HYBRIDOMA
Volume 20, Number 4, 2001
Mary Ann Liebert, Inc.

# Immunogenetic Analysis of Variable Regions Encoding AB1 and γ-Type AB2 Antibodies from the NeuGc-Containing Ganglioside Family

ALEXIS PEREZ,[1] JOSEFA LOMBARDERO,[1] CRISTINA MATEO,[1] GERAUDIS MUSTELIER,[1] MAURO ALFONSO,[1] ANA M. VAZQUEZ,[1] and ROLANDO PEREZ[1]

## ABSTRACT

**The variable regions from P3, a murine monoclonal antibody (MAb) against NeuGc-containing gangliosides, and two anti-idiotype MAbs directed to P3 MAb were cloned and sequenced. Comparisons with previously reported sequences showed that P3 is a germline antibody encoded by genes from the $V_H$Q52 and $V_\kappa$19 families. Analysis of nucleotides at the heavy chain CDR3 (H-CDR3) showed the presence of an extensive 3′ N region that contains almost 50% of the nucleotides of this CDR. In addition, amino acid sequence analysis of the H-CDRs of this MAb revealed the presence of three arginines, two of which are present in the H-CDR3, that could be involved in the interaction of P3 MAb with its electronegative epitope on gangliosides. Anti-idiotype 1E10, which seems to define a "regulatory" idiotope on P3 MAb (it induces Id+ Ab3), represents a germline Ab2 that belongs to the $V_H$J558 and $V_\kappa$10 gene families. By contrary, the anti-idiotype 3B11 is an extensively mutated antibody that belongs to the $V_H$3660 and $V_\kappa$4/5 gene families, defining a "private" idiotope on P3 MAb. Even when different V genes contribute to the variable regions of 1E10 and 3B11 MAbs, they share an acidic motif E/D-D-Y/D-Y-D in H-CDR3, suggesting that both Ab2s recognize paratope positive residues on the Ab1. Therefore, complementary electrostatic interactions involving H-CDR3 from both Ab1 and Ab2, might provide a clue to understand the molecular basis for the generation of γ-type anti-idiotype antibodies to V regions recognizing glycolylated ganglioside antigens.**

## INTRODUCTION

GANGLIOSIDES are sialic acid-containing glycolipids that are normal components of mammalian tissues. They are considered as attractive targets for cancer immunotherapy in humans due to changes that occur in their expression patterns during oncogenic transformation. Whereas *N*-glycolyl neuraminic acid (NeuGc) is absent in normal human tissues,[1] the expression of NeuGc-containing gangliosides in some human tumors has been demonstrated.[2–4] Murine monoclonal antibody (MAb) P3 (IgM; k), generated by our group in BALB/c mice, reacts specifically with a broad battery *of* N-glycolyl-containing gangliosides, sulfated glycolipids, and with antigens expressed in human breast tumors.[5] Immunochemical experiments and molecular modeling have led to the delineation of the epitope recognized by this MAb, which includes the carboxyl and glycolic hydroxyl groups, together with the nitrogen function of sialic acid.[6] Previous work in syngeneic model has shown the capacity of this MAb to induce an IgG auto-anti-idiotypic response even when injected alone, results that were not obtained when other murine IgM MAbs were used.[7]

Looking for internal image anti-idiotypic MAbs (Ab2), we immunized BALB/c mice with P3 MAb coupled to KLH in the presence of Freund's adjuvant. Different MAbs were selected and cloned for further studies. All these antibodies inhibit the binding of P3 MAb to GM3(NeuGc), and induced anti-anti-idiotypic antibodies (Ab3) in syngeneic animals, but were unable to generate Ab3 with the same specificity of P3 MAb.[7] These results suggested that these antibodies were not internal image Ab2 and they were classified as γ-type Ab2. These γ-type Ab2s

[1]Department of Antibody Engineering, Center of Molecular Immunology, P.O. Box 16040, Havana 11600, Cuba.

could be distinguished by their different abilities to induce Ab3 sharing idiotopes with P3 MAb in syngeneic animals (Id+ Ab3).

In this paper we try to relate those immunochemical properties with the immunogenetic analysis of the primary structures of P3 MAb and two of the anti-P3 Ab2s previously reported: 1E10 Ab2, which was able to induce Id+ Ab3 response in syngeneic model, and 3B11 Ab2, that showed the highest binding affinity for P3 MAb but was unable to induce Id+ Ab3.[7] We determined the nucleotide sequences of the variable regions of the heavy and light chains ($V_H$, $V_L$) of these antibodies by cloning cDNAs derived from the RNA extracted from hybridoma cells. The frequency of complementary charged residues on both Ab1 and Ab2, suggests that electrostatic interactions contribute to the specificity and high affinity of the $\gamma$-type idiotypic recognition, in the V-region repertoire recognizing gangliosides.

## MATERIALS AND METHODS

### *Ab1 and Ab2 MAbs*

P3 MAb (IgM, $\kappa$) (Ab1) was generated by immunization of BALB/c mice with lyposomes containing NeuGc-GM3 and tetanus toxoid, as previously described.[5] Syngeneic Ab2 MAbs 1E10 ($IgG_1$, $\kappa$) and 3B11 ($IgG_1$, $\kappa$) were obtained from BALB/c mice immunized with purified P3 MAb coupled to KLH (Keyhole Limpet Haemocyanin) in Freund's adjuvant.[7]

### *Primers*

The following primers were used:

(1) CM1FOR (3′ primer for the heavy chain IgM): 5′-GGAAGCTTAAGACATTTGGGAAGGACTGAC-TCTC-3′;
(2) CG1FOR (3′ primer for the heavy chain IgG1); 5′-GGAAGCTTAGACAGATGGGG GTGTCGTTTTG-3′;
(3) CK2FOR (3′ primer for the kappa light chain): 5′-GGAAGCTTGAAGATG GATACAGTTGGTG-CAGC-3′;
(4) McKAs.XBA (3′ primer for the kappa light chain): 5′-GCGTCTAGAACTGGATGGTGGGAAGATGG-3′;
(5) VH1BACK (5′ primer for the heavy chain framework region 1 (FR1)): 5′-AGGT(G/C)(A/C)A(A/G)CTGCAG(G/C)AGTC(A/T)-GG-3′;
(6) VK8BACK (5′ primer for the $V_K8$ light chain FR1): 5′-C(A/T)GAGAAATTCAGCTGACCCAGTCTC-3′;
(7) VK1BACK (5′ primer for the $V_K1$ light chain FR1): 5′-GACATTCAGCTGACCCAGTCTCCA-3′;
(8) MLALT4.RV (5′ primer for the light chain signal sequence) 5′-GGGGATATCCACCATGAGG(G/T)CCCC(A/T)-(G/A)CTCAG(C/T)T(C/T)C(T/G)GT-3′.

Primers 1-3 contain a *Hind* III restriction site; primers 5–7 contain a *Pst* I restriction site; primer 4 contains a *Xba* I restriction site; primer 8 contains an *Eco* RV restriction site.

### *cDNA synthesis*

Total cellular RNA was extracted by Trizol reagent (Gibco-BRL, Paisley, Scotland). First strand cDNA was synthesized from 5 $\mu$g of total RNA, using a first-strand cDNA synthesis kit (Gibco BRL), essentially as described by the manufacturer. Specific primers complementary to the coding strand of the 5′ end of the constant regions (CH1 heavy and kappa light chains) were used. Primers 1 and 2 were used in the VH cDNA synthesis of P3 and Ab2 MAbs, respectively; primer 3 was used in the VL cDNA synthesis of P3 and 3B11 MAbs; primer 4 was used in the VL cDNA synthesis of 1E10 MAb. Before polymerase chain reaction (PCR), samples were heat-treated to inactivate reverse transcriptase.

### *PCR amplification of V-genes*

cDNAs were amplified by PCR as described by Orlandi et al.,[8] using one unit of Vent DNA polymerase, 200 $\mu$M dNTP and 0.5 $\mu$M of each primer in a final volume of 50 $\mu$L of reaction buffer essentially as described by the manufacturer (New England Biolabs, Inc., Beverly, MA). Primers 1 and 5 were used for the amplification of VH P3. Primers 3 and 6 were used for the amplification of VL P3. Primers 2 and 5 were used for the amplification of VH of Ab2. Primers 3 and 7 were used for the amplification of VL 3B11. Primers 3 and 8 were used for the amplification of VL 1E10.

Each reaction mixture was subjected to 30 cycles of amplification using a thermal minicycler PTC-150 (MJ Research Inc, Watertown, MA). Each cycle consisted of the following steps: 94°C for 1 min (except first cycle: 4 min), 55°C for 1 min and 72°C for 1 min with a final extension of 10 min at 72°C. Amplified variable regions were analyzed on a 1.8% low melting point agarose/Tris acetate-EDTA (TAE) gel and visualized with ethidium bromide. The band of the expected size was excised and purified by phenol extraction and ethanol precipitation.

### *Cloning and sequencing of PCR products*

For cloning, purified products were ligated into the M13mp19 vector (Pharmacia Biotech, Uppsala, Sweden). The amplified cDNAs encoding the heavy chain variable regions of P3 and Ab2 MAbs were digested with *Hind*III and *Pst*I and ligated into the M13mp19 vector, which was previously digested with *Hind* III and *Pst* I. The amplified cDNA encoding the light chain variable regions of P3 and 3B11 were digested with *Hind* III and *Pvu* II and ligated into M13mp19 vector, previously digested with *Hind* III and *Hinc* II. The amplified cDNA encoding the light chain variable region of 1E10 was digested with *Hind* III and *Eco* RV and ligated into M13mp19 vector, previously digested with *Hind* III and *Hinc* II. A standard ligation protocol[9] was performed using T4 DNA Ligase (Gibco-BRL, Gaithersburg, MD). XL1-Blue cells were transformed and plated for blue-white selection on X-Gal-IPTG plates. White colonies where picked and single strand M13mp19 DNA was purified by phenol extraction after precipitation of phage particles from cultures in liquid media.

Sequencing of the variable regions was performed by the dideoxy method using the T7 sequencing kit (Pharmacia Biotech). The M13 universal primer was used for sequencing. In addition, each variable region was separately amplified,

cloned, and sequenced at least twice to exclude possible errors introduced by Vent DNA polymerase.

*Assignment of the origin of the variable gene segments utilized*

Individual VH and VL sequences were searched against the EMBL/Genbank database for sequence homology with known murine genes. The germline gene with the closest homology was analyzed as a possible donor sequence of the rearranged gene. However, in some cases it was impossible to find a germline gene close enough to the analyzed sequence. In those cases we searched for the closest variable region sequence reported in EMBL/Genbank. The $J_K$ and $J_H$ regions sequenced were compared with the five known murine $J_K$ genes and with the four known murine $J_H$ genes.[10] The analysis of the junctional sequences in VH sequences was performed as Bangs.[11] D region nucleotides were identified after sequence alignment of heavy chain CDR3 (H-CDR3) nucleotide sequences with known murine D minigenes.[10] When junctional nucleotides could have been contributed by either D or the adjacent $V_H$ or $J_H$ segment, they were assigned to $V_H$ or $J_H$. In cases in which no $V_H$ or $J_H$ nucleotides were deleted during recombination with the D gene, possible P additions[12] were identified.

## RESULTS AND DISCUSSION

*Analysis of VH gene usage*

The nucleotide sequences and the deduced amino acid sequences of the heavy chain variable region of P3 MAb (Ab1) and Ab2 antibodies to P3, 1E10 and 3B11, are shown in Figs. 1 and 2, respectively, and their gene segment usage is summarized in Table 1.

VH P3 (sequence shown in Fig. 1) is a member of the Q52 (VH II) gene family and is 96% identical in nucleotide sequence to the related germline PJ14 gene from the same family.[13] The most related sequence to VH P3 is the cDNA sequence of asws1 (IgG2a, k), a mouse antibody generated in A.SW/SnJ (H-$2^S$) strain specific to U3 and U8 ribonucleoproteic particles.[14] There are only two nucleotide differences between P3 and asws1 (VH from codons 2 to 94), leading to only one amino acid replacement (56: Ser → Asn). This suggests that both

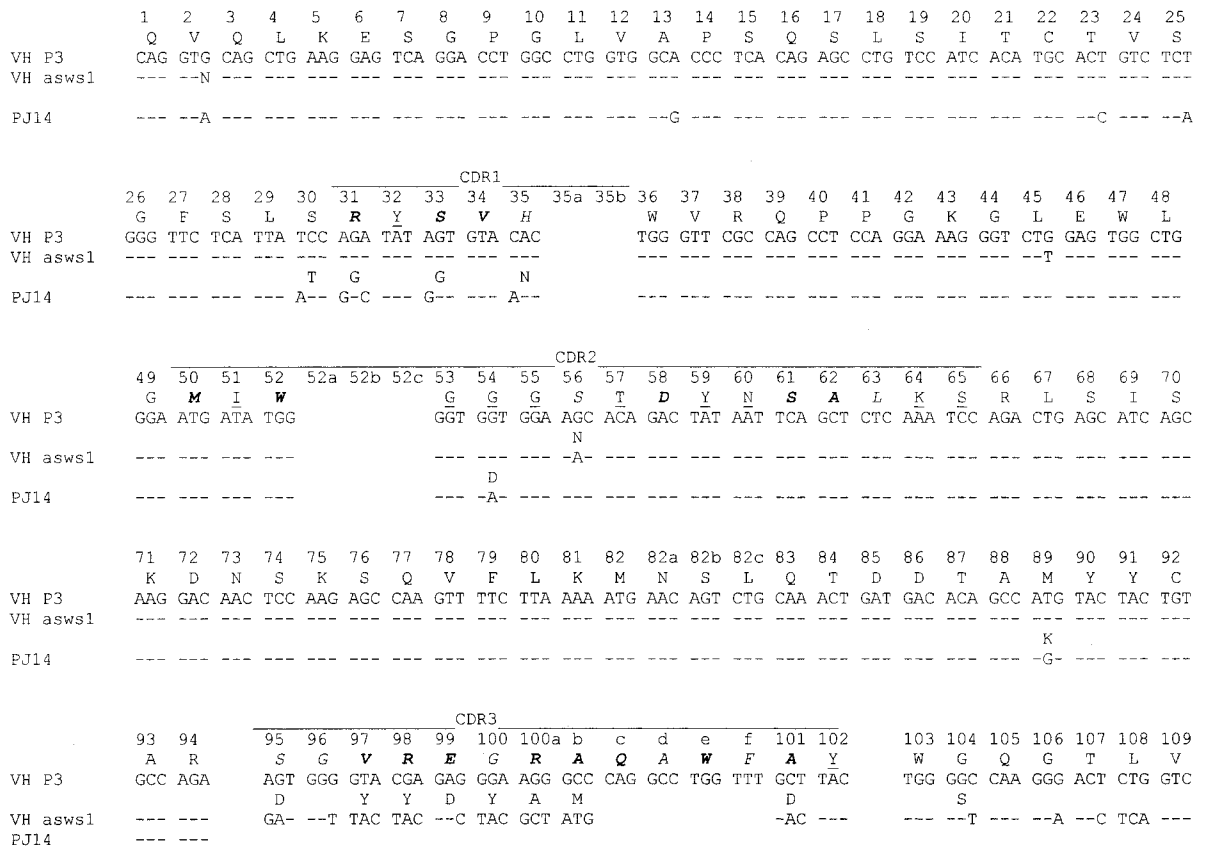


**FIG. 1.** Nucleotide and deduced amino acid sequences of the mRNA encoding the heavy chain variable region of P3 MAb (VH P3) compared with the most closely related germline and cDNA sequences (see text). Amino acids are numbered according to Kabat et al.[10] Spaces have been introduced to maximize alignment. Dashed lines indicate nucleotide sequence identities. N signifies undetermined bases. The amino acid residue encoded by each codon is given above the nucleotide sequence. CDR amino acids are represented according to the frequency of appearance in each position in a compilation of 1474 murine $V_H$ sequences extracted from Kabat database. Frequent amino acids (i.e., with frequencies >20%) are underlined, amino acids with frequencies between 10 and 20% are represented in italics; infrequent amino acid (i.e., frequencies <10%) are represented in bold italics.


**FIG. 2.** Nucleotide and deduced amino acid sequences of the mRNA encoding the heavy chain variable regions of anti-idiotypes 1E10 (VH 1E10) and 3B11 (VH 3B11) compared with the most closely related germline and cDNA sequences (see text). (For details, see legend to Fig. 1. X means unknown amino acid.)

TABLE 1. GENE SEGMENT USAGE IN P3 MAb AND Ab2 ANTIBODIES

| *MAb* | *Isotype* | $V_H$ | *D* | $J_H$ | $V_\kappa$ | $J_\kappa$ |
|---|---|---|---|---|---|---|
| P3 | IgM, κ | Q52 (II) | DSP2 | $J_H3$ | $V_\kappa19$ | $J_\kappa1$ |
| 1E10 | $IgG_1$, κ | J558 (I) | DSP2.2+ DSP2 | $J_H2$ | $V_\kappa10$ | $J_\kappa1$ |
| 3B11 | $IgG_1$, κ | 36–60 (III) | DSP2.2 | $J_H3$ | $V_\kappa4/5$ | $J_\kappa5$ |

$V_H$P3 and $V_H$asws1 may be encoded by a germline element other than PJ14. The amino acid $Ser_{56}$ encoded by VH P3 is a conserved amino acid that appears in all the germline VH sequences from Q52 family reported until 1997.[15] Furthermore, divergence nucleotides in codons 45 and 56 of VH P3 are also present in a high percent (95.6 and 82%, respectively) of the closest 23 sequences (94% or higher identity) extracted from Genbank database with the only exception of VH asws1. This result strongly suggests that VH P3 could represent a new undescribed gene from the Q52 family in BALB/c mice.

$V_H$Q52 gene family has been involved in the antibody response against some gangliosides. Zenita,[16] in a Northern blot analysis of the $V_H$ segments used by a large panel of anti-ganglioside antibodies, found that $V_H$ Q52 genes were preferably used by antibodies directed to gangliosides with sialic acid in $\alpha(2 \rightarrow 6)$ linkage. However, in addition to the present report, the usage of genes from the $V_H$ Q52 family have been previously observed in other autoantibodies directed to $\alpha(2 \rightarrow 3)$ sialic acid-containing gangliosides like GM1[17] and GM2 (G. Mustelier, manuscript in preparation). In BALB/c mice, the $V_H$ Q52 genes are highly represented in the fetal B-cell repertoire, due to the preferential usage of these genes by $CD5^+$ B lymphocytes (B-1 cells), the predominant B-cell subpopulation at this period.[18,19] Furthermore, it is known that this preference for D-proximal gene families like Q52 and 7183 is also observed in adult's B-1 cells.[19,20] The frequent appearance of $V_H$ genes from this family in anti-ganglioside antibodies might be due to the involvement of B-1 cells in their production.

VH 1E10 (Fig. 2) is a member of the J558 (VH I) family, the most $J_H$-distal group of $V_H$ genes in BALB/c mice. Although it has 93% identity from codons 4 to 94 to V102.1, a germline $V_H$ from the V NP group of genes in the same family,[21] its $V_H$ segment is almost identical to the one present in H35-C7, a hybridoma generated in BALB/c that produces an antibody to the influenza virus hemagglutinin.[22] There is only one amino acid difference at the position 94 (Arg → Asp), that possibly was originated during the recombination process. This strongly suggests that 1E10 and H35-C7 use the same germline $V_H$ gene from J558 gene family, but different to the V102.1 gene, defining most likely a new $V_H$ gene segment.

When compared with germline $V_H$ gene segments, the cDNA sequence of VH 3B11 showed 89% nucleotide identity from codons 1 to 94 with VH EA7 (Fig. 2), a member of the 3660 (VH III) gene family.[23] There are 21 nucleotide positions that differ between both sequences that appear all along the VH segments. The most homologous sequence to VH 3B11 is the cDNA sequence from MRB9, a polyrreactive murine MAb that binds to purified total histones.[24] Both sequences share a 95% homology, having differences in 11 nucleotide positions (codons 1 to 94). According to this analysis $V_H$ 3B11 belongs

```
                1   2   3   4   5   6   7   8   9  10  11  12  13  14  15  16  17  18  19  20  21  22  23  24  25
                 D   I   V   M   T   Q   S   H   K   F   M   S   T   S   V   G   D   R   V   S   I   T   C   K   A
VkP3            GAC ATT GTG ATG ACC CAG TCT CAC AAA TTC ATG TCC ACA TCA GTA GGA GAC AGG GTC AGC ATC ACC TGC AAG GCC
VkH250-6        --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---

                    CDR1
               26  27  28  29  30  31  32  33  34  35  36  37  38  39  40  41  42  43  44  45  46  47  48  49  50
                S   Q   D   V   S   T   A   V   A   W   Y   Q   Q   K   P   G   Q   S   P   K   L   L   I   Y   S
VkP3            AGT CAG GAT GTG AGT ACT GCT GTA GCC TGG TAT CAA CAG AAA CCA GGA CAA TCT CCT AAA CTA CTG ATT TAC TCG
VkH250-6        --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---

                    CDR2
               51  52  53  54  55  56  57  58  59  60  61  62  63  64  65  66  67  68  69  70  71  72  73  74  75
                A   S   Y   R   Y   T   G   V   P   D   R   F   T   G   S   G   S   G   T   D   F   T   F   T   I
VkP3            GCA TCC TAC CGG TAC ACT GGA GTC CCT GAT CGC TTC ACT GGC AGT GGA TCT GGG ACG GAT TTC ACT TTC ACC ATC
VkH250-6        --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---

                                                                                    CDR3
               76  77  78  79  80  81  82  83  84  85  86  87  88  89  90  91  92  93  94  95  96  97  98  99 100
                S   S   V   Q   A   E   D   L   A   V   Y   Y   C   Q   Q   H   Y   S   T   P   W   T   F   G   G
VkP3            AGC AGT GTG CAG GCT GAA GAC CTG GCA GTT TAT TAC TGT CAG CAA CAT TAT AGT ACT CCG TGG ACG TTC GGT GGA
VkH250-6        --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---

               101 102 103 104
                G   T   K   L
VkP3            GGC ACC AAG CTG
VkH250-6        --- --- --- ---
```

**FIG. 3.** Nucleotide and deduced amino acid sequences of the mRNA encoding the light chain variable region of P3 MAb (VkP3) compared with the most closely related cDNA sequence (VkH250-6). (For details, see legend to Fig. 1.)

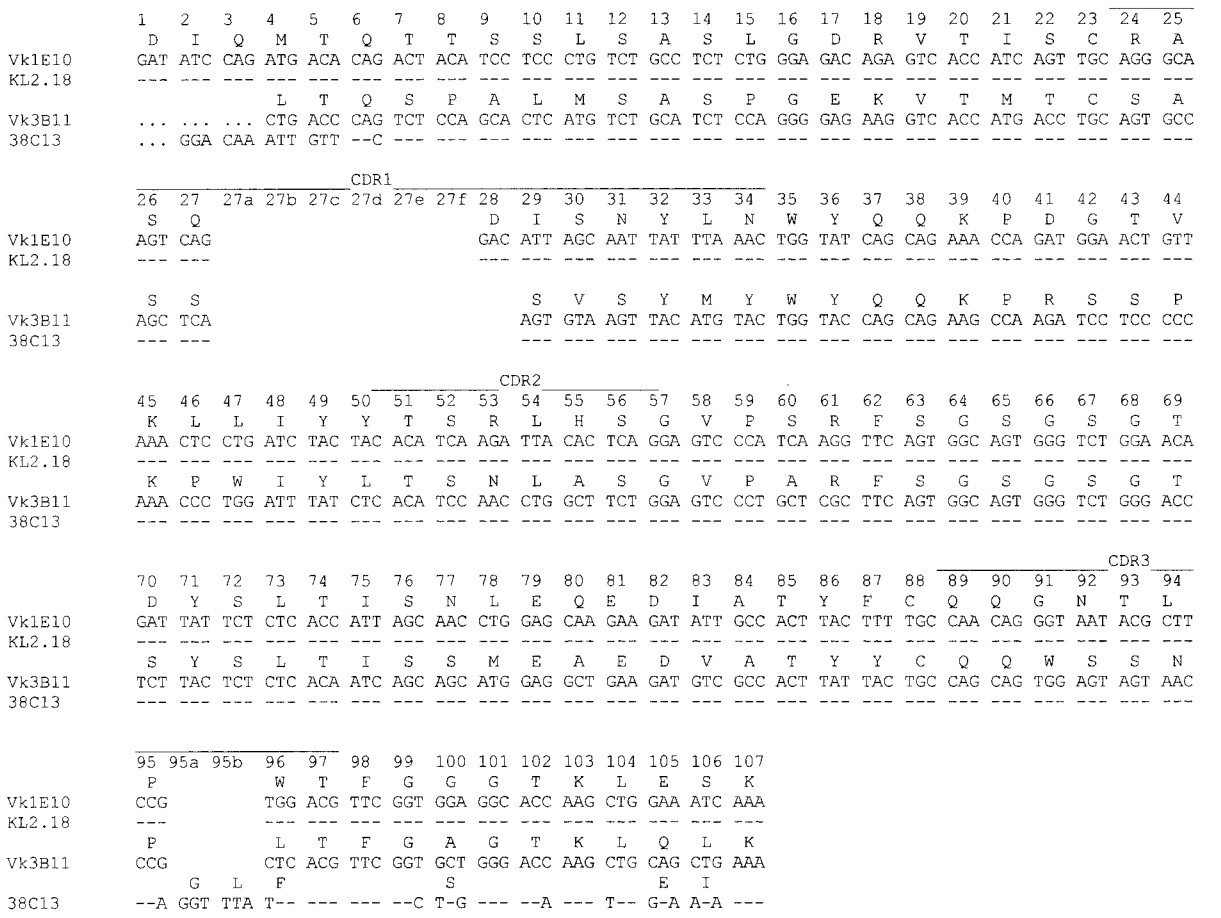


**FIG. 4.** Nucleotide and deduced amino acid sequences of the mRNA encoding the light chain variable regions of anti-idiotypes 1E10 (Vk1E10) and 3B11 (Vk3B11) compared with the most closely related germline and cDNA sequences (see text). [For details, see legend to Fig. 1.)

to the 3660 ($V_H$III) family. Therefore, γ-type anti-idiotypic antibodies to the same Ab1 (P3 MAb) can be obtained from different $V_H$ gene families.

### *Analysis of VL gene usage*

The light chain variable region cDNA sequences of the three MAbs presented here are derived from different $V_\kappa$ gene families (Table 1). The VL sequence of P3 belongs to the $V_\kappa$19 family wand is 100% identical to one of its members, H250-6, an anti-influenza virus hemagglutinin MAb originated from a BALB/c mouse[22] (Fig. 3). This identity involves both the use of the same $V_\kappa$ gene segment (codons 1–95) and of the same $J_\kappa$ gene segment (Jκ1, codons 96–107, Fig. 5A). The fact that the light chain variable region of P3 MAb is fully identical to the one corresponding to H250-6 MAb, suggests that P3 H chain probably contributes more than the L chain to the combining site to gangliosides.

The VL 1E10 sequence is a member of the *V*κ10 family and uses the V-IdCR gene segment. VL 1E10 sequence is identical to that present in KL2.18, KL4C11, and KL4B10 MAbs (M63608, M63617, and M63615, respectively, in Genbank database). This VL sequence is associated with the antibody response against *p*-azophenylarsonate in A mice (CAL-20xA/J)F1.[25] VL 1E10 is 100% identical to VL KL2.18 in codons 1 to 107, implying the use of the same $V_\kappa$ and $J_\kappa$ ($J_\kappa$1) gene segments.

The VL sequence of 3B11 belongs to Vk-Ox1 superfamily.[26] It is 99.3% identical in nucleotides to the VL of 38C13, an antibody with unknown specificity produced by a murine B-cell lymphoma of the same name.[27] The $V_\kappa$ segment sequences (codons 1–95) of both antibodies differ only in two nucleotides that do not lead to amino acid replacements. In VL 3B11, $J_\kappa$5 is used (Fig. 5), and FR4 sequence differs from the sequence of this gene by one nucleotide leading to one amino acid replacement ($Glu_{105} \rightarrow Gln_{105}$). In 38C13, $J_\kappa$4 is used.

In conclusion, based on the identity of the light chain variable regions either from 1E10 or 3B11 MAbs to previously reported sequences, one might expect a greater contribution of VHs than VLs to the Ab2 binding sites to P3 MAb.

### *Quantitative and qualitative analysis of somatic hypermutations*

Sequence analysis presented above of the heavy and light chain variable regions demonstrated the germline nature of P3

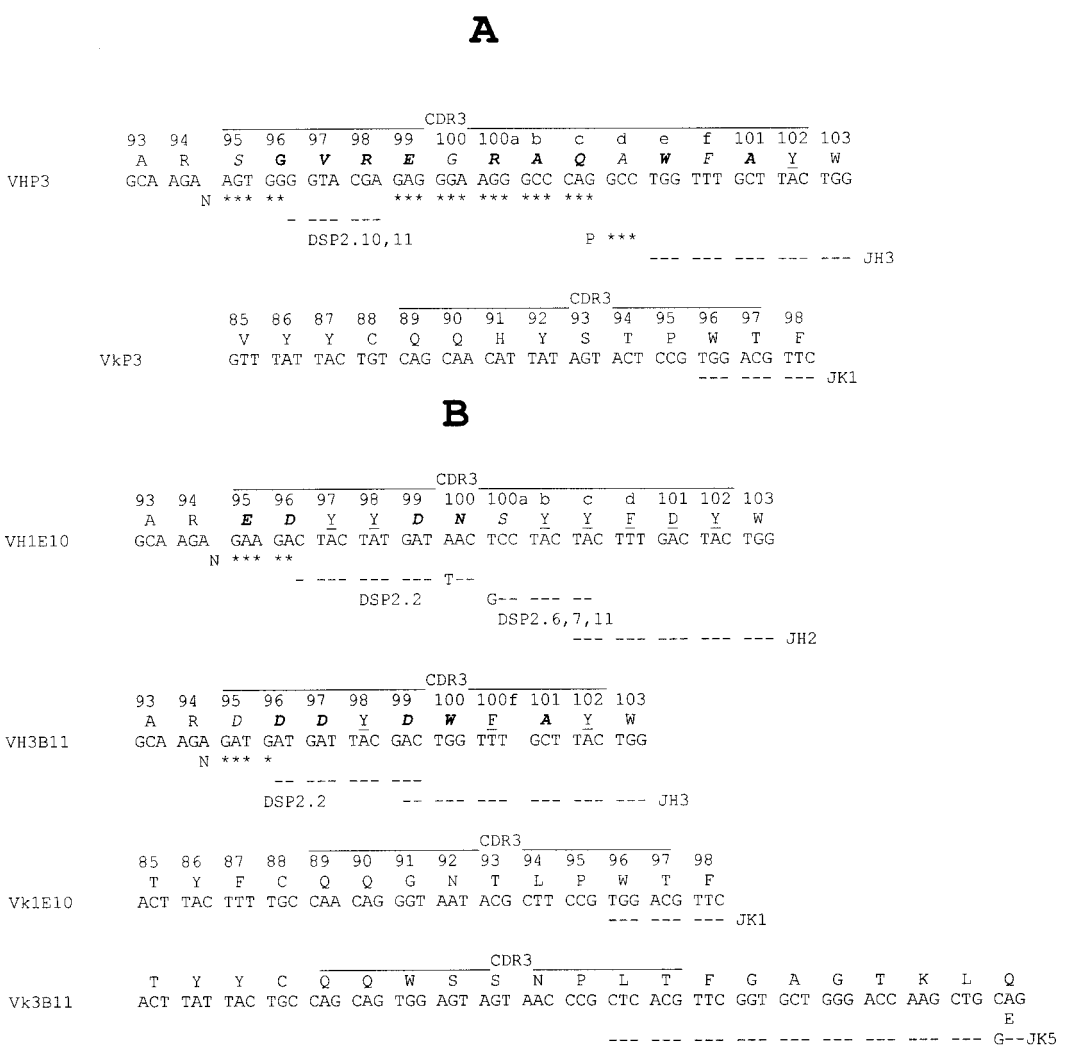


**FIG. 5.** Possible origins of heavy and light chain CDR3 of P3 (**A**) and anti-idiotypic MAbs 1E10 and 3B11 (**B**). P and N nucleotides are represented with asterisks. Dashed lines indicate homology with known D and J genes. The germline sequences of DSP2.2, DSP2.6 and DSP2.7, together with all the germline J sequences, are from Kabat et al.[10] The germline sequences DSP2.10 and DSP2.11 are from the International Immunogenetics Database (IMGT: *http://www.ebi.ac.uk/imgt/*). H-CDR amino acids are represented according to their frequencies in each position in a compilation of murine $V_H$ from Kabat database (http://immuno.bme.nwu.edu). (For details, see legend to Fig. 1.)

and 1E10 MAbs. On the other hand, 3B11 MAb showed features of extensive somatic mutations. As a replacement to silent mutation ratio (R/S) greater than 2.9 has been proposed to indicate a positive selection of nucleotide changes that affect amino acid sequence,[28] this ratio was calculated for the CDR and framework regions of $V_H$ 3B11. A quantitative and qualitative analysis of R and S mutations in $V_H$ 3B11 showed a pattern which is characteristic for hypermutated antibodies. There is a clear tendency to cluster replacement (and nonconservative) mutations in the H-CDRs, whereas all mutations in the frameworks tend to be silent or conservative. Whereas FR regions of $V_H$ 3B11 showed an R/S ratio of 2 ($4/2 < 2.9$), this ratio was 4 ($4/1 > 2.9$) for the CDRs. Furthermore, replacement mutations in FRs involved conservative amino acid changes, with the only exception of $Ser_{94} \rightarrow Arg_{94}$, most likely originated during $V_H$–D–$J_H$ recombination process. By contrary, all the replacement mutations that appear in the H-CDRs 1 and 2 involved nonconservative amino acid changes.

Other data supporting the evidence of somatic mutations in $V_H$ 3B11 came from comparisons with the closest homologous sequences reported in Genbank. Nucleotides in eight out of ten putative mutated positions of $V_H$ 3B11 were found exclusively in this sequence and only two nucleotides, those in codons 54 and 88, were found to be shared by most VH sequences (90% and 80% of the sequences with 94% or higher identity, respectively), suggesting their germline origin.

In addition, a replacement mutation in FR4 of VL 3B11, confirms the evidence of somatic mutations given above. However, it is striking to find such a great difference between the levels of mutations in both chains. Previous studies have shown that, in general, somatic mutations in H- and L-chain variable regions tend to occur concurrently.[29] Nevertheless, in some re-

ports it is possible to find differences in the levels of mutations between the heavy and light chains.[30,31] This has been interpreted as an evidence of independent activity of the hypermutation machinery in each chain.[31]

Although 3B11 and 1E10 MAbs were generated in the same fusion experiment and both are able to inhibit P3 MAb binding to gangliosides, immunochemical studies have shown that both Ab2s can be differentiated by their binding affinities to Ab1. 3B11 MAb, the one that showed the higher affinity, probably as a consequence of an affinity muturation process, seems to define a "private" idiotope. In fact, 3B11 was unable to induce in syngeneic model anti-anti-idiotype antibodies that shared idiotopes with P3 MAb, in contrast to the germline-encoded 1E10 MAb that seems to define a "regulatory" idiotope.[7]

### *Analysis of junctional regions*

Only 7 nucleotides out of 42 that codify the H-CDR3 of P3 MAb belong to a D region (Fig. 5A). Either DSP2.10 or DSP2.11 gene segments used in reading frame 3 could have contributed to this region. N nucleotides constitute almost 50% of the sequence of this H-CDR3, and they appear at both sides of the D segment. In particular, the 3′N region, which is more frequently absent or shorter than 5′N region[11,36] is unusually extensive in this H-CDR3 with 15 nucleotides. Even when this number is higher than average additions catalyzed by TdT enzyme (4–5),[32] the observed (G+C)/(A+T) ratio (3) is characteristic of the activity of this enzyme.[33] The last five amino acids in H-CDR3 of P3 MAb are codified by P nucleotides[12] and $J_H3$.

The D segment of 1E10 has 21 nucleotides and is more likely to be derived from fusion of thirteen nucleotides of DSP2.2 and seven nucleotides from another member of DSP2 family (either DSP2.6, DSP2.7, or DSP2.11, could have contributed). Both minigenes are used in reading frame 1. The H-CDR3 of 1E10 has a five nucleotides long 5′N region, whereas 3′N region is absent.

The D segment of 3B11 can be readily assigned to the DSP2.2 germline gene (Fig. 5B), which also contributed to the D region in 1E10. Nine nucleotides in H-CDR3 of 3B11 belong to this D region, for which reading frame 1 is used. As it happens in H-CDR3 of 1E10, the H-CDR3 of 3B11 lacks 3′N nucleotides. There are 4 nucleotide additions at 5′ of the D segment, and D and JH segments overlap two nucleotides at junction.

### *Analysis of CDR amino acid sequences of P3 MAb*

To characterize the residues found in the CDRs of our antibodies, two approaches were used. First, we classified residues in each position in the H-CDRs according to their frequencies of appearance in a compilation of murine VH sequences extracted from Kabat database (http://immuno.bme.nwu.edu). These frequencies can be retrieved from web site http://bioinfo.cim.sld.cu/kabat.html. In the second approach, we calculated and compared the frequency of certain residues in H- and L-CDRs between our antibodies and a compilation of sequences obtained and processed by Brard[34] from Kabat database.[10] The major observation is that several key positions, mainly distributed among the H-chain CDRs, present unusual charged residues.

Previously published studies suggested the leading role of electronegative groups from gangliosides (carboxyl and hydroxyl groups in *N*-glycolyl neuraminic acid) and sulfated glycolipids ($SO_3^-$ group) in the binding of P3 MAb to these antigens.[6] These findings first suggested the existence of complementary positive-charged residues in the CDRs of P3 MAb able to interact with those electronegative determinants. In fact, VH P3 has a slightly higher content of basic (R:3, K:1) than acidic (D:1, E:1) amino acid residues (Fig. 1). Analysis of positional frequencies of residues in H-CDRs reveals that arginine is infrequently found in positions 31(4.8%), 98(4.2%), and 100a(2.8%). The presence of lysine at position 64 (CDR2) does not constitute an unusual feature of P3 MAb because it is observed in 81% of murine VH sequences in Kabat's compilation, and this amino acid could possibly play a structural role.[35] The presence of arginines at unusual positions and with higher frequencies than those observed in Kabat's compilation (Table 3), has a great relevance in the comprehension of the interaction of P3 MAb and its ligands. Arginines in H-CDR1 and 3, and L-CDR2 of P3 MAb able to establish ionic interactions as well as hydrogen bonds could be possibly involved in the recognition of these structures.

Another interesting feature of P3 MAb amino acid sequence is its H-CDR3. This region contains a high number of positions with infrequent amino acids (57% of the sequence), given in part by random N nucleotides. In addition, whereas several studies have shown the preferred use of D segments in the reading frame 1, which generates tyrosine-rich sequences,[36,37] the short D segment used in H-CDR3 P3 is translated in reading frame 3, and codifies two of the infrequent amino acids found at this region. The infrequent amino acid-rich sequence of the H-CDR3 in a germline antibody like P3 MAb, could be possibly related to the ability of this MAb to induce in syngeneic model an IgG anti-idiotypic response in the absence of carrier protein or adjuvant.[7]

### *Analysis of CDR amino acid sequences of Ab2 MAbs*

The major finding obtained from 1E10 and 3B11 sequence analysis was the presence of homologous acidic amino acid-rich sequences in their H-CDR3 regions (Fig. 2), in contrast with the low homology observed in their variable region gene segments. The acidic motif, E/D–D–Y/D–Y–D, is codified by N nucleotides between $V_H$ and D, and also by the first D region nucleotides. Both Ab2 share the DSP2.2 minigene, which codifies the last residues in the motif. This is the only murine D minigene that could contribute with two aspartic residues to the CDR3 sequence. During assembly of the heavy chain variable region, a selected D gene element is usually trimmed by an exonuclease to give rise to a core sequence that represents only part of the initial minigene, and random nucleotides are added to one or both ends of this core sequence. The presence of two codons for aspartic acid in DSP2.2 could therefore increase the probability of at least one codon remaining after mutilation. This could explain the appearance of this minigene in both H-CDR3 regions, in view of the importance that acidic residues seem to have in the binding sites of 3B11 and 1E10 MAbs.

As it is shown in Table 2, acidic residues in H-CDR3 of Ab2 antibodies constitute the 22.7%, percentage higher than the av-

TABLE 2. FREQUENCIES OF SELECTED AMINO ACIDS IN KABAT CDRs Ab2[a]

| | $CDR1_H$ | | | $CDR2_H$ | | | $CDR3_H$ | | |
|---|---|---|---|---|---|---|---|---|---|
| | *Ab2* | *Kabat* | Δ | *Ab2* | *Kabat* | Δ | *Ab2* | *Kabat* | Δ |
| R/K | 0.0 | 1.7 | −1.7 | 15.1 | 10.6 | 4.5 | 0.0 | 5.0 | −5.0 |
| D/E | 9.1 | 8.1 | 1.0 | 12.1 | 8.3 | 3.8 | 38.1 | 15.4 | 22.7 |
| Y | 27.3 | 20.9 | 6.4 | 6.1 | 11.6 | −5.5 | 33.3 | 23.7 | 9.6 |
| F/W | 9.1 | 8.4 | 1.3 | 9.1 | 4.6 | 4.5 | 14.3 | 4.5 | 9.8 |
| Q/N | 18.2 | 7.8 | 10.4 | 12.1 | 12.3 | −0.2 | 4.8 | 3.6 | 1.2 |
| S/T | 18.2 | 17.8 | 0.4 | 15.1 | 18.9 | −3.8 | 4.8 | 13.0 | −8.2 |
| | $CDR1_L$ | | | $CDR2_L$ | | | $CDR3_L$ | | |
| | *Ab2* | *Kabat* | Δ | *Ab2* | *Kabat* | Δ | *Ab2* | *Kabat* | Δ |
| R/K | 4.3 | 6.9 | −2.6 | 7.1 | 14.0 | −6.9 | 0.0 | 2.7 | −2.7 |
| D/E | 4.3 | 4.1 | 0.2 | 0.0 | 7.0 | −7.0 | 0.0 | 4.5 | −4.5 |
| Y | 13.0 | 7.4 | 5.6 | 7.1 | 3.0 | 3.9 | 0.0 | 8.1 | −8.1 |
| F/W | 0.0 | 0.8 | −0.8 | 0.0 | 2.6 | −2.6 | 7.1 | 7.4 | −0.3 |
| Q/N | 13.0 | 12.3 | 0.7 | 7.1 | 9.3 | −2.2 | 42.8 | 18.8 | 24 |
| S/T | 34.8 | 36.6 | −1.8 | 42.9 | 33.4 | 9.5 | 21.4 | 25.0 | −3.6 |

[a]Ab2 stands for 1E10 and 3B11 MAbs. Data in Kabat were taken from Brard et al.[38] Kabat stands for 1450 sequences compiled by Kabat et al.[10] The third column (Δ) corresponds to the difference between the first two columns.

erage representation in Kabat's H-CDR3s, whereas basic residues are slightly under-represented. In addition, acidic residues in this CDR are not usually found in their positions in Kabat's compilation (with the only exception of $Asp_{101}$ in VH 1E10, which is codified by the $J_H$ gene segment). Acidic and basic amino acid residues are found in CDRs 1 and 2 of Ab2 antibodies with levels close to those in Kabat's compilation. However, the presence of unusual acidic residues like $Asp_{31}$ (CDR1) and $Asp_{58}$ (CDR2) in VH 1E10 and VH 3B11, respectively, and $Asp_{54}$, which appears in a shared motif DGS (Fig. 3) is noticeable.

In contrast, the light chain CDRs (L-CDRs) of Ab2 present few unusual features. There are no differences in charged residue levels between these L-CDRs and Kabat's compilation (Table 2), supporting the previous suggestion of the major contribution of the Ab2 heavy chains to the binding site to P3 MAb. The only striking deviation from Kabat's compilation appears in the levels of asparagine and glutamine residues in L-chain CDR3s (24% higher than L-chain CDR3 frequency in Kabat's compilation).

Early studies in the field of antigen–antibody interactions have shown that cationic antigens develop anionic antibody re-

TABLE 3. FREQUENCIES OF SELECTED AMINO ACIDS IN KABAT CDRs P3[a]

| | $CDR1_H$ | | | $CDR2_H$ | | | $CDR3_H$ | | |
|---|---|---|---|---|---|---|---|---|---|
| | *P3* | *Kabat* | Δ | *P3* | *Kabat* | Δ | *P3* | *Kabat* | Δ |
| R/K | 20.0 | 1.7 | 18.3 | 6.25 | 10.6 | −4.3 | 14.3 | 5.0 | 9.3 |
| D/E | 0.0 | 8.1 | −8.1 | 6.25 | 8.3 | −2.0 | 7.1 | 15.4 | −8.3 |
| Y | 20.0 | 20.9 | −0.9 | 6.25 | 11.6 | −5.3 | 7.1 | 23.7 | −17 |
| F/W | 0.0 | 8.4 | −8.4 | 6.25 | 4.6 | 1.7 | 14.3 | 4.5 | 9.8 |
| Q/N | 0.0 | 7.8 | −7.8 | 0.0 | 12.3 | −12 | 7.1 | 3.6 | 3.5 |
| S/T | 20.0 | 17.8 | 2.2 | 25.0 | 18.9 | 6.1 | 7.1 | 13.0 | −5.9 |
| | $CDR1_L$ | | | $CDR2_L$ | | | $CDR3_L$ | | |
| | *P3* | *Kabat* | Δ | *P3* | *Kabat* | Δ | *P3* | *Kabat* | Δ |
| R/K | 9.1 | 6.9 | 2.2 | 14.3 | 14.0 | 0.3 | 0.0 | 2.7 | −2.7 |
| D/E | 9.1 | 4.1 | 5.0 | 0.0 | 7.0 | −7.0 | 0.0 | 4.5 | −4.5 |
| Y | 0.0 | 7.4 | −7.4 | 28.6 | 3.0 | 25.3 | 14.3 | 8.1 | 6.2 |
| F/W | 0.0 | 0.8 | −0.8 | 0.0 | 2.6 | −2.6 | 0.0 | 7.4 | −7.4 |
| Q/N | 9.1 | 12.3 | −3.2 | 0.0 | 9.3 | −9.3 | 28.6 | 18.8 | 9.6 |
| S/T | 27.3 | 36.6 | −9.3 | 42.9 | 33.4 | 9.5 | 28.6 | 25.0 | 3.6 |

[a]P3 stands for P3 MAb. Data in Kabat were taken from Brard et al.[38] Kabat stands for 1450 sequences compiled by Kabat et al.[10] The third column (Δ) corresponds to the difference between the first two columns.

sponses, whereas the immunization with anionic antigens induces cationic antibodies.[38] Variable region sequences analysis of anti-histones[34,39] and anti-DNA antibodies,[40] have strongly supported these initial observations. Therefore, the finding of an acidic motif in H-CDR3 of 1E10 and 3B11 Ab2s, and the general excess in acidic amino acids in their H-CDRs suggest that electrostatic interactions with complementary positive residues in the hypervariable loops of P3 MAb, are important in $\gamma$-type Ab2–Ab1 interactions.

Our results prompt us to suggest that H-CDR3 residues from 1E10 and 3B11 Ab2 antibodies bind to residues in the H-CDR3 of P3 MAb. This could seem contradictory in view of the lack of mimicry shown by both Ab2 in syngeneic model.[7] However, the study of the crystal structure of the complex between the Fv fragments of the anti-lysozyme antibody D1.3 and the anti-D1.3 antibody E5.2 (Ab2$\beta$),[41] has shown that the mimicking is functional and involves similar binding interactions rather than exact topological replicas. This concept is particularly relevant to explain the ability of certain antibodies to mimic gangliosides and other carbohydrate-bearing molecules. There are several plausible explanations to be considered for the failure of our anti-idiotypic antibodies to mimic *N*-glycolyl-containing gangliosides. As Fields pointed out,[41] molecular mimicking depends on several factors like the structure of the antigen, the choice of the antibody (Ab1) and of anti-idiotypic antibody, in addition to other numerous complexities of immune responses. Finally, it could be also the case that even when our Ab2 carried the internal image of these antigens, they could not break the tolerance established against these widespread self-structures in mouse. Interestingly, in cancer patients, where *N*-glycolyl-containing gangliosides are nonself antigens, the immunizations with 1E10 MAb induced a specific response against these antigens (M. Alfonso, manuscript in preparation).

## REFERENCES


1. Higashi H, Naiki M, Matuo S, and Okouchi K: Antigen of "serum sickness" type of heterophile antibodies in human sera: Identification as ganglioside with N-glycolylneuraminic acid. Biochem Biophys Res Commun 1977;79:388–395.
2. Hirabagashi Y, Kasabura H, Matsumoto M, Higashi H, Kato S, Kasai N, and Naiki M: Specific expression of unusual GM2 ganglioside with Hanganutziu-Deicher antigen activity on human colon cancer. Jpn J Cancer Res (Gann) 1987;78:251–260.
3. Hirabagashi Y, Higashi H, Kato S, Taniguchi M, and Matsumoto M: Occurrence of tumor associated ganglioside antigens with Hanganutziu-Deicher antigenic activity on human melanomas. Jpn J Cancer Res (Gann) 1987;78:614–620.
4. Marquina G, Waki U, Fernández LE, Kon K, Carr A, Valiente O, Pérez R, and Ando S: Gangliosides expressed in human breast cancer. Cancer Res 1996;56:5165–5170.
5. Vázquez AM, Alfonso M, Lanne B, Karlsson KA, Carr A, Barroso O, Fernández LE, Rengifo E, Lanio ME, Alvarez C, Zeuthen J, and Pérez R: Generation of a murine monoclonal antibody specific for N-glycolylneuraminic acid-containing gangliosides that also recognizes sulfated glycolipids. Hybridoma 1995;14:551–556.
6. Moreno E, Lanne B, Vázquez AM, Kawashima I, Tai T, Fernández LE, Karlsson K-A, Ångström J, and Pérez R: Delineation of the epitope recognized by an antibody specific for N-glycolylneuramini c acid-containing gangliosides. Glycobiology 1998;8:695–705.
7. Vázquez AM, Pérez A, Hernández AM, Macías A, Alfonso M, Bombino G, and Pérez R: Syngeneic anti-idiotypic monoclonal antibodies to an anti-NeuGc-containing ganglioside monoclonal antibody. Hybridoma 1998;17:527–534.
8. Orlandi R, Gussow DH, Jones PT, and Winter G: Cloning immunoglobulin variable domains for expression by the polymerase chain reaction. Proc Natl Acad Sci U S A 1989;86:3833–3837.
9. Sambrook J, Fritsch EF, and Maniatis T: *Molecular cloning: A laboratory manual*, 2nd ed. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York, 1989.
10. Kabat EA, Wu TT, Peery HM, Gottesmann KS, and Foeller C: *Sequences of proteins of immunological interest*. 5th ed. U.S. Department of Health and Human Services, U.S. Government Printing Office, Washington, 1991.
11. Bangs LA, Sanz IE, and Teale JM: Comparison of D, $J_H$ and junctional diversity in the fetal, adult, and aged B cell repertoires. J Immunol 1991;146:1996–2004.
12. Lafaille JJ, DeCloux A, Bonneville M, Takagaki Y, and Tonegawa S: Junctional sequences of T cell receptor $\gamma\delta$ genes: implications for $\gamma\delta$ T cell lineages and for a novel intermediate of V-(D)-J joining. Cell 1989;59:859–870.
13. Sakano H, Maki R, Kurosawa Y, Roeder W, and Tonegawa S: Two types of somatic recombination are necessary for the generation of complete immunoglobulin heavy-chain genes. Nature 1980;286: 676–683.
14. Monestier M, Losman MJ, Novick KE, and Aris JP: Molecular analysis of Mercury-induced antinucleolar antibodies in H-$2^S$ mice. J Immunol 1994;152:667–675.
15. Almagro JC, Hernández I, Ramírez MC, and Vargas-Madrazo E: The differences between the structural repertoires in germline gene segments of mice and humans: implications for the molecular mechanism of the immune response. Mol Immunol 1997;34: 1199–1214.
16. Zenita K, Hirashima K, Shigeta K, Hiraiwa N, Takada A, Hashimoto K, Fujimoto E, Yago K, and Kannagi R: Northern hybridization analysis of $V_H$ gene expression in murine monoclonal antibodies directed to cancer-associated ganglioside antigens having various sialic acid linkages. J Immunol 1990;144:4442–4451.
17. Weng NP, Ritter E, Yucel E, Zhang D, Ritter G, and Marcus DMJ: Specificity and structure of murine monoclonal antibodies against GM1 ganglioside. Neuroimmunol 1994;55:61–68.
18. Jeong HD and Teale JM: Contribution of the CD5+ B cell to D-proximal VH family expression early in ontogeny. J Immunol 1990;145:2725–2729.
19. Casali P, Kasaian MT, and Haughton G: B-1 (CD5 B) cells. In: *Autoimmunity Physiology and Disease*. Coutinho A, and Kazatchkine MD (Eds.). Wiley-Liss Inc., New York, 1994, p. 57.
20. Kantor AB, Merrill CE, Herzenberg LA, and Hillson JL: An unbiased analysis ov V(H)-D-J(H) sequences from B-1a, B-1b, and conventional B cells. J Immunol 1997;158:1175–1186.
21. Blankenstein T, Zoebelein G, and Krawinkel U: Analysis of immunoglobulin heavy chain V-region genes belonging to the V NP-gene family. Nucleic Acid Res 1984;12:6887–5900.
22. Kavaler J, Caton AJ, Staudt LM, Schwartz D, and Gerhard W: A set of closely related antibodies dominates the primary antibody response to the antigenic site CB of the A/PR/8/34 influenza virus hemagglutinin. J Immunol 1990;145:2312–2321.
23. Robbins PF, and Nisonoff A: Characteristics of two idiotypic subfamilies of murine anti-p-azobenzenearsonate antibodies. J Immunol 1987;139:3868–3877.
24. Novick KE, Fasy TM, Losman MJ, and Monestier M: Polyreactive IgM antibodies generated from autoimmune mice and selected for histone-binding activity. Int Immunol 1992;4:1103–1111.
25. Wysocki LJ, Gridley T, Huang S, Grandea AG, and Gefter ML: Single germline VH and V kappa genes encode predominating antibody variable regions elicited in strain A mice by immunization with p-azophenylarsonate. J Exp Med 1987;166:1–11.

26. Kaartinen M, Pelkonen E, Even J, and Makela O: V genes of the primary antibody response of C57BL/10 mice to the hapten phenyloxazolone. Eur J Immunol 1988;18:1095–1100.
27. Roth MS, Weiner GJ, Allen EA, Terry VH, Harnden CE, Boehnke M, Kaminsk MS, and Ginsburg D: Molecular characterization of anti-idiotype antibody resistant variants of a murine B cell lymphoma. J Immunol 1990;145:768–777.
28. Shlomchik MJ, Marshak-Rothstein A, Wolfowicz CB, Rothstein TL, and Weigert MG: The role of clonal selection and somatic mutation in autoimmunity. Nature 1987;328:805–811.
29. Milstein C: From the structure of antibodies to the diversification of the immune response. Scand J Immun 1993;37:386–398.
30. Chen J, Borden P, Liao J, and Kabat EA: Variable region cDNA sequences of three mouse monoclonal anti-idiotypic antibodies specific for anti-$\alpha(1 \rightarrow 6)$ dextrans with groove- or cavity-type combining sites. Mol Immunol 1992;29:1121–1129.
31. David D, and Zouali M: Variable region light chain genes encoding human antibodies to HIV-1. Mol Immunol 1995;32:77–88.
32. Mak TW, and Simard JJL: B-Lymphocyte genes. In: *Handbook of Immune Response Genes*. Plenum Press, New York 1998, pp. 109–70.
33. Meek K, Hasemann C, Pollock B, Alkan SS, Brait M, Slaoui M, Urbain J, and Capra JD: Structural characterization of antiidiotypic antibodies: Evidence that Ab2s are derived from the germline differently than Ab2s. J Exp Med 1989;169:519–533.
34. Brard F, Petit S, Coquerel G, Gilbert D, Koutouzov S, Pérez G, and Tron F: Modeling of anti-nucleosome immunoglobulin Fv domains: Analysis of electrostatic interactions. Mol Immunol 1997; 34:793–807.
35. Padlan EA: Anatomy of the antibody molecule. Mol Immunol 1994;31:169–217.
36. Ichihara Y, Hayashida H, Miyazawa S, and Kurosawa Y: Only DFL16, DSP2, and DQ52 gene families exist in mouse immunoglobulin heavy chain diversity gene loci, or which DFL16 and DSP2 originate from the same primoridal DH gene. Eur J Immunol 1989;19:1849–1854.
37. Gu H, Foster I, and Rajewsky K: Sequence homologies, N sequence insertion and JH gene utilization in VHDJH joining: implications for the joining mechanism and the ontogenetic timing of Lyl B cell and B-CLL progenitor generation. EMBO J 1990;9:2133–2140.
38. Sela M, and Mozes E: Dependence of the chemical nature of antibodies on the net electrical charge of antigens. Proc Natl Acad Sci USA 1967;55:445.
39. Monestier M, Fasy JM, Losman MJ, Novick KE, and Muller S: Structure and binding properties of monoclonal antibodies to core histones from autoimmune mice. Mol Immunol 1993;30: 1069–1075.
40. Eilat D, and Anderson WF: Structure-function correlates of autoantibodies to nucleic acids. Lessons from immunochemical, genetic and structural studies. Mol Immunol 1994;31:1377–1390.
41. Fields BA, Goldbaum FA, Ysern X, Poljak RJ, and Mariuzza RA: Molecular basis of antigen mimicry by an anti-idiotope. Nature 1995;374:739–742.

Address reprint requests to:
*Alexis Pérez*
*Department of Antibody Engineering*
*Center of Molecular Immunology*
*P.O. Box 16040*
*Havana 11600*
*Cuba*

*E-mail:* alexis@ict.cim.sld.cu

Received for publication October 23, 2000. Accepted after revisions April 21, 2001.