1060–1067 Nucleic Acids Research, 1995, Vol. 23, No. 6

Coding sequence composition flanking either signal element alters V(D)J recombination efficiency

Nikolai V. Boubnov, Zachary P. Wills and David T. Weaver*

Division of Tumor Immunology, Dana-Farber Cancer Institute and Department of Microbiology and Molecular Genetics, Harvard Medical School, Boston, MA 02115, USA

Received August 29, 1994; Revised and Accepted January 4, 1995

ABSTRACT

Lymphoid V(D)J rearrangement is targeted by recombination signal sequences (RSS) bordering V, D or J exons. We demonstrate that the DNA composition of flanking coding positions, particularly poly(A) or poly(T) stretches at one or both RSS, diminishes V(D)J recombination up to 100-fold. Positionally correct cleavages occur in the inhibited reactions, since the junctions formed show the same frequency of precision as uninhibited reactions. Open/shut cleavage/rejoining is not increased at a normal RSS in substrates containing inhibitory A/T homopolymers versus random sequence at a second RSS. Thus recombinase action at both cleavage sites is severely disrupted by modified coding sequences.

INTRODUCTION

V(D)J gene rearrangement is a site-specific lymphoid cell recombination mechanism. Each variable (V), diversity (D) and joining (J) gene segment is bordered by a recombination signal sequence (RSS) consisting of a highly conserved heptamer separated by a spacer from a conserved nonamer. The RSS elements direct the mechanism of V(D)J recombination in pairs where the RSS are of different spacer lengths, with 12 or 23 bp spacers [RSS(12) and RSS(23)] (1). Mutations in one of the two RSS heptamers result in a 100-fold inhibition of V(D)J recombination (2). Although these elements are essential for the reaction, the requirements for RSS of different spacer length are not understood.

V(D)J gene rearrangement occurs by a multistep process involving the synapsis of two gene segments and cleavages at two RSS, followed by rearrangement and resolution of two new junctional products (for a review see 3). The outside border of the RSS heptamer with flanking coding sequence is the putative site of cleavage based on the structure of the signal and coding junctions normally formed (4,5). The actual cleavage structures are undescribed, although some type of double-strand break is thought to form in one or more cleavage steps. Evidence has recently been put forward for the formation of blunt signal ends and coding ends in hairpin configurations (6,7).

In other site-specific recombination mechanisms, roles for DNA sequences flanking the cleavage sites in adjusting the efficiency of the reaction have been noted (8–10). These effects have most frequently been attributed to alterations in the DNA structure that impact on the efficiency of cleavages and recombination. Variations in coding DNA sequences can alter the efficiency of V(D)J product formation (11,12), but resolution of whether compositional differences or structural changes were important was not determined. Here we have investigated a large number of coding end DNA sequences that strongly inhibit the V(D)J recombination reaction. A/T homopolymers flanking either RSS most dramatically reduce the recombination efficiency. RSS(12) and RSS(23) appear to be equally sensitive to the inhibitory effects of coding sequences. In addition, cleavage/rejoining at the unmodified RSS were also blocked, suggesting that the inhibitory influence extends to both RSS in the pair. An important determinant of the recombinase complex may involve structural properties of the coding DNA proximal to the RSS. Interaction with this coding DNA region, or a structure generated by it, either affects the stability of recombination initiation complexes or product formation.

MATERIALS AND METHODS

Plasmid substrates

V(D)J recombination plasmid substrates were derived from pJH290 (13) by deletion of the pJH290 RSS to facilitate oligonucleotide substitution as previously described (11). RSS orientation in these recombination templates is such that deletional V(D)J recombination occurs with retention of the RSS–RSS joint on recombined plasmids. The RSS(12) for these constructions was 5′-CACAGTGCGGCCGACTGGAACAAAAACC and the RSS(23) was 5′-CACAGTGCTCGAGCTCCACTGTCTGGCTGTACAAAAACC (heptamer and nonamer elements are underlined). Thus the recombination templates in this study differed only in the terminal 10 bp of their coding DNA sequences.

Modified V(D)J recombination templates were prepared with G_{10}, A_{10}, T_{10} or C_{10} polymers flanking the heptamer elements (Fig. 1). For example, $pT_{10}H/A_{10}H$-Sg contains one coding end of T_{10} flanking RSS(12) = [5′-*Sal*I-T_{10}-CACAGTG] and one coding end of A_{10} flanking RSS(23) = [5′-*Bam*HI-A_{10}-CA-

*To whom correspondence should be addressed at: Dana-Farber Cancer Institute, 44 Binney Street, Boston, MA 02115, USA

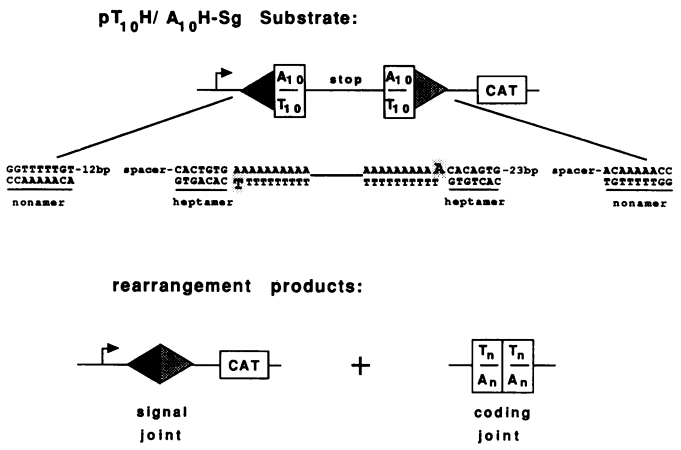

Figure 1. Plasmid substrates for V(D)J recombination. The example of the $pT_{10}H/A_{10}H$-Sg substrate is shown. The substrates also contain an AMP^r gene and the polyoma *ori* and T antigen for DNA replication in mouse cells (not shown). RSS elements in triangles differ by containing 12 bp (black triangle) or 23 bp (stippled triangle) spacers with the orientation and DNA sequence of heptamer and nonamer elements as specified. Boxes indicate the positions of oligonucleotide insertions to generate altered coding sequences flanking the positions of one or both RSS heptamers. Stippled nucleotides at the terminal coding sequence bordering each RSS are used to identify the plasmid substrates (see Tables 1 and 2). The two rearrangement products of V(D)J recombination assays (signal and coding junctions) are diagrammed; for these substrates signal junctions are retained on the plasmid.

CAGTG]. Additional coding end modifications were as follows: $(CG)_4C\underline{G}H$ = [5′-*Sal*I-GCGCGCGCG-CACAGTG]; $G_3CAG\ \underline{G}H$ = [5′-*Sal*I-GGGCAGG-CACAGTG]; $A_3TA_2\underline{A}H$ = [5′-*Bam* HI-AAATAAA-CACAGTG]; $A_3TGA\underline{A}H$ = [5′-*Bam*HI-AAAT-GAA-CACAGTG]; $(TA)_4T\underline{A}H$ = [5′-*Bam*HI-TATATATATA-CA CAGTG]. Also, $AGT\underline{T}H$ = [5′-*Sal*I-AGT-CACAGTG], $CG\underline{C}H$ = [5′-*Bam*HI-GATCCAGGCGC-CACAGTG], $GTCG\underline{A}H$ = [5′-TCGAGTCGA-CACAGTG-3′].

Cell culture and V(D)J recombination assay

HDR37A B lymphoma cells contain stably integrated RAG1 and RAG2 genes expressed under the control of the *Drosophila* HSP70 heat shock inducible promoter (14). 18-8 is an Abelson murine leukemia virus-transformed pre-B cell line that is recombinase-positive without induction. Cells were maintained and transfected as previously described (11).

V(D)J recombination of the signal junction (Sg) substrates was scored in a bacterial transformation assay developed by Gellert and colleagues (13). Introduction of rearranged plasmid DNA into bacteria allows transcription of the flanking *CAT* gene as described by Hesse *et al.* and Petrini *et al.* (13,15).

Precision of RSS junctions was measured by digestion of DNA of individual recombinant clones with the restriction enzyme *Apa*LI. A perfect fusion of the two RSS heptamers without base loss or addition generates a new *Apa*LI site in the plasmid. Imperfect products ($ApaL1^r$) were also tested for deletional V(D)J recombination by *Hin*dIII + *Bgl*II digestion, revealing a diagnostic 0.3 kb shift of the fragment containing a rearrangement region. The imperfect signal junctions were subsequently characterized by DNA sequencing.

Open/shut assay

Open/shut recombinase-dependent events are characterized by cleavage and rejoining without rearrangement (5). To monitor open/shut products, we prepared a plasmid substrate where the double-stranded oligonucleotide $GTCG\underline{A}H$ (5′-GTCGA-CACAGTG-3′) was introduced at RSS(12). The $GTCG\underline{A}H$ sequence contains five of the six residues of a *Sal*I restriction site. The sixth residue of the *Sal*I site is formed by the RSS heptamer. This oligonucleotide was introduced as described above into templates with either $CG\underline{C}H$, A_{10} or T_{10} coding sequences at RSS(23).

The strategy for quantitating open/shut events is as follows. $pGTCG\underline{A}H/CG\underline{C}H$-Sg, $pGTCG\underline{A}H/A_{10}H$-Sg and $pGTCG\underline{A}H/T_{10}H$-Sg templates were recovered from transiently transfected HDR37A cells by Hirt lysis and DNAs were digested with *Dpn*I to determine recombination frequencies as described above. Next, an aliquot of the DNA was digested with *Sal*I and ATP-dependent DNase (exonuclease V; US Biochemicals) to eliminate unrearranged DNA. Unrearranged substrates contain a unique *Sal*I restriction site. Open/shut products, as well as standard V(D)J rearrangements, are $SalI^r$ (Fig. 3). The two $SalI^r$ products are distinguished by bacterial cell plating as follows. One hundred AMP^r colonies from each sample were streaked on CAM + AMP selection plates (100 μg/ml AMP + 80 μg/ml CAM) to identify those that were specifically CAM^s (potential open/shut products). Plasmid miniprep DNAs from AMP^r + CAM^s colonies were tested for $SalI^r$.

RESULTS

Inhibition of V(D)J recombination by coding DNA sequences flanking RSS

We have prepared a series of V(D)J recombination substrates containing specific compositions of coding DNA sequences flanking the RSS. For these experiments we measured the efficiency of formation of RSS–RSS joining, a product of the reaction not directly involving coding DNA sequences (Fig. 1). With these substrates, the coding junction product is deleted from the plasmid and not scored. Recombination substrates are named from the nucleotide composition of coding DNA that is 5′ to the CACAGTG of the RSS heptamer (H) for both RSS(12) and RSS(23). These substrates only differ from each other in the terminal 10 nt of the coding DNA sequences. Thus the templates of this study are a means to evaluate the role of specific DNA sequences in the V(D)J reaction.

We compared the V(D)J recombination efficiencies of substrates containing homopolymers flanking each RSS with other substrates containing a mixed nucleotide composition at the coding DNA termini (Table 1). Recombination efficiencies of substrates in this study were normalized relative to the recombination efficiency of pJH200 in HDR37A cell transfections completed the same day. First we examined substrates with A/T or G/C 10mers flanking both RSS. Addition of A/T 10mers bordering each RSS strikingly reduced the V(D)J recombination reaction efficiency. $pA_{10}H/T_{10}H$-Sg and $pT_{10}H/A_{10}H$-Sg yielded recombination frequencies that were 2% of the level for pJH200 (Table 1). These two substrates had the A/T 10mers in two different orientations relative to the RSS. Therefore, A or T 10mers in either orientation relative to RSS(12) or RSS(23) both significantly diminished recombination potential. In contrast,

G/C 10mers flanking both RSS only decreased V(D)J recombination ~2-fold. $pG_{10}H/C_{10}H$-Sg and $pC_{10}H/G_{10}H$-Sg each yielded average corrected recombination values of 52% of pJH200 values. Therefore, G/C 10mers had considerably less influence on recombination efficiency in either orientation flanking the RSS. Because these stretches of 10 bp are the only differences amongst the substrates, we can conclude that coding DNA composition had a significant impact on the reaction. The impairment of V(D)J rearrangement observed with A/T modified templates is as striking as that for RSS heptamer point mutations in related plasmid substrates and transient transfection assays (13). These effects are more inhibitory than mutations in the well-conserved nonamer sequence or subtle variations in RSS spacer lengths.

Table 1. V(D)J recombination of substrates containing homopolymers in the HDR37A and 18-8 cell lines

Substrate		Expt	No. Amp^r	No. Amp^r + Cam^r	R (%)[a]	R_{corr} (%)[b]	R_{av} (% ± SD)[c]	Correct joins (%)[d]
HDR37A:	pGACH/TCCH-Sg	1	790 000	27 600	3.5	100	100	98 (49/50)
	(pJH200)	2	1 900 000	89 700	4.7	100		
		3	1 080 000	44 000	4.1	100		
		4	840 000	26 200	3.1	100		
		5	1 440 000	56 100	3.9	100		
		6	2 070 000	140 000	6.8	100		
		7	2 440 000	81 000	3.3	100		
	$pG_{10}H/C_{10}H$-Sg	4	800 000	16 700	2.1	68	52	86 (32/37)
		6	2 230 000	54 100	2.4	35	(±17)	
		7	1 540 000	26 000	1.7	52		
	$pC_{10}H/G_{10}H$-Sg	4	1 060 000	8300	0.78	25	52	96 (24/25)
		6	1 110 000	45 100	4.1	60	(±24)	
		7	1 090 000	25 000	2.3	70		
	$pG_{10}H/A_{10}H$-Sg	1	870 000	4200	0.48	14	17	NA
		2	1 310 000	14 600	1.1	23	(±4.9)	
		3	1 400 000	8800	0.63	15		
	$pA_{10}H/G_{10}H$-Sg	1	690 000	8100	1.2	33	37	NA
		2	1 360 000	32 300	2.4	51	(±12)	
		3	2 040 000	22 200	1.1	27		
	$pC_{10}H/T_{10}H$-Sg	1	1 240 000	1670	0.13	3.7	6.0	NA
		2	910 000	2300	0.25	5.3	(±2.7)	
		3	1 030 000	3800	0.37	9.0		
	pT_{10}/C_{10}-Sg	1	2 130 000	2650	0.12	3.4	7.6	NA
		2	1 300 000	8200	0.63	13	(±4.9)	
		3	1 580 000	4100	0.26	6.3		
	$pA_{10}/T_{10}H$-Sg	4	1 530 000	510	0.033	1.1	2.0	92 (22/24)
		5	1 320 000	1010	0.077	2.0	(±0.9)	
		6	450 000	890	0.20	2.9		
	$pT_{10}H/A_{10}H$-Sg	4	1 420 000	320	0.023	0.74	1.9	80 (28/35)
		5	1 100 000	850	0.077	2.0	(±1.2)	
		6	1 490 000	3100	0.21	3.1		
18-8 (TdT$^+$):	pGACH/TCCH-Sg	1	18 000	395	2.2	100	100	98 (49/50)
	(pJH200)	2	1 360 000	24 000	1.8	100		
	$pG_{10}H/C_{10}H$-Sg	1	15 000	258	1.7	77	56	94 (50/53)
		2	664 000	3380	0.5	28		
	$pC_{10}H/G_{10}H$-Sg	1	14 400	270	1.9	86	88	96 (53/55)
		2	554 000	8780	1.6	89		
	$pA_{10}H/T_{10}H$-Sg	1	576 000	65	0.011	0.5	0.8	100 (15/15)
		2	800 000	160	0.020	1.1		
	$pT_{10}H/A_{10}H$-Sg	1	217 000	35	0.016	0.7	0.6	72 (21/29)
		2	1 200 000	125	0.010	0.6		

[a]Calculated as percentage Amp^r Cam^r colonies from Amp^r colonies.
[b]Recombination frequency corrected for the recombination frequency of pJH200 in the same experiment.
[c]Average corrected recombination frequency.
[d]Calculated as percentage *Apa*LI^s signal joins from total signal joins.
H, heptamer (5′-CACAGTG); a nucleotide adjacent to the 5′ side of the heptamer is underlined.

The above results were documented in the HDR37A cell line, where V(D)J recombination is induced by heat shock (14). We also examined V(D)J recombination in a murine pre-B cell line (18-8) isolated from the recombinationally active stage of mouse B cell lymphopoiesis. 18-8 cells have constitutively active RAG1 and RAG2 gene expression and recombinase activity (11) and are therefore an important control for whether the coding DNA sequence effects are also present when Rag1 and Rag2 protein levels are physiologically relevant and in the appropriate cell stage for V(D)J recombination. We also examined the same templates in 18-8 cells where A/T or G/C 10mers were placed adjacent to both RSS. We found that homopolymers of A/T flanking both RSS strikingly inhibited the reaction. $pA_{10}H/T_{10}H$-Sg and $pT_{10}H/A_{10}H$-Sg yielded recombination frequencies that were 0.8 and 0.6% of the level for pJH200 (Table 1). Likewise, $pG_{10}H/C_{10}H$-Sg and $pC_{10}H/G_{10}H$-Sg each yielded only slightly reduced R_{av} of 56 and 88% respectively. As observed for HDR37A, either orientation of A/T homopolymers was inhibitory. Therefore, coding sequence composition changes profoundly affect the V(D)J recombination reaction efficiency; because these results were observed in cell lines with constitutively expressed or induced RAG1/RAG2, we conclude that the properties we are measuring are intrinsic to the basal V(D)J recombination machinery.

Coding sequence A/T 10mers flanking one RSS inhibit V(D)J recombination

Considering that A/T residues flanking RSS were significant in dictating the reaction efficiency, we next combined A/T homopolymers at one RSS with G/C homopolymers at the other RSS in a group of substrates. $pC_{10}H/T_{10}H$-Sg and $pT_{10}H/C_{10}H$-Sg reduced recombination to an R_{av} of 6.0 and 7.6% respectively (Table 1). Similarly, two other substrates, $pG_{10}H/A_{10}H$-Sg and $pA_{10}H/G_{10}H$-Sg, were found to give an R_{av} of 17 and 37% respectively relative to pJH200. Therefore, the influence of an A/T homopolymer flanking one RSS dramatically impacted V(D)J recombination efficiency. These results suggested that the $T_{10}H$ orientation relative to the heptamers may have the most severe effect.

We next tested substrates containing homopolymers of 10 residues placed at RSS(23) that had a mixed composition coding sequence (AGTH; Materials and Methods) at RSS(12). These substrates were used to directly compare the effects of A, T, G or C 10mers at the same position. We found that two of the substrates dramatically inhibited V(D)J recombination, $pAGTH/A_{10}H$-Sg (R_{av} = 11%) and $pAGTH/T_{10}H$-Sg (R_{av} = 3.6%) were strikingly reduced compared with the value for pJH200 (Table 2). In contrast, substrates with G/C coding termini had little effect on the reaction efficiency; $pAGTH/C_{10}H$-Sg (R_{av} = 43%) and $pAGTH/G_{10}H$-Sg (R_{av} = 101%). These effects are attributed to the 10mers at RSS(23), rather than the mixed composition coding sequence at RSS(12), because reaction efficiency is high in other AGTH-containing templates (Table 3 and data not shown). We also found that A/T 10mers flanking RSS(12) significantly inhibited the reaction, but not G/C 10mers (data not shown).

Table 2. V(D)J recombination of substrates containing a homopolymer at one of the RSS[a]

Substrate		Expt	No. Amp^r	No. Amp^r +Cam^r	R (%)	R_{corr} (%)	R_{av} (% ± SD)
HDR37A:	pGACH/TCCH-Sg	1	670 000	32 400	4.8	100	100
	(pJH200)	2	740 000	10 800	1.5	100	
		3	2 440 000	81 000	3.3	100	
		4	1 130 000	12 500	1.1	100	
		5	730 000	8100	1.1	100	
		6	580 000	19 400	3.3	100	
	pAGTH/G_{10}H-Sg	1	1 900 000	51 600	2.7	56	101
		2	690 000	9800	1.4	93	(±50)
		3	960 000	49 000	5.1	155	
	pAGTH/C_{10}H-Sg	1	780 000	15 100	1.9	40	43
		2	500 000	1920	0.38	25	(±20)
		3	550 000	11 600	2.1	64	
	pAGTH/A_{10}H-Sg	1	1 070 000	1710	0.16	3.3	11
		2	1 050 000	1660	0.16	11	(±7.9)
		3	730 000	4600	0.63	19	
	pAGTH/T_{10}H-Sg	1	1 010 000	310	0.031	0.65	3.6
		2	610 000	880	0.140	9.3	(±5.0)
		3	960 000	220	0.023	0.70	
18-8 (TdT$^+$):	pGACH/TCCH-Sg	1	18 000	395	2.2	100	100
	(pJH200)	2	1 360 000	24 000	1.8	100	
	pAGTH/G_{10}H-Sg	1	18 900	155	0.8	36	59
		2	1 000 000	13 100	1.3	72	
	pAGTH/C_{10}H-Sg	1	9100	280	3.1	141	112
		2	604 000	8760	1.5	83	
	pAGTH/A_{10}H-Sg	1	152 000	131	0.086	3.9	3.6
		2	219 000	135	0.062	3.4	
	pAGTH/T_{10}H-Sg	1	78 000	22	0.028	1.3	1.5
		2	740 000	225	0.030	1.7	

[a]For details, see footnotes to Table 1.

Table 3. V(D)J recombination of substrates with varying sequences flanking the RSS in the HDR37A cell line[a]

Substrate	Expt	No. Amp^r	No. Amp^r + Cam^r	R (%)	R_{corr} (%)	R_{av} (% ± SD)
pGACH/TCCH-Sg	1	840 000	26 200	3.1	100	100
(pJH200)	2	580 000	19 400	3.3	100	
	3	1 440 000	56 100	3.9	100	
	4	1 900 000	89 700	4.7	100	
	5	1 080 000	44 000	4.1	100	
	6	730 000	8100	1.1	100	
pAGTH/CGCH-Sg	1	1 080 000	12 600	1.2	39	70
	2	430 000	8100	1.9	58	(±38)
	3	580 000	25 700	4.4	113	
pAGTH/A_3TAAAH-Sg	4	1 190 000	32 500	2.7	57	56
	5	1 030 000	29 000	2.8	68	(±13)
	6	810 000	3800	0.47	43	
p$(CG)_4$CGH/$(TA)_4$TAH-Sg	3	1 750 000	68 500	3.9	100	81
	5	2 110 000	47 700	2.3	56	(±23)
	6	850 000	8200	0.96	87	
pG_3CAGGH/A_3TGAAH-Sg	4	1 420 000	102 000	7.2	153	99
	5	1 580 000	46 000	2.9	71	(±47)
	6	1 640 000	13 100	0.80	73	

[a]For details, see footnotes to Table 1.

In 18-8 cells, the recombination efficiencies of pAGTH/A_{10}H-Sg and pAGTH/T_{10}H-Sg were strikingly reduced relative to the pJH200 control (R_{av} of 4.5 and 2.5% respectively; Table 2). These reductions in recombination potential were as low as was found with HDR37A cell transfections. G/C 10mers at RSS(23) had considerably less impact on the reaction in 18-8 cells as well. pAGTH/G_{10}H-Sg (R_{av} = 55%) and pAGTH/C_{10}H-Sg (R_{av} = 115%) yielded recombination values close to the normal level (Table 2). Therefore, the inhibitory effect of a single homopolymer of A/T residues was not cell type-specific.

Inhibitory effects of A/T 10mers flanking one or both RSS were not distinguishably different for the T 10mer orientation. Comparing three substrates with T_{10} at RSS(23), pAGTH/T_{10}H-Sg, pC_{10}H/T_{10}H-Sg and pA_{10}H/T_{10}H-Sg, we found R_{av} values of 3.6, 6.0 and 2.0% respectively (Tables 1 and 2). These values are not statistically different ($p > 0.6$). Similar results were shown for these substrates in 18-8 cells (Tables 1 and 2). With A 10mers we found some increase in inhibition with compositional changes at RSS(12), but not in both cell lines. Three substrates with A_{10} at RSS(23) were compared: pAGTH/A_{10}H-Sg, pG_{10}H/A_{10}H-Sg and pT_{10}H/A_{10}H-Sg had R_{av} of 11, 17 and 1.9% respectively. In 18-8 cells, pAGTH/A_{10}H-Sg and pT_{10}H/A_{10}H-Sg yielded R_{av} of 3.6 and 0.6% respectively, values that are not statistically different. Similarly, addition of G/C 10mers at either one or both RSS did not significantly influence the recombination potential of these substrates (Table 1 and 2). Overall, it appears that placement of an A/T 10mer flanking one RSS severely reduces recombination efficiency and that addition of an A/T 10mer at the other RSS has little increased inhibitory influence.

Mixed A/T coding DNA does not inhibit V(D)J recombination

Variations in nucleotide composition were examined to elaborate the requirement for A/T residues flanking the RSS. Placement of A_3TA_2AH flanking RSS(23) produced only a slightly reduced recombination frequency relative to pJH200 (pAGTH/A_3TA_2AH-Sg, R_{av} = 56%; Table 3). Likewise, the dinucleotide polymer (alternating T and A residues) yielded a normal recombination efficiency [p$(CG)_4$CGH/$(TA)_4$TAH-Sg, R_{av} = 81%]. The interspersion of a TG dinucleotide in the midst of an A 5mer (AAATGAA-CACAGTG) gave a normal recombination frequency (pG_3CAGGH/A_3TGAAH-Sg, R_{av} = 99%). Thus, AT-rich composition alone is insufficient to inhibit the reaction. Instead, A/T homopolymers may be promoting a structural change in the helix that influences V(D)J recombination (see Discussion).

We also compared potential effects of the terminal nucleotide closest to RSS. AGTH, CGCH and CGAH place a T, C or A residue respectively immediately flanking the heptamer. AGTH-, CGCH- and CGAH-containing templates all produced recombination in the normal range when matched with other mixed composition coding ends (Tables 3 and 4). Likewise, pG_3CAGG/A_3TGAA-Sg, p$(CG)_4$CGH/$(TA)_4$TAH-Sg and pG_5/CGCH-Sg each place a G residue flanking the heptamer and had a high recombination value (R_{av} = 99 and 81% respectively and data not shown). pJH200 also has a coding DNA C next to both heptamers and yields high recombination values (Table 1). Thus there is no critical nucleotide immediately adjacent to the RSS heptamer, in disagreement with the postulate of Gerstein and Lieber (12). The nucleotide composition of the terminal 10 residues better dictates the magnitude of inhibitory influences to be expected for these substrates.

V(D)J recombination products from the reduced efficiency reactions retain the correct cleavages

RSS junctions are ordinarily the precise fusions of heptamers from the two RSS with different spacer lengths, suggesting that cleavage occurs on at least one strand at the 5′-heptamer border (5′-↓CACAGTG-3′). We tested whether the inhibitory effect of adjacent A/T coding DNA could generate an altered RSS recognition, such that the residual cleavage would occur at a changed position relative to the heptamer. Despite a significantly reduced recombination efficiency of A/T 10mer substrates, the RSS junctions were structurally correct and formed with nearly

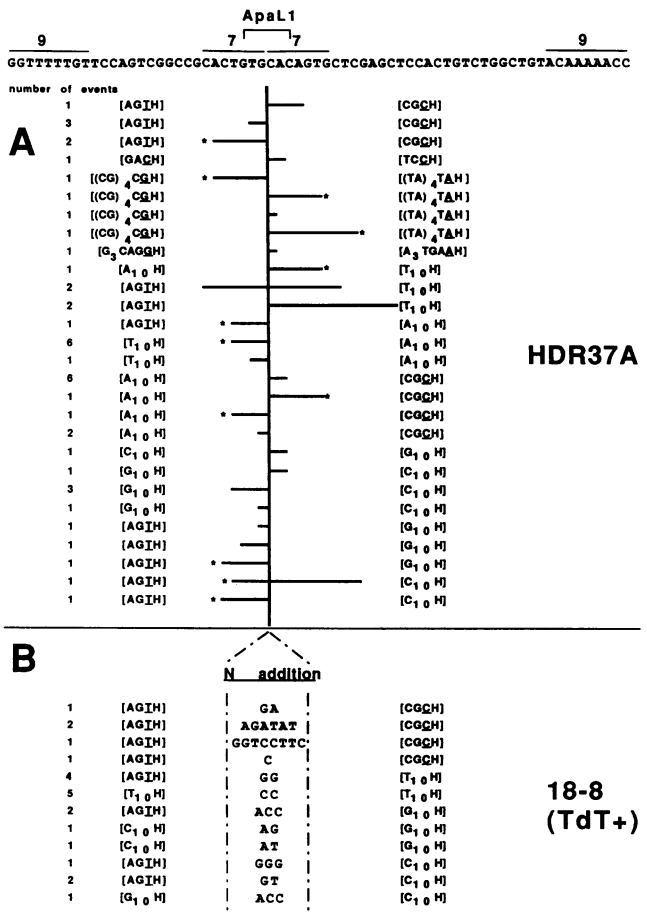

Figure 2. Imperfect RSS–RSS junctions that arose as a consequence of V(D)J recombination. The top line shows the DNA sequence of a perfect RSS–RSS junction. Bracketed residues are coding ends of the substrates as described in the text. (**A**) The extent of junctional deletion for substrates in HDR37A cells is illustrated with a horizontal line. *deletions where the residue assignment of the junction is ambiguous, but can be aligned to be consistent with deletion from one of the two heptamers only. (**B**) The extent of nucleotide addition in RSS junctions of 18-8 cells.

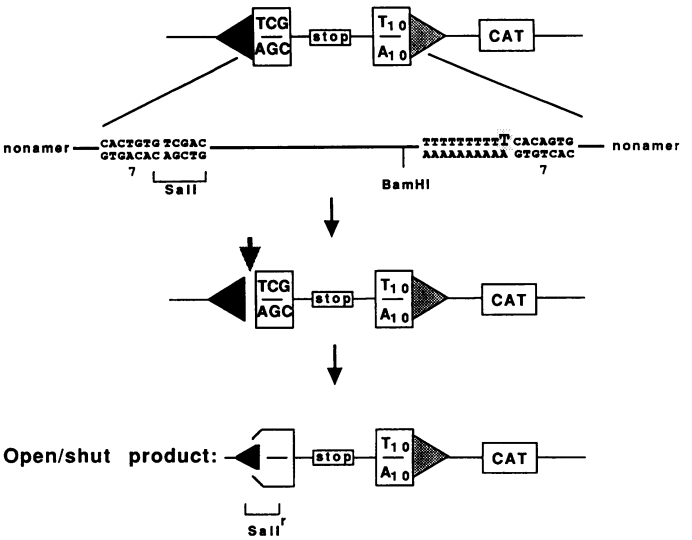

Figure 3. Open/shut recombinase events of pGTCGAH/T_{10}H-Sg. Cleavage at the mixed composition coding DNA–heptamer border is designated by a stippled arrow. Inaccurate rejoining of this break leads to inactivation of a *SalI* restriction enzyme site.

equal proficiency as wild-type controls (Table 1). A similar outcome was observed for both the HDR37A and 18-8 cell lines with each of the modified substrates inhibiting V(D)J recombination. Thus the specificity of the V(D)J recombination cleavages is normal, even for inhibited reactions.

Although most of the RSS junctions formed were precise, the lower frequency aberrant products that were *Apa*L1^r were also analyzed by DNA sequencing. We found RSS junction deletions of 1–14 nt from either RSS end (Fig. 2A). The sizes of these deletions are reminiscent of normal coding end deletions. The same extent of deletion was observed for substrates with mixed nucleotide coding ends, coding ends with G/C homopolymers or coding ends with A/T homopolymers, as shown (Fig. 2). A nucleotide redundancy in the RSS junctions restricted the possibility of unambiguously identifying the signal end contributing the junction nucleotides in all cases. However, all but three of the events were consistent with all of the deletion occurring from one RSS only. No particular preference of signal junction deletion was associated with RSS(12) or RSS(23) flanked by A/T homopolymers.

We also investigated the aberrant RSS junctions (*Apa*LI^r) for these substrates in 18-8 cells. An additional feature of 18-8 pre-B cells is the presence of the enzyme terminal deoxynucleotidyl transferase (TdT), implicated in the addition of N residues in the V(D)J recombination pathway of lymphoid progenitors (11,16,17). These junctions contained nucleotide insertions precisely at the heptamer–heptamer junction (Fig. 2B). No deletion of nucleotides from either RSS was observed. The nucleotides added were of variable lengths (1–8 residues) and differing composition. Also, no unique added nucleotides were observed for the A/T, G/C and mixed nucleotide coding end substrates that would have indicated an alternative cleavage pattern. We conclude that the V(D)J recombination cleavage specificity is not altered by adjacent A/T homopolymer coding sequences, even though these sequences dramatically diminish the reaction efficiency.

Open/shut products are not increased in the presence of inhibitory A/T coding DNA

Open/shut recombinase-mediated events are detected by the cleavage and imperfect rejoining at one RSS without rearrangement (4,5). These events require a second RSS of opposite spacer length. If the second RSS is composed of A/T coding sequences that inhibit recombination, then open/shut events may or may not be inhibited at the normal RSS. An increased open/shut frequency would be indicative of independent cleavage at the normal RSS in the reaction. Alternatively, a low level of open/shut events in these substrates would be supportive of a mechanism where cleavage at one RSS was dependent on the coding structure of the other RSS.

We compared three substrates sharing the same coding sequence (GTCGAH) at RSS(12), but differing in the coding sequence flanking RSS(23) (pGTCGAH/CGCH-Sg, pGTCGAH/A_{10}H-Sg and pGTCGAH/T_{10}H-Sg) (Fig. 3; Materials and Methods). In each case open/shut products at the GTCGA–RSS region were quantitated. The GTCGAH sequence contains five of six nucleotides of a *SalI* restriction enzyme site where the final residue of the site, C, is also the first residue in the RSS(12) heptamer (Fig. 3). Therefore cleavage at the heptamer border in the GTCGAH–RSS with imperfect rejoining will lead to inactivation of the *SalI* site.

Table 4. Frequencies of V(D)J recombination signal junction formation and open/shut events

Substrate	Expt	No. Amp^r	No. Amp^r + Cam^r	R (%)[a]	R_{av} (% ± SD)	O/S (%)[b]	O/S$_{av}$ (% ± SD)
pGTCGAH/CGCH-Sg	1	162 000	6300	3.9	3.7	0.022	0.063
	2	95 000	2200	2.3	(±0.9)	0.075	(±0.034)
	3	380 000	15 900	4.2		0.102	
	4	372 000	16 100	4.3		0.053	
pGTCGAH/A$_{10}$H-Sg	1	239 000	462	0.19	0.24	0.024	0.026
	2	360 000	510	0.14	(±0.10)	0.007	(±0.014)
	3	1 105 000	3950	0.36		0.030	
	4	1 295 000	3650	0.28		0.041	
pGTCGAH/T$_{10}$H-Sg	1	207 000	363	0.17	0.15	0.018	0.011
	2	360 000	475	0.13	(±0.02)	0.015	(±0.007)
	3	1 950 000	2950	0.15		0.006	
	4	1 290 000	2050	0.16		0.003	

[a]Calculated as percentage Amp^r Cam^r colonies from Amp^r colonies.
[b]Calculated as percentage Amp^r Cam^s *Sal*I^r colonies from Amp^r colonies.

Measuring the overall V(D)J recombination reaction efficiency, pGTCGAH/A$_{10}$H-Sg and pGTCGAH/T$_{10}$H-Sg (R = 0.17 and 0.15% respectively) were severely inhibited relative to pGTCGAH/CGCH-Sg (R = 3.1%), as expected. These values are ~1.6% of the level of recombination measured in the control, pGTCGAH/CGCH-Sg, for the same experiments. When open/shut events were measured from the same experiments for these three substrates, we found that the level of open/shut events was not increased in rearrangement-inhibited reactions (Table 4). In fact, the two substrates containing inhibitory sequences *in trans*, pGTCGAH/A$_{10}$H-Sg and pGTCGAH/T$_{10}$H-Sg, had open/shut values that were reduced relative to the control levels (0.026 and 0.011% compared with 0.063%). Low numbers of open/shut events for all of these experiments precluded an accurate estimation of the level of inhibition. In any case, the open/shut frequency is not increased by the presence of inhibitory coding sequences *in trans*. DNA sequencing indicates that each of these events corresponds to deletions at RSS(12) (data not shown), in accordance with open/shut recombinase events scored previously (5). Therefore, the impairment of V(D)J recombination is consistent with inhibition of the cleavage reaction at both RSS sites, even though only one is altered by coding sequence changes.

DISCUSSION

Here we have demonstrated that terminal coding DNA sequences have a significant impact on V(D)J recombination efficiency, but not specificity. These effects are consistent with interference with either initiation or joining events in V(D)J rearrangement. Whether the stage of the reaction that is inhibited is pre- or post-cleavages, our data is supportive of the hypothesis that inhibitory coding DNA sequences affect recombinational complexes. V(D)J recombination in the *xrs*, *XR-1* and *sxi-3* mutant cell lines shows that both coding and RSS products are severely diminished (18–20). These mutants are also ionizing radiation-sensitive and are defective in double-strand break repair (reviewed in 3). It is interesting that alterations in coding DNA composition can have the same influence on V(D)J rearrangement products and quantitatively reduce RSS joining as profoundly as these cell mutants (Tables 1 and 2). The paradoxical effects of coding DNA sequences on RSS products may occur by dissolution of a protein–DNA recombination synapse that has to include gene products for joining. One of these gene products could interact with coding DNA regions and be sensitive to the DNA composition.

On the other hand, several arguments can be raised to support an effect of inhibitory coding sequences prior to product formation in the reaction. These earlier stages might include the recognition of RSS, recombinational synapsis or cleavage; inhibition of any of these steps would also result in low yields of both coding and signal joining products.

Our observations are that the level of cleavage/rejoining at the unmodified RSS is not increased relative to that of a control template only differing in the terminal 10 nt of the other coding end (Table 4). Clearly, the presence of A/T 10mers at one RSS interferes with a high rate of cleavage/rejoining at the other RSS. The open/shut cleavage/rejoining assay we used is not yet able to discriminate between all of the possible stages where V(D)J recombination may be impacted on by coding DNA composition. Although open/shut product level decreases without the presence of a second RSS (5), there is as yet no definitive data to show that these events are measuring normal V(D)J recombination cleavage intermediates. Thus, even though we cannot yet determine whether initiation or joining steps are inhibited by A/T coding DNA, it is likely that the inhibition is manifested in less stable recombinase complexes.

The nucleotide requirements for coding sequence inhibition of V(D)J recombination are still unresolved. Our data show that A/T homopolymers block V(D)J recombination most dramatically (Tables 1 and 2). Three templates with an A/T-rich coding end yielded proficient recombination (Table 3). The coding end sequence T$_3$A$_2$T$_2$T near RSS(12) inhibited RSS junction formation and it was concluded that the T residue flanking the heptamer blocked the reaction (12). Since A/T 10mers of either orientation flanking RSS(12) or RSS(23) are strongly inhibitory (Tables 1 and 2 and data not shown), we argue that the nucleotide immediately adjacent to the heptamer is not of primary importance. Also, A/T 10mers previously showed a 4–6-fold greater inhibitory effect than A/T 5mers (11) where the same residue is always proximal to the RSS. These results suggest that nucleotides adjacent to the RSS are less significant than the overall composition of the coding end positions.

Coding DNA structure, rather than composition, may determine V(D)J recombination efficiency. Biophysical measurements of DNA duplexes containing homopolymers of A/T (>4mer) show local deformations in the helix yielding a narrowing of the minor groove (reviewed in 21). Proton NMR measurements of A/T base pairs indicate a helical alteration when A/T base pairs occur in $A_{n>4}$ tracts as opposed to control sequences with $A_{n<4}$ (22,23). These abnormal NMR lifetimes are increased with higher n values (22). Consistent with biophysical measurements, decreased sensitivity to hydroxyl radicals and endonuclease cleavage occurs in A tracts (24,25).

Our data are strikingly consistent with this structural information. We did not find inhibitory effects of coding ends containing A_1, A_2 or A_3 flanking the heptamer, while A_5 or A_{10} had significant inhibitory effects (Tables 1 and 2). A/T 10mers also had a 4–6-fold stronger inhibitory effect than A/T 5mers with coding junction substrates (11). Interestingly, A tracts that are 5′-A_2T_2-3′, but not 5′-T_2A_2-3′, give altered helical structures (22). Inhibitory coding sequences actually correlate well with this additional observation, since the coding sequence A_2T_2TH (our nomenclature) inhibits V(D)J recombination even though the size of the polymer was three (T_3) (12). Where A/T tracts were interrupted by other base pairs, as with the sequences A_3TA_2AH and A_3TGAAH, there was no appreciable inhibition of V(D)J recombination (Table 3). Thus, 5′-A_2T_3-3′, but not 5′-TA_3-3′, also inhibits V(D)J recombination.

Local distortions of DNA may ordinarily be innocuous. However, in conjunction with specific placing and synapsis in a site-specific recombination event, these distortions may influence the ability of recombination protein(s) to hold this region of DNA in a synaptic complex. Distortions created by coding sequences could affect interactions between proximal coding DNA sites and parts of the recombinase complex that are not necessarily mediated by RSS binding. A simple model for explaining the inhibitory influence of coding DNA sequences is that different portions of the recombinase complex may associate with different DNA components in the reaction. Some portion of the recombinase may interact directly with RSS, directing both synapsis and cleavage. In addition, another part of the recombinase may associate with the proximal coding DNA sequences to stabilize the synaptic complex and to hold the coding ends during processing and joining. This association would be relatively sequence non-specific, in keeping with our observations that a number of different coding sequences are inhibitory and that many sequences do not stimulate or diminish reaction efficiency. This other factor could be an accessory protein(s) or another domain of the same recombinase protein interacting with RSS.

The profound coding sequence inhibition observed here restricts the range of possible V, D or J coding exon sequences utilizable by the immune system. In fact, analysis of GenBank (release 82) sequences for V, D and J gene segments shows that all analyzed sequences present acceptable coding DNA ends for efficient V(D)J recombination. Possibly, inhibitory sequences have been lost due to evolutionary selection because of a poor ability to undergo V(D)J rearrangement.

ACKNOWLEDGEMENTS

We thank members of the Weaver laboratory for many helpful discussions and critical reading of the manuscript. We thank Tim Ernst for oligonucleotide synthesis. This research was supported in part by grants NIH CA54326, NIH CA52694 and NIH GM39312 and a Sandoz Drug Discovery Program grant to D.T.W.

REFERENCES

1 Tonegawa,S. (1983) *Nature*, **302,** 575–581.
2 Hesse,J.E., Lieber,M.R., Mizuuchi,K. and Gellert,M. (1989) *Genes Dev.*, **3,** 1053–1061.
3 Weaver,D. (1994) *Adv. Immunol.*, **58,** 29–85
4 Hendrickson,E.A., Liu,V.F. and Weaver,D.T. (1991) *Mol. Cell. Biol.*, **11,** 3155–3162.
5 Lewis,S.M. and Hesse,J.E. (1991) *EMBO J.*, **10,** 3631–3639.
6 Schlissel,M., Constantinescu,A., Morrow,T., Baxter,M. and Peng,A. (1993) *Genes Dev.*, **7,** 2520–2532.
7 Roth,D.B., Zhu,C. and Gellert,M. (1993) *Proc. Natl. Acad. Sci. USA*, **90,** 10788–10792.
8 Huisman,O., Faelen,M., Girard,D., Jaffe,A., Toussaint,A. and Rouviere-Yaniv,J. (1989) *J. Bacteriol.*, **171,** 3704–3712.
9 Bender,J. and Kleckner,N. (1992) *Proc. Natl. Acad. Sci. USA*, **89,** 7996–8000.
10 Wu,Z. and Chaconas,G. (1992) *J. Biol. Chem.*, **267,** 9552–9558.
11 Boubnov,N.V., Wills,Z.P. and Weaver,D.T. (1993) *Mol. Cell. Biol.* **13,** 6957–6968.
12 Gerstein,R.M. and Lieber,M.R. (1993) *Genes Dev.*, **7,** 1459–1469.
13 Hesse,J.E., Lieber,M.R., Gellert,M. and Mizuuchi,K. (1987) *Cell.*, **49,** 775–783.
14 Oltz,E.M., Alt,F.W., Lin,W.-C., Chen,J., Taccioli,G., Desiderio,S. and Rathbun,G. (1993) *Mol. Cell. Biol.*, **13,** 6223–6230.
15 Petrini,J.H.J., Donovan,J.W., DiMare,C. and Weaver,D.T. (1994) *J. Immunol.*, **152,** 176–183.
16 Komori,T., Okada,A., Stewart,V. and Alt,F.W. (1993) *Science*, **261,** 1171–1175.
17 Gilfillan,S., Dierich,A., Lemeur,M., Benoist,C. and Mathis,D. (1993) *Science*, **261,** 1175–1178.
18 Taccioli,G.E., Cheng,H.-L., Varghese,A. J., Whitmore,G. and Alt,F.W. (1994) *J. Biol. Chem.*, **269,** 7439–7442.
19 Pergola,F., Zdzienicka,M.Z. and Lieber,M.R. (1993) *Mol. Cell. Biol.*, **13,** 3464–3471.
20 Boubnov,N.V., Hall,K.T., Wills,Z., Lee,S.E., He,D.M., Benjamin,D.M., Pulaski,C.R., Band,H., Reeves,W., Hendrickson,E.A. and Weaver,D. (1995) *Proc. Natl. Acad. Sci. USA*, **92,** in press.
21 Hagerman,P.J. (1990) *Annu. Rev. Biochem.*, **59,** 755–781.
22 Leroy,J.-L., Charretier,E., Kochoyan,M. and Gueron,M. (1988) *Biochemistry*, **27,** 8894–8898.
23 Nadeau,J.G. and Crothers,D.M. (1989) *Proc. Natl. Acad. Sci. USA*, **86,** 2622–2626.
24 Burkhoff,A.M. and Tullius,T.D. (1987) *Cell*, **48,** 935–943.
25 Drew,H.R. and Travers,A.A. (1984) *Cell*, **37,** 491–502.