BEN, a surface glycoprotein of the immunoglobulin superfamily, is expressed in a variety of developing systems

(HNK-1/cell adhesion)

OLIVIER POURQUIÉ*, CATHERINE CORBEL*, JEAN-PIERRE LE CAER†, JEAN ROSSIER†, AND NICOLE M. LE DOUARIN*

*Institut d'Embryologie Cellulaire et Moléculaire, 49 bis Avenue de la Belle Gabrielle, 94736 Nogent sur Marne, Cedex, France; and †Laboratoire de Physiologie Nerveuse, Centre National de la Recherche Scientifique, 91198 Gif sur Yvette, Cedex, France

ABSTRACT We have previously identified a 95- to 100-kDa cell surface glycoprotein, which we named BEN (for bursal epithelium and neurons), that is widely expressed during chicken embryonic development. In the central nervous system, it is restricted to subsets of neurons including the motoneurons and the inferior olivary nucleus neurons (which provide the cerebellum with the climbing fibers) where its expression occurs during the phase of axonogenesis and synaptogenesis. In the present work, we show that BEN expression extends to a variety of tissues originating from the three embryonic germ layers. We have found that BEN immunopurified from neural, epithelial, and hemopoietic tissues is differently glycosylated and may or may not carry the HNK-1 epitope. We then cloned a full-length cDNA encoding this protein. Analysis of its sequence reveals that BEN is a member of the immunoglobulin superfamily. Two molecules with an identical cDNA sequence were recently reported: DM-GRASP [Burns, F. R., Von Kannen, S., Guy, L., Raper, J. A., Kamholz, J. & Chang, S. (1991) Neuron 7, 209–220] and SC1 [Tanaka, H., Matsui, T., Agata, A., Tomura, M., Kubota, I., Mcfarland, K. C., Kohr, B., Lee, A., Phillips, H. S. & Shelton, D. L. (1991) Neuron 7, 535–545]. Their pattern of expression and structural properties are consistent with those reported for BEN. Therefore BEN, DM-GRASP, and SC1 are likely to be the same molecule of the immunoglobulin superfamily.

By raising a mouse monoclonal antibody (mAb) against epithelial cells of the chicken bursa of Fabricius, we discovered a surface glycoprotein well represented on the bursal epithelium that provides the stromal cells on which B-lymphocyte differentiation takes place in the avian class of vertebrates. Unexpectedly though, as we investigated the degree of tissue specificity of the antigen thus disclosed, we found that this mAb reveals an antigenic determinant borne by a 95- to 110-kDa membrane glycoprotein (1) also present on various neuronal subsets in both the central and the peripheral nervous system. We therefore designated this molecule by the acronym BEN, which stands for bursal epithelium and neurons. Thus, in the spinal cord, it is present on motoneurons during neurite extension and synaptogenesis, after which its synthesis is down-regulated (1). In the brain, we have described a similar phenomenon for the climbing fibers originating in the inferior olivary nucleus at the time when they are growing to the cerebellar cortex (2). One particularity of BEN was therefore its association with axonal growth and establishment of neural connections. In contrast, in the sympathetic and parasympathetic ganglia and fibers and in the enteric intrinsic innervation, BEN expres-

sion remains constant, at least during embryogenesis. In the course of our investigation of BEN expression in the chicken embryo, we also found that it is present on hemopoietic cells. BEN is present on thymocytes and on peripheral activated T cells (3) as well as on subsets of bone marrow myeloid and erythroid progenitors (C.C., O.P., and F. Cormier, unpublished data). On the whole, BEN is restricted to certain differentiating hemopoietic precursor cells and is absent from the B-cell lineage at all stages as well as from mature hemopoietic cells except activated T cells. We now report on the further study of the BEN expression pattern in various developing systems, such as epithelial and skeletal structures, and we describe its biochemical and molecular characteristics. BEN was purified from various types of tissues: epithelia, central nervous system, and hemopoietic cells. This revealed that each developing system carries differently glycosylated molecular forms of BEN. We cloned a full-length cDNA encoding this protein and found, from the sequence analysis, that BEN is a member of the immunoglobulin superfamily.‡ Together with the human melanoma marker MUC18 (4), we propose that it constitutes an additional structural subfamily. While we were completing our sequence analysis, we became aware of the publication of the sequences of two molecules identified in chicken, DM-GRASP involved in neurite–neurite interactions (5) and SC1, which was shown to be a homophilic adhesion molecule (6), to which BEN is almost completely identical. Therefore, BEN, DM-GRASP, and SC1 are very likely to represent the same adhesion molecule.

MATERIALS AND METHODS

Embryos, Antibodies, and Immunocytochemistry. Fertilized eggs from chicken (Gallus gallus) and quail (Coturnix coturnix japonica) were obtained from commercial sources and were incubated in a rotary incubator at 38°C. Stages of development of the embryos were expressed according to Hamburger and Hamilton (7) or in days of incubation for later stages (Ex = embryonic day x). Immunocytochemistry using anti-BEN mAb was performed as described (1, 2).

Cell Culture and Confocal Microscopy. Cell suspensions of the ventral part of the hindbrain of E13 chicken embryos were obtained by trypsin dissociation. BEN-positive cells were sorted out with anti-BEN mAb by using the panning procedure described by Bloch-Gallego et al. (8). Staining with BEN mAb was carried out as described (1) and analyzed using a confocal microscope (Molecular Dynamics, Sunnyvale, CA).

Abbreviations: mAb, monoclonal antibody; Ex, embryonic day x; V, variable; C, constant; BISANCE, base informatique sur les séquences d'acides nucléiques pour les chercheurs Européens.
‡The cDNA sequence reported in this paper has been deposited in the GenBank data base (accession no. X64301).

Immunopurification of BEN Glycoprotein and Microsequencing. Anti-BEN mAb was purified from ascites on a DEAE-Trisacryl column as described (9). The protocols reported by Schneider *et al.* (10) and Bastiani *et al.* (11) were used for the fabrication of the anti-BEN mAb affinity columns and the subsequent purification of BEN protein. The standard purification factor was 10^6 with a yield of 20 μg of recovered protein for 20 g of E5 embryos. For N-terminus sequencing, 10 μg of affinity-purified protein was blotted on a poly(vinylidene difluoride) membrane and directly sequenced by using a gas-phase microsequencer (Applied Biosystems). For peptide analysis, 30 μg of protein electroeluted with a Bio-Rad eluter was deglycosylated with 300 ng of endoglycosidase F (Boehringer Mannheim) in 50 mM NH$_4$HCO$_3$ overnight at 37°C and digested by adding 2 μg of endoproteinase Asp-N (Boehringer Mannheim); the reaction was incubated for 6 h at 37°C. Peptides were purified on a reverse-phase HPLC and sequenced. The N terminus of the protein was sequenced several times, and the sequence obtained was LYTVNAVYGDTITMPXSLV. Three internal sequences of peptides obtained by two independent digestions were 1,DDVPDYK; 2,DVSAISIPE; and 3,DL-GNIEENKKLEEN. None of those sequences was found in the National Biomedical Research Foundation (release no. 29) or Genpro (release no. 68) data bases using the Kanehisa program for alignment of a protein sequence with a data base (12) from the base informatique sur les séquences d'acides nucléiques pour les chercheurs Européens (BISANCE) facility.

Molecular Biology and Sequence Analysis. The cloning strategy was derived from the mixed oligonucleotide-primed amplification of cDNA procedure described by Lee *et al.* (13). Degenerative oligonucleotides derived from the amino acid sequence of the proteolytic fragments were used as primers to amplify a cDNA fragment by PCR. Sense primers BEN4 and BEN2 were derived from overlapping sequences from the N-terminus portion. Antisense primers BI1, BI2, and BI3 were derived, respectively, from internal sequences 1, 2, and 3. cDNA was generated from 40 μg of total RNA from E16 spinal cord, E5 embryo, and the IIIC5 T-cell line (14) by using BI1, BI2, or BI3 as primers for the reverse transcriptase. The PCR reaction was carried in a thermal cycler with the following program: 94°C for 45 sec, 40°C for 1 min, 72°C for 3 min for 30 cycles. A first amplification step using BEN4 and either BI1, BI2, or BI3 as primers was carried out. Analysis of the reaction product on agarose gel showed a smear with several bands for each amplification reaction. A second amplification step was then carried out on this PCR product by using the BEN2 primer overlapping with BEN4 and the BI1, BI2, and BI3 primers. A distinct band of ≈200 base pairs was observed in the three tissues examined for only the BEN2/BI1 combination. This band hybridized in Southern blot to 197, an internal probe derived from the C-terminal part of the N-terminal sequence. This fragment, called BEN160, was then sequenced and proved to contain both the N-terminal sequence and the sequence of the proteolytic fragment DDVPDYK in an open reading frame. We have then hybridized BEN160 to a Northern blot containing mRNA from E16 bursa of Fabricius and adult brain, liver, and thymus. It strongly recognized a doublet of 3.5–4.0 kilobases in the bursa of Fabricius, where the protein is abundantly expressed, and a fainter signal of the same size in the brain and thymus but not in the liver (data not shown). Total RNA was purified from 60 E16 bursa of Fabricius (15), and mRNA was prepared by two purification cycles on an oligo(dT) (Pharmacia) column. Twenty micrograms of this mRNA was sent to Invitrogen where a cDNA library was constructed in the pCDM8 vector. This library was screened with BEN160, and three clones were isolated. Fragments of the clones were subcloned in the pGEM3Z vector and

sequenced by using specific oligonucleotides with Sequenase (United States Biochemical). Standard molecular biology techniques were performed as described by Sambrook *et al.* (16). Sequence analysis was performed by using the PC-GENE facility (IntelliGenetics) and the BISANCE facility (17). The sequences of the oligonucleotides used for the cloning are shown below: BEN4,5'-GGGGAATTCTA(T-C)AC(IC)GT(ACGT)AA(TC)GC(ACGT)GT-3'; BEN2, 5'-GGGGAATTCTAGAA(TC)GCIGTITA(TC)GGIGA(T-C)AC-3'; BI1, 5'-GAACTAGTCGACTGCAGTT(AG)-TA(AG)TC(ACGT)GG(ACGT)AC(AG)TC(AG)TC-3'; 197, 5'-GGCATATIGT(AGT)AT(ACGT)GT(AG)TC-3'.

RESULTS

Expression of BEN on Epithelial and Preosteogenic Cells. BEN expression was studied on serial sections of whole embryos up to E10 and on isolated organs at later stages. BEN was detected on epithelia of various embryonic origin. The peritoneal epithelium, the external layer of the epidermis, and the amnios were strongly BEN positive until E10, the last time point of our observation. A strong expression was correlated with placode-associated morphogenesis. For example, defined patches of reactivity were observed on placodal epithelium during ear, nose, or head ganglia formation, particularly in the neurogenic regions. A striking feature of BEN expression on epithelial cells was its regionalized and transient pattern. BEN was also transiently expressed on various epithelia of endodermal origin, like the gut (Fig. 1A) and the bronchial epithelium but was never seen on others, such as hepatocytes. Reactivity with anti-BEN mAb was detected on cells localized in certain osteogenic areas. For example, in the head, BEN reactivity was first found at E4 in membrane bone primordia. Later, as ossification proceeded, BEN expression became localized at the periphery of the forming bone—i.e., the periosteum (Fig. 1B). BEN expression was also detected in E4–E6 embryos around the notochord where the vertebral body is being formed, but not in the rest of the vertebra. Reactivity was not observed in the limb bones.

Subcellular Localization Analyzed by Confocal Microscopy in Inferior Olivary Nucleus Cultures. The developmental pattern of BEN expression and the fact that it bears the HNK-1 epitope, which is thought to be implicated in adhesion phenomena in the nervous system (18), led us to speculate that BEN could be implicated in cell adhesion events such as axonal fasciculation or cell aggregation. We examined the distribution of BEN protein on the cell surface of E13 inferior olivary nucleus neurons cultured for 24 h by using confocal microscopy. BEN was rarely uniformly distributed at the cell surface but was observed on areas involved in cell interactions such as growth cones, neurites, or regions of cell–cell contact. Thus, as shown in Fig. 2, BEN immunoreactivity is strikingly concentrated in the area of contact between cells, as expected for an homophilic cell adhesion molecule.

Glycosylation Analysis. In this work, BEN protein purified from various organs—E13 brain, posthatching day 21 bursa of Fabricius, posthatching day 21 thymus, and E18 spleen—was characterized (Fig. 3). SDS/PAGE analysis followed by silver staining of the various glycoproteins revealed heterogeneous molecular masses. BEN exists at least under three different forms: a heavy epithelial form of 110 kDa, an intermediate hematopoietic form of 100 kDa, and a neural form of 95 kDa (Fig. 3A). Treatment with endoglycosidase F yields a product with the same apparent molecular mass (75 kDa) in the four tissues examined, showing that this heterogeneity is mostly due to N-glycosylation (Fig. 3B). BEN is one of the surface molecules reacting with the HNK-1 mAb (1), which recognizes a glycosylated epitope differentially

expressed on a variety of cells depending on the developmental stage (19–21). The HNK-1 epitope was revealed by Western blotting on the BEN protein extracted from thymus and brain but not from spleen and bursa of Fabricius (Fig. 3C).

Isolation of BEN-Encoding cDNA and Sequence Analysis. Three clones of 1, 2, and 3.5 kilobases were obtained from the screening of an E16 bursa of Fabricius cDNA library with the BEN160 probe and were sequenced. The longest contained an open reading frame coding for 589 amino acids (Fig. 4) including proteolytic sequences obtained from microsequencing and a poly(T) tail. The hydrophobicity profile of the deduced amino acid sequence (not shown) predicts a first hydrophobic stretch of 32 amino acids, which is likely to correspond to a signal peptide. A second hydrophobic sequence of 25 amino acids is observed between positions 499 and 523. It is likely to correspond to a transmembrane domain. The predicted molecular mass of the mature protein sequence is 62 kDa, which is lower than the molecular mass observed after endoglycosidase F treatment. Moreover, the search for consensus sites on the protein sequence using the Prosite data base (release no. 7) indicates that there are eight potential sites of N-glycosylation and one site for glycosaminoglycan attachment (Fig. 4), which can account for the high percentage of glycosylation observed. No RGD motif was detected in the protein sequence. No consensus sequences such as tyrosine, serine, or threonine phosphorylation sites or a cytoskeleton attachment site were detected in the cytoplasmic domain using the Prosite data base.

BEN Belongs to the Immunoglobulin Superfamily. Comparison of this sequence with the National Biomedical Research Foundation and Genpro data bases using the Kanehisa alignment program (12) shows that this protein is related to the
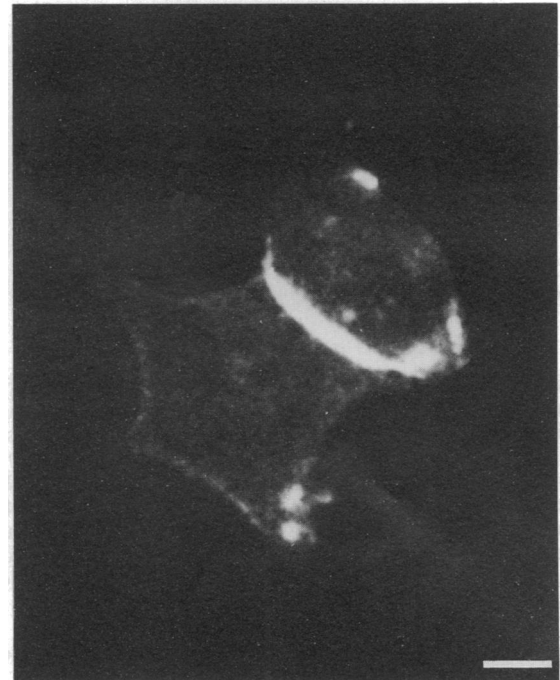
FIG. 2. Confocal image of cells of the inferior olivary nucleus cultured for 24 h and stained with anti-BEN mAb and a secondary antibody coupled to fluorescein isothiocyanate. The detail of two cells showing the distribution of the reactivity in the zone of intercellular contact is shown. (Bar = 1 μm.)

immunoglobulin superfamily. But no homology >25% was found in the data bases. The extracellular domain of the BEN
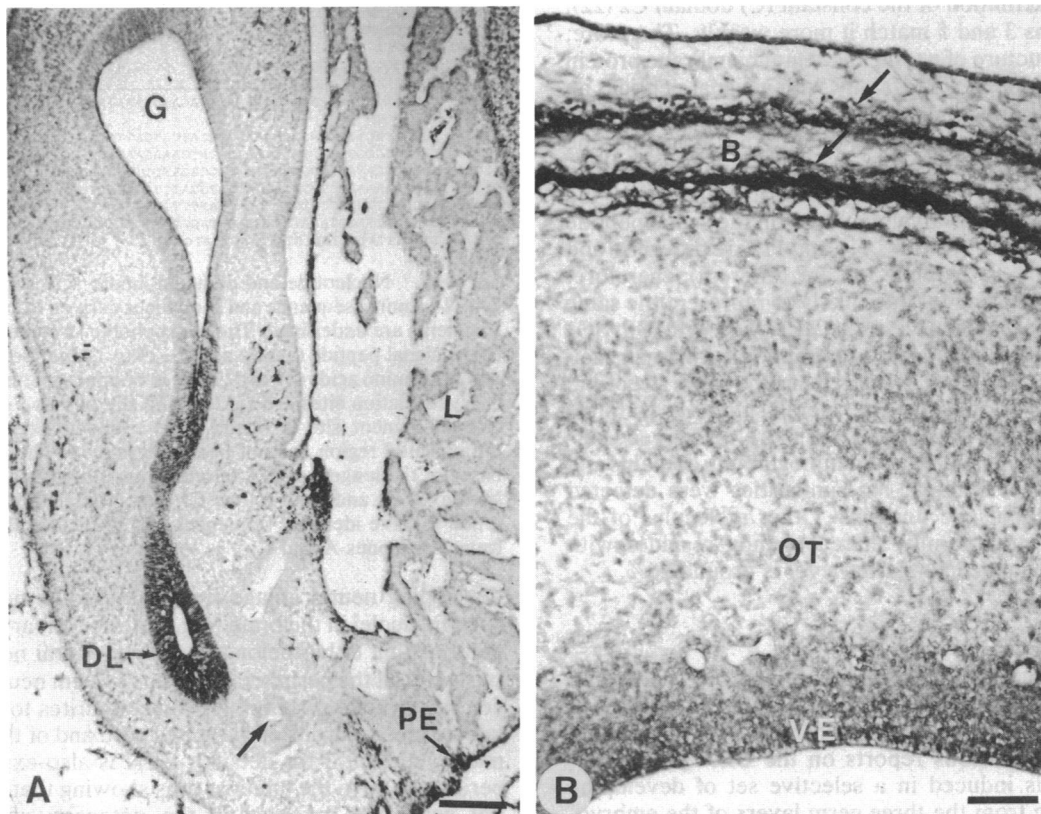
FIG. 1. Expression of BEN on various cell types as evidenced with the anti-BEN mAb visualized with the peroxidase reaction. (A) Transverse section of an E5 chicken embryo. BEN is expressed on the portion of the endoderm that will give rise to the duodenal loop (DL). Gizzard (G) and liver (L) are devoid of reactivity, whereas the peritoneal epithelium (PE) is positive. (Bar = 10 μm.) (B) Sagittal section of the head of an E9 chicken embryo at the level of the optic tectum (OT). Immunoreactivity is concentrated on both sides of the forming bone (B). Arrows point to the osteogenic zone, which is strongly reactive. VE, ventricular epithelium. (Bar = 5 μm.)
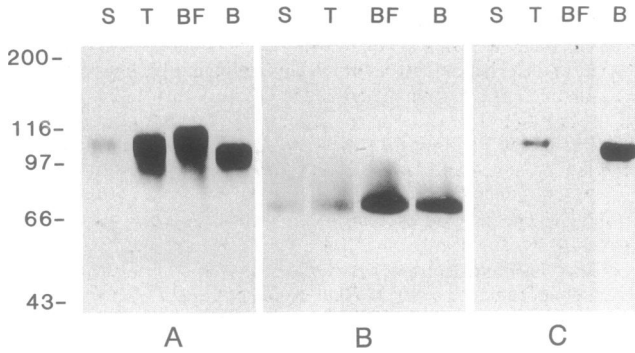
FIG. 3. SDS/PAGE analysis, under reducing conditions, of molecular heterogeneity between BEN isoforms. The same quantities of protein were loaded on the gels in *A*, *B*, and *C*: 1 μg for thymus (T), bursa of Fabricius (BF), and brain (B) and 100 ng for spleen (S). (*A*) Silver-stained gel showing the protein immunopurified from different tissues. (*B*) Silver-stained gel with purified BEN proteins from different tissues after endoglycosidase F treatment. (*C*) Western blot of the same gel as in *A* revealed with HNK-1 mAb. The lower quantity of protein loaded for the spleen may account for the absence of observed reactivity.

protein was found, by manual alignment based on cysteine position, to be composed of five repeated subunits of the immunoglobulin type. The first two domains have their cysteines spaced by 69 and 62 amino acids, respectively, whereas the third, fourth, and fifth domains have their cysteines spaced by 42, 37, and 49 amino acids. On the basis of domain size and residue conservation, domains 1 and 2 match the definition of the variable (V)-set domains of the immunoglobulin superfamily, even though they miss the usual salt bridge between the D and E β-sheets. Domain 4 closely fits the definition of the constant (C) domain C2 (22), whereas domains 3 and 5 match it more weakly. Therefore, BEN has the structure of a typical integral membrane protein of the immunoglobulin superfamily with a signal peptide, an extracellular domain composed of two V-type domains and three C2-type domains, followed then by a transmembrane domain and a short cytoplasmic tail (33 amino acids). The BEN sequence was found to be almost identical to the sequence of DM-GRASP (5 amino acids in the signal peptide and the amino acids in positions 111 and 112 are different) and to SC1 (amino acids 296 and 368 are different); therefore, BEN, DM-GRASP, and SC1 are likely to represent the same molecule. The closest relative of the BEN/DM-GRASP/SC1 protein in the immunoglobulin superfamily is a human melanoma marker called MUC18, which has the same structural arrangement V-V-C2-C2-C2. The overall similarity with this molecule is 25%, but some parts are more conserved, particularly in the second and third immunoglobulin domains and the transmembrane domain. No similarities were detected between the cytoplasmic domains. Other molecules of the immunoglobulin superfamily share common residues with BEN, but the sequence similarity level is around 20%.

DISCUSSION

In this work, we have further documented the spatial and temporal variation of BEN expression in different developing systems. The data confirm the two main aspects already disclosed in our previous reports on the BEN protein expression. (*i*) It is induced in a selective set of developing systems deriving from the three germ layers of the embryo. (*ii*) Its temporal pattern of expression is strictly regulated during development. No general rule can be drawn about the nature of the systems in which BEN is expressed during ontogeny, as far as their embryonic origin is concerned. BEN has been found to be abundantly present in parts of the

```
   1    CTGCTCCGGTGCGGAGTCGCTCCTCGCCGCCGCCGCCGAGAGGGAGAGCAGCCGCCCCGCCAC
  64    TTGCCGCGCCCGCGCACGGCCGCGCGAGGAGGAGAAGTTGCCTCCGCGCTCCGCAGCCGGAGGG
 127    AAGAGCCCCCGGGCTGCCGCCGCCGCCTCGCCGCCTCCCCGTCGCCTCCCGGTCCCTCCGCGG
 -33                   M  M  E  P  P  A  A  A  A  R  A  S  C  R  R  R  P  L
 190    GCAGGCGGGATGATGGAGCCCCCCGCCGCCGCCGCCCGCGCCTCCTGCCGCCGCCGCCCGCTG
 -15     L  C  L  L  L  A  A  L  C  M  P  P  A  L  G  L  Y  T  V  N  A
 253    CTCTGCCTGCTGCTCGCAGCCCTCTGCATGCCGCGCCAGCTCTGGGATTGTATACAGTAAATGCA
   7     V  Y  G  D  T  I  I  T  M  P  C  R  L  E  V  P  D  G  L  M  F  G
 316    GTATATGGAGACACTATTACTATGCCTTGTCGTCTAGAAGTACCAGATGGTCTGATGTTTGGA
  28     K  W  K  Y  E  M  P  N  G  S  P  V  F  I  A  F  R  S  S  T  K
 379    AAATGGAAATATGAAATGCCAAACGGATCTCCAGTATTTATCGCCTTCAGATCATCAACAAAA
  49     K  N  V  Q  Y  D  D  V  P  D  Y  K  D  R  L  S  L  S  E  (N) Y
 442    AAGAATGTACAGTATGACGATGTACCAGACTACAAAGACAGGTTGAGTCTCTCAGAAAACTAT
  70     T  L  S  I  K  N  A  R  I  S  D  E  K  R  F  V  C  M  L  V  T
 505    ACATTATCCATCAAGAATGCAAGAATCAGCGATGAAAAGAGATTTGTATGCATGCTAGTAACA
  91     E  D  D  V  S  E  E  P  T  V  V  K  V  F  K  Q  P  S  Q  P  E
 568    GAAGATGATGTTTCCGAGGAGCCTACGGTAGTCAAAGTCTTCAAACAACCCTCTCAACCAGAA
 112     I  L  H  Q  A  D  F  L  E  T  E  K  L  K  M  L  G  E  C  V  V
 631    ATCTTGCATCAAGCAGACTTCTTAGAAACAGAGAAGCTAAAAATGCTCGGGGAGTGTGTTGTG
 133     R  D  S  Y  P  E  G  (N) V  T  W  Y  K  N  G  R  V  L  Q  P  V
 694    AGGGATAGCTATCCTGAAGGCAATGTTACGTGGTACAAAAATGGGAGAGTTTTGCAGCCTGTG
 154     E  E  V  V  V  I  N  L  R  K  V  (N) R  S  T  G  L  F  T  M
 757    GAAGAAGTTGTGGTCATAAATTTGCGAAAGGTTGAAAACAGATCAACTGGACTCTTCACCATG
 175     T  S  S  L  Q  Y  M  P  T  K  E  D  A  N  A  K  F  T  C  I  V
 820    ACATCATCTCTTCAGTACATGCCAACAAAAGAAGATGCAAATGCTAAATTTACTTGCATTGTG
 196     T  Y  H  G  P  S  G  Q  K  T  I  Q  S  E  P  V  V  F  D  V  H
 883    ACGTATCATGGACCATCTGGCCAGAAAACAATACAGTCTGAACCAGTTGTCTTTGATGTTCAC
 217     Y  P  T  E  K  V  T  I  R  V  L  S  Q  S  S  T  I  K  E  G  D
 946    TATCCAACAGAGAAGGTGACTATTCGAGTGCTGTCACAAAGTAGTACCATTAAAGAAGGTGAT
 238     (N) V  T  L  K  C [S  G  N  G] N  P  P  Q  E  F  L  F  Y  I
1009    AACGTCACTCTGAAGTGCTCAGGGAATGGCAACCCTCCTCCACAGGAGTTCCTGTTTTACATT
 259     P  G  E  T  E  G  I  R  S  S  D  T  Y  V  M  T  D  V  R  R  (N)
1072    CCAGGAGAAACAGAAGGTATAAGAAGCTCAGATACTTATGTAATGACAGATGTGAGACGAAAT
 280     A  T  G  E  Y  K  C  S  L  I  D  K  S  M  M  D  T  T  T  I  T
1135    GCAACAGGAGAATACAAGTGTTCTCTGATTGACAAAAGCATGATGGACACAACCACCATCACT
 301     V  H  Y  L  D  L  Q  L  T  P  S  G  E  V  T  K  Q  I  G  E  A
1198    GTGCACTATCTGGATTTGCAGCTTACTCCTAGTGGAGAAGTCACAAAACAGATTGGAGAGGCC
 322     L  P  V  S  C  T  I  S  S  R  (N) A  T  V  F  W  I  K  D  N
1261    CTGCCTGTGTCGTGCACAATTTCTTCTAGTAGAAATGCAACTGTGTTTTGGATAAAGGACAAT
 343     T  R  M  K  T  S  P  S  F  S  S  L  Q  Y  Q  D  A  G  N  Y  I
1324    ACCAGAATGAAAACAAGTCCGTCATTTTCAAGTCTGCAGTATCAAGATGCTGGAAACTACATC
 364     C  E  T  T  H  K  E  V  E  G  L  K  K  R  K  T  L  K  L  I  V
1387    TGTGAAACTACTCACAAGGAAGTTGAAGGGCTAAAGAAAAGAAAGACACTTAAACTTATAGTG
 385     E  G  K  P  Q  I  K  M  T  K  K  T  N  T  N  K  M  S  K  T  I
1450    GAAGGAAAAACCTCAAATCAAAATGACAAAGAAAACCAACACAAATAAAATGTCTAAAACAATT
 406     V  C  H  V  E  G  F  P  K  P  A  V  Q  W  T  V  T  G  S  G  S
1513    GTCTGTCATGTGGAAGGATTTCCCAAACCAGCAGTGCAATGGACAGTTACTGGTAGTGGAAGC
 427     L  I  N  K  T  E  E  T  K  Y  V  N  G  K  F  S  S  K  I  I  I
1576    CTCATAAATAAGACAGAGGGAGACCAAATATGTTAATGGAAAATTTTCCAGTAAAATTATAATT
 448     A  P  E  E  (N) V  T  L  T  C  I  A  E  N  E  L  E  R  T  V  T
1639    GCTCCTGAGGAAAATGTAACATTAACATGCATTGCAGAAAATGAGCTAGAAAGGACTGTGACC
 469     S  L  (N) V  S  A  I  S  I  P  E  Y  D  E  P  E  D  R  N  D  D
1702    TCTCTGAATGTCTCTGCTATAAGTATTCCAGAATATGATGAACCGGAGGATAGAAATGATGAC
 490     N  S  E  K  V  N  D  Q  A  K  L  I  V  G  L  V  G  L  L  L
1765    AACAGCGAAAAGGTTAATGACCAGGCAAAGCTAATCGTGGGGATTGTAGTCGGTCTTCTGCTG
 511     V  A  L  V  A  G  V  V  Y  W  L  Y  V  K  K  S  K  T  A  S  K
1828    GTTGCCCTGGTTGCAGGTGTTGTCTACTGGCTCTATGTGAAGAAATCAAAGACAGCATCAAAA
 532     H  V  D  K  D  L  G  N  I  E  E  N  K  K  L  E  E  N  N  H  K
1891    CATGTAGACAAAGATCTCGGAAATATAGAAGAAAATAAAAAACTAGAAGAAAACAATCATAAA
 553     S  E  T  *
1954    TCTGAGACTTAAGAGACAAAACGTGCATCATCATTCCAGAGAATACATATAGACCAATTGAAG
2017    CATGACGTGGATTGTATTTAAACGTATCAAAGAAGTGACAGCTATTCATGGTTCAAGTATTAA
2080    ACAGATCATACTACCAAGTTTTCAGCGGATTTTAGACACAATTATCTCAAGCTAAAAACAAAA
2143    AAAAAAAAAAACTTTAGAGCACAATGGATCTCGAGGGTCTTGCAACACCAGTGATAATAGGTC
2206    ACCATGTTGTGTTCTGTTTAAACTATTTTTTGTAAATGTCTGCACTGAAGATTTCTTTTTGT
2269    TTGCCTTTATGTAAATTTTTACGTAGCTATTTTTATACACTGTAAGGCTTTGTTCTGGGAGT
2332    TGCTGTTATTCTGATGTATAGTGTAATGG
```
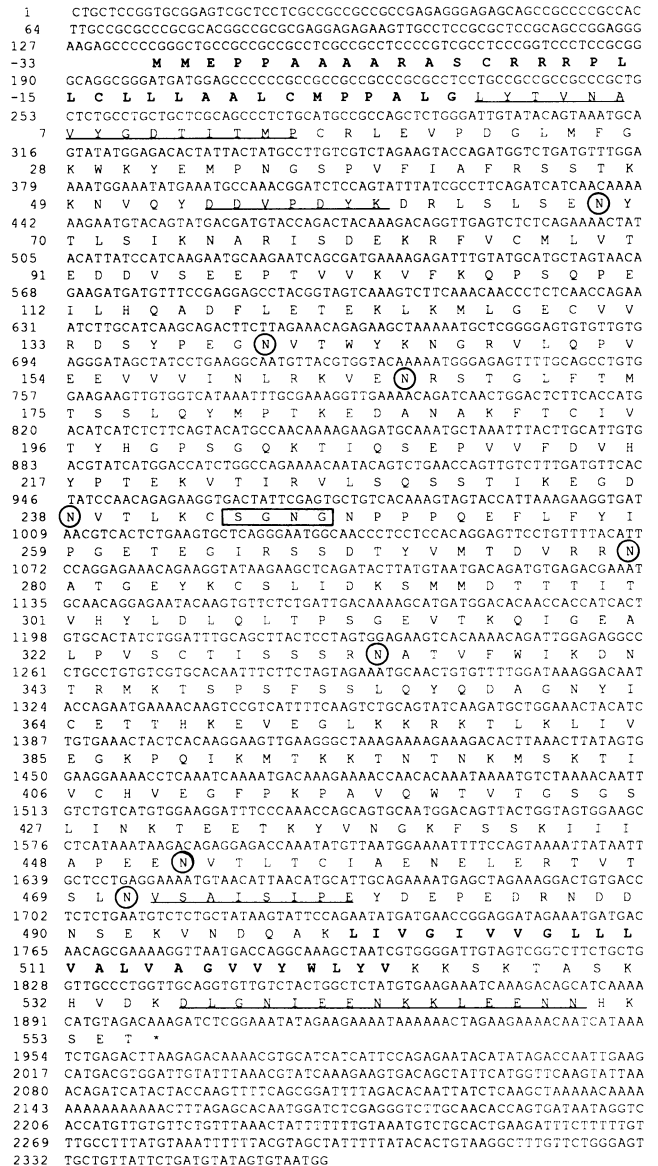
FIG. 4. Nucleotide and deduced amino acid sequence of BEN. The N-terminal sequence and sequences derived from the proteolytic fragments are underlined. The hydrophobic stretches corresponding to the signal peptide (amino acids −33 to 1) and the transmembrane domain (amino acids 499 to 523) are in boldfaced letters. The putative N-glycosylation sites are circled, and the putative glycosaminoglycan attachment site is boxed. The complete sequence of the 3′ untranslated region has not been determined. The first 5′ 2.3 kilobases of the longest cDNA (clone A) and the two other clones of 1 kb (clone H) and 2 kb (clone C) were completely sequenced and found to have identical sequences and to include the open reading frame for clones A and C.

developing neural anlage from the ectodermal placodes to specific nuclei in the brain (e.g., inferior olivary nucleus) and neurons that either belong to the peripheral nervous system (virtually all the peripheral nervous system neurons) or to the central nervous system but project neurites to the periphery (all the motoneurons of the spinal cord and of the motor brain nuclei; see ref. 1 for details). BEN is also expressed in the periderm and in the amnios, thus showing that it is largely (if not generally) involved in the ontogeny of a variety of ectodermal derivatives. This protein is also transiently expressed in the gut epithelium and on a selected set of organs of endodermal origin along the digestive tract as the gall bladder or the lung. BEN is involved in the development of a strikingly limited set of mesodermally derived tissues.

These are the hematopoietic lineage in which BEN is expressed during a well-defined stage of the ontogeny of myeloid and erythroid precursors in the bone marrow or the paraaortic region of the embryo. It is not involved in B-lineage development but is expressed during T-cell differentiation. As far as BEN expression in skeletal structures is concerned, it has so far been observed in the cartilage of the vertebral body and in membrane bone osteogenesis taking place in the head but not in any of the limb bones.

In the nervous system, only TAG-1 (23–25), a cell surface glycoprotein identified in the rat, shares a restricted specificity comparable to that seen for BEN. Interestingly, TAG-1 and BEN have several other features in common. They are expressed by restricted axonal populations (e.g., the motoneurons in the spinal cord and the sensory neurons of the dorsal root ganglion); moreover, they have the same expression dynamics since, after the beginning of neurite extension, they become restricted to the nerve fiber when synaptogenesis is in progress. In invertebrates, the establishment of stereotyped axonal networks has been shown to involve selective expression of various homophilic adhesion molecules known as fasciclins (26, 27). They have been shown to play a role in routing the axons of defined neurons to the appropriate fascicle and ultimately to their targets. BEN and TAG-1 constitute good candidates to be functional vertebrate homologs of the fasciclins. BEN, like some fasciclins [as fasciclin I and II (11)], is expressed in various cell types, including epithelial cells, in a developmentally restricted fashion. More interesting in this respect is the role of BEN in neurite fasciculation in the cerebellum, the white matter of which is the site of growth of a number of different fiber tracts that reach varied and well-defined targets. The selective expression of BEN on one of these systems exclusively—the climbing fibers—obeys the same developmental strategy as that involving fasciclins in insect neurogenesis (28).

SDS/PAGE analysis of the purification product showed that the molecular masses were different depending on the cell type considered, with a heavy epithelial form of 110 kDa, an intermediate hematopoietic form of 100 kDa, and a neural form of 95 kDa. N-Glycosylation appeared to be mostly responsible for this molecular heterogeneity. This was confirmed by the differential presence of the HNK-1 carbohydrate epitope on the protein extracted from two out of four tissues. The sequence of BEN protein thus identified is almost 100% similar to the recently published sequences of two proteins designated DM-GRASP (7-amino acid difference) (5) and SC1 (2-amino acid difference) (6). DM-GRASP is able to support neurite extension in culture, and SC1 was shown to be a homophilic adhesion molecule. Moreover, their distribution in the spinal cord is similar (29). The fact that the cDNA encoding DM-GRASP was isolated from brain, whereas that encoding BEN was isolated from E16 bursa of Fabricius from which neurons are absent, indicates that the same protein is expressed by neural and epithelial cells. BEN, DM-GRASP, and SC1 have been found to be related to a human molecule called MUC18 (4). The overall homology of only 25% between BEN and MUC18 and MUC18 distribution on melanocytes makes it unlikely to correspond to a human equivalent of BEN. Nonetheless, some common features between BEN and MUC18 suggest that they might define an additional subgroup of the immunoglobulin superfamily. First, they are the only two molecules having the V-V-C2-C2-C2 arrangement of immunoglobulin-like domains. Second, they both possess a short cytoplasmic tail: 33 amino acids for BEN and 26 for MUC18. Third, they are highly glycosylated proteins with eight potential glycosylation sites.

1. Pourquié, O., Coltey, M., Thomas, J. L. & Le Douarin, N. M. (1990) *Development* 109, 743–752.
2. Pourquié, O., Hallonet, M. E. R. & Le Douarin, N. M. (1992) *J. Neurosci.* 12, 1548–1557.
3. Corbel, C., Bluestein, H., Pourquié, O., Vaigot, P. & Le Douarin, N. M. (1992) *Cell. Immunol.* 41, 99–110.
4. Lehmann, J. M., Riethmuller, G. & Johnson, J. P. (1989) *Proc. Natl. Acad. Sci. USA* 86, 9891–9895.
5. Burns, F. R., Von Kannen, S., Guy, L., Raper, J. A., Kamholz, J. & Chang, S. (1991) *Neuron* 7, 209–220.
6. Tanaka, H., Matsui, T., Agata, A., Tomura, M., Kubota, I., Mcfarland, K. C., Kohr, B., Lee, A., Phillips, H. S. & Shelton, D. L. (1991) *Neuron* 7, 535–545.
7. Hamburger, V. & Hamilton, H. L. (1951) *J. Morphol.* 88, 49–92.
8. Bloch-Gallego, E., Huchet, M., El M'Hamdi, H., Xie, F. K., Tanaka, H. & Henderson, C. E. (1991) *Development* 111, 221–232.
9. Corthier, G., Boschetti, E. & Charley-Poulain, J. (1984) *J. Immunol. Methods* 66, 75–79.
10. Schneider, C., Newman, R. A., Sutherland, D. R., Asser, U. & Greaves, M. F. (1982) *J. Biol. Chem.* 257, 10766–10769.
11. Bastiani, M. J., Harrelson, A. L., Snow, P. M. & Goodman, C. S. (1987) *Cell* 48, 745–755.
12. Kanehisa, M. I. (1984) *Nucleic Acids Res.* 12, 203–213.
13. Lee, C. C., Wu, X., Gibbs, R., Cook, R., Muzny, D. & Caskey, T. (1988) *Science* 239, 1288–1291.
14. Corbel, C. & Thomas, J. L. (1990) *Dev. Comp. Immunol.* 14, 439–446.
15. Chomczynski, P. & Sacchi, N. (1987) *Anal. Biochem.* 162, 156–159.
16. Sambrook, J., Fritsch, E. F. & Maniatis, T. (1989) *Molecular Cloning: A Laboratory Manual* (Cold Spring Harbor Lab., Cold Spring Harbor, NY), 2nd Ed.
17. Dessen, P., Fondrat, C., Valencien, C. & Mugnier, C. (1990) *Comput. Appl. Biosci.* 6, 355–356.
18. Künemund, V., Jungalwala, F. B., Fischer, G., Chou, D. K., Keilhauer, G. & Schachner, M. (1988) *J. Cell Biol.* 106, 213–223.
19. Abo, T. & Balch, C. M. (1981) *J. Immunol.* 127, 1024–1029.
20. Kruse, J., Mailhammer, R., Wernecke, H., Faissner, A., Sommer, I., Goridis, C. & Schachner, M. (1984) *Nature (London)* 311, 153–155.
21. Tucker, G. (1984) Doctoral thesis (University of Paris).
22. Williams, A. F. & Barclay, A. N. (1988) *Annu. Rev. Immunol.* 6, 381–405.
23. Dodd, J., Morton, S. B., Karagogeos, D., Yamamoto, M. & Jessell, T. M. (1988) *Neuron* 1, 105–116.
24. Furley, A. J., Morton, S. B., Manalo, D., Karagogeos, D., Dodd, J. & Jessell, T. M. (1990) *Cell* 61, 157–170.
25. Karagogeos, D., Morton, S. B., Casano, F., Dodd, J. & Jessell, T. M. (1991) *Development* 112, 51–67.
26. Snow, P. M., Bieber, A. J. & Goodman, C. S. (1989) *Cell* 59, 313–323.
27. Elkins, T., Hortsch, M., Bieber, A. J., Snow, P. M. & Goodman, C. S. (1990) *J. Cell Biol.* 110, 1825–1832.
28. Grenningloh, G., Rehm, E. J. & Goodman, C. S. (1991) *Cell* 67, 45–57.
29. Tanaka, H. & Obata, K. (1984) *Dev. Biol.* 106, 26–37.