Joining of immunoglobulin heavy chain gene segments: Implications from a chromosome with evidence of three D–JH fusions

(Fredrick W. Alt* and David Baltimore)

Whitehead Institute for Biomedical Research, and Center for Cancer Research and Department of Biology, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139

Contributed by David Baltimore, March 8, 1982

ABSTRACT A chromosomal segment with a unique structure around the immunoglobulin heavy chain joining region (JH) has been molecularly cloned from an Abelson murine leukemia virus-transformed cell line. Attached to JH3 in the cloned DNA, in inverted sequence, is the DNA from JH2 to the JH1 recognition sequence. The inverted segment is attached at its other end to the 5′ recognition sequence of a diversity segment (D). To form this structure, three joining events must have occurred on the same chromosome. One of these events could have been a normal D–JH joining but the others must have been irregular events including ones that result in inversions. One of the joining events left fused recognition elements from JH2 and a D whose sequence shows that, during joining, reciprocal joinings of the recognition elements must occur to fuse the heptameric elements back to back. Because joined D and JH undergo deletion of terminal coding sequence during recombination but the joined heptameric recognition sequences do not contain the deleted sequence, joining must be a nonreciprocal event. Also, extra nucleotides are inserted between D and JH as part of the joining process; it is suggested that this added sequence is a product of the activity of terminal deoxynucleotidyltransferase at the D/JH (and probably the VH/D) joints and that it represents a new element of heavy chain gene structure, the N region.

The variable region of immunoglobulin light and heavy chains is encoded in multiple germ-line DNA elements which are rearranged somatically to form the complete variable region gene (1–5). Formation of a complete heavy chain variable region gene involves at least two recombinational events: joining of a variable gene segment (VH) to a diversity segment (D) and joining of the D to a joining segment (JH) which produces the complete VH-D-JH heavy chain variable region several kilobases (kb) 5′ from the most proximal constant region (CH) gene (4, 5). The recombination processes involved with variable region gene formation are apparently mediated by a set of highly conserved recognition sequences which consist of a palindromic heptamer and a characteristic nonamer separated by a spacer region (4, 5). A complete recognition sequence starts with the heptamer, lies flush with the 5′ border of each VH and D and the 5′ border of each D and JH (4–7). The spacer region is characterized by 23 base pairs (bp) long for VH and JH recognition sequences (4, 5) and 12 bp for the 3′ and 5′ D recognition sequences (6, 7). Apparently, the recombination process can only occur between recognition sequences containing 12- or 23-bp spacers, (4, 5). Thus the VH-to-JH joining process for heavy chains appears to be obligately mediated by VH/D and D/JH joinings.

Most of our current understanding of this joining process has come from studies that compared the organization and structure of the various elements in myeloma cell DNA to those in embryonic DNA. We have been exploring the possibility of using Abelson murine leukemia virus (A-MuLV)-transformed cells as a model system for studying aspects of the early stages of B-lymphoid cell differentiation. These studies have shown that some A-MuLV transformants continue immunodifferentiative events during growth in culture (8–11). Of particular interest are lines derived from transformation of early fetal liver cells. In culture, such lines usually exhibit continued heavy chain rearrangement at or near the JH region, in most cases probably in the form of continued rearrangement of a preexisting D–JH complex (8). In this report we describe a novel series of rearrangements at a single JH of such a fetal liver line and discuss the possible significance of these rearrangements in terms of the general mechanism of VH–D–JH recombination.

EXPERIMENTAL

The 38B9 cell line was derived by A-MuLV transformation of 13-day murine fetal liver cells (8). It was apparent that the 38B9 line underwent JH-associated rearrangements in culture because the parental line contained four distinct, nonembryonic JH-containing EcoRI fragments (8) whereas subclones derived from the line contained one or two of these fragments and often additional JH-containing EcoRI fragments of novel size (not shown). To define the events of this rearrangement process we have molecularly cloned rearranged JH-associated DNA fragments from this and other lines. Below, we describe a single EcoRI fragment containing a uniquely rearranged JH derived from a 38B9 subclone.

The 38B9-7 subclone had two JH-containing EcoRI fragments, one of 5.2 kb that comigrated with a fragment found in the parental clone and a new fragment of 5.6 kb. Total EcoRI-digested DNA from this subclone was cloned into the EcoRI site of Charon 16A λ phage, and recombinants containing DNA that hybridized to a JH-specific probe were purified and amplified as described (9). The 5.6-kb insert from one such clone was subcloned into the EcoRI site of pBR322 to yield plasmid pJH38B9-7.1 (the insert of which is hereafter referred to as "insert 7.1"). Comparison of a partial restriction map of insert 7.1 and the 6.2-kb EcoRI embryonic DNA fragment that contains the JH indicated that a new sequence had been rearranged into JH3 (Fig. 1). To analyze the nature of this rearrangement further, we prepared a specific probe from the rearranged portion of the clone

Abbreviations: V, variable segment; D, diversity segment; J, joining segment; C, constant segment; kb, kilobase(s); bp, base pair(s); A-MuLV, Abelson murine leukemia virus.

* Present address: Dept. of Biochemistry, College of Physicians and Surgeons, Columbia Univ., New York, NY 10027.
consisting of the 5'-terminal 2.0- and 0.7-kb BamHI fragments of insert 7.1. This probe hybridized to a set of embryonic EcoRI fragments that was nearly identical to the set containing the previously described families of genomic Ds (12). This result suggested that insert 7.1 contained a D in the absence of a V, because appending a V to a D would delete the 5' flanking sequence that provides the hybridization signal from a D.

To analyze the rearranged portion of this clone further, we determined the sequence of the 800 bases directly 5' to the JH3 joint as described in Fig. 2. A new sequence was rearranged into JH3 at point 3 bases 3' from the beginning of the JH3 coding sequence (point A in Fig. 2). The JH1-JH2 region of the chromosome had been inverted and attached to JH3. Specifically, the JH1 segment starting at point 5 bases from the start of its coding sequence and the entire JH1-JH2 spacer sequence had been inverted and linked to JH3 via the 4-base-pair (bp) sequence TA-C-C (point A). The inverted segment extended through the inverted JH2 nonamer recognition sequence and

---

**Fig. 1.** Partial restriction map of clone pJH38B9-7.1. Genomic DNA from cellular subclone 38B9-7.1 was digested to completion with EcoRI and cloned into the EcoRI site of Charon 16A λ phage. Phage containing inserts that hybridized to a JH-specific probe were then purified and amplified, and their inserts subcloned into the EcoRI site of pBR322. For restriction mapping, the subcloned inserts were excised from pBR322 by digestion with EcoRI, labeled with [32P] by nick-translation, and digested with various combinations of enzymes, and the products were analyzed by electrophoresis through agarose gels. All details of these methods have been reported (9). No sites are indicated between the most 5' BamHI and EcoRI sites in either map. (A) Partial restriction map of 6.2-kb embryonic EcoRI fragment which contains JH (5). (B) Partial restriction map of clone pJH38B9-7.1; thick line indicates the region of nonidentity between the two clones.

---

**Fig. 2.** Partial nucleotide sequence of insert 7.1. The 250-bp HindIII-Bgl II fragment spanning the JH3 joint and the adjacent 5' 600-bp Bgl II-BamHI fragment were labeled with [γ-32P]ATP at either end and subjected to sequence determination by the method of Maxam and Gilbert (13) as described (9). Only sequence around the novel joints that are found in insert 7.1 is indicated. Point A (lower part) indicates the 4-bp insert at the joint of JH3 and the inverted JH1 sequence. Point B indicates the fusion between the 5' flanking sequence of JH2 and a D element. For comparison, part of the published embryonic sequences of JH1, -2, and -3 plus their immediate 5' flanking sequence (5) and the 5' flanking sequence of DFL16.1 (12) are shown in italics above and below the corresponding sequences in insert 7.1. The 5'-to-3' orientation of the embryonic sequences, relative to the normal direction of transcription, is indicated. Differences between the sequence of insert 7.1 and the corresponding embryonic segments—including eight differences in the JH2-JH1 spacer region—are indicated with asterisks. Characteristic recognition heptamers and nonamers in insert 7.1 are bracketed.

**DISCUSSION**

Model of rearrangement

A schematic version of the structure described above is shown in Fig. 3C. Although this structure could have been generated by various rearrangement events, to explain it we have chosen only to consider events mediated by the D and JH recognition sequences and therefore subject to the 12-bp/23-bp spacer rule (4, 5). Given this assumption, insert 7.1 must have resulted from at least three events—one each occurring between a D and JH<sub>5</sub> 1, 2, and 3. Although a unique pathway of rearrangement cannot be derived from the data shown here, we will outline one series of rearrangements that leads to the final structure and present further details about individual steps or alternative pathways below.

The three steps involved in generating insert 7.1 could be as follows. (i) Starting with the embryonic configuration (Fig. 3A), a normal D-JH joining event fuses a D (indicated as D<sub>1</sub>) to JH<sub>1</sub> (Fig. 3B). (ii) The D<sub>1</sub>-JH<sub>1</sub> structure then loops around (Fig. 3C) and uses the 5' D<sub>1</sub> recognition sequence to promote fusion of the D<sub>1</sub>-JH<sub>1</sub> to JH<sub>3</sub> (Fig. 3D). Such a recombination event inverts the DNA segment from JH<sub>1</sub> to the end of the JH<sub>3</sub> heptamer recognition sequence. In the process, the inverted JH<sub>1</sub> joins to JH<sub>3</sub> by a 4-bp sequence T-A-C-G-G derived from D<sub>1</sub> and the JH<sub>3</sub> heptamer recognition sequence becomes fused to the 5' D<sub>1</sub> heptamer sequence (Fig. 3D). (iii) The final rearrangement fuses the 5' D heptamer sequence of a second D (indicated as D<sub>2</sub>) to the inverted heptamer sequence of JH<sub>2</sub> (Fig. 3E). This rearrangement event generates two products: one is a circle containing a fusion of D<sub>2</sub> to JH<sub>2</sub> (Fig. 3F); the other is the observed structure in insert 7.1 (Fig. 3G).

**Types of D-JH Recombination**

The proposed mechanism for the generation of insert 7.1 involves three types of D-JH recombinations which we will term "normal," "inverted," and "pseudo-normal" (Fig. 4).

**Normal D-JH Joining.** This type of joining involves the fusion of the 3' side of a D to a JH<sub>3</sub> (6). During this process the intervening DNA sequence is presumably excised as a circle (see below) by fusion of 3' D and JH recognition sequences. The first proposed rearrangement in the formation of insert 7.1 was a normal joining of D to JH<sub>1</sub> (Fig. 3).

**Inverted D-JH Joining.** We use the term "inverted D-JH joining" to describe the process by which a D uses its 5' recognition sequence to mediate joining with a JH<sub>1</sub> (Fig. 4). This type of recombination theoretically should be possible because a 5' recognition element

\[
\]

read in the inverted direction is nearly identical to a 3' recognition element

\[
5' \text{ C-A-C-A-G-T-C-G} (12 \text{ bp}) \text{ A-C-A-C-A-C-G-T-G-C} 3'
\]

and therefore could be appropriately recognized by recombination proteins. Whereas joining events in the normal orientation lead to deletion of the intervening sequence, such an event in the inverted configuration would invert the DNA sequence lying between the recombination points.

In our example, inverted recombination occurs between JH<sub>3</sub> and the 5' side of a D already joined on its 3' side to JH<sub>1</sub>. Thus, the 4-bp sequence T-A-C-G-G which links JH<sub>1</sub> and JH<sub>3</sub> (Fig. 2) is presumably the core sequence of a D that was deleted on both its 5' and 3' sides during the two recombination events it underwent. The occurrence of the sequence C-G-T-A-C in the core of many Ds (12) supports this proposal. Deletion of bases from both the D and JH<sub>3</sub>, as observed in the JH<sub>1</sub>-D-JH<sub>3</sub> joint in insert 7.1, apparently is a normal part of this recombination system (3-7, 12, 14, 15). Curiously, the T-A-C-G-G sequence is a direct repeat of the final 4 bases of the attached JH<sub>1</sub> segment, but we assume this to be coincidental. Although pathways leading to the generation of insert 7.1 could be postulated in which-
Inverted joining was the first or last step, all models subject to
the 23-bp/12-bp joining rule require an inverted joining event.

Inverted joining is possible because of the bifunctional nature
of Ds which must fuse with both a JH and a VH to mediate
the formation of a complete VDHJH heavy chain variable region
gene (4–7). No evidence for such inverted D–JH (or VH–D) join-
ing has been reported and it remains to be determined whether
such inverted recombination events can mediate the construc-
tion of complete VDHJH genes. It could be that the joining pro-
cess is organized in such a way as to avoid primary inverted
joints but that secondary inverted joints (Fig. 3) occur with some
frequency because joining of the 5’ side of a D to a JH “activates”
the ability of the 5’ side to join or because of the proximity of
a 5’ D recognition element to JH recognition elements after an
initial D joining has occurred.

**Pseudonormal D–JH Recombination.** We use this term to
describe the process by which the 5’ side of a D is linked to a
previously inverted JH (Fig. 4). The joining is referred to as
pseudonormal because it involves D and JH recognition ele-
ments oriented in their normal configuration, but the 5’ D rec-
ognition sequence (rather than the usual 3’ one) couples to a
previously inverted JH recognition sequence. The reverse ori-
etation of the process would cause the D–JH complex to be lost
from the chromosome as a circle; the fused recognition elements
would be retained. It seems unlikely that pseudonormal joining
plays any significant biological role. It should also be noted that
other recombination pathways can be constructed that lead to
insert 7.1 but involve only normal and inverted D–JH joining
(although one normal joint would have to involve a previously
inverted D).

**Retention of fused recognition elements**

An important aspect of the structure of insert 7.1 is that it retai-
ns the reciprocal product of a normal joining event—fused back-
to-back recognition heptamers. Recognition elements at a D–JH
joint usually are lost from the cell’s DNA. The ligation of the
fused recognition elements after the D–JH joining in insert 7.1
suggests that these elements are also ligated during a normal
D–JH joining. Thus, if the recombination process were intra-
chromosomal, the intervening sequence would be lost as a circle
(Fig. 4). Intrachromosomal recombination must have been in-
volved in the generation of insert 7.1 because inverted joinings
between sister chromatids would have resulted in one dicentric
chromosome and another lacking a centromere. In addition, the
general absence of sequence 5’ to JH in cells that have under-
gone D–JH joining (8, 16) and of Ds in cells with two VHD–D–JH
rearrangements (unpublished data) suggests that sister chroma-
matid D → JH or VH → DJH joinings are extremely rare events.

**Deletion and insertion during joining**

It is notable that, although the point of recombination between
D and JH is variable and almost never occurs at the boundaries
of both sequences (4–7, 9, 12, 15), in the fused recognition hepta-
mers the joint is a perfect back-to-back fusion of the heptamers
(Fig. 2). Similarly, perfectly fused Vc and Jc recognition se-
quencies are often found as reciprocal joining fragments in cells
that have undergone Vc-to-Jc joining (17, 18). Again, however,
the Vc-to-Jc joining is almost never precise (2, 3, 14). These
observations suggest that D-to-JH (Vc-to-Jc, etc.) joining is not
the result of a reciprocal recombination event but is a multistep
process. In Fig. 5 we present such a multistep process. Al-
though its details could easily be different, the overall format
of the process is likely to occur in this fashion.

At stage I, the partners are shown as being nicked DNA
strands because all four chains must be opened before recom-
bination can occur. At stage II, the recognition signals are
joined. Because the products of a joining event are not recip-
rocal, joining of the signals must be a separate event from the
joining of the coding elements. The production of a covalently
linked circle by this joining event is indicated because the re-
cover of the joined signals in insert 7.1 implies the occurrence
of such a product.

At stage II, the coding sequences do not join but they must
be held by protein (dashed box) so that their proximity is re-
tained. D or JH coding sequence must be removed to account
for the loss of bases at the joint (6, 7, 12). This is most easily
diagrammed as the activity of a double-strand exonuclease al-
though other possibilities could be imagined.

One of the most puzzling characteristics of the D–JH joining
process is that at the joint there are often extra nucleotides pres-
ent that do not occur in any known D or in any of the Js. It has
been suggested that they may arise from D–D joining (12), but
no evidence of such a process has yet been found. The structure
at stage III suggests that these nucleotides could be added to
the 3’ ends available after exonuclease digestion. An enzyme
capable of such an addition and present in many A-MuLV trans-
formants is terminal deoxynucleotidyltransferase (19). This en-
zyme, found at highest concentration in thymus but also present
Allelic exclusion

We have previously proposed that heavy chain allelic exclusion might be related in part to a high rate of nonproductive rearrangement but would ultimately rely on a cessation of the rearrangement process upon production of a heavy chain protein (or $C_n$ region) (8, 9). This model relied on the assumption that $D-J_H$ rearrangements were intermediate in the $V_H$ gene assembly process. The data in this paper support the proposal that $D-J_H$ rearrangements are not abortive because, once formed, a $D-J_H$ can clearly undergo further recognition sequence-mediated recombination.

The rearrangements involved in the generation of insert 7.1 also indicate that a $D$ joining to one $J_H$ does not render the other $J_H$ inactive with respect to recombination. However, it should be noted that, once a $V_HJ_H$ joining has occurred, functional or not, that allele will probably be terminally rearranged because no further $D$s would be available to mediate further $V_H$ joining (unpublished data). The possibility of continued rearrangement could be more significant in the generation of light chain genes, in which a nonfunctional $V_LJ_L$ joint potentially could be replaced by joining of an upstream $V_L$ to a downstream $J_L$.

We thank Elise Thomas for expert technical assistance and Dr. Michael Boss and Ms. Nancy Andrews for their advice and assistance. We are also indebted to Drs. Y. Kurosawa and S. Tonegawa for making available unpublished data and for their stimulating discussions—particularly their observation that no bases are added to joints between $D$s and $J$s that contain overlapping bases. This work was supported by Grant MV-34M from the American Cancer Society, Grant CA14051 (core grant to S. E. Luria) from The National Cancer Institute, and by a contribution from The Whitehead Charitable Foundation. D.B. is an American Cancer Society Research Professor.