Journal of Neurochemistry
Raven Press, Ltd., New York

Structure of the Mouse Myelin P_2 Protein Gene

Vinodh Narayanan, *Klaus H. Kaestner, and †Gihan I. Tennekoon

*Department of Pediatrics, The University of Pittsburgh, Pittsburgh, Pennsylvania; *Department of Biological Chemistry, The Johns Hopkins University School of Medicine, Baltimore, Maryland; and †Departments of Pediatrics and Neurology, The University of Michigan School of Medicine, Ann Arbor, Michigan, U.S.A.*

Abstract: Myelin P_2 protein is a small (14,800 Da) protein found in peripheral and central nervous system myelin. To investigate the regulation of expression of this myelin protein, a mouse genomic library was screened with a rabbit P_2 cDNA (pSN2.2-2), and a single positive phage clone containing a 20-kb insert was obtained. This insert contained a single internal *Sal*I restriction site and several *Eco*RI sites. The *Eco*RI fragments from this insert were subcloned into Bluescript. The rabbit P_2 cDNA (pSN2.2-2) hybridized with a 4-kb *Eco*RI fragment, and this 4-kb fragment was then sequenced after the creation of nested deletions. The mouse gene contained four exons: exon 1 coded for amino acids 1–24, exon 2 for amino acids 24–81, exon 3 for amino acids 82–115, and exon 4 for amino acids 116–131. The three introns were 1.2 kb, 150 bp, and 1.3 kb in length. Primer extension analysis revealed two transcription start sites at +1 and +47, consistent with the presence of two P_2 mRNAs, with the larger transcript appearing more abundant. The amino acid sequences predicted from the mouse DNA indicate that the mouse protein differs from the rabbit protein at 16 different positions, with most of the differences located in exon 3. When the gene structure of fatty acid binding protein (FABP) genes were compared, the P_2 gene had the same overall structure as others in the FABP family, suggesting a common ancestral gene for members of this family. The 5′-flanking region contains candidate TATA and CAAT sequences, as well as two AP-1 binding sites that may be important in modulation of the expression of this gene. **Key Words:** Myelin P_2 protein gene—*Eco*RI fragments—Fatty acid binding protein. **Narayanan V. et al.** Structure of the mouse myelin P_2 protein gene. *J. Neurochem.* **57,** 75–80 (1991).

Myelin, a structure unique to the nervous system, is a specialization of the plasma membrane of Schwann cells and oligodendrocytes. It is composed predominantly of lipids, with proteins accounting for about 30% of its wet weight. Myelin obtained from the central and peripheral nervous systems contains many different proteins. Although the function of many of these proteins has been open to speculation, only P_0 protein, myelin basic proteins, proteolipid protein, and myelin-associated glycoproteins appear to play important roles in establishing and maintaining the lamellar structure of compact myelin (Lemke, 1988; Filbin et al., 1990).

P_2 is a small (14,800 Da) basic protein found in myelin and in the cytosol of Schwann cells and oligodendrocytes and whose expression increases during myelination (Narayanan et al., 1988). The function of this protein may be inferred from its amino acid sequence and tertiary structure determined by x-ray crystallography. P_2 bears a striking similarity in amino acid sequence to a family of fatty acid binding proteins (FABPs) from various tissues, including the mouse adipocyte FABP (aP_2/422); rat liver, intestinal, and heart FABPs; and the cellular retinol and retinoic acid binding proteins (Narayanan et al., 1988). Crystallographic studies have revealed that the tertiary structures of bovine myelin P_2 protein (Jones et al., 1988) and rat intestinal FABP (Sacchettini et al., 1988) are very similar to each other, each consisting of two layers of β-sheet structure enveloping a hydrophobic pocket in which a fatty acid ligand is carried. Studies demonstrating that bovine P_2 protein specifically binds oleic acid, retinol, and retinoic acid support this idea (Uyemura et al., 1984).

Unlike other myelin proteins, the myelin P_2 protein varies considerably in abundance in different species; P_2 accounts for <1% of myelin proteins in rat or mouse sciatic nerves but makes up 5–12% of myelin proteins in human, bovine, and rabbit nerve roots (Greenfield

Received August 24, 1990; revised manuscript received December 10, 1990; accepted December 11, 1990.

Address correspondence and reprint requests to Dr. G. Tennekoon at The University of Michigan Medical School, Division of Pediatric Neurology, R6060, Kresge II, Ann Arbor, MI 48019-0570, U.S.A.

Abbreviations used: FABP, fatty acid binding protein; FSE, fat-specific element.

et al., 1973, 1982). Another fact that remains unexplained is the tissue distribution of P_2 protein, with P_2 protein most abundant in peripheral nerves and nerve roots, less abundant in spinal cord, and least abundant in the brainstem and cerebrum (Trapp et al., 1983). Whether these species differences and the tissue distribution within a single species reflect variations in fatty acid metabolism is uncertain.

To investigate the mechanisms mediating the differential expression of this protein in various parts of the nervous system and the variability in protein abundance among different species, we have cloned the mouse myelin P_2 protein gene. In this report, we present the structural characteristics of this gene and compare it to other genes in the FABP family.

METHODS AND RESULTS

A genomic library was constructed in the bacteriophage vector, λ EMBL3, representing nuclear DNA extracted from cultured mouse 3T3-L1 adipocytes (Ntambi et al., 1988). In brief, genomic DNA isolated from cultured cells was partially digested with *Mbo*I and size-fractionated by sucrose density gradient centrifugation to separate fragments about 20 kb long (Maniatis et al., 1982). λ EMBL3 arms, digested with *Bam*HI, were obtained from Stratagene. Vector fragments and *Mbo*I-digested, genomic DNA were ligated, packaged in vitro with Gigapack Gold packaging extract (Stratagene), and titered with P2392 indicator cells. The library was amplified once before screening.

This genomic library was screened with a pSN2.2-2, P_2 cDNA probe (Narayanan et al., 1988) at reduced stringency (0.1× SSC/0.1% sodium dodecyl sulfate washed at 55°C; 1× SSC is 0.15 *M* NaCl, 0.015 *M* Na_3 citrate). A single positive phage clone, designated L25-1, containing a 20-kb insert was identified out of about 1×10^6 recombinants. The insert contained a single internal *Sal*I restriction site and several *Eco*RI and *Bam*HI sites (Fig. 1). Each of the fragments that resulted when the phage DNA was digested with *Bam*HI, *Eco*RI, *Bam*HI and *Sal*I, or *Eco*RI and *Sal*I was subcloned into a plasmid (Bluescript, Stratagene). By a series of cross-hybridization experiments with these subclones, overlapping segments were identified, and the restriction sites were mapped (Fig. 1). The rabbit P_2 cDNA clone (pSN2.2-2) hybridized to a single 4-kb *Eco*RI fragment of the L25-1 phage DNA. This segment, GPR-10, was sequenced entirely by creating nested deletions with exonuclease III (Henikoff, 1984), starting at either end of the insert (Fig. 1). Comparing the nucleotide sequence of the GPR-10 with the rabbit P_2 cDNA sequence and a partial sequence of a murine P_2 cDNA clone (Monuki et al., 1989 and sequence information provided by G. Lemke) allowed us to distinguish between the coding region and introns. The mouse gene codes for a 131-amino acid protein, differing from the rabbit P_2 at 16 residues, with most of these differences occurring in exon 3. These amino acid differences are indicated in Fig. 2. The nucleotide sequences at the splice junctions at the 5′ and 3′ ends of introns 1, 2, and 3 are all consistent with the 5′-GT . . . AG-3′ rule (Breathnach and Chambon, 1981).

We have identified *two* sites of transcription initiation by primer extension analysis. Total RNA was extracted from sciatic nerves of 10-day-old mice and adult mouse brain by using the acid guanidinium isothiocyanate–phenol–chloroform extraction method (Chomczynski and Sacchi, 1987). A synthetic oligonucleotide MP2Pr-Ex (5′-CTTTCATGT-AGTCATCGA-3′) complementary to the protein coding region of exon 1 of the mouse P_2 gene (Fig. 2) was end labeled to a specific activity of 5×10^6 cpm/pmol using [γ-^{32}P]ATP (>6,000 Ci/mmol) and T_4 polynucleotide kinase (Maniatis et al., 1982). Conditions for annealing the primer to the RNA and for extension were as described previously by Ntambi et al. (1988). Twenty micrograms each of yeast tRNA, mouse sciatic nerve RNA, or mouse brain RNA was annealed to 0.1 pmol of ^{32}P-labeled primer and extended with avian myeloblastosis virus reverse transcriptase. The reaction products were analyzed on a 6% polyacrylamide, 7 *M* urea sequencing gel, together with Sanger dideoxy sequencing reaction products of the genomic clone, GPR-10, as template and the MP2Pr-Ex oligonucleotide as primer (Fig. 3). These data indicate two transcription start sites and, thus, the presence of two P_2 mRNAs, with the larger transcript perhaps being more abundant, because it always gave a darker band on autoradiograms after primer extension analysis. However, additional studies will be required to confirm whether there are indeed two P_2 mRNAs.

The two sites of transcription initiation lie within a 53-bp direct repeat that immediately precedes the translation initiation ATG (Fig. 2). As indicated in Fig. 2, the 5′-flanking

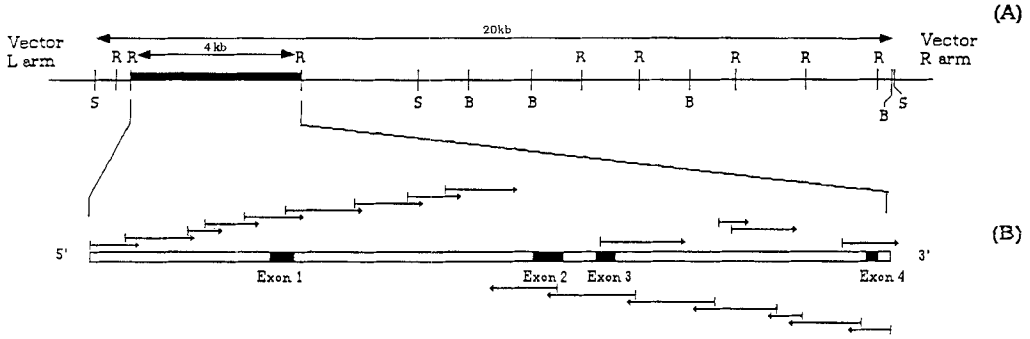

FIG. 1. Restriction map of L25-1 phage DNA and the mouse P_2 exon/intron structure. **A:** Shown on the top line is the restriction map of the L-25 phage DNA. R, *Eco*RI; B, *Bam*HI; S, *Sal*I. Shaded and labeled GPR-10 is the 4-kb *Eco*RI fragment that contained the protein coding segment of the gene. **B:** Details of the structure of the GPR-10 segment are shown here. Arrows indicated the various segments of this clone that were sequenced. The exons and introns were identified by comparison with the rabbit P_2 cDNA.

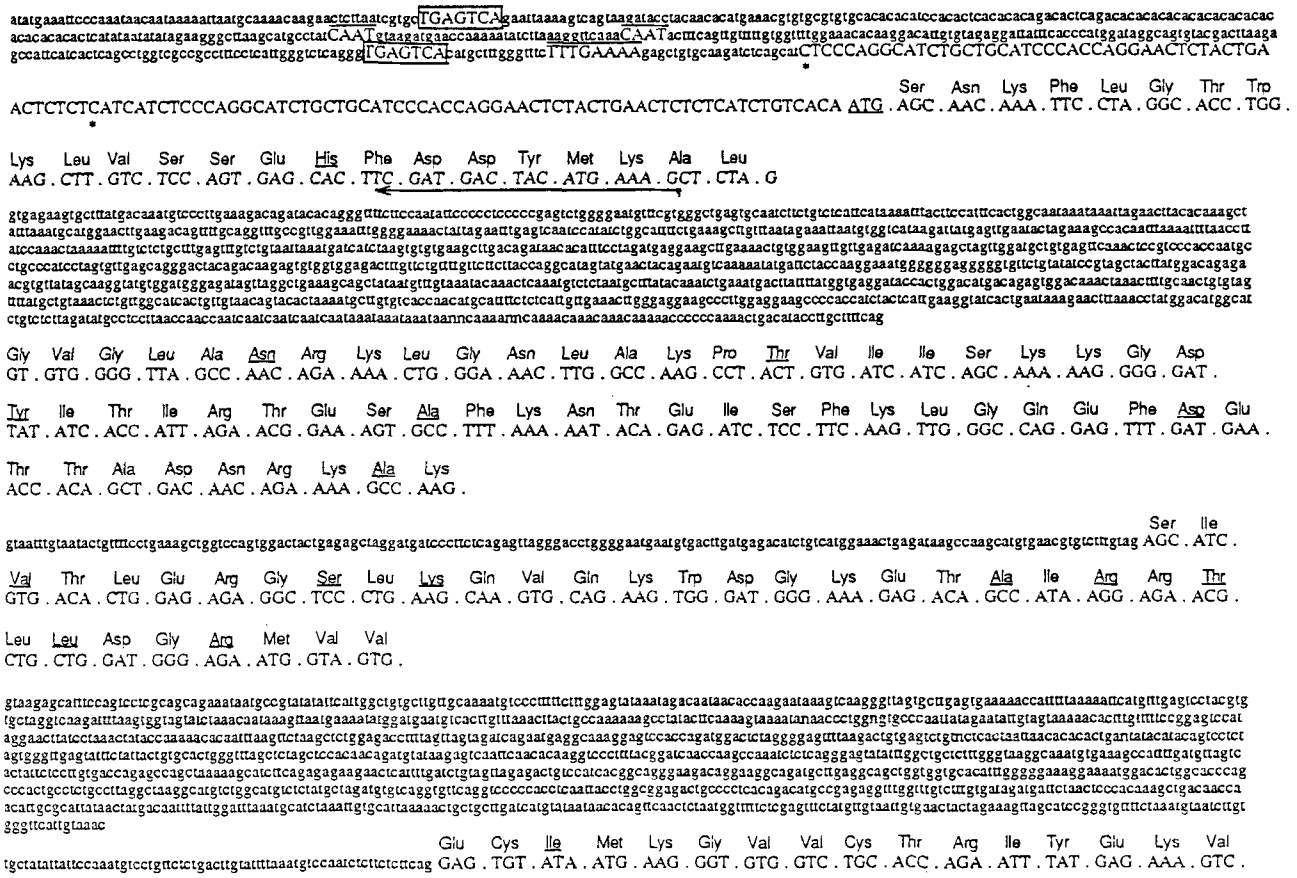

FIG. 2. Nucleotide sequence of the mouse P_2 gene. Part of the nucleotide sequence of the GPR-10 clone is shown above. The translation initiation ATG and termination TGA are underlined. Amino acid residues at which this differs from the rabbit P_2 protein are also underlined. The position of the oligonucleotide primer, MP2Pr-Ex, used for primer extension analysis is shown by an arrow beneath the complementary sequence in exon 1. The two sites of transcription initiation are marked by an asterisk under the corresponding nucleotide. The TATA and CAAT sequences are overlined. Two potential AP-1 binding sites are boxed. n, ambiguous nucleotide.

sequence of this gene contains consensus CAAT-box elements and an AT-rich segment, TTGAAA, which might serve as a TATA box. Lying within 400 bp of this TATA element are two potential AP-1 binding sites, 5′-TGAGTCA-3′.

DISCUSSION

One goal of our research is a molecular biological characterization of the events underlying myelination. These events include (a) the surface interaction between Schwann cell or oligodendrocyte and the axon; (b) the extracellular signals that induce the synthesis of lipids and myelin-specific proteins, which are then assembled into the vast surface area of myelin membrane; and (c) the intracellular messengers that mediate these changes. These processes involve regulation of gene transcription, mRNA translation, or posttranslational events. For a better understanding of these events, we have begun an analysis of the myelin P_2 protein gene and its regulatory elements. We had earlier characterized a cDNA encoding rabbit myelin P_2 protein (Narayanan et al., 1988). By using this as a probe, we have isolated and characterized the protein coding segment of the mouse P_2 gene.

We believe the gene we have characterized is the myelin P_2 gene and not the closely related aP_2/422 gene. Direct comparison of the nucleotide sequences of the two genes (Phillips et al., 1986; Hunt et al., 1986), examination of the sequence of the murine P_2 cDNA clone (Monuki et al., 1989) with the exon sequence of the mouse P_2 gene, and the results of the primer-extension experiments done with mouse sciatic nerve RNA and a primer complementary to a segment in the first exon in GPR-10 further support the contention that this gene is the mouse myelin P_2 gene. The exon/intron structure of this gene has been assigned by comparing the rabbit P_2 cDNA and the murine P_2 sequence with the genomic sequence. This mouse P_2 gene is divided into four exons by three introns, 1.2 kb, 150 bp, and 1.3 kb in length. We have compared the sequence of the mouse P_2 gene with that of the mouse aP_2/422

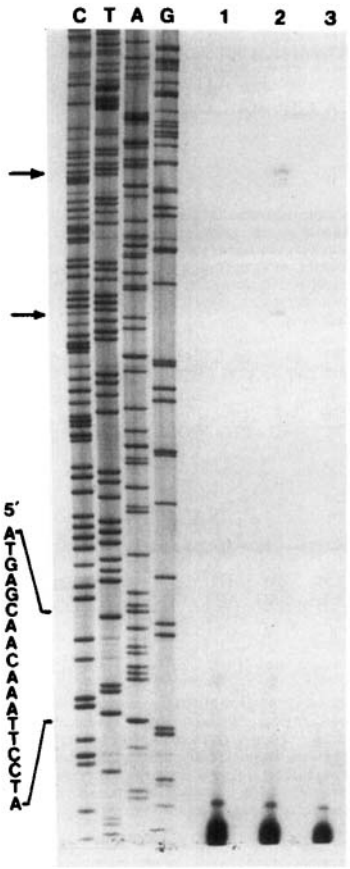

FIG. 3. Mapping the transcription initiation site. Shown above are the results of primer extension analysis (see Methods and Results) done with 20 μg of tRNA (lane 1), 20 μg of total RNA from mouse sciatic nerve (lane 2), and 20 μg of total RNA from adult mouse brain (lane 3). The oligonucleotide, MP2Pr-Ex, that was used in the primer extension experiment, was used to sequence the genomic clone, GPR-10, and the reaction products analyzed on the same polyacrylamide gel. The sites of transcription initiation are indicated by arrows. Shown to the left is the nucleotide sequence of a segment of GPR-10 starting with the translation initiation site.

gene, the rat liver FABP (Sweetser et al., 1986), the cellular retinol binding protein-II gene (Demmer et al., 1987), and the cellular retinoic acid binding protein gene (Shubeita et al., 1987) (Fig. 4). Each of these genes has the same overall structure containing four exons, but the sizes of the introns are quite different. The sites of the exon/intron boundaries occur at similar positions in the amino acid sequence in each of these genes. For instance, the first splice site occurs at a glycine residue (23, 24, or 25 depending on the numbering of the amino acids), and the second occurs at a lysine residue (80, 81, 82, or 84) in four of the five genes (Fig. 4). This degree of similarity strongly supports the idea that these various genes evolved from a common ancestral gene.

We have also sequenced about 1 kb of the 5'-flanking region of this gene, part of which is shown in Fig. 2. This 5'-flanking sequence contains an AT-rich region, which is a candidate TATA element, and two CAAT-box sequences, elements that are found in most eukaryotic promoters (Dynan and Tjian, 1985; Maniatis et al., 1987). Also interesting, a 53-bp direct repeat occurs immediately before the translation initiation site (Fig. 2). The two transcription initiation sites for P_2 mRNA lie at the 5' end of this directly repeated sequence, and by primer extension analysis, the larger transcript is perhaps the more abundant message. Another interesting sequence motif, 5'-TGAGTCA-3', occurring twice in this region, is a potential binding site for proteins in the AP-1 family (Fig. 2). These proteins that alter gene transcription are sequence-specific DNA-binding proteins that interact with enhancer elements in various genes (Curran and Franza, 1988; Turner and Tjian, 1989). The induction of specific genes by phorbol esters appears to be mediated by the binding of AP-1 proteins to these target sequences (Angel et al., 1987; Lee et al., 1987). Also, at least two oncogene products, Jun and Fos, belong to a family of transcriptional factors that bind directly or indirectly to the consensus AP-1 binding site (Curran and Franza, 1988; Turner and Tjian, 1989) and alter gene transcription. When investigating gene expression in adipocyte differentiation, Spiegelman and his colleagues identified similarities in the sequences of the 5'-flanking regions of three genes (Phillips et al., 1986). One of these sequences, 5'-CATGACTCAGAGGAAAAC-ATAC-3', called the fat-specific element 2 (FSE-2), occurs in the 5'-flanking region of the aP_2/422 gene and the glycerol-3-phosphate dehydrogenase gene. In the aP_2/422 gene, the FSE-2 element contains a consensus AP-1 binding site and binds a protein complex containing Fos (Distel et al., 1987; Rauscher III et al., 1988). This protein complex appears to regulate aP_2/422 gene activity in preadipocytes (Rauscher et al., 1988; Christy et al., 1989), although there is no similar sequence to FSE-2 in the mouse P_2 promoter. Identification of two consensus AP-1 binding sites in the mouse P_2 gene is of interest, because Monuki et al. (1989) have shown that, in rat Schwann cells in vitro, the induction of myelin P_0 and P_2 gene expression by forskolin is preceded by a fall in intracellular c-*jun* expression. This finding led them to speculate that the Jun protein down-regulates the expression of myelin genes by binding AP-1 binding sites. In preliminary studies with nuclear extracts containing c-*jun*, the mouse myelin P_2 promoter containing the AP-1 binding sites resulted in a shift of the labeled DNA fragment by DNA mobility shift assay (data not shown, Narayanan and Tennekoon). Whether this binding results in altering transcription is yet to be determined. Furthermore, comparison of sequence homologies in the 5'-flanking region with other genes coding for FABPs or for myelin proteins reveals no homologies. Currently, we are attempting to study how the myelin P_2 gene is regulated during myelinogenesis.

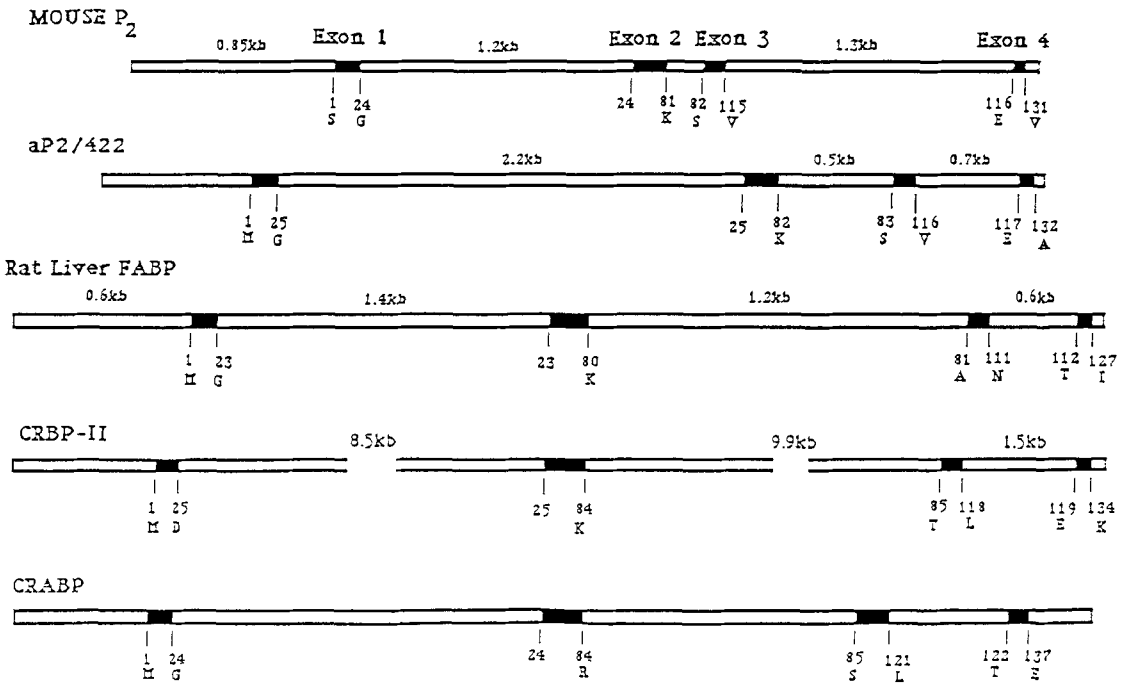

FIG. 4. Comparison of the structures of FABP genes. The exon/intron structures of the mouse P_2, aP_2/422, rat liver FABP, cellular retinol binding protein-II (CRBP-II), and the cellular retinoic acid binding protein (CRABP) genes are shown above. The locations of the splice sites are indicated by the number of the amino acid residue at which each occurs. Shown below are the residues on either side of the splice site, in single letter amino acid code. The lengths of the introns, where known, are as indicated.

Acknowledgment: We would like to thank Dr. James M. Ntambi for advice on library screening and Drs. Pamela Talalay and Keith Peden for help in the preparation of this manuscript. This work was supported by National Institutes of Health Grant NS21700. Dr. Vinodh Narayanan is a recipient of a Clinical Investigator Development Award NS1282.

REFERENCES

Angel P., Imagawa M., Chiu R., Stein B., Imbra R. J., Rahmsdorf H. J., Jonat C., Herrlich P., and Karin M. (1987) Phorbol ester-inducible genes contain a common cis element recognized by a TIPA modulated trans-acting factor. *Cell* **49,** 729–739.

Breathnach R. and Chambon P. (1981) Organization and expression of eucaryotic split genes coding for proteins. *Annu. Rev. Biochem.* **50,** 349–383.

Chomczynski P. and Sacchi N. (1987) Single-step method of RNA isolation by acid guanidium thiocyanate–phenol–chloroform extraction. *Anal. Biochem.* **162,** 156–159.

Christy R. J., Yang V. W., Ntambi J. M., Geiman D. E., Landschulz W. H., Friedman A. D., Nakabeppu Y., Kelly T. J., and Lane M. D. (1989) Differentiation-induced gene expression in 3T3-L1 preadipocytes: CCAAT/enhancer binding protein interacts with and activates the promoter of two adipocyte-specific genes. *Genes Dev.* **3,** 1323–1335.

Curran T. and Franza B. R. Jr. (1988) Fos and jun: the AP-1 connection. *Cell* **55,** 395–397.

Demmer L. A., Birkenmeier E. H., Sweetser D. A., Levin M. S., Zollman S., Sparkes R., Mohandas T., Lusis A. J., and Gordon J. I. (1987) The cellular retinol binding protein II gene. *J. Biol. Chem.* **262,** 2458–2467.

Distel R. J., Ro H.-S., Rosen B. S., Groves D. L., and Spiegelman B. M. (1987) Nucleoprotein complexes that regulate gene expression in adipocyte differentiation: direct participation of c-*fos*. *Cell* **49,** 835–844.

Dynan W. S. and Tjian R. (1985) Control of eukaryotic mRNA synthesis by sequence specific DNA binding protein. *Nature* **316,** 774–778.

Filbin M. T., Walsh F. S., Trapp B. D., Pizzey J. A., and Tennekoon G. I. (1990) The role of myelin P_0 protein as a homophilic adhesion molecule. *Nature* **344,** 871–872.

Greenfield S., Brostoff S., Eylar E. H., and Morell P. (1973) Protein composition of myelin of the peripheral nervous system. *J. Neurochem.* **20,** 1207–1216.

Greenfield S., Weise M. J., Gantt G., Hogan E. L., and Brostoff S. W. (1982) Basic proteins of rodent peripheral nerve myelin: immunochemical identification of the 21.5K, 18.5K, 17K, 14K, and P_2 proteins. *J. Neurochem.* **39,** 1278–1282.

Henikoff S. (1984) Unidirectional digestion with exonuclease III creates targeted breakpoints for DNA sequencing. *Gene (Amst.)* **28,** 351–359.

Hunt C., Ro J. H.-S., Dobson D. E., Min H. Y., and Spiegelman B. M. (1986) Adipocyte P_2 gene: developmental expression and homology of 5′-flanking sequences among fat cell-specific genes. *Proc. Natl. Acad. Sci. USA* **83,** 3786–3790.

Jones T. A., Bergfors T., Sedzik J., and Unge T. (1988) The three-dimensional structure of the P_2 myelin protein. *EMBO J.* **7,** 1597–1604.

Lee W., Mitchell P., and Tjian R. (1987) Purified transcriptional factor AP-1 interacts with TPA-inducible enhancer element. *Cell* **49,** 741–752.

Lemke G. (1988) Unwrapping the genes of myelin. *Neuron* **1,** 535–543.

Maniatis T., Fritsch E. F., and Sambrook J. (1982) *Molecular Cloning: A Laboratory Manual,* pp. 126–127, 280–281. Cold Spring Harbor Laboratory, Cold Spring Harbor, New York.

Maniatis T., Goodbourn S., and Fischer J. A. (1987) Regulation of

the inducible and tissue specific gene expression. *Science* **236,** 1237–1245.

Monuki E. S., Weinmaster G., Kuhn R., and Lemke G. (1989) SCIP: a glial pou domain gene regulated by cyclic AMP. *Neuron* **3,** 783–793.

Narayanan V., Barbosa E., Reed R., and Tennekoon G. (1988) Characterization of a cloned cDNA encoding rabbit myelin P_2 protein. *J. Biol. Chem.* **263,** 8332–8337.

Ntambi J. H., Buhrow S. A., Kaestner K. H., Christy R. J., Sibley E., Kelly T. J. Jr., and Lane M. D. (1988) Differentiation-induced gene expression in 3T3-L1 preadipocytes. *J. Biol. Chem.* **263,** 17291–17300.

Phillips M., Dijan P., and Green H. (1986) The nucleotide sequence of three participating in adipose differentiation of 3T3 cells. *J. Biol. Chem.* **261,** 10821–10827.

Rauscher F. J. III, Sambucetti L. C., Curran T., Distel R. J., and Spiegelman B. M. (1988) Common DNA binding site for fos protein complexes and transcription factor AP-1. *Cell* **52,** 471–480.

Sacchettini J. C., Gordon J. I., and Banaszak L. J. (1988) The structure of crystalline *Escherichia coli*-derived rat intestinal fatty acid-binding protein at 2.5-Å resolution. *J. Biol. Chem.* **263,** 5815–5819.

Shubeita H. E., Sambrook J. F., and McCormick A. M. (1987) Molecular cloning and analysis of functional cDNA and genomic clones encoding bovine cellular retinoic acid-binding protein. *Proc. Natl. Acad. Sci. USA* **84,** 5645–5649.

Sweetser D., Lowe J. B., and Gordon J. I. (1986) The nucleotide sequence of the rat liver fatty acid binding protein gene. *J. Biol. Chem.* **261,** 5553–5561.

Trapp B. D., Itoyama Y., MacIntosh T. D., and Quarles R. H. (1983) P_2 protein in oligodendrocytes and myelin of the rabbit central nervous system. *J. Neurochem.* **40,** 47–54.

Turner R. and Tjian R. (1989) Leucine repeats and an adjacent DNA binding domain mediate the formation of functional cFos-cJun heterodimers. *Science* **243,** 1689–1694.

Uyemura K., Yoshimura K., Suzuki M., and Kitamura K. (1984) Lipid binding activities of the P_2 protein in peripheral nerve myelin. *Neurochem. Res.* **9,** 1509–1514.