CGC GGT GCT CAG ATC GCT AGC ACG GGC CAG CTA TAC GCG GCA TGG AAT GGC GGC TTG GAC AGA TGT	1150	
C L E R G A Q I A S T G Q L Y A A W N G G L D R C	300	
AGC CCT GGC TGG CTG GCT GAT GGC AGT GTG CGG TAC CCC ATC ATC ACG CCC AGC CAA CGC TGT GGG GGA GGC CTG	1225	
S P G W L A D G S V R Y P I I T P S Q R C G G G L	325	
CCA GGA GTC AAG ACC CTC TTC CTC TTT CCC <u>AAC CAG ACT</u> GGC TTC CCC AGC AAG CAG AAC CGC TTC AAT GTC TAC	1300	
P G V K T L F L F P <u>N Q T</u> G F P S K Q N R F N V Y	350	
TGC TTC CGA GAC TCT GCC CAT CCC TCT GCC TTC TCT GAG CCT CCA GCC CAG CCT CTG ATG GACTAGAGGCCATTGTCAC	1379	
C F R D S A H P S A F S E P P A Q P L M D	371	
AGTGACAGAGAAGCTGGAGGAAGTGCAGTTGCCTCAGGAAGCTGTGGAGAGCGAGTCTCGTGGGGCGATCTACTCCATCCCCATCACAGAAGATGGGG	1478	
AGGAGGAAGCTCTACCCAGAAAGACCCAGCAGAGGCCCCAG	1520	

**Figure 2.** Nucleotide and deduced amino acid sequences of rat BEHAB. The nucleotide sequence was determined from the overlapping cDNA clones shown in Fig. 1 A. The deduced amino acid sequence contains a signal peptide, which is *underlined*, and a putative cleavage site, indicated by an *arrow*. The sequence also contains two potential N-glycosylation sites, which are *boxed*. These sequence data are available from EMBL/GenBank/DBJ under accession number Z28366.

ney, spleen, lung or muscle, even with longer film exposures. In situ hybridization to whole embryos (Fig. 10) also shows that BEHAB expression is restricted to the CNS (described further below).

The spatial distribution of BEHAB mRNA within the nervous system was determined at higher resolution by in situ hybridization on tissue sections from P21 rat forebrain, brainstem, spinal cord, and cerebellum. Near adjacent sections were probed with an antisense cRNA probe of rat clone

B1 (see Fig. 1 A), NF as a positive control for neuronal mRNA expression, and a sense probe of clone B1 as a negative control.

BEHAB expression is widely distributed in the brain, in both gray and white matter. The cortex exhibits diffuse hybridization with no laminar specificity (Fig. 7 A). Hybridization is detected in white matter tracts, including the corpus callosum (CC), the fimbria of the hippocampus (F), and the anterior commissure (AC). In the hippocampus (H) the most

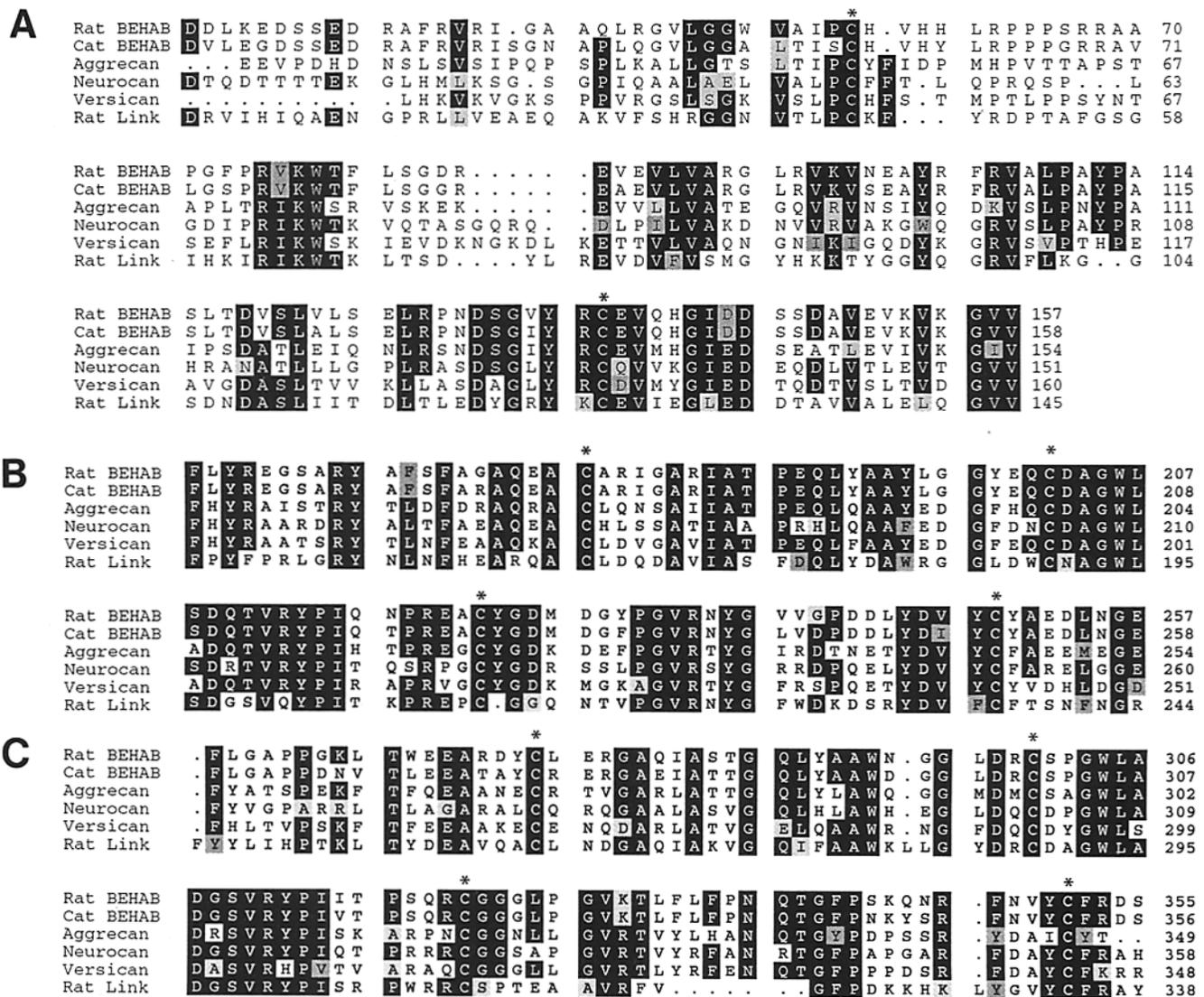
CGGCACGAGCTCGTGC CGAATTCGGCCACAGAGGGACCGAGCGTGGACCCGGAGGAGAGCCCGGAGGAGAGC	171
CCGGAGGAGGGCGCAAACCTGGCGGTGCGCACCCCTAGCCCCGGCCCTCGGCCTGCCGGAAGAAAACAAAGGCCCTGAGAGCTTAAGGAACCTGCAGCAAG	170
TTGACTAGCGCCAGGTCCTGGTTCGAGGAGGAATCCTGGTGGGAGACAGGATCAGAAGCCGAGGGTGTAAACAGTGAAGTCTCCAGCAGCCCTGAGC	269
ATG GCC CCA CTG TTC CTG CCC CTG CTG ATA GCC CTG GCC CTG GCC CCG GGC CCC ACG GCC TCA GCT GAT GTC CTG	344
<u>M A P L F L P L L I A L A L A P G P T A S A</u> D V L	25
GAA GGG GAC AGC TCA GAG GAC CGG GCC TTC CGC GTG CGC ATC TCG GGC AAC GCG CCG CTG CAG GGC GTG CTG GGC	419
E G D S S E D R A F R V R I S G N A P L Q G V L G	50
GGC GCC CTC ACC ATC TCG TGC CAC GTT CAC TAC CTG CGG CCG CCG CCG GGC CGC CGG GCC GTG CTG GGC TCC CCG	494
G A L T I S C H V H Y L R P P P G R R A V L G S P	75
CGG GTC AAG TGG ACC TTC CTG TCC GGG GGC CGG GAG GCC GAG GTG CTG GTG GCG CGG GGG CTG CGC GTC AAG GTG	569
R V K W T F L S G G R E A E V L V A R G L R V K V	100
AGC GAG GCC TAC CGG TTC CGC GTG GCG CTG CCC GCC TAC CCG GCG TCC CTC ACC GAC GTC TCC CTG GCA CTG AGC	644
S E A Y R F R V A L P A Y P A S L T D V S L A L S	125
GAG CTG CGG CCC <u>AAC GAC TCT</u> GGC ATC TAC CGC TGC GAG GTC CAG CAC GGC ATA GAC GAC AGC AGC GAC GCC GTG	719
E L R P <u>N D S</u> G I Y R C E V Q H G I D D S S D A V	150
GAG GTC AAG GTC AAA GGG GTC GTC TTT CTC TAC CGG GAG GGC TCT GCC CGC TAC GCT TTC TCC TTC GCC CGG GCC	794
E V K V K G V V F L Y R E G S A R Y A F S F A R A	175
CAG GAG GCC TGT GCC CGC ATC GGA GCC CGC ATC GCC ACC CCG GAG CAG CTC TAC GCT GCC TAC CTC GGG GGC TAT	869
Q E A C A R I G A R I A T P E Q L Y A A Y L G G Y	200
GAG CAG TGC GAT GCT GGC TGG CTG TCC GAC CAA ACC GTG AGG TAT CCC ATC CAG ACC CCA CGG GAG GCC TGT TAT	944
E Q C D A G W L S D Q T V R Y P I Q T P R E A C Y	225
GGA GAC ATG GAT GGC TTC CCT GGG GTC CGG AAC TAT GGC CTG GTG GAC CCG GAT GAC CTC TAC GAT ATC TAC TGC	1019
G D M D G F P G V R N Y G L V D P D D L Y D I Y C	250
TAT GCT GAA GAC CTA AAT GGA GAG CTG TTC CTG GGC GCC CCT CCA GAC <u>AAC GTG ACG</u> CTG GAG GAG GCT ACG GCA	1094
Y A E D L N G E L F L G A P P D <u>N V T</u> L E E A T A	275
TAC TGC CGT GAG CGG GGT GCA GAG ATT GCT ACC ACG GGC CAG CTG TAT GCA GCC TGG GAT GGC GGC CTG GAC CGC	1169
Y C R E R G A E I A T T G Q L Y A A W D G G L D R	300
TGC AGC CCC GGC TGG CTG GCC GAT GGC AGC GTG CGC TAC CCC ATC GTC ACG CCC AGC CAG CGC TGC GGT GGG GGC	1244
C S P G W L A D G S V R Y P I V T P S Q R C G G G	325
CTG CCT GGC GTC AAG ACT CTC TTC CTC TTC CCC <u>AAC CAG ACC</u> GGC TTC CCC AAC AAG TAC AGC CGC TTC AAC GTC	1319
L P G V K T L F L F P <u>N Q T</u> G F P N K Y S R F N V	350
TAC TGC TTC CGA GAC TCT GGC CAG CCC TCC ACC ACC CCT GAG GCC TCT GAC CAG CCT CTG ACG GGC TGG AGG CCA	1394
Y C F R D S G Q P S T T P E A S D Q P L T G W R P	375
TTG TCA CAGTGACAGAGACCCTAGAGGAGCTCCACGTGCCGCGGAAGCTGTGGAGAGCGAGTCCCGGGGAGCCATCTACTCCGTCCTCCCATTTGTGGA	1492
L S Q	378
GGATGGGGGAGGTGCAAGGTCCCTCCA	1519

**Figure 3.** Nucleotide and deduced amino acid sequences of cat BEHAB. The nucleotide sequence was determined from the overlapping cDNA clones shown in Fig. 1 B. The deduced amino acid sequence contains a signal peptide, which is *underlined*, with the putative cleavage site, indicated by an *arrow*. Three potential N-glycosylation sites are *boxed*. These sequence data are available from EMBL/GenBank/DBJ under accession number Z28367.

intense hybridization is present in the CA1 subfield, with less intense hybridization in subfields CA2 and CA3, and the dentate gyrus. Examination of emulsion-coated slides counter stained with cresyl violet reveal that the labeling in CA1 is associated with neurons (Figs. 8, A and B). The pattern of NF hybridization in the hippocampus is essentially reciprocal to that of BEHAB (Fig. 7 E); the NF probe hybridizes most intensely in hippocampal subfields CA2, CA3, and in

the dentate gyrus. BEHAB hybridization is also seen throughout the inferior colliculus (IC) and less intensely in the superior colliculus (SC). In addition to the hippocampus, BEHAB hybridization in gray matter is most intense in the substantia nigra (SN). The BI sense probe generates almost no signal in most of the brain, but a low level of hybridization is seen in the hippocampus and dentate gyrus (Fig. 7 I).

In the brainstem, BEHAB is expressed throughout the re-



**Figure 4.** BEHAB has homology to the PTR family of HA-binding proteins. A database search showed that the deduced amino acid sequence of rat and cat BEHAB have a high degree of homology to several PTR HA-binding proteins. The PTR proteins contain three functional domains: an immunoglobulin fold (A), and two domains thought to be involved in HA binding, PTR 1 (B) and PTR 2 (C). The homology among all of the HA-binding proteins is greatest in PTR 1 and PTR 2, followed by the Ig fold (see Table I). Identical amino acids are shown in black. Intensity of gray counter-shading, indicating amino acid similarity, was determined by the method of Schwartz and Dayhoff (1979). Sequence alignment was produced using the GCG PILEUP function and counter-shaded using PRETTYBOX.

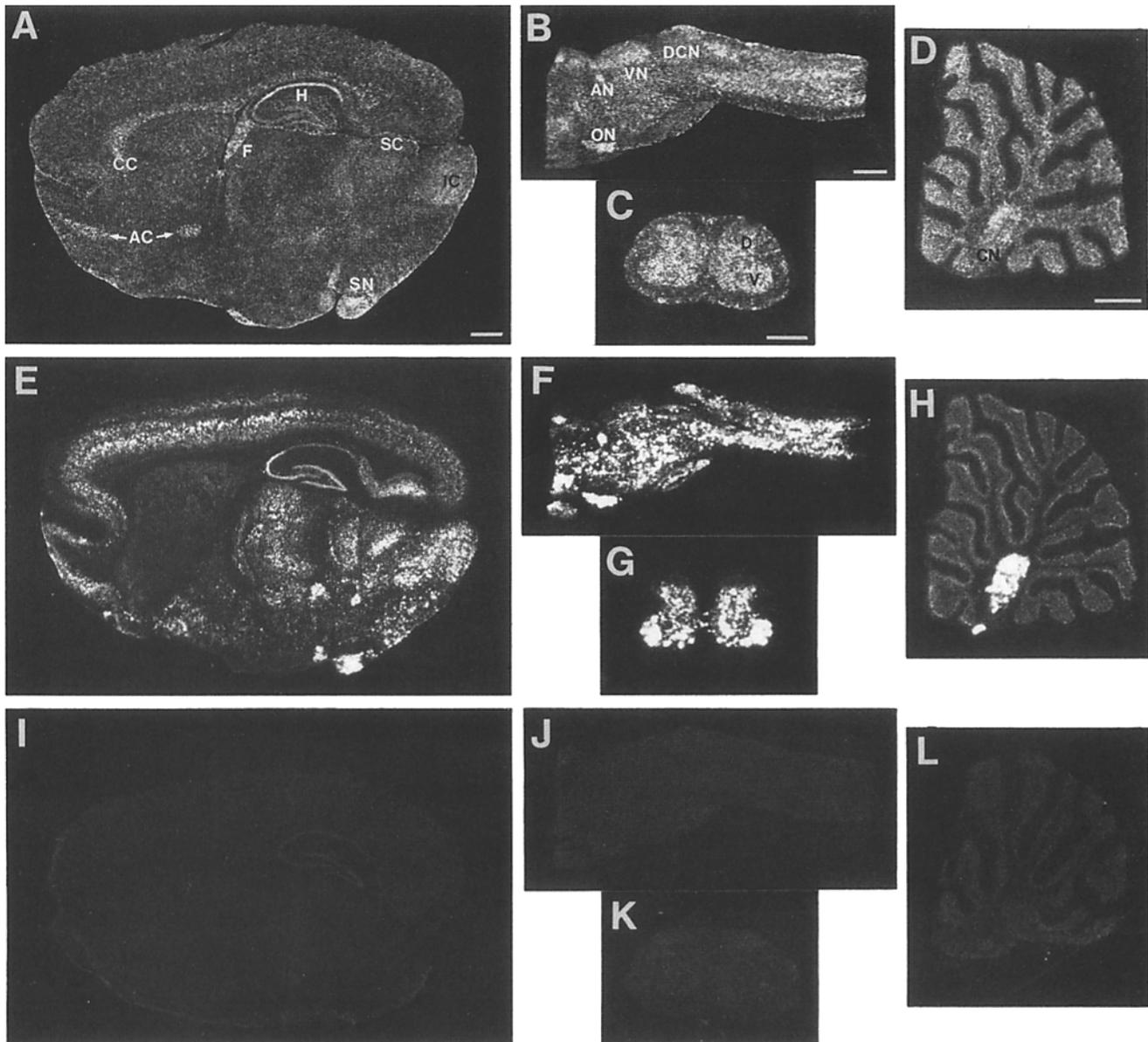
ticular formation (Fig. 7 B). Several brainstem nuclei also express BEHAB, including the superior olivary nucleus (ON), the vestibular nuclei (VN), the abducens nucleus (AN) and the dorsal column nuclei (DCN). A similar hybridization pattern is observed with NF (Fig. 7 F), while no hybridization signal is detected with the sense B1 probe (Fig. 7 J).

BEHAB expression in the spinal cord is greater in the gray matter than white matter (Fig. 7 C). In the gray matter, BEHAB expression is slightly greater in the ventral than in the dorsal horn. BEHAB hybridization is lacking in the substantia gelatinosa (layer 2), but is present in layer 1. In the ventral horn hybridization is seen over motor neurons (Fig. 8, C and D). In the spinal cord white matter, the size of labeled cells and their distribution indicates that BEHAB is ex-

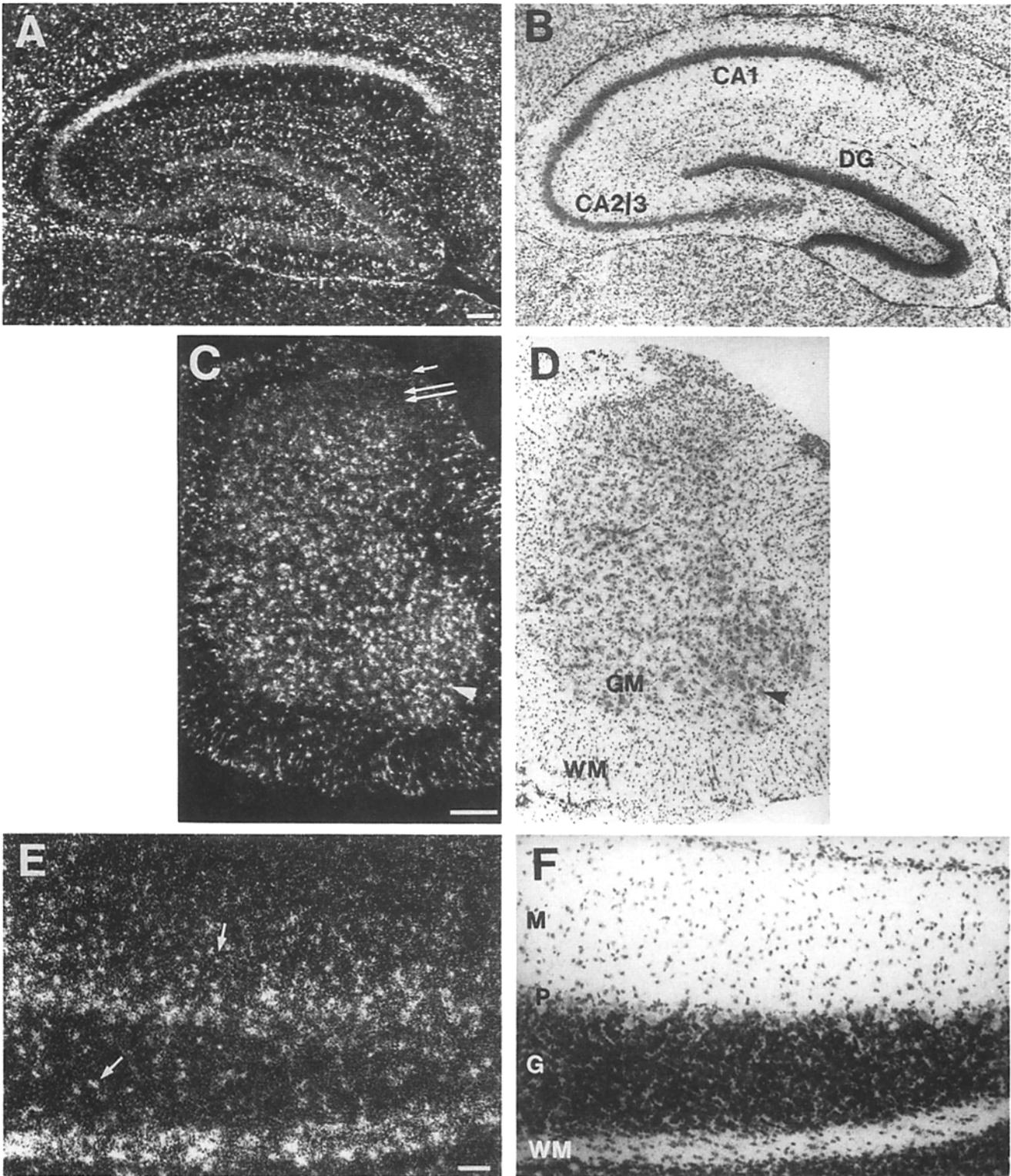
pressed by glial cells. Like BEHAB, NF expression is greater in the ventral horn than in the dorsal horn; however, unlike BEHAB, NF is not detected in the spinal white matter (Fig. 7 G). As observed in the brainstem, no hybridization signal is detected in the spinal cord with the B1 sense probe (Fig. 7 K).

In the cerebellum, BEHAB expression is greatest in the deep cerebellar nuclei (CN) (Fig. 7 D). In the cerebellar cortex, labeling is detected in all three cortical layers. In the molecular layer the distribution of

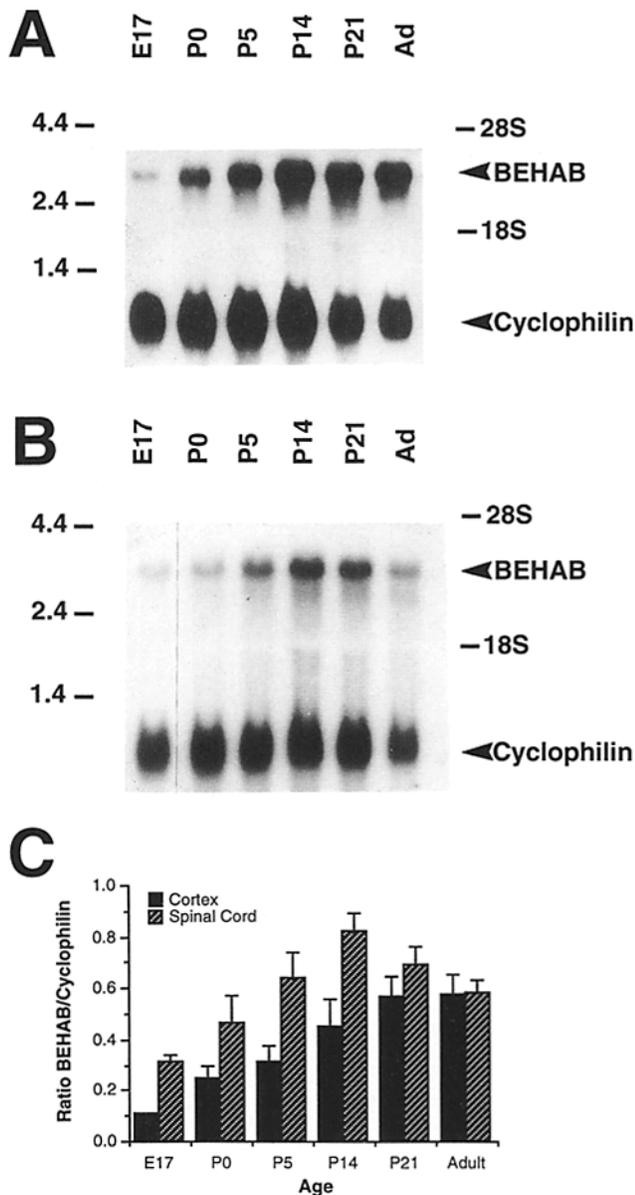




**Figure 7.** BEHAB mRNA is expressed throughout the central nervous system. In situ hybridization was performed on sections from P21 rat forebrain, brainstem, spinal cord, and cerebellum using a <sup>35</sup>S-labeled cRNA probe to rat clone B1 (A-D) (see Materials and Methods). Near adjacent sections were probed with NF (neurofilament, middle) as a marker for neuronal mRNA expression (E-H) and a sense probe of clone B1 as a negative control (I-L). (A-D) BEHAB hybridization is detected in both gray and white matter. A diffuse hybridization signal is detected throughout the gray matter at all levels. In the forebrain (A), the highest levels of hybridization are detected in the hippocampus (H) and substantia nigra (SN). In the hippocampus the most intense hybridization is present in the CA1 subfield (see Figs. 8, A and B). The inferior colliculus (IC) has a higher hybridization signal than the superior colliculus (SC). A high level of hybridization is seen in white matter tracts, including the corpus callosum (CC), the fimbria of the hippocampus (F), and the anterior commissure (AC). (B) BEHAB is expressed in several brainstem nuclei, including the superior olivary nucleus (ON), the vestibular nucleus (VN), the abducens nucleus (AN), and the dorsal column nuclei (DCN). (C) BEHAB expression in the spinal cord is greater in the gray matter than white matter. In the gray matter, BEHAB hybridization is slightly greater in the ventral horn (V) than in the dorsal horn (D). In the dorsal horn, BEHAB is present in layer 1, but is absent in the substantia gelatinosa (see Fig. 8, C and D). (D) BEHAB expression in the cerebellum is greatest in the deep cerebellar nuclei (CN). Labeling is also detected in all three cortical layers and the white matter (see Fig. 8, E and F). (E) NF hybridization is very high in many gray matter regions, but is virtually absent from white matter tracts. In the hippocampus NF localization is reciprocal to that of BEHAB; NF hybridizes most intensely in subfields CA2, CA3, and in the dentate gyrus. (F) NF hybridization in the nuclei of the brainstem is similar to BEHAB, but it is not seen in white matter tracts. (G) In the spinal cord, NF mRNA is restricted to the gray matter, where, like BEHAB, it is present expressed at greater levels in the ventral horn. (H) In the cerebellar cortex NF is primarily expressed by Purkinje cells; it is also expressed at high levels by the neurons in the deep cerebellar nuclei. (I-L) The sense probe generates almost no hybridization signal, except a low level seen in the hippocampus, dentate gyrus, and granule cell layer of the cerebellum. Bar, 1 mm.



**Figure 8.** BEHAB mRNA is expressed by neurons and glial cells. Emulsion-coated sections were counter stained with cresyl violet and photographed under dark field optics to visualize silver grains (A, C, and E) and with transmitted illumination to visualize cell bodies (B, D, and F). (A and B) In the hippocampus BEHAB hybridization is most intense in the CA1 subfield, and is much less intense in subfields CA2 and CA3, and in the dentate gyrus (DG). Observation of cresyl violet counter stained sections shows that the BEHAB hybridization pattern is not attributable to differences in cell density. (C and D) In the spinal cord BEHAB expression is seen in gray matter (GM) and white matter (WM). In the gray matter, BEHAB expression is slightly greater in the ventral than in the dorsal horn. In the ventral horn hybridization is seen over motor neurons (arrowhead). In the dorsal horn hybridization is present in layer 1 (single arrow), but is absent from the substantia gelatinosa (double arrow). In the white matter the size and distribution of labeled cells suggests BEHAB synthesis by astrocytes. (E and F) In the cerebellar cortex, intense BEHAB hybridization is detected in the Purkinje cell layer (P) and in the white matter (WM). The distribution of silver grains in the molecular (M) and granule cell (G) layer indicates expression by interneurons (arrows). Bars: (A-D) 200  $\mu$ m; (E and F) 50  $\mu$ m.



**Figure 9.** BEHAB mRNA is developmentally regulated. Northern analysis was performed with 25  $\mu$ g total RNA from E17, P0, P5, P14, P21, and adult (*Ad*) rat cortex (*A*) and spinal cord (*B*). Hybridization was performed simultaneously with  $^{32}$ P-labeled probes of rat clone H1 and the non-developmentally regulated gene cyclophilin (see Materials and Methods). A single 3.9-kb mRNA transcript is detected in both cortex and spinal cord. To determine the developmental expression of BEHAB, densitometry was performed, and standardized by calculating the ratio of BEHAB absorbance relative to cyclophilin at each age (*C*). In the cortex, BEHAB mRNA is detected at E17 and its expression gradually increases with development to reach adult levels by P21. ( $N = 5$ , except at E17 where  $N = 1$ ) In the spinal cord, BEHAB is also detected at E17, gradually increases to reach maximal expression by P14 and then declines slightly. ( $N = 6$ , except at E17 where  $N = 2$ ).

Unlike actin and tubulin, whose abundance varies with development (Schmitt et al., 1977), cyclophilin maintains a constant relative abundance in the CNS throughout development (Lenoir et al., 1986). The Northern blots were ana-

lyzed by densitometry, and band intensity of BEHAB was standardized by calculating a ratio of the abundance of BEHAB to cyclophilin at each developmental age (Fig. 9 *C*).

In the cortex, BEHAB recognizes a single 3.9-kb mRNA transcript (Fig. 9 *A*). BEHAB expression is detected at embryonic day 17 (E17) and gradually increases to attain adult levels by P21 (Fig. 9 *C*). In the spinal cord, BEHAB also recognizes a 3.9-kb mRNA transcript (Fig. 9 *B*). At all ages, except the adult, BEHAB expression is greater in the spinal cord than cortex. Like the cortex, BEHAB is present in the spinal cord at E17 and gradually increases with age, but reaches maximal levels by P14. Unlike the cortex, BEHAB expression in the spinal cord then declines slightly (Fig. 9 *C*).

The expression of BEHAB in the embryo, like in the postnatal animal, is restricted to the CNS (Fig. 10 *A*). BEHAB expression is absent in dorsal root ganglia (DRG), a peripheral nervous system structure (Fig. 10, *B* and *C*). Tissues in the embryo that express high levels of closely related genes, such as the cartilage (which expresses aggrecan), also show no hybridization signal for BEHAB (Fig. 10, *B* and *C*). The distribution of BEHAB expression in the embryonic CNS differs slightly from the postnatal brain. The highest levels of BEHAB expression are found in regions that contain mitotically active cells, such as the ventricular zone of the medulla, midbrain, and spinal cord (Fig. 10 *A*). Expression of BEHAB is heterogeneous in the developing brain, the details of which are beyond the scope of this report and will be described in detail elsewhere (manuscript in preparation).

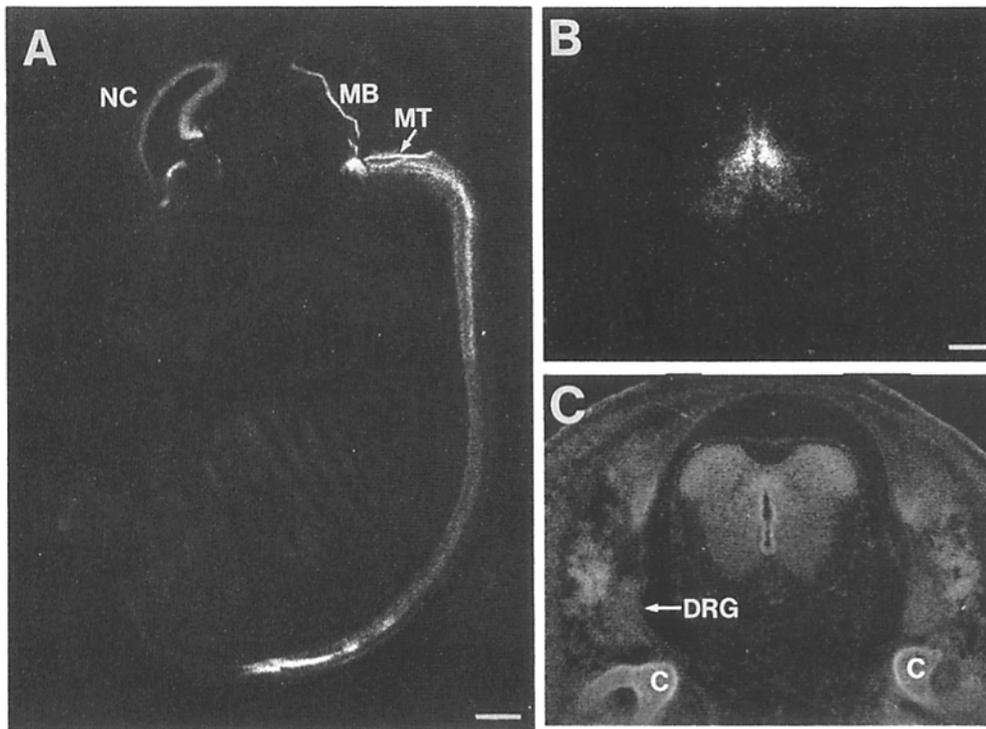
## Discussion

### *BEHAB Encodes a New Member of the PTR Family of Proteins*

The complete coding sequence of the BEHAB gene from rat and cat is reported here. The high degree of sequence identity between rat and cat suggests that the protein encoded by BEHAB is a functionally important element of the mammalian CNS. BEHAB shows sequence homology with a subset of HA-binding proteins, the proteoglycan tandem repeat (PTR) family, which includes aggrecan, versican, link protein, neurocan, and CD44.

The PTR family of HA-binding proteins is characterized by several conserved domains in the NH<sub>2</sub> terminus (see Fig. 11). The first 350 amino acids of these proteins contain ten conserved cysteine residues that generate three disulfide-bonded loops. The first loop, formed by a single disulfide bond, has homology to proteins in the immunoglobulin superfamily (Williams and Barclay, 1988; Bonnet et al., 1986). The second and third loops, referred to as PTR 1 and PTR 2 (Perkins et al., 1989), or B and B' (Doerge et al., 1990), are each formed by a pair of disulfide bonds. The PTR 1 loop is the most highly conserved of the three loops, followed by the PTR 2 loop, and then the Ig loop. The most conserved regions of BEHAB are, as for the other family members, the PTR 1 loop, and then the PTR 2 loop and the Ig loop. The molecular and histological data presented here clearly show that BEHAB is closely related to, but distinct from, all previously described members of the PTR family.

Among the members of the PTR protein family, the size of the mRNA recognized by BEHAB is most similar to link



**Figure 10.** In the embryo, as in the adult, BEHAB mRNA is restricted to the central nervous system. In situ hybridization was performed on E17 rat embryos using a  $^{35}\text{S}$ -labeled cRNA probe for rat clone B1. (A) Parasagittal sections of whole embryos demonstrates that BEHAB expression is restricted to the CNS. In the fore-brain, BEHAB hybridization is present in the neocortex (NC). BEHAB is also expressed in the developing brainstem and spinal cord. In the midbrain, pons and medulla, the highest levels of expression are found in regions that contain actively mitotic cells, the ventral ventricular zone of the midbrain (MB) and the medial tegmental neuroepithelium (MT) of the medulla. (B and C) Cross sections of the spinal cord photographed as an autoradiogram (B) or a negative image of the same section stained with cresyl violet (C)

show in more detail the distribution of BEHAB. BEHAB hybridization is restricted to the ventral portion of the developing spinal cord, with highest levels of expression in the ventral half of the ependymal layer, a region of mitotic activity. There is no detectable BEHAB expression in the peripheral nervous system, as illustrated here for the dorsal root ganglia (DRG). This pair of photomicrographs also demonstrates that cartilage (C), which expresses aggrecan, an mRNA with 60% identity to BEHAB, does not hybridize with the BEHAB probe. The negative cresyl violet stained image was generated by projection printing directly to photographic film. Bars: (A) 1 mm; (B and C) 200  $\mu\text{m}$ .

protein. However, several lines of evidence demonstrate that BEHAB encodes a protein distinct from link protein. First, on Northern blots, at low stringency BEHAB fails to detect a message in chondrosarcoma RNA (data not shown). Second, the reported location of link protein in brain is histologically distinct from that of BEHAB (Ripellino et al., 1989). Third, we have isolated a PCR clone from cat brain that is almost identical to cartilage link protein and is significantly different from cat BEHAB. Therefore, BEHAB is unlikely to encode link protein.

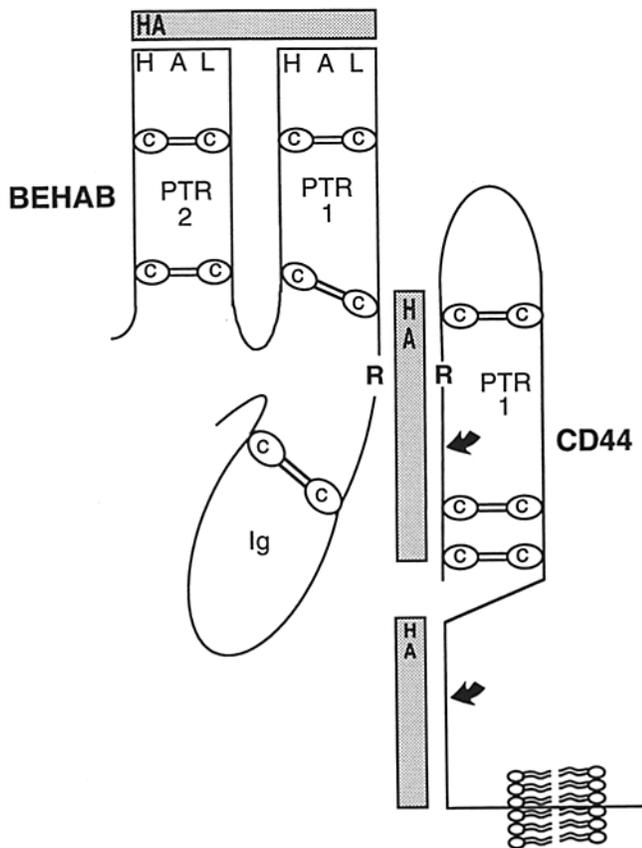
Two HA-binding proteins of approximately the same size as that predicted for BEHAB have been reported in the brain, glial hyaluronate binding protein (GHAP) and hyaluronectin. BEHAB is also distinct from these neural HA-binding proteins. GHAP is a 60-kD glycoprotein isolated from human brain white matter (Perides et al., 1989). The amino acid sequence of GHAP is identical to the  $\text{NH}_2$ -terminal portion of versican, leading to the suggestion that GHAP may be a proteolytic fragment of versican (Zimmermann and Ruoslahti, 1989). Hyaluronectin is a 68-kD glycoprotein purified from CNS white matter, for which only partial amino acid sequence is available. The  $\text{NH}_2$ -terminal 20 amino acids of hyaluronectin are virtually identical to GHAP (Delpech et al., 1991). The sequence of BEHAB clearly differentiates it from GHAP and hyaluronectin. Also, in contrast to BEHAB, which is only detected in the CNS, hyaluronectin is expressed in the peripheral nervous system and in

the kidney, as well as in the CNS (Bignami and Dahl, 1986; Delpech and Delpech, 1984).

#### *The Sequence of BEHAB Predicts an Extracellular HA-binding Protein*

The PTR family of HA-binding proteins can be divided into two classes, HA receptors and extracellular HA-binding proteins (Wight et al., 1991). The PTR proteins that have a transmembrane domain, CD44 and TSG-6, act as HA receptors and have only a single PTR loop. Those that lack a transmembrane domain, and are extracellular proteins, have two PTR loops. The BEHAB gene encodes both a PTR 1 and a PTR 2 loop, and lacks the COOH-terminal hydrophobic domain present in transmembrane proteins, indicating that BEHAB encodes an extracellular HA-binding protein (see Fig. 11).

A HA-binding consensus sequence has not been definitively identified. Binding to HA has been demonstrated for aggrecan (Fryer et al., 1992), versican (LeBaron et al., 1992), GHAP (Perides et al., 1991), link protein (Goetinck et al., 1987), CD44 (Miyake et al., 1990), TSG-6 (Lee et al., 1992), T1 (Iwata et al., 1993), and the Cat-301 antigen (Fryer et al., 1992). HA binding by neurocan (Rauch et al., 1992) is predicted from deduced amino acid sequence. Based on sequence conservation, it has been suggested that the HA-binding region resides in the PTR loops (Doege et



**Figure 11.** Schematic representation of the hypothesized interaction of PTR proteins with HA. Two types of PTR HA-binding protein have been described. Those with a single PTR loop (illustrated here for CD44) and a transmembrane domain (shown at bottom right) are called HA receptors. In contrast, extracellular HA-binding proteins (illustrated here for BEHAB) lack a transmembrane domain and contain two PTR loops. Sequence analysis of BEHAB suggests that it is an extracellular HA-binding protein. Binding to HA (shown as a stippled bar) is mediated either through sequences at the tip of the PTR loops (the HAL sequence), or by clusters of basic amino acids, including an arginine (R) in PTR1. The two clusters of basic amino acids which have been shown to be critical for HA binding in CD44 are indicated with arrows. The Ig loop of the extracellular HA binding proteins may permit interaction with other Ig loop containing proteins (see text for details).

al., 1991; Fosang and Hardingham, 1989; Goetinck et al., 1987). The most conserved region among these proteins is a 15-amino acid sequence CDA GWL(A/S)D(Q/G)(T/S)-VRYPI, referred to as the HAL peptide (using the nomenclature of Iwata and Carlson, 1993), which appears both in loops PTR 1 and PTR 2, of aggrecan, versican, neurocan, and link protein. Predictions of the structure of the PTR proteins suggest that the HAL sequence forms the apex of the PTR loops (see Fig. 11), consistent with the idea that the HAL sequence could have a role in HA binding. BEHAB also has two copies of the HAL sequence, one in each of the predicted PTR loops.

The HAL sequence does not exclusively define proteins that bind HA, because several proteins that bind HA lack the HAL sequence, including T1 (Iwata and Carlson, 1993), CD44 (Stamenkovic et al., 1989), RHAMM (Hardwick et

al., 1992), and TSG-6 (Lee et al., 1992). Data from several reports suggest that HA binding may not be conferred by a specific amino acid sequence, but rather, may be determined by a clustering of basic amino acids (Lyon, 1986; Hardingham et al., 1976): peptides that inhibit link protein binding to HA contain clusters of basic amino acids (Goetinck et al., 1987), and the HA-binding region of CD44 contains two clusters of basic amino acids (Peach et al., 1993). Replacement of basic amino acids in CD44 with alanine using site-directed mutagenesis reduces HA binding. Substitution of even a single basic residue reduces HA binding, while multiple substitutions have an additive effect. One substitution, at Arg-41, almost completely abolishes HA binding. An Arg residue is present at an equivalent position in all PTR proteins, and is also present in BEHAB (Arg-166 in Fig. 4).

These sequence analyses show that BEHAB is similar to all members of the PTR family, including CD44. BEHAB is structurally similar to the PTR proteins, contains the HAL sequence, and has a basic residue (Arg) at the position that has been shown to be critical for HA binding by CD44. These data indicate that BEHAB is a new member of the PTR family of proteins and predict that BEHAB encodes an extracellular HA-binding protein.

### **BEHAB Represents a Neural Specific HA-binding Protein**

We show here by Northern analysis that BEHAB is expressed in the mammalian CNS, but is not expressed in liver, kidney, spleen, lung, or muscle. Many of these non-neural tissues contain high levels of HA and express previously characterized HA-binding proteins. Our developmental studies using in situ hybridization demonstrate that BEHAB expression in the embryo is, similarly, restricted to the CNS.

The distribution of BEHAB in the mature CNS demonstrated by in situ hybridization parallels that reported for HA (Fryer et al., 1992; Fryer, H. J. L., G. M. Kelly, R. Kalb, and S. Hockfield, unpublished results). Biochemical assays of HA show that the corpus callosum contains three times the concentration of HA as the cerebral cortex. Also, the brainstem and spinal cord have higher concentrations of HA than the cerebral cortex (Delpech et al., 1989). Similar to the distribution of HA, a higher level of BEHAB hybridization is detected in the corpus callosum than in the cortex, and the brainstem and spinal cord have higher hybridization signals than the cortex (Fig. 7). In the white matter, clusters of silver grains from BEHAB hybridization are found over astrocytes, which synthesize HA (Delpech et al., 1989). This colocalization provides further support for the possibility that BEHAB is a HA-binding protein in brain.

While HA is abundant in the mammalian CNS, no brain specific HA-binding protein has been reported to date. Several HA-binding proteins are expressed in the brain as well as in other tissues. Versican is expressed in endoneural and endomysial contacts, kidney medulla, precartilagenous mesenchyma, and brain (Bignami et al., 1993). The T1 antigen, which is distributed throughout the brain, is also present in the electoplax and at neuromuscular junctions (Carlson et al., 1986). CD44 is expressed in lymphocytes, epithelium, and brain (Picker et al., 1989). The HA-binding protein neurocan, isolated from brain, is not expressed in liver (Rauch et al., 1992), but the presence of neurocan in tissues that contain high levels of HA and HA-binding proteins, such as

kidney, spleen, lung and muscle, has not been investigated. BEHAB represents the first reported, brain specific HA-binding protein.

Hyaluronan is found virtually in all tissues, so the existence of a tissue-specific HA-binding protein, like BEHAB, might not have been predicted. However the high degree of cellular complexity in the mammalian CNS, as well as the heterogeneity in extracellular matrix proteins in brain (see below), might require a greater diversity in HA-binding proteins than that required in other tissues. The identification of this brain specific HA-binding protein could suggest that other tissues also express tissue- or cell-type specific HA-binding proteins.

### ***The Developmental Regulation of BEHAB Parallels the Elaboration of the Mature CNS Extracellular Matrix***

The 3.9-kb BEHAB transcript is detected in the rat CNS at embryonic day 17 and its level of expression increases over the first two postnatal weeks. Adult levels of expression are reached at P21 in the spinal cord and cortex. In the spinal cord a transient peak occurs at P14; a similar peak may also occur in the cortex between P21 and adulthood. This would be consistent with the delayed maturation of cortex relative to spinal cord.

The existence of an extracellular matrix in brain has been debated. The extracellular space in brain is electron lucent and contains a very low abundance of laminin and collagen, major components of most extracellular matrices. However, studies over the last several years have shown that some elements of extracellular matrices are found in brain, including HA, heparan sulfate proteoglycans, chondroitin sulfate proteoglycans, keratan sulfate proteoglycans, and fibronectin (Margolis and Margolis, 1993; Fryer et al., 1992; Sheppard et al., 1991; Rauch et al., 1991; Hockfield et al., 1990; Herndon and Lander, 1990; Stewart and Pearlman, 1987; Margolis et al., 1975).

The composition of the brain extracellular matrix changes over the course of development. During the embryonic period, newly born neurons migrate to their final destinations and undergo initial dendritic and axonal outgrowth. The cell motility and synaptic refinement that occurs in this period requires a relatively fluid extracellular environment, consistent with our observations of high levels of BEHAB in regions of mitotic activity. During later stages in development the mature structure and synaptic relationships of neurons are established and are maintained, in large measure, throughout the life of an animal. Over the course of the early postnatal period an insoluble extracellular matrix is elaborated in the CNS. Hyaluronan becomes increasingly associated with a water insoluble fraction of brain during the early postnatal period. Ninety percent of HA is water extractable at P7 while only 15% is water extractable in the adult (Margolis et al., 1975). Several components of this insoluble matrix, including the neuronal cell surface proteoglycan recognized by monoclonal antibody Cat-301, and a variety of other cell surface proteoglycans, are initially detected during the first postnatal week and do not reach adult levels until the third or fourth postnatal week (Maeda et al., 1992; Hockfield et al., 1990; Watanabe et al., 1989; Aquino et al., 1984; Hockfield et al., 1983; Hockfield, S., S. Zaremba, and G. Kelly, unpublished results).

While the CNS contains many extracellular matrix components, the interaction of these components needs to be explored further. The extracellular matrix of cartilage contains large insoluble aggregates, formed by the interaction of aggrecan (a chondroitin sulfate proteoglycan) and HA, which are stabilized by link protein (Neame et al., 1986). The interaction between aggrecan and link protein is mediated by the immunoglobulin loops of the two proteins. As in cartilage, chondroitin sulfate proteoglycans in brain could form large aggregates through multivalent interactions with HA and link protein or a link-like protein. The size and sequence of BEHAB are similar, but not identical, to link protein, suggesting that BEHAB could serve a link-like function in the stabilization of proteoglycan-HA aggregates in CNS extracellular matrix. The cell surface proteoglycans in the CNS appear more molecularly heterogeneous than in cartilage and are expressed on distinct populations of neurons (Bertolotto et al., 1991; Watanabe et al., 1989; Fujita et al., 1989; Hockfield and McKay, 1983). Cell-type specific localization (and possibly function) of these proteoglycans could be mediated through immunoglobulin loops. Many classes of extracellular and transmembrane neural proteins contain immunoglobulin loops, including cell adhesion molecules (e.g., N-CAM) and signal transducing molecules (e.g., tyrosine phosphatases). BEHAB could associate with other immunoglobulin loop containing proteins and participate in cell-cell or cell-matrix interactions.

There is a considerable degree of cellular diversity in the mammalian CNS. The recent identification of chondroitin sulfate proteoglycans present on different populations of neurons has shown that the brain extracellular matrix is heterogeneous and suggest that this matrix may be an important determinant of neuron specific traits. The matrix heterogeneity in the CNS could include the expression of small HA-binding proteins that might serve link-like functions. The gene described here, BEHAB, encodes a neural specific protein that, based on sequence analysis, binds HA and lacks a transmembrane domain. These data, along with the time course of expression, suggest that the BEHAB protein is an important component of the CNS extracellular matrix.

We wish to thank Dr. Kurt Doege (Shriner's Hospital for Crippled Children, Portland, OR) for generously supplying aggrecan cDNA clones and rat chondrosarcoma RNA, Dr. John Marshall (Yale University Department of Pharmacology) for the P12 rat brain cDNA library, and Lynna Stone-Infeld (Yale University Biomedical Computing) for GCG training. We also thank Hugh Fryer and Mustafa Sahin for critical review of the manuscript.

This work was supported by EY06511 (S. Hockfield) and EY06451 (D. M. Jaworski).

Received for publication 3 December 1993, and in revised form 26 January 1994.

### ***References***

- Aquino, D. A., R. U. Margolis, and R. K. Margolis. 1984. Immunocytochemical localization of a chondroitin sulfate proteoglycan in nervous tissue. II. Studies in developing brain. *J. Cell Biol.* 99:1130-1139.
- Bertolotto, A., G. Rocca, G. Canavese, A. Migheli, and D. Schiffer. 1991. Chondroitin sulfate proteoglycan surrounds a subset of human and rat CNS neurons. *J. Neurosci. Res.* 29:225-234.
- Bignami, A., and D. Dahl. 1986. Brain-specific hyaluronate-binding protein. A product of white matter astrocytes? *J. Neurocytol.* 15:671-679.
- Bignami, A., G. Perides, and F. Rahemtulla. 1993. Versican, a hyaluronate-binding proteoglycan of embryonal precartilaginous mesenchyma, is mainly expressed postnatally in rat brain. *J. Neurosci. Res.* 34:97-106.
- Bomdareff, W. 1967. An intercellular substance in rat cerebral cortex: submicroscopic distribution of ruthenium red. *Anat. Rec.* 157:527-536.

- Bonnet, F., J. Perin, F. Lorenzo, J. Jolles, and P. Jolles. 1986. An unexpected homology between link proteins of the proteoglycan complex and immunoglobulin-like proteins. *Biochim. Biophys. Acta* 873:152-155.
- Carlson, S. S., P. Caroni, and R. B. Kelly. 1986. A nerve terminal anchorage protein from electric organ. *J. Cell Biol.* 103:509-520.
- Castejon, H. V. 1970. Histochemical demonstration of sulphated polysaccharides at the coat of nerve cells in the mouse central nervous system. *Acta Histochem.* 38:55-64.
- Church, G. M., and W. Gilbert. 1984. Genomic sequencing. *Proc. Natl. Acad. Sci. USA* 81:1991-1995.
- Cole, G. J., and C. F. McCabe. 1991. Identification of a developmentally regulated keratan sulfate proteoglycan that inhibits cell adhesion and neurite outgrowth. *Neuron* 7:1007-1018.
- Culty, M., K. Miyake, P. W. Kincade, E. Silorski, E. C. Butcher, and C. Underhill. 1990. The hyaluronate receptor is a member of the CD44 (H-CAM) family of cell surface glycoproteins. *J. Cell Biol.* 111:2765-2774.
- Danielson, P., S. Forss-Petter, M. A. Brow, L. Calavetta, J. Douglass, R. J. Milner, and J. G. Sutcliffe. 1988. p1B15: a cDNA clone of the rat mRNA encoding cyclophilin. *DNA* 7:261-267.
- Delpech, A., and B. Delpech. 1984. Expression of hyaluronic acid-binding glycoprotein, hyaluronectin, in the developing rat embryo. *Dev. Biol.* 101:391-400.
- Delpech, B., A. Delpech, G. Bruckner, N. Girard, and C. Maingonnat. 1989. Hyaluronan and hyaluronectin in the nervous system. In *The Biology of Hyaluronan*, D. Evered and J. Whelan, editors. John Wiley & Sons, Chichester. 208-232.
- Delpech, B., C. Maingonnat, A. Delpech, P. Maes, N. Girard, and P. Bertrand. 1991. Characterization of a hyaluronic acid-binding protein from sheep brain. Comparison with human brain hyaluronectin. *Int. J. Biochem.* 23:329-337.
- Devereux, J., P. Haerberli, and O. Smithies. 1984. A comprehensive set of sequence analysis programs for the VAX. *Nucleic Acid Res.* 12:387-395.
- DeYoe, E. A., S. Hockfield, H. Garren, and D. Van Essen. 1990. Antibody labeling of functional subdivisions in visual cortex: Cat-301 in V1, V2, V3 and MT of the Macaque monkey. *Visual Neurosci.* 5:67-81.
- Doerge, K., J. R. Hassell, B. Caterson, and Y. Yamada. 1986. Link protein cDNA sequence reveals a tandemly repeated protein sequence. *Proc. Natl. Acad. Sci. USA* 83:3761-3765.
- Doerge, K., M. Sasaki, E. Horigan, J. R. Hassell, and Y. Yamada. 1987. Complete primary structure of the rat cartilage proteoglycan core protein deduced from cDNA clones. *J. Biol. Chem.* 262:17757-17767.
- Doerge, K., C. Rhodes, M. Sasaki, J. R. Hassell, and Y. Yamada. 1990. Extracellular Matrix Genes. Academic Press, Inc., New York. 137-155.
- Doerge, K., M. Sasaki, T. Kimura, and Y. Yamada. 1991. Complete coding sequence and deduced primary structure of the human cartilage large aggregating proteoglycan, aggrecan. *J. Biol. Chem.* 266:894-902.
- Feinberg, A. P., and B. Vogelstein. 1983. A technique for radiolabeling DNA restriction endonuclease fragments to high specific activity. *Anal. Biochem.* 132:6-13.
- Fosang, A. J., and T. E. Hardingham. 1989. Isolation of the NH<sub>2</sub> terminal globular protein domains from cartilage proteoglycans. *Biochem. J.* 261:801-809.
- Fryer, H. J. L., G. M. Kelly, L. Molinaro, and S. Hockfield. 1992. The high molecular weight Cat-301 chondroitin sulfate proteoglycan from brain is related to the large aggregating proteoglycan from cartilage, aggrecan. *J. Biol. Chem.* 267:9874-9883.
- Fujita, S. C., Y. Tada, F. Murakami, M. Hayashi, and M. Matsumura. 1989. Glycosaminoglycan related epitopes surround different subsets of mammalian central neurons. *Neurosci. Res.* 7:117-130.
- Goetinck, P. F., N. S. Stirpe, P. A. Tsonis, and D. Carlone. 1987. The tandemly repeated sequences of cartilage link protein contain the sites for interaction with hyaluronic acid. *J. Cell Biol.* 105:2403-2408.
- Hardingham, T. E., R. J. F. Ewins, and H. Muir. 1976. Cartilage proteoglycans. Structure and heterogeneity of the protein core and the effects of specific protein modifications on the binding to hyaluronate. *Biochem. J.* 157:127-143.
- Hardingham, T. E., and A. J. Fosang. 1992. Proteoglycans: many forms and many functions. *FASEB (Fed. Am. Soc. Exp. Biol.)* 6:861-870.
- Hardwick, C., K. Hoare, R. Owens, H. P. Hohn, M. Hook, D. Moore, V. Cripps, L. Austen, D. M. Nance, and E. A. Turley. 1992. Molecular cloning of a novel hyaluronan receptor that mediates tumor cell motility. *J. Cell Biol.* 117:1343-1350.
- Herdon, M. E., and A. D. Lander. 1990. A diverse set of developmentally regulated proteoglycans is expressed in the rat central nervous system. *Neuron* 4:949-961.
- Hockfield, S. 1990. Proteoglycans in neural development. *Sem. Dev. Biol.* 1:55-63.
- Hockfield, S., and R. D. McKay. 1983. A surface antigen expressed by a subset of neurons in the vertebrate central nervous system. *Proc. Natl. Acad. Sci. USA* 80:5758-5761.
- Hockfield, S., and M. Sur. 1990. Monoclonal antibody Cat-301 identifies Y-cells in the dorsal lateral geniculate nucleus of the cat. *J. Comp. Neuro.* 300:320-330.
- Hockfield, S., R. McKay, S. H. C. Hendry, and E. G. Jones. 1983. A surface antigen that identifies ocular dominance columns in cortical area 17 and laminar features of the lateral geniculate nucleus. *Cold Spring Harbor Symp. Quant. Biol.* 48:877-889.
- Hockfield, S., S. Zaremba, R. G. Kalb, and H. Fryer. 1990. Expression of neural proteoglycans correlates with the acquisition of mature neuronal properties in the mammalian brain. *Cold Spring Harbor Symp. Quant. Biol.* 55:505-514.
- Iwata, M., and S. S. Carlson. 1993. A large chondroitin sulfate proteoglycan has the characteristics of a general extracellular matrix component of adult brain. *J. Neurosci.* 13:195-207.
- Iwata, M., T. N. Wight, and S. S. Carlson. 1993. A brain extracellular matrix proteoglycan forms aggregates with hyaluronan. *J. Biol. Chem.* 268:15061-15069.
- Kozak, M. 1984. Compilation and analysis of sequences upstream from the translational start site in eukaryotic mRNAs. *Nucleic Acids Res.* 12:857-872.
- Kozak, M. 1991. An analysis of vertebrate mRNA sequences: intimations of translational control. *J. Cell Biol.* 115:887-903.
- Laurent, T. C., and J. R. E. Fraser. 1992. Hyaluronan. *FASEB (Fed. Am. Soc. Exp. Biol.)* 6:2397-2404.
- LeBaron, R. G., D. R. Zimmermann, and E. Ruoslahti. 1992. Hyaluronate binding properties of versican. *J. Biol. Chem.* 267:10003-10010.
- Lee, T. H., H.-G. Wisniewski, and J. Vilcek. 1992. A novel secretory tumor necrosis factor-inducible protein (TSG-6) is a member of the family of hyaluronate binding proteins, closely related to the adhesion receptor CD44. *J. Cell Biol.* 116:545-557.
- Lenoir, D., E. Battenberg, M. Kiel, F. E. Bloom, and R. J. Milner. 1986. The brain-specific gene 1B236 is expressed postnatally in the developing rat brain. *J. Neurosci.* 6:522-530.
- Letourneau, P. C., A. M. Madsen, S. L. Palm, and L. T. Furcht. 1989. Immunoreactivity for laminin in the developing ventral longitudinal pathway of the brain. *Dev. Biol.* 125:135-144.
- Lyon, M. 1986. Specific chemical modifications of link protein and their effect on binding to hyaluronate and cartilage proteoglycan. *Biochim. Biophys. Acta* 881:22-29.
- Maeda, N., F. Matsui, and A. Oohira. 1992. A chondroitin sulfate proteoglycan that is developmentally regulated in the cerebellar mossy fiber system. *Dev. Biol.* 151:564-574.
- Mandel, M., and A. Higa. 1970. Calcium-dependent bacteriophage DNA infection. *J. Mol. Biol.* 53:159-162.
- Margolis, R. K., and R. U. Margolis. 1993. Nervous tissue proteoglycans. *Experientia* 49:429-445.
- Margolis, R. U., R. K. Margolis, L. B. Chang, and C. Preti. 1975. Glycosaminoglycans of brain during development. *Biochemistry* 14:85-88.
- Marshall, R. D. 1974. The nature and metabolism of the carbohydrate-peptide linkage of glycoproteins. *Biochem. Soc. Symp.* 40:17-26.
- Martin, K. A., S. G. N. Grant, and S. Hockfield. 1992. The mas proto-oncogene is developmentally regulated in the rat central nervous system. *Dev. Brain Res.* 68:75-82.
- McLoon, S. C., L. K. McLoon, S. L. Palm, and L. T. Furcht. 1988. Transient expression of laminin in the optic nerve of the developing rat. *J. Neurosci.* 8:1981-1990.
- Miyake, K., C. B. Underhill, J. Lesley, and P. W. Kincade. 1990. Hyaluronate can function as a cell adhesion molecule and CD44 participates in hyaluronate recognition. *J. Exp. Med.* 172:69-75.
- Nakanishi, S. 1983. Extracellular matrix during laminar pattern formation of neocortex in normal and reeler mutant mice. *Dev. Biol.* 95:305-316.
- Neame, P. J., J. E. Christner, and J. R. Baker. 1986. The primary structure of link protein from rat chondrosarcoma proteoglycan aggregate. *J. Biol. Chem.* 261:3519-3523.
- Peach, R. J., D. Hollenbaugh, I. Stamenkovic, and A. Aruffo. 1993. Identification of hyaluronic acid binding sites in the extracellular domain of CD44. *J. Cell Biol.* 122:257-264.
- Perides, G., F. Biviano, and A. Bignami. 1991. Interaction of a brain extracellular matrix protein with hyaluronic acid. *Biochim. Biophys. Acta* 1075:248-258.
- Perides, G., W. S. Lane, D. Andrews, D. Dahl, and A. Bignami. 1989. Isolation and partial characterization of a glial hyaluronate-binding protein. *J. Biol. Chem.* 264:5981-5987.
- Perkins, S. J., A. S. Nealis, J. Dudhia, and T. E. Hardingham. 1989. Immunoglobulin fold and tandem repeat structures in proteoglycan NH<sub>2</sub> terminal domains and link protein. *J. Mol. Biol.* 206:737-753.
- Perris, R., and S. Johansson. 1990. Inhibition of neural crest cell migration by aggregating chondroitin sulfate proteoglycans is mediated by their hyaluronan-binding region. *Dev. Biol.* 137:1-12.
- Picker, L. J., M. Nakache, and E. C. Butcher. 1989. Monoclonal antibodies to human lymphocyte homing receptors define a novel class of adhesion molecules on diverse cell types. *J. Cell Biol.* 109:927-937.
- Prieto, A. L., C. Andersson-Fisone, and K. L. Crossin. 1992. Characterization of multiple adhesive and counter adhesive domains in the extracellular matrix protein cytotactin. *J. Cell Biol.* 119:663-678.
- Rauch, U., P. Gao, A. Janetzko, A. Flaccus, L. Hilgenberg, H. Tekotte, R. K. Margolis, and R. U. Margolis. 1991. Isolation and characterization of developmentally regulated chondroitin sulfate and chondroitin/keratan sulfate proteoglycans of brain identified with monoclonal antibodies. *J. Biol. Chem.* 266:14785-14801.

- Rauch, U., L. Karthikeyan, P. Maurel, R. U. Margolis, and R. K. Margolis. 1992. Cloning and primary structure of neurocan, a developmentally regulated, aggregating chondroitin sulfate proteoglycan of brain. *J. Biol. Chem.* 267:19536-19547.
- Reichardt, L. F., and K. J. Tomaselli. 1991. Extracellular matrix molecules and their receptors: functions in neural development. *Annu. Rev. Neurosci.* 14:531-570.
- Ripellino, J. A., R. U. Margolis, and R. K. Margolis. 1989. Immunoelectron microscopic localization of hyaluronic acid-binding region and link protein epitopes in brain. *J. Cell Biol.* 108:1899-1907.
- Sambrook, J., E. F. Fritsch, and T. Maniatis. 1989. Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory Press, Cold Spring Harbor. 1.90-1.104.
- Sanes, J. R. 1989. Extracellular matrix molecules that influence neural development. *Annu. Rev. Neurosci.* 12:491-516.
- Sanger, F., S. Nicklen, and A. R. Coulson. 1977. DNA sequencing with chain-terminating inhibitors. *Proc. Natl. Acad. Sci. USA.* 74:5463-5467.
- Schmitt, H., I. Gozes, and U. Z. Littauer. 1977. Decrease in levels and rates of synthesis of tubulin and actin in developing rat brain. *Brain Res.* 121:327-342.
- Schwartz, R. M., and M. O. Dayhoff. 1979. Atlas of protein sequence and structure. National Biomedical Research Foundation, Washington, DC. 353-358.
- Sheppard, A. M., S. K. Hamilton, and A. L. Pearlman. 1991. Changes in the distribution of extracellular matrix components accompany early morphogenetic events of mammalian cortical development. *J. Neurosci.* 11:3928-3942.
- Snow, D. M., D. A. Steindler, and J. Silver. 1990. Molecular and cellular characterization of the glial roof plate of the spinal cord and optic tectum: a possible role for a proteoglycan in the development of an axon barrier. *Dev. Biol.* 138:359-376.
- Stamenkovic, I., M. Amiot, J. M. Pesando, and B. Seed. 1989. A lymphocyte molecule implicated in lymph node homing is a member of the cartilage link protein family. *Cell.* 56:1057-1062.
- Stewart, G. R., and A. L. Pearlman. 1987. Fibronectin-like immunoreactivity in the developing cerebral cortex. *J. Neurosci.* 7:3325-3333.
- Toole, B. P. 1991. Proteoglycans and Hyaluronan in Morphogenesis and Differentiation. In *Cell Biology of Extracellular Matrix*. E. D. Hay, editor. Plenum Press, New York. 305-341.
- Watanabe, E., S. C. Fujita, F. Murakami, M. Hayashi, and M. Matsumura. 1989. A monoclonal antibody identifies a novel epitope surrounding a sub-population of the mammalian central neurons. *Neurosci.* 29:645-657.
- Wight, T. N., D. K. Heinegard, and V. C. Hascall. 1991. Proteoglycans: structure and function. In *Cell Biology of Extracellular Matrix*. E. D. Hay, editor. Plenum Press, New York. 45-78.
- Williams, A. F., and A. N. Barclay. 1988. The immunoglobulin superfamily-domains for cell surface recognition. *Ann. Rev. Immunol.* 6:381-405.
- Zimmermann, D. R., and E. Ruoslahti. 1989. Multiple domains of the large fibroblast proteoglycan, versican. *EMBO (Eur. Mol. Biol. Organ.). J.* 8:2975-2981.
- von Heijne, G. 1986. A new method for predicting signal sequence cleavage sites. *Nucleic Acids Res.* 14:4683-4690.