*Proc. Natl. Acad. Sci. USA*
Vol. 84, pp. 886–890, February 1987
Neurobiology

# Identification of a cDNA coding for a fifth form of myelin basic protein in mouse

(mRNA splicing)

SAMMYE NEWMAN, KUNIO KITAMURA, AND ANTHONY T. CAMPAGNONI*

Mental Retardation Research Center, University of California at Los Angeles School of Medicine, 760 Westwood Plaza, Los Angeles, CA 90024

*Communicated by John Garcia, October 1, 1986 (received for review August 8, 1986)*

**ABSTRACT** **The primary sequences of four molecular mass variants (14, 17, 18.5, and 21.5 kDa) of the mouse myelin basic protein (MBP) have recently been determined through analysis of cDNA clones of their mRNAs. The mRNAs coding for the four MBP variants are thought to arise by differential splicing of two exons (exons 2 and 6) from a single gene. In contrast, exons 2 and 5 may be spliced out in the posttranscriptional processing of the human MBP gene. To investigate the possibility that a third exon (exon 5) may also be differentially spliced out in the processing of the mouse MBP gene transcript, a mouse cDNA library was screened to search for cDNAs missing exon 5. A MBP cDNA was isolated whose coding region specified a fifth mouse MBP variant with a molecular mass of ≈17 kDa. The mass of this variant (17,257 Da) is so close to that of the other 17-kDa mouse MBP (17,224 Da) that the two would be indistinguishable on NaDodSO$_4$/polyacrylamide gels. Analysis of the sequence of the cDNA clone indicates that excision of exons 2 and 5 of the mouse MBP gene would produce the mRNA encoding this newly described 17-kDa MBP, whereas excision of exon 6 would produce the mRNA for the other 17-kDa MBP variant. Thus, the "17-kDa" mouse MBP consists of at least two molecular forms with very similar molecular masses but markedly different primary sequences. Of five full-length or near full-length cDNAs representing 17-kDa MBPs, one was missing exons 2 and 5 and four were missing exon 6.**

The integrity of the myelin sheath is a major determinant in the maintenance of a functionally normal brain. Thus, investigations into the regulation of myelin protein gene expression are important for understanding myelinogenesis and the molecular mechanisms governing myelin degeneration and reformation in neurodegenerative diseases. The myelin basic proteins (MBPs) have been the most extensively studied myelin proteins, although the exact structural relationships among the various forms of the mouse MBPs have been established only recently. It was originally thought that mammalian brain contained either one or two MBPs, until it was discovered that mouse myelin contained four structurally related forms of the basic proteins having apparent molecular masses of 14 kDa, 17 kDa, 18.5 kDa, and 21.5 kDa (1). Since these four MBP variants were shown to be the products of four separate mRNAs (2), it was assumed that these mRNAs either arose as different splice products of one or two genes or were coded for by separate genes. Recently, the structural relationships among the four mouse proteins (shown in Fig. 1) were determined through an analysis of the gene and of the cDNAs that correspond to the mouse MBP mRNAs (3, 4). The MBP gene appears to be very long [30–35 kilobases (kb)], consisting of seven relatively short exons interrupted by very long introns (3, 4). The 21.5-kDa protein

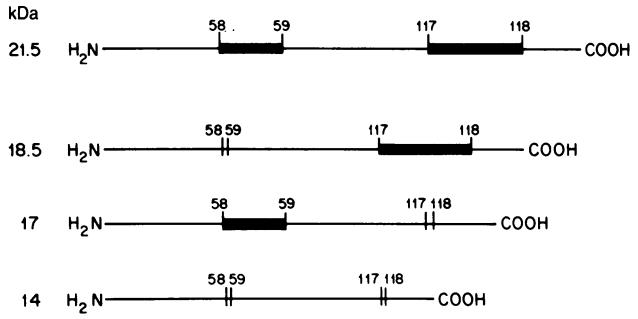


FIG. 1. Schematic representation of four forms of the mouse MBP. Thick lines represent peptide sequences encoded by exons 2 and 6, which can be deleted to give rise to the three smaller MBP variants. The deleted sequences between amino acids 58 and 59 and between amino acids 117 and 118 correspond to those encoded by exons 2 and 6, respectively, of the MBP gene.

contains peptide sequence encoded by all seven exons of the MBP gene and represents the longest of the four variants. The 18.5-kDa variant is identical in structure to the 21.5-kDa MBP except for the deletion of 26 amino acids encoded by exon 2 of the MBP gene. The 17-kDa protein described by de Ferra *et al.* (3) is missing a sequence of 41 amino acids encoded by exon 6, and the 14-kDa protein is missing the amino acid sequences encoded by both exons 2 and 6. Thus, the four previously described mouse MBPs are structurally related to each other through the presence or absence of two internal peptide sequences corresponding to two exons of the MBP gene (4).

In the human, the 21.5-kDa and 18.5-kDa MBPs are comparable in structure to the corresponding mouse variants (5). Recent work in this laboratory and elsewhere has identified a human MBP cDNA encoding a novel form of the 17-kDa MBP in which the nucleotide sequences corresponding to exons 2 and 5 are absent (5, 6). Deibler *et al.* (7) reported the isolation and identification of a human MBP variant corresponding to the protein predicted by that cDNA. These results suggested that the predominant splicing pathway in the human might be different from that in the mouse. However, the ability to splice out exon 5 might also exist in the mouse as well as in the human. Thus, the mouse might synthesize a 17-kDa MBP variant analogous to the human protein in addition to the 17-kDa variant already identified. The results reported here confirm that possibility.

## MATERIALS AND METHODS

**Preparation of cDNA Libraries and Screening Procedures.** Polyribosomes were isolated from the brains of 18-day-old C57BL/6J mice, and poly(A)$^+$ RNA was prepared as described (2, 8, 9). The poly(A)$^+$ RNA was used to prepare a

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.

Abbreviation: MBP, myelin basic protein.

*To whom requests for reprints should be addressed.

cDNA library in the phage λgt11 essentially as described by Gubler and Hoffman (10), with minor modifications (6). In brief, first-strand cDNA synthesis was accomplished by conventional oligo(dT)-primed reverse transcription (11), and second-strand synthesis was achieved by the procedure described by Gubler and Hoffman (10). The double-stranded cDNA was methylated and ligated, by blunt-end addition, to commercially prepared linkers (New England Biolabs), containing an internal *Eco*RI site, and the mixture was subjected to *Eco*RI digestion. The cDNAs were fractionated by size, and fragments longer than 1.5 kb were selected for cloning into λgt11. The cDNAs were then ligated to phage arms and packaged into phage particles with a commercial "phage-packaging" mixture (Promega Biotec, Madison, WI), and the entire library was amplified.

The amplified library was screened with a 540-base-pair (bp) mouse MBP cDNA probe (PP535) complementary to the 5′ end of the MBP mRNA, including the entire coding region of the mRNA for mouse 14-kDa MBP. The sequence of PP535 has been published (6). The PP535 probe was radioactively labeled by nick-translation (11) and hybridized to the mouse cDNA phage library grown in *Escherichia coli* strain Y1090. Positive plaques were identified by autoradiography and purified by three rounds of plating and screening with the PP535 probe. From a screen of 300,000 plaques, 50 plaques identified as containing cDNA inserts corresponding to the MBP mRNA coding region were chosen for further characterization. The number of positive MBP clones in the library (1 per 1000) was greater than the number chosen for analysis and was comparable to the abundance reported by others (3).

**Blot Analysis.** Phage DNAs containing mouse MBP cDNA inserts were purified from plaques (11), and insert sizes were determined by excision of the inserts with *Eco*RI followed by Southern blot analysis (12) using the PP535 probe. Restriction enzyme analysis of the inserts was performed by cleaving the MBP cDNA clones with *Eco*RI, *Ava* II, and *Dde* I (International Biotechnologies, New Haven, CT), separating the DNA fragments by electrophoresis in 5% polyacrylamide gels (13), electrophoretically transferring the fragments to a nylon membrane (New England Nuclear GeneScreen*Plus*), and probing the immobilized DNA with radioactively labeled PP535 cDNA.

In some cases an exon 1-specific probe was used to identify clones containing exon 1. This probe, corresponding to 154 bases at the 5′ end of the MBP mRNA, was prepared as a "deletion" fragment during the sequencing of a human MBP cDNA clone, KK36 (6). The sequence of the exon 1 probe exhibits >90% homology with that of the mouse exon 1 coding sequence. When used as a radioactive probe, this fragment was radioactively labeled by the random-priming procedure (14).

**DNA Sequence Analysis.** Sequence analysis of the cDNAs was performed using the dideoxy nucleotide chain-termination procedure of Sanger *et al.* (15) with [α-[$^{35}$S]thio]dATP as the radioactive label, after subcloning the cDNA into M13 mp19 phage. The entire cDNA sequence was determined by preparing deletion fragments of the cDNAs in the M13 phage vector (using the Cyclone System kit of International Biotechnologies), sequencing each fragment, and deducing the arrangement of fragments through overlapping regions (6).

## RESULTS

Approximately 3 × 10$^5$ members of a mouse brain cDNA library were screened with an MBP cDNA probe specific for

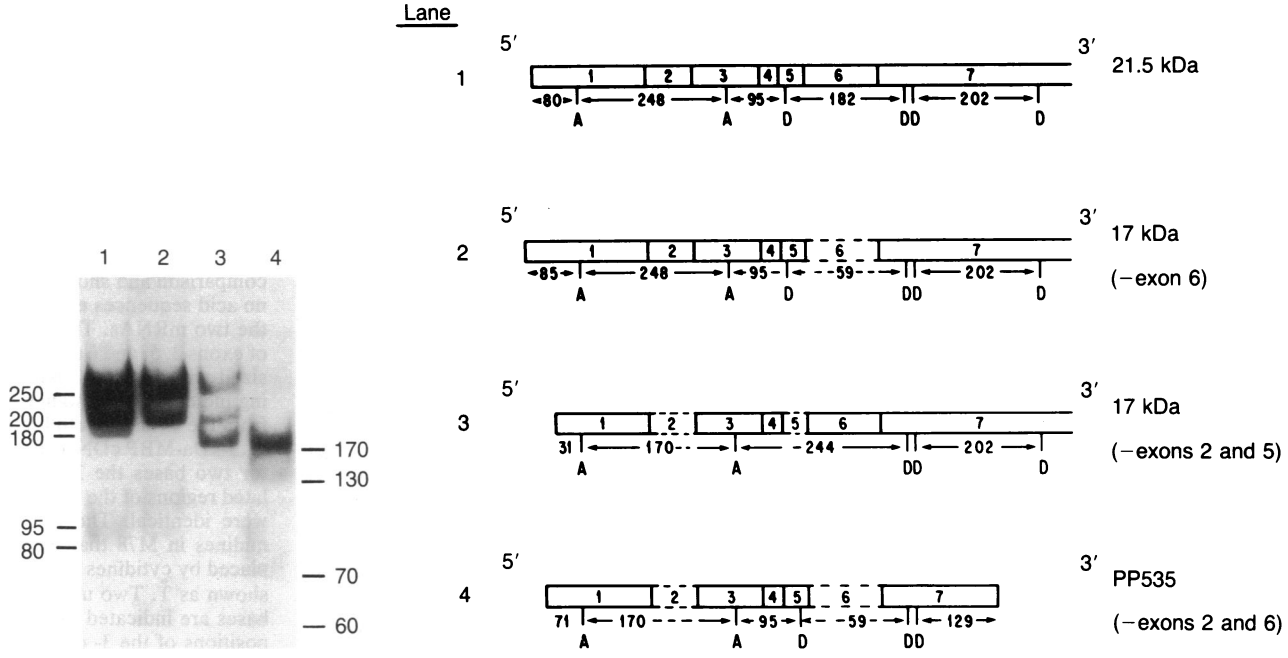


FIG. 2. (*Left*) Blot-hybridization analysis of cDNA clones after treatment with the restriction enzymes *Ava* II and *Dde* I and electrophoretic transfer of the fragments to a nylon membrane. Each lane contained 1.25 μg of digested DNA, and the membrane was probed with radiolabeled PP535. Lane 1: clone M44 (21.5-kDa-MBP cDNA). Lane 2: clone M72 (putative 17-kDa-MBP cDNA missing exon 6). Lane 3: clone M78 (putative 17-kDa-MBP cDNA missing exons 2 and 5). Lane 4: clone PP535 (14-kDa-MBP cDNA). Numbers at the sides represent DNA fragment lengths (bp). (*Right*) Diagram illustrating the arrangement of exons 1–7 within the structures of four MBP cDNA clones. *Ava* II (A) and *Dde* I (D) restriction sites at the 5′ ends of the clones and the fragment lengths (in bp) between the restriction sites are indicated below the structures. Broken lines represent missing exons. The sizes of the bands observed in the electroblot (*Left*) are consistent with those predicted from the structures of the cDNAs (*Right*). The 3′ and 5′ ends of PP535 are incomplete and do not extend either to the 5′ end of the MBP mRNA or to the *Dde* I restriction site at the far 3′ end shown on the map, thereby giving shorter fragments than would be expected with a full-length clone. Also note that the 5′ ends of the cDNA clones were not 97 bp, as predicted if the clones were full-length, but varied from 31 to 85 bp depending upon how far short they fell from the 5′ end. A 16-bp fragment resulting from digestion of exon 7 with *Dde* I was not evident in the blots under the conditions employed for this experiment.

the coding region (PP535), and 50 plaques were selected that appeared to contain MBP coding region. To facilitate the screening of these clones, a scheme was devised to select for MBP clones that did not contain sequences corresponding to exon 2 or 5. Using the sequence of the mouse 21.5-kDa MBP cDNA as a guide, two restriction enzymes were selected that would provide significantly different fragmentation patterns in the coding region of the MBP cDNAs depending upon the exons present in the clone. That is, MBP cDNAs missing one or more of the coding region exons 2–6 would yield predictable and distinct fragmentation patterns, thereby allowing a tentative identification of the clone. The restriction enzymes *Ava* II (which cleaves on either side of exon 2) and *Dde* I (which cleaves within exon 5) were used for the restriction analysis. The fragments were electrophoretically transferred to a nylon membrane and probed with radioactively labeled PP535, thereby limiting the analysis to those fragments at the 5′ (coding) end of the cDNAs. Results from this type of analysis for four clones are shown in Fig. 2 *Left*. The restriction pattern obtained with clone M78 was different from that of a 21.5-kDa-MBP cDNA (clone M44) and a 14-kDa-MBP cDNA (PP535), and it was also different from that of clone M72, which matched the pattern expected of a 17-kDa-MBP cDNA missing exon 6 (Fig. 2 *Right*). The sizes of the fragments generated from clone M78 were consistent with those expected of a cDNA missing both exons 2 and 5.

Based upon this restriction analysis, clones M72 and M78 were chosen for further examination by sequence analysis. The results of the sequence determinations are presented in Fig. 3, along with the deduced amino acid sequences of the proteins encoded by the two mRNAs. Clone M72, which contained exons 1–5 and exon 7, corresponded to an mRNA encoding the 17-kDa mouse MBP described by de Ferra *et al.* (3). In contrast, clone M78 contained exons 1, 3, 4, 6, and 7 and encoded a second 17-kDa mouse MBP analogous to the human MBP variant described in a recent report (6). The coding regions of M72 and M78 were identical, with the exception of the deleted exons, to that published by de Ferra *et al.* (3) for a mouse 21.5-kDa-MBP cDNA. However, there were 11 single-base differences, two 4-bp deletions, and one 3-bp deletion noted in the 3′ noncoding region of the clones reported in this study compared to that published by de Ferra *et al.* (3) (Fig. 3). This is probably due to strain differences between the NIH Swiss mice used by de Ferra *et al.* (3) and the C57BL/6J mice used in this study for the preparation of the cDNA libraries. In addition to these base differences, two single-base discrepancies existed between the 3′ noncoding regions of M72 and M78 (Fig. 3). In both cases a cytidine was substituted in M72 at positions where thymidine was found in both M78 and in the sequence reported by de Ferra *et al.* (3). Reverse transcriptase, which was used in the preparation of these libraries, is known to "misread" bases as frequently as once every 600 bases (16). Such misreading may be responsible for these differences between the two clones.

We attempted to obtain a crude estimate of the approximate ratios of the various MBP mRNAs among the clones

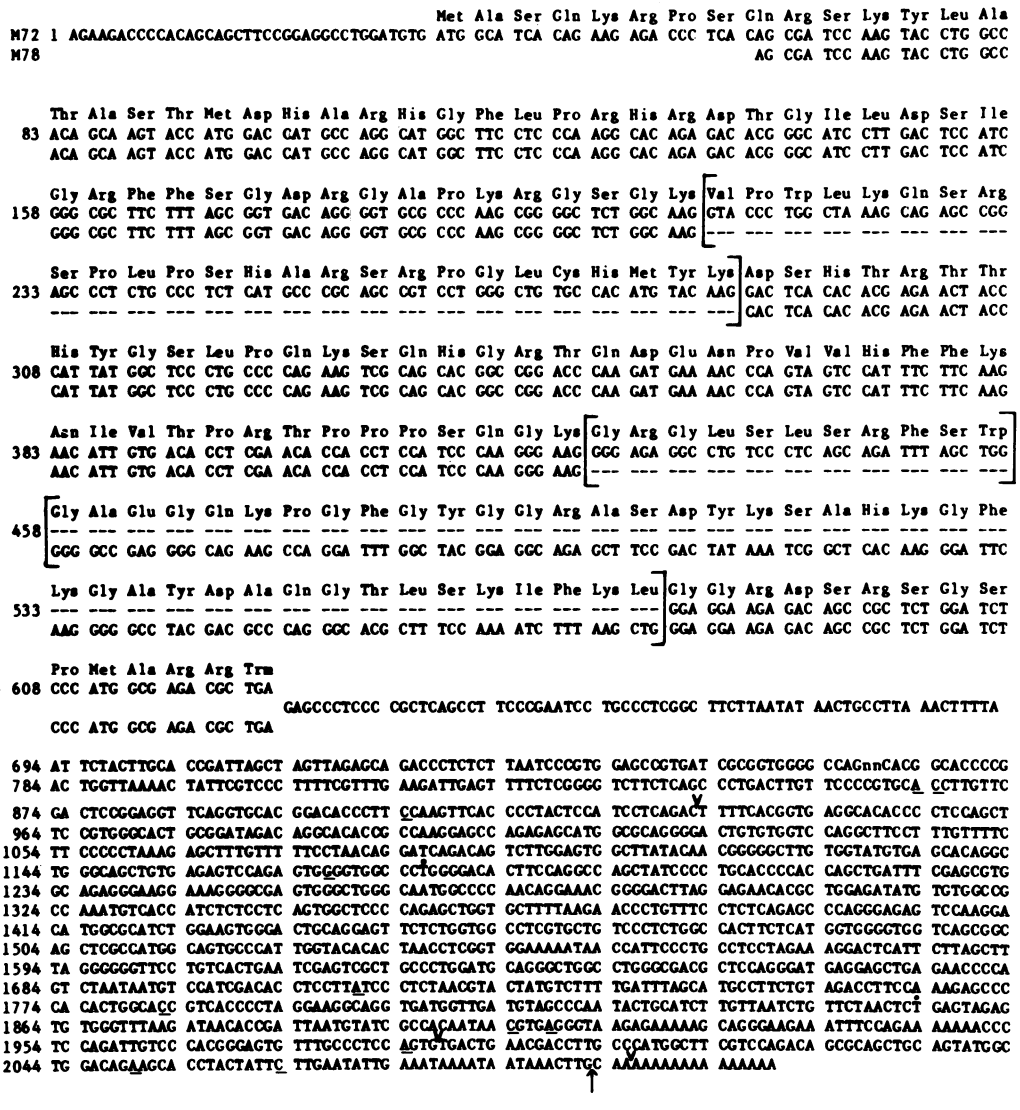

FIG. 3. Sequences of the 17-kDa mouse MBP clones M72 and M78. The DNA sequences are aligned for ease of comparison and show the amino acid sequences encoded by the two mRNAs. The borders of exons 2, 5, and 6 are emphasized with brackets. Numbering indicates the distance (bp) from the 5′ end of a full-length 21.5-kDa-MBP cDNA. Except for two bases the 3′ untranslated regions of the two clones were identical. The two thymidines in M78 that were replaced by cytidines in M72 are shown as T. Two unidentified bases are indicated by nn; the positions of the 3- or 4-bp deletions found in this study compared to that of de Ferra *et al.* (3) are indicated by ∨, and single base substitutions are underlined. Arrow indicates position of 3′ end of clone M78. The probe used to isolate these clones (PP535) is a cDNA containing the 14-kDa MBP coding region (missing exons 2 and 6) and is complementary to nucleotides 27–768 in this figure.

that we isolated. Of the 50 clones isolated, 25 contained exon 1 and were at least 1.7 kb long as determined from a Southern blot-hybridization of the inserts with an exon 1-specific probe. These 25 full-length or near full-length clones were then subjected to the restriction mapping analysis described above. The numbers of cDNAs representing the 14-, 17-, 18.5-, and 21.5-kDa-MBP mRNAs were 15, 5, 4, and 1, respectively. Obviously, the small numbers involved in this analysis preclude a conclusive extrapolation of these results to the *in vivo* state; however, they are in line with estimates of the individual MBP mRNA levels based upon analyses of MBPs synthesized in cell-free systems using mouse brain mRNA preparations (ref. 17; A.T.C. and M. J. Hunkeler, unpublished results). Of the five clones found to belong to the 17-kDa-MBP cDNA class, four were missing exon 6 and one was missing exons 2 and 5. Although these numbers are only approximate, these studies suggest that within the 17-kDa-MBP mRNA class in 18-day-old mice, the 17-kDa-MBP mRNAs missing exons 2 and 5 are less abundant than the 17-kDa-MBP mRNAs missing exon 6.

## DISCUSSION

The results presented here demonstrate the existence of a mouse MBP mRNA that is specifically lacking the nucleotide sequence corresponding to exons 2 and 5 of the MBP gene. The existence of this cDNA implies the presence of the 17-kDa MBP variant for which it codes in mouse central nervous system. An analogous protein appears to be a relatively abundant MBP variant in humans (7, 18), whereas this study suggests that in the mouse this MBP is probably a minor form at 18 days. This "17-kDa" variant (molecular mass = 17,257 Da) would not be easily distinguishable from the other "17-kDa" MBP (molecular mass = 17,224 Da) by $NaDodSO_4$ polyacrylamide gel electrophoresis, the conventional method of separating and identifying these proteins.

The ratios of synthesis of the mouse MBP variants have been shown to change rather significantly with development (17, 19) and our observation that there are two 17-kDa mouse MBP variants raises the question of whether there is a shift during development from one 17-kDa form to another, or whether both are synthesized in constant proportion throughout development.

If, as the data presented here indicate, a functional splicing of exon 5 from MBP mRNAs does occur, then other MBP mRNAs lacking this exon might also exist. For example, exon 5 could be spliced alone or in combination with exon 6 or with both exons 2 and 6, to give rise to MBPs with molecular masses of approximately 20 kDa, 15.6 kDa, and 13 kDa, respectively. There is, in fact, evidence for the existence of such MBPs. We and others have observed the presence of a 12- to 13-kDa MBP-related polypeptide in translations of mRNA preparations from mouse, guinea pig, and human (refs. 6, 19, and 20; A.T.C. and M. J. Hunkeler, unpublished observations), and such a polypeptide has been noted by others in cell-free translations of mouse brain mRNA (21). Synthesis of an ≈15-kDa polypeptide immunoprecipitable by anti-MBP has been observed in cell-free translations using mouse (20) and guinea pig (A.T.C. and M. J. Hunkeler, unpublished observations) mRNA preparations. Greenfield *et al.* (22) identified two faintly staining bands, one of which migrated between the 14-kDa and 17-kDa MBPs and the other of which migrated between the 18.5-kDa and 21.5-kDa MBPs in immunoblots of mouse brain proteins. Moreover, a protein band with slightly faster mobility than that of the 21.5-kDa MBP has been noted in both rabbit (23) and mouse (20). Taken together, these findings support the idea that the differential splicing of exons 2, 5, and 6 from MBP pre-mRNA may generate as many as eight distinct molecular forms of MBP in the mouse.

Alternative splicing events have been described for a number of systems in which primary RNA transcripts can give rise to several variant polypeptides. For example, differential splicing of human fibronectin pre-mRNA may generate as many as 10 polypeptides from a single gene (24). In the α-crystallins of the mouse, the presence or absence of a 22 amino acid insert in the mature protein is determined by differential splicing of a single pre-mRNA (25). Variation in the pattern of splicing at the 3′ end of a pre-mRNA appears to account for the formation of three forms of myosin heavy chain gene transcript in *Drosophila* (26), and an intricate pattern of differential splicing in both the 3′ and 5′ region of the rat skeletal fast troponin T gene accounts for the production of at least 10 distinct troponin T mRNAs (27). In the mammalian nervous system, the bovine preprotachykinin gene is processed in a tissue-specific manner to generate two distinct mRNAs encoding the neuropeptide substance P alone or with substance K (28). Two distinct mRNAs that encode either calcitonin or calcitonin gene-related peptide are generated in rat neuronal tissue from one gene by the selective use of alternative polyadenylylation sites (29). Thus, a growing number of mammalian systems exist in which two or more related polypeptides are known to be generated by alternative splicing. The MBPs of mouse brain now stand as an example of this group in that at least five and possibly as many as eight distinct MBPs may be generated by selective excision of exons 2, 5, and 6 from the mouse MBP gene transcript.

We thank M. J. Hunkeler for guidance in preparing the cDNA library, H. J. Roth for helpful discussions and for providing the MBP exon 1 probe, and T. Anzalone for assistance with the preparation of the manuscript. This work was supported, in part, by National Institutes of Health Grants NS23022, NS23322, and RR5756.

1. Barbarese, E., Braun, P. E. & Carson, J. M. (1977) *Proc. Natl. Acad. Sci. USA* **74,** 3360–3364.
2. Yu, Y. T. & Campagnoni, A. T. (1982) *J. Neurochem.* **39,** 1559–1568.
3. de Ferra, F., Engh, H., Hudson, L., Kamholz, J., Puckett, C., Molineaux, S. & Lazzarini, R. A. (1985) *Cell* **43,** 721–727.
4. Takahashi, N., Roach, A., Teplow, D. B., Prusiner, S. B. & Hood, L. (1985) *Cell* **42,** 139–148.
5. Kamholz, J., de Ferra, F., Puckett, C. & Lazzarini, R. (1986) *Proc. Natl. Acad. Sci. USA* **83,** 4962–4966.
6. Roth, H. J., Kronquist, K., Pretorius, P. J., Crandall, B. F. & Campagnoni, A. T. (1986) *J. Neurosci. Res.* **16,** 227–238.
7. Deibler, G. E., Krutzsch, H. C. & Kies, M. W. (1986) *J. Neurochem.* **47,** 1219–1225.
8. Roth, H. J., Hunkeler, M. J. & Campagnoni, A. T. (1985) *J. Neurochem.* **45,** 572–580.
9. Campagnoni, A. T., Campagnoni, C. W., Bourre, J. M., Jacque, C. & Baumann, N. (1984) *J. Neurochem.* **42,** 733–739.
10. Gubler, U. & Hoffman, B. J. (1983) *Gene* **25,** 263–269.
11. Maniatis, T., Fritsch, E. F. & Sambrook, J. (1982) *Molecular Cloning: A Laboratory Manual* (Cold Spring Harbor Laboratory, Cold Spring Harbor, NY).
12. Southern, E. M. (1975) *J. Mol. Biol.* **98,** 503–517.
13. Laemmli, U. K. (1970) *Nature (London)* **227,** 680–685.
14. Feinberg, P. & Vogelstein, B. (1983) *Anal. Biochem.* **132,** 6–13.
15. Sanger, F., Nicklen, S. & Coulson, A. R. (1977) *Proc. Natl. Acad. Sci. USA* **74,** 5463–5467.
16. Battula, N. & Loeb, L. A. (1974) *J. Biol. Chem.* **249,** 4086–4093.
17. Campagnoni, C. W., Carey, G. D. & Campagnoni, A. T. (1978) *Arch. Biochem. Biophys.* **190,** 118–125.
18. Kerlero de Rosbo, N., Carnegie, P. R., Bernard, C. C. A. & Linthicum, D. S. (1984) *Neurochem. Res.* **9,** 1359–1369.
19. Carson, J. H., Nielson, M. L. & Barbarese, E. (1983) *Dev. Biol.* **96,** 485–492.
20. Sorg, B. J. A., Agrawal, D., Agrawal, H. C. & Campagnoni, A. T. (1986) *J. Neurochem.* **46,** 379–387.

21. Carnow, T. B., Carson, J. H., Brostoff, S. W. & Hogan, E. L. (1984) *Dev. Biol.* **106,** 38–44.
22. Greenfield, S., Weise, M. J., Gantt, G., Hogan, E. L. & Brostoff, S. W. (1982) *J. Neurochem.* **39,** 1278–1282.
23. Agrawal, H. C., Randle, C. L. & Agrawal, D. (1981) *J. Biol. Chem.* **256,** 12243–12246.
24. Kornblihtt, A. R., Umezawa, K., Vibe-Pedersen, K. & Baralle, F. E. (1985) *EMBO J.* **4,** 1755–1759.
25. King, C. R. & Piatigorsky, J. (1983) *Cell* **32,** 707–712.
26. Rozek, C. E. & Davidson, N. (1983) *Cell* **32,** 23–34.
27. Breitbart, R. E., Nguyen, H. T., Medford, R. M., Destree, A. T., Mahdavi, V. & Nadal-Ginard, B. (1985) *Cell* **44,** 67–82.
28. Nawa, H., Kotani, H. & Nakanishi, S. (1984) *Nature (London)* **312,** 729–734.
29. Amara, S. G., Evans, R. M. & Rosenfeld, M. G. (1984) *Mol. Cell. Biol.* **4,** 2151–2160.