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 synopsis 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 recombinated plasmids. The RSS(12) for these constructions was 5'-CACAGTGCGGCCGACTGGAACAAACC and the RSS(23) was 5'-CACAGTGCTCGAGCTCACGTGTCTGGCTGTACAAAAACC (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 G10, A10, T10 or C10 polymers flanking the heptamer elements (Fig. 1). For example, pT10H/A10H-Sg contains one coding end of T10 flanking RSS(12) = [5'-SalI-T10-CACAGTG] and one coding end of A10 flanking RSS(23) = [5'-BamHI-A10-CAGTG].
diagnostic 0.3 kb shift of the fragment containing a rearrangement.

HindIII + BgIII digestion, revealing a V(D)J recombination by

of individual recombinant clones with the restriction enzyme

acterized by DNA sequencing.

Imperfect products (ApaI) were also tested for deletional

loss or addition generates a new

ApaI site in the plasmid.

et al. and Petrini et al. (13,15). described by Hesse

cAT gene as

scored in a bacterial transformation assay developed by Gellert

and transfected as previously described (11).

recombinase-positive without induction. Cells were maintained

TCGAGTCGA-CAC AGTG-3’].

Also, AGTH = [5’-5a/I-AGT-CACAGTG], CG_H = [5’-5a/I-GCGCGCGC-CACAGTG]; G

(junctions are retained on the plasmid.

assays (signal and coding junctions) are diagrammed; for these substrates signal

were recovered from transiently transfected HDR37A cells by Hirt lysis and DNAs were digested with

dermine recombination frequencies as described above. Next, an aliquot of the DNA was digested with SalI and

ATP-dependent DNase (exonuclease V; US Biochemicals) to eliminate unarranged DNA. Unrearranged substrates contain a unique SalI restriction site. Open/shut products, as well as standard V(D)J rearrangements, are SalI (Fig. 3). The two SalI products are distinguished by bacterial cell plating as follows. One hundred AMP 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+ (potential open/shut products). Plasmid miniprep DNAs from AMP + CAM− colonies were tested for SalI.

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 homopolymer 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 pH200 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. pA10/H/T10-H-Sg and pT10/A10-H-Sg yielded recombination frequencies that were 2% of the level for pH200 (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 pH200 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

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

^aCalculated as percentage Amp' Cam' colonies from Amp' colonies.

^bRecombination frequency corrected for the recombination frequency of pH200 in the same experiment.

^cAverage corrected recombination frequency.

^dCalculated as percentage Amp' 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 lymphopoesis. 18-8 cells have constitutively active RAG1 and RAG2 gene expression and recombine 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. pA10H/T10H-Sg and pT10H/A10H-Sg yielded recombination frequencies that were 0.8 and 0.6% of the level for pJH200 (Table 1). Likewise, pG10H/C10H-Sg and pC10H/G10H-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. pC10H/T10H-Sg and pT10H/C10H-Sg reduced recombination to an $R_{av}$ of 6.0 and 7.6% respectively (Table 1). Similarly, two other substrates, pG10H/A10H-Sg and pA10H/G10H-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 T10H 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/A10H-Sg ($R_{av}$ = 11%) and pAGTH/T10H-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/C10H-Sg ($R_{av}$ = 43%) and pAGTH/G10H-Sg ($R_{av}$ = 101%) remained almost unchanged. 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**

<table>
<thead>
<tr>
<th>Substrate</th>
<th>Expt</th>
<th>No. Amp$^a$</th>
<th>No. Amp$^b$ +Cam$^b$</th>
<th>$R$ (%)</th>
<th>$R_{corr}$ (%)</th>
<th>$R_{av}$ (± SD)</th>
</tr>
</thead>
<tbody>
<tr>
<td>HDR37A:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>pGACH/TCCH-Sg</td>
<td>1</td>
<td>670 000</td>
<td>32 400</td>
<td>4.8</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td>(pJH200)</td>
<td>2</td>
<td>740 000</td>
<td>10 800</td>
<td>1.5</td>
<td>100</td>
<td></td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>2 440 000</td>
<td>81 000</td>
<td>3.3</td>
<td>100</td>
<td></td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>1 130 000</td>
<td>12 500</td>
<td>1.1</td>
<td>100</td>
<td></td>
</tr>
<tr>
<td></td>
<td>5</td>
<td>730 000</td>
<td>81000</td>
<td>1.1</td>
<td>100</td>
<td></td>
</tr>
<tr>
<td></td>
<td>6</td>
<td>580 000</td>
<td>19 400</td>
<td>3.3</td>
<td>100</td>
<td></td>
</tr>
<tr>
<td>pAGTH/G10H-Sg</td>
<td>1</td>
<td>1 900 000</td>
<td>51 600</td>
<td>2.7</td>
<td>56</td>
<td>101</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>690 000</td>
<td>98000</td>
<td>1.4</td>
<td>93</td>
<td>(±50)</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>960 000</td>
<td>49 000</td>
<td>5.1</td>
<td>155</td>
<td></td>
</tr>
<tr>
<td>pAGTH/C10H-Sg</td>
<td>1</td>
<td>780 000</td>
<td>15 100</td>
<td>1.9</td>
<td>40</td>
<td>43</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>500 000</td>
<td>19200</td>
<td>0.38</td>
<td>25</td>
<td>(±20)</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>550 000</td>
<td>11 600</td>
<td>2.1</td>
<td>64</td>
<td></td>
</tr>
<tr>
<td>pAGTH/A10H-Sg</td>
<td>1</td>
<td>1 070 000</td>
<td>17100</td>
<td>0.16</td>
<td>3.3</td>
<td>11</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>1 050 000</td>
<td>16600</td>
<td>0.16</td>
<td>11</td>
<td>(±7.9)</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>730 000</td>
<td>46000</td>
<td>0.63</td>
<td>19</td>
<td></td>
</tr>
<tr>
<td>pAGTH/T10H-Sg</td>
<td>1</td>
<td>1 010 000</td>
<td>310</td>
<td>0.031</td>
<td>0.65</td>
<td>3.6</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>610 000</td>
<td>8800</td>
<td>0.140</td>
<td>9.3</td>
<td>(±5.0)</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>960 000</td>
<td>220</td>
<td>0.023</td>
<td>0.70</td>
<td></td>
</tr>
<tr>
<td>18-8 (Td$^+$):</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>pGACH/TCCH-Sg</td>
<td>1</td>
<td>18 000</td>
<td>395</td>
<td>2.2</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td>(pJH200)</td>
<td>2</td>
<td>1 360 000</td>
<td>24 000</td>
<td>1.8</td>
<td>100</td>
<td></td>
</tr>
<tr>
<td>pAGTH/G10H-Sg</td>
<td>1</td>
<td>18 900</td>
<td>155</td>
<td>0.8</td>
<td>36</td>
<td>59</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>1 000 000</td>
<td>13 100</td>
<td>1.3</td>
<td>72</td>
<td></td>
</tr>
<tr>
<td>pAGTH/C10H-Sg</td>
<td>1</td>
<td>9100</td>
<td>280</td>
<td>3.1</td>
<td>141</td>
<td>112</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>604 000</td>
<td>8760</td>
<td>1.5</td>
<td>83</td>
<td></td>
</tr>
<tr>
<td>pAGTH/A10H-Sg</td>
<td>1</td>
<td>1 52 000</td>
<td>131</td>
<td>0.086</td>
<td>3.9</td>
<td>3.6</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>219 000</td>
<td>135</td>
<td>0.062</td>
<td>3.4</td>
<td></td>
</tr>
<tr>
<td>pAGTH/T10H-Sg</td>
<td>1</td>
<td>78 000</td>
<td>22</td>
<td>0.028</td>
<td>1.3</td>
<td>1.5</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>740 000</td>
<td>225</td>
<td>0.030</td>
<td>1.7</td>
<td></td>
</tr>
</tbody>
</table>

$^a$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 pH200 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_TA_{2}AH flanking RSS(23) produced only a slightly reduced recombination frequency relative to pH200 (pAGTH/A_{2}TA_{2}AH-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}TA_{4}H-Sg, R_{av} = 81%]. The interspersion of a TG dinucleotide in the midst of an A 5mer (AAATGAA-CACAGTG) gave a normal recombination frequency (p_{G2}CAGGH/A_{3}TGAAH-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-containing templates all produced recombination in the normal range when matched with other mixed composition coding ends (Tables 3 and 4). Likewise, p_{G2}CAGGH/A_{3}TGAAH-Sg and p_{G2}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). pH200 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
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(DJ) recombination. Thus the specificity of the V(DJ) recombination cleavage is normal, even for inhibited reactions.

Although most of the RSS junctions formed were precise, the lower frequency aberrant products that were ApaLV 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 (ApaLV) 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(DJ) 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 length (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(DJ) 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 recombination-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 is 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) (GTCGAH/C/GC/HA10H-Sg, GTCGAH/A10H-Sg and GTCGAH/T10H-Sg) (Fig. 3; Materials and Methods). In each case open/shut products at the GTCGAH–RSS region were quantitated. The GTCGAH sequence contains five of the six nucleotides of a SaII restriction enzyme site.
The paradoxical effects of coding DNA sequences on RSS products may occur by dissolution of a protein−DNA recombinational 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 synopsis 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 T3A2T2T1 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.

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

<table>
<thead>
<tr>
<th>Substrate</th>
<th>Expt</th>
<th>No. Ampf</th>
<th>No. Ampf + Camf</th>
<th>R (%)^a</th>
<th>Rsuv (% ± SD)</th>
<th>O/S (%)^b</th>
<th>O/Ssuv (% ± SD)</th>
</tr>
</thead>
<tbody>
<tr>
<td>pGTCG(\Delta)/CGCH-Sg</td>
<td>1</td>
<td>162 000</td>
<td>6300</td>
<td>3.9</td>
<td>3.7</td>
<td>0.022</td>
<td>0.063</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>95 000</td>
<td>2200</td>
<td>2.3</td>
<td>(20.9)</td>
<td>0.075</td>
<td>(±0.034)</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>380 000</td>
<td>15 900</td>
<td>4.2</td>
<td>0.102</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>372 000</td>
<td>16 100</td>
<td>4.3</td>
<td>0.053</td>
<td></td>
<td></td>
</tr>
<tr>
<td>pGTCG(\Delta)/A_{10}H-Sg</td>
<td>1</td>
<td>239 000</td>
<td>462</td>
<td>0.19</td>
<td>0.24</td>
<td>0.024</td>
<td>0.026</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>360 000</td>
<td>510</td>
<td>0.14</td>
<td>(±0.10)</td>
<td>0.007</td>
<td>(±0.014)</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>1 105 000</td>
<td>3950</td>
<td>0.36</td>
<td>0.030</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>1 295 000</td>
<td>3650</td>
<td>0.28</td>
<td>0.041</td>
<td></td>
<td></td>
</tr>
<tr>
<td>pGTCG(\Delta)/T_{10}H-Sg</td>
<td>1</td>
<td>207 000</td>
<td>363</td>
<td>0.17</td>
<td>0.15</td>
<td>0.018</td>
<td>0.011</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>360 000</td>
<td>475</td>
<td>0.13</td>
<td>(±0.02)</td>
<td>0.015</td>
<td>(±0.007)</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>1 950 000</td>
<td>2950</td>
<td>0.15</td>
<td>0.006</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>1 290 000</td>
<td>2050</td>
<td>0.16</td>
<td>0.003</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

^a Calculated as percentage Ampf Camf colonies from Ampf colonies.
^b Calculating as percentage Ampf Camf colonies from Ampf colonies.

Measuring the overall V(D)J recombination reaction efficiency, pGTCG\(\Delta\)/A_{10}H-Sg and pGTCG\(\Delta\)/T_{10}H-Sg (R = 0.17 and 0.15% respectively) were severely inhibited relative to pGTCG\(\Delta\)/CGCH-Sg (R = 3.1%), as expected. These values are ~ 1.6% of the level of recombination measured in the control, pGTCG\(\Delta\)/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, pGTCG\(\Delta\)/A_{10}H-Sg and pGTCG\(\Delta\)/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 xni-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 recombinational 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 synopsis 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 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 T3A2T2T1 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 \( A_n < 4 \) 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 \( 5'-A_2T_3-3' \), but not \( 5'-T_3A_2-3' \), give altered helical structures (22). Inhibitory coding sequences actually correlate well with this additional observation, since the coding sequence \( A_2T_3 \) (our nomenclature) inhibits V(D)J recombination even though the size of the polymer was three (T3) (12). Where A/T tracts were interrupted by other base pairs, as with the sequences \( A_3T_2A_2H \) and \( A_3TGA_2AH \), there was no appreciable inhibition of V(D)J recombination (Table 3). Thus, \( 5'-A_2T_3-3' \), but not \( 5'-T_3A_2-3' \), also inhibits V(D)J recombination.

Local distortions of DNA may ordinarily be innocuous. However, in conjunction with specific placing and synopsis 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