Proc. Natl. Acad. Sci. USA
Vol. 93, pp. 2333–2338, March 1996
Immunology

Genetic evidence that the RAG1 protein directly participates in V(D)J recombination through substrate recognition

CHRISTOPHER A. J. ROMAN AND DAVID BALTIMORE

Department of Biology, Massachusetts Institute of Technology, Cambridge, MA 02139

Contributed by David Baltimore, October 18, 1995

ABSTRACT **RAG1 protein is essential for the activation of V(D)J recombination in developing lymphocytes (V, variable; D, diversity; J, joining). However, it has not been determined whether its role involves substrate recognition and catalysis. A single amino acid substitution mutation in the *RAG1* gene has now been identified that renders its activity sensitive to the sequence of the coding region abutting the heptamer site in the recombination signal sequence. These results strongly imply that RAG1 interacts directly with DNA.**

V(D)J recombination—the joining of DNA segments in early lymphocytes that creates the antigen receptor genes—requires activity of the recombination activating genes *RAG1* and *RAG2* (1, 2) (V, variable; D, diversity; J, joining). These genes are expressed almost exclusively in developing lymphocytes that are undergoing antigen receptor gene assembly and are the only known lymphoid-specific proteins required for the process (1). The simultaneous expression of *RAG1* and *RAG2* in nonlymphoid cells is sufficient to activate V(D)J recombination activity (1). Furthermore, mice that lack either RAG1 or RAG2 fail to develop mature lymphocytes due to their inability to initiate rearrangement of the antigen receptor genes (3–5).

An essential step in recombination is the recognition of specific DNA sequences (recombination signal sequences; RSS) adjacent to the variable gene segments which appropriately target the recombination events (6). The core RSS is composed of conserved heptamer and nonamer sequences separated by 12 or 23 nucleotides (7). Furthermore, the nucleotide sequence of the coding ends can profoundly influence the efficiency of RSS usage (8–10). During the initial steps of recombination, chromosomally distal RSSs must be juxtaposed and double-strand breaks introduced at the heptamer/coding sequence junctions. Analysis of broken DNA ends in lymphoid cells from normal and *Scid* (severe combined immunodeficiency) mice have shown that the heptamer ends are blunt and 5′ phosphorylated (11–14); in contrast, the coding ends in *Scid* lymphocytes are retained as hairpins, which most likely represent normal intermediates in the catalytic pathway (13, 15). The subsequent processing and joining of ends is dependent on factors involved in double-strand break repair (16).

Although the exact role of RAG1 and RAG2 has remained elusive, recent evidence has suggested that RAG1 directly activates recombination. *In vitro* studies have shown that recombinant RAG1 can promote the formation of DNA cleavage products associated with the initial steps of V(D)J recombination when supplemented into nuclear extracts that contain RAG2 (17). However, it is not known whether RAG1 participates directly in the catalytic mechanism and/or activates the recombinase. Here we provide genetic evidence that RAG1 is directly involved in the recognition of the DNA substrate.

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

MATERIALS AND METHODS

Plasmids. The RAG1 and RAG2 CDM8 eukaryotic expression vectors are described by Oettinger *et al.* (1). The RAG1 CDM8 expression constructs either contained or lacked a short 5′ untranslated region of the *RAG1* cDNA. This extra sequence decreased RAG1 protein expression ≈5-fold but had no impact on recombination activity, as described in *Results*. Site-directed mutations in RAG1 were generated by oligonucleotide-mediated mutagenesis (Bio-Rad) of a single-stranded phagemid Bluescript (Stratagene) that contained the *RAG1* cDNA. For cloning convenience, we created an artificial R1H609L allele for expression from the CDM8 vector, which lacked the 5′ untranslated region; this construct contains a new *Hin*dIII site at amino acid position 609 and lacks the *Ava* I site present in the original R1H609L (M6) allele.

The V(D)J recombination substrates pJH200 (18), pJH288, and pJH289 (19) were the gifts of M. Gellert and J. Hesse. Mutations in the V(D)J recombination substrates were created by two- and three-step overlap PCR-mediated mutagenesis (20). In creating the pJH288 derivatives, a *Bgl* II fragment containing an M13 origin of replication was removed. This had no effect on recombination activity. Sequencing of all recombination substrates and RAG1 derivatives was done with an automated sequencing apparatus (Applied Biosystems) to confirm their integrity. The identity of the recombination substrates is shown in Table 1.

Cell Lines and Transfections. DNA was transiently introduced into the fetal human kidney carcinoma cell line 293 cells by calcium phosphate-mediated transfection as described (21). RAG1 and RAG2 CDM8 expression vectors (6 μg each) were cotransfected with 6 μg of the indicated V(D)J recombination plasmid substrate into cells plated at 2×10^6 cells per 60-mm dish 16 hr earlier. Cells were harvested after 44–48 hr. Transfected plasmid DNA was recovered by Hirt extraction of cell lysates as described (18, 22). Less than a 2-fold variation in transfection efficiency was routinely observed.

Recombination Assay. One percent of the recovered plasmid was subject to PCR amplification. The following oligonucleotides were used to direct the amplification of signal-joint products: RA-CR2, 5′-TTTGTTCCAGTCTGTAG**CACTGCGCAC**-3′; RA-14, 5′-TCCAGCTGAACGGTCTGG-3′. The oligonucleotide RA-CR2 spans most of the 12-RSS and the signal joint generated after recombination. It contains a mismatch (underlined) within the sequence complementary to the 12-heptamer site (boldface), which prevents the amplification of fragments from the unrearranged vector. RA-14 is complementary to sequences within the chloramphenicol-resistance gene. The PCR fragment sizes from the pJH200 and pJH288 series are 256 and 236 bp, respectively. The following oligonucleotides were used to detect circular (closed) excision products for pJH200 and its derivatives: RA-CR3, 5′-ATTGGTGAGAATCGCAGCAACTTGT-3′; OOP2, 5′-CGGCAACCGAGCGTTCTGAAC-3′. The PCR fragment size is 190 bp.

Abbreviations: V, variable; D, diversity; J, joining; RSS, recombination signal sequence.

Table 1. List of recombination substrates used in this study

Plasmid	5′-Heptamer coding sequence
pJH200 (SJ-DEL)	5′-AACCTGCACATGT-(23)-nonamer Jκ1
	5′-AGACTGCACATGT-(12)-nonamer VκL8
pJH200J	5′-GGATCC Jκ1 (*Bam*HI)
pJH200V(ApaLI)	5′-GTGCAC VκL8 (*Apa*LI)
pJH200V(SalI)	5′-GTCGAC VκL8 (*Sal* I)
pJH200JV(SalI)	
pJH200JV(ApaLI)	
pJH288 (INV)	5′-GGATCC Jκ1 (*Bam*HI)
pJH289 (SJ-DEL)	5′-GTCGAC VκL8 (*Sal* I)
pJH288J	5′-GGAGTG Jκ1
pJH288V	5′-GTCGTG VκL8
pJH288JV	
pDGR (INV)	5′-GTCCAC Jκ1
	5′-TCCTCC Vκ21-C

Plasmid name is indicated in the left column, with parental vectors listed in boldface type. Coding-end sequences 5′ of the indicated heptamer for specified plasmids are listed in the right column; relevant sequences from pJH200 are shown more completely as a model for those below them. The 12-RSS is from VκL8, and the 23-RSS is from Jκ1 (see ref. 18 for details). Heptamer sequence is underlined, and the spacer number is given in parentheses. Two nucleotides in boldface are the two nucleotides 5′ to the heptamer. SJ-DEL, signal-joint retaining deletion substrate; INV, inversion substrate. For the pJH200 and pJH288 derivatives, a J and/or V is added to the plasmid name to indicate at which heptamer the coding-end changes were made. The two coding-end changes at the 12-heptamer in pJH200 are indicated by the restriction site used to create the coding-end changes in parentheses preceded by a V. Restriction sites for the pJH200 derivatives are adjacent to the heptamer sites, as they are in pJH288. The *Apa*LI sites in the pJH288 derivatives span the coding end and heptamer (GTGCAC). The coding-ends of pDGR are shown for comparison.

A 25-μl PCR amplification reaction mixture contained 10 mM Tris·HCl (pH 8.3), 50 mM KCl, 5.5 mM $MgCl_2$, 0.1 mg of acetylated bovine serum albumin per ml, 100 ng of each oligonucleotide, 20 μM each dNTP, 2 μCi of [α-^{32}P]dCTP (3000 Ci/mmol; 1 Ci = 37 GBq; Amersham), and 1 unit of *Taq* DNA polymerase (Boehringer Mannheim) and was subject to the following amplification profile: 94°C for 30 sec and 72°C for 2 min × 22 cycles; 72°C for 10 min. The radiolabeled products were resolved by electrophoresis through a 5% polyacrylamide gel in 0.5 × TBE. Bands were visualized by autoradiography and quantitated by PhosphorImaging (Molecular Dynamics).

R1WT and mutant constructs were always transfected in parallel. To obtain relative values for recombination activity, the amount of the fragment amplified from the R1WT sample, as reflected in the band intensity detected by PCR and PhosphorImaging, was defined as one. Relative values were determined by the simultaneous PCR amplification of a serially diluted positive control sample (either R1WT or R1H609LΔC1023-1040, whichever was higher) from which a standard curve was derived. Values obtained in this way were linear over a 100-fold range for amplification products from both pJH200 and pJH288. The relative values between two samples was extremely reproducible when determined from independent PCR runs with standard curves derived from different dilution standards (e.g., run I: sample A set at 1.0, sample B = 0.36; run II: sample A set at 1.0, sample B = 0.34).

The recombination frequency of pJH200 was 1–3% with R1WT as determined by bacterial transformation and ≈0.1% for pJH288. The relative amounts of PCR fragments amplified from these V(D)J recombined plasmids and their derivatives reflected these differences.

Protein Analysis. 293 cells plated as described above were transfected with 20 μg of the indicated expression vectors. Nuclei were isolated by incubation of transfected cells (3–4 × 10^6 cells at the time of harvest) in 0.5 ml of lysis buffer (1% Nonidet P-40, 10 mM Tris·HCl, pH 8.0/10 mM NaCl). Nuclei were then lysed by sonication in 0.5 ml of lysis buffer plus 0.4 M NaCl, and the insoluble fraction was collected by centrifugation. The insoluble pellets were resuspended by sonication in 100 μl of SDS sample buffer and 1/10th of the samples were resolved through 10% polyacrylamide gels via SDS/PAGE by standard procedures (23). Proteins were then transferred to nitrocellulose membranes, probed with a mouse anti-RAG1 monoclonal antibody (Pharmingen), and visualized with horseradish peroxidase-conjugated sheep anti-mouse secondary antibody (Fisher) developed with ECL (Amersham).

RESULTS

The *RAG1* genomic locus was originally isolated by its ability to activate recombination in the fibroblast line 3TGR (2, 24). 3TGR harbors an integrated retroviral recombination substrate (pDGR; ref. 25) that contains a neomycin (neo)-resistance gene, which is dependent on inversion via V(D)J recombination for its transcriptional activation. Two murine *RAG1* cDNAs, M2 and M6, were originally isolated (2), but only M2 promoted the formation of neo-resistant colonies of 3TGR; in contrast, the M6 cDNA was inactive in this assay. Nevertheless, sequence analysis of M6 failed to reveal any obvious coding sequence flaw, such as a nonsense mutation. We compared the complete sequences of M2 and M6 and found that they differed at only one nucleotide, which resulted in a missense mutation changing histidine (M2) to leucine (M6) at position 609 (Fig. 1*A*). Although the published sequence of murine RAG1 is of M6 (2), sequence comparison of RAG1 from different species showed that a histidine residue was present at this position in all other species (compiled in ref. 26). Furthermore, we have independently derived murine RAG1 sequences that contain histidine at this site. Thus, the M2 allele is most likely the true wild-type murine gene (herein referred to as RAG1 wild type; R1WT) and M6 a serendipitous variant (herein referred to as R1H609L).

We tested both alleles to compare their ability to recombine plasmid substrates in fibroblasts by transient cotransfection with RAG2. We first tested R1WT and R1H609L with the plasmid substrate pJH200, a deletion substrate that retains signal joints after recombination (18) (Table 1). We measured

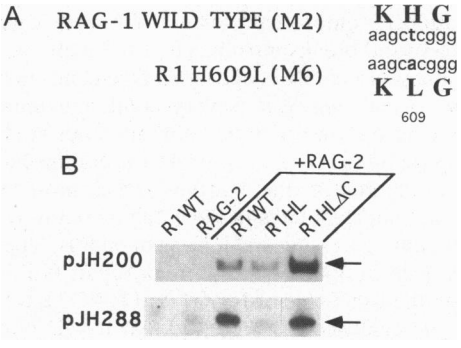

FIG. 1. Two alleles of *RAG1*, R1WT and R1H609L, have different relative activities with different plasmid substrates. (*A*) Sequence of murine wild-type *RAG1* (R1WT, M2) and R1H609L (M6) from nucleotides 1916–1924. Nucleotide at position 1920 is indicated in boldface. R1H609L contains an *Ava* I restriction site here; this is lost in the wild-type allele. (*B*) PCR amplification of fragments derived from recombination plasmid substrates pJH200 and pJH288. An autoradiograph of fragments amplified from Hirt plasmid preparations from cells transfected with either pJH200 (*Upper*) or pJH288 (*Lower*) as described. Band representing the fragment created by V(D)J recombination is indicated with an arrow. RAG expression vectors included in the transfections are indicated above each lane. R1HL, R1H609L; R1HLΔC, R1H609LΔC1023–1040.

the relative abundance of recombined pJH200 recovered from transiently transfected cells by PCR amplification of DNA fragments whose formation depended on V(D)J recombination (see *Materials and Methods*). We found that in contrast to their relative behavior in the 3TGR colony-forming assay, both R1WT and R1H609L activated recombination of pJH200 (Fig. 1*B*). As measured by the PCR assay, the relative activity of R1H609L was ≈55% that of R1WT (Table 2). We independently confirmed the PCR assay results by measuring the proportion of recombinant plasmids by transformation of bacteria (data not shown). In parallel, we tested a deletion mutant of R1H609L, R1H609LΔC1023-1040, from which the C-terminal 17 amino acids was removed (27). This truncation activated the recombination frequency of the H609L allele by ≈7.5-fold and was on average 4-fold more active than full-length R1WT as measured with pJH200 (Fig. 1*B* and Table 2; see below). We carried this allele through with all of the later recombination substrates because it acts as an independent test of activity.

In contrast, we observed a striking deficiency of R1H609L with respect to R1WT in its ability to activate recombination of pJH288, an inversion substrate that retains both coding and signal joints (ref. 19; Fig. 1*B* and Table 2; see below). The ability of R1H609L to activate recombination in this assay was only 7% of the activity of R1WT. We also observed a dramatic reduction in the relative activity of R1H609LΔC1023-1040. This variant activated recombination less efficiently than R1WT (80% R1WT activity) with pJH288 (Fig. 1*B* and Table 2; see below). The relative activities of R1WT and R1H609L as measured by pJH288 mirrored the activity observed using the 3TGR colony-forming assay.

The different relative activities of the RAG1 alleles were not due to variations in protein levels. We initially observed that H609L was expressed at an ≈5-fold lower steady-state amount compared to R1WT when both alleles were expressed from the same expression vector (Fig. 2*A*). However, we determined that under our transfection conditions the recombination assay was indifferent to a >20-fold decrease in the steady-state amount of wild-type RAG1 and was limiting only for RAG2 (Fig. 2*B*; data not shown). We confirmed this observation directly by varying the transfection parameters either by lowering the amount of RAG-1WT expression vector or by changing the expression vectors such that both proteins were expressed at comparable amounts. These changes did not have a significant effect on the relative activities of the mutant and wild-type alleles with a panel of recombination substrates (Fig. 2*B*). Furthermore, indirect immunofluorescence analysis of R1H609L showed that the protein localized normally within the nucleus (27). Therefore, the behavior of these RAG1 variants reflected intrinsic differences in their activities. These results strongly suggested that RAG1 acts directly in the V(D)J recombination reaction mechanism because different alleles differentially activated the recombination of different plasmid substrates.

Since both plasmids tested here contained identical 12- and 23-RSSs, we were interested in understanding what other

Table 2. Comparison of activity of *RAG1* alleles as measured with different plasmid V(D)J recombination substrates

RAG1 allele	Recombination substrate		
	pJH200	pJH288	pJH288JV
R1WT	1.000	1.000	1.000
R1H609L	0.553 ± 0.082	0.077 ± 0.041	0.712 ± 0.122
R1H609LΔC	4.153 ± 1.043	0.840 ± 0.253	8.06 ± 1.336

Relative activities of the different alleles of *RAG1* as determined by PCR assay, with the value of R1WT set as one. Data represent values obtained from 3 (pJH200), 6 (pJH288), or 5 (pJH288JV) experiments ± SD. These values are reiterated in bar-graph format in Figs. 3 and 4.

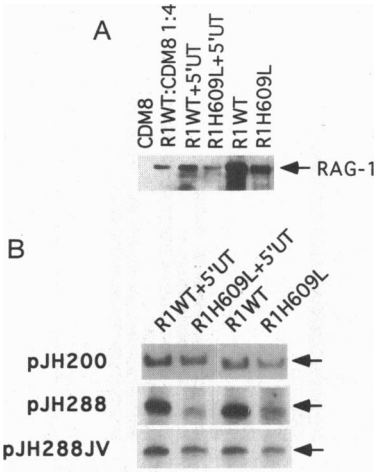

FIG. 2. Modest changes in RAG1 protein levels do not affect recombination activity. (*A*) Western blot analysis of RAG1 proteins from cells transfected with different RAG1 CDM8 expression vectors. All RAG1 constructs were in the CDM8 backbone and a total of 20 μg was used per transfection. +5′UT, CDM8 vector, which contains the 5′ untranslated region of RAG1, was used. Lane 1, empty CDM8 vector was transfected; lane 2, one-fifth the amount (5 μg) of CDM8-R1WT was transfected and the remainder (15 μg) was empty CDM8. (*B*) Recombination activity of *RAG1* alleles expressed from different constructs. Autoradiograph of fragments amplified from the indicated samples from one experiment in which they were transfected in parallel. All transfections included RAG2.

differences between pJH288 and pJH200 could account for the phenotype of R1H609L. Three major differences in these plasmids could have been responsible for the effect: (*i*) the DNA topology of inversion versus deletion, (*ii*) the requirement of coding joint formation for inversion, and (*iii*) the coding-end nucleotide sequence adjacent to the RSS. We believed the last possibility most likely because it is known that the coding nucleotides immediately adjacent to the heptamer could significantly affect recombination frequency by the wild-type V(D)J recombinase by several orders of magnitude (8–10). Furthermore, the choice between inversion versus deletion in one class of plasmid substrates was influenced by the content of the coding sequences adjacent to the heptamer sites (9). We noted that pJH288 and pDGR contain the same sequences abutting the heptamer sites (5′-AC-heptamer or CC-heptamer; see Table 1) but are different from pJH200 where the sequence is 5′-TG-heptamer.

We therefore changed the coding-end sequence at the heptamer junctions of pJH288 to those present in pJH200 (Table 1). We first challenged the RAG1 variants with the plasmid substrate pJH288JV, which contains TG 5′ to both heptamers. Simply changing the sequences of the coding ends in pJH288 activated the ability of R1H609L to recombine this substrate >10-fold compared to R1WT (75% the activity of R1WT; compare Fig. 3 *A* and *B* and Table 2). The relative recombination frequencies of R1WT and R1H609L with this altered pJH288 substrate were comparable to those with pJH200. This was reflected in the relative activity of R1H609LΔC1023-1040; this form of RAG1 was also activated 10-fold, which resulted in an activity 8 times higher than R1WT (Fig. 3 *A* and *B* and Table 2). The C-terminal truncation thus provided independent confirmation of the effect of coding-end sequence on recombination activity by the H609L allele. As measured by both pJH288 and pJH288JV, the difference in the activities of R1H609L and R1H609LΔC1023-1040 was >11-fold. It is evident that the R1H609L variant is sensitive to the coding sequence adjacent to the heptamer site.

To assess the individual contribution of each coding end to the recombination activity, we tested the ability of R1WT and

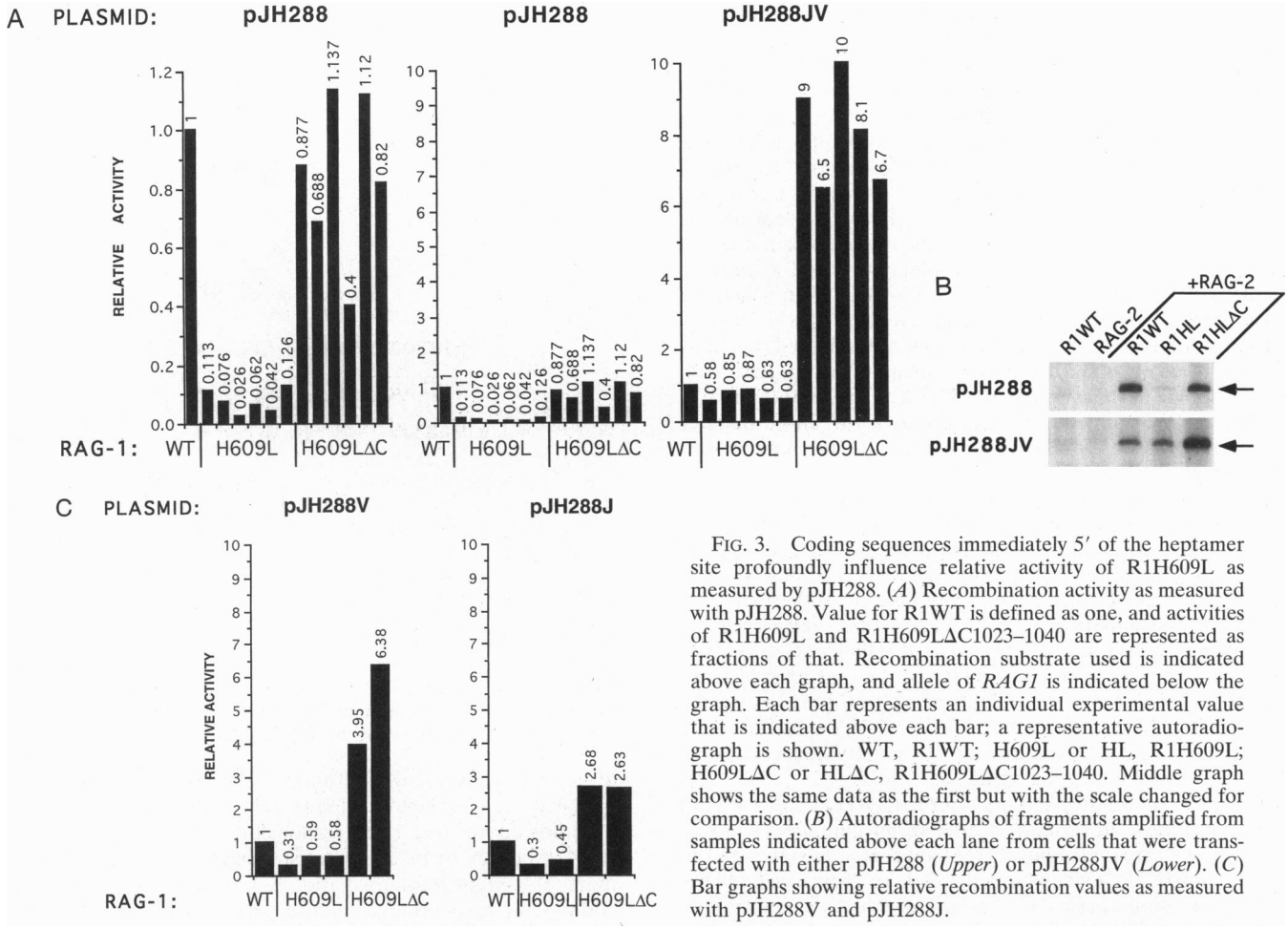

FIG. 3. Coding sequences immediately 5′ of the heptamer site profoundly influence relative activity of R1H609L as measured by pJH288. (*A*) Recombination activity as measured with pJH288. Value for R1WT is defined as one, and activities of R1H609L and R1H609LΔC1023–1040 are represented as fractions of that. Recombination substrate used is indicated above each graph, and allele of *RAG1* is indicated below the graph. Each bar represents an individual experimental value that is indicated above each bar; a representative autoradiograph is shown. WT, R1WT; H609L or HL, R1H609L; H609LΔC or HLΔC, R1H609LΔC1023–1040. Middle graph shows the same data as the first but with the scale changed for comparison. (*B*) Autoradiographs of fragments amplified from samples indicated above each lane from cells that were transfected with either pJH288 (*Upper*) or pJH288JV (*Lower*). (*C*) Bar graphs showing relative recombination values as measured with pJH288V and pJH288J.

R1H609L to activate recombination of pJH288 derivatives that contained base substitutions at only the 12-heptamer or 23-heptamer coding junctions (pJH288V and pJH288J; see Table 1). Changes at single coding ends allowed substantial but incomplete rescue of R1H609L activity relative to R1WT (Fig. 3*C*). The effect seemed to be additive because each mutation allowed 50–60% of the total possible rescue. We also noted that mutations at the 12-heptamer coding junction seemed to have a slightly stronger effect on the relative activities of the two *RAG1* alleles than those at the 23-heptamer. Both of these observations were reiterated in the activities of the C-terminal truncation of R1H609L.

Because changing the coding sequences adjacent to the heptamer sites in pJH288 was sufficient to render it a significantly better substrate for R1H609L, we anticipated that reciprocal mutations in pJH200 would have a conversely deleterious effect. We first tested the pJH200 derivative pJH200JV(ApaLI), which contains substitutions of 5′-TG at both the 12-heptamer and 23-heptamer coding junctions. As shown in Fig. 4, these changes did indeed repress the recombination activities of R1H609L and R1H609LΔC1023–1040 with respect to R1WT. However, in marked contrast to pJH288, we observed only an ≈2-fold repressive effect. Thus, as measured with this plasmid substrate, the activity of R1H609L was only slightly affected by the two coding-end nucleotides. Similarly, mutations at either the 12- or 23-heptamer [plasmids pJH200V(ApaLI) and pJH200J] had little if any effect (data not shown). Thus, there are contextual effects that we have yet to understand.

One explanation for this difference is that the initial steps of recombination—i.e., binding, synapsis, and cleavage—were not strongly affected by these nucleotide changes. In this scenario, coding-joint formation with those ends may be hindered by R1H609L and the manifestation of this deficiency would arise only in pJH288, which must retain coding joints. To address this issue, we examined the excision products of pJH200 and its derivatives. We expected that if coding joint formation were disproportionately affected, the relative amounts of excision products that contain coding joints should be significantly different than the plasmid products. We therefore designed a PCR-based assay to detect the excised circular products. As shown in Fig. 5*A*, excision products that retain coding joints can be detected from Hirt plasmid preparations recovered from cells transfected either with pJH200 and pJH200JV(ApaLI). The relative abundance of these products paralleled that observed with the respective parental plasmid substrates (see Fig. 5*B*). We therefore conclude that coding-joint formation was not differentially affected by R1H609L.

This observation implied that either inversion of the intervening sequence in pJH288 or other sequence differences between the plasmid families were affecting recombination activity. To address the first possibility, we determined the recombination profile of pJH289, a signal-joint-retaining deletion substrate of the same plasmid family as pJH288 (19). This plasmid had a recombination frequency profile indistinguishable from that observed with pJH288 (data not shown), and thus inversion was not responsible for the differences in recombination activity observed.

To address the second possibility, we created the substrate pJH200JV(SalI) to test whether heptamer-proximal nucleotides beyond the two at the coding end might influence recombination activity of pJH200; pJH288 contains a *Sal* I rather than an *Apa*LI site adjacent to the 12-heptamer (see Table 1). Although these restriction sites differ only by the

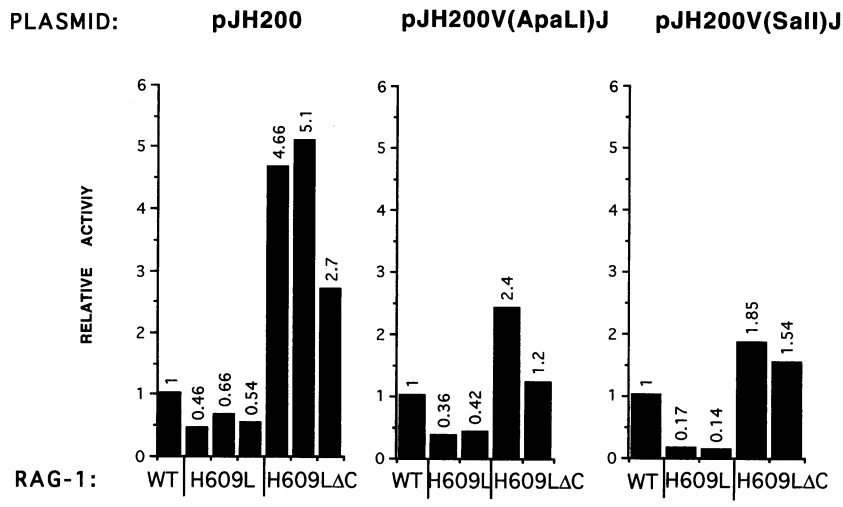

FIG. 4. Coding sequences immediately 5′ of the heptamer site have a modest effect on relative activity of R1H609L as measured with pJH200. pJH200 derivative used to measure activity is indicated above each graph. Otherwise the description and derivation of the graphs are the same as in Fig. 3.

orientation of the central 2 nucleotides, we did observe an additional but small (≈2-fold) effect by the presence of the *Sal* I site (Fig. 4). This effect was more apparent with full-length R1H609L. Nevertheless, a direct comparison of pJH289 and pJH200JV(SalI) still revealed on average a 2- to 3-fold difference in the recombination efficiency profiles (Fig. 6). This observation suggests that RAG-1 is also sensitive to the third coding nucleotide and sequences beyond those since they also subtly influence recombination efficiency.

DISCUSSION

We have serendipitously identified a RAG1 variant, R1H609L, the activity of which is sensitive to the coding nucleotides adjacent to the heptamer site of the RSS. The behavior of this mutant strongly implicates RAG1 as a component of the V(D)J recombinase that is important for substrate recognition and reiterates the important influence of coding-end content on recombination efficiency.

We do not yet know how the H609L mutation affects RAG1 structure or what aspect of DNA recognition is involved. These parameters include the efficiency of DNA binding, synapsis, and/or cleavage. Although we have shown an effect of coding-end sequences and not the RSS proper on the recombination activity of RAG1, the coding ends can be considered an extension of the RSS since they are an established part of the DNA sequence relevant to determining recombination efficiency. The effect of coding end sequence on the recombination activity by R1H609L reported here may be related to but is distinct from that observed with the wild-type V(D)J recombinase. In previous studies, plasmid recombination substrates that contained A or T adjacent to the heptamer site were shown to have a much lower frequency of recombination than those with G or C (8, 10, 28). We did observe a 2- to 3-fold inhibitory effect on the absolute recombination activity of R1WT by the presence of 5′-TG that replaced 5′-AC and 5′-CC when we compared pJH288 with pJH288JV and pJH200

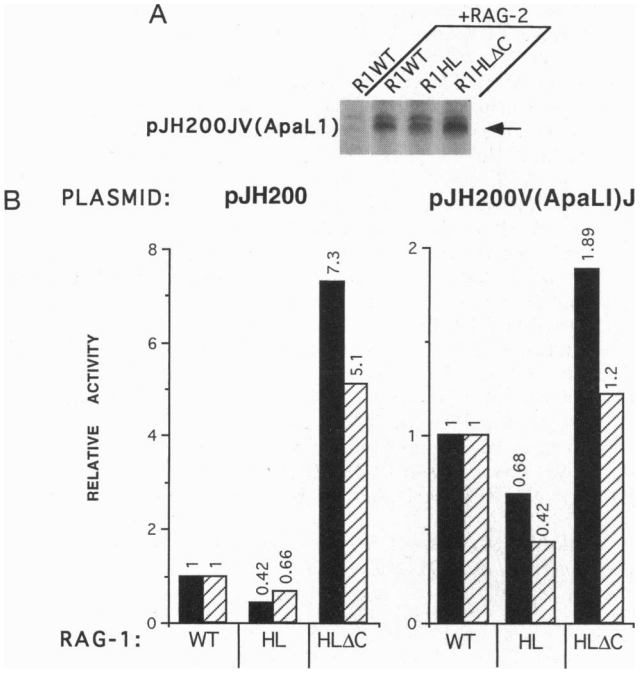

FIG. 5. Formation of excision products that contain coding joints is not differentially affected by R1H609L. (*A*) Autoradiograph of fragments amplified from excision products by PCR of Hirt supernatants from cells transfected with pJH200JV(ApaLI). (*B*) Bar graph comparing relative amounts of plasmid (▨) and excision (■) products of pJH200JV(ApaLI) and pJH200 as determined by PhosphorImaging.

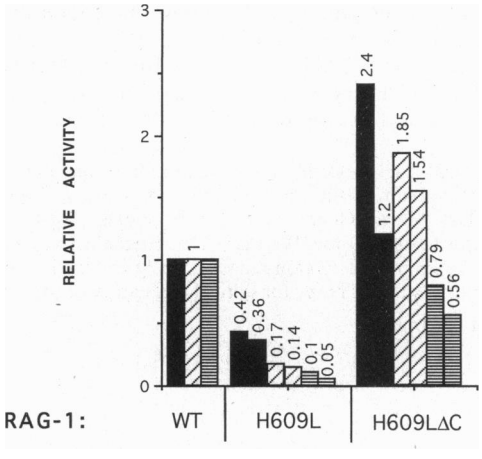

FIG. 6. Comparison of RAG1 activity with pJH289 and pJH200 derivatives. Bar graph comparing relative recombination profiles of pJH289 (▤), pJH200JV(ApaLI) (■), and pJH200JV(SalI) (▨). These plasmids share identical RSSs and the two most heptamer-proximal coding nucleotides.

with pJH200JV(ApaLI) (data not shown). Our data suggest that wild-type RAG1 may be responsible for the sequence sensitivity to those substrates.

One explanation for the behavior of the R1H609L variant is that this portion of RAG1 may directly contact DNA and that the leucine moiety is compatible with a different spectrum of nucleotide sequences than the histidine. Alternatively, the mutation may affect the conformation of the putative DNA binding/catalytic domain of RAG1 or interactions with another protein which itself is important for mediating these activities. Notably, a His-to-Ala substitution at this position did not elicit the same phenotype as R1H609L and behaved instead like R1WT (data not shown). Other alanine substitutions near the histidine residue either had no effect (E607A; data not shown) or reduced the activity of RAG1 with all substrates (K608A; data not shown).

However, while this manuscript was in preparation, a mutant of RAG1 was reported by Sadofsky *et al.* (29) that contains a 2 amino acid replacement of 4 amino acids that span positions 606–611. This mutant had a phenotype similar to but much more severe than the R1H609L variant. Nevertheless, the coding-end sequence dependence of the two mutants appears to be similar and therefore most likely affects the same aspect of substrate recognition. These two independently derived RAG1 mutants are so far unique and underscore the importance of this region in influencing DNA recognition.

Although the coding sequence composition had a 10-fold effect on the relative recombination activity elicited by R1H609L, we observed this strong a manifestation of coding-sequence alteration only within the context of pJH288. Reciprocal mutations in pJH200, which extended its coding-end identity with pJH288 and pJH289 to 6 nucleotides, did have a repressive effect on the recombination frequency of pJH200 by R1H609L. However, the effect was not as pronounced as with pJH288. Since the differences in recombination activities are relative, it is possible that the parental pJH200 vector itself may not contain the optimal sequences required for its full recombination potential and the effect of the nucleotide changes we introduced may not able to exert any more severe an effect in that context. Nevertheless, our results suggest that other DNA sequences outside of the RSS and the coding-end nucleotides can have a modest impact on recombination efficiency.

Our data provide strong evidence to support a model in which RAG1 is important for RSS recognition. RAG1 may directly contact the DNA molecule, or it may be a component of a DNA binding complex even though it may not contact DNA directly. We consider this latter possibility less likely given the subtlety of the mutations in both RAG1 and the recombination substrates. Although we have identified a role for RAG1 in DNA target site specificity, it does not preclude the possibility that RAG1 also contributes to other functions, such as DNA cleavage and activation or organization of other proteins involved in the complete recombination reaction.

We especially acknowledge the contribution of Dr. Patricia Cortes for an early version of the recombination PCR assay. We would like to thank Drs. Martin Gellert and Moshe Sadofsky for sharing information prior to publication. We thank Drs. Eugenia Spanopoulou and P. Cortes for immunodetection reagents, technical advice, and discussions. We thank Sara Cherry for critical evaluation of the manuscript; Ben Chen and Dr. Xiaolu Yang for valuable suggestions; and Drs. Dina Alexandropoulos, Joshy Jacob, and members of the Baltimore laboratory for discussions. We are also grateful to Pam Svec for excellent technical assistance. C.A.J.R. was supported by the Damon Runyon–Walter Winchell Cancer Research Fund and by the American Cancer Society. D.B. is an American Cancer Society Research Professor. This work was supported by National Institutes of Health Grant CA51462-08 to D.B.

1. Oettinger, M. A., Schatz, D. G., Gorka, C. & Baltimore, D. (1990) *Science* **248,** 1517–1523.
2. Schatz, D. G., Oettinger, M. A. & Baltimore, D. (1989) *Cell* **59,** 1035–1048.
3. Shinkai, Y., Rathbun, G., Lam, K.-P., Oltz, E. M., Stewart, V., Mendelsohn, M., Charron, J., Datta, M., Young, F., Stall, A. M. & Alt, F. W. (1992) *Cell* **68,** 855–867.
4. Spanopoulou, E., Roman, C. A. J., Corcoran, L. M., Schlissel, M. S., Silver, D., Nussenzweig, M., Shinton, S., Hardy, R. & Baltimore, D. (1994) *Genes Dev.* **8,** 1030–1042.
5. Mombaerts, P., Iacomini, J., Johnson, R. S., Herrup, K., Tonegawa, S. & Papaioannou, V. (1992) *Cell* **68,** 869–877.
6. Tonegawa, S. (1983) *Nature (London)* **302,** 575–581.
7. Hesse, J. E., Lieber, M. R., Mizuuchi, K. & Gellert, M. (1989) *Genes Dev.* **3,** 1053–1061.
8. Gerstein, R. M. & Lieber, M. R. (1993) *Genes Dev.* **7,** 1459–1469.
9. Gauss, G. H. & Lieber, M. R. (1992) *Genes Dev.* **6,** 1553–1561.
10. Ezekiel, U. R., Engler, P., Stern, D. & Storb, U. (1995) *Immunity* **2,** 381–389.
11. Roth, D. B., Nakajima, P. B., Menetski, J. P., Bosma, M. J. & Gellert, M. (1992) *Cell* **69,** 41–53.
12. Roth, D. B., Zhu, C. & Gellert, M. (1993) *Proc. Natl. Acad. Sci. USA* **90,** 10788–10792.
13. Roth, D. B., Menetski, J. P., Nakajima, P. B., Bosma, M. J. & Gellert, M. (1992) *Cell* **70,** 983–991.
14. Schlissel, M., Constantinescu, A., Morrow, T., Baxter, M. & Peng, A. (1993) *Genes Dev.* **7,** 2520–2532.
15. Zhu, C. & Roth, D. B. (1995) *Immunity* **2,** 101–112.
16. Taccioli, G., Rathbun, G., Shinkai, Y., Oltz, E. M., Cheng, H., Whitmore, G., Stamato, T., Jeggo, P. & Alt, F. W. (1992) *Curr. Top. Microbiol. Immunol.* **182,** 107–114.
17. Gent, D. C. v., McBlane, J. F., Ramsden, D. A., Sadofsky, M. J., Hesse, J. E. & Gellert, M. (1995) *Cell* **81,** 925–934.
18. Hesse, J. E., Lieber, M. R., Gellert, M. & Mizuuchi, K. (1987) *Cell* **49,** 775–783.
19. Lieber, M. R., Hesse, J. E., Lewis, S., Bosma, G. C., Rosenberg, N., Mizuuchi, K., Bosma, M. J. & Gellert, M. (1988) *Cell* **55,** 7–16.
20. Higuchi, R., Krummel, B. & Saiki, R. K. (1988) *Nucleic Acids Res.* **16,** 7351–7367.
21. Pear, W. S., Nolan, G. P., Scott, M. L. & Baltimore, D. (1993) *Proc. Natl. Acad. Sci. USA* **90,** 8392–8396.
22. Hirt, B. (1967) *J. Exp. Biol.* **26,** 365–369.
23. Sambrook, J., Fritsch, E. F. & Maniatis, T. (1989) *Molecular Cloning: A Laboratory Manual* (Cold Spring Harbor Lab. Press, Cold Spring Harbor, NY), 2nd Ed.
24. Schatz, D. G. & Baltimore, D. (1988) *Cell* **53,** 107–115.
25. Landau, N., Schatz, D. G., Rosa, M. & Baltimore, D. (1987) *Mol. Cell. Biol.* **7,** 3237–3243.
26. Sadofsky, M. J., Hesse, J. E., McBlane, J. F. & Gellert, M. (1993) *Nucleic Acids Res.* **21,** 5644–5650.
27. Silver, D. P., Spanopoulou, E., Mulligan, R. C. & Baltimore, D. (1993) *Proc. Natl. Acad. Sci. USA* **90,** 6100–6104.
28. Gerstein, R. M. & Lieber, M. R. (1992) *Nature (London)* **363,** 625–627.
29. Sadofsky, M., Hesse, J. E., van Gent, D. C. & Gellert, M. (1995) *Genes Dev.* **9,** 2193–2199.