ANALYSIS OF THE STRUCTURE, EXPRESSION AND EVOLUTION OF THE SHARK MYELIN PROTEINS AND GENES

Thesis by

Lance Fors

In Partial Fulfillment of the Requirements

for the Degree of Doctor of Philosophy

California Institute of Technology

Pasadena, California

1990

(Submitted December 14, 1989)

© 1990

Lance Fors

All rights reserved

To my mother, father, grandmother and
grandfather for their love, support and encouragement

ACKNOWLEDGEMENTS

I would first like to thank my advisor Lee Hood for being a role model and for providing a nearly ideal scientific environment and my committee members, Drs. Scott Emr, Mark Tanouye, Ellen Rothenberg and Barbara Wold. Next, I want to acknowledge my collaborators Raul Saavedra, Ruedi Aebersold, Bernhard Arden and Suzzanna Horvath with special thanks to Raul who is a great human being and a pleasure to work with. Third, I want to thank my friends, Paul Mueller, Bernard Vernoij, Terresa Geffroy, Carol Readhead, Carmie Puckett, Chris Gomez, Joan Kobori and Jane Sanders who are still at Caltech. Fourth, to all of the people who have gone on to bigger and better things, Lloyd Smith, Doug Fisher, Rick Barth, Mitch Kronenberg, Roger Perlmeter, David Woo, Henry Sun, Paul Tempst, Steve Kent and Steve Hunt who made it so much fun to play the game. Fifth, I am very grateful to the staff, Bertha, Ria and Jessie for all their help.

Lastly, I want to thank those closest to me, my parents and grandparents whose love, support and assistance were essential, and Shari who has and will continue to make it all worthwhile.

ABSTRACT

Myelin is a compacted multilamellar membrane which encases axons to provide electrical insulation and facilitates the rapid transmission of nerve impulses. The major myelin structural proteins produced by oligodendrocyctes in the central nervous system (CNS) of mammals are proteolipid protein (PLP) and myelin basic protein (MBP). In contrast, the major myelin structural proteins produced by the Schwann cells in the peripheral nervous system (PNS) of mammals are protein zero (Po) and MBP. Sharks (class Chondrichthyes) are the oldest living vertebrates that have a concentric multilamellar "mammalian-like" myelin structure around axons. In addition to this structural similarity, the shark and mammalian myelin proteins appeared to be distantly related biochemically and immunologically even though they diverged from each other about 400 million years ago. Logically those regions in the shared proteins, genes and promoters which are most similar between sharks and mammals are likely to be functionally important to both. Therefore by analyzing these elements in shark myelin and comparing them to what is already known about mammalian myelin we could learn about shark myelin, its evolution and what regions are essential for the proper function and expression of mammalian myelin. This thesis contains an analysis of the structure, expression and evolution of the shark myelin proteins and genes.

The first chapter (Saavedra, R., Fors, L., Aebersold, R., Arden, B., Horvath, S., Sanders, J., and Hood, L. J. **Mol. Evol. 29**:149) describes the isolation and sequencing of the two major shark CNS proteins Po and MBP and their corresponding cDNAs. This study shows that the myelin proteins of the shark brain are similar to the myelin proteins of the mammalian peripheral nervous system in both primary and secondary structures.

The second chapter (Fors, L., Saavedra, R., and Hood, L. **Nuc. Acids Res., Submitted**) contains a novel genomic walking technique that was developed to clone the shark Po and MBP promoters. Using this technique it was possible to clone approximately 400 nucleotides immediately upstream of the shark Po and MBP transcription initiation sites. This genomic walking technique will be generally useful for cloning promoters or other sequences of interest without the need for constructing or screening genomic libraries.

The third chapter presents and discusses the similarity between these shark Po and MBP promoters, the JC virus enhancer (which directs tissue-specific expression in oligodendrocytes), and the mouse Po and MBP promoters. The implications of these findings on nervous system specific and CNS vs. PNS specific gene expression are discussed.

Lastly, the appendix describes the current status of Shiverer transgenic mouse experiments in which constructs bearing the shark MBP gene are injected into mouse eggs. These transgenic experiments are testing if the structural

similarity between shark and mammalian MBPs translates into any measurable functional similarity in vivo.

TABLE OF CONTENTS

INTRODUCTION

Myelin is a multilamellar membranous structure arising from the spiral folding of oligodendrocytes in the central nervous system (CNS) or Schwann cells in the peripheral nervous system (PNS) around axons. Myelin is unique to vertebrates above the cyclostomates (Tai and Smith 1984). Phylogenetically, it is in sharks that compact multilamellar myelin first appears (Bakey and Lee 1966). This appearance of myelin in sharks was an essential evolutionary step in the development of the vertebrate nervous system since it facilitated the rapid transmission of electric signals over long distances (Waehneld et al. 1985). In this introduction I will review the current state of knowledge of mammalian myelin and its constituent proteins and genes, summarize what was known about shark myelin prior to my work, and discuss how this thesis work on shark myelin could provide valuable information about the structure, function, and evolution of the nervous system.

MYELIN

Myelin is one of the obvious specific structural features of vertebrate nervous tissue making up about 25 to 30% of the total dry weight. Its sheath is a greatly extended and modified plasma membrane which is wrapped around the nerve in a spiral fashion (Raine 1984). The myelin membranes originate from a single membrane process sent out by Schwann cells in the PNS or multiple processes sent out by oligodendroglial cells in the CNS (Figure 1A).

Both of these myelin generating cells produces myelin for only one segment of any given axon. The periodic interruptions where short portions of the axon are left uncovered, the nodes of Ranvier, are critical to the functioning of myelin (Ritchie 1984).

Function. Myelin is an excellent electrical insulator. In myelinated axons, the excitable axonal membrane is exposed to the extracellular space only at the node of Ranvier. The remainder is covered by the myelin sheath which has a much higher resistance and much lower capacitance than the axonal membrane. The sheath's high resistance and low capacitance allows the wave of depolarization to jump from node (saltatory conduction) to node with much greater speed and using much less energy. An example of the advantages of myelination is obtained by a comparison of two different nerve fibers both of which conduct at 25 m/sec. at 20^{o} C. The 500 um diameter unmyelinated giant axon of the squid requires 5,000 times more energy and occupies about 1,500 times more space than a 12 um diameter myelinated nerve in frog (Ritchie 1984). Thus, myelin facilitates conduction while conserving space and energy.

Ultrastructure. Electron microscope and x-ray diffraction studies have shown that myelin has a characteristic ultrastructure (Kirschner et al. 1984). These techniques visualize myelin as a series of alternating dark and less dark lines (protein layers) separated by

unstained zones (the lipid hydrocarbon chains). The less dark, or intraperiod, line represents the closely apposed outer protein coats of the original cell membrane. The dark, or major period, line is the fused, inner protein coats of the cell membrane (Figure 1B). In the CNS the minimum observed diameter of myelinated fibers is about 0.3 um, close to the theoretically critical minimum size to make myelination useful (Ritchie 1984). In the PNS, the minimum observed diameter is about 1 um (due possibly to the longer internodal distances created during Schwann cells myelination)(Hirano and Dembitzer 1967).

Formation. The myelination process involves an extension of the cell membrane. In PNS myelination the large axons (greater than 1 um) are enclosed singly, one cell per axon per internode. The Schwann cells line up along the axon with intervals between them. These intervals become the nodes of Ranvier. The Schwann cell membrane then surrounds the axon and interacts to form a double membrane structure that communicates with the axon's cell surface (Geren 1954). This structure, the mesaxon, then elongates in a spiral "jelly roll" fashion and condenses into a compact myelin sheath with the cytoplasmic surfaces forming the major dense line and the external surfaces forming the intraperiod line. Myelination in the CNS is similar to the PNS except that CNS nerve fibers are not separated by connective tissue, nor are they surrounded by cell cytoplasm, and specific glial cell nuclei are not associated

with particular myelinated fibers. In addition, one glial cell can myelinate many (40 or more) separate axons whereas a Schwann cell can myelinate only one (Bunge et al. 1962). Our understanding of the process of myelin deposition is still poor. Even though a single axon can have up to 100 myelin layers all we know for sure is that the membrane system is able to expand and contract, and that the layers slip over each other.

Ontogenetic Development. The process of myelination has been shown to follow the order of phylogenetic development. During nervous system development portions of the PNS myelinate first, then the spinal cord and lastly the brain. Even within the brain different regions myelinate at different rates. Throughout the nervous system there are many small fibers that never myelinate. Generally speaking, the pathways in the nervous system become myelinated before they become completely functional. Although myelination may or may not be the terminal step in preparing the nervous system for function it is certainly an essential step (Raine 1984).

Composition. Myelin has a characteristic compostion. Myelin in situ has a water content of about 40%. Both CNS and PNS myelin are characterized by a high proportion of lipids in their dry mass (70-85%) and a low proportion of protein (15-30%) (Norton and Cammer 1984). Even though there are no "myelin-specific" lipids, cerebroside is most typical of myelin (Miller et al. 1977). Its concentration

in the nervous system during development is directly proportional to the amount of myelin present. In addition to cerebroside, the major myelin lipids are ethanolamine containing plasmalogens, lecithin, and cholesterol which is the dominate lipid in myelin. Interestingly, the lipid composition of brain myelin from all species studied to date is nearly equivalent.

THE CNS MYELIN PROTEINS

The protein composition of CNS myelin (Table 1) is relatively simple compared to that of other biological membranes with proteolipid protein (PLP) and myelin basic protein (MBP) making up 60 to 80 percent of the total in most vertebrate species (Benjamins et al. 1984). These proteins are isolated from whole brain or spinal cord by homogenization of the tissue, density banding of the myelin in sucrose, osmotic shock to release impurities, delipidation in acetone and either chloroform-methanol extraction to yield PLP or acid extraction for MBP. Several other proteins and glycoproteins are also present to lesser extents. Sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) of myelin reveals these other "high molecular weight" proteins which include the myelin-associated glycoprotein (MAG) and several enzymes.

Proteolipid Protein. Myelin PLP is the major component of the organic solvent soluble portion of whole brain. Proteolipid protein, originally named by Folch and Lees (Folch and Lees 1951), has been the subject of much study

because of its solubility in chloroform even after nearly all of its bound lipids have been removed. Protein sequence analysis of PLP shows that it has a molecular weight of around 30,000. The protein has 60% nonpolar and 40% polar amino acids and is very hydrophobic. Interestingly, PLP is 99% conserved between bovine, rat and human demonstrating an essential need to preserve its primary sequence. PLP contains around 3 moles of fatty acids per mole of protein in ester linkages to hydroxy amino acids. Although the biological role of the covalently bound fatty acids is unknown, this ability to increase the hydrophobicity of PLP could be important for membrane condensation and organization, and myelin stability. Protein sequence analysis and structural data show PLP to be a transmembrane protein with several external domains, three membrane spanning domains, and one internal domain plus the C-Terminal tail (Figure 1C). The insertion of PLP into the membrane is another area of interest because no classic signal sequence is identifiable that could direct the insertion of the 3 internal hydrophobic membrane spanning regions. In addition to PLP there is a minor protein species found in the organic phase designated DM-20 that reacts strongly with anti-PLP antibodies and originates by an alternative splicing of the RNA which gives rise to the major PLP. Sequencing data shows that the structure of DM-20 is related to that of PLP by a deletion of 40 amino acids that deletes the only significant stretch of basic residues

in the PLP extracellular domain (Nave et al. 1987a). The functional significance of the DM-20 varient is unknown. PLP has also been shown to have a possible ionophoric function suggested by its ability to transport ions across artificial bilayers (Ting-Beal et al. 1979). This observation indicates that proton movement may occur across the myelin lamallae which is consistent with the current concept of dynamic processes associated with myelin.

Myelin Basic Protein. MBPs comprise approximately 30 to 40% of the protein content of CNS myelin (Lees and Brostoff 1984). They comprise a family of basic proteins of different molecular weights (21.4 kD, 20.1 kD, 18.4 kD, 17.2 kD, 17.0 kD, and 14.1 kD) but with similar amino acid sequences with the 18.4 kD form being the most abundant in most species (Lemke 1988; Table 1). This major form is 169 amino acids long (in bovine) and has an unusually high percentage (24%) of basic residues distributed apparently randomly along the polypeptide chain, which produces an isoelectric point (pI) greater than 10. MBP has long been studied with great interest because it is the antigen that, when injected into an animal, elicits EAE (experimental allergic encephalomyelitis), a CNS autoimmune disease (Brostoff 1984). Also, MBP is a partially unfolded protein with little tertiary structure in solution (Kurscgber et al. 1984; Figure 1C). It can be extracted from myelin or brain with dilute acid and salt solutions and with 52% polar and

48% nonpolar amino acids and, once extracted, is very soluble in water.

MBP is found on the cytoplasmic surface of the myelin membrane. Whether MBP has a functional as well as a structural role is currently unknown. However, both methylation of arginine and phosphorylation of serine and threonine residues are post-translational modifications that could affect MBP function. Although the ratio of unmethylated to mono- to dimethylarginine varies from species to species, the sequence of amino acids around the methylated arginine is similar in all mammalian species as well as in the chicken basic protein (Small and Carnegie 1982). Rapid turnover of phosphate groups present in about a quarter of MBP molecules has also been detailed and might influence the close apposition of cytoplasmic faces of the membrane (DesJardins and Morell 1982).

Myelin-Associated Glycoprotein. Another well studied high molecular weight myelin protein is the 100 kD myelin-associated glycoprotein (MAG) (Steinberger et al. 1979; Table 1). MAG is an integral membrane glycoprotein with approximately one-third of its weight resulting from N-linked oligosaccharides. It accounts for about 1% of the total myelin protein and is found in myelin from all regions of the nervous system, and in all species studied. Biochemical and immunocytochemical studies have demonstrated that MAG is concentrated in the inner periaxonal membrane of myelin sheaths (i.e., it is absent from the outer myelin

loops and the bulk of compact myelin), resulting in much higher MAG concentrations in the area that interacts with the axon. Sequence analysis of MAG cDNA clones indicate that the predicted protein sequence has a single membrane spanning domain separating its C-terminal cytoplasmic tail from its heavily glycosylated N-terminus. This N-terminus has five domains related in sequence to each other, and to immunoglobulin-like molecules (Arquint et al. 1986; Figure 1C). Therefore, MAG seems to be a member of the immunoglobulin superfamily that includes neural cell adhesion molecule (N-CAM), another nervous system molecule involved in cell-cell interaction (Hunkapiller and Hood 1989.) A further relationship between MAG, N-CAM and the immune system was demonstrated by the antibody HNK-1 (anti-leu 7) which binds a common carbohydrate determinant on MAG, N-CAM, other adhesion molecules of the nervous system, and a subset of human lymphocytes (McGarry et al. 1983).

Other CNS Proteins. There are several other high molecular weight proteins present in CNS myelin in addition to PLP, MBP and MAG. The amounts vary from 10 to 30% of the total myelin proteins depending upon the species and the degree of maturity of the animal (younger animals having more high molecular weight proteins). A few of these proteins have been characterized (Lees and Brastoff 1984). Two of these protein bands at 44 and 50 kD referred to as W1 and W2, respectively, (named after F. Wolfgram) have 2′:3′ - cyclic nucleotide 3′ phosphodiesterase (CNP) enzymatic

activity. The CNPs (or Wolfgram proteins) constitute 2% of the myelin protein and are greatly enriched in myelin and in cells committed to form myelin compared with other types of cell membrane surfaces. The physiological significance of such a high concentration of CNP activity in myelin is unknown (Sprinkle et al. 1985). CNP also has a high isoelectric point. Thus, many of the major proteins of myelin, namely, MBP, PLP and CNP all have high isoelectric points and thus contain charged groups ready to interact with acidic lipids at physiological pH.

PNS MYELIN PROTEINS

The protein composition of PNS myelin is quite different from that of the CNS (Smith 1983). MAG is present at approximately the same concentration (1%). MBP and CNP are present in much lower amounts and PLP, although expressed in myelinating Schwann cells, is not incorporated into peripheral myelin sheaths (Puckett et al. 1987). PLP appears to be replaced in PNS myelin by protein zero (Po).

Protein Zero. The major protein of PNS myelin is Po, a 30.0 kD glycoprotein that accounts for over 50% of PNS myelin protein (Table 1). Po is a transmembrane protein with a highly basic cytoplasmic domain, a single transmembrane domain, and an extracellular domain which is thought to interact homo-typically (with itself) during myelin compaction (Lemke and Axel 1985; Figure 1C). This hypothesis is supported by the observation that the sequence of Po, like that of MAG (previously discussed), is related

to several members of the immunoglobilin superfamily, including N-CAM which are known to participate in homotypic interactions (Edelman 1986). The lower content of MBP in the PNS varying from about 5 to 15% of the total protein contrasts with the 30 to 40% seen in the CNS. It is logical to speculate that MBP may not play as critical a role in myelin structure in the PNS as it does in the CNS. This hypothesis is born out by the mouse mutant, shiverer, which has a mutation effecting the MBP gene and is unable to synthesize CNS myelin. In shiverer mice, that small amount of CNS myelin which is present has no dense line whereas the PNS has nearly normal myelin amounts and structure while lacking MBP (Chernoff 1981).

P2 Protein. The P2 protein is an additional basic protein found predominantly in PNS myelin (Sedzik et al. 1988). Sequence data show it to be a 131 amino acid protein which is difficult to separate electrophoretically from the 14 kD form of MBP (Table 1). The amount of P2 is quite variable from species to species comprising about 15% of the total PNS protein in bovine myelin, 5% in humans, and less than 1% in rodents and it has recently been shown to be expressed in small amounts in the CNS myelin sheaths of some species (Sedzik et al. 1988). It appears to be localized to the cytoplasmic side of myelin where it seems to play a role similar to MBP. There is substantial interest in P2 because it is the antigen for experimental allergic neuritis (EAN), the PNS counterpart of EAE (Brostoff 1984).

In addition to Po, MAG and the basic proteins MBP and P2 there is at least one other protein in PNS myelin. This additional protein is a glycoprotein of about 170 KD that is specific for PNS myelin. This glycoprotein accounts for about 5% of the total myelin protein and appears to be evenly distributed throughout the myelin sheath.

MYELIN GENES

The Proteolipid Protein Gene. Transcription of the single rat PLP gene begins postnatally and coincides with the onset of myelination. Transcription initiates at one of three possible sites within 75 bp of each other and polyadenylation can also occur at any of three possible sites to produce two abundant families of mRNAs around 3200 and 1600 nucleotides and a less abundant family around 2400 nucleotides. It is still unclear whether there are any functional differences in the protein produced from these three families of mRNA (Milner 1985).

Another form of PLP called DM20 is translated from an alternately spliced mRNA which lacks sequences as a result of the unusual utilization of a secondary splice site within the third exon of the PLP gene (Nave et al. 1987b). Thus, both MBP and PLP use alternative splicing to produce internal variants with altered coding capacities. Although no functional purpose for the DM20 variant is known, it has been shown that the internal splice in this variant removes the only significant stretch of basic residues in the PLP extracellular domain (Lemke 1988).

It was suspected that PLP might be the product of the X-linked jimpy locus in the mouse, another mutation affecting CNS myelination (Sidman et al. 1964). Mice hemizygous for jimpy exhibit severe CNS hypomyelination along with greatly reduced levels of PLP and DM20 in developing brains. The clinical symptoms are similar to those of shiverer, although possibly more acute. Structurally, jimpy mice have almost no CNS myelin and the small amount which is present has intraperiod line irregularities.

Recently, a defect in the PLP gene was shown to be the basis of the jimpy mutation (Nave et al. 1986). The PLP mRNA in jimpy mice carries a 74 nucleotide deletion that results in a reading frame shift generating a PLP protein with an altered C-terminus. This deletion has been shown to result from a point mutation that inactivates a splice acceptor site (AG to GG) before the fifth exon of the PLP gene (Nave et al. 1987b). The alteration of this single base in the jimpy mutation causes the splicing apparatus to skip over this exon, which alters the reading frame and produces a premature stop codon.

The Myelin Basic Protein Gene. The single 32 kb MBP gene encodes a family of at least six closely related proteins in mice (Lemke 1988). Transcription of this gene begins during the first week of postnatal development in mice and rats and produces the MBP family of proteins by a complex pattern of alternative RNA splicing events during

which the genes seven exons are mixed and matched (Newman 1987). To date, six different MBP mRNAs have been characterized: one in which all seven exons are present, one each in which either exon II, V, or VI is excluded, and one in which both exon 11 and either exon V or Vl are excluded (Lemke 1988). Thus, these alternatively spliced mRNAs encode MBPs which differ from each other by one or more internal deletions. Previous studies (Carson et al. 1983) have shown that the larger forms of MBP with exon 11 (17.1 and 21.4 kD) are abundant early in myelination and rare later on. This suggests that these forms may play an important role in initial myelin wrapping and provides our first insight into the possible functional significance of this complex RNA splicing pattern.

As discussed earlier, shiverer mice have low MBP levels, equivalent to less than 1% of wild-type levels in both the PNS and CNS. This autosomal recessive mutation maps to chromosome 18 in the mouse. It produces a generalized action tremor initially, followed by increasingly violent frequent convulsions and results in premature death between 50 and 100 days after birth (Brostoff 1984). Compacted myelin is largely absent from the CNS in homozygous shiverer mutants. Occasionally some abnormally formed membranes are observed which have a tightly compacted intraperiod line, but are uncompacted at the major dense line (Kirschner and Ganser 1980) thus providing a clue to the function of MBP. No MBP mRNA is

detectable in homozygous shiverers (Roach et al. 1985). Subsequent detailed mapping experiments on the MBP gene revealed that about 20 kb of this gene is deleted from the mutant genome (Roach et al. 1985). Transgenic experiments in which the cloned wild-type MBP gene was reintroduced into the shiverer genome and effectively restored myelination provided the ultimate proof that the severe phenotypic consequences of this mutation resulted from the inability to normally express MBP (Readhead et al. 1987).

Another mutant affecting MBP, called myelin-deficient (mld), was recently shown to involve a second alteration of MBP gene structure (Popka et al. 1987). In mld mice the MBP gene has been duplicated resulting in two tandem MBP genes, the first of which contains an extensive inversion of its 3′ end. Steady state levels of MBP mRNA are only 2% of wild-type in mld mutants for reasons which are still unclear. As mentioned earlier, one interesting feature of the shiverer mutation is the fact that it renders the CNS non-functional while only subtly altering peripheral myelin. This suggests that some protein found in the PNS and not the CNS might be capable of substituting for MBP in its absence. This component is likely to be the cytoplasmic domain of Po (Lemke and Axel 1985).

The Myelin-Associated Glycoprotein Gene. MAG transcription begins postnatally in the rat brain at a GC-rich promoter from seven sites within 70 nucleotides (Sutcliffe 1987). MAG, like MBP and PLP, has a lower

molecular weight form. This variant, derived by alternate mRNA splicing has been termed "adult" mRNA because it is preferentially expressed late in neural development (Lai et al. 1987). Conversely, the larger variant which contains additional N-terminus protein sequence and an altered C-terminus was termed "early" mRNA. It can be detected at the very earliest stages of myelination, before any of the major myelin proteins are detectable (Trapp et al. 1981). The functional significance of the "early" verses "late" mRNA expression is unknown. The fact that MAG is the earliest detectable protein during myelination combined with its localization exclusively at the periaxonal margin of the myelin membrane immediately adjacent to the axon suggest that it may play a role in mediating axon-glial recognition events that proceed myelination and specify the initial path of myelin deposition (Lemke 1988).

The Protein Zero Gene. The Po gene is expressed exclusively in myelinating Schwann cells and not in CNS myelin or the nonmyelinating Schwann cells that are present in the peripheral nerves (Brockes et. al. 1980). Transcription of the single Po gene increases rapidly in the rat sciatic nerve after birth producing an approximately 2,000 nucleotide mRNA (Lemke and Axel 1985). The observed differential expression of Po between the CNS and the PNS was used by Lemke to clone the Po cDNA. Nucleotide sequence analysis of this clone showed that Po is synthesized as a precursor with a N-terminal signal sequence, which is

cleaved off, followed by a relatively hydrophobic and glycosylated extracellular domain, a single transmembrane domain, and a highly basic intracellular domain.

Shark Myelin. Sharks (class chondrichthyes) are the oldest living vertebrates that have myelin organized as a concentric multilamellar membrane around axons (Bullock et al. 1984). Modern sharks diverged from other vertebrate class about 400 million years ago (Romer 1970). However, the shark myelination process appears to be similar to that of higher vertebrates (Kemali and Miralto 1983). Previous biochemical and immunological studies suggested that shark myelin may have proteins similar to mammalian Po and MBP (Wacheldt et al. 1985). The fact that shark and mammalian myelin are similar ultrastructually, biochemically and immunologically yet are separated by 400 million years, means that those regions in the shared proteins, genes and promoters which are most similar between shark and mammals are likely to be functionally important. Therefore, by analyzing these elements in shark myelin and comparing them to what is already known about mammalian myelin we could learn not only about shark myelin and evolution, but also about those regions which are essential for the proper function and expression of mammalian myelin.

Thesis. To investigate the expression, structure and evolution of the shark myelin genes, I have isolated and sequenced the major shark CNS proteins, MBP and Po, their corresponding cDNAs and promoters, and performed

phylogenetic comparison studies using this data (Chapters 1, 2 and 3). In addition, a new polymerase chain reaction (PCR) application was devised for genomic walking and was used to clone the previously uncloneable shark MBP and Po promoters (Chapter 2). Finally, an in vivo evolutionary analysis was initiated to study the effects of shark MBP on the Shiverer phenotype by constructing Shiverer transgenic mice bearing the shark MBP gene (Appendix I).

The first chapter describes the molecular characterization of the major shark CNS proteins, MBP and Po, their corresponding genes, and their evolutionary relationship to mammalian myelin. This work was initiated to determine which features of myelin structure have remained constant and which have become specifically altered as myelin composition changed during evolutionary development. In the analysis of the data, it was found that the myelin proteins of the shark CNS are similar to the myelin proteins of the mammalian PNS and that MBP and Po in sharks and mammals are similar in both primary and secondary structures.

The second chapter describes the cloning of the shark Po and MBP promoters using a genomic walking technique based on the polymerase chain reaction. After multiple unsuccessful attempts to clone these promoters by conventional techniques a novel application of the ligation mediated single-sided PCR method was developed. This genomic walking technique will be generally useful for

cloning promoters, insertion sites, or other sequences of interest without the need for constructing and sequencing genomic libraries. Using this method approximately 400 nucleotides of shark genomic DNA sequence immediately upstream from the 5′ end of our full-length MBP and Po cDNAs was cloned. These promoters are directly linked to their first exons, contain transcription initiation sites and sequences commonly found in eukaryotic promoters. In addition, similarity was discovered between the shark and mouse MBP promoters.

Chapter 3 presents and discusses the similarity between the shark MBP and Po promoters, the JC virus enhancer and the mouse MBP and Po promoters. The shark and mouse MBP promoters share several highly conserved regions that could effect their tissue-specific gene expression. One of these regions is shown here to be similar to the JC virus enhancer which directs its tissue-specific expression in mammalian oligodendrocytes. A concensus JC virus enhancer-like sequence is created by comparing the viral enhancer with its corresponding extremely conserved sequences in the mouse and shark MBP promoters. This concensus sequence is then located on the non-coding strand of shark Po. The implications of these findings on nervous system specific and CNS vs. PNS specific gene expression are discussed. In addition, several important issues raised in Chapters 1 and 2 are discussed in further detail.

Finally, Appendix 1 describes the current status of the constructs of Shiverer transgenic mice bearing the shark MBP gene. The transgenic construct contains the mammalian MBP promoter driving transcription of the shark MBP cDNA and a 3' fragment containing the prerequisite mammalian splicing site and polyadenylation site to ensure proper processing. These transgenics will test whether the striking similarity in predicted secondary structure between the shark and mammalian MBPs translates into any measurable functional similarity.

In this thesis, I have attempted to initiate studies on shark myelin, with particular emphasis on the expression, structure and evolution of the shark MBP and Po proteins, genes and promoters upon which very few studies have been performed. I expect that the work presented in this thesis will provide the motivation for additional experiments on shark myelin and stimulate additional phylogenetic studies aimed at unwrapping the myelin genes and their tissue-specific gene expression.

References

Arquint, Roder, J., Chai, L., Down, J., Wilkinson, D., Bayley, J., Brown, P., and Dunn, R. Molecular cloning and primary structure of myelin - associated glycoprotein. **Proc. Natl. Acad. Sci. USA 83:**1-5, 1986.

Barkey, L., and Lee, J. Ultrastructural changes in the edematous central nervous system. **Arch. Neurol. 14:**644 -660, 1966.

Benjamins, J., Morell, P., Hartman, B.,and Argrawal, H. CNS Myelin. In A. Lajtha (ed.), Handbook of Neurochemistry Plenum, New York, 1984, pp. 361-415.

Brockes, J., Raff, M., Nishiguchi, D., and Winter J. Studies on cultured rat Schwann cells. 111 Assays for peripheral myelin proteins. **J. Neurocytol. 9:**67-77, 1980.

Brostoff, S. Immunological Responses to Myelin. In P. Morell (ed.) Myelin-2nd ed., Plenum, New York, 1984, pp. 405-430.

Bullock, T., Moore, J., and Fields, D. Evolution of myelin sheaths: Both lamprey and hagfish lack myelin. **Neurosci. Letters 48:**145-148, 1984.

Bunge, M., Bunge, R., and Pappas, G. Electron microscopic demonstrations of connections between glial and myelin sheats in the developing mammalian central nervous system. **J. Cell Biol. 12:**448-457, 1962.

Carson, J., Nielson, M., and Barlarese, E. Developmental regulation of myelin basic protein expression in mouse brain. **Dev. Biol. 96:**485-492, 1983.

Chernoff, G. Shiverer: an autosomal recessive mutant mouse with myelin deficiency. **J. Hered. 72:**128-132, 1981.

DesJardins, K., and Morell, P. The phosphate groups modifying myelin basic proteins are metobolically labile; the methyl groups are stable. **J. Cell Biol. 97 :**438-446, 1982.

Edelman, G. Cell adhesion molecules in the regulation of animal form and tissue pattern **Ann. Rev. Cell Biol. 2:** 81-116, 1986.

Folch, J., and Lees, M. Proteolipids, a new type of tissue lipoproteins, **J. Biol. Chem. 191:**807, 1951.

Geren, B. The formation of the Schwann cell surface of myelin in the peripheral nerves of chick embryos. **Exp. Cell Res 7:**558-562, 1954.

Hirano, A., and Dembitser, H. A structural analysis of the myelin sheath in the central nervous system. **J. Cell Biol. 34:**555-567, 1967.

Hunkapiller, T., and Hood, L. Diversity of the immunoglobulin gene superfamily. **Adv. Immunology 44:**1-63, 1989.

Kemali, M., and Miralto, A. The habenular nuclei of the elosmobranch Scyllium stellare myelinated perikarya. **Z MikroskAnat Forsch, Leipzig 97:**3-14, 1983.

Kirschner, D., Ganses, A., and Casper, D. Diffraction Studies of Molecular Organization and Membrane Interaction in Myelin. Im P. Morell (ed) Myelin-2nd ed., Plenum, New York, 1984, pp. 51-96.

Kirschner, D., and Ganser, A. Compact Myelin exists in the absence of myelin basic protein in the shiverer mouse. **Nature 283:**207-210, 1980.

Lai, C., Brow, M., Nave, K., Noronha, A., Quarles, R., Bloom, F., Milner, R., and Sutcliffe, J. Two forms of 1B236/myelin-associated glycoprotein, a cell adhesion molecule for postnatel development, are produced by alternate splicing. **Proc. Natl. Acad. Sci. USA 84:**4227-4341, 1987.

Lees, M., and Brostoff, S. Proteins of Myelin. In P. Morell (ed.) Myelin - 2nd ed., Plenum, New York, 1984, pp. 197-217.

Lemke, G., and Axel, R. Isolation and sequence of a cDNA encoding the major structural protein of peripheral myelin. **Cell 40:**501-508, 1985.

Lemke, G. Unwrapping the Genes of Myelin. **Neuron 1:**535-543, 1988.

McGarry, R., Helfand, S., Quarles, R., and Roder, J. Recgnition of myelin-associated glycoprotein by the monoclonal antibody HNK-1. **Nature 306:**376-378, 1983.

Miller, S., Benjamins, J., and Morell, P. Metabolism of glycerophospholipids of myelin and microsomes in rat brain. **J. Biol. Chem. 252:**4025-4037, 1977.

Milner, R., Lai, C., Nave, K., Lenoir, D., Ogata, J., and Sutcliffe, J. Nucleotide sequence of two mRNAs for rat brain myelin proteolipid protein **Cell 42:**931-939, 1985.

Nave, K., Lai, C., Bloom, F., and Milner, R. Jimpy mutant mouse : a 74 base deletion in the mRNA for myelin proteolipid protein and evidence for a primary defect in RNA splicing. **Proc. Natl. Acad. Sci. USA 83:**9264-9268, 1986.

Nave, K., Lai, C., Bloom, F., and Milner, R. Splice site selection in the gene for proteolipid protein (PLP) and primary structure of the DM20 protein of CNS myelin. **Proc. Natl. Acad. Sci. USA 84:**5665-5669, 1987a.

Nave, K., Bloom, F., and Milner, R. A single nucleotide difference in the gene for myelin proteolipid protein defines the jimpy mutation in the mouse. **J. Neurochem 49:**1873-1877, 1987b.

Newman, S., Kitamura, K., and Campagnoni, A. Identification of a cDNA coding for a fifth form of myelin basic protein in mouse. **Proc. Natl. Acad. Sci. USA 84:**886-890, 1987.

Norton, W., and Cammer, W. Isolation and Characterization of Myelin. In P. Morell (ed.) Myelin-2nd ed., Plenum, New York, 1984, pp. 147-196.

Popka, B., Puckett, C., Lai, E., Shine, H., Readhead, C., Takohashi, N., Hunt, S., Sidman R., and Hood, L. Myelin deficient mice : expression of myelin basic protein and generation of mice with varying levels of myelin. **Cell 48:**713-721, 1987.

Puckett, C., Hudson, L., Ono, K., Friedrich, V., Benecke, J., Dubois-Dalcg, M., and Lazzarini, R. Myelin-Specific Proleolipid Protein is expressed in myelinating Schwann cells but is not incorporated into myelin sheaths. **J. Neurosci. Res. 18:**511-518, 1987.

Raine, C. Morphology of myelin and myelination. In P. Morell (ed.), Myelin - 2nd ed., Plenum, New York, 1984, pp.
1-41.

Readhead, C., Popka, B., Takahashi, N., Shine, H., Sasvedra, R., Sidman, R., and Hood, L. Expression of a myelin basic protein gene in transgenic shiverer mice: correction of the dysmyelinating phenotype. **Cell 48:** 703-712, 1987.

Ritche, J. Physiological basis of conduction in myelinated nerve fibers. In P. Morrell (ed.), Myelin - 2nd ed., Plenum, New York, 1984, pp. 117-141.

Roach, A., Takahashi, N., Provtcheva, D., Ruddle, F., and Hood, L. Chromosomal mapping of mouse myelin basic protein gene and structure and transcription of the partially deleted gene in shiverer mutant mice. **Cell 42:**149-155, 1985.

Romer, A. In: The Vertebrate Body, ed 4. Saunders Co., Philadelphia, 1970.

Sedzik, J., Bergfors, T., Aluryn Jones, T., and Weise, M. Bovine P2 myelin basic protein crystalizes in 3 different forms. **J. Neurochem. 50:**1908-1913, 1988.

Sidman, R., Dickie, M., and Appel, S. Mutant mice (quaking and jimpy) with deficient myelination in the central nervous system. **Science 144:**309-310, 1964.

Small, D., and Carnegie, P. In vivo methylation of an arginine in chick myelin basic protein. **J. Neurochem. 38:**184, 1982.

Smith, M. Peripheral nervous system myelin properties and metabolism. In A. Lajtha (ed.), Handbook of Neuro - chemistry (2nd ed.), Plenum, New York, 1983, pp. 201-223.

Sprinkle, T., McMorris, F., Yoshino, J. and De Vries, G. Differential expression of 2′:3′- cyclic nucleotide 3′- phospho-diesterase in cultured central, peripheral, and extroneural cells. **Neurochem. Res. 10:** 919-931, 1985.

Sternberger, N., Quarles, R., Itoyokma, Y., Webster, H. Myelin-associated glycoprotein demonstrated immunochemically in myelin & myelin forming cells of developing rat. **Proc. Natl. Acad. Sci. USA 76:**1510, 1979.

Sutcliffe, G. The genes for myelin. **Trends in Neurosci. 3:** 73-76, 1987.

Tai, F., and Smith, R. Comparison of the major proteins of shark myelin with the proteins of higher vertebrates. **J. Neurochemistry 42:**426-433, 1984.

Ting-Beal, H., Lees, M., and Robertson, J. Interaction of Folch-Lees proteolopid apoprotein with planar lipid bilayers. **J. Membrane Biol. 51:**33, 1979.

Tropps, B., Quarles, R., and Griffin, J. Myelin-associated glycoprotein and myelinating Schwann Cell - axon interaction in chronic B, B' - immunodipro - piontrile neuropathology. **J. Cell Biol. 98:**1272-1278, 1981.

Waehneldt, T., Malotka, J., Karvin, N., and Matthieu, J. Phylogentic examination of vertebrae central nervous system myelin proteins by electro-immunoblotting. **Neurosci. Letters 57:**97-102, 1985.

Figure 1. **The Molecular Structure of Myelin.** Figure 1A is a schematic cross section of a CNS and PNS myelinated axon. The compacted myelin sheath that is boxed is amplified in Figure 1B. Figure 1B shows the amplified myelin exhibiting the characteristic intraperiod line (extracellular surface) and major dense line (cytoplasmic surface). Figure 1C illustrates the currently prevailing model of the individual proteins within the bilayer. MAG is separated from the other proteins to indicate that it is not found within the compacted region of the sheath. Abbreviations: N, amino terminus; C, carboxy terminus; A, Fatty acid acylation region of PLP; +, concentrated region of basic amino acids; -, concentrated region of acidic amino acids. See text for explanations of additional details. Figure adapted from Lemke et al. (1988).

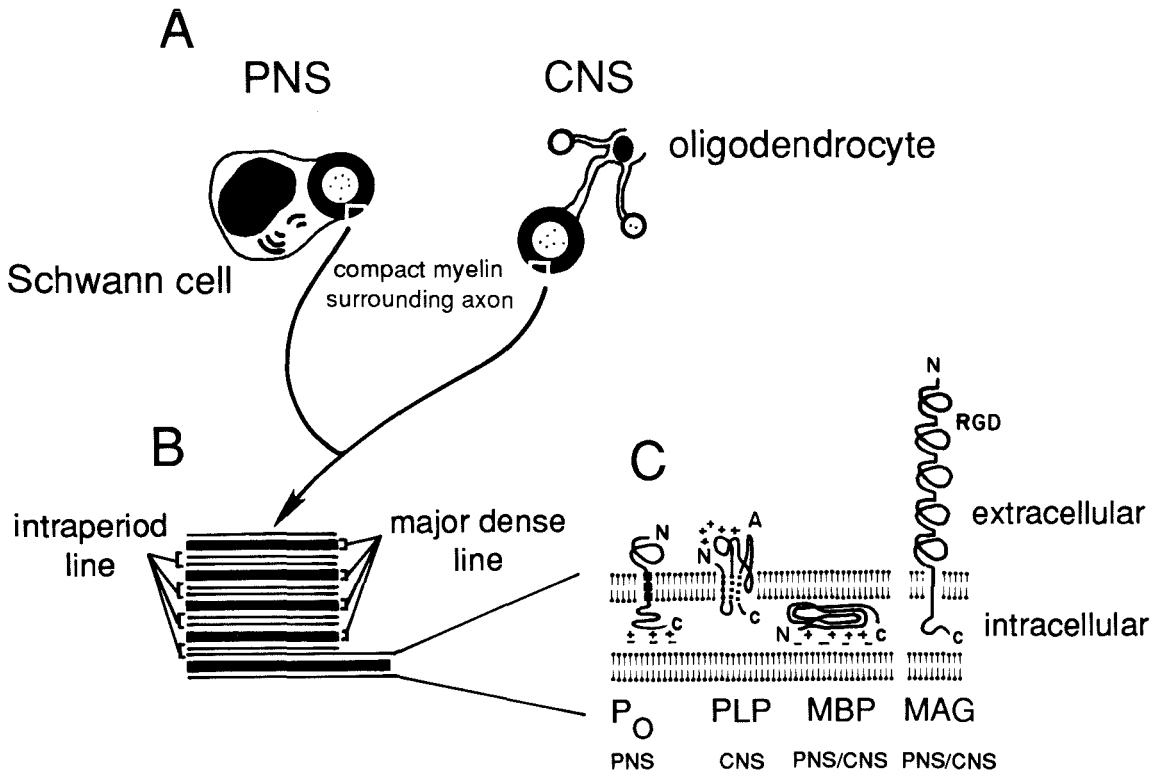

A
PNS
CNS
oligodendrocyte
Schwann cell
compact myelin
surrounding axon
B
intraperiod
line
major dense
line
C
N
RGD
extracellular
intracellular
A
N
C
N
C
C
N
P O
PLP
MBP
MAG
PNS
CNS
PNS/CNS
PNS/CNS

Table 1. Myelin Specific Proteins
(Adapted From Lemke et. al. 1988)

Myelin Protein		CNS Myelin	PNS Myelin	Alternatively Spliced Forms	Post-translational Modifications
Protein zero	(Po)	-	>50%	24.9 kD	Glycosylated and phosphorylated
Proteolipid protein	(PLP) (DM-20)	50%	-	30.0 kD 26.5 kD	Acylated
Myelin basic protein	(MBP)	30%-40%	5%-15%	21.4 kD 20.1 kD 18.4 kD 17.1 kD (two forms) 14.1 kD	Phosphoylated and methylated
Myelin-associated glycoprotein	(MAG)	1%	<1%	68.9 kD 64.0 kD	Glycosylated (HNK-1) and phosphorylated
P_2 basic protein	(P_2)	-	Wide variability between species.	14.4 kD	Amidated

Chapter 1

THE MYELIN PROTEINS OF THE SHARK BRAIN ARE SIMILAR TO THE MYELIN PROTEINS OF THE MAMMALIAN PERIPHERAL NERVOUS SYSTEM

By Raul Saavedra, Lance Fors, Reudi Aebersold*, Bernhard Arden+, Suzanna Horvath, Jane Sanders, and Leroy Hood

Division of Biology

California Institute of Technology

Pasadena, California 91125

*Biomedical Research Centre

Vancouver, Brirish Columbia

V6T IW5, Canada

+Max-Plank-Institute for Immunology

Stubewg 51, 7800 Freiburg, West Germany

Journal of Molecular Evolution **29**, pp. 149-156.
Copyright 1989 by Springer-Verlag New York, Inc.

The Myelin Proteins of the Shark Brain Are Similar to the Myelin Proteins of the Mammalian Peripheral Nervous System

Raul A. Saavedra, Lance Fors, Ruedi H. Aebersold,* Bernhard Arden,† Suzanna Horvath, Jane Sanders, and Leroy Hood

Biology Division, 147-75, California Institute of Technology, Pasadena, California 91125, USA

Summary. The two major structural proteins in the shark CNS are similar to the structural proteins, P_o and myelin basic protein (MBP), found in the mammalian peripheral nervous system (PNS). Shark P_o is 46% similar to its mammalian counterpart. The extracellular domain of shark P_o also appears to be organized as an immunoglobulin-like domain that mediates homotypic interactions. The intracellular domain of shark P_o also is very basic and may play a role in myelin condensation analogous to that of MBP. Shark MBP is 44% similar to mammalian MBP. Both MBPs show conserved interspersed regions and are present in multiple forms that arise by alternative splicing of a single transcript. These structural analyses indicate that the complexities seen in mammalian myelin arose early during vertebrate evolution.

Key words: Myelin proteins — Structure — Function — Genes — Splicing — Molecular evolution

Introduction

The evolution of the vertebrate nervous systems requires a structure that allows rapid transmission of electric signals over long distances. Myelin is the dynamic structure that provides electric insulation and thus facilitates the rapid conduction of nerve impulses along axons. In the mammalian CNS the myelinating cells are the oligodendrocytes, whereas in the peripheral nervous system (PNS) they are the Schwann cells. During development, each oligodendrocyte sends multiple processes that surround many nerve fibers, whereas each Schwann cell sends one process to surround one nerve fiber. In both cases, the vast majority of cytoplasmic and extracellular materials are subsequently extruded leaving a compact structure of concentric membranes that contains the myelin proteins.

The pattern of myelin protein expression in mammals is different in the oligodendrocytes of the CNS and the Schwann cells of the PNS. The major structural proteins in the CNS of mammals are proteolipid protein (PLP) and myelin basic protein (MBP), whereas in the PNS the major proteins are P_o and MBP (Lees and Brostoff 1984; de Ferra et al. 1985; Readhead et al. 1987). Both, PLP and P_o are integral membrane proteins with extracellular and cytoplasmic domains, whereas MBP is largely hydrophilic and is found associated with the cytoplasmic sides of the membranes (Lees and Brostoff 1984; Lemke and Axel 1985; Lemke et al. 1988). Limited information is available on the evolution of myelin proteins because these proteins from older vertebrate species have not been studied in sufficient detail.

Sharks (class Chondrichthyes) are the oldest known living vertebrates that have myelin organized as a concentric multilamellar structure around the nerve fibers (Bullock et al. 1984). Modern sharks diverged from other vertebrate classes about 400 million years ago; however, their myelination process appears to be similar to that of higher vertebrates (Bakay and Lee 1966; Romer 1970; Kemali

Offprint requests to: R.A. Saavedra
* *Present address:* Biomedical Research Centre, Vancouver, British Columbia V6T 1W5, Canada
† *Present address:* Max-Planck-Institute für Immunobiologie, Stubeweg 51, 7800 Freiburg, West Germany

and Miralto 1983). These observations make shark a system of choice to study the evolution of myelin proteins. Previous studies suggested that shark myelin may have proteins similar to P_o and MBP (Tai and Smith 1983; Waehneldt et al. 1985; Tai et al. 1986). In this article, we demonstrate that the major structural proteins in shark CNS myelin are similar to the mammalian PNS proteins P_o and MBP in both primary and secondary structure.

Materials and Methods

Isolation of Myelin Proteins. Brains and spinal cords of the shark *Heterodontus francisci* were dissected in toto, and the myelin fraction was isolated by the sucrose gradient method of Norton and Poduslo (1973) and delipidated with 90% acetone (Readhead et al. 1987). A fraction of delipidated total myelin proteins was extracted with 0.3 M HCl to further purify basic proteins. Total and basic myelin proteins were separated on a Thomas and Kornberg's (1975) 18% SDS-PAGE system and stained with Coomassie blue. Myelin proteins extracted from mouse, rat, and cow brains served as molecular weight standards (not shown). Gels were photographed using Type 55 Kodak film, and the negatives were scanned with an integrator densitometer. The proteins shown in Fig. 1 were extracted in the absence of protease inhibitors, because such inhibitors would interfere with the protein sequencing procedure. Isolation of shark myelin proteins performed in the presence of a mixture of protease inhibitors (phenylmethylsulfonyl fluoride, ethylenediamine tetraacetic acid, and Aprotinin, from Sigma Co., St. Louis, Missouri) rendered a pattern of bands indistinguishable from the one obtained in the absence of inhibitors (data not shown).

Protein Sequencing. Amino-terminal sequences were obtained from shark P_o after electrophoretic transfer of the myelin proteins from polyacrylamide gels onto chemically modified glass fiber sheets. The bands of interest were identified by staining with a coumarin derivative and sequenced using a Caltech-developed gas-phase sequenator as described by Aebersold et al. (1986, 1987). To obtain internal fragment sequences from shark P_o and MBP, the appropriate bands electroblotted onto nitrocellulose were cut and digested in situ with trypsin and/or *Staphylococcus aureus* V8 (Aebersold et al. 1987, 1988). The resulting fragments then were separated by HPLC and sequenced in the gas-phase sequenator.

Oligonucleotide Probes. The underlined protein sequences shown in Figs. 2 and 3 were reverse translated and the corresponding oligonucleotide probes were synthesized by using the phosphoramidite chemistry of Beaucage and Caruthers (1981) and Horvath et al. (1987). A 23-mer and a 26-mer for shark P_o, and a 20-mer, a 23-mer, and a 33-mer for shark MBP were synthesized. At those positions of high degeneracy inosine, was used to increase the probability of base pairing, because inosine can pair with adenine, thymine, and cytosine (Ohtsuka et al. 1985).

Construction and Screening of a cDNA Library from Shark Brain. Polyadenylated RNA from shark brain was used as a template to generate the first strand complementary DNA (cDNA) with reverse transcriptase. The second strand was synthesized using RNase H and DNA polymerase (Gubler and Hoffman 1983). The double-stranded (ds) cDNA was ligated to a 50-fold molar excess of EcoRI linkers. After digestion with the restriction enzyme EcoRI, ds cDNA molecules equal to or larger than 1 kilobase were separated on a Bio-Gel A 50-m column, and ligated into the EcoRI restriction site of lambda gt10 phage (Huynh et al. 1985). The phage was packaged in vitro and the library screened with the above oligonucleotides under hybridization conditions of increasing stringency. The initial screening was done in 5× saline sodium citrate and 0.1% sodium dodecyl sulfate (SDS), at 37°C, and the final screening was done at 45°C.

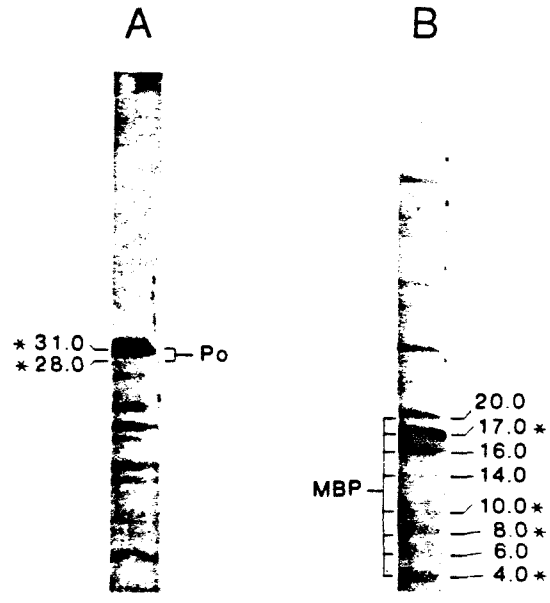


Fig. 1. Shark myelin proteins. **A** Total shark myelin proteins. **B** Shark basic myelin proteins. The asterisks indicate those bands from which protein sequences were obtained.

Sequencing of cDNA Clones. Lambda clones P_o81 and MBP137 were subcloned into M13-mp18 phage and pUC18 plasmid and sequenced by the dideoxynucleotide method of Sanger et al. (1977; SEQUENASE from U.S. Biochemical Co. Cleveland, Ohio) using synthetic oligonucleotides as primers (Beaucage and Caruthers 1981; Horvath et al. 1987). The cDNA clone P_o81 also was sequenced at selected regions using the dideoxynucleotide method (Sanger et al. 1977) in conjunction with the 370A DNA automatic sequenator (Applied Biosystems, Foster City, California; Smith et al. 1986). The sequences obtained by the manual and automated methods were in excellent agreement with each other, above 99%.

Sequence Similarities. The sequence similarities between shark and mammalian proteins were determined by computing the number of identical amino acids after manual alignment of the corresponding sequences.

Results and Discussion

Shark Myelin Contains Two Major Structural Proteins

In Fig. 1A proteins extracted from shark CNS myelin were separated electrophoretically on an 18% SDS-PAGE (Thomas and Kornberg 1975). The most prominent protein had a molecular weight (Mr) of 31,000. Below this major band migrated a minor

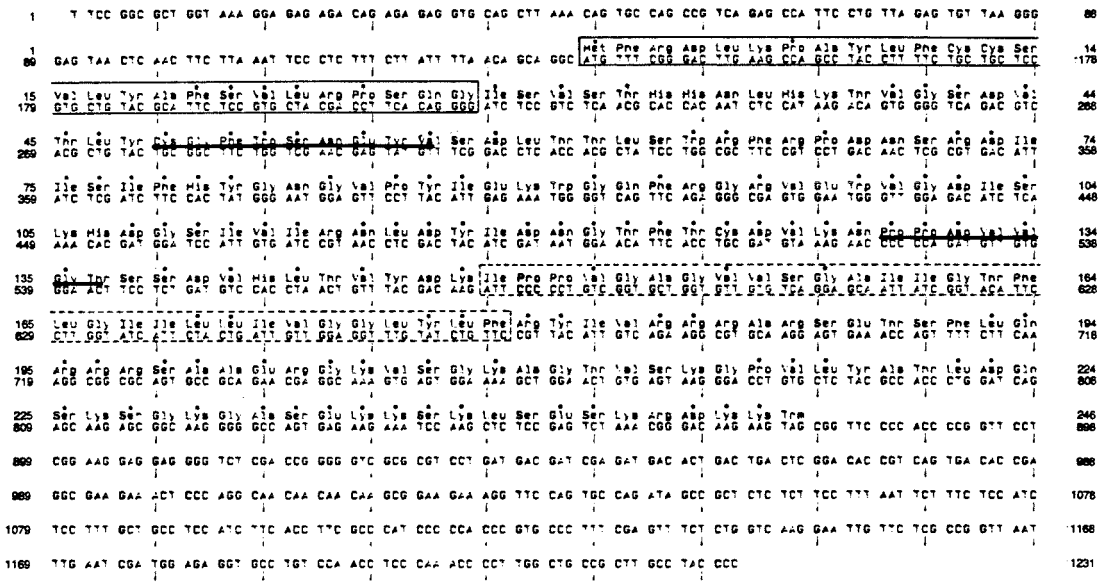


Fig. 2. Shark P_o cDNA clone 81 and translated protein sequences. The signal peptide of P_o is indicated by a solid line box and the transmembrane domain is indicated by a dashed line box. The underlined residues show the regions that were reverse translated to generate oligonucleotide probes. A dot marks those residues that are identical on the shark and rat proteins.

protein species of Mr 28,000. The N-terminal and five internal fragments from the large protein and two internal fragments from the smaller protein were sequenced. This yielded a total of 87 residues. The two internal fragments from the smaller protein were identical to two of the internal fragments from the larger protein. These shark sequences were similar to five distinct segments on the external domain of rat P_o (see Materials and Methods; Lemke and Axel 1985; Lemke et al. 1988). Therefore, we called these proteins shark P_o. The gels were scanned with an integrating densitometer, and the two bands account for about 45% of the total protein in shark CNS myelin. The difference in size between the two shark P_o could be due to differential glycosylation or proteolysis.

The shark myelin proteins shown in Fig. 1A then were extracted with 0.3 M HCl to purify the basic proteins (Norton and Poduslo 1973; Readhead et al. 1987). These basic proteins also were separated using the SDS-PAGE system described above (Thomas and Kornberg 1975). Figure 1B shows a prominent shark basic protein of Mr 17,000. Figure 1B also shows less prominent bands of Mr 20,000, 16,000, 14,000, 10,000, 8000, 6000, and 4000. These shark basic proteins were blocked at their amino terminus, therefore four internal fragments from the Mr 17,000, three from the Mr 10,000, four from the Mr 8000, and two from the Mr 4000 bands were sequenced. This yielded a total of 53 amino acids. These shark sequences were similar to five distinct segments on mouse MBP (see Materials and Methods; Lees and Brostoff 1984). In addition, all the above bands were recognized by antibodies made against the shark Mr 17,000 band (data not shown). Therefore, we called these proteins shark MBP, and together they also were shown to constitute about 45% of the total myelin proteins. The various MBP bands seen in shark may be due to differential splicing, posttranslational modifications, and/or proteolytic products.

The above observations, taken together, unequivocally demonstrate that shark CNS myelin is composed mainly of two structural proteins similar to mammalian P_o and MBP. Therefore, the myelin proteins from shark CNS are similar to those extracted from mammalian PNS. It still remains to be determined why P_o expression is restricted to mammalian PNS, and PLP is expressed in mammalian CNS. Possibly, PLP could provide lamellar interactions different from those provided by P_o, and thus provide distinct functional properties to the mammalian CNS.

Cloning the Shark P_o and MBP Genes

In order to clone the P_o and MBP genes from shark, selected sequences of both shark proteins were reverse translated and low degeneracy oligonucleotide probes were synthesized (Beaucage and Caruthers 1981; Horvath et al. 1987). Two P_o and three MBP probes (indicated on Figs. 2 and 3) were constructed with inosine inserted at appropriate degenerate sites (see Materials and Methods; Ohtsuka et al. 1985).

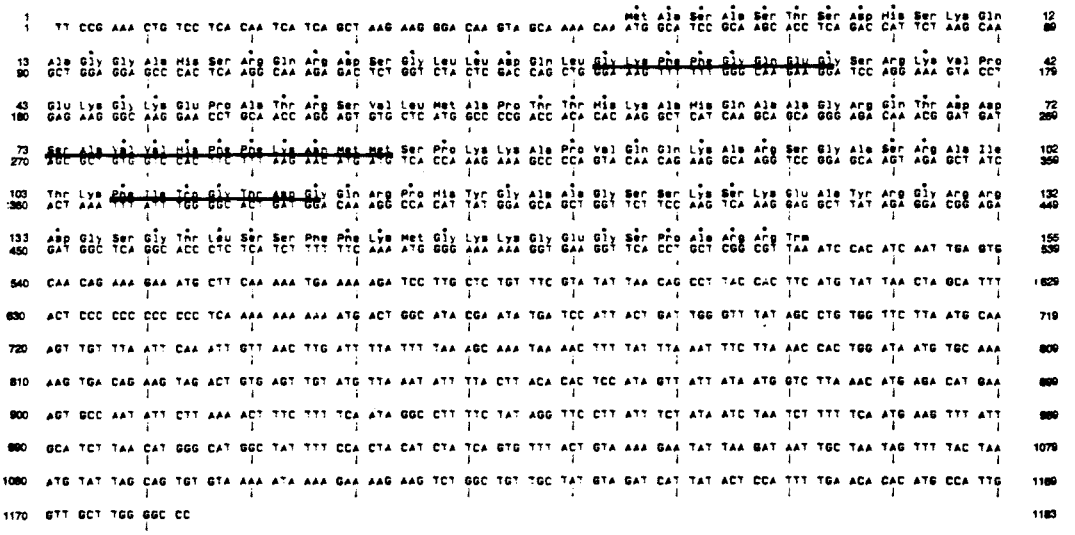


Fig. 3. Shark MBP cDNA clone 137 and protein sequences. The underlined residues show the regions that were reverse translated to generate oligonucleotide probes. A dot indicates those residues that are identical on the shark and mouse proteins.

These probes were used to screen a cDNA library made from shark brain (Gubler and Hoffman 1983; Huynh et al. 1985), and four clones for P_o and seven clones for MBP were obtained. Full-length cDNA clones of 2.8 kb for P_o and 2.2 kb for MBP were sequenced within the coding region. An additional 350 bp of the P_o clone and 660 bp of the MBP clone also were sequenced beyond the 3′ end of the respective coding regions. Moreover, 350 bp at the 3′ end of each of the two clones were also sequenced. Both cDNA clones had putative polyadenylation sites and polyadenylated tails.

The shark cDNA clone P_o81 has an open reading frame of 242 codons that translated into a protein of calculated molecular weight of 27,300 (Fig. 2). The mature shark protein (24,200), like that of rat P_o (Lemke and Axel 1985; Lemke et al. 1988), shows an extracellular, a transmembrane, and an intracellular domain (see Fig. 4A). The overall similarity between the shark and rat proteins is 46% (see Materials and Methods; Lemke and Axel 1985; Lemke et al. 1988). However, the similarities are different for the various domains: the extracellular domains are 58%, the transmembrane domains are 22%, and the intracellular domains are 55% similar. This observation suggests that the most conserved regions play essential roles in myelin compaction.

The shark cDNA clone MBP137 has an open reading frame of 155 codons that translated into a protein of calculated molecular weight of 16,600 (Fig. 3). There is a 44% sequence similarity between shark MBP and the Mr 18,500 form of mouse MBP (see Materials and Methods; de Ferra et al. 1985). The shark protein has a net charge of +16, whereas the Mr 18,500 form of mouse MBP has a net charge of +20. Shark and mouse MBPs also contain regions that are highly conserved.

The available evidence is consistent with a model of myelin where the transmembrane protein has its external domain interacting in a homotypic fashion with a protein on the adjacent membrane. Moreover, the cytoplasmic domains of the transmembrane protein and the hydrophilic protein may interact among themselves and/or with the charged lipid heads on the cytoplasmic sides of adjacent membranes (Braun 1984; Kirschner et al. 1984). All these protein–protein and protein–lipid interactions result in the compact multilamellar structure characteristic of mature myelin.

The Extracellular Domain of Shark P_o May Be Organized as an Immunoglobulin-like Domain

It has been proposed that the extracellular domain of mammalian P_o may be organized as an immunoglobulin-like domain, and therefore may interact homotypically in the condensation of myelin in a manner similar to immunoglobulins (Williams 1987; Lemke et al. 1988). An immunoglobulin domain consists of seven or eight antiparallel beta sheets (A–G) alternating with beta turns located at the end of loops of various lengths. To form the domain, four and three or four and four antiparallel beta sheets are juxtaposed to each other and are firmly held in place by a cysteine disulfide bridge and other interactions. We compared the putative secondary structure of the extracellular domains of shark and rat P_o. The Kyte–Doolittle (1982) hydrophilicity analysis and the Chou–Fasman (1978) algorithm for

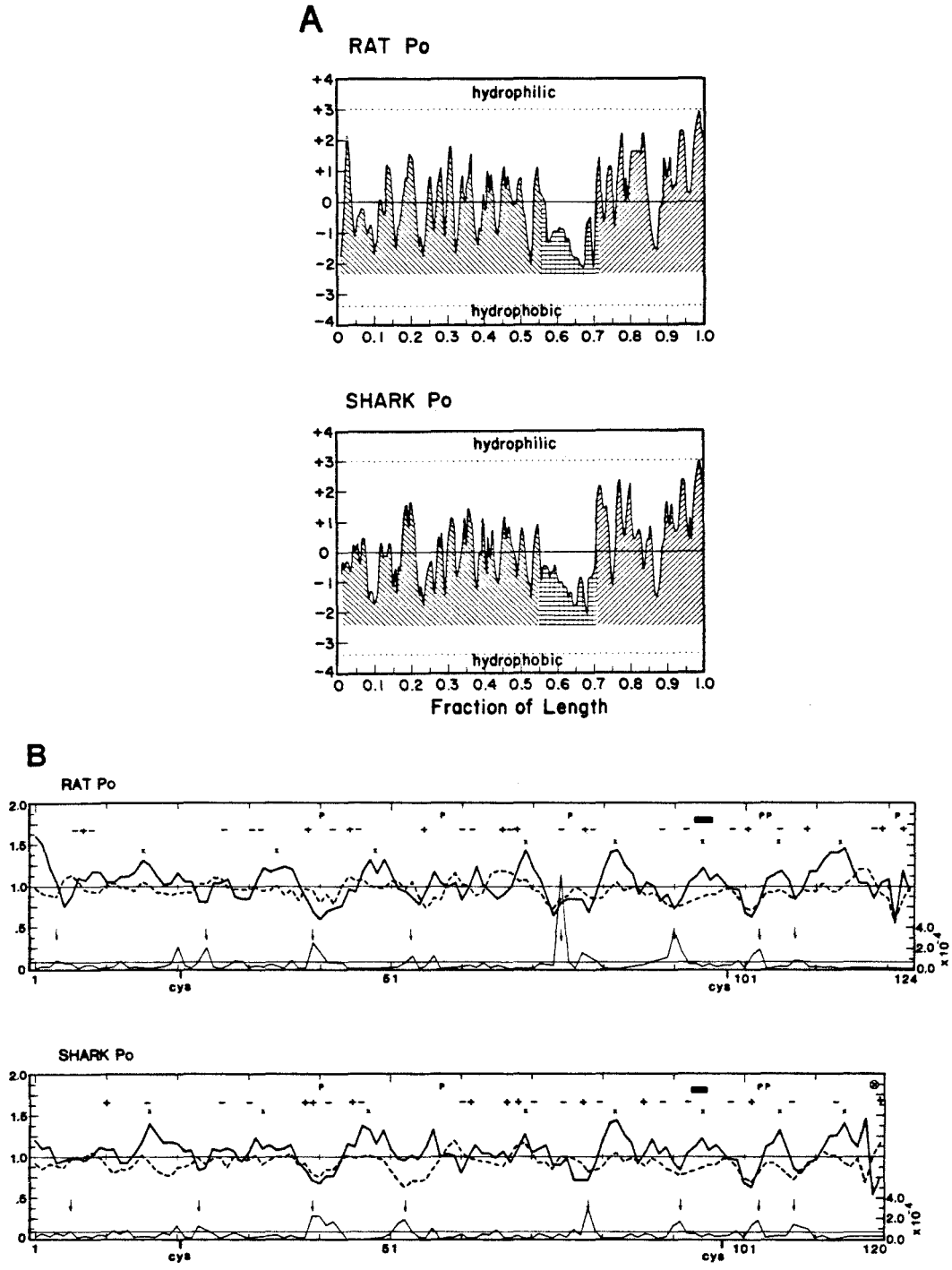


Fig. 4. Hydrophilicity and alpha helix and beta sheet structure of rat and shark P_o. **A** Kyte–Doolittle (1982) analysis of mature rat and shark P_o. Regions of similarity are marked by the same patterns. The profiles clearly show three domains on each protein, i.e., extracellular (lines inclining to the left), transmembrane (cross hatched), and intracellular (lines inclined to the right). **B** Chou–Fasman (1978) analysis for the extracellular domain of rat and shark P_o. Dashed line, alpha helix potential; solid line, beta sheet potential. Crosses indicate regions of high beta sheet potential. Arrows show regions of potential beta turns. Positive and negative signals indicate positively and negatively charged amino acids. The position of proline residues is marked by P, and the relative positions of cysteine residues thought to form disulfide bonds are indicated by cys. Solid bars indicate putative glycosylation sites.

alpha and beta structure of the two P_o extracellular domains were superimposable and consistent with the proposed structure of eight antiparallel beta sheets alternating with beta turns (see Fig. 4A and B). A feature common to all members of the immunoglobulin superfamily is the presence of conserved residues in beta strands B, C, E, and F (Williams 1987). The two cysteine residues are located in beta strands B and F, respectively, and serve as reference points for the sequence alignments. These cysteines also have the same relative positions in rat and shark P_o (Lemke and Axel 1985; Williams 1987; Lemke et al. 1988). Other sequences like aspartic–X–glycine (where X represents any amino acid) found adjacent to the beta strand F and the residue arginine found at the base of the beta strand D are common to all members of the variable set of immunoglobulin proteins (Williams 1987). These sequences and residues also are conserved in rat and shark P_o (Lemke and Axel 1985; Williams 1987; Lemke et al. 1988). The conserved nature of these residues is consistent with the notion that these proteins are homologous (see Williams 1987). These observations raise the possibility that the outer domain of shark P_o also may be organized as an immunoglobulin-like domain and interact in a homotypic fashion.

It also has been proposed that the intracellular domain of P_o could play a role similar to that of MBP, bridging the intracellular faces of two adjacent membranes by means of electrostatic interactions between these charged proteins and the phospholipid heads (Lemke and Axel 1985; Lemke et al. 1988). To see if shark P_o has the same potential, we compared the nature and distribution of charged residues along the intracellular domain of shark and rat P_o. Both molecules show a similar distribution of positively and negatively charged residues on the intracellular domain (data not shown). These structural resemblances suggest that these intracellular domains play similar roles; however, the actual mechanism remains to be determined.

The Structural Analysis of Shark MBP cDNA Clones Shows Alternative Splicing from a Single Transcript

There are five forms of MBP in mouse of Mr 21,000, 18,500, 17,200, 17,000, and 14,400, respectively (de Ferra et al. 1985; Newman et al. 1987). These forms are encoded by a single copy gene that is 32 kb long and contains seven exons. The various MBP bands seen in our shark myelin preparations, and the precedent of differential splicing for mammalian MBPs led us to sequence the coding region of the other six MBP cDNA clones. The nucleotide sequence of shark cDNA clone MBP137 included coding regions similar to mouse exons I, III, IV, V, VI, and VII, but not to exon II (see Materials and Methods; Fig. 5A). Thus, this cDNA codes for a shark MBP of calculated molecular weight 16,600, which is most similar to the Mr 18,500 form of mouse MBP (de Ferra et al. 1985). Four additional cDNA clones, MBP88, MBP90A, MBP92, and MBP129, were identical to cDNA clone MBP137 within the coding region. Meanwhile, the nucleotide sequence of shark cDNA clones MBP90B and MBP126 included coding regions similar to mouse exons I, III, IV, VI, and VII, but not to exons II or V (see Materials and Methods; Fig. 5A). These cDNA clones, therefore, code for a shark MBP of calculated molecular weight 15,000 that is most similar to the Mr 17,200 form of mouse MBP (Newman et al. 1987). Coded by the first exon of the mouse MBP gene, there is a sequence of 13 amino acids (residues 4–16) that could not be found on any of the shark cDNAs. None of our MBP cDNA clones shows similarity to exon II. This could be due to the low abundance of mRNA-bearing exon II or to the formal possibility that exon II is not present in shark. Southern blot analysis indicated that there is only one copy of the MBP gene in shark (data not shown). These facts indicate that the alternative splicing of exons in sharks is similar to that found in mammals, and that the mechanism used to generate the various forms of MBP appeared early during vertebrate evolution.

Secondary Structure Analysis of MBP

Models for MBP secondary structure have been proposed based solely on the amino acid sequence of the mammalian proteins (Martenson 1981; Stoner 1984). We compared the putative secondary structures of the mouse and shark MBPs to assess whether the shark protein sequence also is consistent with the proposed models. Comparison of the Kyte–Doolittle (1982) hydrophilicity profiles shows that about 75% of both proteins have similar profiles (Fig. 5B). There are two dissimilar regions, one located near the N-terminus and the other near the C-terminus. The N-terminus region (residues 4–16; see above) is present in mammals but absent in sharks. The absence of this sequence in shark may result in the loss of the first beta sheet found in mouse MBPs (data not shown; see Stoner 1984). The remaining shark and mouse MBP sequences show similar alpha helix and beta sheet potentials, including the region near the C-terminus that exhibits different hydrophilicity profiles. Within those sequences of similar potential, shark and mouse MBPs show four regions each of which could adopt beta sheet conformations. The putative beta sheet regions of mouse MBP that are thought to form the

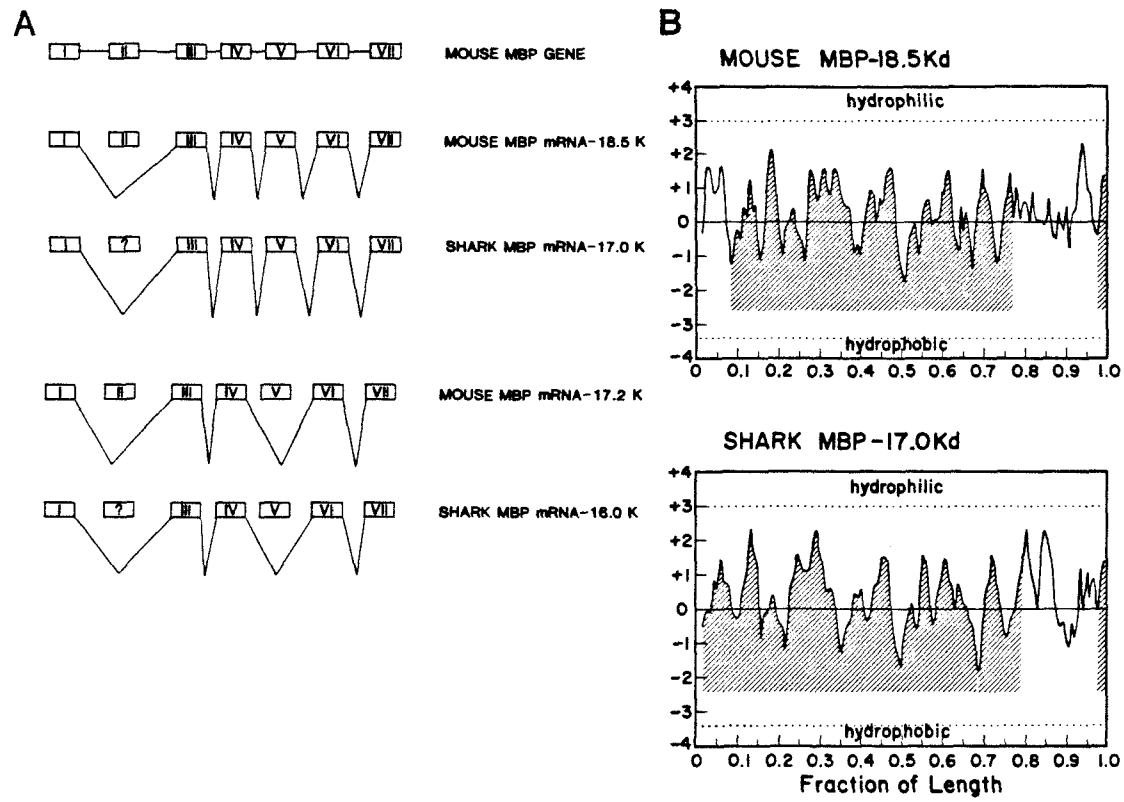


Fig. 5. Alternative splicing and structural comparison between mouse and shark MBP. **A** Schematic representation of the mouse MBP gene, mouse MBP mRNA, and shark MBP mRNA. Alternative splicing thought to generate two different MBP forms (Mr 17,000 and 16,000) in shark is illustrated and compared to the analogous splicing pattern that generates the Mr 18,500 and 17,200 MBP forms in mouse. The boxes represent exons. Roman numbers indicate known exons in mouse and putative exons in shark. The question mark indicates a hypothetical exon II in shark. **B** Comparison between the Kyte–Doolittle (1982) hydrophobicity profiles of mouse and shark MBPs. Regions of similarity are marked by the same pattern (lines inclining to the right).

central core and mediate the folding of the protein (see Stoner 1984) are on average 55% similar to those of shark MBP, whereas the regions between beta sheets (putative loops) are 34% similar.

We have shown that the myelin proteins from sharks and mammals appear to have similar structures, although (1) they are less than 50% similar in their sequences, and (2) these organisms diverged from each other about 400 million years ago (Romer 1970). These structural similarities will allow us to proceed further in our approach toward a better understanding of myelin proteins. It will allow us, for instance, to ask whether shark myelin proteins can function in a transgenic mouse, and if their expression there could improve demyelinating phenotypes. Moreover, our observations and those of others (Tai and Smith 1983; Waehneldt et al. 1985; Tai et al. 1986) raise questions on the nature of the myelinating cells in sharks and their subsequent evolution into the dual system—Schwann cell/oligodendrocyte—found in mammals (see Bakay and Lee 1966).

Acknowledgments. We are indebted to our colleagues Carol Readhead, Brian Popko, Carmie Puckett, Stephen Kent, Gary Pipes, Teresa Geffroy, and Joyce Liao for their valuable contributions and essential insights. This work was supported by grants from the National Institute of Health and the National Multiple Sclerosis Society (N.M.S.S.) to L.E.H. R.A.S. was a postdoctoral Fellow of the N.M.S.S., and R.A. was a Fellow of the Swiss National Science Foundation.

References

Aebersold RH, Teplow DB, Hood LE, Kent SBH (1986) Electroblotting onto activated glass: high efficiency preparation of proteins from analytical sodium dodecyl sulfate-polyacrylamide gels for direct sequence analysis. J Biol Chem 261: 4229–4238

Aebersold RH, Leavitt J, Saavedra RA, Hood LE, Kent SBH (1987) Internal amino acid sequence analysis of proteins separated by one- or two-dimensional gel electrophoresis after in situ protease digestion on nitrocellulose. Proc Natl Acad Sci USA 84:6970–6974

Aebersold RH, Pipes G, Nika H, Hood L, Kent SBH (1988) Covalent immobilization of proteins for high sensitivity se-

quence analysis electroblotting onto chemically-activated glass from SDS-polyacrylamide gels. Biochemistry 27:6860–6974

Bakay L, Lee JC (1966) Ultrastructural changes in the edematous central nervous system. Arch Neurol 14:644–660

Beaucage SL, Caruthers MH (1981) Deoxynucleoside phosphoramidites—a new class of key intermediates for deoxypolynucleotide synthesis. Tetrahedron Lett 22:1859–1862

Braun PE (1984) Molecular organization of myelin. In: Morell P (ed) Myelin, ed 2. Plenum Press, New York, pp 97–116

Bullock TH, Moore JK, Fields RD (1984) Evolution of myelin sheaths: both lamprey and hagfish lack myelin. Neurosci Lett 4:145–148

Chou PY, Fasman GD (1978) Prediction of the secondary structure of proteins from their amino acid sequence. Adv Enzymol 47:45–148

de Ferra F, Engh H, Hudson L, Kamholz J, Puckett C, Molineaux S, Lazzarini R (1985) Alternative splicing accounts for the four forms of myelin basic protein. Cell 43:721–727

Gubler U, Hoffman BJ (1983) A simple and very efficient method for generating cDNA libraries. Gene 25:263–269

Horvath SJ, Firca JR, Hunkapiller T, Hunkapiller MW, Hood L (1987) An automated DNA synthesizer employing deoxynucleoside 3'-phosphoramidites. Meth Enzymol 154:314–326

Huynh TV, Young RA, Davis RW (1985) Constructing and screening cDNA libraries in lambda gt10 and lambda gt11. In: Glover DM (ed) DNA cloning techniques: a practical approach, vol. 1. JRL Press, Oxford, pp 49–78

Kemali M, Miralto A (1983) The habenular nuclei of the elasmobranch *Scyllium stellare* myelinated perikarya. Z Mikrosk-Anat Forsch, Leipzig 97:3–14

Kirschner DA, Ganser AL, Caspar DLD (1984) Diffraction studies of molecular organization and membrane interactions in myelin. In: Morell P (ed) Myelin, ed 2. Plenum Press, New York, pp 51–95

Kyte J, Doolittle RF (1982) A simple method for displaying the hydropathic character of a protein. J Mol Biol 157:105–132

Lees MB, Brostoff SW (1984) Proteins of myelin. In: Morell P (ed) Myelin, ed 2. Plenum Press, New York, pp 197–294

Lemke G, Axel R (1985) Isolation and sequence of a cDNA encoding the major structural protein of peripheral myelin. Cell 40:501–508

Lemke G, Lamar E, Patterson J (1988) Isolation and analysis of the gene encoding peripheral myelin protein zero. Neuron 1:73–83

Martenson RE (1981) Prediction of the secondary structure of myelin basic protein. J Neurochem 36:1543–1560

Newman S, Kitamura K, Campagnoni AT (1987) Identification of a cDNA coding for a fifth form of myelin basic protein in mouse. Proc Natl Acad Sci USA 84:886–890

Norton WT, Poduslo SE (1973) Myelination in rat brain: method of myelin isolation. J Neurochem 21:749–757

Ohtsuka E, Matsuki S, Ikehara M, Takahashi Y, Matsubara K (1985) An alternative approach to deoxyoligonucleotides as hybridization probes by insertion of deoxyinosine at ambiguous codon positions. J Biol Chem 280:2606–2608

Readhead C, Popko B, Takahashi N, Shine HD, Saavedra RA, Sidman RL, Hood L (1987) Expression of a myelin basic protein gene in transgenic shiverer mice: correction of the dysmyelinating phenotype. Cell 48:703–712

Romer AS (1970) In: The vertebrate body, ed 4. WB Saunders Co., Philadelphia

Sanger F, Nicklen S, Coulsen A (1977) DNA sequencing with chain terminating inhibitors. Proc Natl Acad Sci USA 74: 5463–5467

Smith LM, Sanders JZ, Kaiser RJ, Hughes P, Dodd P, Connell CR, Heiner C, Kent SBH, Hood LE (1986) Fluorescence detection in automated DNA sequence analysis. Nature 321: 674–679

Stoner GL (1984) Predicted folding of beta-structure in myelin basic protein. J Neurochem 43:433–447

Tai FL, Smith R (1983) Shark CNS myelin contains four polypeptides related to the PNS protein P_o of higher classes. Brain Res 278:350–353

Tai FL, Smith R, Bernard CCA, Hearn MWT (1986) Evolutionary divergence in the structure of myelin basic protein: comparison of Chondrichthye basic proteins with those from higher vertebrates. J Neurochem 46:1050–1057

Thomas JO, Kornberg RD (1975) An octamer of histones in chromatin and free in solution. Proc Natl Acad Sci USA 72: 2626–2630

Waehneldt TV, Malotka J, Karin NJ, Matthieu J-M (1985) Phylogenetic examination of vertebrate central nervous system myelin proteins by electro-immunoblotting. Neurosci Lett 57: 97–102

Williams AF (1987) A year in the life of the immunoglobulin superfamily. Immunol Today 8:298–303

Received October 20, 1988/Revised and accepted January 9, 1989

Note added in proof. Recently, we used total shark brain RNA and oligonucleotides complementary to the sequences flanking the coding region of MBP in a polymerase chain reaction (PCR) to amplify its entire coding region. When these PCR products were sequenced they yielded the same splicing variants of MBP (bearing and lacking exon V) previously described. Thus, using the PCR technique and the cDNA sequence analysis mentioned in the text, we have been able to find only two classes of MBP transcripts in shark.

Chapter 2

CLONING OF THE SHARK PO AND MBP PROMOTERS USING A GENOMIC WALKING TECHNIQUE BASED ON THE POLYMERASE CHAIN REACTION

Lance Fors, Raul A.Saavedra[1],[2] and Leroy Hood

Division of Biology

California Institute of Technology

Pasadena, California 91125

[1]Present Address: Laboratory for the Study of Skeletal Dissorders

The Children's Hospital

Harvard Medical School

300 Longwood Avenue

Boston, MA 02115

[2]To whom correspondence should be addressed

ABSTRACT

We have cloned the putative shark promoters of protein zero (Po) and myelin basic protein (MBP) using a novel application of the ligation mediated single-sided polymerase chain reaction (PCR) method. This method uses linker ligation and subsequent amplifications with a linker primer and multiple specific primers to generate specificity. The method allowed us to amplify approximately 400 base pairs of shark genomic DNA sequence immediately upstream from the 5′ end of our full-length Po and MBP cDNAs. Both promoters were shown to be directly linked to their first exons, contain transcription initiation start sites and sequences commonly found in eukaryotic promoters. This genomic walking technique will be useful for cloning promoters, insertion sites, and other sequences of interest without the need for constructing and screening genomic libraries.

INTRODUCTION

Myelin of the vertebrate nervous system is a dynamic structure that provides electric insulation and facilitates the rapid conduction of nerve impulses along axons. The major myelin proteins in the mammalian central nervous system(CNS) are proteolipid protein(PLP) and myelin basic protein(MBP); whereas, in the peripheral nervous system(PNS) they are protein zero(Po) and MBP. Sharks diverged from other vertebrates about 400 million years ago and are the most primitive living vertebrates that show myelin organized as a concentric multilamellar structure (Saavedra et al. 1989. We have previously isolated and sequenced the major CNS myelin proteins and their corresponding cDNAs from the shark **Heterodontus francisci** (Saavedra et al. 1989). The sequence data showed that the major proteins of shark CNS myelin are similar to mammalian PNS proteins Po and MBP. This differential distribution of myelin proteins between sharks and mammals, the fact that Po expression is restricted to the mammalian PNS and the unconventional promoter sequence found in mammalian Po and MBP (Saavedra et al. 1989; Takahashi et al. 1985; Lemke et al. 1988) led us to clone the shark Po and MBP promoters.

Our inability to clone these promoters by conventional techniques combined with the large size of the shark genome (about two times the size of the mammalian genome; Hinegardner 1976) and the time and cost associated with constructing and screening of genomic libraries motivated us

to find an alternative strategy to clone the shark Po and MBP promoters. The polymerase chain reaction (PCR) (Saiki et al. 1988) allows the amplification of a region of DNA between two sequence-specific oligo-nucleotides as primers, and appropriately modified could provide the needed alternative. The PCR technique takes advantage of a DNA-polymerase from **Thermus aquaticus** (Taq DNA-polymerase), which is functional at high temperatures. This thermostability permits one to denature the template, hybridize primers and carry out the polymerization reaction at high temperatures. The high reaction temperature makes possible artifacts due to spurious hybridization less likely. Conventional PCR requires one to know the sequences complementary to both oligonucleotide primers. A modification called "inverted PCR"(IPCR) allows the amplification of unknown sequences located outside of the known DNA region (Triglia et al. 1988). This procedure involves cutting the DNA with restriction endonucleases and ligation the resulting free ends under conditions that favor circularization of the DNA. As a result of this manipulation, oligonucleotides complementary to the 5′ and 3′ ends of the known sequences can now be used to amplify the unknown sequence of interest located at the 5′ and/or 3′ ends.

Recently a ligation mediated single-sided PCR strategy has been developed (Mueller and Wold 1989) for in vivo footprinting. This procedure randomly cleaves DNA with

dimethyl sulfate/piperdine, uses a 5′ specific primer in an extension reaction to synthesize the sequence of interest, attaches a linker at the newly created 3′ blunt end and amplifies that region through 16 cycles of PCR using a linker primer and a second specific 5′ oligonucleotide.

In this article, we describe an adaptation of the ligation mediated single-sided PCR method that allowed us to amplify and clone the putative shark Po and MBP promoters.

MATERIALS AND METHODS

Oligonucleotides. For each gene (Po and MBP) a set of four oligonucleotides were synthesized in a 380B automatic DNA synthesizer (Applied Biosytems Inc., Foster City, Ca), using the chemistry of Caruthers (Beaucage and Caruthers 1981). Three of these oligonucleotides, the specific primers, were used for the required extensions and amplification reactions. The specific primers for Po were: a 20-mer (specific primer Po1; 5′-TAAGAAGTTGAGTTACTCCC-3′), a 31-mer (specific primer Po2; 5′-AGGAATGGCTCTGACGGCTGGCACTGTTTAA-3′), and a 25-mer (specific primer Po3; 5′-CACTGTTTAAGCTGC ACCTCTCTCT-3′). The specific primers for MBP were: a 20-mer (specific primer MBP1;5′-GTGGGCTCCTCCAGCTTGCT-3′), a 31-mer (specific primer MBP2;5′-GGTGCTTGCGGATGCCATTTGTTTTGCTACT-3′), and a 25-mer (specific primer MBP3;5′-TTTTGCTACTTGTCC CTTCTTAGCT-3′). The fourth oligonucleotide (probe) was synthesized from the 5′ most sequence of exon I and used to identify the specific amplification products at each stage of the procedure. In the case of Po a 25-mer(Po6;5′-GTCTCTCTCCTTTACCAGCGCCGGAA-3′) was synthesized, while for MBP a 27-mer (MBP6;5′-GCTGATGATTGTGAGGACAGTTTCGGA-3′) was synthesized.

Restriction Enzyme Cleavage. Fifteen ug of liver DNA from the shark **Heterodontus francisci** was cleaved with a mixture of 40 units each of Bg1II, EcoRI, and HindIII for two hours at 37^{o}C. The DNA mixture was then placed at 68^{o}C for 10 minutes, extracted with phenol/chloroform/isoamyl

alcohol (50:48:2) mix, ethanolprecipitated and washed. The resulting pellet, which contained genomic fragments with protruding or recessed ends and with a mean size of 1 to 2 kb, was resuspended in 20 ul of 10 mm Tris, pH 8.0, 0.1 mM EDTA (TE buffer) and used in the primer extensions.

Primer extensions. The extensions were done by mixing 3 ug of restricted shark liver DNA, 0.15 pm of specific primer 1 in a modified SEQUENASE buffer (40 mm Tris, pH 7.7, 50 mM NaCl) to a final volume of 15 ul. Denaturation at 95^{o}C for 3 minutes was followed by annealing at Tm - 5^{o}C (47^{o}C for Po and 61^{o}C for MBP) for 30 minutes. The mixture was then placed on ice and 7.5 ul of 20 mm $MgCl_2$, 20 mM dithiohreitol (DTT) and 0.1 mM of each deoxynucleotide triphosphate (dNTP) was added followed by 1.5 ul of freshly diluted (1:4 in ice cold TE buffer) SEQUENASE (modified T7 DNA-polymerase from United States Biochemical Co., Cleveland, OH). The extension reaction was done at 45^{o}C for 5 minutes. The reaction tube was then transferred to ice and 6 ul of ice cold 310 mM Tris, pH 7.7 was added followed by 15 min incubation at 67^{o}C to inactivate the enzyme.

Linker. Two complementary oligonucleotides to serve as a linker were synthesized using the procedure described above (Beaucage and Caruthers 1981; Horvath et al. 1987). One was a 25-mer (5′-GCGGTGACCCGGGAGATCTGAATTC-3′) the other was a 11-mer (5′-GAATTCAGATC-3′) (Mueller and Wold 1989). These two oligonucleotides were hybridized by boiling and slowly cooling them to room temperature over a period of

several hours. The result is a partially double-stranded linker with one blunt and one recessed or protruding end.

Ligation Reaction. The blunt end of the linker was ligated to the blunt end of the specific primer 1 extended DNA. The 25-mer linker oligonucleotide (linker primer) can then be used with specific primer 2 in the subsequent PCR amplification. To each extension reaction 20 ul of ice cold (17.5 mm $MgCl_2$, 43 mm DTT and 125 ug/ml bovine serum albumin, BSA) and 25 ul of ice cold ligation mix were added to yield final concentrations of 50 mm Tris, pH 7.7, 10 mm $MgCl_2$, 20 mm DTT, 3 mm ATP, 50 ug/ml BSA, 100 pM linker and 3 units of T4 DNA-ligase. The ligation reaction was incubated at 22°C for 2 hours followed by 10 minutes at 70°C to inactivate the enzyme. The ligation reaction was precipitated after adding 0.1 volume sodium acetate, 10 ug tRNA and 2.5 volumes of 100% ethanol, washed with 75% ethanol and resuspended in 50 ul of water.

Polymerase Chain Reactions. PCR was performed in a Cetus-Perkins Thermocycler using a standard protocol (Gene Amp™, Perkin-Elmer Cetus Instruments catalogue, Norwalk, CT). Each reaction mixture (a total volume of 100 ul) contained the reaction buffer (100 mm Tris-HC1, pH 8.3 at room temperature, 500 mm KC1, 15 mm $MgC1^2$ and 0.1 % gelatin), 200 um each dNTP, template DNA (for the first amplification: 1 ug total shark DNA, which had been restricted, extended and linker ligated; for the second amplification: approximately 1 ng obtained from a region

excised from a low melting point agarose, LMP-agarose, gel), 825 ng of each primer, and 2.5 units of Taq DNA-polymerase (Perkin-Elmer Cetus Instruments, Norwalk, CT). The samples were denatured for 7 minutes at 94^{o}C, followed by 25 or 30 amplification cycles (1 minute denaturation at 94^{o}C, 2 minutes annealing at 66^{o}C, and 3 minutes extension at 76^{o}C).

Agarose Gel Electrophoresis and Southern Blots. The PCR products were analyzed by agarose gel electrophoresis and transferred onto Zeta-Probe Blotting membranes(Bio-Rad Laboratories, Richmond CA) as described by Southern (Southern 1977). The membranes were hybridized for 18 hours in 5 X saline sodium citrate(SSC), 10 mm EDTA, 1% sodium dodecyl sulphate(SDS) with the appropriate oligonucleotide probe at T_m-10^{o}C. The membranes were washed in 5X SSC and 0.1% SDS, two times for 10 minutes at 22^{o}C, and two times for 30 minutes at 37^{o}C, and exposed for 18 hours to identify the bands of interest. When indicated LMP-agarose was used to facilitate the subsequent step of the procedure.

Cloning of Amplified Sequences. It is probable that the amplified products contain incompletely polymerized ends. Therefore, the amplified fragments were first treated with T4 DNA-polymerase in ligation buffer, 10 um of each dNTPs, and 5% polyethylene glycol for 20 minutes at 37^{o}C, to create blunt ends, followed by 5 minutes at 65^{o}C to inactivate the enzyme. The resulting insert was ligated directly into Sma I digested M13mp18 and pUC18 using a 3:1 molar excess of insert for 18 hours at room temperature and

transformed into DH-alpha cells (Bethesda Research Laboratories, Bethesda MD).

DNA Sequencing. DNAs from cloned plasmid preparations or amplified PCR fragments were sequenced using the dideoxynucleotide method (Sanger et al. 1989) with SEQUENASE in conjunction with appropriate oligonucleotides primers (Beaucage and Caruthers 1981; Horvath et al. 1987). The gels were autoradiographed using XR-5 film from Kodak (Rochester,NY).

Linkage of the Amplified Sequences to Exon I. A PCR experiment was performed using shark genomic DNA as the template and oligonucleotides complementary to the promoter and the 3'-most sequence amplified from exon I. In the case of Po specific primer Po2 and a 28-mer (PoC;5'-GACCAGTAGGCA GTTGCATCTGCAAAC-3'), while for MBP specific primer MBP2 and a 28-mer (MBPC;5'-GCCATCCAAGAGGATAGCTCAGTTCCTG-3') were used. The production of the expected amplification products would show that the promoter was located adjacent to exon I in the shark genome.

Determination of the Transcription Initiation Sites. Extension experiments were done using 10 ug of total shark brain RNA (Maniatis et al. 1982). The RNA was hybridized with 3 X 10^5 cpm of MBP and Po specific oligonucleotides 1.1 and 2 (purified by gel electrophoresis after end labeling with ^{32}P-gamma-adenosine triphosphate and T4 polynucleotide-kinase), respectively, in 10 mM Tris, pH 8.0, 250 mM KCl and 10 mM EDTA for 45 min at 62^{o}C. Samples were allowed to cool

to room temperature and 2.5 volumes of 50 mM Tris, pH 8.0, 10mM $MgCl_2$, 5 mM DTT, 325 ug each dNTP, 10 ug/ml actinomycin and 8 to 10 units of reverse transcriptase (AMV 007 from Life Science, Plainview, NY) were added. The samples were incubated for 45 minutes at 42°C. The reactions were stopped by putting them on ice, and 2 volumes of a 400 mM ammonium acetate, 10 mM EDTA mix were added. The samples were extracted with phenol/ chloroform/isoamyl alcohol, ethanol precipitated, resuspended in dye buffer, and loaded on a sequencing gel adjacent to a M13mp18 sequence ladder that served as a size marker.

Sequence Similarities. The sequence similarities between the shark and mammalian promoters were determined by computing the number of identical base pairs (bp) after manual alignment of the corresponding sequences from their respective transcription initiation sites.

RESULTS

The shark Po and MBP cDNA cloned that we previously isolated contained 5′ untranslated regions, coding regions, relatively long 3′ untranslated regions, putative polyadenylation signals and polyadenosine tails similar to their mammalian counterparts (Saavedra et al. 1989). RNA blot analysis showed that both cDNA cloned were approximately the same size as their respective transcripts. Thus, we considered these clones to be full-length cDNAs. Next, we attempted to isolate genomic clones containing the shark Po and MBP promoters by screening shark genomic libraries generated by SauIIIA partial digestions (constructed by us or kindly provided by G. Litman, Florida) using our shark cDNA clones as probes. After constructing three and unsuccessfully screening a total of five shark genomic libraries, we decided to attempt to clone the shark Po and MBP promoters using alternative methods. We chose the IPCR (Triglia et al. 1988) and the single-sided PCR method developed for in vivo footprinting (Mueller and Wold 1989) to permit us to walk from a known region of a genome (i.e., from exon I) toward an unknown region (i.e., the promoter).

Specific Primers, Probes and Linker. A set of three specific primers for each gene was synthesized as described in Materials and Methods (Figure 1A). Specific primer 1 is separated by 10 bases from specific primer 2, which in turn overlaps 10 bases with specific primer 3. An additional

oligonucleotide (probe) was also synthesized for each gene from the 5′ most sequence of exon I. The specific primers and the probe were all complementary to sequences previously identified as the 5′ most end of exon I on either the shark Po or MBP cDNAs (Saavedra et al. 1989).

The Genome Walking Scheme. Figure 1B shows our genome walking scheme. High molecular weight shark liver DNA (greater than 50 kilobases (Kb), as determined by agarose gel electrophoresis; data not shown) was digested with a mix of three different restriction endonucleases with six bp recognition sites (see Materials and Methods), all of which restrict the DNA leaving recessed or protruding ends. The restricted DNA was then denatured by heating it to 95^{o}C for 2 minutes and specific primer 1 was allowed to anneal by lowering the temperature to Tm-5^{o}C (Po1 at 47^{o}C, and MBP1 at 61^{o}C) for 30 min. SEQUEASE was used to extend for 5 minutes at 45^{o}C from specific primer 1 up to the nucleotide sequence where the restriction enzymes had cut (unknown up to this step of the procedure). Theoretically, only those fragments which are primed could produce a partially double stranded DNA molecule blunted at the unknown end. A DNA linker with a single blunt end was ligated directionally onto the double stranded blunt end of the extension product using T4 DNA-ligase. This linker has no 5′ phosphates and is staggered to eliminate self-ligation and provide directionality. Also, the duplex between the 25-mer and 11-mer is stable at

the ligation temperature, but denatures easily during subsequent PCR.

The resulting ligation product was then denatured by heating at 94^{o}C for 1 minute and specific primer 2 was allowed to anneal by lowering the temperature to 66^{o}C for 2 min. Specific primer 2 was extended using Taq DNA-polymerase, and a first PCR (25 cycles) was carried out using specific primer 2 and the linker primer (25-mer), both of which have a comparable Tm. The amplification products were electrophoresed on an agarose gel. Ethidium bromide staining of this gel reveled a faint smear instead of a discrete band (Figure 2A). Two possible reasons for this smear are: i) specific primer 1 (a 20-mer) in combination with the 45^{o}C primer extension temperature could not provide sufficient specificity and/or ii) unwanted blunt ends capable of accepting the linker were produced by contaminants in the restriction enzymes used. The DNA was transferred onto a membrane, hybridized with a ^{32}P-labeled oligonucleotide probe, and autoradiographed (Figure 2B). The probe hybridized with a single band located near the top of the smear. Bands of about 400 and 535 bp were obtained for the Po and MBP genes, respectively. The amplification products were then run on a LMP-agarose gel electrophoresis and the regions from 325 to 445 bp for Po, and from 460 to 580 bp for MBP were excised, melted and used as template for a new round of PCR (25 cycles). This time, specific primer 3 and the linker primer were used. The products of this

second round of amplification were discrete bands of 385 bp for Po (clone Po 385) and 520 bp for MBP (clone MBP520) with little or no background (Figure 2C). The identity of these bands was determined by hybridization to ^{32}P-labeled oligonucleotide probes complementary to the 5′ most end of the corresponding exon I (data not shown).

The Sequence of the Amplification Products. The amplification products Po 385 and MBP520 were sequenced (see Materials and Methods), and each of these clones contained three regions: specific primer 3 with the remaining 5′ sequence of exon I at their 3′ ends (Saavedra et al. 1989), an unknown sequence representing the putative promoters in the middle, and the linker primer sequence at their 5′-ends. In addition, the 5′ most end of the unknown regions contained one of the expected restriction endonuclease recognition sites (EcoRI for Po and HindIII for MBP; Figures 3 and 4).

The Known and the Unknown Regions of Both Clones Po385 and MBP520 Are Located Adjacent to each Other in the Shark Genome. We next tested whether the known and unknown regions of the amplification products are linked in the shark genome, using PCR (30 cycles). In this experiment, an oligonucleotide complementary to the known region and the unknown region, facing each other, served as primers with shark genomic DNA as template. The oligonucleotides used for shark Po were specific primer Po2 and PoC, and for shark MBP were specific primer MBP2 and MBPC. The expected sizes

of the amplification products were 315 bp for Po and 131 bp for MBP. The observed amplification products were a band of about 315 bp for Po and a band of about 130 bp for MBP (Figure 5A). These results show that the known and unknown regions which are adjacent to each other in the Po 385 the MBP520 clones are also adjacent to each other in the shark genome.

Determination of the Transcription Initiation Sites. To determine the transcription initiation sites for shark Po and MBP, primer extension experiments were performed using total shark brain RNA, in conjunction with primers Po2 and MBP1.1 for Po and MBP, respectively (see Materials and Methods). The primers were annealed and extended, producing discrete products whose end points identify potential transcription start sites. The initiation of transcription for Po lies around the sequence 5'-GCTCGACT-3', where C is the most probable site (Figure 5B). The initiation of transcription for MBP lies around sequence 5'-CTCAGTTC-3', where A is the most probable site (Figure 5B). Both sites were located 15 to 40 bp upstream from the 5' end of our original Po and MBP cDNA clones (Saavedra et al. 1989). These data also support the idea that our original cDNA clones are full-length copies of their respective mRNA's.

DISCUSSION

The sequence analysis of clones Po385 and MBP520 identified priviously unknown DNA regions which we called the shark Po and MBP promoters for the following reasons. First, these DNA sequences were shown to be adjacent to the 5′ end of their corresponding cDNAs in the shark genome. Second, the priviously unknown regions each contain transcription initiation sites, as determined by primer extension experiments. A weak concensus sequence (5′-PyPyCAPyPyPyPy-3′;Py=pyrimidine) for the initiation of transcription has been identified in mammalian genes, with transcription starting at A (Smale and Baltimore 1989). Our proposed initiation site for MBP (5′-CTCAGTTC-3′) is 88% similar to this concensus sequence, contains the two specified bases and initiates at the A. Third, the previously unknown regions show other sequence features that are often found in eukaryotic promoters. Clone Po385 shows a TATA box (5′-TTAAAA-3′), a CAT box (5′-CAATTT-3′), and sequence of 26 alternated Py/G (5′-TGTGCGTGTGTGCGTGCGTGTGTGTG-3′) which could potentially form left handed Z-DNA (Rich and Wang 1984; Bourbon et al. 1988). These sequences are located -26 to -31, -113 to -118 and -250 to -276 bp from the transcription initiation site, respectively. The proposed shark Po TATA and CAT boxes are 67% and 83% similar to their mammalian counterparts, respectively (Lemke et al. 1988). Clone MBP520 shows a TATA box (5′-ATCCAA-3′) and a CAT box (5′-CCATCG-3′). These

sequences are located -12 to -17 and -39 to -44 bp from the initiation site of transcription, respectively. The proposed shark MBP TATA and CAT boxes are 67% and 50% similar to their mammalian counterparts, respectively (Takahashi et al. 1985). Additional TATA and CAT homologies were identifies at -51 to -56 and -103 to -110 bp, respectively, which could drive transcription from two other weak transcription initiation sites (-22 and -33 bp, respectively) seen in Figure 5B. This positional information and sequence similarity taken against an overall similarity of less than 25% between the shark and mammalian promoter sequences (aligned from their rsspective transcription initiation sites suggests that these are functionally important features of true promoters.

We have recently found additional similarity between the shark and the mammalian Po and MBP promoters that is described in Chapter 3. This newly discovered sequence similarity should help explain the differential expression exhibited by these genes.

The large size (6×10^9 bp[4]) of the shark haploid genome could explain why some single copy genes - like Po and MBP (Saavedra et al. 1989) - were absent from our genomic libraries. Therefore, the shark provides an excellent system to test this genome walking strategy. A major focus in the design of any amplification strategy is specificity. We have solved this problem by adapting, as follows, the ligation mediated PCR technique of Mueller and

Wold (Mueller and Wold 1989). First, the linker is ligated to the only blunt ended DNA molecule that should be available after the first DNA denaturation and primer extension. Since the DNA was cut with enzymes which produce only protruding or recessed ends, then only those molecules specifically extended should have a single blunt end which can directionally accept the single blunt end of the linker. Second, each of the subsequent amplification steps has been carried out using the linker primer and a new specific primer that is fully contained within the DNA sequence of the product of the previous step. Third, the specific amplification products from the first round of PCR were significantly enriched by size fractionation on LMP-agarose before the final roun of amplification.

We have shown that the above strategy can be used to walk from a known region towards an unknown region in a large genome. There are several possible technical improvements which could increase its effectiveness. First, the length of the cloned products could be increased by adjusting the number and nature of restriction endonucleases used to digest the genomic DNA. In our experiments the length of the promoter regions we amplified was limited by the EcoRI site for Po and the HindIII site for MBP found at their 5′ ends, respectively (see Figures 3 and 4). Second, the first specific primer could be longer (a 25-mer or more), and Taq DNA-polymerase could be used instead of SEQUENASE in order to perform the initial extension step at

a temperature higher than 45°C. Third, the length of each cycle or the number of cycles could be altered. Modifications of this kind, however, must take into consideration the size limit of the PCR products (now up to 5 or 8 kb) and the probability of obtaining amplification artifacts. The appropriate manipulation of some or all of these variables should result in a higher yield of products of the desired length.

The advent of new biotechnologies has made the cloning and sequencing of large genomes a manageable enterprise. The large size and sequence organization of eukaryotic genomes makes it possible that some regions of these genomes will not yield to present cloning techniques. It would also be advantageous to clone smaller regions of genomic DNA without having to construct and screen a genomic library. Our genome walking technique is independent of conventional cloning steps, and therefore provides a method of choice to fill in genomic library gaps or to clone any other region of interest.

ACKNOWLEDGEMENTS

We wish to thank Paul Mueller and Barbara Wold for providing us with the linker oligonucleotides, essential theoretical insights and practical suggestions. We also wish to thank Teresa Geffroy for outstanding technical support and Bernard Vernooy for critically reading the manuscript. The oligonucleotides used in this work were synthesized at the Caltech microchemical facilities by Suzanna Horvath and coworkers. This work was supported by grants from the National Institutes of Health and the National Multiple Sclerosis Society (NMSS) to LH. The data contained in this article is a part of LF's thesis, in partial fulfillment for the degree of doctor of philosophy from Caltech; he also acknowledges the support from NIH grant NS14069. RAS was a post - doctoral fellow of the NMSS.p

References

Beaucage, S.L. and Caruthers, M.H. Deoxynucleoside phosphoramidites - a new class of key intermediates for deoxypolynucleotide synthesis. **Tetrahedron Lett. 22:** 1859-1862, 1981.

Bourbon, H.-M., Prudhomme, M. and Amalric, F. Sequence and structure of the nucleolin promoter in rodents: characterization of a strikingly conserved CpG island **Gene 68:**73-84, 1988.

Hinegardner, R. The cellular DNA content of sharks, rays and some other fishes. **Comp. Biochem. Physiol. 55B:**367-370, 1976.

Horvath, S.J., Firca, J.R., Hunkapiller, T., Hunkapiller, M.W. and Hood, L. An automated DNA synthesizer employing deoxynucleoside 3'-phosphoramidites. **Meth. Enzymol. 154:**314-326, 1987.

Lemke, G., Lamar, E. and Patterson, J. Isolation and analysis of the gene encoding peripheral myelin protein zero. **Neuron 1:**73-83, 1988.

Maniatis, T., Fritsch, E.F. and Sambrook, J. **Molecular cloning:** A Laboratory Manual, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, 1982.

Mueller, P.R. and Wold, B. In vivo footprinting of a muscle specific enhancer by ligation mediated PCR. **Science 246:**780-786, 1989.

Rich, A., Nordheim, A. and Wang, A. The Chemistry and Biology of left-handed Z-DNA. **Annu. Rev. Biochem.53:** 791-846, 1984.

Saavedra, R.A., Fors L., Aebersold, R.H., Arden, B. Horvath, S., Sanders, J. and Hood, L. The myelin proteins of the shark brain are similar to the myelin proteins of the mammalian peripheral nervous system. **J. Mol. Evol. 29:** 149-156, 1989.

Saiki, R.K., Gelfand, D.H., Stoffel, S., Scharf, S.T., Higuchi, R., Horn, G.T., Mullis, K.B. and Erlich, H.A. Primer-directed enzymatic amplification of DNA with a thermostable DNA polymerase. **Science 239:**487-491, 1988.

Sanger, F., Nicklen, S. and Coulsen, A. DNA sequencing with chain-terminating inhibitors. **Proc. Natl. Acad. Sci. USA, 74:**5463-5467, 1977.

Smale, S.T. and Baltimore, D. The "Initiator" as transcriptional control element. **Cell 57:**103-113. 1989

Southern, E. Detection of specific sequences among DNA fragments separated by gel electrophoresis. **J. Mol. Biol. 98:**503-518, 1977

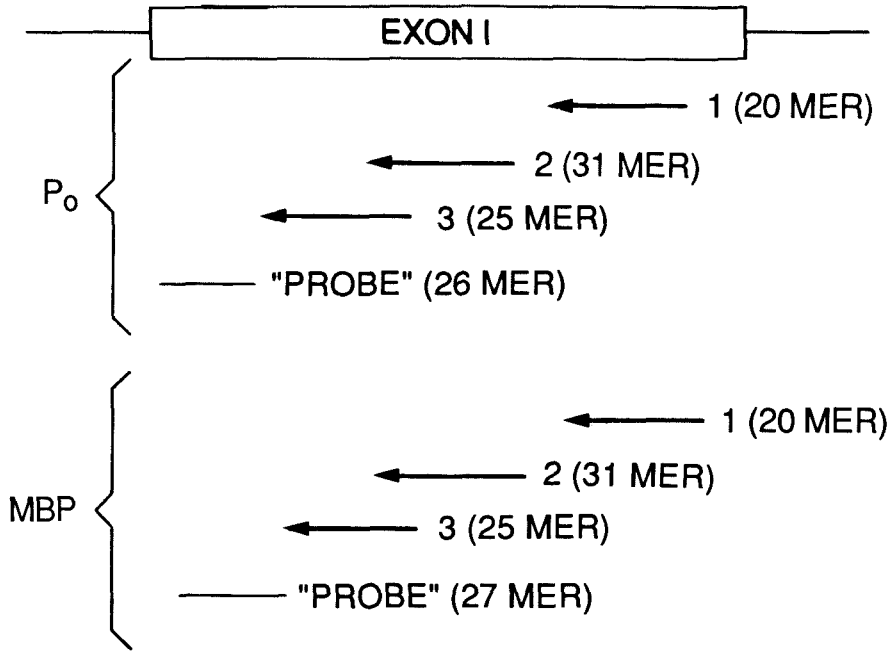
Takahashi, N., Roach A., Teplow, D.B., Prusiner, S.B. and Hood, L.E. Cloning and characterization of the myelin basic protein gene from mouse: One gene can encode both 14 kD and 18.5 kD MBPs by alternative use of exons. **Cell 42:**139-148, 1985.

Triglia, T., Peterson, M.G. and Kemp, D.J. A procedure for in vitro amplification of DNA segments that lie outside the boundaries of known sequences. **Nuc. Acids Res. 16:** 8186, 1988.

FIGURE 1. The chromosomal walking procedure. A. The specific oligonucleotides and the linker (including restriction sites). B. A schematic of the chromosomal walking procedure (for additional details on the initial steps see Mueller and Wold, 1989).

A

SPECIFIC PRIMERS



LINKER

5' – GCGGTGACCCGGGAGATCTGAATTC – 3' 25 MER
CTAGACTTAAG – 5' 11 MER

RESTRICTION SITES :
Bgl II , Bst E II
EcoR I, Sma I

B

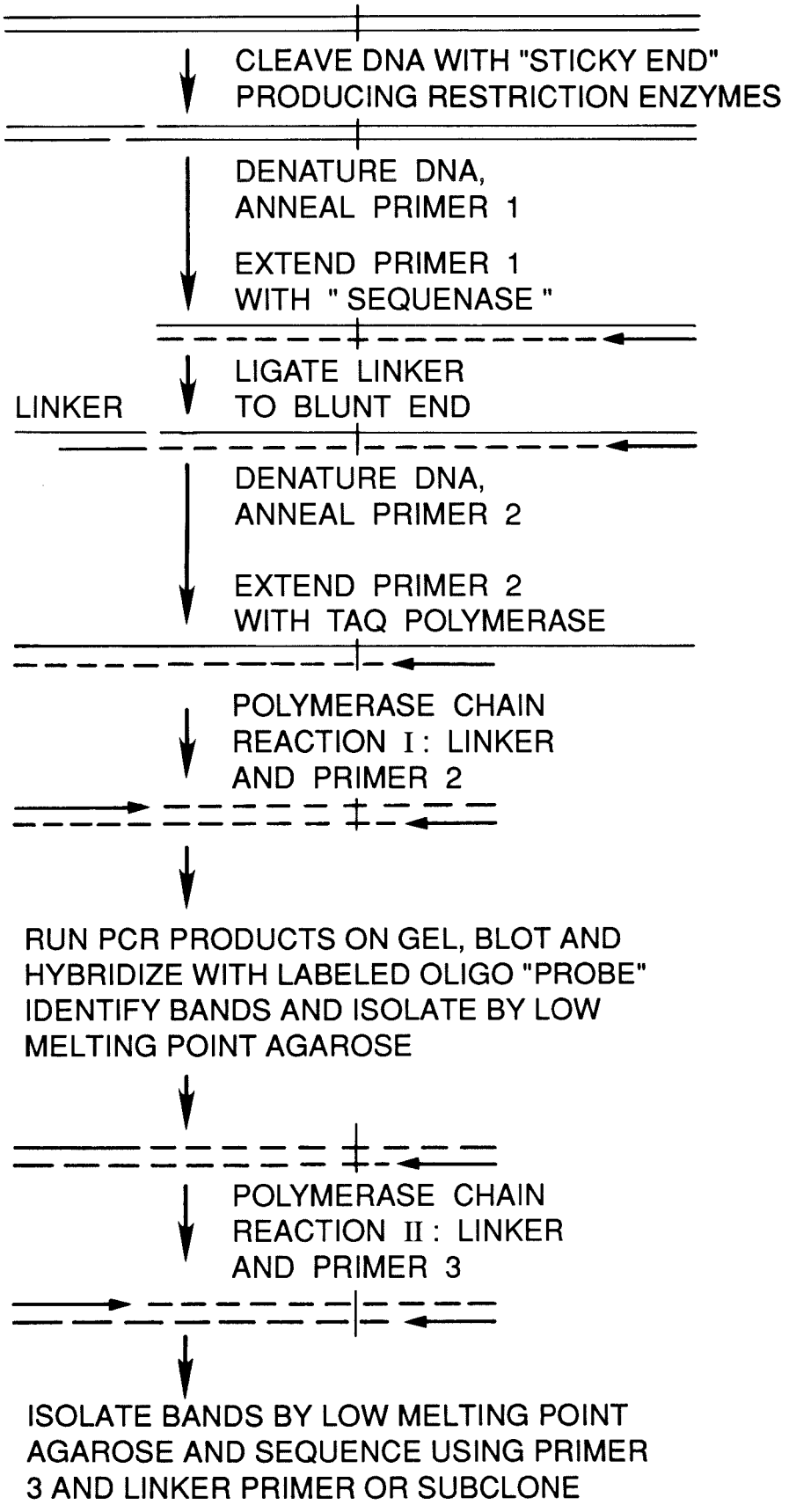
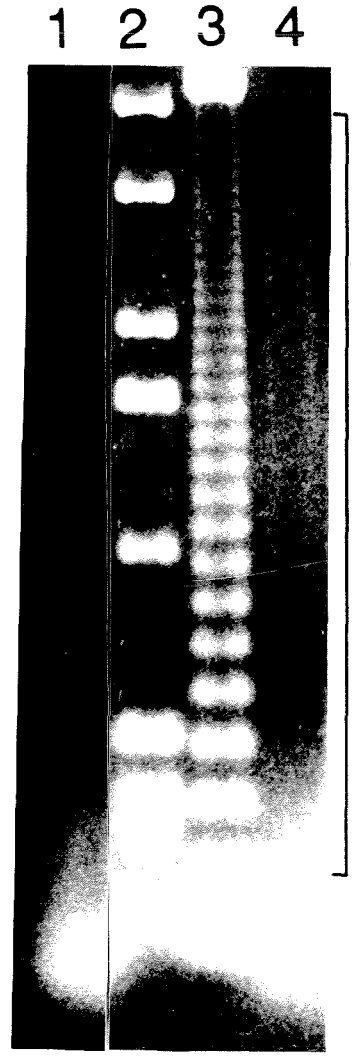
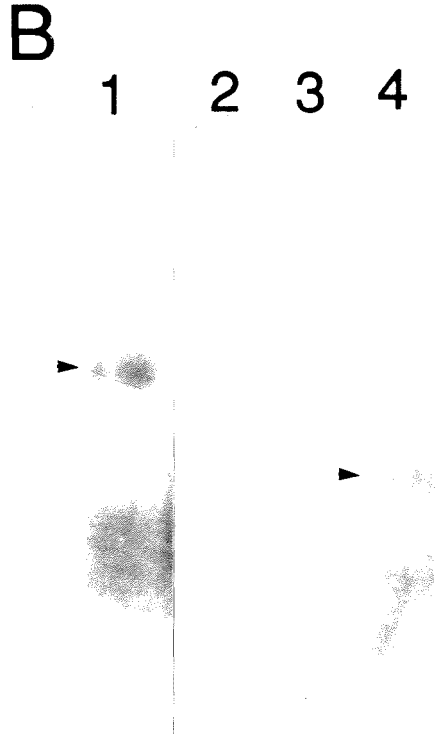

CLEAVE DNA WITH "STICKY END"
PRODUCING RESTRICTION ENZYMES
DENATURE DNA,
ANNEAL PRIMER 1
EXTEND PRIMER 1
WITH " SEQUENASE "
LIGATE LINKER
TO BLUNT END
LINKER
DENATURE DNA,
ANNEAL PRIMER 2
EXTEND PRIMER 2
WITH TAQ POLYMERASE
POLYMERASE CHAIN
REACTION I : LINKER
AND PRIMER 2
RUN PCR PRODUCTS ON GEL, BLOT AND
HYBRIDIZE WITH LABELED OLIGO "PROBE"
IDENTIFY BANDS AND ISOLATE BY LOW
MELTING POINT AGAROSE
POLYMERASE CHAIN
REACTION II : LINKER
AND PRIMER 3
ISOLATE BANDS BY LOW MELTING POINT
AGAROSE AND SEQUENCE USING PRIMER
3 AND LINKER PRIMER OR SUBCLONE

FIGURE 2. Results from specific steps of the procedure. A. Ethidium bromide picture of an agarose gel after the first round of amplification. Lane 1, PCR products (faint smear) from 25 cycles of amplification with linker primer and specific oligonucleotide MBP2, lane 2, a 1 kb ladder marker; lane 3, 123 bp ladder marker; lane 4, PCR products (faint smear) from 25 cycles of amplification with linker primer and specific oligonucleotide Po2, [demarks a region where a faint smear could be seen. B. Southern blot of the gel shown in A. Lane 1, MBP band (at about 535 bp, marked by an arrow head) detected by probe MBP6; lanes 2 and 3, 1 kb and 123 ladder markers (not visible here); lane 4, Po band (at about 400 bp, marked by an arrow head) detected by probe Po6. C. Ethidium bromide picture of an agarose gel after the second round of amplification. Lane 1, PCR product (a band of about 520 bp, marked by an arrow head) from 25 cycles of amplification with linker primer and specific oligonucleotide MBP3; lane 2, 123 bp ladder marker; lane 3, PCR product (a band of about 385 bp, marked by an arrow head) from 25 cycles of amplification with linker primer and specific oligonucleotide Po3. In this Figure, part A, lanes 1-4 are from the same gel; part B, lanes 1-4 are from the same autoradiogram.

A




B
1 2 3 4

C

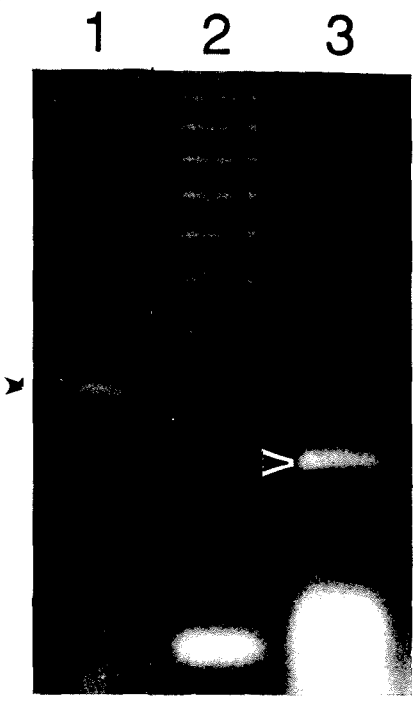


FIGURE 3. Nuceotide sequence of the shark Po promoter. In upper-left box (light line), the linker; in bottom-right box (heavy line), the 5′ end of shark Po cDNA (Saavedra et al. 1989), underlined (light), putative shark Po TATA and CAT sequences; underlined (heavy), a sequence of alternated Pu/Py (with Z-DNA formation potential).

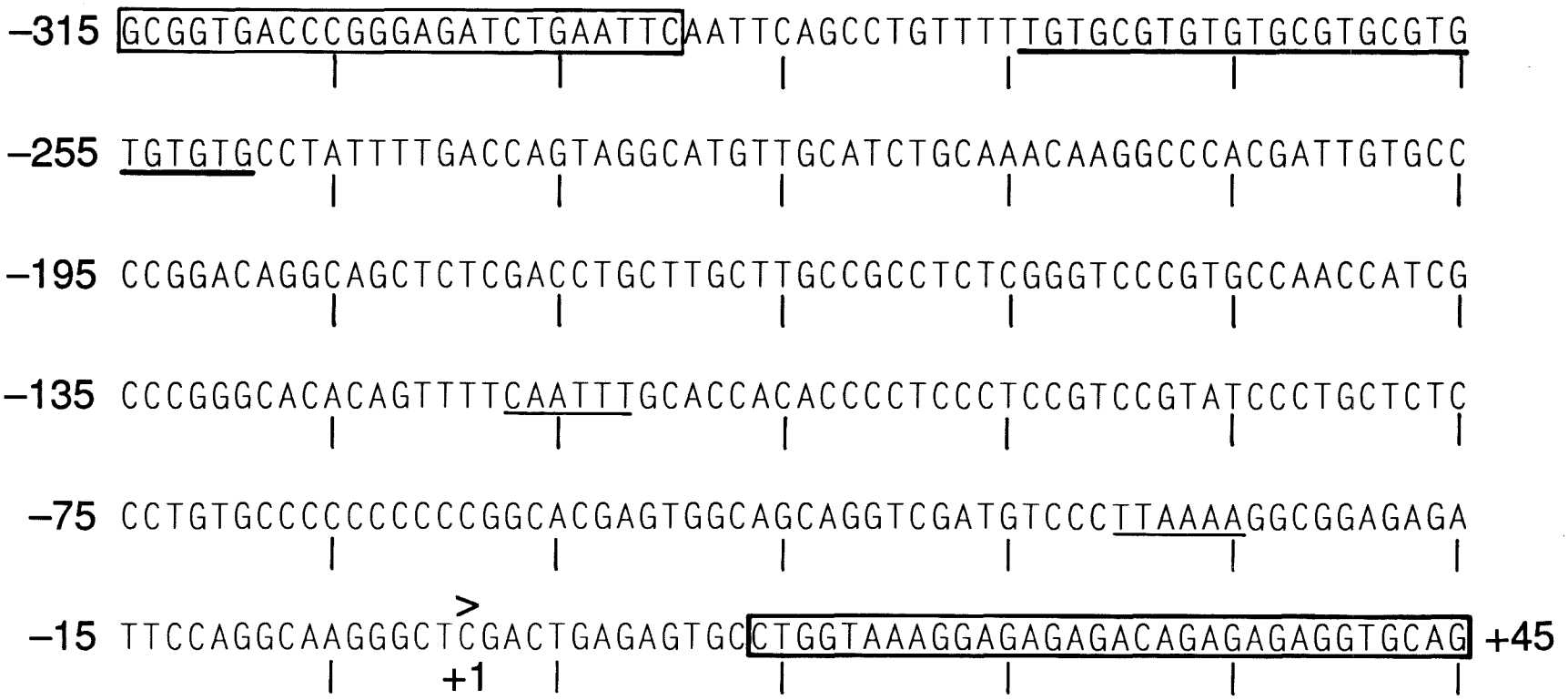
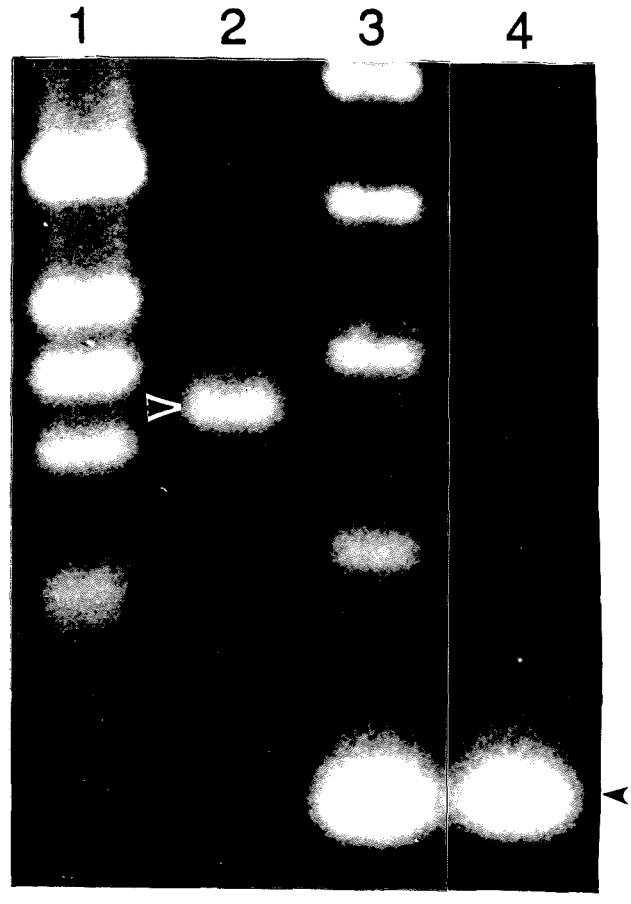

−315 GCGGTGACCCGGGAGATCTGAATTCAATTCAGCCTGTTTTTGTGCGTGTGTGCGTGCGTG
−255 TGTGTGCCTATTTTGACCAGTAGGCATGTTGCATCTGCAAACAAGGCCCACGATTGTGCC
−195 CCGGACAGGCAGCTCTCGACCTGCTTGCTTGCCGCCTCTCGGGTCCCGTGCCAACCATCG
−135 CCCGGGCACACAGTTTTCAATTTGCACCACACCCCTCCCTCCGTCCGTATCCCTGCTCTC
−75 CCTGTGCCCCCCCCCCGGCACGAGTGGCAGCAGGTCGATGTCCCTTAAAAGGCGGAGAGA
−15 TTCCAGGCAAGGGCTCGACTGAGAGTGCCTGGTAAAGGAGAGAGACAGAGAGAGGTGCAG +45
+1

FIGURE 4. Nucleotide sequence of the shark MBP promoter. In upper-left box (light line), the linker; in bottom-right box (heavy line), the 5′ end of shark MBP cDNA (Saavedra et al. 1989); underlined (light), shark MBP TATA and CAT sequences.

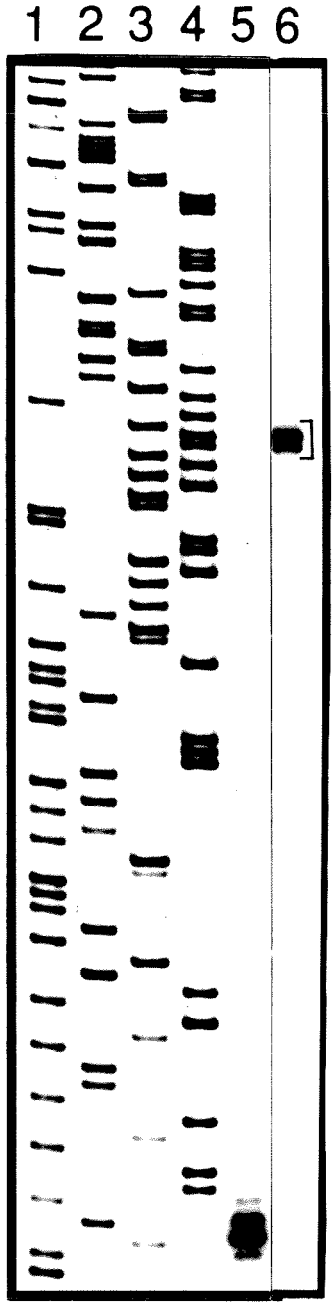

-421 GCGGTGACCCGGGAGATCTGAATTCAGCTTTGAATTAACAGAAAGGTAACTGCTGTAAAA
-361 TTTAAAATTTGATTTGCTTCTGTGAAATTGAATTGGTGAAATGACTACGGATGTCTTTGT
-301 GAAAAACAGCATTGTTCTGAGATGTGTTCTGTATTCATTGCTTGTGCTTGTTATTTTCAG
-241 CAGAGGCAGATGTTAATCAGAACAATGGATCTTCCTCAAAGCAGAAGAAGATGGTGGCGA
-181 CAGGCTCCAAGAGTGCAGAGGACCCAAAGAATGCCTGGAAGTCGCCCAGCACAGCTGAGC
-121 CAGCTGTTCGCCCACACTTAGTCCGCCTCTTTTCTCGTGACGCACCAGGACGTGAGGACA
-61 ACACCTTCAAAGACCGCCCATCGGAATCAGAAGACCTTCAGGCCATCCAAGAGGATAGCT
-1 CAGTTCCTGTGGAAGAGCTCAGATTTGATGGAAGTGCAGAAACTGTCCTCACAATCATCA +59
+1

FIGURE 5. The promoter sequences are linked to the 5′ end of the shark Po and MBP cDNA's, and each contains a transcription initiation site. A. Lane 1, 1 kb ladder marker; lane 2, the 315 bp band, marked by an arrowhead, from Po using Po2 (complementary to the 5′ end of Po cDNA[1]) and PoC (complementary to the promoter region) as primers; lane 3, 123 bp ladder marker; lane 4, the 130 bp band, marked by an arrowhead, from MBP using MBP2 (complementary to the 5′ end of cDNA[1]) and MBPC (complementary to the promoter region) as primers. B. Lanes 1-4, the A, C, G and T sequencing reactions of M13mp18 using the universal primer to serve as size markers; lane 5, total RNA extended with Po2 (77 bp); lane 6, total RNA extended with MBP1.1 (137 bp, bracketed).

A



B



Chapter 3

DISCUSSION

ANALYSIS OF THE SHARK AND MOUSE PO AND MBP PROMOTERS

Myelinogenenis is one of the essential processes in the nervous system development of the vertebrates (Lemke et al. 1988a). Myelin surrounds and insulates the neuronal axons to facilitate the conduction of electrical impulses. The myelinating cells in the mammalian central nervous system (CNS) are oligodendrocytes which produce the major myelin structural proteins, proteoliped protein (PLP) and myelin basic protein (MBP). In contrast, the myelinating cells in the mammalian peripheral nervous system are Schwann cells which produce myelin protein zero (Po) and MBP. The fact that mammalian MBP is expressed both in the CNS and PNS whereas Po is expressed only in the PNS (Lemke et al. 1988b) make these promoters excellent examples for the study of nervous tissue-specific (CNS and PNS) and CNS vs. PNS specific gene expression. Chapter 2 describes the cloning and the sequence of the shark Po and MBP promoters. These shark promoters are particularly interesting for (i) their evolutionary separation from mammals by 400 million years (Romer 1970) and (ii) the expression of shark Po in the brain. Initially we identified no significant similarity between the shark and mammalian promoters (see Chapter 2). Here I will discuss the newly identified similarity between our shark and the mammalian promoters, recent in vitro transcription results from the literature on the mouse Po and MBP promoters, the implications of these data on CNS and

PNS specific gene expression, and future experiments aimed at further unwrapping the control of myelin gene expression.

Shark and Mouse MBP Promoters. Transgenic mouse experiments have already specified the 1.3 kilobase (kb) region of the MBP promoter immediately upstream from the transcription initiation site as necessary and sufficient for nervous tissue-specific expression of the MBP gene in vivo (Katsuki et al. 1988). Experiments with brain nuclear extracts reduced this essential region from 1.3 kb to 253 nucleotides (Tamura et al. 1989). Recent deletion studies using the mouse MBP promoter to transiently express B-galactosidase in a neuronal/glial hybrid cell line have further reduced this region to 208 nucleotides upstream from the transcription initiation sites that are sufficient to produce wild-type transcription levels (Miura et al. 1989). Interestingly, we observe no similarity between the mouse MBP promoter -208 to -175 sequence and the shark MBP promoter. It is possible that these 34 nucleotides are not essential for proper tissue-specific expression in the mouse and that a deletion to -175, had it been done, would also have yielded wild-type transcription levels.

Figure 1 has an alignment of the shark and mouse MBP promoters from -174 to +1 in the mouse. These sequences are colinear except for the C at position -121 in the mouse, not found in the shark, and the G in the shark between mouse -110 and -111. Including the above insertion and deletion shark and mouse MBP promoters are 76% similar (134

identities out of 176). By examining these sequences in closer detail it is possible to delineate several highly conserved regions which, due to their conservation between shark and mouse, could be essential for proper MBP promoter function. The first such regions, on the top line of Figure 1, are Boxes 1 and 2.

Boxes 1 and 2. The shark and mouse Box 1s are identical at 11 of 12 nucleotides and is separated by 10 poorly conserved nucleotides (approximately one complete turn of the DNA helix) from Box 2s which are identical at 12 of 12 nucleotides. The 12 nucleotide sequences that make up Boxes 1 and 2 also appear to be related since they both begin with GAC and have AA at or near the end. The overall similarity of these two boxes between shark and mouse (96%), their separation by a single turn of the helix spacer and their internal similarity to each other makes them possible cis-acting elements. A mouse MBP promoter deletion to -139, which deletes Box 1 and sequences downstream up to but not including Box 2, yields only 60% of wild-type transcription in a neuronal/glial hybrid cell extract (Miura et al. 1989).

Enhancer. The region labeled mouse MBP enhancer from mouse -127 to -106 is defined by DNase I footprinting analysis that reveals a cellular factor(s) that protects these sequences in neuronal/glial cell extracts. Also, the mouse MBP promoter region from -139 to -106, which includes Box 2 and the protected region, activates the SV40 promoter equally well in either orientation. In addition, a mouse

MBP promoter deletion to -118, which deletes box two and 42% of the enhancer, reduced transcription to 18% of wild-type in neuronal/glial cell extracts (Miura et al. 1989). Similar transcription experiments with the mouse MBP promoter in HeLa cell extracts show no effect with deletions up to -139 but only 10% or less of wild-type transcription levels with the deletion up to -118. The region protected in HeLa cells extracts in DNase I footprinting experiments is slightly larger on the 5′ end extending from -130 to -106 (Tamura et al. 1988). Tamura and coworkers went on to show that this same general region (-135 to -103) could function as a heterologous enhancer for the adenovirus major late promoter irrespective of its orientation. In addition, they demonstrated that the mouse MBP enhancer has one of the strongest nuclear factor one (NFI) binding motifs known among cellular genes and showed that purified NFI gave the same footprinting patterns as and competed with, crude HeLa cell extracts. NFI was originally found to stimulate replication of adenovirus DNA in vitro in HeLa cell extracts (Gronostajski 1987). To date, many promoters have been documented to contain NFI-binding sites as transcription elements including chick lysozyme, human c-myc, human IgM, albumin, mouse mammary tumor virus and cytomegalovirus (Tamura et al. 1988). Although these HeLa cell extract experiments do not address nervous tissue restricted MBP gene transcription, it is possible to combine their results

and with the results of tissue-specific experiments in the following manner.

While it has been shown that the mouse MBP enhancer binding factor, as identified in HeLa cells, is a ubiquitous protein among the vertebrates it is possible that a nervous system-specific NFI-like protein may interact with this enhancer in the CNS and/or PNS. Consistant with this notion, the JC virus (JCV) enhancer completed the MBP enhancer-footprinting in HeLa cells (Tamura et al. 1988). JCV is known to cause a variety of diseases in human and animal brains (Narayon et al. 1973). Tissue-specific expression of the JCV in glial cells (i.e., CNS oligodendrocytes) is directed by the viral enhancer-promoter (Kenney et al. 1984). Khalili and coworkers have demonstrated that a 27 nucleotide DNA sequence (GGATGGCTGCCAGCCAAGCATGAGGTC; Figure 2) in the viral enhancer binds to a cellular protein(s) in a tissue-specific manner (Khalili et al. 1988). The 14 base pair underlined sequence in the JCV enhancer is the consensus NFI binding motif which has strong NFI affinity (Gronostajski 1987). Figure 2 shows a comparison between the mouse MBP promoter, the corresponding shark MBP promoter sequence, and the JCV enhancer. The JCV similarity starts at -123 in mouse MBP and runs to -98 with an overall similarity of 63% to both the shark and mouse MBPs. However, JCV appears to be more closely related to the shark MBP promoter than to the mouse since both the JCV and shark MBP enhancers are missing a C

at -121 in the mouse and both have a G between -111 and -110 of the mouse sequence. This mouse insertion upstream and deletion downstream which are separated by 10 base pairs, effectively cancel each other out and have no effect on nervous tissue-specific expression, since both JCV and mouse MBP have a strong CNS tropism.

This quandary of how the general transcription factor NFI could direct tissue-specific expression of MBP in the CNS was clarified by Khalili and coworkers (Khalili et al. 1988). They identified the JCV enhancer binding protein by UV cross-linking experiments. The 27 nucleotide JCV enhancer interacts with an 85 kD protein in HeLa cell extracts and a 45 kD protein in human fetal brain extracts. Also, base substitution experiments with this enhancer using brain or HeLa cell extracts showed that several mutations completely abolished binding of the 45 kD brain-specific factor while having no effect on the binding of 85 kD HeLa-specific factor. Thus, the JCV enhancer which has similar footprinting patterns in HeLa cells and brain extracts recognizes distinct protein factors in these cells. It is probable that the mouse MBP enhancer and its corresponding sequence in the shark bind a CNS specific transcription factor in the brain but can enhance transcription in HeLa and other cells when separated from its surrounding sequences. The mouse MBP promoter deletion, footprinting and enhancement experiments and the JCV cross-linking experiments are summarized in Table 1.

GC and CAT Boxes. The regions downstream of the enhancer are the GC Box and the CAT Box which are both similar at 5 of 6 nucleotides (Figure 1). In the HeLa cells extracts the GC Box is essential, since alteration of the 2 Cs at -92 and -91 in the mouse promoter yield only 30% of wild-type transcription. However, mutations within the CAT box have little effect (Tamura et al. 1988). In neuronal/glial cell extracts the situation again appears to be different. Comparing deletions of the mouse MBP promoter tested in these glial hybrid extracts to -118 (which eliminates 40% of the enhancer sequence and its function) with those to -89 (which eliminates the critical CC for HeLa extracts from the GC Box and the downstream G) shows no additional reduction in transcription efficiency (Miura et al. 1989). The only difference between the mouse and shark GC Boxes is that first C to T substitution. An additional deletion from -89 to -75 (which eliminates to 4 nucleotides downstream of the CAT box) reduces transcription sevenfold (from 21% to 3% of wild-type) in the neuronal/glial cell extracts, but has little additional effect in HeLa cell extracts (see Table 1).

TATA Box Region. The next region is an area extending for 21 base pairs from -46 to -26 in the mouse MBP promoter which is 100% similar between shark and mouse and which contains the TATA box (Figure 1). It is interesting that mouse MBP promoter deletions to -53 still directed transcription in human fetal brain extracts 2.2 times

stronger than that observed in liver extracts (Tamura et al. 1989). This region from -53 to +1 in the mouse that directs preferential transcription in brain extracts is 80% similar to the shark sequence. No known factor binding sites have yet been identified between -53 and +1. However, there is some evidence suggesting core promoter-driven tissue-specific transcription (Dierich et al. 1987). It has also been suggested that eukaryotic TATA box factors are heterogeneous (Chen and Struhl 1988). The TATA box sequence of the MBP promoter does not fit the common consensus sequences (Roach et al. 1985). Thus, the extraordinary conservation of the shark and mouse MBP TATA box regions and their non-classical TATA sequence argues for a specific TATA box factor that appears before sharks split from the mammalian line of evolution over 400 million years ago.

Wobble Zone Region. The next to last region of interest which runs from -101 to -5 in the mouse and the corresponding sequence in the shark is termed the "wobble zone" (Figure 1). The term derives from 15 triplets out of 28 (54%) from -88 to -5 in the mouse where the only non-identities over 83 base pairs (except for -79, -49 and -14 in the mouse) between shark and mouse is in the third position of these 15 triplets. The wobble, normally seen in exons, is a function of the degeneracy of the genetic code where variations in the third base pair of a triplet could occur and still code for the same amino acid thereby conserving the protein during evolution. This wobble is not

in frame with exon 1 of mouse or shark MBP and neither protein has any known forms not coded for by known exons. Yet the idea that an exon for some protein could be located within the MBP promoter core seems extreme. However, translating and comparing the shark and mouse sequence within this wobble zone showed only 3 amino acid changes (Ile to Val, Asp to Glu and Glu to Asp), all conservative. Also, the wobbles are absent from regions of putative importance for promoter function (CAT and TATA boxes) where conservation is expected. It is interesting that both shark and mouse have a consensus splice acceptor site (-101 to -97 in mouse, underlined in Figure 1); (Mount 1982) at the putative 5′ end of this wobble zone. The mouse has a consensus donor site near the 3′ end of the wobble zone beginning at -14 (underlined) while the shark has 2 possible donor sites approximately 35 base pairs farther downstream starting at shark +2 and +8 (underlined). This wobble zone is currently being compared against Genebank for possible similarity to any known protein and/or gene.

Duplicated Region. The last region of interest in the MBP promoter is the shark promoter sequence from -19 to +1 which seems to be a duplication of the previous 25 base pairs in the shark (Figure 1). If one allows for two 1 base pair gaps in the -19 to +1 shark sequence (see Figure 1) this sequence is 61% similar (14 of 23) to mouse -24 to +1 and 70% similar (16 of 23) to shark -46 to -20. We also found a weak initiation around -20 in the shark (equivalent

to +1 in the mouse). The actual significance of this is unknown.

Po Promoter. While the MBP promoter activates transcription in the CNS and PNS, the Po promoter in the mouse is the only myelin gene exclusively transcribed in the Schwann cells of the PNS (Lemke et al. 1988). However, the Po promoter in the shark directs transcription in the shark CNS (Saavedra et al. 1989). Therefore, a comparison of the shark and mouse Po promoters could yield insights into CNS vs. PNS specific gene expression. Unfortunately, little similarity was found between the shark and mouse Po promoters. This is not surprising since the specificity of this promoter must have changed between the transition from water-based (bony fish) to land-based (amphibians) vertebrates when Po expression was turned off in the CNS and the Po protein was replaced by PLP (Waehneldt et al. 1985). Figure 3 shows those regions of the mouse and shark MBP promoters and the JCV enhancer that exhibit similarly to the shark and mouse Po promoters. Most notable is the 14 base pair core of the JCV enhancer (underlined) that makes up the consensus NFI binding site. This region has been shown to bind a brain-specific factor in human fetal brain (Khalili et al. 1988). Thirteen base pairs of the core are found in the shark Po promoter in approximately the same position as the shark MBP promoter but on the non-coding strand (-158 to -170). The shark Po promoter also has the G inserted between mouse MBP -110 and -111 seen in shark MBP and JCV.

Only one base pair (A at a position equivalent to -112 in mouse MBP) of the 13 is a variant not found mouse or shark MBP or JCV. The location of the core of this enhancer in the inverse orientation on the non-coding strand should, have little or no effect on its enhancement of transcription (Miura et al. 1989). The presence of this core sequence in shark Po and not mouse Po, and its similarity to the JCV enhancer core which binds a brain-specific factor further supports the case for a NFI-like brain (CNS) specific factor. This hypothesis will soon be tested by screening lambda gt11 cDNA expression libraries made from human fetal brain and 18 day old mouse brain mRNA with oligonucleotide probes representing the mouse MBP enhancer, the JCV enhancer and the shark MBP and Po enhancers.

A-Rich Region. There is also an 11/12 base pair region in shark and mouse Po, respectively, which includes two thirds of Box 2 and which ends one base pair before the JCV enhancer similarity to the MBPs begins (-135 to -125 mouse MBP; Figure 3). This 11 base pair region in shark Po is similar at 9 of 11 base pairs to shark MBP and at 7 of 11 base pairs to mouse MBP. This sequence in mouse Po is similar at 8 of 12 base pairs to mouse MBP, 9 of 12 base pairs to shark MBP and at 8 of 11 to shark Po. Interestingly in both shark and mouse Po this 11/12 base pair region is not continuous with the enhancer as it is in the MBP. For shark Po this 11 base pair region is located 47 base pairs upstream of the enhancer core on the coding

strand (-217 to -207 shark Po). For mouse Po this 12 base pair region is located 853 base pairs upstream of the transcription start site on the coding strand (-853 to -842 mouse Po). There is no enhancer identified in the mouse Po promoter yet.

GC Box. The last region of interest in shark Po is the GC box (-138 to -133 shark Po) which is identical in sequence and nearly so in its relative position downstream of the enhancer core to shark MBP.

The most interesting findings from the comparison of the promoters of shark Po and mouse Po were (i) the presence of only the enhancer core on the non-coding strand in the shark and its complete absence in mouse and (ii) the presence of the 11/12 base pair region of similarity in both but much further upstream than seen in the MBPs. Why is the core enhancer sequence in shark Po on the non-coding strand and absent in mouse Po? The simplest explanation for this would assume that the enhancer confers brain or CNS specificity on these genes. This would explain its presence in shark Po which is the most abundant shark CNS myelin protein and its absence in mouse Po whose promoter has lost its ability to direct transcription in the CNS. In contrast, the presence in both shark and mouse Po of this 11/12 base pair region, also seen in the MBPs, would argue that this is a nervous tissue-specific (both CNS and PNS) sequence. The validity of this hypothesis will soon be tested in promoter chimera experiments that will mix and

match elements of the shark and mouse Po and MBP promoters in Schwann cell and glial cell extracts to determine what cis-acting elements direct CNS vs. PNS expression.

THE PO TO PLP TRANSITION IN THE CNS

Glycosylated protein zero (Po) is the major hydrophobic myelin protein of the central nervous system (CNS) of both bony and cartilaginous fish and the peripheral nervous system (PNS) of land-based vertebrates (Waehneldt et al. 1989). In contrast, the unglycosylated proteolipid protein (PLP) is the major hydrophobic myelin protein in the CNS of all tetrapod classes (mammalia, aves, reptilia and amphibia)(Waehneldt et al. 1985). Although serving apparently similar structural functions, PLP and Po are dissimilar in their protein sequence, the number of transmembrane domains and in their exon-intron arrangements (Lemke et al. 1988b; Diehl et al. 1986). Therefore, there is a change in the major intrinsic myelin CNS protein at the transition of vertebrates from a water-based (bony fish) to a land-based (amphibian) existence. In Chapters 1, 2 and 3 showed that the Po protein was conserved in primary and putative secondary structureSduring this CNS transition, but that the promoter diverged greatly. Here I will examine the pertinent data presented in this thesis and the current literature on the structure, function and expression of Po and PLP to determine why Po was replaced by PLP as the major myelin CNS protein.

The Transition. Did the appearance of the first amphibians exactly coincide with the acquisition of PLP in their CNS myelin or was this a multistep evolutionary process during which pretetrapodal fishes were already endowed with a PLP-like protein instead of Po? Waehneldt and coworkers have identified two apparent intermediate evolutionary forms between our glycosylated Po seen in the shark CNS and the fatty acid esterified PLP seen in tetropods (Waehneldt et al. 1986 and 1989). The transition is from a glycosylated Po in sharks and bony fish to a glycosylated PLP-immunoreactive protein in African Lungfish (Protopterus Dipnoi), then to a unglycosylated PLP-immunoreactive protein in coelocanths (Latimeria chalumnae) to the fatty acid esterified PLP seen in tetrapods.

PLP vs. Po. A most interesting question is why nature should have dispensed with a protein as versatile as Po during the evolution of oligodendrocytes and CNS myelination. As we discussed in the introduction, the Po cytoplasmic domain is much like MBP in function and its extracellular domain is an immunoglobulion-like domain that facilitates self-adhesion. The significance of Po's exceptionally basic cytoplasmic domain is seen in shiverer mice, which do not make MBP. Here the Po domain accounts for the vast majority of cytoplasmic protein seen in this "normal" PNS myelin. In contrast, the CNS in shiverer mice shows PLP tightly compacting the extracellular space, but doing little or nothing for intracelluular compaction

(Readhead et al. 1987). Po also has a unique position within the immuunoglobulin superfamily. It is the only member of this family that contains a single immunoglobulin domain that is hypothesized to function through self-recognition. This suggests that the Po extracellular domain is minimally diverged from the primordial recognition molecules from which both the immune and nervous system are thought to have evolved (Hunkapiller and Hood 1989).

The PLP protein seems to have brought to the CNS of tetropods at least four possible benefits over Po. First, PLP containing myelin has a narrower extracellular apposition (by 5 angstroms or 15% of the total width) than Po-containing myelin (Kirschner et al. 1989). This may be due, in part, to the significantly higher hydrophobicity of PLP which would allow more residues to be immersed in the lipid bilayer thus conferring a more compact confirmation (Kirschner et al 1989). Interestingly, shark Po has extra, mostly hydrophobic amino acids at the N-terminus in its extracellular domain. However, the extracellular separation in shark CNS is actually 6 to 10 angstroms smaller than that seen in mouse sciatic nerve myelin. This tighter membrane apposition found in shark Po and further enhanced by the replacement of Po with PLP in tetrapods probably has some positive effect on electrophysiology and the insulating properties of myelin.

Second, PLP appears to have a more complex regulatory role. Po in mouse and shark has possible phosphorylation

sites for cyclic AMP-dependent kinase (Arg-Arg-X-Ser)(Krebs and Beavo 1979). This suggests that Po may have a signal transduction role across the myelin membrane. However, PLP has a complex regulatory role apparent from the mouse mutant jimpy. Jimpy mice have a point mutation preceding the fifth exon of the PLP gene (Nave et al. 1987). This single base mutation results in the splicing machinery skipping over this exon which in turn alters the reading frame, generating a premature stop codon and eventually leads to oligodendrocyte cell death. Jimpy mice tend to be more severely incapacitated than shiverer mice. This lack of PLP has broad pleiotropic effects on other proteins, lipid levels and on the MBP splicing pattern (Ikenaka et al. 1988).

Third, PLP may have an active role in the transport of materials in and out of the axon as evidenced by its proton ionophore properties in artificial bilayers (Lees and Brostaff 1984). Lastly, PLP is able to interact with MBP and could possibly create a translamellar "scaffold" comprised of both proteins (and possibly others) that could play a role in myelin compaction and maintenance (Edwards et al. 1989). Table 2 summarizes the comparison of PLP and Po highlighting the benefits of having PLP in the CNS. These potential functions associated with PLP involve all of its domains (extracellular, transmembrane and cytoplasmic) and help explain the extraordinary conservation observed between bovine, mouse and human PLP (99%) versus the other myelin

proteins (90% or less). For these benefits oligodendrocytes gave up Po, with its self-adhesive external domain and its cytoplasmic domain capable of substituting for MBP during pre-tetrapodal evolution. This was accomplished by alteration of the promoters CNS-specific cis-acting sequences while conserving the immunoglobulin-like external and the MBP-like cytoplasmic domains. It should be possible to follow this loss of oligodendrocyte-specific elements by using our promoter cloning technique (Chapter 2) to isolate the Po promoter from the transitional species, lungfish and coelocanths, for sequencing and comparative analysis.

THE SHARK AND MOUSE MBPS

In Chapter 1 we showed that the mouse 21.5 kD MBP can be generated from the shark 17 kD MBP by a 12-residue deletion at the N-terminus and a 25-residue deletion corresponding to mouse exon 2. The conserved regions between mouse and shark are moderately hydrophobic and have higher propensities for the beta strand and alpha helix, whereas the variable regions are mostly hydrophilic and have a higher propensity for beta-turn and coil (Saaveda et al. 1989). In mouse and rat the 14 kD form of myelin MBP predominates. The conserved regions of the rodent 14 kD form of MBP and shark MBP are in the first and third exons. Thus, these regions may contain the functional core of myelin. Based on concensus sequences of the phosphorylation site with c-AMP-dependent kinase (Krebs and Beavo 1979) or calmodulin-dependent kinase (Pearson et al. 1985), i.e.,

Arg-X-X-Ser(Thr), three regions in mouse MBP (residues 41-44, 130-133 and 186-189 in 21.5 kD) and two regions in shark (residues 100-103 and 132-135) may be possible phosphorylation sites (Inouye and Kirschner 1990). The residues 130-133 in mouse and 100-103 in shark are next to the encephalitogenic determinant Trp-X-X-X-X-Gln-Arg (Lys)(Carnegie 1971) which is also conserved in both animals. The Pro-Pro-Pro bridge sequence in mouse (122-124) is not present in shark. I will soon have some idea of the importance of these similarities and differences from the results of our shark MBP shiverer transgenic's (see Appendix).

As discussed earlier, the sequences related to mouse exon II are not present in shark. In mouse, exon II is present in the larger forms (17 and 21.5 kD) of MBP. These larger exon II containing forms may play a more important role in the early stages of myelin wrapping, since they are relatively abundant at this time and relatively rare or absent later on (Carson et al. 1983). In fact, mouse MBPs follow a general gradient from larger to smaller isoforms during development. The smaller forms of MBP are hypothesized to produce tighter compaction of the cytoplasmic space and hence better insulation and electrophysiology (Kirshner et al. 1989). Interestingly, there are 8 isoforms of shark MBP (20,17,16,14,10,8,6 and 4 kD) all of which react with both anti-shark and anti-mouse MBP antibodies (unpublished results). Four of these

isoforms (17,10,8 and 4 kD) were also sequenced and shown to be similar to mouse MBP (see Chapter 1). It is possible that these 8 MBP forms seen in shark may also be developmentally regulated. This is possible because all of our shark brain preparations were made from 5 or more animals at a time which varied in size and age. Therefore, any differences in developmental expression would not be observable. It would be interesting to look developmentally at shark MBP to see if the transition from larger to smaller MBP isoforms during development arose with the appearance of compacted "mammalian-like" myelin seen in sharks or evolved after sharks branched off from the mammalian line of evolution.

MYELIN GENE EXPRESSION IN OLIGODENDROCYTES AND SCHWANN CELLS

The expression of myelin genes differs depending upon the cells in which these genes are expressed. In oligodendrocytes, a terminally differentiated glial cell type, whose only apparent function is to elaborate myelin, the developmental expression of myelin genes parallels the differentiation of these cells (ffrench-constant and Raff 1986). The expression of the myelin genes in oligodendrocytes is largely unaffected by the presence or absence of neurons and shows little plasticity (Zeller et al. 1985). However, Schwann cell expression of myelin genes is almost entirely controlled by neurons and shows remarkable plasticity. The initial induction of these genes seems to require axonal contact. In fact, if this contact

is disrupted anytime after initial induction, the steady tate levels of the myelin gene products drop sharply (Politis et al. 1982). This effect was studied in vivo by transecting rat peripheral nerves and examining Po and MBP expression in Schwann cells distal to the transection site (where axons degenerate)(Trapp et al. 1988). Trapp and coworkers showed that transections done on 35 day old rats, which are past peak myelination, produced a 40-fold decrease in steady state Po and MBP levels distal to the cut, most of which occured within 5 days of transection. Interestingly, oligodendrocytes cultured together with Schwann cells in the absence of neurons continue to express the myelin genes while Schwann cell expression of these components is dramatically reduced (Mirsky et al. 1980). Therefore, oligodendrocytes exhibit constitutive whereas Schwann cells exhibit induced expression of the same myelin genes.

The molecular basis of the control of Schwann cell gene expression by axons is unknown. It has been proposed that an inducer molecule, expressed on the surface of a subset of axons, could interact with a Schwann cell-specific receptor and associated second messenger system and thereby trigger differentiation of the Schwann cell (Lemke 1988a). There is some evidence to suggest that cAMP may be the second messanger. Both the Po and MBP genes are inducable in cultured Schwann cells by cAMP-elevating drugs (dibutyl cAMP, forskolin and isobutyl methylxanthine). These drugs exhibit dose-response properties that mirror the relative

levels of expression of Po and MPB in myelinating Schwann cells (Gandelman et al. 1989). The verification that cAMP is the second messenger is an important area of future research.

In addition to the trans-acting regulatory mechanism already described, cis-acting enhancer/promoter elements on the myelin genes themselves are probably involved in their restricted expression in oligodendrocydes and Schwann cells. My comparative analysis (see Table 1 and Figures 1, 2 and 3) of the regulatory regions of the shark and mouse Po and MBP promoters and the JC virus enhancer defines several possible candidates. The presence of some of these elements on all the promoters except mouse Po (the only one of these genes not expressed in oligodendrocytes) suggests they are oligodendrocyte-specific. Whereas, the presence of other elements on all the promoters including mouse Po suggests that these are Schwann cell-specific. Having identified these elements by their conservation we can now begin to analyze the specificity of each of these possible transcriptional control elements in oligodendrocytes and Schwann cells using the appropriate promoter/enhancer varient attached to a reporter gene. In addition, the strong conservation between the CNS expressed genes (shark Po and MBP, mouse MBP and JC virus) means that the "oligodendrocytes" of the shark already possessed many of the essential cis-acting and most likely the trans-acting factors found in mammals.

References

Carnegie, P. Properties, structure and possible neuroreceptor role of the enaphalitogenic protein of human brain. **Nature 229:**25-28, 1971.

Carson, J., Nielson, M., and Barbarese, E. Developmental regulation of myelin basic protein in mouse brain. **Dev. Biol. 96:**485-492,1983.

Chen, W. and Strudhl, K. Saturation mutagenesis of yeast His3 "TATA Elements": genetic evidence for a specific TATA-binding protein. **Proc. Natl. Acad. Sci. USA 85:**2691-2695, 1988.

Diehl, H., Schaich, M., Budzinski, R., and Staffel, W. Individual exons encode the integral membrane domains of human myelin proteolipid protein. **Proc. Natl. Acad. Sci. USA 83:**9807-9811, 1986.

Dierich, A., Gaub, M., LePennic, J., Astinotle, D. and Chembon,P. Cell-specificity of the chicken ovalbumin and conalbumin promoters. **EMBO J. 6:**2305-2312, 1987.

Edwards, A., Ross, N., Ulmer, J., and Braun P., Interaction of myelin basic protein and proteolipid protein. **J. Neurosci. Res. 22:**97-102, 1989.

ffrench-Constant, C. and Raff, M. The oligodendrocyte type-2 astrocyte cell linage is specialized for myelination. **Nature 323:** 333-338, 1986.

Gandelman, K., Pfeiffer, S. and Carson, J. cAMP regulation of Po glycoprotein and myelin basic protein gene expression in semi-differentiated peripheral neurinoma cell line D6P2T. **Development 106:**389-398, 1989.

Gronostojski, R. Site-specific DNA binding of nuclear factor I: effect of the spacers region. **Nuc. Acids Res. 15:**5545-5558, 1987.

Hunkapiller, T. and Hood, L. Diversity of the immunoglobulin gene superfamily. **Adv. Immunology 44:**1-63, 1989.

Ikenaka, K., Furuichi, T., Iwasaki, Y., Moriguchi, A., Okano, H. and Mikoshiba, K. Myelin proteolipid protein gene structure and its regulation of expression in normal and jimpy mutant mice. **J. Molec. Biol. 199:**587-596, 1988.

Inouye, H. and Kirschner, D. Phylogentic aspects of myelin structure. In G. Jeserich, T. Waehneldt, H. Althaus

(ed.) Cellular and Molecular Biology of Myelination. NATO Advanced Research Workshop, Springer-Verlag, in press, 1990.

Katsuki, M., Sato, M., Kimura, M., Yokoyama, M., Kolayoshi, K. and Nomura, T. Conversion of normal behavior to shiverer by myelin basic protein antisense cDNA in transgenic mice. **Science 241:** 593-595, 1988.

Kenney, S., Natorojan, V., Strike, D., Khoury, G. and Salzman, N. JC virus enhancer-promoter active in human brain cells. **Science 226:**1337-1339, 1984.

Khalili, K., Rappaport, J. and Khoury, G. Nuclear factors in human brain cells bind specifically to JCV regulatory region. **EMBO J. 7:**1205-1210, 1988.

Kirshchner, D., Inouye, H., Ganser, A. and Mann, V. myelin membrane structure and composition correlated: A phylogenetic study. **J. Neurochem. 53:**1599-1609, 1989.

Krebis, E. and Beavo J., Phosphorylation-dephosphorylation of enzymes. **Ann. Rev. Biochem 48:**923-959, 1979.

Lees, M. and Brostoff, S. Proteins of Myelin. In P. Morell (ed.) Myelin-2nd ed., Plenum, New York, 1984, pp. 197-217.

Lemke, G. Unwrapping the genes of myelin. **Neuron 1:**535-543, 1988a

Lemke, G., Lamar, E. and Patterson, J. Isolation and analysis of the gene encoding peripheral myelin protein zero. **Neuron 1:**73-83, 1988b.

Mirsky, R., Winter, J., Abney, E., Pruss, R., Gavrilovic, J. and Raff, M. Myelin-specific proteins and glycolipids in rat Schwann cells and oligodendrocytes in culture. **J. Cell Biol. 84:** 483-494, 1980.

Miura, M. et al., The promoter element of mouse myelin basic protein gene function efficiently in NG 108-15 neuronal glial cells. **Gene 75:**31-38 1989.

Mount, S. A catalog of splice junction sequences. **Nuc. Acids Res.10:**459-472, 1982.

Narayan, O., Penny, J., Johnson, R., Herndon, R. and Weiner, L. **New Engl. J. Med. 289:**1278-1282, 1973.

Nave, K., Bloom, F. and Milner, R., A single nucleotide difference in the gene for myelin proteolipid protein defines the jimpy mutation in the mouse. **J. Neurochem. 49:**1873-1877, 1987.

Pearson, R., Woodgett, J., Cohen, P. and Kemp, B. Substrate specificity of a multifunctional calmodulin-dependent protein kinase. **J. Biol. Chem. 260:**14471-14476, 1985.

Politis, M., Sternberger, N., Ederle, K. and Spencer, P. Studies on the control of myelogenesis. **J. Neurosci. 2:** 1252-1266, 1982.

Readhead, C., Popka, B., Takahashi, N., Shine, H., Saavedra, R., Sidman, R. and Hood, L. Expression of a myelin basic protein gene in transgenic shiverer mice: Correction of the dysmyelinating phenotype. **Cell 48:** 703-712, 1987.

Roach, A., Takahashi, N., Provtcheva, D., Ruddle, F. and Hood, L. Chromosomal mapping of mouse myelin basic protein gene and structure and transcription of the partially deleted gene in shiverer mutant mice. **Cell 42:**149-155, 1985.

Romer, A. In: The Vertebrate Body, Ed 4. Saunders Co., Philadelphia, 1970.

Tamura, T., Masayuki, M., Ikenaka, K. and Mikoshibu, K. Analysis of transcription control elements of the mouse myelin basic protein gene in HeLa cell extracts: Demonstration of a strong NF1-binding motif in the upstream region. **Nuc. Acids Res. 16 (24):**11441, 1988.

Tamura, T., Aoyama, A., Inoue, T., Miura, M., Okano, H. and Mikashiba, K. Tissue-specific in vitro transcription from the mouse myelin basic protein promoter. **Molec. and Cell Biol. 9 (7):**3122-3126 1989.

Trapp, B., Hauer, P. and Lemke, G. Axonal regulation of myelin protein mRNA levels in actively myelinating Schwann cells. **J. Neurosci. 8:**3515-3521, 1988.

Waehneldt, T., Matthieu, J. and Jeserich, G. Major central nervous system myelin glycoprotein of the African lung-fish (Protopterus dolloi) cross-reacts with myelin proteolipid protein antibodies, indicating a close phylogenetic relationship with amphibians. **J. Neurochem 46:**1387-1391, 1986.

Waehneldt, T., Molotha, J., Karin, N. and Molthieu, J. Phylogenetic examination of vertebrate central nervous system myelin proteins by electroimmunobloting. **Neurosci. Lett. 57:**97-102, 1985.

Waehneldt, T. and Molotka, J. Presence of proteolipid protein in coelacanth brain myelin demonstrates tetrapod affinities and questions a chondrichthyan association. **J. Neurochem. 52:**1941-1943, 1989

Zeller, N., Behor, T., Dubois-Dalcg, M. and Lazzarini, R. Timely expression of myelin basic protein gene in rat brain oligodendrocytes cultured in the presence and absence of neurons. **J. Neurosci. 5:**2955-2962, 1985.

Table 1 - Mouse MBP Promoter Experiments

	Tissue-Specific	Non-Specific
Cell Extracts Used	Neuronal/glial Hybrid (1)	HeLa (2)
Enhancer Region Sequence Protected in Footprints	-127 to -106	-130 to -106
Sequence Used to Measure Enhancement of Heterologous Enhancer-less Promoters	-139 to -106 (to SV40)	-135 to -103 (to Ad2-MLP)
Deletion Effects (% Wild-type)		
to -139	60%	100%
to -118	20%	10%
to -89	20%	1 to 2%
to -75	3%	>1%
to -53	2.2X Perferential Transcription in Brain vs. Liver Extracts (3)	Not Done
Protein Cross-linked to JCV Virus Enhancer (4)	45 kD	85 kD
Effect of 2 Different Point Mutations on JC Virus Enhancer Binding (4)	Abolishes	None

References:

1. Tamura et al. **Nuc. Acids Res. 16**:11441-11461, 1988.
2. Miura et al. **Gene 75**:31-38, 1989.
3. Tamura et al. **Molec. Cell Biol. 9**:3122-3126, 1989.
4. Khalili et al. **EMBO J. 7**:1205-1210, 1988.

Table 2 - Comparison of Po and PLP

	Po	PLP
Extracellular Domain	Immunoglobulin Domain Similarity Mouse and Shark (55% Similarity) (36A) Extracellular Space	----- ----- (31A) Extracellular Space
Transmembrane Domain	Single Membrane Spanning Domain Has Signal Sequence Mouse and Shark (22% Similarity)	3 Putative Membrane Crossings No Signal Sequence Proton Ionophore Properties in Artificial Bilayers
Cytoplasmic Domain	Mouse and Shark (55% Similarity) Substitutes for MBP in Shiverer Highly Basic 5 to 15% of Total Myelin as Cytoplasmic MBP	----- Associates with MBP Not Basic 30 to 40% of Total Myelin as Cyoplasmic MBP
Post-Translational Modifications	Glycosylated and Phosphorylated	Acylated
Conservation Between Bovine, Mouse and Human	Approximately 90%	99%
Found in the CNS	No	Yes
Found in the PNS	Yes	Yes (Not Incorporated into Myelin Membrane)
Regulatory Function	Possible Prosphorylation Site for c-AMP Dependent Kinase	Jimpy - Pleiotropic Effect on Other Proteins and Lipid Levels and MBP Splicing Pattern

Figure 1. **Comparison of the Mouse and Shark MBP Promoters.** The sequences are colinear except for a C in mouse at position -121 and G in shark between mouse positions -111 and -110. The mouse MBP sequence is on top. The (*) in the shark sequence denote identities at that position with the mouse. The regions of strongest similarity (Boxes 1 and 2, Enhancer, GC Box, CAT Box and TATA Box) are defined and bracketed above the mouse sequence and underlined below the shark sequence. The mouse sequences underlined are the GAC's (-161 to -159 and -139 to -137) and AA's (-151 to -150 and -130 to -129) that begin and end Boxes 1 and 2, respectively, a concensus shared splice aceptor site (-101 to -97), a concensus mouse splice donor site (-14 to -10) and two potential shark splice donor sites (+2 to +6 and +7 to +11). The wobble zone is defined by a bracket underneath the shark sequence (from -88 to -5 of the mouse sequence). See the text for additional details.

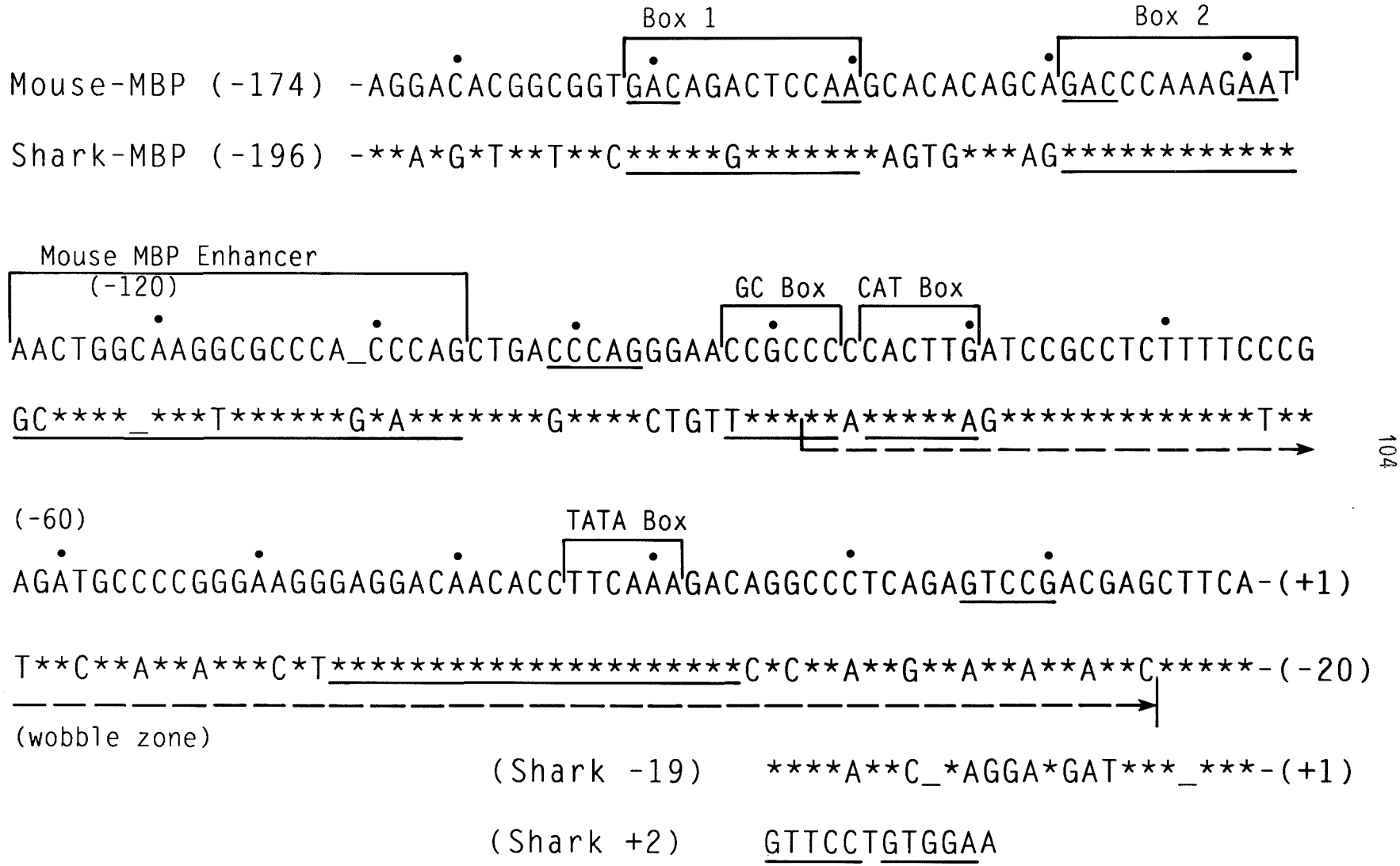

Box 1
Box 2
Mouse-MBP (-174) -AGGACACGGCGGTGACAGACTCCAAGCACACAGCAGACCCAAAGAAT
Shark-MBP (-196) -**A*G*T**T**C*****G*******AGTG***AG**********
Mouse MBP Enhancer
(-120)
GC Box
CAT Box
AACTGGCAAGGCGCCCA_CCCAGCTGACCCAGGGAACCGCCCCCACTTGATCCGCCTCTTTTCCCG
GC****_***T******G*A*******G****CTGTT******A*****AG**********T**
(-60)
TATA Box
AGATGCCCCGGGAAGGGAGGACAACACCTTCAAAGACAGGCCCTCAGAGTCCGACGAGCTTCA-(+1)
T**C**A**A***C*T**********C*C**A**G**A**A**A**C*****-(-20)
(wobble zone)
(Shark -19) ****A**C_*AGGA*GAT***_***-(+1)
(Shark +2) GTTCCTGTGGAA

Figure 2. **Comparison of the JC Virus Enhancer with the Mouse MBP Promoter Enhancer and its Corresponding Shark Promoter Sequence.** The (*) mark identifies the mouse MBP sequence. Gaps in shark MBP and JC virus at mouse -121, in mouse between -111 and -110, and in mouse and shark between mouse -105 and -104 are added to maintain alignment. The 14 base pair sequence of JC virus that is underlined is a NFI concensus binding site.

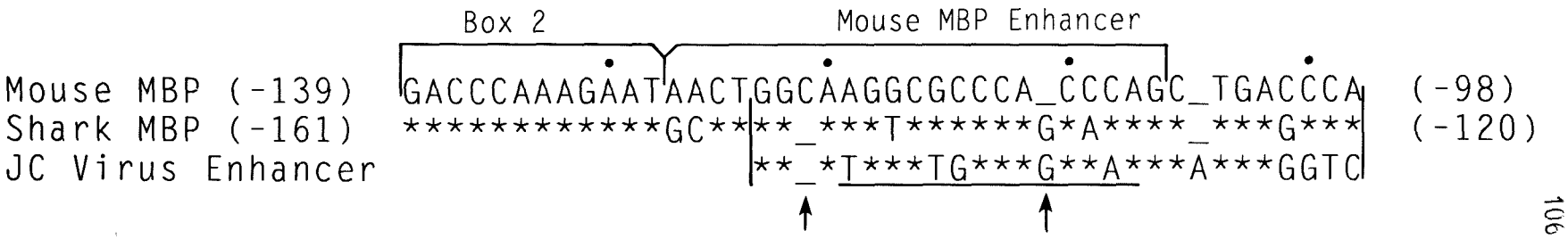
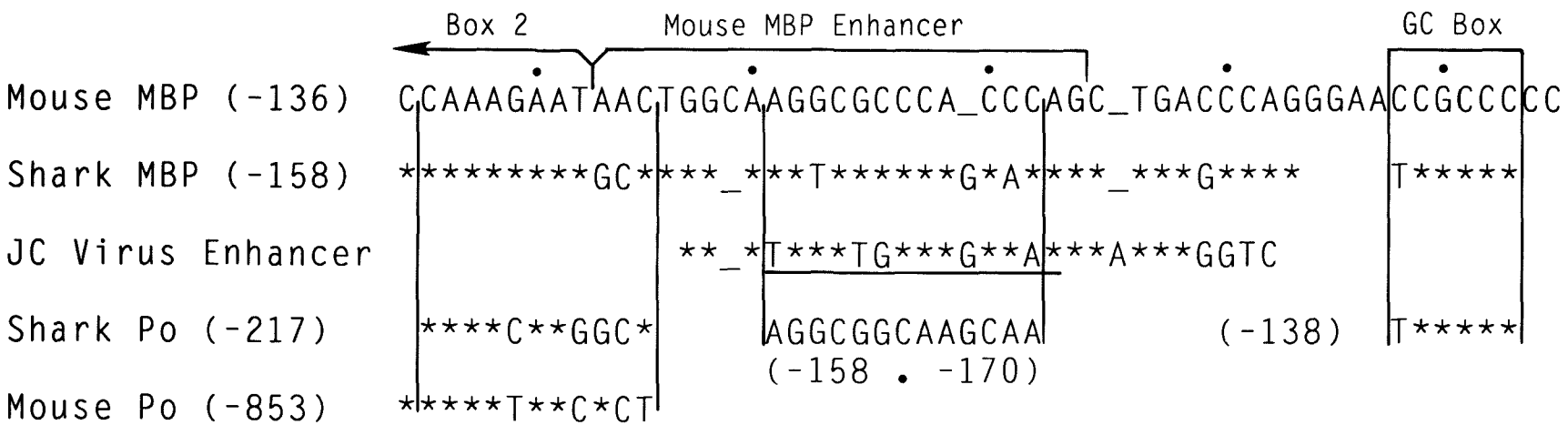

Box 2
Mouse MBP Enhancer
Mouse MBP (-139)
GACCCAAAGAATAACTGGCAAGGCGCCCA_CCCAGC_TGACCCA
(-98)
Shark MBP (-161)
**********GC**
_*T******G*A****_***G***
(-120)
JC Virus Enhancer
_*T*TG***G**A***A***GGTC

Figure 3. **Comparison of the Mouse and Shark MBP Promoters, the JC Virus Enhancer and the Mouse and Shark Po Promoters.** The (*) indicates identities to the mouse MBP promoter sequence. Gaps were added to maintain alignment (See Figure 2 legend for details). The 14 base pair NFI concensus binding site is underlined in the JC virus enhancer sequence. From upstream to downstream the 11/12 base pair region, the enchancer core and the GC Box are aligned with vertical lines to facilitate comparison.


Box 2
Mouse MBP Enhancer
GC Box
Mouse MBP (-136) CCAAAGAATAACTGGCAAGGCGCCCA_CCCAGC_TGACCCAGGGAACCGCCCCC
Shark MBP (-158) *********GC*****_****T******G*A*****_***G*****T*****
JC Virus Enhancer **_*T***TG***G**A***A***GGTC
Shark Po (-217) ****C**GGC* AGGCGGCAAGCAA (-138) T*****
(-158 . -170)
Mouse Po (-853) *****T**C*CT

Appendix I

SHIVERER TRANSGENIC MICE BEARING THE SHARK MYELIN BASIC PROTEIN GENE

Myelin in the vertebrate nervous system is a multilamellar membrane structure that wraps around axons to provide electrical insulation and facilitate the rapid conduction of nerve impulses. The major myelin proteins in the mammalian central nervous system (CNS) are proteolipid protein (PlP) and myelin basic protein (MBP). In contrast, the major myelin proteins in the mammalian peripheral nervous system (PNS) are protein zero (Po) and MBP. Sharks (class chondrichthyes) are the most primitive vertebrates that express compacted mammalian-like myelin. Yet sharks diverged from other vertebrates about 400 million years ago. Therefore, by analyzing shark myelin and comparing that data to what is already known about mammalian myelin it should be possible to learn not only about shark myelin and its evolution, but also about what is important for the proper function and expression of mammalian myelin. Based on this reasoning, we extracted and sequenced the two major CNS proteins from the shark **Heterodontus francisci**. The sequence data showed that the two major proteins of the shark CNS are similar to mammalian Po and MBP. We cloned and sequenced the cDNAs that code for shark Po and MBP. The translated sequences of these shark proteins are approximately 45% similar overall to their mammalian counterparts. However, both shark proteins share domains or regions with their mammalian counterparts that exhibit much greater similarity. In addition, the predicted secondary structures of the shark and mammalian proteins is strikingly

similar. These conservations in primary and secondary structures raised the interesting possibility that those regions or domains in the shark myelin proteins may be sufficiently conserved over 400 million years of evolution to function in mammals. The ability to introduce these shark functional genes into mice by germ-line transformation would allow us to test this hypothesis.

The most prevalent technique for introducing genes into the germ line is direct microinjection of cloned DNA into pronuclei of fertilized eggs (Palimiter and Brinster 1986). The animals are generally called "transgenic" and the foreign genes are referred to as transgenes. In these mice integration usually occurs prior to DNA replication with about 70% of the transgenic mice carrying the transgene in all of their cells, including the germ cells. In the remaining 30% of the transgenic mice, integration appears to occur after one or more rounds of replication, thus the transgenes are only in a fraction of the cells. In general, the foreign DNA is stably transmitted for many generations with no apparent rearrangement.

Shiverer mice have a deletion in the structural gene for MBP from the second intron to 2 kb 3′ of the last exon (Roach et al. 1985). Phenotypically, the MBP deletion in homozygous shiverer mice produces animals which initially appear normal, but by 2 weeks after birth begin to develop a shivering motion that increases in severity and is eventually accompanied by paralysis and seizures. These

mice, which produce little or no MBP cDNA or protein and cannot properly condense their myelin, are fertile but die prematurely between 50 and 100 days (Bird et al. 1977). Therefore, if one were able to appropriately express the shark MBP gene in shiverer mice one could test whether the observed similarities in primary and secondary structure between the shark and mouse MBPs translates into any functional similarity in vivo.

The DNA construct used for this experiment contains the mammalian MBP promoter driving transcription of the shark MBP cDNA followed by a SV40 fragment which contains a splice and a polyadenylation site all inserted into the vector pUC18 (Figure 1). The approximately 3.2 kilobase (kb) Eco R1 mouse MBP promoter fragment contains all the essential sequences necessary for the tissue-specific expression in the mouse of the shark MBP cDNA sequence. In fact, it was recently shown the only 1.3 kb of this promoter is required for biologically normal expression in transgenic mice (Kimura et al. 1989). The approximately 1.7 kb shark cDNA fragment contains the full 5'-untranslated region, the entire coding region for the most abundant 17 kilodalton (kD) shark MBP protein and a large 3' untranslated region without the shark polyadenylation site. The final downstream fragment contains a SV40 splice site which has been shown to be essential for the proper expression of some genes (Buchman and Berg 1988) followed by a SV40 polyadenylation site.

To date using this construct, one founder (first generation) shiverer mouse which contained the integrated construct was produced in August. Unfortunately, this mouse died before it was able to be bred. In mid November nine new transgenic mice were born which are now being analyzed using the polymerase chain reaction (PCR) to verify that the construct is integrated.

Based on what is known about shiverer transgenic mice and shark and mammalian myelin, it is possible to speculate on the expected results. First, this construct will almost certainly exhibit tissue-specific expression and developmental regulation as have several other transgenic constructs driven by this exact mouse MBP promoter fragment (Carol Readhead, personal communication). Second, only one form of shark MBP (17 kD) should be produced by this shark cDNA, as all the variation yet seen in mouse MBPs arises solely from the differential splicing of genomic DNA. Third, the level of shark MBP protein expression in these shiverer transgenics should comparable to that seen using mouse MBP. In the work by Readhead and coworkers (Readhead et al. 1987) the mouse MBP transgene produced about 8.5% of wild type MBP levels in heterozygous shiverers and about 20% of wild-type levels in homozygous shiverers. Fourth, shark MBP has an excellent chance of exhibiting some functionality. This is because the shark and mouse MBPs have similar hydrophilicity profiles, alpha helix and beta sheet potentials, distribution of positively charged amino

acids, and share a large central core of putative beta sheets which are believed to mediate the folding of the protein (Stoner 1984). This central core is 55% similar compared with the regions between the beta sheets (putative loops) which are only 34% similar (Saavedra et al. 1989). In addition, the fact that shiverer mice exhibit a "functional" PNS with Po and without MBP means that the cytoplasmic domain of Po can substitute, to a first approximation, for MBP (Lemke 1988). It is therefore possible to see how shark MBP which is significantly more related to mouse MBP could perform this function in the CNS of the shiverer mice. Lastly, although a full phenotypic recovery of the shiverer mouse is possible some intermediate results could be obtained. In the shiverer transgenic mouse a loss of the shivering phenotype and a tightly compacted CNS myelin is observed. In the shiverer shark transgenic some gradiation of the phenotype between a fully shivering and normal may be observable which should correspond to the extent and degree of myelin compaction seen through the electron microscope. Whatever the results, we will learn how the conservation of the primary and secondary structure of these MBPs, over 400 million years of evolution, affects their function in vivo.

References

Bird, T., Farrel, D., and Sumi, S., Genetic developmental myelin defects in "shiverer" mice. **Trans. Am. Soc. Neurochem. 8:**153, 1977.

Buchman, A., and Berg, P., Comparison of intron-dependent and intron-independent gene expression. **Molec. Cell. Biol. 8:**4395-4405, 1988.

Kimura, M., Sata, M., Akatsuka, A., and Katsuki, M. Restoration of myelin formation by a single type of myelin basic protein in transgenic shiverer mice. **Proc. Natl. Acad. Sci. 85:**5661-5665, 1989.

Lemke, G. Unwrapping the myelin genes. **Neuron 1:**535-543, 1988.

Palmiter, R., and Brinster, R., Germ-line transformation of mice. **Ann. Rev. Genet. 20:** 465-499, 1986.

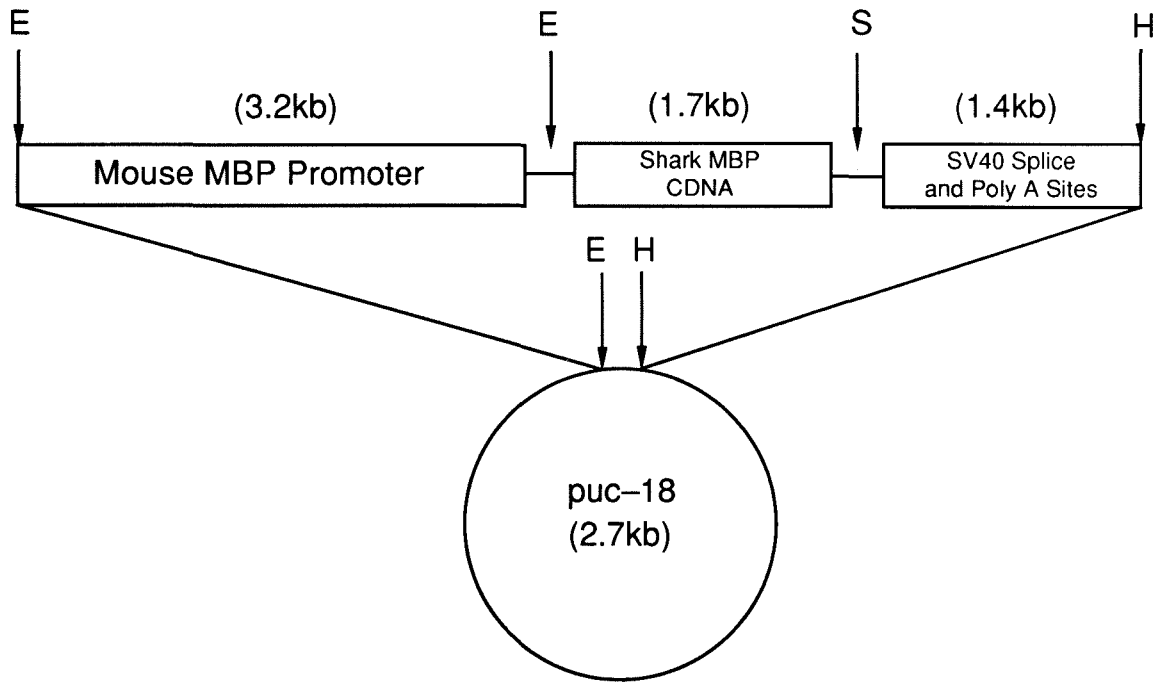
Readhead, C., Popko, B., Takahashi, N., Shine, D., Saavedra, R., Sidman, R., and Hood, L. Expression of myelin basic protein gene in transgenic shiverer mice: Correction of the dysmyelinating phenotype. **Cell 48:** 703-712, 1987.

Roach, A., Takahashi, N., Provtcheva, D., Ruddle, F., and Hood, L. Chromosomal mapping of mouse myelin basic protein gene and structure and transcription of the partially deleted gene in shiverer mice. **Cell 42:** 149-155, 1985.

Saavedra, R., Fors, L., Aebersold, R., Arden, B., Horvath, S., Sanders, J. and Hood, L. The myelin proteins of the shark brain are similar to the myelin protein of the mammalian peripheral nervous system. **Jour. Molec. Evol. 29:**149-156, 1989.

Stoner, G. Predicted Folding of B-structure in myelin basic protein. **J. Neurochem. 43:**433-447, 1984.

FIGURE 1. DNA constructs used for expression of shark MBP in Shiverer mice. The restriction site used to generate this construct are shown: (E) - EcoR1, (H) - HindIII and (S) - Sph1


E
E
S
H
(3.2kb)
(1.7kb)
(1.4kb)
Mouse MBP Promoter
Shark MBP
CDNA
SV40 Splice
and Poly A Sites
E H
puc–18
(2.7kb)