91. R. T. Maguire, T. S. Robins, S. S. Thorgersson, C. A. Heilman, *ibid.* **80**, 1947 (1983).
92. P. H. Hamlyn and T. H. Rabbitts, *Nature (London)* **304**, 135 (1983).
93. R. Taub *et al.*, *Cell* **36**, 339 (1984).
94. K. Kelly, B. H. Cochran, C. D. Stiles, P. Leder, *ibid.* **35**, 603 (1983).
95. C. B. Thompson, P. B. Challoner, P. E. Neiman, M. Groudine, *Nature (London)* **314**, 363 (1985); S. R. Hann, C. B. Thompson, R. N. Eisenman, *ibid.*, p. 366.
96. T. H. Rabbitts, P. H. Hamlyn, R. Baer, *ibid.* **306**, 760 (1983).
97. T. H. Rabbitts, A. Forster, P. Hamlyn, R. Baer, *ibid.* **309**, 592 (1984).
98. L. W. Stanton, P. D. Fahrlander, P. M. Tesser, K. B. Marcu, *ibid.* **310**, 423 (1984).
99. C. Gazin *et al.*, *EMBO J.* **3**, 383 (1984).
100. N. G. Copeland and G. M. Cooper, *J. Virol.* **33**, 1199 (1980).
101. J. A. Lautenberger, R. A. Schulz, C. F. Garon, P. H. Tsichlis, T. S. Papas, *Proc. Natl. Acad. Sci. U.S.A.* **78**, 1518 (1981).
102. K. Quade, *Virology* **98**, 461 (1979).
103. G. Goubin, D. S. Goldman, J. Luce, P. E. Neiman, G. M. Cooper, *Nature (London)* **302**, 114 (1983).
104. H. Rubin, *ibid.* **309**, 518 (1984).
105. J. Erikson, J. Finan, Y. Tsujimoto, P. C. Nowell, C. Croce, *Proc. Natl. Acad. Sci. U.S.A.* **81**, 4144 (1984).
106. J. J. Yunis, M. D. Oken, M. E. Kaplan, K. M. Ensurd, R. R. Howe, A. Theologides, *New Engl. J. Med.* **307**, 1231 (1982).
107. P. J. Fialkow and J. W. Singer, in *Leukemia*, I. L. Weissman, Ed., Dahlem Konferenzen Berlin (Springer Verlag, Berlin, 1985), p. 203; P. J. Fialkow *et al.*, *Blood* **57**, 1068 (1981).
108. J. C. Der, T. G. Krontiris, G. M. Cooper, *Proc. Natl. Acad. Sci. U.S.A.* **79**, 3637 (1982).
109. R. W. Ellis, D. R. Lowy, E. M. Scolnick, *Advances in Viral Oncology*, G. Klein, Ed. (Raven, New York, 1982), vol. 1, p. 107.
110. D. J. Capon, E. Y. Chen, A. D. Levinson, P. H. Seeburg, D. V. Goeddel, *Nature (London)* **302**, 33 (1983).
111. T. Finkel, J. D. Channing, G. M. Cooper, *Cell* **37**, 151 (1984).
112. E. P. Reddy, R. K. Reynolds, E. Santos, M. Barbacid, *Nature (London)* **300**, 149 (1982).
113. A. P. Feinberg, B. Vogelstein, M. J. Droller, S. B. Baylin, B. D. Nelkin, *Science* **220**, 1175 (1983).
114. E. Santos *et al.*, *ibid.* **223**, 661 (1984).
115. R. Muschel and G. Khoury, personal communication.
116. R. Sager, K. Tanaka, C. C. Lau, Y. Ebina, A. Anisowicz, *Proc. Natl. Acad. Sci. U.S.A.* **80**, 7601 (1983).
117. J. J. Harvey and J. East, *Int. Rev. Exp. Pathol.* **10**, 265 (1971).
118. J. A. Levy, *J. Nat. Cancer Inst.* **46**, 1001 (1973).
119. S. A. Aaronson and G. I. Todaro, *Nature (London)* **225**, 458 (1970).
120. S. A. Aaronson and C. A. Weaver, *J. Gen. Virol.* **13**, 245 (1971).
121. V. Klement, M. Friedman, R. McAllister, W. Nelson-Rees, R. J. Huebner, *J. Natl. Cancer Inst.* **47**, 65 (1971).
122. L. M. Pfeffer and L. Kopelvich, *Cell* **10**, 313 (1977).
123. J. A. Levy, *Nature (London)* **253**, 140 (1975).
124. Y. Yuasa *et al.*, *ibid.* **303**, 775 (1983).
125. J. Fujita *et al.*, *ibid.* **309**, 464.
126. A. Balmain, M. Ramsden, G. T. Bowden, J. Smith, *ibid.* **307**, 658 (1984).
127. M. Barbacid, personal communication.
128. S. Sukumar, V. Notario, D. Martin-Zanca, M. Barbacid, *Nature (London)* **306**, 658 (1983).
129. E. Farber, *Cancer Res.* **44**, 5463 (1984).
130. R. Holliday, *Nature (London)* **306**, 742 (1983).
131. J. Fogh, Ed., *Human Tumor Cells in Vitro* (Plenum, New York, 1975); S. E. Salmon, *Cloning of Human Tumor Stem Cells* (Liss, New York, 1980).
132. M. Wigler *et al.*, in *Cancer Cells: Oncogenes and Viral Genes*, G. F. Vande Woude, A. J. Levine, W. C. Topp, J. D. Watson, Eds. (Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y., 1984), pp. 419–423.
133. D. J. Capon *et al.*, *Nature (London)* **304**, 507 (1983).
134. S. Pulciani *et al.*, *ibid.* **300**, 539 (1982).
135. A. P. Albino, R. Le Strange, A. I. Oliff, M. E. Furth, L. J. Old, *ibid.* **308**, 69 (1984).
136. K. M. Vousden and C. J. Marshall, *EMBO J.* **3**, 913 (1984).
137. M. A. Tainsky, C. S. Cooper, B. C. Giovanella, G. F. Vande Woude, *Science* **225**, 643 (1984).
138. M . Ruta, personal communication.
139. J. Maisel, V. Klement, M. M-C. Lai, W. Ostertag, P. H. Duesberg, *Proc. Natl. Acad. Sci. U.S.A.* **70**, 3536 (1973).
140. R. W. Ellis, D. Defeo, M. E. Furth, E. M. Scolnick, *Mol. Cell. Biol.* **2**, 1339 (1982).
141. L. F. Parada, C. Tabin, C. Shih, R. A. Weinberg, *Nature (London)* **297**, 474 (1982).
142. E. M. Scolnick, W. C. Vass, R. S. Howk, P. H. Duesberg, *J. Virol.* **29**, 964 (1979).
143. H. E. Varmus, *Annu. Rev. Gen.* **18**, 553 (1984).
144. R. Dahr *et al.*, *Science* **217**, 934 (1982).
145. M. Lusky and P. H. Duesberg, unpublished.
146. D. A. Spandidos and N. M. Wilkie, *Nature (London)* **310**, 469 (1984).
147. C. J. Tabin and R. A. Weinberg, *J. Virol 53*, 260 (1985).
148. H. M. Temin, *Nature (London)* **302**, 656 (1983).
149. T. A. Stewart, P. K. Pattengale, P. Leder, *Cell* **38**, 627 (1984).
150. I specifically thank my collaborators N. Kan, M. Nunn, T. Papas, P. H. Seeburg, and D. Watson and also my colleagues M. Botchan, M. Carey, G. S. Martin, S. Pfaff, W. Phares, C. Romerdahl, and H. Rubin for encouragement and many critical comments and L. Brownstein for typing numerous drafts of this manuscript. The work from my laboratory is supported by NIH grant CA 11426 from the National Cancer Institute and by grant CTR 1547 from the Council for Tobacco Research—U.S.A., Inc.

RESEARCH ARTICLE

DNA Elements Are Asymmetrically Joined During the Site-specific Recombination of Kappa Immunoglobulin Genes

Susanna Lewis, Ann Gifford, David Baltimore

During B-cell differentiation, a series of site-specific recombinations assemble the variable (V) region exons of immunoglobulin genes (*1–3*). Recombinational assortment of gene segments expands the coding capacity of a locus, enabling a large repertoire of gene products to be generated from a relatively small number of components. This appears to be a common feature of the immune system loci that encode antigen-binding proteins; the variable region exons of T-cell–specific antigen receptor genes also assemble recombinationally (*4, 5*) as do those in another locus that is also thought to be involved in immune recognition (*6*).

The gene segments targeted for rearrangement at these loci are flanked by characteristic DNA sequences. The putative "joining signal" consists of a heptamer, a spacer region, and a nonamer (*1–5*). The heptamer and nonamer elements are evolutionarily conserved, being similar in different vertebrate classes (*7*), as well as in genes that rearrange in different cell lineages (*1–5*). Joining signals always have one of two forms; the heptamer and nonamer sequences are separated by an approximately 12-base spacer, or by an approximately 23-base spacer. Gene segments linked to joining signals with 12-base spacers appear to recombine only with those linked to joining signals containing 23-base spacers and vice versa (*1–5*). The presence of joining signals and the adherence to the 12–23 spacer rule have been taken to indicate that similar, perhaps identical, enzymes catalyze recombination at all six loci that are known to rearrange during immunodifferentiation (*4*).

The kappa immunoglobulin locus provides a simple and well-characterized system in which to study the details of the recombination process. Only two component parts, V_κ and J_κ, recombine in forming a complete kappa variable region exon. Rearrangement of the kappa locus generates two distinct classes of recombinant junctions (*8–13*). One product of rearrangement is a "coding joint," which is the junction between V_κ and J_κ coding sequences in an assembled variable region exon. The other product we refer to as a "reciprocal joint": it consists of the two joining signals derived from V_κ and J_κ fused to one another at their (formerly) coding-proximal borders (*8*) (Table 1). The existence of reciprocal joints at the kappa locus indicates that gene rearrangement may be a reciprocal process, yet reciprocal joints apparently

The authors are on the staff of the Whitehead Institute for Biomedical Research, 9 Cambridge Center, Cambridge, Massachusetts 02142. Susanna Lewis is a graduate studet and David Baltimore is a professor in the Department of Biology, Massachusetts Institute of Technology, Cambridge 02139. The present address for Susanna Lewis is Department of Biochemistry, Stanford University School of Medicine, Stanford, California 94305.

Abstract. *Immunoglobulin K genes are constructed during lymphocyte differentiation by the joining of two DNA elements, V_K and J_K, to form both a V_KJ_K coding unit and a reciprocal recombination product. The two products formed in single V_K-to-J_K joining events can be directly isolated through the use of a retrovirally introduced recombination substrate. The structural analysis of a number of recombinants and the derivation of secondary recombination products define some of the basic features of the mechanism of immunoglobulin gene assembly.*

do not accumulate at the lambda or heavy chain gene loci with the same frequency with which they are found at kappa; only one example has been reported (*14*). The reason the kappa locus should be unique in this regard is debated, but it is very likely that the presence or absence of reciprocal joints at a given locus is primarily a result of its topography (*10, 15*).

To characterize possible intermediates in the joining reaction, it is of benefit to examine the fine structure of recombinant junctions. The sequences of these junctions is informative because the component gene segments are joined in a region devoid of homology (*16*), enabling the crossover sites to be precisely defined. We thus can identify the sites at which bonds in the parental duplexes must have been broken before ligation created the recombinant linkage. The new connections define some of the basic parameters of the rearrangement process. In taking this approach, we have found it particularly useful to examine a coding joint pairwise with the corresponding reciprocal joint because then all the products of a single recombination event are represented.

Much of what is known about the rearranged kappa locus has come from the study of transformed B-lineage cell lines such as myeloma lines. Nevertheless, reciprocal pairs of recombinant junctions—a coding joint along with its reciprocal—have never been demonstrated in the DNA of rearranged myelomas and hybridomas (*8, 11–13*). It has been suggested that nonreciprocal junctions in myelomas may have been produced as a result of multiple recombination events (*9, 10*) as is discussed below.

To isolate reciprocal recombination products, and to facilitate experimental manipulation of the system, we designed a recombination substrate that could be introduced into an actively rearranging cell (*15*). The construct could be infected as a retrovirus into the Abelson murine leukemia virus transformant PD. The substrate DNA, in a site-specifically recombined form, can be recovered from the DNA of PD cells as proviral clones in bacteriophage λ (*15*). Because the V_κ and J_κ sequences in the substrate were arranged so that a coding joint would be formed via an inversion, the coding joint and a reciprocal recombination product are both present at the inversion junctions of a single proviral insert. We used this approach to isolate and analyze the products of numerous individual rearrangement events. It is clear from this analysis of the products of gene rearrangement that they are qualitatively quite different from those of any other site-specific recombination system, and we expect that this reflects a novel mechanism.

Isolation and analysis of recombinants. The substrate that we used to detect V_κ-to-J_κ recombination was based on a retrovirus vector that had been engineered to contain a murine germline V_κ gene segment and the five germline J_κ segments. These gene segments are each linked to their flanking joining signals. The LTR-to-LTR region (LTR, long terminal repeat) of the substrate is represented in Fig. 1. [The pBR322-based construct, pVJG, and the preparation of the viral form of the substrate are described in (*15*)]. Reciprocal joining of V_κ to a J_κ element should cause an inversion of the region between them. This design was intended to allow the recovery of both a coding joint and a reciprocal joint at either border of the inversion. In addition, pVJG contained a selectable marker, the *Escherichia coli gpt* gene (*17*). The *gpt* gene is not linked to a eukaryotic promoter and is transcriptionally inactive in the substrate, unless an inversion such as that resulting from site-specific V_κ-to-J_κ recombination reorients the *gpt* gene so that it is under the transcriptional control of the 5′ viral LTR (Fig. 1). Mycophenolic acid may be used to select for cells that express the bacterial gene (*17*).

Initially we isolated 19 mycophenolic acid–resistant PD cell lines that had been infected with the substrate, most of which were independent. We generated a restriction map of the integrants in 14 of these lines by probing restriction enzyme–cleaved samples of their DNA for

Table 1. Terms used.

Term	Definition	Abbreviation	Symbol
Germline V_κ	Unrearranged V_κ	V_κ	
Germline J_κ	Unrearranged J_κ	J_κ	
V_κ coding sequences		V	
J_κ coding sequences		J	
V_κ joining signal	The heptamer, 12-base spacer, nonamer element flanking a germline V_κ	H_v	
J_κ joining signal	The heptamer 23-base spacer nonamer of germline J_κ	H_1*	
Coding joint	Fusion of V_κ and J_κ coding sequences	V/J_1*	▽
Reciprocal joint	Fusion of V_κ and J_κ joining signals	H_v/H_1*	▼
Pseudo-coding joint	Fusion of J_κ joining signal to J_κ coding sequences	H_1/J_2*	▽
PD-A (-D,-I, and other)	Infected, mycophenolic acid–resistant isolate of cell line PD		
Recombinant A (D and other)	Recombinant form of provirus after cloning out of PD-A DNA into phage λ		

*These are examples.

gpt gene sequences after electrophoresis through agarose gels and transfer to nitrocellulose. We chose one cell line, PD-A, for sequencing studies because our analysis indicated that the substrate in this cell line had acquired a single internal inversion consistent with a V_κ-to-$J_{\kappa 1}$ recombination. We isolated a λ phage clone of the integrant in PD-A and demonstrated that the DNA sequences of the inversion junctions in the recombinant (designated A in Figs. 1 and 2) do in fact represent a V_κ–$J_{\kappa 1}$ coding joint and its corresponding H_V–H_1 reciprocal joint (Table 1) (*15*).

All of the other PD isolates that were mapped by hybridizing variously digested samples of cellular DNA to *gpt* probes as described above, harbored a recombinant substrate, except for one which neither expressed the bacterial gene by direct enzyme assay nor showed evidence for any introduced *gpt* gene sequences in its DNA. All recombined substrate sequences appeared to have been site-specifically rearranged. No evidence for activation of the *gpt* gene by any means other than site-specific V_κ-to-J_κ recombination was detected, although a sizable fraction of the PD isolates contained recombined substrates that appeared to have undergone more complex rearrangements than the inversion observed in PD-A. To examine additional examples of V_κ-to-$J_{\kappa 1}$ recombination, we selected isolates that appeared identical to PD-A. Because we sought to investigate the variants that were detected by our initial mapping study as well, we also selected several isolates that were recombined to another J_κ segment and several examples of the more complex alterations.

A genomic library was prepared from each of eight mycophenolic acid–resistant PD derivatives included in our survey and screened (*15*). Recombinant substrate sequences were recovered as Sac I inserts in λgtwes · λB (Fig. 1).

Properties of coding joints. The recombinant junctions in variable region exons are not precisely specified (*1–3*); the same V gene segment or J element can recombine at a variety of sites. This property of the joining reaction has not been systematically investigated because it is difficult to identify the precursors of a given endogenous rearrangement, or to isolate more than a few examples of the same VJ (or VDJ) segments in fused form without imposing a selection for a particular protein product. In the present system, we can examine junctional diversity where the precursors are known, and, most important, without selection for expression of kappa protein. To this end, we isolated five independent $V_\kappa/J_{\kappa 1}$ recombinants (A, I, O, N, and X) and three V_K/J_{K4} recombinants (D, M, and P). The sequences of these junctions are shown in Fig. 2.

One consequence of junctional diversity was immediately apparent. Only three of the eight coding joints (those in recombinants O, P, and X) occurred in the same translational reading as a functional kappa immunoglobulin protein (Fig. 3). This ratio was close to the value expected if the recombination process fused V_κ and J_κ elements randomly with respect to the coding properties of the product (*18, 19*). The frequency with which rearrangement can generate functional coding joints is an important factor toward understanding how the rearrangement process might be regulated (*20*). While our data are limited, they represent an essentially unbiased measurement of the coding fidelity of V_κ-to-J_κ gene rearrangement and fully confirm earlier proposals (*18, 19*).

The variability of V/J coding joints can be examined more closely by comparing recombination sites. Among the eight

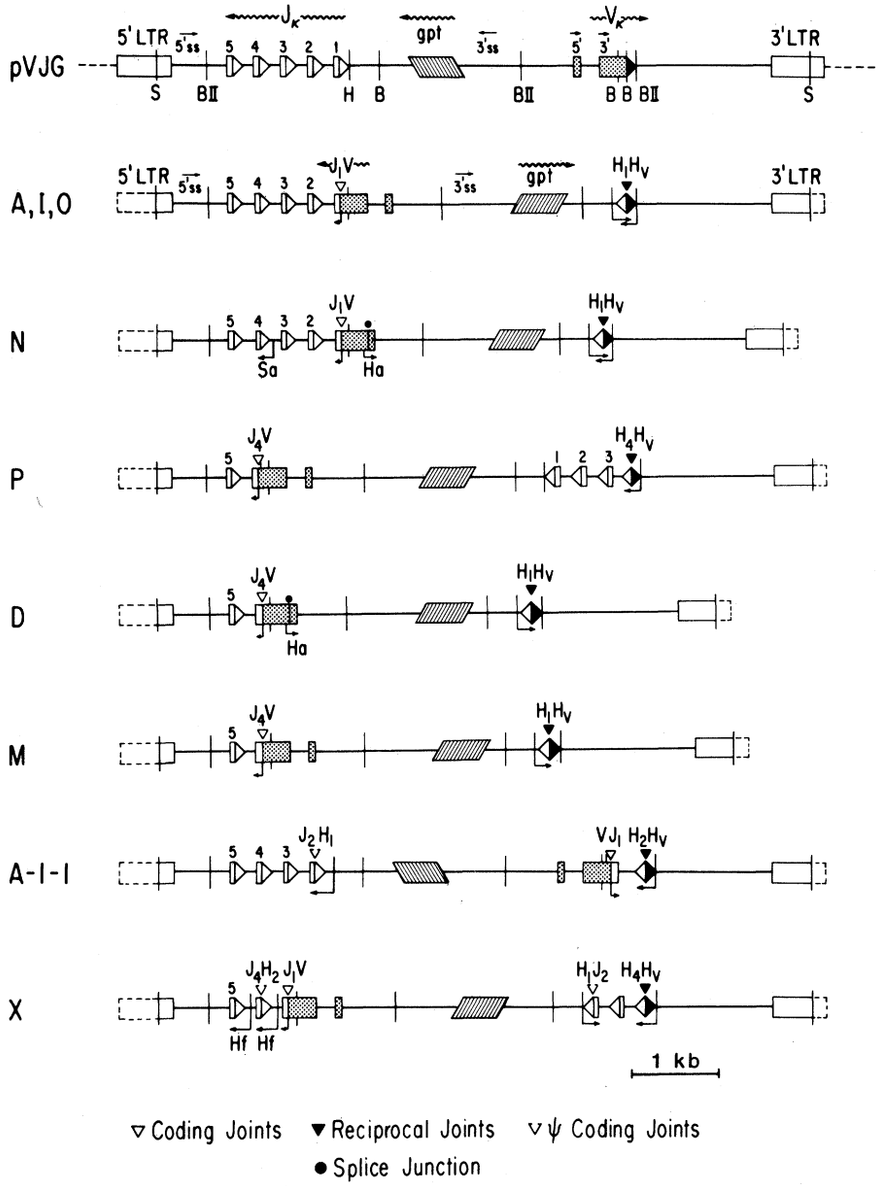


Fig. 1. Structures of VJG recombinants. Structures of the region from Sac I to Sac I of each recombinant after cloning into phage λ are diagramed. Each recombinant junction is designated by one of the symbols listed in the key (Table 1). The sequenced regions are shown below each diagram. Restriction endonuclease sites are abbreviated as follows: S, Sac I; BII, Bgl II; H, Hind III; B, Bam HI; Ha, Hae III; Hf, Hinf I; and Sa, Sau 3AI. Recombinants A, N, O, and P were isolated from PD31-infected derivatives as described (*15*), recombinants D and I were from PD40 derivatives, and recombinants M and X from PD34-28-1 derivatives. The phenotypes and gene structures of the PD subclones, with the exception of PD34-28-1, have been described (*10*). PD34-28-1 is a derivative of PD34 (*10*) that has deleted both endogenous kappa constant region genes.

coding joints presented in Fig. 2, no two are alike. Recombination sites can be unambiguously assigned in most cases. Where there is fortuitous redundancy in the DNA sequence at the junction between the target elements, as in recombinants I and M, precise assignment is not possible. The sites of crossing-over in the coding joints are indicated by open triangles in Fig. 2. In Fig. 3 (open triangles) these data are compiled relative to unrearranged V_κ and J_κ elements.

Recombination sites can be seen to cover a region on the V_κ element ranging from one to three bases away from the heptamer border. On the J_κ element they are located either directly at the heptamer border or up to four bases in toward the coding sequences. As suggested by the frequency and distribution of sites shown in Fig. 3, the positioning of a coding joint crossover on either component element (V_κ or J_κ) appears to be random within a small region. We searched for other possible patterns that would not be evident in this compilation. It was hoped that such patterns might give some clue as to the geometry of the protein-DNA interactions during joining. Specifically, a pairwise comparison of the displacements of the crossover sites on V_κ and J_κ as measured from either their heptamer borders or from other landmarks such as the last codon, revealed no positive or negative correlation. In addition, a variable number of base pairs is missing relative to the precursor sequences from recombinant junctions (Fig. 2). It would appear from this analysis that if there is any restriction on the location of a given V_κ crossover site with respect to a given J_κ crossover site, there is no hint of it in this collection. Any pattern that exists must be fairly subtle, and we conclude from this that there is a component of the rearrangement mechanism that can generate variability in an essentially random fashion.

Junctional insertion is found in the coding joint of one recombinant. The junction in recombinant D shows an unexpected feature. It contains two bases at the recombinant joint that could not be derived from either of the two precursor gene segments. We have confirmed that these two bases are not in the sequence of the V_κ and $J_{\kappa 4}$ elements in the substrate (data not shown). Therefore, barring a mutation in the substrate at some time after infection but before rearrangement (a remote possibility that cannot be excluded), it is apparent that the A and T residues (A, adenine; T, thymine) at the junction were acquired during the recombination process itself.

The appearance of extra bases in recombinant D raises the issue of whether they may represent an N region. N regions consist of short stretches of nucleotides of fairly random base composition appearing as apparent insertions at V_H/D and D/J_H junctions of heavy chain genes (*21–23*) and in assembled β-chain variable regions of the T-cell receptor (*24, 25*). N regions have not been reported to be present in joined light chain genes [although one junction is ambiguous in this regard (*19*)]. The formation of N regions has been correlated with the presence of terminal deoxynucleotidyl transferase (*23*).

Unfortunately, there is no way to know on the basis of a single example if the extra base pairs observed in recombinant D should be considered an N region or whether they have an unrelated origin. Short insertions, similar in length to N regions, are sometimes observed at the junction sites of transfected DNA (*26–28*), and have been demonstrated at the breakpoints of *myc* gene translocations as well (*29*). Because of this, no obvious criteria distinguish N regions from other types of insertions. The junctional insert in recombinant D does not seem to have been caused by terminal transferase, because enzymatic assays of cellular extracts (*23*) from PD-A, -X and -D showed no measurable activity above background in any of these lines (*30*). Whatever their cause, the two-base insertion in the junction of recombinant D demonstrates that the ends that form kappa gene coding joints are accessible to modification.

Properties of reciprocal joints. Five of the recombinants analyzed (A, I, O, N, and P) contain junctions that reciprocally correspond to their coding joints. In striking contrast to the variability of the co-isolated V/J coding joints, the H_v/H_1 (Table 1) reciprocal joints found in the recombinants A, I, O, N and the H_v/H_4

Coding Joints		Reciprocal Joints		# Δ b.p.
A -GGATCCGGACGTT-	VJ_1	-CTAC[CACTGTG]CACAGTG]CTCC-	H_1H_v	5
I -GGATCCTGGACGTT-	VJ_1	-CTAC[CACTGTG]CACAGTG]CTCC-	H_1H_v	3
N -GGATCCTCGTGGACGTT-	VJ_1	-CTAC[CACTGTG]CACAGTG]CTCC-	H_1H_v	1
O -GGATCCTTGGACGTT-	VJ_1	-CTAC[CACTGTG]CACAGTG]CTCC-	H_1H_v	4
P -GGATCCTACGTTCGG-	VJ_4	-GAAT[CACTGTG]CACAGTG]CTCC-	H_4H_v	6
Coding Joints		**"Reciprocal" Joints**		
D -GGATCCaTATTCACGTT-	VJ_4	-CTAC[CACTGTG]CACAGTG]CTCC-	H_1H_v	
M -GGATCCTCACGTT-	VJ_4	-CTAC[CACTGTG]CACAGTG]CTCC-	H_1H_v	
Coding Joints		**"Reciprocal" Joints**		**ψ Coding Joints**		# Δ b.p.
A-H -GGATCCGGACGTT-	VJ_1	-ACAC[CAGTGTG]CACAGTG]CTCC-	H_2H_v	-CTAC[CACTGT]CACGTT-	H_1J_2	6
X -GGATCCGTGGACGTT-	VJ_1			-CTAC[CACTGT]ACACGTT-	H_1J_2	
		-GAAT[CACTGT]CACAGTG]CTCC-	H_4H_v	-ACAC[CA]TTCACGTT-	H_2J_4	7(?)

Fig. 2. Sequences of recombinant junctions. Recombinant junctions were sequenced by the method of Maxam and Gilbert (*53*). The junctions are symbolized as in Fig. 1. Heptameric elements of the joining signals are boxed; the spacer regions and nonamer elements of the joining signals are not included in the sequence shown. Sequencing gels were read for 20 bases and more on either side of the junctions in each case, except for the coding joints, which were sequenced by labeling at a Bam HI site in the V_κ coding sequences eight bases 5′ to the joining signal. Ambiguous crossover sites due to redundancy between the recombining sequences are indicated by multiple triangles. The translational reading frame of the V and J gene segments is indicated by the wavy lines in Fig. 3. The lowercase letters in the coding joint of recombinant D represent a 2-bp insertion. The total number of base pairs missing from each reciprocal pair of junctions is shown in the right-hand columns. The number of base pairs lost on formation of the reciprocal junctions in recombinant X is based upon the scheme in Fig. 5, but alternative derivations would give a different value. The sequences of recombinant A have been reported previously (*15*).

junction in P are identical types of fusions. In every case, the joining signals located at the 3′ flank of V_κ and at the 5′ flank of J_κ were linked to one another exactly at the borders of their heptamers proximal to the coding regions (Fig. 2).

The invariant structure of reciprocal joints was first noted in endogenously arising recombinants (*8, 11, 13*). However, in these examples, all were $J_{\kappa 1}$-derived and all were nonreciprocal to the coding joints in the same cell. The possibility therefore arose that these endogenous $J_{\kappa 1}$-derived reciprocal joints were not directly formed in functional V_κ-to-J_κ joining events but represented by-products of a related reaction [(*10*) and see be low]. The recombinants discussed above, which include repeated isolations of reciprocal V_κ-to-$J_{\kappa 1}$ joining products as well as reciprocal products of a V_κ-to-$J_{\kappa 4}$ rearrangement, definitively demonstrate that precise recombination sites are a general feature of reciprocal joints.

A consequence of the qualitative difference between reciprocal joints and coding joints is that the two products of a reciprocal V_κ-to-J_κ joining event are not exact reciprocals of one another at the nucleotide level. Because the coding joint crossover sites that we have analyzed here are always displaced to the coding side of the joining signals on one or both elements, a variable loss of bases has resulted from every recombination event. In our collection, from one to six bases that existed in the precursor sequences are missing from the products (Fig. 2). Whether the bases disappeared during formation of the coding joint, the reciprocal joint, or both cannot be determined. Nevertheless, the microscopically nonconservative nature of immunoglobulin gene rearrangement is a provocative feature. Analogy can be made to the integration of many eukaryotic and prokaryotic transposons (*31–33*), retroviruses (*34*) and dispersed, processed pseudogenes (*35*), as well as other types of interspersed, repetitive DNA elements (*36, 37*), all of which have the common feature that they apparently insert into new sites through recombination events that are grossly reciprocal but nonconservative in detail. In these cases, a small number of bases are commonly duplicated in the process of integration. Terminal repeats of sequences at the insertion site of transposable elements have been suggested to result from the introduction of staggered breaks on the target duplex (*32*). Although our work is an instance where base pairs are lost, not added, it is conceivable that the nonconservative aspect of V-to-J recombination similarly reflects the way in which the participating duplexes are cut and reconnected (see Fig. 7).

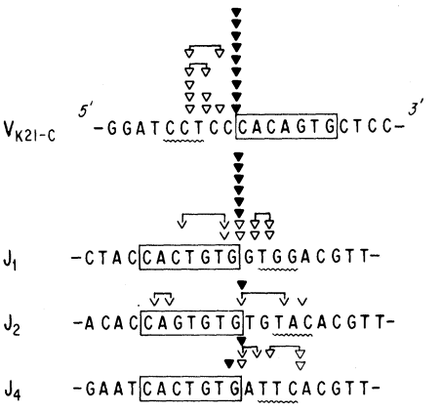


Fig. 3. Distribution and frequency of crossover sites. The DNA sequence of unrearranged V_κ21-C, $J_{\kappa 1}$, and $J_{\kappa 4}$ elements in VJG were determined directly [(*15*) and this article]. The unrearranged $J_{\kappa 2}$ sequence is from (*54*). Symbols above the crossover sites are as in Fig. 1. Ambiguous junction sites are delimited by a bar with triangles at the extreme right and extreme left. Codon 95 of V_κ and codon 96 of each J_κ element (*55*) are underlined.

Substrate sequences may frequently undergo secondary recombination events. Our substrate was designed to detect recombination that occurred by an inversion. When the joining signals that recombine to one another are present in a different orientation, however, rearrangement might also be imagined to produce deletions. We believe that this is the case in several of the cell lines that we isolated. As was mentioned previously, our initial survey of various mycophenolic acid–resistant PD derivatives indicated that the integrated substrate in some of these lines may have undergone multiple recombination events mediated by reoriented joining signals. We detected deletions that mapped to a region of the substrate that suggested their formation specifically via a secondary recombination of a reciprocal joint with the $J_{\kappa 1}$ element.

To examine directly the structure of two integrants that appeared to have undergone secondary rearrangement events, we cloned the rearranged substrate from the lines PD-D and PD-M. In both cases, the coding joints were fusions of V_κ and $J_{\kappa 4}$, but the reciprocal joints had been derived from V_κ and $J_{\kappa 1}$ (Fig. 2). Because recombinants D and M were missing all of the J_κ cluster lying 3′ to the $J_{\kappa 1}$ heptamer and 5′to $J_{\kappa 4}$ coding sequences, their structures were consistent with the secondary recombination reaction diagramed in Fig. 4. As indicated in Fig. 4, the recombination of a reciprocal joint with a reoriented $J_{\kappa 1}$, if it occurs in a fashion analogous to primary recombination events, will have two consequences: the formation of an H_I/H_V reciprocal joint and the deletion of a portion of the J_κ cluster. These alterations had evidently occurred in both recombinants D and M.

Mapping data indicated that deletions of portions of the J_κ cluster similar to those observed in D and M—every one of which was accompanied by the presence of an H_I/H_V reciprocal joint—existed in 5 of the 14 original PD isolates. Because we have confirmed the structures that we inferred from the in situ maps of the integrated substrate in the cases of PD-M and PD-D, we would argue from these observations that (i) secondary reciprocal joint recombinations occur fairly frequently, and (ii) H_I/H_V reciprocal joints are overrepresented as a consequence.

The deletive secondary rearrangement of the substrate provides a simple and consistent paradigm for both the nonreciprocal junctions and the H_I/H_V bias that is generally observed in the endogenous rearranged kappa alleles of myelomas (*8, 9, 11–13*). An obvious explanation for the H_I/H_V bias found among the putative secondary recombinants in our collection is that H_I/H_V junctions represent a limit recombination product; in the process of rearranging to $J_{\kappa 1}$, all other similarly oriented J_κ elements are eliminated. By analogy, repeated reciprocal joint recombination events might well underlie the parallel H_V/H_I bias and nonreciprocal junctions observed in endogenous myeloma genes. In contrast to our previous suggestion (*10*), however, reciprocal joint recombinations do not involve intact V_κ gene segments and thus are related, but not identical, to V_κ-to-J_κ joining events. As a direct demonstration of secondary rearrangement of a reciprocal joint, an example of an individual integrated proviral substrate was isolated both before and after it had recombined.

Experimental proof of secondary recombination. The secondary deletions that occurred in recombinants D and M removed a segment from the substrate which presumably would have contained one of the two junctions formed in the event. A more complete analysis of secondary recombination requires the examination of both products of a secondary rearrangement, not just one. To this end, we chose to isolate an example of a secondary inversion rather than a secondary deletion. Accordingly, we started with the line PD-A (Fig. 1) because the substrate in this line would have to invert if it were to rearrange a second time, due

to the orientation of all the available target J_κ segments.

To isolate secondary recombinations of the substrate in PD-A, we subcloned PD-A in selective medium and then passaged one subclone, PDA-1, in nonselective medium. It was necessary to passage cells in nonselective medium because inversion of the *gpt* gene would have been lethal to cells maintained in mycophenolic acid (deletions such as those observed in M and D, in contrast, do not interfere with *gpt* expression). After about a month, we subcloned PDA-1 in nonselective medium and mapped the substrate in these lines as before. The substrate appeared to have rearranged relative to the PDA-1 precursor in 3 of 15 subclones analyzed. The line PDA-1-1 was chosen for further study.

Recombinant A-1-1 was cloned into λgtwes · λB (Figs. 1, 2, and 5a). The V_κ joining signal of the original reciprocal joint in PDA-1 had recombined with $J_{\kappa 2}$. The DNA sequences of the recombinant junctions in A-1-1 are presented in Fig. 2. As expected we found that (i) the coding joint in A-1-1 was the identical $V_\kappa/J_{\kappa 1}$ junction as in recombinant A; (ii) the $J_{\kappa 1}$ heptamer that had been incorporated into the original reciprocal joint was now linked to $J_{\kappa 2}$ coding sequences, forming an H_1/J_2 junction; and (iii) a new reciprocal joint had been formed from the $J_{\kappa 2}$ heptamer sequences and that of V_κ (resulting in an H_2/H_V joint).

We conclude that reciprocal joints, once formed, are not recombinationally inert, and can participate in secondary recombinations with unrearranged J_κ elements. The enzymes that rearrange immunoglobulin genes can evidently recognize not only intact gene segments, but also joining signals that have been unlinked from their corresponding coding sequences.

Asymmetry is an integral feature of the joining process. The special case of secondary rearrangement of a reciprocal joint provided the opportunity to determine whether the asymmetry of the joining process—whereby coding joints are variable but reciprocal joints are invariant—depends on the sequence of the joined elements or instead reflects an underlying asymmetry in the enzymology. Secondary recombination parallels primary recombination (Fig. 6) with the exception that V_κ coding sequences are replaced by a $J_{\kappa 1}$ joining signal (H_1). This replacement can reveal whether the sequence of a joining signal in itself specifies a recombination site at the heptamer border, or whether it is the context of a sequence in the overall reaction that determines how it is processed. In the first case, we would expect the H_1/J_2 "pseudo-coding joint" to contain an intact $J_{\kappa 1}$-derived joining signal; in the second, we would expect the pseudo-coding joint to contain an H_1 heptamer lacking one or more bases as its coding-proximal border.

It is evident from the sequence of the pseudo-coding joint in recombinant A-1-1 (Figs. 2 and 6) that during the joining process a G-C (G, guanine; C, cytosine) base pair was lost from the $J_{\kappa 1}$ heptamer. Thus in a novel context, a joining signal can be treated similarly to a coding sequence and is no longer recombined at the heptamer border.

The loss of the terminal G from the $J_{\kappa 1}$ heptamer in recombinant A-1-1 was not adventitious. We have two other examples of pseudo-coding joints, both occurring in junctions found in recombinant X (Fig. 2). One of these represents a fusion of the same two elements that were joined together in A-1-1; the heptamer 5′ to $J_{\kappa 1}$ was joined to the coding sequences of $J_{\kappa 2}$. The other pseudo-coding joint was formed form the 5′ flank of $J_{\kappa 2}$ and the coding sequences of $J_{\kappa 4}$. Both of these pseudo-coding joints were like the pseudo-coding joint in A-1-1 in that the J_κ heptamers were not preserved intact. In addition, it can be seen that the independently isolated pseudo-coding joints in A-1-1 and X represent variable fusions of the same two elements (H_1 and J_2). Thus pseudo-coding joints are qualitatively analogous to coding joints; the presence of a J_κ joining signal in place of V_κ coding sequences does not perturb the process of joining in any detectable way.

Recombinant X has an unusual reciprocal junction. The structure of recombinant X is shown in Fig. 2 and its probable derivation is diagramed in Fig. 5b. The origin of X is of interest because, in addition to the V/J_1 coding joint, the H_1/J_2 pseudo-coding joint, and the H_2/J_4 pseudo-coding joint mentioned in preceding sections, it has an H_4/H_V recipro-

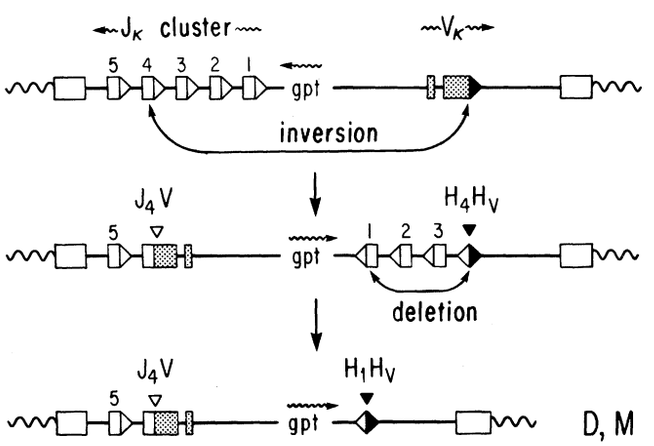
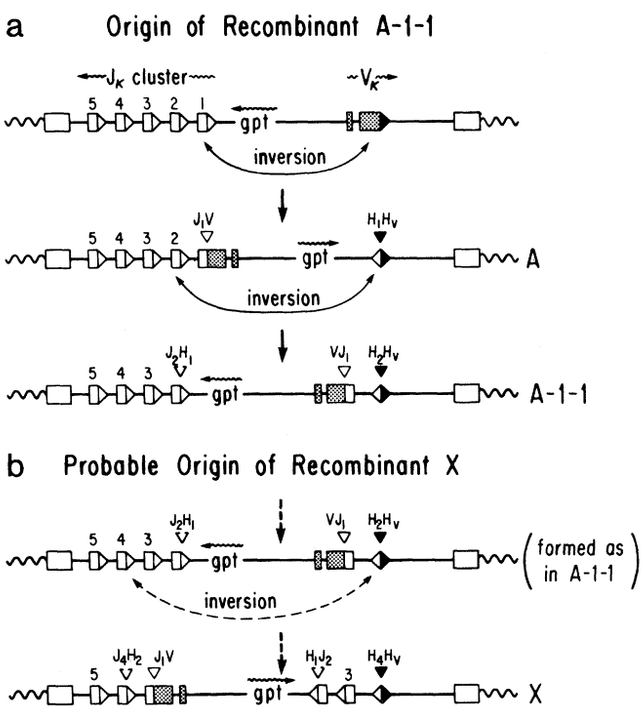


Fig. 4 (left). Derivation of recombinants D and M. Schematic representation shows the probable origin of D and M. Coding joints and reciprocal joints are designated as in Fig. 1. Fig. 5 (right). Derivations of recombinants A-1-1 and X. (a) Origin of recombinant A-1-1. The rearrangements that are known to have led to the final structure of recombinant A-1-1. (b) Probable origin of recombinant X. At least three stepwise recombinations must have occurred to form recombinant X. If no more than three rearrangements took place, an unambiguous order can be deduced from the final structure of X. The first two recombinations are as shown for recombinant A-1-1 (a); the last rearrangement is shown in (b).



cal joint. The H_4/H_V reciprocal joint in X is exceptional in that it is not a direct union between the heptamer of $J_{\kappa 4}$ and that of V_κ. Instead, the $J_{\kappa 4}$ heptamer is missing one base at the site of joining.

The significance of the imprecise reciprocal joint in X is difficult to evaluate because recombinant X was not isolated by serial subcloning of a precursor cell line. We cannot know for certain how it originated, although the simplest series of recombinations is a scheme that proceeds identically to the derivation of A-1-1 (Fig. 5a) and then in addition includes one final inversion (Fig. 5b). The remaining J_κ elements on the substrate were not recombinant, as determined by sequencing the $J_{\kappa 5}$ element and inferred from the map to be true of $J_{\kappa 3}$ as well, and the other regions of X indicated in Fig. 1 and 5b, appeared intact.

The junction in recombinant X is the only imprecise reciprocal joint among either the nine examples in our collection (see Fig. 2), or the nine endogenously derived sequences reported previously (*8, 11–13*). Although the terminal G of the heptamer may have been missing from the $J_{\kappa 4}$ heptamer prior to joining, this seems unlikely. (GTG)/(CAC) appears at the coding-proximal border of every heptamer known to be functional. Our tentative conclusion is that the G residue was eliminated during recombination, not before, and that the junction in recombinant X signifies little more than that the rule of precise reciprocal joint junctions is not absolute. Even so, as is evident from the compilation of crossover sites shown in Fig. 3 (closed triangles), variation occurs rarely and does not therefore undermine the general conclusions described in previous sections.

Splicing of retrovirally introduced sequences. In the initial analysis of the mycophenolic acid–resistant PD derivatives, we found several instances (in independent infections) of a deletion of approximately 200 base pairs that did not appear to be related to secondary rearrangements because it was located within the internal Bam HI fragment of the substrate. Two independent recombinants, D and N, were analyzed to characterize the structure of the deleted region in more detail. We found that the deletion precisely excised the intron between the V_κ leader peptide exon and the V_κ coding sequences proper (Fig. 1) (*38*). This presumably occurred while the substrate was in RNA form and had to have occurred before rearrangement because the leader splice sites are incorrectly oriented in the transcripts made after inversion.

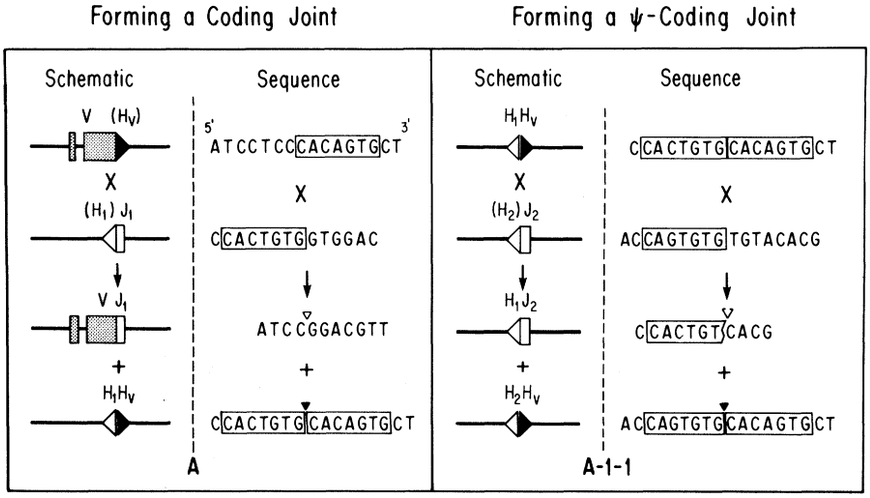


Fig. 6. Parallel reactions form pseudo-coding and coding joints. On the left is a schematic of coding joint formation, as well as the partial DNA sequence of the precursors and the products of the rearrangement that occurred in recombinant A. On the right, pseudo-coding joint formation, which is an analogous recombination, is shown. In the secondary rearrangement that formed the A-1-1 pseudo-coding joint, sequences designated H_1 (right) replace those designated V (left).

Value of retroviral introduction of recombination substrates. The analysis of the splice-induced deletions completed our investigation of all the structural variations revealed by our initial mapping experiments. The substrate may have acquired other alterations that we have not detected; on a gross level, however, the integrity of the substrate appears to be remarkably well maintained. As a result, all of the complex recombinations we analyzed (with the exception of the splicing event) were site-specific rearrangements. We attribute the low level of nonspecific rearrangement to our use of a retroviral delivery system. By comparison, a D- and J_H-containing substrate that was transfected into a similar Abelson transformant (*39*), apparently acquired a large D segment duplication. Illegitimate and partially homologous recombination events typically occur during transfection (*27, 28, 40*) and could potentially limit the interpretation of data obtained in this manner.

Mechanism of immunoglobulin gene rearrangement. While our data represent the most extensive complication of precursor and product sequences from V_κ-to-J_κ joining events, they are not readily interpretable in terms of a specific model for variable region exon formation. It is useful, however, to consider the information provided by these data in light of schemes that have been suggested and to formulate the basic implications.

Two features of immune system gene segment assembly imply that an unusual mechanism is at work. One is the flexibility of coding joint formation. This seems to be an intrinsic property of the process, and is exhibited even though crossing-over takes place in a region of no homology. The other is that reciprocal products are not exact reciprocals of one another. The two products of V_κ-to-J_κ joining do not conserve all of the sequence present in the precursor elements.

The suggestion has been made that joining of gene segments is initiated by a double-strand scission at the borders of the joining signals, at the edge between the conserved heptamers and the coding sequences they abut (*13, 14*). Although extracts prepared from cells that undergo joining have failed to exhibit an activity with exactly this property, a nuclease has been identified that can make double-stranded cuts in vitro near J_κ joining signals (*41, 42*). In addition, duplex breaks are thought to initiate recombination in other systems (*43, 44*).

If, in our example, recombination is initiated by simultaneous double-strand cuts at the heptamer borders, then the four ends that are created by these cleavages have to be differentially processed in the ensuing steps. The joining signals must be fused without alteration, whereas the coding sequences must be variably trimmed before they are ligated. Two suggestions for how this could occur have been offered. One proposal is that reciprocal joints are ligated directly after cutting, but coding joint ends are not joined immediately. Instead, coding joints are held in proximity to one another while an exonuclease removes some of their sequences (*14*). The other proposal is that an exonuclease that is spe-

cific for coding joint sequences can trim coding joint ends and will spare reciprocal joint ends, presumably acting at a stage when all four duplex ends are equally exposed (*13*). These suggestions are similar in that the coding joint junction is not formed until after the termini have been trimmed.

Our data rule out the second version (*13*) in its simplest form. Joining signal sequences must not be intrinsically resistant to a putative nuclease (nor are coding sequences specifically sensitive), because pseudo-coding joints can be formed in which heptamers do not remain intact (Figs. 2 and 6).

In both of the studies cited above, the structure of reciprocal recombination products was inferred from the structures of randomly isolated, nonreciprocal junctions. Since then, we have directly demonstrated that bases are indeed missing from the products of a single recombination event (*15*). While it may seem reasonable to assume, as implied in both of the above proposals, that (i) the bases that are missing from the recombinant products are subtracted from the coding joint ends only and (ii) the missing bases are eliminated before ligation forms recombinant connections, these assumptions are not demanded by the data. For example, as in the proposed mechanism for transposition of some prokaryotic elements (*32, 45*), the cuts that occur before strand exchange need not be directly across from one another on each duplex but might be staggered. If such were the case for V-to-J recombination, schemes can be invented (Fig. 7) whereby the bases that are missing in the final products are (i) lost during repair of both coding joints and reciprocal joints and (ii) are lost in reactions that occur after the principal connections have already been made. This is a qualitatively different view from the proposition that an exonuclease is an obligatory participant at a stage after breaks are introduced into the target duplexes and before strand exchange occurs (*13, 14*).

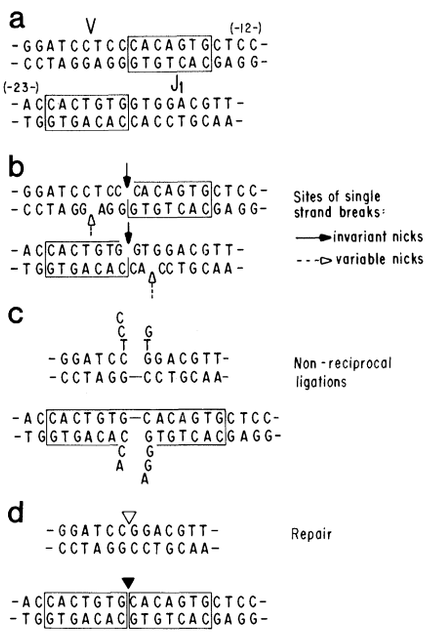


Fig. 7. Scission and ligation reactions in V_κ-to-J_κ joining could precede exonucleolytic trimming. One of several possible patterns in which single-strand interruptions and recombinant connections could result in the observed V-J rearrangement products is shown. The loss of nucleotides that accompanies V_κ-to-J_κ joining need not necessarily be a removal from the coding joint alone (*20, 21*), nor is it required that an exonuclease act as an integral, intermediate participant in the process (*20, 21*). (a) V_κ and J_κ precursor sequences. Heptamer sequences are boxed. Sequences to the left of the heptamer (V) or to the right of the heptamer (J_1) are the coding sequences. The remainder (including the boxed heptamers) in each line partially represent the joining signals; the length of the spacer region of each joining signal is indicated in parentheses, but the nonamer sequences are not shown. (b) Location of four single-strand scissions. Each target duplex has an invariant break on one strand and a variable break on the other. The polarity of the strand in each duplex that receives an invariant nick is the same relative to the coding sequences. (c) Two nonreciprocal ligations are indicated. The 3′ end of an invariantly nicked strand from the V_κ joining signal is joined to the 5′ end of the invariantly nicked strand of the J_κ joining signal. This ligation precisely fuses the heptamer borders. There are two other termini created by the invariant nicks that are not reciprocally joined. Another covalent bond similarly forms the coding joint—it occurs between two of the four termini that were created at the sites of the variant nicks. In this second ligation, the same target duplex donates a 5′ terminus as in the reciprocal joint fusion. The net result of cutting and joining as specified are two recombinant duplexes, each ligated on only one of their two strands. Each duplex contains noncomplementary tails opposite the newly formed junction. (d) Maturation of the recombinant junctions. The structures in (c) could be trimmed and sealed by nonspecific repair activities in the cell, resulting in the final products shown in panel d. Rather than removing nucleotides from only one of the recombinant junctions (the coding joint), exonucleolytic trimming would have operated on both recombinant joints equally. A variety of hypothetical pathways involving asynchronous or tightly coupled steps could be assembled that would lead to the structures in (c). The main purpose of the outline is to show that the formation of the principal recombinant connections might occur prior to exonucleolytic events. The invariant nicks are situated at chemically nonequivalent sites; whether this is feasible is probably best considered when more information about the enzymology of the process becomes available. One direct implication of the scheme as shown, however, is that the two targets, V_κ and J_κ, must be differentiated from one another according to which donates the 5′ end and which donates the 3′ end to each of the two junctions shown in (c). One can imagine that, during the nonreciprocal ligation steps the target V and J elements are designated 3′- or 5′-end donors by the spacer lengths of their joining signals.

The scheme in Fig. 7 is not, of course, the only alternative to the previous proposals. In its broadest outline, the actual mechanism might be fundamentally different from any break-join-nibble proposition. Because the products of immunoglobulin gene rearrangement are not like those of any other site-specific recombination reaction, clues as to a possible mechanism may not emerge by comparison. However, nicking/closing enzymes have been suggested to catalyze strand exchange in a variety of both site-specific and illegitimate recombination systems (*45–51*). It seems reasonable to at least consider the possible function of a type I topoisomerase in the present example. Therefore, we have suggested one conceptual approach to a mechanism that could rely principally on single-stranded break/join reactions.

Summary. The features of V_κ-to-J_κ recombination (and by extension, recombination of the other antigen-receptor genes in the immune system) are as follows. (i) Coding joints display variable junctions, with no apparent restriction on the association of one particular V_κ crossover with one particular J_κ crossover site; (ii) reciprocal joints have generally invariant crossover sites, specified by the coding proximal border of the heptamer element; (iii) the recombinase can reproducibly, site-specifically recombine DNA elements that do not include V_κ coding sequences; (iv) joining elements can recombine in either orientation; (v) a J_κ heptamer element can be either precisely joined or imprecisely joined depending on its context in the overall recombination reaction; (vi) the mechanism although macroscopically reciprocal, is nonconservative—a variable loss of base pairs occurs overall; and (vii) some of the time, base pairs can be added into coding joints.

In a more general sense, we have directly tested an inversion-deletion model for kappa gene joining (*15*) by introducing a substrate that was designed to reflect, in miniature, the way the endogenous kappa locus itself may be arranged—that is, with V_κ "backward" relative to J_κ. We show here that such a substrate recombines to reproduce most of the recombinant forms of the kappa locus observed in myelomas (coding joints that are not related to the accompanying reciprocal joints, deletion of J_κ gene segments, and a $J_{\kappa 1}$ bias among reciprocal joints). While this is

not a critical test, it firmly establishes that our model is feasible and can account for the phenomena we sought to explain. Other models such as unequal sister chromatid exchange (*9*) or reintegration of DNA that is deleted in the initial joining step (*8*) have not been similarly supported.

Note added in proof: It has recently been shown that rearrangements in endogenous K genes have properties that imply an inversional joining mechanism (*52*).

References and Notes

1. S. Tonegawa, *Nature (London)* **302**, 575 (1983).
2. R. Joho, C. Nottenburg, R. Coffman, I. Weissman, *Curr. Top. Dev. Biol.* **18**, 15 (1983).
3. T. Honjo, *Annu. Rev. Immunol.* **1**, 499 (1983).
4. M. Davis, Y-h Chien, N. Gascoigne, S. Hedrick, *Immunol. Rev.* **81**, 235 (1984).
5. T. Mak and Y. Yanage, *ibid.*, p. 221.
6. H. Saito, D. Kranz, Y. Takagaki, A. Hayday, H. Eisen, S. Tonegawa, *Nature (London)* **309**, 757 (1984).
7. G. Litman *et al.*, *ibid.* **303**, 349 (1983).
8. M. Steinmetz, W. Altenburger, H. Zachau, *Nucleic Acids Res.* **8**, 1709 (1980).
9. B. Van Ness, C. Coleclough, R. Perry, M. Weigert, *Proc. Natl. Acad. Sci. U.S.A.* **79**, 262 (1982).
10. S. Lewis, N. Rosenberg, F. Alt, D. Baltimore, *Cell* **30**, 807 (1982).
11. J. Hochtl, C. Muller, H. Zachau, *Proc. Natl. Acad. Sci. U.S.A.* **79**, 1383 (1982).
12. J. Hochtl and H. Zachau, *Nature (London)* **302**, 260 (1983).
13. E. Selsing, J. Voss, U. Storb, *Nucleic Acids Res.* **12**, 4229 (1984).
14. F. Alt and D. Baltimore, *Proc. Natl. Acad. Sci. U.S.A.* **79**, 4118 (1982).
15. S. Lewis, A. Gifford, D. Baltimore, *Nature (London)* **308**, 425 (1984).
16. O. Bernard, N. Hozumi, S. Tonegawa, *Cell* **15**, 1133 (1976).
17. R. Mulligan and P. Berg, *Proc. Natl. Acad. Sci. U.S.A.* **78**, 2072 (1981).
18. W. Altenburger, M. Steinmetz, H. Zachau, *Nature (London)* **287**, 603 (1980).
19. E. Max, J. Seidman, H. Miller, P. Leder, *Cell* **21**, 793 (1980).
20. C. Coleclough, *Nature (London)* **303**, 23 (1983).
21. H. Sakano, Y. Kurosawa, M. Weigert, S. Tonegawa, *ibid.* **290**, 562 (1981).
22. Y. Kurosawa and S. Tonegawa, *J. Exp. Med.* **155**, 201 (1982).
23. S. Desiderio *et al.*, *Nature (London)* **311**, 752 (1984).
24. J. Kavaler, M. Davis, Y-h Chien, *ibid.* **310**, 421 (1984).
25. G. Siu, M. Kronenberg, E. Strauss, R. Haars, T. Mak, L. Hood, *ibid.* **311**, 344 (1984).
26. J. Wilson, P. Berget, J. Pipas, *Mol. Cell. Biol.* **2**, 1258 (1982).
27. R. Anderson, S. Kato, R. Camerini Ortero, *Proc. Natl. Acad. Sci. U.S.A.* **81**, 206 (1984).
28. C. Wake, T. Gudewicz, T. Porter, A. White, J. Wilson, *Mol. Cell. Biol.* **4**, 387 (1984).
29. S. Gerondikas, S. Cory, J. Adams, *Cell* **37**, 973 (1984).
30. S. Desiderio, personal communication.
31. M. Calos and J. Miller, *Cell* **20**, 579 (1980).
32. N. Kleckner, *Annu. Rev. Genet.* **15**, 341 (1981).
33. H.-P. Doring and P. Starlinger, *Cell* **39**, 253 (1984).
34. H. Varmus, *Science* **216**, 812 (1982).
35. D. Cleveland, *Cell* **34**, 330 (1984).
36. R. Thayer and M. Singer, *Mol. Cell. Bol.* **3**, 967 (1983).
37. R. Hawley, M. Shulman, N. Hozumi, *ibid.* **4**, 2565 (1984).
38. S. Lewis and D. Baltimore, unpublished data.
39. K. Blackwell and F. Alt, *Cell* **137**, 105 (1984).
40. M. Calos, J. Lebkowski, M. Botchan, *Proc. Natl. Acad. Sci. U.S.A.* **80**, 3015 (1984).
41. S. Desiderio and D. Baltimore, *Nature (London)* **308**, 860 (1984).
42. T. Kataoka, S. Kondo, M. Nishi, M. Kodairo, T. Honjo, *Nucleic Acids Res.* **12**, 5995 (1984).
43. J. Strathern *et al.*, *Cell* **31**, 183 (1982).
44. J. Szostak, T-Orr-Weaver, R. Rothstein, F. Stahl, *ibid.* **33**, 25 (1983).
45. Y. Kikuchi and H. Nash, *Proc. Natl. Acad. Sci. U.S.A.* **76**, 3760 (1979).
46. M. Been and J. Champoux, *ibid.* **78**, 2883 (1981).
47. R. Reed, *Cell* **25**, 713 (1981).
48. B. Halligan, J. Davis, K. Edwards, L. Liu, *J. Biol. Chem.* **257**, 3995 (1982).
49. M. Krasnow and N. Cozzarelli, *Cell* **32**, 1313 (1983).
50. P. Bullock, W. Forrester, M. Botchan, *J. Mol. Biol.* **174**, 55 (1984)
51. K. Mizuuchi, *Cell* **39**, 395 (1984).
52. R. M. Feddersen and B. G. Van Ness, *Proc. Natl. Acad. Sci. U.S.A.*, in press.
53. A. Maxam and W. Gilbert, *ibid.* **74**, 560 (1977).
54. H. Sakano *et al.*, *Nature (London)* **280**, 288 (1979).
55. Numbered according to E. Kabat *et al.*, "Sequences of Proteins of Immunological Interest," U.S. Department of Health and Human Services, Washington, D. C., 1983.

56. Supported by American Cancer Society grant IM-355P (to D.B.). We thank K. Kirkegaard for useful criticisms.

31 December 1984; accepted 27 February 1984

RESEARCH ARTICLE

Function and Autoregulation of Yeast Copperthionein

Dean H. Hamer, Dennis J. Thiele, Jeffrey E. Lemontt

Heavy metals play a dual role in biology. Ions such as copper and zinc are essential trace nutrients for all life forms because of their participation in oxidation, electron transfer, and various enzymatic reactions. In contrast, inappropriately high concentrations of the same ions, or of elements such as cadmium and mercury, act as potent inhibitors of cell growth and development. How do living organisms acquire sufficient heavy metal ions to survive yet protect themselves against metal poisoning?

The metallothioneins, and metallothionein-like proteins, are thought to play an important role in this homeostatic process. These small, cysteine-rich polypeptides chelate heavy metal ions through thiolate complexes and are present in many kinds of organisms including vertebrates, invertebrates, plants, fungi, and even prokaryotes. The amino acid sequences of the metallothioneins from higher eukaryotes have been strongly conserved during evolution, while those from lower forms show less homology and may represent the products of either convergent or divergent evolution. In vertebrates, metallothioneins are expressed in various tissues and cell types, the greatest accumulation occurring in the liver and kidney. Metallothionein synthesis is inducible, at the transcriptional level, by the same heavy metal ions to which the proteins bind. The metal content of a metallothionein thus depends upon the species, heavy metal exposure, and cell type from which it is isolated (*1*).

The precise role of metallothioneins in heavy metal metabolism has been debated ever since their discovery more than a quarter of a century ago. It is clear that they can protect cells against heavy metal poisoning since cells that overproduce metallothionein, as a result of gene amplification, are unusually resistant to heavy metal poisoning, whereas cells that synthesize low levels of these proteins are unusually sensitive (*2*). The role of metallothioneins in normal metal metabolism is less certain. It has been suggested that they may participate in heavy metal storage, transport, or the activation of metalloenzymes. However, the evidence rests largely on experiments in which metallothionein synthesis is altered by nonspecific treatments, such as food restriction or injection with actinomycin D, that could affect metal metabolism by other routes. It has also been speculated that metallothioneins might participate in sulfur or nucleotide metabolism, control of the intracellular redox potential, amino acid transport, cellular differentiation, or the regulation of their own expression (*1, 3*). None of these possibilities has been experimentally tested.

One direct way to determine the actual function or functions of metallothionein is to study the behavior of organisms that

D. H. Hamer and D. J. Thiele are with the Laboratory of Biochemistry, National Cancer Institute, National Institutes of Health, Bethesda, Maryland 20205. J. F. Lemontt is with Integrated Genetics, 31 New York Avenue, Framingham, Massachusetts 01701.