EARLY ONSET OF SOMATIC MUTATION IN IMMUNOGLOBULIN VH GENES DURING THE PRIMARY IMMUNE RESPONSE

By NINA S. LEVY, URSULA V. MALPIERO, SERGE G. LEBECQUE, AND PATRICIA J. GEARHART

From the Department of Biochemistry, The Johns Hopkins University School of Hygiene and Public Health, Baltimore, Maryland 21205

Somatic mutation of Ig variable genes occurs by a unique mechanism that creates diversity in antibodies after antigen stimulation. Mutation is characterized by a high frequency of nucleotide substitutions within a localized region surrounding the rearranged variable gene (1, 2). To date, the mechanism of mutation is not well understood. Progress has been hampered by the lack of experimental systems to study mutation, in that most B cell lines do not mutate in vitro, nor is it known when B cells mutate in vivo. Previous work has indicated that mutation appears in antibody variable genes 2 wk after primary immunization of mice (3), but it is not known when the process begins. As a first step towards understanding the mechanism of somatic mutation, we have studied the dynamics of mutation in vivo in order to define a population of B cells undergoing mutation for future studies.

Preliminary experiments demonstrated that several antibody specificities were produced in response to a primary injection of phosphorylcholine (PC)1 conjugated to a protein carrier with Bordetella pertussis adjuvant. RIAs of supernatants from hybridomas made 1–2 wk after immunization showed that 6% were specific for PC, 14% were specific for the carrier, and 30% were specific for the adjuvant. The specificity of the other 50% was unknown but may have included the constant region of self-IgG (4) and other autoantigens, or the hybridomas may have been derived from nonspecifically activated B cells. Nucleotide sequence analysis showed that the PC-specific antibodies were encoded by the VnV1 gene segment of the VnS107 subfamily, and the other antibodies were encoded by various gene segments from the Vn7183, Vn3660, and VnJ558 subfamilies. Gene segments from three of the subfamilies, Vn7183, Vn3660, and VnS107, were further analyzed because many of their germline sequences were published and could be used to identify mutated sequences.

An experimental approach to identify mutation was used that is more rapid and efficient than preparing hybridomas. Splenic mRNA from immunized mice was annealed to γ constant region primers to make cDNA transcripts encoding Vn genes, and the cDNA clones were sequenced. Only transcripts from IgG-producing B cells were studied since mutation is found more frequently in IgG antibodies than in

---

1Abbreviations used in this paper: FWR, framework regions; HVR, hypervariable regions; KLH, keyhole limpet hemocyanin; PC, phosphorylcholine; R, replacement; S, silent.
IgM antibodies (5). The results showed that V\textsubscript{\mu} genes had a low level of mutation by day 5 after immunization and accumulated more mutation by day 7 at a rate of 10\textsuperscript{-3} mutations per nucleotide per generation. However, by day 13, the number of mutations per gene did not increase, and mutations were found clustered in the hypervariable regions. We conclude that mutations were generated mainly during the first week of the immune response and were selected for binding to antigen during the second week.

Materials and Methods

*Immunization.* BALB/cj males 8 wks old (The Jackson Laboratory, Bar Harbor, ME) were immunized with 100 \( \mu \)g i.p. of alum-precipitated PC-keyhole limpet hemocyanin (KLH) (generously provided by J. Kenny, Frederick Cancer Research Facility, Frederick, MD) and 10\textsuperscript{9} heat-killed *Bordetella pertussis* (Calbiochem, La Jolla, CA). Mice were killed 5, 7, and 13 d after immunization. The number of mice used on each day was 12, 14, and 4, respectively.

*Preparation of RNA.* All of the spleens from each time point were pooled to make one cDNA library, with the exception of day 7 when one cDNA library was made from two mice and a second cDNA library was made from 12 mice. Spleens were removed, frozen immediately in liquid \( N_2 \), and crushed into a fine powder while suspended in liquid \( N_2 \). The suspension was transferred to a 50-ml conical tube at 4\textdegree C and then resuspended in 12.5 ml/spleen of 4 M guanidine thiocyanate (Fluka Chemical Corp., Ronkonkoma, NY), 0.5% sarkosyl, 25 mM sodium citrate, pH 7, and 0.1 M \( \beta \)-mercaptoethanol (6). The mixture was homogenized extensively and then vortexed at high speed for 30 s. 25 ml of homogenate was layered onto a 10-ml cushion of 5.7 M cesium chloride, 0.1 M EDTA, pH 7, in a polycyliner SW27 tube and spun at 28,500 rpm for 30 h at 15\textdegree C. The supernatant was discarded and the RNA pellet was solubilized in 180 \( \mu \)l of guanidine thiocyanate without sarkosyl and precipitated in 0.2 M potassium acetate and 2 vol of ethanol. The RNA pellet was resuspended in 1% sarkosyl, 20 mM EDTA, pH 7, extracted with phenol and phenol:chloroform 1:1, and precipitated with ethanol. Poly(A)+ RNA was obtained by chromatography over an oligo-dT-cellulose column (Collaborative Research, Lexington, MA).

*cDNA Library Construction, Screening, and Sequencing.* cDNA was made according to a cDNA synthesis System Plus (code RPN 1256; Amersham Corp., Arlington Heights, IL) using Eco RI linkers. The cDNA was size selected on a 4% polyacrylamide gel, and molecules in the 400-600-bp range were electroeluted and cloned into the Eco RI site of the Agt10 vector. Constant region primers complementary to the 5' end of \( \gamma_1, \gamma_2a, \gamma_2b, \gamma_3, \text{and} \ C_\gamma \) regions were used to initiate synthesis. The \( C_\gamma \) primer was included to increase recovery of cDNA. cDNA libraries were screened separately with the following probes: a 210-bp Eco RI–Pst I fragment encoding the \( V_\gamma 81X \) gene from the \( V_\gamma 7183 \) subfamily (7); a 635-bp Xba I–Eco RI fragment encoding the \( V_\gamma SB32 \) gene from the \( V_\gamma 3660 \) subfamily (8); and a 450-bp Hind III–Bam HI fragment encoding a cDNA copy of \( V_\gamma V1 \) from the \( V_\gamma S107 \) subfamily rearranged to DFL16.1 and J\textsubscript{\lambda}=1. Filters were washed at high stringency (9). Positive plaques were isolated and cDNA inserts were subcloned into M13 vectors for sequence analysis by the dideoxy chain termination method. Approximately 10 cDNA clones from each subfamily were sequenced at each time point.

*Assignment of Mutations.* Mutations were identified by comparing each sequence to germ-line or consensus sequences known for that subfamily. 13 of the 20 gene segments known for the \( V_\gamma 7183 \) subfamily were identified among our cDNA clones (Fig. 1), and seven were not found: \( V_\gamma MOPC21 \) (7), \( V_\gamma 81X \) (7), \( V_\gamma E4.PSI \) (7), \( V_\gamma E4.15 \) (7), \( V_\gamma 3660 \) (15), and \( V_\gamma 50.1 \) (16). For the \( V_\gamma 3660 \) subfamily, four of the five known gene segments were identified among the cDNA clones (Fig. 1), and one was not found: \( V_\gamma 1210 \) (17). \( V_\gamma IB43 \) of the \( V_\gamma 3660 \) subfamily is identical to \( V_\gamma VGam3-2 \) (18) with the exception of a T to C change at position II of amino acid 28. Since we have sequenced this gene twice with the same substitution, \( V_\gamma VGam3-2 \) may contain a mutation at this site. For the \( V_\gamma S107 \) subfamily, one of the four germ line gene segments (12), \( V_\gamma V1 \), was used, and there were not found: \( V_\gamma V3, V_\gamma V11, \text{and} V_\gamma V13 \). D and J\textsubscript{\lambda} sequences were assigned to known germ-line-encoded gene seg-
Figure 1. Germline and consensus DNA sequences of \( V_\alpha \) genes from the \( V_{\alpha}7183, V_{\alpha}3660, \) and \( V_{\alpha}8107 \) subfamilies. Sequences are numbered according to amino acid codons, and a dash indicates identity with the top sequence shown for each subfamily. Sequences were derived from the following sources: pre-B cell, \( V_{\alpha}6.96 \) (reference 7 and this study), and \( V_{\alpha}14.29 \) (10); germline, \( V_{\alpha}283 \) (11), \( V_{\alpha}8107 \) (8), and \( V_{\alpha}11 \) (12); neonatal hybridoma, \( V_{\alpha}76 \) (13); hybridoma consensus, \( V_{\alpha}V2 \) (14), \( V_{\alpha}V3 \) (14), \( V_{\alpha}V4 \) (14), and \( \alpha RF \) (reference 4 and this study); cDNA consensus, \( V_{\alpha}345, V_{\alpha}914, V_{\alpha}5-84, V_{\alpha}1843, V_{\alpha}M315, \) and \( V_{\alpha}735 \) (all from this study), and cDNA, \( V_{\alpha}2,76-84 \) and \( V_{\alpha}7.39 \) (this study).
ments. However, variability at the site of joining of V, D, and J, gene segments could be
due to N regions or mutations. As a general rule, base differences at these junctions were
called mutations only when they were located three or more nucleotides into germline-encoded
gene segments. When determining whether base changes were silent or replacement muta-
tions, two changes occurring in the same codon were counted independently, and silent changes
were considered to have occurred first.

Results

Sequences of V, Genes. Mutations were identified by comparing V, sequences from
sequential time points after immunization to known germline or consensus sequences.
The 18 V, gene segments found in this study included both previously published
and new sequences (Fig. 1). In the V,7183 subfamily, eight gene segments were pre-
viously identified and five are new. The new gene segments have >85% nucleotide
homology to other members of the subfamily. Of the five new gene segments,
V,345, V,914, and V,5-84 were found repeatedly in independent rearrangements,
and V,5-76 and V,7-39 were found only once. The latter two sequences have 17
and 13 differences respectively from the nearest homologous gene segments, V,5V2
and V,57.1. Because the number of differences exceeds the highest reported fre-
cency of mutation in V, gene segments (4%, reference 2), it is likely that these
are new V, gene segments that may contain mutations. Both V,5-76 and V,7-39
sequences are included in Fig. 1, but were not used to calculate mutation frequency.
In the V,3660 subfamily, three gene segments were previously identified and one,
V,733, is new. V,733 has >85% homology to other members of the subfamily, but
is unique in that it has three extra bases coding for amino acid 35B and does not
have an Eco RI site at amino acid 82. In the V,S107 subfamily, the V,S1 gene seg-
ment was exclusively used. An unusual hybrid gene, clone 5-1 of the V,S107 sub-
family (Table I), was isolated that contained sequences from V,RF of the V,7183
subfamily and V,S1 of the V,S107 subfamily; the origin of this clone will be dis-
cussed in a separate paper (Levy, N., and P. Gearhart, manuscript in preparation).
The gene segments encoding 103 cDNA clones from immunized mice are sum-
marized in Table I. Some clones were encoded by certain combinations of gene seg-
ments, which suggests they encode an antibody specific for a single determinant.
For example, the vast majority of clones in the V,S107 subfamily were encoded by
V,S1, DFL16.1, and J,1 gene segments, which is characteristic of antibodies specific
for PC (19), and many clones from the V,3660 subfamily used a combination of
V,733, DQ52, and J,2 gene segments. Other clones used a variety of gene segments
and may be derived from antigen-stimulated or nonspecifically activated B cells.

Time Course of Mutation. Mutations were scored 5, 7, and 13 d after primary im-
munization. The position, number, and type of mutation in each cDNA clone are
shown in Fig. 2 and Table II. All of the mutations were single or double nucleotide
substitutions; no insertions or deletions were observed. V, genes from all three sub-
families showed a similar increase in mutation frequency over time. As seen in Fig.
3, the frequency was lowest on day 5, increased more than twofold on day 7, and
slightly increased on day 13. An exception to this pattern were the clones in the
V,S107 subfamily, which decreased in mutation frequency on day 13.

To ascertain whether mutation occurred sequentially on a gene after cell division,
we determined the number of mutations per gene over time. Many V, sequences
### Table I

**Description of Gene Segments Encoding cDNA Clones**

| VH subfamily | Clone | VH | D | JH | IgG | Clone | VH | D | JH | IgG | Clone | VH | D | JH | IgG |
|--------------|-------|----|---|----|-----|-------|----|---|----|-----|-------|----|---|----|-----|
| 7183         | 76    | 5-76| SP2.2| 4 | 2b | 11 | V3  | FL16.1| 4 | 2a | 16 | 283 | FL16 | 3 | 2a |
|              | 77    | RF  | Q52 | 4 | 2b | 12 | 345 | SP2.9 | 4 | 2a | 17 | 345 | Q52 | 4 | 3  |
|              | 78    | 345 | SP2 | 1 | 2b | 13 | V2  | SP2.3 | 4 | 2  | 18 | V2  | FL16.2| 4 | 1,2* |
|              | 79    | RF  | Q52 | 1 | 2b | 14 | V4  | SP2 | 4 | 2b | 21 | 914 | FL16.1| 2 | 1  |
|              | 80    | RF  | SP2.2| 1 | 2b | 15 | 76  | SP2.2| 4 | 2b | 23 | RF  | FL16.1| 4 | 1  |
|              | 81A   | 283 | FL16.1| 3 | 2b | 36 | 283 | SP2 | 1 | 2b | 25 | 345 | FL16.1| 4 | 1,2* |
|              | 81B   | 14.29| SP2.2| 2 | 2b | 37 | 345 | SP2 | 4 | 3  | 26 | 345 | SP2.2| 4 | 2a |
|              | 82    | 5-84| SP2 | 1 | 2a | 38 | V2  | FL16.2| 4 | 2b | 29 | 345 | SP2 | 4 | 1,2* |
|              | 83    | 345 | SP2.2| 2 | 2a | 39 | 7-39| Q52 | 4 | 2b | 30 | 345 | Q52 | 4 | 2b |
|              | 84    | 5-84| SP2 | 4 | 3  | 44A| 914 | SP2.2| 3 | 1  | 33 | 6.96| SP2  | 4 | 2a |
|              | 85    | 345 | SP2.9| 4 | 3  | 44B| 283 | SP2.2| 1 | 3  | 34 | RF  | SP2.2| 4 | 1  |
|              | F7    | 914 | SP2 | 1 | 2b | 45 | 345 | FL16.1| 2 | 2b | 35 | 345 | SP2 | 4 | 1,2* |
|              | F9    | 345 | FL16.2| 1 | 3  | 13-2| 345 | SP2.6| 1 | 2a | 13-4| 283 | SP2.9| 1 | 2b |
|              |       |     |     |   |    |    |     |     |    |    |    |     |     |    |    |
|              | 3660  | 68  | M315| 11 | SB32| 20  | SB32| 23  | 1B43 |
|              | 70    | 733 | Q52 | 1 | 2b | 12 | M315| 23  | 1B43 |
|              | 75    | SB32| 28  | M315| Q52| 4  | 3 | 24 | 1B43 |
|              | 76    | 733 | Q52 | 2 | 3  | 30 | SB32| 25  | M315 |
|              |       |     |     |   |    |    |    |    |    |    |     |     |    | 32 |
|              |       |     |     |   |    |    | 33 | 733 | Q52 | 2 | 1  |
|              |       |     |     |   |    |    | 35 | 733 | Q52 | 2 | 3  |
|              |       |     |     |   |    |    | 36 | 733 | Q52 | 2 | 3  |
|              |       |     |     |   |    |    | 37 | 733 | Q52 | 2 | 3  |
|              | S107  | 5-1 | RF/V1| FL16.1| 1 | 1 | 7-1 | V1  | FL16.1| 1 | 1 | 13-3 | V1  | FL16.1| 1 | 2b |
|              |      | 5-3 | V1  | FL16.1| 1 | 2a| 7-2 | V1  | FL16.1| 1 | 1 | 13-4 | V1  | FL16.1| 1 | 2,1* |
|              |      | 5-4 | V1  | FL16.1| 3 | 2b| 7-3 | V1  | FL16.1| 1 | 2b | 13-5 | V1  | FL16.1| 1 | 3  |
|              |      | 5-5 | V1  | FL16.1| 1 | 3 | 7-4 | V1  | FL16.1| 1 | 1 | 13-7 | V1  | FL16.1| 1 | 2b |
|              |      | 5-6 | V1  | FL16.1| 1 | 2b| 7-5 | V1  | FL16.1| 1 | 3 | 13-13| V1  | FL16.2| 1 | 1  |
|              |      | 5-7 | V1  | SP2  | 1 | 3 | 7-6 | V1  | FL16.1| 1 | 1 | 13-16| V1  | FL16.1| 1 | 3  |
|              |      | 5-8 | V1  | FL16.1| 1 | 1 | 7-7 | V1  | FL16.1| 1 | 1 | 13-18| V1  | FL16.1| 1 | 2a |
|              |      | 5-9 | V1  | FL16.1| 1 | 3 | 7-9 | V1  | FL16.1| 1 | 3 | 13-22| V1  | FL16.1| 1 | 1  |
|              |      | 5-10| V1  | FL16.1| 1 | 3 | 7-10| V1  | FL16.1| 1 | 3 | 13-26| V1  | FL16.1| 1 | 1  |
|              |      |     |     |   |    |    | 7-11| V1  | SP2.2| 1 | 3 | 13-32| V1  | FL16.1| 1 | 1  |
|              |      |     |     |   |    |    | 7-11A| V1  | FL16.1| 1 | 3 | 13-33| V1  | FL16.1| 1 | 3  |
|              |      |     |     |   |    |    | 7-12| V1  | FL16.1| 1 | 3 | 13-34| V1  | SP2.2| 1 | 3  |
|              |      |     |     |   |    |    | 7-13| V1  | FL16.1| 1 | 3 | 13-35| V1  | FL16.1| 1 | 1  |
|              |      |     |     |   |    |    | 7-14| V1  | FL16.1| 1 | 2a| 13-36| V1  | FL16.1| 1 | 3  |
|              |      |     |     |   |    |    | 7-15| V1  | FL16.1| 1 | 1 | 13-40| V1  | FL16.1| 1 | 1,2* |
|              |      |     |     |   |    |    | 7-17| V1  | FL16.1| 1 | 1 | 13-42| V1  | FL16.1| 1 | 1,2* |
|              |      |     |     |   |    |    | 7-18| V1  | FL16.1| 1 | 3 | 13-42| V1  | FL16.1| 1 | 1,2* |
|              |      |     |     |   |    |    | 7-19| V1  | FL16.1| 1 | 3 | 13-42| V1  | FL16.1| 1 | 1,2* |
|              |      |     |     |   |    |    | 7-20| V1  | FL16.1| 1 | 2 | 13-42| V1  | FL16.1| 1 | 1,2* |

SP2 in the D column signifies the sequence was too short to be uniquely identified. The SP2.9 sequence is reported in reference 13. All of the gene segments were productively joined to encode heavy chain proteins.

* Sequence was too short to assign to either the γ1 or γ2 constant gene.

† M315, SB32, and 1B43 VH gene segments had an Eco RI site at amino acid 82 which prevented recovery of the D and JH gene segments. Clone 3660-28 contained a mutation that removed the Eco RI site.
did not contain any mutations, suggesting that not all B cells had undergone mutation. Indeed, statistical analysis showed that the number of mutations per clone on day 7 did not fit a Poisson distribution expected for mutation occurring randomly in all B cells ($p < 0.01$); rather, only a portion of the B cells appeared to be mutating. To calculate the number of mutations per gene, only those clones with at least one mutation were considered. The number of mutations found on each day was divided by the number of nucleotides sequenced and multiplied by 350 bp, which is the average length of a rearranged $V_\mu$ gene. This analysis showed that the number of mutations per gene increased from 2.0 on day 5 to 3.4 on day 7, and decreased to 3.0 on day 13. A mutation rate of $0.3 \times 10^{-3}$ was calculated between days 5 and 7 based on the increase in unselected, silent mutations and a generation time of 18 h. Assuming the silent changes represented one-third of total mutations, the overall rate of mutation was $\sim 10^{-3}$ mutations/nucleotide/generation, which is consistent with the rate calculated by others (15, 20, 21).
| V_H sub-family | Clone | Amino acid position | Base change in codon | Type | V_H sub-family | Clone | Amino acid position | Base change in codon | Type |
|---------------|-------|---------------------|---------------------|------|---------------|-------|---------------------|---------------------|------|
| 5             | 7183  | 76                  | 100C               | G→A  | 13            | 7183  | 16                 | T→C               | I     |
|               |       |                     |                     | R    |               |       |                     |                     | R    |
| 78            | 27    | T→C                | 59                 | G→A  |               | 50    | T→A                | I     |
| 81A           | 85    | G→A                | 2A                 | A→G  |               | 17    | T→C                | I     |
| 67            | 31    | G→A                | 2A                 | A→G  |               | 52    | T→C                | I     |
| 56            | C→T   | III                | S                  | 76   | A→C          | II    | R                  |
| 57            | C→T   | II                 | R                  | 79   | T→A          | I     | R                  |
|               |       |                     |                     |      |               |       |                     |                     |      |
| 81            | A→T   | II                 | R                  | 102  | T→C          | I     | R                  |
| 3660          | 76    |                     |                     | 9    | T→G          | I     | R                  |
| 7             | 7183  | 14                 | 31                 | G→C  |               | 53    | G→A                | I     |
|               |       |                     |                     | R    |               |       |                     |                     | R    |
| 52A           | A→G   | I                  | R                  | 33   | G→A          | III   | S                  |
| 55            | G→C   | II                 | R                  | 32   | A→T          | II    | R                  |
| 60            | C→A   | II                 | R                  | 55   | G→A          | II    | R                  |
| 75            | A→G   | II                 | R                  | 98   | G→A          | III   | S                  |
| 100           | C→G   | II                 | R                  | 102  | A→T          | I     | R                  |
| 15            | 82A   | A→G                | I                  | 56   | A→T          | I     | R                  |
| 86            | C→T   | III                | S                  | 89   | G→C          | II    | R                  |
| 100F          | G→C   | I                  | R                  | 32   | G→A          | III   | S                  |
| 36            | 2     | G→A                | I                  | 31   | G→A          | III   | S                  |
|               |       |                     |                     | 32   | A→T          | II    | R                  |
|               |       |                     |                     | 55   | G→A          | II    | R                  |
|               |       |                     |                     | 98   | G→A          | III   | S                  |
|               |       |                     |                     | 102  | A→T          | I     | R                  |
| 38            | 34    | G→A                | III                | 35   | T→C          | III   | S                  |
| 3660          | 28    |                     |                     | 82A  | A→G          | III   | S                  |
| 37            | 34    | T→C                | I                  | 33   | G→A          | III   | S                  |
|               |       |                     |                     | 32   | A→T          | II    | R                  |
|               |       |                     |                     | 55   | G→A          | II    | R                  |
|               |       |                     |                     | 98   | G→A          | III   | S                  |
| 44B           | 79    | A→T                | II                 | 112  | C→T          | III   | S                  |
| 3660          | 28    |                     |                     | 82A  | A→G          | III   | S                  |
| 37            | 34    | T→C                | I                  | 59   | T→C          | I     | R                  |
|               |       |                     |                     | 64   | A→G          | II    | R                  |
| 43            | A→G   | II                 | R                  | 34   | G→A          | II    | R                  |
| 74            | T→C   | III                | S                  | 24   | G→A          | II    | R                  |
| 81            | G→C   | III                | R                  | 25   | T→A          | I     | R                  |
| 83            | A→G   | II                 | R                  | 52   | G→A          | I     | R                  |
| 88            | T→C   | III                | S                  | G→A  | II            |
| 100A          | C→A   | I                  | R                  | 56   | G→A          | II    | R                  |
| 102           | C→T   | I                  | R                  | 27   | A→G          | II    | R                  |
| 33            | 21    | A→G                | I                  | 58   | G→A          | II    | R                  |
| 108           | C→A   | II                 | R                  | 93   | C→T          | II    | R                  |
| 109           | C→T   | I                  | R                  | 17   | T→C          | III   | S                  |
| S107          | 7-1   | -3                 | A→C                | R    |               |       |                     |                     |      |
|               |       |                     |                     | 52C² | G→C          | I     | R                  |
|               |       |                     |                     | 106  | G→A          | II    | R                  |
|               |       |                     |                     | 101  | G→A          | I     | R                  |
|               |       |                     |                     | 64³  | A→G          | II    | R                  |
|               |       |                     |                     | 74   | G→A          | III   | S                  |
|               |       |                     |                     | 80   | G→C          | III   | R                  |
|               |       |                     |                     | 5¹   | G→A          | III   | S                  |
|               |       |                     |                     | 28¹  | G→G          | III   | S                  |
|               |       |                     |                     | 7-7  | T→C          | I     | R                  |
|               |       |                     |                     | 7-10 | 82B          | C→T   | II                 |

* Replacement (R) or silent (S) mutation.
† Mutation shared by clones 7-1, 7-2, 7-6, 7-7, and 7-17 of the S107 subfamily.
§ Mutation shared by clones 7-2 and 7-17 of the S107 subfamily.
\* Mutation shared by clones 7-6 and 7-7 of the S107 subfamily.
Distribution of Mutations in the Hypervariable Regions (HVR) and Framework Regions (FWR). To determine if mutations had been selected for binding to antigen, we analyzed the distribution of mutations within the variable region. Since the HVR encode antigen-binding sites, one would expect that mutants selected for higher affinity would have both a higher number of mutations and more replacement (R) than silent (S) mutations in the HVR than expected for a random distribution. Alternatively, in the FWR where structure needs to be conserved to maintain function, one would expect to find a lower number of mutations and fewer R substitutions than expected under random conditions. As shown in Table III, this is indeed the case, in that the proportion of mutations and the ratio of R to S increased in the HVR between days 7 and 13. The results indicate that some selection for high affinity antibodies had occurred by day 7 and was more pronounced by day 13.

Transitions and Transversions. The types of nucleotide substitutions may provide information concerning the mechanism of somatic mutation. Only the silent mutations were analyzed because they are not selected at the protein level. Mutations were recorded from the sense strand of DNA such that a C→T substitution indicates that a C:G pair was replaced by a T:A pair; however, the mutation may have taken place on either strand. Out of 21 silent mutations, 17 were transitions, including eight C→T, five G→A, three T→C, and one A→G. Transversion base changes included two T→G, one C→G, and one C→A.

Clonal Relationship. The presence of V<sub>H</sub> genes with identical mutations suggests that the original precursor may have been a secondary B cell. Five cDNA clones

| Table III | Increased Selection of Mutations with Time |
|-----------|------------------------------------------|
|           | Percent mutations* | R/S ratio<sup>1</sup> |
|           | HVR | FWR | HVR | FWR |
| Random    | 20  | 80  | 3.9 | 3.1 |
| Day 7     | 40  | 60  | 7.5 (15:2) | 1.8 (16:9) |
| Day 13    | 65  | 35  | 12.0 (24:2) | 1.8 (9:5) |

* Random percentages were calculated based on length of region. Experimental data consist of 40 mutations for day 7, and 42 mutations for day 13.

<sup>1</sup> Random ratios were calculated based on the codons used by genes in this study. The number of substitutions found on days 7 and 13 is shown in parentheses.
from the V₆S107 subfamily on day 7 appeared to be derived from the same precursor because they shared the same V₆-D and D-J₆junctions and three mutations in the V₆ and D gene segments. In addition to the shared mutations, each clone had unique mutations. A genealogical tree can be constructed to describe the relationship between these cDNA clones. The initial events included three mutations at amino acids 52C, 100, and 101. Although a cDNA containing these three mutations was not found, it is the logical progenitor of the later members of this tree. After cell division, four new mutations at amino acids -3, 5, 28, and 32 occurred among three daughter cells, giving rise to clones 1, 6, and 17. After further cell division, another three mutations at amino acids 64, 74, and 80 occurred in clones 6 and 17, giving rise to daughter cells 7 and 2, respectively.

Discussion

Mutation Appears Early after Immunization. We have determined the onset of mutation in vivo in order to define a population of cells actively undergoing mutation. To date, studies focusing on the primary immune response have reported a high frequency of mutation in hybridomas made 14 d after immunization, but little or no mutation after 3-9 d (22-24). However, our results of sequences from cDNA clones of splenic mRNA show that the mutational mechanism is activated as early as day 5 after primary immunization and continues through day 7, giving rise to an increase in both the frequency and number of mutations per gene. The high level of mutation on day 7 was observed in two independently generated cDNA libraries and is unlikely to be an artifact. One reason for the discrepancy between our data and previous reports may be that cells undergoing mutation are refractory to cell fusion and would not be detected as hybridomas.

An analysis of the generation of mutation must take into account the contribution of secondary, IgG-producing B cells. Secondary B cells may have accumulated mutation before immunization with PC-KLH, and the data would not represent newly occurring mutation. Alternatively, some or all of the mutations may have occurred after restimulation of secondary B cells. In our data five cDNA clones were found from the V₆S107 subfamily on day 7 that appear to be derived from one secondary B cell precursor. All of the cDNA clones shared at least three mutations, yet each clone had a unique overall pattern of mutation. One possibility is that the shared mutations occurred before injection of PC-KLH, perhaps by contact with PC-containing bacteria in the environment, and the unique mutations occurred after deliberate immunization. Reports differ as to whether secondary B cells can undergo mutation after antigen stimulation. Siekevitz et al. (25) found very little mutation occurring in secondary B cells stimulated during an adoptive response. In contrast, Berek and Milstein compared sequences obtained at primary, secondary, and tertiary stages of the immune response and showed that the number of mutations per gene increased from 2.4 to 5.2 to 7.7 (calculated from reference 20), respectively, suggesting that mutation also occurs in memory B cells. The problem of determining whether the mutation seen in this study was due to newly occurring mutation versus mutation that took place before immunization was addressed by looking at the time course and pattern of mutation.

The increase in frequency and number of mutations per gene with time strongly suggests that most of the mutations occurred de novo after immunization. Between
days 5 and 7, the frequency of mutations increased greater than twofold at a rate of $10^{-3}$ mutations per nucleotide per generation. Assuming a constant rate of mutation and cell division for the first 7 d, the results suggest that mutation started on day 3. Thus, the mutational process appears to be most active during the early stages of B cell activation when resting B cells are stimulated to divide and differentiate into plasma cells and/or secondary cells. This developmentally regulated process appears to coincide with B cell switching from IgM to IgG, and most likely requires T cell factors to initiate mutation (26).

**Mutation May Occur over a Brief Period.** It would appear, however, that mutation does not occur at a significant rate between days 7 and 13. Although the frequency of mutation in genes from the V_{7183} and V_{3660} subfamilies increased from days 7 to 13, the distribution of these later mutations suggested that the increased frequency was due to selection of existing mutations rather than to generation of new mutations. The number of mutations increased in the HVR with time and correspondingly decreased in the FWR, indicating a greater degree of selection on day 13 than 7. Similarly, the R:S ratio of mutations in the HVR increased from 7.5 to 12 on day 13, whereas the ratio of changes in the FWR was 1.8 at both time points. These results are consistent with the hypothesis that early mutations are randomly inserted into the gene, and subsequently selected for both accumulation and replacement amino acids in the HVR. Furthermore, the average number of mutations per gene did not increase beyond 3.4 on day 13. This result is probably not due to the inability of the variable gene to tolerate more than three mutations since about one-third of the mutated clones in this study had five or more mutations. Rather, it suggests that most of the mutation occurred during the first week of the primary immune response.

Previous reports support the idea that mutation is not a continuously active process during clonal expansion (27, 28). In those studies, several hybridomas were independently isolated from a single mouse and shown to contain the same pattern of mutation, implying that a mutated B cell can divide without further mutation. We propose that mutation occurs during a short time period, about four to five cell divisions, soon after antigen stimulation, and mutations are inserted randomly throughout the gene. The mutational machinery is then shut off. B cells expressing mutated antibodies are selected by antigen, and the initial random pattern shifts to a clustered distribution of mutations and selection of replacement mutations in the HVR. Further rounds of mutation may occur at a later time, for example, in memory B cells restimulated after a secondary immunization.

**Mutation May Be DNA Strand Directed.** To gain insight into the mechanism of mutation, the types of nucleotide changes producing silent substitutions were analyzed. Of 21 silent mutations found, 8 were C$\rightarrow$T transitions on the sense strand. An excess of C$\rightarrow$T mutations can result from the methylation of a cytosine residue followed by spontaneous deamination to yield thymidine (29). The same event occurring on the opposite strand would lead to a G$\rightarrow$A mutation on the sense strand. However, we observed more C$\rightarrow$T mutations than G$\rightarrow$A changes, suggesting that there is a strand preference for the incorporation of mutations. Although the sample size is small, it is consistent with a similar study we performed on 172 silent and flanking mutations in antibody variable genes (30), where twice as many C$\rightarrow$T mutations as G$\rightarrow$A substitutions were found ($p < 0.05$). In contrast, among pseudogenes that do not encode immunoglobulins, there is an equal frequency of C$\rightarrow$T and G$\rightarrow$A
substitutions produced during meiosis (31). One explanation for this asymmetrical phenomenon in variable genes is that mutations are preferentially incorporated into one of the DNA strands. An outcome of such a mechanism is that, after subsequent division, only one of the daughter cells is mutated while the other is not, preventing loss of germline information.

Unequal Frequencies of Mutation Observed Among cDNA Clones. By day 13, the frequency of mutation in the VHV1 gene from the VHV107 subfamily was only 0.04% compared with 0.65% mutation in genes from the VHV3660 and VHV7183 subfamilies. The low level of mutation in VHV1 is not due to an intrinsic property of the nucleotide sequence or protein structure that prevents mutation, since the gene has been shown to undergo as much as 4% mutation and still produce a functional antibody (2). Rather, the low frequency may be due to negative selection of mutations that cause decreased affinity for PC. The affinity of the germline-encoded antibody for PC appears to be better than most mutated sequences since approximately half of secondary antibodies have no mutation (32-34). Another possibility is that PC-specific B cells are infrequently stimulated to mutate, perhaps by unknown interactions between B and T cells. Such disparate frequencies of mutation among IgG antibodies illustrates the need to look at more than one specificity before making conclusions about the general frequency of mutation. As noted earlier, PC-specific B cells comprised only a minor proportion of the immune response to PC-KLH. B cells secreting antibodies encoded by genes in the VHV7183 and VHV3660 subfamilies encompassed a larger fraction of the immune response and most likely reflect the general level of mutation. Thus, splenic B cells taken during the first week after immunization may be a good source of cells for elucidating the mechanism of mutation at the biochemical level.

Summary

The dynamics of somatic mutation in Ig variable genes was investigated in order to define a population of B cells undergoing mutation. BALB/cJ mice were injected with PC-KLH, and splenic RNA was prepared 5, 7, and 13 d later. The mRNA was annealed to γ constant region primers to make cDNA transcripts encoding VH genes. 103 cDNA clones corresponding to 18 different genes from the VHV107, VHV3660, and VHV107 subfamilies were sequenced to identify mutation. VH genes had a low level of mutation on day 5 after immunization and accumulated more mutation by day 7 at a rate of $10^{-3}$ mutations per nucleotide per generation. However, by day 13, the number of mutations per gene did not increase, and most of the substitutions encoded replacement amino acid changes that were clustered in the hypervariable regions, indicating that the mutational process was less active during the second week and that antigen selection had occurred. The data are consistent with a developmentally regulated mechanism in which mutation is activated during the first week of the primary immune response for a limited time period, after which selection acts to preserve the beneficial mutants.

We thank Ann M. Lawler for critical reading of the manuscript and D'Ann M. Caron for technical assistance.

Received for publication 30 January 1989.
ONSET OF SOMATIC MUTATION IN IMMUNOGLOBULIN V\(_n\) GENES

References

1. 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.

2. Kim, S., M. Davis, E. Sinn, P. Patten, and L. Hood. 1981. Antibody diversity: somatic hypermutation of rearranged V\(_n\) genes. *Cell.* 27:573.

3. Griffiths, G. M., C. Berek, M. Kaartinen, and C. Milstein. 1984. Somatic mutation and the maturation of immune response to 2-phenyl oxazolone. *Nature (Lond.)* 312:271.

4. Shlomchik, M., D. Nemazee, J. Van Snick, and M. Weigert. 1987. Variable region sequences of murine IgM anti-IgG monoclonal autoantibodies (rheumatoid factors). II. Comparison of hybridomas derived by lipopolysaccharide stimulation and secondary protein immunization. *J. Exp. Med.* 165:970.

5. 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.

6. Chirgwin, J. M., A. E. Przybyla, R. MacDonald, and W. J. Rutter. 1979. Isolation of biologically active ribonuclease acid from sources enriched in ribonuclease. *Biochemistry.* 18:5294.

7. Yancopoulos, G. D., S. V. Desiderio, M. Paskind, J. F. Kearney, D. Baltimore, and F. W. Alt. 1984. Preferential utilization of the most J\(_{\alpha}\)-proximal V\(_n\) gene segments in pre-B cell lines. *Nature (Lond.)* 311:727.

8. Dzierzak, E. A., C. A. Janeway Jr., N. Richard, and A. Bothwell. 1986. Molecular characterization of antibodies bearing Id-460. I. The structure of two highly homologous V\(_n\) genes used to produce idioype positive immunoglobulins. *J. Immunol.* 136:1864.

9. Brodeur, P. H., and R. Riblet. 1984. The immunoglobulin heavy chain variable region (IgH-V) locus in the mouse. I. One hundred IgH-V genes comprise seven families of homologous genes. *Eur. J. Immunol.* 14:922.

10. Lawler, A. M., P. S. Lin, and P. J. Gearhart. 1987. Adult B-cell repertoire is biased toward two heavy-chain variable-region genes that rearrange frequently in fetal pre-B cells. *Proc. Natl. Acad. Sci. USA.* 84:2454.

11. Ollo, R., J.-L. Sikorav, and F. Rougen. 1983. Structural relationships among mouse and human immunoglobulin V\(_n\) genes in the subgroup III. *Nucleic Acids Res.* 11:7887.

12. Crews, S., J. Griffin, H. Huang, K. Calame, and L. Hood. 1981. A single V\(_n\) gene segment encodes the immune response to phosphorylcholine: somatic mutation is correlated with the class of antibody. *Cell.* 25:59.

13. Riley, S. C., S. J. Connors, N. R. Klinman, and R. T. Ogata. 1986. Preferential expression of variable region heavy chain gene segments by predominant 2,4-dinitrophenyl-specific BALB/c neonatal antibody clonotypes. *Proc. Natl. Acad. Sci. USA.* 83:2589.

14. Caton, A. J., G. G. Brownlee, L. M. Staude, and W. Gerhard. 1986. Structural and functional implications of a restricted antibody response to a defined antigenic region on the influenza virus hemagglutinin. *EMBO (Eur. Mol. Biol. Organ.)* J. 5:1577.

15. Clarke, S. H., K. Huppi, D. Ruzinsky, L. Staude, W. Gerhard, and M. Weigert. 1985. Inter- and intronal diversity in the antibody response to influenza hemagglutinin. *J. Exp. Med.* 161:687.

16. Hartman, A. B., and S. Rudikoff. V\(_n\) genes encoding the immune response to β-(1,6)-galactan: somatic mutation in IgM molecules. *EMBO (Eur. Mol. Biol. Organ.)* J. 3:3023.

17. Near, R. I., E. C. Juszack, S. Y. Huang, S. A. Sicari, M. N. Margolis, and M. L. Gelfer. 1984. Expression and rearrangement of homologous immunoglobulin V\(_n\) genes in two mouse strains. *Proc. Natl. Acad. Sci. USA.* 81:2167.

18. Winter, E., A. Radbruch, and U. Krawinkel. 1985. Members of novel V\(_n\) gene families are found in VDJ regions of polyclonally activated B-lymphocytes. *EMBO (Eur. Mol. Biol. Organ.)* J. 4:2861.

19. Perlmutter, R. M., S. T. Crews, R. Douglas, G. Sorenson, N. Johnson, N. Nivera, P. J.
Gearhart, and L. Hood. 1984. The generation of diversity in phosphorylcholine-binding antibodies. Adv. Immunol. 35:1.
20. Berek, C., and C. Milstein. 1987. Mutation drift and repertoire shift in the maturation of the immune response. Immunol. Rev. 96:23.
21. Sablitzky, F., G. Wildner, and K. Rajewsky. 1985. Somatic mutation and clonal expansion of B cells in an antigen-driven immune response. EMBO (Eur. Mol. Biol. Organ.) J. 4:345.
22. Kaartinen, M., G. M. Griffiths, A. F. Markham, and C. Milstein. 1983. mRNA sequences define an unusually restricted IgG response to 2-phenyloxazolone and its early diversification. Nature (Lond.). 304:320.
23. Cumano, A., and K. Rajewsky. 1985. Structure of primary anti-(4-hydroxy-3-nitrophenyl) acetyl (NP) antibodies in normal and idiotypically suppressed C57BL/6 mice. Eur. J. Immunol. 15:512.
24. Wysocki, L., T. Manser, and M. L. Gefter. 1986. Somatic evolution of variable region structures during an immune response. Proc. Natl. Acad. Sci. USA. 83:1847.
25. 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.
26. Maizels, N., and A. Bothwell. 1983. The T-cell-independent immune response to the hapten NP uses a large repertoire of heavy chain genes. Cell. 43:715.
27. Claflin, J. L., J. Berry, D. Flaherty, and W. Dunnick. 1987. Somatic evolution of diversity among anti-phosphocholine antibodies induced with Proteus morganii. J. Immunol. 138:3060.
28. Shlomchik, M. J., A. Marshak-Rothstein, C. B. Wolfovicz, T. L. Rothstein, and M. G. Weigert. 1987. The role of clonal selection and somatic mutation in autoimmunity. Nature (Lond.). 328:805.
29. Couloudre, C., J. H. Miller, P. J. Farabaugh, and W. Gilbert. 1978. Molecular basis of base substitution hotspots in Escherichia coli. Nature (Lond.). 274:775.
30. Golding, G. B., P. J. Gearhart, and B. W. Glickman. 1987. Patterns of somatic mutations in immunoglobulin variable genes. Genetics. 115:169.
31. Li, W. H., C. I. Wu, and C. C. Luo. 1984. Nonrandomness of point mutation as reflected in nucleotide substitutions in pseudogenes and its evolutionary implications. J. Mol. Evol. 21:58.
32. Gearhart, P. J., A. M. Lawler, N. S. Levy, S. G. Lebeque, and U. V. Malipiero. 1988. Generation of antibody diversity before and after immunization. Ann. NY Acad. Sci. 546:57.
33. Claflin, J. L., and J. Berry. 1988. Genetics of the phosphocholine-specific antibody response to Streptococcus pneumoniae. Germ-line but not mutated T15 antibodies are dominantly selected. J. Immunol. 141:4012.
34. Feeney, A. J., S. H. Clarke, and D. E. Mosier. 1988. Specific H chain junctional diversity may be required for non-T15 antibodies to bind phosphorylcholine. J. Immunol. 141:1267.