Germ Line Variable Regions That Match Hypermutated Sequences in Genes Encoding Murine Anti-Hapten Antibodies

Vivek David,[1] Nancy L. Folk and Nancy Maizels

Department of Molecular Biophysics and Biochemistry, Yale University School of Medicine, New Haven, Connecticut 06510

Manuscript received June 11, 1992
Accepted for publication July 10, 1992

ABSTRACT

We asked whether there are germ line immunoglobulin variable (V) segments that match sites of hypermutation in V regions encoding murine antibodies. Murine germ line DNA was probed with a panel of short deoxyoligonucleotides identical in sequence to segments of hypermutated V regions from hybridomas generated in the BALB/c response to the hapten 2-phenyloxazolone (Ox). Germ line sequences that match mutations in both heavy and κ light chain V regions were identified, and clones of some of these germ line V segments were obtained. Comparison of these clones with hypermutated V regions revealed regions of identity ranging in size from 7 to over 50 nucleotides. In an effort to separate the effects of antigen selection from the mutagenic process, we also searched for matches to a panel of silent mutations in V_H regions from germinal center B cells. Fourteen silent mutations occur among a collection of 36 hypermutated V_H regions from two separate germinal centers of C57BL/6 mice stimulated with the hapten 4-hydroxy-3-nitrophenyl. Matches to nine of these silent mutations can be found among published sequences of C57BL/6 V_H regions of the J558 family. Taken together, these data are consistent with the possibility that a template-dependent mutational process, like gene conversion, may contribute to somatic hypermutation.

SOMATIC hypermutation is a targeted process of mutagenesis which introduces single base changes into the rearranged variable regions encoding immunoglobulin heavy and light chain genes. This mutational process occurs following B cell activation by antigen and introduces single base changes at a rate approaching 10^{-3} per base per generation (McKEAN *et al.* 1984; SABLITZKY, WILDNER and RAJEWSKY 1985; ALLEN *et al.* 1987; BEREK and MILSTEIN 1987) which is 10^5–10^6-fold higher than the typical rate of mutation in mammalian somatic cells. Somatic hypermutation gives rise to clones that produce antibody molecules with 10–50-fold increased affinity for antigen, thus enhancing the specificity and efficiency of the immune response.

Two very different models have been advanced to explain the mechanism of hypermutation. One postulates that an error-prone DNA polymerase targeted to the immunoglobulin loci introduces the observed single base changes. Initially proposed to explain antibody variability (LEDERBERG 1959; BRENNER and MILSTEIN 1966), this model was later applied specifically to somatic hypermutation (see, for example, CREWS *et al.* 1981; GEARHART and BOGENHAGEN 1983). The other model ascribes hypermutation to a templated process of segmental recombination akin to gene conversion (reviewed by MAIZELS 1989, 1991; DAVID and MAIZELS 1989). This model originates in proposals by SMITHIES (1967a,b) and EDELMAN and GALLY (1967, 1970) that some sort of segmental recombination might contribute to antibody structure. SEIDMAN *et al.* (1978) suggested that shared homology among V regions might facilitate such intergenic recombination in somatic cells, and BALTIMORE (1981) noted that multiple rounds of gene conversion could explain the patchy sequence homology observed among members of immunoglobulin germ line heavy and light chain V region families. The idea that gene conversion might mediate somatic hypermutation fell into disfavor, however, when donors for particular mutations in murine immunoglobulin genes could not be found at allelic or highly homologous loci (BERNARD, HOZUMI and TONEGAWA 1978; CREWS *et al.* 1981; reviewed by TONEGAWA 1983).

In a variety of organisms, specific genomic sequences undergo targeted diversification in response to developmental or environmental stimuli. Examples include mating type loci in the yeasts *Saccharomyces cerevisiae* and *Schizosaccharomyces pombe* (HABER, ROGERS and MCCUSKER 1980; KLAR *et al.* 1980; KLAR 1987; reviewed by STRATHERN 1988), genes encoding variant cell-surface glycoproteins in trypanosomes (LONGACRE and EISEN 1986; ROTH *et al.* 1986), heavy and light chain V regions in the preimmune chicken (REYNAUD *et al.* 1987; THOMPSON and NEIMAN 1987; reviewed by MAIZELS 1987; REYNAUD *et al.* 1989), and heavy chain V regions in the preimmune rabbit (BECKER and KNIGHT 1990; KNIGHT and BECKER

[1] Present address: Department of Radiology, Brigham and Women's Hospital, 75 Francis Street, Boston, Massachusetts 02115.

1990). In all these cases, germ line sequences template the diversification of the targeted locus in a process that is commonly referred to as gene conversion.

Because targeted gene conversion occurs in such varied organisms, and because gene conversion is the only molecular mechanism that has been shown to diversify targeted regions of DNA in a regulated manner, we reasoned that it would be useful to test whether the sequences of hypermutated mammalian immunoglobulin genes are consistent with the existence of a templated mutational process. While the absence of matches for certain mutations has been interpreted as invalidating this possibility (BERNARD, HOZUMI and TONEGAWA 1978; CREWS *et al.* 1981; CHIEN *et al.* 1988), recent experiments have shown that, at least in certain instances, gene conversion can induce or be accompanied by untemplated mutations (THOMAS and CAPECCHI 1986; REYNAUD *et al.* 1987). This raises the possibility that a single molecular mechanism could generate both templated and untemplated mutations, presumably by a pathway that involves formation of a recombinational intermediate that carries both donor and recipient sequences, and its subsequent repair. Hence the absence of a match for any particular subset of mutations would not exclude a templated mechanism of mutagenesis.

If mutation is templated, even in part, then there will be matches in germ line DNA for segments of hypermutated variable regions. We asked in two different ways whether such matches exist. First, we probed murine germ line DNA with a panel of short deoxyoligonucleotides identical in sequence to segments of hypermutated V regions. We found matches for hypermutated positions in both heavy and κ light chain V regions, and we cloned germ line V regions that match some of these probes. Comparison of the hypermutated and germ line V_H and $V\kappa$ regions revealed regions of sequence identity ranging in size from 7 to over 50 nucleotides (nt). Second, in an effort to distinguish the effects of antigen selection from the mechanism of mutation, we also searched for germ line matches to silent mutations in V_H regions from germinal center B cells of C57BL/6 mice stimulated with the hapten 4-hydroxy-3-nitrophenyl (NP). Among the published J558 family V_H region sequences, we found matches to 9 of the 14 silent mutations in a panel of 36 hypermutated V regions sequenced from two separate germinal centers (JACOB *et al.* 1991). Our results are consistent with templated hypermutation, but clearly do not constitute proof of mechanism. Our data suggest, however, that it will be useful to pursue further experiments directly addressing the possible contribution of templated mutagenesis to somatic hypermutation of immunoglobulin genes.

MATERIALS AND METHODS

Oligonucleotide sequences:

OL5: 5′d(TGCTGGAACCAGTTCATGTA), $V\kappa$ codons 32–38
OL6: 5′d(TGCTGGAACCACTGCATGTA), $V\kappa$ codons 32–38
OL16: 5′d(ATCATTCCCAGCCACTCCAG), V_H codons 45–51
OL21: 5′d(CATTAACCAACTATGGTGTA), V_H codons 28–34
OL22: 5′d(TACACCATAGGTGGTTAATG), V_H codons 28–34
OL23: 5′d(GTGCTTCCACCAGTCCATAT), V_H codons 51–57
OL24: 5′d(CTTCTGCTGGAACCACTGCA), $V\kappa$ codons 31–38
OL25: 5′d(TGGAACCACTGCATGTAACT), $V\kappa$ codons 33–39
OL29: 5′d(GGGTCTGGAATGGCTGGGAG), V_H codons 43–50
OL33: 5′d(TGCTGGTACCAGTGCATGTA); $V\kappa$Ox-1 codons 32–38
OLJH4: 5′d(GAGGTTCCTTGACCCCAGTA); JH4 region
1804: 5′d(AGTCTGTCGAATTCCCACTGCCACTGAA); RI-$V\kappa$, 3′ end in codon 62
3417: 5′d(AGCAAGTTCAGCCTGGTTAAG) λgt10 forward
3418: 5′d(CTTATGAGTATTTCTTCCAGGGTA); λgt10 reverse
4422: 5′d(CGGAATTCAGCTGACCCAGTCTCCA); RI-$V\kappa$, 3′ end in codon 7
4423: 5′d(CGGAATTCGGTSMARCTGCAGSAGTCWGG); RI-V_H, 3′ end in codon 8
4424: 5′d(CGGAATTCATANAYSGCWGTGTCNTCAG); RI-V_H, 3′ end in codon 84

Notation for degenerate positions: N = A, T, C or G; Y = C or T; R = A or G; S = G or C; W = A or T; M = A or C. Oligonucleotides 1804, 4422, 4423 and 4424 were designed to amplify and clone V regions. Their 3′ sequences correspond to conserved positions in $V\kappa$ and V_H regions (ORLANDI *et al.* 1989; SCHLISSEL and BALTIMORE 1989); their 5′ sequences include *Eco*RI sites to facilitate λgt10 cloning.

Oligonucleotide labeling: Oligonucleotides were labeled for 45 min at 37° in 7-μl reactions containing 5–10 pmol oligonucleotide, 0.75 μl [γ-^{32}P]ATP (NEN,135 mCi/μl), 70 mM Tris, pH 7.6, 10 mM $MgCl_2$, 5 mM dithiothreitol, with 10 units T4 polynucleotide kinase (Biolabs). For gel hybridization probes were purified on a 15% acrylamide, 7 M urea gel as described by THIEN and WALLACE (1986); for hybridization to nylon or nitrocellulose membranes, probes were purified on a Sephadex G-50 spun column. Specific activities were 2–8 × 10^6 cpm/pmole.

Gel hybridization: Genomic DNA was isolated from the livers of BALB/c and C57BL/6 mice (STEFFEN *et al.* 1979). Following digestion with restriction enzymes (Biolabs), about 10 μg of DNA was loaded per 6-mm lane of a 0.75% agarose gel in Tris-acetate buffer, pH 7.8. Following electrophoresis, the DNA was not transferred but prepared for hybridization in the gel itself, since this increases the intensity of oligonucleotide hybridization signals about 3-fold (WALLACE and MIYADA 1987). The gel was soaked in 3 volumes of 0.5 M NaOH, 1.5 M NaCl for 2 × 20 min, rinsed with distilled water, then soaked in 3 volumes of 2 M NaCl, 0.5 M Tris, pH 7.4 for 2 × 20 min, rinsed with distilled water, placed on Whatman 3mm paper and dried for 1 hr

under vacuum at ambient temperature, then for 1 hr under vacuum at 60° (WALLACE and MIYADA 1987). The dried gel was floated off the Whatman paper into a dish of distilled water, placed in a plastic bag, and prehybridized at 37° for 4–6 hr in hybridization buffer (50 mM Hepes, pH 7.0, 0.9 M NaCl, 0.09 M sodium citrate, 0.1% sodium dodecyl sulfate (SDS), 0.2 mg/ml Ficoll, 0.2 mg/ml polyvinylpyrrolidone, 0.2 mg/ml BSA, 0.2 mg/ml sheared, boiled salmon sperm DNA) that had been passed through a 0.25-μm HA filter (Millipore). Prehybridization buffer was removed and fresh hybridization buffer (5 ml per 72 cm^2 dried gel) added along with 4–6 × 10^6 cpm of labeled oligonucleotide. After overnight hybridization, the gel was washed in 6 × SSC (200 ml per 72 cm^2 dried gel) for 2 × 30 min at room temperature, and then TMACS (100 ml per 72 cm^2 dried gel) for 2 × 20 min at room temperature. TMACS is 3 M tetramethylamonium chloride, 2 mM EDTA, 50 mM Tris, pH 8.0, containing 0.2% SDS. The tetramethylammonium chloride (Aldrich) is prepared as a 5 M stock solution and filtered through a Whatman #1 filter. The molar concentration (C) is determined from the refractive index η by the formula $C = (\eta - 1.331)/0.18$. Following a rinse in TMACS, the gel was sealed in a plastic bag containing prewarmed TMACS and incubated for 15 min at 56°, with agitation. The gel was rinsed for 10 min at room temperature in TMACS, then autoradiographed without further drying for 6–10 days with two intensifying screens at −70°.

In TMAC the contribution of G-C and A-T base pairs to duplex stability is equalized, so the melting profile is sharpened and largely independent of sequence (WOOD *et al.* 1985). We performed pilot experiments with control plasmids and oligonucleotides to determine empirically the sensitivity of hybridization to sequence mismatch. We found that a duplex formed between an oligonucleotide 20 bases in length and a much longer sequence is stable at 56° in the gel hybridization conditions described above, provided that the two sequences contain no internal mismatches; mismatches at the end of an oligonucleotide probe contribute less to destabilization than do internal mismatches.

Cloning and sequencing: Genomic V region libraries were produced by amplifying 20 ng BALB/c liver DNA in 50-μl reactions containing 1 pmol each primer, 400 μM dNTP, 50 mM KCl, 20 mM Tris, pH 8.3, 20 μg/ml gelatin, 2–3 mM $MgCl_2$, 80 units AmpliTaq polymerase (Cetus), in a 25-cycle reaction, annealing temperature 44°. Sequences of V_H (4423, 4424) and Vκ (1804, 4422) primers containing sites for *Eco*RI cleavage are shown above. Following amplification, DNA was extracted once with phenol-chloroform, once with ether, ethanol precipitated, *Eco*RI digested, and then cloned into the λgt10 vector and plated on *Escherichia coli* Hfl. Plaques were blotted to Zetabind nylon disks (AMF/Cuno), and the disks floated on solutions of 0.5 M NaOH, 1.5 M NaCl for 2 min, of 1.5 M NaCl, 0.5 M Tris, pH 7.4, for 2 min, rinsed in 5 × SSC, baked at 80° for 30 min, soaked in 5 × SSC, 0.5% SDS for 30 min at 65°, then prehybridized and hybridized in conditions similar to those for gel hybridization, except that solutions were not filtered, SDS was omitted from the TMACS, hybridization was at 56° to permit approximately one mismatch per probe, and rinses after hybridization were in 5 × SSC. Phage that gave positive hybridization signals were purified and grown in 1-ml liquid cultures on *E. coli* Hfl. The V region insert in each λgt10 clone was amplified with λgt10 flanking primers, purified by Centricon 30 centrifugation (Amicon), and sequenced with Sequenase (USB).

RESULTS

Hypermutation of the VκOx-1 and V_HOx-1 genes: The process of hypermutation can be studied separately from V region repertoire use because, in certain strains of inbred mice, a single combination of heavy and light chain germ line V regions encodes the majority of antibodies produced in response to immunization with a specific hapten coupled to a protein carrier. One of the best characterized examples of such a strain-specific response is the BALB/c response to the hapten 2-phenyloxazolone (Ox), in which most antibodies are encoded by the germ line V region combination V_HOx-1 and VκOx-1. MILSTEIN and collaborators have published extensive sequence and affinity analysis of hypermutated antibodies from the Ox response (KAARTINEN *et al.* 1983a,b; GRIFFITHS *et al.* 1984; EVEN *et al.* 1985; BEREK, GRIFFITHS and MILSTEIN 1985; BEREK and MILSTEIN, 1987, 1988; RADA *et al.* 1991). We took advantage of these published sequences and generated a panel of short oligonucleotides that matched regions of hypermutated V_HOx-1 and VκOx-1 genes, as shown in Figure 1. Most of the oligonucleotide probes span only a single hypermutated base; however, if mutation is templated, neighboring mutations may derive from a single donor and appear to be genetically linked. Some probes were therefore designed to test for matches to two mutations a few bases apart.

Several Vκ probes were designed to assay for matches to mutations at codons 34 and 36 of VκOx-1 (Figure 1). Mutations at both these codons are found in a number of independently isolated hybridomas from the Ox response (BEREK and MILSTEIN 1987), although kinetic studies suggest that appearance of these mutations is not usually temporally coupled (RADA *et al.* 1991). These residues should be under strong antigen selection, since structural analysis has shown that they are within the antigen binding region of the molecule (ALZARI *et al.* 1990). Chain recombination experiments have demonstrated that either of the two replacements at codon 34 [His (CAC) to Gln (AAC) or Asn (CAG)] increases binding affinity about 10-fold, so the repeated occurrence of mutations at this position could be ascribed to selection for high affinity antibodies. In contrast, the Tyr to Phe (TAC to TTC) mutation at codon 36 does not affect antibody affinity directly (BEREK and MILSTEIN 1987), although amino acids near to Gln and Asn residues can influence the rate of deamination and could be subject to selection on that basis (ROBINSON and ROBINSON 1991).

One of the five V_H region probes was designed to test for matches for two neighboring mutations: OL16 matches two single base changes that result in a replacement mutation at codon 50 of one V_HOx-1 derived V region (GRIFFITHS *et al.* 1984). The other V_H

```
                          34      36
                        S   Y   M   H   W   Y   Q   Q   K
VκOx-1                  AGT TAC ATG CAC TGG TAC CAG CAG AAG
OL4                         --- --- --- --- --- --- --
OL5                         --- --- A-- --- -T- --- --
OL6                         --- --- --G --- -T- --- --
OL24                            --  --G --- -T- --- --- ---
OL25                    --- --- --- --G --- -T- --

                            31
                        S   L   T   S   Y   G   V
VHOx-1                  TCA TTA ACC AGC TAT GGT GTA
OL21                     -- --- --- -A- --- --- ---
OL22                     -- --- --- -C- --- --- ---

                                46              50          53
                        K   G   L   E   W   L   G   V   I   W   A   G   G   S   T
VHOx-1                  AAG GGT CTG GAG TGG CTG GGA GTA ATA TGG GCT GGT GGA AGC ACA
OL29                     - --- --- --A --- --- --- -
OL16                        --- --- --- --- --- A-G --
OL23                                                --- --- A-- --- --- --- --
```

FIGURE 1.—Probes for hypermutated regions of VκOx-1 and V_HOx-1. The sequences of portions of the coding regions of VκOx-1 (codons 31–39) and V_HOx-1 (codons 28–34 and 43–57) are shown, with translated sequence shown above each codon and the 20 nt hypermutated region targeted by each synthetic probe shown below the appropriate portion of the sequence. Dashes indicate identity with the germ line sequence. A line is drawn above codons that encode amino acids in complementarity-determining regions (CDRs) of the antibody. All oligonucleotides except OL29 probe replacement mutations. The sequence of each probe is shown in MATERIALS AND METHODS.

oligonucleotides, OL21, OL22, OL23 and OL29, were all designed to probe single base changes that have been observed in two or more independently isolated hybridomas (BEREK and MILSTEIN 1987). One of these five V_H oligonucleotides, OL29, probes for a silent mutation, and the rest probe for replacement mutations.

Probing germ line DNA for matches with regions of VκOx-1 mutated sequence: To establish that the TMAC hybridization protocol (see MATERIALS AND METHODS) can distinguish sequences that match the probe perfectly from sequences that differ at one or more bases, we tested whether a 20-nt probe that matches the germ line VκOx-1 region hybridizes only with correct sequences in genomic blots. EVEN *et al.* (1985) cloned and sequenced 13 germ line Vκ regions that hybridize with a VκOx-1 cDNA probe, including one (H3, a 3.3-kb *Hin*dIII fragment) that carries the VκOx-1 region itself, and three others (H6, H9 and R9; 1.7-, 3.6- and 4.3-kb *Hin*dIII fragments, respectively) identical in sequence to VκOx-1 within the region probed. When *Eco*RI and *Hin*dIII digests of BALB/c liver DNA are probed with OL4, which matches the VκOx-1 germ line sequence, and then washed under stringent conditions with TMAC, the oligonucleotide hybridizes as expected to *Hin*dIII fragments of 4.3, 3.6, 3.3 and 1.7 kb (Figure 2, lanes 1 and 2). Two cloned germ line Vκ regions have been shown to differ from the probe at single positions (R2, a 2.1-kb *Hin*dII fragment, and R13, a 4.0-kb *Eco*RI fragment; EVEN *et al.* 1985). These bands are only faintly apparent in the autoradiogram, indicating that hybridization discriminates perfect and imperfect matches.

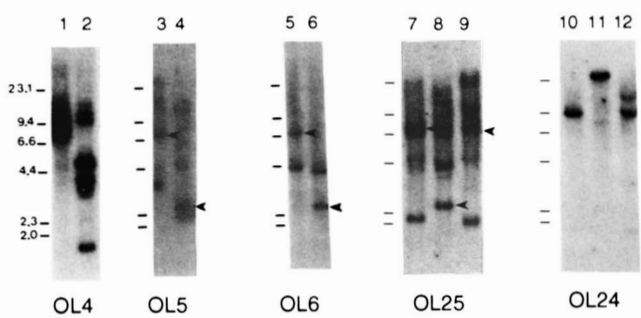

FIGURE 2.—Mouse germ line DNA probed for sequences that match hypermutated segments of VκOx-1. BALB/c liver DNA digested with *Eco*RI (lanes 1, 3, 5, 7 and 10) or *Hin*dIII (lanes 2, 4, 6, 8 and 11) or C57BL/6 DNA digested with *Eco*RI (lanes 9 and 12) was probed with the labeled oligonucleotide OL4 (VκOx-1 germ line sequence), OL5, OL6, OL25 and OL24. Hybridizations shown in lanes 3–9 include as control the OLJH4 oligonucleotide, which hybridizes to a single band in each digest (marked by an arrowhead). Mobility of marker fragments is indicated on the left of each set of lanes. DNA preparation, oligonucleotide labeling, hybridization and washes were performed as described in MATERIALS AND METHODS.

Oligonucleotides OL5 and OL6 were used as probes to assay for the presence of germ line sequences that match the hypermutated VκOx-1 regions shown in Figure 1. Figure 2 shows the results of hybridization with these oligonucleotides. These hybridizations included oligonucleotide OLJH4, complementary to the single-copy JH4 region, as a positive internal control for gene copy number and the stringency of hybridization. The hybridizing JH4 fragments in germ line DNA digests are marked by arrows in Figure 2. Following hybridization with OL5 (Figure 2, lanes 3 and 4), four bands are apparent in both the *Eco*RI and *Hin*dIII digestion pattern, but none of them is as intense as the band hybridizing with the

control oligonucleotide, OLJH4. This suggested that four germ line regions are near but imperfect matches to this oligonucleotide probe. Hybridization of OL6 to *Eco*RI and *Hin*dIII digests of genomic DNA reveals a single band in each lane (Figure 2, lanes 5 and 6); comparison with the internal control suggests that the hybridizing band is probably a perfect match present in single copy. The results show that there is a single genomic sequence that matches the 20-nt OL6 probe.

To ask whether the fragment that matches OL6 is identical in sequence to VκOx-1 beyond the 20 nt probed by this oligonucleotide, genomic digests were probed with oligonucleotides displaced 3-nt 5′ (OL25) or 4-nt 3′ (OL24) of the OL6 sequence (see Figure 1 for oligonucleotide sequences). As shown in Figure 2 (lanes 7 and 8, 10 and 11), both these oligonucleotides hybridize to digests of BALB/c liver DNA, but neither hybridizes to the same band as OL6, showing that the match to the VκOx-1 sequence does not extend much beyond the boundaries of the OL6 probe. Cloning experiments (see below) provide a more quantitative test of the extent of sequence match between hypermutated and genomic V regions.

The probes described above were designed to assay for matches with hypermutated positions in BALB/c mice, but since there appears to be fairly good conservation of V region sequence between BALB/c and C57BL/6, we also tested hybridization to DNA from this strain. The BALB/c and C57BL/6 hybridization patterns of OL25 are very similar, although there are several bands evident in C57BL/6 DNA that are not apparent in BALB/c (compare Figure 2, lanes 7 and 9). OL24 hybridizes to a 9.4-kb RI fragment in both C57BL/6 and BALB/c DNA, and somewhat less intense hybridization to a larger fragment is evident in the C57BL/6 digest (Figure 2, lanes 10 and 12). These data suggest that both oligonucleotides hybridize to sequences or sequence families that are relatively well conserved in the two strains.

Cloned sequences that match hypermutated VκOx-1 regions: We used OL5 as a probe to screen libraries of BALB/c liver DNA for matching sequences. Two separate Vκ region libraries were generated by amplifying germ line DNA with degenerate PCR primers complementary to conserved sequences in Vκ, and cloning into λgt10 (see MATERIALS AND METHODS). The screening produced two different Vκ region clones that match OL5 at codon 34, but neither of these matches the probe at codon 36. Figure 3 shows these V region sequences, compared with VκOx-1. One of the cloned Vκ regions is identical to R2, described by EVEN *et al.* (1985); the other, which we call Vκ5-5, appears not to have been previously described. A recurrent change among a collection of VκOx-1-derived sequences generated by directly amplifying DNA from germinal center B cells is replacement of both Ser-31 and His-34 with Asn (RADA *et al.* 1991). Vκ5-5 differs from VκOx-1 at both these positions and could serve as a template for segmental mutation of these positions, which are underlined in Figure 3. In control experiments we identified another germ line Vκ region, Vκ33-1, that is identical to Vκ5-5 at codon 31. Furthermore, both Vκ5-4 and Vκ5-5 carry three matches for hypermutated sites that were not probed in the screen (the legend to Figure 3 notes the positions of these matches). Thus it is unlikely that the sequence matches are due to polymerase chain reaction (PCR) artifact.

Probing germ line DNA for matches with hypermutated V_HOx-1 sequences: We used five different oligonucleotide probes to assay BALB/c germ line DNA for matches with hypermutated regions in the V_HOx-1 heavy chain. Figure 1 shows the sequences of portions of the V_HOx-1 germ line V region and corresponding sequences from hypermutated V_HOx-1 regions. Figure 4 shows autoradiograms of digests of germ line DNA hybridized with these 5 probes; the OLJH4 probe was included as a positive control in reactions with OL21, OL22 and OL23. OL16 hybridized to two fragments in both *Eco*RI and *Hin*dIII digests of BALB/c DNA, and hybridization to a 6.2 kb fragment in each digest was particularly intense (lanes 1 and 2). OL29 hybridized to about six fragments in *Eco*RI and *Hin*dIII digests of BALB/c DNA (Figure 4, lanes 4 and 5). OL23 hybridizes to four fragments in the BALB/c *Eco*RI digest and to three fragments in the *Hin*dIII digest (lanes 7 and 8). In contrast, there appears to be no match in germ line DNA for the 20-nt region spanned by either OL21 or OL22, although these hybridizations showed the expected bands hybridizing to the OLJH4 probe included as a positive control (lanes 9–11 and 12–14).

As with the Vκ-specific probes, we also tested for hybridization to digests of C57BL/6 DNA. While OL16 did not hybridize to an *Eco*RI digest of C57BL/6 DNA (lane 3), five of the bands in the C57BL/6 *Eco*RI digest are of similar mobility with bands in RI-digested BALB/c DNA (lane 6). Thus both the V_H and Vκ probes hybridize to sequences or sequence families that are somewhat but not completely conserved between C57BL/6 and BALB/c mice.

Cloned sequences that match hypermutated V_H regions: To identify cloned V regions that match hypermutated V_H sequences, probes OL16, OL23 and OL29 were used to screen two separate λgt10 libraries produced by PCR amplification of BALB/c DNA with degenerate V_H region primers (see MATERIALS AND METHODS). Sequences of clones are shown in Figure 5. One clone, 16-5, exactly matches the mutations at codon 50 (GTA to ATG) probed by OL16, and three clones (29-4, 29-8 and 29-44) match a silent mutation (GTA to GTG) frequently observed

```
                          10                  15                  20
VκOx-1      CCA GCA ATC ATG TCT GCA TCT CCA GGG GAG AAG GTC ACC ATG ACC TGC AGT
 5-4 = R2   --- --- --- --A G-- --- --- --T --- --- --- --- --C --- --- ---
 5-5        --- --- --- --- --- --- --- --- --- C-- --A --- --- --A --- --- ---
 33-1       --- --- --- --- --- --- --- -T- --- --- --- --- --- --- -G- --- --G

        25       CDR1        31              35                  40
VκOx-1 GCC AGC TCA AGT GTA AGT TAC ATG CAC TGG TAC CAG CAG AAG TCA GGC ACC TCC
 5-4    --- --- --- --- --- --- --- --- A-- --- --- --- --- --A C-- --A T-- ---
 5-5    --- -T- --- --- --- -A- --- --- A-- --- --- --- --- --A C-- --A T-- ---
 33-1   --- --- --- --- --- -A- --- --- --- --- --- --- --- --- --- A-- --- ---
 OL5                            --- --- A-- --- -T- --- --

                45                  50      CDR2        55              60
VκOx-1 CCC AAA AGA TGG ATT TAT GAC ACA TCC AAA CTG GCT TCT GGA GTC CCT GCT CGC
 5-4    --- --- -T- --- --- --- -GT -T- --- --C --- --- --- --- --x --- --- ---
 5-5    --- --- -T- --- --- --- --- --- --- --- --- --- --x --- --- --- --- --- ---
 33-1   --- --- CTC --- --- --- --- --- --- --- --- --- --- --- --- --- --A -G- ---
```

FIGURE 3.—Sequences of cloned Vκ regions that match probes for hypermutated sequences. The top line shows the sequences of VκOx-1, the germ line light chain V region used in the majority of antibodies in the BALB/c response to Ox. A line is drawn above codons that encode amino acids in complementarity-determining regions (CDRs) of the antibody. Sequences of 3 Vκ regions cloned by PCR amplification of germ line DNA are shown below the VκOx-1 sequence. Dashes indicate identity with the germ line sequence; x indicates an uncertainty in the sequence. We note that differences from VκOx-1 at positions 21, 31, 34, 42, 50 and 61 all correspond to mutations found in VκOx-1-derived hypermutated Vκ regions. Vκ5-5 differs from VκOx-1 at codons 31 and 34, and segmental transfer of this region of sequence would replace both Ser-31 and His-34 with Asn, a linked change that occurs among several members of a collection VκOx-1-derived sequences generated by directly amplifying DNA from germinal center B cells (RADA *et al.* 1991); these differences at codons 31 and 34 are underlined. The sequence of probe oligonucleotide OL5 is shown below the sequences of the cloned V regions. Vκ5-4 is identical in sequence to R2, previously described by EVEN *et al.* (1985).

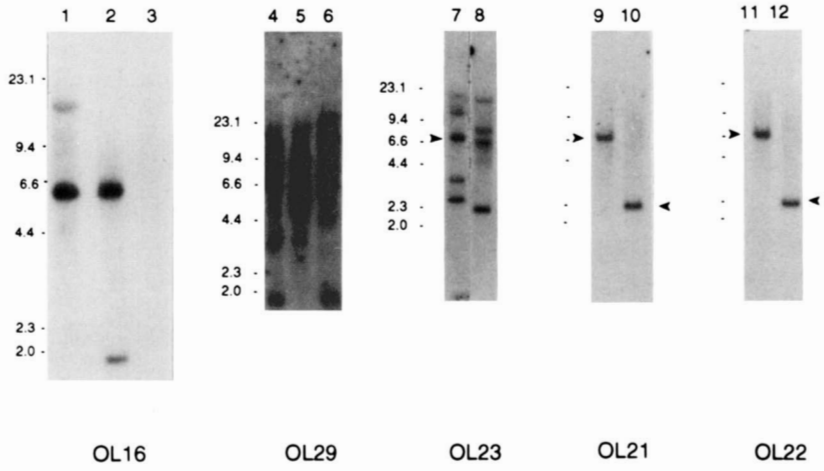

FIGURE 4.—Mouse germ line DNA probed for sequences that match hypermutated segments of V_HOx-1. BALB/c liver DNA digested with *Eco*RI (lanes 1, 4, 7, 9 and 11) or *Hin*dIII (lanes 2, 5, 8, 10 and 12) or C57BL/6 DNA digested with *Eco*RI (lanes 3 and 6) was probed with labeled oligonucleotide OL16, OL29, OL23, OL21 or OL22, as indicated. Hybridizations shown in lanes 7–12 include as control the OLJH4 oligonucleotide, which hybridizes to a single band in each digest (marked by an arrowhead). Mobility of marker fragments is indicated on the left of each set of lanes.

at codon 50. Two clones (29-4 and 29-44) match the mutation at codon 53 probed by OL23. We also isolated a clone that carries the mutation probed by OL23 from a conventional EMBL3 library, and part of the sequence of this cloned V region, G23-5, is shown in Figure 5.

While TMAC hybridization should not produce false positives, failure to hybridize does not necessarily mean that a mutation is not present in germ line DNA; rather, because the TMAC hybridization protocol is capable of distinguishing perfect matches from single base mismatches, failure to hybridize may simply mean that the germ line DNA region containing the mutation does not match the entire length of the oligonucleotide probe. For example, as shown in Figure 4, OL21, an oligonucleotide that probes an AGC to AAC mutation at codon 31, did not hybridize to germ line DNA. Nonetheless, we isolated six clones that carry matches to the G to A change at codon 31 that OL21 was designed to probe. All of these clones differ in sequence from OL21 at one or more positions, consistent with the absence of hybridization of OL21 to germ line DNA on blots.

Homology between germ line V_H regions and hypermutated regions: Templated mutation requires some degree of homology between donor and target sequences. Figure 6 compares the homologies of portions of some of the cloned germ line V_H regions with hypermutated V_H regions from anti-Ox hybridomas (BEREK, GRIFFITHS and MILSTEIN 1985). When two

	20				25					30	CDR1			
V_HOx-1	CTG	TCC	ATC	ACT	TGC	ACT	GTC	TCT	GGG	TTT	TCA	TTA	ACC	AGC	TAT	GGT	GTA
16-5	---	---	---	--A	---	--C	---	--A	---	--C	---	---	---	G--	---	---	---
21-4											---	---	---	-A-	-C-	---	---
29-4	---	---	---	--C	---	--A	---	---	--A	--C	---	---	--T	---	---	---	---
29-8	---	---	--x	--A	---	---	---	--A	---	--C	---	---	--x	GA-	---	---	---
29-44	---	---	---	--C	---	--A	---	---	C--	--C	---	---	--T	---	---	---	---
29-5	---	---	---	--A	---	---	---	---	--T	--C	---	---	---	-A-	-C-	---	---
29-1	---	---	---	--A	---	---	---	---	--T	--C	---	---	---	-A-	-C-	---	---
29-47	---	---	---	--A	---	---	---	---	--T	--C	---	---	---	-A-	-C-	---	---
29-12	---	---	---	--A	---	---	---	---	--T	--C	---	---	---	-A-	-C-	---	---
G23-5	---	---	---	--A	---	---	---	---	---	--C	---	---	---	---	---	-C-	A--
OL21											--	---	---	-A-	---	---	---

	35				40					45					50	
V_HOx-1	CAC	TGG	GTT	CGC	CAG	CCT	CCA	GGA	AAG	GGT	CTG	GAG	TGG	CTG	GGA	GTA	ATA
16-5	A--	---	---	---	---	---	---	---	---	---	---	---	---	---	---	A-G	---
21-3	xxx	xxx	xxx	T--	---	T--	---	---	---	---	---	---	---	---	--x	---	---
21-4	---	---	---	---	---	T--	---	---	---	---	---	---	---	---	--x	---	---
29-4	---	---	---	---	---	T--	---	---	---	---	---	---	---	---	---	--G	---
29-8	AG-	---	A--	---	---	---	---	---	---	---	---	---	---	---	---	--G	---
29-44	---	---	---	---	---	T--	---	---	---	---	---	---	---	---	---	--G	---
29-5	---	---	---	---	---	T--	---	---	---	---	---	---	---	---	---	---	---
29-1	---	---	---	-a-	---	T--	---	---	---	---	---	---	---	---	---	---	---
29-47	---	---	---	---	---	T--	---	---	---	---	---	---	---	---	---	---	---
29-12	---	---	---	---	---	TA-	---	---	---	---	---	---	---	---	---	---	---
G23-5	---	---	---	---	---	--A	---	---	---	---	---	---	---	--T	---	---	---
OL29										-	---	---	--A	---	---	---	-
OL16											---	---	---	---	---	A-G	--
OL23																	---

	CDR2	55					60						65				
V_HOx-1	TGG	GCT	GGT	GGA	AGC	ACA	AAT	TAT	AAT	TCG	GCT	CTC	ATG	TCC	AGA	CTG	AGC
16-5	---	-G-	-A-	---	---	---	G-C	---	---	--A	--T	--G	AAA	---	---	---	---
21-3	---	-G-	-A-	---	---	---	---	---	---	--A	---	T--	-AA	---	---	---	---
21-4	---	---	---	---	---	---	---	---	---								
29-4	---	AG-	---	---	---	---	G-C	--C	---	G-A	C--	T--	--A	---	---	---	---
29-8	---	-G-	---	---	---	---	T-C	---	---	--A	---	---	-A-	---	---	---	---
29-44	---	AG-	---	---	---	---	G-C	---	---	G--	xxx	xxx	--A	---	---	---	---
29-5	---	-G-	---	---	---	---	T-C	---	---	--A	---	---	-A-	---	---		
29-1	---	-G-	-A-	---	---	---	---	---	---	--A	---	T--	-AA	---	---	---	---
29-47	---	-G-	-A-	---	---	---	---	---	---	--A	---	T--	-AA	---	---	---	---
29-12	---	xxx	xxx	---	---	---	---	---	---	--A	---	T--	-AA	---	---	---	---
G23-5	---	A--	---	---	G--	---	---	---	---	--A	---	---	-AA	---	---	---	---
OL23	---	A--	---	---	---	--											

FIGURE 5.—Sequences of cloned V_H regions that match probes for hypermutated sequences. The top line shows part of the sequence of V_HOx-1, the germ line heavy chain V region used in the majority of antibodies in the BALB/c response to Ox. A line is drawn above codons that encode amino acids in complementarity-determining regions (CDRs) of the antibody. Sequences of portions of 10 V_H regions cloned by PCR amplification of germ line DNA and one V_H region cloned from a conventional library (V_HG23-5) are compared to the V_HOx-1 sequence. A dash indicates identity with the germ line sequence, x indicates uncertainty in the sequence. The match in V16-5 to linked mutations at codon 50 is underlined. Sequences of probe oligonucleotides are shown below V region sequences. For brevity, parts of V region sequences are omitted from the figure.

or more germ line V_H regions match the hypermutated sequence, one example each of minimal and maximal identity is shown. From Figure 6 it is evident that sequences of 7 to over 50 nt could serve as targets for templated mutation. These numbers are consistent with a similar analysis of the cloned Vκ sequences (not shown).

Matches for silent mutations in the C57BL/6 response to NP: B cells undergo intense selection by antigen *in vivo*. Since most published sequences of hypermutated V regions derive from cells with significantly increased affinity for antigen, mutations that enhance affinity will be overrepresented in this data set. In order to separate the effects of antigenic selection from the mechanism of mutation, it seemed important to search for matches in a panel of hypermutated sequences that had undergone minimal selection. The hapten NP induces a strain-specific immune response in C57BL/6 mice, so sites of hypermutation can be readily identified. JACOB *et al.* (1991) have sequenced 36 V_H regions that were PCR-amplified directly from DNA of germinal center B cells of C57BL/6 mice immunized with NP. These sequences seemed especially appropriate for analysis because they do not carry mutations typical of affinity selection (JACOB *et al.* 1991) and because the activated B cells were never selected as hybridomas. Among these 36 V_H regions, all of which derive from V_H186.2, there are a total of 14 silent mutations at 14 different positions (excluding sites adjacent to the V-D junction). We searched for matches to these 14 mutations among approximately 30 published sequences of unmutated V_H regions of the J558 family in C57BL/6 mice (BOTHWELL *et al.* 1981; MAIZELS and BOTHWELL

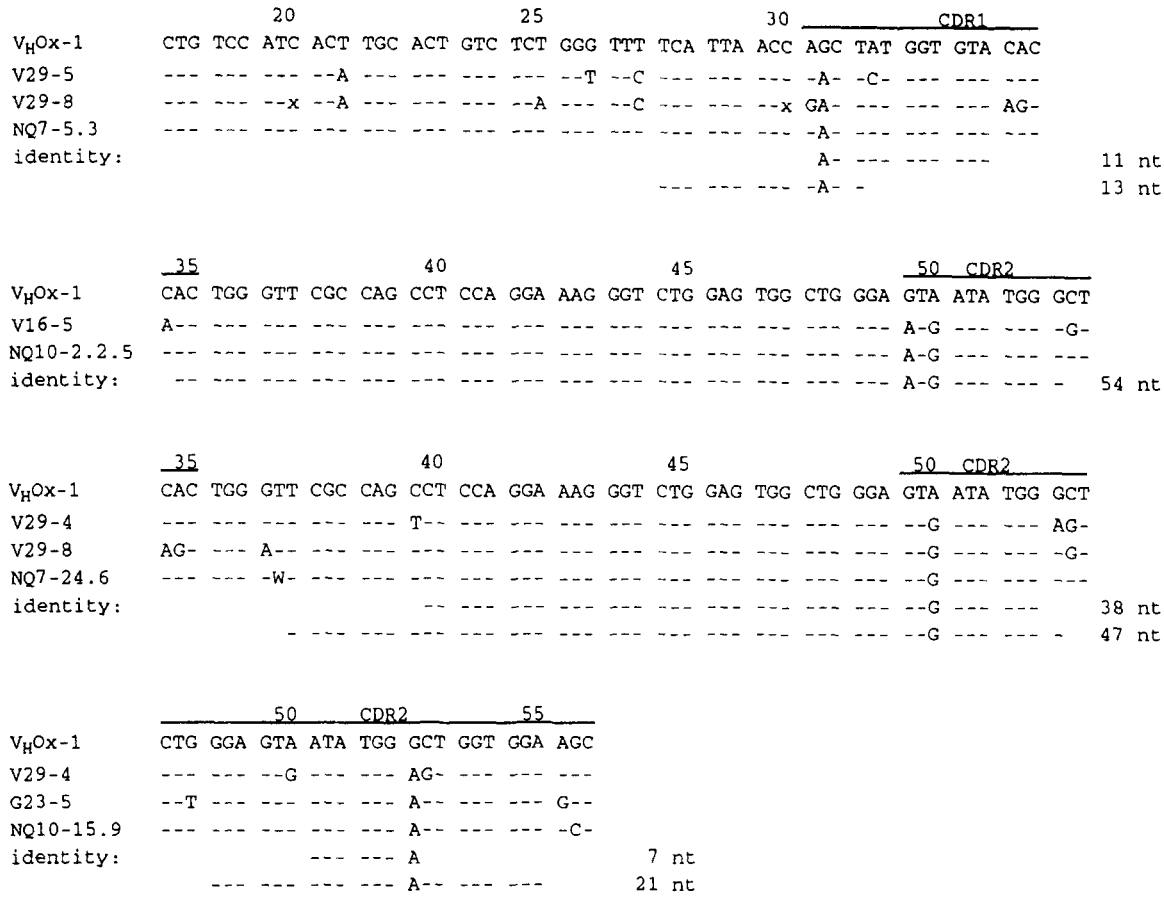

FIGURE 6.—Extent of homology between hypermutated antibodies and germ line V_H regions at positions at which templated mutation may have occurred. Portions of germ line V_H regions are compared with hypermutated V regions from anti-Ox hybridomas (BEREK, GRIFFITHS and MILSTEIN 1985) throughout the region of homology; notations are as in Figure 3. When more than one germ line V_H region matches the hypermutated region, an example of minimal and maximal identity is shown.

1985; SIEKEVITZ *et al.* 1987; FORSTER, GU and RAJEWSKY 1988). Matches for 9 of the 14 silent mutations are apparent among these published sequences, as shown in Figure 7. The germ line sequences that match the silent mutations at positions 10 and 73 also carry fortuitous matches to linked mutations, which are underlined in Figure 7. CUMANO and RAJEWSKY (1986) have noted another example of a single germ line V region that matches several linked sites of hypermutation in an anti-NP antibody.

DISCUSSION

We have searched for matches in germ line DNA to hypermutated regions of sequence, and have found matches among cloned germ line V regions for about two-thirds of the hypermutated regions tested. The germ line sequences characterized include matches for replacement mutations, and silent mutations, as well as several examples of linked mutations.

Technical considerations: The matches to oligonucleotide probes in genomic DNA were demonstrated initially by genomic blotting using a stringent TMAC hybridization protocol which could be shown, by appropriate pilot experiments and internal controls, to discriminate perfect matches from single base mismatches. We cloned germ line sequences that matched some of the probed hypermutated sites from V region libraries generated by PCR. We turned to PCR-based libraries for V region cloning in order to streamline the screening process and to ensure that V region sequences were not lost during cloning (see below). A potential concern about PCR-generated libraries is that Taq polymerase errors could create an apparent match where none really exists. However, several groups have shown that the error rate of PCR is actually quite low: it is well below 10^{-3}, and probably closer to 10^{-4} under amplification conditions resembling our own (TINDALL and KUNKEL 1988; KEOHAVONG and THILLY 1989; MEYERHANS *et al.* 1989). Two other kinds of data make it unlikely that matches in the genes we have cloned are due to artifacts of PCR amplification: (1) In all cases in which we probed for a single base change, two or more different clones were found that matched the position probed but that were quite distinct throughout the rest of their sequences; (2) several V regions initially identified with

```
                                    10
VH186.2              CAG CCT GGG GCT GAG CTT GTG AAG
GC8:2; GC24:3        --- --- --- --- --A --- --- ---
GC24:2               --- --- --- A-- --A --- --- ---
V23                  --- --- --- A-- --A --G --- ---

                                    20
VH186.2                  TCA GTG AAG CTG TCC TGC AAG
GC8:1; GC24:1            --- --- --- T-- --- --- ---
132.16, MS23.6.18        --- --- --- T-- --- --- ---

                                    37
VH186.2                  ATG CAC TGG GTG AAG CAG AAG
GC24:1                   --- --- --- --A --- --- ---
119.13                   --- --- --- --A --A --- ---
MS23.6.18                --- --- --- -xA --- --- ---

                                    53
VH186.2                  ATT GAT CCT AAT AGT GGT GGT
GC24:1                   --- --- --- --C --- --- -A-
133.25                   --- A-- --- --C -A- --- ---
C1A4                     --- A-- --- -GC --- --- TA-

                                    59
VH186.2                  GGT ACT AAG TAC AAT GAG AAG
G8:1                     --- --- --- --T --- --- -G-
G24:1                    --- -G- --- --T --- --- ---
V130,133.23              AA- --- --A --T GCC CC- ---

                                    60
VH186.2                  ACT AAG TAC AAT GAG AAG TTC
GC24:1                   --- --- --- --C --- --- ---
205.12                   --- -GC --- --C C-- --- ---

                                    69
VH186.2                  AAG GCC ACA CTG ACT GTA GAC
GC8:1                    --- --- --- T-- --- --- ---
V3                       --- --- --- T-- --- --- ---

                                    73
VH186.2                  ACT GTA GAC AAA CCC TCC AGC
GC8:6                    --- --- --- --G --- --- ---
GC8:2                    --- --G --- --G --- --- -A-
GC8:2                    --- --- --- --G T-- --- ---
205.12, 219.23           --- --- --- --G T-- --- ---

                                    81
VH186.2                  GCC TAC ATG CAG CTC AGC AGC
GC8:1                    --- --- --- --A --- --- ---
132.16,133.16            --- --- --- --A --- --- ---
C1A4                     --- --- --- --A --G --- ---
```

FIGURE 7.—Silent mutations in germinal center V_H regions of NP-immunized C57BL/6 mice and matching regions of J558 family V segments. Silent mutations in V_H regions of two different germinal centers (JACOB *et al.* 1991) are compared with the progenitor V_H186.2 region and with matching regions in J558 family germ line V_H segments. The number in each portion of the figure denotes the position of the codon at which the silent mutation occurs. The matches to the silent mutations at codons 10 and 73 that include fortuitous matches to linked mutations are underlined. The two germinal centers, denoted GC8 and GC24, comprise two separate cell lineages. The number of recurrences of each silent mutation within each lineage is shown after the colon, so in the first example there were 2 V_H regions in GC8 and 3 V_H regions in GC24 that carried the GAG to GAA mutation at codon 10, but no mutation at neighboring positions. Only 7–8 codons are shown in each example, for the sake of brevity; some matches extended farther than shown in the figure. The published sources of V region sequences are: V23, V130 (BOTHWELL *et al.* 1981); 119.13, 132.16, 133.16, 133.23, 133.25, 219.23 (MAIZELS and BOTHWELL 1985); MS23.6.18, C1A4, V3 (SIEKEVITZ *et al.* 1987); 205.12 (FORSTER, GU and RAJEWSKY 1988).

one probe turned out upon sequence analysis to match a second, unprobed and unselected mutation.

Another reason we turned to PCR-generated libraries is that published data as well as our own observations suggest that V region sequences may be difficult to clone by traditional techniques. Several laboratories have attempted extensive screens for murine V regions, but the number of V_H sequences actually cloned is about an order of magnitude lower than the minimum number of 500–1000 V_H sequences estimated by hybridization (LIVANT, BLATT and HOOD 1986). In addition, only a fraction of V_H segments could be cloned from the chicken genome in an exhaustive screen (REYNAUD *et al.* 1989), and only about two-thirds of Vλ segments could be cloned from sheep genomic DNA (REYNAUD *et al.* 1991). It may be that the genomic organization of V regions makes them difficult to clone. Members of a highly repetitive gene family, organized as direct or inverted repeats, may tend to be lost during propagation on standard host strains. In fact, in a model experiment where we attempted to clone homologous V regions in head-to-head organization, we found that nearly 1 kb of spacer between the V regions was necessary to produce a stable plasmid clone (T. TSUKUDA, N. BATCHVAROVA and N. MAIZELS, unpublished data).

We have thus far been unable to isolate clones that match the mutations probed by OL25 and OL29, both of which hybridize to a number of fragments in germ line DNA. This may reflect a biased representation of V region sequences in the PCR-based libraries. The major limitation of PCR-based cloning of repetitive gene families is that the sequences amplified will depend on the exact choice of primer, and that highly divergent V segments will not be represented in the library. The primers we used matched framework sequences of V_H and Vκ that appear to be highly conserved (ORLANDI *et al.* 1989; SCHLISSEL and BALTIMORE 1989), but it would be very surprising if all V segments were amplified equally well by these primers, and there are undoubtedly some truncated pseudo-V segments that will not be amplified at all. Alternatively, some hybridization in the genomic blots may be to sequences that do not derive from V regions, and matches to such probes would not be retrieved from PCR-based V region libraries. In genomic libraries we have identified clones that carry as much as 18 bases of contiguous match to probe sequences, but do not derive from V region families;

and similar results have been obtained by others attempting to screen conventional genomic libraries for matches to hypermutated regions of sequence (WYSOCKI, GEFTER and MARGOLIES 1990).

Germ line matches for hypermutated sequences: We find regions of identity between germ line and hypermutated V_HOx sequences ranging from 7 to 51 nt in length. While the regions of sequence that undergo gene conversion on an evolutionary time scale are relatively long (*e.g.*, see MAEDA and SMITHIES 1986), in both chickens and trypanosomes patches of sequence as short as 4 nt in length undergo templated diversification in response to developmental stimuli. The lengths of the matches that we observe are therefore consistent with a templated mechanism of mutation.

If hypermutation is not templated, what are some alternative explanations for the existence of matches between hypermutated and germ line sequences? The pool of germ line V regions is very large: hybridization studies estimate that there are 500–1000 V_H segments in the BALB/c J558 family alone (LIVANT, BLATT and HOOD 1986). A question often raised is whether one could expect to find any given single base change within this large sequence pool. However, a rough calculation shows that this is unlikely. A V segment is about 300 nt in length. Assuming that sequence is random, and that all but 50 positions are invariant due to structural constraints, one can estimate how many V regions there would have to be in order to find all possible changes at all the 50 positions free to undergo variation. The number approaches 4^{50}. The fact that V region sequences are not random but related would tend to decrease this number; nonetheless, a pool of about 1000 genes would contain all possible base changes at only 5 positions ($4^5 = 1024$).

Another possibility is that matches between hypermutated and germ line DNA might mean that natural selection of germ line V regions over evolutionary time resembles antigen selection of expressed immunoglobulin genes during an immune response. This intriguing possibility could explain replacement mutations, but not silent mutations, and we readily found matches for silent mutations in the NP response among a rather small pool of published V region sequences (Figure 7). Preferential codon usage is most unlikely to account for selection of all these mutations, particularly since the hypermutated V regions we searched were cloned directly from amplified germinal center DNA, in a protocol that did not involve production or selection of hybridomas (JACOB *et al.* 1991).

Templated mutation that involves segmental transfer of information can result in multiple single base changes at neighboring positions; the simultaneous appearance of these mutations reflects their genetic linkage in the donor. We find four examples of matches that are consistent with such segmental mutation. One of these matches (VH16-5, Figure 5) was cloned by hybridization with a probe (OL16) that matched both mutated positions, while the others were encountered while searching for matches to single mutated positions (Vκ5-5, Figure 3 and positions 10 and 73, Figure 7). Could mechanisms other than templated mutation produce apparently linked mutations such as these? One possibility is that neighboring replacement mutations might be coselected in certain circumstances. For example, as discussed in Results, mutations are often found in both codons 34 and 36 of VκOx-1, although only the mutation at codon 34 increases antibody affinity (BEREK and MILSTEIN 1987). The two common mutations at codon 34 replace His (CAC) with Gln (AAC) or Asn (CAG). Both Gln and Asn can undergo spontaneous posttranslational deamination in certain protein microenvironments, so neighboring residues that influence the rate of deamination may therefore be subject to selection (ROBINSON and ROBINSON 1991).

Others have addressed the question of whether gene conversion might contribute to somatic hypermutation of mammalian immunoglobulin genes and found their data incompatible with a mechanism that involves templating. CHIEN *et al.* (1988) reported that they were unable to find matches in germ line DNA for two hypermutated sequences in murine anti-phosphorylcholine antibodies, and concluded from these data that gene conversion plays no role in somatic hypermutation at the immunoglobulin locus. However, these experiments are open to technical criticism because there were no internal controls for hybridization stringency, which can be sensitive to minor variations in temperature or salt concentrations, and only a single probe was used to test for a germ line match to each hypermutated region. WYSOCKI, GEFTER and MARGOLIES (1990) screened genomic libraries for matches to recurrent mutations in CDR2 of the V_H region of anti-*p*-azophenylarsonate antibodies; remarkably, three hybridomas have been identified which carry the same four contiguous replacement mutations in codons 52–59. However, the clones these authors identified did not derive from V region sequences, and these authors concluded that it was unlikely that all four replacement mutations were inserted in one or two templated mutational events. They suggested that intense selection probably accounts for the recurrent appearance of these four mutations, a conclusion supported by the observation that these mutations can appear independently (FISH *et al.* 1989) and significantly increase antibody avidity (SHARON *et al.* 1989; PARHAMI-SEREN *et al.* 1990). REYNAUD *et al.* (1991) compared sequences of expressed light chain V regions from sheep Peyer's patch

cells with germ line light chain V regions. The absence of obvious long conversion tracts led these authors to conclude that gene conversion is not the mechanism of hypermutation in the sheep, but two features of these data suggest this conclusion should be reconsidered. First, a number of matches were in fact evident, particularly in regions of the gene that are not subject to intense selection, but were discounted as being too short to be significant. Second, a substantial fraction of germ line light chain V regions were not represented in the germ line library (REYNAUD *et al.* 1991), and matches could exist among these uncloned sequences.

There is considerable evidence that as an immune response progresses, hypermutation accumulates not only in V regions, but also in J regions and in the J-C intron (LEBECQUE and GEARHART 1990; BOTH *et al.* 1990; STEELE, ROTHENFLUH and BOTH 1992). Although apparently untemplated, these changes could be caused by the same molecular mechanism responsible for templated hypermutation earlier in the immune response. There is evidence to suggest that, while chromosomal gene conversion events are not normally associated with mutation (STACHELEK and LISKAY 1988), a highly error-prone heteroduplex repair process can operate in certain instances in mammalian cells. THOMAS and CAPECCHI (1986) found that introduction of homologous DNA sequences into a cell can result in mutation of the corresponding chromosomal locus which apparently occurs when the heteroduplex formed between the two sequences is incorrectly repaired. Error-prone repair has also been observed to affect extrachromosomal heteroduplex DNAs (AYARES *et al.* 1987; WEISS and WILSON 1988).

Gene conversion may involve formation of a heteroduplex and its subsequent repair. If donor and recipient sequences are homologous but not identical, accurate heteroduplex formation followed by faithful repair will transfer new information from donor to recipient, and all the diversified sequences will have an obvious template in germ line DNA. If donor and recipient sequences are misaligned, or if repair is inaccurate, then both templated and untemplated mutations will appear in the diversified recipient sequence. In fact, there may be circumstances in which gene conversion results *exclusively* in untemplated mutations: If identical sequences (for example, from sister chromatids) undergo heteroduplex formation, incorrect resolution of the heteroduplex can only produce *un*templated mutations. This may be relevant to the observed mutations in the Vλ regions, J regions, and the J-C intron. If an expressed V(D)J region underwent heteroduplex formation not with a nearby V region segment, but with the identical rearranged V(D)J region on a sister chromatid, incorrect resolution of this duplex would produce mutations despite the fact that there are no homologous but divergent donor sequences. Faulty repair of other sorts of recombinational intermediates (gapped DNAs, or duplex molecules carrying segments of information from different donors) could similarly produce untemplated mutation as the sole outcome of a molecular process that was also capable of templated mutagenesis.

The existence of germ line matches to hypermutated sequences is consistent with a templated mutational mechanism, but does not of course constitute proof of mechanism. However, these results suggest that one useful direction of future research will be to create transgenic mice carrying constructs that are designed to test the role of templated mutation. Studies on hypermutation are currently limited by the fact that no existing cell line undergoes efficient hypermutation in culture, and in the longer run definitive analysis of molecular mechanism will require developing such a cell line.

We thank our neighbors in C-wing basement for their good advice and encouragement. This work was supported by National Institutes of Health grant GM41712 to N.M. V.D. was the recipient of a Howard Hughes Medical Institute predoctoral fellowship.

LITERATURE CITED

ALLEN, D., A. CUMANO, R. DILDROP, C. KOCKS, K. RAJEWSKY, N. RAJEWSKY, J. ROES, F. SABLITZKY and M. SIEKEVITZ, 1987 Timing, genetic requirements and functional consequences of somatic hypermutation during B-cell development. Immunol. Rev. **96:** 5–22.

ALZARI, P. M., S. SPINELLI, R. A. MARIUZZA, G. BOULOT, R. J. POLJAK, J. M. JARVIS and C. MILSTEIN, 1990 Three-dimensional structure determination of an anti-2-phenyloxazolone antibody: the role of somatic mutation and heavy/light chain pairing in the maturation of an immune response. EMBO J. **9:** 3807–3814.

AYARES, D., D. GANEA, L. CHEKURI, C. R. CAMPBELL and R. KUCHERLAPATI, 1987 Repair of single-stranded DNA nicks, gaps, and loops in mammalian cells. Mol. Cell. Biol. **7:** 1656–1662.

BALTIMORE, D., 1981 Gene conversion: some implications for immunoglobulin genes. Cell **24:** 592–594.

BECKER, R. S., and K. L. KNIGHT, 1990 Somatic diversification of immunoglobulin heavy chain VDJ genes: evidence for somatic gene conversion in rabbits. Cell **63:** 987–997.

BEREK, C., G. M. GRIFFITHS and C. MILSTEIN, 1985 Molecular events during maturation of the immune response to oxazalone. Nature **316:** 412–418.

BEREK, C., and C. MILSTEION, 1987 Mutation drift and repertoire shift in the maturation of the immune response. Immunol. Rev. **96:** 23–41.

BEREK, C., and C. MILSTEIN, 1988 The dynamic nature of the antibody repertoire. Immunol. Rev. **105:** 5–25.

BERNARD, O., N. HOZUMI and S. TONEGAWA, 1978 Sequences of mouse immunoglobulin light chain genes before and after somatic changes. Cell **15:** 1133–1144.

BOTH, G. W., L. TAYLOR, J. W. POLLARD and E. J. STEELE, 1990 Distribution of mutations around rearranged heavy-chain antibody variable-region genes. Mol. Cell. Biol. **10:** 5187–5196.

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:** 625–637.

BRENNER, S., and C. MILSTEIN, 1966 Origin of antibody variation. Nature **211:** 242–243.

CHIEN, N. C., R. R. POLLOCK, C. DESAYMARD and M. D. SCHARFF, 1988 Point mutations cause the somatic diversification of IgM and IgG2a antiphosphorylcholine antibodies. J. Exp. Med. **167:** 954–973.

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:** 59–66.

CUMANO, A., and K. RAJEWSKY, 1986 Clonal recruitment and somatic mutation in the generation of immunological memory to the hapten NP. EMBO J. **5:** 2459–2468.

DAVID, V., and N. MAIZELS, 1989 Templated, targeted sequence diversification displays a similar choreography in different organisms. Cold Spring Harbor Symp. Quant. Biol. **54:** 227–231.

EDELMAN, G. M., and J. A. GALLY, 1967 Somatic recombination of duplicated genes: an hypothesis on the origin of antibody diversity. Proc. Natl. Acad. Sci. USA **57:** 353–358.

EDELMAN, G. M., and J. A. GALLY, 1970 Arrangement and evolution of eukaryotic genes, pp. 962–972 in *The Neurosciences: Second Study Program*, edited by F. O. SCHMITT. Rockefeller University Press, New York.

EVEN, J., G. M. GRIFFITHS, C. BEREK and C. MILSTEIN, 1985 Light chain germ-line genes and the immune response to 2-phenyloxazolone. EMBO J. **4:** 3439–3445.

FISH, S., E. ZENOWICH, M. FLEMING and T. MANSER, 1989 Molecular analysis of original antigenic sin. I. Clonal selection, somatic mutation, and isotype switching during a memory B cell response. J. Exp. Med. **170:** 1191–1209.

FORSTER, I., H. GU and K. RAJEWSKY, 1988 Germline antibody V regions as determinants of clonal persistence and malignant growth in the B cell compartment. EMBO J. **7:** 3693–3703.

GEARHART, P. J., and D. F. BOGENHAGEN, 1983 Clusters of point mutations are found exclusively around rearranged antibody variable genes. Proc. Natl. Acad. Sci. USA **80:** 3439–3443.

GRIFFITHS, G. M., C. BEREK, M. KAARTINEN and C. MILSTEIN. 1984 Somatic mutation and the maturation of the immune response to 2-phenyloxazolone. Nature **312:** 271–275.

HABER, J. E., D. T. ROGERS and H. J. MCCUSKER, 1980 Homothallic conversions of yeast mating-type genes occur by intrachromosomal recombination. Cell **22:** 277–289.

JACOB, J., G. KELSOE, K. RAJEWSKY and U. WEISS, 1991 Intraclonal generation of antibody mutants in germinal centres. Nature **354:** 389–392.

KAARTINEN, M., G. M. GRIFFITHS, P. H. HAMLYN, A. F. MARKHAM, K. KARJALAINEN, J. L. T. PELKONEN, O. MÄKELÄ and C. MILSTEIN, 1983a Anti-oxazolone hybridomas and the structure of the oxazolone idiotype. J. Immunol. **130:** 937–945.

KAARTINEN, M., G. M. GRIFFITHS, A. F. MARKHAM and C. MILSTEIN, 1983b mRNA sequences define an unusually restricted IgG response to 2-phenyloxazolone and its early diversification. Nature **304:** 320–324.

KEOHAVONG, P., and W. G. THILLY, 1989 Fidelity of DNA polymerases in DNA amplification. Proc. Natl. Acad. Sci USA **86:** 9253–9257.

KLAR, A. J. S., 1987 Differentiated parental DNA strands confer developmental asymmmetry on daughter cells in fission yeast. Nature **326:** 466–470.

KLAR, A. J. S., J. MCINDOO, J. N. STRATHERN and J. B. HICKS, 1980 Evidence for a physical interaction between the transposed and the substituted sequences during mating type gene transposition in yeast. Cell **22:** 291–298.

KNIGHT, K. L., and R. S. BECKER, 1990 Molecular basis of the allelic inheritance of rabbit immunoglobulin VH allotypes: implications for the generation of antibody diversity. Cell **60:** 963–970.

LEBECQUE, S. G., and P. J. GEARHART, 1990 Boundaries of somatic mutation in rearranged immunoglobulin genes: 5'boundary is near the promoter, and 3' boundary is approximately 1 kb from V(D)J gene. J. Exp. Med. **172:** 3652–3665.

LEDERBERG, J., 1959 Genes and antibodies. Science **129:** 1649–1652.

LIVANT, D., C. BLATT and L. HOOD, 1986 One heavy chain variable region gene segment subfamily in the BALB/c mouse contains 500–1000 or more members. Cell **47:** 461–470.

LONGACRE, S., and H. EISEN, 1986 Expression of whole and hybrid genes in *Trypanosoma equiperdum* antigenic variation. EMBO J. **5:** 1057–1063.

MAEDA, N., and O. SMITHIES, 1986 The evolution of multigene families: human haptoglobin genes. Annu. Rev. Genet. **20:** 81–108.

MAIZELS, N., 1987 Diversity achieved by diverse mechanisms: gene conversion in developing B cells of the chicken. Cell **48:** 359–360.

MAIZELS, N., 1989 Might gene conversion be the mechanism of somatic hypermutation of mammalian immunoglobulin genes? Trends Genet. **5:** 4–8.

MAIZELS, N., 1991 Templated, targeted mutation of mammalian immunoglobulin genes, pp. 129–136 in *Somatic Hypermutation in V-Regions*, edited by E. O. STEELE. CRC Press, Boca Raton, Fla.

MAIZELS, N., and A. BOTHWELL, 1985 The T-cell-independent immune response to the hapten NP uses a large repertoire of heavy chain genes. Cell **43:** 715–720.

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:** 3180–3184.

MEYERHANS, A., R. CHEYNIER, J. ALBERT, M. SETH, S. KWOK, J. SNINSKY, L. MORFELDT-MANSON, B. ASJO and S. WAIN-HOBSON, 1989 Temporal fluctuations in Hiv quasispecies in vivo are not reflected by sequential HIV isolations. Cell **58:** 901–910.

ORLANDI, R., D. H. GUSSOW, P. T. JONES and G. WINTER, 1989 Cloning immunoglobulin variable domains for expression by the polymerase chain reaction. Proc. Natl. Acad. Sci. USA **86:** 3833–3837.

PARHAMI-SEREN, B., L. J. WYSOCKI, M. N. MARGOLIES and J. SHARON, 1990 Clustered H chain somatic mutations shared by anti-*p*-azophenylarsonate antibodies confer enhanced affinity and ablate the cross-reactive idiotype. J. Immunol. **145:** 2340–2346.

RADA, C., S. K. GUPTA, E. GHERARDI and C. MILSTEIN, 1991 Mutation and selection during the secondary response to 2-phenyloxazolone. Proc. Natl. Acad. Sci. USA **88:** 5508–5512.

REYNAUD, C. A., V. ANQUEZ, H. GRIMAL and J.-C. WEILL, 1987 A hyperconversion mechanism generates the chicken light chain preimmune repertoire. Cell **48:** 379–388.

REYNAUD, C. A., A. DAHAN, V. ANQUEZ and J.-C. WEILL, 1989 Somatic hyperconversion diversifies the single V_H gene of the chicken with a high incidence in the D region. Cell **59:** 171–183.

REYNAUD, C.-A., C. R. MACKAY, R. G. MULLER and J.-C. WEILL, 1991 Somatic generation of diversity in a mammalian primary lymphoid organ: the sheep ileal Peyer's patches. Cell **64:** 995–1005.

ROBINSON, A. B., and L. R. ROBINSON, 1991 Distribution of glutamine and asparagine residues and their near neighbors in peptides and proteins. Proc. Natl. Acad. Sci. USA **88:** 8880–8884.

ROTH, C. W., S. LONGACRE, A. RAIBAUD, T. BALTZ and H. EISEN,

1986 The use of incomplete genes for the construction of a *Trypanosoma equiperdum* variant surface glycoprotein gene. EMBO J. **5:** 1065–1070.

SABLITZKY, F., G. WILDNER and K. RAJEWSKY, 1985 Somatic mutation and clonal expansion of B cells in an antigen-driven immune response. EMBO J. **4:** 345–350.

SCHLISSEL, M. S., and D. BALTIMORE, 1989 Activation of immunoglobulin kappa gene rearrangement correlates with induction of germline kappa gene transcription. Cell **58:** 1001–1007.

SEIDMAN, J. G., A. LEDER, M. NAU, B. NORMAN and P. LEDER, 1978 Antibody diversity. Science **202:** 11–17.

SHARON, J., M. L. GEFTER, L. H. WYSOCKI and M. N. MARGOLIES, 1989 Recurrent somatic mutations in mouse antibodies to *p*-azophenylarsonate increase affinity for hapten. J. Immunol. **142:** 596–601.

SIEKEVITZ, M., C. KOCKS, K. RAJEWSKY and R. DILDROP, 1987 Analysis of somatic mutation and class switching in naive and memory B cells generating adoptive primary and secondary responses. Cell **48:** 757–770.

SMITHIES, O., 1967a Antibody variability. Science **157:** 267–273.

SMITHIES, O., 1967b The genetic basis of antibody variability. Cold Spring Harbor Symp. Quant. Biol. **32:** 161–168.

STACHELEK, J. L., and R. M. LISKAY, 1988 Accuracy of intrachromosomal gene conversion in mouse cells. Nucleic Acids. Res. **16:** 4069–4076.

STEELE, E. J., H. S. ROTHENFLUH and G. W. BOTH, 1992 Defining the nucleic acid substrate for somatic hypermutation. Immunol. Cell Biol. (in press).

STEFFEN, D., S. BIRD, W. P. ROWE and R. A. WEINBERG, 1979 Identification of DNA fragments carrying ecotropic proviruses of AKR mice. Proc. Natl. Acad. Sci. USA **76:** 4554–4558.

STRATHERN, J. N., 1988 Control and execution of homothallic switching in Saccharomyces cerevisiae, pp. 445–464 in *Genetic Recombination*, edited by R. KUCHERLAPATI and G. R. SMITH. American Society for Microbiology, Washington, D.C.

THIEN, S. L., and R. B. WALLACE, 1986 The use of synthetic oligonucleotides as specific hybridization probes in the diagnosis of genetic disorders, pp. 33–50 in *Human Genetic Disease, a Practical Approach*, edited by K. DAVIES. IRL Press, Oxford, England.

THOMAS, K. R., and M. R. CAPECCHI, 1986 Introduction of homologous DNA sequences into mammalian cells induces mutations in the cognate gene. Nature **324:** 34–38.

THOMPSON, C. B., and P. E. NEIMAN, 1987 Somatic diversification of the chicken immunoglobulin light chain gene is limited to the rearranged variable gene segment. Cell **48:** 369–378.

TINDALL, K. R., and T. A. KUNKEL, 1988 Fidelity of DNA synthesis by the *Thermus aquaticus* DNA polymerase. Biochemistry **27:** 6008–6013.

TONEGAWA, S., 1983 Somatic generation of antibody diversity. Nature **302:** 575–581.

WALLACE, R. B., and C. G. MIYADA, 1987 Oligonucleotide probes for the screening of recombinant DNA libraries, pp. 432–442 in *Guide to Molecular Cloning Techniques*, edited by S. L. BERGER and A. R. KIMMEL. Academic Press, New York.

WEISS, U., and J. H. WILSON, 1988 Heteroduplex-induced mutagenesis in mammalian cells. Nucleic Acids Res. **16:** 2313–2322.

WOOD, W. I., J. GITSCHIER, L. A. LASKY and R. M. LAWN, 1985 Base composition-independent hybridization in tetramethylammonium chloride. A method for oligonucleotide screening of highly complex gene libraries. Proc. Natl. Acad. Sci. USA **82:** 1585–1588.

WYSOCKI, L. J., M. L. GEFTER and M. N. MARGOLIES, 1990 Parallel evolution of antibody variable regions by somatic processes: consecutive shared somatic alterations in V_H genes expressed by independently generated hybridomas apparently acquired by point mutation and selection rather than by gene conversion. J. Exp. Med. **172:** 315–323.

Communicating editor: N. A. JENKINS