Chicken Ig<sub>L</sub> variable region gene conversions display pseudogene donor preference and 5' to 3' polarity

Wayne T. McCormack and Craig B. Thompson

Howard Hughes Medical Institute, Departments of Internal Medicine and Microbiology and Immunology, University of Michigan Medical Center, Ann Arbor, Michigan 48109-0650 USA

Chicken immunoglobulin variable region diversity is generated during B-cell development in the bursa of Fabricius by intrachromosomal gene conversion, resulting in the replacement of sequence blocks within the unique rearranged V<sub>L</sub> and V<sub>HI</sub> genes with homologous sequences derived from V region pseudogene segments (ψV<sub>L</sub>). In this report, the nucleotide sequences of 217 gene conversion events in 52 random Ig<sub>L</sub> clones were analyzed to characterize the molecular mechanism of gene conversion. The frequency of ψV<sub>L</sub> usage as gene conversion donors is shown to depend on the proximity of the ψV<sub>L</sub> segment to V<sub>LR</sub>, extent of homology with V<sub>LR</sub>, and relative orientation of the ψV<sub>L</sub> segments. Gene conversion events are not observed in the 5' region of homology between ψV<sub>L</sub> segments and V<sub>L</sub>, but are distributed throughout the remainder of the V<sub>L</sub> exon. The 5' ends of individual gene conversion events always begin in regions of homology between the donor ψVL and recipient VL<sub>R</sub> gene, whereas the 3' ends can occur in regions of nonhomology and often have nucleotide insertions or deletions. These results suggest a 5' to 3' polarity in the gene conversion mechanism. The implications of our data are discussed in relation to current molecular models of gene conversion.

[Key Words: Bursa of Fabricius; chicken; diversity; gene conversion; immunoglobulin genes; pseudogene]

Received December 8, 1989; revised version accepted February 5, 1990.

The chicken has a novel mechanism for the generation of an antibody repertoire, which involves genetically programmed sequence modification of VL and VH gene segments during B-cell development in the bursa of Fabricius [for review, see Weill and Reynaud 1987; McCormack and Thompson, in press]. The chicken immunoglobulin (Ig) loci encode single functional V<sub>HI</sub>, I<sub>L</sub>, V<sub>LR</sub>, and J<sub>L</sub> gene segments [Reynaud et al. 1985, 1989], which recombine in B cells to generate the heavy- and light-chain genes required for Ig production. Despite the lack of combinatorial diversity, which creates the primary repertoire in mammals [for review, see Alt et al. 1986; Hunkapiller and Hood 1989], chickens are able to produce Ig molecules bearing structurally diverse heavy and light chains and to mount functional immune responses [Benedict and Yamaga 1976; Jalkanen et al. 1984]. Chicken V(D)J joining introduces only limited junctional diversity and occurs not as an ongoing process, but as a brief programmed event in many embryonic tissues during early embryonic development [Weill et al. 1986; McCormack et al. 1989a, b]. Most chicken Ig gene diversity is generated after V(D)J joining in the bursa of Fabricius, beginning between day 15 and day 18 of embryonic development [Reynaud et al. 1987; Thompson and Neiman 1987]. The somatic diversification of the chicken Ig<sub>L</sub> variable region is targeted to the rearranged V<sub>L</sub> gene segment only, as demonstrated by the successive loss of restriction sites only within rearranged V<sub>L</sub> segments during embryonic development, and is linked to transcriptional activation of the rearranged Ig<sub>L</sub> allele and the appearance of unique DNase I hypersensitive sites in the rearranged, but not the unarranged, Ig<sub>L</sub> allele [Thompson and Neiman 1987].

The genomic organizations of the chicken Ig<sub>L</sub> and Ig<sub>HI</sub> loci have been determined by Reynaud et al. [1987, 1989]. The Ig<sub>L</sub> locus consists of 25 V<sub>L</sub> pseudogene segments (ψV<sub>L</sub>) within a 19-kb region located 5' of the single functional V<sub>L</sub> and J<sub>L</sub> gene segments. Sequence comparisons of diversified V<sub>L</sub> cDNA sequences with ψV<sub>L</sub> segments revealed that blocks of nucleotide substitutions found in the cDNA sequences were also present in the ψV<sub>L</sub> segments [Reynaud et al. 1985, 1987]. Similarly, the chicken Ig<sub>HI</sub> locus encodes a family of ψV<sub>HI</sub> segments spanning 60–80 kb, which is located 5' of a single functional V<sub>HI</sub> gene, approximately 15 D<sub>HI</sub> segments, and one I<sub>HI</sub> gene segment [Reynaud et al. 1989]. Blocks of nucleotide substitutions in diversified V<sub>HI</sub> cDNA sequences are shared with the ψV<sub>HI</sub> segments and, surprisingly, extend into the D<sub>HI</sub> region. Although the mechanism of this somatic diversification has not been formally addressed, the sequence data of Reynaud et al. [1987, 1989] are consistent with the possibility that it may be gene conversion.

Gene conversion was described originally to explain the unequal segregation of alleles during meiosis in lower eukaryotes [for review, see Radding 1978;
Hastings 1988). The "conversion" of one allele by another occurs by a nonreciprocal DNA recombination event in which a copy of donor sequence is transferred into the recipient gene, leaving the donor gene unchanged. This unidirectionality of the recombination event distinguishes gene conversion from other modes of sequence exchange, such as double homologous recombination. Gene conversion events have been studied in mammalian cells using selectable markers, and some of the molecular requirements for efficient intrachromosomal gene conversion to occur in these systems include an overall nucleotide sequence match of >80% (Waldman and Liskay 1986) and at least 200–300 bp of homology (Liskay et al. 1987). As little as 14–25 bp of identical sequence may be sufficient for extrachromosomal gene conversion events to occur (Rubnitz and Subramani 1984; Ayares et al. 1986).

As described in the accompanying report (Carlson et al. 1990), chicken IgL gene diversification occurs by intrachromosomal gene conversion. Detailed restriction mapping of donor ψVL segments and recipient VL1 genes from a panel of clonal v-rel-transformed chicken B-cell lines, and direct sequencing of ψV1 donors segments used in defined gene conversion events, demonstrated that no modification of ψV1 segments occurs as a result of gene conversion. Nucleotide sequence polymorphisms between the alleles of the VL1 segments and ψV1 family of the SC strain of chicken (an F1 cross between two inbred strains, G4 and S3) allowed the assignment of donor and recipient gene segments in selected gene conversion events to the same allele, indicating that gene conversion uses donor sequences derived from the same chromosome.

In this report, we analyze the nucleotide sequences of 217 gene conversion events from 52 random rearranged IgL clones to elucidate the molecular mechanism of somatic gene conversion in chicken IgL genes. The frequency of ψV1 segment usage as a gene conversion donor appears to be determined by its proximity to VL1, its extent of homology with VL1, and its orientation relative to VL1. Polarity in the gene conversion mechanism is suggested by the distribution of gene conversions within VL1 and by the degree of homology and precision of the 5' and 3' ends of 139 gene conversion events with identifiable ends. Gene conversions are not observed within the 5' region of homology between ψVL and VL1, but are found throughout the remaining VL1 exon. The 5' ends of gene conversion events are characterized by regions of high sequence homology between VL1 and the donor ψVL segments, whereas the 3' ends display less homology and often have nucleotide insertions or deletions. These data suggest several novel features of the molecular mechanism of IgL gene conversion, and are discussed in relation to current gene conversion models.

Results

Nucleotide sequence analysis of gene conversion events

Gene conversion events have been analyzed using the nucleotide sequences of 42 random clones of the rearranged IgL locus isolated from the bursae of SC strain chickens at day 18 of development and ten clones from v-rel-transformed clonal B-cell lines derived from bursal tumors isolated at the day of hatching. A summary description of all 52 IgL clones sequenced is shown in Table 1. An equal number of clones was obtained from each parental allele. Six of the clones were out-of-frame for translation, four at the VL1/VL2 junction and two as a result of internal sequence modification. This is consistent with previously reported percentages of out-of-frame V–J joints observed during this developmental period (McCormack et al. 1989a,b). In addition, the recovery of out-of-frame joints suggests that these sequences reflect the results of the gene conversion mechanism, rather than the results of selection for light-chain expression.

By use of the published nucleotide sequences for the ψVL segments of the CB chicken strain (Reynaud et al. 1987), the nongerm-line base pairs of each clone were assigned to ψVL donors where possible. The total number of nucleotide substitutions found in the set of 52 clones was 952. Of these, only 16 (1.7%) could not be accounted for by gene conversion events from the published ψVL sequences of the CB strain. Fourteen of these untemplated substitutions were transitions relative to the germ line VL1 sequence, and two were transversions. Some of these may have been due to unidentified ψVL nucleotide sequence polymorphisms unique to the SC strain parental alleles. Alternatively, some of the substitutions not accounted for by gene conversion events may be artifacts that result from misincorporation by Taq polymerase (Saiki et al. 1988) or may result from an error-prone feature of the gene conversion mechanism.

A complete sequence analysis of two representative clones is shown in Figure 1. The sequence of clone #5 (Fig. 1A) reveals two nonoverlapping gene conversion events. The first gene conversion event, involving ψVL18, extends at least from the G substitution in codon 6 through the 15-bp insertion between codons 25 and 26. Regions of nucleotide sequence identity at each end of the gene conversion event prevent precise assignment of the ends of the recombination event, and are shown in symbolic form as shaded blocks (Fig. 1C). The second gene conversion event used ψVL7 as the sequence donor. Whereas the 5' end of this event has a region of 32-bp

| Table 1. Summary of the 52 rearranged IgL clones isolated from the bursa of Fabricius of the SC chicken |
|-------------------------------------------------|
| IgL allele representation | 26 G4 clones | 26 S3 clones |
| Number of in-frame clones | 46 |
| Number of out-of-frame clones | 6 (2 by gene conversion) |
| Total number of nucleotide substitutions | 952 |
| Number of substitutions with potential ψVL donor | 936 (98.3%) |
| Total number of gene conversion events identified | 217 |
| Gene conversions assigned to specific ψVL donors | 180 (83%) |
McCormack and Thompson

identity between \( \psi_{V_L} \) and \( V_{L1} \), the 3' end of the gene conversion event is precisely defined between the first and second nucleotides of codon 91. The sequence of clone #64 (Fig. 1B) reveals five gene conversion events, including one with multiple (>3) potential donors, and one with two potential donors \( \psi_{V_L6} \) and \( \psi_{V_L11} \). The overlap of gene conversion events involving \( \psi_{V_L4} \) and \( \psi_{V_L5} \) prevents an accurate description of the 3' and 5' ends, respectively. Such ends are illustrated by brackets in the symbolic representation (Fig. 1C).

Gene conversion occurs in cis

The total number of gene conversion events identified was 217, of which 179 (82%) could be assigned to 1–3 specific donor \( \psi_{V_L} \) segments. Of these, 141 could be assigned to a single potential donor, 20 to 2 possible donors, and 18 to 3 possible donors. The remaining 38 gene conversion events had more than three potential donors. In order to determine the \( I_{\theta L} \) allele from which donor \( \psi_{V_L} \) sequence was transferred during gene conversion, partial nucleotide sequences of 13 \( \psi_{V_L} \) segments were obtained for the G4 and S3 parental strains of the SC chicken after polymerase chain reaction (PCR) amplification and cloning of the \( \psi_{V_L} \) segments from erythrocyte DNA. All of the sequenced \( \psi_{V_L} \) segments were found to be polymorphic (\( \psi_{V_L2}, 4, 5, 7, 8, 10, 11, 12, 14, 17, 18, 20, \) and 24; sequence data not shown), with 1–10 nucleotides of sequence polymorphism each. On the basis of these \( \psi_{V_L} \) donor sequence polymorphisms, and sequence polymorphisms in the intron and \( V_{L1} \) exon of the recipient \( V_{L1} \) gene, the donor and recipient gene segments were assigned to one allele or the other to determine whether sequence transfer occurs within a chromosome (cis) or between homologous chromosomes (trans). Gene conversion events that were long enough to encompass multiple \( \psi_{V_L} \) polymorphisms, a \( \psi_{V_L} \) polymorphism flanked on both sides by unique \( \psi_{V_L} \) sequence markers, or both were scored. On the basis of these criteria, 24 gene conversion events were found to be in cis, involving sequence donors \( \psi_{V_L2}, 4, 5, 7, 10, 12, \) and 24. No events were found to be in trans.

Locations of gene conversion events within \( V_{L1} \)

Symbolic representations of the 139 gene conversion events for which one or both ends of the recombination event could be identified are shown in Figure 2. The smallest gene conversion event observed in this data set is no longer than 8 bp in length, and the largest event is at least 249 bp in length. The average minimum size of the gene conversion events is 26.7 ± 36.3 bp (mean ± sd). On the basis of the positions of minimum gene conversion lengths illustrated in Figure 2, the cumulative number of gene conversion events involving each base pair within the region of \( \psi_{V_L} \) homology with \( V_{L1} \) was determined (Fig. 3). Gene conversions were not observed at the 5' region of the \( V_{L1} \) gene segment, even though donor sequences homologous to the leader portion of the \( V_{L1} \) exon and an average of 25 bp of homolo-

gous leader intron sequence were present in 17 of the \( \psi_{V_L} \) donor segments (Reynaud et al. 1987). Numerous polymorphisms exist in the leader intron sequences of the \( \psi_{V_L} \) donor segments, which would allow detection of gene conversions in this region. Gene conversions are observed throughout the remainder of the \( V_{L1} \) exon, including every nucleotide position of codon 1 through the V-J junction, and are observed more frequently in CDR1, at the FR2/CDR2 boundary, and in CDR3 (Fig. 3). The actual 5' recombination sites for the gene conversion events beginning at nucleotide position 20 may be distributed along a positional “gradient,” rather than beginning precisely at that position (Fig. 3), because of the blocks of sequence identity shared between \( V_{L1} \) and donor \( \psi_{V_L} \) segments that extend 5' for various distances before an informative polymorphism is reached (Fig. 2).

Frequency of \( \psi_{V_L} \) usage as a sequence donor depends on three variables

The frequency of \( \psi_{V_L} \) segment usage as sequence donors was determined by counting the number of times each \( \psi_{V_L} \) segment was used in gene conversion events with only one or two potential donors. Events in which two \( \psi_{V_L} \) segments were potential donors were counted as half an event for each \( \psi_{V_L} \) segment. The bar graph shown in Figure 4 illustrates the \( \psi_{V_L} \) segment usage based on this analysis. Reynaud et al. (1987) have noted previously that the \( \psi_{V_L} \) segments located most proximal to the rearranged \( V_{L1} \)–\( I_{\lambda} \) locus are used more frequently than the more distal \( \psi_{V_L} \) segments and this pattern is evident in Figure 4. Superimposed on this pattern, however, is a striking increase in the usage of \( \psi_{V_L} \) segments 24, 20, 18, 14, 10, 7, and 5. These \( \psi_{V_L} \) segments are in the inverted or antisense orientation with respect to the \( V_{L1} \)–\( I_{\lambda} \) transcriptional unit. As shown in Table 2, the preferential usage of inverted \( \psi_{V_L} \) segments is statistically significant, and is independent of their location and overall homology with \( V_{L1} \).

Another variable that appears to influence \( \psi_{V_L} \) usage as donors is the extent of homology with \( V_{L1} \), which is determined by both the length of the \( \psi_{V_L} \) segment and the percent nucleotide sequence identity with \( V_{L1} \). The frequency of usage for truncated \( \psi_{V_L} \) segments is lower than for full-length \( \psi_{V_L} \) segments (Fig. 4). Several \( \psi_{V_L} \) segments are truncated at the 5' end (\( \psi_{V_L3}, 6, \) and 9), the 3' end (\( \psi_{V_L16} \) and 21), or at both ends (\( \psi_{V_L15}, 22, \) and 25) (Reynaud et al. 1987).

Sequence identity at the ends of gene conversion events

The exact borders of gene conversion events cannot be determined precisely in most cases, because of the 75–92% sequence identity between the \( \psi_{V_L} \) segments and \( V_{L1} \). Most gene conversion events are flanked on the 5' and 3' sides by blocks of nucleotide sequence of variable length at which the \( V_{L1} \) gene segment and the \( \psi_{V_L} \) donor segment are identical (Figs. 1 and 2). The distribution of the lengths of identical sequence bordering the 5' and 3' ends of the gene conversion events summarized
Figure 1. Nucleotide sequences of representative gene conversion events. The nucleotide sequences of day-18 bursa clones #5 (A) and #64 (B), and \( \psi V \) donor segments are compared to the germ line sequence of the V\(_4\) exon and part of IL for the G4 allele \( \text{top line} \). Identity to the germ line sequence is indicated by dashes. Codons are numbered according to Reynaud et al. (1987), and complementarity-determining regions \( \text{CDR} \) are overlined. \( \text{C} \) Symbolic representations of gene conversion events. Nucleotide positions are numbered from the 5' end of the V\(_L\) exon, and subregions \( \{ \text{LI leader; FR) framework regions; and CDR}\) are indicated at the top of the figure. The nucleotide boundaries of the V\(_4\) exon in rearranged genes (position 1 and the V\(_L\)--JL junction at position 294) are marked by solid vertical lines. For most \( \psi V \) donor segments, homology with V\(_4\) extends from position \(-25\) in the leader intron to position \(+294\) at the 3' end of the V\(_4\) exon. The minimum gene conversion length \( \text{line} \) is flanked by blocks of identical sequence shared by the donor \( \psi V \) segment and recipient V\(_4\) gene \( \text{shaded blocks} \). Insertion of sequence is shown by an inverted triangle. Precise ends are denoted by a vertical line, and ends of gene conversion events that cannot be localized because they overlap with other events are shown as open brackets.
in Figure 2 have been compared. The range for the length of sequence identity 5' of gene conversion events \(n = 102\) is 1–61 bp, with an average of 18.4 ± 1.8 bp \(\text{mean} \pm \text{SEM}\). None of the 5' ends occurs at precise end points (i.e., with a shared block of 0 identical nucleotides). The range for the length of sequence identity at the 3' ends of gene conversion events \(n = 110\) is 0–41 bp, with an average of 10.1 ± 0.9 bp \(\text{mean} \pm \text{SEM}\). The average length of identity between donor \(\psi\)VL segments and \(V_{\text{L}}\) at the 5' ends of gene conversion events is therefore significantly greater than the length of identity observed at 3' ends of gene conversion events \(p < 0.001\) by Student's t test.

**Precision of 5’ and 3’ ends of gene conversion events**

Differences were also observed in the precision with which the 5' and 3' ends of gene conversion events could be localized. Of the 102 gene conversion events with identifiable 5' ends, none had 5' ends that could be localized precisely. In contrast, of the 110 gene conversion events with identifiable 3' ends, eleven \(10\%\) had 3' ends that ended precisely between two specific nucleotides. The nucleotide sequences of two representative gene conversion events with precise 3' ends are shown in Figure 5A.

Other features observed at the 3' ends of gene conversion events, but not at 5' ends, are nucleotide insertions relative to the \(V_{\text{L}}\) gene segment. Although a number of \(\psi\)VL segments have codon insertions in CDR1 when aligned to \(V_{\text{L}}\) by homology (e.g., see Fig. 1A), we observed 12 gene conversion events that led to nucleotide insertion in regions where no insertion is apparent in the germ line \(\psi\)VL donor segment (two examples are shown in Fig. 5B). All 12 of these insertions occurred at 3' ends of gene conversion events, and alignments of the donor \(\psi\)VL and recipient \(V_{\text{L}}\) sequences suggest that they result from imprecise joining of \(\psi\)VL and \(V_{\text{L}}\) sequence. Two of the insertions led to 3' junctions between \(\psi\)VL and \(V_{\text{L}}\) sequence which placed the variable region sequence out of frame. Two other insertions resulted in the addition of sequence at the \(V_{\text{L}}\) junction \(V_{\text{L}}\) (Fig. 5C), which adds diversity to CDR3. Nucleotide deletions were observed at the 3' ends of two gene conversion events (Fig. 5D). All of these structural features unique to the 3' ends of gene conversion events suggest a polarity in the molecular mechanism of gene conversion.

**Discussion**

Diversity of the chicken immunoglobulin variable region gene segments is generated in the bursa of Fabricius by the transfer of nucleotide sequence blocks from families of \(\psi\)V segments into the unique rearranged \(V_{\text{L}}\) and \(V_{\text{L}}\) gene segments. Recent experiments have demonstrated that this recombination process is intrachromosomal gene conversion \(\text{Carlson et al. 1990}\). Organization of other avian \(\text{Ig}_{\text{L}}\) loci similar to the chicken \(\text{Ig}_{\text{L}}\) locus suggests that gene conversion may be a common strategy of \(V\) gene diversification among avian species \(\text{McCormack et al. 1989c}\). The present studies were undertaken to determine whether the sequence analysis of a large number of random gene conversion events may provide insights into the molecular mechanism by which intrachromosomal gene conversion occurs. All of the \(\text{Ig}_{\text{L}}\) clones we analyzed were obtained prior to antigen exposure, either at day 18 of embryogenesis or at the day of hatch, thus eliminating the potential influence of antigen selection on the sequences recovered. Furthermore, the observation that six of the clones are out-of-frame for translation suggests that these clones reflect the products of the gene conversion mechanism, rather than the result of selection at the protein level.

Gene conversion events could be identified to account for nearly all \(98.2\%\) of the 952 nucleotide substitutions observed in the nucleotide sequences of the 52 \(\text{Ig}_{\text{L}}\) clones. Nucleotide sequence polymorphisms within the \(\psi\)VL segments and \(V_{\text{L}}\) genes of the parental alleles of the SC chicken strain allowed assignment of donor \(\psi\)VL segments and recipient \(V_{\text{L}}\) genes to parental alleles. Twenty-four definitive \(\text{cis}\) events, but no \(\text{trans}\) events, were identified, consistent with the hypothesis that gene conversion occurs intrachromosomally. Selected gene conversion events spanning polymorphisms in parental \(\psi\)VL donor segments have also identified exclusively \(\text{cis}\) events \(\text{Carlson et al. 1990}\).

Sixteen nucleotide substitutions could not be derived from known \(\psi\)VL templates. Most of the untemplated substitutions observed in the 52 \(\text{Ig}_{\text{L}}\) clones and the cDNA sequences of \text{Reynaud et al. 1987} are transitions rather than transversions. Although \(\text{Taq}\) polymerase errors could have introduced some of these nucleotides \(\text{Saiki et al. 1988}\), the untemplated substitutions observed in clonal \(v\)-rel cell lines are unlikely to have occurred by that mechanism. As suggested by \text{Reynaud et al. 1987}, untemplated substitutions often occur at the ends of gene conversion events. However, we also observe some untemplated substitutions at sites distant from gene conversion events and others that are internal to gene conversion events (data not shown). This suggests that additional sequence heterogeneity may be generated either by an error-prone feature of gene conversion or by an independent mechanism of sequence diversification.

The frequency with which individual \(\psi\)VL segments are used in gene conversion events depends on three factors. [1] The \(\psi\)VL segments closest to the rearranged \(V_{\text{L}}\) gene are used more frequently than those farther upstream, as originally suggested by \text{Reynaud et al. 1987}. [2] \(\psi\)VL segments with the greatest homology with the \(V_{\text{L}}\) gene are used more frequently as donors. For example, full-length \(\psi\)VL segments are found more frequently in gene conversion events than truncated \(\psi\)VL gene segments \(\text{Fig. 4}\). [3] The \(\psi\)VL segments that are found in the inverted transcriptional orientation with respect to the \(V_{\text{L}}\) gene are used preferentially \(\text{Fig. 4 and Table 2}\). The simplest explanation for the bias in usage of inverted \(\psi\)VL segments as donors may be that it is easier for the chromosome to fold back on itself to align the \(\psi\)VL segment and \(V_{\text{L}}\) gene, than it is to form a complete loop required to align direct repeats.

The observation that homology is an important deter-
Figure 2. Summary of gene conversion events. Symbolic representations of 139 chicken $V_{L1}$ gene conversion events are illustrated, with symbols as described in Fig. 1C. Nucleotide positions are numbered from the 5' end of the $V_{L1}$ exon. The nucleotide boundaries of the $V_{L1}$ exon in rearranged genes (position 1 and the $V_{L1}-L$ junction at position 294) are marked by solid vertical lines. For most $\psi V_{L}$ donor segments, homology with $V_{L1}$ extends from position -25 in the leader intron to position +294 at the 3' end of the $V_{L1}$ exon.
Intronic FR1 CDR1 FR2 CDR2 FR3 CDR3 JL

-25 0 25 50 75 100 125 150 175 200 225 250 275 300

Nucleotide Position

Figure 3. Location of gene conversion events within VLI. The minimum gene conversion lengths for the 139 events illustrated in Fig. 2 were used to tabulate the number of times gene conversion occurred at each nucleotide position of the VLI gene. Nucleotide positions are numbered from the 5′ end of the VLI exon and subregions are labeled as in Fig. 2. The mean and ±1 SD calculated for the number of occurrences of gene conversion at positions -25 through 298 are indicated by horizontal lines.

Dominant of gene conversion has been made in other experimental systems (Waldman and Liskay 1986; Liskay et al. 1987), and is consistent with a variety of models of sequence transfer between a ψVLI segment and the VLI gene. However, the observation that relative orientation is a determinant of gene conversion donor frequency has important implications to our understanding of the molecular mechanism by which sequence transfer occurs. The orientation bias we observed suggests that sequence substitution does not occur through a diffusible intermediate, such as a DNA fragment or sterile transcript from the ψVLI region. Instead, the orientation bias supports a gene conversion model in which there is a physical interaction between the chromosomal ψVLI and VLI gene segments.

The gene conversion events we observed also demonstrate a 5′ to 3′ polarity. This polarity is revealed by the positions of gene conversion events within VLI and by differences in the homology and structure of the 5′ and 3′ ends of individual gene conversion events. As shown in Figures 2 and 3, only one gene conversion event is observed within the most 5′ region of homology between ψVLI donors and the VLI gene. There are 26 additional sequence polymorphisms between the ψVLI segments and the VLI gene within this region (Reynaud et al. 1987) that are not detected in our set of diversified IgL sequences. Gene conversions have occurred throughout the remaining VLI exon, however. When individual gene conversion events are analyzed, the 5′ ends are characterized by regions of high sequence identity between donor and recipient gene segments (Fig. 2). In contrast, the 3′ ends of gene conversion events have less sequence identity between donor and recipient gene segments, and often have precise ends, insertions, or deletions (Figs. 2 and 5). These observations support the hypothesis that gene conversion is initiated at the 5′ end in a region of sequence homology. Once initiated, gene conversion events may extend in the 3′ direction past the region of VLI homology into the JLI region (Reynaud et al. 1987), and in the heavy-chain gene past the region of VH homology into the D region (Reynaud et al. 1989).

A variation of single-strand break repair models for gene conversion (Meselson and Radding 1975; Radding et al. 1982) could account for the apparent polarity of IgL gene conversion we have observed. As shown in Figure 6, a single-stranded break in the nontranscribed strand of the rearranged VLI gene may initiate gene conversion during transcription. The accessibility of the nontranscribed strand to endonuclease cleavage is suggested by the “open” chromatin configuration of the rearranged VLI gene (Thompson and Neiman 1987), which may be caused by local changes in DNA topology associated with transcription (Liu and Wang 1987; Osborne and Reysenbach).
Table 2. Utilization of full-length $\psi V_L$ segments as gene conversion donors based on orientation

| Pseudogene orientation |  |  |
|-------------------------|---|---|
|                         | sense ($n=8$) | antisense ($n=7$) |
| Position                | 13.4 ± 6.1$^a$ | 14.0 ± 7.0$^a$ |
| Percent identity with $V_L$ | 83 ± 4 | 87 ± 3 |
| Length of identity      | 323 ± 23 | 312 ± 23 |
| Number of identical nucleotides | 269 ± 28$^a$ | 270 ± 21$^a$ |
| Times used as gene conversion donor | 4.5 ± 3.6$^b$ | 12.8 ± 4.8$^b$ |

Values are expressed as the mean ± the standard deviation. Means were compared for significant differences using the Student’s t test.

$^a$Not significant.

$^b$p < .01

Guarente 1988; Brill and Stern glanz 1988), or nicks left by incomplete religation by a type II topoisomerase associated with the transcription complex (Thomas and Rothstein 1989). A free 3' end generated by the nick may then interact with the duplex DNA of $\psi V_L$ segments. The transfer of a single DNA strand from the $V_L$ gene segment into the $\psi V_L$ duplex region may be mediated by a molecule similar to the rec1 protein of the lower eukaryote Ustilago (Kmiec and Holloman 1982), or the strand-transferase activities partially purified from human cells (Hsieh et al. 1986; Cassuto et al. 1987; Ganea et al. 1987). Unlike Escherichia coli recA protein (Radding 1989), these eukaryotic enzymes have been shown to insert a free 3' end into a homologous duplex and create a D loop by strand displacement in a 3' to 5' direction (Kmiec and Holloman 1983; Hsieh et al. 1986). The 3' end of the invading strand could then be used as a primer to initiate DNA synthesis using the $\psi V_L$ antisense strand as a template. Resolution of the extended $V_L$ nontranscribed strand from the $\psi V_L$ duplex without crossovers would leave an extended strand that overlaps with the preexisting strand for some length, depending

Figure 5. Molecular features unique to 3' ends of gene conversion events. Gene converted clones and $\psi V_L$ segments are aligned with the germ line sequence (top lines) according to flanking homologies (Reynaud et al. 1987). Sequence comparisons are illustrated as in Fig. 1. (A) Precise 3' ends of gene conversion events. (●) Junctions between germ line and $\psi V_L$ sequence. (B) Insertion/duplication at 3' ends of gene conversion events. (Insertions/duplications are underlined.) Imprecision at the 3' junction of the germ line and $\psi V_L$ sequence in the gene-converted clones has resulted in insertion of the underlined sequences. (C) Nucleotide insertions that add sequence to the $V_L$-JL junction are underlined. The CA nucleotides shown at the germ line $V_L$-JL junction for clone #67 represent specific nucleotide additions that occur during $V$-J joining in chicken B cells (McCormack et al. 1989a). (D) Imprecision at the 3' junction of the germ line and $\psi V_L$ sequence in the gene-converted clones may also result in the deletion of germ line sequence, as indicated by the parentheses.
on the length of 3' strand extension and possible exonuclease activity at the 5' end. The donated sequence may be incorporated into the V_{L1} duplex molecule by branch migration or by the rewinding action caused by the passage of RNA polymerase II (Gamper and Hearst 1982) during continued transcription. Structural features unique to 3' ends of gene conversion events may be caused by imprecision in strand ligation, resulting in occasional nucleotide insertion or deletion at this site. The final gene conversion product may be determined by the mechanism of heteroduplex DNA repair in bursal lymphocytes. A strand bias that uses the nontranscribed strand as the repair template would result in gene conversion in all events. Examples of preferential repair of the transcribed strand have been observed for some forms of DNA damage in mammalian cells (Mellon et al. 1987; Hanawalt 1989; Vrieling et al. 1989).

Chicken V_{L} gene conversion is also consistent with the double-strand break repair model of yeast gene conversion (Szostak et al. 1983; Sun et al. 1989), if polarity of gene conversion can be accounted for. Bidirectional strand extension from a double-strand break instead of a single-strand break would result in new DNA synthesis on the transcribed strand also (Fig. 6, #1). However, the homology and structural polarity of observed gene conversion events suggest that extensions of the transcribed strand are not recovered in the final product. One mechanism of excluding this strand from the gene conversion event may be the inability of the transcription complex to wind in this strand as a result of the presence of a nick or gap (Fig. 6, #2). Alternatively, if the extension of the transcribed strand is incorporated into heteroduplex DNA, preferential repair of the transcribed strand may remove donated sequence on that strand (Fig. 6, #3).

It has been suggested that gene conversion may contribute to the somatic hypermutation process observed in mammalian V_{H} and V_{L} genes (Maizels 1989; Wysocki and Getfer 1989). However, the complexity of mammalian Ig gene families complicates the search for definitive gene conversion recipient and donor sequences. In contrast, the unique genomic organization of the chicken Ig gene families appears to have evolved to promote this strategy for the somatic generation of Ig gene diversity by gene conversion, by providing a large number of pseudogene donor sequences within relatively small sites near the target V_{H} and V_{L} gene segments. The molecular mechanism by which gene conversion diversifies the chicken Ig variable region genes may now be approached experimentally because of the tissue specificity and developmental regulation of the gene conversion process in the chicken bursa of Fabricius.

**Methods**

**Chicken strains and cell lines**

Chickens used in these experiments were Hyline SC birds, an F_{2} cross between two inbred B'0 chicken strains, designated G4 and S3. Blood samples of the parental strains were obtained...
from Hylinc. The parental strains differ in several restriction enzyme sites around the Igκ locus (Thompson and Neiman 1987) and have nucleotide sequence polymorphisms within the leader intron, the V_L exon, and ψV_L segments (Thompson and Neiman 1987, Carlson et al. 1990, and this report). B-cell lines were derived from SC chickens by v-rel transformation as described by Carlson et al. (1990).

**Nucleotide sequencing of rearranged Igκ clones**

The isolation of rearranged Igκ genes from embryonic bursa DNA libraries and v-rel-transformed cell lines by cDNA cloning and by amplification using PCR (Saiki et al. 1988) has been described (McCormack et al. 1989a,b). Germ line ψV_L segments were amplified from erythrocyte DNA of each SC parental strain by PCR, using sense primers located in FR1 and antisense primers located in CDR3 or in 3'-flanking regions based on the published ψV_L sequences of the CB chicken strain (Reynaud et al. 1987). Dideoxynucleotide sequencing was performed using double-stranded plasmid template DNA with a Sequenase kit according to the supplier’s protocols (U.S. Biochemical). Each clone was sequenced completely on both strands using oligonucleotide primers specific for the SP6 and T7 promoter sites of the vector pGEM-3Z (Promega), and 17- to 20-mer primers at sites within the Igκ locus. The nucleotide sequences of the 52 Igκ clones will be submitted to the GenBank database.

**Statistical comparisons**

Means, standard deviations [s], standard errors of the means [SEM], and analysis of the difference between two independent sample means using the Student’s t test were calculated according to standard formulas (Bahn 1972).

**Acknowledgments**

We thank Louise Carlson and Christina Postema for technical assistance, Beverly Burck and Kurt Paffitt for preparation of figures, and Danie Osborne and Sheila Norton for oligonucleotide syntheses. We are grateful to Drs. Eric Humphries, Jeff Leiden, John Lowe, and Jim Wilson for review of the manuscript and helpful discussions. This work was supported by grant CA-48023 from the National Institutes of Health to C.B.T.

**References**

Alt, F.W., T.K. Blackwell, R.A. DePinho, M.G. Reth, and G.D. Yancopoulos. 1986. Regulation of genome rearrangement events during lymphocyte differentiation. *Immune Rev.* 89: 5–30.

Ayares, D., L. Chekuri, K.-Y. Song, and R. Kucherlapati. 1986. Sequence homology requirements for intermolecular recombination in mammalian cells. *Proc. Natl. Acad. Sci.* 83: 5199–5203.

Bahn, A.K. 1972. *Basic Medical Statistics.* Grune & Stratton, New York.

Benedict, A.A. and K. Yamaga. 1976. Immunoglobulins and antibody production in avian species. In *Comparative immunology* (ed. J.J. Marchalonis), pp. 335–375. Blackwell Scientific, New York.

Brill, S.J. and R. Sterniglaz. 1988. Transcription-dependent DNA supercoiling in yeast DNA topoisomerase mutants. *Cell* 54: 403–411.

Carlson, L.C., W.T. McCormack, C.E. Postema, E.H. Humphries, and C.B. Thompson. 1990. Templated insertions in the rearranged chicken Igκ V gene segment arise by intrachromosomal gene conversion. *Genes Dev.* 4: 1041–1047.

Cassuto, E., L.-A. Lightood, and P. Howard-Flanders. 1987. Partial purification of an activity from human cells that promotes homologous pairing and the formation of heteroduplex DNA in the presence of ATP. *Mol. Gen. Genet.* 208: 10–14.

Gamper, H.B. and J.E. Hearst. 1982. A topological model for transcription based on unwinding angle analysis of E. coli RNA polymerase binary, initiation and ternary complexes. *Cell* 29: 81–90.

Hanawalt, P.C. 1989. Concepts and models for DNA repair: from Escherichia coli to mammalian cells. *Environ. Mol. Mutagen.* [suppl. 16] 14: 90–98.

Hastings, P.J. 1988. Conversion events in fungi. In *Genetic recombination* (ed. R. Kucherlapati and G.R. Smith), pp. 397–428. American Society of Microbiology, Washington, D.C.

Hsieh, P., S. Meyn, and R.D. Camerini-Otero. 1986. Partial purification and characterization of a recombinase from human cells. *Cell* 44: 885–894.

Hunkapiller, T. and L. Hood. 1989. Diversity of the immunoglobulin gene superfamilies. *Adv. Immunol.* 44: 1–63.

Jalkanen, S., K. Granfors, M. Jalkanen, and P. Toivanen. 1983. Immune capacity of the chicken bursectomized at 60 hours of incubation: Failure to produce immune, natural, and autoantibodies in spite of immunoglobulin production. *Cell. Immunol.* 80: 363–373.

Jalkanen, S., M. Jalkanen, K. Granfors, and P. Toivanen. 1984. Defect in the generation of light-chain diversity in bursectomized chickens. *Nature* 311: 69–71.

Kmic, E.B. and W.K. Holloman. 1982. Homologous pairing of DNA molecules promoted by a protein from Ustilago. *Cell* 29: 367–374.

Liskay, R.M., A. Letsou, and J.L. Stachelek. 1987. Homology requirement for efficient gene conversion between duplicated chromosomal sequences in mammalian cells. *Genetics* 115: 161–167.

Liu, L.F. and J.C. Wang. 1987. Supercoiling of the DNA template during transcription. *Proc. Natl. Acad. Sci.* 84: 7024–7027.

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

McCormack, W.T. and C.B. Thompson. 1990. Somatic diversification of the chicken immunoglobulin light chain gene. *Adv. Immunol.* [in press].

McCormack, W.T., L.W. Tjoelker, L.M. Carlson, B. Petryniak, C.F. Barth, E.H. Humphries, and C.B. Thompson. 1989a. Chicken Igκ gene rearrangement involves deletion of a circular episome and addition of single nonrandom nucleotides to both coding segments. *Cell* 56: 785–791.

McCormack, W.T., L.W. Tjoelker, C.F. Barth, L.M. Carlson, B. Petryniak, E.H. Humphries, and C.B. Thompson. 1989b. Selection for B cells with productive Igκ gene rearrangements occurs in the bursa of Fabricius during chicken embryonic development. *Genes Dev.* 3: 838–847.

McCormack, W.T., L.M. Carlson, L.W. Tjoelker, and C.B.
McCormack and Thompson

Thompson. 1989c. Evolutionary comparison of the avian Ig locus: Combinatorial diversity plays a role in the generation of the antibody repertoire in some avian species. *Int. Immunol.* 1: 332–341.

Mellon, I., G. Spivak, and P.C. Hanawalt. 1987. Selective removal of transcription-blocking DNA damage from the transcribed strand of the mammalian DHFR gene. *Cell* 51: 241–249.

Meselson, M.S. and C.M. Radding. 1975. A general model for genetic recombination. *Proc. Natl. Acad. Sci.* 72: 358–361.

Osborne, B.I. and L. Guarente. 1988. Transcription by RNA polymerase II induces changes of DNA topology in yeast. *Genes Dev.* 2: 766–772.

Radding, C.M. 1978. Genetic recombination: Strand transfer and mismatch repair. *Annu. Rev. Biochem.* 47: 847–880.

——. 1989. Helical RecA nucleoprotein filaments mediate homologous pairing and strand exchange. *Biochim. Biophys. Acta* 1008: 131–145.

Radding, C.M., J. Flory, A. Wu, R. Kahn, C. DasGupta, D. Gonda, M. Bianchi, and S.S. Tsang. 1982. Three phases in homologous pairing: polymerization of recA protein on single-stranded DNA, synapsis, and polar strand exchange. *Cold Spring Harbor Symp. Quant. Biol.* 47: 821–828.

Reynaud, C.-A., V. Anquez, A. Dahan, and J.-C. Weill. 1985. A single rearrangement event generates most of the chicken immunoglobulin light chain diversity. *Cell* 40: 283–291.

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.

Rubnitz, J. and S. Subramani. 1984. The minimum amount of homology required for homologous recombination in mammalian cells. *Mol. Cell. Biol.* 4: 2253–2258.

Saiki, R.K., D.H. Gelfand, S. Stoffel, S.J. Scharf, R. Higuchi, G.T. Horn, K.B. Mullis, and H.A. Erlich. 1988. Primer-directed enzymatic amplification of DNA with a thermostable DNA polymerase. *Science* 239: 487–491.

Sun, H., D. Treco, N.P. Schultes, and J.W. Szostak. 1989. Double-strand breaks at an initiation site for meiotic gene conversion. *Nature* 338: 87–90.

Szostak, J.W., T.L. Orr-Weaver, and R.J. Rothstein. 1983. The double-strand-break repair model for recombination. *Cell* 33: 25–35.

Thomas, B.I. and R. Rothstein. 1989. Elevated recombination rates in transcriptionally active DNA. *Cell* 56: 619–630.

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.

Vrieling, H., M.L. Van Rooijen, N.A. Groen, M.Z. Zdzienicka, J.W.I.M. Simons, P.H.M. Lohman, and A.A. van Zeeland. 1989. DNA strand specificity for UV-induced mutations in mammalian cells. *Mol. Cell. Biol.* 9: 1277–1283.

Waldman, A.S. and R.M. Liskay. 1987. Differential effects of base-pair mismatch on intrachromosomal versus extrachromosomal recombination in mouse cells. *Proc. Natl. Acad. Sci.* 84: 5340–5344.

Weill, J.-C., C.-A. Reynaud, O. Lassila, and J.R.L. Pink. 1986. Rearrangement of chicken immunoglobulin genes is not an ongoing process in the embryonic bursa of Fabricius. *Proc. Natl. Acad. Sci.* 83: 3336–3340.

Weill, J.-C. and C.-A. Reynaud. 1987. The chicken B cell compartment. *Science* 238: 1094–1098.

Wysocki, L.J. and M.L. Gefter. 1989. Gene conversion and the generation of antibody diversity. *Annu. Rev. Biochem.* 58: 509–531.
Chicken IgL variable region gene conversions display pseudogene donor preference and 5' to 3' polarity.

W T McCormack and C B Thompson

*Genes Dev.* 1990, 4:
Access the most recent version at doi:10.1101/gad.4.4.548

References

This article cites 39 articles, 13 of which can be accessed free at:

http://genesdev.cshlp.org/content/4/4/548.full.html#ref-list-1

License

Email Alerting Service

Receive free email alerts when new articles cite this article - sign up in the box at the top right corner of the article or click here.