

The human myelin basic protein gene is included within a 179-kilobase transcription unit: Expression in the immune and central nervous systems

(multiple sclerosis/thymus/immunology/experimental allergic encephalomyelitis)

T. M. PRIBYL*, C. W. CAMPAGNONI*, K. KAMPF*, T. KASHIMA*, V. W. HANDLEY*, J. McMAHON†, AND A. T. CAMPAGNONI*

*Developmental Biology Group, Mental Retardation Research Center, University of California at Los Angeles School of Medicine, Los Angeles, CA 90024-1759; and †The Eve Clinic, Los Angeles, CA 90034

Communicated by Charles H. Sawyer, August 16, 1993

ABSTRACT Two human Golli (for gene expressed in the oligodendrocyte lineage)-MBP (for myelin basic protein) cDNAs have been isolated from a human oligodendroglioma cell line. Analysis of these cDNAs has enabled us to determine the entire structure of the human Golli-MBP gene. The Golli-MBP gene, which encompasses the MBP transcription unit, is ≈ 179 kb in length and consists of 10 exons, seven of which constitute the MBP gene. The human Golli-MBP gene contains two transcription start sites, each of which gives rise to a family of alternatively spliced transcripts. At least two Golli-MBP transcripts, containing the first three exons of the gene and one or more MBP exons, are produced from the first transcription start site. The second family of transcripts contains only MBP exons and produces the well-known MBPs. In humans, RNA blot analysis revealed that Golli-MBP transcripts were expressed in fetal thymus, spleen, and human B-cell and macrophage cell lines, as well as in fetal spinal cord. These findings clearly link the expression of exons encoding the autoimmunogen/encephalitogen MBP in the central nervous system to cells and tissues of the immune system through normal expression of the Golli-MBP gene. They also establish that this genetic locus, which includes the MBP gene, is conserved among species, providing further evidence that the MBP transcription unit is an integral part of the Golli transcription unit and suggest that this structural arrangement is important for the genetic function and/or regulation of these genes.

The myelin basic protein (MBP) is a potent autoencephalitogen that plays a role in experimental allergic encephalomyelitis and postvaccinal encephalomyelitis (1, 2). Its action as an encephalitogen and the presence of MBP-reactive T cells in the blood of multiple sclerosis patients (3, 4) have led to considerable interest in its potential role in multiple sclerosis (5, 6). It has also been reported (4, 7) that normal individuals possess MBP-reactive T cells, although the mechanism by which these T cells are produced is obscure.

The structure of the human MBP transcription unit has been determined to consist of seven exons (see Fig. 1) distributed over a length of 45 kb on chromosome 18 (8, 9). Exons 2 and 5 of the MBP gene are alternatively spliced to produce four MBP mRNAs encoding polypeptides that range in size from 17 to 21.5 kDa (10, 11). These MBP mRNAs and proteins begin to be expressed in the developing human fetal spinal cord between 10 and 20 weeks after conception (12).

It has recently been shown in mice that the MBP transcription unit is part of a complex genetic locus, called Golli (for gene expressed in the oligodendrocyte lineage)-Mbp (13). This locus is 105 kb in length and contains two transcription

start sites, one of which gives rise to the MBP mRNAs and another that produces at least three, alternately spliced, Golli mRNAs that contain exons from the MBP transcription unit. Both the Golli and MBP families of transcripts are under independent developmental regulation. The predicted proteins encoded by some of the Golli transcripts contain MBP polypeptide sequences because the Golli exons splice into the MBP exons such that the protein coding regions are in-frame. In the normal developing mouse brain, Golli transcripts of 5.1 kb and 2.6 kb have been identified in cells within the oligodendrocyte lineage, and their expression partially overlaps, but significantly precedes, the expression of the MBP transcripts.

Here we report the structure of the Golli-MBP locus in humans.‡ We find that Golli-MBP mRNAs are expressed within the human fetal spinal cord, thymus, and spleen, as well as in cell lines derived from the human immune system. These data indicate that expression of Golli-MBP products is not confined to the nervous system and provide a “molecular” link between the expression of exons of the MBP gene (and the MBP epitopes they encode) in the central nervous system and in cells and tissues of the immune system. Hence, depending upon the context in which these epitopes may be expressed, these findings might provide insight into the mechanism whereby either MBP-reactive T cells are produced or deleted in a normal individual.

MATERIALS AND METHODS

Cell Lines and Tissue Samples. The cell lines used in this study were two human adult oligodendroglioma cell lines, HOG and TC 620 (14); a human macrophage cell line (U-937); a human B-cell line (Raji); and a human T-cell line (MOLT-4). HOG and TC620 cell lines were grown as described (14). The U-937 cell line was grown in suspension at 37°C and 5% CO₂ in RPMI 1640 medium/10% heat-inactivated fetal calf serum. The Raji and MOLT-4 cell lines were grown under the same conditions, except that Iscove's medium/10% heat-inactivated fetal calf serum was used. Human fetal tissue samples were obtained, usually immediately after elective or spontaneous abortions, and frozen in liquid nitrogen and stored at -80°C. The conceptional ages of the fetuses were determined from measurements of fetal foot length (15). RNA isolations and blot analyses were done as described (13, 14). cDNA and λ genomic library construction and screening and pulse-field gel electrophoreses (PFGEs) were done as described (13), with the following modifications. cDNA librar-

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked “advertisement” in accordance with 18 U.S.C. §1734 solely to indicate this fact.

Abbreviations: MBP, myelin basic protein; PFGE, pulsed-field gel electrophoresis; Golli, gene expressed in the oligodendrocyte lineage.

‡The sequences reported in this paper have been deposited in the GenBank data base (accession nos. L18861–L18866).

RESULTS

Isolation and Analysis of Golli-MBP Transcripts from a Human Oligodendrogloma Cell Line. As part of the characterization of two human oligodendrogloma cell lines, HOG and TC620 (14), we noted that a human MBP cDNA probe hybridized to two mRNAs of ≈ 2.6 and ≈ 5.1 kb on RNA blots, sizes inconsistent with known sizes of the MBP transcripts. Therefore, we prepared cDNA libraries in λ gt11 from the HOG cell line mRNA and screened with probes specific to the MBP gene. From these screens, we isolated cDNA clones that could be divided into two categories and are represented by clones HOG 5 and HOG 7 (see Fig. 1A for exon structures; for reference, a map of the Golli-MBP transcription unit is also shown, with the exons belonging to the MBP gene numbered below the map). These two cDNA clones each contained exon 1 of the human MBP gene (shown as Golli-MBP exon 5b) and the same 5' sequence upstream of exon 1, but they differed in their 3' ends. At their 3' ends, HOG 7 contained sequences identified as exons 3–7 of the human MBP gene, whereas HOG 5 contained a sequence that was identical to intron 1 of the human MBP gene (9), now referred to as Golli-MBP exon 5c. A portion (213 nt) of the 5' ends of the two cDNAs was identical to the genomic sequence immediately upstream of MBP exon 1, as reported by Streicher and Stoffel (9). This segment is denoted as exon 5a and corresponds to the Golli-MBP exon 5a/MBP exon 1a of the mouse gene (13, 16). Interestingly, this is also a region of high sequence conservation in the genomes of mouse, human, and shark (17).

Subclones containing the specific sequences from the 5' end of the HOG cDNAs were used to screen a human λ genomic library. Subsequent mapping, subcloning, and sequence analysis of sections of the positive λ clones indicated that the distal portion of the 5' end of the HOG 5 and HOG 7 cDNAs was derived from three distinct exons. The sequences of the additional human exons are shown in Fig. 1B (in uppercase letters) along with some of the flanking genomic regions (in lowercase letters). The sizes and sequences of these human exons correspond to the analogous Golli-Mbp exons recently identified in the mouse (13). Further, a comparison of these sequences and exon composition with those obtained from the mouse indicated that HOG 5 is the human equivalent of mouse clone BG21 (13). HOG 7 contains two more MBP exons (i.e., exons 5 and 6) than does the mouse clone J37, to which it is clearly related (13). We have portrayed Golli-MBP exon 4/MBP exon 0 by a dashed box and a ? in Fig. 1A because it was not formally identified in this study. We include it because

hybridization of a mouse exon 0-specific probe (described in refs. 13 and 16) detected a distinctive band by RNA blot analysis of human fetal spinal cord mRNA, suggesting that it is both conserved and expressed in human (V.W.H. and A.T.C., unpublished observations).

Fine Mapping of the Human Golli-MBP Locus. The physical location of the three exons with respect to exon 1 of the human MBP gene was determined through a series of PFGE analyses. The resulting map is shown in Fig. 2A. For clarity, the exon numbering of the human Golli-MBP gene is shown above the map, and the numbering of the human MBP gene is shown below the map. Sequence and mapping data on the human genomic subclones established the presence of a *Not* I restriction enzyme site <100 bp from Golli-MBP exon 1 and another ≈ 9 kb upstream of MBP exon 1/Golli-MBP exon 5. Human genomic DNA was digested with *Not* I and subjected to PFGE; the DNA was transferred to nylon. Genomic DNA probes corresponding to the human Golli-MBP exon 1 and the region 9 kb upstream of MBP exon 1 both hybridized to a 125-kb *Not* I fragment (Fig. 2B). This procedure established the total physical distance separating Golli-MBP exon 1 from MBP exon 1 as 134 kb.

The intervening distances between the Golli-MBP exons 1–3 and MBP exon 1 were determined with a combination of single and double digests of human genomic DNA with the indicated restriction enzymes (i.e., *Bst*BI, *Not* I, and *Sma* I). After PFGE and transfer, the resulting Southern blots were hybridized with genomic DNA probes that spanned the restriction enzyme sites. Positions of the probes used to establish the map are shown in Fig. 2A. The use of probes that spanned adjacent fragments permitted us to establish the sizes of the two fragments adjoining the restriction site. The results of these analyses (an example is shown in Fig. 2C) demonstrated that the distances between Golli-MBP exons 1 and 2, and exons 2 and 3, were ≈ 23 and ≈ 80 kb, respectively. The distance between Golli-MBP exon 3 and MBP exon 1/Golli-MBP exon 5 was determined to be 31 kb. This total of 134 kb is almost twice the 77 kb that we determined previously for the mouse gene.

The transcription start site for the Golli-MBP gene was determined by both S1 nuclease protection and ribonuclease protection assays (data not shown). The analyses indicated the presence of two possible start sites located within ≈ 20 bp of each other, and the positions corresponding to these sites are indicated by arrows in Fig. 1B. The presence of multiple transcription start sites within short distances of one another has been noted in other genes (18) and might be predicted from the sequence of the upstream promoter region of the human

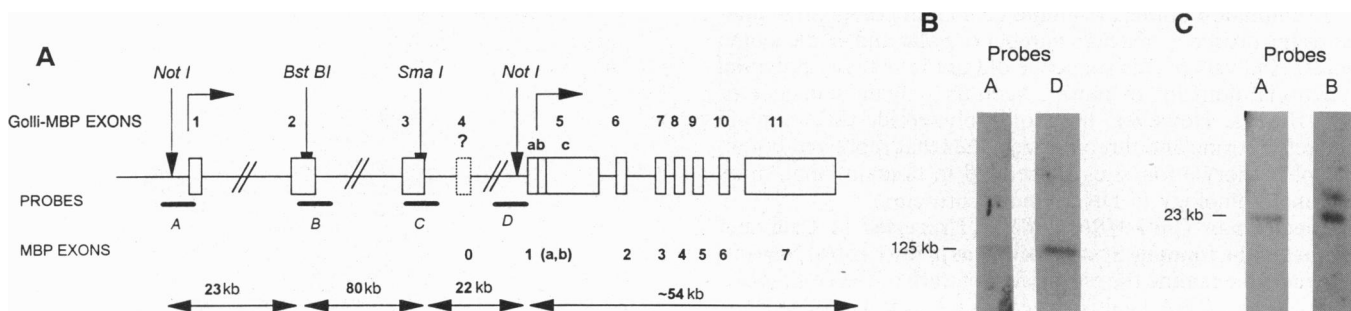


FIG. 2. PFGE and physical map of Golli-MBP gene. (A) A physical map depicting arrangement of the Golli exons and the distances separating them from each other as well as from the MBP gene is shown. The exon-numbering system with respect to the Golli gene is shown above the exons, whereas the former MBP exon numbers are shown below the map. Positions of the relevant restriction enzyme sites and the genomic probes used for physical mapping are indicated. Bent arrows mark the start sites for the Golli and MBP transcription units. Golli-MBP exon 4 is drawn with a broken line to indicate that it was not formally identified (note that map is not drawn to scale). (B) Southern blot of human genomic DNA digested with *Not* I, subjected to PFGE, and hybridized to genomic probes A and D. Both probes hybridize to a 125-kb band. (C) Southern blot of human genomic DNA digested with a combination of *Not* I/*Bst*BI, subjected to PFGE, and hybridized to genomic probes A and B. Both probes hybridize to a 23-kb band. The larger band in lane B is due to the presence of another *Bst*BI site in the 80-kb intron located 35 kb downstream of the *Bst*BI site shown in A.

Golli-MBP gene. Sequence analysis of the region upstream of the start sites indicates that there are no canonical promoter elements such as a TFIID binding site (TATA box) or a CCAAT box. However, the promoter contains sequences homologous to binding sites for the transcription factors AP-2 and Sp1 (underlined and overlined sequences in exon 1 of Fig. 1B, respectively), which partially overlap. This unusual arrangement is similar to that described for the juxtaposed binding of AP-2 and Sp1 in the 21-bp repeats of the simian virus 40 enhancer (19). The presence of transcripts with slightly variable 5' ends may be expected from promoters that lack a TATA box, which tend to be less precise in their transcription start site (20). Also, the Golli-MBP transcription unit produces at least two different mRNA transcripts (see Fig. 1A). The two start sites could represent alternative initiation sites for these individual mRNAs. Clarification of this point requires the identification of a tissue or cell type that exclusively expresses only one Golli-MBP gene product.

Comparison of Human and Mouse Golli-MBP Genes. The MBP protein-encoding exons are highly conserved between the human and mouse (10, 11, 21) and so also are the Golli exons. The nucleotide sequence of these human exons (i.e., Golli-MBP exons 1, 2, 3, and 5a) was found to be 79% identical to the mouse.

Both the human Golli-MBP and mouse Golli-Mbp transcripts have an AUG initiation codon in exon 2 (underlined and in italics in Fig. 1B). In the HOG 5 and HOG 7 cDNAs, this codon initiates an open reading frame that will read in-frame into the MBP-coding regions present within both cDNAs. In each case, this predicts a hybrid protein that has Golli amino acid sequences linked to MBP amino acid sequences. For HOG 5, the predicted protein contains a 133-amino acid Golli peptide fused to the first 58 amino acids of the human MBP with 5 additional amino acids at its C terminus derived from exon 5c. For HOG 7, the predicted polypeptide is a hybrid of the 133-amino acid Golli peptide linked to the 18.5-kDa MBP (a 170-amino acid protein). The amino acid sequences of the predicted polypeptides encoded by HOG 5 and the homologous mouse BG21 cDNAs are shown in Fig. 3. The MBP amino acid sequences are overlined and shown in italics. At the amino acid level, the predicted Golli peptides share 79% homology between the human and the mouse, again demonstrating a high degree of conservation. The polypeptides predicted from analysis of the open reading frames of the two clones were confirmed by transcribing the cDNAs into complementary RNA, followed by translation of the complementary RNAs in a reticulocyte lysate system. The products from both translations were immunoprecipitable with anti-MBP antiserum (data not shown).

Examination of the 133-amino acid Golli polypeptide indicates the presence of a high number of polar and acidic amino acids. Analysis of this sequence did not reveal any potential functional domains or motifs, such as a signal sequence or zinc fingers. However, the Golli polypeptide does contain clustered serine and threonine residues that represent potential phosphorylation sites embedded in domains that share regional homology to DNA-binding proteins.

The Human Golli-MBP Gene Is Expressed in Cells and Tissues of the Immune System as well as in the Central Nervous System. To examine the expression pattern of the Golli-MBP gene *in vivo*, RNA analyses were done on poly(A)⁺ mRNA isolated from a number of human tissues as well as the HOG and TC620 oligodendrogloma cell lines. An example of such a RNA blot is shown in Fig. 4A. This blot was probed with a cDNA subclone corresponding to Golli-MBP exons 1-3. The use of this Golli-specific probe eliminates any signal from MBP mRNAs in samples derived from the nervous system. Two bands at 5.1 kb and 2.6 kb were detected in each sample. Expression of the two mRNAs in the spinal cord was expected because at this age of fetal development, myelina-

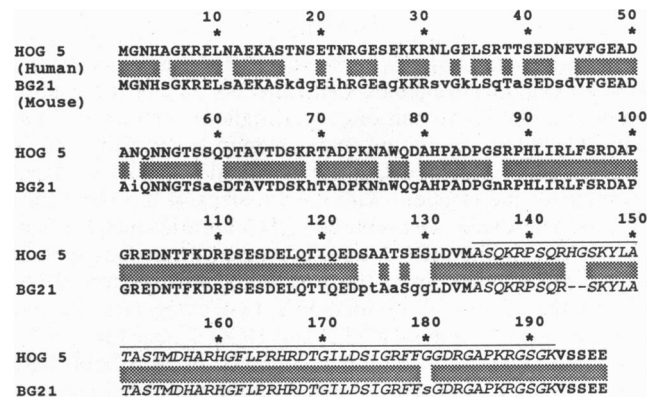


FIG. 3. Comparison of predicted amino acid sequences encoded by human Golli-MBP and mouse Golli-Mbp cDNAs. The predicted protein product encoded by the human HOG 5 cDNA is aligned above the predicted product of the mouse BG21 cDNA. Identical residues are indicated by hatched boxes between the two sequences. Amino acid sequences corresponding to MBP exon 1 are overlined and shown in italics. The last five residues, shown in boldface type, are encoded by Golli exon 5c.

tion is commencing (11, 12). Further, this developmental pattern agrees with the expression of these Golli-MBP gene transcripts in the mouse brain at approximately postnatal days 2-10 (13). In addition to revealing the 5.1-kb and 2.6-kb Golli-MBP mRNAs in the fetal spinal cord and human oligodendrogloma cell lines, these mRNAs were also de-

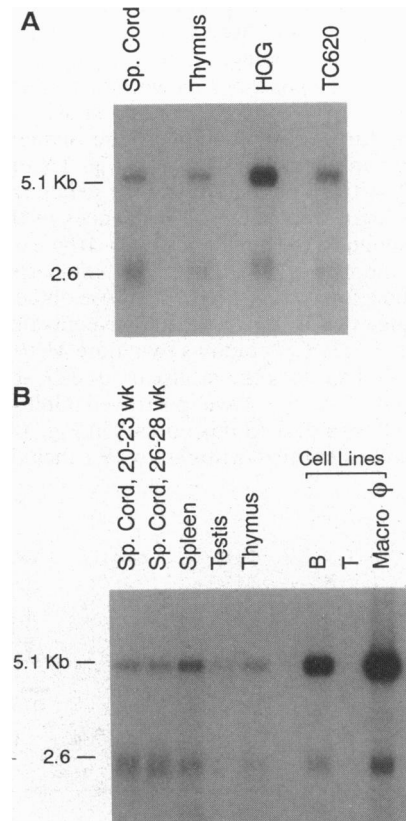


FIG. 4. RNA analysis of human Golli-MBP gene expression. (A and B) Four micrograms of poly(A)⁺ mRNA isolated from the indicated human fetal tissues and cell lines was electrophoresed, blotted to nylon, and hybridized to a human Golli probe containing exons 1-3. Sp. Cord, spinal cord; Macro, macrophage; Golli transcripts (i.e., 5.1 kb and 2.6 kb) were detected in spleen, thymus, and two of the cell lines derived from the immune system, as well as in central nervous system tissues and oligodendrogloma cell lines.

tected in the human fetal thymus sample. Because expression of the Golli-MBP mRNAs in the thymus was unexpected, we examined additional fetal tissues and cell lines derived from the human immune system.

The Golli exon-1 to -3 cDNA subclone was used to probe another RNA blot that contained poly(A)⁺ mRNA isolated from several human tissues as well as human B-cell, T-cell, and macrophage cell lines (Fig. 4B). As before, both the Golli-MBP 5.1-kb and 2.6-kb messages were expressed in the spinal cord samples (lanes 1 and 2). Very faint hybridization was detected to both bands in the testis sample, and no hybridization could be detected in the T-cell line. Various levels of hybridization to both Golli-MBP transcripts were detected in the fetal thymus, fetal spleen, B-cell, and macrophage cell lines; the highest levels of expression were seen in the macrophage cell line. The abundant appearance of both Golli-MBP transcripts in nonneural tissues implies that their biological role is not limited to glial-specific functions.

DISCUSSION

Identification of the Golli-MBP gene in humans (*i*) establishes that this complex genetic locus is not unique to the mouse and is conserved among species, (*ii*) provides further evidence that the MBP transcription unit is an integral part of the larger Golli transcription unit, and (*iii*) suggests that this structural arrangement is important for the genetic function and/or regulation of these genes.

The arrangement of the Golli-MBP gene is distinctive in that it consists of two integrated transcription units: (*i*) one in which the 3' end of the Golli transcription unit giving rise to HOG 5 (i.e., exon 5c) overlaps only the 5' end of the MBP transcription unit (i.e., exon 5b) and (*ii*) one in which the Golli transcription unit giving rise to HOG 7 encompasses the entire MBP transcription unit. In addition to expressing the family of alternatively spliced MBP-encoding transcripts, this atypical array of two transcription units also expresses at least two Golli-MBP mRNAs, encoding "hybrid" proteins, each possessing the same Golli exons linked to one or more MBP exons.

The two human transcripts are arranged such that the Golli coding sequences splice in-frame with MBP coding sequences. HOG 5 is analogous to the mouse transcript BG21. The second human transcript, HOG 7, is related to the mouse J37 cDNA (13) in that they both contain the Golli exons 1-3 and 5a, but they each contain a different array of downstream MBP exons. Because the alternative splicing pattern of the MBP transcription unit differs in the two species (for review, see ref. 22), it may not be surprising that these two clones differ in that regard.

The pattern of expression of the human Golli-MBP gene is interesting in that it is expressed in tissues and cell types outside the nervous system in the developing fetus. Both the 5.1- and 2.6-kb mRNAs were expressed in the fetal thymus and spleen as well as in the developing fetal spinal cord. The observation of significant Golli-MBP expression in the human fetal thymus and macrophage and B-cell lines is interesting in view of the presence of MBP-reactive T cells in normal individuals (4, 7). Until now, it has generally been assumed that expression of the MBP gene is confined to the nervous system and that the immune system is naive to central nervous system proteins, such as MBP, by virtue of the blood-brain barrier. Our data suggest that such an assumption may not be valid because the MBP exons in the Golli-MBP transcripts (which presumably encode MBP epitopes) are expressed in cells and tissues within the immune system. Although the immunological consequences of this observation are not yet clear, it could explain both the acquisition of overall tolerance to MBP as well as the origin of MBP-reactive T cells in normal individuals. Cells within the thymus could express the Golli-

MBP gene and the protein product processed (e.g., by thymic macrophages) and presented (e.g., by thymic macrophages, epithelial, or dendritic cells) to developing T cells. Depending upon the major histocompatibility complex context and the cell type presenting the various epitopes, the developing T cells could be subjected to clonal deletion (negative selection), which would lead to acquisition of tolerance to the MBP epitope presented. Alternatively, in a different thymic microenvironment, other developing T cells could undergo positive selection that would result in sensitization to MBP epitopes. Aberrant activation of these T cells could result in a highly specific immune response directed against MBP, ultimately leading to demyelination and subsequent nerve dysfunction (such as in patients with multiple sclerosis). Although this is clearly speculation, the present work establishes a link between the expression of the MBP gene in the nervous system and in the immune system. It should provide a basis for further investigations into the regulation of the expression of the predicted Golli proteins in both the immune and nervous systems and the role such expression may play in establishing MBP immunogenicity and tolerance.

We thank Ms. Andrea Viczian for her help in isolating and sequencing the HOG clones, Jon Hernandez for assistance with PFGE analysis, and Tommy Phan for growing the cells used in this study. This work was supported, in part, by National Institutes of Health Grants NS23022, NS23322; and HD25831 and by Grant RG2233A1 from the National Multiple Sclerosis Society.

- Alvord, E. C., Kies, M. W. & Suckling, A. (1984) *Experimental Allergic Encephalomyelitis: A Useful Model for Multiple Sclerosis* (Liss, New York), pp. 229-325.
- Hafler, D. A., Benjamin, D. S., Burks, J. & Weiner, H. L. (1987) *J. Immunol.* **139**, 68-72.
- Allegretta, M., Nicklas, J. A., Sriram, S. & Albertini, R. J. (1990) *Science* **247**, 718-721.
- Pette, M., Fujita, K., Kitz, B., Whitaker, J. N., Albert, E., Kappos, L. & Wekerle, H. (1990) *Neurology* **40**, 1770-1776.
- Weiner, H. L., Mackin, G. A., Matsui, M., Orav, E. J., Khoury, S. J., Dawson, D. M. & Hafler, D. A. (1993) *Science* **259**, 1321-1324.
- Bell, R. B. & Steinman, L. (1991) *Semin. Immunol.* **3**, 237-245.
- Saruhan-Direskeneli, G., Weber, F., Meinel, E., Pette, M., Giegerich, G., Hinkkanen, A., Epplen, J. T., Hohlfeld, R. & Wekerle, H. (1993) *Eur. J. Immunol.* **23**, 530-536.
- Kamholz, J., Toffenetti, J. & Lazzarini, R. A. (1988) *J. Neurosci. Res.* **21**, 62-70.
- Streicher, R. & Stoffel, W. (1989) *Biol. Chem. Hoppe-Seyler* **370**, 503-510.
- Kamholz, J., deFerra, F., Puckett, C. & Lazzarini, R. (1986) *Proc. Natl. Acad. Sci. USA* **83**, 4962-4966.
- Roth, H. J., Kronquist, K. E., Kerlero de Rosbo, N., Crandall, B. F. & Campagnoni, A. T. (1987) *J. Neurosci. Res.* **17**, 321-328.
- Kronquist, K. E., Crandall, B. E., Macklin, W. B. & Campagnoni, A. T. (1987) *J. Neurosci. Res.* **18**, 395-401.
- Campagnoni, A. T., Pribyl, T. M., Campagnoni, C. W., Kampf, K., Amur-Umarjee, S., Landry, C. F., Handley, V. W., Newman, S. L., Garbay, B. & Kitamura, K. (1993) *J. Biol. Chem.* **268**, 4930-4938.
- Kashima, T., Tiu, S. N., Merrill, J. E., Vinters, H. V., Dawson, G. & Campagnoni, A. T. (1993) *Cancer Res.* **53**, 170-175.
- Moore, K. L., ed. (1982) *The Developing Human: Clinically Oriented Embryology* (Saunders, Philadelphia), p. 95.
- Kitamura, K., Newman, S. L., Campagnoni, C. W., Verdi, J. M., Mohandas, T., Handley, V. W. & Campagnoni, A. T. (1990) *J. Neurochem.* **54**, 2032-2041.
- Fors, L., Hood, L. & Saavedra, R. A. (1993) *J. Neurochem.* **60**, 313-321.
- Gonzalez-Crespo, S. & Boronat, A. (1991) *Proc. Natl. Acad. Sci. USA* **88**, 8749-8753.
- Mitchell, P. J., Wang, C. & Tjian, R. (1987) *Cell* **50**, 847-861.
- Smale, S. T. & Baltimore, D. (1989) *Cell* **57**, 103-113.
- deFerra, F., Engh, H., Hudson, L., Kamholz, J., Puckett, C., Molineaux, S. & Lazzarini, R. A. (1985) *Cell* **43**, 721-727.
- Campagnoni, A. T. (1988) *J. Neurochem.* **51**, 1-14.