## Two forms of 1B236/myelin-associated glycoprotein, a cell adhesion molecule for postnatal neural development, are produced by alternative splicing

(immunoglobulin super family/myelinogenesis/platelet-derived growth factor/P<sub>0</sub> protein)

Cary Lai<sup>†‡</sup>, Mary Ann Brow<sup>§</sup>, Klaus-Armin Nave<sup>†‡</sup>, Antonio B. Noronha<sup>¶</sup>, Richard H. Quarles<sup>¶</sup>, Floyd E. Bloom<sup>†</sup>, Robert J. Milner<sup>†</sup>, and J. Gregor Sutcliffe<sup>§</sup>

<sup>†</sup>Departments of Basic and Clinical Research and <sup>§</sup>Molecular Biology, Research Institute of Scripps Clinic, 10666 North Torrey Pines Road, La Jolla, CA 92037; <sup>‡</sup>Departments of Biology and Neuroscience, University of California, San Diego, La Jolla, CA 92093; and <sup>§</sup>Developmental and Metabolic Neurology Branch, National Institute of Neurological and Communicative Disorders and Stroke, National Institutes of Health, Bethesda, MD 20892

Contributed by Floyd E. Bloom, March 13, 1987

ABSTRACT The structures of two rat brain-specific 1B236 mRNAs, alternative splice products from a single gene regulated differently during postnatal brain development, were deduced from full-length cDNA clones. The 626- and 582amino acid-long encoded proteins are indistinguishable from two forms of myelin-associated glycoprotein, a cell adhesion molecule involved in axonal-glial and glial-glial interactions in postnatal brain development, particularly in myelination. The two proteins share a single membrane-spanning domain and a glycosylated N terminus but differ in the structures of their C termini. The N terminus consists of five domains related in sequence to each other and to immunoglobulin-like molecules, especially the neural cell adhesion molecule N-CAM, suggesting a common structure for cell adhesion molecules.

Rat brain protein 1B236 was originally defined by nucleotide sequence analysis of randomly selected cDNA clones of mRNAs expressed in adult rat brain but not detectable in liver or kidney (1, 2). Antisera to synthetic peptides corresponding to nonoverlapping regions of the partial 1B236 sequence detected a postnatally expressed 100-kDa rat brain protein containing  $\approx 30\%$  N-linked carbohydrate and proteolytic fragments derived from its C-terminal regions (2-5). During early postnatal development, 1B236 is expressed predominantly by oligodendrocytes in myelinating fiber tracts. Subsequently, there is gradual elaboration of the adult pattern in which 1B236 mRNA is detected predominantly in subsets of neurons in grey matter regions and the 1B236 protein is restricted to specific neuronal cell bodies and fibers (refs. 2, 4, and 6; G. A. Higgins, H. Schmale, F.E.B., M. C. Wilson, R.J.M., unpublished data).

We have now analyzed 1B236 expression more completely and report here the structures of two differently regulated, alternatively spliced 1B236 mRNAs that encode proteins with alternative C-terminal tails. The shared N-terminal region consists of five domains that are related in sequence to each other and to proteins of the immunoglobulin super family (8), especially the neural cell adhesion molecule N-CAM (9). We show that the 1B236 protein is indistinguishable from myelin-associated glycoprotein (MAG), a nervous system-specific glycoprotein of 100 kDa that contains  $\approx 30\%$ carbohydrate (10). MAG has been implicated in the interactions between myelinating cells and axons (10) and the formation and maintenance of the periaxonal space (11). In the peripheral nervous system, MAG appears to take over Schwann cell-axon and Schwann cell-Schwann cell interactions initiated by N-CAM and L1/neuron-glia cell adhesion

molecule (12). Thus two cell adhesion molecules that act at successive stages of neural development also share significant structural properties.

## EXPERIMENTS

Primary Structure of the 1B236 mRNA: Alternative Splicing Produces Two Forms. We described a 1500-nucleotide (nt) partial cDNA clone of the 2500-nt 1B236 mRNA (2). Two additional clones (p1B236-18 and p1B236-20) with apparently full-length inserts were isolated from a rat brain cDNA library and their nucleotide sequences were determined. The sequence of p1B236-18 (nt 1-2389) was identical where it overlapped (nt 1085-2389) with the published partial sequence (2), except for a minor difference (Fig. 1). Primer extension analyses show that clone p1B236-18 is a full-length clone (data not shown). The single open reading frame begins at nt 163, extends to the termination triplet TGA at nt 2043, and can be translated into a 626-residue protein (Fig. 1). The sequence of most of the single rat 1B236 gene has been determined and, when aligned with the p1B236-18 cDNA sequence, reveals the positions of 11 introns (Fig. 1).

The second cDNA (p1B236-20) was identical in sequence with p1B236-18 except that it included an additional 45 nt (Fig. 2A) between positions 1878 and 1879 and lacked nt 20-71. The additional 45 nt in clone 20 are flanked in the genomic sequence by introns (Fig. 2A). Thus the mRNAs corresponding to clones 18 and 20 differ by the alternate inclusion (Fig. 2B) of exons 2 or 12. Exclusion of exon 2 does not affect the coding region since the exon is located in the 5'-noncoding region. However, inclusion of exon 12 introduces a different coding sequence and a termination codon. This produces a 582-residue sequence with the last 9 amino acids distinct from the last 53 residues of the p1B236-18 form.

A cDNA fragment containing exon 12 and flanking regions was subcloned into pT7/T3-19 and used to generate a labeled minus-strand copy (Fig. 3). The inserts from p1B236-18 and p1B236-20 were subcloned into pGEM-4 and used to produce full-length copy RNAs with and without exon 12. When these RNAs were mixed with the labeled probe and the resulting hybrids were treated with RNase A, the clone 20 transcript protected fragments of  $\approx$ 230 and 120 nt and the clone 18 transcript protected fragments of  $\approx$ 158 and 147 nt (Fig. 3). No fragments were protected by tRNA. Fragments that comigrated with both the clone 20- and clone 18-protected RNAs were protected by brain RNAs (Fig. 3). Thus mRNAs lacking exon 12 (presumably encoding the 626-residue protein) and mRNAs containing exon 12 (presumably encoding the 582residue form) are both stable gene products.

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.

Abbreviations: N-CAM, neural cell adhesion molecule; MAG, myelin-associated glycoprotein; nt, nucleotide(s).

CAGAAGCCAGACCATCCAACCTTCTGTATCA	GTGCTCCTCGTCGCCTCACTGTACTTCACGG/		GAGCGGAATCAGGAGACATTCCCAACTCAGG	5 120
AGACTGAGGTGAGGGCCCTAGCTCGCCCACT	MetIlePheLeuThrThrLeu	uProLeuPheTrpIleMetIleSerA	LaSerArgGly3lyHisTrpGlyAlaTrpMet	26
	IgCTGGACAAGATGATATTCCTTACCACCCIC	GCCTCTGTTTTGGATAATGATTICAG	CTTCTCGAGGGGGGGCACTGGGGTGCCTGGAT	240
ProSerSerIleSerAlaPheGluGlyThrC	ysValSerIleProCysAryFheAspPhePro	oAspGluLeuArgProAlaValValH	isGlyValTrpTyrPheAsnSerProTyrPro	66
CCCTCGTCCATCTCAGCCITCGAGGGCACGT	GTGTCTCCATCCCC「GCCGTTTCGACTTCCC	GGATGAGCTCAGACCGGCTGTGGTAC	ATGGCGTCTGGTATTTCAACAGTCCCTACCCC	360
LysAsnTyrProProValValPheLysSerA	rgThrClnValValHisGluSerPheGlnGly	yArgSerArgLeuLeu3lyAspLeuG	LyLeuArgAsnCysThrLeuLeuSerTh	106
AAGAACTACCCGCCAGTGGTCTTCAAGICCC	GCACACAAGTGGTCCACGAGAGCTTCCAGGG	CCGTAGCCGCCTGTTGGGAGACCTGG	GCCTACGAAACTGCACCCTGCTTCTCAGCACC	480
LeuSerProGluLeuGlyGlyLy&TyrTyrP.	he <u>ArqGlyAsp</u> LeuGlyGlyTyrAsnGlnTyr	rThrPheSerGluHisSerValLeuA	BPILeILeABnThrProAbnIleValValPro	2 146
CTGAGCCCTGAGCTGGGAGGGAAATACTATT	TCCGAGGTGACCTGGGCGGCTACAACCAGTA(	CACCTICTCGGAGCACAGCGTCCTGG	ACATCATCAACACCCCCAACATCGTGGTGCCC	600
ProGluValValAlaGlyThrGluValGluV	alSerCyBMetValProAspAsnCysProGli	uLeuArgProGluLeuSerTrpLeuG	lyHisGluGlyLeuGlyGluÞroThrValLeu	186
CCAGAAGTGGTGGCAGGAACGGAAGTAGAGG	TCAGCTGCATGGTGCCGGACAACTGCCCAGAG	GCTGCGCCCTGAGCTGAGCTGGCTGG	GCCACGAGGGGCTAGGGGAGCCCACIGITCTG	720
GlyArgLeuArgGluAspGluGlyThrTrpV	alGlnValSerLeuLeuHisPheValProThr	rArgGluAlaAsnGlyHisArgLeuG	lyCysGlnAlaAlaPheProAsnThrThrLeu	226
GGTCGGCTGCGGGAGGATGAAGGCACCTGGG	TGCAGGTGTCACTGCTACACTTCGTGCCTACI	FAGAGAGGCCAACGGCCACCGTCTGG	GCTGTCAGGCTGCCTTCCCCAACACCACCTTG	840
GlnPheGluGlyTyrAlaSerLeuAspValL	ysTyrProProValIleValGluMetAsnSer	rSerValGluAlaIleGluGlySerH	isValSerLeuLeuCysGlyAlaAspSerAs	266
CAGTTCGAGGGTTACGCCAGTCTGGACGTCA	AGTACCCCCCGGTGATTGTGGAGATGAATTCC	CTCTGTGGAGGCCATTGAGGGCTCCC	ATGTCAGCCTGCTCTGTGGGGCTGACAGCAAC	960
ProProProLeuLeuThrTrpMetArgAspG	lyMetValLeuArgCluAlaValAlaCluSer	rLeuTyrLeuAspLeuGluGluValT)	hrProAlaGluAspGlyIleTyrAlaCysLes	306
CCGCCACCGCTGCTGACTTGGATGCGGGATG	GGATGGTGTTGAGGGAGGCAGTTGCTGAGAGG	CCTGTACCTGGATCTGGAGGAGGTGA	CCCCAGCAGAGGGGGGCATCIATGCTTGCCTG	1080
AlaGluA&AlaTyrGlyGlnA&pA&ArgT	hrValGluLeuSerValMetTyrAlaProTrp	oLysProThrValAsnGlyThrValVe	alAlaValGluGlyGluThrValSerIleLeu	346
GCAGAGAATGCCTATGGCCAGGACAACCGCA	CGGTGGAGCTGAGCGTCAIGTATGCACCTTGC	GAAGCCCACAGTGAATGGGACGGTGG	TGGCGGTAGAGGGGGGAGACAGTCTCCATCCTC	1200
CysSerThrGlnSerAsnProAspProIleL	euThrIlePheLyeGluLysGlnIleLeuAld	zThrValIleTyrGluSerGlnLeuG	In LouGlu LouPro AlaVal ThrProGlu Aoj	386 386
TGTTCCACACAGAGCAACCCGGACCCTATTC	TCACCATCTTCAAGGAGAAGCAGATCCTGGCC	CACGGTCATCTATGAGAGTCAGCTGC	NGC TGGAACTCCC TGCAG TGACGCCCGAGGAC	320
CysSerThrGinSerAenProAspProIleL	suthrILePheLyeCluLyeClnILeLenAld	a <b>Thrval IletyrGluSerGlnLeuG</b>	L <b>RLENGLULENPROALAVALThrProGLUABJ</b>	<b>386</b>
TGTTCCACACAGAGCAACCCGGACCCTATTC	TCACCATCTTCAAGGAGAAGCAGATCCTGGCC	ACGGICATCTATGAGAGICAGCIGG	AGCTGGAACTCCCTGCAGTGACGCCCCGAGGAC	<b>1320</b>
AspClyCluTyrTrpCysValAlaCluAenG	InTyrClyClnArgAlaThrAlaPheAsnLeu	USErValGluPheAlaProIleIleL	EULEUGLUSERHİBCYBALAALAALAALAARGABJ	426
GATGGGGAGTACTGGTGTGTAGCTGAGAACC	AGTATGGCCAGAGAGCCACCGCCITCAACCT(	STCTGTGGGAGTTIGCICCCATAATCC	ITCTGGAATCGCACTGTGCAGCGGCCAGAGAG	1440
CyeSerThrGinSerAenProAppro IleL	euthrItePhelyeGluLyeGlnIteLeuAld	aThrval IletyrcluserClaleuc	L <b>RLENGLULENPROALAVALThrProGluAe</b>	<b>386</b>
TGTTCCACACAGAGCAACCCGGACCCTATTC	TCACCATCTTCAAGGAGAAGCAGATCCTGGCO	ACGGICATCTATGAGATCAGETGCI	INCTEGAACTCCCTGCAGTGACGCCCCGAGGAC	<b>1320</b>
AspGlyGluTyrTrpCysValAlaGluAenG	InTyrGlyGlnArgAlaThrAlaPheAsnLeu	Iserval GluPhe Alapro Ile Ile Li	euLeuGluSerHiðCyðAlaAlaAlaAlaArgAø	426
GATGGGGAGTACTGGTGTGTAGCTGAGAACC	AGTATGGCCAGAGAGCCACCGCCITCAACCTO	STCTGTGGAGTTTGC ICCCATAATCC	ITCTGGAATCGCACTGTGCAGCGGCCAGAGAC	1440
ThrValGlnCysLeuCysValValLysSerA	enProGluProSerValAlaPheGluLeuPro	DSer Arg Aënval Thrval AënGluTi	hrGluArgGluPheValTyrSerGluArgSer	466
ACCGTGCAGTGCCTGTGTGTGGGTAAAATCCA	ACCCGGAACCCTCCGIGGCCTTTGAGCTGCCT	ITCCCGCAACG IGACTGTGAACGAGA	CAGAGAGGGGAGTTTGTGTACTCAGAGCGCAGC	1560
CyeSerThrGInSerAemProAmproa	euthrILePhelyeClulyeClnIleLenAld	27hrValiletyrGluserGinleuG	LnLeuGluLeuProAlaValThrProGluAeg	386      1320      426      1440      466      1560      506      1680
TGTTCCACACAGAGCAACCCGGACCCTATTC	TCACCATCTTCAAGGAGAAGCAGATCCTGGCC	CACGGTCATCTATGAGAGTCAGCTGC	AGCTGGAACTCCCTGCAGTGACGCCCCGAGGAC	
AspGlyGluTyrTrpCysValAlaGluAemG	InTyrGlyGlnArgAlaThrAlaPheAsnLeu	USerValGluPheAlaProileileL	euLeuGluSerHieCyeAlaAlaAlaAlaArgAeg	
GATGGGGAGTACTGGTGTGTAGCTGAGAACC	AGTATGGCCAGAGAGCCACCGCCITCAACCTC	GTCTGTGGAGTTTGCTCCCATAATCC	TTCTGGAATCGCACTGTGCAGCGGGCCAGAGAC	
ThrValGlnCysLeuCysValValLysSerA	enPro3luProSerValAlaPheGluLeuPro	DSerArgABnValThrValABnGluT	hrGluArgGluPheValTyrSerGluArgSer	
ACCGTGCAGTGCCTGTGTGTGGTAAAATCCA	ACCCGGAACCCTCCGIGGCCTTTGAGCTGCCT	ITCCCGCAACGIGACTGTGAACGAGA	CAGAGAGGGAGTTTGTGTACTCAGAGCGCAGG	
GlyLeuLeuLeuThrSerIleLeuThrLeuA	rgGlyGlnAlaGlnAlaProProArgYalIle	ECy8ThrSerArgA8nLeuTyrGlyT	hrGlnSerLeuGluLeuProPheGlnGlyAla	
GGCCTCCTGCTCACCAGCATCCTCACGCTCC	GGGGTCAGGCCCAAGCCCCACCCGCGTCAT	ITGTACCTCCAGGAACCTCTACGGCA	CCCAGAGCCTCGAGGTGCCTTTCCAGGGAGCA	
CyeSerthrGinSerAenProAepProIleL TGTTCCACACAGAGCAACCCGGACCCTATTC AepGlyGluTyrTrpCyeValAlaGluAenG GATGGGGAGTACTGGTGTAGCTGAGAAACC ThrValGlnCyeLeuCyeValValLyeSerA ACCGTGCAGTGCCTGTGTGTGGTAAAATCCA GlyLeuLeuLeuThrSerIleLeuThrLeuA GGCCTCCTGCTCACCAGCATCCTCACGCTCC DG- HieArgLeuMetTrpAlaLye <u>IleGlyProV</u> CACCGACTGATGTGGGGCCAAAATCGGCCCTG	euthrilePhelyeolulyeolnileLenAla TCACCATCTTCAAGGAGAAGCAGATCCTGGCO InTyrClyGlnArgAlaThrAlaPheAenLeu AGTATGGCCAGAGAGCCACCGCCITCAACCTO enProSluProSerValAlaPheGluLeuPra ACCCGGAACCCTCCGIGGCCTTTGAGCTGCCT rgGlyGlnAlaGlnAlaProProArgYalIle GGGGTCAGGCCCAAGCCCCACCCGCGTCATI alGlyAlaValValAlaPheAlaIleLeuILe TGGGTGCTGTGGGICGCCTTTGCCATCCTGATI	aThrval IletyrGluserGlnleuG CACGGTCATCTATGAGAGTCAGCTGC USErValGluPheAtaProIleIleL GTCTGTGGAGTITGCTCCCATAATCC DSErAryAenValThrValAenCluT ITCCCGCAACGIGACTGTGAACGAGA CyeThrSerArgAenLeuTyrGlyT ITGTACCTCCAGGAACCTCTACGGCA eAlaIleValCysTurIleThrGlnT IGCCATTGTCIGCTACATCACCCCAGA	LALGUGU LEUPROALAVAITAPPROGUABE AGCTGGAACTCCCTGCAGTGACGCCCCGAGGAC euLeuGluSerHibCysAlaAlaAlaAlaAlaArgAbe ITCTGGAATCGCACTGTGCAGCGGCCAGAGAC hrGluArgGluPheValTyrSerGluArgSer CAGAGAGGGAGTTTGTGTACTCAGAGCGCAGG hrGlnSerLeuGluLeuProPheGInGlyAla CCCAGAGCCTGCCTTTCCAGGGAGCA DT A A A A A A A A A A A A A A A A A A A	386      1320      426      1440      466      1560      506      1680      546      1800
CyeSerthrGinSerAenProAepProIleL TGTTCCACACAGAGCAACCCGGACCCTATTC AspGlyGluTyrTrpCysValAlaGluAenG GATGGGGAGTACTGGTGTAGCTGAGAACC ThrValGlnCysLeuCysValValLysSerA ACCGTGCAGTGCCTGTGTGTGGGTAAAATCCA GlyLeuLeuLeuThrSerIleLeuThrLeuA GGCCTCCTGCTCACCAGCATCCTCAGGCTCC DG- HisArgLeuMetTrpAlaLys <u>IleGlyProV</u> CACCGACTGATGTGGGGCCAAAATCGGCCCTG SerPheSerAlaGlyAspAenProHieValL AGCTTCTCAGCGGGGAGACAACCCTCATGTCC	euthr II ePheLyeoTuLyeoTuTieLeuAlo TCACCATCTTCAAGGAGAAGCAGATCCTGGCO InTyrGlyGlnArgAlaThrAlaPheAsnLeu AGTATGGCCAGAGAGCCACCGCC ITCAACCTO enProJluProSerValAlaPheGluLeuPro ACCCGGAACCCTCCG IGGCCTTTGAGCTGCCT rgGlyGlnAlaGlnAlaProProArgValILe GGGGTCAGGCCCAAGCCCCACCCGCGTCATI alGLyALaVaLVaLALaPheALaILeLeuILe TGGGTGCTGTGGGICGCCTTTGCCATCCTGAT eutyrSerProGluPheArgILeSerGlyALo TGTACAGCCCCGAATCTCGGATCTTGGGGC	2Thrval IletyrGluserGlnleuG CACGGTCATCTATGAGAGTCAGCTGC USErValGluPheAlaProIleIleL GTCTGTGGAGTTTGCTCCCATAATCC DSErArg Aenvalthrval Aenglut TTCCCGCAACGIGACTGTGAACGAGA eCyethrserArgAenLeutyrGlyt TTGTACCTCCAGGAACCTCTACGGCA <u>eAlaIleValCystyrIlethr</u> Glnt TGCCATTGTCTGCTACATCACCCAGA 2ProAepLyetyrGluSerGluLyeA ACCTGATAAGTATGAGAGTGAGAAAGC	LALENGULLENProALaValThrProGLAGE AGCTGGAACTCCCTGCAGTGACGCCCCGAGGAC euLeuGluSerHibCysAlaAlaAlaAlaArgAbg TTCTGGAATCGCACTGTGCAGCGGCCAGAGAC hrGluArgGluPheValTyrSerGluArgSer CAGAGAGGGAGTTTGTGTACTCAGAGCGCAGG hrGlnSerLeuGluLeuProPheGlnGlyAla CCCAGAGCCTGCAGCGCTGCTGCCTGGGGAGCA hrArgArgLysLySAsnValThrGluSerPro CAGAGAGAAAAAAGAACGTCACAGAGAGGCCCC rgLeuGlySerGluArgArgLeuLeuGlyLeu GCCTGGGGTCCGAGAGGAGGGAGGCTGCTGGGCCTT	386      1320      426      1440      466      1560      506      1680      546      1800      586      1920
CyeSerthrGinSerAenProAepProIIeL TGTTCCACACAGAGCAACCCGGACCCTATTC AspClyCluTyrTrpCysValAlaCluAenG GATGGGGAGTACTGGTGTAGCTGAGAACC ThrValClnCysLeuCysValValLysSerA ACCGTGCAGTGCCTGTGTGTGGGTAAAATCCA ClyLeuLeuLeuThrSerIleLeuThrLeuA GGCCTCCTGCTCACCAGCATCCTCACGCTCC DG HisArgLeuMetTrpAlaLys <u>IleGlyProV</u> CACCGACTGATGTGGGGCCAAAATCGGCCCTG SerPheSerAlaClyAspAenProHisValL AGCTTCTCAGCGGGAGACAACCCTCATGTCC ArgClyCluProProCluLeuAspLeuSerT AGGGGGGAACCCCCAGAACTGGACCTCAGT	euthrilePhelyeGlulyeGlnileLeuAld TCACCATCTTCAAGGAGAAGCAGATCCTGGCC InTyrGlyGlnArgAlaThrAlaPheAsnLeu AGTATGGCCAGAGAGCCACCGCCITCAACCTC enProJluProSerValAlaPheGluLeuPro ACCCGGAACCCTCCGIGGCCTTTGAGCTGCCT rgGlyGlnAlaGlnAlaProProArgValIle GGGGTCAGGCCCAAGCCCCCCGCGICATI alGlyAlaValValAlaPheAlaIleLeuIle TGGGTGCTGTGGICGCCTTTGCCATCCTGATI euTyrSerProGluPheArgIleSerGlyAle TGTACAGCCCCGAATTCCGAATCTCTGGGGC YrSerHieSerAepLeuGlyLyeArgProThe ATTCCCACTCAGACCTGGGGAAACGACCCACC	27hrval IletyrGluserGlnleuG CACGGTCATCTATGAGAGTCAGCTGC UservalGluPheAlaProIleIleL STCTGTGGAGTTTGCTCCCATAATCC DSerAry Aonvalthrval AonGlut TTCCCGCAACGTGACTGTGAACGAGA CCVBThrSerArg AonLeuTyrGlyT TTGTACCTCCAGGAACCTCTACGGCA <u>eAlaIleValCusturIleThr</u> Glnt TGCCATTGTCTGCTACATCACCCAGA 2ProAopLysTyrGluSerGluLysA ACCTGATAAGTATGAGAGTGAGAAGC rLysAopSerTyrThrLeuThrGluG CAAGGACAGCTACACCCTGACAGAGG	LALENGULLENPROALAVAITAPPROGUAS AGCTGGAACTCCCTGCAGTGACGCCCCGAGGAC EULEUGUSERHIBCY8ALAALAALAALAARGAB ITCTGGAATCGCACTGTGCAGCGGCCAGAGAC hrGluArgGluPheValTyrSerGluArgSer CAGAGAGGGAGTTTGTGTACTCAGAGCGCAGC hrGlnSerLeuGluLeuProPheGlnGlyALa CCCAGAGCCTGCAGCTGCCTTTCCAGGGAGCA hrHrgArgLyELySABANValThrGluSerPro CAAGAAGAAAAAAAGAACGTCACAGAGAGCCCC rgLeuGlySerGluArgArgLeuLeuGlyLeu SCCTGGGGTCCGAGAGGAGGCTGCTGGGCCT LuLeuALaGLUTYrALaGLYILeArgValLyet AGCTGGCTGAGTACGCAGAAATCCGAGTCAA	386      1320      426      1440      466      1560      506      1680      546      1800      586      1920      626      2040
CyeScrftrGInSerAenProAspProfleC TGTTCCACACAGAGCAACCCGGACCCTATTC AspGlyGluTyrTrpCysValAlaGluAenG GATGGGGAGTACTGGTGTGTAGCTGAGAACC ThrValGlnCysLeuCysValValLysSerA ACCGTGCAGTGCCTGTGTGTGGGTAAAATCCA GlyLeuLeuLeuThrSerIleLeuThrLeuA GGCCTCCTGCTCACCAGCATCCTCACGCTCC DB	suthrilePhelysGluLysGlnIleLeuAld TCACCATCTTCAAGGAGAAGCAGATCCTGGCC InTyrGlyGlnArgAlathrAlaPheAsnLeu AGTATGGCCAGAGAGCCACCGCC ITCAACCTG snPro3luProSerValAlaPheGluLeuPro ACCCGGAACCCTCCGIGGCCTTTGAGCTGCCT rgGlyGlnAlaGlnAlaProProArgYalIle GGGGTCAGGCCCAAGCCCCACCCGCGTCATI alGlyAlaValValAlaPheAlaIleLeuIl TGGGTGCTGTGGIGGCCCTTGGACCACCTGGTCA TGTACAGCCCCGAATCTCGAATCTCTGAGCC yrSerProGluPheArgIleSerGlyAld TGTACAGCCCCGAATCCTCGAATCTCTGGAGCA yrSerHisSerAspLeuGlyLysArgProTh ATTCCCACTCAGGACCCTCGCTTGGCCCCCAC	IThrval IletyrGluserClaleuG ACGGTCATCTATGAGAGTCAGCTGL UserValGluPheAlaProileileL STCTGTGGAGTTTGCTCCCATAATCC oserArgAønValThrValAønGluT ITCCCGCAACGIGACTGTGAACGAGA eCyeThrSerArgAønLeuTyrGlyT ITGTACCTCCAGGAACCTCTACGGCA eAlaILøValCustyrIleThrGInT IGCCATTGTCIGCTACATCACCCAGA AProAspLysTyrGluSerGluLysA ACCTGATAAGTATGAGAGTGAGAAAGC rLysAøpSerTyrThrLeuThrGluG CAAGGACAGCTACACCCTGACAGAGG IGGCCGTGGGCTCCCTTTCTCIGGAG	La LeuGlu LeuProAlaVal ThrProGlu Asp AGC TGGAAC TCCCTGCAGTGACGCCCCGAGGAC euLeuGluSerHisCysAlaAlaAlaAlaArgAsp TTCTGGAATCGCACTGTGCAGCGGCCAGAGAC hrGluArgGluPheValTyrSerGluArgSer CAGAGAGGGGAGGTTGTGTACTCAGAGCGCAGC hrGlnSerLeuGluLeuProPheGlnGlyAla CCCAGAGCCTCGAGCTGCCTTTCCAGGGAGCA hrArg Arg LysAsonValThrGluSerPro CAGGAGAAAAAAGAACGTCACAGAGAGCCCC hrArg Arg LysAsonValThrGluSerPro CAGGAGAAAAAAGAACGTCACAGAGAGCCCC hrArg Arg LysAsonValThrGluSerPro CAGGAGAAAAAAGAACGTCACAGAGAGCCCC hrArg Arg LysEySasonValThrGluSerPro CAGGAGAAAAAAGAACGTCACAGAGAGCCCCC hrArg Arg LysErGluArg ArgLeuLeuGlyLeu SCCTGGGGTCCGAGAGGAGGCTGCTGGCCCT LuLeuAlaGluTyrAlaGluTieArgValLysA AGCTGGCTGAGTACGCAGAAAATCCGAGTCAAC	386      1320      426      1440      466      1560      506      1680      546      1800      586      1920      626      2040      2160
CyeSertrGinSerAenProAepProIleL TGTTCCACACAGAGCAACCCGGACCCTATTC AspClyCluTyrTrpCysValAlaCluAsnG GATGGGGAGTACTGGTGTGTAGCTGAGAACC ThrValClnCysLeuCysValValLysSerA ACCGTGCAGTGCCTGTGTGTGGGTAAAATCCA ClyLeuLeuLeuThrSerIleLeuThrLeuA GGCCTCCTGCTCACCAGCATCCTCACGCTCC DG HisArgLeuMetTrpAlaLys <u>IleGlyProV</u> CACCGACTGATGTGGGGCCAAAATCGGCCTG SerPheSerAlaClyAspAsnProHisValL AGCTTCTCAGCGGGAGACAACCCTCATGTCC ArgClyCluProProCluLeuAspLeuSerT AGGGGGGAACCTGGGGGCTGGCCCTGTGGCTCA CAGGAAACCAGTGAGGTCTTAGGGGCCCGGCC	euthrILePhelyeCluLyeCluLyeClnILeLeuAld TCACCATCTTCAAGGAGAAGCAGATCCTGGCC InTyrClyClnArgAlaThrAlaPheAsnLeu AGTATGGCCAGAGAGCCACCGCCITCAACCTC enProJuProSerValAlaPheCluLeuPro ACCCGGAACCCTCCGIGGCCTTTGAGCTGCCT rgClyClnAlaClnAlaProProArgValIle GGGGTCAGGCCCAAGCCCCACCCCGCGICATI alClyAlaValValAlaPheAlaIleLeuIle TGGGTGCTGIGGICGCCTTTGCCATCCTGAT eutyrSerProGLuPheArgIleSerClyAld TGTACAGCCCCGAATCCCGGGAAACGACCCACC yrSerHieSerAepLeuGlyLyBargProTh ATTCCCACTCAGGACCCTCGCTTGGCCCCCACT	arhrval IletyrGluserClaleuG CACGGTCATCTATGAGAGTCAGCTGC UservalGluPheAlaProJleIleL STCTGTGGAGTTTGCTCCCATAATCC DSerArg Aonvalthrval AonGlut TTCCCGCAACGTGACTGTGAACGAGA CONSTRUE CONSTRUE CONSTRUE CALALEVALCUSTUFILETHFOLM TGCCATTGTCTGCTACATCACCCAGA APPOASPLysTyrGLUSErCluLySA ACCTGATAAGTATGAGAGTGAGAAGGC TLYSASPSerTyrThrLeuThrGluG CAAGGACAGCTACACCCTGACAGGAGG TGGCCGTGGGCTCCCTTTCTCTTGAG	LALENGLULENPROALAVAITAPPROGLUAE AGCTGGAACTCCCTGCAGTGACGCCCCGAGGAC EULEUGLUSERHIBCYBALAALAALAALAARGAB TTCTGGAATCGCACTGTGCAGCGGGCCAGAGAC hrGluArgGluPheValTyrSerGluArgSer CAGAGAGGGGAGGTTGTGTTACTCAGAGCGCAGC hrGlnSerLeuGluLeuProPheGlnGlyALa CCCAGAGCCTCGAGCTGCCCTTTCCAGGGAGCC AGTAGAAGAAAAAGAACGTCACAGAGAGGCCCC rgLeuGlySerGluRrgArgLeuLeuGlyLeu GCCTGGGGTCCGAGAGGGGGGGGGGCGGGC AGTGGTAGGGGTGGGGGCGGGCAGGGCGGGC CTCTCCCCTCTTTTAACCTCAGCTGTTGAG	386      1320      426      1440      466      1560      506      1680      586      1920      626      2040      626      2040      2160      2280
CyeScrftrGInSerAemProAmproAmpro TGTTCCACACAGAGCAACCCGGACCCTATTC AspGlyGluTyrTrpCysValAlaGluAemG GATGGGGAGTACTGGTGTGTAGCTGAGAACC ThrValGlnCysLeuCysValValLysSerA ACCGTGCAGTGCCTGTGTGTGGGTAAAATCCA GlyLeuLeuLeuThrSerIleLeuThrLeuA GGCCTCCTGCTCACCAGCATCCTCACGCTCC DB- HisArgLeuMetTrpAlaLys <u>IleGlyProV</u> CACCGACTGATGTGGGGCCAAAATCGGCCCTG SerPheSerAlaGlyAspAemProHisValL AGCTTCTCAGCGGGAGACAACCCTCATGTCC ArgGlyGluProProGluLeuAspLeuSerT AGGGGGGAACCCCCAGAACTGGACCTCA CAGGAAACAGTGGGGGCTGGCCCTGTGGCTCA	euthrILePhelyeCluLyeCluLieClanIleLeuAld TCACCATCTTCAAGGAGAAGCAGATCCTGGCC InTyrGlyGlnArgAlaThrAlaPheAsnLeu AGTATGGCCAGAGAGCCACCGCC ITCAACCTG enProJluProSerValAlaPheGluLeuPro ACCCGGAACCCTCCGIGGCCTTTGAGCTGCCT rgGlyGlnAlaGlnAlaProProArgYalIle GGGGTCAGGCCCAAGCCCCACCCGCGTCATI alGlyAlaValValAlaPheAlaIleLeuIl TGGGTGCTGTGGICGCCTTTGGCACCCGAT eutyrSerProGluPheArgIleSerGlyAld TGTACAGCCCCGAATTCCGATCTCTGGAGCA yrSerHieSerAepLeuGlyLyeArgProTh ATTCCCACTCAGGACCCTCGCTTGGCCCCACT tCCCCCCATCAGGACCCTCGCTGGCCCCCACT	athrval flatyrGluserClaleuG CACGGTCATCTATGAGAGTCAGCTGL USErValGluPheAlaProileileL STCTGTGGAGTTTGCTCCCATAATCC DSErArgAönValThrValAönGluT TTCCCGCAACGIGACTGTGAACGAGA ECYBThrSerArgAanLeuTyrGlyT TTGTACCTCCAGGAACCTCTACGGCA CACGTAAGTATGAGAGTGAGAAGCC CAAGGACAGCTACACCCTGACAGAGG CAAGGACAGCTACACCCTGACAGAGG ACCTGGACAGCTACACCCTGACAGGGA CCAGGACCTGTACAAAAGGGACATGA	LALENGLULENPROALAVALTAPPROGLUAE AGCTGGAACTCCCTGCAGTGACGCCCCGAGGAC EULEUGLUSERHIBCYBALAALAALAATGAB TTCTGGAATCGCACTGTGCAGCGGCCAGAGAC hrGluArgGLuPheValTyrSerGluArgSer CAGAGAGGGGAGTTTGTGTACTCAGAGCGCAGC hrGlnSerLeuGLULeuProPheGlnGLyALA CCCAGAGCCTCGAGCTGCCTTTCCAGGGAGCA hrarg Arg LyBLyBABNValThrGLuSerPro CAAGAAGAAAAAAGACGTCACAGAGAGGCCCC AGTGGGGTCCGAGAGGAGGCTGCTGGGCCTT LULEUALAGLUTYRALAGLUILEARGVALLYBA AGCTGGCTGAGTACGCAGAAATCCGAGTCAAC AGTGGTAGGGGTGGGGGGCGGGAAGGGGCGGGC CTCTCCCCTCCTTTTAACCTCAGCTGTTGAGA	386      1320      426      1440      466      1560      506      1680      546      1800      586      1920      626      2040      2160      2280      2400
CyeSertrGinSerAenProAspProfleC TGTTCCACACAGAGCAACCCGGACCCTATTC AspClyCluTyrTrpCysValAlaCluAsnG GATGGGGAGTACTGGTGTGTAGCTGAGAACC ThrValClnCysLeuCysValValLysSerA ACCGTGCAGTGCCTGTGTGTGGGTAAAATCCA ClyLeuLeuLeuThrSerIleLeuThrLeuA GGCCTCCTGCTCACCAGCATCCTCACGCCCC DG HisArgLeuMetTrpAlaLys <u>IleGlyProV</u> CACCGACTGATGTGGGGCCAAAATCGGCCCTG SerPheSerAlaClyAspAsnProHisValL AGCTTCTCAGCGGGAGACAACCCTCATGTCC AgGGGGGAACCCGGGGGGCACAACCGTGGCTCA CAGGAAACCAGTGGGGCCTGGCCCTGTGGCTCA CAGGAAACAGTGGGGGCTGGCCCTGTGGCTCA CAGGAAACAGTGAGGTCTTAGGGGCCCGGCC GGGGTGCTCTGTCTGTCCATGTTATTTATTG ACCCATCCTTTGGAGGAAAGGGGCATATTAG	euthrILePhelyeGLuLyeGLnILeLeuAld TCACCATCTTCAAGGAGAAGCAGATCCTGGCG InTyrGLyGlnArgAlaThrAlaPheAsnLeu AGTATGGCCAGAGAGCCACCGCCITCAACCTG enProJLuProSerValAlaPheGluLeuPro ACCCGGAACCCTCCGIGGCCTTTGAGCTGCCT rgGLyGlnAlaGLnAlaProProArgValILe GGGGTCAGGCCCAAGCCCACCCCGCGICATI alGLuALaVaLVaLALaPheALaILeLeuILe TGGGTGCTGTGGICGCCTTTGCCATCCTGAT euTyrSerProGLuPheArgILeSerGLyALe TGTACAGCCCCGAATCCCGGGAAACGACCCACC CCCCCCATCAGGACCCTCGCTTGGCCCCCACT TCCCCTCCTGCGGCCTCGCGCAAACGACCCACC CCCCCCATCAGGACCCTCGCTTGGCCCCCACT	arhrval IletyrGluserCinleuG CACGGTCATCTATGAGAGTCAGCTGC UservalGluPheAlaProIleIleL STCTGTGGAGTTTGCTCCCATAATCC DSerArg AonvalThrval AonGluT TTCCCGCAACGTGACTGTGAACGAGA CUBThrSerArg AonLeuTyrCiyT TTGTACCTCCAGGAACCTCTACGGCA CALaILevalCusturIleThrCinT TGCCATTGTCTGCTACATCACCCAGA AProAopLysTyrGluSerCluLysA ACCTGATAGTATGAGAGTGAGAGAGC TLys AspSerTyrThrLeuThrGluG CAAGGACAGCTACACCCTGACAGGGG TGGCCGTGGGCTCCCTTTCTCTTGAG ACATCCTGCACCTATGTTACAGCTCC CCAGGACCTGTACAAAAGGGACATGA TGGACAAAAACAC	LALENGLULENPROALAVALTAPPROGLUAE AGCTGGAACTCCCTGCAGTGACGCCCCGAGGAC EULEUGLUSERHIBCYBALAALAALAALAARGAB TTCTGGAATCGCACTGTGCAGCGGGCCAGAGAC hrGluArgGluPheValTyrSerGluArgSer CAGAGAGGGGAGTTTGTGTACTCAGAGCGCAGC hrGlnSerLeuGluLeuProPheGlnGlyALa CCCAGAGCCTCGAGCTGCCCTTTCCAGGGAGCCA DTATAGAGCCTCGAGCTGCCCTTCCAGGGAGCCCC hrArgArgLyELYBABNValThrGluSerPro CAAGAAGAAAAAAGAACGTCACAGAGAGCCCC rgLeuGlySerGluRrgArgLeuLeuGlyLeu GCCTGGGGTCCGAGAGGGGGGGGGGCCGGC AGTGGTAGGGGTGGGGGGCGGGCCGGGC AGTGGTAGGGGTGGGGGGGGGGGGGGGGGGGGCGGGC CTCTCCCCTCCTTTTAACCTCAGCTGTTGAGA	386      1320      426      1440      466      1560      506      1680      586      1920      626      2040      2160      2280      2474

FIG. 1. Nucleotide sequence of cDNA clone p1B236-18 and the translated 1B236 amino acid sequence. Arrowheads under the sequence indicate positions of introns. D0-D8 represent sections of protein defined by intron-exon boundaries. Stars indicate putative sites for N-glycosylation. Underlined residues 118-120 are a possible site for interaction with cell attachment substrates (13). Underlined residues 514-534 represent the putative membrane spanning segment. Vertical arrows near C terminus represent proteolytic cleavage sites (5). In a report (2) nt 1089 and 1090 were clerically transposed, and hence two adjacent amino acids were mistranslated. The poly(A) tail for clone 1B236-18 appears following nt 2389; the poly(A) tail in the original clone appeared after residue 2474.

Fragments that comigrate with those protected by both clone 18 and 20 transcripts are detected with RNA samples from all postnatal ages (Fig. 3). Those corresponding to the clone 18 transcript appear more abundant in the hindbrain at earlier ages and increase to their highest level with RNA from days 17-29 and decrease with day-50 RNA. Fragments comigrating with those protected by the clone 20 transcript are present at low levels in the samples protected by RNA from the younger rats and increase continuously when protected with later hindbrain RNA samples. Thus, the mRNA lacking exon 12 peaks during early postnatal development and decreases, whereas the mRNA containing exon 12, a minor form at early times, becomes the predominant adult species. The two forms do not correlate in a simple way with the oligodendrocyte to neuronal shift in 1B236 expression because antisera to two peptides whose sequence is exclusive to the product of the early mRNA react with neurons in adult rats (6). Similar RNase protection experiments suggest that the inclusion of exon 2 is not coordinately regulated with splicing of exon 12.

**1B236 Shares Primary Sequence with MAG.** The properties of the 1B236 glycoprotein led us to compare this molecule with MAG, a 100-kDa glycoprotein known to consist of two polypeptides of 67 kDa and 72 kDa (14). Antisera against MAG (15) and 1B236 peptide P5 (2) both detect a 100-kDa band in a purified MAG sample and in brain but not liver or kidney extracts (Fig. 4). Antisera to 1B236 synthetic peptides P6 and P7 (2) gave similar results. With all three anti-peptide sera, reactivity was blocked by the synthetic peptide. Thus molecules defined as MAG contain 1B236 antigenic determinants.

The insert of p1B236-18, subcloned into pGEM-4, was used to generate an RNA copy of the sense strand. Two truncated 1B236 RNAs lacking different lengths of the C-terminal coding regions were generated by cleaving the template with restriction endonucleases (Fig. 5A). The copy RNAs were translated *in vitro*, and the products were immunoprecipitated with polyclonal anti-MAG sera (15), with monoclonal anti-MAG antibodies (17), or with anti-P5 sera directed against the C terminus of 1B236 (2) (Fig. 5B). The full-length transcript was translated to give a 70-kDa product that was Neurobiology: Lai et al.



FIG. 2. Structures of alternatively spliced 1B236 mRNAs. (A) Comparison of cDNA and encoded amino acid sequences of clones 1B236-20 and 1B236-18, which differ by 45 nt located between positions 1878 and 1879 of the clone 18 sequence (Fig. 1). Sequence of genomic DNA is aligned below, demonstrating that the additional 45 nt (uppercase letters) form an exon flanked by introns (lowercase letters). D7 and D8 refer to the domains defined in Fig. 1; the different portion encoded by 1B236-20 is designated D8\*. (B) The 1B236/MAG gene contains 13 exons spanning 16 kilobases. Clones 1B236-18 ("early" mRNA) and 1B236-20 ("late" mRNA, see Fig. 3) differ by alternate inclusion of exons 2 and 12. Protein-coding portions of exons are shaded; domains D0-D8 are encoded by exons 4-11 and 13; exon 12 of the "late" mRNA which encodes D8\* is shaded differently to indicate the alternative C terminus. Dashed lines in exon 13 indicate that at least two sites are used for polyadenylylation. Sizes of the exons are exaggerated and are not strictly proportional.

precipitated by anti-P5 serum and the polyclonal and monoclonal anti-MAG antibodies, but not by control antiserum. The truncated RNAs were translated into products of 53 kDa and 36 kDa: each of these reacted with anti-MAG polyclonal but not with anti-P5 antibodies (Fig. 5B). Both anti-MAG monoclonal antibodies reacted with the 53-kDa product but not the 36-kDa product. Hence, the polyclonal anti-MAG antibodies react with antigenic determinants at least in the N-terminal region of the protein encoded by the cDNA clone 1B236-18, while both monoclonal antibodies react with determinants in regions D4 or D5 (defined below). Thus the 1B236 protein contains MAG epitopes not dependent on secondary modifications such as carbohydrate addition. Analogous results were obtained with the p1B236-20 insert subcloned into the pGEM-4 vector: as expected, the translation product was precipitated with anti-MAG and anti-P5 but not anti-P6 or -P7 antibodies.

## DISCUSSION

We propose the following structure for the 1B236 proteins (Fig. 6). The predominant early postnatal form of the 1B236 polypeptide is 626 amino acids long; a later developmental product of the same gene has a length of 582 residues. Introns divide the 1B236 protein into several sections, arbitrarily designated as domains D0 to D8 (Fig. 1) and D8\* (Fig. 2A). The two forms share D0-D7 and differ by either D8 or D8\*. Both forms include up to eight N-linked carbohydrate side chains (the site in the putative cytoplasmic tail is probably not used). The N-terminal domain D0 resembles a signal peptide sequence (18); cleavage after alanine-16 would result in polypeptides of 610 and 566 residues with  $M_{\rm rs}$  67,710 and 62,680. A presumptive internal transmembrane domain (21 uncharged amino acids of D6 underlined in Fig. 1) separates the glycosylated N-terminal region from alternative C-terminal tails of 92 and 48 residues. The C-terminal tails contain sets of tandem basic amino acids that are targets for proteolysis (5). The sequence of D8 suggests that it could possibly be a substrate for other modifications such as phosphorylation.

The N-terminal region is likely to be extracellular. Thus our hypothesis (2) that peptides produced by C-terminal proteolysis of 1B236 might have "neurotransmitter-like" functions seems unlikely, given that the C terminus is probably cytoplasmic. Nevertheless, these peptides might have intracellular signaling functions. Alternatively, modifications of the C terminus by proteolytic cleavage or phosphorylation may modulate binding activity of the N terminus or alter the accessibility of the C terminus to cytoplasmic components. Generation of different C termini by alternative splicing (Fig. 2) emphasizes the potential functional importance of this region of 1B236/MAG.

Arquint et al. (19) have isolated a partial cDNA clone encoding a protein that reacts with anti-MAG antibodies and have shown that its sequence overlaps that of 1B236. The identity of the clone with MAG was established by amino acid sequence analysis of fragments from purified MAG. Those data and the results presented here demonstrate that, using the available immunological reagents, 1B236 and MAG are indistinguishable. Thus, 1B236 and MAG are highly related if not identical molecules. There are differences, however, in the known cellular distributions of 1B236 and MAG. In young rats, MAG is predominantly oligodendrocytic (20), and 1B236 protein and mRNA are expressed predominantly in oligodendrocytes (ref. 1 and 4; G. A. Higgins, H. Schmale, F.E.B., M. C. Wilson, and R.J.M., unpublished data). However, in adults, both in situ hybridization (G. A. Higgins, H. Schmale, F.E.B., M. C. Wilson, and R.J.M., unpublished data) and immunohistochemistry (6) show that 1B236 is largely, although not exclusively, neuronal, whereas by immunohistochemistry and radioimmunoassay MAG has been found only in oligodendrocytes and myelin sheaths (20-23). One possible explanation for this apparent paradox is as yet unrecognized RNA splicing events that affect the N-terminal region and generate a neuronal molecule lacking the epitopes recognized by some MAG antibodies. A second possibility is differential accessibility in neuronal and glial cells of these N-terminal MAG epitopes in immunohistochemistry and RIA experiments. Thus our work-



FIG. 3. Developmental expression of alternatively spliced forms of 1B236 mRNAs. Probe (lane P) was made by excising the 231-base-pair Rsa I restriction fragment containing exon 12 and flanking regions from p1B236-20 and subcloning it into the Sma I site of pT7/T3-19 (Bethesda Research Laboratories). The construct was linearized with EcoRI that cuts within the insert, 9 base pairs from its 5' end. Transcription in the presence of [32P]GTP using T7 polymerase gave a predicted 270-nt species (the insert and 39 nt of vector) that was mixed with RNA samples, digested with RNase A (50  $\mu$ g/ml, 37°C for 30 min), and loaded onto 8% denaturing acrylamide gel. As controls, complete inserts of p1B236-18 and p1B236-20 were isolated by cleavage with BamHI, subcloned into the BamHI site of pGEM-4 (Promega Biotec, Madison, WI), linearized by digestion with Asp 718, and incubated with T7 RNA polymerase to produce sense-strand transcripts. Lane A shows two RNase A-protected species resulting from hybridization of probe with 1B236-20 RNA (contains exon 12) transcribed in vitro. A large band of  $\approx$ 230 nt represents the fully-protected probe; the small band (120 nt) has not been accounted for. Lane B shows the protected species resulting from hybridization of probe with 1B236-18 RNA (no exon 12) transcribed in vitro. Lane T shows protection of probe by 5  $\mu$ g of tRNA. Remaining lanes each show protection of probe by 1  $\mu$ g of cytoplasmic poly(A)<sup>+</sup> RNA derived from either hindbrain (pons/medulla) postnatal days 5-50 or cortex (CTX) (telencephalon) days 29 and 50. Apparent low day-13 signal is due to less RNA in the hybridization.

ing hypothesis is that 1B236 and MAG are identical or nearly identical products of the single rat 1B236 gene; in mice, this gene has been mapped to within 1 centimorgan of the *Gpi-1* locus on chromosome 7 (C. Blatt, L. Weiner, J.G.S., M. N. Nesbitt, M. I. Simon, unpublished data). Until the exact relationship of 1B236 and MAG is known, we propose the provisional name 1B236/MAG.

The shared N-terminal region of 1B236/MAG (Fig. 6) consists of five, roughly equal-sized domains (D1–D5) that are related to each other in sequence and show highly significant alignments with immunoglobulin sequences, with the sequences of other members of the immunoglobulin super family (8) and with three molecules not previously recognized as immunoglobulin super family members: platelet-derived growth factor receptor (25), the product of the oncogene v-fms (26), and the peripheral myelin protein P<sub>0</sub> (27). The most striking similarities are found between 1B236/MAG and N-CAM (9). The short Cys-Cys distances of the five 1B236/MAG domains suggest that they fold into the 7-chain, disulfide-bonded, two-layer  $\beta$ -sheet structure characteristic of immunoglobulin constant regions rather than the 9-chain structure of variable regions (8).



FIG. 4. Immunoblot analysis of 1B236 and MAG. Aliquots of purified MAG (lanes M, 150 ng) made as described (16) and of NaDodSO<sub>4</sub> extracts (50  $\mu$ g) of adult rat brain (lanes B), liver (lanes L), and kidney (lanes K) separated on 7.5% polyacrylamide gels, transferred to nitrocellulose, and assayed with rabbit anti-MAG antiserum (15) (1:500 dilution), anti-P5 serum (2) (1:500 dilution), with or without peptide P5 at 50  $\mu$ g/ml. Reactive band in all lanes of anti-MAG blot probably reflects detection of a common antigen by other antibodies in the serum and is unrelated to anti-MAG activity.

In particular, 1B236/MAG is highly related in primary sequence to N-CAM, which plays a critical role in neural development (28). N-CAM, L1/neuron–glia cell adhesion mol-



FIG. 5. Immunoprecipitation of in vitro-translated 1B236 proteins. (A) Schematic drawing illustrating structures of the protein products encoded by the T7 polymerase transcripts from 1B236 transcription templates. Cleavage with Asp 718 generates a transcript encoding the entire open reading frame. Cleavage with FnudII or Sty I generates truncated transcripts encoding proteins of the indicated molecular weight. NC, noncoding; TM, transmembrane; kd, kDa. (B) Transcripts were translated in vitro using a rabbit reticulocyte lysate system (Promega Biotec) with [35S]methionine. Lanes a, g, and l; products from the full-length transcript, the FnudII truncate, and the Sty I truncate, respectively. Lanes b-f, h-k, m-p; immunoprecipitation of the products shown in lanes a, g, and l, respectively, when incubated with negative control rabbit serum to an unrelated synthetic peptide (lane b), polyclonal anti-MAG(R-31, ref. 15) (lanes c, h, and m), anti-MAG monoclonal B11F7 (17) (lanes d, i, and n), anti-MAG monoclonal D7E10 (17) (lanes e, j, and o), and anti-1B236-P5 (2) (lanes f, k, and p). Samples were fractionated on a 10% NaDodSO<sub>4</sub>/polyacrylamide gel; positions of molecular size standards in kDa are indicated at left.



FIG. 6. Models of early and late expressed forms of the 1B236/MAG protein. Five immunoglobulin-like domains are indicated as disulfide-bonded loops and positions of carbohydrate side chains by branched structures. The Arg-Gly-Asp (RGD) sequence may be either in a loop that protrudes from "binding pocket" of domain 1 or in a linker region separating D1 from D2. Transmembrane domain D6 anchors the molecule in the cellular membrane. The two forms differ in the structures of their C termini. The early form contains a potential site (P) for tyrosine kinase phosphorylation within domain D8. Sites for proteolytic cleavage in the C-terminal tails are indicated by arrows.

ecule, and 1B236/MAG share carbohydrate determinants (7, 24, 29) and appear to act sequentially during peripheral myelination as cell adhesion molecules (12). In other neural cell interactions, N-CAM is thought to utilize homotypic reactions in mediating cell adhesion (28). As a cell adhesion molecule, 1B236/MAG may participate in similar homotypic interactions between oligodendrocytes or Schwann cells, particularly between adjacent surfaces of the myelinating cell membranes as these wrap around axons (10, 12). The ability of immunoglobulin domains to interact with each other in a homotypic or heterotypic manner (8) provides a possible structural explanation for the biology of these cell adhesion molecules. Po protein is also thought to react homotypically during peripheral myelin compaction (27). The finding that  $P_0$  protein is immunoglobulinlike supports this hypothesis and suggests that a single-domain molecule such as  $P_0$  protein could be similar in structure to the ancestral cell interaction molecule (8). 1B236/MAG may also be involved in heterotypic interactions: other members of the immunoglobulin family-for example, neuronally expressed OX-2, Thy-1, or N-CAM-are potential ligands. The sequence Arg-Gly-Asp, which has been specifically implicated in cell attachment (13), is located in D1 (Fig. 6) and could contribute to such interactions. The finding that 1B236/MAG is expressed

in neurons suggests that it may be involved in neuronal as well as glial cell interactions.

We thank Alan Williams and Neil Barclay for advice in comparing 1B236/MAG to the immunoglobulin super family; Ann-Ping Tsou, Dominique Lenoir, and Mary Kiel for reagents; Deeanne Bahr for technical assistance; Joseph Watson, Gerry Higgins, Hartwig Schmale, Ron Ogata, Peter Staeheli, Ute Hochgeschwender, Doug Feinstein, Freda Miller, Gabriel Travis, Patria Danielson, and Sonja Forss-Petter for helpful discussions; and Linda Elder for preparing the manuscript. This work was supported in part by Grants NS20728, GM32355, and NS22111 from the National Institutes of Health; Grant AA06420 from the NIAAA Alcohol Research Center; and McNeil Pharmaceuticals.

- 1. Milner, R. J. & Sutcliffe, J. G. (1983) Nucleic Acids Res. 11, 5497-5520.
- Sutcliffe, J. G., Milner, R. J., Shinnick, T. M. & Bloom, F. E. (1983) Cell 33, 671–682.
- Sutcliffe, J. G., Milner, R. J. & Bloom, F. E. (1983) Cold Spring Harbor Symp. Quant. Biol. 48, 477-484.
- Lenoir, D., Battenberg, E., Kiel, M., Bloom, F. E. & Milner, R. J. (1986) J. Neurosci. 6, 522-530.
- Malfroy, B., Bakhit, C., Bloom, F. E., Sutcliffe, J. G. & Milner, R. J. (1985) Proc. Natl. Acad. Sci. USA 82, 2009–2013.
- Bloom, F. E., Battenberg, E. L. F., Milner, R. J. & Sutcliffe, J. G. (1985) J. Neurosci. 5, 1781–1802.
- Friedlander, D. R., Grumet, M. & Edelman, G. M. (1986) J. Cell Biol. 102, 413-419.
- Williams, A. F., Barclay, A. N., Clark, M. J. & Gagnon, J. (1985) in *Gene Expression During Normal and Malignant Differentiation*, eds. Anderson, L. C., Gahmberg, C. G. & Ekblom, P. (Academic, London), pp. 125-138.
- Hemperly, J. J., Murray, B. A., Edelman, G. M. & Cunningham, B. A. (1986) Proc. Natl. Acad. Sci. USA 83, 3037-3041.
- 10. Quarles, R. H. (1984) Dev. Neurosci. 6, 285-303.
- Trapp, B. D., Quarles, R. H. & Suzuki, K. (1984) J. Cell Biol. 99, 594-606.
- 12. Martini, R. & Schachner, M. (1986) J. Cell Biol. 103, 2439-2448.
- Ruoslahti, E. & Pierschbacher, M. D. (1986) Cell 44, 517-518.
  Frail, D. E. & Braun, P. E. (1984) J. Biol. Chem. 259,
- 14857–14862.
- Quarles, R. H., Johnson, D., Brady, R. O. & Sternberger, N. H. (1981) Neurochem. Res. 6, 1115–1127.
- Quarles, R. H., Barbarash, G. R., Figlewicz, D. A. & McIntyre, L. J. (1983) *Biochim. Biophys. Acta* 757, 140–143.
- Dobersen, M. J., Hammer, J. A., Noronha, A. B., MacIntosh, T. D., Trapp, B. D., Brady, R. O. & Quarles, R. H. (1985) *Neurochem. Res.* 10, 499-513.
- 18. von Heijne, G. (1985) J. Mol. Biol. 184, 99-105.
- Arquint, M., Roder, J., Chia, L.-S., Down, J., Wilkinson, D., Bayley, H., Braun, P. & Dunn, R. (1987) Proc. Natl. Acad. Sci. USA 84, 600-604.
- Sternberger, N. H., Quarles, R. H., Itoyama, Y. & Webster, H. DeF. (1979) Proc. Natl. Acad. Sci. USA 76, 1510-1514.
- Itoyama, Y., Sternberger, N. H., Webster, H. deF., Quarles, R. H., Cohen, S. R. & Richardson, E. P. (1980) Ann. Neurol. 7, 167-177.
- Johnson, D., Quarles, R. H. & Brady, R. O. (1982) J. Neurochem. 39, 1356-1362.
- Johnson, D., Sato, S., Quarles, R. H., Inuzuka, T., Brady, R. O. & Tourtellotte, W. W. (1986) J. Neurochem. 46, 1086–1093.
- 24. Noronha, A. B., Ilyas, A., Antonicek, H., Schachner, M. & Quarles, R. H. (1986) Brain Res. 385, 237-244.
- Yarden, Y., Escobedo, J. A., Kuang, W.-J., Yang-Feng, T. L., Daniel, T. O., Tremble, P. M., Chen, E. Y., Ando, M. E., Harkins, R. N., Francke, U., Fried, V. A., Ullrich, A. & Williams, L. T. (1986) Nature (London) 323, 226-232.
- Hampe, A., Gobet, M., Sherr, C. J. & Galibert, F. (1984) Proc. Natl. Acad. Sci. USA 81, 85-89.
- 27. Lemke, G. & Axel, R. (1985) Cell 40, 501-508.
- 28. Edelman, G. M. (1986) Annu. Rev. Cell Biol. 2, 81-116.
- Kruse, J., Mailhammer, R., Wernecke, H., Faissner, A., Sommer, I., Goridis, C. & Schachner, M. (1984) Nature (London) 311, 153-155.