*Proc. Natl. Acad. Sci. USA*
Vol. 84, pp. 600–604, January 1987
Neurobiology

# Molecular cloning and primary structure of myelin-associated glycoprotein

(cDNA/neural cell adhesion molecule/p1B236/jimpy mutant)

MONIQUE ARQUINT*, JOHN RODER*†, LOO-SAR CHIA*, JIM DOWN†, DAVE WILKINSON*†, HAGAN BAYLEY‡, PETER BRAUN§, AND ROBERT DUNN†¶

*Division of Molecular Immunology and Neurobiology, Mt. Sinai Hospital Research Institute, 600 University Avenue, Toronto, ON, Canada M5G 1X5; †Department of Medical Genetics, Medical Sciences Building, University of Toronto, Toronto, ON, Canada M5S 1A8; ‡Howard Hughes Medical Institute, Center for Neurobiology and Behavior, Columbia University, New York, NY 10032; and §Department of Biochemistry, McGill University, 3655 Drummond Street, Montreal, PQ, Canada H3G 1Y6

*Communicated by Norman Davidson, October 1, 1986 (received for review August 28, 1986)*

ABSTRACT **Myelin-associated glycoprotein (MAG) may play a role in the cellular interactions leading to myelination. Using monoclonal antibodies and conventional antisera against MAG, we have isolated a cDNA clone from an expression library prepared from rat brain mRNA. The identity of the clone was confirmed by the exact match between its nucleotide sequence and two peptide sequences of 13 and 9 amino acids that we obtained by Edman degradation of two CNBr fragments of MAG. The cDNA clone hybridized to two size species of mRNA in rat approximately 3.5 kilobases in length. These mRNAs were present in brain but not liver and were expressed most abundantly at the time of active myelination (day 14). The mRNA for MAG was present at barely detectable levels in hypomyelinating jimpy mice compared to normal littermate controls. Therefore the MAG cDNA clone is both brain and myelin specific. DNA sequence analysis revealed that our MAG cDNA was derived from the same mRNA as clone p1B236, a randomly selected, brain-specific, partial cDNA isolated by Sutcliffe *et al.* [Sutcliffe, J. G., Milner, R. J., Shinnick, T. M. & Bloom, F. E. (1983) *Cell* 33, 671–682]. Analysis of the predicted protein sequence suggests that MAG has a long extracellular domain (499 amino acids), followed by a short transmembrane segment (20 amino acids) and an intracellular carboxyl-terminal domain (90 amino acids). The molecule has several glycosylation sites, three internal repeats homologous to a repeat in the neural cell adhesion molecule (N-CAM), and sites for phosphorylation near the carboxyl terminus. The primary structure reported here provides a molecular framework for further investigations into the function of the MAG molecule.**

Myelin-associated glycoprotein (MAG) is a 100-kDa integral membrane protein of low abundance (<1% in brain) that is found in both central and peripheral nervous system myelin (1). During development, MAG appears on oligodendrocytes and Schwann cells just prior to the formation of the myelin sheath (2). The localization of MAG to the periaxonal region in adult myelin (2, 3), although controversial (4), suggests that MAG may participate in neuron–glial cell interactions that lead to myelination. In tissue culture, MAG and other surface molecules do seem to provide an adhesion system for neuron–oligodendrocyte interactions (5). This hypothesis is supported by the observations that (*i*) anti-MAG autoantibodies, which appear in some plasma cell dyscrasia patients, may lead to demyelination (6), and (*ii*) mouse mutants that are defective in the production or maintenance of myelin also have reduced levels of MAG protein (7).

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

To examine the genetic control over MAG expression and its role in the development of the myelin sheath, we report here on the isolation and identification of a cDNA clone for MAG. Since the abundance of MAG mRNA in brain is low (8) and since no protein sequence was available at the outset, we chose a strategy of expression cloning in the phage vector λgt11, using polypeptide-specific, anti-MAG monoclonal antibodies for selection (9).

## EXPERIMENTAL PROCEDURES

**Protein Isolation and Sequencing.** Myelin was prepared from frozen rat brains and crude MAG was extracted by the lithium/iodosalicylate/phenol procedure (10). Purified MAG was obtained by preparative electrophoresis of crude MAG on $NaDodSO_4$/10% polyacrylamide slab gels followed by electroelution. Purified MAG (100 μg) was dissolved in 60% (wt/vol) formic acid and allowed to react with 30% (wt/vol) CNBr (Eastman) overnight at room temperature in the dark. After lyophilization, fragments were purified by preparative electrophoresis on $NaDodSO_4$/15% polyacrylamide gels, electroeluted, and sequenced by using the gas-phase protein sequencing technique (11) on an Applied Biosystems model 470A sequenator (Applied Biosystems, Foster City, CA).

**Isolation of MAG Clones.** Two cDNA expression libraries in the λgt11 vector were prepared from rat brain, one from membrane-bound polysomal RNA and the other from total poly(A)$^+$ RNA (V. Auld, J. Marshall, and R.D., unpublished data). Approximately 600,000 phages from each library (which contained $>10^6$ independent cDNAs) were screened with a mixture of the two monoclonal antibodies GenS1 and GenS3, generously provided by N. Latov (Columbia University, New York) (3), using the procedures described by Young and Davis (9). Antibody-reactive phages were plaque purified. Fusion proteins were tested for reaction with polyclonal anti-MAG sera preabsorbed with *Escherichia coli* lysate coupled to Sepharose 4B.

**DNA Sequence Analysis.** DNA was prepared from purified phage and digested with *Eco*RI, and the resulting fragments were ligated into M13mp18 phage (12). DNAs with sequential deletions of the DNA inserts were prepared (13) and the DNA sequences were obtained by using the chain-termination procedure (14).

**RNA Analysis.** RNA was isolated from rat tissues by lysis in guanidinium hydrochloride followed by precipitation with sodium acetate, glyoxylation, and electrophoresis (3.5 V/cm) on 1% agarose gels. The RNA was blotted onto nitrocellulose membranes and hybridized to $^{32}$P-labeled probes ($10^6$ cpm/ml) followed by washing and exposure to x-ray film for

Abbreviation: MAG, myelin-associated glycoprotein.

¶To whom reprint requests should be addressed.

16 hr at −70°C with enhancing screens. The size of RNA was estimated from 18S and 28S ribosomal RNA markers electrophoresed in parallel as indicated.

## RESULTS

**Isolation of MAG cDNA Clones.** Two murine monoclonal antibodies, GenS1 and GenS3 (both IgG1κ), directed against different polypeptide epitopes on the human MAG protein, were kindly provided by N. Latov (3). In our hands these antibodies reacted with a single 94-kDa protein in crude extracts of rat brain, and no binding was evident in liver extracts. On the basis of these results, a mixture of the GenS1 and GenS3 monoclonal antibodies was chosen as the specific probe for the MAG protein, and the screening system was optimized to detect as little as 0.5 ng of purified MAG. Fifteen positive phage plaques were identified upon screening the two rat brain cDNA libraries. Phages were plaque purified and fusion proteins were retested with two different rabbit polyclonal antisera that had been raised against MAG. Fusion proteins from only two of the phages, pMAG1.2 and pMAG6.2, reacted with the polyclonal antisera. Both cDNA clones cross-hybridized to one another but not to any of the other clones tested.

**Identification of MAG mRNAs.** Radiolabeled probes were prepared from cDNA clones and hybridized to blots of rat brain RNA. pMAG1.2 reacted with two closely migrating rat brain mRNAs approximately 3.5 kilobases (kb) in size as shown in Fig. 1, lane 1. No hybridization was observed to liver RNA (lane 3). Similar results were obtained with pMAG6.2. The abundance of RNA hybridizing to pMAG1.2 was greater in the developing rat brain at the time of active myelination (day 14) compared to mature rat brain where myelination has ceased (lane 2). These results demonstrate that pMAG1.2 and pMAG6.2 hybridized to brain-specific mRNAs of appropriate size and abundance to code for MAG. Other clones detected by the monoclonal or polyclonal antibodies did not meet this selection criterion.

To further delineate the specificity of our clones for myelin, we examined the effect of the hypomyelinating jimpy mutation on the levels of the putative MAG mRNA. As shown in Fig. 1, the pMAG1.2 cDNA probe detects a single 3.5-kb mRNA (lane 4) in normal myelinating mouse brain (day 14), and the amount of this mRNA is dramatically reduced in the homozygous jimpy mutant (lane 5), compared to its normal littermate control (lane 6). When these blots were stripped and rehybridized with a β-actin cDNA probe (15), no significant difference in the level of actin mRNA was detected (data not shown), thereby confirming that similar amounts of mRNA were present in each lane.

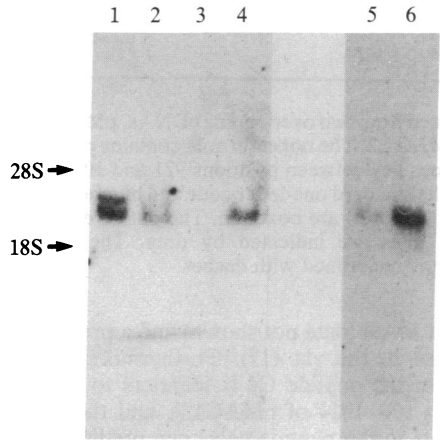


FIG. 1. RNA blot analysis of cDNA clones. Glyoxylated total RNA (25 μg) from 14-day-old rat brain (lane 1), mature rat brain (lane 2), rat liver (lane 3), 14-day-old mouse brain (lane 4), jimpy (*jp/Y*) mouse brain (lane 5), or normal-appearing littermate mouse brain (lane 6) was electrophoresed, blotted onto nitrocellulose, and hybridized with $^{32}$P-labeled pMAG1.2. The positions of 18S and 28S ribosomal RNA markers are shown.

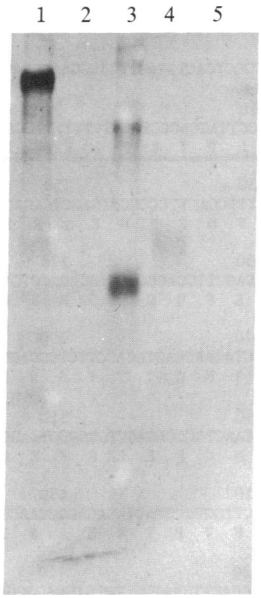


FIG. 2. Immunoreactive CNBr fragments of MAG. Gel-purified MAG was treated with CNBr and four fragments (C1–C4) were electroeluted from gels and subjected to a second electrophoresis followed by immunoblotting with monoclonal anti-MAG antibody (GenS3 or GenS1). Lane 1, gel-purified intact MAG; lane 2, C4; lane 3, C3; lane 4, C2; and lane 5, C1.

**Determination of MAG Peptide Sequences.** To provide an unequivocal identification of our cDNAs it was important to determine the amino acid sequence of at least a portion of the MAG protein and to relate this to nucleotide sequences in our clones. For the peptide analysis, gel-purified rat MAG protein was cleaved with CNBr and four of the resulting fragments, C1 (60 kDa), C2 (40 kDa), C3 (35 kDa), and C4 (22 kDa), were purified by preparative gel electrophoresis and were shown to react with rabbit anti-MAG sera on immunoblots. As shown in Fig. 2, two fragments, C2 and C3, also reacted with the two anti-MAG monoclonal antibodies, GenS1 and GenS3, used to isolate the cDNA clones. The four purified fragments were collected by electroelution and subjected to sequence analysis by Edman degradation. Fragments C2 and C4 yielded short stretches of reliable amino acid assignments. C2 gave the sequence Tyr-Ala-Pro-Xaa-Lys-Pro(Thr)Val-Xaa-Gly(Thr)Val-Val-Ala(Val), whereas C4 gave the sequence (Pro or Gly)(Ser)(Ser)Ile-Ser-Ala-Phe-Glu-Gly. The less definitive assignments are shown in parentheses.

**DNA Sequence of Phage 1.2.** The 1200-base-pair (bp) insert in pMAG1.2 was excised with *Eco*RI to yield two fragments, which were subcloned in phage M13mp18 and sequenced. A

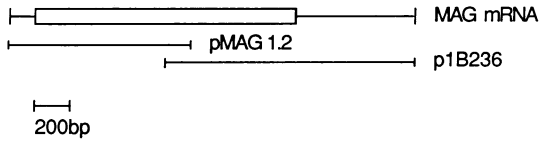


FIG. 3. Structure of MAG mRNA. The DNA inserts of pMAG1.2 and p1B236 (16) are shown below the deduced MAG mRNA. The 1878-nucleotide open reading frame is indicated by the open box. The mRNA is 2446 nucleotides in length, excluding the 3′ poly(A) tail.

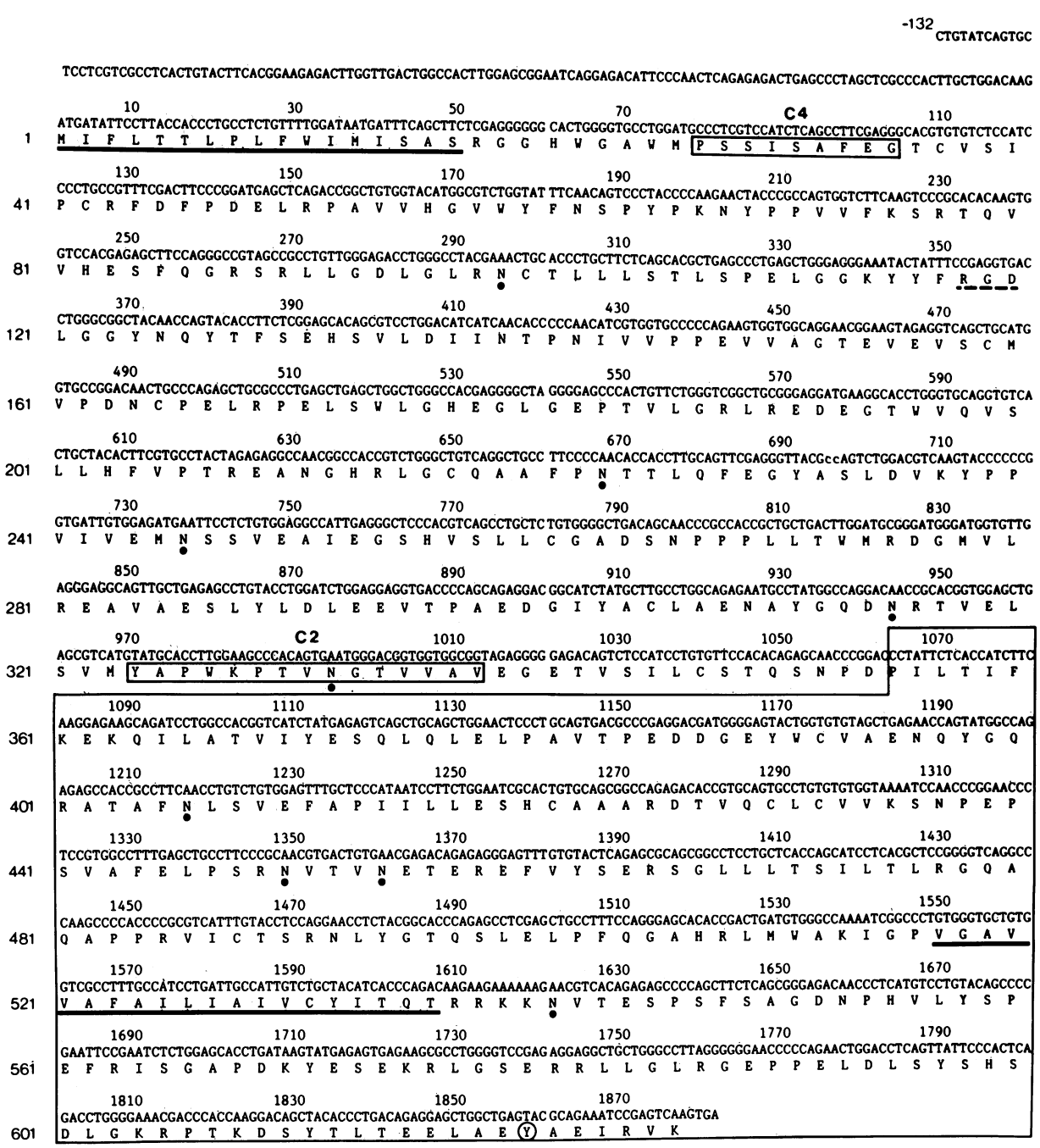


FIG. 4. Nucleotide sequence of the coding region for MAG. The sequence was obtained from two overlapping cDNAs, pMAG1.2 and p1B236, as shown in Fig. 3. The upper half of the figure contains the entire sequence of clone pMAG1.2. The boxed-in area contains part of the published sequence of the random cDNA clone p1B236 (16). The overlap between the two clones lies between positions 921 and 1062. The numbering of nucleotides starts at the putative translation start site. Amino acids are presented in the standard one-letter code and numbered in the left-hand margin. The amino acid sequences corresponding to the sequenced CNBr peptides C2 and C4 are boxed in. The putative signal peptide and transmembrane regions are underlined. Potential asparagine-linked glycosylation sites are indicated by dots. The potential tyrosine phosphorylation site is circled and the fibronectin recognition sequence Arg-Gly-Asp is underlined with dashes.

search of the Genbank‖ sequence data bank indicated that pMAG1.2 overlapped a previously isolated, randomly cloned, partial cDNA (1B236) (16) as shown in Fig. 3. Clone pMAG1.2 and random clone 1B236 overlap by 141 nucleotides between positions 921 and 1062. The combined sequences from the two clones provide a 1878-nucleotide open reading frame (Fig. 4). The amino acid composition of the predicted sequence was in approximate agreement with the composition obtained from our purified rat MAG (data not shown) and a previously published composition in the rat (17). Furthermore, the amino acid sequence of the peptide C2 is identical to that predicted by nucleotides 970–1014 of pMAG1.2, and the sequence of C4 corresponds to the peptide sequence predicted by nucleotide positions 79–105 (Fig. 4). The correlation between the amino acid and cDNA sequence data provides a definitive proof of the identity of the MAG cDNA.

‖National Institutes of Health (1986) Genetic Sequence Databank: Genbank (Research Systems Div., Bolt, Beranek, and Newman, Inc., Boston), Tape Release 40.0.

## DISCUSSION

We report here the isolation of cDNAs encoding part of the gene for MAG. Three lines of evidence support the identifi-

cation of our cDNA clones. First, the fusion proteins encoded by two clones, pMAG1.2 and pMAG6.2, reacted with two highly MAG-specific monoclonal antibodies recognizing different epitopes on the MAG polypeptide and with two hyperimmune rabbit sera, raised against gel-purified MAG. These two cDNA clones cross-hybridized. Second, the MAG cDNAs hybridized to two brain mRNAs, whose sizes were greater than the 2000 nucleotides required to code for the 70-kDa protein portion of the MAG glycoprotein. These two mRNAs were detected in brain but not liver and were more abundant in the actively myelinating rat brain than in the mature rat brain. From the observation that *in vitro* translation of MAG mRNAs yielded two proteins, p67 and p72, Frail and Braun (8) suggested that the MAG mRNA might be subject to differential splicing. Our observation of two MAG mRNAs in the blots of rat RNA and preliminary sequence data on additional cDNA clones support this suggestion.

A further test for the myelin specificity of our cDNAs was carried out by blot hybridization analysis of mRNAs from hypomyelinating jimpy mice. Although the primary genetic lesion for this defect probably lies in the gene for proteolipid protein (18) it has been shown by *in vitro* translation that MAG mRNA is poorly expressed in jimpy brain (17). This MAG deficiency is thought to result from a generalized lack of myelin mRNAs. Our blotting studies showed a dramatic reduction of mRNA hybridizing to cDNA pMAG1.2 in the jimpy strain (Fig. 1). This is consistent with the myelin origin of our cDNA clone.

The third and critical piece of evidence for the identification of the cDNAs relied on the determination of the amino-terminal amino acid sequences of two MAG peptides, C2 and C4, derived from MAG by CNBr cleavage. Confirmation that the peptides originated from MAG was provided by demonstrating that they reacted with anti-MAG polyclonal antisera and that one of these peptides (C2) also reacted with an anti-MAG monoclonal antibody (Fig. 2). Phenylthiohydantoin derivatives of amino acids were identified at 13 of the first 15 cycles of Edman degradation for C2 and 9 of 9 cycles for C4. These sequences matched the predicted amino acid sequence of the MAG cDNA at all assigned positions. The lack of assignment for the asparagine in peptide C2 was likely due to N-glycosylation because it lies within a canonical Asn-Xaa-Ser sequence for N-glycosylation sites. We also note that the DNA sequence correctly predicts a methionine residue positioned one residue ahead of the peptide sequence for both C2 and C4, as expected for CNBr cleavage. In summary, these data unambiguously demonstrate that the MAG peptides C2 and C4 are encoded by the MAG cDNA clone 1.2.

## OVERVIEW OF FEATURES

We have examined the structure of the predicted protein for possible insights into function. Translation of our cDNA clones and the clone from Sutcliffe yielded a single open reading frame of 1878 nucleotides, encoding 626 amino acids with a predicted molecular weight of 69,274, which matches estimates from the electrophoretic mobility of the nonglycosylated MAG precursor protein in polyacrylamide gels (8). The start codon ATG was selected since all reading frames in the upstream sequence generated stop codons. A hydrophobicity plot (19) revealed a hydrophobic region between residues 517 and 536 that could be a transmembrane segment. This segment contains a single cysteine residue, which may be the site of fatty acid attachment as described for the heavy chains of some HLA proteins (20). Since the amino terminus has a 17-residue hydrophobic leader segment for transmembrane insertion, this end of the molecule probably represents the extracellular portion of MAG. Several carbohydrate attachment sites and one tripeptide recognition sequence (Arg-Gly-Asp) for the fibronectin receptor (21) can be found in this region. The carboxyl-terminal 90 residues of MAG may be intracytoplasmic and contain several potential phosphorylation sites. Comparison of the human epidermal growth factor (EGF) receptor to the MAG sequence revealed a strong homology around the MAG tyrosine residue located 7 residues from the carboxyl terminus with the major site of autophosphorylation in the EGF receptor (tyrosine-1137) (22). In this sequence, 7 of 11 residues are conserved if a 2-residue gap is permitted. In the case of the EGF receptor, there is direct evidence that phosphorylation at this site regulates receptor function (22). We also note that potential sites for cAMP- and cGMP-dependent kinases and protein kinase C occur in the proposed MAG cytoplasmic tail at positions 575–582, 604–608, and 537–543 (Fig. 4).

A search of the gene bank revealed that part of our MAG cDNA between nucleotides 921 and 1062 was identical to a randomly isolated brain-specific cDNA clone designated 1B236 (16). Sutcliffe and coworkers have characterized protein 1B236 by using antibodies produced to synthetic peptides as deduced from the DNA sequence. These studies show that protein 1B236, like MAG, is a 100-kDa glycoprotein found in the nervous system (23).

Scanning the translated MAG sequence revealed three adjacent regions of internal homology in the extracellular domain (Fig. 5). Two of these regions, 246–331 and 332–416, were highly homologous, sharing 45% amino acid identity, whereas the third domain, 417–513, shared homology mainly in the amino-terminal 24 amino acid residues. Two cysteine residues, spaced approximately 45 residues apart, were conserved in all three repeats and could conceivably form intradomain disulfide bridges analogous to those seen in immunoglobulin domains (25). As shown in Fig. 5, the first two repeated domains in the MAG sequence show 29% and 25% amino acid identity, respectively, to an analogous repeat sequence found in the neural cell adhesion protein (N-CAM) (24). The occurrence of homologous, immunoglobulin-like domains in the extracellular regions of these molecules suggests a common adhesive mechanism for MAG and N-CAM. These polypeptide homologies indicate that the two proteins may have arisen from a common progenitor by gene duplication events. Hence seemingly diverse cell adhesion molecules in the nervous system might belong to a common cell-adhesion-molecule superfamily.

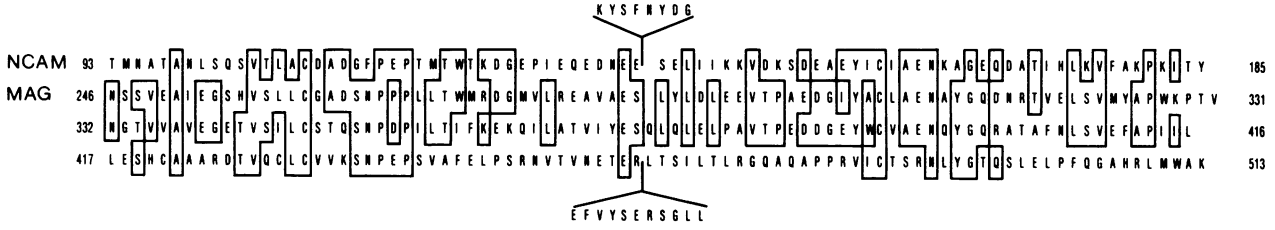


FIG. 5. Homology between repeated domains in MAG and neural cell adhesion molecule (N-CAM). Line 1 shows one of the four internal repeats of N-CAM (24). Lines 2, 3, and 4 contain the contiguous amino acid sequence of MAG between amino acids 246 and 513, aligned to show maximum internal homology. Amino acid positions are indicated in the left and right margins. Amino acid identities are boxed.

**Note Added in Proof.** The nucleotide sequence of an independent clone, pMAG6.2, confirms the sequence shown in Fig. 4 between positions −68 and 1716.

We thank Martha Garrett for technical assistance, Vanessa Auld for cDNA libraries, Norman Latov for monoclonal antibodies GenS1 and GenS3, Robert Lazzarini for rabbit anti-MAG antisera, and Don Frail, Vanessa Auld, and James Dunn for valuable discussions. This work was supported by grants from the Multiple Sclerosis Society of Canada. L.-S.C. and J.D. are fellows of the Multiple Sclerosis Society and Medical Research Council of Canada, respectively, and M.A. was supported by a McLaughlin Scholarship from Queen's University.

1. Quarles, R. H. (1984) *Dev. Neurosci. (Basel)* **6,** 285–303.
2. Steinberger, N. H., Quarles, R. H., Itoyama, Y. & Webster, H. F. (1979) *Proc. Natl. Acad. Sci. USA* **76,** 1510–1514.
3. Nobile-Orazio, E., Hays, A. P., Latov, N., Perman, G., Colier, J., Shy, M. E. & Freddo, L. (1984) *Neurology* **34,** 1336–1342.
4. Webster, H., Palkovits, C., Stoner, G. L., Favella, J., Frail, D. E. & Braun, P. (1983) *J. Neurochem.* **41,** 1469–1479.
5. Riopelle, R. J., McGarry, R. C. & Roder, J. C. (1986) *Brain Res.* **367,** 20–25.
6. Braun, P. E., Frail, D. E. & Latov, N. (1982) *J. Neurochem.* **39,** 1261–1265.
7. Frail, D. E. & Braun, P. E. (1985) *J. Neurochem.* **45,** 1071–1075.
8. Frail, D. E. & Braun, P. E. (1984) *J. Biol. Chem.* **259,** 14857–14862.
9. Young, R. A. & Davis, R. W. (1983) *Proc. Natl. Acad. Sci. USA* **80,** 1194–1198.
10. Quarles, R. M. & Pasnak, C. L. (1977) *Biochem. J.* **163,** 635–637.
11. Hunkapiller, M. W., Hewick, E. M., Dreher, W. J. & Hood, L. E. (1983) *Methods Enzymol.* **91,** 399–413.
12. Messing, J. (1983) *Methods Enzymol.* **101,** 20–78.
13. Henikoff, S. (1984) *Gene* **28,** 351–359.
14. Sanger, F., Nicklen, S. & Coulson, A. R. (1977) *Proc. Natl. Acad. Sci. USA* **74,** 5463–5467.
15. Nudel, V., Zakut, R., Shani, M., Newman, S., Levy, Z. & Yaffe, D. (1983) *Nucleic Acids. Res.* **11,** 1759–1771.
16. Sutcliffe, J. G., Milner, R. J., Shinnick, T. M. & Bloom, F. E. (1983) *Cell* **33,** 671–682.
17. Quarles, R. H., Barbarash, G. R., Figlewicz, D. A. & McIntyre, L. J. (1983) *Biochim. Biophys. Acta* **757,** 140–143.
18. Dautigny, A., Mattei, M. G., Morello, D., Alliel, P. M., Pham-Dinh, D., Amar, L., Arnaud, D., Simon, D., Mattei, J. F., Guenet, J. L., Jolles, P. & Avner, P. (1986) *Nature (London)* **321,** 867–869.
19. Kyte, J. & Doolittle, R. F. (1982) *J. Mol. Biol.* **157,** 105–132.
20. Kaufman, J. F., Krangel, M. S. & Strominger, J. L. (1984) *J. Biol. Chem.* **259,** 7230–7238.
21. Ruoslahti, E. & Pierschbacher, M. D. (1986) *Cell* **44,** 517–518.
22. Downward, J., Parker, P. & Waterfield, M. D. (1984) *Nature (London)* **311,** 483–485.
23. Malfroy, B., Bukhit, C., Bloom, F. E., Sutcliffe, J. G. & Milner, R. J. (1985) *Proc. Natl. Acad. Sci. USA* **82,** 2009–2013.
24. Hemperly, J. J., Murray, B. A., Edelman, G. & Cunningham, B. A. (1986) *Proc. Natl. Acad. Sci. USA* **83,** 3037–3041.
25. Hood, L., Kronenberg, M. & Hunkapiller, T. (1985) *Cell* **40,** 225–229.