The EMBO Journal vol.10 no.12 pp.3631 - 3639, 1991

Cutting and closing without recombination in V(D)J joining

Susanna M.Lewis and Joanne E.Hesse[1]

Division of Biology, California Institute of Technology, Pasadena, CA 91125, and [1]Laboratory of Molecular Biology, National Institute of Diabetes and Digestive and Kidney Diseases, National Institutes of Health, Bethesda, MD 20892, USA

Communicated by P.A.Sharp

Open and shut junctions are rare V(D)J joining products in which site-specific recognition, cleavage and re-ligation of joining signals has been uncoupled from recombination. Here, we investigate the relationship of opening and shutting to recombination in two ways. First, we have tested a series of substrates containing one or two joining signals in an *in vivo* assay. Opening and shutting can be readily observed in substrates that have only one consensus joining signal. Thus, unlike recombination, the majority of open and shut events do not require interactions between two canonical joining signals. Next we examined two-signal substrates to investigate the effect of signal proximity on the frequency of dual open and shut events. These experiments indicate that at least some of the time opening and shutting can be a two-signal transaction. Together these results point to two mechanistically related, but distinct origins for open and shut joining events. In one case, cutting and closing may occur without interaction between two signals. In the other, we suggest that interaction of a canonical signal with 'cryptic' signal-like elements whose sequence is extensively diverged from canonical signals, may bias the V(D)J recombination machinery towards opening and shutting rather than recombination. Open and shut operations could in this way provide a means whereby mistakes in target recognition by the V(D)J recombination machinery produce a non-recombinant outcome, avoiding deleterious chromosomal rearrangements in lymphoid tissues.

Key words: Ig and TCR gene assembly/site-specific recombination/V(D)J joining

Introduction

The genes that encode antigen receptors are assembled from non-contiguous DNA segments in T and B cells before they can be expressed. Although the general features of this process have been known for years (reviewed in Tonegawa, 1983), only fairly recently, with the development of more refined joining substrates, has it been possible to make systematic incursions into its more fundamental properties. The present study represents a characterization of a rare type of V(D)J joining product, of interest as it bears on the mechanics of V(D)J recombination and on how fidelity might be maintained during V(D)J joining.

Immunoglobulin and T cell receptor genes are assembled through the fusion of two or more gene segments termed V, D and J; the total number and identity of the segments involved depends on the locus in question. Adjacent to these gene segments are joining signals, each consisting of a conserved heptamer, a 12 or 23 bp variable spacer, and a nonamer motif. It has been shown that joining signals alone provide sufficient recognition elements for the V(D)J joining machinery. In one round of recombination, two segments (V and D, for example) are detached from their joining signals and then connected to one another. The joining signals themselves can become reciprocally fused (reviewed in Lewis and Gellert, 1989).

It has been found that the standard pattern of V(D)J joining described above, resulting in the formation of 'coding joints' and 'signal joints', can in fact vary (Lewis *et al.*, 1988). As shown in Figure 1A, two other outcomes exist. In one, a coding segment is detached from its own signal and then connected to the partner signal to form a 'hybrid' joint (Figure 1A, middle; Lewis *et al.*, 1988; Morzycka-Wroblewska *et al.*, 1988). In another (the focus of this paper) a coding segment is detached and then reconnected to its own signal without recombination. These are termed 'open and shut' joints (Figure 1A, bottom; Lewis *et al.*, 1988). Non-standard junction products are detected most readily with introduced recombination substrates, but a small number of endogenously derived hybrid and open and shut joints have also come to light (Roth *et al.*, 1988, 1989 and cited in Lewis *et al.*, 1988). In all types of V(D)J junction, whether standard, hybrid or open and shut, signal elements are typically joined at the edges of their heptamers, coding elements have usually lost a small and variable number of residues, and junctional inserts are observed (reviewed in Lewis and Gellert, 1989).

A significant difference between opening and shutting and recombination is that in the former, there is no *a priori* requirement that two sites must come together. Indeed in previous work with a plasmid substrate, open and shut junctions were found at one signal or the other more often than at both, raising the possibility that opening and shutting may not be a two-signal transaction (Lewis *et al.*, 1988). With respect to the mechanism of V(D)J joining, the ability to carry out virtually all necessary operations (i.e. those that accomplish site-specific recognition, cleavage at the borders of the signal elements, base subtraction/addition, and ligation of ends) on one signal alone would strongly imply that the synapsis of two joining signals is not central to the joining reaction.

Here we have investigated this issue in two ways. We measured open and shut junctions in substrates lacking a consensus partner, as well as in a substrate in which both 12- and 23-signals are present but are prevented from interacting. We found that open and shut junctions can indeed be generated at solitary joining signals but that, when two joining signals are proximate, no doubly opened and shut molecules are recovered. Thus the possibility that opening and shutting is exclusively a one-signal operation is ruled

out. Further data indicate that interactions between a joining signal and a cryptic signal-like sequence may underlie some of the putatively single open and shut events. An implication of these observations is that opening and shutting could be one means whereby the fidelity of V(D)J joining is maintained during the differentiation of T and B lymphocytes.

Results

Detection of open and shut junctions

Our *in vivo* assay for V(D)J joining makes use of extrachromosomally replicated shuttle plasmids containing consensus joining signals, as has been previously described in detail (Hesse *et al.*, 1987, 1989; Lieber *et al.*, 1987, 1988). Briefly, substrates are introduced into virally transformed murine pre-B cells active for V(D)J recombination. Joining events are scored after the plasmid DNA has been recovered from the eukaryotic cells and used to transform *Escherichia coli*. Site-specific rearrangement activates a chloramphenicol acetyl transferase gene included in the substrate. Recombinants (whether they contain standard or hybrid junctions) are doubly resistant to chloramphenicol and ampicillin while non-recombinant molecules confer resistant to ampicillin only.

Here, in order to facilitate isolation of both recombinant and non-recombinant joining products, the substrates (slightly modified from those described in Hesse *et al.*, 1989) contain unique *Sal*I and *Bam*HI restriction sites at positions that overlap the borders of the 12- and 23-signals respectively (see Figure 1B). After transfection of a rearranging cell line, 1-8, the recovered plasmid DNA can be enriched for those molecules that contain site-specific junctions by digestion with *Sal*I or *Bam*HI. V(D)J joining products are then identified using oligonucleotide probes in a filter hybridization assay of transformant *E.coli* colonies (grown on ampicillin-containing plates; i.e. without selection for recombinants). Non-hybridization to either of two 'recombination site' probes specific for the input sequences identifies molecules of interest (Lewis *et al.*, 1988). These 15 base oligonucleotides span the junction sites at the borders of the 12- or 23-signals (Figure 2). All negative colonies, among which are those containing V(D)J joining products, are further categorized according to junction class on the basis of (i) whether or not they grow on chloramphenicol-containing plates and (ii) positive or negative hybridization to a collection of oligonucleotides shown in Figure 2. Details are provided in Materials and methods.

Opening and shutting at solitary signals

Previously we isolated a number of examples in which an open and shut joint was found at only one of two joining signals in a substrate (Lewis *et al.*, 1988). This observation indicated that either open and shut junctions were created in the absence of any interaction between the two partner signals or, despite an interaction, the sequences at the border of only one joining signal became changed. In the latter instance, one could imagine that singly opened and shut

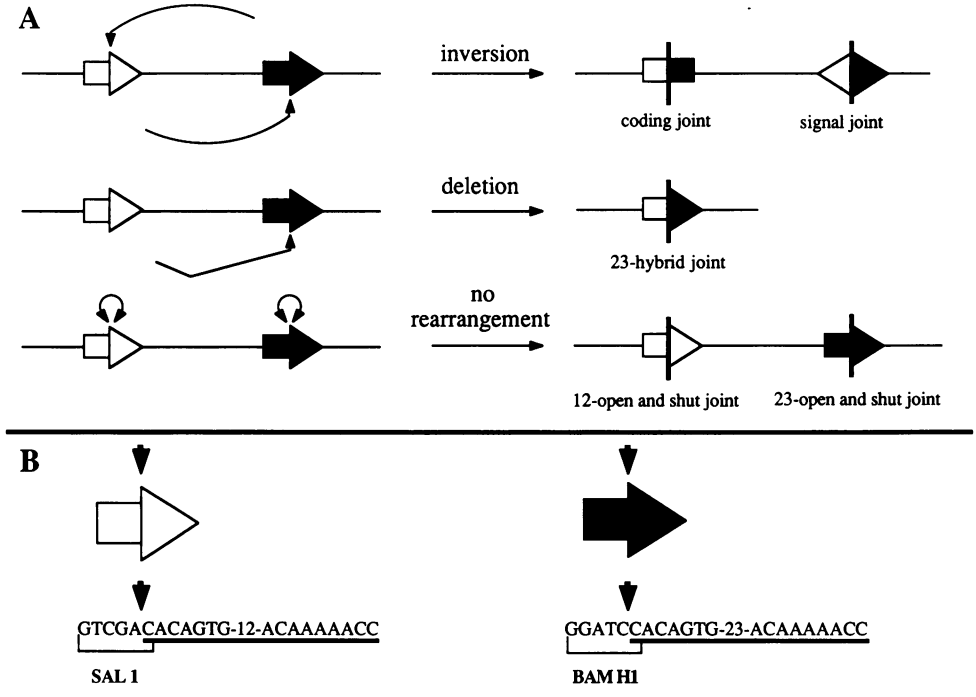


Fig. 1. (**A**) Three outcomes of V(D)J joining. Shaded triangles represent 23-spacer signals, open triangles represent 12-spacer signals, while shaded and open boxes represent the abutting sequences. Recombination sites are indicated by arrowheads and the positions of product junctions are indicated by heavy vertical bars. Top: standard recombination generates two products: coding joints in which coding ends are joined, and signal joints in which signal ends are fused. (The term coding end is used for simplicity throughout the text; in the present substrates, coding elements such as V, D or J gene segments are replaced by restriction sites.) Middle: hybrid recombination in the form of a site-specific deletion joins the 23-signal with the coding end that was originally abutting the 12-signal. Bottom: open and shut junctions are generated in a non-recombinant outcome. (**B**) Sequences at the borders between coding ends and heptamers in p12×23 and its derivatives. The six-base recognition sites for *Sal*I and *Bam*HI are in brackets, the sequences of the signal elements are underlined. Note the overlap; the last base of the restriction site is also the first base of the heptamer. Arrowheads indicate the positions at which the input sequences are interrupted when any of the three types of junctions depicted in panel A (above) are formed.

isolates would arise if, after being brought together, one of the two signals either re-closed without base loss or addition or, alternatively, were to be incorporated into a synaptic complex without ever becoming cleaved.

To investigate the requirement for both signals in opening and shutting, products arising from substrates containing either one or two consensus joining signals were compared (Figure 3). The first (designated p12×23) contains one 12-signal and one 23-signal. This arrangement is consistent with the '12/23 rule', a requirement that one signal have a 12 base spacer and the other a 23 base spacer for efficient V(D)J recombination (Tonegawa *et al.*, 1983; Hesse *et al.*, 1989). The signals in p12×23 are oriented in such a way

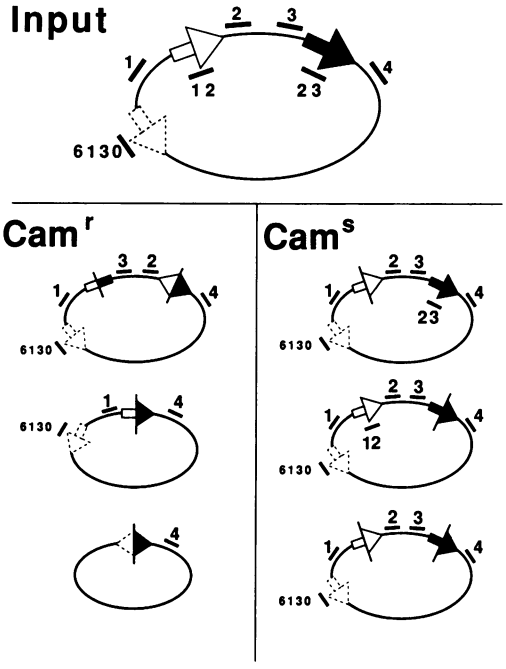


Fig. 2. The structure of p12×23 and V(D)J joining products. Shaded and open triangles and boxes represent joining signals and the flanking DNA as described in the legend to Figure 1. Dashed lines indicate the cryptic site 6130 (see text). (**A**) The substrate (not to scale) showing the locations of the oligonucleotides used in the hybridization assay. Unrearranged p23 and p23×23 plasmids (not shown) hybridize to all of the oligonucleotides indicated with the exception of that specific for the 12-signal; p12 (not shown) hybridizes to all except those specific for the 23-signal and for site 6130. The sequences of the three recombination-site oligonucleotides are as follows: '12' indicates 12-SIG (AGGTCGACACAGTGG); '23' indicates 23-SIG (GAGGATCCACAGTGA); and '6130' designates 6130-SIG (TATGTTGTGTGGAATTGT). Flanking oligonucleotides are indicated as follows: 1 (LAC-1), described in Lewis *et al.* (1988); 2 (TER-1), described in Lewis *et al.* (1988); 3 (TER-2) lies within λ phage-derived terminator sequences on the constructs (Hesse *et al.*, 1987), corresponding to bases 38502–38522; and 4 (JH33, sequence: AGCTCCTGAAAATCTCGCCA). (**B**) Chloramphenicol-resistant recombinants. Plasmids containing standard, hybrid and site 6130 junctions can be discriminated from non-recombinant products because they confer chloramphenicol resistance on bacterial hosts. This is a feature of the original constructs from which the present substrates were derived (Hesse *et al.*, 1987, 1989). Each recombinant junction gives a different pattern of hybridization to test oligonucleotides as shown. (**C**) Chloramphenicol-sensitive plasmids. This category includes molecules singly opened and shut at the 12- or the 23-signal, and dual open and shut products. Not shown are examples of recombinants involving sites other than site 6130, or non-specifically deleted products. Open and shut products are distinguished from various aberrantly rearranged products because they retain hybridization to all oligonucleotides except 12-SIG, and/or 23-SIG.

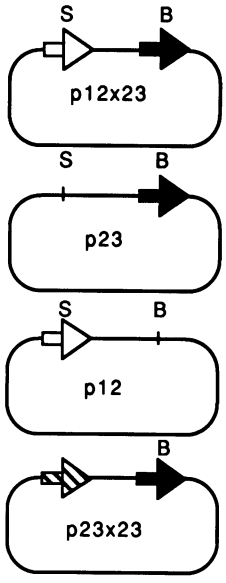
that standard recombination is an inversional rearrangement (forming both a coding joint and a signal joint), and hybrid junction formation is a deletional event (resulting in the formation of '23-hybrids') (Figures 1A and 2). Opening and shutting, as is always the case, occurs without any gross reorganization of the input substrate. A second plasmid lacks the 12-signal (p23), but is otherwise identical, a third lacks the 23-signal (p12), and the last contains two 23-signals (p23×23). These four constructs were individually introduced into the A-MuLV transformed pre-B cell line, 1-8, as described previously (Hesse *et al.*, 1987; Lieber *et al.*, 1987) and analyzed for joining events.

Each line of Figure 3 gives the number of products detected in equal aliquots of transfected plasmid DNA after recovery from 1-8 cells. With p12×23, we found twelve 23-signal open and shut junctions and five 12-signal open and shut junctions (lines 1 and 2, column 5). As was noted previously, only some of the open and shut product molecules had opened and shut at both the 12- and 23-signals (numbers in parentheses, Figure 3); here, about one-third of the isolates were doubly opened and shut.

When the other three substrates were similarly tested for joining events, opening and shutting was found to occur at solitary signals as well as at the 'mis-paired' signals of p23×23, and these junctions were recovered at frequencies comparable with that found with p12×23 (Figure 3).

Sequence analysis of the open and shut junctions confirmed that they had the expected characteristic structure (Lewis *et al.*, 1988); signals were usually intact, coding ends were 'nibbled', and junctional inserts were prevalent (Figure 4). Two of the doubly opened and shut isolates recovered in the *Bam*HI-digested sample were again recovered in the *Sal*I-digested aliquot verifying that our methods were sensitive enough to detect a large fraction of the very low frequency open and shut junctions present in a given transfection. Further, the occurrence of 'P nucleotides' was evident in this collection. As described for coding joints by McCormack *et al.* (1989) and Lafaille *et al.* (1989), one or two non-random bases can sometimes be found in junctions where no residues have been lost from one or both joined coding segments. The P nucleotide insert abutting a non-nibbled end appears as the inverse complement of that segment. In the present case, candidate P nucleotides are G or GA residues adjacent to any intact 23-coding ends, and likewise T or TC abutting intact 12-coding ends. Such inserts occur within open and shut junctions at a frequency above that expected by chance (Figure 4). The presence of P nucleotides further affirms the relationship of non-recombinant open and shut junctions to recombinant V(D)J joining products.

Although open and shut junctions are rare products and the total numbers isolated were small, the conclusion that the numerical values obtained with each substrate are indeed comparable is supported by various observations (Figure 3, columns 1, 6 and 7). First, replication of the substrates (indicating the number of DNA molecules that reached the nucleus and thus the transfection efficiency) was quantified by digesting the recovered plasmid DNAs with *Dpn*I prior to *E.coli* transformation (Lieber *et al.*, 1987). Figure 3, column 1 shows that similar numbers of replicated molecules were present in equal aliquots; no transfection had worked markedly better or worse than another. Second (in all cases save p12), we made use of the cryptic site (designated site 6130) as an internal control. Site 6130 is a sequence within the plasmid backbone that is situated 357 bases leftward of



1 REPLICATED MOLECULES	2 DIGEST	3 STANDARD (INVERSION)	4 HYBRID (DELETION)	5 OPEN AND SHUT	6 SITE 6130 (DELETION)	7 % RECOVERY
3.2×10^4	Bam H1	823	133	12(3)	153	76
3.2×10^4	Sal 1	600	101	5(3)	136	57
3.0×10^4	Bam H1	NA	NA	10	40	30
3.0×10^4	Sal 1	NA	NA	8	NA	NA
2.6×10^4	Bam H1	12	1	20(2)	103	83

Fig. 3. Quantification of V(D)J junctions isolated from the p12×23 series of substrate. The left side of the figure shows a cartoon of each of the four plasmids. Open and shaded triangles and boxes represent 12- and 23-signals and their adjacent DNA sequences as described in the legend to Figure 1. The crosshatched element in p23×23 represents a 23-signal that has been inserted at the *SalI* site, as described in Materials and methods. Values given for each substrate were determined for 1/5 of the sample as harvested after transfection into 1-8 cells. Column 1: number of colonies arising from (eukaryotically) replicated plasmids in each sample, as determined by small scale transformations of *DpnI*-digested aliquots. Column 2: restriction enzyme used (in conjunction with *DpnI*) in each case to enrich for junction products prior to transformation. Columns 3, 4, 5 and 6: numbers of each type of junction product, determined as described in Materials and methods. Column 7: an estimate of the efficiency with which junctions were detected in each sample determined by measuring the ratio of chloramphenicol-resistant colonies scored on the chloramphenicol replicas to the number arising after transformation of an aliquot of undigested DNA. Non-specific losses occurring during the restriction enzyme digestion and repurification of DNA cause the ratio to be less than one. NA, not applicable. (Because it does not have a 23-signal, p12 cannot form standard, hybrid or site 6130 junctions and thus does not yield significant numbers of chloramphenicol-resistant colonies.)

the 23-signal. It functions as a 12-signal, forming signal joints (by deletion) upon recombination with the 23-signal at a measurable frequency (Figure 3, column 6). By comparison with a canonical 12-signal, its sequence (CACAACA-12-GCATAAAGT) has a four out of seven match in its heptamer and a five of nine match in its nonamer. Site 6130 recombinants were identified as described in Materials and methods, and served to establish that substrates encountered the recombination machinery within the transfected cells. Finally, the numbers of chloramphenicol-resistant colonies detected after digestion, transformation and screening procedures were counted and compared with the numbers expected from an undigested sample (see legend to Figure 3). As given in column 7, this calculation gives an indication of the overall efficiency with which junction products were recovered for each substrate (again excepting p12, for which such a control is not applicable).

Based on these observations, it is clear that opening and shutting at a given joining signal occurs undiminished in the absence of its appropriate 'partner' signal. Thus the notion that undetected opening and shutting occurs at one of two signals in a two-signal substrate does not adequately account for the origin of singly opened and shut molecules.

Open and shut interactions involving non-canonical partners

While interactions between canonical joining signals do not appear to underlie the majority of open and shut events, it remained possible that perhaps some opening and shutting may involve interactions between a joining signal and a *non*-canonical partner. Two preliminary experiments were performed to investigate this possibility. As mentioned above, the cryptic site 6130 functions as a 12-signal in recombination with the consensus 23-signal, becoming incorporated into both standard signal joints (Figure 3) and hybrid junctions (data not shown). To test the possibility that some of the putative single open and shut isolates had in fact also opened and shut at a non-consensus site, all 47 of the 23-signal open and shut molecules that were isolated in the present study were screened for non-hybridization of the 6130-SIG oligonucleotide (Figure 2). Two of these were found to have also opened and shut at site 6130 (confirmed by sequence analysis; data not shown). The fact that these few open and shut isolates contained a hidden junction at site 6130, suggests that other apparently single open and shut molecules may likewise have opened and shut at a second, unidentified site.

Another way to examine the effect of non-canonical interactions upon opening and shutting is to measure the ratio of dual opening and shutting to recombination in a substrate in which signal−signal interactions occur that violate the 12/23 rule. A 'conventional' substrate, p12×23, will yield standard, hybrid and double open and shut joining products at relative percentages of ~86, 14 and 0.4 respectively (Figure 3, lines 1 and/or 2). With p23×23, the three types of products (recombinants, hybrid junctions and doubly open and shut joints) occur at relative percents of 80, 6 and 13. Both standard recombination and hybrid junction formation

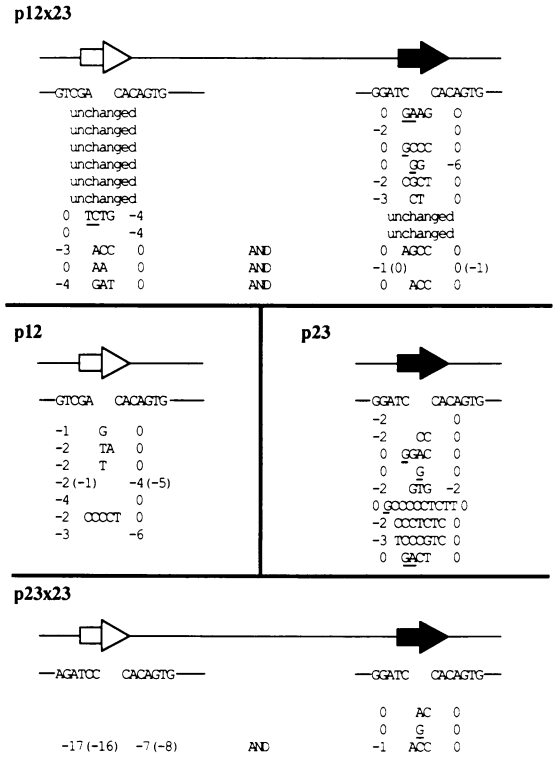


Fig. 4. Sequences of open and shut junctions. Beneath each cartoon are the sequences of the flanking restriction site (the 'coding end') and the heptamer portion of the signal. The sequence is gapped to show the crossover site at the heptamer border. Each of these non-recombinant junctions is represented by two numbers, indicating the number of residues absent on either side of the crossover site. Letters indicate junctional inserts. In cases where sequence redundancy prevents an unambiguous assignment, the alternative is indicated in parentheses. P nucleotides (see text) are underlined. The number of junctions listed here is smaller than the numbers given in Figure 2 because some junctions were isolated more than once, and in the case of p23×23, not all of the open and shut junctions were sequenced.

between two 'like' signals is reduced (to 2% or lower) relative to the levels observed for a 12- and 23-signal pair, indicating that these processes both follow the 12/23 rule (Figure 3). In contrast, dual open and shut junctions were found in numbers that were not significantly different from those obtained using the p12×23 construct (Figure 3, bottom line, number in parentheses). The relative frequency of doubly opened and shut products among the joining products increased >30-fold in a substrate containing mis-matched signals.

These data provide circumstantial evidence that some opening and shutting may actually be a two-signal operation. Key to this interpretation is the assumption that double opening and shutting (whether between two consensus joining signals or between a signal and a cryptic site) is indicative of a signal−signal interaction. To investigate this point, we determined whether or not double open and shut junctions could still be detected in a substrate where there is a barrier to signal−signal interaction.

Absence of dual open and shut events in a substrate where interaction of signals is hindered

According to experiments of Shore *et al.* (1981), Shore and Baldwin (1983) and Horowitz and Wang (1984), increasing energy is required to twist and/or bend shorter lengths of

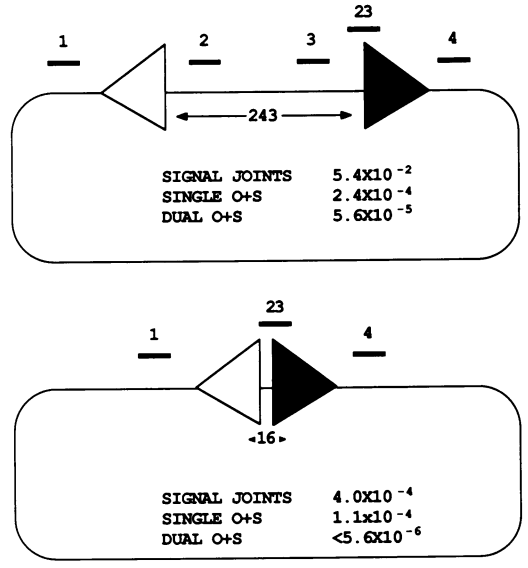


Fig. 5. The effect of signal proximity on double opening and shutting. pD243 (top) and pD16 (bottom) are shown, along with the oligonucleotide probes to which they hybridize. 1, 2, 3 and 4 are as described in the legend to Figure 2. Open triangles represent 12-signals, shaded represent 23-signals. (Boxes representing flanking DNA sequences have been omitted for simplicity.) '23' is the recombination site oligonucleotide D23-SIG (GATGGATCCACAG-TG). Frequencies of standard recombinants (in this case, signal joints), of single 12-signal open and shut joints, and of double open and shut isolates are given for each substrate. Actual numbers detected are given in the text.

covalently closed, duplex DNA, an effect that becomes especially pronounced below 0.5 kb. Thus the closer two sites are located to one another, the greater the energetic barrier to their interaction.

For comparison, we measured the dual open and shut frequency in two substrates. Both contain the consensus 12- and 23-signals, however, in the first, the two signals are separated by 243 bases (pD243); in the second, by only 16 (pD16). In these substrates, the signals were oriented deletionally (Figure 5) so that standard recombination would result in signal joint formation.

Each construct was transfected in duplicate into the 1-8 cell line. Aliquots of harvested DNA were enriched for junction products by restriction digestion and then analyzed by colony filter hybridization as described in Materials and methods. The frequency of each class of junction among the total number of replicated molecules screened is presented in Figure 5. Both single and double open and shut junctions are readily detected in the pD243 substrate (eight single open and shut events were detected in 3.4×10^4 replicated molecules and four dual open and shut products were found upon screening 7×10^4 replicated molecules). In contrast, even though 13 signal open and shut examples were discovered among $\sim 1.5 \times 10^5$ molecules (Figure 5) a search through nearly 2×10^5 pD16 products failed to yield any doubly open and shut isolates. Thus dual open and shut events are at least 10-fold reduced in pD16 relative to pD243. Recombination was measured for both substrates as well, and was found to be reduced 130-fold in pD16 (Figure 5).

Close proximity of two joining signals thus has a negative effect not only upon recombination but also upon dual opening and shutting. This result provides evidence that dual

opening and shut junctions originate in a process that involves some form of signal–signal interaction.

Discussion

Although open and shut junctions constitute a minor fraction of all V(D)J joining products (~1% by our assay), they highlight potentially important features of the joining mechanism. Open and shut junctions are formed by cutting and closing target DNA in a site-specific fashion. Because recombination does not result from this type of operation, one cannot presuppose that opening and shutting involves a signal–signal interaction. This issue of whether or not a second site is necessary is a key to understanding how the normal joining reaction might work. Further, as discussed below, the nature of these cutting and closing events bears on how fidelity is maintained in the system.

This study has shown the following: first, in the majority of cases, opening and shutting does not require the interaction of two canonical joining signals; second, double open and shut events are reduced at least 10-fold (to below measurable levels) when sites are in close proximity; third, at least some apparently singly open and shut molecules have in fact opened and shut at a second, cryptic site as well; and fourthly, when a signal–signal interaction is suboptimal for recombination the dual open and shut fraction of joining products can increase.

These findings can be understood in terms of a model in which open and shut junctions are formed in two distinct but related ways. Single open and shut junctions may result from events in which single site recognition, cleavage, and re-ligation of the DNA ends occurs in the absence of any interaction with a second site. Such a situation would strongly suggest that in the normal V(D)J joining reaction, cleavage precedes, rather than follows, synapsis. In a somewhat different fashion, double open and shut junctions most probably arise after synapsis between two cleaved signal sequences. In this second case, reclosure of cut ends without strand exchange represents a failed attempt at recombination.

In the remainder of this discussion, our results will be considered in the context of this model.

Opening and shutting in one-signal constructs

Opening and shutting is not abrogated when either one of two signals is removed from a recombination substrate (compare p12×23, p23 and p12, Figure 3). This result shows that two consensus signals are not required in order to form an open and shut joint. One of two possibilities must therefore hold: either the reaction can occur at a single signal in the absence of synapsis with another signal, or, if synapsis is necessary, it occurs between the consensus signal and a cryptic, non-consensus signal present elsewhere in the substrate. We have identified one cryptic site (site 6130), the sequence of which is only a 54% match to a canonical joining signal, that appears to have been involved in a small number of open and shut events. Whether any of our other 'single' open and shut isolates have arisen through interactions with as yet undetected, cryptic signals remains an open question, and is under investigation.

Our detection of open and shut events in one-signal substrates contrasts with the results obtained by Hendrickson *et al.* (1991), who failed to find open and shut junctions in an integrated substrate containing a solitary joining signal. The reasons for this discrepancy are difficult to ascertain with certainty because a quantitative basis for comparing our two studies is lacking. Possibly, their assay is not as sensitive as ours. A number of less likely explanations will be considered briefly.

One trivial possibility is that the opening and shutting we detect with one-signal substrates is an artifact caused by sequences elsewhere in the plasmid that have perfect or near-perfect homology to a canonical 12- or 23-signal. Because the sequence of our substrates is known, we could search directly for such matches, and none was found. Partial homologies (the best one having an overall identity of 75%) were found in numbers not significantly different from their predicted occurrence in DNA of random sequence (see Materials and methods). Similar numbers of such sites would be expected to exist in the Hendrickson substrates, which are of comparable sizes to those used here. A second possibility is that the cryptic site 6130 (used as an internal control in some of these experiments), is the direct cause of the opening and shutting observed in one-signal substrates. However, we have indication of such involvement in only two cases (see Results). Moreover, open and shut junctions are readily detected after transfection of the substrates p12, pD16 and pD243; site 6130 is not present in any of these plasmids (see Materials and methods). In sum, as discussed in a later section (below), although we have been able to identify a subset of open and shut junctions that appear to arise from signal–signal interactions, the majority of open and shut junctions are not immediately attributable to such transactions.

We conclude that opening and shutting can occur in one-signal substrates. A small number of the observed events can be accounted for on the basis of an interaction with a cryptic, non-consensus signal. Either the apposition/interaction of two signals is not an obligate intermediate step in opening and shutting, or such interactions take place between a consensus signal and a second non-standard partner.

Opening and shutting as a result of signal–signal interactions

An interaction between canonical 12- and 23-signals is not the major pathway by which open and shut junctions are generated (as evidenced by the results summarized in Figure 3); however, certain aspects of our data are at odds with a conclusion that opening and shutting is exclusively a one-signal operation. In particular, the frequency of p12×23 isolates in which both signals opened and shut was far higher than would be anticipated if they were to have arisen in two distinct (uncoupled/independent) operations. In the case of two unrelated open and shut events one might expect to encounter double-junction molecules at a frequency which is the product of the two single open and shut frequencies. Instead, the double open and shut frequency in p12×23 was some 1500-fold higher (9×10^{-5} versus 6×10^{-8}; calculated from the values given in lines 1 and 2, columns 1 and 5, Figure 3).

To investigate further, we compared double opening and shutting in a substrate containing two proximate sites with one in which the two signals were more distantly spaced. The relative position of two joining signals should have little effect on the double open and shut frequency if no signal–signal interaction need occur: if anything, a processive machinery might generate more double junctions when two

signals are very close together. In contrast, doubly opened and shut isolates should decrease in the signals-proximate substrate if joining signals must interact. This is because whether signals are brought together before strand scission takes place or while cut ends are constrained by the joining apparatus, the inherent inflexibility of short segments of duplex DNA (Shore *et al.*, 1981; Shore and Baldwin, 1983; Horowitz and Wang, 1984) would impede their interaction. Closely positioned joining signals might also suffer mutual steric occlusion if, in order to open and shut both sites, the joining machinery must bind both simultaneously.

We have found that proximate signals do not open and shut together as readily as do more distantly spaced sites. No doubly opened and shut molecules were found with pD16 (Figure 5). In these experiments, double open and shut events occur at frequencies near to the detection limits of our assay, and we cannot therefore measure the absolute magnitude of the effect; however, the reduction is at least 10-fold. This result is inconsistent with the view that opening and shutting, when observed at both signals in a two-signal substrate, involves two fundamentally unlinked events. Instead, a signal−signal transaction is indicated.

The mechanism of V(D)J joining

An ordered synaptic structure that incorporates components of the recombination machinery along with two uncleaved recombination signals is formed early in many site-specific recombination systems (reviewed in Craig, 1988). The possibility that some opening and shutting is a one-signal operation suggests, in contrast, that for V(D)J recombination, cleavage at one or both signals may occur prior to the point at which the sites are brought together for strand exchange. Clearly, the present data indicate that the V(D)J recombination machinery can target a joining signal, cut at the recombination site, carry out base addition and subtraction and rejoin the cut ends [in short, can recreate every aspect of V(D)J recombination except actual strand exchange] in the absence of a consensus partner signal. It may well be the case that no second signal (or substitute thereof) is involved in many open and shut reactions. If so, synapsis cannot be an important first step in the reaction mechanism. A second type of evidence, from our proximate-signal experiment, supports this notion as well. The fact that recombination, although reduced, is not eliminated in a substrate with signals only 16 bases apart (Figure 5) implies that the DNA is broken before sites are brought together.

Biological implications

With regard to substrate specificity, the V(D)J recombination machinery must accomplish two things: it must be able to recombine authentic joining signals, and it must be able to avoid rearrangement of fortuitous signal-like sequences that occur at random elsewhere in the genome. Consensus heptamer and nonamer sequences appear to be the best targets for V(D)J joining: they will recombine at higher frequencies than any variant tested to date (Hesse *et al.*, 1989). The (authentic) physiological substrates for joining, however, most often do not exactly match the canonical sequence (Hesse *et al.*, 1989 and cited therein). Thus in order effectively to assemble antigen receptor genes, the V(D)J joining apparatus cannot be highly stringent. Such non-stringent targeting has attendant risks; for example, there is evidence that the molecular lesion underlying some T-cell leukemias is a rearrangement caused by the V(D)J joining machinery as it operates on inappropriate sites lying outside antigen receptor loci (Aplan *et al.*, 1990; Brown *et al.*, 1990).

Fidelity in V(D)J joining is likely to be maintained in several ways. At one level, incorrect targets may be screened out either by poor recognition/binding by the joining machinery, or perhaps by limiting access of the machinery to certain regions of the chromosome. At another level, incorrect sequences even if recognized, may not be efficiently cut. Opening and shutting could provide a third and last opportunity to prevent mistakes in joining. Inappropriate interactions between a 'real' joining signal and a cryptic site could be aborted at a late stage by rejoining cut and modified ends in a non-recombinant (open and shut) configuration. Mistakes that have proceeded as far as cleavage may be rare, but it is precisely such events that have the potential to cause aberrant chromosomal rearrangements. The ability of the V(D)J joining machinery to correct mistakes after cleavage and without gross rearrangement, may thus be critical.

The notion of a late-stage checkpoint in the V(D)J joining reaction is supported by our results with p23×23. In the suboptimal type of interaction represented by the pairing of two 23-signals, the fraction of open and shut junctions among the three possible outcomes in V(D)J joining goes up 50-fold. Taking this observation a step further, the existence of sequences that will promote non-recombinant over recombinant outcomes is predicted if opening and shutting plays a significant role in maintaining fidelity in the joining process. Such sequences, if they can be demonstrated, might provide insight into the nature of the events that generate some translocations, in addition to establishing the corrective function of open and shut operations. The efficency with which mistakes in recognition/cleavage are or are not circumvented by this mechanism may be a key to understanding the origins of certain lymphoid cancers.

Materials and methods

Plasmids

All substrates used in the present work (Figures 3 and 5) were derived from either pJH288 or pJH298 (Lieber *et al.*, 1988), and were identical to one another except as noted. p12×23 was constructed from pJH288 by stepwise removal, and replacement of the *Bam*HI fragment and *Sal*I fragment, each of which contains a joining signal. A new synthetic 23-spacer signal (CACAGTGATTCATATCACTGCGCCCCCGTTACAAAAACC) was inserted at the *Bam*HI site. Both of its ends were *Bam*HI compatible (not shown), but the sequence at the nonamer end was such that it would not reconstruct a *Bam*HI recognition site upon insertion. Similarly, a new synthetic 12-spacer signal with *Sal*I compatible ends (CACAGTGGTA-CAGACTGGAACAAAAACC) was introduced at the *Sal*I site, reconstructing a *Sal*I site at the heptamer end only. Both inserts contain consensus heptamer and nonamer sequences (Hesse *et al.*, 1989). p23 was derived in the same fashion except that no 12-signal was inserted at the *Sal*I site. p12 was derived from pJH298 by removal of its 23-signal without replacement, and the new synthetic 12-signal (above) was inserted at the *Sal*I site. [The pJH298 backbone is altered relative to pJH288 by 2 bp. The change inactivates cryptic site 6130 for recombination (Lieber *et al.*, 1988 and unpublished data) but is inconsequential with regard to p12, where the consensus 23-signal is absent.] For p23×23, the original *Bam*HI fragment in pJH288 was removed, attached to *Sal*I linkers and used to replace the original *Sal*I fragment, thus positioning a 23-signal in the usual 12-signal location. Next, the new synthetic 23-signal (above) was inserted at the *Bam*HI site. The two 23-signals in this plasmid are non-identical in their spacer and flanking sequences, and the sole *Bam*HI site is located at the heptamer of the signal at the original *Bam*HI site. The use of non-identical 23-signals was necessary to eliminate homologous recombination between the two

23-signals, the results of which would be indistinguishable from site-specifically derived hybrid joints lacking inserts or base loss.

pD16 was constructed from pJH298 by replacement of the 12- and 23-signals with the above signals, oriented in such a way as to form signal joints (by deletion) upon recombination. A synthetic linker connected the *Sal*I site at the edge of the 12-signal to the *Bam*HI site at the edge of the 23-signal, with an intervening *Cla*I site such that the heptamer-to-heptamer distance was a total of 16 bp. pD16 confers chloramphenicol resistance whether or not it is rearranged, due to the absence of the transcription terminator located between its joining signals. pD243 was derived by introduction of the *Cla*I fragment containing the transcription terminator from pJH298 into the corresponding site of pD16.

Cell culture and recombination assay

The cell line used throughout this study was the Abelson murine leukemia virus transformant 204-1-8 (referred to here as 1-8), which was grown in RPMI, 10% fetal bovine serum and 1×10^{-5} M β-mercaptoethanol. This cell line has been used extensively with the plasmid assay by us (Lieber *et al.*, 1987; Lewis *et al.*, 1988; Hesse *et al.*, 1989). p12×23, p12, p23 and p23×23 constructs were transfected into the same culture of 1-8 cells on the same day, and were subsequently processed in parallel. pD243 and pD16 were transfected in duplicate in a separate experiment, again in parallel. Substrates were introduced by a modified DEAE−dextran transfection procedure employing osmotic shock. Plasmid DNAs were recovered from the transfected cells 48 h later by alkaline lysis and co-digested with either *Dpn*I and *Bam*HI, *Dpn*I and *Sal*I or (in the case of some pD16 and pD243 samples) with all three. Digested DNAs were phenol extracted, chloroform extracted twice, and then re-precipitated before being used to transform *E.coli*. Transformed cells were selected on plates containing 100 μg/ml ampicillin.

Measurement of standard, hybrid and open and shut junctions

All junctions could be uniquely identified using some or all of the criteria shown in Figure 2. Colony transfers to nitrocellulose and hybridizations were carried out as described previously (Lewis *et al.*, 1988). Oligonucleotide probes are described in the legend to Figure 2. The exact procedure used to isolate product junctions depended upon the construct and the preliminary digestion as outlined below. These methods of identification of standard, hybrid and site 6130 recombinants, were verified by diagnostic restriction digests of DNA isolated from a sampling of colonies (>80 total). Where carried out, standard double-stranded Sanger DNA sequencing techniques were employed, with oligonucleotides 1 or 4 (Figure 2) serving as primers.

Although a battery of tests is used to ultimately identify open and shut junctions, an important safeguard against the elimination of rare products through technical irregularities is provided by the fact that the first three steps in their isolation involve negative criteria (i.e. lack of a restriction site, non-hybridization of recombination site oligonucleotides, and non-growth on chloramphenicol). Because any colony that for technical reasons did not transfer well onto filters or failed to replicate onto chloramphenicol was provisionally designated an open and shut candidate and selected for further analysis, a certain number were later found to belong to one of the recombinant junction classes. For all except the p12×23 samples (where such junction classes were calculated as described), these numbers were added to the totals.

p12×23 (BamHI and DpnI-digested sample)

Transformants were plated at a density pre-determined to give several hundred or fewer colonies per plate. Colonies were transferred to nitrocellulose and filter lifts were prepared as previously described (Lewis *et al.*, 1988). After allowing for regrowth, colonies were replica plated onto plates containing 40 μg/ml chloramphenicol. Filters were probed with the recombination site oligonucleotide, 23-SIG (Figure 2). All non-hybridizing, chloramphenicol-resistant colonies were scored as recombinant, and further sorted into categories as described below. All non-hybridizing, chloramphenicol-sensitive colonies were picked onto ampicillin plates in grid arrays. Filter lifts were prepared and again probed with 23-SIG. Duplicates were also probed with oligonucleotides 3 and 4 (which hybridize to either side of the 23-signal) in order to screen out those molecules that were 23-SIG(−) due to gross deletions unrelated to opening and shutting. DNA was prepared from all colonies that were confirmed to be negative for 23-SIG hybridization but positive for hybridization to the other two probes. Plasmids were then restricted with *Hgi*AI to determine which had the same gross structure as the input DNA, tested with *Bam*HI to confirm the presence of an open and shut junction at the 23-signal, and finally digested with *Sal*I to determine whether or not they had also opened and shut at the 12-signal. All identified junctions were then sequenced.

The 1107 chloramphenicol-resistant colonies recovered in the sample contained plasmids bearing hybrid, standard or site 6130 deletion junctions. Of these colonies 400 were picked to grid arrays on chloramphenicol-containing plates. Filter lifts were prepared and hybridization to the end-labeled oligonucleotides 1 and 3 was tested. Standard junctions were positive with both probes, hybrid junctions were positive with probe 1 only, and site 6130 deletions were positive for neither (Figure 2). The total numbers of standard recombinants, hybrid junctions and site 6130 deletions reported in Figure 2 were calculated based on their occurrence in these 400.

p12×23 (SalI and DpnI digested)

The analysis of the *Sal*I-digested sample was as described above except that hybridizations to oligonucleotide 12-SIG (Figure 2) were performed, and open and shut junctions were identified using the flanking oligonucleotides 1 and 2.

Of a total of 836 chloramphenicol-resistant colonies 240 were picked to grid arrays and tested with the 6130-SIG oligonucleotide and probes 1 and 2. Standard inversions hybridized to all three, hybrid junctions to 6130-SIG and oligonucleotide 1, and site 6130 deletions were negative with all three.

p23 (BamHI and DpnI digested)

Open and shut junctions were identified as described for the *Bam*HI-digested p12×23 sample (above). Filter lifts from the chloramphenicol replicas were tested for hybridization to oligonucleotide 3 and negative colonies were designated site 6130 deletions.

p23×23 (BamHI and DpnI digested)

Open and shut junctions were identified as for the *Bam*HI-digested p12×23 sample, except that only about one-third of the isolates were sequenced. Filter lifts from chloramphenicol replicas were probed with oligonucleotide 3. DNA was prepared from all positive (or uncertain) colonies and standard inversions were identified by restriction analysis. The colonies that were negative for hybridization (site 6130 deletions or, potentially, hybrid junctions) were picked to grid arrays and were retested with probe 3 and also with probe 1. The single hybrid junction recovered with p23×23 was sequenced.

Dual open and shut junctions involving the substituted 23-signal in p23×23 were scored by testing for non-hybridization to the end-labeled oligonucleotide JH23-SIG described previously (Lewis *et al.*, 1988). Southern blot transfers of the entire collection of single open and shut isolates were hybridized, and the open and shut isolates thus identified were sequenced.

p12 (SalI and DpnI digested)

Ampicillin-resistant transformants were tested for hybridization to 12-SIG. DNA was prepared from all negative colonies and open and shut isolates were identified through diagnostic *Hig*AI, *Sal*I and *Bam*HI digestions.

pD243 (BamHI and DpnI digested or BamHI, SalI and DpnI digested)

Open and shut junctions were identified by first probing with oligonucleotide 2 (Figure 2). All positive colonies were picked to grid arrays and probed with oligonucleotides 2 and 4 in addition to the recombination site oligonucleotide specific for the pD243 and pD16 constructs (D23-SIG, see legend to Figure 5). DNA was prepared from all D23-SIG-negative colonies that were found to be positive with probes 2 and 4, and open and shut isolates identified through diagnostic restriction analysis. In some experiments, hybridization to probe 1 was also tested.

Colonies negative with oligonucleotide probe 2 were scored as standard recombinants (pilot studies confirmed the accuracy of this designation).

pD16 (BamHI and DpnI digested or BamHI, SalI and DpnI digested)

Open and shut junctions and standard joint deletions were identified by first probing with D23-SIG, and then picking all negative colonies to grid arrays. Colonies were re-probed with D23-SIG and also with probes 1 and 4. DNA was prepared from all colonies that remained negative with D23-SIG but were positive with the other oligonucleotides. Diagnostic digests identified a subset of the standard recombinants among the isolates that, lacking junctional inserts or deletions at the signal joint, generated a new *Hgi*AI site. The remaining isolates were sequenced.

Homology search

The sequence of pJH299 was searched for homology to 12- and 23-spacer signals, in either orientation using the DNANALYZE collection of programs (Scwindinger and Warner) on a MacII personal computer (software ported to the Apple Macintosh by Dr P.Markiewicz, Department of Biology, University of California, Los Angeles, and kindly provided by him). All sites with >62% homology to CACAGTG (N12 or N23) ACAAAAACC (in either orientation) were listed. The expected number of sites in 8 kb

of random DNA was calculated using the binomial distribution. For example, the number of sites matching 11 of the 16 consensus residues (taking both orientations as well as two-spacer lengths into account) was expected to be nine (15 were found), while a match of 12 out of 16 was expected once (two were found). In addition, we searched for elements with 100% homology to the more conserved and functionally significant portions of the joining signals (Hesse *et al.*, 1989). 119 sites were found to contain the four most critical residues (124 sites expected for random sequence). Sequences with CACA within the heptamer and AA at positions 6 and 7 of the nonamer were also listed (again, both orientations were tested). Eight such sites would be expected to occur by chance in 8 kb of random sequence, and 17 were found.

Acknowledgements

This work was initiated in the laboratory of Martin Gellert (NIH), whose criticisms and stimulating discussions are gratefully acknowledged. S.L. would like to thank H.D.Lipshitz (California Institute of Technology) for support in the final stages of the study and for advice in evaluating the homology search. We thank David Brown (NIH) for preparing a file of the sequence of pJH299. This research was funded in part by research grant #IM-599 from the American Cancer Society (S.L.) and an award from the Searle Scholars Program of the Chicago Community Trust to Howard D.Lipshitz. We thank Joanne Topol and Howard Lipshitz for their comments on the manuscript.

References

Aplan,P.D., Lombardi,D.P., Ginsberg,A.M., Cossman,J., Bertness,V.L. and Kirsch,I.R. (1990) *Science*, **250**, 1426–1429.

Brown,L., Cheng,J.-t., Chen,Q., Siciliano,M.J., Crist,W., Buchanan,G. and Baer,R. (1990) *EMBO J.*, **9**, 3343–3351.

Craig,N.L. (1988) *Annu. Rev. Genet.*, **22**, 77–105.

Hendrickson,E.A., Liu,V. and Weaver,D.T. (1991) *Mol. Cell. Biol.*, **11**, 3155–3162.

Hesse,J.E., Lieber,M.R., Gellert,M. and Mizuuchi,K. (1987) *Cell*, **49**, 775–783.

Hesse,J.E., Lieber,M.R., Mizuuchi,K. and Gellert,M. (1989) *Genes Dev.*, **3**, 1053–1061.

Horowitz,D.S. and Wang,J.C. (1984) *J. Mol. Biol.*, **173**, 75–91.

Lafaille,J.J., DeCloux,A., Bonneville,M., Takagaki,Y. and Tonegawa,S. (1989) *Cell*, **59**, 859–870.

Lewis,S.M., Hesse,J.E., Mizuuchi,K. and Gellert,M. (1988) *Cell*, **55**, 1099–1107.

Lewis,S. and Gellert,M. (1989) *Cell*, **59**, 585–588.

Lieber,M.R., Hesse,J.E., Mizuuchi,K. and Gellert,M. (1987) *Genes Dev.*, **1**, 751–761.

Lieber,M.R., Hesse,J.E., Lewis,S., Bosma,G.C., Rosenberg,N., Mizuuchi,K., Bosma,M.J. and Gellert,M. (1988) *Cell*, **55**, 7–16.

McCormack,W.T., Tjoelker,L.W., Carlson,L.M., Petryniak,B., Barth,C.F., Humphries,E.H. and Thompson,C.B. (1989) *Cell*, **56**, 785–791.

Morzycka-Wroblewska,E., Lee,F.E.H. and Desiderio,S. (1988) *Science*, **242**, 261–263.

Roth,M.E., Lacy,M.J., Klis McNiel,L. and Kranz,D.M. (1988) *Science*, **241**, 1354–1358.

Roth,M.E., Lacy,M.J., Klis McNiel,L. and Kranz,D.M. (1989) *Science*, **245**, 1032.

Shore,D. and Baldwin,R.L. (1983) *J. Mol. Biol.*, **170**, 957–981.

Shore,D., Langowski,J. and Baldwin,R.L. (1981) *Proc. Natl. Acad. Sci. USA*, **78**, 4833–4837.

Tonegawa,S. (1983) *Nature*, **302**, 575–581.

Received on July 19, 1991; revised on September 3, 1991