

V(D)J Recombination: In Vitro Coding Joint Formation

FRANCES WEIS-GARCIA,¹ EVA BESMER,² DENNIS J. SAWCHUK,¹ WONG YU,¹ YUN HU,¹
SYLVANIE CASSARD,¹ MICHEL C. NUSSENZWEIG,² AND PATRICIA CORTES^{1*}

*Laboratory of Molecular Immunology¹ and Howard Hughes Medical Institute Research Laboratory,²
The Rockefeller University, New York, New York 10021*

Received 13 May 1997/Returned for modification 6 July 1997/Accepted 18 August 1997

Antigen receptor genes are assembled through a mechanism known as V(D)J recombination, which involves two different joining reactions: signal and coding joining. Formation of these joints is essential for antigen receptor assembly as well as maintaining chromosomal integrity. Here we report on a cell-free system for coding joint formation using deletion and inversion recombination substrates. In vitro coding joint formation requires RAG1, RAG2, and heat-labile factors present in the nuclear extract of nonlymphoid cells. Both inversion- and deletion-mediated coding joint reactions produce diverse coding joints, with deletions and P nucleotide addition. We also show that deletion-mediated coding joint formation follows the 12/23 rule and requires the catalytic subunit of DNA-dependent protein kinase.

B and T lymphocytes produce diverse antigen receptors utilizing the V(D)J recombination reaction. The DNA sequence requirements for V(D)J recombination consist of highly conserved heptamer and nonamer DNA motifs (recombination signal sequences [RSSs]) separated by a spacer of 12 or 23 bp (12 RSS and 23 RSS), which flank recombining regions of both the immunoglobulin and T-cell receptor loci. Efficient recombination occurs almost exclusively between RSSs with different spacer lengths. This restriction, known as the 12/23 rule (25, 49), has been shown to be regulated during the cleavage step of V(D)J recombination (7, 46, 52). Orientation of the RSSs that are to be recombined (head-to-head or head-to-tail orientation) determines whether rearrangement follows a deletional or inversional mechanism (25).

The first step of the V(D)J recombination reaction involves specific recognition and cleavage at the RSSs by RAG1 and RAG2 (15, 28, 50). Competition assays have suggested that the nonamer motif may play a fundamental role in sequence-specific recognition (4, 34), and direct binding of RAG1 to the nonamer motif has been demonstrated (5, 44). Interestingly, this binding activity was mapped to a domain which shows significant homology to the DNA binding domain of the Hin family of bacterial invertases (5, 44). A stable complex of RAG1 and RAG2 with a RSS has recently been isolated, suggesting that RAG1 and RAG2 may function as a complex during V(D)J recombination. This conclusion is supported by immunoprecipitation studies (23, 45).

Cleavage occurs at the heptamer/coding border in a characteristic manner. First, RAG1 and/or RAG2 introduce a nick at the heptamer/coding border, followed by a nucleophilic attack by the free hydroxyl group on the bottom strand of DNA, resulting in the generation of a double-strand break at the signal end and a hairpin at the coding end (28). These two types of ends are generated through a one-step transesterification reaction similar to that used by both Mu transposase and human immunodeficiency virus integrase (51). The hairpin coding ends produced are joined imprecisely to form a coding joint, and the blunt 5' phosphorylated signal ends are also

joined (in a precise manner) to form signal joints. The imprecision with which coding ends are processed and joined serves to add further to antigen receptor diversity.

Several proteins that mediate V(D)J recombination in vivo have been identified. The lymphoid-specific components of the recombination machinery are RAG1, RAG2, and terminal deoxynucleotidyltransferase (TdT) (11, 20, 21, 31, 43). The non-lymphoid-restricted components have not all been identified but include a number of proteins that are involved in DNA double-strand break repair. Tissue culture cell lines deficient in Ku 80, the XRCC4 protein, or the catalytic subunit of the DNA-dependent protein kinase (DNA-PK_{cs}) are unable to repair DNA double-strand breaks and fail to mediate V(D)J recombination in transfection assays (16, 18, 26, 33, 36, 38, 47, 48). Furthermore, mice defective in DNA-PK_{cs} or deficient in the p80 subunit of Ku possess a severely immunodeficient phenotype (12, 19, 30, 55). All of these non-lymphoid-specific components are likely to participate in the joining step of the reaction, but their specific architectural and catalytic roles remain unclear.

Despite rapid progress in understanding the recognition and cleavage steps of the V(D)J recombination reaction, the molecular mechanisms that govern joining of the ends remain largely obscure. In addition to the genetically defined non-lymphoid-specific factors, several other activities, including hairpin opening, polymerase, 3'-5' exonuclease, and ligase activities, are predicted to be necessary to complete the recombination reaction. Identification of all of the components required for V(D)J recombination and a complete understanding of the V(D)J joining mechanism will require in vitro reconstitution of the reaction. As an important step toward reaching this goal, we have developed a cell-free system that mediates coding joint formation in both deletion and inversion substrates.

MATERIALS AND METHODS

Purification of truncated GST-RAG1 and GST-RAG2 fusion proteins. Truncated versions of RAG1 (amino acids 330 to 1040) and RAG2 (amino acids 1 to 383) were expressed as glutathione S-transferase (GST) fusion proteins under the transcriptional control of the elongation factor 1 α promoter (29). 293T cells (293 cells expressing the simian virus 40 large T antigen) were transiently transfected with both RAG constructs by calcium phosphate precipitation (32). Two days after transfection, the cells were harvested and the proteins were purified as described previously (40). RAG1 and RAG2 proteins were expressed at levels corresponding to approximately 1 to 2 μ g of RAG1 and RAG2 from one

* Corresponding author. Mailing address: The Rockefeller University, 1230 York Ave., Box 220, New York City, NY 10021. Phone: (212) 327-8069. Fax: (212) 327-8370. E-mail: cortes@rockvax.rockefeller.edu.

100-mm-tissue culture dish. Although RAG1 and RAG2 were highly purified (as determined by Coomassie blue staining [data not shown]), we cannot rule out the possibility that our preparations contained additional factors that contributed to the biochemical activities described.

HeLa cell nuclear extract preparation and fractionation. Nuclear extracts were prepared by the Dignam protocol (6). Fifty milliliters of nuclear extract was loaded onto a P11 ion-exchange column, and step elutions with 0.1, 0.3, 0.5, and 0.85 M KCl were performed. Fractions containing the peak protein content were subsequently pooled and dialyzed against low-salt buffer C containing 100 mM KCl as described previously (10).

In vitro recombination. Fifty nanograms of pJH200 or pJH299 was incubated with 1 μ l of copurified RAG proteins (approximately 50 ng of RAG1 and 100 ng of RAG2 per μ l) and 1 μ l of HeLa cell nuclear extracts from P11 fractions (0.85 M fraction = 0.68 μ g/ μ l) in the presence of 25 mM HEPES (pH 7.5), 12.5 mM Tris-HCl (pH 8.0), 150 mM sodium acetate, 25 mM NaCl, 10 mM magnesium acetate, 1 mM EGTA, 10% glycerol, and 1 mM dithiothreitol. The total reaction volume equaled 20 μ l. After a 5-h incubation at 37°C, the samples were digested with proteinase K for 1 h at 55°C and then extracted with phenol-chloroform. The DNA was recovered by ethanol precipitation, using 3 μ g of poly(dI-dC) (Boehringer Mannheim) as the carrier, and resuspended in 20 μ l of 0.1 \times Tris-EDTA. Five percent of the recovered DNA was used as the template in the following PCR detection assay.

PCR analysis. pJH200 recombined DNA was amplified on a Perkin-Elmer 9600 apparatus, using 30 cycles of 94°C for 5 s, 60°C for 15 s, and 72°C for 15 s. The 20- μ l PCR mixture contained 5 mM Tris-HCl (pH 9.0), 25 mM KCl, 2.5 mM MgCl₂, 5 ng of each primer (O2 [GGC AAC CGA GCG TTC TGA AC] and R3 [GAG AAT CGC AGC AAC TTG TCG]), 10 μ M deoxynucleoside triphosphates (dNTPs), 0.5 μ Ci of [α -³²P]dTTP (3,000 Ci/mmol; Amersham), and 0.5 U of *Taq* DNA polymerase (Boehringer Mannheim). Recombined pJH299 was amplified with primers R3 and RA2 (GGA ATT GTG AGC GGA TAA CAA TTT CAC AC) under similar reaction conditions, using 30 cycles of 94°C for 5 s, 60°C for 15 s, and 72°C for 15 s. ³²P-labeled products were analyzed by electrophoresis in 6% denaturing polyacrylamide-1 \times Tris-borate-EDTA (TBE) gels. As a loading control, a 247-bp fragment from the chloramphenicol acetyltransferase (*CAT*) gene (present both in pJH200 and pJH299) was amplified by using primers RA1 (TCA CTG GAT ATA CCA CCG TTG ATA TAT CC) and RA6 (CTA TCC CAT ATC ACC AGC TCA CCG) in a two-step PCR (30 cycles of 94°C for 10 s and 72°C for 30 s). Reaction conditions were as described above. The template DNA was first titrated, and 1 to 10 pg of template was found to be in the linear range for amplification; therefore, 5 pg of DNA was used as the template for amplification in each control reaction. The 247-bp ³²P-labeled fragment was resolved by electrophoresis in an 8% polyacrylamide-1 \times TBE gel.

In vivo recombination. 293T cells (a transformed human embryonic kidney cell line) were transiently transfected at 2 million cells/plate with 2 μ g of the recombination substrate, either alone or together with GST-truncated RAG1 and GST-truncated RAG2 constructs (2 μ g of each), by calcium phosphate precipitation (32). Plasmid DNA was recovered from the cells 2 days later by a standard alkaline lysis protocol.

Subcloning and sequencing of coding joints. In vitro and in vivo recombined pJH200 12/23 constructs were individually amplified on a Perkin-Elmer 9600 apparatus, using 35 cycles of 94°C for 5 s, 65°C for 10 s, and 72°C for 20 s. Each 100- μ l PCR mixture contained 3 μ l of the sample, 10 mM Tris-HCl (pH 9.0), 50 mM KCl, 5 mM MgCl₂, 20 μ M dNTPs, 100 ng of each primer (O2NEcoRI [CCG GAA TTC CGT TCT GAA CAA ATC CAG ATG G] and R3NXhoI [CCG CTC GAG CGC CAA TCG AGC CAT GTC G]), and 3.75 U of *Taq* DNA polymerase (Boehringer Mannheim). Following amplification, the DNA was purified, digested for 2 h with *Eco*RI and *Xho*I, and resolved on a 2% agarose gel. PCR fragments were gel purified with the Qiaquick kit (Qiagen), ligated into pBluescript (Stratagene), transformed, and miniprep by using the Qiaprep Spin Miniprep kit (Qiagen); 5 μ l of each miniprep was prepared for sequencing with a Dye Terminator Cycle Sequencing kit (PE Applied Biosystems). The sequencing reactions were run and analyzed on a Perkin-Elmer/Applied Biosystems 310 Genetic Analyzer. In vitro recombined pJH299 (27) was amplified with primers RA2 and R3NXhoI, using identical reaction conditions and 30 cycles of 94°C for 5 s, 60°C for 10 s, and 72°C for 20 s. The PCR-amplified products were gel purified and subcloned by using Topo TA cloning (Invitrogen). DNA preparation and sequencing were as indicated for pJH200.

RESULTS AND DISCUSSION

To measure coding joint formation, we used the deletion substrate pJH200 (13) and detected recombination with a sensitive PCR assay (37). Upon V(D)J recombination, two unequal products are generated: a large plasmid which contains the signal joint, and a small DNA circle carrying the coding joint (Fig. 1A). Primers R3 and O2 were used to amplify fragments of approximately 190 bp containing the coding joints. In the nonrecombined substrate, these primers are too far apart to allow any product to be amplified. As a control for

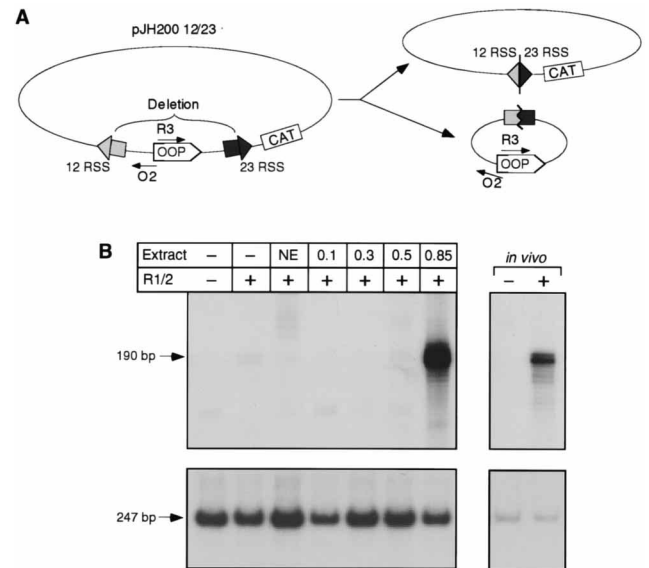


FIG. 1. In vitro coding joint formation by deletion requires RAG1, RAG2, and a single fraction from HeLa nuclear extract. (A) Schematic representation of the exogenous deletion substrate pJH200 12/23 (13). The products generated upon V(D)J recombination and the primers (R3 and O2) used to PCR amplify the coding joint are shown. (B) pJH200 12/23 was incubated with the indicated combinations of copurified RAG1 and RAG2 (R1/2), HeLa cell nuclear extract (NE), or various salt elutions (0.1 M to 0.85 M) of HeLa cell nuclear extract eluted from a P11 column. DNA was purified and amplified in the presence of [³²P]dTTP, and the ³²P-labeled PCR products were then resolved on a 6% denaturing polyacrylamide gel. The right panel shows the ³²P-labeled products amplified in an identical manner from in vivo recombined pJH200 12/23; the bottom panel shows a loading control in which primers RA1 and RA6 were used to amplify a 247-bp fragment from the *CAT* gene present in pJH200.

coding joint formation, pJH200 was transiently transfected into 293T cells alone or along with RAG1 and RAG2. A fragment of the *CAT* gene present in all the recombination substrates was amplified and used as a DNA loading control.

Neither RAG1 nor RAG2, alone or in combination, was sufficient to allow recombination of pJH200 in vitro (Fig. 1B and data not shown). To determine whether non-lymphoid-restricted factors present in nuclear extracts could complement the purified RAG proteins, we added HeLa nuclear extract to the reaction mixture. Unfractionated HeLa nuclear extract was unable to complement the RAGs in generating coding joints (Fig. 1B). Since inhibitory factors present in crude extracts are known to mask other complex in vitro reactions, we assayed HeLa nuclear extract that had been fractionated by chromatography on phosphocellulose. Complementation of RAGs with different phosphocellulose fractions from HeLa nuclear extract resulted in coding joint formation being observed only when the 0.85 M KCl fraction was present (Fig. 1B). When different fractions were combined and used to complement RAGs, none of the combinations exhibited a level of activity higher than that of the 0.85 M fraction alone (data not shown). Furthermore, addition of the 0.1 M flowthrough to the 0.85 M fraction dramatically inhibited the reaction (data not shown), demonstrating the existence of factors that are capable of blocking coding joint formation. In similar complementation and fractionation experiments performed with 293T nuclear extract, the levels of coding joint formation activity were lower than those observed with fractionated HeLa extract (data not shown). Based on these initial experiments, only the 0.85 M phosphocellulose fraction was used for further experimentation. Interestingly, low levels of coding joint formation were

A *In vivo coding joints*

	Coding seq	12 RSS	23 RSS	Coding seq
	CCTGCTGAACCTG	<i>CACAGTG</i>	<i>CACTGTG</i>	CAGGTCTCCAGTA
55	CCTGCTGAACCTG	CAGGTCTCCAGTA
02	CCTGCTGAACCTGNCN.....GTCTCCAGTA
01	CCTGCTGAACCTGCTCCAGTA
02	CCTGCTGAACCTG	<i>C</i>TC.....	CAGGTCTCCAGTA
03	CCTGCTGAACCTG	<i>C</i>AGTA
01	CCTGCTGAACCTG	<i>CA</i> Δ 13
01	CCTGCTGAACCT.GGTCTCCAGTA
01	CCTGCTGAACCT.GTCTCCAGTA
02	CCTGCTGAACCC..AGGTCTCCAGTA
03	CCTGCTGAACCC..TCCAGTA
01	CCTGCTGAACCC..CAGTA
01	CCTGCTGAAC..	CAGGTCTCCAGTA
05	CCTGCTGAAC..AGGTCTCCAGTA
01	CCTGCTGAAC..GGTCTCCAGTA
01	CCTGCTGAAC..GTGCNC.....GTCTCCAGTA
03	CCTGCTGAAC..CTCCAGTA
01	CCTGCTGAA..GGTCTCCAGTA
01	CCTGCTGA..	CAGGTCTCCAGTA
02	CCTGCTGA..GGTCTCCAGTA
01	CCTGCT..	CAGGTCTCCAGTA
02	CCTGC.....	CAGGTCTCCAGTA
01	CCT.....	CAGGTCTCCAGTA
01	C.....AC.....GGTCTCCAGTA
01	Δ 15..... Δ 27
93	Total clones			

B *In vitro coding joints*

	Coding seq	12 RSS	23 RSS	Coding seq
	CCTGCTGAACCTG	<i>CACAGTG</i>	<i>CACTGTG</i>	CAGGTCTCCAGTA
29	CCTGCTGAACCTG	CAGGTCTCCAGTA
01	CCTGCTGAACCTG	CAGGTCTCCAGTA
01	CCTGCTGAACCTG <i>ACCTG</i>	CAGGTCTCCAGTA
01	CCTGCTGAACCTG <i>GAGACCTG</i>	CAGGTCTCCAGTA
01	CCTGCTGAACCTG	<i>C</i>CTCCAGTA
09	CCTGCTGAACCTG	<i>CA</i>	CAGGTCTCCAGTA
01	CCTGCTGAACCTG	<i>CA</i>T.....GTCTCCAGTA
01	CCTGCTGAACCTG	<i>CAC</i>GGTCTCCAGTA
01	CCTGCTGAACCTG	<i>CACAG</i> <i>TG</i>	CAGGTCTCCAGTA
01	CCTGCTGAACCTG	<i>CACAGTG</i>	CAGGTCTCCAGTA
01	CCTGCTGAACCTG	<i>CACAGTG+3</i>	CAGGTCTCCAGTA
01	CCTGCTGAACCTG	<i>CACAGTG+4</i>	CAGGTCTCCAGTA
01	CCTGCTGAACCTG	<i>CACAGTG+9</i> <i>CACTGTG</i>	CAGGTCTCCAGTA
01	CCTGCTGAACCTG	<i>CACAGTG+15</i> .GAA.....	CAGGTCTCCAGTA
01	CCTGCTGAACCTG	<i>CACAGTG+15</i> .G.....	CAGGTCTCCAGTA
01	CCTGCTGAACCTGCAGTA
01	CCTGCTGAACCT.	CAGGTCTCCAGTA
02	CCTGCTGAACCT.GGTCTCCAGTA
01	CCTGCTGAACCC..	CAGGTCTCCAGTA
01	CCTGCTGAACCC..AGGTCTCCAGTA
03	CCTGCTGAACCC..CAGTA
01	CCTGCTGAAC..TCCAGTA
01	CCTGCTGAAC..CTCCAGTA
01	CCTGCTGAAC..TCTCCAGTA
01	Δ 16..... <i>+1CACTGTG</i>	CAGGTCTCCAGTA
01	Δ 22.....CTCCAGTA
02	Δ 26.....AGGTCTCCAGTA
03	Δ 27.....	CAGGTCTCCAGTA
01	Δ 36.....	CAGGTCTCCAGTA
01	Δ 58.....GTCTCCAGTA
04	Δ 81..... <i>+46CACTGTG</i>	CAGGTCTCCAGTA
81	Total clones			

observed with some preparations of copurified RAG proteins (Fig. 1B). Since RAG proteins were purified from 293T cells, it is possible that cellular factors important for V(D)J recombination can be copurified with RAGs.

To verify that the 190-bp PCR product amplified from the *in vivo* and *in vitro* reactions represented authentic coding joints, we subcloned the PCR fragments and sequenced the putative coding junctions. Coding joint products from five independent *in vivo* and five independent *in vitro* reactions were analyzed (Fig. 2). Coding joints produced *in vivo* and *in vitro* were classified into five partially overlapping categories as follows: (i) precise, (ii) deletions, (iii) P elements, (iv) N nucleotides, and (v) anomalous. Due to the sequence at the coding ends in pJH200, the number of precise joints may be overestimated because the presence of a P element of 1 to 5 bp could be masked by a deletion of equal length at the opposite coding end. In addition, proper cleavage at the heptamer border followed by resolution of the hairpin into a P element of 1 or 2 bp would result in a sequence identical to aberrant cleavage within the heptamer. Thus, some of the joints scored as precise may in fact represent P elements, and some of the joints scored as 1- to 2-bp P elements may be due to aberrant cleavage.

When *in vivo* joining reaction products were amplified, cloned, and sequenced (Fig. 2A), we found that 59% of the joints from transfected cells were precise, 39% showed deletions, 12% contained P elements, and 6% had N nucleotides (Table 1). The level of precise joints seen in our transfection system is high but not inconsistent with previous reports. The level of coding joint diversity reported by other groups using different cell lines varies from 1% of precise joints up to 41% (8, 17). This high degree of variation may in part be accounted for by intrinsic differences between the cell lines used. Finally, the fact that 6% of the coding junctions produced in our transfected cells contained N nucleotides is also in agreement with published *in vivo* studies (42, 48). However, these N nucleotides must have been incorporated by a TdT-independent mechanism, since 293T cells do not express this enzyme (8).

Coding joints produced in the *in vitro* reaction resemble the *in vivo* controls in that 36% were precise, 40% exhibited deletions, 28% had P elements, and 7% possessed one to two N nucleotides (Fig. 2B and Table 1). In addition to these standard features, another element was observed. Seventeen percent of the coding joints were anomalous and represent the joining of products presumably cleaved at various positions inside the 12 or 23 RSS. These anomalous joints were not present in the transfected controls but have been found in a different cell-free V(D)J recombination system using a similar deletion substrate (41). In conclusion, the sequences found at the coding junctions produced by our *in vitro* reactions were very similar to those produced in tissue culture lines transfected with RAG1 and RAG2.

In vivo and *in vitro* experiments have shown that the 12/23

FIG. 2. *In vitro* and *in vivo* recombined coding joints are diverse. Coding joints generated *in vivo* (A) and *in vitro* (B) were PCR amplified, subcloned, and sequenced as described in Materials and Methods. P elements are underlined, nucleotides from the 12 or 23 RSS that have been retained are represented in italics, and N nucleotides are placed in the center of each line. The number following or preceding the partial RSS sequence indicates the number of additional template nucleotides present in the coding joint that are not included here. The top line of sequence shows some of the relevant bases from the unrecombined substrate. The number of times that every coding joint was found is indicated to the left of the sequence. Sequences for the *in vitro* recombined coding joints were obtained from five samples and three different experiments, while those for the junctions generated *in vivo* represent five samples from two independent experiments.

TABLE 1. Summary of coding joints

Type of joint	pJH200			pJH299, % in vitro
	In vivo	In vitro	In vitro + TdT	
Precise	59	36	17	35
Deletions	39	40	45	44
P elements	12	28	36	29
N nucleotides				
Total	6	7	24	12
=1	0	6	2	
>1	6	1	21	
Aberrant cleavage	0	17	29	0

rule is established at the level of cleavage and is mediated by RAG1, RAG2, and additional unidentified cellular activities (7, 40, 46, 52). To determine the extent to which the in vitro V(D)J joining reaction obeys the 12/23 rule, a mutant pJH200 deletion plasmid containing two 12 RSSs was assayed. As observed in Fig. 3, the level of coding joint formation in the 12/23 substrate was remarkably higher than the activity observed with a 12/12 substrate. The low levels of recombination found with the 12/12 substrate is consistent with published transfection experiments (14, 24).

To further characterize the in vitro reaction, we analyzed the energy and metal ion requirements (Fig. 4A). In vitro 12/23-regulated coding joint formation was dependent on the addition of the divalent cation Mg^{2+} , as in the presence of Mn^{2+} , coding joints were observed both with a 12/23 and a 12/12 substrate (Fig. 4 and data not shown). Coding joint formation was not dependent on any exogenous source of ATP, and the use of nonhydrolyzable ATP analogs had no effect on the reaction, suggesting that ATP is not required. Incubation of the 0.85 M KCl fraction at 65°C for 15 min abolished all activity, suggesting that some of the factors may have a protein component. These components could be the Ku p70/p80 heterodimer or the DNA-PK_{cs}, since both of these proteins are

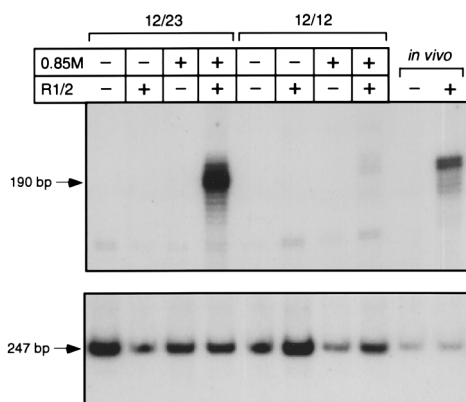


FIG. 3. In vitro recombination adheres to the 12/23 rule. pJH200 12/23 and pJH200 12/12 were individually incubated with the indicated combinations of copurified RAG1 and RAG2 (R1/2) and the 0.85 M nuclear extract fraction. An autoradiograph of the ^{32}P -labeled PCR products amplified from one such experiment and resolved on a 6% denaturing polyacrylamide gel is presented. The two lanes on the right show the ^{32}P -labeled products amplified in an identical manner from DNA isolated from in vivo recombined pJH200 12/23. The bottom panel is a DNA loading control as described in the legend to Fig. 1.

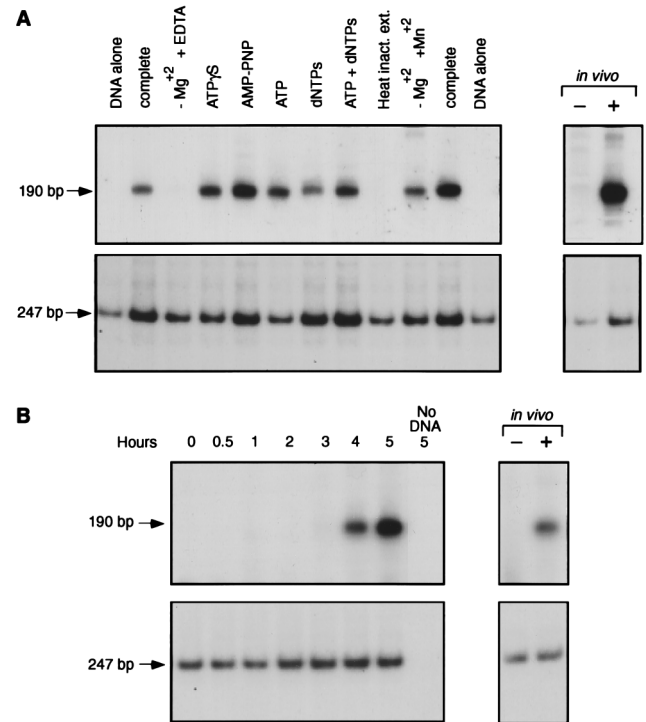


FIG. 4. Characteristics of in vitro coding joint formation. (A) Requirements of in vitro coding joint formation. The "complete" in vitro recombination reaction mixtures containing pJH200 12/23, copurified RAG1/RAG2, the 0.85 M fraction, and buffer system were incubated and processed as described in Materials and Methods. The PCR products were resolved on an 8% polyacrylamide-1× TBE gel and detected by autoradiography. "DNA alone" reaction lanes lack RAGs and fractionated extract. The "-Mg²⁺+EDTA" reaction mixture lacked magnesium and contained 1 mM EDTA. The next five lanes contained the complete reaction mixture plus 0.5 mM ATPγS (Boehringer Mannheim), 0.5 mM adenylyl-imidodiphosphate (Boehringer Mannheim), 0.5 mM ATP (Boehringer Mannheim), and/or 50 μM dNTPs (Boehringer Mannheim), as indicated. In the "Heat inact. ext." lane, the 0.85 M fraction was heated to 65°C for 15 min prior to addition to the reaction mixture. The "-Mg²⁺+Mn²⁺" lane shows coding joints amplified from a reaction mixture in which magnesium was substituted with 10 mM Mn²⁺. As a positive control, in vivo recombined substrate was analyzed in parallel. (B) Time course analysis of the in vitro recombination reaction. Here the "complete" in vitro recombination reaction was stopped and processed at the indicated times. The lane marked "No DNA" represents a sample that contained all components except for the pJH200 deletion substrate. The bottom portions of panels A and B are DNA loading controls.

present in the 0.85 M KCl fraction (data not shown) and are necessary for V(D)J recombination in vivo. A ligase activity was also present in the 0.85 M KCl fraction (data not shown). This ligase activity could be important for joining of the coding ends. Other, as yet unidentified components important for coding joint formation may be present in this fraction. Studies to further characterize the 0.85 M KCl fraction are in progress.

A time course of the reaction showed that coding joint formation was relatively slow and detectable only after 4 h at 37°C (Fig. 4B). Although the time course of coding joint formation in vivo has not been directly measured, processing of the coding ends in vivo is thought to be rapid since the hairpin intermediate is difficult to detect (35, 39, 55). It is possible that some of the components required for joining in vivo are in limiting concentrations in this in vitro system.

In vivo, TdT is responsible for the incorporation of N nucleotides (25). To examine the role of TdT in vitro, we complemented our cell-free reaction with purified enzyme. The addition of TdT did not significantly modify the level of V(D)J joining activity (Fig. 5A). However, a significant increase in N

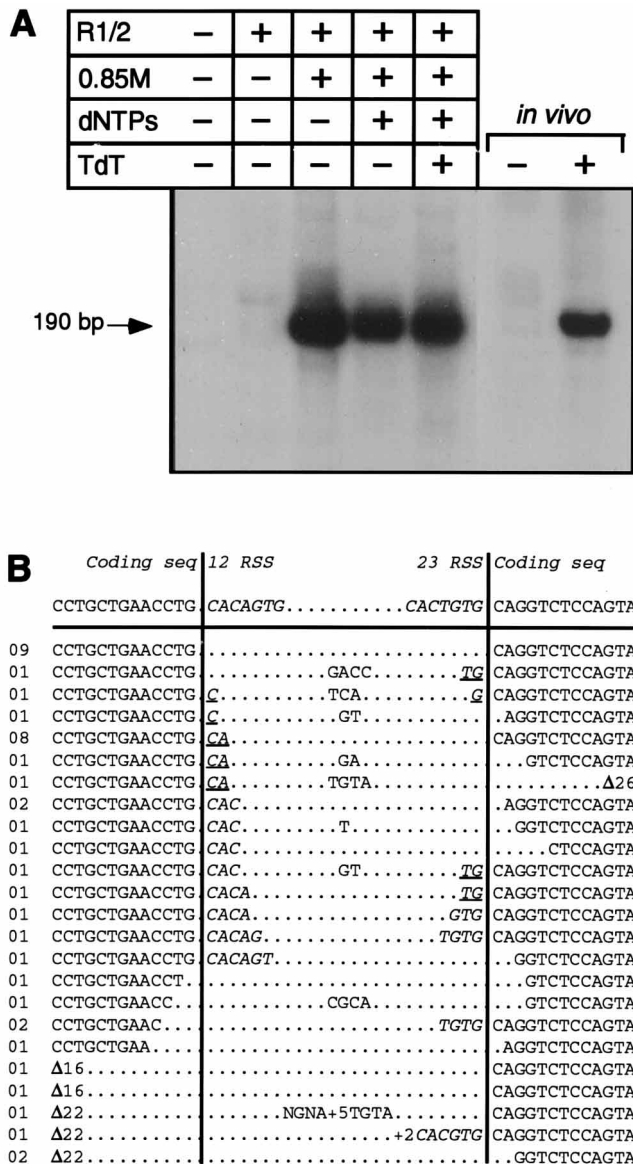


FIG. 5. Addition of N nucleotides by TdT in vitro. (A) pJH200 12/23 was subjected to in vitro recombination, coding joints were amplified by PCR, and the products were resolved on an 8% polyacrylamide-1× TBE gel. Each reaction mixture contained the indicated components: copurified RAG1 and RAG2 (R1/2), 0.85 M fraction of the HeLa nuclear extract, 50 μM dNTPs, and 25 U of TdT (Promega). (B) Coding joints from in vitro reaction performed in the presence of TdT were PCR amplified, subcloned, and sequenced as described in Materials and Methods. Every coding joint identified is presented with the number of times it was found listed to the left of the sequence. P elements are underlined, nucleotides from the 12 or 23 RSS that have been retained are represented in italics, and N nucleotides are placed in the center of each line. The number following or preceding the partial RSS sequence indicates the number of additional template nucleotides present in the coding joint that are not included here. The top line of sequence shows some of the relevant bases from the unrecombined substrate. These sequences were obtained from duplicate samples processed in one experiment.

nucleotide addition was observed. The frequency of N nucleotides was increased from 7% in the absence of TdT to 24% when TdT was included in the reaction mixture (Fig. 5B and Table 1). More importantly, in the presence of TdT, most of the N nucleotide additions were more than one nucleotide in length, as opposed to just one nucleotide added in the absence

of TdT. We conclude that TdT mediates N nucleotide addition in our in vitro coding joint reaction.

Genetic experiments have demonstrated that factors other than RAG1 and RAG2 are required for efficient V(D)J recombination in vivo (2, 12, 18, 19, 22, 26, 33, 36, 38, 47, 48). One of the factors implicated is the DNA-PK_{cs}. To determine whether in vitro coding joint formation mimics in vivo recombination in this respect, antibodies against DNA-PK_{cs} were added to the in vitro system in an attempt to block the reaction. Coding joint formation was specifically inhibited by an anti-DNA-PK_{cs} monoclonal antibody and not by an isotype-matched control (Fig. 6). Thus, the in vitro reaction resembles V(D)J joining in vivo in that it requires the catalytic component of the DNA-PK_{cs}. However, immunodepletion experiments with a monoclonal antibody that recognizes the Ku p70/p86 heterodimer shows only 50% inhibition (data not shown). The reasons for this incomplete inhibition are still unclear to us and are under study.

Inversion reactions are characteristic of authentic V(D)J recombination. To determine whether inversion-mediated coding joint formation occurs in our in vitro reactions, we used a PCR-based assay similar to that described above. The primers R3 and RA2 were designed to amplify a product of 125 bp from the pJH299 substrate only after a coding joint has been generated. As shown in Fig. 7, both RAGs and the 0.85 M HeLa phosphocellulose fraction were required for inversion-mediated coding joint formation. During inversion- and deletion-mediated coding joint formation in vitro, 1 to 5% of the input substrate undergoes coding joint formation, as determined by a semiquantitative PCR analysis (data not shown).

Inversion-mediated coding joint products were characterized by sequence analysis. As shown in Fig. 7B, 34 coding joints were sequenced, and their distribution was as follows: 35% precise, 44% deletions, 29% P elements, and 12% N nucleotides (summarized in Table 1). In contrast to the pJH200 deletion substrate, no aberrant coding joints were observed. Among the deletions, we found three coding joints which pos-

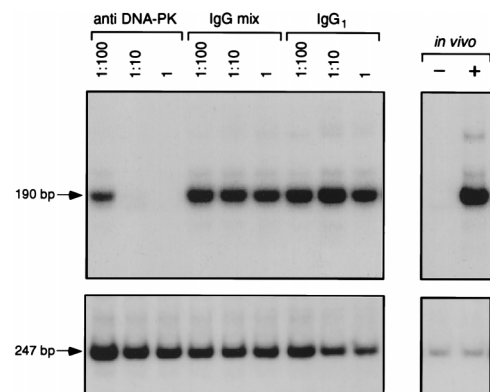


FIG. 6. A monoclonal antibody raised against DNA-PK inhibits in vitro coding joint formation. Copurified RAGs (1 μl) and 0.85 M extract fraction (1 μl) were preincubated for 15 min at room temperature with 1, 1/10, or 1/100 μl of either DNA-dependent protein kinase Ab-1 (clone 18-2; 20 ng/μl; NeoMarkers), a mixture of immunoglobulin G (IgG) isotypes (20 ng/μl; Sigma Immuno Chemicals), or an isotype-matched monoclonal antibody (20 ng/μl; NeoMarkers) in the presence of 25 mM Tris-HCl (pH 8.0), 50 mM NaCl, 20% glycerol, and 2 mM dithiothreitol. After addition of pJH200 12/23 and the rest of the buffer system (see Materials and Methods), the samples were incubated at 37°C for 5 h, and the DNA was purified and analyzed by PCR. An autoradiograph of the ³²P-labeled PCR products amplified from one such experiment resolved on an 8% polyacrylamide gel-1× TBE is presented. The two right lanes show the ³²P-labeled products amplified in an identical manner from in vivo recombined pJH200 12/23. The bottom panel shows the loading control.

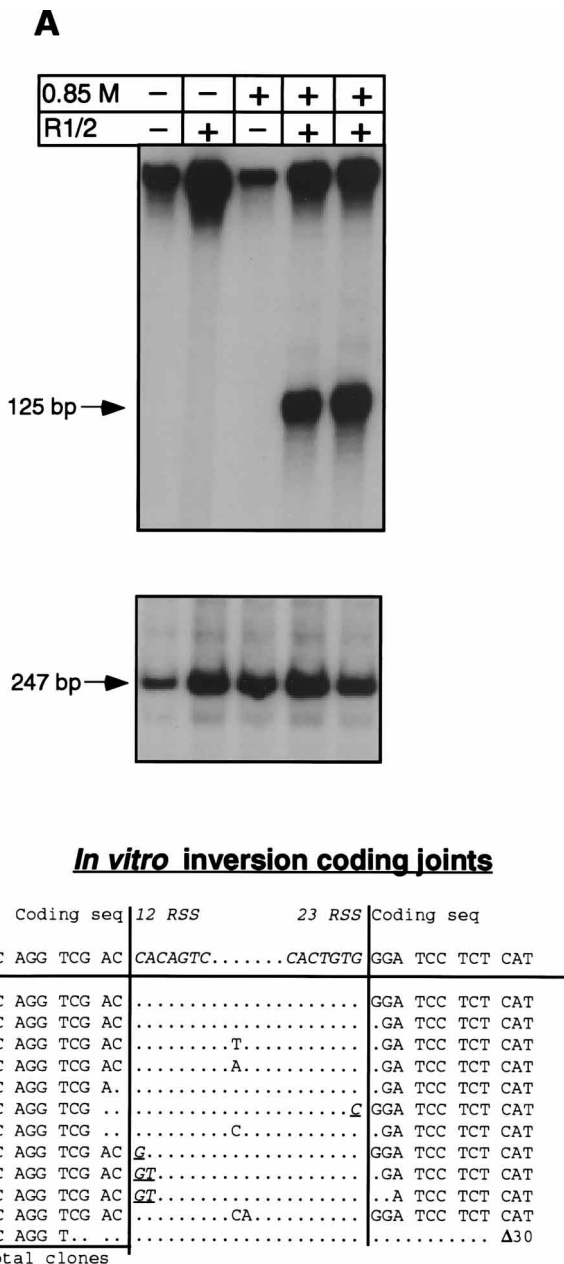


FIG. 7. In vitro coding joint formation by inversion requires RAGs and a single fraction from HeLa nuclear extract. (A) pJH299 12/23 was incubated with the indicated combinations of copurified RAG1 and RAG2 (R1/2) and the 0.85 M nuclear extract fraction as indicated. DNA was purified and analyzed by PCR, and the ^{32}P -labeled PCR products were then resolved on a 6% denaturing polyacrylamide gel. The arrow in the top panel indicates the 125-bp fragment which is amplified only from the recombined DNA (primers RA2 and CR3XhoI point in the same direction in the unrecombined DNA). The arrow in the bottom panel represents the DNA loading control. (B) Inversion-mediated coding joint formation in vitro generates diverse coding joints. Coding joints generated in vitro were PCR amplified, subcloned, and sequenced as described in Materials and Methods. Every coding joint identified is presented here with the number of times it was found listed to the left of the sequence. P elements are shown in italics and underlined, and N nucleotides are placed in the center of each line. The top line of sequence shows some of the relevant bases from the unrecombined substrate. Sequences were obtained from four samples and two different experiments.

essed a deletion of 30 nucleotides at the coding end closer to the 23 RSS. This large deletion correlates with the presence of a 24-bp palindromic sequence at this position in the pJH299 substrate. The difference between the structures of the coding joints produced from the recombination substrate pJH200 and pJH299 may be the consequence of the difference in any or all of the following: sequences present next to the RSS, the distance between the RSS, or the mechanism of coding joint formation (deletion versus inversion). In pJH299, there is only one nucleotide of ambiguity at the ends, which can represent a precise joint or a deletion of one nucleotide combined with the addition of a P element of one nucleotide. Even with this ambiguity, it seems that our in vitro coding joint formation system, which uses RAGs and nuclear extract from nonlymphoid cells, produces a large number of precise coding joints. It will be interesting to investigate whether reconstitution of this reaction using lymphoid extracts will give a different degree of diversity.

Signal joint formation was not detected with either the deletion or the inversion substrate in this in vitro system (data not shown). One possible explanation for this finding comes from a recent study by Agrawal and Schatz (1). They showed that a stable complex of RAG1, RAG2, and HMG-1 persists at the signal ends after cleavage, protecting the ends from degradation possibly inhibiting their religation. Since our reaction conditions are similar to those described in their study, the signal ends may not be available for religation. We have published a cell-free system that does mediate the generation of precise signal joints in a Ku-dependent but not 12/23-regulated manner (3). One important difference between the coding reaction and the signal joint reaction is that in vitro signal joint formation occurs in Mn^{2+} and not in Mg^{2+} . Studies directed at obtaining an in vitro system capable of efficiently generating coding and signal joints in a 12/23-dependent manner are in progress.

In summary, copurified RAG1 and RAG2 in combination with fractionated HeLa nuclear extract recapitulate four essential aspects of deletion-mediated coding joint formation: (i) the 12/23 rule; (ii) diverse coding joints, containing P elements and deletions; (iii) TdT-dependent N nucleotide addition; and (iv) a requirement for DNA-PK $_{cs}$ in the reaction. This in vitro assay for coding joint formation provides the means to study the role of the known components of the reaction, to identify new factors, and to investigate the molecular mechanism of deletion- and inversion-mediated coding joint formation.

ACKNOWLEDGMENTS

We thank David Baltimore and Juan Cárcamo for comments on the manuscript and very helpful discussions. We also thank David Schatz for communicating results before publication and Christopher Roman for providing the oligonucleotides and PCR conditions used to assay coding joint formation. F.W.-G. is supported by a Cancer Research Institute fellowship. D.J.S. is supported by a Natural Sciences and Engineering Research Council of Canada PGS-B fellowship. This work was supported by grants from the National Institutes of Health to M.C.N. M.C.N. is an associate investigator of the Howard Hughes Medical Institute.

REFERENCES

1. Agrawal, A., and D. Schatz. 1997. RAG1 and RAG2 form a stable postcleavage synaptic complex with DNA containing signal ends in V(D)J recombination. *Cell* **89**:43–53.
2. Blunt, T., N. J. Finnie, G. E. Taccioli, G. C. M. Smith, J. Demengeot, T. M. Gottlieb, R. Mizuta, A. J. Varghese, F. W. Alt, P. A. Jeggo, and S. P. Jackson. 1995. Defective DNA-dependent protein kinase activity is linked to V(D)J recombination and DNA repair defects associated with the murine *scid* mutation. *Cell* **80**:813–823.
3. Cortes, P., F. Weis-Garcia, Z. Misulovin, A. Nussenzweig, J.-S. Lai, G. Li,

- M. C. Nussenzweig, and D. Baltimore. 1996. *In vitro* V(D)J recombination: signal joint formation. Proc. Natl. Acad. Sci. USA **93**:14008–14013.
4. Cuomo, C. A., C. L. Mundy, and M. A. Oettinger. 1996. DNA sequence and structure requirements for cleavage of recombination signal sequences. Mol. Cell. Biol. **16**:5683–5690.
 5. Difilippantonio, M. J., C. J. McMahan, Q. M. Eastman, E. Spanopoulou, and D. Schatz. 1996. RAG1 mediates signal sequence recognition and recruitment of RAG2 in V(D)J recombination. Cell **87**:253–262.
 6. Dignam, J. D., P. L. Martin, B. S. Shastry, and R. G. Roeder. 1983. Eukaryotic gene transcription with purified components. Methods Enzymol. **101**:582–598.
 7. Eastman, Q. M., T. M. L. Leu, and D. G. Schatz. 1996. Initiation of V(D)J recombination *in vitro* obeying 12/23 rule. Nature **380**:85–88.
 8. Gauss, G. H., and M. R. Lieber. 1996. Mechanistic constraints on diversity in human V(D)J recombination. Mol. Cell. Biol. **16**:258–269.
 9. 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.
 10. Ge, H., E. Martinez, C.-M. Chiang, and R. G. Roeder. 1996. Activator-dependent transcription by mammalian RNA polymerase II: *in vitro* reconstitution with general transcription factors and cofactors. Methods Enzymol. **274**:57–71.
 11. Gilfillan, S., A. Dierich, M. Lemeur, C. Benoist, and D. Mathis. 1993. Mice lacking TdT: mature animals with an immature lymphocyte repertoire. Science **261**:1175–1178.
 12. Hendrickson, E. A., X.-Q. Qin, E. A. Bump, D. G. Schatz, M. Oettinger, and D. T. Weaver. 1991. A link between double-strand break-related repair and V(D)J recombination: the *scid* mutation. Proc. Natl. Acad. Sci. USA **88**:4061–4065.
 13. Hesse, J. E., M. R. Lieber, M. Gellert, and K. Mizuuchi. 1987. Extrachromosomal DNA substrates in pre-B cells undergo inversion or deletion at immunoglobulin V-(D)-J joining signals. Cell **49**:775–783.
 14. 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**:1053–1067.
 15. Hiom, K., and M. Gellert. 1996. A stable RAG1-RAG2-DNA complex that is active in V(D)J cleavage. Cell **88**:65–72.
 16. Jackson, S. P., and P. A. Jeggo. 1995. DNA double-strand break repair and V(D)J recombination—involve ment of DNA-PK. Trends Biochem. Sci. **20**:412–415.
 17. Kallenbach, S., N. Doyen, M. Fanton d'Andon, and F. Rougeon. 1992. Three lymphoid-specific factors account for all junctional diversity characteristic of somatic assembly of T-cell receptor and immunoglobulin genes. Proc. Natl. Acad. Sci. USA **89**:2799–2803.
 18. Kemp, L. M., S. G. Sedgwick, and P. A. Jeggo. 1984. X-ray sensitive mutants of Chinese hamster ovary cells defective in double-stranded break rejoining. Mutat. Res. **132**:189–196.
 19. Kirchgessner, C. U., C. K. Patil, J. W. Evans, C. A. Cuomo, L. M. Fried, T. Carter, M. A. Oettinger, and J. M. Brown. 1995. DNA-dependent kinase (p350) as a candidate gene for the murine *scid* defect. Science **267**:1178–1183.
 20. Komori, T., A. Okada, V. Stewart, and F. W. Alt. 1993. Lack of N regions of antigen receptor variable region genes of TdT-deficient lymphocytes. Science **261**:1171–1175.
 21. Landau, N. R., D. G. Schatz, M. Rosa, and D. Baltimore. 1987. Increased frequency of N-region insertion in a murine pre-B-cell line infected with a terminal deoxynucleotidyl transferase retroviral expression vector. Mol. Cell. Biol. **7**:3237–3243.
 22. Lees-Miller, S. P., R. Godbout, D. W. Chan, M. Weinfeld, R. S. Day III, G. M. Barron, and J. Allalunis-Turner. 1995. Absence of p350 subunit of DNA-activated protein kinase from a radiosensitive human cell line. Science **267**:1183–1185.
 23. Leu, T. M., and D. G. Schatz. 1995. RAG-1 and RAG-2 are components of a high-molecular-weight complex and association of RAG-2 with this complex is RAG-1 dependent. Mol. Cell. Biol. **15**:5657–5670.
 24. Lewis, S. M., and J. E. Hesse. 1991. Cutting and closing without recombination in V(D)J joining. EMBO J. **10**:3631–3639.
 25. Lewis, S. M. 1994. The mechanism of V(D)J joining: lessons from molecular, immunological and comparative analysis. Adv. Immunol. **56**:27–150.
 26. Li, Z., T. Otevrel, Y. Gao, H.-L. Cheng, B. Seed, T. D. Stamoto, G. E. Taccioli, and F. Alt. 1995. The XRCC4 gene encodes a novel protein involved in DNA double-strand break repair and V(D)J recombination. Cell **83**:1079–1089.
 27. 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 coding sequences in V(D)J recombination. Cell **7**:7–16.
 28. 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.
 29. Mizushima, S., and S. Nagata. 1990. pEF-BOS, a powerful mammalian expression vector. Nucleic Acids Res. **18**:5322.
 30. Nussenzweig, A., C. Chen, V. da Costa Soares, M. Sanchez, K. Sokol, M. C. Nussenzweig, and G. C. Li. 1996. Requirement for Ku80 in growth and V(D)J recombination. Nature **382**:551–555.
 31. Oettinger, M. A., D. G. Schatz, C. Gorka, and D. Baltimore. 1990. RAG-1 and RAG-2, adjacent genes that synergistically activate V(D)J recombination. Science **248**:1517–1523.
 32. Pear, W., G. Nolan, M. Scott, and D. Baltimore. 1993. Production of high titer helper-free retrovirus by transient transfection. Proc. Natl. Acad. Sci. USA **90**:8392–8395.
 33. Pergola, F., M. Z. Zdzienicka, and M. R. Lieber. 1993. V(D)J recombination in mammalian mutants defective in DNA double-strand break repair. Mol. Cell. Biol. **13**:3464–3471.
 34. Ramsden, D. A., J. F. McBlane, D. C. van 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.
 35. 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.
 36. Rathmell, W. K., and G. Chu. 1994. A DNA end-binding factor involved in double-strand break repair and V(D)J recombination. Mol. Cell. Biol. **14**:4741–4748.
 37. Roman, C. A., and D. Baltimore. 1996. Genetic evidence that the RAG1 protein directly participates in V(D)J recombination through substrate recognition. Proc. Natl. Acad. Sci. USA **93**:2333–2338.
 38. Roth, D. B., T. Lindahl, and M. Gellert. 1995. Repair and recombination. How to make ends meet. Curr. Biol. **5**:496–499.
 39. 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.
 40. Sawchuk, D. J., F. Weis-Garcia, S. Malik, E. Besmer, M. Bustin, M. C. Nussenzweig, and P. Cortes. 1997. V(D)J recombination: modulation of RAG1 and RAG2 cleavage activity on 12/23 substrates by whole cell extract and DNA bending proteins. J. Exp. Med. **185**:2025–2032.
 41. Schatz, D. Personal communication.
 42. Schatz, D., M. Oettinger, and M. Schlissel. 1992. V(D)J recombination: molecular biology and regulation. Annu. Rev. Immunol. **10**:359–383.
 43. Schatz, D. G., M. A. Oettinger, and D. Baltimore. 1989. The V(D)J recombination activating gene, RAG-1. Cell **59**:1035–1048.
 44. Spanopoulou, E., F. Zaitseta, F.-H. Wang, S. Santagata, D. Baltimore, and G. Panayotou. 1996. The homeodomain region of Rag-1 reveals the parallel mechanisms of bacterial and V(D)J recombination. Cell **87**:263–276.
 45. Spanopoulou, E., P. Cortes, C. Shih, E. Huang, D. P. Silver, P. Svec, and D. Baltimore. 1995. Localization, interaction and RNA binding properties of the V(D)J recombination activating proteins Rag-1 and Rag-2. Immunity **3**:715–726.
 46. Steen, S. B., L. Gomelsky, and D. B. Roth. 1996. The 12/23 rule is enforced at the cleavage step of V(D)J recombination *in vivo*. Genes Cells **1**:543–553.
 47. Taccioli, G. E., T. M. Gottlieb, T. Blunt, A. Priestly, J. Demengeot, R. Mizuta, A. R. Lehmann, F. W. Alt, S. P. Jackson, and P. A. Jeggo. 1994. Ku80: product of the XRCC5 gene and its role in DNA repair and V(D)J recombination. Science **265**:1442–1445.
 48. Taccioli, G. E., G. Rathbun, E. Oltz, T. Stamato, P. A. Jeggo, and F. W. Alt. 1993. Impairment of V(D)J recombination in double-strand break repair mutants. Science **260**:207–210.
 49. Tonegawa, S. 1983. Somatic generation of antibody diversity. Nature **302**:575–581.
 50. van Gent, D. C., J. F. McBlane, D. A. Ramsden, M. J. Sadofsky, J. A. Hesse, and M. Gellert. 1995. Initiation of V(D)J recombination in a cell-free system. Cell **81**:825–934.
 51. van Gent, D. C., K. Mizuuchi, and M. Gellert. 1996. Similarities between initiation of V(D)J recombination and retroviral integration. Science **271**:1592–1594.
 52. van Gent, D. C., D. A. Ramsden, and M. Gellert. 1996. The RAG1 and RAG2 proteins establish the 12/23 rule in V(D)J recombination. Cell **85**:107–113.
 53. Weaver, D. T. 1995. What to do at the end—DNA double-strand-break repair. Trends Genet. **11**:388–392.
 54. Zhu, C., M. A. Bogue, D.-S. Lim, P. Hasty, and D. B. Roth. 1996. Ku86-deficient mice exhibit severe combined immunodeficiency and defective processing of V(D)J recombination intermediates. Cell **86**:379–389.
 55. 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.