Junctional Diversity in Signal Joints from T Cell Receptor Beta and Delta Loci via Terminal Deoxynucleotidyl Transferase and Exonucleolytic Activity

By Serge Candéias,* Kathrin Muegge,* and Scott K. Durum‡

From the *Intramural Research Support Program, SAIC Frederick, National Cancer Institute (NCI)-Frederick Cancer Research Center and Development Center; and ‡Laboratory of Molecular Immunoregulation, NCI, Frederick, Maryland, 21702-1201

Summary

The site-specific V(D)J recombination reaction necessary to assemble the genes coding for immunoglobulin (Ig) and T cell receptor (TCR) variable regions is initiated by a precise double strand cut at the border of the recombination signals flanking the genes. Extensive processing of the coding ends before their ligation accounts for most of the Ig and TCR repertoire diversity. This processing includes both base additions to and loss from the coding ends. On the other hand, it has generally been thought that signal ends are not modified before they are fused, and that signal joints consist of a perfect head-to-head ligation of the recombination signals. In this study, we analyzed signal joints created during the rearrangement of different TCR-β and TCR-δ genes in thymocytes. We show that a significant fraction (up to 24%) of these signal joints exhibits junctional diversity. This diversity results from N nucleotide additions for TCR-β signal joints, and from N additions and exonucleolytic digestion for TCR-δ joints. Altogether, our findings suggest that: (a) signal ends can undergo some of the same modifications as coding ends, (b) inversional rearrangement generates more diversity than deletional events, and (c) fine differences exist in the recombinase/DNA complexes formed at each rearranging locus.

The exons coding for the variable regions of Ig and TCR subunits are assembled during lymphocyte development from different V, D, and J gene segments by a site-specific recombination reaction. This rearrangement is specifically targeted at the Ig and TCR genes by the presence upstream and/or downstream of the V, D, and J segments of recombination signals (RSS)[1]. RSS are short nucleotidic motifs composed of conserved heptamers and nonamers separated by 12 or 23 bases; they are sufficient to direct the V(D)J recombinase activity (1). The first step of the reaction is the introduction of double strand breaks exactly between the V, D, or J coding sequences and their flanking RSS. These DNA ends are then resolved and there is formation of a coding joint (junction of the coding ends) and of a signal joint (junction of the RSS heptamers).

Using artificial recombination substrates, it has been recently shown that the products of the recombination activating genes, RAG-1 and RAG-2, are sufficient to perform the cleavage reaction in vitro (2). The resulting coding ends are sealed in a hairpin structure, whereas the signal ends are open and blunt. Hairpin structures have also been found in vivo, in recombination-deficient SCID mice (3), and in wild-type cell lines (4), and are believed to represent an intermediate of the rearrangement mechanism. If the rearranged genes are in the same transcriptional orientation, the intervening DNA is excised from the chromosome, and the signal joint is on a circular, extra-chromosomal piece of DNA (rearrangement by deletion). These circular DNA molecules can be found in thymocytes and mature lymphocytes (5–7). If the genes are in opposite transcriptional orientations, there is inversion of the intervening DNA and the signal joint is retained on the chromosome (rearrangement by inversion) (8, 9).

The assembly of the coding sequences is highly imprecise: nucleotides are frequently lost from the gene extremities, and extra bases can be added at the junctions, either nontemplated nucleotides or short inverted repeats (N and P nucleotides, respectively). Whereas the former result from the activity of the terminal deoxynucleotidyl transferase (TdT) (10, 11), the latter are most likely due to the opening and processing of the hairpin structure during the re-

combination process. In sharp contrast, the fusion of signal ends is generally thought to happen without any modification of the RSS extremities, so that the heptamers are ligated together without insertions and/or deletions, whether the rearrangement is by inversion or by deletion. This conclusion results essentially from the analysis of signal joints formed in B lymphocytes after the recombination of endogenous Igκ genes (12, 13) or after the rearrangement of exogenous artificial recombination substrates introduced in transformed pre–B cell lines (14). However, one of the key participants in the generation of junctional diversity, namely TdT, is usually not expressed in pre–B cells (15), and N nucleotides are generally absent from Igκ rearranged genes (16). Accordingly, the lack of N nucleotides observed in signal joints may merely reflect the absence of TdT activity in these cells rather than a true mechanistic constraint. Indeed, TdT-based diversity can be found at the signal joints after the recombination of artificial substrates if they are introduced in transformed pre–B cell lines that express TdT at a sufficient level (17). In contrast with B cell differentiation, TdT is expressed during all stages of T lymphocyte development in adult animals and N nucleotides can be found in TCR-β, -α, -δ, and -γ coding joints.

Although it has generally been thought that signal joints are formed by the ligation of perfect signal ends, there are examples in the literature of modified signal joints (18–21). To establish definitively whether N nucleotide insertion is a normal feature of signal joints created during the rearrangement of endogenous loci or a special feature of the artificial substrate assay, we analyzed the junctional diversity of signal joints formed after deletional and inversional rearrangement at TCR-β and -δ loci in thymocytes. Our results show that recombination by inversion leads to significantly more junctional diversity than recombination by deletion, and that different mechanisms may operate during TCR-δ versus TCR-β gene rearrangement.

Material and Methods

Mice. Male and female C57BL/6 mice were mated overnight and separated in the morning. Pregnant females were killed at 16 d of gestation. Newborn mice were killed within 24 h after birth.

DNA Preparation and Amplification of Signal Joints. The thymi were harvested from 16-d old fetuses, newborn mice, mice of 7 d and 3 wk of age, and adult animals. Thymocytes were digested overnight at 37°C in a 10 mM Tris-HCl, pH 7.5, 0.3 M sodium acetate, 1% SDS, and 100 μg/ml proteinase K. DNA was extracted once with phenol/chloroform, once with chloroform, and then ethanol-precipitated. DNA from TdT$^{-/-}$ animals and their $^{+/-}$ littermates (11) was kindly provided by Dr. S. Gilfillan (Institut de Génétique et de Biologie Moléculaire et Structurale, Strasbourg, France). 200 ng of DNA was amplified in 50 mM KCl, 20 mM Tris-HCl, pH 8.3, 2.5 mM MgCl$_2$, 0.1 mg/ml BSA in the presence of 0.2 pmol of each primer, 0.1 mM dNTPs, and 1 U of AmpliTaq (Perkin-Elmer Corp., Norwalk, CT) for 32 cycles, each cycle consisting of 30 s of denaturation at 94°C, 30 s of annealing at 55°C, and 1 min of elongation at 72°C. The samples were denatured for 5 min at 94°C prior amplification, and the last cycle was followed by 10 min of elongation at 72°C. The oligo-

nucleotide primers used to amplify Dβ1/Vβ3, Dβ1/Vβ8, and Dβ1/Vβ14 signal joints are 5′Dβ1 (5′-AGAGGAGCAGCT-TATCTGGTGG-3′) and 3′Vβ3 (5′-GGTTTTGCACTGGAG-AGAATGCAC-3′); 5′ Dβ1 and 3′Vβ8 (5′-TWGTTYTCCY-WATTYMCTTTCTGTGCA-3′) (W = AT; M = AC; Y = CT); and 5′Dβ1 and 3′Vβ14 (5′-CTTTGGTGACTTCTGACT-TGA-3′), respectively. To amplify Dδ2/Vδ2 and Dδ2/Vδ5 signal joints, we used the primers 5′Dδ2 (5′-TGGCTTGACATGCA-GAAAACACCTG-3′) and 3′ Vδ2 (5′-TAGTCCTGCCTCAG-GTGGGCTCT-3′) and 5′ Dδ2 and 3′Vδ5 (5′-CCTGGCAT-GTGACTTTCTGTG-3′), respectively.

Analysis of Signal Joint Diversity. After amplification, PCR products were extracted, ethanol precipitated, and resuspended in 20 μl of water. 2–8 μl of DNA were then digested overnight at 37°C with 20 U of ApaL1 in the manufacturer's buffer. Only unmodified signal joints are digested with this enzyme, as its recognition site (CACGTG) is created by RSS ligation if no bases are added and/or lost from the heptamers. The DNA fragments were separated on a 3.5% Metaphor agarose gel (FMC Bioproducts, Rockland, ME), and the gel blotted overnight onto a Nytran membrane (Schleicher and Schuell, Keene, NH). The prehybridization and hybridization were then performed in Rapid-Hyb buffer (Amersham Life Sciences Inc., Arlington Heights, IL) according to the manufacturer's instruction with the indicated oligonucleotide probes. To reveal signal joints including Dβ1 RSS, we used as a probe the oligonucleotide 5′Dβ1p (5′-AGGGGTAGAC-CAATGGGAGGG-3′). For TCR-δ PCR products, the probe was either 5′Dδ2p (5′-GACACGTGATACAAAGCCCAGGGAA-3′) or 3′Vδ5p (5′-AGGGCACTCCTAAACTGTGCTCC), as indicated in the figure legends.

For cloning, PCR products were blunted with Klenow, phosphorylated with T4 polynucleotide kinase and ligated into the EcoRV site of pZERO (Invitrogen, San Diego, CA). After transformation, bacterial colonies were screened by colony lifting for the presence of signal joint inserts. Plasmid DNA was prepared from each positive colony. Each plasmid was then individually digested with ApaL1. pZERO contains three sites for this enzyme. Thus, the insertion in the plasmid of a signal joint with no modifications creates a new ApaL1 site in the construct and the digest yields four bands on an agarose gel. Modified signal joints were identified by the presence of only three bands after ApaL1 digestion. Plasmids showing unexpected patterns after digestion, either in the number or the size of DNA fragments, were excluded from further analysis. All the modified signal joints were sequenced using a Sequenase kit (U.S. Biochemical Corp., Cleveland, OH). The sequences were then analyzed for junctional diversity by aligning them with the germline sequences of the RSS. Since we were unable to determine whether the multiple occurrence of a given sequence represented independent events or overamplification of one signal joint, we counted repeated sequences only once in Table 1. Sequences were generated and analyzed from either one C57BL/6 mouse and two TdT$^{+/-}$ animals or from two TdT$^{-/-}$ mutant mice. As similar distributions of modified vs unmodified signal joints were observed for each animal within each group, the data from all animals within one group were then pooled for analysis and comparison.

Results

Junctional diversity of coding joints results from both P and N nucleotide insertions and base loss from genes ends. Extensive base loss has not so far been reported at sig-
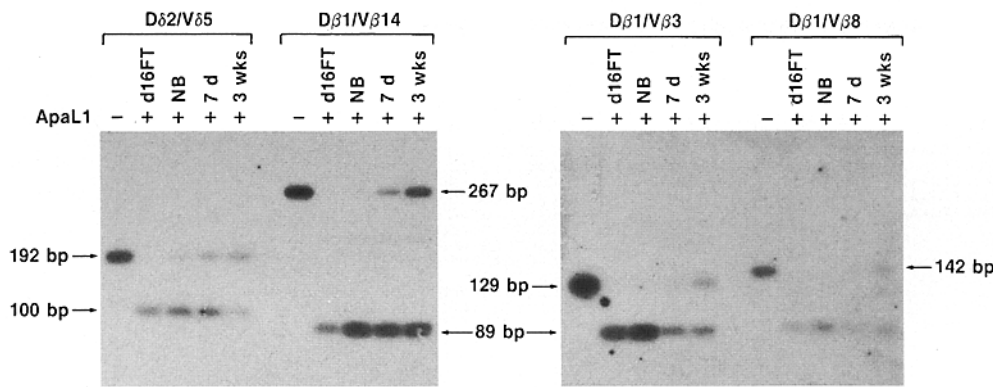
Figure 1. Ontogeny of diversity of signal joints at TCR-β and -δ loci Dδ2/Vδ5, Dβ1/Vβ14, Dβ1/Vβ3, and Dβ1/Vβ8 signal joints were amplified from DNA extracted from thymocytes at the indicated times of development. The PCR products were then extracted and digested with ApaL1 to detect the presence of modified junctions. Digested DNA fragments were separated on an agarose gel, blotted, and hybridized with oligonucleotide probes 3'Vδ5p and 5'Dβ1p to reveal the digestion products. We do not know the identity of the intermediate band in Dβ1/Vβ14 samples. (*d16FT*) Fetal thymus at 16 d of gestation, (*NB*) newborn thymocytes, (*7 d*) 7-d-old thymocytes; (*3 wks*) 3-wk-old thymocytes; (*bp*) basepairs.

nal joints in normal mice, and P nucleotide additions probably occur during the processing of the hairpin structure present only at coding ends. Consequently, if diversity exists at the signal joints, it should essentially be N nucleotide insertion polymerized by TdT. This enzyme is not constitutively expressed throughout lymphocyte ontogeny: it is absent from fetal and neonatal thymocytes, and substantial levels of TdT mRNA are not detected before 4 or 5 d of life (22). The pattern of N nucleotide insertions at TCR coding joints reflects this regulation: they are rare during the fetal/neonatal period and increase sharply 6 to 7 d after birth.

We analyzed signal joints produced during fetal life, at birth, and at 7 d and 3 wk of age by the recombination of Dβ1 and Vβ3, Dβ1 and Vβ8, Dβ1 and Vβ14, Dδ2 and Vδ5, and Dδ2 and Vδ2, to examine examples of deletional (Vβ3, Vβ8, Vδ2) and inversional (Vβ14, Vδ5) rearrangement at two different loci. In all cases, a new site for the restriction enzyme ApaL1 is created after fusion of intact RSS, allowing an easy distinction of modified and unmodified signal joints. Thus, after amplification with adequate primers, PCR products were digested with ApaL1 and the resulting DNA fragments separated on a high resolution agarose gel. Modified signal joints are not cut by ApaL1 whereas the digestion of unmodified signal joints generates smaller fragments.

Signal joints are readily detected from day 16 of gestation onward (Fig. 1). At this early time in ontogeny, virtually all are digestible by ApaL1, and thus nonmodified. In the case of Dβ1/Vβ14 and Dδ2/Vδ5 signal joints, ApaL1-resistant junctions become faintly detectable after birth and their proportion increases dramatically until 3 wk of age. In contrast, the appearance of ApaL1-resistant junctions seems delayed for Dβ1/Vβ3 and Dβ1/Vβ8, and reaches significant levels only between 7 d and 3 wk after birth. Thus, the appearance of ApaL1-resistant signal joints after inversional recombination closely parallels the kinetic of expression of TdT activity, suggesting that N nucleotides can be added to RSS by this enzyme before they are ligated together. In the case of deletional recombination, appearance of modified signal joints is delayed until after TdT expression

reaches its adult level, suggesting that these RSS ends are less susceptible (available) to TdT activity.

To confirm the involvement of TdT in the generation of modified signal joints, we then analyzed this phenomenon in adult (>6 wk) mutant mice lacking TdT activity (TdT$^{-/-}$) and age-matched heterozygous littermates (TdT$^{+/-}$). Modifications were easily detected in TdT$^{+/-}$ mice at all the signal joints examined. But in TdT$^{-/-}$ animals, we were unable to detect any ApaL1-resistant junctions after the amplification of Dβ1/Vβ3 and Dβ1/Vβ14 signal joints (data not shown). This finding definitively establishes that TdT is responsible for signal joint diversity within the TCR-β locus.

Signal Joints at the TCR-β Locus. To quantify the frequency of modifications, we cloned the amplified signal

Table 1. *Frequency of Modified Junctions within Signal Joint Populations*

	TdT $^{+/+}$ or $^{+/-}$		TdT$^{-/-}$	
	R	S	R	S
Dβ1/Vβ3	5	44	0	69
Dβ1/Vβ14	26	84	0	59
Dδ2/Vδ5	19	64	9	54

Indicated signal joints were amplified from thymocytes DNA extracted from TdT-expressing and TdT$^{-/-}$ mice. The PCR products were then cloned and signal joints containing plasmids isolated and individually digested with ApaL1 to identify modified and unmodified signal joints as described in Materials and Methods. For TdT-expressing mice, data are pooled from one C57BL/6 and two TdT$^{+/-}$ animals. For TdT$^{-/-}$ animals, the results presented here were obtained from two different mice. The higher frequency of modified vs unmodified junctions observed for Dβ1/Vβ14 signal joints when compared with Dβ1/Vβ3 joints is statistically significant (*P* <0.05 when compared by Pearson's chi-squared test without Yate's continuity). R, plasmids containing inserts resistant to ApaL1 digestion; S, plasmids containing inserts sensitive to ApaL1 digestion.

A TCRβ Locus Signal Joints

TDT +/+/+/-   TDT +/+/+/-

```
Dβ1                          Vβ3        Dβ1                              Vβ14
TGTAACATTGTG          CACAGCATTGAAA     TGTAACATTGTG              CACACTGAGTAGG

TGTAACATTGTG   GG     CACAGCATTGAAA     TGTAACATTGTG   GGG        CACACTGAGTAGG
TGTAACATTGTG   GC     CACAGCATTGAAA     TGTAACATTGTG   GCC        CACACTGAGTAGG
TGTAACATTGTG   CC     CACAGCATTGAAA     TGTAACATTGTG   GG         CACACTGAGTAGG
TGTAACATTGTG   C      CACAGCATTGAAA     TGTAACATTGTG   AA         CACACTGAGTAGG
TGTAACATTGTG   AGTT   CACAGCATTGAAA     TGTAACATTGTG   A          CACACTGAGTAGG
                                        TGTAACATTGTG   CG         CACACTGAGTAGG
                                        TGTAACATTGTG   C          CACACTGAGTAGG
                                        TGTAACATTGTG   ACC        CACACTGAGTAGG
                                        TGTAACATTGTG   CC         CACACTGAGTAGG
                                        TGTAACATTGTG   AG         CACACTGAGTAGG
                                        TGTAACATTGTG   TC         CACACTGAGTAGG
                                        TGTAACATTGTG   AGAAAC     CACACTGAGTAGG
                                        TGTAACATTGTG   ACT        CACACTGAGTAGG
                                        TGTAACATTGTG   T          CACACTGAGTAGG
                                        TGTAACATTGTG   CTCT       CACACTGAGTAGG
                                        TGTAACATTGTG   GG         CACACTGAGTAGG
                                        TGTAACATTGTG   ATC        CACACTGAGTAGG
                                        TGTAACATTGTG   GGCAG      CACACTGAGTAGG
                                        TGTAACATTGTG   GGG        CACACTGAGTAGG
                                        TGTAACATTGTG   AG         CACACTGAGTAGG
                                        TGTAACATTGTG   CCCC       CACACTGAGTAGG
                                        TGTAACATTGTG   AAA        CACACTGAGTAGG
                                        TGTAACATTGTG   AAC        CACACTGAGTAGG
                                        TGTAACATTGTG   AGT        CACACTGAGTAGG
                                        TGTAACATTGTG   ACG        CACACTGAGTAGG
                                        TGTAACATTGTG   TT         CACACTGAGTAGG
```

B TCRδ Locus Signal Joints

TDT +/+/+/-   TDT −/−

```
Dδ2                          Vδ5        Dδ2                              Vδ5
CTGTAGCACCGTG        CACTGTGGTGCAG      CTGTAGCACCGTG            CACTGTGGTGCAG

CTGTAGCACCGTG G      CACTGTGGTGCAG      CTGTAGCACCGTG              CTGTGGTGCAG
CTGTAGCACCGTG AC     CACTGTGGTGCAG      CTGTAGCACCG             CACTGTGGTGCAG
CTGTAGCACCGTG AGGGGG CACTGTGGTGCAG      CTGTAGCACCG              TGTGGTGCAG
CTGTAGCACCGTG TT     CACTGTGGTGCAG      CTGTAGCACC               TGTGGTGCAG
CTGTAGCACCGTG CCG    CACTGTGGTGCAG      CTGTAGCACC     T          TGGTGCAG
CTGTAGCACCGTG TGA    CACTGTGGTGCAG      CTGTAGCA                CACTGTGGTGCAG
CTGTAGCACCGTG GG     CACTGTGGTGCAG      CTGTAG                 CACTGTGGTGCAG
CTGTAGCACCGTG C      CACTGTGGTGCAG      CTGTA          T         TGTGGTGCAG
CTGTAGCACCGTG AG     CACTGTGGTGCAG      CTGTA                    TGTGGTGCAG
CTGTAGCACCGTG GGGGG  CACTGTGGTGCAG
CTGTAGCACC    TTTGACC CACTGTGGTGCAG
CTGTAGCAC     GA           CAG
CTGTAGCACCGTG GG        CTGTGGTGCAG
CTGTAGCACCGTG T        ACTGTGGTGCAG
```

```
Dδ1                          Vδ5
GTGTTTCACTGTG        CACTGTGGTGCAG

GTGTTTCACTGTG ATGGCTC  CTGTGGTGCAG
```

Figure 2. Sequences of ApaLl-resistant junctions. The signal joints found to be modified were sequenced to determine the exact molecular nature of the modifications. (A) Dβ1/Vβ3 and Dβ1/Vβ14 modified signal joints. No base loss from the RSS were observed. The genomic sequences of the different elements are shown at the top. The heptamers are underlined. All the sequences shown here were isolated from TdT-expressing mice, as no modified signal joint were found in TdT−/− animals. (B) Dδ2/Vδ5 modified signal joints in TdT-expressing and TdT−/− mice. Note the last sequence in TdT-expressing mice is a Dδ1/Vδ5 signal joint, probably amplified by cross-reaction of the 5βDδ2 primer. This sequence shows both deletion and N nucleotides insertion, and was included in our calculations (Table 1). In TdT−/− mice, two sequences acquired a single T during the fusion of signal ends. Such TdT-independent insertions are known to happen at a low frequency during V(D)J recombination in absence of TdT either in vivo or in cell lines (10, 26).

joints so that each junction can be analyzed individually for the presence of an ApaL1 site and sequenced to know precisely the nature of the modification(s). 11% (5 out of 44) of Dβ1/Vβ3 and almost 24% (26 out of 110) Dβ1/Vβ14 signal joints were found to be modified in TdT-expressing mice; none were found in TdT−/− mice (Table 1). Thus, inversional recombination leads to significantly more modifications at signal joints than deletional rearrangement ($P < 0.05$, see legend to Table 1). The sequencing of all the modified junctions reveals that no exonucleolytic loss from RSS ends occurred before they were ligated: in all cases the heptamer sequences are intact (Fig. 2 A). This finding, together with the fact that no modified junctions were found in TdT−/− mice, shows that N nucleotide addition is the only mechanism generating signal joint diversity at the TCR-β locus. In both inverted and deleted signal joints, the level of N nucleotide additions is similar (2.69 and 2.20 nucleotides per sequence for Vβ14 and Vβ3, respectively), and the insertions are GC rich (58 and 72% of GC for Vβ14 and Vβ3, respectively) (see Table 2 for Vβ14).

Signal Joints at the TCR-δ Locus. The recombination of

V δ5 and Dδ2 also occurs by inversion, and as for Dβ1/Vβ14 signal joints, we found an elevated frequency (almost 23%, 19 out of 83) of modified junctions after this rearrangement in TdT-expressing mice. However, contrasting with our findings concerning TCR-β gene rearrangement, we also found ApaL1-resistant junctions in TdT−/− mice, albeit at a lower level (14%, 9 out of 63) than in TdT-expressing mice (Table 1). In TdT-expressing mice, all the modified Dδ2/Vδ5 signal joints contained N nucleotides. In addition, four of them have lost nucleotides from one or both signal ends. A fifth example is the Dδ1/Vδ5 junction amplified by cross-reaction and cloned. In TdT−/− animals, all nine junctions show such deletions, and two of them have acquired an extra nucleotide (one T in each case). In both cases, the deletions range from 1 to 10 nucleotides (Fig. 2 B). Thus, diversity can be generated independently of TdT at the TCR-δ locus, by loss of bases from the signal ends. This mechanism does not seem to operate at the TCR-β locus, as we did not find any deletion in Dβ1/Vβ3 and Dβ1/Vβ14 signal joints. To examine other differences between TCR-δ and TCR-β rearrangements, we compared

Table 2. *Composition of N Nucleotide Insertions in Dβ1/Vβ14 and Dδ2/Vδ5 Signal Joints*

	G	A	T	C
Dβ1/Vβ14	20	20	9	21
Dδ2/Vδ5	22	7	9	8

Figure 3. ApaL1-resistant Dδ2/Vδ2 and Dδ2/Vδ5 signal joints are present in TdT$^{-/-}$ thymocytes. The indicated signal joints were amplified from TdT$^{+/-}$ and TdT$^{-/-}$ thymocytes, extracted, digested with ApaL1 and electrophoresed on an agarose gel, blotted, and hybridized with 5' Dδ2p, a probe specific for the 92-bp Dδ2 part of the signal joints. ApaL1-resistant junctions are present in both TdT$^{+/-}$ and TdT$^{-/-}$ mice, showing that signal joints modifications by TdT-independent mechanisms, probably deletions, take place both during rearrangement by deletion and by inversion at the TCR-δ locus. (*bp*) Basepairs.

the relative frequencies of each nucleotide within the N additions (Table 2). They were not significantly different and were in both cases GC rich (68 and 58% for δ and β, respectively). But for δ, this is mainly the result of a high representation of G, whereas for β, there are equal percentages of G and C. Given the preference of TdT for polymerization of dGTP, this difference suggests that both Vβ14 and Dβ1 signal ends are accessible to TdT during TCR-β rearrangement, whereas Vδ5 signal end is far less accessible than its Dδ2 counterpart. Thus, TCR-β and TCR-δ signal joints differ both in regard to base loss and in regard to the composition of N insertions.

As we found that inversional (compared with deletional) rearrangements have more junctional diversity at the TCR-β locus, we then wanted to determine if this relationship also applies to the TCR-δ genes. We amplified Dδ2/Vδ2 signal joints (which are generated by deletional rearrangement) from TdT-expressing and TdT$^{-/-}$ mice and analyzed them for the presence of modified junctions by ApaL1 digestion together with Dδ2/Vδ5 signal joints (generated by inversion) (Fig. 3). We can clearly detect ApaL1-resistant PCR products in TdT$^{-/-}$ mice after rearrangement of both Vδ2 and Vδ5 (Fig. 3, and data not shown), even if they are less abundant than in presence of TdT. For both Vδ2 and Vδ5, ApaL1-resistant PCR products present in TdT$^{-/-}$ appear to run a little bit faster than their TdT-expressing counterparts, and it is likely that they represent junctions in which bases have been lost from signal ends before ligation. Thus, TdT-independant diversity can be generated at the TCR-δ locus whether the recombination takes place by inversion or by deletion.

Discussion

Diversity of Signal Joints. N nucleotide additions were believed to occur in coding, but not in signal joints. In this report, we show that a sizable fraction (between 11 and 24%) of signal joints generated in thymocytes in vivo during the recombination of endogenous TCR-β and -δ loci are not monomorphous, but exhibit junctional diversity. At the TCR-β locus, this diversity is due exclusively to the insertion by TdT of N nucleotides between the RSS. At the TCR-δ locus, additional diversity is generated by the loss of up to 10 nucleotides from the signal ends. N nucleotide additions at signal joints have already been described during the rearrangement of TCR-δ and -γ genes (18–21), of IgH genes (23), and of artificial recombination substrates undergoing rearrangement after introduction in pre-B cell
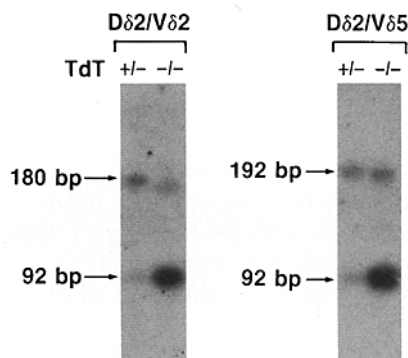
lines (17). We now show that these additions also take place during the rearrangement of endogenous TCR-β genes and that TCR signal ends can be processed by some of the mechanisms operating at coding ends, N nucleotide additions by TdT and/or exonuclease activity.

These findings contrast with previous studies in which it was found that signal ends at TCR-δ in BALB/C (24) and SCID mice (25) are almost exclusively blunt ended; only a minor fraction (<5%) of Dδ2 signal ends showed nucleotide loss or addition (24). Part of this discrepancy may stem from the fact that the substrate used in these experiments was thymocyte DNA extracted from newborn animals, in which TdT is not yet fully active. Indeed, modified junctions were shown to be essentially absent from newborns (18, 21), a conclusion confirmed here by our experiments (Fig. 1). They appear gradually with a kinetic reminiscent of that of TdT expression and N additions in TCR-β coding joints (22). These studies also failed to detect significant nucleotide loss at Dδ2 signal ends in adult animals, even by the ApaL1 digestion assay. This difference is unlikely to be due to the genetic background of the mice, because we observed this phenomenon in two TdT$^{-/-}$ mice with different MHC haplotypes (H-2^k and H-2^b). To reconcile these results one must postulate that the exonucleolytic activity acting on signal ends is developmentally regulated, as is TdT expression, a hypothesis also supported by the analysis of a limited set of TCR-δ signal joint sequences reported by Caroll et al. (20). But we cannot exclude that the ApaL1-resistant Dδ2/Vδ5 signal joints present in newborn thymus (Fig. 1) are due to a low level of such an activity around birth. It will be interesting to study the ontogeny of nucleotide loss at Dδ2 signal ends at different times during the development in TdT$^{-/-}$ mice. It should, however, be noted that the extent of coding end deletions does not vary between fetal/neonatal and adult Ig and TCR repertoires (26).

Blunt-ended signal ends were also reported at IgH and Igκ loci in newborn and adult bone marrow (27). But in this study, no sensitive assay like ApaL1 digestion was used to probe the exact molecular structure of signal ends, and the assays used may not have been sensitive enough to detect nucleotide loss in a fraction, even sizable, of RSS ends. Moreover, in light of our results showing that different mechanisms may operate at different TCR loci in T cells (see below), we cannot exclude that unlike TCR signal ends, Ig signal ends in B lymphocytes are truly all blunt ended.

Signal Joint Diversity at TCR-β Locus. We analyzed one example each of deletional (Dβ1/Vβ3) and inversional (Dβ1/Vβ14) rearrangement at the TCRβ locus. Diversity was found significantly more at the latter. Although some of the nucleotides found in these junctions could theoretically be P nucleotides or arise from imprecise cutting by the V(D)J recombinase, no diversity was found in $TdT^{-/-}$ thymocytes. Thus, the only mechanism operating at this locus is N nucleotide addition by TdT.

TdT is a potentially "harmful" and mutagenic enzyme: it can add nontemplated nucleotides to free DNA ends, and it seems unlikely that such an activity could wander freely within the nucleus, where it could modify DNA ends arising from various types of lesions before they are repaired. Thus, it seems likely, although no data have shown it so far, that TdT is tightly linked to the recombinase in developing lymphocytes to control its activity, and the only place where signal joints can acquire N nucleotides is within the recombinase/DNA complex. Thus, our results suggest that the V(D)J recombinase can ligate signal ends as well as coding ends. We believe that the higher percentage of modified junctions after inversional recombination at the β locus reflects a difference in the structure of the recombinase complex, when compared with deletional rearrangement. After inversion, the signal joint is retained on the chromosome. In this case, the joining of signal ends generated by the recombinase is crucial to maintain the genome integrity and to the future development of the T cell in which this event takes place. On the other hand, the formation of a signal joint is dispensable for the survival of the cell after rearrangement by deletion. It is possible that in the former case, signal and coding ends are tightly held within the recombinase complex, so that the formation of coding and signal joints can be achieved simultaneously, by the same ligation mechanism. According to this model, signal ends would be available for modifications by TdT, and we found that nearly one-quarter of Dβ1/Vβ14 show such modifications. For rearrangements by deletion, the DNA separating the recombined genes is excised from the chromosome, and signal ends have been shown to be quite abundant in thymocytes (7). It is possible that they are liberated from the recombinase complex shortly after they are generated, and that their ligation occurs outside this complex (4). Thus, in the majority of the cases, signal ends generated by deletional rearrangement are not accessible to TdT activity. The likelihood for a signal end to be modi-

fied by TdT could depend on its rate of release from the recombinase/DNA complex.

Signal Joint Diversity at TCR-δ Locus. At the TCR-δ locus, we found that signal joint diversity is generated both by N nucleotides additions to and deletions from the signal ends. Extensive nucleotide loss from signal ends has been shown during the rearrangement of artificial substrates in SCID cells (28) and more rarely in endogenous TCR loci (18, 20, 21). The data presented here indicate that this loss can also occur at significant levels during the rearrangement of TCR-δ genes in normal thymocytes. These deletions are not specific for Dδ2/Vδ5 inversional rearrangement, because ApaL1-resistant Dδ2/Vδ2 junctions are also present in $TdT^{-/-}$ thymocytes. They could theoretically be generated from an imprecise cutting by the recombinase. But this would generate "reciprocal" junctions in which some nucleotides from the coding ends are retained, and we did not find any in $TdT^{-/-}$ mice. Thus, our results are in agreement with the current model for V(D)J rearrangement, in which this reaction is initiated by a cut exactly at the RSS border (2, 29). Recently, it has been proposed that deletion of coding ends could arise exclusively from the processing of the hairpin structure sealing these ends (30). So far, no such hairpin has been described for signal ends. Furthermore, the presence of a hairpin at signal ends would probably lead to the presence of P nucleotides in signal joints, and we did not find any in $TdT^{-/-}$ mice, neither at TCR-δ nor TCR-β loci. Thus, the deletions observed at TCR-δ signal joints are probably created by a true exonucleolytic activity and our results suggest that such an activity can be associated with the recombinase complex.

Deletions were found only at TCR-δ, but not TCR-β signal joints. This finding suggests that different mechanisms operate during the recombination of these loci. Subtle differences between two rearranging systems have already been described. For example, the addition of short inverted repeats at processed coding ends occurs only in human cell lines (even in human fibroblasts transfected with murine RAG-1, RAG-2 and TdT), but not in cells of murine origin (31). Another example is the fact that in human, but not murine, lymphoid cell lines, signal joint formation is by far more efficient than coding joint formation and consequently, inversional rearrangement efficiency is much lower than for deletional rearrangement (32). Our results could be explained in either of two ways: (a) this effect is due to the nature of the locus (TCR-β vs TCR-δ), or (b) this effect is specific for the cell type (α/β vs γ/δ T cells) in which the rearrangement takes place. We used DNA extracted from unseparated thymocytes to amplify TCR-δ signal joints and γ/δ-expressing T cells represent only a small fraction of total thymocytes. Moreover, TCR-δ rearrangement products can be retained and are easily detected in α/β T cells (7, 33). Thus, it is likely that the vast majority of Dδ2/Vδ5 and Dδ2/Vδ2 signal joints analyzed arise from TCR-δ rearrangement in α/β expressing cells, and we favor our first hypothesis. To prove this point, it will be necessary to analyze and compare the junctional di-

versity at TCR-β and TCR-δ signal joints in purified α/β and γ/δ T lymphocytes.

The presence of N diversity in addition to nucleotide loss suggests that this exonucleolytic activity is linked to the recombinase complex, but we observed no deletions of TCR-β signal ends. Moreover, TCR-β and TCR-δ signal joints also differ in the composition of their N nucleotide additions and we can only speculate on the origins of these different patterns of processing. Such differences may indicate that signal ends created during TCR-β and TCR-δ rearrangement are differentially accessible to the DNA-ends processing mechanisms, and that the overall architecture of V(D)J recombinase/DNA complexes is different at TCR-β and TCR-δ loci. It is possible that the complex formed during rearrangement of TCR-δ genes has a different three-dimensional structure (or is less stable) than at the TCR-β locus, and allows more (or a different) exposure of signal ends to the processing mechanisms. Alternatively, we cannot exclude that more exonuclease(s) activity(ies) is

(are) recruited during TCR-δ rearrangement by the factor(s) mediating TCR-δ genes accessibility to the recombinase.

In conclusion, signal joints in murine thymocytes display considerably more junctional diversity than previously known. TdT is responsible for adding N bases in signal joints of both β and δ loci. An exonuclease deletes bases in δ signal joints but not in β. At the TCR-β locus, diversity is more frequent after rearrangement by inversion than after rearrangement by deletion. This finding suggests that signal ends are associated more tightly with the recombinase complex in the former case than in the latter, perhaps to maximize the probability that a signal joint will be created in order to avoid chromosomal damage and to promote cell survival. Overall, these results indicate that both coding and signal ends can be processed before they are ligated, and that fine differences in the structure and/or composition of the recombinase complex may exist between different loci.

References

1. Lewis, S. 1994. The mechanism of V(D)J joining: lessons from molecular, immunological and comparative analyses. Adv. Immunol. 56:27–150.

2. McBlane, J.F., D.C. van Gent, D.A. Ramsden, C. Romeo, C.A. Cuomo, M. Gellert, and M.A. Oettinger. 1995. Cleavage at a V(D)J recombination signal requires only RAG1 and RAG2 proteins and occurs in two steps. Cell. 83:387–395.

3. Roth, D.B., J.P. Menetski, P.B. Nakajima, M.J. Bosma, and M. Gellert. 1992. V(D)J recombination: broken DNA molecules with covalently sealed (hairpin) coding ends in scid mouse thymocytes. Cell. 70:983–991.

4. Ramsden, D.A., and M. Gellert. 1995. Formation and resolution of double-strand break intermediates in V(D)J rearrangement. Genes Dev. 9:2409–2420.

5. Okasaki, K., D.D. Davis, and H. Sakano. 1987. T cell receptor β gene sequences in the circular DNA of thymocyte nuclei: direct evidence for intramolecular DNA deletion in V-D-J joining. Cell. 49:477–485.

6. Toda, M., S. Fujimoto, T. Iwasato, S. Takeshita, K. Tezuka, T. Ohbayashi, and H. Yamagishi. 1988. Structure of extrachromosomal circular DNAs excised from T-cell antigen receptor alpha and delta loci. J. Mol. Biol. 202:219–231.

7. Livak, F., and D.G. Schatz. 1996. T-cell receptor α locus V(D)J recombination by-products are abundant in thymocytes and mature T cells. Mol. Cell. Biol. 16:609–618.

8. Malissen, M., C. McCoy, D. Blanc, J. Trucy, C. Devaux, A.-M. Schmitt-Verhulst, F. Fitch, L. Hood, and B. Malissen. 1986. Direct evidence for chromosomal inversion during T-cell beta-gene rearrangements. Cell. 319:28–33.

9. Korman, A.J., J. Maruyama, and D.H. Raulet. 1989. Rearrangement by inversion of a T-cell receptor δ variable region gene located 3' of the δ constant region gene. Proc. Natl. Acad. Sci. USA. 86:267–271.

10. Komori, T., A. Okada, V. Stewart, and F.W. Alt. 1993. Lack of N regions in antigen receptor variable genes of TdT-deficient lymphocytes. Science (Wash. DC). 261:1171–1175.

11. Gilfillan, S., A. Dierich, M. Lemeur, C. Benoist, and D. Mathis. 1993. Mice lacking TdT: mature animals with an immature lymphocyte repertoire. Science (Wash. DC). 261: 1175–1178.

12. Selsing, E., J. Voss, and U. Storb. 1984. Immunoglobulin gene "remnant" DNA-implications for antibody gene recombination. Nucleic Acid Res. 12:4229–4246.

13. Steinmetz, M., W. Altenburger, and H.G. Zachau. 1980. A rearranged DNA sequence possibly related to the translocation of immunoglobulin gene segments. Nucleic Acid Res. 8: 1709–1720.

14. Lewis, S., A. Gifford, and D. Baltimore. 1985. DNA elements are asymmetrically joined during the site-specific recombination of kappa immunoglobulin genes. Science (Wash.

DC). 228:677–685.

15. Hardy, R.R., C.E. Carmack, S.A. Shinton, J.D. Kemp, and K. Hayakawa. 1991. Resolution and characterization of pre-B and pre-pro-B cell stages in normal mouse bone marrow. *J. Exp. Med.* 173:1213–1225.

16. Hiramatsu, R., K. Akagi, M. Matsukoa, K. Sakumi, H. Nakamura, L. Kingsbury, C. David, R.R. Hardy, K.-I. Yamamura, and H. Sakano. 1995. The 3' enhancer region determines the B/T specificity and pro-B/pre-B specificity of immunoglobuli V_κ-J_κ joining. *Cell.* 83:1113–1123.

17. Lieber, M.R., J.E. Hesse, K. Mizuuchi, and M. Gellert. 1988. Lymphoid V(D)J recombination: nucleotide insertion at signal joints as well as coding joints. *Proc. Natl. Acad. Sci. USA.* 85:8588–8592.

18. Iwasato, T., and H. Yamagishi. 1992. Novel excision products of T cell receptor γ gene rearrangements and developmental stage specificity implied by the frequency of nucleotide insertions at signal joints. *Eur. J. Immunol.* 22:101–106.

19. Takeshita, S., M. Toda, and H. Yamagishi. 1989. Excision products of the T cell receptor gene support a progressive rearrangement model of the α/δ locus. *EMBO (Eur. Mol. Biol. Organ.) J.* 8:3261–3270.

20. Carroll, A.M., J.K. Slack, and X. Mu. 1993. V(D)J recombination generates a high frequency of nonstandard TCR Dδ-associated rearrangements in thymocytes. *J. Immunol.* 150:2222–2230.

21. Carroll, A.M., J.K. Slack, and W.-T. Chang. 1993. Biased T-cell receptor δ element recombination in scid thymocytes. *Mol. Cell. Biol.* 13:3632–3640.

22. Bogue, M., S. Gilfillan, C. Benoist, and D. Mathis. 1992. Regulation of N-region diversity in antigne receptors through thymocyte differenciation and thymus ontogeny. *Proc. Natl. Acad. Sci. USA.* 89:11011–11015.

23. Shimizu, T., and H. Yamagishi. 1992. Biased reading frames of pre-existing D_H-J_H coding joints and preferential nucleotide insertions at V_H-DJ_H signal joints of excision products of immunoglobulin heavy gene rearrangements. *EMBO (Eur. Mol. Biol. Organ.) J.* 11:4869–4875.

24. Roth, D.B., C. Zhu, and M. Gellert. 1993. Characterization of broken DNA molecules associated with V(D)J recombination. *Proc. Natl. Acad. Sci. USA.* 90:10788–10792.

25. Zhu, C., and D.B. Roth. 1995. Characterization of coding ends in thymocytes of *scid* mice: implications for the mechanism of V(D)J recombination. *Immunity.* 2:101–112.

26. Nadel, B., S. Tehanchi, and A.J. Feeney. 1995. Coding end processing is similar throughout ontogeny. *J. Immunol.* 154:6430–6436.

27. Schlissel, M., A. Constantinescu, T.B. Morrow, M. and A. Peng. 1993. Double-strand signal sequence breaks in V(D)J recombination are blunt, 5'-phosphorylated, RAG-dependent, and cell cycle regulated. *Genes Dev.* 7:2520–2532.

28. Lieber, M.R., J.E. Hesse, S. Lewis, G.C. Bosma, N. Rosenberg, K. Mizuuchi, M.J. Bosma, and M. Gellert. 1988. The defect in murine severe combined immune deficiency: joining of signal sequences but not of coding segments in V(D)J recombination. *Cell.* 55:7–16.

29. van Gent, D.C., J.F. McBlane, D.A. Ramsden, M.J. Sadofsky, J.E. Hesse, and M. Gellert. 1995. Initiation of V(D)J recombination in a cell-free system. *Cell.* 81:925–934.

30. Nadel, B., and A.J. Feeney. 1995. Influence of coding-end sequence on coding-end processing in V(D)J recombination. *J. Immunol.* 155:4322–4329.

31. Gauss, G.H., and M.R. Lieber. 1996. Mechanistic constraints on diversity in human V(D)J recombination. *Mol. Cell. Biol.* 16:258–269.

32. Gauss, G.H., and M.R. Lieber. 1993. Unequal signal and coding joint formation in human V(D)J recombination. *Mol. Cell. Biol.* 13:3900–3906.

33. Livak, F., H.T. Petrie, I.N. Crispe, and D.G. Schatz. 1995. In-frame TCR-δ gene rearrangements play a critical role in $\alpha\beta/\gamma\delta$ T lineage decision. *Immunity.* 2:617–627.