HEAVY CHAIN VARIABLE REGION

Multiple Gene Segments Encode Anti-4-(hydroxy-3-nitrophenyl)acetyl Idiotypic Antibodies

By MARTINA E. BOERSCH-SUPAN, SADHANA AGARWAL, MARY E. WHITE-SCHARF, AND THEREZA IMANISHI-KARI

From the Center for Cancer Research and Department of Biology, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139

The antigen binding specificity of an antibody molecule is determined by the amino acid sequences of the variable (V)[1] regions of the heavy (H) and light (L) chains. The antibody repertoire of an individual is large and complex. The studies of antibody polypeptides and genes in mice have suggested three sources of diversity: multiple germline V genes (1–4), rearrangement of three sets of H chain (V_H, D [diversity], and J_H [joining region of the H chain]) and two sets of L chain (V_L and J_L) germline segments (5–7); and two types of somatic alterations (7–9 and 10–12). Several questions, however, remain to be answered: (*a*) What are the relative contributions of the germline, combinatorial, and somatic variation mechanisms to the functional repertoire of antibody molecules? (*b*) Is somatic diversification a random process happening at the base somatic mutation rate, or is it a consequence of a V gene–specific mutation mechanism (13, 14)? (*c*) At what stage of B cell differentiation does somatic mutation take place?

A better understanding of this problem requires a more detailed study of the heterogeneity of the response to specific antigens. As amino acid sequences of antibodies (15, 16) and nucleotide sequences (2, 11, 17, 18) of V_H genes of restricted heterogeneity (idiotypes) have been obtained, the notion has emerged that somatic point mutation and combinatorial diversity of a single germline V_H gene are responsible for the diversity observed among idiotype-crossreactive–antibody-producing cells.

We have undertaken a systematic analysis of the immune response to the hapten 4-(hydroxy-3-nitrophenyl)acetyl (NP) at early stages of the response, using nucleotide sequence analysis of expressed V region genes.

This work was supported by grant 2P30-CA14051 from the National Cancer Institute, awarded to S. E. Luria (Cancer Center Support [Core] Grant), and grants AI 19248 and PO1 CA 28900 from the National Institutes of Health. M. Boersch-Supan is supported by a Department Fellowship from the Department of Biology, MIT. S. Agarwal is a participant in the Undergraduate Research Opportunity Program (UROP) at MIT. M. White-Scharf is the recipient of a Special Fellowship from the Leukemia Society of America.

[1] *Abbreviations used in this paper:* Ars, *p*-azophenylarsonate; C, constant region of Ig; cDNA, complementary DNA; CDR, complementarity-determining region; D, diversity-generating region of Ig; GAC, group A streptococcal carbohydrate; GAT, oligo–glutamine-alanine-tyrosine; H, heavy chain of Ig; J, joining region of Ig; L, light chain of Ig; mAb, monoclonal antibody; mRNA, messenger RNA; NIP, 5-iodo-NP; NP, (4-hydroxy-3-nitrophenyl)acetyl; OX, oxazolone; PC, phosphorylcholine; TBE, Tris-borate-EDTA buffer; V, variable region of Ig.

 J. Exp. Med. © The Rockefeller University Press · 0022-1007/85/6/1272/21 $1.00
Volume 161 June 1985 1272–1292

The immune response of C57BL/6 and BALB/c mice to this hapten has been studied in detail, and the following characteristics have been established: (*a*) the primary response consists of λ L chain-bearing antibodies (19, 20); (*b*) in C57BL/6 mice, these anti-NP antibodies bear the NP^b idiotype, while BALB/c mice express NP^a; (*c*) these idiotypes are inherited as a single genetic unit, in close linkage to the Ig H chain constant region (Ig-C_H) allotype (21, 22); (*d*) serological analysis of hybridoma proteins reveals that NP^b and NP^a antibodies form two families of closely related, λ-bearing antibodies (23–26 and Imanishi-Kari, unpublished results); (*e*) the NP^b family of antibodies is composed of six serologically defined subgroups, in which subgroups I–IV share more determinants than subgroups V and VI, which appear to be quite distinct; (*f*) cloning and sequencing of complementary DNA (cDNA) from the hybridoma B1.8, which synthesizes a strongly NP^b-positive antibody and belongs to subgroup II, revealed that this hybridoma expresses unmutated germline V_H and V_L genes (2, 27); (*g*) the NP^a family is composed of crossreactive but nonidentical antibodies (25) that seem to fall into a single serological group; (*h*) these NP^a-positive monoclonals antibodies (mAb) have determinants that crossreact strongly with antisera against members of subgroup VI, but not subgroups I–V (26).

Differences within the NP^b idiotype, which led to its subdivision into six subgroups, could result either from combinatorial rearrangement or somatic mutation involving a single germline V_H gene, or from the expression of several related germline genes. The finding that the two mAb (B1.8 and B1.48) represent the expression of a single germline V_H gene (186.2), and that the third (S43) arose from the expression of an homologous gene, likely a somatic variant of 186.2, led to the premature conclusion that a single germline V_H gene is responsible for encoding the entire family of anti-NP antibodies (2). It is important to test this conclusion by looking at the germline V_H genes actually employed in a large collection of anti-NP hybridomas. From our sequences and Southern blot analyses, we present strong evidence that multiple germline V_H genes are expressed in the NP^b family of anti-NP antibodies.

Materials and Methods

Mice. Female BALB/c, C57BL/6, and C57BL/6 × BALB/c F1 mice were obtained from The Jackson Laboratory, Bar Harbor, ME.

Hybridomas. All anti-NP hybridomas were obtained from fusions of NP–chicken gammaglobulin–immunized BALB/c or C57BL/6 spleen cells, with the nonsecreting myeloma line X63.Ag8.6.5.3, fused 7 d after immunization (26). For preparation of DNA and RNA, the hybridomas were grown as solid tumors subcutaneously in BALB/c or C57BL/6 × BALB/c F1 mice.

Isolation of Purification of Cellular DNA and RNA. Cellular DNA was prepared from mouse kidney, hybridomas, and myelomas, by methods previously described (29).

Total cellular RNA was isolated from 6 g (wet weight) tumor (29) by a combined guanidinium and hot phenol extraction, yielding 25–40 mg total RNA. Enrichment for poly-A^+ RNA was achieved with two passes through an oligo(dT)-cellulose column (Type II, Collaborative Research Inc., Lexington, MA) (29).

Northern Blot Analysis. RNA was checked to insure that it was intact, and for the presence of antibody-encoding transcripts using a northern blot analysis. 1–2 μg poly-A^+ RNA were electrophoresed through a formaldehyde-containing 1% agarose gel (29). Nick-translated J_H, V_H, or C region probes (γ2b and μ) were used for hybridization (the

γ2b probe was given to us by Dr. S. Gilles, Damon Biotech, Waltham, MA and the μ probe was a gift of Dr. R. Near, Massachusetts General Hospital, Boston, MA).

Southern Blot Analysis. Southern blot analysis was conducted according to the modified procedure of Wahl et al. (30), by fractionating ~15 μg of Eco RI-digested DNA in a 0.7% agarose gel. The DNA was transferred to Zetabind blotting matrix (AMF, Microfiltration Products Division, Meriden, CT) and hybridized for at least 18 h with nick-translated J_H probes (Fig. 3*B*) using 10^6 cpm/ml hybridization mixture. Washes were performed as recommended by AMF for Zetabind Blotting Matrix.

J_H probes (Fig. 3*B*) were obtained with double digests (Bam H1 and Hind III or Eco RI and Xba I) of a plasmid (pBR322) containing the 6.4 kilobase (kb) germline J_H segment (kindly provided by Dr. F. Alt). Digests were separated on an 8% acrylamide gel containing 12% glycerol. The appropriate DNA fragments were detected by ethidium bromide staining (final concentration, 5 μg/ml) and were electroeluted. Ethidium bromide was removed by multiple butanol extractions, followed by two ether extractions. The DNA was precipitated with ethanol, and washed twice in 70% ethanol. Restriction enzymes were purchased from New England Biolabs, Inc., Beverly, MA.

The 17.2.25 V_H probe (a gift of Dr. David Weaver, the Whitehead Institute, Cambridge, MA) was used as a 230 basepair (bp) Pst I fragment inserted in pBR322, or cut out and purified as described above.

Oligonucleotide Primers. Three synthetic oligonucleotide primers, complementary to specific sites in heavy chain messenger RNA (mRNA) were used in the direct sequencing technique. A γ-chain primer, (Cγ-17), d(GGGGCCACGTGGATAGC), hybridizes to the γ_1, γ_{2a}, and γ_{2b} C regions 22–38 bases 3′ to the V region. Our μ-chain primer, (Cμ-17), d(GCAGGAGACGAGGGGGA), hybridizes 40–57 bases from the 5′ end of the C_μ region. The H chain NP idiotype primer, (NPV_H-18), d(GTTCGGAGACGAAGTGGG), hybridizes between the second base of tryptophan codon 36 and the first base of glutamic acid codon 42 of the NP idiotype V_H regions shown in Fig. 1. The Cγ-17 and Cμ-17 primers were gifts of Dr. M. Weigert, Institute for Cancer Research, Philadelphia, PA, and were also purchased from the City of Hope Research Center, Duarte, CA. The NPVH-18 primer was purchased from the City of Hope Research Center.

0.3 μg (50 pmoles) of a primer were kinase-labeled (28), using 0.5 mCi γ[^{32}P]ATP (5,000 Ci/mmol; New England Nuclear, Boston, MA). The kinase reaction mixture was passed over a Sep pak (C18) Cartridge (Waters Associates, Milford, MA) to separate unincorporated γ[^{32}P]ATP from reacted product.

Generation of cDNA by Primer Extension. Primer extension was performed as described (28), using reverse transcriptase and 20–40 μg poly-A^+ RNA, after denaturation for 20 min with CH_3HgOH. Reaction mixtures were electrophoresed on a 5% acrylamide sequencing gel (37 cm × 20 cm × 0.75 mm) at 40 W for ~2 h. Exposure to X-ray film for 30–45 min was used to locate the full-length transcript on the gel. The full-length cDNA band, as well as shorter transcripts, were excised and electroeluted for 30 min to 1 h at 3 W in a electrophoretic sample concentrator (Isco, Inc., Lincoln, NE) in 0.2× TBE (10× TBE is 1 M Tris, 1 M boric acid, and 20 mM EDTA). Sample wells were rinsed several times with 0.2× TBE and Maxam-Gilbert elution buffer (31).

DNA Sequencing. The purified cDNA samples were subjected to base-specific chemical cleavage, as described by Maxam and Gilbert (31).

Results

NP^b- and NP^a-positive Hybridomas. ~60 hybrids of C57BL/6 origin, and 13 hybrids of BALB/c origin were selected on the basis of NP binding and the presence of λL chain, and were established as cell lines. Further serological analysis revealed that most of the λ-bearing antibodies of C57BL/6 origin were positive with polyspecific anti-NP^b idiotypic reagents. These NP^b-positive anti-NP antibodies could be further ordered into six serologically defined subgroups (Table I). Subgroups I–IV share more determinants with each other than with

TABLE I
Subgroups of NP^b-positive, C57BL/6 Anti-NP mAb

I	II	III	IV	V	VI
P5.29.1 (γ1)	B1.8 (μ)	P5.57.1 (γ1)	P7.39.1 (γ1)	P3.6.5 (γ1)	P9.37.1 (μ)
P5.70.11 (γ2b)	P4.21.5 (γ1)	P1.8.50 (γ1)	P6.61.1 (γ1)	P4.6.1 (γ1)	P5.40.1 (γ1)
P6.41.1 (γ1)	P5.84.1 (γ1)	P6.43.1 (γ1)	P8.44.18 (γ2a)	P4.16.25 (γ1)	P8.49.3 (μ)
P6.55.1 (γ1)	P5.83.1 (γ1)	P6.27.1 (γ1)	P10.52.17 (γ1)	P5.18.7 (γ1)	P8.89.4 (γ2b)
P6.31.1 (γ1)	P5.7.4 (γ2b)	P6.49.1 (γ1)	S24.63.12 (γ3)	P5.89.29 (γ1)	P10.15.1 (γ1)
P7.57.2 (γ1)	P5.43.6 (μ)	P7.52.1 (γ1)		P5.41.E (γ2b)	P10.20.8 (γ1)
P8.14.3 (γ2a)	B1.48 (γ1)	P7.53.1 (γ1)		P8.56.7 (γ2a)	
	P8.86.9 (γ1)	P8.56.8 (γ1)			
		P8.90.2 (γ1)			

Idiotypically, there are six subgroups of anti-NP antibodies, I–VI. Most of the hybridomas were obtained from a primary response to NP–chicken gammaglobulin. They represent 12 different fusion experiments. For instance, P5-29.1 means fusion 5 of a primary response, hybrid 29, subclone 1. B1-8 and B1-48 hybridomas are products of a primary response, and S24-63.12 is a product of hyperimmunization with NP coupled to heat-killed group A streptococci (23).

subgroups V and VI. (manuscript in preparation). Hybrid cell lines of BALB/c origin have been established and analyzed (25, 26) to evaluate the genetic basis of the phenotypic differences in the anti-NP response. The λ anti-NP antibody response of BALB/c, defined as the NP^a idiotype, seemed to be homogeneous, and could not be grouped further (25). NP^a-positive hybridomas were shown to share idiotopes with members of subgroup VI (26).

cDNA Sequences of NP^b- and NP^a-positive Hybridomas. We analyzed the H chain mRNA from NP^b-positive hybridomas of subgroups V and VI of C57BL/6 origin, and NP^a-positive hybridomas of BALB/c origin by complete V region sequencing of cDNA using the primer extension method (28).

Sequences of three members of subgroup V (P4.6.1, P4.16.25, and P3.6.5) and subgroup VI (P5.40.1, P10.15.1, and P9.37.1), and two NP^a-positive hybridomas (18.1.16 and 20.1.43) are shown in Fig. 1. In the figure we also include the sequence of 17.2.25, a NP^a-positive hybridoma that we reported previously (28), as well as the sequence of the germline gene V130 of C57BL/6 origin, published by Bothwell (3).

The nucleotide sequence comparison in Fig. 1 indicates ≥90% homology between the expressed V_H genes of different hybridomas. There are 17 positions where exchanges are nonrandom, that is, frequent exchanges are taking place at those positions, and 11 positions undergoing random exchanges, 4 of which are silent and are scattered over the V_H gene, from position −19 to position 98.

The nonrandomness of certain nucleotide exchanges seems to correlate with strain differences or idiotypic grouping. These exchanges are boxed in Fig. 1. For instance, exchanges from BALB/c to C57BL/6 at codons −14, 61, and 66 are GT*T* to GT*C* or *A*TC, G*A*C to G*C*C, and *G*GC to *A*TC or G*T*C. These three positions seem to be strain specific. Position −14 also seems to indicate a group-specific exchange, where *GTT* to *G*T*C* or *A*TC is specific for BALB/c subgroups VI and V, respectively. Constant nucleotide exchange, that is, nonrandom within a group, is also observed. There is a constant nucleotide exchange within subgroup VI at positions 8, 31, and 93. Subgroup V seems to have more

```
BALB/c      -19                                  -10                                                      1                                          10                                          20
20.1.43     ATG AAA TGC AGC TGG GTT ATG TTC TTC CTG ATG GCA GTG GTT ACA GGG GTC AAT TCA GAG GTT CAG CTG CAG CAG TCT GGG GCA GAG CTT GTG AGG CCA GGG GCC TCA GTC AAG TTG
17.2.25     --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---
18.1.16     --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- -A- --- --- --- --- --- --- ---
C57BL/6 group VI
P5.40.1     --- --- T   --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---
P10.15.1    --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- -G- ---
P9.37.1     --- --- T   --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---
V130        --- --- T   --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --C --- --- --- --- --A --- --- --- --- --- --- --- --- --- --- --- --- ---
C57BL/6 group V
P4.6.1      --- --- --- --- --- A   --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- -A- --- --- --- --- --- --- ---
P4.16.25    --- --- --- --- --- A   --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---
P3.6.5      --- --- --- --- --- A   --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---

                                                30    CDR I          FRAMEWORK PRIMER   40                                        50          CDR II                   60
BALB/c
20.1.43     TCC TGC ACA GCT TCT GGC TTC AAC ATT AAA GAC ACC TAT ATG CAC TGG GTG AAG CAG AGG CCT GAA CAG GGC CTG GAG TGG ATT GGA AGG ATT GAT CCT GCG AAT GGT AAT ACT AAA TAT
17.2.25     --- --- --G --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---
18.1.16     --- --- --G --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---
C57BL/6 group VI
P5.40.1     --- --- --- --- --- --- --- --- --- --- A   --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- -G- --- --G ---
P10.15.1    --- --- --- --- --- --- --- --- --- --- A   --- --- --- --- --- --- --- --- --- --- --- --A --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---
P.9.37.1    --- --- --- --- --- --- --- --- --- --- A   --- --- --C --- --- --- --- --- --- --- --- --- --- --- --- --- --- --C --- --- --- --- --- --- --- --- --- --- ---
V130        --- --- --- --- --- --- --- --- --- --- A   --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --A --- --- --- --- --- --- --- --- --- --- --- ---
C57BL/6 group V
P4.6.1      --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- A-- --- --- --- --- --- --- --- --- --- --- --- --- -A- G-- --- GA- --- --- ---
P4.16.25    --- --- --- --- --- --- --- --- --- --- --- TA- --- --- --- --- --- --- --- --- A-- --- --- --- --- --- --- --- --- --- --- --- --- -A- G-- --- GA- --- --- ---
P3.6.5      --- --- --G --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- A-- --- --- --- --- --- --- --- --- --- --- --- --- -A- G-- --- GA- --- --- ---

                                             70                                      80                                      90                                  D  100
BALB/c
20.1.43     GAC CCG AAG TTC CAG GGC AAG GCC ACT ATA ACA GCA GAC ACA TCC TCC AAC ACA GCC TAC CTG CAG CTC AGC AGC CTG ACA TCT GAG GAC ACT GCC GTC TAT TAC TGT GCT AGT TAT AGG
17.2.25     --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- A-- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- TAC TAT
18.1.16     --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --A TAC CTC
C57BL/6 group VI
P5.40.1     --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- A   --- --- --- --- --G TAT TAC
P10.15.1    --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- A   --- --- --- --C CC- TAT TAC
P9.37.1     --- --- --- --- --- --- --- --- --- --- --T --- --- --- --- --- --- --- A-- --- --- --- --- --- --- --- --- --- --- --- --- --- A   --- --- --- --C CTC TAT TAC
V130        --- --- --- --- --- --- --- --- --- --- G-T --- --- --- --- A-- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- A   --- --- --- --- --A
C57BL/6 group V
P4.6.1      --- --- A   --- -T- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --A TAC GAT
P4.16.25    --- --- A   --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --A TAT TAC
P3.6.5      --- --- A   --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --C TAT TAC

                     D                                          J
BALB/c
20.1.43     TAC GAG AGG ... ... ... ... GCC TGG TTT GCT TAC TGG GGC CAA GGG ACT CTG GTC ACT GTC TCT GCA
17.2.25     AGG TAC CCT ... ... ... TAC TAT GCT A-G -AC --- --- --T --- --A --C TCA --- --C --- --C T--
18.1.16     TAT TAC TAC GGT AGT AGC ... ... -AC --- -AC --- --- --C --- --C --C ACT C-- --A --- --C T--
C57BL/6 group VI
P5.40.1     TAC GAG AGT AGC CT. ... ... ... ..G --- --- --- --- --- --- --- --- --- --- --- --- --- ---
P10.15.1    TAC GGG GGG ... ... ... ... ... -AC --C -A- GTC --- --- GC- --- -CC AC- --C --- --T --C T--
P9.37.1     TAC GTT AGT AGC TAC G.. ... ... .-- --- --- --- --- --- --- --- --- --- --- --- --- --- ---
C57BL/6 group V
P4.6.1      TAT TAC TAC GGG T.. ... ... .-- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---
P4.16.1     TAC GGT GC. ... ... ... ..T GAC -AC --- -AC --- --- --- --- --C --C ACT C-- --A --- --C T--
P3.6.5      TAC GGT AGT AGC TTC ... ... ... ... --- --- --- --- --- --- --- --- --- --- --- --- --- ---
```

TABLE II

Comparison of Sequences of the V_H Gene Segment Coding for Anti-NP

Idiotype and subgroup:	NP[b] II	NP[a]			NP[b] VI			Germline	NP[b] V		
	BI-8	20.1.43	18.1.16	17.2.25	P5.40.1	P10.15.1*	P9.37.1*	V130	P4.6.1	P4.16.25	P3.6.5
					Amino acids from position −19 to 98 (117 aa)						
BI-8		37	36	37	37	39	40	39	35	35	37
20.1.43	77 (16)	—	2	2	8	6	8	6	10	9	7
18.1.16	74 (16)	3 (1)	—	2	8	7	9	6	8	9	9
17.2.25	76 (16)	2 (1)	3 (1)	—	8	7	9	6	10	9	9
P5.40.1	77 (17)	11 (2)	13 (4)	13 (3)	—	4	4	2	11	11	10
P10.15.1*	78 (16)	8 (2)	10 (3)	11 (3)	6 (3)	—	4	4	10	10	9
P9.37.1*	80 (16)	11 (3)	13 (4)	13 (4)	7 (4)	7 (3)	—	4	13	12	12
V130	74 (14)	11 (4)	12 (5)	11 (5)	9 (5)	9 (5)	8 (5)	—	11	11	11
P4.6.1	76 (16)	14 (2)	13 (3)	16 (3)	16 (3)	15 (4)	18 (3)	19 (5)	—	4	4
P4.16.25	76 (16)	13 (2)	14 (3)	15 (3)	17 (3)	14 (4)	17 (3)	18 (5)	5 (0)	—	2
P3.6.5	76 (16)	12 (4)	12 (3)	12 (2)	16 (3)	13 (5)	15 (4)	18 (6)	5 (1)	4 (1)	—
					Nucleotides from position −19 to 98 (351 bp)						

Differences between two V_H segments are presented at the intersection of the respective row and column. The number of different amino acids is shown above the diagonal; all nucleotide differences are reported below the diagonal with "silent" substitutions indicated parenthetically.

* P10.15.1 and P9.37.1 were only counted through the second nucleotide of codon 97, based on the assumption that, in the process of joining (V to D–J) a deletion occurred at the 3′ end of the V_H gene segment, and that these residues were filled in by de novo synthesis.

nonrandom exchanges than subgroup VI or NP[a], especially in CDR II (complementarity-determining region II). These positions are −14 (*A*TC), 41 (*A*CT), 54 (G*A*G), 55 (*G*AT), 57 (*G*A*A*), and 63 (AA*A*). These exchanges, as seen in Fig. 1, are in all members of the same subgroup or strain. Positions −17 (G to T), 13 (G to A), 23 (A to G), and 66 (*G*G*C* to *A*T*C*) contain nonrandom exchanges, but are not in all members of the subgroup. Three out of four of these exchanges are transitions from G to A or A to G. Two exchanges result in amino acid substitutions. In position −17 there is a transversion from G to T where two members of subgroup VI (P5.40.1 and P9.37.1) share the codon TTC, while P10.15.1 has TGC like the other V_H mRNA hybridomas of BALB/c and subgroup V origin. Interestingly, the germline V130 of C57BL/6 mice has TTC at this position.

The V_H sequences of these hybridoma mRNA are extremely homologous. Table II indicates the number of nucleotide exchanges within a subgroup or between subgroups at the nucleotide level and at the translated amino acid level. From this comparison, we observe that the number of nucleotide exchanges within a subgroup is significantly lower than the number of nucleotide differences between subgroups. Specifically, while the differences within a subgroup range from two or three (NP[a]) to seven (subgroup VI), differences between subgroups

FIGURE 1. cDNA sequences of H chains of eight anti-NP hybridomas. The codons are numbered sequentially from the leader at position −19. Dashes in sequences indicate identity to the top line. Grey, black, and white boxes indicate C57BL/6-, subgroup VI-, and V-specific positions, respectively. The positions of first and second hypervariable regions (CDR I and CDR II, respectively) and D and J regions are marked with bars. (.) indicates a gap; used to separate the D–J junction. H chain sequences are compared with the sequences of 17.2.25, NP[a]-positive hybridoma (32), and V130 germline V_H gene (3). The hybridomas presented in this figure originate from eight different fusion experiments. P9.37.1 is the only hybridoma of a μ H chain subclass, all the other are γ1.

range from eight to eighteen. In particular, it seems that NP^a and subgroup VI are more similar than either NP^a and subgroup V, or subgroups V and VI. These results are consistent with our serological analysis of these hybridoma proteins using antiidiotypic reagents.

The nonrandom pattern of the nucleotide exchanges that parallels the idiotypic subgrouping of these anti-NP antibodies suggests that each group represents the products of a distinct germline V_H gene, a BALB/c germline gene encoding the NP^a V_H, and two germline genes encoding the NP^b-positive subgroups V and VI, respectively.

The germline gene V130 of C57BL/6 origin (2, 3) has a sequence that is highly homologous to the expressed V_H genes in NP^a and NP^b hybridomas. However, there are four nucleotide differences (codon 2, 7, 48, and 71), all silent, which distinguish V130 from all the hybridoma genes studied (Fig. 1). In fact, the hybridoma genes, while representing different subgroups, still share identical sequences at three of these codons (2, 7, 48).

Contribution from the V_H-D-J_H Joining Region. In attempting to align the different H chain sequences with what appears to be the corresponding germline sequences (see Fig. 2) it is clear that the sites of joining between germline V_H and D and between D and J_H segments are not precisely fixed. These alignments are provisional, because we do not know with certainty which germline segment is involved in each rearrangement, nor the sequences of that V_H segment. Nonetheless, several striking consensus features emerge. For example, codon 98, within a V_H gene, is AGX. This position seems to be constant in almost all NP^b-like germline sequences so far published (2, 32, and this report), 10 have AGX at position 98.

It is striking to see that all of the anti-NP antibody H chain V regions sequenced so far have TAT or TAC at position 99 (this report, 2, 3, and 32), which codes for tyrosine. The germline D gene segment, DFL16.1, or a very similar gene is expressed in all members of subgroups V and VI, as well as in the BALB/c hybridoma, 18.1.16. An interesting feature has been observed in the case of the P4.6.1 and 18.1.16 hybridomas, where two additional codons seem to be inserted 5′ of the DFL16.1 exon. The first of these codons is TAC, conserving position 99 as tyrosine. The D sequences of the 20.1.43 and 17.2.25 hybridomas do not readily agree with the published germline D gene segments. There is partial homology with DFL16.1 and DSP2.3/4 germline D sequences. A similar D sequence was observed (18) in the hybridoma NQU2.6.1, which binds oxazalone. Nucleotide differences in the middle of the D regions compared to the published germline BALB/c D genes might be a consequence of somatic mutation or expression of as yet unidentified germline D genes in the C57BL/6 genome. The D segments are of variable length (5–8 amino acids) and are joined directly, or in one case (P10.15.1) over a short de novo N sequence (8) to a J_H gene. The NP^b and NP^a hybridomas seem to use all four different J_H genes. In group V and VI antibodies, we noticed a preferential expression of J_H3. All of the J_H segments sequenced correspond exactly to the published germline J_H genes (5). The length of the expressed J_H is variable from hybridoma to hybridoma, and is inversely proportional to the size of the D region (Fig. 1).

Analysis of Hybridoma DNA by Southern Blot Hybridization with Specific J_H

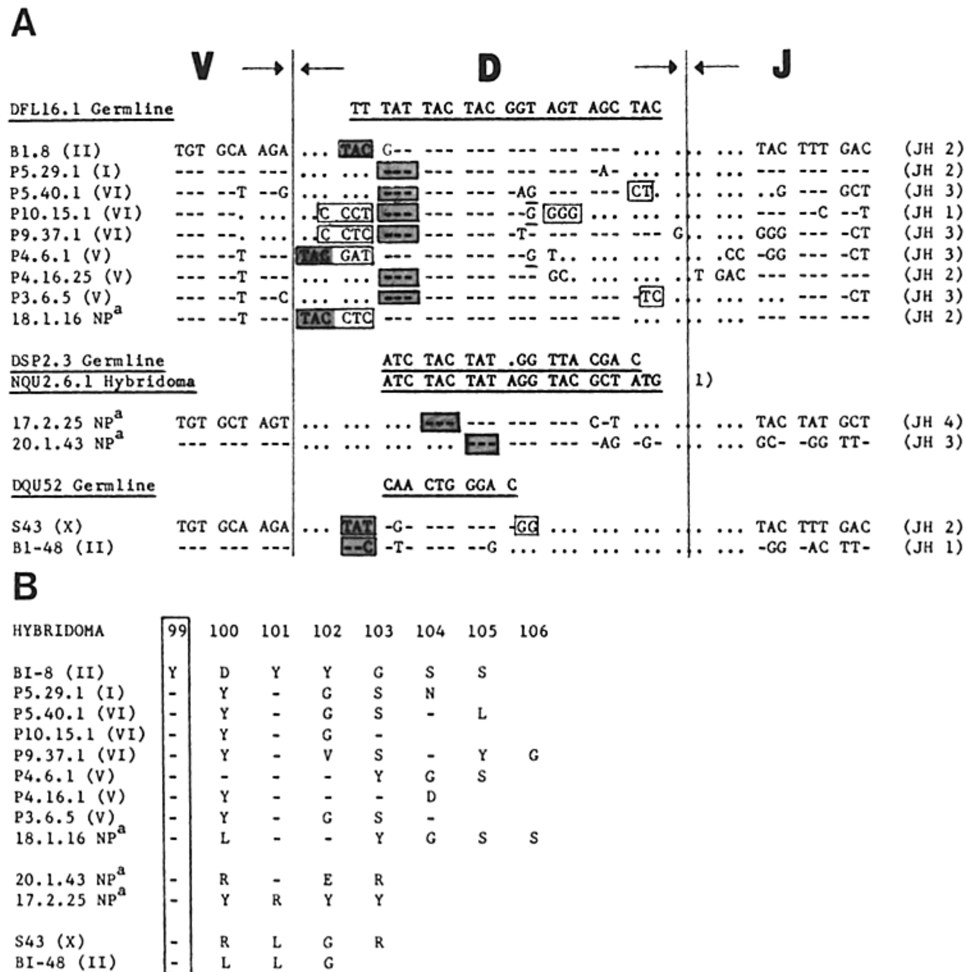


FIGURE 2. (*A*) Nucleotide sequences of germline D segments and D regions of H chain genes of anti-NP hybridomas. To determine the boundaries between V_H and D, as well as between D and J_H in the assembled V gene with complete certainty, it is necessary to compare the nucleotide sequences of three germline regions, namely the 3′ end of the V gene, the D segment, and the 5′ end of the J_H segment (5, 9). B1-8 and P5-29.1 are encoded by the 186.2 germline gene (2), thus the 3′ end of this V_H is known. Several other germline V_H genes have been sequenced (2–4, 11, 17, 32), and the 3′ ends of different V_H genes are highly homologous. We assume the other germline V genes also share sequence homology in this region, and infer codon AGX as the germline V_H 3′ end. The sequences of DFL.16.1, DSP2.3, and DQU52 (9) are used as reference sequences for each D group, as shown: dashes indicate identity to the relevant germline D segment. Sequences that appear to be products of the recombination process, in that they are not accounted for by germline segments (N sequences [8]) are boxed. (.) indicate gaps introduced into the diagram for the purpose of maximally aligning sequences with each other and with the respective germline sequences. Shaded boxes represent codon 99, in all cases TAC or TAT, which codes for tyrosine. (*B*) The predicted amino acid sequences of D regions of anti-NP antibodies. Identity of B1-8 sequence is indicated as a dash.

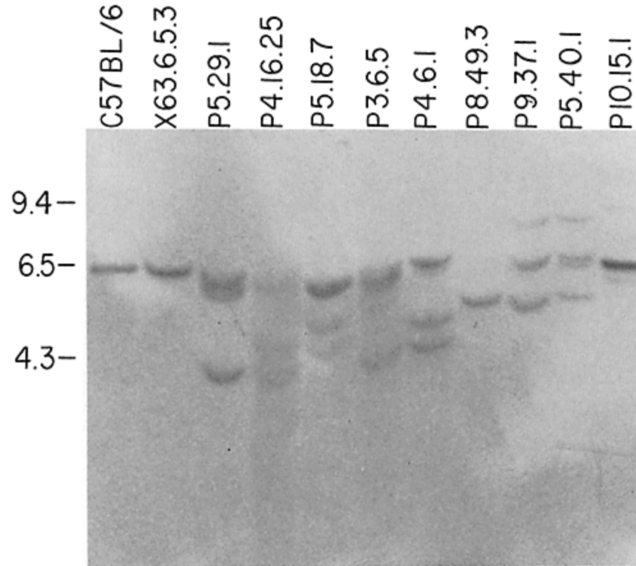
Probes. To test for the presence of the distinct germline V_H genes suggested from our sequence analysis, we performed Southern blot hybridization analysis (30). If the hybridomas use the same V_H germline gene, we should find a rearranged V_HDJ_H gene fragment of uniform size in every hybridoma, using J_H

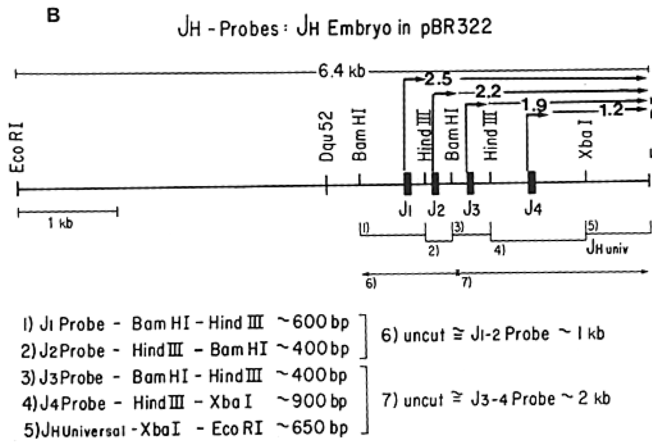
hybridization probes. Minor size differences could be due to the expression of different J_H in different hybridomas. Our sequence analysis had already revealed which J_H gene was expressed in each hybridoma (see below). On the other hand, different V_H genes would likely yield differently sized rearranged V_HDJ_H fragments. We predicted that similarities in rearranged fragment sizes would be limited within subgroups. Mouse kidney and hybridoma DNA were digested with Eco RI, separated in agarose gels, blotted onto filters and hybridized with J_H-specific probes (Fig. 3*B*). The kidney DNA of C57BL/6 and BALB/c mice gave fragments of 6.4 kb containing the four J_H DNA segments, as expected from previous reports (5). The parental cell line X63.Ag8.6.5.3 also gave a fragment of 6.4 kb, however, it hybridized only with J_H3- and J_H4-specific probes, which suggests that X63.Ag8.6.5.3 has a rearranged fragment containing J_H3, again in agreement with previous reports (33). Eco RI digests of total DNA from the different NP^b- and NP^a-positive hybrid cell lines gave one or two fragments in addition to those seen in the fusion parent, and differing in size from the fragment containing the germline J_H gene segments (Fig. 3*A*).

The C57BL/6-derived hybridomas, P5.29.1, produced anti-NP antibodies that are idiotypically classified as belonging to subgroup I. As shown in Fig. 3*A* and Table II, P5.29.1 Eco RI-restricted DNA contains three fragments of 6.4 kb, 6.0 kb, and 4 kb that hybridize with J_H probes. The rearranged fragments of P5.29.1 hybridoma DNA were cloned and sequenced (Maplethorpe and Imanishi-Kari, unpublished results). From this analysis, we found that the functional Eco RI fragment containing V_HDJ_H was 4 kb and had the J_H2 segment expressed (Table III). Since the J_H2 segment lies 2.2 kb upstream of an Eco RI site (Fig. 3*B*), this maps another Eco RI site 1.8 kb upstream of J_H2. The coding region of V_H spans ~0.4 kb, so we can tentatively place an Eco RI site 1.4 kb 5′ to the P5.29.1 V_H gene (Fig. 3*C*).

We have analyzed four hybridomas of subgroup V by Southern blot hybridization (P4.6.1, P4.16.25, P3.6.5, and P5.18.7), Fig. 3*A*. By sequence analysis we found that P4.6.1 and P3.6.5 use J_H3, and P4.16.25 uses J_H2 segments. P3.6.5 hybridoma DNA contained three distinct fragments hybridizing with the J_H3 and J_H4 probes (6.4 kb, 5.8 kb, and 4.3 kb). The 5.8 kb band hybridized with the J_H1 and J_H2 probes, and we therefore assume it is a nonfunctional rearrangement. Similarly, the 6.4 kb band is likely an unexpressed parental fragment. Thus the 4.3 kb band must represent the functionally rearranged H chain gene. Mapping

FIGURE 3. (*A*) Southern blot analysis of 15 μg of kidney DNA from C57BL/6 mice, compared with the pattern obtained using 15 μg of DNA from the parental myeloma cell line X63.Ag8.6.5.3 and hybridomas P5-29.1 of subgroup I, P4.16.25, P5.18.7, P3.6.5, and P4.6.1 of subgroup V, and P8.49.3, P9.37.1, P5.40.1, and P10.15.1 of subgroup VI. The DNA samples were digested to completion with Eco RI, and fragments were separated by electrophoresis in 0.7% agarose. The probe used was the J_H3 and J_H4 segment shown in *B*. Positions of standard DNA size markers are shown to the left of the figure (in kb). (*B*) Partial restriction map of recombinant plasmid clone (pBR322-6.4 J_H) containing the Eco RI fragment of germline J_H region (5). The different probes used are outlined. Distances (in kb) of each J_H gene to the 3′ end Eco RI site are indicated between arrows. (*C*) Mapping of predicted Eco RI site at the 5′ end and the actual 3′ end of V_H was obtained from the Southern blot analysis of different hybridoma DNA digested with Eco RI enzyme shown in *A* (5). These sites are indicated (in kb) between arrows. Partial restriction map of 186.2 and VH130 germline V_H genes were obtained from references 2 and 3.


Southern with C57BL/6 Hybridomas
C57BL/6
X63.6.5.3
P5.29.1
P4.16.25
P5.18.7
P3.6.5
P4.6.1
P8.49.3
P9.37.1
P5.40.1
P10.15.1
9.4–
6.5–
4.3–
A


B
JH - Probes: JH Embryo in pBR322
6.4 kb
2.5
2.2
1.9
1.2
Eco RI
Dqu52
Bam HI
Hind III
Bam HI
Hind III
Xba I
Eco RI
J1 J2 J3 J4
1 kb
JH univ
1) J1 Probe - Bam HI - Hind III ~600 bp
2) J2 Probe - Hind III - Bam HI ~400 bp
3) J3 Probe - Bam HI - Hind III ~400 bp
4) J4 Probe - Hind III - Xba I ~900 bp
5) JH Universal - Xba I - Eco RI ~650 bp
6) uncut ≅ J1-2 Probe ~ 1 kb
7) uncut ≅ J3-4 Probe ~ 2 kb

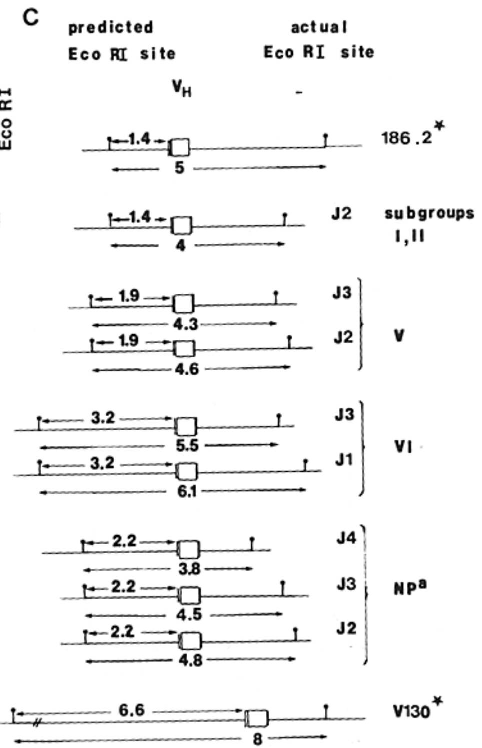

C
predicted Eco RI site
actual Eco RI site
VH
186.2*
J2 subgroups I,II
J3
J2
V
J3
J1
VI
J4
J3
J2
NPa
V130*

TABLE III

Eco RI Fragments Hybridizing to J_H Probes

Source of DNA	Number of bands	Size (in kb) of expressed J_H segment
C57BL/6 Kidney	1	6.4 germline J_H, 1–4
BALB/c kidney	1	6.4 germline J_H, 1–4
X63.Ag8.6.5.3 (myeloma)	1	6.4 (J_H3)*
Group I hybridoma		
P5.29.1	3	6.4 (J_H3)*, 6 (J_H, 1 or 2), 4.0 (J_H2)
Group V hybridomas		
P4.6.1	3	6.4 (J_H3)*, 4.8 (J_H3), 4.3 (J_H3)
P4.16.25	3	6.4 (J_H3)*, 4.6 (J_H2), 4.0 (J_H3)
P3.6.5	3	6.4 (J_H3)*, 5.8 (J_H1), 4.3 (J_H3)
P5.18.7	3	6.4 (J_H3)*, 4.8 (J_H2), 4.5 (J_H2)
Group VI hybridomas		
P5.40.1	3	9.3 (J_H4), 6.4 (J_H3)*, 5.5 (J_H3)
P10.15.1	3	9.4 (J_H2), 6.1 (J_H1)
P9.37.1	3	9.3 (J_H4), 6.4 (J_H3)*, 5.5 (J_H3)
P8.49.1	1	5.5 (J_H3)
BALB/c hybridomas		
18.1.16	3	6.7 (J_H3), 6.4 (J_H3)*, 4.8 (J_H2)
20.1.43	3	6.4 (J_H3)*, 5.8 (J_H2), 4.5 (J_H3)
17.2.25	3	6.4 (J_H3)*, 6.2 (J_H1), 3.8 (J_H4)
20.1.21	2	6.4 (J_H3)*, 5.8 (J_H2)

The sizes of DNA fragments that hybridize with different J_H probes are shown in this table. The J_H probes used are shown in Fig. 3*B*. The J_H shown in parentheses is the J_H present in the fragment of that particular size. Underlined is the size of the Eco RI fragment containing functional, rearranged V_H–D–J_H genes.

* Fragment of X63.Ag8.5.3 origin.

the upstream restriction site as we did for P5.29.1, we identify a new V_H gene fragment with an Eco RI site 1.9 kb upstream of the V_H sequence (Fig. 3*C*). Using this same analysis, we find that P4.16.25 has a functionally rearranged Eco RI fragment of 4.6 kb, containing J_H2. This maps the Eco RI site 1.9 kb upstream of V_H, just as with the P3.6.5 hybridoma. We can tentatively conclude that the P3.6.5 and P4.16.25 V_H genes are identical (Fig. 3*C*). P4.6.1 has two rearranged fragments, containing J_H3 segments of 4.8 kb and 4.3 kb, while P5.18.7 has two rearranged fragments of 4.8 kb and 4.6 kb, containing J_H2 (Fig. 3*A*). This ambiguity precludes identification of the functional fragment by this analysis. However, it is striking that both hybridomas contain bands identical to functional fragments from other J_H-related members of the same subgroup. Thus our Southern data analysis is compatible with the assumption that all of the hybridomas belonging to subgroup V may utilize a unique V_H gene (Fig. 3*C*).

The closely related hybridomas of subgroup VI, P5.40.1, P9.37.1, and P8.49.3, share a related hybridization pattern, all containing a single 5.5 kb rearranged fragment hybridizing with a J_H3 probe. P10.15.1, which is idiotypically strongly crossreactive, and very similar to the former antibodies in sequence, shows a higher molecular weight fragment of 6.1 kb, which hybridizes to the J_H1

probe. Our sequence analysis revealed that P5.40.1 and P9.37.1 express J_H3, and P10.15.1 expresses J_H1. Since the same germline V_H gene, when rearranged to J_H1, would generate a fragment ~0.6 kb larger than when rearranged with J_H3 (Fig. 3*B*), these results allow us again to identify a new anti-NP V_H gene, with an Eco RI site 3.2 kb upstream of the V_H sequence used uniquely by the members of subgroup VI.

Four BALB/c hybridomas: 18.1.16, 20.1.43, 17.2.25, and 20.1.21 were subjected to the same analysis. The Southern blots are not shown, but the results are summarized in Table III and Fig. 3*C*.

Discussion

Recent molecular analysis of antigen-specific immune responses in well-defined idiotype systems has suggested that a process of hypermutation superimposed on combinatorial joining are the mechanisms by which antibody molecules are diversified. In the A/J response to *p*-azophenylarsonate (Ars), it appears (14, 34) that the H chains of antibodies bearing the crossreactive idiotype are derived from a single germline V_H segment with superimposed somatic mutation. Similarly, in the BALB/c antibody response to phosphorylcholine (PC), the antibodies differ from each other and yet all are encoded by a single germline V_H gene segment (35). Similar results have been suggested for antibodies against oligo–glutamine-alanine-tyrosine (GAT) (36) and oxazalone (OX) (18, 37).

To understand further the relative contributions of the germline repertoire, combinatorial events, and somatic mutation in the generation of the functional repertoire of antibody molecules, we analyzed the heterogeneous response of mice to the hapten NP at the molecular level.

Molecular analysis of the V_H region genes expressed by three different hybridomas (B1-8 and B1-48, from a primary response, and S43 [2, 3]) revealed that expression of NP^b determinants reflected the expression of unmutated germline sequences; differences between members of the same subgroup represented combinatorial variations; and the lack of expression of NP^b determinants probably represented somatic point mutations of V_H and V_L gene segments, as seen in hybridoma S43, which originated from a hyperimmune response. Analysis of 15 germline V_H genes revealed that B1-8 and B1-48 express the ummutated germline V_H 186.2 gene (2, 3). The sequence of the expressed V_H gene segment, S43, is not identical to any of the germline V_H sequences to date; however, it resembles most the V_H 186.2 germline gene. Thus it was concluded that S43 is the product of somatic mutation of this gene (2). Furthermore, it was prematurely concluded that the NP^b family of anti-NP antibodies, as in the immune response to PC, OX, Ars, or GAT, was a product of a single germline V_H gene, with superimposed somatic mutations and recombinatorial differences (2). Thus the serological subgrouping would represent members having similar mutations or similar residues due to recombinatorial events. However, the molecular analyses presented in this report indicate that serological subgroups represent expression of distinct V_H germline genes. Thus the diversity of functional anti-NP antibodies is greatly contributed by diversity at the germline V_H level.

Multiple V_H Gene Segments Encode λ-bearing Anti-NP Antibodies. Comparison of nucleotide and amino acid sequences obtained from hybridomas of subgroup

V and VI (Figs. 4 and 5, and Table II) to those of B1-8 indicate that different germline V_H genes are expressed in these hybridomas. They differ from B1-8 by 40 amino acids. These results explain why subgroups V and VI, though NP^b-positive, do not crossreact with the anti-B1.8 idiotypic reagents. From the sequence data presented in Fig. 1, and from the Southern blot analysis, we suggest that the subgroups V and VI are encoded by different germline V_H gene segments, which have constant nucleotide exchanges that are shared within the subgroup, some of which are silent. In contrast to the results obtained in PC, Ars, GAT, and OX systems, we have shown that some of the H chain diversity of the NP^b-positive antibodies results from the use of more than one V_H gene segment. Another explanation would be that each H chain is the product of somatic mutation of a single germline V_H gene, and that strong antigenic selection has occurred. However, the hybridomas analyzed were products of different fusion experiments and of different B cell clonal origins. Further distinguishing groups V and VI are two silent mutations at positions 57 and 63. It seems unlikely that consistent mutation patterns would occur independently in the V_H segment as well as in flanking nontranslated (unselected) regions. Detailed analysis of members of other subgroups (I–IV), will enable us to decide whether the NP^b response is a product of expression of a family of highly homologous germline V_H genes. Similarly, the existence of multiple, highly homologous germline V_H gene segments encoding A/J anti-GAC (group A streptococcal carbohydrate) antibodies was suggested (38).

Expansion of the V_H NP Gene Family. Studies (2, 3) have shown that the NP V_H family of germline genes is large. Sequence analysis of seven germline V_H genes most closely related to 186.2 indicated that they are extremely homologous. Most of the differences are in CDR II, which could be explained by minigene insertion (39) during evolution. The question that arises is what role all these extremely homologous ($\geq$90%) genes play in the antibody repertoire.

Bothwell et al. (3) have isolated a weakly crosshybridizing V_H gene (V130) using a B1.8 V_H cDNA probe. The V130 gene (shown in Figs. 1 and 5, and Table II), is most homologous in sequence to members of subgroup VI. It differs at most by nine nucleotides, with four amino acid substitutions. At four of these positions (2, 7, 48, and 71), the exchanges are silent. All members of subgroups V and VI, as well as NP^a, have identical sequences at these positions. These results suggest that, in spite of a high degree of homology, V130 is not the V_H gene segment coding for subgroup VI H chains. The comparison of our Southern analysis with the V130 restriction map (3) further corroborates this suggestion (Fig. 3*C*).

We report here that very homologous, but different germline V_H genes are expressed and used to bind the same antigenic determinant. Moreover, related V_H genes can be used to bind different antigens. The NP^a hybridoma, 18.1.16, differs from the G5.BB2.2 BALB/c anti-GAT V_H segment at only three positions (14, 23, and 98) (36). This level of similarity could suggest that the same V_H segment that encodes anti-NP H chains can also encode H chains for anti-GAT. However, all of the anti-GAT antibodies have the same silent exchange at position 14, strongly suggesting the presence, in the germline, of yet another V_H gene segment coding for anti-GAT. The silent G $\rightarrow$ A exchange at position 23 is also

```
          -19                                 -10                                    -1  1                                 10                                    20
186-2 (B1-8) ATG GGA TGG AGC TGT ATC ATG CTC TTC TTG GCA GCA ACA GCT ACA GGT GTC CAC TCC CAG GTC CAA [CTG CAG] CAG [CCT] GGG GCT GAG CTT GTG AAG CCT GGG GCT TCA GTG AAG CTG
20.1.43      --- AA- --C --- --G G-T --C T-- --- C-- ATG --- GTG -T- --- --G --- A-T --A G-- --T --G [--- ---] --- [T--] --- --A --- --- --- -G- --A --- --C --- --C --- T--
P10.15.1     --- AA- --C --- --G G-- --C T-- --- C-- ATG --- GTG -T- --- --G --- A-T --A G-- --T --G [--- ---] --- [T--] -T- --A --- --- --- -G- --A --- --C --- --C -G- T--
P3.6.5       --- AA- --C --- --G --- --C T-- --- C-- ATG --- GTG -T- --- --G --- A-T --A G-- --T --G [--- ---] --- [T--] --- --A --- --- --- -G- --A --- --C --- --C --- T--

          21                                30  CDR I                                   40  41                                      50          CDR II                      60
186-2     TCC TGC AAG GCT TCT GGC TAC ACC TTC ACC AGC TAC TGG ATG CAC TGG GTG AAG CAG AGG CCT GGA CGA GGC CTT GAG TGG ATT GGA AGG ATT GAT CCT AAT AGT GGT GGT ACT AAG TAC
20.1.43   --- --- -CA --- --- --- -T- -A- ATT -AA GA- AC- -AT --- --- --- --- --- --- --- --- -A- -AG --- --G --- --- --- --- --- --- --- --- GCG -A- --- AA- --- --A --T
P10.15.1  --- --- -CA --- --- --- -T- -A- ATT -AA AA- AC- -AT --- --- --- --- --- --- --- --- -A- -A- --- --G --- --- --- --- --- --- --- --- CCC -A- --- AA- --- --A --T
P3.6.5    --- --- -C- --- --- --- -T- -A- ATT -AA GA- AA- -AT --- --- --- --- --- --- --- A-- -A- -AG --- --G --- --- --- --- --- --- --- --- G-G GA- --- -A- --- --A --T

          61                                70                                      80  81                                      90                              98
186-2     AAT GAG AAG TTC AAG AGC AAG GCC ACA CTG ACT GTA GAC AAA CCC TCC AGC ACA GCC TAC [ATG CAG] CTC AGC AGC CTG ACA TCT GAG GAC TCT GCG GTC TAT TAT TGT GCA AGA
20.1.43   G-C CC- --- --- C-- G-- --- --- --T A-A --A -C- --- -C- T-- --- -A- --- --- --- [C-- ---] --- --- --- --- --- --- --- --- A-- --C --- --- --C --- --T --T
P10.15.1  GCC CC- --- --- C-- GT- --- --- --T A-A --A -C- --- -C- T-- --- -A- --- --- --- [C-- ---] --- --- --- --- --- --- --- --- A-- --C A-- --- --C --- --C CCT
P3.6.5    GCC CC- --A --- C-- GT- --- --- --T A-A --A -C- --- -C- T-- --- -A- --- --- --- [C-- ---] --- --- --- --- --- --- --- --- A-- --C --- --- --C --- --T ---
```

FIGURE 4. Comparison of the V_H cDNA sequences of a member of each group. (20.1.43 [NP^a], P10.15.1 [VI], and P3.6.5 [V]) are compared with the sequence of the germline gene 186.2 coding for B1.8 hybridoma (2). The open boxes represents Pst I restriction sites. The Pst I site at positions 81 and 82 is characteristic of NP^a-like sequences. The shaded box represents position 7, which, in 186.2-like sequences, is a CCT codon for proline, while in NP^a-like sequences, it is TCT, codon for serine. CDR I and CDR II are marked with bars.

```
             -1    -1         1         2         3        4       5
              9     0    1    0         0         0        0       0
                                                        CDRI
B1-8       -GW-CI-L--A-TA--H-Q----P----K-------K--YTFTSYW-[ ]-----GR-----[ ]
20.1.43    MKCSWVMFFLMAVVTGVNSEVQLQQSGAELVRPGASVKLSCTASGFNIKDTYM[H]WVKQRPEQGLEWIG[R]
18.1.16    ----------------------------K------------------------[-]-------------[-]
17.2.25    -----------------------------------------------------[-]-------------[-]

P5.40.1    --F--------------------V---------------------N-------[-]-------------[-]
P10.15.1   -----------------------V----------R----------N-------[-]-------------[-]
P9.37.1    --F--------------------V---------------------N--I----[-]-------------[-]
V130       --F--------------------V---------------------N-------[-]-------------[-]

P4.6.1     -----I----------------------K------------------------[-]----T--------[-]
P4.16.25   -----I-----------------------------------------Y-----[-]----T--------[-]
P3.6.5     -----I-----------------------------------------------[-]----T--------[-]
                                                        CDRI

                                                  1
                6      7      8      9            0
                0      0      0      0            0
           CDRII                                     D          J
B1-8       [-]-NS-G---NE--KS---L-N-KP-N---M--------S------R[D]-YGSS...Y-D------TL---S
20.1.43    [D]PANGNTKYDPKFQGKATITADTSSNTAYLQLSSLTSEDTAVYYCAS[Y]RYER....AWFAYWGQGTLVTVSA
18.1.16    [-]---------------------------------------------R[L]-YYGSS..Y-D------TL---S
17.2.25    [-]--------------------T------------------------R[Y]RYY...YYAMD------S-----

P5.40.1    [-]-----S---A---I--------------------------I----R[Y]-GSSL....------------
P.10.15.1  [-]---------A---V--------------------------I----R[Y]-GG.....Y-DV--A--T----S
P.9.37.1   [-]---------A---V--------------T-----------I----L[Y]-VSSYG...------------
V130       [-]---------A------------------------------I----R[ ]

P4.6.1     [-]-ED-E---A--LI-------------------------------R[D]-YYGS...------------
P.4.16.25  [-]-ED-E---A---V--------------------------------R[Y]-YGD....Y-D------TL---S
P3.6.5     [-]-ED-E---A---V---------------------------------[Y]-GSS.....------------
           CDRII
```

FIGURE 5. Comparisons of the predicted amino acid sequences of 10 V_H regions of NP-binding hybridomas, and of germline V_H V130. B1-8 and S43 sequences are from reference 2, 17.2.25 is from 32, and V_H V130 is from 3. Those residues common to 20.1.43 are indicated as dashes. The top line displays the amino acid sequence of B1-8. Positions 35, 50, 52, and 99 are boxed. These residues were suggested to be involved in antigen binding (42).

observed in 20.1.43 (NP[a]), all members of subgroup VI, and in P3.6.5 (subgroup V). Differences at position 98 reflect the J region mutations discussed in detail later.

The apparent polymorphic exchanges at positions −14, 61, and 66 (Fig. 1) are the same for subgroups V and VI, and different in NP[a]. Given that subgroup V antibodies are encoded by a different germline V_H gene from antibodies of subgroup VI, it would be fair to assume that gene duplication has occurred since the time of strain divergence in mice. The study of the V_H NP gene family may provide us with insight into the genetic mechanisms involved in the process of gene expansion and contraction.

Limited Heterogeneity of Antibodies of Subgroups V, VI, and NP[a]. The nucleotide sequences of V_H cDNA derived from hybridomas that are members of subgroup V, VI, and NP[a] are quite similar: the two most divergent sequences, P9.37.1 and P4.6.1, differ by 18 exchanges within the V_H segment (Fig. 1 and Table II). In agreement with idiotype data, hybridomas of subgroup VI are more closely related in V region gene sequence to NP[a] than to subgroup V. It is interesting to note that NP[a] V_H gene sequences, in spite of the fact that they are extremely homologous to those of the NP[b] subgroups V and VI, show nucleotide exchanges that are probably polymorphic, such as at positions −14, 61, and 66 (Fig. 1).

This result may indicate that the gene(s) encoding the H chains of subgroup VI are the C57BL/6 allelic counterpart to the gene coding for NP^a antibodies.

The nucleotide sequences of the three hybridomas belonging to NP^a (17.2.25 [32], 18.1.16, and 20.1.43) differ by at most three nucleotides in the entire V_H gene segment. They are probably encoded by a single germline gene with few somatic point mutations. So far, NP^a seems to resemble antibody responses to PC, Ars, OX, and GAT, where somatic mutations in a single germline V_H amplify the diversity of expressed H chains. (11, 18, 34, 36). However, a limited number of nucleotide exchanges may not always indicate somatic point mutations of a particular V_H germline gene. Distinct germline V_H were found where there was a single nucleotide exchange in the coding region (2, and D. Loh, Washington University, St. Louis, MO, personal communication).

Contribution of Genetic Diversity, Somatic Point Mutation, and Combinatorial Joining to the Functional Antibody Repertoire. The studies presented in this report suggest that genetic diversity contributes more to the functional antibody repertoire than believed so far. As suggested previously (40, 41), we observe that antibodies with similar sequences have similar fine specificity. The anti-NP antibodies analyzed all use λ1 L chains, and are encoded both by extremely homologous, and clearly different V_H sequences. The NP^a-like antibodies have very similar fine specificity; that is, they are heteroclitic (25). The degree of heteroclicity, however, is less than that observed in B1-8 or P5-29.1 antibodies. The difference in fine specificity between NP^a-like and B1.8 or P5.29.1 antibodies is probably a result of the use of different V_H genes.

Comparison of each group of V_H genes studied indicates that, in NP^a V_H gene segments, they differ by at most three nucleotide exchanges, in subgroup V, by five, and subgroup VI, by seven. The exchanges are scattered throughout the V_H segment, and no somatic point mutations were observed in any of the J_H segments sequenced. Nucleotide exchanges observed at positions 13, 23, 71, and 93 are nonrandom (Fig. 1). These nonrandom exchanges can be explained by the existence of yet other germline V_H genes, or by assuming that these sites are hot spots of mutation. Since we do not have a germline sequence, we cannot compare their affinities. Consequently, we are not certain whether the possible somatic point mutations did affect the specificity.

In spite of differences in V_H sequence and fine specificity between NP^a-like antibodies and antibody B1-8, they all have positions 35, 52, and 99 in common, all of which are found in CDR regions. These positions were suggested to be important for NP/NIP (5-iodo-NP) binding (42). The amino acid residues at positions 35 and 52 are also the same in anti-GAT antibodies. In all the anti-NP antibodies analyzed so far, however, position 99 is tyrosine, while it is glycine in the anti-GAT antibodies (Table IV) (35).

The codon TAC or TAT for tyrosine at position 99 is a product of the recombination event between V_H and D. It is, in some instances, a direct joining of V_H to the germline DFL16.1 (9) or NQU2.6.1 (18). In other cases, the codon TAC or TAT is inserted, conserving position 99 as tyrosine (Fig. 2, *A* and *B*) (2, 3).

In idiotype systems such as PC (11, 16), OX (18), GAT (36), Ars (43, 44), and α-1.3-dextran (15), position 99 is also invariant, with few exceptions (Table IV).

TABLE IV
Constancy of Position 99 in Different Idiotype Systems

Antigen	Number of V_H sequences analyzed	Number of V_H sequences with invariant position 99	Amino acid at position 99	V_H gene family*	Exceptions
NP	12	12	Tyr	J558	—
Ars	9	9	Ser	J558	—
GAT	4	4	Gly	J558	—
α-1.3 dextran	13	13	Asp	J558	—
OX	15	12	Asp	Q52	3 (Id^-)
PC	18	16	Asp	S107	2 (Id^-)

Data for NP are reported herein, and in 2 and 3; OX, 18; ARS, 43 and 44; PC, 11 and 16; α-1.3 dextran in 15.
* V_H gene family is designated according to 48.

This position may be invariant because of the specific interaction with L chains. Anti-NP and anti–α-1.3-dextran, however, express an unmutated λ1 L chain gene, but they differ at position 99. The invariance of this position in the H chain J region may be due to the fact that all antibodies analyzed were selected to bind the antigen (45). A single amino acid residue at this position, in combination with a specific L chain seems to be crucial for antigen binding. We isolated a BALB/c anti-NP antibody, C8-5, which uses λ1 L chain, has a different amino acid at position 99 (tryptophan instead of tyrosine), and only weakly binds NIP (Boersch-Supan, White-Scharf, and Imanishi-Kari, unpublished results). Another possibility is that position 99 is a target for T or B cell recognition (idiotypic-specific T or B cells?). In this context, it is interesting to note that three anti-OX and two anti-PC antibodies, shown in Table IV, contained variant amino acids at position 99 and were idiotype-negative or of uncertain idiotype (16, 37). Recently (46, 47), two other instances have been reported where idiotypic selection seems to be used by regulatory cells. It will be important to distinguish between antigenic and idiotypic selection in order to understand the regulatory mechanisms involved in the creation of antibody repertoire. A combination of idiotype- and antigen-selection would explain idiotype dominance in certain immune responses.

Summary

The hapten (4-hydroxy-3-nitrophenyl)acetyl (NP), when conjugated to carrier proteins, elicits a characteristic idiotypic response (NP^b) in C57BL/6 mice. The response can be divided serologically into two distinct NP^b-positive groups of antibodies. The first group consists of four crossreacting subgroups (I–IV), the second of two subgroups (V, VI). Some antibodies of subgroups I and II have been shown to express the unmutated heavy chain variable region (V_H) germline gene 186.2. Antibodies of subgroups V and VI crossreact extensively with the NP^a-positive antibodies of BALB/c mice.

We sequenced heavy chain complementary DNA from eight hybridomas

producing anti-NP antibodies. Six of these belong to subgroups V and VI, and two were NP^a-positive hybridomas of BALB/c origin.

All sequences were homologous to each other, and differed by ~80 basepairs from the 186.2 C57BL/6 germline V_H gene. From our sequence and Southern blot analyses we suggest: (*a*) the NP^b idiotypic response is the product of several V_H germline genes, (*b*) some of these genes are very homologous to the gene coding for the BALB/c NP^a idiotype, and might represent the C57BL/6 allelic forms of this gene, (*c*) the diversity regions of NP^b and NP^a-positive antibodies are diverse in length and amino acid composition, except for the first residue, which is always tyrosine, (*d*) all four heavy chain joining region gene segments are expressed without mutation.

We discuss our data in terms of diversity in the germline V_H gene repertoire, as well as diversity created by gene segment-joining events and somatic mutation.

We thank Christopher Albanese for technical assistance, Drs. Edward Reilly and Koichi Tamoto for assisting us with the primer extension and DNA sequencing, and Charles Maplethorpe for permission to cite unpublished sequence and Southern data on P5.29.1. We are also thankful to Drs. Tim Manser, David Ucker, and Henry Wortis for critically reviewing the manuscript.

Received for publication 19 November 1984 and in revised form 12 February 1985.

References

1. Tonegawa, S. 1983. Somatic generation of antibody diversity. *Nature.* (*Lond.*). 302:575.
2. Bothwell, A. L. M., M. Paskind, M. Reth, T. Imanishi-Kari, K. Rajewsky, and D. Baltimore. 1981. Heavy chain variable region contribution to the NP^b family of antibodies; somatic mutation evident in a γ2a variable region. *Cell.* 24:625.
3. Bothwell, A. L. M. 1984. The genes encoding anti-NP antibodies in inbred strains of mice. *In* The Biology of Idiotypes. M. I. Greene and A. Nisonoff, editors. Plenum, New York. 19–34.
4. Cohen, J. B., and D. Givol. 1983. Allelic immunoglobulin V_H genes in two mouse strains: possible germline gene recombination. EMBO 2:2013.
5. Sakano, H., R. Maki, Y. Kurosawa, W. Roeder, and S. Tonegawa. 1980. Two types of somatic recombination are necessary for the generation of complete immunoglobulin heavy-chain genes. *Nature* (*Lond.*). 286:676.
6. Early, P., H. Huang, M. Davis, K. Calame, and L. Hood. 1980. An immunoglobulin heavy chain variable region is generated from three segments of DNA: V_H, D and J_H. *Cell* 19:981.
7. Weigert, M., R. Perry, and D. Kelley. 1980. The joining of V and J gene segments creates antibody diversity. *Nature* (*Lond.*). 283:497
8. Alt, F. W., and D. Baltimore. 1982. Joining of immunoglobulin heavy chain gene segments: Implications from a chromosome with evidence of three D-J_H fusions. *Proc. Natl. Acad. Sci. USA.* 79:4118.
9. Kurosawa, Y., and S. Tonegawa. 1982. Organization, structure, and assembly of immunoglobulin heavy chain diversity of DNA segments. *J. Exp. Med.* 155:201.
10. McKean, D., K. Huppi, M. Bell, L. Staudt, W. Gerhard, and M. Weigert. 1984. Generation of antibody diversity in the immune response of BALB/c mice to influenza virus hemagglutinin. *Proc. Natl. Acad. Sci. USA.* 81:3180.

11. Crews, S., J. Griffin, H. Huang, K. Calame, and L. Hood. 1981. A single V_H gene segment encodes the immune response to phosphorylcholine: somatic mutation is correlated with the class of the antibody. *Cell.* 25:59.
12. Pech, M., J. Hochtl, H. Schnell, and H. G. Zachau. 1981. Differences between germline and rearranged immunoglobulin V coding sequences suggest a localized mutation mechanism. *Nature (Lond.).* 291:668.
13. Kim, S., M. Davis, E. Sinn, P. Patten, and L. Hood. 1981. Antibody Diversity: Somatic Hypermutation of Rearranged V_H Genes. *Cell.* 27:573.
14. Sims, J., T. H. Rabbitts, P. Estess, C. Slaughter, P. W. Tucker, and J. D. Capra. 1982. Somatic mutation in genes for the variable portion of the immunoglobulin heavy chain. *Science (Wash. DC).* 216:309.
15. Schilling, J., B. Clevinger, J. M. Davie, and L. Hood. 1980. Amino acid sequence of homogeneous antibodies to dextran and DNA rearrangments in heavy chain V-region gene segments. *Nature (Lond.).* 283:35.
16. Gearhart, P. J., N. D. Johnson, R. Douglas, and L. Hood. 1981. IgG antibodies to phosphorylcholine exhibit more diversity than their IgM counterparts. *Nature (Lond.).* 291:29.
17. Near, R. I., E. C. Juszczak, S. Y. Huang, S. A. Sicari, M. N. Margolies, and M. L. Gefter. 1984. Expression and rearrangement of homologous immunoglobulin V_H genes in two mouse strains. *Proc. Natl. Acad. Sci. USA.* 81:2167.
18. Kaartinen, M., G. M. Griffiths, A. F. Markham, and C. Milstein. mRNA sequences define an unusually restricted IgG response to 2-phenyloxazolone and its early diversification. *Nature (Lond.).* 304:320.
19. Imanishi, T., and Mäkelä, O. 1974. Inheritance of antibody specificity. I. Anti(4-hydroxy-3-nitrophenyl)acetyl of the mouse primary response. *J. Exp. Med.* 140:1498.
20. Imanishi-Kari, T., E. Rajnavolgyi, T. Takemori, R. S. Jack, and K. Rajewsky. 1979. The effect of light chain gene expression on the inheritance of an idiotype associated with primary anti-(4-hydroxy-3-nitrophenyl)acetyl (NP)-specific antibodies. *Eur. J. Immunol.* 9:324.
21. Jack, R. S., T. Imanishi-Kari, and K. Rajewsky. 1977. Idiotypic analysis of the response of C57BL/6 mice to the (4-hydroxy-3-nitrophenyl)acetyl group. *Eur. J. Immunol.* 7:559.
22. Karjalainen, K. 1980. Two major idiotypes in mouse anti-(4-hydroxy-3-nitrophenyl) acetyl (NP) antibodies are controlled by "allelic" genes. *Eur. J. Immunol.* 10:132.
23. Reth, M., G. J. Hammerling, and K. Rajewsky. 1978. Analysis of the repertoire of anti-NP antibodies in C57BL/6 mice by cell fusion. *Eur. J. Immunol.* 8:393
24. Reth, M., T. Imanishi-Kari, and K. Rajewsky. 1979. Analysis of the repertoire of anti-(4-hydroxy-3-nitro-phenyl)acetyl (NP) antibodies in C57BL/6 mice by cell fusion. *Eur. J. Immunol.* 9:1004.
25. White-Scharf, M. E., and T. Imanishi-Kari. 1981. Characterization of the NP^a idiotype through the analysis of monoclonal BALB/c (4-hydroxy-3-nitrophenyl) acetyl (NP) antibodies. *Eur. J. Immunol.* 11:897.
26. White-Scharf, M. E., and T. Imanishi-Kari. 1982. Cross-reactivity of the NP^a and NP^b idiotypic responses of BALB/c and C57BL/6 mice to (4-hydroxy-3-nitrophenyl)acetyl (NP). *Eur. J. Immunol.* 12:935.
27. Bothwell, A. L. M., M. Paskind, M. Reth, T. Imanishi-Kari, K. Rajewsky, and D. Baltimore. 1982. Somatic variants of murine immunoglobulin light chains. *Nature (Lond.).* 298:380.
28. Reilly, E. B., R. M. Reilly, R. M. Breyer, R. T. Sauer, and H. N. Eisen. 1984. Amino

acid and nucleotide sequences of variable regions of mouse immunoglobulin light chains of the λ3-subtype. *J. Immunol.* 133:471.

29. Maniatis, T., E. F. Fritsch, and T. Sambrook. 1982. Extraction, purification, and analysis of mRNA from eukaryotic cells. *In* Molecular Cloning. Cold Spring Harbor Laboratory, New York. 194–203.
30. Wahl, G. M., M. Stern., and G. R. Stark. 1980. Efficient transfer of large DNA fragments from agarose gels to diazobenzyloxymethyl paper and rapid hybridization using dextran sulfate. *Proc. Natl. Acad. Sci. USA.* 76:3683.
31. Maxam, A. M., and W. Gilbert. 1980. Sequencing end-labeled DNA with base specific chemical cleavage. *Methods Enzymol.* 65:499.
32. Loh, D. Y., A. L. M. Bothwell, M. E. White-Scharf, T. Imanishi-Kari, and D. Baltimore. 1983. Molecular basis of a mouse strain-specific anti-hapten response. *Cell* 33:85.
33. Neuberger, M. S. and F. Calabi, 1983. Reciprocal chromsome translocation between *c-myc* and immunoglobulin γ2b genes. *Nature* (*Lond.*). 305:243.
34. Siekevitz, M., S. Y. Huang, M. L. Gefter. 1983. The genetic basis of antibody production: a single heavy chain variable region gene encodes all molecules bearing the dominant anti-arsonate idiotype in the strain A mouse. *Eur. J. Immunol.* 13:123.
35. Perlmutter, R. M., S. T. Crews, R. Douglas, G. Sorensen, N. Johnson, N. Nivera, P. J. Gearhart, and L. Hood. 1984. The generation of diversity in phosphorylcholine-binding antibodies. *Adv. Immunol.* 35:1.
36. Rocca-Serra, J., H. W. Matthes, M. Kaartinen, C. Milstein, J. Theze, and M. Fougereau. Analysis of antibody diversity: V-D-J mRNA nucleotide sequence of four anti-GAT monoclonal antibodies. A paucigene system using alternate D-J recombinations to generate functionally similar hypervariable regions. *EMBO.* 2:867.
37. Kaartinen, M., G. M. Griffiths, P. H. Hamlyn, A. F. Markham, K. Karjalainen, J. L. T. Pelkonen, O. Makela, and C. Milstein. 1983. Anti-oxazolone hybridomas and the structure of the oxazolone idiotype. *J. Immunol.* 130:937.
38. Perlmutter, R. M., J. L. Klotz, M. W. Bond, M. Nahm, J. M. Davie, and L. Hood. 1984. Multiple V_H gene segments encode murine antistreptococcal antibodies. *J. Exp. Med.* 159:179.
39. Kabat, E. A., T. T. Wu, and H. Bilofsky. 1979. Evidence supporting somatic assembly of the DNA segments (minigenes), coding for the framework and complementary-determining segments of immunoglobulin variable regions. *J. Exp. Med.* 149:1299.
40. Kunkel, H. G., V. Agnello, G. Joslin, R. J. Winchester, and J. D. Capra. 1973. Cross-idiotypic specificity among monoclonal IgM proteins with anti-γ-globulin activity. *J. Exp. Med.* 137:331.
41. Weigert, M., W. C. Raschke, D. Carson, and M. Cohn. 1974. Immunochemical analysis of the idiotypes of mouse myeloma proteins with specificity for levan or dextran. *J. Exp. Med.* 139:137.
42. Reth, M., A. L. M. Bothwell, and K. Rajewsky. 1981. *In* Immunoglobulin Idiotypes. C. Janeway and E. E. Sercarz, editors. Academic Press, Inc., New York. 20:169.
43. Slaughter, C. A., and J. D. Capra. 1983. Amino acid sequence diversity within the family of antibodies bearing the major antiarsonate cross-reactive idiotype of the A strain mouse. *J. Exp. Med.* 158:1615.
44. Gridley, T., M. N. Margolies, and M. L. Gefter. 1984. The association of various D elements with a single immunoglobulin V_H gene segment: influence on the expression of a major cross-reactive idiotype. *J. Immunol.* In press.
45. Azuma, T., V. Ingras, E. B. Reilly, and H. N. Eisen. 1984. Diversity at the variable–

joining region boundary of λ light chains has a pronounced effect on immunoglobulin ligand-binding activity. *Proc. Natl. Acad. Sci. USA.* 81:6139.

46. Meek, K., D. Jeske, M. Slaoui, O. Leo, J. Urbain, and J. D. Capra. 1984. Complete amino acid sequence of heavy chain variable regions derived from two monoclonal anti-*p*-azophenylarsonate antibodies of BALB/c mice expressing the major cross-reactive idiotype of the A/J strain. *J. Exp. Med.* 160:1070.
47. Pollok, P. A., J. F. Kearney, M. Vakil, and R. P. Perry. 1984. A biological consequence of variation in the side of D-J_H gene rearrangement. *Nature* (*Lond.*). 311:376.
48. Brodeur, P. H., and R. Riblet. 1984. The immunoglobulin heavy chain variable region (*Igh-V*) locus in the mouse. I 100 *Igh-V* genes comprise 7 families of homologous genes. *Eur. J. Immunol.* In press.