Sequence of the Spacer in the Recombination Signal Sequence Affects V(D)J Rearrangement Frequency and Correlates with Nonrandom Vκ Usage In Vivo

By Bertrand Nadel, Alan Tang, Guia Escuro, Geanncarlo Lugo, and Ann J. Feeney

The Scripps Research Institute, Department of Immunology, La Jolla, California 92037

Summary

Functional variable (V), diversity (D), and joining (J) gene segments contribute unequally to the primary repertoire. One factor contributing to this nonrandom usage is the relative frequency with which the different gene segments rearrange. Variation from the consensus sequence in the heptamer and nonamer of the recombination signal sequence (RSS) is therefore considered a major factor affecting the relative representation of gene segments in the primary repertoire. In this study, we show that the sequence of the spacer is also a determinant factor contributing to the frequency of rearrangement. Moreover, the effect of the spacer on recombination rates of various human Vκ gene segments in vitro correlates with their frequency of rearrangement in vivo in pre-B cells and with their representation in the peripheral repertoire.

Diversity of the Ig and TCR primary repertoires is generated during the V(D)J recombination process, in which the variable (V), diversity (D), and joining (J)[1] gene segments are somatically assembled to form a complete variable region (1, 2). However, the various V, D, and J gene segments do not contribute equally to the Ig and TCR repertoires (3–13). One main factor contributing to this biased representation is antigenic or ligand selection and subsequent clonal expansion, a process that begins as early as at the pre-B/T stage of differentiation, and continues in the periphery (14–17). But even before this stage nonrandom gene usage occurs, due to differences in the frequency with which the different gene segments rearrange. One factor that could influence the frequency of rearrangement is the relative proximal or distal location of the gene segment on the locus, as exemplified in extreme in the human kappa locus, where the entire distal half of the locus is rarely used (18–20). However, a major influence is likely to be variation in the efficiency of individual recombination signal sequences (RSSs) to promote rearrangement (21–23). The recombination process requires the presence of RSSs flanking each gene segment (24, 25). Each RSS consists of a conserved heptamer motif (consensus sequence: CACAGTG) and a conserved nonamer motif (consensus sequence: ACAAAAACC) separated by a spacer sequence of 12 or 23 bp. Alterations in the RSS can be extremely deleterious for recombination. Changes in the first three nucleotides of the heptamer have been shown to dramatically reduce the frequency of recombination, and changes in the other four positions have more varied effects (24–26). Similarly, the presence of three consecutive A's in the nonamer sequence is required for efficient recombination, and changes of >1 nucleotide in the spacer length result in a severe drop in the recombination frequency (25, 26). However, changes in the sequence of the spacer are thought to be unimportant (24, 26, 27), although two studies have suggested that this may not always be the case (28, 29). Most V, D, and J gene segments are flanked by RSSs containing some degree of polymorphism in their sequence, and consequently, variation from consensus in the heptamer and nonamer sequences and in spacer length could be a major factor affecting the representation of the gene segments in the primary repertoire.

In the course of a study aimed to assess the role of the RSS on the representation of the Vκ gene segments in the human κ repertoire, we used a sensitive competition substrate assay to compare and quantitate the relative recombination frequencies of a VκII gene, A2, and several members of the VκIII family. Since the A2 RSS has an A to G change in the fourth position of the nonamer, as do all VκII genes, whereas all gene segments from the VκIII family are flanked by consensus heptamers and nonamers, we were anticipating a lower rate of recombination of the A2 segment compared with a VκIII RSS. Surprisingly, we found that rearrangement of A2 is favored compared with the VκIII genes. Detailed analysis of the various RSSs showed that the A2 spacer mediates recombination at a

[1]*Abbreviations used in this paper:* D, diversity; J, joining; RSS, recombination signal sequence; V, variable.

J. Exp. Med. © The Rockefeller University Press • 0022-1007/98/05/1495/09 $2.00
Volume 187, Number 9, May 4, 1998 1495–1503
http://www.jem.org

consistently higher frequency than all the other Vκ spacers tested, thus compensating for the defect in the A2 nonamer. We show that each natural spacer flanking the VκIII segments analyzed mediates recombination at a different rate, and that two nucleotide changes in the spacer sequence can cause a sixfold decrease in recombination. Furthermore, to determine the effect of the spacer on the representation of the Vκ gene segments in the initial and peripheral repertoires, we analyzed the relative frequencies of rearrangement of the VκIII segments in vivo, and compared them with the corresponding in vitro data. We find a correlation between the effect of the spacer and the relative representation of VκIII segments both in the initial pre-B cell repertoire and in the peripheral repertoire.

Materials and Methods

Construction of the Competition Plasmids. The organization of the competition substrate plasmid is shown in Fig. 1 *A*. The coding ends to be joined, flanked by their RSSs, are separated by a termination signal. The upstream P_{tac} promoter will transcribe the downstream CAT gene only when the termination signal is deleted by V(D)J recombination. Two Vκ segments are located on the 5′ side of the termination signal and are therefore competing for rearrangement with the Jκ1 segment located on the 3′ side. Coding ends flanked by their RSSs were obtained by PCR from human genomic DNA, with primers containing the appropriate restriction sites to enable insertion into the vector backbone (Fig.1 *A* and see below). PCR primers were designed so that PCR products would contain >100 bp of the genomic DNA upstream of the RSS and 10 (P38b, P40b, P42b, and P43b) to 100 bp (P32b, P41b, AF126, and AF166) downstream of the RSS. Subsequent modifications of the RSSs were introduced by PCR from the appropriate clones using mismatched oligonucleotide primers (P38b, P40b, P42b, and P43b). All cloned PCR products were confirmed by sequencing.

Competition Plasmids and PCR Primers. Plasmids were constructed as follows (parentheses indicate primers and restriction sites, and see also Tables 1–3): Comp6: (AF125, AF126) (P32a,P32b); Comp20: (P32a, P32b) (AF125, AF166); Comp22: (P39a, P38b) (AF125, AF166); Comp32: (P41a, P41b) (P42a, P42b); Comp25: (P39a, P40b) (AF125, AF166); Comp24: (P39a, P38b) (P39a, P40b); Comp31: (P39a, P38b) (P39a, P40b); Comp26: (P41a, P41b) (AF125, AF126); Comp34: (P41a, P41b) (AF125, P43b). 5′ primers: P32a: 5′ TCTCACGCGTGCCAGATCGATTCAG 3′; P39a: 5′ CGGCCACGCGTCAGGGACAGATTTC 3′; P41a: 5′ CTTCACGCGTACCATCAGCAGACTGG 3′; P42a: 5′ TTCAGCGGCCGCGGGTCAGGGACAGA 3′; and AF125: 5′ TTCAGCGGCCGCGGGTCTGGGACA 3′. 3′ primers: P32b: 5′ AGTCCGCGGCCGCAAGCCCGTCGACTGAG 3′; P38b: 5′ CACCGCGGCCGCGAGGTTTTTGTATTGGTC 3′; P40b: 5′ CACCGCGGCCGCGAGGTTTTTGTTTCAAGCTGAATCACTGTGGG AGGAAG 3′; P41b: 5′ CATTGCGGCCGCTATCTGTAAAGGAAG 3′; P42b: 5′ GGACAGTCGACCAGGGAGGTTTTTGATTG 3′; P43b: 5′ GATGGTCGACTTGAGGTTTTTGTTTCATGTTGAATCACTG 3′; AF126: 5′ CCAAGTCGACCATCGATGCTAACACTGAGTGG 3′; and AF166: 5′ TAAAGTCGACTTGTCATTGCAGCAGCTA. PCR screening primers: AF74: 5′ TATCCCTGATGGATACAAGG 3′; and P4b: 5′ AACACATCTAGGCTTGC 3′.

Recombination Assay. 18.8 Abelson murine leukemia virus–transformed pre-B cells (2×10^7) are transiently transfected with 20 μg of Qiagen-purified plasmid (Qiagen, Chatsworth, CA) by electroporation (960 μF, 0.3 Kv) and resuspended in 10 ml of RPMI 1640 supplemented with 10% FCS, 2 mM glutamine, and 1 mM caffeine. After 48 h, plasmids are recovered from transfected cells by alkaline lysis, followed by DpnI/SpeI digestion. Digested plasmids are then electroporated into JM109 *Escherichia coli* and plated on plates containing ampicillin (100 μg/ml) and chloramphenicol (5 μg/ml). The status of the recombination was assessed by PCR, using primers located upstream and downstream of the most distal gene segments (AF74 and P4; Fig.1 *B*). PCR screens are performed from the colonies by resuspending bacterial cells containing plasmid directly in the PCR mixture, and amplifying with AF74 and P4 primers for 30 cycles (1 min at 94°C, 1 min at 55°C, and 2 min at 72°C). In this assay, recombination of the external gene segment with Jκ1 generates a 370-bp PCR product, whereas recombination of the internal gene generates a 520–630-bp PCR product, depending on the coding ends (Fig. 1 *B*). Rare background unrearranged clones can be observed as 1,580-bp PCR products. The reliability of this assay was confirmed by simultaneous sequencing of random clones.

Statistical Analysis. For each substrate, two to five independent transfections were performed. The ratio of rearrangements to the external versus the internal competing RSSs was determined for each transfection that yielded >15 colonies, and the mean and SEM were calculated and shown in the tables. Of the 31 transfections, only 3 gave a such a low yield of colonies. In general, 30–60 colonies were sampled per transfection. In addition, the data was also calculated by totaling all of the colonies from all of the transfections using each substrate, and then determining the ratio of rearrangements to the external versus internal RSSs based on the sum (calculations not shown). In all but three cases, the two values (ratio determined on the sum of all colonies versus ratio of each independent transfection calculated, and a mean determined) were the same or only one percentage point different. The three exceptions differed by only 3–4%.

Preparation of Cells. Bone marrow pre-B cells were isolated from marrow aspirates. Both bone marrow and peripheral blood mononuclear cells were prepared by Ficoll-Hypaque purification. Bone marrow cells were stained with FITC-conjugated CD10 (Becton Dickinson, San Jose, CA) and biotin-labeled anti–human IgM (PharMingen, San Diego, CA), followed by PE-labeled streptavidin. $CD10^+$ surface Ig^- cells (sIg^-) were isolated by cell sorting on a FACS® Vantage. All peripheral blood and bone marrow samples were collected by the Scripps General Clinical Research Center.

Analysis of VkIII Rearrangements. DNA was prepared as previously described (30). VκIII PCRs were performed with a VκIII leader primer and a set of Jκ primers as previously described (20). Cloning and sequencing was done as previously described (31).

Results

Relative Frequency of Recombination of VkA2 Compared with Other Vk Gene Segments. To analyze the potential role of the RSS on the representation of the VκII genes, all of which are flanked by nonconsensus nonamers, relative to VκIII gene segments, all of which are flanked by consensus RSSs, we used an in vitro competition substrate assay in which two Vκ segments are in competition for rearrangement with a single Jκ1 gene segment (Fig. 1 *A*). Since both Vκ segments are on the same plasmid, this strategy allowed

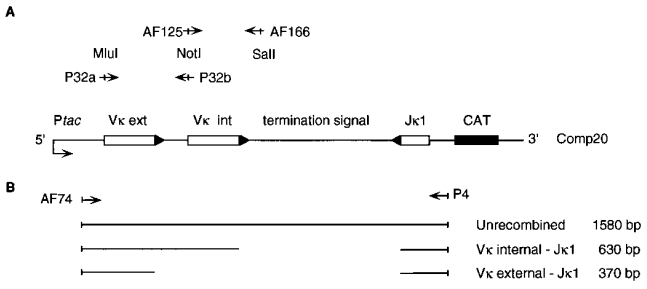

Figure 1. (*A*) Competition substrate map (not to scale). Comp20 is given as an example. The two sets of primers used to generate the coding ends (*open boxes*) flanked by their RSSs (*black triangles*) are shown with the restriction sites used for cloning. (*B*) PCR screen, used to determine the frequency of rearrangements to the internal versus external Vκ coding end. Expected sizes of the PCR products are indicated.

us to directly compare the frequency of rearrangement of one RSS relative to the other with great precision and reliability. We have shown in another study that the external or internal position of the various gene segments in the plasmid has no effect on the relative frequencies of recombination.[2] We first constructed a competition substrate containing the A2 segment flanked by its RSS in competition for rearrangement with the L6 segment, a member of the VκIII family (Comp6; Table 1). Since the VκIII L6 RSS contains a consensus heptamer and nonamer, we were anticipating a bias favoring the recombination of the L6 segment compared with the A2 segment. Surprisingly, results in Table 1 show a dramatic under-use of the L6 segment compared with the A2 segment in Comp6 (14:86; Table 1). To test the possibility that the L6 segment would have an unsuspected defect in its recombination abilities, we replaced the L6 segment with an A27 segment, which is one of the most frequently observed Vκ genes in vivo. Although not as dramatic as for the L6 segment, results with Comp20 still show a significant bias of rearrangement toward the A2 segment (67:33, or 2-fold decrease; Table 1). In another study, we showed that the A2 nonamer change does result in a 2.6-fold reduction in the rate of recombination compared with the same segment with a consensus nonamer,[2] and similar results were obtained by Akamatsu et al. (26). Altogether, these results suggest that the A2 segment is better at promoting recombination than are other Vκ gene segments, despite the reduction in recombination frequency due to its defective nonamer. In addition, these results indirectly show that the L6 and A27 segments rearrange with different frequencies, although they bear the same consensus heptamers and nonamers. This indicates that frequency of recombination is controlled not only by the sequence of the heptamer and nonamer, but that another parameter in the A2 segment compensates for the deleterious effect of the nucleotide change in the A2 nonamer.

The RSS Spacer Plays a Determinant Role in the Frequency of Recombination. To assess the difference in the frequency of

Table 1. *Relative Recombination Frequencies of the A2 Segment Compared with VκIII Segments in Competition Substrates*

Plasmid	External Vκ gene				Internal Vκ gene				Results		
	Coding end	Heptamer	Spacer	Nonamer	Coding end	Heptamer	Spacer	Nonamer	No. of colonies	No. of transfections	Distal/proximal (± SE)
Reference*	AGCTTCCTCC	CACAGTG	GTACAGACCAAT	ACAAAAACC	AGCTTCCTCC	CACAGTG	GTACAGACCAAT	ACAAAAACC			
Comp 6 L6	-CTGG-----	-------	A-T-CACATG-A	---------	A2 ----------	-------	----------	---G-----	83	3	14:86 (± 5.7)
Camp20 A2	----------	-------	----------	---G-----	A27 GCTCA-----	-------	A-T---CTTG-A	----------	166	5	67:33 (± 4.7)

*The reference sequence consists of the VκA2 coding end and spacer, with consensus heptamer and nonamer.

[2]Nadel, B., A. Tang, G. Lugo, V. Love, G. Escuro, and A.J. Feeney, manuscript in preparation.

recombination between A27 and A2 segments in the same context of consensus heptamer and nonamer sequences, we designed a construct similar to Comp20, in which we restored the consensus sequence in the A2 nonamer (Comp22; Table 2). We therefore analyzed the relative frequencies of recombination of the A27 segment versus an A2 segment that was flanked by a consensus RSS, but which retained the A2 spacer. As expected, the bias towards the A2 segment was even greater, now recombining in 73% of the cases (Table 2). Identical results were obtained for Comp32, in which the position of the two competing segments are inverted (Table 2), thus confirming that position of the segment does not influence the relative frequencies of rearrangement.

The only differences between the A27 and A2 segments in Comp22 and 32 are the coding end sequence and the RSS spacer sequence. To accurately determine the effect of the spacer difference alone, we designed a plasmid containing two A2 segments, flanked by consensus RSSs containing either the A2 or the A27 spacer (Comp24; Table 2). The two segments are therefore identical in all respects except for the spacers. In this plasmid, the recombination frequency of the segment containing the A2 spacer was significantly increased (1.8-fold) compared with that of the segment containing the A27 spacer (64:36, different from a 50:50 distribution, normal approximation to the exact binomial test: $P < 0.001$).

Although the terminal five base pairs of A2 and A27 coding ends are identical, there are changes further 5′ (Table 2). To test the potential effect of those nucleotide changes on the recombination frequency, we designed a construct containing the A27 segment (internal) and a modified A2 segment (external) containing the A2 coding end flanked by the A27 RSS (Comp25; Table 2). The only difference between the two segments is the sequence of the coding end. Results show a minimal favoring effect of the A2 coding end segment compared with A27 (57:43, or 1.3-fold increase), although the relative frequencies of recombination are not significantly different from a 50:50 distribution (normal approximation to the exact binomial test: $P = 0.24$). These results suggest that the presence of the A2 spacer induces a significant increase in the recombination frequency compared with the A27 spacer, whereas the differences between A2 and A27 coding ends have a small influence on the frequency of recombination.

To test the general character of the role of the spacer in the frequency of recombination in the context of the same coding ends, we also tested an artificial variant of Comp24, containing two random base pair changes in the A27 spacer at positions 2 and 10 (Comp31, internal segment; Table 2). These substitutions gave rise to a dramatic drop in the recombination of the gene segment containing the variant spacer compared with the A2 spacer (86:14, or 6-fold decrease), and also compared with the original A27 spacer (3-fold decrease), indicating that the spacer alone can have a major effect on the frequency of recombination.

Role of the RSS Spacer in the Representation of the Vκ gene Segments in the Human κ Repertoire. Since all VκIII genes

Table 2. *Relative Recombination Frequencies in Competition Substrates Containing Various Spacers and Coding Ends in the Context of Consensus Heptamers and Nonamers*

Plasmid	External Vκ gene				Internal Vκ gene				Results		
	Coding end	Heptamer	Spacer	Nonamer	Coding end	Heptamer	Spacer	Nonamer	No. of colonies	No. of transfectants	Distal/proxmal (± SE)
Reference*	AGCTTCCTCC	CACAGTG	GTACAGACCAAT	ACAAAAACC	AGCTTCCTCC	CACAGTG	GTACAGACCAAT	ACAAAAACC			
Comp22	A2 ----------	-------	----------	---------	A27 GCTCA-----	-------	A-T---CTTG-A	---------	150	5	73:27 (± 3.8)
Comp32	A27 GCTCA-----	-------	A-T---CTTG-A	---------	A2 ----------	-------	----------	---------	111	2	27:73 (± 3.1)
Comp24	A2 ----------	-------	----------	---------	A2 ----------	-------	A-T---CTTG-A	---------	245	3	64:36 (± 1.5)
Comp25	A2 ----------	-------	A-T---CTTG-A	---------	A27 GCTCA-----	-------	A-T---CTTG-A	---------	108	2	57:43 (± 3.3)
Comp31	A2 ----------	-------	----------	---------	A2 ----------	-------	AGT---CTTC-A	---------	114	2	86:14 (± 3.7)

*The reference sequence consists of the VκA2 coding end and spacer, with consensus heptamer and nonamer.

Table 3. *Relative Recombination Frequencies in Competition Substrates Containing VκIII Gene Segments*

Plasmid	External Vκ gene				Internal Vκ gene				Results		
	Coding end	Heptamer	Spacer	Nonamer	Coding end	Heptamer	Spacer	Nonamer	No. of colonies	No. of transfections	Distal/proximal (± SE)
Reference*	GCTCACCTCC	CACAGTG	ATTCAGCTTGAA	ACAAAAACC	GCTCACCTCC	CACAGTG	ATTCAGCTTGAA	ACAAAAACC			
Comp26 A27	----------	-------	------------	---------	L6 G--GG-----	-------	----CA-A----	---------	155	4	73:27 (± 3.8)
Comp34 A27	----------	-------	------------	---------	L2 G--GG-----	-------	-----A-A----	---------	224	5	67:33 (± 4.3)

*The reference sequence consists of the VκA27 coding end and spacer, with consensus heptamer and nonamer.

have the consensus heptamer and nonamer sequence, we previously assumed that their RSS would support the same frequency of recombination. However, since the data presented here show that the spacer can significantly affect the recombination frequency, we compared the relative rates of recombination of the actual A27, L6, and L2 gene segments flanked by their natural RSSs (Comp26 and Comp34, Table 3). Since all of those segments contain consensus nonamers and heptamers, and have highly conserved coding end sequences, these substrates allowed us to assess the role of the different spacers on the relative recombination rates. The L6 gene segment has three nucleotide changes at positions 5, 6, and 8 of the spacer compared with A27 (Table 3). Table 1 shows a 2-fold decrease in the recombination of the A27 segment compared with A2, and a 6.1-fold decrease in the recombination of the L6 segment compared with A2. We therefore expected a 3-fold difference in the rates of recombination of the A27 segment compared with L6. Comp26 directly compares those two segments and confirms a 2.7-fold decrease in the rearrangement of the L6 segment compared with the A27 segment (73:27%, Table 3). A27 and L2 segments differ at only two nucleotides, positions 6 and 8 in the spacer, and L2 and L6 segments have only one nucleotide difference, position 5 in the spacer (Table 3). Results show a 2-fold reduction in the rearrangement frequencies of L2 compared with A27 (67:33%, Comp34), and therefore a 1.35-fold reduction of the L6 segment compared with L2. These results confirm further that a few or even a single nucleotide substitution(s) in the spacer can have a significant effect on the frequency of recombination.

To test if the reductions in recombination frequencies due to the changes in the spacer would correlate to the representation of the different segments in the actual human κ repertoire, we determined the relative representation of the VκIII gene segments in vivo. To amplify the Vκ genes in an unbiased manner, the Vκ PCR primer must match the sequence exactly in all genes to be amplified. For each Vκ family, the leader sequence contains these requisite stretches of identity, and thus we compared the relative rates of rearrangement of the VκIII family members using a primer to a conserved region in the VκIII leader. However, such an analysis precludes comparison of the relative rates of rearrangement of the VκII gene, A2, to any of the VκIII genes. Analysis of nonproductive rearrangements from sIg^- $CD10^+$ adult bone marrow pre-B cells provides a good representation of the initial frequency of κ rearrangements before selection, whereas productive sequences from adult PBLs represent the distribution of the VκIII gene segments in the peripheral repertoire. A total of 163 nonproductive sequences were obtained from three bone marrow pre-B cell samples, and 42 productive sequences were obtained from two peripheral blood donors. Relative recombination frequencies for the A27, L2, and L6 gene segments are shown in Fig. 2. As observed in the in vitro data, the A27 segment is the most highly represented of the VκIII segments in the kappa repertoire in vivo. The frequency of rearrangement of the A27 segment as observed in pre-B cells is 1.8-fold

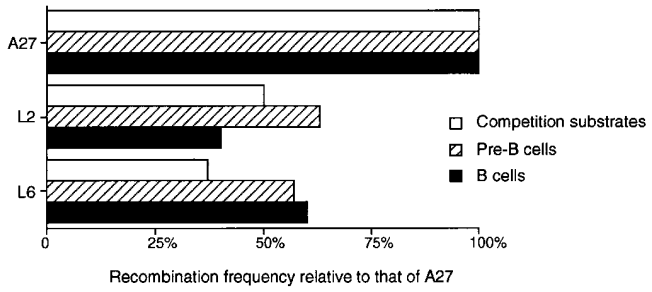

Figure 2. Relative recombination frequencies of A27, L2, and L6 VκIII gene segments in vitro and in vivo. All recombination frequencies are calculated relative to A27. 163 nonproductive sequences from $CD10^{+}sIg^{-}$ bone marrow pre-B cells, and 42 productive sequences from adult peripheral blood cells, were obtained by VκIII-specific PCR. The L10 pseudogene rearranged with the same frequency as L2 in pre-B cells, and the three distal Vκ genes, duplicates of A27, L2, and L6, were rarely observed (9 out of 163 sequences in pre-B cells and 2 out of 42 sequences in peripheral blood), consistent with previous observations of the paucity of rearrangements of the Vκ genes in the distal half of the Vκ locus (18, 20, 35).

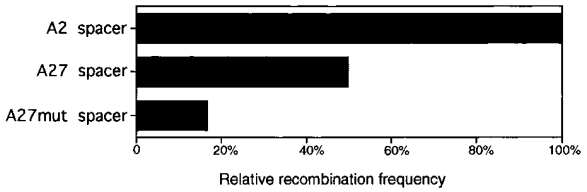

Figure 3. Effect of the spacer sequence on relative frequencies of recombination in the context of A2 coding ends and consensus heptamers and nonamers. Results are issued from Comp24 and Comp31 (Table 2) and are expressed relative to the gene segment containing the A2 spacer.

higher than the L2 segment and 1.6-fold higher than the L6 segment. This is in agreement with our previous observation in nonproductive sequences from cord blood that A27 rearranges approximately twice as often as L2 or L6 (20). Also, this relative frequency of initial rearrangements is similar to that observed in the peripheral repertoire by us (Fig. 2), and by others (19, 32–35). Thus, the relative representation of the A27 and L2 segments within the VκIII family correlates with their frequency of rearrangement in vitro, and therefore suggests a direct influence of the spacer sequence on their representation in the primary repertoire. Although the frequency of rearrangement of the L6 segment is also reduced compared with A27 (1.6-fold), the fact that its representation is higher than expected from the in vitro data (2.7-fold decrease) suggests that factors other than the RSS spacer sequence may play an additional role in vivo. It is interesting to note that the gene segment representation in the initial repertoire has a major influence in shaping the peripheral repertoire, as shown in Fig. 2, since the VκIII gene segments have a very similar relative distribution in nonproductive sequences from pre-B cells and productive sequences from peripheral blood. These results show the fundamental importance of the initial shaping of the primary repertoire by the recombination process.

Discussion

The RSS Spacer Sequence Affects the Outcome of Recombination. In this study, we clearly show that the nucleotide sequence of the spacer is a determinant factor contributing to the recombination frequency of the gene segments (Fig. 3). Using a sensitive recombination substrate assay, in which two V genes are in competition for rearrangement, we were able to quantify up to a sixfold decrease in recombination rates merely due to nucleotide changes in the spacer sequence. We found that as little as one nucleotide difference in the spacer sequence can result in a significant decrease of recombination frequency, and that the spacer sequence flanking natural gene segments contributes significantly to their relative rearrangement frequencies and representation in the repertoire.

These results are in contrast with earlier reports suggesting that only the sequence of the heptamer and nonamer motifs or the length of the spacer, but not the sequence of the spacer, would significantly affect the frequency of recombination (24, 26, 27). This paradigm was based on the following considerations. First, the fact that there is no apparent conservation in the sequence of the RSS spacer suggested that the sequence would not play any role in the outcome of the recombination. However, after compiling an extensive database of published RSSs, Ramsden et al. observed that RSS spacers do in fact possess moderate conservation of sequence (36). Although the level of conservation of the spacer sequence is not nearly as high as found for the heptamer and nonamer motifs, the lack of strong conservation does not provide evidence for the absence of influence on recombination, and it is conceivable that sequence departure from consensus signal might represent a way to control the representation of gene segments in the primary repertoire (27). In this case, the spacer sequence might act as a "fine tuning" of the recombination levels. Second, the initial classic experiments that analyzed the role of the different RSS motifs on the frequency of recombination found drastic effects of some changes in the RSSs, resulting in a >100-fold decrease in recombination, but only minimal effects of differences in the spacer sequence (ranging from a 1.2- to 2.2-fold decrease) (24, 26, 27). However, these minor effects of changes in the spacer sequence were concluded to be insignificant because they fell below the resolution power of the assay, since the experiments were conducted with traditional recombination substrates. On the other hand, competition substrates contain all segments to be recombined within the same plasmid, and, due to the elimination of the variability in transfection efficiency of two different substrates, are the method of choice to quantify small differences in relative frequencies of recombination with great precision and reliability. Using such a competition substrate, Fanning et al. have recently shown a significant 1.6-fold decrease in the recombination frequency upon a single nucleotide change (A to C) at the 5th position of a mouse Vκ spacer (28). The presence of an A in this fifth position has been shown to be particularly

conserved (present in 67% of 12-bp spacers and 64% of 23-bp spacers) (36). In agreement with this observation, we show here that L6, which differs from L2 only at this conserved position 5 in the spacer (A5 to C5), displays a 1.35-fold reduction in its recombination frequency compared with L2. In addition, we show changes in the spacer which can result in up to a 6-fold decrease in recombination rates, a range comparable to that of some changes in the heptamer and nonamer motifs.

There is a previous in vivo example where a single nucleotide difference in the spacer between two alleles of a TCR BV gene resulted in a 6.7-fold difference in their expression level in the peripheral blood (29). In this case, there were several explanations for this effect. The change created a CpG sequence, which might be methylated in vivo, thus inhibiting rearrangement (37, 38). Also, this 23-bp spacer has a heptamer-like sequence located 13 bp from the nonamer, and so the allelic change could result in inappropriate cleavage of this RSS, with the internal heptamer-like sequence mimicking a 12-bp RSS. However, the data presented here suggest that it is possible that the spacer sequence itself may be responsible for this decreased recombination independent of methylation. All together, these data indicate that spacer sequences do affect the frequency of recombination, shown here in a range from 1.35–6-fold. It is possible that a systematic search of favorable or unfavorable positions in the spacer could reveal further deleterious effects. Of relevance is the recent study that showed that the footprint of RAG1 extended beyond the nonamer into the adjacent region of the spacer (39). Thus, there may be some direct contacts of the recombinase with at least part of the spacer sequence, which could partially explain our observed modulation of recombination frequencies by differences only in the spacer.

The drastic effects described for some nucleotide changes in the heptamer and nonamer motifs initially led to the concept that more modest decreases in the recombination frequency would be biologically nonsignificant. However, there is little physiological basis for this idea, and, in fact, we have shown that a four- to fivefold decrease in the recombination rate of the A2 gene segment is genetically associated with increased susceptibility to *Haemophilus influenzae* type b disease in Navajos (40).[2] It is therefore important not to underestimate the biological relevance of such modest decreases in recombination frequency. This is especially relevant since we show in this study that the frequency with which all three functional VκIII gene segments rearrange initially correlate very well with their representation in the peripheral repertoire.

Several studies have shown that various factors other than the RSS can affect the frequency of recombination of gene segments. We found in this study that the effect of the spacer on the rate of recombination might also be dependent on the interaction with other surrounding sequence elements, such as the coding ends (Comp25; Table 2). Studies on the effect of the coding end sequence on the frequency of recombination have shown that the presence of As or Ts immediately adjacent to the heptamer markedly diminish the recombination efficiencies, probably by affecting the efficiency of the initial nick at the coding end/heptamer border (41–44). In addition, Ezekiel et al. showed that the influence of these nucleotides is dependent on the location relative to the 12-bp or 23-bp spacer, suggesting an asymmetrical positioning of the recombinase machinery (43). In our study, the last five nucleotides of all coding ends are identical, and therefore more internal nucleotides would have to account for the small difference observed in the frequency of recombination of the A27 versus A2 coding ends (Comp25). Although internal nucleotides are unlikely to participate directly in the initial nicking, it is conceivable that internal sequence motifs influence the binding of some components of the recombinase machinery. This is in agreement with recent experiments suggesting that flanking sequences outside coding regions can also affect the outcome of recombination (45, 46). All together, these results suggest that, although the appropriately spaced heptamer and nonamer are the only *cis*-elements necessary for the recombination to take place, numerous other factors in *cis* outside the RSS are involved in the fine modulation of the efficiency of recombination.

Functional Role of the RSS Spacer in the Recombination. To what extent is the spacer sequence influencing the recombination frequency, and how does it affect the representation of the κ gene segments in the initial and peripheral repertoires? To answer this question, we compared the relative frequency of recombination of the VκIII genes A27, L2, and L6 in competition substrates in vitro with that in the nonproductive rearrangements of $sIg^{-}CD10^{+}$ sorted bone marrow pre-B cells in vivo (Fig. 2). Results show that relative frequencies of recombination of the A27 and L2 gene segments in vitro correlate with their relative frequency of recombination in bone marrow pre-B cells in vivo. Although still rearranging at a lower frequency than the A27 segment in vivo, the L6 gene segment seems to rearrange slightly more often and be expressed more often in vivo than in our competition substrates. This indicates that other factors in addition to the RSS, i.e., promoter activity, relative location of the gene segment within the locus, or proximity to yet undescribed enhancers or locus control regions, might also play an important role in shaping the repertoire. Since we were not able to compare in vivo rearrangement rates of the VκIII genes with the VκII gene A2 using our VκIII family–specific PCR, we do not have direct data addressing the relative in vivo rearrangement frequency of this gene. However, Foster et al. reported results of single cell PCR on peripheral blood cells where, by analyzing the frequency of non-productive rearrangements, they could estimate the relative rates of rearrangement of all Vκ genes (35). They observed that A18, the proximal region copy of A2, rearranged with a very high frequency, thus supporting our in vitro data. All genes in the distal region of the Vκ locus, such as A2, rearrange with a vastly reduced frequency, so the A18 gene, which is almost identical in sequence to A2 except for a small number of differences

spread throughout the coding region, is the appropriate gene for this comparison (18, 20, 35).

Furthermore, we, and many others, have shown that the Vκ gene A27 is present at approximately twice the frequency of L2 and L6 in the peripheral repertoire of both adults and newborns (19, 20, 32–35), suggesting minimal variation in relative representation due to antigenic expansion on the basis of the κ chain. Thus, the effect of the spacer is directly reflected in the representation of some gene segments in the peripheral repertoire, showing the importance of the initial recombination event. These results are in agreement with previous observations showing that over or under use of certain segments in vivo correlates with the ability of their RSS to support V(D)J recombination in vitro (22, 23).

In conclusion, we found that the sequence of the spacer can significantly affect the outcome of recombination, at the same or higher levels than some changes in the heptamer and nonamers. In addition, we show that the effect of the spacer on recombination rates of various Vκ gene segments as measured in vitro correlates with their frequency of rearrangement in vivo and with their representation in the peripheral repertoire.

This work was supported by National Institutes of Health (NIH) grants AI-37098 and AI-29672. B. Nadel was supported by postdoctoral fellowship 1-33-96B from the American Cancer Society, California Division, and G. Lugo was supported by training grant GM-08303 from NIH. The Scripps General Clinical Research Center is supported by NIH grant M01 RR-00833. This is manuscript No. 11297-IMM from The Scripps Research Institute.

Address correspondence to Ann J. Feeney, The Scripps Research Institute, Department of Immunology, IMM-22, 10550 North Torrey Pines Rd., La Jolla, CA 92037. Phone: 619-784-2979; Fax: 619-784-9190; E-mail: feeney@scripps.edu

Received for publication 8 January 1998 and in revised form 27 February 1998.

References

1. Alt, F.W., T.K. Blackwell, and G.D. Yancopoulos. 1987. Development of the primary antibody repertoire. *Science*. 238: 1079–1087.
2. Tonegawa, S. 1983. Somatic generation of antibody diversity. *Nature*. 14:575–581.
3. Yancopoulos, G.D., 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.
4. Dildrop, R., U. Krawinkel, E. Winter, and K. Rajewsky. 1985. V_H-gene expression in murine lipopolysaccharide blasts distributes over the nine known V_H-gene groups and may be random. *Eur. J. Immunol.* 15:1154–1156.
5. Perlmutter, R.M., J.F. Kearney, S.P. Chang, and L.E. Hood. 1985. Developmentally controlled expression of immunoglobulin V_H genes. *Science*. 227:1597–1600.
6. Wu, G.E., and C.J. Paige. 1986. V_H gene family utilization in colonies derived from B and pre-B cells detected by the RNA colony blot assay. *EMBO (Eur. Mol. Biol. Organ.) J.* 5:3475.
7. Reth, M.G., S. Jackson, and F.W. Alt. 1986. VHDJH formation and DJH replacement during pre-B differentiation: nonrandom usage of gene segments. *EMBO (Eur. Mol. Biol. Organ.) J.* 5:2131–2138.
8. Schulze, D.H., and G. Kelsoe. 1987. Genotypic analysis of B cell colonies by in situ hybridization: stoichiometric expression of three V_H families in adult C57BL/6 and BALB/c mice. *J. Exp. Med.* 166:163–172.
9. Jeong, H.D., and J.M. Teale. 1988. Comparison of the fetal and adult functional B cell repertoires by analysis of VH gene family expression. *J. Exp. Med.* 168:589–603.
10. Yancopoulos, G.D., B.A. Malynn, and F.W. Alt. 1988. Developmentally regulated and strain-specific expression of murine V_H gene families. *J. Exp. Med.* 168:417–435.
11. Freitas, A.A., M.-P. Lembezat, and A. Coutinho. 1989. Expression of antibody V-regions is genetically and developmentally controlled and modulated by the B lymphocyte environment. *Int. Immunol.* 1:342–354.
12. Malynn, B.A., 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.
13. Freitas, A.A., L. Andrade, M.-P. Lembezat, and A. Coutinho. 1990. Selection of VH gene repertoires: differentiating B cells of adult bone marrow mimic fetal development. *Int. Immunol.* 2:15–23.
14. Gu, H., D. Tarlinton, W. Müller, K. Rajewsky, and I. Förster. 1991. Most peripheral B cells in mice are ligand selected. *J. Exp. Med.* 173:1357–1371.
15. Melchers, F., H. Karasuyama, D. Haasner, S. Bauer, A. Kudo, N. Sakaguchi, B. Jameson, and A. Rolink. 1993. The surrogate light chain in B-cell development. *Immunol. Today*. 14:60–68.
16. Groettrup, M., and H. von Boehmer. 1993. A role for a pre-T-cell receptor in T-cell development. *Immunol. Today*. 14: 610–614.
17. ten Boekel, E., F. Melchers, and A.G. Rolink. 1997. Changes in the VH gene repertoire of developing precursor B lymphocytes in mouse bone marrow mediated by the pre-B cell receptor. *Immunity*. 7:357–368.
18. Klein, R., R. Jaenichen, and H.G. Zachau. 1993. Expressed human immunoglobulin κ genes and their hypermutation. *Eur. J. Immunol.* 23:3248–3271.
19. Cox, J.P.L., I.M. Tomlinson, and G. Winter. 1994. A directory of human germ-line V_κ segments reveals a strong bias in their usage. *Eur. J. Immunol.* 24:827–836.
20. Feeney, A.J., G. Lugo, and G. Escuro. 1997. Human cord

blood κ repertoire. *J. Immunol.* 158:3761–3768.
21. Suzuki, H., and H. Shiku. 1992. Preferential usage of J_{H2} in D-J joinings with D_{Q52} is determined by the primary DNA sequence and is largely dependent on recombination signal sequences. *Eur. J. Immunol.* 22:2225–2230.
22. Ramsden, D.A., and G.E. Wu. 1991. Mouse κ light-chain recombination signal sequences mediate recombination more frequently than do those of λ light chain. *Proc. Natl. Acad. Sci. USA*. 88:10721–10725.
23. Connor, A.M., L.J. Fanning, J.W. Celler, L.K. Hicks, D.A. Ramsden, and G.E. Wu. 1995. Mouse V_H7183 recombination signal sequences mediate recombination more frequently than those of V_HJ558. *J Immunol.* 155:5268–5272.
24. 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.
25. 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.
26. Akamatsu, Y., N. Tsurushita, F. Nagawa, M. Matsuoka, K. Okazaki, M. Imai, and H. Sakano. 1994. Essential residues in V(D)J recombination signals. *J. Immunol.* 153:4520–4529.
27. Wei, Z., and M.R. Lieber. 1993. Lymphoid V(D)J recombination. *J. Biol. Chem.* 268:3180–3183.
28. Fanning, L., A. Connor, K. Baetz, D. Ramsden, and G.E. Wu. 1996. Mouse RSS spacer sequences affect the rate of V(D)J recombination. *Immunogenetics*. 44:146–150.
29. Posnett, D.N., C.S. Vissinga, C. Pambuccian, S. Wei, M.A. Robinson, D. Kostyu, and P. Cancannon. 1994. Level of human TCRBV3S1 (Vβ3) expression correlates with allelic polymorphism in the spacer region of the recombination signal sequence. *J. Exp. Med.* 179:1707–1711.
30. Chukwuocha, R.U., B. Nadel, and A.J. Feeney. 1995. Analysis of homology-directed recombination in VDJ junctions from cytoplasmic Ig^- pre-B cells of newborn mice. *J. Immunol.* 154:1246–1255.
31. Feeney, A.J., and D. Thuerauf. 1989. Sequence and fine specificity analysis of primary 511 anti-phosphorylcholine antibodies. *J. Immunol.* 143:4061–4068.
32. Bridges, S.L., Jr., S.K. Lee, M.L. Johnson, J.C. Lavelle, P.G. Fowler, W.J. Koopman, and H.W. Schroeder, Jr. 1995. Somatic mutation and CDR3 lengths of immunoglobulin κ light chains expressed in patients with rheumatoid arthritis and in normal individuals. *J. Clin. Invest.* 96:831–841.
33. Weber, J.-C., G. Blaison, T. Martin, A.-M. Knapp, and J.-L. Pasquali. 1994. Evidence that the VκIII gene usage is nonstochastic in both adult and newborn peripheral B cells and that peripheral $CD5^+$ adult B cells are oligoclonal. *J. Clin. Invest.* 93:2093–2105.
34. Klein, U., R. Küppers, and K. Rajewsky. 1993. Human IgM^+IgD^+ B cells, the major B cell subset in the peripheral blood, express $V_κ$ genes with no or little somatic mutation throughout life. *Eur. J. Immunol.* 23:3272–3277.
35. Foster, S.F., H.-P. Brezinschek, R.I. Brezinschek, and P.E. Lipsky. 1997. Molecular mechanisms and selective influences that shape the kappa gene repertoire of IgM^+ cells. *J. Clin. Invest.* 99:1614–1627.
36. Ramsden, D.A., K. Baetz, and G.E. Wu. 1994. Conservation of sequence in recombination signal sequence spacers. *Nucleic Acids Res.* 22:1785–1796.
37. Hsieh, C.-L., and M.R. Lieber. 1992. CpG methylated minichromosomes become inaccessible for V(D)J recombination after undergoing replication. *EMBO (Eur. Mol. Biol. Organ.) J.* 11:315–325.
38. Engler, P., A. Weng, and U. Storb. 1993. Influence of CpG methylation and target spacing on V(D)J recombination in a transgenic substrate. *Mol. Cell. Biol.* 13:571–577.
39. Nagawa, F., K.-I. Ishiguro, A. Tsuboi, T. Yoshida, A. Ishikawa, T. Takemori, A.J. Otsuka, and H. Sakano. 1998. Footprint analysis of the RAG protein recombination signal sequence complex for V(D)J recombination. *Mol. Cell. Biol.* 18:655–663.
40. Feeney, A.J., M.J. Atkinson, M.J. Cowan, G. Escuro, and G. Lugo. 1996. A defective $V_κ$A2 allele in Navajos which may play a role in increased susceptibility to *Haemophilus influenzae* type b disease. *J. Clin. Invest.* 97:2277–2282.
41. Gerstein, R.M., and M.R. Lieber. 1993. Coding end sequence can markedly affect the initiation of V(D)J recombination. *Genes Dev.* 7:1459–1469.
42. Boubnov, N.V., Z.P. Wills, and D.T. Weaver. 1993. V(D)J recombination coding junction formation without DNA homology: processing of coding termini. *Mol. Cell. Biol.* 13: 6957–6968.
43. Ezekiel, U.R., P. Engler, D. Stern, and U. Storb. 1995. Asymmetric processing of coding ends and the effect of coding end nucleotide composition on V(D)J recombination. *Immunity*. 2:381–389.
44. Ezekiel, U.R., T. Sun, G. Bozek, and U. Storb. 1997. The composition of coding joints formed in V(D)J recombination is strongly affected by the nucleotide sequence of the coding ends and their relationship to the recombination signal sequences. *Mol. Cell. Biol.* 17:4191–4197.
45. VanDyk, L.F., T.W. Wise, B.B. Moore, and K. Meek. 1996. Immunoglobulin DH recombination signal sequence targeting. Effect of DH coding and flanking regions and recombination partner. *J. Immunol.* 157:4005–4015.
46. Roch, F.A., R. Hobi, M.W. Berchtold, and C.C. Kuenzle. 1997. V(D)J recombination frequency is affected by the sequence interposed between a pair of recombination signals: sequence comparison reveals a putative recombinational enhancer element. *Nucleic Acids Res.* 25:2303–2310.