Identification of cis-Regulatory Elements in the Myelin Proteolipid Protein (PLP) Gene*

Jo Ann Berndt, Jin G. Kim‡, and Lynn D. Hudson§
From the Laboratory of Viral and Molecular Pathogenesis, National Institute of Neurologic Disorders and Stroke, National Institutes of Health, Bethesda, Maryland 20892

Regulatory elements of the proteolipid protein (PLP) gene were identified physically by footprinting and gel mobility shift assays and functionally by transfecting glial cell lines with PLP-chloramphenicol acetyltransferase chimeric genes. In both human and rat glial cells, only several hundred base pairs of upstream sequence are sufficient for high level activity of the human PLP promoter. This region contains five sites that contact nuclear proteins in vitro. More distal recognition sites may exist, as regions upstream of -524 displayed silencing activity indicative of a negative regulatory element. A series of site directed mutations revealed one essential positive element (ATGGA at -118) which is found in other genes encoding myelin proteins. Our combined biochemical and functional analyses indicate that the key cis sites for maximal tissue-specific expression of PLP in cultured glial cells are clustered near the promoter. Within this cluster are several conserved motifs that may coordinate the regulation of myelin-specific genes.

Coordinate gene regulation may be most simply achieved by the action of one or more common trans-acting factors with preserved cis elements that are located within the selected genes, as originally proposed by Britten and Davidson (1971). In muscle a single trans factor, MyoD, can activate, via a sequence found in most muscle-specific genes, the battery of genes required to turn an undifferentiated cell or even a differentiated cell like a fibroblast into a muscle cell (Weintraub et al., 1991). Other tissues appear to employ more complex patterns of regulation that may still originate from a master regulator, but one that behaves in a more restricted fashion, such as by activating only a set of other transcription factors whose further activity eventually culminates in the expression of the differentiated phenotype. For example, in the liver, no one transcription factor has yet been described that can singlehandedly activate liver-specific genes; nor has the same cis element been found in each hepatic gene, nor is the distribution of the transcription factors that act in hepatocyte-specific gene expression limited to liver cells (Xanthopoulis et al., 1991). Thus, at the regulatory level in which the final products of a differentiated cell are synthesized, the hepatocyte must employ a number of transcription factors that work in concert with one another.

While features of tissue-specific controls are emerging, little is known about the stage-specific controls that serve to temporally regulate a set of genes in a differentiated cell. Myelinating glial cells offer a suitable system for examining the molecular basis of coordinate controls, since for a brief period these cells must synthesize vast quantities of myelin membrane. The synthesis of myelin proteins, which are expressed exclusively in glial cells, appears to be coordinately controlled at the transcriptional level. Both in the central and peripheral nervous systems, glial cells are influenced to myelinate both by the neuronal targets that ensheathe and by a range of hormones and growth factors (reviewed by Mirsky and Jessen (1990), Lemke (1990), and Richardson et al. (1990)). Only a limited repertoire of proteins are incorporated into myelin, with myelin basic protein (MBP)* and proteolipid protein (PLP) contributing 50% of total protein. A number of myelin-specific genes have been cloned, including those encoding the structural proteins MBP, PLP, and PO, proteins that may play a role in neuronal-glial cell adhesion, such as myelin-associated glycoprotein and possibly oligodendrocyte-myelin glycoprotein, and proteins that may function in the transport of myelin constituents (P2 basic protein) or the assembly of the myelin sheath (2',3'-cyclic-nucleotide 3'-phosphodiesterase) (reviewed in Hudson (1990)). Regulatory regions, including several promoter elements that govern brain-specific transcription, have been identified in the MBP gene (Tamura et al., 1988, 1989, 1990; Miura et al., 1989; Aoyama et al., 1990; Devine-Beach et al., 1990; Leshigari et al., 1990).

We have chosen the PLP gene as a prototype myelin gene to identify the trans-acting factors that regulate myelination. While PLP is abundantly expressed with the cohort of myelin genes during the peak of myelination, the expression of this gene can also be uncoupled from the coordinate controls exerted on myelin genes. The PLP locus encodes at least two, and very likely a third, proteolipid proteins: PLP, DM-20, and a novel proteolipid with an amino terminus identical to PLP/DM-20 (Schindler et al., 1990). Transcripts originating from the PLP locus are detectable prior to the onset of myelination (Gardinier and Macklin, 1988; LeVine et al., 1990) and long before oligodendrocytes are generated (E15 in mouse brain), suggesting that some of the alternatively

* The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

The nucleotide sequence(s) reported in this paper has been submitted to the GenBank™/EMBL Data Bank with accession number(s) M95057.

‡ Supported by the National Research Council and the National Multiple Sclerosis Society.
§ To whom correspondence should be addressed: Bldg. 36, Rm. 5D04, NIH, Bethesda, MD 20892. Tel.: 301-496-9660; Fax: 301-496-0899.

1 The abbreviations used are: MBP, myelin basic protein; PLP, proteolipid protein; CAT, chloramphenicol acetyltransferase; EMSA, electrophoretic mobility shift assay; kb, kilobase pair; bp, base pair, HEPES, 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid.

2 J. Kim and L. Hudson, unpublished results.
spliced isoforms of the PLP locus may have a role other than that of a structural protein of the myelin sheath. A second function for the PLP locus in early nervous system development was also suggested by the pleiotropic phenotype of PLE mutations, which includes an excessive proliferation of oligodenodrocyte progenitors and a profound loss of myelin, and oligodenodoncytes (reviewed in Hudson and Nadon (1992)). In this report the cis elements that bind these putative trans-acting proteins are characterized functionally by transferting portions of the PLP promoter/upstream region into cultured glial cells and physically by a combination of gel shift and footprinting assays. Maximal tissue-specific expression requires a small evolutionarily conserved region which contains several binding motifs that may participate in either the glial- or stage-specific expression of myelin genes.

**EXPERIMENTAL PROCEDURES**

**Construction of Deletion and Site-directed Mutations in the PLP Promoter**—The vector for all deletion constructs was the promoterless, enhancerless p18CAT plasmid (McKinnon et al., 1986), which expresses CAT levels and levels of chlamomphenicol acetylase activity when transfected in glial cells. Other vectors such as the original one devised for testing promoter constructs, pSVOCAT (Gorman et al., 1982), produced unacceptably high levels of CAT activity in glial cells; in a typical experiment in NG-108 cells, pSVOCAT gave only slightly higher levels (245 cpm/µg of protein) and p18CAT gave background levels (0.4 cpm/µg of protein).

Restriction fragments of the human PLP upstream region were subcloned into p18CAT in both orientations from pB010, a plasmid derived from b2 (Hudson et al., 1989) and which contains the 4.2-kb human PLP upstream region, promoter and exon. Each promoter construct extended to 85 of the human PLP sequence and contained an upstream region that ranged in size from 18 to 4200 bp and were named accordingly. The -524, -312, -274, and -254 constructs were prepared by polymerase chain reaction using pB010 as template, then sequenced to confirm the absence of polymerase chain reaction artefacts.

Site-directed mutagenesis was carried out by a polymerase chain reaction-based method (Jones and Howard, 1990) using the plasmid with 1086 bp of the PLP upstream region as the template. For each construct, approximately 150 bp of DNA encompassing the site of the mutation were sequenced by the dideoxy method with the modified phage T7 DNA polymerase (Sequenase; U.S. Biochemical).

Plasmids which served as positive controls included pSV2CAT, in which the SV40 promoter was cloned upstream of the CAT reporter (Gorman et al., 1982), and pSVOCAT-luciferase, which used the Rous sarcoma virus promoter to drive firstly luciferase expression (De Wet et al., 1987).

**Transfection of Glial and Nonglial Cells**—Two cell lines which displayed some glial characteristics were used for transfections: SVG cells, which were established by transfecting human fetal glial cells with an origin-defective SV40 (Major et al., 1985), and B103 cells, which were cloned from nitrosoethylurea-induced brain tumors in rats (Schubert et al., 1986). SVG cells appear to be of an astroglial lineage, as they express the astrocyte marker glial fibrillary acidic protein (Major et al., 1985). B103 cells have both neuronal and glial properties (Schubert et al., 1986), but they express oligodenodrocyte-specific messages such as PLP (Monuki et al., 1989), indicating that these cells have the complement of transcription factors necessary to transcribe myelin genes.

Cells were seeded at low density the day prior to transfection so that at the start of the transfection they were approximately 60% confluent. In some experiments with B103 cells, 20 µM forskolin (Calbiochem) was added at the time of plating, and replenished during the transfection. For B103 and SVG cells, transfections were carried out by the calcium phosphate method as reviewed by McKinnon and Graham (1986) except no carrier DNA was included. Each plasmid was tested by gel electrophoresis to ensure that over 90% of the preparation was in a supercoiled form, and then transfected in 10 mm Tris-HCl, 1 mm EDTA (pH 8.0) were coprecipitated with the pRSV-luciferase plasmid as an internal control for standardizing transfection efficiency as a factor of pH 7.1 containing a final concentration of 0.125 M CaCl2, 21 mM HEPES, 135 mM NaCl, 5 mM KCl, 0.7 mM Na2HPO4, 5.5 mM D-glucose. For each 100-mm plate, 10 µg of test plasmid with either 10-100-fold less pRSV-luciferase plasmid in a total volume of 1 ml were directly added to the media (10 ml of Dulbecco's modified Eagle's medium (GIBCO) containing 10% fetal calf serum). DNA was incubated with cells for 17-19 h in 10% CO2, then the media were changed and the cells were cultured for an additional 48 h before harvest. CAT activity was directly quantitated by the scintillation method (Neurath et al., 1987) that is based on the [3H]acetyl-CoA substrate (NET-390L, 200 mCi/mmol, Du Pont-New England Nuclear) and the solubility of the acetylated chlamomphenicol product in the organic scintinating (Econofluor, Du Pont-New England Nuclear). Only 0.5 µCi final concentration 10 µl of [3H]acetyl-CoA was used in each reaction, as the luminescence linear with as little as 0.1 µCi (de Wet and Neurath, 1987). Luciferase activity was measured according to de Wet (1987) using the potassium salt of d-luciferin (Analytical Luminescence Laboratory) on a Monolight 2010 luminometer (Analytical Luminescence Laboratory).

For HeLa cells, transfections were carried out with a modified DEAE-dextran method (Hulten et al., 1989) which were briefly exposed to DEAE-dextran (M, 500,000; Pharmacia) then washed with Dulbecco's modified Eagle's medium and incubated with DNA in the presence of 0.1 mM chloroquine (Sigma).

For the transfection results, a minimum of three independent observations were used to calculate the mean and standard deviation. Significance of promoter expression was based on the fold stimulation in each condition (cell passage number, cell density and serum lot), and many constructs were transfected in over six experiments with each cell line and more than one plasmid preparation was tested for each construct. In each experiment, the results obtained when comparing constructs by calculating the CAT specific activity (see Figures 1 and II) were similar to those obtained by normalizing CAT activities to the amount of cotransfected luciferase.

**EMSA and Methylation Interference Footprinting of the PLP Promoter**—Nuclear extracts were prepared from HeLa cells by the method of Shapiro et al. (1986) and from glial cell lines by the "mini-extractor procedure of Schreiber et al. (1988). Extracts from the mouse brain, nuclei were isolated essentially as described by Gorski et al. (1986), then resuspended in 20 mM Tris-HCl (pH 7.9), 25 mM KCl, 0.15 mM spermine, 0.5 mM spermidine, 1 mM EDTA, 2 µM sucrose, 10% glycerol. Nuclei were disrupted with a glass B Dounce and the supernatant from a 25,000 × g, 30-min centrifugation was dialyzed against 20 mM Tris-HCl (pH 7.9), 20% glycerol, 0.1 M KCl, 0.2 mM EDTA, 0.5 mM dithiothreitol, 0.5 mM pheynethylsulfonyl fluoride and recentrifuged to remove debris. A typical yield was 6.2 mg of nuclear protein from 15.2 g of brain.

For the EMSA, binding reactions consisted of 100,000 cpm of 32P-end-labeled, double-stranded oligonucleotide (usually 30-mers; approximately 0.2 ng of DNA), 4 µg of poly-d(1-dC)poly(d1-dC) (Pharmacia) as carrier and 5-10 µg of nuclear protein in a final volume of 20 µ containing 10 mM Tris-HCl (pH 8.0), 50 mM KCl, 100 mM β-mercaptoethanol, 0.1% Triton X-100, 4% glycerol. For both the EMSA and footprinting experiments, a range of protein concentrations was used in preliminary experiments. The mixture was incubated on ice for 10 min with or without a 100-fold molar excess of unlabeled competitor and for 20 min with the labeled oligonucleotide. DNA-protein complexes were resolved at room temperature on a 6% acrylamide (30:1, acrylamide:acrylamide) gel containing 25 µM Tris-HCl (pH 8.3), 100 mM glycine, 1 mM EDTA.

For methylation interference footprinting, a 32P-end-labeled restriction fragment or synthetic nucleotide was partially methylated by dimethyl sulfate, then incubated with nuclear extract and electrophoresed on a nondenaturing gel as above to separate the DNA-protein complexes from the "free" (not bound to protein) DNA. The gel was transblotted onto DB1 paper and the multiple radioactive bands were then cut out and dried. DNA-protein complexes were hydrolyzed to nucleotides in 0.3 N NaOH, then separated by thin-layer chromatography. The unbound DNA was eluted directly and scanned for bands or footprinting (only oligo 4; see Fig. 4, site 4) included the following. As noted by underlining, in some cases the sequences of the oligonucleotide termini were not colinear with the DNA sequence of Fig. 2, as some oligonucleotides were used for preparing catenates of a binding site in order to screen expression libraries.
Transcription factors to test putative DNA binding regions of the cells were readily transfected with a luciferase plasmid activity from an SV40 promoter. HeLa cells were chosen as constructs to the same extent as the strong viral promoter (SV40) amphenicol acetyltransferase activity was present in the non-glial lines. As shown in Table I, CAT activity following transfection with the myelin gene constructs (Table I, Fig. 1).

### RESULTS

Expression of PLP Promoter Activity Is Restricted to Glial Cells—A series of promoter constructs containing successively shortened regions of the PLP promoter subcloned upstream of the CAT reporter were transfected into both glial and nonglial cell lines. As shown in Table I, no significant chloramphenicol acetyltransferase activity was present in any of the PLP constructs, although the cells were readily transfected with a luciferase plasmid used as an internal standard and displayed high levels of CAT activity from an SV40 promoter. HeLa cells were chosen as the nonglial line because of their human origin (derived from a cervical carcinoma) and their use as a rich source of general transfection assays.

Limited transfection of nonglial and glial cell lines with human PLP promoter constructs. To determine relative promoter strength, CAT activity in those constructs with over 1 kb of 5' sequence. Transcription of PLP in glial cells are in the close vicinity of the TATA box, leaving only 18 bp before the initiation of transcription, markedly reduced but did not abolish activity. Thus, most of the regulatory sites necessary for optimal transcription of PLP in glial cells are in the close vicinity of the PLP promoter. CAT activity in those constructs with over 1 kb of 5' sequence. Transcription is diminished 3-fold from a putative "silencer" (s) located between -1088 and -2000 in B103 cells and between -524 and -2000 in SVG cells (Fig. 1).

Contacts between the PLP Promoter and Nuclear Proteins Detected by Gel-shift Analysis—A combination of EMSA and dimethyl sulfate footprinting assays defined five sites of interaction with nuclear proteins (boxed regions in Fig. 2) within the 312 bp that confer maximal PLP promoter activity in transfection assays.

The binding side which encompasses the initiation of transcription and is related to a NF-I site (labeled oligo 1 in Fig. 2) displayed a simple EMSA pattern (Fig. 3A). The same two protein-DNA complexes were present in mouse brain, SVG and B103 cells (lanes 2, 5, and 8, Fig. 3A), but complex II was more abundant in the cell lines than in brain. These complexes were specific, as they could be competed with a 100-fold molar excess of unlabeled oligo 1 (Fig. 3B). No competition was detected when a mutant oligonucleotide (labeled mut 1a in Fig. 2) was added, indicating that the nuclear protein which binds oligo 1 is unable to recognize the altered oligo 1.

The binding site which lies within an unusually long, exclusively purine stretch of the PLP promoter, oligo 2, formed 5 complexes that could be specifically competed with the cognate unlabeled oligonucleotide, but not with an unrelated one (lane 3 versus 6, Fig. 3B). Mutant oligonucleotides that maintain the purine character but scramble the sequence of oligo 2 (mut 2a) or include pyrimidines to disrupt the string of purines (mut 2b) could only form significant amounts of complex III (lanes 9 and 12, respectively, Fig. 3B).
Functional Analysis of the PLP Gene

Fig. 2. Sequence of protein-binding sites in the human PLP promoter region. The five sites in the PLP promoter that bind nuclear proteins are boxed: sites I–IV, which displayed reproducible footprints, are shaded. Site 5, which bound protein only by EMSA, is represented by an open box. The purine stretch encompassing site 2 is in italics. Asterisks denote the guanosine residues on either strand that define DNA-protein contact sites as determined by dimethyl sulfate footprinting (Fig. 4). The TATA box is underlined, and the initiation site start site marked by +1. On the basis of nucleic acid protection experiments (Macklin et al., 1987), PLP has a single major transcription initiation site; the additional minor bands often seen as primer extension experiments most likely represent prematurely terminated transcripts or extensions of partially degraded PLP transcripts. The sequence extends to +85, which is the 3' end of all constructs subcloned into the pBluescript vector. An inverted CCAAT element is marked by an arrow at −145 to −140. Inverted triangles indicate the positions of the deletion constructs. Below each sequence are shown the mutations that were tested in EMSA and transfection experiments. The thyroid hormone half-site AGGACA found in site 4 (at −258) and at +26 are noted by dots; both the −258 site (Fig. 4) and the +26 site (Nave and Lemke, 1991) bind nuclear proteins in footprinting assays.

from both glial lines resembled the brain extracts shown in Fig. 3B, but the B103 line displayed less intense binding. The third binding site formed four complexes on a gel shift with brain or glial cell extracts (lanes 1, 5, 9, Fig. 3C), but these complexes were expressed in different proportions from the three sources of extracts. Complex II was prominent in the glial cell lines while barely detectable in brain, while complex III and IV were diminished in B103 cells. Forskolin treatment of B103 cells did not change the pattern of complexes I–IV (Fig. 3C, lane 11) but could be competed with oligo 3 or mut 3a (lanes 4, 6, and 8) and, when radioactively labeled, could form complexes I–IV (Fig. 3C, lanes 4, 8, and 12). The other mutated binding site (labeled mut 3b in Fig. 2) did not compete for binding of complexes I–IV and, when labeled, was not able to form complexes I–IV but did form several complexes of altered mobility in SVG and B103 extracts (Fig. 3C, lanes 7 and 11). These results suggest that the recognition site for nuclear proteins on oligo 3 includes the two guanosines at positions −116 and −115 and excludes positions −112 through −110. Further delineation of the sequences involved in forming complexes I–IV arose fortuitously from the construction of the mutant oligonucleotide 4b, which happened to form the same pattern of complexes as oligo 3 (Fig. 3D, lane 9). These complexes could not be competed with oligo 4 (Fig. 3D, lane 10) or mut 3a (lanes 13) but could be competed with oligo 3 or mut 3b (lanes 11 and 12). As mut 4b shares the sequence ATGGA with oligo 3, and the replacement of guanosines in this sequence with cytosines (in mut 3a) eliminates the gel shift bands, ATGGA must be the basis of the sequence-dependent binding of oligo 3.

The fourth binding site showed a different pattern of complexes with each extract. Of the four complexes detected with SVG extracts (Fig. 3E, lane 5), only one (complex II) appeared to overlap with those expressed by B103 cells (Fig. 3E, lane 9) and none were in common with mouse brain (Fig. 3E, lane 1). Both mutated oligonucleotides competed to some extent with the unaltered oligo 4 in forming complexes I–IV (Fig. 3E, lanes 7 and 8), and both mutant oligonucleotides when radioactively labeled were able to form complexes III and IV (Fig. 3F, lanes 7 and 8).

Contacts between the PLP Promoter and Nuclear Proteins Identified by in Vitro Footprinting—Of the five sites that bind nuclear proteins by EMSA, three were also found by footprinting methods to contact protein. As shown in Fig. 4, both HeLa and brain extracts protected guanosine residues in sites 1 (lane 2 for top strand, lanes 4 and 5 for bottom) and 3 (lanes 2) in a similar pattern. The region of purine asymmetry, site 2, was faintly detectable by dimethyl sulfate footprinting, but could be observed by DNase I footprinting. Site 5 formed two complexes from brain extracts that were distinct from Hela by EMSA, but these interactions were undetectable by either method of footprinting. The remainder of the PLP promoter, extending to approximately −600, did not display significant binding of nuclear proteins by dimethyl sulfate footprinting assays.

The proteins that bound to sites 1–3 were further characterized by a Southwestern analysis, in which nuclear extracts of mouse brain or HeLa cells were electrophoresed on denaturing gels, blotted, reratured, and probed with labeled oligonucleotides. Oligo 1 bound to a single band of approximately 66 kDa in both brain and HeLa extracts, oligo 2 bound predominantly to a 60 kDa protein in both brain and HeLa extracts, and oligo 3 bound a 120-kDa protein unique to brain extracts in addition to faintly recognizing six other bands which were also present in HeLa extracts.

Site-directed Mutagenesis Reveals One Site Essential for PLP Promoter Activity in Transfected Cells—Of the sites in the PLP upstream region that could bind nuclear proteins, only mutation in site 3 abolished promoter activity in transfected glial cells (Table II). Two different mutations were created in site 3 by replacing the guanosine residues that contact nuclear proteins as identified by dimethyl sulfate footprinting (Fig. 4). By gel mobility shift analysis, only the substitution of cytosine for guanosine at position −116 and −115 (Fig. 2, mut 3a) resulted in an oligonucleotide which no longer formed the same DNA-protein complexes and was unable to compete for binding with the wild-type oligonucleotide (Fig. 3C). Incorporation of these changes into a PLPCAT plasmid that contained 1088 bp of upstream region

3 J. Kim, unpublished results.
resulted in nearly a 100-fold reduction of reporter activity in B103 cells (Table II, mut 3a).

For the remaining site-directed mutations, little change in activity was noted in SVG or B103 cells. Mutations in the purine stretch of site 2 were not evaluated by transfection, as the site 2 mutant oligonucleotides shown in Fig. 2 retained binding of several specific complexes by EMSA (Fig. 3).

Conservation of the Five Protein-binding Sites in the PLP Promoter—The PLP promoter region, which consists of a cluster of five protein-binding sites that together confer high levels of reporter activity in transient transfection assays, is highly conserved between rodents and man. Sites 1–3 are included in a stretch (to -203) that is 86% homologous between mouse and man: none of the nucleotide differences fall within the binding regions of these sites, as shown in Fig. 5. Further upstream, in the region encompassing sites 4 and 5 (-311 to -203), the homology drops to 65%. Like sites 1–3, site 5 is well conserved, with only a single base change distinguishing mouse from the rat or human sequence.

Only site 4, which includes at least two different protein binding sites, shows variation among species. Site 4 is also the region that shows the greatest interspecies differences on EMSA (Fig. 3, E and F): no common bands were observed between mouse and other species, while only one of over four complexes migrated similarly between rat (B103) and human (SVG) extracts.

This can be explained in part by the redundancy of site 4, as the human site 4 contains one section (CATCTTC) that is also present in an inverted orientation in site 3, and within the palindromic arrangement of site 4 is another motif (TCCA) that forms the core of the well conserved sites 3 and 5.

Sites 1–5 are represented to varying extents in the two other myelin promoters for which sequence information is available. One or more purine stretches like that of site 2 are found in all myelin promoters, as is the core of site 3 (ATGGA) and 5 (CCACCC). Site 4 displays a striking similarity to two juxtaposed elements of the PO promoter, one with an 11-base identity with PLP and the other in the inverse orientation with a 5-base identity to the remaining segment of site 4. The
**Functional Analysis of the PLP Gene**

**Fig. 4. Dimethyl sulfate footprinting of two regions in the human PLP promoter that bind nuclear proteins.** Contacts between guanine residues and nuclear proteins as revealed by dimethyl sulfate footprinting are illustrated by filled circles next to the sequence (refer to Fig. 2 for the complete sequence and relative positions of these sites). For site 1, a 130-bp SacI-StuI fragment was used for the top strand: lane 1, G ladder of the 130-bp SacI-StuI fragment; lane 2, HeLa extract; lane 3, mouse brain extract; lane 4, Roc-1 extract; lane 5, free (unbound) DNA. For the bottom strand of site 1, a 118-bp HindIII-StuI fragment was labeled: lane 1, G ladder of the 118-bp HindIII-StuI fragment; lanes 2 and 3, mouse brain extract; lanes 4 and 5, HeLa extract. When separating free DNA from protein-DNA complexes, often multiple bands of bound complexes were detected on non-denaturing gels; each band corresponding to a complex was eluted and subsequently analyzed on a denaturing gel (e.g., lanes 2 and 3 represent two different complexes formed when nuclear proteins interacted with the 118-bp HindIII-StuI fragment). For site 3, a 150-bp SacI-HaeIII fragment was labeled for the top strand: lane 1, G ladder; lane 2, HeLa extract; lane 3, mouse brain extract; lane 4, free (unbound) DNA. For the bottom strand of site 3, a 171-bp PstI-HaeIII fragment was end-labeled: lane 1, free (unbound) DNA; lane 2, HeLa extract; lanes 3 and 4, mouse brain extract; lane 5, G ladder of 171-bp PstI-HaeIII.

**Table II. Transfection of glial cell lines with PLP promoter mutations**

Upper case letters, wild-type sequence; lower case letters, altered bases. The relative CAT activities are expressed as indicated under "Experimental Procedures." Values are means ± S.D.

| Site-directed mutations | CAT activity |
|-------------------------|--------------|
| | SVG | B103 | % wild type |
| site 4 | 100 | 100 | 100 |
| site 5 | 170 (±22) | 162 (±18) | |
| site 3 | 202 (±79) | 125 (±14) | |
| site 1 | 53 (±17) | 47 (±10) | |
| mut4a | 9 (±12) | 1 (±2) | |
| mut4b | 177 (±38) | 169 (±89) | |

MBP promoter does not share this extended identity, although it has individual blocks that comprise the two core elements of site 4.

Site 1 resembles a NF-I site (TGGCTNNAGCCCA) and functions as a recognition element for purified NF-I *in vitro* and is necessary for optimal MBP activity (Tamura et al., 1988, 1989; Aoyama et al., 1990). Moreover, site 1 is identical in 10/14 positions to an NF-I binding site (TGGCGTCCAGCCCA) characterized by Amemiya and co-workers (1989) in the JC virus, a human polyomavirus whose growth in the brain is limited to glial cells. A recently cloned transcription factor that binds the NF-I-related sequence of the JC virus is structurally unrelated to NF-I (Kerr and Khalili, 1991) and may not recognize myelin gene promoters, since an oligonucleotide containing the NF-I-like sequence of the JC virus migrates differently from the PLP site 1 on EMSA.

A putative thyroid hormone response element present in both site 4 of the human PLP gene and in the highly conserved +26 region of all PLP genes (Fig. 2) consists of the same half-site (AGGACA) as has been identified in the MBP gene (Farsetti et al., 1991). Nuclear runoff assays and transfection
Functional Analysis of the PLP Gene

Fig. 5. PLP-regulatory motifs found in other glial promoters. Sites 1–4 refer to the sequences found to bind nuclear proteins both by EMSA and footprinting analysis (illustrated in Fig. 2). Positions contacted by proteins in dimethyl sulfate footprinting are marked by asterisks; regions protected in DNase I footprinting experiments are boxed. The site number 5 was detectable in EMSA but not in footprinting assays. Sites 1, 2, 3, and 5 are highly conserved in the PLP promoter regions of all species examined. Sequences shown are from Hudson et al. (1989) (human PLP), Gence and Hudson (1989) (mouse PLP), Nave and Lemke (1991) (rat PLP), Kambholz et al. (1988) (human MBP), Miura et al. (1989) (mouse MBP), Lemke et al. (1988) (rat PO). The numbering of each sequence represents the distance from the transcription start site, with parentheses used to indicate sequences in the reverse orientation.

The combined biochemical and functional analyses indicate that most of the key sites necessary for maximal PLP expression in cultured glial cells are clustered near the promoter. In two different glial cell lines, only a few hundred bases of the PLP upstream region both strongly drove the CAT reporter for glial-specific transcription factors, which sites are instrumental in the coordinate control of myelin genes, and which are subject to ubiquitous factors? Based on sequence comparisons (Fig. 5), binding sites 2 and 3 are among the most conserved between the three myelin gene promoters, thus recommending these cis elements for a role in myelination. Moreover, most of the conserved sites (1–5 in Fig. 5) fall within upstream regions shown to be functionally important for the MBP (Tamura et al., 1989; Miura et al., 1989; Devine-Beach et al., 1990), PO (Lemke and Chao, 1988), and PLP (this study) genes. Several sites in the PLP promoter may be contacted by either ubiquitous or brain-specific proteins, as suggested by the expression cloning of both a novel brain-specific DNA-binding protein and a ubiquitous transcription factor that recognize the purine-rich site 2. The NF-1-like site of PLP is unlikely to represent a point of coordinate control for myelin genes, based on the low degree of homology between the NF-1-related sites (Fig. 5) which implies the existence of an array of transcription factors, including alternatively spliced NF-1 polypeptides (Santoro et al., 1988), brain-specific members of the NF-1 family (Aoyama et al., 1990), or factors structurally unrelated to NF-1 (Kerr and Khalili, 1991) that recognize these sites. Sites 4 and 5 are preserved in PO but not as well in MBP. If these elements influence glial-specific transcription, then glial cells may have a multiplicity of "tissue-specific" transcription factors, with differing combinations acting on genes exclusively expressed by glia.

Common hormonal controls on myelin genes may exist, since both MBP (Farsetti et al., 1991) and PLP display a form of the thyroid hormone response element. In the steroid hormone superfamily, selective recognition of a particular binding site depends in part on the spacing between half-sites, in which direct repeats of the core sequence (AGGTCA) separated by 3, 4, or 5 nucleotides are recognized selectively by the vitamin D3, thyroid hormone, or retinoic acid receptors, respectively (Umesono et al., 1991; Naar et al., 1991). These relationships predict that other steroid family members may fill the remaining spacing slots (Umesono et al., 1991). Given the modest effect of thyroid hormone on MBP expression (Farsetti et al., 1991), it is conceivable that MBP and PLP are regulated by another member of the steroid receptor family that recognizes myelin promoters more efficiently than thyroid hormone.

Both glial cell lines displayed a similar pattern of expression for the panel of PLP promoter constructs, despite the fact that one of the glial lines transcribes the endogenous PLP gene (B103 cells) and the other (SVG cells) does not. Also, the human PLP promoter was just as efficiently transcribed in rodent (B103) as human (SVG) cells. Nor did the nature of the transformation of these glial cells lines (chemically induced for B103 versus virally induced with an origin-defective SV40 for SVG cells) affect the pattern of PLP expression. Finally, differences in the complement of transcription factors assayed by gel shifts (Fig. 3) had little impact on the transfection efficiencies of the panel of progressively shortened PLP promoter constructs, an observation consistent with the cooperative nature of transcription factors in regulating gene expression. Site-directed mutagenesis identified one site (ATGGG; site 3) that was essential for activating reporter gene expression in each cell line. Since both lines expressed the same DNA binding protein(s) that recognize the normal site, the inability of this common trans-factor(s) to bind the mutated site 3 must result in inefficient transcription, a situation that could not be compensated by the remaining complement of transcription factors binding to PLP regulatory regions. That HeLa cells do not express the PLP gene is most likely due to the absence of several glial-specific transcription factors, since neither site 4 or 5 produce complexes with HeLa cells of the same mobility as seen in brain extracts.

Interspecies variation may be responsible for differences observed between the rat (Nave and Lemke, 1991) and human (this study) PLP regulatory regions, especially for sites located some distance from the highly conserved promoter. The rat PLP gene appears to have a silencer at -480 to -391 (Nave and Lemke, 1991), while negative regulatory elements in the human gene are located much further upstream (Fig. 1). Two of the sites footprinted in the rat PLP gene were deemed important for myelination based on sequence comparisons.
with other myelin promoters (Nave and Lemke, 1991); however, neither of these sites in the human PLP promoter binds protein in vitro, and the upstream site (~342) was not necessary for maximal activation of the human PLP construct in transfection experiments. In both studies considerable variation in transfection efficiencies was observed, a problem which was not reduced by the use of an internal standard. This variation limits interpretation of the transfection data, in that 2-fold differences may not be biologically significant. Nonetheless, for both the human and the rat PLP gene, most of the cis elements required for maximal expression of the PLP gene in transfected glial cell lines are located in close proximity to the promoter, within 312 bp for the human gene and 394 bp for the rat.

The transfection paradigm does not always reflect the in vivo state. While 4.2 kb of the human PLP promoter are sufficient for endogenous levels of tissue-specific expression in transgenic mice (Nadon et al., 1989), preliminary results indicate that 1088 bp of the human PLP promoter are not active in transgenic mice (Gout et al., 1991) despite the considerable promoter strength of this construct in glial cell lines (Fig. 1, Table 1). Disparity between transfected cells and the gene in transfected glial cell lines are located in close proximity to the promoter, within 312 bp for the human gene and 394 bp for the rat. Aoyama, E. (1990) J. Neurosci. Res. 12, 235-250. Macklin, W., Miller, A., Mournier, P., Traub, D., Redd, B. (1985) Proc. Natl. Acad. Sci. U.S.A. 12, 253-255.

References

Amemiya, K., Devine-Beach, K., Haas, S., and Khalili, K. (1990) J. Neurosci. 10, 1171-1177.

Aoyama, A., Tamura, T.-a., Aoyama, A., and Mikoshiba, K. (1988) Gene 75, 31-38.

Bos, E., Weinmaster, G., Kuhn, R., and Lemke, G. (1989) Neuron 3, 199-212.

Brown, S., and Hudson, L. (1990) J. Neurochem. 55, 3665-3676.

Burnell, J., Bloom, F., and Milner, A. (1987) Proc. Natl. Acad. Sci. U.S.A. 84, 5659-5669.

Cai, K., Yue, J., and You, B. (1988) Science 243, 120-123.

Clark, G., and Jessen, K. R. (1990) J. Biol. Chem. 265, 1451-1453.

Farr, T. M., Simon, S., and Hirota, S. (1988) J. Neurosci. 8, 3833-3842.

Furutani, Y. (1990) J. Biol. Chem. 265, 14777-14784.

Goto, O., and Lemke, G. (1988) J. Neurosci. Res. 21, 375-385.

Huang, L. D. (1989) J. Neurosci. 9, 8233-8242.

Kawamura, K., and Sato, T. (1988) J. Biol. Chem. 263, 2061-2064.

Kawamura, K., and Sato, T. (1989) J. Biol. Chem. 264, 16363-16368.

Kawamura, K., and Sato, T. (1990) J. Biol. Chem. 266, 3519-3524.

Kawamura, K., and Sato, T. (1991) J. Biol. Chem. 266, 3519-3524.

Kawamura, K., and Sato, T. (1992) J. Biol. Chem. 267, 277-281.

Kawamura, K., and Sato, T. (1993) J. Biol. Chem. 268, 3519-3524.

Kawamura, K., and Sato, T. (1994) J. Biol. Chem. 269, 3519-3524.

Kawamura, K., and Sato, T. (1995) J. Biol. Chem. 270, 3519-3524.

Kawamura, K., and Sato, T. (1996) J. Biol. Chem. 271, 3519-3524.

Kawamura, K., and Sato, T. (1997) J. Biol. Chem. 272, 3519-3524.

Kawamura, K., and Sato, T. (1998) J. Biol. Chem. 273, 3519-3524.

Kawamura, K., and Sato, T. (1999) J. Biol. Chem. 274, 3519-3524.

Kawamura, K., and Sato, T. (2000) J. Biol. Chem. 275, 3519-3524.

Kawamura, K., and Sato, T. (2001) J. Biol. Chem. 276, 3519-3524.

Kawamura, K., and Sato, T. (2002) J. Biol. Chem. 277, 3519-3524.

Kawamura, K., and Sato, T. (2003) J. Biol. Chem. 278, 3519-3524.

Kawamura, K., and Sato, T. (2004) J. Biol. Chem. 279, 3519-3524.

Kawamura, K., and Sato, T. (2005) J. Biol. Chem. 280, 3519-3524.

Kawamura, K., and Sato, T. (2006) J. Biol. Chem. 281, 3519-3524.

Kawamura, K., and Sato, T. (2007) J. Biol. Chem. 282, 3519-3524.

Kawamura, K., and Sato, T. (2008) J. Biol. Chem. 283, 3519-3524.

Kawamura, K., and Sato, T. (2009) J. Biol. Chem. 284, 3519-3524.

Kawamura, K., and Sato, T. (2010) J. Biol. Chem. 285, 3519-3524.

Kawamura, K., and Sato, T. (2011) J. Biol. Chem. 286, 3519-3524.

Kawamura, K., and Sato, T. (2012) J. Biol. Chem. 287, 3519-3524.

Kawamura, K., and Sato, T. (2013) J. Biol. Chem. 288, 3519-3524.

Kawamura, K., and Sato, T. (2014) J. Biol. Chem. 289, 3519-3524.

Kawamura, K., and Sato, T. (2015) J. Biol. Chem. 290, 3519-3524.

Kawamura, K., and Sato, T. (2016) J. Biol. Chem. 291, 3519-3524.

Kawamura, K., and Sato, T. (2017) J. Biol. Chem. 292, 3519-3524.

Kawamura, K., and Sato, T. (2018) J. Biol. Chem. 293, 3519-3524.

Kawamura, K., and Sato, T. (2019) J. Biol. Chem. 294, 3519-3524.

Kawamura, K., and Sato, T. (2020) J. Biol. Chem. 295, 3519-3524.

Kawamura, K., and Sato, T. (2021) J. Biol. Chem. 296, 3519-3524.

Kawamura, K., and Sato, T. (2022) J. Biol. Chem. 297, 3519-3524.

Kawamura, K., and Sato, T. (2023) J. Biol. Chem. 298, 3519-3524.

Kawamura, K., and Sato, T. (2024) J. Biol. Chem. 299, 3519-3524.

Kawamura, K., and Sato, T. (2025) J. Biol. Chem. 300, 3519-3524.

Kawamura, K., and Sato, T. (2026) J. Biol. Chem. 301, 3519-3524.