

## Postcleavage Sequence Specificity in V(D)J Recombination

EMILY A. AGARD AND SUSANNA M. LEWIS\*

*Program in Genetics and Genomic Biology, Hospital for Sick Children Research Institute, Toronto, Ontario, Canada M5G 1X8, and Department of Immunology, University of Toronto, Toronto, Ontario, Canada M5S 1A8*

Received 16 March 2000/Returned for modification 7 April 2000/Accepted 24 April 2000

**Unintended DNA rearrangements in a differentiating lymphocyte can have severe, oncogenic consequences, but the mechanisms for avoiding pathogenic outcomes in V(D)J recombination are not well understood. The first level at which fidelity is instituted is in discrimination by the recombination proteins between authentic and inauthentic recombination signal sequences. Nevertheless, this discrimination is not absolute and cannot fully eliminate targeting errors. To learn more about the basis of specificity during V(D)J recombination, we have investigated whether it is possible for the recombination machinery to detect an inaccurately targeted sequence subsequent to cleavage. These studies indicate that even postcleavage steps in V(D)J recombination are sequence specific and that noncanonical sequences will not efficiently support the resolution of recombination intermediates in vivo. Accordingly, interventions after a mistargeting event conceivably occur at a late stage in the joining process and the likelihood may well be crucial to enforcing fidelity during antigen receptor gene rearrangement.**

A critical process in B- and T-lymphocyte development is the assembly of antigen receptor genes. In this developmentally regulated DNA rearrangement, variable (V), diversity (D), and joining (J) gene segments become connected to one another to create the variable exon of an immunoglobulin (Ig) or T-cell receptor (TCR) gene. V(D)J recombination entails the site-specific recombination of specific DNA motifs termed recombination signal sequences (RSSs) in a process that can be conceptually as well as biochemically divided into two stages: stage 1, where RSS recognition, synapsis, and cleavage takes place, and stage 2, where DNA ends are modified by nucleotide addition and subtraction and become rejoined. Necessary DNA transactions are accomplished through a collaboration between DNA sequence-specific proteins, RAG-1 and RAG-2, and non-sequence-specific nucleases, polymerases, ligases, and structural components (the list includes terminal deoxynucleotidyltransferase, DNA ligase IV, DNA-PKcs, Ku70, Ku80, and XRCC4; reviewed in reference 15). Ultimately, this multicomponent recombination machinery accomplishes the precisely localized cut-and-paste operations that can convert dispersed V, D, and J gene segments into a functional antigen receptor gene.

Central to stage 1 is the site-specific recognition of two RSSs, each comprised of a heptamer (CACAGTG), a spacer of 12 or 23 bp, and a nonamer (ACAAAAACC). Although a consensus RSS is evident from examination of natural joining signals (30, 44), typically only a small minority of RSSs at an Ig or TCR locus exactly match this canonical sequence (for example, see reference 27). The V(D)J recombination machinery therefore is constrained in two opposing ways: it must have sufficient flexibility to recognize naturally occurring RSS variations, but at the same time it must be able to avoid recombination of inappropriate DNA sequences. Such sequences, termed cryptic RSSs, happen to resemble real RSSs, but if joined will promote unintended genome rearrangement.

The recombination machinery is not able to discriminate

between target authentic RSSs and cryptic RSSs with absolute success. This is illustrated by the fact that DNA sequences that match a canonical RSS at only about 50% of the heptamer- or nonamer-equivalent positions will still recombine when tested in an extrachromosomal V(D)J recombination assay (24). The number of cryptic RSSs that can be documented within artificial recombination substrates suggests that a site with functionally relevant similarities to an RSS can be expected to occur at least once every 600 bp in the mammalian genome (24). Even though the intrinsic joining proficiency of most such cryptic RSSs is several orders of magnitude lower than that of an authentic RSS, the number of cryptic sites is overwhelming—fidelity must certainly involve biological strategies that go beyond target site discrimination alone.

Accuracy in V(D)J recombination can be envisioned to rely upon the regulation of the accessibility of genomic sequences to recombination proteins (reviewed in references 35 and 40). Features of chromatin structure are known to allow recombination of appropriate Ig and TCR loci during B- and T-cell differentiation (for recent discussions, see references 8, 13, 45, and 46), and it is reasonable to suppose, by extension, that an inaccessible chromatin configuration may protect every other site in the genome from illegitimate rearrangement. However, the molecular basis of accessibility is only beginning to be elucidated, and how much genome alteration is actually prevented by such a mechanism has not been investigated. Here we have examined an additional, complementary possibility, which is that the recombination enzymes can detect the participation of a cryptic RSS even after cleavage of the mistargeted sequence has already taken place. Postcleavage sequence specificity may provide an important means of error avoidance by allowing for late-stage interventions when mistargeting occurs.

The need for a midcourse correction mechanism is underscored by the observation that site-specific cryptic RSS rearrangement can occur between cryptic signal-like sequences at locations completely removed from any Ig or TCR locus. The threat this type of event poses, and the necessity for safeguards is illustrated for T-cell acute lymphocytic leukemia (T-ALL). In a subset of T-ALL cases, a pair of cryptic RSSs at the *scl/tal* locus have recombined, resulting in the unscheduled expression of the *tal* gene and contributing to leukemogenesis (3, 7,

\* Corresponding author. Mailing address: Program in Genetics and Genomic Biology, Hospital for Sick Children Research Institute, 555 University Ave., Toronto, ON, Canada M5G 1X8. Phone: 416-813-8980. Fax: 416-813-8883. E-mail: susanna@sickkids.on.ca.

9). Cryptic RSSs have also been shown to mediate a less-disastrous site-specific deletion at the *hpri* locus, detectable in T cells sampled from healthy adults (11, 34). A key feature illustrated by these examples is that the site-specific cryptic rearrangements can take place without interfering with the concomitant recombination of the Ig or TCR loci in the affected cell. Thus, this type of aberrant outcome is unlikely to be screened out by any of the highly specialized cellular selective forces operating in T- and B-cell development, all of which focus on the products of antigen receptor gene rearrangement. A cell that harbors a potentially harmful rearrangement would be just as likely as any other to undergo proper recombination at its antigen receptor loci and thereby to receive further proliferative signals.

In this regard, clearly a mechanism that can detect mistargeting even though the joining process has already proceeded as far as RSS cleavage, would help to maintain genome stability during V(D)J recombination. To discover whether such a mechanism exists, we have refined an *in vivo* plasmid assay so that the capacity of a particular RSS to support postcleavage operations can be distinguished from its function in precleavage and cleavage phases of the V(D)J recombination process. By applying this assay to various cryptic RSSs, our results indicate that the events that occur after cleavage in V(D)J recombination are indeed sequence specific. These observations provide *in vivo* support for suggestions based upon *in vitro* evidence that RAG1 and RAG2 (RAG1/2) and its maintenance within a site-specific protein-DNA complex has an important role to play in the later steps of Ig and TCR gene recombination (for example, see references 5, 18, 20, 33, and 37). In the context of fidelity and genome stability, the postcleavage sequence specificity observed here raises the possibility that an interruption of V(D)J recombination and/or reduced cellular survival results as a late-stage consequence of target misrecognition.

#### MATERIALS AND METHODS

Substrates pWT-SJ $\Delta$ , pWT-Inv (designated p12x23 in reference 25), p6131-SJ $\Delta$ , p6131-Inv, pCA-SJ $\Delta$ , pCA-Inv, pT4Anon-SJ $\Delta$ , pT4Anon-Inv, p6131hWTnon-SJ $\Delta$ , p6131hWTnon-Inv, pNOnon-SJ $\Delta$ , and pNOnon-Inv were derived from p23 (25) and differ from one another only in the identity and orientation of the RSS occupying site 1. Twelve-spacer RSSs (12-RSS) were inserted in the form of a double-stranded oligonucleotide cassette with *Sa*I compatible ends at a unique *Sa*I site in the common p23 plasmid backbone. pWT-CJ $\Delta$  was derived from p12 (25) after introducing a canonical 23-spacer RSS (23-RSS) that is identical to the 23-RSS in p23 and its derivatives, at a unique *Bam*HI site. An intermediate plasmid, pEA23flip, was constructed from pJH298 (26) after removal of its 23-RSS and replacement with our 23-RSS used in the desired orientation. p6131-CJ $\Delta$  and pT4Anon-CJ $\Delta$  were then constructed by insertion of the appropriate 12-RSS oligonucleotide cassette at the *Sa*I site of pEA23flip. All modifications were confirmed by dideoxy sequencing (Sequenase). p6131-Invno2 was constructed from pJH298 and p6131-Inv. The two plasmids were double digested with *Sa*I and *Aat*II, and the small fragment of the former was ligated to the large fragment of the latter. p6131-Inv(506) was constructed from p6131-Invno2 by duplication of the 227-bp *Cla*I fragment between the two RSSs; p6131-CJ $\Delta$ (546) was similarly constructed from p6131CJ $\Delta$ . The control site 2 plasmid is p6912 described elsewhere (24) and is identical to the SJ $\Delta$  constructs except that it contains a nonfunctional sequence at site 1. In summary, all of the plasmids used in the present study are identical to one another save only for (i) the identity and orientation of the sequence in site 1, and/or (ii) the orientation of the sequence in site 3, and/or (iii) the presence or absence of a functional RSS at site 2.

**Cell culture.** A pre-B-like cell line, 204-1-8, derived by Abelson virus transformation was used for all experiments (32). Cells were maintained in RPMI-10% heat-inactivated fetal bovine serum-50  $\mu$ M  $\beta$ -mercaptoethanol.

**Extrachromosomal recombination assay.** Transfections of 204-1-8 cells were performed with 150 ng of substrate DNA and included 1 mM caffeine as described previously (24). Transfected DNA was harvested, and recombinant colonies were selected as described earlier. Every construct was transfected 4 to 10 times, and at minimum 500 colonies were picked in grid arrays to nitrocellulose filters. No more than 20% of a given transfection was sampled. For experiments

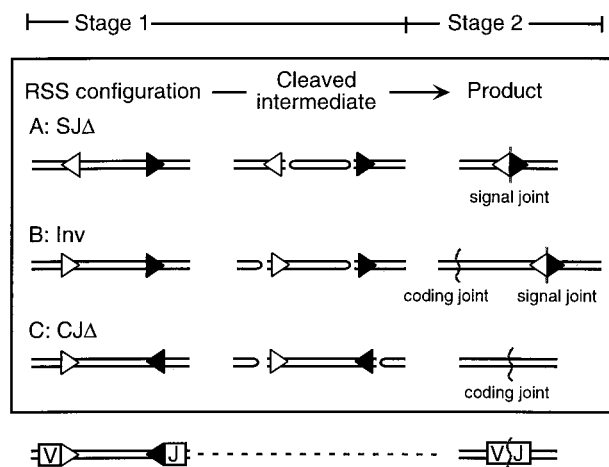


FIG. 1. Configuration-specific differences in recombination outcome. RSS orientation dictates the type of junction that will be recovered after recombination. Three different outcomes, a signal joint via deletion, a signal and coding joint via inversion, and a coding joint via deletion via represent the recovered products resulting from the configurations shown in sections A, B, and C, respectively. Regardless of configuration, stage 1 operations involve cleavage at the heptamers of each RSS (shown as an open triangle for the 12-RSS and as a solid triangle for the 23-RSS). At the end of stage 1, as shown in the middle column ('Cleaved intermediate'), two blunt signal ends and two hairpin-terminated coding ends are formed in every case. Subsequently, during stage 2, there are different operations required for resolution of the coding and/or signal ends. The identity of stage 1 operations and the divergence of stage 2 operations with different configurations forms the basis for distinguishing the postcleavage effects of RSS sequence *in vivo*.

involving cotransfection with the site 2 control plasmid, 75 ng of each plasmid was used.

**Screening of recombinants.** Chloramphenicol-resistant colonies were picked to replicate grids and transferred to nitrocellulose filters for hybridization to <sup>32</sup>P-end-labeled oligonucleotides. The diagnostic hybridization patterns are given in Fig. 2. As shown in the figure, probe A is LACPRIME (CTCATTAGGCAC CCCAGGCT), probe B is LACMER (TATGTTGTGTGGAATTGT), probe C is PLAC20 (TCCTAACAGCTATGACCATG), and probe D is TER-2 (TCCA AAGTCTCAATGCTGC). Probe X represents the RSS-specific oligonucleotides used in the screening of CJ $\Delta$  and Inv substrates: PD12MER (AGGTCGA CACAGTGG) was used for WT, T4Anon, and NOnon; O/6131INV (AGGTC GACACAACAT) was used for 6131 and 6131hWTnon; and O/CAINV (AGG TCGACACACACA) was used for CA. Any colony that gave an ambiguous signal by hybridization was reanalyzed. The accuracy of the recombinant designations for all experiments was tested by subjecting DNA samples from random, typed colonies to diagnostic restriction enzyme analysis; in all cases there was 100% concordance.

In experiments with a cotransfected control site 2 plasmid, the ratio of plasmids within the transfected cells was determined by treating an aliquot of the recovered plasmid DNA with *Dpn*I prior to transformation of *Escherichia coli*. Nonrecombinant colonies were selected on media without chloramphenicol and were randomly picked and typed according to whether or not they hybridized to the oligonucleotide B probe (see Fig. 2). This gives the ratio of successfully transfected (replicated) molecules in the cells, and the number was used to normalize the score obtained for each class of recombinant according to the sample calculation shown in Fig. 4B.

#### RESULTS

**Experimental design.** Stage 2 sequence discrimination is detected by the configuration-specific test illustrated in Fig. 1. This test measures the number of complete recombinants recovered when the orientation of the RSSs (and only this) is varied in an artificial recombination substrate. The configuration-specific test (described in more detail below) avoids complications associated with attempts to quantify cut recombination intermediates *in vivo*. Though seemingly more direct, any measurement of cleaved intermediates must be related in some fashion to rates of formation, degradation, and conversion to

product before being meaningfully interpreted, and this is not possible at present for any in vivo V(D)J recombination system.

In brief, a pair of 12- and 23-RSSs will recombine when present in any of three possible configurations, here designated S $\Delta$ , Inv, and C $\Delta$ . Regardless of configuration, identical stage 1 operations must take place; that is, the engagement of two RSSs and the cleavage of four DNA strands. In all cases, the products of RAG-mediated cleavage are one pair of blunt signal ends and one pair of hairpin-terminated coding ends (Fig. 1). However, after stage 1, each of the three configurations results in a different product (reviewed in references 15 and 22). For the S $\Delta$  configuration, a signal joint is created by simple ligation of two signal ends. For C $\Delta$ , DNA ends with hairpin termini must be opened and variably modified in the process of becoming connected to form a coding joint. For the Inv configuration, coding and signal ends are both joined, and the coding and signal joint products must be formed in a temporally correlated fashion (Fig. 1). If all steps in stage 1 are the same and all products in stage 2 are different, it follows by simple logic that configuration-related differences in the joining proficiency of an RSS should be attributed to the processes that take place after stage 1 and during stage 2.

We have employed an extrachromosomal V(D)J recombination assay to detect configuration-related differences in cryptic RSS recombination, and one caveat to the line of reasoning outlined above is that topological differences in our recombination substrates may create configuration-specific stage 1 effects. By altering the relative orientation of the RSSs, it is possible that physical limitations to DNA flexibility introduce different degrees of strain during stage 1 at the time of synapsis (discussed, for example, in reference 36). For the sake of clarity, a detailed description of the experiments performed to address this possibility has been deferred to the end of the Results section.

The stage 2 competence of a given RSS was determined by inserting the sequence at a position termed site 1 within the plasmid construct. The RSS was tested for joining with a canonical 23-RSS positioned at site 3 (Fig. 2) in up to three different configurations (S $\Delta$ , Inv, and C $\Delta$ ). For standardization, the recombination of an invariant control 12-RSS in site 2 with the 23-RSS in site 3 was also measured in parallel. The ratio of site 1 and site 2 recombinants gave the relative joining proficiency (RJP) for a given construct.

To establish the RJP values for the canonical sequence, three substrates, pWT-S $\Delta$ , pWT-Inv, and pWT-C $\Delta$ , were created bearing a canonical sequence, called WT, in site 1. Each construct was transfected multiple times into the 204-1-8 cell line, which is an Abelson murine leukemia virus-transformed pre-B-like cell active for V(D)J recombination (see Materials and Methods). The RJP of the WT sequence when measured in the S $\Delta$  plasmid was 90 (units are arbitrary and defined by the site 2 control RSS, which is a noncanonical sequence [24, 25]). With pWT-C $\Delta$ , a similar RJP of 62 was obtained. Finally, the RJP for WT in the pWT-Inv plasmid was 5.0, a value much lower than the RJP for WT in either deletional construct (Fig. 3).

The fact that the RJP of a consensus RSS in an inversional configuration is not on par with the values measured for the deletional configuration (S $\Delta$  or C $\Delta$ ) demonstrates, as observed by others, that the joining proficiency of an RSS is configuration sensitive when measured in the extrachromosomal assay (12). Configuration-specific differences in joining proficiency have been attributed to the fact that inversion requires the coordinate formation of two joints, whereas for deletion only one junction need be created. Having established

a canonical standard in each configuration, we can compare the behavior of other, noncanonical DNA sequences tested similarly. That is, with a noncanonical RSS, the RJP for all constructs may be proportionally depressed, in which case there would be no evidence for sequence discrimination during stage 2. Alternatively, the RJP for certain configurations might be disproportionately reduced. The latter situation would indicate that sequence specificity pertains even to the postcleavage phase of V(D)J recombination and would additionally provide clues as to the affected postcleavage process.

**Cryptic RSSs show configuration-specific deficits in RJP.** Two cryptic signals were tested for joining proficiency in each of three configurations (Fig. 2). One cryptic RSS, termed 6131 (24), differs from the WT RSS at multiple positions within both the heptamer and nonamer. In all, 7 of 16 nucleotides in the 6131 sequence are noncanonical (Fig. 3). (The 6131 sequence also served as the control against which the RJP of each test RSS was determined.) A second cryptic RSS, termed CA, was comprised of a CA dinucleotide repeated 14 times. CA repeats are highly represented in both mouse and human genomes, occurring once every 30 kb according to one estimate (41), and are known to possess cryptic signal function in an endogenous chromosomal context as well as by the extrachromosomal assay (11, 24). Compared to a consensus RSS, the pure CA repeat diverges at six positions (Fig. 3), and again these differences are distributed between both the heptamer and the nonamer.

Six constructs (representing each of three configurations for the two cryptic RSSs) were tested in 204-1-8 cells. The numbers of site 1 and site 2 recombinants were determined (details are given in Materials and Methods). As expected, both 6131 and CA were much weaker than the consensus WT sequence, regardless of orientation. For example, in an S $\Delta$ -type construct, 6131 and CA both exhibited RJP of about 0.5, well below the value of 90 obtained for the canonical RSS (pWT-S $\Delta$ ; see Fig. 3).

The key result however was that both 6131 and CA were considerably more compromised for inversion than they were for deletional signal joint formation (Fig. 3). For comparison, the ratios of the RJP obtained with pInv-type plasmids versus the RJP obtained with S $\Delta$ -type plasmids are presented in Fig. 3. Thus, as mentioned above, the WT sequence supports inversion less readily than deletional signal joint formation, giving an Inv/S $\Delta$  ratio of 1:18. If stage 2 operations were unaffected by substitution of a cryptic RSS, we would observe the same Inv/S $\Delta$  ratio for the cryptic sequences despite an overall depression in recombination proficiency. However, the cryptic RSS 6131 gave a much lower ratio of 1:167, and CA gave a ratio of 1:124. Thus, 6131 and CA are excessively impaired in their ability to support inversion.

The diminished Inv/S $\Delta$  ratio for CA or 6131 does not of itself reveal whether it is the nature of the joints resulting from inversion (i.e., a coding joint and a signal joint) or the number of joints that must form (i.e., two instead of one) that becomes problematic after cryptic RSS cleavage (Fig. 1). This question can be explored through testing C $\Delta$  constructs, where coding joints are formed in a one-joint (deletional) rather than the two-joint (inversional) context. An impediment that relates specifically to coding end joining at stage 2 will be revealed if joining proficiency for a cryptic RSS is excessively reduced for the C $\Delta$  configuration.

It is necessary to reverse the orientation of the 23-RSS at site 3 in the C $\Delta$  configuration (Fig. 1C). Consequently, the 23-RSS is in the wrong orientation with respect to the control RSS at site 2 to measure site 2 deletion formation as before (see Fig. 2C). C $\Delta$  plasmids were therefore cotransfected with a separate control plasmid (Fig. 2C). This control plasmid was

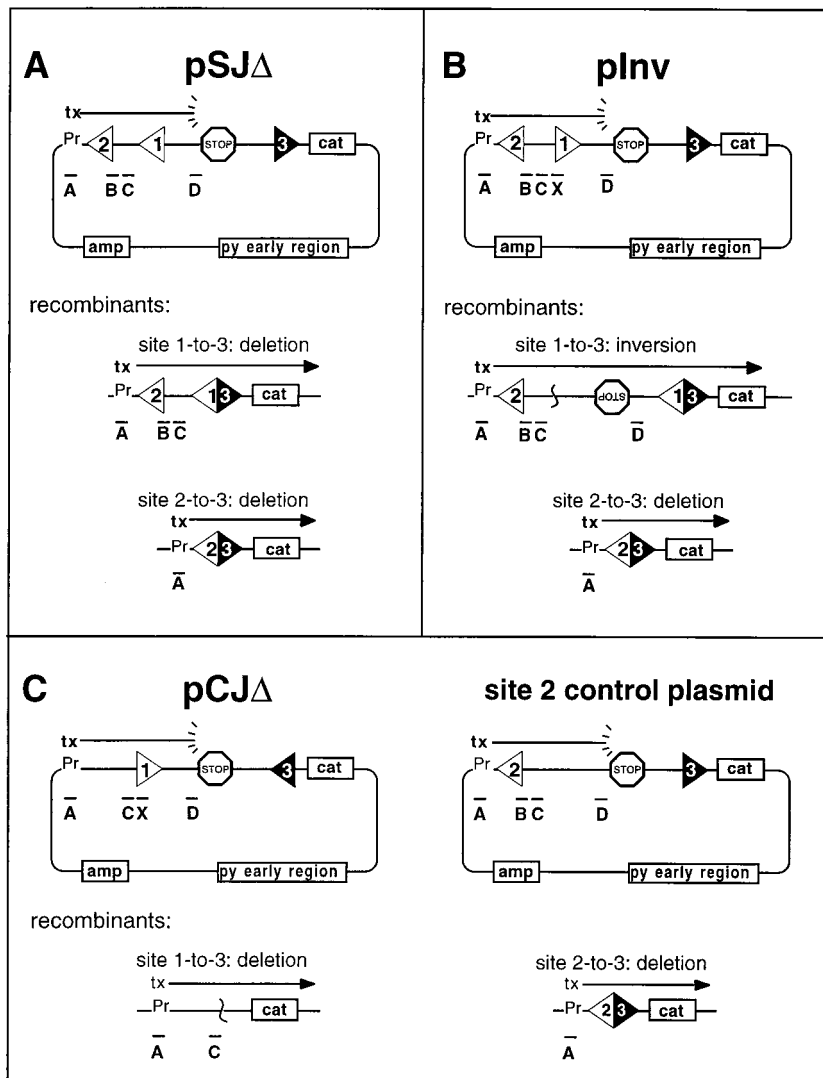


FIG. 2. Recombination substrates. In all substrates, 12-RSSs occur at sites 1 or 2, and the canonical 23-RSS is inserted at site 3. The frequency with which a given test RSS positioned at site 1 recombines is tested relative to the frequency with which the invariant sequence at site 2 is rearranged. In each section of the figure, labeled bars represent the oligonucleotide probes that are used in the typing of recombinants (see Materials and Methods). (A) The SJΔ substrate and two possible recombinant deletion products, involving either the site 1 or site 2 RSS. (B) Inv substrate with the inversional recombinant resulting from site-1-to-site-3 rearrangement and site 2 deletion product. (C) The CJΔ substrate is shown with the deleted coding joint product. This plasmid is cotransfected with the site 2 control plasmid, which gives rise to the site 2 deletion product.

identical to an SJΔ plasmids except that at site 1 the RSS was replaced by an inert sequence. For transfection, the CJΔ plasmid was mixed in a 1:1 ratio with the control site 2 plasmid. The effective ratio (of molecules that successfully reached the nucleus) was determined by assessing the representation of each plasmid among replicated DNA molecules after transfection, as described in Materials and Methods. This result was used to normalize the numbers of site 1 or site 2 recombinants measured in the usual manner by filter hybridization of chloramphenicol-resistant transformants (see Fig. 4B).

Whereas the canonical RSS had given similar RJPs when tested in either the SJΔ or the CJΔ configuration (90 and 62, respectively; Fig. 3), this was not the case for either 6131 or CA. The RJP for p6131-SJΔ was 0.50, while the RJP for p6131-CJΔ was only 0.064 (Fig. 3). Similarly, pCA-SJΔ gave an RJP of 0.46, while pCA-CJΔ gave an RJP of only 0.031. Thus, it was

clear that both of the cryptic signals, 6131 and CA, were particularly ineffective in supporting coding-joint formation.

**The RJP for a given construct is reproducible.** The reproducibility of RJP measurements was good, as shown in Fig. 4A. The construct p6131-SJΔ contains the same sequence (i.e., 6131) at both sites 1 and 2, each being oriented to promote a deletional signal joint formation. Nine separate transfections were performed several years apart by different authors. After analysis by transformation of *E. coli* with DNA from each of nine transfections, it can be seen that the RJP for any single determination is close to the RJP calculated from the pooled numbers (0.50), and the mean for the nine trials was  $0.52 \pm 0.15$ . Based upon the fact that there was little variation in the RJP from one transfection to the next, we have pooled the total number of site 1 and site 2 recombinants to give the RJP for all constructs, as shown in Fig. 3. To assure representative

	Joining proficiency			Stage 2 impact
	site1	site2	RJP	Inv:SJΔ
<b>WT</b>				
CACAGTG-12a-ACAAAAACC				1:18
SJA	1254	14	90	
Inv	972	195	5.0	
CJA	2179	35	62	
<b>6131</b>				
CACAaca-12b-gCAT AAAgt				1:167
SJA	318	632	0.50	
Inv	5	1693	0.0030	
CJA	109	1692	0.064	
<sup>1</sup> SJAno2	345	1020	0.34	
<sup>2</sup> Invno2	8	1358	0.0059	
<sup>3</sup> Inv(506)	0	804	<0.0012	
<sup>4</sup> CJA(546)	66	936	0.071	
<b>CA</b>				
CACAcac-12c-ACAcAcAcAc				1:124
SJA	332	728	0.46	
Inv	3	817	0.0037	
CJA	32	1045	0.031	
<b>T4Anon</b>				
CACAGTG-12a-t Cttttt CC				1:15
SJA	1003	105	9.6	
Inv	276	418	0.66	
CJA	1913	713	2.7	
<b>NOnon</b>				
CACAGTG-12a-tgt c tct ga				1:55
SJA	360	131	2.7	
Inv	19	386	0.049	
<b>6131h/WTnon</b>				
CACAaca-12b-ACAAAAACC				1:150
SJA	494	21	24	
Inv	62	389	0.16	

FIG. 3. Variant RSSs and configuration-specific RJP. The sequence of each tested RSS is given, indicating matches to the canonical RSS in uppercase letters. Spacer sequences were changed according to the dictates of a given experiment. The 12a spacer is an arbitrary sequence as given previously (25). The 12b spacer is identical to the spacer of the RSS appearing in site 2, and the 12c spacer is a continuation of the CA repeat. The total number of recombinants typed as either site 1 or site 2 rearrangements is given for each of the SJA and Inv constructs. For experiments where the site 2 control was introduced on a separate plasmid, the values shown are calculated as described in Materials and Methods (a sample calculation is given in Fig. 4). The RJP is defined as the ratio of site 1 to site 2 recombinants. The ratio of inversion to SJA provides a means by which to compare the stage 2 impact of a different RSS variants. Superscript numbers: 1, cotransfection of p6131-SJAno2 and the site 2 control plasmid (see Fig. 2C); 2, cotransfection of p6131-Invno2 and the site 2 control plasmid; 3, cotransfection of p6131-Inv(506) (in which the distance between sites 1 and 3 was increased from 279 to 506 bp) and the site 2 control plasmid; and 4, cotransfection of p6131-CJA(546) (in which the distance between sites 1 and 3 was increased from 319 to 546 bp) and the site 2 control plasmid.

samplings, no more than one-fifth of a given transfection was transformed into *E. coli*. DNA sequence analysis of coding-joint recombinants indicated that the methods employed were not susceptible to “jackpot” effects (that is, multiple reisolation of a single recombinant due to vector replication in the transfected 204-1-8 cells), even in cases where screening one-fifth of a transfection yielded only a few, rare site 1 recombinants.

**The fine-structure features of the products of cryptic RSS recombination are normal.** Our data suggested that engagement of a cryptic RSS by the V(D)J joining machinery affects the efficiency of coding-joint formation. It was therefore of interest to investigate whether the fine structure of coding joints resulting from cryptic RSS recombination was altered.

Abnormal loss of DNA sequence up to a limit of 111 bp from the site 1-associated coding end and up to 23 nucleotides from the canonical site 3-associated coding end is detected by the present approach. In addition, it is also informative to score P nucleotide insertion, which reflects details of the hairpin opening step (6, 23, 28).

Analysis of 44 recombinant coding joints, incorporating all variant RSSs, is presented in Fig. 5. There were no gross anomalies. Nucleotide loss as well as insertion at the joint were within normal ranges and fully consistent with prior analyses of joints formed in the same cell line with canonical extrachromosomal recombination substrates (21, 24, 25, 28). Coding joints formed from the CA RSS included bits of the CA repeat itself at the junction, but this was not unexpected given that a CA repeat can be targeted in multiple ways, by shifting the recognized sequence by 2-bp increments (see Fig. 5B for examples). DNA sequence analysis of signal joints recovered from pCA-Inv and pCA-SJA were consistent with this interpretation and exhibited recombination sites at the CAC delimiting the 12th or 13th repeat in addition to that appearing at the 14th repeat (data not shown).

Additionally, at least two signal joints were sequenced for each of the tested RSSs. Dozens more from various experiments were assessed by digestion with *Apa*LI, (a restriction enzyme specific for the GTGCAC sequence formed at a precise signal joint). No irregularity with respect to signal joint structure was revealed.

Our analysis indicated that a cryptic RSS causes a depression in coding joint recovery without affecting the properties of the junctions that are observed. The apparent interruption of recombination events involving a cryptic RSS is not circumvented in any obvious way by alternative resolution functions.

**Both the nonamer and heptamer elements of the RSS are recognized during rejoining of cleaved ends.** To explore the importance of the nonamer motif of the RSS in stage 2 operations, we determined configuration-specific RJP values for the NOnon RSS. As its name implies, NOnon possessed a noncanonical substitution at each position of the nonamer. Otherwise, the heptamer (as well as spacer) sequence of NOnon was identical to that of WT. The result, as shown in Fig. 3, was that inversion was excessively depressed for NOnon, (the Inv/SJA ratio was 1:55), indicating that the nonamer portion of the RSS motif is recognized during postcleavage, stage 2, joining operations.

To determine whether stage 2 operations require site-specific interactions with the heptamer as well, we tested an RSS variant, termed 6131h/WTnon (Fig. 3). Here only the three nonamer-proximal nucleotides of the heptamer deviate from the canonical sequence (the heptamer being CACAaca instead of CACAGTG). We found that whereas signal joint formation was only mildly affected by the variant RSS (the RJP was 24 as opposed to the RJP of 90 for a canonical RSS), there was nevertheless a significant stage 2 impact. The inversional construct gave 0.16 for p6131h/WTnon-Inv as opposed to 5.0 for pWT-Inv. In other words, the Inv/SJA ratio was 1:150 for the RSS with a mutant heptamer compared to the 1:18 value for pWT (Fig. 3).

The 6131h/WTnon result demonstrates that a configuration-related depression of joining proficiency is seen upon modification of the heptamer only. The apparent stage 2 sequence specificity is revealed upon introduction of a minimal change in the RSS, for which, as seen for the SJA construct, joining function has not been greatly compromised. Thus, there is no obvious correlation between the strength of a cryptic signal and its stage 2 impact.

Taken together, the results obtained with 6131h/WTnon and

A				B							
p6131SJA				p6131CJA							
Transfection	Site 1	Site 2	Site 1:Site 2 (RJP)	Transfection	Recombinants		Plasmids		Adjusted Counts		Site 1:Site 2 (RJP)
					Site 1 A	Site 2 B	Site1 C	Site 2 D	Site 1 A(C+D) C	Site 2 B(C+D) D	
9449	40	60	0.67	117	0	16	99	33	0	64	0
9450	43	64	0.67	118	10	126	78	54	17	308	0.055
9480	38	76	0.50	119	4	30	86	46	6	86	0.070
94128	39	74	0.53	121	2	27	85	47	3	76	0.039
116	31	63	0.49	122	17	34	118	14	19	321	0.059
125	25	94	0.27	123	14	79	96	36	19	290	0.066
154	39	74	0.53	126	4	35	95	37	6	125	0.048
155	44	65	0.68	127	6	25	97	35	8	94	0.085
156	19	62	0.31	157	8	66	76	56	14	156	0.090
	<b>318</b>	<b>632</b>	<b>0.50</b>	158	10	73	76	56	17	172	0.099
									<b>109</b>	<b>1692</b>	<b>0.064</b>

FIG. 4. Reproducibility of the measured RJP. (A) Nine transfections of p6131-SJA were conducted and screened. The top four transfections were performed 4 years earlier than the bottom five. These numbers correspond to the pooled data reported in Fig. 3. (B) Ten transfections of p6131-CJA cointroduced with the site 2 control plasmid were analyzed. Colonies selected on chloramphenicol were typed to determine the number of site 1 and site 2 recombinants as usual (columns A and B). Colonies isolated on ampicillin after *DpnI* digestion of the harvested DNA were typed to determine the numbers of each parental plasmid successfully transfected as described in Materials and Methods (columns C and D). Colony counts were normalized as shown (two subsequent columns) to give the RJP (rightmost column).

NONon highlight differences between RSS features that have been predicted to have an impact upon stage 1 binding and cleavage and those that are of more consequence in postcleavage, stage 2 operations. Based upon studies that measure cleavage or binding of RSSs by RAG1/2 (2, 10, 19, 31), the nonamer is more critical than the heptamer for binding and, in particular, the nonamer-proximal nucleotides of the heptamer are relatively unimportant for either precleavage binding or cleavage itself.

**Differences between stage 1 and stage 2 sequence specificity are further established with T4Anon.** Previous studies have shown that A-to-T transversions within the A tract of the nonamer result in an RSS that is bound no better than a nonamerless (or randomized-nonamer) RSS by RAG1/2 (2, 10). To determine if A-to-T transversions similarly affect RSS function during stage 2, a final variant, T4Anon, was tested. This sequence was identical to WT except that every A nucleotide within the nonamer had been replaced by a T (i.e., TCT TTTCC). Overall, the T4Anon RSS is noncanonical at a total of 6 of 16 positions.

Oddly enough, the results with T4Anon in the configuration-sensitive assay contrasted with those of every other variant described above. For T4Anon, the Inv/SJA ratio gave no indication of an impairment of stage 2 functions. The ratio was 1:15, a value actually somewhat higher than the Inv/SJA ratio for a canonical RSS and much higher than that for the nonamerless RSS (NONon). T4Anon was also tested in the CJA configuration, where the RJPs for signal joints and coding joints were 9.6 versus 2.7. By these tests, the T4A sequence was found to have only a mild, if any, effect on stage 2 joining operations. This is exceptional and indicates that T-to-A transversions in the nonamer, which as noted above have a large impact on binding and cleavage by RAG1/2 *in vitro*, do not adversely affect stage 2 joining.

**The configuration-specific assay is not sensitive to topological and topographical differences between test plasmids.** In the present study experimental variation is strictly limited to

RSS orientation, stage 2 effects being deduced on the basis of configuration-sensitive effects on joining proficiency (Fig. 1). The distance between the interacting RSSs had been chosen according to the structure of the original plasmid upon which the present substrates are based (17) and is within a range (roughly 250 to 300 bp) for which it is conceivable that synapsis of the two RSSs might be impeded by the difficulty of bending the intervening DNA in certain configurations but not others. Should this be the case, then configuration-sensitive changes in RJP could reflect a need for strong RSS interactions during synapsis (i.e., during stage 1) and not later, as outlined in Fig. 1. Additionally, a related concern is that although the distance between the midpoint of the RSSs in site 1 and that of site 3 are maintained in all substrates without exception, the distances between the points of DNA cleavage (located at the heptamer edge of each RSS) must differ with orientation. Lastly, a feature that might confound the analysis is that the CJA recombinants have to be normalized against a cointroduced control plasmid.

To determine which, if any, of the above factors should be taken into account in the interpretation of the results, several additional experiments were performed. We again measured the RJP for the 6131 RSS in both the SJA and the Inv configurations, having first removed site 2 from the test constructs. The new constructs (p6131-SJA $\Delta$ no2 and p6131-Invno2) were then cotransfected with the site 2 control plasmid, and results were normalized as described for the CJA series of constructs (Fig. 4). The values obtained whether or not the site 2 RSS was located on the same substrate molecule as the test RSS were quite close: where the RJP of p6131-Inv was 0.0030, that of p6131-Invno2 was 0.0059. Similarly, the RJP of p6131-SJA was 0.50, while that of p6131-SJA $\Delta$ no2 was 0.34 (Fig. 3). These results indicated that the relative joining proficiency is roughly the same even if the site 2 control is measured on a separate, cotransfected substrate. Thus, for example, the possibility of secondary rearrangements (between the rearranged 23-RSS in an inversionsal site 1 recombinant and the site 2 RSS) can be

## A

construct	"coding" end (site 1)	N	"coding" end (site 3)
input: Inv	TGCAGGTCGA		GATCCTCTCA
PWT-Inv	TGCAGGTCG TGCAGG		CCTCTCA <u>TCCTCTCA</u>
p6131-Inv	TGCAGG TGCAGG TGCA TGCAGGTCGA TGCAGGTCGA	GG	<u>GATCCTCTCA</u> CCTCTCA ATCCTCTCA TCCTCTCA CCTCTCA
PCA-Inv	TGCAGGT TGCAGGTCGA TGCAGGTCGA	CC ACATA	GATCCTCTCA CTCTCA CCTCTCA
pT4Anon-Inv	TGCAGG TGCAGG TGCAGGTC TGCAGG TGCAGGTC TGCAGG TGCAGG	C CTC	<u>GATCCTCTCA</u> <u>TCCTCTCA</u> ATCCTCTCA <u>TCCTCTCA</u> <u>GATCCTCTCA</u> ATCCTCTCA TCCTCTCA
pNOnon-Inv	TGCAGGTC TGCAGG TGCAGG TGCAGGTCGA	GC C	<u>GATCCTCTCA</u> GATCCTCTCA <u>GATCCTCTCA</u> CA
p6131h/WTnon-Inv	TGCAGG TGCAGG -12 TGCA		<u>TCCTCTCA</u> <u>TCCTCTCA</u> -23 <u>GATCCTCTCA</u>
input: CJA	TGCAGGTCGA		GATCCCCGGG
pWT-CJA	TGCAGG TGCAGG TGCAGGTC TGCAGG TGCAGG	T G AAC	<u>GATCCCCGGG</u> ATCCCCGGG GG CCCCGGG GATCCCCGGG
p6131-CJA	TGCA TGCAGGTCGA TGCAGG		<u>GATCCCCGGG</u> CCCCGGG CCCCGGG
pCA-CJA	TGCAGGTCGACACACA TGCAGGTCGACACACACA TGCAGGTCGACACACA		<u>CCCCGGG</u> GATCCCCGGG <u>CCCCGGG</u>
PT4Anon-CJA	TGCAGGTCGA TGCAGG TGCAGG TGCA		TCCCCCGGG CGGG <u>TCCCCCGGG</u> <u>GATCCCCGGG</u>
p6131-CJA(546)	TGCAGGTC TGC TGCAGG TGCAGG		<u>GATCCCCGGG</u> <u>ATCCCCGGG</u> <u>TCCCCCGGG</u> <u>TCCCCCGGG</u>

## B

GTCGA **CACAcac** ACACACACACAC **ACAcAcAcA**  
 GTCGACA **CACAcac** ACACACACACAC **ACAcAcAtC**  
 GTCGACACA **CACAcac** ACACACACACAC **ACAcATcgA**  
 GTCGACACACA **CACAcac** ACACACACACAC **ACAtcgACC**  
 GTCGACACACACA **CACAcac** ACACACACACAC **AtcgAcctg**

discounted, and it is valid to compare the CJA results with results obtained for other configurations.

Concerns about how the varied inter-RSS distances and the different topologies might affect synapsis were also addressed. If the difficulty of bending covalently closed DNA to bring the two RSSs together is a significant factor in these experiments, configuration-specific joining deficits may actually reflect poor RSS binding during stage 1 and not stage 2. Although differences in DNA flexibility do not cause marked differences in ring closure for DNA fragments in the size range separating the cleavage and joining sites in the present substrates (251 bp for SJA, 319 bp for CJA, and 279 bp for Inv) (38, 39), shortening the intersignal distance still further in V(D)J recombination substrates has been shown to reduce recombination frequencies (36). To determine whether configuration-specific joining deficits are a manifestation of a hindered synapsis at stage 1 caused by close spacing between the RSSs, we looked for an amelioration of the deficits upon increasing the intersignal distance. In particular, for the 6131 RSS, we tested whether nearly doubling the distance between the 12 and 23 signals could improve the RJP for this sequence in either the Inv or the CJA configuration.

The constructs p6131-Inv(506) and p6131-CJA(546) were each cotransfected with the control site 2 plasmid, and the RJP values were determined. Results for both of the larger substrates were fully consistent with those obtained using the original, standard versions (Fig. 3). No site 1 recombinants were found among the 805 recombinants screened for p6131-Inv(506), which compares well to, and is certainly not better than, the 8 site 1 recombinants per 1,358 obtained with the construct for which the intersignal distance was 279 bp (Fig. 3). Reconstruction experiments confirmed that the recombinant form of p6131-Inv should have been detected (data not shown). The plasmid p6131-CJA(546) also had an RJP value like its standard counterpart: 0.071 compared to 0.064. Thus, there was no evidence that precleavage strain due to the different topologies of the variously configured substrates could account for the observed configuration-sensitive depressions in RJP.

## DISCUSSION

Cells in which cryptic RSSs have been recognized and cleaved can threaten survival of an animal, even when the breaks are generated at a very low frequency. This is because the proliferative and highly selective nature of the lymphoid tissue is such that a cell with a preleukemic rearrangement caused by illegitimate V(D)J recombination is especially at risk of becoming clonally expanded. We know that the stringency of target site recognition is not so refined as to eliminate mistargeting altogether (discussed in reference 24), and it is of interest to learn what biological mechanism might exist to recognize and rectify recombination errors once the process has proceeded as far as cleavage.

Two cryptic RSSs, 6131 and CA, were chosen for a detailed analysis. Although neither sequence closely matches a canonical RSS, both sequences are able to function at a low level in

FIG. 5. Coding joints. (A) Sequences of coding joints from randomly selected inversions (Inv) and deletions (CJA). Underlined nucleotides can be assigned to either end. (B) Variant target recognition of a CA repeat. As seen in analysis of coding joints in panel A, as well as of signal joints (not shown), a CA repeat is recognized in alternative frames displaced by 2 bp. Heptamer and nonamer elements are in boldface, and capital letters indicate matches to the canonical sequence.

V(D)J recombination (24). 6131 is a sequence discovered to contribute to an unwanted background in the extrachromosomal assay (17). The other cryptic RSS, CA, corresponds to an abundant dinucleotide repeat in the mouse and human genomes (41) and has been seen to function as a cryptic RSS in the human *hpvt* locus (11). 6131 and CA are representative of millions of naturally occurring RSS-like sequences in the genome with respect to their resemblance to the canonical RSS, being similarly divergent in both the heptamer and the nonamer (24).

To learn whether sequence discrimination occurs not only in the initial binding and cleavage of a V(D)J recombination target (stage 1) but also during the second, rejoining phase of the gene assembly process (stage 2), we tested 6131 and CA for evidence of configuration-sensitive differences in V(D)J joining proficiency. In both cases, the results reveal a defect in carrying out stage 2 operations, indicating that there is indeed sequence recognition late in the joining reaction. These findings extend previous studies that implicate RAG1/2 in postcleavage operations (5, 20, 29, 33, 37, 42) by providing evidence that the protein-DNA interactions are both sequence specific and relevant *in vivo*.

Significantly, despite the fact that by the second stage of V(D)J recombination the coding ends have become unlinked from their RSSs, the CA and 6130 sequences quite clearly exhibited a deficit in coding-joint formation. This raises the logical question as to why operations that effect coding-end joining should depend upon site-specific DNA recognition between recombination proteins and the signal ends.

One likely explanation for postcleavage sequence specificity is a requirement for RAG1/2 to be retained within a protein-DNA complex in order to carry out an enzymatic role in the postcleavage processing of coding ends. The notion that later steps in V(D)J recombination take place in the context of a specific postcleavage complex has had experimental support from several quarters (1, 14, 18, 20) and, although the complex itself (along with all associated components) is not yet biochemically defined, several stable postcleavage complexes containing RAG1/2 and cleaved DNA ends have been described. The various complexes contain RAG1/2, HMG-1 (or -2) and all four ends (18), RAG1/2 and the cut coding and signal end derived from a single RSS (4), and complexes comprised of RAG1/2 together with two cut signal ends (1). The present results are consistent with an *in vivo* role for a postcleavage complex minimally involving RAG protein and both types of cut end (signal end as well as coding end) in the resolution of recombination intermediates.

A more specific suggestion that fits well with our data is that RAG1/2 itself is responsible for opening hairpin coding-end intermediates and/or for further processing of the coding ends during their ligation (5, 33, 37). If the signal ends help to anchor RAG1/2 in the vicinity of the hairpin coding ends through RSS-specific interactions and if RAG1/2 provides essential hairpin DNA processing functions, then it follows that sequence-specific interactions involving the RSS should influence the joining of coding ends even though coding ends have, by this time, become unlinked from the RAG recognition motifs.

Further, characterization of stage 2 sequence specificity has revealed the importance of both heptamer and nonamer elements in postcleavage joint formation. Of particular interest, a marked postcleavage effect was revealed with the RSS-like sequence, 6131H/WTnon, in which the only noncanonical alteration was in the three nonamer-proximal nucleotides of the heptamer (i.e., CACAGTG to CACAaga). Mutational analyses of the nonamer-proximal nucleotides in the heptamer have

shown relatively small effects on the early binding and/or cleavage of RSS-containing DNA by purified RAG1/2 *in vitro* (2, 31). In contrast, these three nucleotides have a clear *in vivo* function during stage 2 of V(D)J joining. It is possible that the interaction in question is fairly weak and/or transient, and it is of particular interest that nonamer-proximal nucleotides of the heptamer have been highlighted as a possible contact site for RAG1/2 in one modification interference study (43).

We suggest that the postcleavage complex serves as a focal point in the control of fidelity in V(D)J recombination. After RSS cleavage, coding ends are excluded from ligation events until the hairpin structure at the DNA ends is opened. For a cryptic RSS, a failure to anchor RAG1/2 in a postcleavage complex through its interactions with the RSS could interfere with efficient hairpin removal. Such a delay in hairpin end opening, perhaps interpreted by cellular control mechanisms as irreparable DNA damage, may trigger an apoptotic response. Given the fact that mistakes in RSS target recognition are not fully avoidable, sequence specificity during stage 2 of V(D)J recombination may provide an essential safeguard against pathogenic rearrangements during pre-B- and pre-T-cell differentiation.

#### ACKNOWLEDGMENTS

We thank Howard Lipshitz and anonymous reviewers for comments on the manuscript.

This work was supported by a grant from the National Cancer Institute of Canada (NCIC). E.A.A. is a recipient of a Doctoral Research Award from the Medical Research Council of Canada. S.M.L. is a Scientist of the NCIC.

#### REFERENCES

- Agrawal, A., and D. B. Schatz. 1997. RAG1 and RAG2 form a stable post-cleavage synaptic complex with DNA containing signal ends in V(D)J recombination. *Cell* **89**:43–53.
- Akamatsu, Y., and M. A. Oettinger. 1998. Distinct roles of RAG1 and RAG2 in binding the V(D)J recombination signal sequences. *Mol. Cell. Biol.* **18**:4670–4678.
- Aplan, P. D., D. P. Lombardi, A. M. Ginsberg, J. Cossman, V. L. Bertness, and I. R. Kirsch. 1990. Disruption of the human SCL locus by "illegitimate" V-(D)-J recombinase activity. *Science* **250**:1426–1429.
- Bailin, T., X. Mo, and M. Sadofsky. 1999. A RAG1 and RAG2 tetramer complex is active in cleavage in V(D)J recombination. *Mol. Cell. Biol.* **19**:4664–4671.
- Besmer, E., J. Mansilla-Soto, S. Cassard, D. J. Sawchuk, G. Brown, M. Sadofsky, S. M. Lewis, M. C. Nussenzweig, and P. Cortes. 1998. Hairpin coding end opening is mediated by the recombination activating genes RAG1 and RAG2. *Mol. Cell* **2**:817–828.
- Binnie, A., S. Olson, G. E. Wu, and S. M. Lewis. 1999. Gamma-irradiation directly affects the formation of coding joints in SCID cell lines. *J. Immunol.* **163**:5418–5426.
- Brown, L., J.-T. Cheng, Q. Chen, M. J. Siciliano, W. Crist, G. Buchanan, and R. Baer. 1990. Site-specific recombination of the *tal-1* gene is a common occurrence in human T cell leukemia. *EMBO J.* **9**:3343–3351.
- Cherry, S. R., and D. Baltimore. 1999. Chromatin remodeling directly activates V(D)J recombination. *Proc. Natl. Acad. Sci. USA* **96**:10788–10793.
- Chervinsky, D. S., X. F. Zhao, D. H. Lam, M. Ellsworth, K. W. Gross, and P. D. Aplan. 1999. Disordered T-cell development and T-cell malignancies in SCL LMO1 double-transgenic mice: parallels with E2A-deficient mice. *Mol. Cell. Biol.* **19**:5025–5035.
- Difilippantonio, M. J., C. J. McMahon, Q. M. Eastman, E. Spanopoulou, and D. G. Schatz. 1996. RAG1 mediates signal sequence recognition and recruitment of RAG2 in V(D)J recombination. *Cell* **87**:253–262.
- Fuscoe, J. C., L. J. Zimmerman, M. J. Lippert, J. A. Nicklas, J. P. O'Neill, and R. J. Albertini. 1991. V(D)J recombinase-like activity mediates *hpvt* gene deletion in human fetal T-lymphocytes. *Cancer Res.* **51**:6001–6005.
- Gauss, G. H., and M. R. Lieber. 1992. The basis for the mechanistic bias for deletional over inversional V(D)J recombination. *Genes Dev.* **6**:1553–1561.
- Golding, A., S. Chandler, E. Ballestar, A. P. Wolffe, and M. S. Schlissel. 1999. Nucleosome structure completely inhibits *in vitro* cleavage by the V(D)J recombinase. *EMBO J.* **18**:3712–3723.
- Grawunder, U., and M. R. Lieber. 1997. A complex of RAG-1 and RAG-2 proteins persists on DNA after single-strand cleavage at V(D)J recombination signal sequences. *Nucleic Acids Res.* **7**:1375–1382.



15. Grawunder, U., R. B. West, and M. R. Lieber. 1998. Antigen receptor gene rearrangement. *Curr. Opin. Immunol.* **10**:172–180.
16. Hammarsten, O., L. G. DeFazio, and G. Chu. 2000. Activation of DNA-dependent protein kinase by single-stranded DNA ends. *J. Biol. Chem.* **275**:1541–1550.
17. Hesse, J. E., M. R. Lieber, K. Mizuuchi, and M. Gellert. 1989. V(D)J recombination: a functional definition of the joining signals. *Genes Dev.* **3**:1053–1061.
18. Hiom, K., and M. Gellert. 1998. Assembly of a 12/23 paired signal complex: a critical control point in V(D)J recombination. *Mol. Cell* **1**:1011–1019.
19. Hiom, K., and M. Gellert. 1997. A stable RAG1-RAG2-DNA complex that is active in V(D)J cleavage. *Cell* **88**:65–72.
20. Leu, R. M. J., Q. M. Eastman, and D. G. Schatz. 1997. Coding joint formation in a cell-free V(D)J recombination system. *Immunity* **7**:303–314.
21. Lewis, S., J. E. Hesse, K. Mizuuchi, and M. Gellert. 1988. Novel strand exchanges in V(D)J recombination. *Cell* **55**:1099–1107.
22. Lewis, S. M. 1994. The mechanism of V(D)J joining: lessons from molecular, immunological, and comparative analyses. *Adv. Immunol.* **56**:27–150.
23. Lewis, S. M. 1994. P nucleotide insertions and the resolution of hairpin DNA structures in mammalian cells. *Proc. Natl. Acad. Sci. USA* **91**:1332–1336.
24. Lewis, S. M., E. Agard, S. Suh, and L. Czyzyk. 1997. Cryptic signals and the fidelity of V(D)J joining. *Mol. Cell. Biol.* **17**:3125–3136.
25. Lewis, S. M., and J. E. Hesse. 1991. Cutting and closing without recombination in V(D)J joining. *EMBO J.* **10**:3631–3639.
26. Lieber, M. R., J. E. Hesse, S. Lewis, G. C. Bosma, N. R. Rosenberg, K. Mizuuchi, M. J. Bosma, and M. Gellert. 1988. The defect in murine severe combined immune deficiency: Joining of signal sequences but not coding segments in V(D)J recombination. *Cell* **55**:7–16.
27. Matsuda, F., K. Ishii, P. Bourvagnet, K. Kuma, H. Hayashida, T. Miyata, and T. Honjo. 1998. The complete nucleotide sequence of the human immunoglobulin heavy chain variable region locus. *J. Exp. Med.* **188**:2151–2162.
28. Meier, J. T., and S. M. Lewis. 1993. P nucleotides in V(D)J recombination: a fine-structure analysis. *Mol. Cell. Biol.* **13**:1078–1092.
29. Ramsden, D. A., T. T. Paull, and M. Gellert. 1997. Cell-free V(D)J recombination. *Nature* **388**:488–491.
30. Ramsden, D. A., K. Baetz, and G. E. Wu. 1994. Conservation of sequence in recombination signal sequence spacers. *Nucleic Acids Res.* **22**:1785–1796.
31. Ramsden, D. A., J. F. McBlane, D. C. V. Gent, and M. Gellert. 1996. Distinct DNA sequence and structure requirements for the two steps of V(D)J recombination signal cleavage. *EMBO J.* **15**:3197–3206.
32. Rosenberg, N., and D. Baltimore. 1976. A quantitative assay for transformation of bone marrow cells by Abelson murine leukemia virus. *J. Exp. Med.* **143**:1453–1463.
33. Santagata, S., E. Besmer, A. Villa, F. Bozzi, J. Allingham, C. Sobacchi, D. B. Haniford, P. Vezzoni, M. C. Nussenzweig, Z.-Q. Pan, and P. Cortes. 1999. The RAG1/RAG2 complex constitutes a 3' flap endonuclease: implications for functional diversity in V(D)J and transpositional recombination. *Mol. Cell* **4**:935–947.
34. Scheerer, J. B., L. Xi, G. W. Knapp, R. W. Setzer, W. L. Bigbee, and J. C. Fuscoe. 1999. Quantification of illegitimate V(D)J recombinase-mediated mutations in lymphocytes of newborns and adults. *Mutat. Res.* **431**:291–303.
35. Schlissel, M. S., and P. Stanhope-Baker. 1997. Accessibility and the developmental regulation of V(D)J recombination. *Semin. Immunol.* **9**:161–170.
36. Sheehan, K. M., and M. R. Lieber. 1993. V(D)J recombination: signal and coding joint resolution are uncoupled and depend on parallel synapsis of the sites. *Mol. Cell. Biol.* **13**:1363–1370.
37. Shockett, P. E., and D. G. Schatz. 1999. DNA hairpin opening mediated by the RAG1 and RAG2 proteins. *Mol. Cell. Biol.* **19**:4159–4166.
38. Shore, D., and R. L. Baldwin. 1983. Energetics of DNA twisting. II. Topoisomer analysis. *J. Mol. Biol.* **170**:983–1007.
39. Shore, D., J. Langowski, and R. L. Baldwin. 1981. DNA flexibility studied by covalent closure of short fragments into circles. *Proc. Natl. Acad. Sci. USA* **78**:4833–4837.
40. Sleckman, B. P., J. R. Gorman, and F. W. Alt. 1996. Accessibility control of antigen-receptor variable-region gene assembly: role of *cis*-acting elements. *Annu. Rev. Immunol.* **14**:459–481.
41. Smider, V., W. K. Rathmell, G. Brown, S. Lewis, and G. Chu. 1998. Failure of hairpin-ended and nicked DNA to activate DNA-dependent protein kinase: implications for V(D)J recombination. *Mol. Cell. Biol.* **18**:6853–6858.
42. Stallings, R. L., A. F. Ford, D. Nelson, D. C. Torney, C. E. Hildebrand, and R. K. Moyzis. 1991. Evolution and distribution of (GT)<sub>n</sub> repetitive sequences in mammalian genomes. *Genomics* **10**:807–815.
43. Steen, S. B., J. O. Han, C. Mundy, M. A. Oettinger, and D. B. Roth. 1999. Roles of the “dispensable” portions of RAG-1 and RAG-2 in V(D)J recombination. *Mol. Cell. Biol.* **19**:3010–3017.
44. Swanson, P. C., and S. Desiderio. 1998. V(D)J recombination signal recognition: Distinct, overlapping DNA-protein contacts in complexes containing RAG1 with and without RAG2. *Immunity* **9**:115–125.
45. Tonegawa, S. 1983. Somatic generation of antibody diversity. *Nature* **302**:575–581.
46. Whitehurst, C. E., S. Chattopadhyay, and J. Chen. 1999. Control of V(D)J recombinational accessibility of the D beta 1 gene segment at the TCR beta locus by a germline promoter. *Immunity* **10**:313–322.
47. Zhong, X. P., J. Carabana, and M. S. Krangel. 1999. Flanking nuclear matrix attachment regions synergize with the T cell receptor delta enhancer to promote V(D)J recombination. *Proc. Natl. Acad. Sci. USA* **96**:11970–11975.