Coding end sequence can markedly affect the initiation of *V(D)J* recombination

Rachel M. Gerstein and Michael R. Lieber¹

Laboratory of Experimental Oncology, Department of Pathology, Stanford University School of Medicine, Stanford, California 94305-5324 USA

In *V(D)J* recombination, two site-specific cuts are made adjacent to *V*, *D*, and *J* subexons to create four DNA ends, two of which (the coding ends) are joined to generate the exon that encodes the variable domain of the antigen receptor. Although deviations from consensus signal sequences have been reported previously to have a large impact on the efficiency of *V(D)J* recombination, coding end sequence has been assumed to be neutral with respect to the efficiency of recombination. We have used extrachromosomal *V(D)J* recombination substrates to undertake a systematic comparison of coding end sequences. Substrates were constructed that contain identical consensus recombination signal sequences, where only the coding ends were varied. Surprisingly, we found that nucleotide sequence at the coding end can affect the efficiency of *V(D)J* recombination >250-fold. Variable initiation of recombination appears to account for most of the effect. This finding has mechanistic implications because it indicates that signal-binding proteins involved in *V(D)J* recombination may have different levels of activity when confronted with coding ends of different sequence. Our results also indicate that coding end sequence must be considered to be among the major factors that shape the antigen receptor repertoire.

[*Key Words*: Site-specific recombination; immunoglobulin; mechanism; T-cell receptor; gene rearrangement]

Received April 16, 1993; revised version accepted June 1, 1993.

Developing mammalian lymphocytes generate an enormous diversity of antigen receptors. Much of this diversity results from the fact that the complete exons encoding the antigen-binding variable domains of immunoglobulins and T-cell receptors (TCRs) do not exist in the germ line. Rather, variable region exons are assembled by *V(D)J* recombination by joining *V* (variable), *D* (diversity), and *J* (joining) region subexons. Thus, the combinatorial diversity of numerous *V*, *D*, and *J* elements, and nucleotide loss and addition at the recombination junctions combine to generate a potential repertoire of 10²² murine antigen receptor molecules (Lieber 1991).

All of these potentially distinct *V(D)J* recombination reactions share the requirement for a recombination signal sequence. Signal sequences lie adjacent to each coding element and consist of a palindromic heptamer and an A/T-rich nonamer (Fig. 1). *V(D)J* recombination occurs between two gene elements: one with a 12-bp spacer between the heptamer and nonamer (12-signal), and the other with a 23-bp spacer between them (23-signal). The sequence directly adjacent to the heptamer, which ultimately forms the recombinant junction within the antigen receptor variable exon, is referred to as the coding end. Coding end sequence has been assumed to be neutral with respect to the efficiency of recombination because (1) the several hundred antigen receptor-coding el-

ement sequences are diverse, and (2) signal sequences, with no additional immunoglobulin or TCR locus sequence, are necessary and sufficient to direct *V(D)J* recombination of exogenous substrates (Akira et al. 1987; Hesse et al. 1987, 1989). Here, we report a systematic comparison of the contribution to recombination efficiency of coding end sequences. Contrary to expectation, we find that nucleotide sequence at the coding end can have a large impact on the efficiency of *V(D)J* recombination. This finding has both mechanistic and developmental relevance for generation of the antigen receptor repertoire.

Results

Experimental strategy

We evaluated the contribution of coding end sequence to *V(D)J* recombination efficiency, as measured by extrachromosomal recombination substrates (Hesse et al. 1987; Lieber et al. 1987). These substrates allow for a quantitative assessment of the fraction of plasmid molecules that have entered the cell and undergone *V(D)J* recombination. The substrates used in this study (Figs. 2–4) were constructed from the same parental plasmid, contain identical consensus recombination signals (Hesse et al. 1989), and vary only at the terminal 1–8 nucleotides of the coding ends. Except where noted, the

¹Corresponding author.

Figure 1. Basic $V(D)J$ reaction and elements of the mechanism. The essential features of the reaction are shown. The order of events before and after the four-ended intermediate is formed is still uncertain. The signals are symbolized as solid rectangles in the top line and as triangles below (the signal with the 12-bp spacer is depicted as an open triangle and the 23-bp spacer signal is shown as a shaded triangle). The template-independent nucleotide addition is due to terminal deoxynucleotidyl transferase (TdT). The coding subexons are symbolized as rectangles [adapted from Lieber 1991].

Abelson murine leukemia virus (AMLV)-transformed pre-B cell line 22D6 was transfected with these substrates. The signal joint retaining substrate pJH200 (Hesse et al. 1987) was always employed in separate, parallel transfections. The extent of recombination was determined after isolation of plasmid DNA from transfected cells. As described previously (Hesse et al. 1987; Lieber et al. 1987), a plating assay was used to assess the percentage of replicated, recovered molecules [*DpnI*-resistant molecules that give rise to ampicillin-resistant colonies (designated DA)] that are recombinants [ampicillin- and chloramphenicol-resistant colonies (designated AC)]. This percentage is referred to as R (Lieber et al. 1987). In addition, DNA from ampicillin- and chloramphenicol-resistant colonies was analyzed by restriction endonuclease digestion to estimate the fraction of colonies that represent actual $V(D)J$ coding joints or signal joints. Thus, R_{corr} is R multiplied by the fraction of actual joints. In this manner, substrates with very low recombination efficiencies, which fall below the false-positive background for the colony assay, can be measured accurately for $V(D)J$ recombination. The data are

presented as R' , which is defined as R_{corr} for a given substrate divided by R for pJH200 obtained in parallel transfections. This normalization of R values allows comparisons of different experiments.

Coding end sequence affects $V(D)J$ recombination efficiency

The most striking aspect of the comparison of substrates with different coding end sequences is that recombination efficiency varies over at least a 260-fold range (Table 1; Fig. 2). pRG21, pRG52, pRG47, and pRG18 are all greatly reduced for R' relative to pRG37 and pRG17. The least efficient recombination was observed with pRG21, with a mean R' value of 0.25. This value is 260 times less than 65, the mean value for pRG37. The pRG37 and pRG21 substrates are identical, except at five of the six terminal base pairs of the coding end attached to the 12-signal. We conclude that the interaction of substrates with the $V(D)J$ recombinase is influenced by the terminal 6 bp of the coding ends.

Figure 2. Comparison of R' values for coding joint substrates. All of the depicted substrates utilize the same plasmid backbone and the same consensus signal sequences (Hesse et al. 1989). Essential features of these recombination substrates have been described previously (Hesse et al. 1987; Lieber et al. 1987). $V(D)J$ recombination signal sequences are separated by a stop sequence, the prokaryotic transcription terminator *oop*, and are placed between the chloramphenicol acetyltransferase gene (*cat*) and a prokaryotic promoter (P_{lac}). A small portion of the coding end nucleotide sequence is displayed. After excision of the signals and *oop* during $V(D)J$ recombination in mammalian cells, the recombinant plasmid expresses *cat* upon transformation of *E. coli*. Substrates contain the β -lactamase gene, which confers ampicillin resistance in *E. coli*. Therefore, recombinants can be selected on ampicillin/chloramphenicol plates. The mean R' values ($\pm$ S.E.M.) for each substrate are presented. The raw data are presented in Table 1.

Coding end sequence alterations appear to affect primarily reaction initiation, not coding joint resolution

The basic elements of the $V(D)J$ recombination mechanism are outlined in Figure 1. The reaction can be discussed in terms of two phases, initiation and resolution. Initiation includes binding of the recombinase to each of the signal sequences, interaction of the sites (synapsis), and cutting at the signal ends. Resolution includes processing of the coding ends, nucleotide addition, and joining of signal joints and coding joints.

If resolution was the basis for the poor recombination efficiency of pRG21, we predicted that coding joint formation would be impaired but that signal joint formation by a substrate with the same signals and coding ends, pRG35 (Fig. 3), would be unaffected. This test to distinguish sequence effects on coding joint resolution versus initiation of the reaction relies on the fact that coding joint and signal joint formation are not coupled (Lieber et al. 1988; Sheehan and Lieber 1993). Hence, inhibition of coding joint resolution does not affect signal joint resolution. Therefore, if a sequence alteration affects both coding joint and signal joint formation, that sequence alteration is likely affecting the efficiency with which the recombination activity initiates the reaction.

The results of such a test are quite clear. Both pRG21 and pRG35, coding joint-retaining and signal joint-retaining substrates with identical coding ends, are $<1\%$ as efficient as pJH200 (Table 2; Fig. 3). Similar results are obtained with pRG18 ($R' = 3$) and its signal joint-retain-

ing analog pRG40 ($R' = 1.1$) (Table 2; Fig. 3); both are impaired for recombination and are not significantly different from each other. Because coding joint-retaining substrates and signal joint-retaining substrates with particular coding ends both recombine inefficiently, we conclude that the major effect of coding end sequence is upon initiation of $V(D)J$ recombination.

The effect of coding end sequence on $V(D)J$ recombination is not restricted to the 22D6 cell line. Experiments done in the AMLV-transformed cell line, 1-8, yielded similar results (Table 3) to those in the 22D6 cell line. R' of pRG21 is very low in 1-8 cells ($R' = 0.59$). This value is not significantly different from the value obtained in 22D6. It is important to note that pRG35 is also relatively low in 1-8 ($R' = 5$). This indicates in a second cell line that initiation of the reaction is also affected by coding end sequence.

$V(D)J$ recombination is sensitive to the nucleotide content of the coding end

How does coding end sequence affect initiation? The largest quantitative effect is associated with the coding end attached to the 12-signal in pRG21. The substrate pRG37, which has the same coding end associated with the 23-signal as pRG21, recombines efficiently ($R' = 65$) (Table 1; Fig. 2). This indicates that the coding end attached to the 23-signal is not responsible for the low level of recombination in pRG21.

The substrate pRG18, with the same 12-signal coding end (defined as the coding end attached to the 12-signal in the substrate) as pRG21, has a low R' value ($R' = 3$) (Table 1; Fig. 2). This suggests that the coding end attached to the 12-signal is primarily responsible for the low R' of pRG21 and pRG18. It should be noted that other than the 23-signal coding end, the rest of the substrate is identical between pRG21 and pRG18. Interestingly, the R' value for pRG18 value is 12-fold higher than pRG21 ($P = 0.005$ by permutation test). This may indicate that the nucleotide sequence of the 23-signal coding end can also affect recombination because the only sequence difference between pRG18 and pRG21 is at the last 7 bp of the coding end attached to the 23-signal. Further experiments will explore the contribution of the 23-signal coding end sequence to the efficiency of recombination, as the majority of substrates used in this study differ only at the 12-signal coding end. Neither of two single-base-pair changes within the 12-signal coding end restores efficient recombination: pRG46 ($R' = 1.2$) has CTT-heptamer and pRG44 ($R' = 2.3$) has GTT-heptamer (Table 4). We conclude that the remarkably low level of $V(D)J$ recombination by these substrates (pRG21, pRG46, and pRG44) cannot be attributed to a continuous 8-bp stretch of AT-rich DNA.

Additional studies continued to demonstrate the sensitivity of $V(D)J$ recombination to the nucleotide content of the coding end proximal to the 12-signal heptamer (Fig. 4; Table 5). The substrate pRG48 places a C next to the heptamer. The mean R' obtained with pRG48 is 1.4, a value 5.6-fold higher than the mean R' value of pRG21.

Table 1. Coding end sequence affects V(D)J recombination efficiency

Substrate	Expt. no.	AC ^a	DA ^b	R (%) ^c	R ₂₀₀ (%) ^d	R' ^e ($R_{\text{corr}}/R_{200} \times 10^2$)	Mean R' $\pm$ s.d.
pRG21	1	52	79,610	0.065	5	<0.044	0.25 $\pm$ 0.29
	2	32	113,886	0.028	5.2	<0.017	
	3	18	104,025	0.017	5.2	0.041	
	4	29	114,266	0.025	2.2	<0.088	
	5	42	112,461	0.037	2.2	0.70	
	6	21	101,688	0.021	2.2	0.62	
	7	69	148,846	0.046	2.2	0.24	
pRG52	1	168	28,557	0.59	5.6	9.6	5.5 $\pm$ 5.5
	2	186	22,667	0.82	5.6	15	
	3	68	27,607	0.25	5.6	4	
	4	678	536,505	0.13	1.9	0.83	
	5	76	106,918	0.071	1.9	0.94	
	6	641	527,779	0.12	1.9	2.8	
pRG47	1	176	49,020	0.36	5.7	6.3	4.4 $\pm$ 3.1
	2	17	3,325	0.51	5.7	9	
	3	59	37,073	0.16	5.7	1.9	
	4	30	13,660	0.22	5.7	3.2	
	5	21	5,605	0.37	5.7	1.8	
pRG18	1	142	70,044	0.2	4	5.0	2.9 $\pm$ 2.1
	2	36	60,370	0.06	4	1.3	
	3	37	66,074	0.056	4	0.14	
	4	141	55,370	0.025	4	0.64	
	5	306	285,508	0.11	1.5	4	
	6	239	193,828	0.12	1.5	4.7	
	7	382	264,786	0.14	1.5	4.3	
pRG37	1	1796	91,108	2	4	49	65 $\pm$ 14
	2	1060	31,716	3.3	4	83	
	3	1605	61,225	2.6	4	65	
	4	1708	70,080	2.4	4	61	
pRG17	1	2220	125,131	1.8	3	59	64 $\pm$ 16
	2	2620	121,546	2.2	3	71	
	3	1329	54,480	2.4	3	81	
	4	1260	93,448	1.3	3.1	43	

^aThe number of transformants arising on ampicillin–chloramphenicol plates (AC) from DNA recovered from transfected 22D6 pre-B cells.

^bThe number of transformants arising on ampicillin plates from *DpnI*-treated DNA recovered from transfected cells (DA).

^cThe recombination value, *R*, was determined from AC divided by DA, multiplied by 100.

^dThe *R* value obtained in parallel transfections with the signal joint-retaining substrate pJH200 (Hesse et al. 1987).

^eThe normalized *R* value, *R'*, is R_{corr} divided by R_{200} , multiplied by 100. The corrected *R* value, R_{corr} , is the *R* value obtained, multiplied by the fraction of actual coding joints (see Materials and methods). *R'* values preceded by a "less than" symbol indicate that no actual coding joints were found among the analyzed samples, and that the fraction of actual coding joints used in the calculation of *R'* was <1 divided by the number of analyzed samples. In these cases, the *R'* value reported may be an overestimate.

This difference is statistically significant ($P = 0.0035$, permutation test). This result indicates that the nucleotide immediately adjacent to the heptamer affects recombination. In addition, because the increase of pRG48 relative to pRG21 is not dramatic, this result suggests that the terminal nucleotide is not the only influence of coding end sequence on recombination efficiency. The 12-signal end of pRG36 has three C's between the 12-signal coding end of pRG21 and the heptamer. The *R'* value is elevated even further, as the mean *R'* of 16 for pRG36 is 63-fold higher than that of pRG21. The result obtained with pRG52 also indicates the importance of

the sequence adjacent to the heptamer. The nucleotide content of the coding end attached to the 12-signal in pRG52 is the same as pRG21, but the DNA strand polarity is reversed so that the coding end is AAA-heptamer. The mean *R'* value obtained with pRG52 is 5.5, which is 22-fold greater than the mean value of pRG21 ($P = 0.00116$) (Table 1; Fig. 4).

A final demonstration of the importance of the coding end nucleotide directly adjacent to the heptamer is demonstrated further by pRG49, where the coding end A next to the heptamer in pRG52 is replaced with a T (Fig. 4). This single-nucleotide change results in a mean *R'*

Figure 3. Comparison of R' values for coding joint- and signal joint-retaining substrates. Pairs of substrates within boxes are coding joint-retaining and signal joint-retaining substrates with identical coding ends. The mean R' values ($\pm$ S.E.M.) for each substrate are presented. The raw data are presented in Tables 1, 2, and 5.

value of 0.87 for pRG49, which is 6.3-fold lower than the mean value obtained with pRG52 ($P = 0.00198$) (Table 5). This result suggests that a T adjacent to the heptamer of the 12-signal is sufficient to lower recombination efficiency significantly.

Does the presence or absence of homology impact on the efficiency of recombination?

One potential reason for a diminished R' for pRG21 could be the inability of the two coding ends to base-pair with each other during joining. In a separate study, we have demonstrated that coding joint formation can use 4 nucleotide blocks of homology located at the coding ends; however, a substrate with 4 nucleotides of homology did not confer a higher R' when compared with a substrate with the potential to pair only 1 or 2 nucleotides between coding ends (Gerstein and Lieber 1993). In this study we have assessed the recombination efficiency of several substrates in which there is no potential for base-pairing between the first 8–9 nucleotides adjacent to the heptamer. As indicated above, we have shown that such substrates have low recombination efficiency, regardless of whether the substrate retains signal joints or coding joints. Thus, we conclude that the effect of coding end sequence is primarily upon initiation of $V(D)J$ recombination and that the lack of potential for base-pairing between the coding ends does not limit recombination.

An additional test of the influence of homology in determining $V(D)J$ recombination efficiency was conducted. The addition of CCC to the coding end of pRG21 to make pRG36 (Fig. 3) changes the nucleotide content next to the heptamer and provides the 12-signal coding end with 3 nucleotides of homology to the 23-signal coding end. If resolution of this substrate is made more efficient by homology at coding joints, we would expect to see a lower R' for signal joint formation than for coding

joint formation. This prediction is not verified: pRG36 has a mean R' value of 16, which is quite similar to the R' value for its signal joint-retaining version, pRG41 ($R' = 29$) (Fig. 3; Tables 2 and 5). In addition, if coding joint resolution depended on homology to align coding ends, we predicted that coding junctions from pRG36 would be highly enriched for homology usage. This is not the case, as sequence analysis of 17 recombinants isolated from 22D6 cells transfected with pRG36 reveals that only three junctions (18%) used 2 or 3 nucleotides of homology at the junction (data not shown).

If homology has a role in dictating recombination frequency, pRG42 might also be expected to recombine more efficiently than pRG21. The substrate pRG42 has 3 T nucleotides added to the 23-signal coding end, providing 3 nucleotides of homology with the 12-signal coding end. Increased recombination did not occur, as the mean R' value obtained with pRG42 is 0.38, compared with 0.25 for pRG21 (not significantly different) (Fig. 4; Table 5). In pRG30, the 23-signal coding end also has more homology to the 12-signal coding end than does pRG21 (Table 4). This substrate has the same 12-signal coding end as pRG21. The R' value obtained with pRG30 (0.6) is also not significantly greater than that of pRG21 (0.25). These results demonstrate that homology is not sufficient to overcome the low recombination efficiency observed with pRG21, pRG30, and pRG42.

In addition to the assessment of recombination efficiency, we determined the nucleotide sequence of coding joints retained by substrates with no potential for base-pairing between coding ends. The sequence analysis demonstrates that these substrates form coding joints that do not differ from coding joints formed by other coding ends (Fig. 5). If homology were essential to coding joint resolution, either coding joints would not form or coding ends would always suffer nucleotide loss sufficient to expose homology. This is not the case among coding joints formed by the two substrates that we have examined in detail. These results indicate that homology between coding ends is not necessary for coding joint formation to occur.

As the cell line in which these coding joints were formed (22D6) has low levels of TdT activity, we cannot rule out the possibility that TdT additions have fortuitously generated homology that was used for resolution. However, if resolution of these coding joints were entirely dependent on N region addition, then we would expect to observe a higher incidence of N regions among the total population of coding joints. This is not the case, as there are N insertions in only 17% of the pRG47 coding joints and in 8% of the coding joints formed by pRG52. These frequencies are similar to the frequency of N region additions observed with other substrates (Gerstein and Lieber 1993). This makes it highly unlikely that N region additions are making homology utilization undetectable.

It is interesting to note that pRG52 and pRG47, which have AAA-heptamer and TAA-heptamer, respectively, are both >10-fold less than pRG37 and pRG17 (Fig. 2; Table 1). (For pRG52 compared with pRG17, $P = 0.0095$;

Table 2. Signal joint-retaining substrates

Substrate	Expt no.	AC	DA	R (%)	R ₂₀₀ (%)	R' (R _{corr} /R ₂₀₀ × 10 ²)	Mean R' ± S.D.
pRG35	1	23	67,488	0.034	4	0.56	1 ± 8
	2	53	84,398	0.063	4	0.69	
	3	68	61,750	0.11	4	2.7	
	4	20	61,826	0.032	4	0.79	
	5	974	567,420	0.17	1.5	1.3	
	6	746	631,610	0.12	1.5	<0.28	
	7	605	505,680	0.12	1.5	0.89	
pRG40	1	161	187,800	0.086	6.5	1.1	1.1
	2	83	109,562	0.076	6.5	0.49	±0.56
	3	186	152,301	0.12	6.5	1.6	
pRG39	1	3321	93,051	3.6	7.8	46	34
	2	2864	72,912	3.9	7.8	38	±14
	3	518	32,438	1.6	7.8	18	
pRG41	1	323	25,872	1.2	7.8	16	30
	2	577	18,816	3.1	7.8	39	±12
	3	1090	42,532	2.6	7.8	33	

for pRG47 compared with pRG17, $P = 0.016$). Additionally, both pRG52 and pRG47 are ~20-fold more efficient than pRG21. We favor the possibility that the sequences at the 12-signal coding end of pRG52 and pRG47 significantly inhibit the initiation of recombination but to a lesser degree than the 12-signal coding end of pRG21. For example, there may be a preference for the terminal nucleotide at coding ends (G,C > A > T). As homology is not sufficient to direct efficient recombination (see above), as indicated by pRG42, and is not essential for coding joint formation, we doubt that the lack of homol-

ogy between the coding ends of pRG52, and between those of pRG47, is likely to be the factor in the determining the lower R' values of these substrates relative to pRG37 and pRG17.

The above experiments indicate that homology between coding ends does not affect the efficiency of $V(D)/J$ recombination. Taken together with the results that demonstrate the sensitivity of recombination to the nucleotide immediately adjacent to the heptamer of the 12-signal without changes to the 23-signal coding end, and also the equivalence in efficiency of signal joint-

Table 3. 1-8 Cell data: coding joint- and signal joint-retaining substrates

Substrate	Expt. no.	AC	DA	R (%)	R ₂₀₀ (%)	R' (R _{corr} /R ₂₀₀ × 10 ²)	Mean R' ± S.D.
pRG21	1	119	111,962	0.11	2.6	1.3	0.59 ± 0.41
	2	21	42,039	0.05	5.3	0.54	
	3	17	19,760	0.086	2.5	1	
	4	42	169,931	0.025	3	0.33	
	5	59	140,945	0.042	3	0.42	
	6	67	198,895	0.034	3	0.33	
	7	60	292,962	0.02	3	0.2	
pRG35	1	63	26,296	0.24	4.6	5.2	5 ± 0.45
	2	44	18,240	0.24	4.6	5.2	
	3	70	34,200	0.2	4.6	4.3	
	4	63	26,448	0.24	4.6	5.2	
pRG47	1	127	108,528	0.12	0.819	11	8.6 ± 1.9
	2	143	183,236	0.078	0.819	8.5	
	3	118	115,520	0.1	0.819	8.1	
	4	119	118,864	0.1	0.819	6.5	
pRG46	1	62	66,272	0.093	0.819	1.4	1.7 ± 0.83
	2	55	127,072	0.043	0.819	1.1	
	3	57	103,892	0.055	0.819	1.5	
	4	76	118,408	0.064	0.819	2.9	

Table 4. Coding joint-retaining substrates

Substrate	Expt. no.	AC	DA	R (%)	R ₂₀₀ (%)	R' (R _{corr} / R ₂₀₀ × 10 ²)	Mean R' ± s.d.
TTTAAGTT▶◀CCCCGGCC							
pRG44	1	81	127,721	0.063	3.9	1.4	2.2 ±1.3
	2	157	141,274	0.11	3.9	1.5	
	3	195	142,884	0.14	3.9	2.4	
	4	208	151,743	0.14	3.9	1.5	
	5	124	93,492	0.13	4.3	1.7	
	6	358	146,804	0.24	4.3	5	
	7	372	173,852	0.21	4.3	2.2	
TTTAAC TT▶◀CCCCGGCC							
pRG46	1	62	28,124	0.22	5.7	3.8	1.2 ±1.8
	2	19	19,920	0.095	5.7	0.56	
	3	14	26,243	0.053	5.7	0.28	
	4	28	47,196	0.059	5.7	0.07	
TTTAATTT▶◀CCCTATAT							
pRG30	1	18	169,050	0.011	5	0.08	0.6 ±0.49
	2	12	100,637	0.02	5	0.099	
	3	243	159,936	0.15	4.3	0.78	
	4	177	114,072	0.15	4.3	0.21	
	5	126	82,320	0.15	4.3	0.19	

retaining substrates and coding joint-retaining substrates, it seems clear that coding end sequence affects primarily initiation of *V(D)J* recombination.

Discussion

Coding ends have previously been assumed to be neutral participants in *V(D)J* recombination. Here, we have tested this assumption, and contrary to our expectation, we found that the nucleotide sequence at the coding ends participating in the reaction can affect the efficiency of *V(D)J* recombination up to 260-fold. A large proportion of the effect of coding end sequence on *V(D)J* recombination appears to affect initiation of the reaction. Consistent with an effect on initiation, we found that homology between coding ends is not essential for

recombination and does not override the reduction in efficiency imparted by certain coding ends. In this study coding ends with TTT-heptamer-12 spacer imparted the largest effect. Substrates with T-heptamer or AAA-heptamer adjacent to the 12-signal also had a significant diminution of recombination.

Initiation of recombination can be divided further into stages: (1) signal-binding protein recognition of signals; (2) synapsis; and (3) cleavage. One potential explanation of sequence-specific effects on initiation is that the signal-binding proteins are sensitive to the nucleotide immediately adjacent to the heptamer such that an overt nucleotide preference exists (such as G,C > A > T). This preference could be manifest at the initial protein-DNA interaction, or at cleavage. An analogy can be drawn to *EcoRI*: The base pairs next to the hexameric recognition sequence contribute to the binding affinity and also to the cleavage rates (Maass 1987 and references within). Once signal-binding proteins have been identified that are clearly shown to be involved in *V(D)J* recombination, it will be of interest to test the binding of these proteins to substrates containing coding ends with varying nucleotide content. Signal-binding proteins that are still under evaluation include those reported by Aguilera et al. (1987), Matsunami et al. (1989), and Shirakata et al. (1991).

Another possibility is that signal-binding proteins do not perceive coding end sequence directly. In this case, the coding end nucleotides may affect the local DNA structure of the heptamer such that binding or cleavage is affected. It has been noted that the sequence GTG has unusual structural attributes (Gellert 1992). This trinucleotide

Figure 4. Comparison of *R'* values for coding joint substrates. The mean *R'* values (±S.E.M.) for each substrate are presented. The raw data are presented in Tables 1 and 5.

Table 5. Coding joint-retaining substrates with different coding ends

Substrate	Expt. no.	AC	DA	R (%)	R ₂₀₀ (%)	R' (R _{corr} /R ₂₀₀ × 10 ²)	Mean R' ± S.D.
pRG48	1	57	74,556	0.076	3	1.4	1.4 ±0.43
	2	179	235,760	0.076	3	1.6	
	3	79	254,084	0.031	2.1	0.53	
	4	173	303,256	0.057	2.1	1.5	
	5	118	168,276	0.07	1.9	1.7	
	6	147	302,009	0.049	1.9	1.5	
pRG36	1	722	95,988	0.75	4.8	16	16 ±2.3
	2	984	133,570	0.74	4.8	15	
	3	1344	161,747	0.83	4.8	17	
	4	678	104,367	0.65	4.8	12	
	5	1358	100,352	0.35	4.3	18	
pRG42	1	22	84,987	0.026	1.5	0.14	0.38 ±0.43
	2	33	39,520	0.083	1.5	<0.37	
	3	64	112,290	0.057	4.2	0.9	
	4	16	80,332	0.02	4.2	0.24	
	5	36	115,102	0.031	4.2	0.28	
	6	65	166,440	0.039	4.2	0.33	
pRG49	1	90	159,056	0.057	3	1.2	0.87 ±0.5
	2	40	137,624	0.029	3	0.48	
	3	65	256,520	0.025	2.1	0.096	
	4	53	134,983	0.039	2.1	1.2	
	5	81	148,540	0.054	1.9	1.6	
	6	94	125,538	0.075	1.9	0.71	
	7	102	251,811	0.04	1.9	0.77	

cleotide in the context of a larger oligonucleotide displays partial base unstacking, lower melting temperature, base-pair tilt, and displacement from the DNA helical axis (Cheung et al. 1984; Patel et al. 1988). These unique structural features may be important for recognition or cleavage by the *V(D)J* recombinase. It is conceivable that the nucleotides next to the heptamer affect this DNA structure, as pairs of adjacent nucleotides have different twist profiles depending on nucleotide content (Dickerson 1992 and references therein).

There are several specific protein–DNA interactions where the nucleotide sequence outside of the region of DNA that is in direct contact with protein affects the interaction. The 4-bp core region of both the bacteriophage 434 and the bacteriophage P22 operators is not in direct contact with its repressor protein. Nonetheless, changing nucleotides within this core region affects the affinity of these repressors for their operator (Koudelka and Carlson 1992; Wu et al. 1992). Experiments indicate that the structure of the unbound operator is affected by the sequence of this core region. Similarly, mutational analysis of the *hisG1* insertion hot spot for Tn10 transposition showed that the 6–9 bp that flank the insertion consensus sequence confer up to 1000-fold changes in insertion frequency (Bender and Kleckner 1992). Because no broadening of the Tn10 consensus sequence was evident from these results, sequence-dependent DNA structure adjacent to the region of transposase–hot spot interaction is likely to be responsible for the observed effect (Bender and Kleckner 1992). Also, the flanking

host sequence directly adjacent to bacteriophage Mu transposons has also been observed to compromise the ability of Mu regions to serve as substrates for cleavage in the transposition reaction (Wu and Chaconas 1992).

This study demonstrates that coding end sequence can be a critical factor in determining *V(D)J* recombination efficiency. The antigen receptor repertoire is not generated in a random manner, with each gene element rearranging with equal frequency. Rather, many biases have been observed (see below). Our results indicate that another parameter, coding end sequence, should be assessed among the major forces that determine these biases. Coding end sequence appears to be important in the mechanistic bias of deletional over inversional D_H/H joining. Examination of the 11 germ-line sequences available for D_H (which has two 12-signal sequences) shows that all 11 end in AC-heptamer at the 3' signal, whereas seven have heptamer-T (5'-CACTGTG-T-3') at the 5' signal. This sequence should be functionally equivalent to A-heptamer (5'-A-CACAGTG-3'), which, in this study, we have observed to be reduced for *V(D)J* recombination efficiency. DFL16.1 is predicted to be affected most by coding end sequence, as it has heptamer-TTTATT at the 5' signal. This coding end is very similar to pRG47, which has AATAA-heptamer, and to pRG52, which has AATTAAA-heptamer. In a separate study, where extrachromosomal *V(D)J* substrates were used to recapitulate D_H/H joining, the deletion/inversion ratios that were observed (Gauss and Lieber 1992) correlate well with the coding end sequence effects documented

Figure 5. Coding joint sequences from pRG47 and pRG52. Coding joints were isolated from 22D6 cells transfected with pRG47 or pRG52. The number of nucleotides lost from each coding end are shown. N insertions are indicated by uppercase letters, and putative P nucleotides (Lafaille et al. 1989) are indicated by lowercase letters.

in this study. For example, a substrate with DFL16.1 joining to an idealized J segment favored deletion 19 : 1. A substrate with the inverse orientation of DFL16.1, which places AATAAA-heptamer at the 3' signal no longer favors deletion (the deletion/inversion ratio is 0.36). Similarly, pD_iJ_i, with AATT-heptamer at its 3' signal, has a deletion/inversion ratio of 0.95, whereas the opposite orientation, with AATT-heptamer at its 5' signal, has a deletion/inversion ratio of 2.4. These results are very consistent with our observation that T-heptamer (same as heptamer-A) is inefficient for initiation of recombination, and that A-heptamer is better, though significantly less than optimal. Endogenous D_H elements may favor deletion, at least in part, because of the sequence content of the coding end associated with the 5' signal, which creates a bias away from frequent initiation at the 5' signal, leading to infrequent inversion.

In light of the relative inefficiency imposed on $V(D)J$ recombination by particular coding ends, it is interesting to note that T-heptamer is significantly under-represented in genomic antigen receptor 12-signal-associated coding ends (V_K , J_β , J_γ , J_δ , D_β , $5'D_\delta$, J_α , D_H). Of 96 such coding ends examined, only 10 have T-heptamer (or hep-

tamer-A; 10.4%). Additionally, 7 of these 10 elements determine the first nucleotide of a codon and, hence, may have other selective factors favoring their evolutionary persistence at the DNA level. Two of the remaining three examples of heptamer-A are at $J_\gamma 1$ and $J_\gamma 2$, where the sequence AT seems to be involved in the generation of a high frequency of certain junctions during T-cell development (Lafaille et al. 1989; Asarnow et al. 1993). It seems apparent that the T nucleotide next to the heptamer has been selected against because of potential inefficiency in $V(D)J$ recombination.

Several biases in the antigen receptor repertoire have been reported previously. In addition to the bias of deletion over inversion at D_H , the V_H repertoire favors J_H proximal V_H genes early in B-cell ontogeny (Yancopoulos et al. 1984; Perlmutter et al. 1985; Malynn et al. 1990), particular TCR combinations are common in fetal development, particularly in anatomic locations (Lafaille et al. 1989 and references therein), and V_K gene usage is non-random (Lawler et al. 1989; Kalled and Brodeur 1990, 1991). The frequency of any given $V(D)J$ recombinant on endogenous chromosomes is subject to a complex combination of many forces, such as cellular selection, chromatin accessibility, recombination signal strength, and the impact of homology on coding joint formation and diversity. The current study adds another major component, the effect of coding end sequence on efficiency of the recombination reaction.

Materials and methods

Recombination substrates

All substrates were derived from pML134, a derivative of pJH298 (Lieber et al. 1988). A cassette design for modification of coding end sequence was utilized as follows: oligonucleotide RG 1,2 was cloned into the *SalI* site of pML134 (this intermediate is referred to as pRG15). This oligonucleotide contains an *XhoI* site within the coding end and also within the 12-bp spacer region. Cassettes with *XhoI*-compatible ends were then inserted to modify the 12-signal coding end. Analogous steps were used to modify the 23-signal coding end: oligonucleotide RG 3,4 was cloned into the *BamHI* site of pRG15 to produce pRG17. This oligonucleotide contains a *MluI* site within the coding end and also within the 23-bp spacer. pRG18 was made by the insertion of RG 5,6 into the *XhoI* site of pRG17. pRG21 has RG 9,10 cloned into the *MluI* site of pRG18. pRG30 has RG 11,12 at the *MluI* site of pRG18. pRG36 has RG 17,18 cloned into the *XhoI* site of pRG21. pRG37 has RG 1,2 cloned into the *SalI* site of pRG21. pRG42 has RG 23,24 at the *MluI* site of pRG18. pRG44 has RG 19,20 at the *XhoI* site of pRG21. pRG46 has RG 27,28 cloned into the *XhoI* site of pRG21. pRG47 has RG 25,26 cloned into the *XhoI* site of pRG21. pRG48 has RG 38,39 cloned into the *XhoI* site of pRG21. pRG49 has RG 40,41 cloned into the *XhoI* site of pRG21. pRG52 has RG 30,31 cloned into the *XhoI* site of pRG21. The 16 nucleotides of coding end most proximal to the signal sequence in oligonucleotide RG 1,2 is identical to the equivalent region of murine D_H element DFL16.1. The 17 nucleotides of coding end most proximal to the signal sequence in oligonucleotide RG 3,4 is identical to the sequence of murine $J_H 1$. No other oligonucleotides used are exactly equivalent to known endogenous antigen receptor coding ends.

Signal joint-retaining substrates were constructed as follows: pRG35 was made by digesting pRG21 with *Sall*, isolating the liberated fragment, and ligating this fragment in the inverted orientation (this intermediate is referred to as pRG34). Next, pRG34 was digested with *Bam*HI, and the *Bam*HI fragment from pRG21 was ligated in the inverse orientation. A similar strategy was used for the other signal joint substrates. pRG39 was made by first inserting RG 1,2 into pRG21 (the intermediate is pRG38) and then cloning the *Bam*HI fragment of pRG21 into pRG38 in the opposite orientation. pRG40 has RG 3,4 cloned into pRG34 at *Bam*HI. pRG41 was made by isolating the *Sall* fragment of pRG36 and ligating this fragment into pRG35 at the *Sall* site.

The sequences of the oligonucleotides used in the constructions of the above substrates are listed below. To facilitate cloning, these oligonucleotides were synthesized (Applied Biosystems 380A synthesizer) and then annealed, leaving *Sall*-, *Bam*HI-, *Xho*I-, or *Mlu*I-compatible 5' overhangs. The heptamer and nonamer sequences are indicated in boldface type. The coding end sequences are underlined. RG1: 5'-TCGACTG-CAGGGTTTTGTTCCAGTCTCGAGCACTGTGGTAGCT-ACTACCGTAGCTCGAG-3'. RG2: 5'-TCGACTCGAGCTA-CGGTAGTAGCTACCAAGTGTCTCGAGACTGGAACAAA-AACCTGACAG-3'. RG3: 5'-GATCCACGCGTCATCGAAG-TACCACTAGCACAGTGGTAGTACGCGTCTGTCTGGCTGTACAAAAACCATG-3'. RG4: 5'-GATCCATGGTTTTTGT-ACAGCCAGACAGACGCGTACTACCACTGTGCTACTGG-TACTTCGATGACGCGTG-3'. RG5: 5'-TCGAGTTTAAATT-CACAGTGC-3'. RG6: 5'-TCGAGCACTGTGAAATTAAC-3'. RG9: 5'-CGCGTACTACCACTGTGCGCGGCA-3'. RG10: 5'-CGCGTGGCGGCGCACAGTGGTAGTA-3'. RG11: 5'-CG-CGTACTACCACTGTGCGCTATATGTTAACA-3'. RG12: 5'-CGCGTGTAAACATATAGGGCACAGTGGTAGTA-3'. RG17: 5'-TCGAGTTTAAATTTCCCCACAGTGG-3'. RG18: 5'-TCG-ACCACTGTGGGAAATTAAC-3'. RG19: 5'-TCGACTTT-AAGTTACAGTGC-3'. RG20: 5'-TCGAGCACTGTGAAC-TAAAG-3'. RG23: 5'-CGCGTACTACCACTGTGTTTCCCG-GCCT-3'. RG24: 5'-CGCGAGGCGCGGAAACACAGTGGTA-GTA-3'. RG25: 5'-TCGACGTAATAAACACAGTGC-3'. RG26: 5'-TCGAGCACTGTG TTTATTACG-3'. RG27: 5'-TCGACT-TTAACTTCACAGTGC-3'. RG28: 5'-TCGAGCACTGTGAA-GTTAAAG-3'. RG30: 5'-TCGACAAATTAACACAGTGC-3'. RG31: 5'-TCGAGCACTGTGTTTAAATTTG-3'. RG38: 5'-TCGACGCGTTTAAATTCACAGTGG-3'. RG39: 5'-TCGAC-CACTGTGGAATTAACGCG-3'. RG40: 5'-TCGACGCGTA-AATTAATCACAGTGG-3'. RG41: 5'-TCGACCACTGTGAT-TAATTTACGCG-3'.

V(D)J recombination assay

The V(D)J recombination assay has been described previously (Hesse et al. 1987) but is summarized briefly below. Plasmid substrates are transfected into the AMLV-transformed murine pre-B-lymphoid cell lines 22D6 or 1-8. These cell lines were obtained from Naomi Rosenberg (Tufts University Medical School, Boston, MA) and have been described previously using extrachromosomal V(D)J recombination substrates (Lieber et al. 1987). While resident in the cells, a fraction of the substrate molecules undergo V(D)J recombination. This recombination results in the deletion of a prokaryotic transcription terminator. In the absence of recombination, the terminator interferes with the expression, in *Escherichia coli*, of a downstream chloramphenicol acetyltransferase (*cat*) gene. Upon recovery from the murine cells by the rapid alkaline/SDS lysis method and transformation into *E. coli*, recombinant plasmids will confer resistance to both ampicillin and chloramphenicol while unrecom-

bined substrates will confer only ampicillin resistance. The ratio of ampicillin/chloramphenicol-resistant (AC) colonies to ampicillin (A) colonies reflects the fraction of substrate that underwent recombination while resident in the eukaryotic cells. This ratio was corrected for DNA entry into the lymphoid cells as follows (Lieber et al. 1987). The substrates used in this study bear a polyoma origin of replication and the polyoma T antigen gene. Thus, the plasmids that enter the cells replicate efficiently, and V(D)J recombinants are detected almost exclusively in the replicated pool of molecules. We distinguished the pool of plasmids that have replicated in the murine cells by virtue of the fact that they have lost their prokaryotic *dam* methylation at the A in GATC sites. *Dpn*I is a restriction enzyme that eliminates plasmid molecules that have not replicated and, therefore, retain the *dam* methylation. We indicate treatment with *Dpn*I using a D. Hence, transformants arising on ampicillin plates from DNA that has been treated with *Dpn*I are designated DA. In this manner, V(D)J recombination (R) is measured as the fraction of recombinant molecules among the replicated molecules. The R values in this study were normalized further for the fraction of AC colonies that represent coding joints or signal joints. This normalization made it possible to determine R values of substrates with very low recombination efficiencies that fall below the false-positive background for the colony assay ($R = 0.01$ - 0.05% for coding joint-retaining substrates). It is noteworthy that the frequency of false-positive AC colonies is significantly lower (0.001%) in fibroblasts that lack RAG1 and RAG2 expression (Pergola et al. 1993). DNA was prepared from a number of AC colonies isolated from transformations of each murine cell transfection (typically 8-12 colonies per transfection, although more colonies were analyzed in most cases where the fraction of true recombinants was $<10\%$ of AC colonies). This DNA was restricted with *Pvu*II in the case of coding joint-retaining substrates and analyzed on a 0.8% agarose gel. True recombinants were scored when a 398-bp (± 15) band was observed, indicating deletion of the recombination signals and *oop* terminator, and joining of the coding ends. In the case of signal joint-retaining substrates, DNA was restricted with *Apa*LI. Recombinants were scored for the presence of a novel *Apa*LI fragment, indicative of a signal joint that contains the *Apa*LI recognition site.

Therefore, R values determined by AC/DA were multiplied by the fraction of colonies that were determined to be true recombinants. This fraction was quite variable, ranging from $<3.0\%$ to 100% . This corrected R value is abbreviated R_{corr} . R values in this study are presented as the ratio of R_{corr}/R determined for pJH200 in parallel experiments. The eukaryotic transfection method and plasmid harvest protocol have been described previously (Hesse et al. 1987). In this study we used electroporation (Bio-Rad Gene Pulser) and electrocompetent *E. coli* DH10B (Hsieh and Lieber 1991).

Sequencing

The coding joints of recombinant plasmids were sequenced using a 20-nucleotide primer complementary to the 5' end of *cat*, using Sequenase according to the manufacturer's (U.S. Biochemical) protocol.

Statistics

The data were analyzed by a permutation test (Miller 1986), using a computer program written by Michael Wolf (Department of Statistics, Stanford University).

Acknowledgments

We thank members of the laboratory for critical review of the manuscript, and Michael Wolf for assistance with statistical analyses. R.M.G. acknowledges support from the Stanford University Program in Cancer Biology and is currently a Fellow of The Jane Coffin Childs Memorial Fund for Medical Research. This investigation has been aided by a grant from The Jane Coffin Childs Memorial Fund for Medical Research. M.R.L. is a Lucille Markey Scholar, and this work was supported in part by a grant from the Lucille P. Markey Charitable Trust. This work was also supported by National Institutes of Health grants. M.R.L. is a Cancer Research Institute Investigator Awardee.

The publication costs of this article were defrayed in part by payment of page charges. This article must therefore be hereby marked "advertisement" in accordance with 18 USC section 1734 solely to indicate this fact.

References

- Akira, S., K. Okazaki, and H. Sakano. 1987. Two pairs of recombination signals are sufficient to cause immunoglobulin V-(D)-J joining. *Science* **238**: 1134–1138.
- Aquilera, R.J., S. Akira, K. Okazaki, and H. Sakano. 1987. A pre-B cell nuclear protein which specifically interacts with the immunoglobulin V-J recombination sequences. *Cell* **51**: 909–917.
- Asanow, D.M., D. Cado, and D.H. Raulet. 1993. Selection is not required to produce invariant T-cell receptor γ -gene junctional sequences. *Nature* **362**: 158–160.
- Bender, J. and N. Kleckner. 1992. Tn10 insertion specificity is strongly dependent upon sequences immediately adjacent to the target-site consensus sequence. *Proc. Natl. Acad. Sci.* **89**: 7996–8000.
- Cheung, S., K. Arndt, and L. Ponzy. 1984. Correlation of *lac* operator DNA imino proton exchange kinetics with its function. *Proc. Natl. Acad. Sci.* **81**: 3665–3669.
- Dickerson, R.E. 1992. DNA structure from A to Z. *Methods Enzymol.* **211**: 67–110.
- Gauss, G.H. and M.R. Lieber. 1992. The basis for the mechanistic bias for deletion over inversional V(D)J recombination. *Genes & Dev.* **6**: 1553–1561.
- Gellert, M. 1992. Molecular analysis of V(D)J recombination. *Annu. Rev. Genet.* **22**: 425–446.
- Gerstein, R.M. and M.R. Lieber. 1993. The extent to which homology can constrain coding exon junctional diversity in V(D)J recombination. *Nature* **363**: 625–627.
- 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.
- 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.
- Hsieh, C.-L. and M.R. Lieber. 1991. CpG methylated minichromosomes become inaccessible for V(D)J recombination after undergoing replication. *EMBO J.* **11**: 315–325.
- Kalled, S.L. and P.H. Brodeur. 1990. Preferential rearrangement of V κ 4 gene segments in pre-B cell lines. *J. Exp. Med.* **172**: 559–566.
- . 1991. Utilization of V κ families and V κ exons: Implications for the available B cell repertoire. *J. Immunol.* **147**: 3194–3200.
- Koudelka, G. and P. Carlson. 1992. DNA twisting and the effects of non-contacted bases on affinity of 434 operator for 434 repressor. *Nature* **355**: 89–91.
- Lafaille, J.J., A. DeCloux, M. Bonneville, Y. Takagaki, and S. Tonegawa. 1989. Junctional sequences of T cell receptor $\gamma\delta$ genes: Implications for $\gamma\delta$ T cell lineages and for a novel intermediate of V(D)J joining. *Cell* **59**: 859–870.
- Lawler, A.M., J.F. Kearney, M. Kuehl, and P.J. Gearhart. 1989. Early rearrangements of genes encoding murine immunoglobulin κ chains, unlike genes encoding heavy chains, use variable gene segments dispersed throughout the locus. *Proc. Natl. Acad. Sci.* **86**: 6744–6747.
- Lieber, M.R. 1991. Site-specific recombination in the immune system. *FASEB J.* **5**: 2934–2944.
- Lieber, M.R., J.E. Hesse, K. Mizuuchi, and M. Gellert. 1987. Developmental stage specificity of the lymphoid V(D)J recombination activity. *Genes & Dev.* **1**: 751–761.
- 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 segments in V(D)J recombination. *Cell* **55**: 7–16.
- Maass, G. 1987. Recognition of DNA sequences by restriction endonucleases. In *Proceedings of a NATO ASI/FEBS course on DNA-ligand interactions: From drugs to proteins* (ed. W. Guschlbauer and W. Saenger), pp. 235–236. Plenum Press, New York.
- Malynn, B., G.D. Yancopoulos, J.E. Barth, C.A. Bona, and F.W. Alt. 1990. Biased expression of J $_H$ -proximal V $_H$ genes occurs in the newly generated repertoire of neonatal and adult mice. *J. Exp. Med.* **171**: 843–859.
- Matsunami, N., Y. Hamaguchi, Y. Yamamoto, K. Kuze, K. Kanagawa, H. Mausuo, M. Kawaichi, and T. Honjo. 1989. A protein binding to the J κ recombination sequence of immunoglobulin genes contains a sequence related to the integrase motif. *Nature* **342**: 934–937.
- Miller, R.G. Jr. 1986. *Beyond ANOVA, basics of applied statistics*. John Wiley, New York.
- Patel, D.J., L. Shapiro, and D. Hare. 1988. NMR-distance geometry studies of helical errors and sequence dependent conformation of DNA in solution. In *Unusual DNA structures* (ed. R.D. Wells and S.C. Harvey), pp. 114–161. Springer-Verlag, New York.
- Pergola, F., M.Z. Zdzienicka, and M.R. Lieber. 1993. V(D)J recombination in mammalian cell mutants defective in DNA double-strand break repair. *Mol. Cell. Biol.* **13**: 3464–3471.
- Perlmutter, R.M., J.F. Kearney, S.P. Chang, and L.E. Hood. 1985. Developmentally controlled expression of immunoglobulin V $_H$ genes. *Science* **227**: 1597–1601.
- Sheehan, K.M. and M.R. Lieber. 1993. V(D)J recombination: Signal and coding joint resolution are uncoupled and depend on parallel synopsis of the sites. *Mol. Cell. Biol.* **13**: 1363–1370.
- Shirakata, M., K. Hüppi, S. Usuda, K. Okazaki, K. Yoshida, and H. Sakano. 1991. HMG1-related DNA-binding protein isolated with V-(D)-J recombination signal probes. *Mol. Cell. Biol.* **11**: 4528–4536.
- Wu, L., A. Vertino, and G.B. Koudelka. 1992. Non-contacted bases affect the affinity of synthetic P22 operators for P22 repressor. *J. Biol. Chem.* **267**: 9134–9139.
- Wu, Z. and G. Chaconas. 1992. Flanking host sequences can exert an inhibitory effect on the cleavage step of the *in vitro* Mu DNA strand transfer reaction. *J. Biol. Chem.* **267**: 9552–9558.
- Yancopoulos, G., S.V. Desiderio, M. Paskind, J.F. Kearney, D. Baltimore, and F.W. Alt. 1984. Preferential utilization of the most J $_H$ -proximal V $_H$ gene segments in pre-B-cell lines. *Nature* **311**: 727–733.

Coding end sequence can markedly affect the initiation of V(D)J recombination.

R M Gerstein and M R Lieber

Genes Dev. 1993, **7**:

Access the most recent version at doi:[10.1101/gad.7.7b.1459](https://doi.org/10.1101/gad.7.7b.1459)

References

This article cites 29 articles, 16 of which can be accessed free at:
<http://genesdev.cshlp.org/content/7/7b/1459.full.html#ref-list-1>

License

Email Alerting Service

Receive free email alerts when new articles cite this article - sign up in the box at the top right corner of the article or [click here](#).