Human immunoglobulin heavy-chain minilocus recombination in transgenic mice: Gene-segment use in μ and γ transcripts

NADINE TUAILLON*, LISA D. TAYLOR†, NILS LONBERG‡, PHILIP W. TUCKER*, AND J. DONALD CAPRA*

*Department of Microbiology, University of Texas Southwestern Medical Center, 5323 Harry Hines Boulevard, Dallas, TX 75235; †GenPharm International, 297 North Bernardo Avenue, Mountain View, CA 94043

ABSTRACT We (N.L. and L.D.T.) have introduced a human heavy-chain minilocus into mice transgenically. Constructs contain 2 heavy-chain variable (VH; \(\psi_{VH}^{21} - 105\) and \(V_{H}^{5-251}\)), 10 diversity (D), 6 heavy-chain joining (JH), and either constant (C)\(\mu\) or C\(\gamma\) gene segments. Several founder lines were established and studied before immunization. Seventy heavy-chain transcripts were cloned and sequenced from murine splenic B lymphocytes, and gene-segment use was assessed before and after class-switching. In general, the repertoire was “fetal” in appearance with little evidence of somatic mutation in any gene segment. The two VH gene segments were found rearranged to \(\mu\)- and \(\gamma\)-chain C segments, with a preference of \(V_{H}^{5-251}\). We observed a preponderance of the most-J-proximal D gene (D\(\text{HO52}\)) segments among the \(\mu\) transcripts (44%). The J\(\text{H}\) gene-segment use mimics most patterns seen in human antibodies. Diversification in CD3 is extensive and included clear examples of D inversions and D-D fusions. These data suggest that a human immunoglobulin minilocus can undergo recombiniator processes in a manner analogous to that seen in the human fetal/preimmune repertoire. This model, in addition to providing a potential source of human monoclonal antibodies, is ideal for the study of further questions concerning immunoglobulin gene-segment recombination.

Recently, Bruggemann et al. (7, 8) described the production of transgenic mice carrying a largely human immunoglobulin minilocus. The transgenic mice were capable of rearranging the exogenous DNA and producing human \(\mu\) chains. Taylor et al. (9) have independently engineered similar transgenic mice to study, among other things, gene segment use. The constructs contain two human VH gene segments originally isolated on the same \(\lambda\) clone. \(V_{H}^{5-251}\) was chosen because it is overexpressed in the fetal repertoire, in patients with acute and chronic lymphocytic leukemia, and during autoimmune diseases (1, 10–13). Moreover, it is expressed as a germ-line transcript in all human B cells (14). \(\psi_{VH}^{3-105}\) is physically linked to \(V_{H}^{5-251}\) and contains a mutation in framework 1 (FR1, codon 28), rendering it a pseudogene (11). Ten human D gene segments were introduced (15–17). All six human J\(\text{H}\) gene segments were also included in the constructs. The constructs used here differ from those described by Bruggemann et al. (7, 8) in that they contain exclusively human coding sequences and all the gene segments are present on the same vector. Moreover, one of our constructs allows an analysis of the class-switching process.

Here, we report the analysis of the expression of two human immunoglobulin heavy-chain miniloci in transgenic mice. This study was initiated in order to analyze \(V_{H}, D, J\) and \(\psi\) gene-segment utilization in \(\mu\) and \(\gamma\) transcripts and to determine whether somatic mutation can occur in the absence of antigen stimulation. Our results indicate that the transgenic minilocus is expressed in mice in a manner analogous to that of the human heavy-chain complex. We found that the repertoire encoded by the minilocus was “fetal” with regard to the restricted gene-segment use and the virtual absence of somatic mutation. We also show that both the pseudogene (\(\psi_{VH}^{3-105}\)) and the functional gene (\(V_{H}^{5-251}\)) are rearranged but at greatly different frequencies. We show that the functional gene segment can undergo a class-switch, suggesting that the appropriate c-terminal elements in the human minilocus can be adequately used by mouse trans-acting factors.

MATERIALS AND METHODS

DNA Constructs and Transgenic Mice. Two different DNA constructs (pIGM1 and pHCl) were used to establish the three founder lines (strains 6, 15, and 57). The preparation of the constructs and mice has been described (9).

RNA Isolation and cDNA Synthesis. Total RNA was prepared from the spleens of 1-mo-old mice by acid guanidium thiocyanate–phenol–chloroform extraction (18). For single-strand cDNA synthesis, 10 \(\mu\)l of total RNA (10–20 \(\mu\)g) was heated at 70°C for 10 min in the presence of 10 pmol of random hexamers. The reaction was then cooled on ice. Reverse

Abbreviations: \(V_{H}\), heavy-chain variable; \(D\), diversity; \(J_{H}\), heavy-chain joining; \(N\) segment, segment containing nucleotides of unknown origin.
transcription was done at 37°C for 1 hr with Maloney murine leukemia virus RNase H+ reverse transcriptase (GIBCO/BRL), as described by the supplier.

PCR Experiments and DNA Sequencing. The PCR was done as described (19). Amplification of 2 μl of single-strand cDNA was done by using 40 cycles of 1-min denaturation (94°C), 1.5 min of annealing (52°C), and 2 min of elongation (72°C) with Taq polymerase (Promega). The final cycle was completed by 7 min elongation at 72°C. Ten picomoles of the primers: VH5-leader (AGT GGG TCA ACC GCC ATC CT), VH5-3 leader (AGT GAG TCT GCC GGT AGC TG), μCG (TGG GGC GGA TGC AGT CCG), and γ1 C (GGT TTT GTC CAA GAT TTT GGC) was used as appropriate. The PCR products were purified using microconcentrators (Centricon Amicon) and were used as templates. A low-melting temperature agarose gel, ligated into EcoRV-digested plasmid (pBluescript KS+). The ligation mixture was used to transform Escherichia coli BSI-72 competent cells. The resulting colonies were screened with a 32P-labeled oligonucleotide. Single-strand DNA was prepared and sequenced with the dideoxynucleotide chain-termination technique (26).

Amplification of Rearranged DNA. Splenic DNA was amplified as described above by using a 5′ oligonucleotide common for VH5-251 and ψVH3-105 (AGT TGC AGC TGG TGS AGT CTG) and a 3′ Jγ consensus (ACC TGA GGA GAC GGT GAC CAG GG) oligonucleotide. The amplified rearranged DNA was purified by using a low-melting temperature agarose gel, ligated into EcoRV-digested plasmids (pBluescript KS+), and grown in BSI-72 competent cells, as described above. The resulting colonies were screened with three different 32P-labeled oligonucleotides: one specific for VH5-251 (VH5) (CTA TCC TGG TGA CTC), a second specific for ψVH3-105 (VH3) (AAATA GTG TCA CGG AAG), or a probe common to both VH gene segments (Cint) (TGT ATT ACT GTG YGA GA). To determine the relative number of ψVH3-105 and VH5-251 rearrangements, >300 colonies resulting from two independent amplifications were screened.

RESULTS

Transgenic Mice with a Human Immunoglobulin Minilocus Produce Human μ and γ Chains. Three different founder lines were produced: strains 6 and 15 with the pIGM1 construct (2 VH, 10 D, 6 JH plus μCγ) and strain 57 with pHCl (2 VH, 10 D, 6 JH, Cμ plus Cγ). The number of transgene copies is estimated to be ~10–50, 5–20, and 10–50 for strains 6, 15, and 57, respectively. The presence of human μ or γ heavy chains in mouse sera was determined by ELISA. The mouse sera contained human μ or μ and γ chains, according to the construct used (9). Human μ heavy chain can also be detected by fluorescent antibody staining and flow cytometry on the surface of 5–10% of the murine splenic and peripheral B lymphocytes.

Both the ψVH3-105 and VH5-251 Gene Segments Are Used in μ Transcripts, but only VH5-251 Is Present in γ Transcripts. Amplification of splenic DNA revealed that the functional VH5-251 gene segment is rearranged more frequently than ψVH3-105. Indeed, among the 397 clones detected by Cint primer, 388 (97.8%) were detected by using VH5 primer, and 9 (2.2%) were detected by using VH3 primer. These experiments did not allow determination of the heavy-chain class. However, mRNA amplification and sequencing indicated that VH5-251 is present in both μ and γ transcripts, whereas ψVH3-105 could only be detected in μ transcripts.

The CD3 sequences of 48 μ and 11 γ transcripts using the VH5-251 gene segment are represented in Figs. 1 and 2. The human μ transcripts derive from the spleens of mice carrying the construct with Cμ (pIGM1, Fig. 1) or the construct with Cμ and Cγ (pHCl, Fig. 2a).

**FIG. 1.** Organization of heavy-chain CD3 of 35 human μ transcripts that use the VH5-251 gene segment. Indicated are sequences from the spleens of four transgenic mice (260, 263, 281, and 289) carrying the Cμ minilocus (pIGM1). Sequences are grouped by their D gene segment in the order of J-proximal (DH502) to J-distal (DH5A1) in the transgenic construct. D gene segments used in an inverted orientation are indicated as D-c (reverse complement) gene segments. Differences from the germ-line sequences are indicated by lowercase letters. Stop codons are underlined, and frame shifts are double-underlined.
D Gene-Segment Use Recapitulates the Human Fetal Repertoire in \( \mu \) Chains. To determine the origin of the D gene segments, the sequences of the transcripts were compared with the germ-line sequences of the 10 human D gene segments present in the constructs and to the known murine D gene segments (20). Due to the limited number of D gene segments within the transgenic construct, we could determine with relative certainty the origin of all CD3\( \delta \)s except one: 281 5M7 (Fig. 1), in which the D-encoded sequence is only 3 nt long, and the sequence TCT is not in any of the 10 D gene segments present in the construct. None of the human transcripts used a murine D gene segment. All the 10 human D gene segments are present but found in at least one transcript. The most-J-proximal D gene segment (DHOQ2) is dramatically underrepresented in the \( \mu \) transcripts (Table 1) with >44% of transcripts using DHOQ2, paralleling the situation in human fetal liver (4–6).

\( J_\delta \) Gene Segment Use Mimics the Human-Expressed \( J_\delta \) Repertoire. All six \( J_\delta \) gene segments were found in at least one transcript (Figs. 2 and 3). A preferential use of the \( J_\delta \)4 gene segment (52.8% of the transcripts) in both \( \mu \) and \( \gamma \) transcripts was seen (Fig. 4). The \( J_\delta \)2, \( J_\delta \)3, and \( J_\delta \)6 gene segments are used in 10%, 17%, and 16% of the transcripts, respectively. The \( J_\delta \)1 and \( J_\delta \)5 gene segments are underrepresented, accounting for <3% of the transcripts each (Fig. 4). These results are similar to \( J_\delta \) gene segment use seen in the human by several groups (5, 21).

The VDJ Junctions Show a Minimal Use of N-Segments or Palindromic Nucleotides. The organization of the CD3 of 59 \( \mu \) and \( \gamma \) transcripts is indicated in Figs. 1 and 2. In 23 sequences (38.9%), no N-segment additions are required to account for the junctions. However, 16 sequences (27.1%) contain N segments at both 5' and 3' ends of the D gene segment, and 10 (16.9%) of the sequences contain N segments at the 5' or 3' end of the D gene segment, respectively. The length of the N segments varies between 1 and 10 nt. The relatively low number of transcripts containing N segments and their short length is comparable to the situation seen in human fetal liver (4, 5) but differs from the reported results in murine fetal liver, in which N segments are rarely found (<5% of the sequences, ref. 22). Among the observed N segments, only eight (17% of the N segments) could be explained by palindromic (P) nucleotides. The length of the potential palindromic segments varies between 1 and 3 nt. Additionally, 33 sequences (58%) contain germ-line-encoded nucleotides deriving from the flanking heptamer sequences at the V\( \delta \)-D or at the D-J\( \delta \) junction. For several transcripts, the length of the N segment causes a frameshift in the J\( \delta \) sequence, leading to a nonproductive rearrangement. The percentage of nonproductive rearrangements is particularly high for the transcripts using the \( \nu \) or 215 5M9 has probably been deleted during V\( \delta \)-DJ\( \delta \) rearrangement. The D gene segment of transcript 215 5M9 (Fig. 2a) probably has been deleted during V\( \delta \)-DJ\( \delta \) rearrangement. All transcripts cannot be explained by a classical V\( \delta \)DJ\( \delta \) rearrangement. The D gene segment of transcript 215 5M9 (Fig. 2a) probably has been deleted during

---

**Table 1.** D gene segment use in \( \mu \) and \( \gamma \) transcripts

<table>
<thead>
<tr>
<th>D gene segment</th>
<th>( \mu ) transcripts, no. (%)</th>
<th>( \gamma ) transcripts, no. (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>DLR1</td>
<td>1 (1.7)</td>
<td>0 (0.0)</td>
</tr>
<tr>
<td>DXP1</td>
<td>6 (10.3)</td>
<td>2 (15.4)</td>
</tr>
<tr>
<td>DXP1'</td>
<td>6 (10.3)</td>
<td>4 (30.7)</td>
</tr>
<tr>
<td>DA1</td>
<td>1 (1.7)</td>
<td>0 (0.0)</td>
</tr>
<tr>
<td>DK1</td>
<td>1 (1.7)</td>
<td>2 (15.4)</td>
</tr>
<tr>
<td>DN1</td>
<td>9 (15.5)</td>
<td>1 (7.7)</td>
</tr>
<tr>
<td>DM2</td>
<td>1 (1.7)</td>
<td>0 (0.0)</td>
</tr>
<tr>
<td>DIR2</td>
<td>5 (8.6)</td>
<td>2 (15.3)</td>
</tr>
<tr>
<td>DLR2</td>
<td>2 (3.4)</td>
<td>0 (0.0)</td>
</tr>
<tr>
<td>DHQ52</td>
<td>26 (44.8)</td>
<td>2 (15.3)</td>
</tr>
</tbody>
</table>

**Fig. 2.** Organization of heavy-chain CDR3 of human \( \nu \)5–251 \( \mu \) and \( \gamma \) transcripts produced in the spleens of mice carrying the C\( \mu \) and C\( \gamma \) minilocus (pHCl). Sequences are annotated as for Fig. 1. (a) V\( \delta \)DJ\( \delta \) junction organization of 13 \( \mu \) transcripts. The D gene segment of transcript 215 5M9 has probably been deleted during V\( \delta \)-DJ\( \delta \) rearrangement. (b) V\( \delta \)DJ\( \delta \) junction organization of 11 \( \gamma \) transcripts.
**DISCUSSION**

There are several issues in V–D–J rearrangement that have been difficult to approach in the human because of the absence of a model system for studying recombination. How and why distinct V, D, and J gene segments are used in specific immune responses could be related to the DNA elements themselves (positional, structural, etc.) or could be selective. We undertook to produce transgenic mice containing a limited number of well-characterized immunoglobulin gene segments to test these various possibilities.

The spleens of the transgenic mice described here contain B cells that produce human heavy chains composed of the gene segments present in the minilocus, probably associated with endogenous light chains. Both μ and γ chains are secreted and can be detected in sera by ELISA or on the surface of B lymphocytes by fluorescent antibody staining and flow cytometry in the absence of immunization. The ability to undergo productive rearrangement did not depend on the chromosomal location of the transgene or on the number of transgene copies. Indeed, the miniloci introduced in our three different founder lines (strains 6, 15, and 57) undergo heavy-chain rearrangement.

The transgenic mice use the exogenous DNA in a polyclonal manner. In the present study we analyzed 70 transcript sequences, and it is noteworthy that two identical sequences were found in the same mouse. The DNA constructs used contained 2 VH, 10 D, and 6 JH gene segments, in addition to the Cμ or Cγ loci. This organization permits the formation of 120 different VDJ combinations, corresponding to 60 different possibilities per VH gene segment. In the sequenced transcripts, 17 different combinations using the Vγ5–251 gene segment are represented. Each particular Vγ5–251–D–J association is present 1–5 times, except for the Vγ5–251–D2H032–JH4 combination, which is found 14 times. However, the frequency of each V–D–J reflects the frequency of use of each gene segment more than a preferential combination of certain gene segments. Indeed, Vγ5–251, D2H032, and JH4 correspond to the gene segments preferentially used in our transcripts.

Splenlic DNA-amplification experiments indicate that although both the Vγ5–105 and Vδ5–251 gene segments are rearranged, the pseudogene is rearranged less frequently (2.2% of the rearranged human VH gene segments). The Vγ5–105 gene segment is a pseudogene because it contains a stop codon in framework 1 (11). Its rearrangement is, however, possible. The amplification of rearranged DNA did not allow determination of the heavy-chain class. The relatively low level of Vγ5–105 rearrangement could be related to the gene-segment position or the flanking sequences. Vγ5–251 is transcribed in a germ-line context in human pre-B and selected B cells (14) and in these transgenic mice (J. Cai and P.W.T., data not shown). Perhaps, the local chromatin structure surrounding Vγ5–251 provides a more “accessible” environment for recombination.

Splenlic mRNA amplification confirmed the existence of Vγ5–105 and Vδ5–251 rearrangements. The two VH gene segments are rearranged to the μ locus, with a higher frequency of the functional VH5–251 gene segment. However, the sequenced γ transcripts contained exclusively the functional VH5–251 gene segment. It has been shown in murine B cells that both productive and nonproductive V–D–J rearrangements are subjected to class-switch (23–28).

In the absence of immunization, class-switch within the minilocus is apparently rare. However, three (27%) of the sequenced γ transcripts were in frame (Fig. 2b), confirming that nonproductive rearrangements can undergo a class-switch. The mice described here are heterozygotic for the minilocus. Therefore, the class-switch probably occurs through an intrachromosomal rearrangement process (29).
The occurrence of a class-switch in the exogenous DNA indicates that the murine factors involved in the class-switch process can identify and use the human switching sequences, such as S segments (27, 28), and are probably able to transcribe the human Cγ locus as a germ-line transcript (30, 31). The possibility of a class-switch involving human Cμ and murine Cγ, as described by Shimizu et al. (32), was not assessed in the present study.

Concerning D gene-segment use, D_{H}Q_{52} is preferentially used in μ transcripts; 44% of the μ transcripts use this gene segment versus only 18% in γ transcripts. The other 56% of μ transcripts use one of the other nine D gene segments without any apparent preference (Table 1). A similar preferential use of D_{H}Q_{52} has been described in the human fetal repertoire but has not been reported in the human adult repertoire (4–6). The D_{H}Q_{52} gene segment is the most-J-proximal D gene segment in our constructs and in the human genome. However, it is unknown whether its preferential use is due to its location or to a particular flanking sequence. The number of γ transcripts in our study is probably too low (n = 11) to draw any conclusions on D gene-segment utilization frequencies at this time.

The J_{H}4 gene segment is preferentially used by the transgenic mice in 56% and 36.4% of the μ and γ transcripts, respectively (Fig. 4). The human J_{H}4 gene-segment selection from the minilocus is similar to that described in the human (5, 21, 33, 34). This selection includes a preferential use of J_{H}4; a relatively high frequency of J_{H}2, J_{H}3, and J_{H}6; and a low frequency of J_{H}1 and J_{H}5. If J_{H}4 gene-segment selection in the human is due to particular flanking sequences, the murine cellular machinery is apparently able to recognize and use these sequences equivalently.

In mammals, antibody diversity is the result of V-D-J recombination as well as several other mechanisms—including somatic mutation, junctional diversity, D-inversion, D-D fusion, and gene replacement (6, 21, 35). The present study indicates that in transgenic mice, these mechanisms also occur in exogenous DNA. The relatively low frequency of mutation in these human transcripts is probably a consequence of the pathogen-free environment in which the mice were kept. Of the mutations seen, there was no evidence that antigenic selection was involved. Functional diversity, which is postulated to be due to the involvement of several enzymes, including exonucleases and terminal deoxynucleotidyltransferase (TdT), was extensive in the human transcripts. In transgenic mice, these enzymes apparently participate in the recombination of the exogenous DNA. Several of the human CD3s only can be explained by invoking D-inverted sequences, and several seem to derive from D-D fusion. Although such D-D and D-inverted D fusions have been postulated, the limited number of D segments in our constructs allows an unambiguous assignment in each case. It is noteworthy that murine B cells can perform exogenous V-D-J recombination by using DIR gene segments, which have no structural murine counterpart. These results suggest that the murine B-cell machinery can rearrange exogenous human DNA by using the same mechanisms as the endogenous immunoglobulin gene locus. All these features are important to the production of human monoclonal antibodies, as they suggest that the transgenic mice will reproduce human gene-segment selection and repertoire maturation.

We are grateful to Virginia Pascual for valuable suggestions throughout this study. J.D.C. holds the Edwin L. Cox Distinguished Chair in Immunology and Genetics at the University of Texas Southwestern Medical Center. This work is supported by grants from the National Institutes of Health (AI-12127, AI-11851, GM-31689). N.L. is supported, in part, by a grant from the National Institutes of Health (AI-31003).