# RECOMBINATION OF ENDOGENOUS AND INTRODUCED KAPPA IMMUNOGLOBULIN GENE SEQUENCES IN THE A-MULV TRANSFORMANT PD

by

SUSANNA MAXWELL LEWIS

B.S., Tufts University
(1976)

SUBMITTED IN PARTIAL FULFILLMENT
OF THE REQUIREMENTS FOR THE
DEGREE OF

DOCTOR OF PHILOSOPHY

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY
January, 1985

Signature of Author _____
Department of Biology (January, 1985)

Certified by _____
David Baltimore
Thesis Supervisor

Accepted by _____
David Botstein, Chairman
Department Committee

# RECOMBINATION OF ENDOGENOUS AND INTRODUCED KAPPA IMMUNOGLOBULIN GENE SEQUENCES IN THE A-MULV TRANSFORMANT, PD.

by

Susanna Lewis

Submitted to the Department of Biology
in partial fulfillment of the requirements for the degree of
Doctor of Philosophy.

## ABSTRACT

To investigate the process by which antigen receptor genes rearrange, I have concentrated on the kappa immunoglobulin locus as a prototypical example. I have established by clonal analysis that the kappa genes in an Abelson murine leukemia virus transformed cell line (PD) undergo frequent and productive rearrangements during growth in culture. Persistant recombination activity is demonstrable even after many subclonings. The frequency with which an embryonic kappa allele rearranges in PD is high--approximately 1% per cell generation. Equivalent rates of rearrangement are not observed at either the heavy chain locus (which is recombinant on both alleles) or the lambda light chain locus (which remains in germline configuration).

The reconfigurations of particular kappa genes in PD can be traced through multiple recombination events. It is possible to observe secondary rearrangement and deletion of previously rearranged alleles. Secondary rearrangements occur at an approximately ten-fold lower rate than primary rearrangements.

Two recombinant products are generated at the kappa locus when a kappa gene rearranges. One represents the variable region exon fusion, and is termed a coding joint, the other represents a by-product of recombination, the reciprocal joint. As isolated from the genomes of kappa light chain-producing myelomas and hybridomas, reciprocal joints and coding joints appear to be non-reciprocal to one another. The examination of these junctions in the PD cell line revealed that both reciprocal joints and coding joints can rearrange secondarily. Thus it was possible that the initial products of recombination might be reciprocal but that secondary rearrangements obscured this relationship. This lead to the proposal that reciprocal joints are generated at the kappa locus because Vκ-to-Jκ recombination is an inversional event at least some of the time.

To pursue this question further, I designed a substrate that would become recombined after infection into cells as a defective retrovirus. It

was possible by this means to directly demonstrate that Vκ-to-Jκ joining could occur inversionally. Secondary inversional and deletional recombination events were also observed. The rate at which the integrated substrate sequences become recombined was measured was measured in one set of experiments and found to be comparable to the rate with which endogenous genes rearrange.

A comparative study of twenty recombinant junctions was undertaken to define some of the basic properties of the mechanism. Several features of recombination could be documented for the first time because it was possible to examine reciprocally-related junctions. This work demonstrated that Vκ-to-Jκ recombination is a reciprocal but non-conservative process. A small number of base pairs are missing from the products of rearrangement relative to the precursors. In addition, the two reciprocal products of rearrangement are qualitatively different; reciprocal joints are almost always invariant in structure whereas coding joints are variable fusions.

Because the introduced substrate sequences can undergo secondary recombination events, a novel type of junction, termed a pseudo-coding joint is formed as a result. Pseudo-coding joints are qualitatively like coding joints, demonstrating that the variable nature of coding joints is independant of the primary sequence of the DNA incorporated into the junction. The mechanistic implications of this structural analysis are discussed.

Thesis supervisor: Dr. David Baltimore

To my family
and to the memory of

Esther Lanman Cushman

and

C.I. Lewis

I wish to acknowledge Ann Gifford and Michael Paskind for their scientific help and for their friendship. They were generous with both.

I wish to thank my advisor for advising:


I owe a debt to Naomi Rosenberg for realizing at the outset that the cell line PD might be interesting. I thank Kathy Berkner, the embodiment of PD, for introducing a new word into my vocabulary. I thank PD.

And this is the time to thank my father.

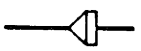
Table of Contents

## List of Figures

## List of Tables

TABLE 1

TERMS, ABBREVIATIONS, AND SYMBOLS

| TERM | DEFINITION | ABBR. | SYMBOL |
|---|---|---|---|
| germline Vκ | unrearranged Vκ | Vκ |  |
| germline Jκ | unrearranged Jκ | Jκ |  |
| Vκ coding sequences | | V |  |
| Jκ coding sequences | | J |  |
| Vκ joining signal | the heptamer-12 base spacer-nonamer element flanking a germline Vκ | Hv |  |
| Jκ joining signal | the heptamer-23 base spacer-nonamer of germline Jκ | eg. H1 |  |
| coding joint | fusion of Vκ and Jκ coding sequences | eg. V/J1 |  |
| reciprocal joint | fusion of Vκ and Jκ joining signals | eg. Hv/H1 |  |
| pseudo-coding joint | fusion of Jκ joining signal to Jκ coding sequences | eg. H1/J2 |  |
| PD-A (-D,-I etc.) | infected, mycophenolic acid resistant isolate of cell line, PD | | |
| recombinant A (D,etc.) | recombinant form of provirus after cloning out of PD-A DNA into phage λ | | |
| southern blot analysis | refers to analysis of cellular DNA sequences by the method of Southern (1975) as described in the Materials and Methods (Chapter V) | | |
| A-MuLV | Abelson murine leukemia virus | | |
| gpt gene | The E. Coli xanthine-guanine phophoribosyl transferase gene | | |

# I.INTRODUCTION

B and T lymphocytes can specifically recognize and respond to a wide array of antigens in the immune environment. Recognition is mediated by antigen binding proteins which exhibit similar domain structures in both B and T cell types (Hedrick et al. 1984, Patten et al 1984, Acuto 1984). In B cells, membrane-bound immunoglobulin is a tetramer composed of two heavy and two light polypeptide chains which are disulfide-linked to form a Y-shaped molecule (reviewed in Kabat et al 1983). In T cells, the immunoglobulin-like antigen receptor (the T cell receptor) is a heterodimer consisting of an $\alpha$ and $\beta$ chain (Allison et al 1982, Kappler et al 1983, Samelson et al 1983). The antigen-binding cavity of an immunoglobulin is composed of the N-terminal domains of one heavy and one light chain. By analogy, the component $\alpha$ and $\beta$ chains of the T cell receptor are thought to fold together into a similar antigen-binding cavity (Hedrick et al, 1984). The sequence homologies and structural similarities between these B and T cell products indicate that they are evolutionarily related, belonging to the same "super gene family" (Patten et al 1984, Mak and Yanagi 1984).

The study of monoclonal immunoglobulins revealed that these molecules posessed two qualitatively distinct regions; a variable region, and a constant region. The first appeared to be encoded by a hundred or more genes, but the second appeared to be a single-copy gene product. A number of years ago, Dreyer and Bennet (1965) proposed that the diversity among antigen-binding molecules might stem from somatic recombination involving one of a number of variable region genes and a single copy constant region gene. This conjecture became affirmed directly with the demonstration that a new recombinant linkage

between a multi-copy V gene segment and a single copy C gene segment was found in immunoglobulin producing cells (Hozumi and Tonegawa 1976). It is now known that the variable, antigen-binding domains of both immunoglobulins and T cell receptors, are encoded by recombinationally assembled exons (reviewed in Tonegawa 1983, Honjo 1983, Joho et al 1983, Davis et al 1984, Mak and Yanagi 1984). A seemingly random shuffling of variant reiterated gene segments expands the genetic endowment many-fold, making a major contribution to the diversity of antigen-binding specificities observed in an animal's immune system.

The recombination events that assemble variable region exons are site-specific, cell-type specific, and developmentally regulated. Site specificity was evident from DNA sequence studies, that showed that there was no homology between joined gene segments in the region in which they recombined (Bernard et al 1976). Instead, gene segments in the germ-line configuration are abutted by distinctive, conserved DNA motifs designated "joining signals" (Sakano et al 1979, Seidman et al 1979), which specify the crossover sites. Cell-type specificity has been demonstrated in studies that show that non-lymphoid cells do not rearrange their antigen receptor loci (Tonegawa et al 1977); further, a B or T cell will only productively rearrange the genes that are needed to template their specific receptor products (Chien et al 1984, Saito et al 1984). Finally, several observations support the notion of developmental regulation. V region exons probably assemble at a specified stage during development: normal B lineage intermediates can be isolated that appear to represent the specific cell types in which gene rearrangement is occurring (Coffman and Weissman 1983); in addition to which, transformed cell lines that are observed to undergo active rearrangement all have "pre-B" characteristics (Lewis et al 1982, Sugiyama et al 1983, Whitlock et al 1983,

Alt et al 1984)-- a more mature B lineage representative with the capacity to undergo rearrangement has never been reported. Immune system gene rearrangement is interesting therefore from more than one perspective; it is a fairly distinctive example of site-specific recombination as well as a landmark of differentiation in the B and T lineages.

The variability of the recombined genes of B and T cells requires that in order to perform biochemical analyses, it is necessary to expand a single clone of cells into a homogeneous population. Usually this is achieved through immortalization, by either infecting cells with transforming viruses, by fusing normal B cells to a permanent cell line, or by studying lymphomas and myelomas directly. To date, virtually all that is known about rearrangements of immune system genes has come not from a study of normal B and/or T cells but from their transformed counterparts (described briefly in Chapter II). A summary of the structures of the rearranging loci, and of the recombinant products that are formed as observed in transformed cells is given below.

## COMPARATIVE ANATOMY OF THE REARRANGING LOCI

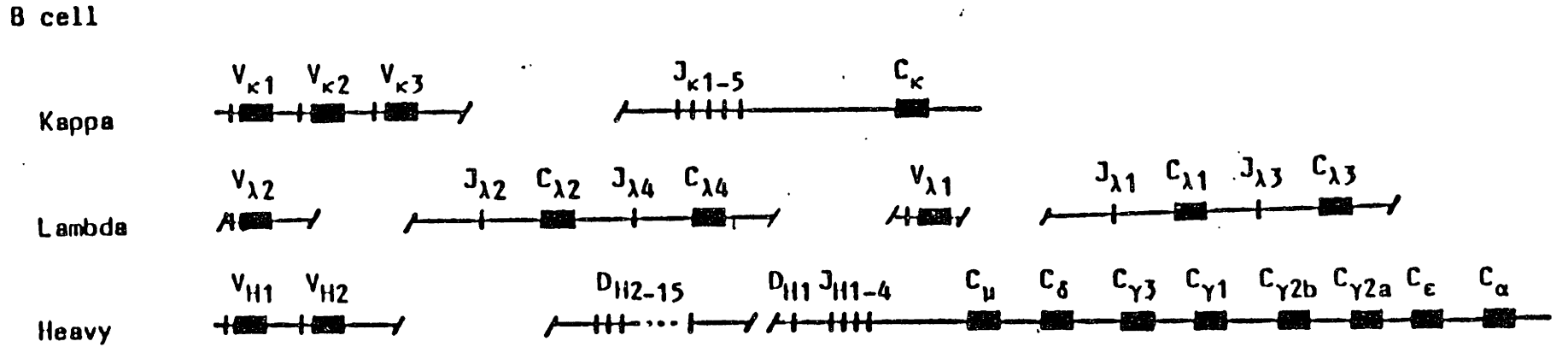
There are a total of six loci that are known to undergo rearrangement in lymphoid cells. The light chain (kappa, and lambda) and heavy chain loci must become recombined in order to be functionally expressed in B cells (reviewed by Tonegawa 1983, Honjo 1983, Joho et al 1983), and the beta, alpha, and a sixth as yet functionally-undefined locus, are observed to rearrange in a clonal T cell-specific fashion (reviewed by Mak and Yanagi, 1984, Davis et al 1984).

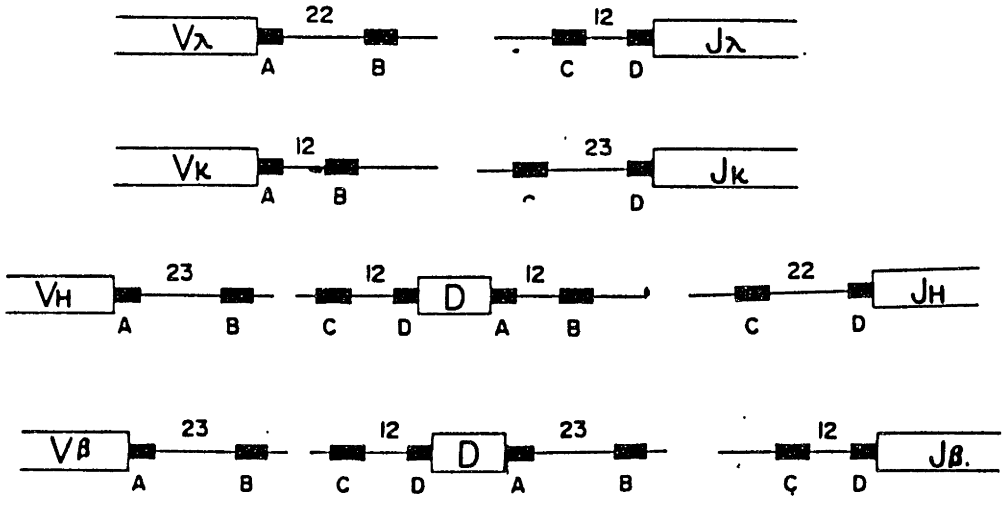
Figure 1

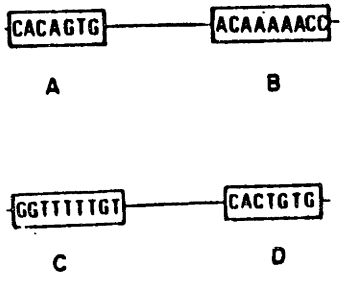
Structures of the rearranging loci.

Upper) Schematic representation of loci expressed in B and T cells (adapted from Malissen et al 1984). Breaks in the lines The indicated clusters of gene segments could be 5' or 3' to one another and in either orientation. Coding sequences are represented by boxes, joining signals are not indicated. The β chain locus is represented according to Siu et al (1984). (The two Jβ-Cβ exons are also referred to as $J_T$-$C_T$ and $J_T$,-$C_T$,--Davis et al 1984).

Lower) Representation of the joining signals linked to the gene segments of the four loci shown above: heptamer elements (designated A,D) are often near-palindromes; the nonamer elements (designated B,C) are AT-rich stretches. The exact sequence of a joining signal can vary, eg, at the β locus, the Jβ consensus is $N^{G}_{A}$CTGTG (Davis et al 1984).


B cell
Kappa
$V_{\kappa 1}$ $V_{\kappa 2}$ $V_{\kappa 3}$
$J_{\kappa 1-5}$ $C_{\kappa}$
Lambda
$V_{\lambda 2}$
$J_{\lambda 2}$ $C_{\lambda 2}$ $J_{\lambda 4}$ $C_{\lambda 4}$
$V_{\lambda 1}$
$J_{\lambda 1}$ $C_{\lambda 1}$ $J_{\lambda 3}$ $C_{\lambda 3}$
Heavy
$V_{H1}$ $V_{H2}$
$D_{H2-15}$ $D_{H1}$ $J_{H1-4}$ $C_{\mu}$ $C_{\delta}$ $C_{\gamma 3}$ $C_{\gamma 1}$ $C_{\gamma 2b}$ $C_{\gamma 2a}$ $C_{\epsilon}$ $C_{\alpha}$
T cell
Beta
$V_{\beta 1}$ $V_{\beta 2}$ $V_{\beta 3}$
$D_{\beta 2-x}$ $D_{\beta}$ $J_{\beta 1-1-y}$ $C_{\beta 1}$ $D_{\beta}$ $J_{\beta 2-1-6}$ $C_{\beta 2}$


Vλ
22
A
B
12
C
D
Jλ
Vκ
12
A
B
23
D
Jκ
VH
23
A
B
12
C
D
D
12
A
B
22
C
D
JH
Vβ
23
A
B
12
C
D
D
23
A
B
12
C
D
Jβ.


CACAGTG
ACAAAAACC
A
B
GGTTTTTGT
CACTGTG
C
D

Component Gene Segments

The antigen receptor genes are built by fusing a total of two or three (and conceivably more) separate units of DNA. In all cases, the basic formulation is V-to-J (Figure 1). A V gene segment (see Table 1) contributes a promoter, a leader peptide exon encoding 17 to 29 amino acids, and the first approximately 95 codons of the variable region exon; a J segment contributes the last 15 or so codons of the exon as well as the 5' splice site that defines its 3' border. The recombinational assembly of V-to-J positions a variable region exon a few kb upstream from the constant region exon(s), representing the full extent of the rearrangements required for expression. At the heavy chain and the beta chain loci, a third element is included in the variable region exon between V and J, which is termed a D segment (Early et al 1980, Sakano et al 1980, Kavaler et al 1984, Clark et al 1984). D segments can introduce up to 17 codons into the completed gene in some cases, but more commonly this number is 6 or fewer (Kabat et al 1983). Each of these gene segments, V,(D, where it occurs), and J is present in more than one copy at most loci (Tonegawa 1983, Honjo 1983, Joho et al 1983).

The process of rearrangement generates variability among assembled genes in three ways. First, reiterated copies of gene segments are different from one another so that the independant assortment of the various V's, (D's), and J's, ensures that a large number of variable region exons can be constructed from a relatively small number of germ-line components. Further, junctional insertions termed N regions (Alt and Baltimore 1982), are frequently found at the recombinant joints of heavy chain and β chain genes (Sakano et al 1981, Kurosawa and Tonegawa 1982, Desiderio et al 1984, Kavaler et al 1984 and Siu et al 1984). It is possible that junctional insertion may be due to an

accessory activity not integral to joining, because at least in the case of kappa and lambda genes, an unambiguous N region has never been reported. But a third source of variability appears to be mechanistically intrinsic to the recombinational process itself because the exact positioning of the junction between any two elements is not fixed (Sakano et al 1979, Max et al 1979). By joining gene segments at variable sites, this "flexibility" can create novel codons at the V/J junction (Weigert et al 1980).

## Joining Signals

The gene segments targeted for rearrangement at immune receptor loci are flanked by characteristic DNA sequences. The putative "joining signals" consist of a heptamer, a spacer region, and a nonamer (Sakano et al 1979, Seidman et al 1979--Figure 1). They are found 3' to V and, in inverted orientation (relative to the coding sequences) 5' to J. The heptamer and nonamer elements are evolutionarily conserved, being similar in different vertebrate classes (Litman et al 1983), as well as in genes that rearrange in different cell lineages (Mak and Yanagi, 1984, Davis et al 1984). Joining signals always have one of two forms; the heptamer and nonamer sequences are separated by an approximately 12 base spacer, or else by a 23 base spacer (Figure 1). Gene segments that are linked to joining signals of the 12-base spacer type appear to recombine only with those that are linked to joining signals containing 23-base spacers and vice versa (Early et al 1980, Sakano et al 1980). The role of joining signals is not known; they may serve principally as recognition elements, they may function as binding sites (Early et al 1980, Wood and Tonegawa 1983), they may specify cleavage sites (Alt and Baltimore 1982, Selsing et al 1984), they may provide homologies that somehow

facilitate proper folding of the substrate (Max et al 1979, Sakano et al 1979), or they may do all of the above.

Heptamer elements in the joining signals of many (but not all) gene segments are near-palindromes. At the kappa locus the sequence

$$5' \begin{matrix} CACAGTG \\ GTGTCAC \end{matrix} 3'$$

is found 3' to many Vκ gene segments and 5' to three of the five Jκ elements (Sakano et al 1979, Seidman et al 1979). The significance of heptamers as either palindromes or as inverted repeats is difficult to asses in light of the observation that the heptamers of fully functional joining signals can vary in both of these properties. At the T cell receptor beta chain locus for example, the heptamers of joining signals are often not palindromes, and further, a pair of evidently compatible joining signals are not necessarily exact inverted repeats of one another (Kavaler et al 1984, Siu et al 1984).

A possible clue to the function of heptamer elements has been suggested by Cheung et al (1984) who suggest that the sequence $\begin{matrix} GTG \\ CAC \end{matrix}$ has special physical properties that might cause a structure variation in the DNA, functioning as a "notch" for proteins in search of their binding sites. Despite the variation in the sequences of heptamer elements, one feature appears to be general; if heptamer elements are compared to one another with their coding-proximal borders in alignment, it is evident that $\begin{matrix} CAC \\ GTG \end{matrix}$ is highly conserved at the coding proximal edge. Following V gene segments, the joining signals almost always begin with CAC, and preceding J segments, joining signals almost always end with GTG. Exceptions to this can be found, but in no case has the gene segment that is linked to such an exceptional joining signal been observed in recombined form. In the absence of direct experimentation, this triplet

nevertheless seems to be strongly implicated as a necessary feature of a functional joining signal.

Nonamer elements are AT-rich, non-palindromic portions of joining signals (Max et al 1979, Sakano et al 1979) (Figure 1). In joining signals that occur 3' to gene segments, the A residues of the nonamer appear on the coding strand, when 5' they are on the non-coding strand. This feature is a constant whether or not the particular joining signal has a 23-base or a 12-base spacer (Figure 1). Thus, when two gene segments recombine, their associated nonamers will be roughly complementary to one another. Because nonamer elements are not stringently conserved, again it is difficult to say how significant the homologies are.

Conceivably, the structurally bipartite joining signals are also functionally bipartite. Occasional aberrant recombinations can be observed where non-immunoglobulin DNA has become joined to Vκ or Jκ sequences, and although in every case a recognizable heptamer could be found adjacent to the aberrant target in its unrearranged form, sequences similar to nonamer elements have not been described (Seidman and Leder 1980, Leder et al 1981, Hochtl and Zachau 1983, Durdick et al 1984). In one instance, significant blocks of inverted homology were observed near the heptamer element, but these were not separated by 12 or 23 bases after the fashion of a typical nonamer (Hochtl and Zachau, 1983). One interpretation of these observations is that the nonamer elements of joining signals increase the probability of recombination without being absolutely required for recognition by the recombinase (Hochtl and Zachau, 1983).

Although it remains to be seen exactly how the recombinase(s) interact with joining signals, the presence of conserved, easily recognizable, joining signals at all six loci that are known to rearrange during

immunodifferentiation indicates that in each case similar, perhaps identical, enzymes are at work (Davis et al 1984).

Arrangement of the component gene segments

V gene segments are located an unknown distance away from J segments (Figure 1). In the mouse, the most distal Vκ gene segment must be around 1,000 kilobases away from the Jκ region if one conservatively estimates that 100 germline Vκ elements exist (Valbuena et al 1978) with an average spacing of 10 kb between them (Wood and Tonegawa 1983). Jκ gene segments are clustered together at intervals of a few hundred bp, within an area several kb from the constant region. The Cκ constant region exon is single copy in mice, although this is not always the case (eg. there are two Cκ sub-isotypes in the rabbit, Heidmann and Rougeon 1983, Emorine et al 1983); and the multiplicities of Vκ and Jκ likewise will vary among species (eg. Bentley and Rabbitts 1981). Generally speaking, the copy number of V is much greater than that of C. Because they are so widely separated, the orientation of the Vκ gene segments relative to the Jκ cluster is not known.

Relative to the kappa locus, other loci display variations on a common theme (Figure 1). For example, in inbred strains of mice, the λ locus contains only two variable region gene segments and four Cλ constant region exons (Blomberg et al 1981). However V, J, and C have the same basic structure, and the intron/exon arrangement of an assembled λ gene is similar to that found with κ. Because each Cλ exon is linked to its own Jλ gene segment, the choice of Jλ segment associates a Vλ gene segment with a different sub-isotype. Vλ2, for example, can be joined with one of three different constant regions (Reilly et al 1984). Although this option at the λ

locus would not appear to have a major impact on the immune repertoire, at the T cell receptor β chain locus, two β-chain constant region sub-isotypes (Malissen et al 1984, Gascoigne et al 1984), each capable of becoming recombined with the same Vβ (Patten et al 1984) have also been found. This observation raised the possibility that one constant region subtype might be expressed in helper T cells and another in cytotoxic T cells (Gascoigne et al 1984), but several studies have failed to support the notion (Davis et al 1984, Royer et al 1984). Therefore there is no example as yet of an instance where the difference in the multiplicities and overall organizational features between one rearranging locus and the next has a functional significance.

D Segments

The variable region exons of both heavy chain and beta chain genes are built from three gene segments, V,D and J, rather than just two. D segments are the source of a small and variable number of codons in a completed V region exon. About twenty elements termed Dh are located 5' to Jh, and an unknown number of Dβ segments occur 5' to the Jβ clusters (Kurosawa and Tonegawa 1982, Wood and Tonegawa,1983, Kavaler et al 1984, Clark et al 1984). The linkage of the D element that is most proximal to each J cluster has been established for both loci, (Sakano et al 1981, Kavaler et al 1984, Clark et al 1984, Siu et al 1984). The closest known D is located approximately 650 bp 5' to the J cluster in each example.

D segments have a joining signal on either side of a short block of coding sequence, and can recombine with a V gene segment at one border, and a J on the other. At the heavy chain locus, the joining signals on either side of Dh are both of the 12-base spacer type, (Figure 1) and are inverted

relative to one another. This, in theory, might be expected to allow Dh segments to recombine to Jh segments in either of two orientations. Against this expectation, no inverted D/J junctions have been documented in any complete Vh exon, nor in any partially assembled D-to-J fusions (Wood and Tonegawa 1983). The structure of an aberrant, inverted J/J junction in an Abelson transformed cell line (Alt and Baltimore, 1982) suggested the possibility that one step in its derivation was an inversional recombination between a D and a J. Because only a single example has been isolated, it is impossible to know how rare inversional joining might actually be. Clearly, a single orientation of D predominates in D/J junctions and this has been taken as evidence for a tracking mechanism (Wood and Tonegawa, 1983).

Chromosomal location

The kappa, lambda, heavy chain and β chain loci have all been assigned to specific chromosomes in both mouse and man. In all cases, V gene segments are on the same chromosome as J gene segments. In addition, all the loci belonging to the group expressed in B cells (that is kappa, lambda and heavy chain) are sequestered onto separate chromosomes in both mouse and man (reviewed by Mak and Yanagi, 1984). The only two rearranging loci that can be assigned to the same chromosome are the kappa and the β chain loci, which are only loosely linked in mice (Davis et al 1984), and are on separate chromosomes in man (Mak and Yanagi, 1984).

What does macroscopic structure imply about the mechanism?

The features of the assembly mechanism that can be directly inferred from these comparisons are few. It appears that rearrangement involves specific recognition of joining signals, and that the recombinase can join elements separated by either small (less than 1 kb in the case of DQ52-to-Jh1 joining) or great (1,000 kb or more) distances. Further, the individual members of the multi-copy gene segments at the various loci must somehow be represented at least semi-randomly in the recombinant products.

A very basic issue is the question of how the recombinase finds its targets. Although nothing conclusive, pro or con, can be said about a tracking mechanism, there is some circumstantial evidence that supports this view. In addition to the orientation-specificity of the recombinant D segments observed at the heavy chain locus, there is evidence that could be taken to indicate that elements that are nearer one another recombine more frequently (Yancopoulos et al 1984). Effects due to position along the DNA seem more in keeping with a tracking mechanism than with an enzyme that freely diffuses in three dimensions between target sites. Also, it may be significant that the various loci that rearrange in a particular cell type are physically unlinked whereas within a locus, V and J are always on the same chromosome; one could imagine that if the recombinase must track along the DNA between target elements, this organisation would serve to ensure that mixed genes would not be created (Early et al 1980). Although the available evidence is open to several interpretations, the issue of tracking and whether or not inter-molecular recombination can occur is mentioned here as it is relevant to the various proposals for V-to-J recombination which will be discussed in later chapters.

A potentially related set of questions pertain to control of V region exon assembly. It is clear that rearrangement is initiated in certain cells and not others because the immunoglobulin genes in non-lymphoid tissues are not recombined (Tonegawa et al 1977). Further, the phenomenon of allelic exclusion (Pernis et al 1976) can been taken as evidence that, having initiated rearrangement, B cells do not recombine their genes continuously but instead will cease this activity in a regulated fashion (reviewed by Coleclough 1983). It seems then, that at the very least, rearrangement must be regulated to the extent that it is specifically activated, and then at some later point, discontinued. Beyond this, there is evidence that gene rearrangement may be differentially regulated at different loci (Perry et al 1980, Maki et al 1981, Alt et al 1981, Korsmeyer 1981, Joho et al 1983) and even perhaps within a given locus (Alt et al 1984). It may be that all of these events are orchestrated by precise regulatory mechanisms (Alt et al 1981,1984) but it is also possible that the ordering merely reflects intrinsically different recombination rates once rearrangement has been activated in a global manner (Perry et al 1981). Structural variations that exist at the different loci could in this instance play a significant role.

## FINE STRUCTURE OF THE RECOMBINANT PRODUCTS

### Sequence variation at the junctions sites.

When the sequences of recombined immunoglobulin genes are compared to their corresponding precursor elements, it is evident that the crossover junction is usually located in a region in which the target DNA segments are not homologous to one another (Bernard, 1976). This is in contrast with every

other site-specific recombination system characterized in either procaryotes or eucaryotes, all of which occur within regions of limited homology (eg. Nash 1981, Simon and Silverman 1983, Broach et al 1982). In addition, mating type-interconversion, illegitimate recombination and transposition have all been suggested to require varying degrees of homology (Strathern et al 1982, Beach 1983, Ruley and Fried 1983, Mizuuchi 1984). The difference between V/J recombination and these other systems suggests that the mechanism of V-to-J joining may be novel.

To those who would study V/J joining, the lack of redundancy in the sequences of the recombining species is an advantage, for by examining the sequence of the recombinants, the site at which strand exchange took place can be identified. This has led to the documentation of a curious feature of V region exon assembly; the precise site at which a given element recombines is not fixed, but will vary from one recombinant to the next (Sakano et al 1979, Max et al 1979, Early et al 1980, Weigert et al 1980, Kurosawa et al 1981, Manser and Gefter 1984, Desiderio et al, 1984). The apparent flexibility in the location of the junction sites may be related to the necessity of recombining non-homologous divergent sequences or, as has been frequently suggested, the rearrangement mechanism may have evolved this property in order to further expand diversity among immunoglobulin genes. Changes created by junction site variation have been demonstrated to affect the binding properties of the resulting protein (Azuma et al 1984, Jeske et al 1984) although it is not known how crucial to survival the expansion of the repertoire that is achieved by this means would be.

One consequence of flexibility is out-of-frame joining (Altenburger et al 1980, Max et al 1980). Recombinant genes with mismatched reading frames are observed not only in myelomas but in normal B cells (Early et al 1982),

suggesting that in part, faulty rearrangements contribute to allelic exclusion (Perry et al 1981).

Reciprocal V-to-J Recombinant Products

Although no V gene segment has been physically linked to a J segment, originally it was thought that it would be possible to distinguish between two possible orientations for these gene segments -direct or inverted- on the basis of the recombinant structures that would be formed from the V-to-J joining event itself. A simple prediction of a conservative, reciprocal, site-specific recombination between inverted V and J sequences is that two recombinant junctions would be formed, both of which would be found in the genome of a rearranged cell (Hozumi and Tonegawa 1976). Conversely, if a V gene segment in direct orientation were to fuse to a J gene segment by the same general mechanism, the result would be a deletion of the sequences between the two elements. Early experiments to detect intervening sequences or the reciprocal recombination products of either Vλ-to-Jλ (Sakano et al 1979), or Vκ-to-Jκ (Seidman et al 1980) fusion in several myeloma cell lines failed to demonstrate such junctions. This result suggested that Vκ and Jκ were in direct orientation, and deletion of the sequences between V and J was a consequence of their joining. Subsequent to these studies, it was found that in the genomes of other myelomas, structures that appeared to be the reciprocal recombination products of Vκ-to-Jκ joining could be occasionally detected. These "reciprocal joints" (see Table I) were not a consistent feature of rearranged alleles, nor were they apparently reciprocally related to the particular Vκ/Jκ fusions in the same cell (Steinmetz et al 1980). Though the presence of reciprocal joints was

mechanistically significant, their inconstant occurence and non-reciprocal relationship to "coding joints" could not be simply explained on the basis of either a deletional or an inversional scheme.

Three models have been proposed to account for the seemingly inconsistent structure and occurrence of reciprocal joints in myelomas. These are presented in detail in Chapter III. One proposal (Lewis et al 1982, Chapter III) which returns to topography for an explanation, is that the murine kappa locus is organized with some (or perhaps all) of the Vκ gene segments inverted relative to the Jκ elements, but that to explain the apparent non-reciprocity between coding joints and reciprocal joints and to explain the observation that kappa gene sequences have been eliminated from some myelomas that have undergone Vκ-to-Jκ joining (Seidman et al 1979), it was suggested that secondary rearrangement between Vκ and Jκ could occur even after a chromosome had already undergone one joining event. By this view, reciprocal joints may not accompany coding joints in a 1:1 ratio because either they are eliminated by secondary rearrangement or because some primary recombination events are between directly-oriented Vκ and Jκ elements. Implicit in the proposal are the requirements ι) that the recombination machinery is capable of recombining gene segments that occur in either orientation relative to one another ιι)that rearrangement is a reciprocal recombination process and ιιι) that secondary recombination of the same chromosome can occur. The fortuitous isolation of a clone containing an aberrant rearrangement of the heavy chain locus provided support for this view (Alt and Baltimore 1982). Although the pathway by which this recombinant was formed was uncertain, its final structure suggested that it originated as a result of three consecutive site-specific recombinations, one of which was an inversion.

## OVERVIEW OF THIS STUDY

To test the validity of an inversion/deletion model, I undertook a systematic analysis of a rearranging cell line, PD. By a series of serial sub-cloning steps, it was possible to trace the fate of a given allele. It was observed that secondary recombination of both coding joints and reciprocal joints could occur (presented in Chapter II). This fulfilled one requirement for the model outlined above.

To demonstrate that Vκ-to-Jκ joining is a reciprocal recombination process, and that the recombinase is able recognize joining signals in either of two orientations, I introduced a substrate with inverted Vκ and Jκ gene segments into PD. The kappa gene sequences on the substrate were observed to undergo both inversional and deletional site-specific rearrangements (Chapter III). The frequency with which inversional joining occured on the introduced substrate was not significantly different from the rate of rearrangement observed to occur in endogenous kappa gene sequences (Chapter II). Taken together, these studies provide substantive support for the inversion/deletion proposal.

The topography of the immunoglobulin loci is of interest principally because it represents the first level at which one might begin to examine the mechanistic attributes of the recombination process. The introduced substrate provided a direct means by which to document some of the basic properties of the joining reaction. Through the use of an introduced substrate, the precursors of gene rearrangement are identified unambiguously, and multiple examples of recombination between the same two elements can be obtained. The results of recombination can be examined in the absence of gene products which may otherwise affect the representation of a rearranged cell in a population.

In addition, both products of reciprocal recombination events could be studied in the present system because they could be co-isolated in a single inverted recombinant. For these reasons, a comparative analysis of a number of independant recombinants was carried out. It was demonstrated that the mechanism of Vκ-to-Jκ joining is non-conservative although grossly reciprocal, and that the reciprocal products of joining are qualitatively different from one another.

To investigate the basis for the qualitative difference between coding joints (formed from coding sequences) and reciprocal joints (formed from joining signals), an integrated substrate was manipulated so that a recombination event would occur in which a joining signal replaced the Vκ coding sequences. The junction that was formed was structurally analogous to a coding joint, demonstrating that the way in which a joining signal becomes recombined is dependant upon its context in the overall reaction. These analyses are presented in Chapter IV.

## II. CHARACTERIZATION OF THE IMMUNOGLOBULIN GENE REARRANGEMENTS OCCURING IN PD

INTRODUCTION: *In vitro* culture systems of pre B analogues.

The developmental stages of normal B cells are classified according to their display of surface antigens, morphology, differentiative potential in vivo and in vitro, and according to their manufacture and localization of immunoglobulins (for reviews, see Melchers et al 1977, Joho et al 1983). Multipotential stem cells in the fetal liver and bone marrow, and in the adult spleen and bone marrow, generate precursors belonging to the B cell lineage which through a series of intermediates, mature into circulating, terminally differentiated plasma cells. The most common classification of intermediates in the pathway from stem cell to plasma cell is based upon the synthesis of and localization of immunoglobulins. According to this scheme, the most primitive cell that is generally considered to be a precursor in the B lineage is a cell that expresses cytoplasmic μ chains, without evidence for secretion or surface expression of either μ chains or whole IgM molecules (Levitt and Cooper, 1980). This cell is termed a "pre-B" cell, but by several criteria, it appears that the category may well embrace several cell types (Coffman 1983). Morphologically, cells with a pre-B pattern of immunoglobulin expression (in which fairly short-lived μ chain are expressed in low levels without light chain in the cytoplasm) can be sorted into two classes, on the basis of size, and certain surface markers (Coffman and Weissman 1983). These correspond to two developmentally distinct classes: small pre B cells give rise to mature B cells (which express surface Ig), *in vitro*, whereas large pre-B cells do

not. Both large and small pre-B cells give rise to B cells in in vivo assays. These findings have been interpreted to mean that the small pre B cells are the more mature of the two types, and that large pre B cells are the precursors of small pre B cells (Joho et al 1983).

Densitometric measurement of the extent of rearrangement of the heavy and light chain genes in sorted large and small pre B cell populations revealed that large pre B cells have less than 1/10 the number of unrearranged heavy chain alleles than is observed in non-lymphoid (unrearranged) tissue. Despite apparent extensive rearrangement at the heavy chain locus, large pre B cells have no evidence of kappa gene assembly. Small pre-B cells (the more mature according to the in vitro assays) display the same degree of heavy chain gene rearrangement, but they also have acquired light chain gene rearrangements. This suggests that during development, immunoglobulin genes are assembled sequentially, with heavy chain gene rearrangement preceding light chain gene rearrangement (Coffman and Weissman, 1983).

Transformed B cell analogues appear to represent a fairly complete array of developmental stages along the proposed pathway from pre-B cell to terminally differentiated plasma cells (in part because such cells have played a large role in defining those developmental stages). The protein production and gene configuration of individual cells in the early phases of development have been elucidated through the use of pre-B leukemias (Vogler et al 1978, 1981 Heiter et al 1981, Korsmeyer 1981), pre-B hybridomas (Burrows et al 1979, Maki et al 1980, Coleclough et al 1981, Perry et al 1981) A-MuLV transformed lymphoid cell lines (Siden et al 1979) and a chemically-induced transformant (Paige et al 1978). The observation that A-MuLV transformants derived from in vitro infection of bone marrow cells were

observed to express μ chains without expression of light chains led to the first suspicions that a " μ-only" cell may be a normal intermediate in B cell development (Siden et al 1979). Likewise, the gene structures of A-MuLV transformants (Alt et al 1981, Sugiyama et al 1982) and pre B cell hybridomas, (Maki et al 1980, Perry et al 1981) provided the first explanation for the μ-only phenotype. In these pre-B analogues, rearrangement at both heavy chain alleles without any alteration of the kappa or lambda light chain loci was commonly observed, whereas the reverse situation (light chain gene rearrangement in the absence of heavy chain gene rearrangement) was not. Thus the pattern of protein expression that typifies a pre B cell seems to be due to the presence or absence of rearrangement of the various loci. The asynchronous assembly of immunoglobulin genes probably underlies the asynchronous onset of μ and κ chain production during ontogeny (Siden et al 1981) as well.

In short, all available evidence indicates that the gene structures and phenotypes of early lymphoid cell types fit into a fairly simple, linear differentiation scheme. The proposed sequence of gene rearrangment in early intermediates of B cell differentiation based on cell-sorting studies (Coffman and Weissman 1983), along with evidence from pre B cell hybridomas and leukemias (Maki et al 1980, Perry et al 1981, Alt et al 1981a,b, Lewis et al 1982, Sugiyama et al 1982, Whitlock at al 1983, Sugiyama et al 1983) is shown in Figure 2. Alongside are indicated some representative A-MuLV transformed lymphoid cell lines.

The cell line PD fits into the hypothetical scheme of B cell development as a transitional cell that is in the last stages of rearrangement and about to pass into the B cell compartment. PD manufactures cytoplasmic μ chains in the abscence of detectable light chain

Figure 2 (cont.)

intermediates has been studied by Coffman and Weissman (1983). "Large pre B cells" express B lineage markers but show no evidence of surface immunoglobulin. These cells have acquired rearrangements of their heavy chain genes and manufacture cytoplasmic μ chains, but are germ-line at the kappa and lambda loci (Coffman and Weissman 1983). Progression to a more mature phenotype is characterized by rearrangement at both the heavy and the light chain loci. These more mature "small pre B cells" can be distinguished on the one hand from their putative precursors by morphology and by the presence of the surface antigen Th.b, and on the other hand from mature B cells by their lack of surface immunoglobulin, and (as assayed on purified populations of cells) by a relatively lower extent of light chain gene rearrangement. Short term in vitro cultivation of small pre B cells results in the appearance of surface-immunoglobulin-positive B cells, which manufacture both heavy and light chain proteins (reviewed in Joho et al 1983).

"G"= germline: "R"= rearranged: "DJ"= Dh to Jh rearrangement: "VDJ"= a completely assembled heavy chain variable region exon:"+" designates a productive rearrangement: "-" designates an unproductive rearrangement: "*" designates continued rearrangement. " μ+", "μ-", "l+" or "l-" refer to production or non-production of cytoplasmic heavy or light chains.

Figure 2

Abelson cell lines may represent intermediates in B cell differentiation.

A. Some representative A-MuLV transformants.

Gene structures of all but three of the cell lines are described in Alt et al 1981, 1984. AT11-2 and AT11-2-24-6 are described in Sugiyama et al (1983) and Akira et al (1983). PD is described in Lewis et al (1982). Lines derived from in vitro infection of fetal liver as well as some of those derived from in vivo infection of newborn mice have the properties shown to the left: no μ chain is manufactured and both alleles have undergone D/J rearrangement. Several such lines have been documented to rearrange further at the heavy chain locus; the rearrangements consist of appending a Vh gene segment onto a pre-existing D/J rearrangement (Sugiyama et al 1983, Desiderio et al 1984, Alt et al 1984). These cell lines do not undergo recombination at their light chain loci at an equivalent rate (Alt et al 1981, Sugiyama et al 1983): both kappa and lambda genes are typically found to be unrearranged. A majority of A-MuLV transformants from in vitro infections of adult bone marrow are of the class shown in the middle: stably rearranged (either productively or unproductively) at the heavy chain locus and germ-line at the light chain locus (Alt et al 1981). Thus, on the basis of protein production and gene organization, these two classes of A-MuLV transformants appear to correspond to the "large pre B cells" described by Coffman and Weissman, (1983) (shown in Panel B). A few A-MuLV transformants produce κ or λ light chains and appear to be able to carry out rearrangement at one or both of these loci (Lewis et al 1982, and this thesis). This class of transformant has already undergone rearrangement of the heavy chain locus. The four cell lines that have been documented to undergo kappa gene recombination after derivation (Lewis et al 1982 and this thesis) and for which detailed information about the heavy chain locus is available (Alt et al 1984) appear to have undergone VDJ recombination on both homologs. All such lines became rearranged at the heavy chain locus prior to light chain gene rearrangement (this thesis). Thus this class of A-MuLV transformant appears to correspond to the more mature, "small pre-B cell" of Coffman and Weissman (1983) (Panel B).

B. Early stages in B cell differentiation (from Coffman 1983).

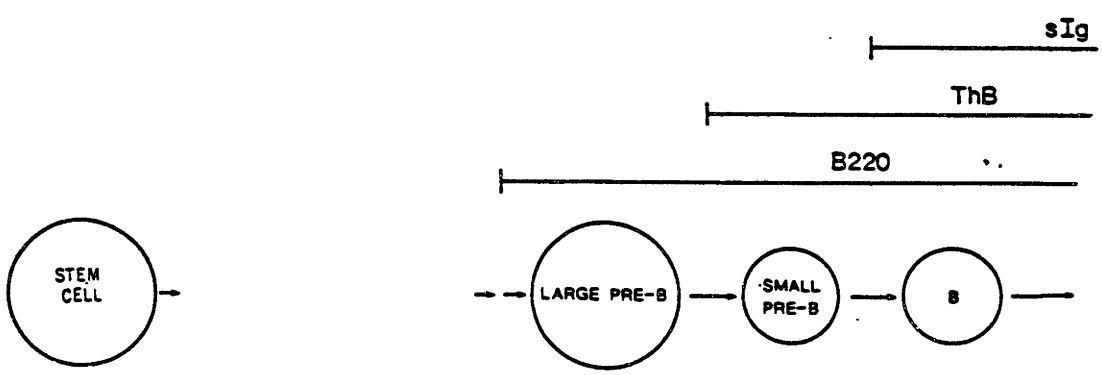
Early B cell differentiation is envisioned as progressing through an ordered series of gene reorganizations with heavy chain gene rearrangement preceding light chain gene rearrangement (references as cited in the text). The gene structure of B cell

A

| | 40E4-2<br>38B9<br>22D6<br>AT11-2 | 22D6-G-2<br>AT11-2-24-6 | 2M3-M<br>18-8<br>PD |
|---|---|---|---|
| IMMUNOGLOBULIN SYNTHESIS | $\mu^-$ | $\mu^+$ | $\mu^+$ I $^+$ |
| HEAVY CHAIN GENE ORGANIZATION | DJ/DJ* | $VDJ^+/DJ$<br>$VDJ^+/VDJ^-$<br>$VDJ^-/VDJ^-$ | $VDJ^+/DJ$ (?)<br>$VDJ^+/VDJ^-$<br>$VDJ^-/VDJ^-$ |
| LIGHT CHAIN GENE ORGANIZATION | G/G | $VDJ^-/DJ$<br>G/G | $VDJ^-/DJ$ (?)<br>G/G* |

B

SURFACE MARKER EXPRESSION




| | | | |
|---|---|---|---|
| IMMUNOGLOBULIN SYNTHESIS | $\mu$ | $\mu$ | $\mu L.\delta L$ |
| HEAVY CHAIN GENE ORGANIZATION | R/R | R/R | R/R |
| LIGHT CHAIN GENE ORGANIZATION | G/G | G/G. G/R and/or R/R | G/R or R/R |

proteins. The most distinctive feature of PD is that it rearranges its kappa locus during growth in culture. Thus PD corresponds to the phenotype and postulated gene structure of the small pre-B cell studied as a sorted cell population by Coffman and Weissman (1983). The relevance of PD or any transformed cell to normal processes that occur during differentiation is an open question. Because the same general constellation of properties that are associated with PD are found in independently-isolated A-MuLV transformants (Wabl et al 1984, Ziegler et al 1984, and see below), either PD is analogous to a real intermediate in B cell differentiation, or else PD presents a systematic, reproducible perturbation of physiological processes. By either view, a close examination of the gene rearrangements in PD cells is justified.

Recently, in vitro systems have been developed that allow the long-term cultivation of normal pre B cells (Whitlock and Witte, 1982, Denis et al 1984, Palacios et al 1984), and these promise to provide alternatives to transformation for the purposes of expanding single precusor cell clones for biochemical analysis. However, A MuLV-transformants have proved to be useful in the past, and will probably continue to be applied to certain studies in the future due to the ease with which they are cultured. Relative to other tumor model systems, A-MuLV transformants have the advantage of being nearly euploid, and stably so (Klein et al 1980); because they are not prone to changes in karyotype or ploidy, they are particularly well-suited to the study of variations in gene structure.

CLONAL ANALYSIS OF PD SHOWS THAT PD ASSEMBLES KAPPA LIGHT CHAIN GENES DURING GROWTH IN CULTURE

The A-MuLV-transformed line PD was derived by in vitro infection of bone marrow cells from an adult NIH/Swiss mouse (Rosenberg and Baltimore, 1976). The transformant was isolated as a single focus growing in soft agar, adapted to liquid culture and carried for about six weeks before being frozen. When PD cells from an early freezing were thawed and subcloned, about half of the derivatives displayed light-chain proteins that had not been detected in the uncloned PD culture (N. Rosenberg, unpublished observation). The light-chain proteins produced by different subclones migrated differently during electrophoresis through SDS-polyacrylamide gels (Figure 3). The species detected in metabolically-labeled lysates by immunoprecipitation with a rabbit antimouse κ serum could be competed by an MPC11 lysate but not by a MOPC104E lysate, confirming the identity of these proteins as the κ light chain isotype (MPC11 contains κ protein but no λ, MOPC104E produces λ light chain but no κ, data not shown).

To examine the structure of the κ light-chain genes in the subclones of PD (which will be referred to as primary subclones), DNA was extracted from each subclone and digested with Bam H1 and Eco R1. The Jκ- and Cκ-coding sequences, if unrearranged, are contained within a 6.8 kb fragment generated by these enzymes (Figure 4). Therefore, Vκ-to-Jκ rearrangement can be detected by southern blot analysis of the size of the band that hybridizes to pHBCκ (Figure 4).

When DNA from the uncloned PD line was probed using nick-translated pHBCκ, a single 6.8 kb band was observed (not shown) that migrated coincidentally with the fragment from BALB/c liver DNA (figure 5, lane L).

Figure 3

Protein production in various PD subclones.

Data is from Lewis et al (1982). Analysis of immunoglobulin production was by Dr. N. Rosenberg, Department of Pathology and Cancer Center, Tufts University School of Medicine. Immunoprecipitation of heavy and light chains from metabolically labeled ($^{35}$S-methionine) lysates was as described in Rosenberg and Witte (1980), using a rabbit antimouse serum reagant. Immunoprecipitates were collected using formalin fixed Staph A, reduced by boiling in gel buffer containing 2-mercaptoethanol, and run on 12.5% polyacrylamide gels under denaturing conditions. MPC11 and MOPC-104E myeloma extracts were included to provide markers for $\gamma$, $\kappa$ and $\mu$ chains as indicated. The species running in the region of the gel indicated by the bracket in the various subclone lines were confirmed in some cases to be kappa chains by competition immunoprecipitation (not shown).

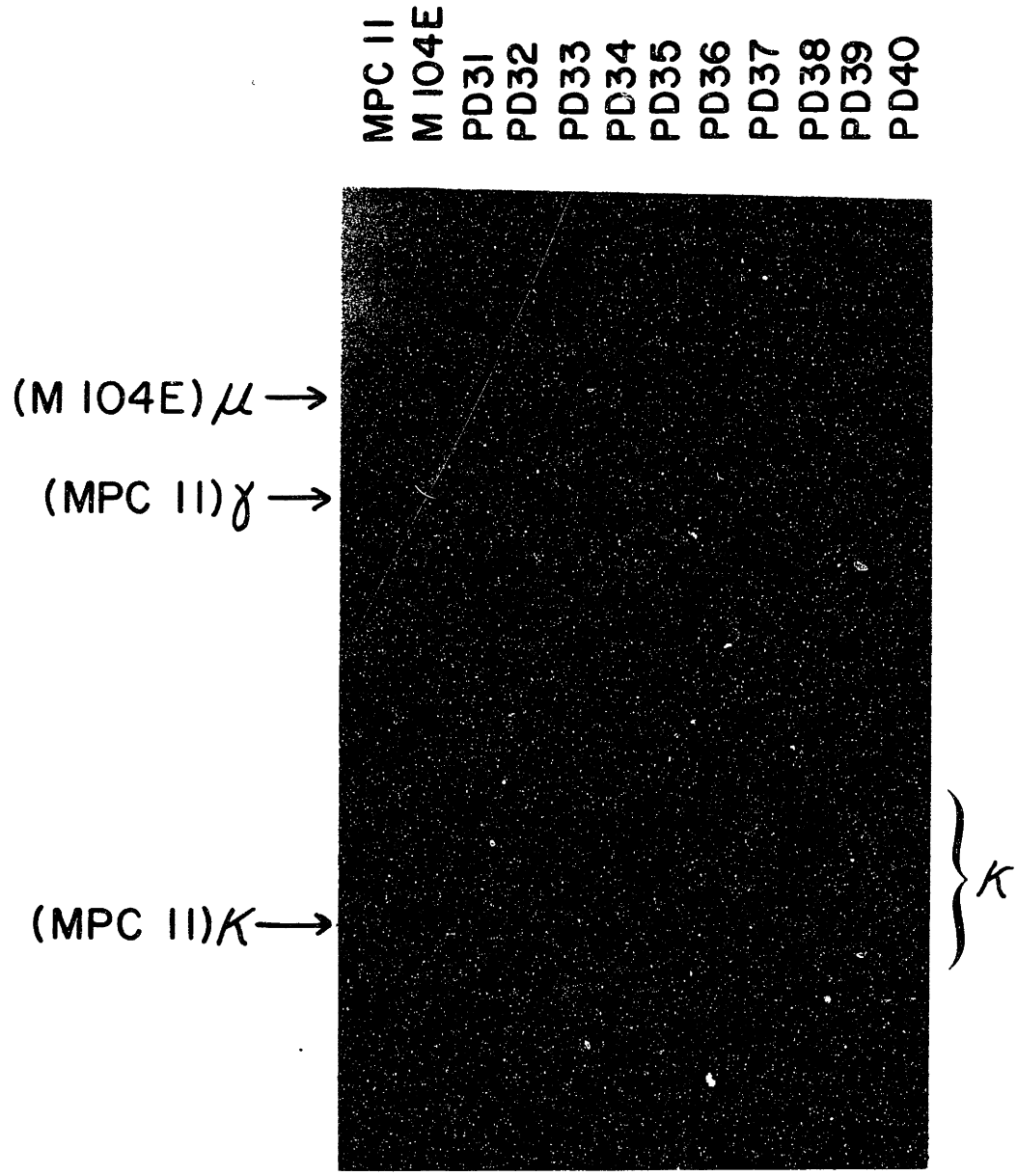

MPC 11
M 104E
PD31
PD32
PD33
PD34
PD35
PD36
PD37
PD38
PD39
PD40
(M 104E) μ
(MPC 11) γ
(MPC 11) κ
κ

Figure 4

Probes used for the analysis of $J_\kappa$ and Jh rearrangement.

A. The Eco RI-to-Bam HI region of the murine kappa locus and the regions that were subcloned for use as more specific probes are shown in alignment. Restriction endonuclease sites are abbreviated as follows: R--Eco RI; B--Bam HI; H--Hind III; Hf--HinfI. All sites within the region for all enzymes except HinfI are indicated. Subcloned regions were named as shown next to each representation.

B. Eco RI sites in the Jh-to-C region of the Balb/c murine heavy chain locus. pRI-Jh was used as a specific probe for alterations of the Jh-containing Eco RI fragment. By this assay, such alterations could represent either D/J or V/D/J rearrangement.

A
R
RH5'
R H
H5'
H
H Hf
Hf J2/3
J1 J2 J3 J4 J5
H
H HJK
H
HBCK
CK
B
RBJCK
1 Kb
KAPPA
B
R
R
R
RI-JH
JH
R
CU
R
1Kb
MU

Unlike liver DNA, the PD line also typically displayed a smear of indistinct lower molecular weight bands. When 13 primary subclones from PD were analyzed, 10 had strongly hybridizing bands at positions different from the 6.8 kb germline band and different from one another. It is evident from this result that the PD line accumulated heterogeneity at its κ locus sometime after the first cell was isolated. Because the three lines that apparently were unrearranged by this assay (PD31, PD35, PD40) also contained no detectable κ proteins (Figures 3 and 5), it appeared that many of the observed rearrangements correspond to in-phase Vκ-Jκ joining.

To determine whether the line was continuing to undergo active kappa gene assembly during in vitro culture, (as opposed to a single burst of recombination at one time after derivation) seven of the primary PD subclones were cloned again: 7 to 20 secondary subclones were derived from each. DNA from these secondary subclones was analyzed for κ gene rearrangement as described above. The same diversity was once again observed in the subclones of PD31 and of PD40 (Figure 6 shows the PD31 series). Like the uncloned PD culture, uncloned PD31 and PD40 each exhibited a single, prominent embryonic band; upon subcloning nearly all of the derivatives of these lines displayed diverse rearrangements at the κ locus. Primary subclones that exhibited different initial configurations were also analyzed in this manner. PD34 (Figure 9), PD36 and PD42 (not shown) similarly showed evidence of continued rearrangment in culture, but the variation was less extreme. The others, PD32 and PD39, showed no variation upon subcloning, and therefore appeared to be stable (not shown). The appearance of a "submolar smear" in these southern blot analyses appears to be a reliable indication of a high level of continued recombination,

Figure 5

Southern blot analysis of the kappa gene structure of the PD primary subclones.

Approximately 12 μg of genomic DNA from each subclone was digested with Eco RI and Bam HI, electrophoresed on 0.7% agarose gels, and transferred to nitrocellulose by the method of Southern (1975), with modifications as described in Chapter 5 (Materials and Methods). The hybridization probe used was nick-translated pHBCκ. The cartoon represents the structure of an unrearranged kappa locus; the bold line indicates the probe. Restriction endonuclease sites are abreviated as in Figure 4.

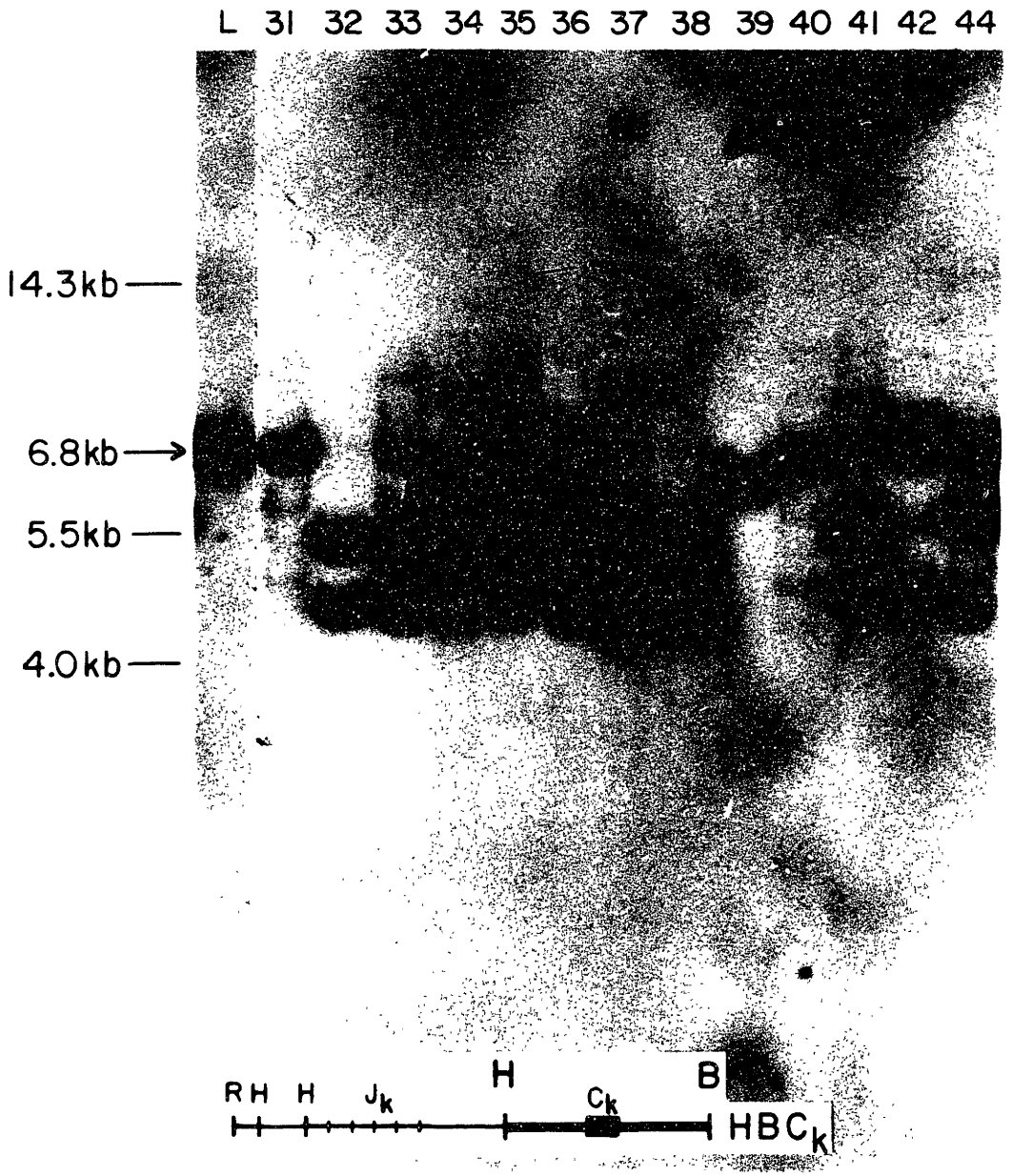
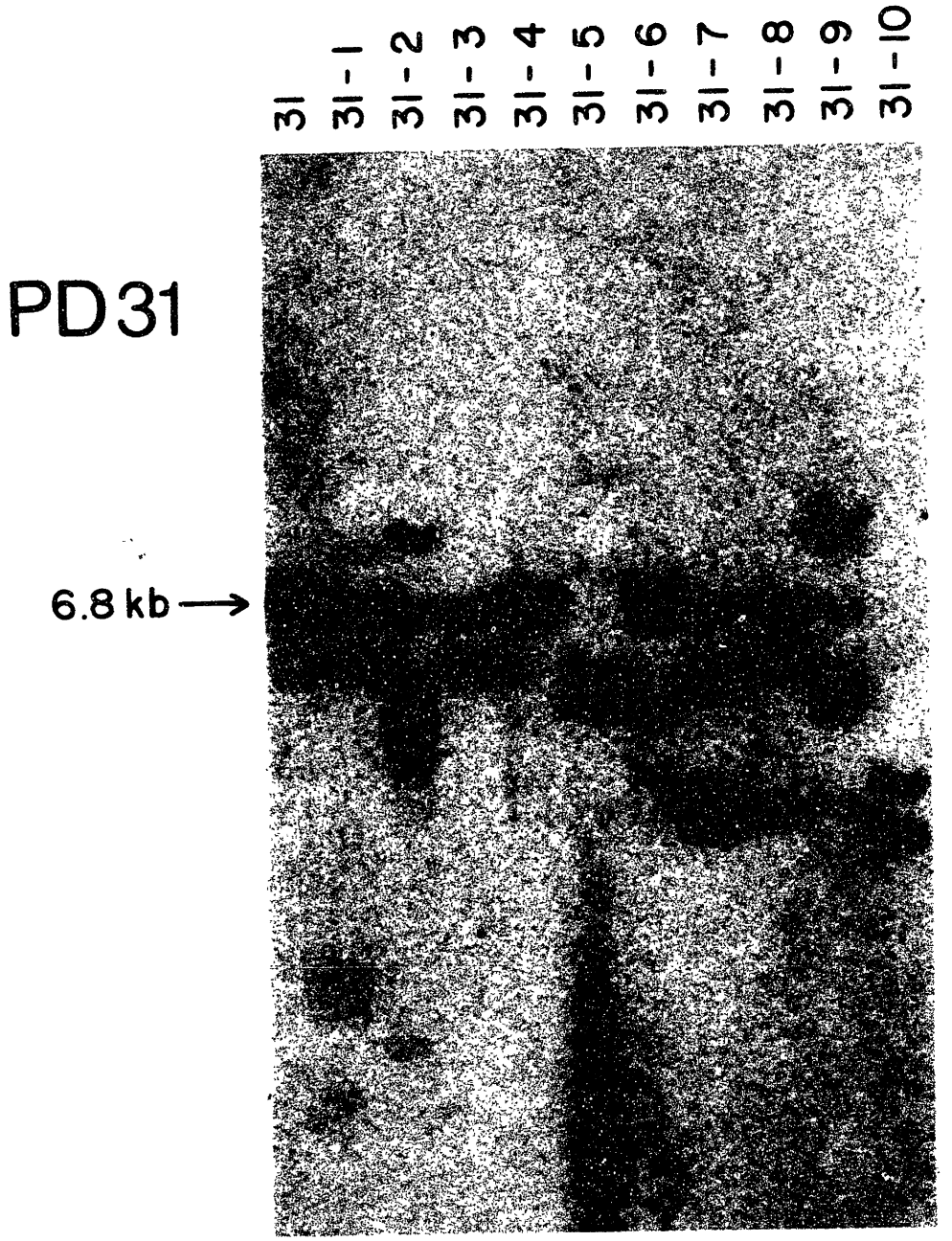

L 31 32 33 34 35 36 37 38 39 40 41 42 44
14.3kb
6.8kb
5.5kb
4.0kb
R H H $J_k$ H $C_k$ B
HB$C_k$

Figure 6

Kappa gene rearrangement in PD31 secondary subclones.

The southern blot analysis was carried out using the same restriction endonucleases (Eco RI and Bam HI) and the same probe (pHBCκ) as in Figure 5.


31
31-1
31-2
31-3
31-4
31-5
31-6
31-7
31-8
31-9
31-10
PD31
6.8 kb

because as seen in Figure 5, of all the primary subclones, only PD32 and PD39 lacked submolar bands.

## A REARRANGED KAPPA ALLELE CAN UNDERGO FURTHER RECOMBINATION

The analysis of several PD sublines suggested that a rearranged allele may be unstable in the PD cell line due to secondary recombination events. To document this in detail, one subclonal lineage was chosen for a more systematic examination. PD31-3 is a secondary subclone of PD31 that bears two rearranged κ alleles. By southern blot analysis, these alleles were detected as 6.8 and 6.0 kb Eco R1-Bam H1 fragments when probed with HBCκ (Figures 6,7). (The 6.8 kb allele is indistinguishabe from an embryonic band in this analysis, but when the PD31-3 DNA was cut with Bam H1 alone, and probed with pHBCκ, or when the pRH5' probe was used, no embryonic genes are detected in PD31-3 DNA).

Because by these analyses, the PD31-3 cell line contained two rearranged alleles, subclones of this line were examined for variants. Using pHBCκ as a probe, most of the PD31-3 tertiary subcones had the κ gene structure evident in the uncloned PD31-3 line (Figure 7, left). However, out of a total of 11 subclones examined, two lacked a 6.0 kb band and had a novel band in its place (PD31-3-4 and PD31-3-13, Figure 7). This result suggested that a rearranged band could rearrange again.

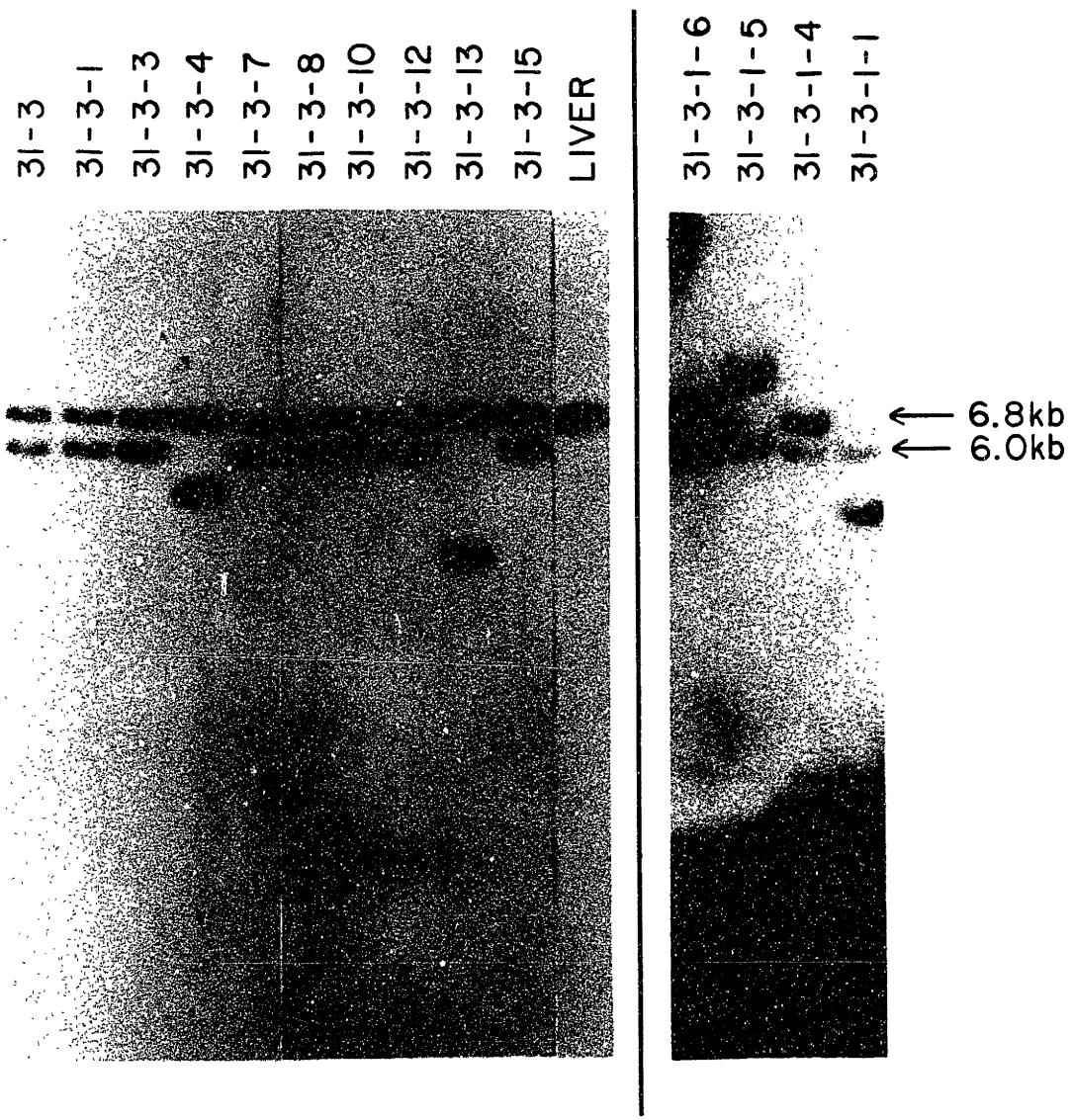
Because both of the PD31-3 variants lacked the same "6.0 kb" allele, it was formally possible that the variant bands did not represent secondary rearrangement of this allele but that each was an example of independent primary recombination of an undetected germline chromosome. In this case when PD31-3 was a single cell, it must have borne the "6.8 kb" allele on one

Figure 7

Kappa gene structure in the PD31-3 sub-lineage demonstrates that a rearranged allele can rearrange again.

Left panel shows the gene structures of the (tertiary) subclones of PD31-3. Southern blot analysis was as described in Figure 5 (cellular DNA samples were digested with Eco RI and Bam HI and the probe used was nick-translated pHBCκ). Both the6.8 kb band and the 6.0 kb band correspond to rearranged genes (see text).

Right panel shows the gene structures of the (quaternary) subclones of PD31-3-1. The analysis was essentially the same as in the left-hand panel except that pRBJCκ was used as a probe instead of pHBCκ (Figure 4). Other analyses have confirmed that all the bands visualized in the autoradiogram in the region of the arrows hybridized to pHBCκ (not shown).


31-3
31-3-1
31-3-3
31-3-4
31-3-7
31-3-8
31-3-10
31-3-12
31-3-13
31-3-15
LIVER
31-3-1-6
31-3-1-5
31-3-1-4
31-3-1-1
6.8kb
6.0kb

homolog and an unrearranged locus on the other. It was possible that, as the result of a "founder effect", the 6.0 kb band detected by southern blot of PD31-3 DNA reflected a predominant rearrangement in the culture, rather than the configuration of the original PD31-3 cell. (This possibility is very unlikely for reasons outlined in a later section.)

To rigorously prove that a rearranged chromosome can rearrange again, it was necessary to subclone a subclone of PD31-3. In this situation the gene structure of the founding cell is known. PD31-3-1 had the identical 6.0 and a 6.8 kb bands as the PD31-3 culture from which it was derived, but by virtue of the fact that it was a single cell derivative of PD31-3, it had been purified from any putative embryonic alleles as well as any independently produced rearrangements. Some of the quaternary PD31-3-1 subclones are shown in Figure 6 (right). Once again variants were observed. This time a precursor-product relationship between a previously rearranged band and the novel variants is established.

Secondary rearrangements are probably V/J joining events

The hybridization properties of the bands that represent secondary rearrangements support the view that these secondary events are rearrangements in which an upstream Vκ gene segment recombines to a Jκ segment lying 3' to the original junction. Such "leapfrog" rearrangements would be expected to delete the first junction, as well as the intervening portion of the Jκ cluster. The precursor 6.8 kb band hybridizes to pJ3/2, (which is specific for an internal protion of the J cluster encompassing Jκ2 and Jκ3, Figure 4) whereas the derivative alleles do not (not shown). The

loss of Jκ hybridization upon secondary rearrangement then, is consistent with a subsequent V-to-J joining event.

## RECIPROCAL JOINTS ARE OBSERVED IN PD

The pRBJCκ probe corresponds to the entire Eco RI-to-Bam HI fragment excised from an unrearranged genome. A strict deletion mechanism for the formation of Vκ-to-Jκ junctions (Sakano et al 1979, Max et al 1979, Seidman et al 1980) would support the prediction that by southern blot analysis, DNA from a stable PD derivative should display no more than one new band per rearranged chromosome if digested with Eco-RI and Bam HI, and probed using pRBJCκ (Figure 4). Contrary to this expectation, PD32 and PD39, showed more than two bands. Because multiple bands in these lines persisted unchanged upon subcloning (data not shown), the trivial explanation of their contribution by variant subpopulations in each culture was ruled out.

Relative to pHBCκ, pRBJCκ appeared to detect "extra" bands in most or the PD primary subclones. By subdividing probe pRBJCκ into smaller regions, we determined the hybridization patterns of all the bands it revealed. The probes used are diagrammed in Figure 4. They subdivided the Eco RI-to-Bam HI region of the kappa locus into three parts: pRH5' hybridizes exclusively to sequences that lie 5' to the Jκ cluster, pHBCκ hybridizes to sequences lying 3' to the Jκ cluster, and pJ3/2 is specific for an internal portion of the Jκ cluster (extending from 3' of Jκ1 to 5' of Jκ4 as shown). These probes were selected to detect the regions lying upstream, downstream and within the area which ought to be disrupted by V-to-J recombination events.

Four filters of parallel sets of Eco R1-Bam H1 digested primary

subclone DNA (with the variations noted), were hybridized to pRBJCκ, pHBCκ, pRH5' and pJ3/2 (Figure 8) . Three patterns of hybridization were found. Some bands hybridized only to pHBCκ (the "downstream probe"), some only to pRH5' (the "upstream probe"), and some hybridized to a combination of one of these plus J3/2. No bands (excepting bands representing unrearranged alleles, and possibly one of the rearranged bands in PD37) hybridized to both HBCκ and to pRH5'. The patterns that were observed were consistent with the notion that the full length pRBJCκ probe detected fragments corresponding to both reciprocal products of Vκ-to-Jκ joining. Rearranged bands hybridizing to HBCκ presumably represented coding joints, whereas those hybridizing to pRH5' were taken to represent reciprocal joints (terminology is defined in Table 1).

To map the recombination sites of these bands more precisely, a single cell lineage, the PD34 secondary subclone series (Figure 9) was analyzed. Because each of the rearranged bands in the PD34 subclone series has a unique distribution among the subclones shown in Figure 9, it was possible to vary the digestion and probe with pRBJCκ without confusing the identity of each junction. The recombination sites for all bands detected by pRBJCκ in the PD34 series could be mapped in this manner to the Jκ region, between the HindIII site 80 base 5' to Jκ1 and the XbaI site 300 bases 3' to Jκ5 (data not shown).

Reciprocal joints were first described by Steinmetz et al (1980). These junctions consist of the fusion of the joining signal lying 5' to Jκ to that lying 3' to Vκ1. As detected in myelomas and hybridomas, reciprocal joints have two peculiar properties: they occur with much less frequency than coding joints, and none has been demonstrated to be reciprocal to a coding joint from the same cell (Steinmetz et al 1980, Hochtl et al 1982,

Van Ness et al 1982, Selsing et al 1984). The PD34 subclone series suggests an explanation for this because it is evident that coding joints can become altered independently of reciprocal joints and vice versa (Figure 9).

## RECIPROCAL JOINTS CAN UNDERGO SECONDARY RECOMBINATION

Because PD34-4and PD34-7 have identical coding joints, it was possible that the differences in their pRH5'-hybridizing band were acquired through secondary rearrangement. Secondary reciprocal joint rearrangement can be visualized as being similar to the secondary rearrangement of coding joints. A pre-existing reciprocal joint might be altered when an unrearranged Vκ gene segment recombines with an unrearranged Jκ gene segment in the portion of the J cluster incorporated into a reciprocal junction. As long as the first reciprocal joint was not formed as a consequence of Jκ1 rearrangement, then it would contain unrearranged Jκ elements and could perhaps be a substrate for further recombination. This would occur far from the Cκ constant region because the process of fusing the 5' flank of a Jκ segment to the 3' flank of a V would have juxtaposed a remnant of the Jκ cluster to within a Vκ gene cluster.

The variant reciprocal joint in PD34-7 could have one of two origins. Either the band detected in PD34-7 was derived by an alteration of the band found for example PD34-4 (or vice versa), or a cell that was a common precursor of PD34-4 and PD34-7 contained two reciprocal joints, one of which was lost subsequently in each of the two subclones.

To prove that secondary reciprocal joint recombination can occur, the PD31-3 sublineage was analyzed. I previously reported that secondary deletion of a reciprocal joint (along with all evidence of hybridization to

Figure 8

Analysis of coding joints and reciprocal joints in PD primary subclones.

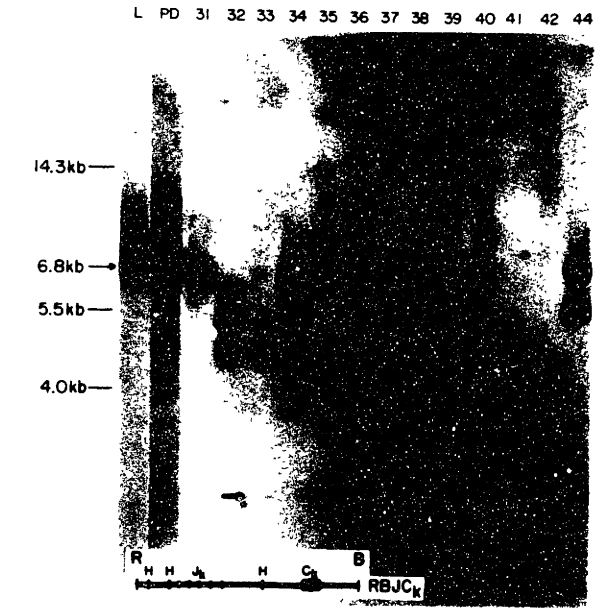
DNA samples of each PD subclone (all of which, with the exceptions noted, were primary subclones) were digested as beforewith Eco RI and Bam HI and assayed for hybridization to the probes indicated in each cartoon (see also Figure 4).

Top Left: Primary subclones analyzed with the probe pRBJCκ. This probe detects both reciprocal joints and coding joints. Arrows indicate bands that reproduced poorly (in addition, two bands in the lane marked 33 one of approximately 6.8kb and the other of higher molecular weight are also only faintly discernible).

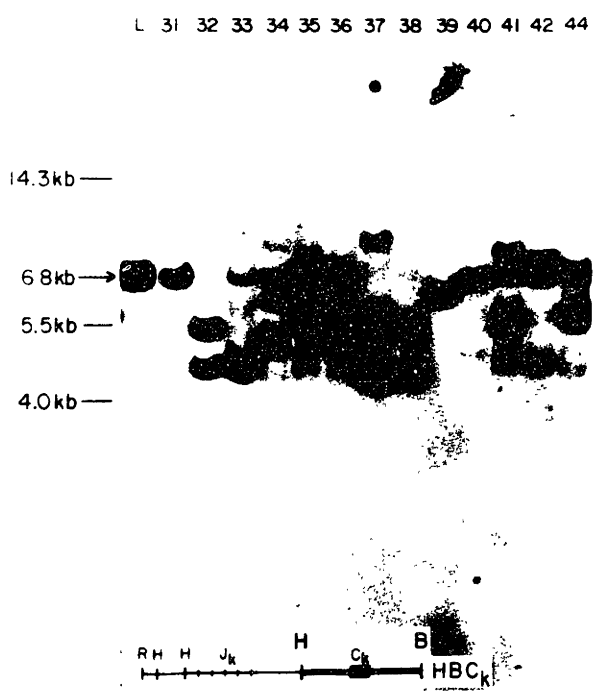
Top Right: Duplicate set of digests analyzed with the probe pRH5'. Among the rearranged bands, the probe pRH5' is specific for reciprocal joints (in addition it will hybridize to non-rearranged fragments). Lanes marked 31 and 32 (containing PD31 and PD32 DNA) were from a separate experiment.

Bottom Left: Samples analyzed with nick-translated pHBCκ (this is the same autoradiogram shown in Figure 5). pHBCκ specifically detects coding joints (in addition to non-rearranged fragments). The autoradiogram was over-exposed to accentuate the "sub-molar smear" (see text).

Bottom Right: PD subclones analyzed with the probe pJ3/2. pJ3/2 by this analysis detects a subset of both coding joints and reciprocal joints in the PD cell line. In some cases secondary subclone DNA was substituted for the corresponding primary subclone DNA. This was done to maximize the possibility of detecting bands corresponding to reciprocal joints (top right), by taking advantage of the increased homogeneity of secondary subclone populations. With one exception, (PD31-5), the secondary subclones chosen had the same reciprocal joints (though not all of the same coding joints) as were detected in the parental primary subclone.


L PD 31 32 33 34 35 36 37 38 39 40 41 42 44
14.3kb
6.8kb
5.5kb
4.0kb
R H H Jk H Ck B
RBJCk


L 31 32 33 34 35 36 37 38 39 40 41 42 44
RH H Jk H Ck B
RH5'


L 31 32 33 34 35 36 37 38 39 40 41 42 44
14.3kb
6.8kb
5.5kb
4.0kb
RH H Jk H Ck B
HBCk


L 31-5 32 33 34-4 35 36-5 37 38 39 40-6 41 42-22 44
RH H Jk H Ck B
J2/3

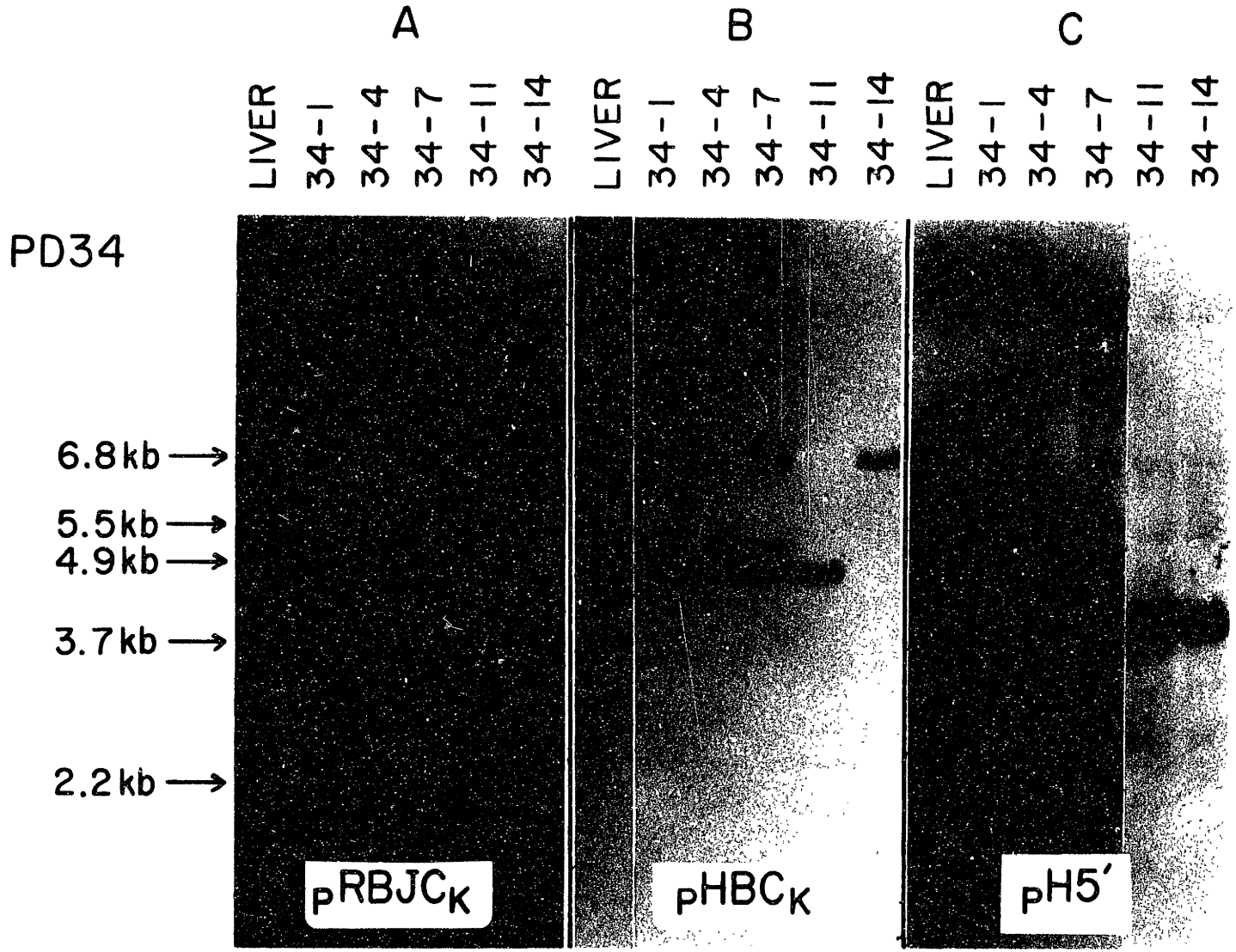
Figure 9

Coding joints can vary independently of reciprocal joints and vice versa.

A. Southern blot analysis of PD34 subclones using the probe pRBJCκ (see Figure 4). pRBJCκ hybridizes to both reciprocal joint- and coding joint-containing fragments. DNA samples were digested with Eco RI and Bam HI.

B. The same series of subclones, analyzed as above, except that the probe pHBCκ, (specific for rearranged coding joints) was used.

C. The same series of subclone DNA samples hybridized to pH5' which, like pRH5' (Figures 4 and 8), will allow specific detection of reciprocal joints.


A
B
C
LIVER
34-1
34-4
34-7
34-11
34-14
PD34
6.8kb
5.5kb
4.9kb
3.7kb
2.2kb
pRBJCK
pHBCK
pH5'

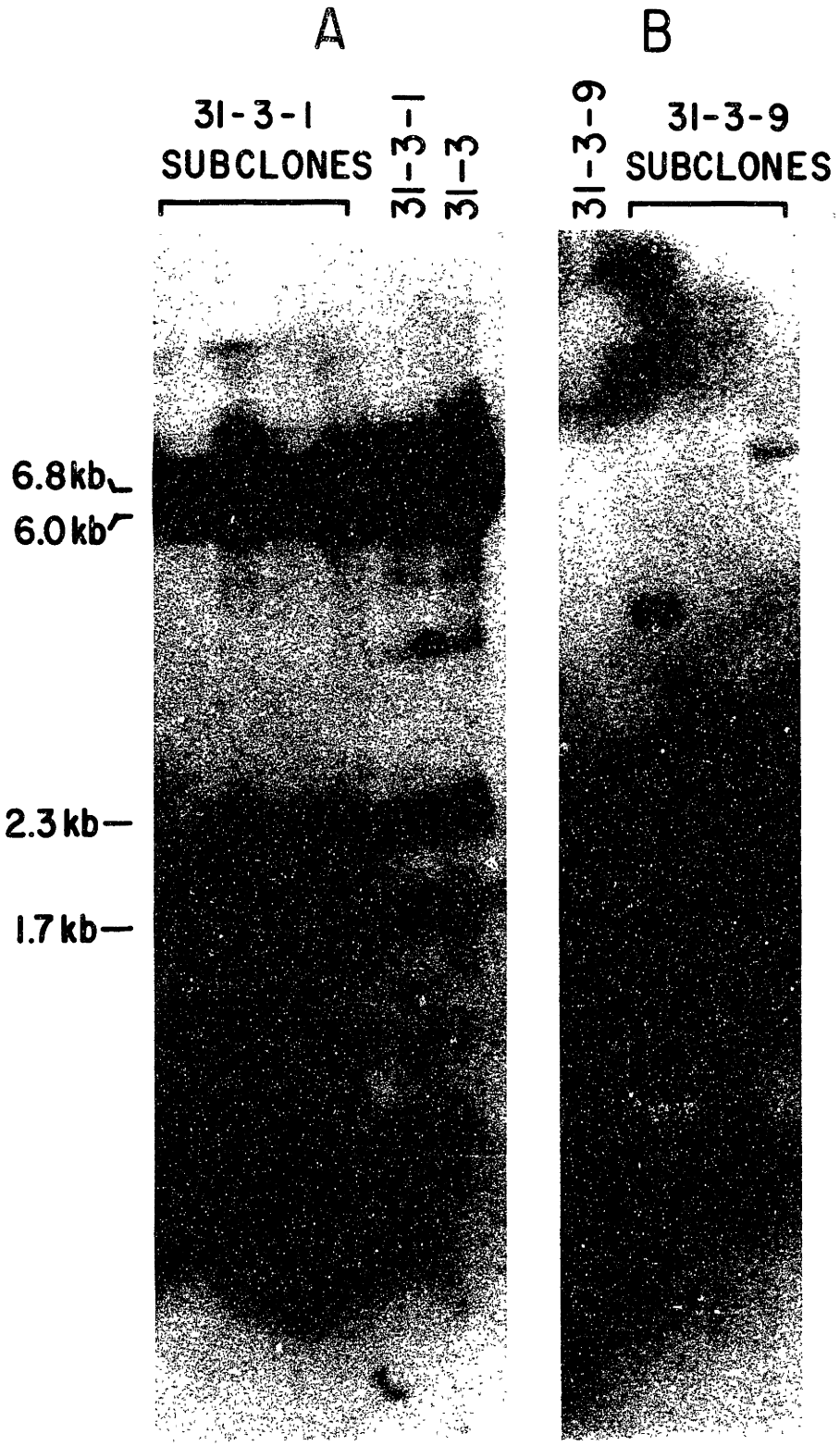
the 5' RH probe) was observable (Lewis et al 1982). My conclusion was based upon the data shown in Figure 10 (left). The single-cell origin and gene structure in PD31-3-1 was established by virtue of its derivation as a subclone of the apparently identical parental line, PD31-3. When PD31-3-1 was subcloned further, one of two reciprocal joints was apparently lost as a secondary event. More recently another explanation for the pattern shown in Figure 10 (left) has emerged, based upon the rearrangement observed to occur in an introduced substrate. It now appears probable that the two low molecular weight bands were not actually allelic reciprocal joints, even though they persisted together after subcloning (eg., in PD31-3, PD31-3-1, and in some PD31-3-1 subclones). More likely the two bands have a highly reproducible precursor-product relationship, and the PD31-3 sublineage has only a single 5'RH-hybridizing reciprocal joint, not two. The analysis of the PD31-3 lineage involved seven subcloning experiments, and reinspection of the data indicated that although the lower molecular weight band could be isolated in some PD31-3 derivatives in the absence of the higher molecular weight band, the reverse configuration was never observed. The higher molecular weight band was always accompanied by the lower molecular weight band, albeit at times the signal could be very faint. The probable nature of this reproducible recombination event will be discussed in the final chapter. At present, I have no direct evidence that reciprocal joints can become deleted as a secondary event.

The data shown in Figures 9 and 10 (left) demonstrate that reciprocal joints are unstable entities, without resolving the issue of whether this instability represents gross deletions, or instead secondary alterations. A direct demonstration of an alteration in which secondary rearrangement can

Figure 10

Reciprocal joints can undergo secondary recombination.

DNA samples were digested with Eco RI and Bam HI and probed with either pRBJCκ (left) or pRH5'(right). pRH 5' was used on the right to specifically detect reciprocal joints. The 2.3 kb and the 1.7 kb bands in the left-most panel hybridize to pRH5' as well (not shown). By analysis with pRBJCκ (not shown) PD31-3-9 and its derivatives exhibited the same 6.8 kb and 6.0 kb coding joints as the other PD31-3 clones shown here (not shown).


A
B
31-3-1
SUBCLONES
31-3-1
31-3
31-3-9
31-3-9
SUBCLONES
6.8kb
6.0kb
2.3kb
1.7kb

be distinguished from deletion is provided by the PD31-3-9 subclones. PD31-3-9 is a PD31-3 derivative that had a pattern of hybridization like that of PD31-3, (except that the intensity of hybridization of the higher molecular weight reciprocal joint band was greater than that of the lower). Figure 10 (right) shows three PD31-3-9 subclones all of which had the same coding joint structure as the parental PD31-3-9 line (not shown). The probe used for this analysis was 5' RH, and it is evident that two of the subclones exhibit fragments that were not detected in either PD31-3, or PD31-3-9. These novel bands were generated by a secondary alteration in the structure of a pre-existing 5'RH-hybridising fragment.

## FREQUENCY OF REARRANGEMENT IN PD CELLS

To get an accurate measurement of the rate at which recombination occurs, it is necessary to measure recombination in single cell progeny, in the absence of selection. To attempt such a measurement in the present system, a single cell must be grown for a known number of generations, and then the number of rearrangements that have occurred in that time must be measured. This measurement might be accomplished by densitometic measurement of the intensity of the germ-line signal in DNA prepared from the cell population, or by analyzing the variabilitylty in the population through a second round of single cell cloning/southern blot analysis. Either of these approaches requires several assumptions (discussed below), but in principle there is no more direct method by which to measure gene rearrangement. Indirect methods (such as inferring the presence of rearrangement on the basis of protein production) require a substantial

amount of preliminary characterization of the system to establish their validity.

Densitometric analyses is probably the preferable method, however I have taken the second approach mentioned above instead because the data were already available. I have selected a set of experiments for which the records are most complete. One drawback is that the experiments upon which these calculations are based were not designed for the purpose--too few cells were analysed, and in most cases the signal attributed to an unrearranged allele was not rigorously confirmed as such by using multiple digests and additional diagnostic probes. These data however seem to represent a typical set, so the expectation is that if an unrearranged allele was misdiagnosed, the error introduced is not critical. The calculations are offered as rough estimates only.

To make the calculation, it is necessary to assume that 1) rearrangement is infrequent enough that the southern blots of DNA made from cells on the average of 14 days after single cell cloning reflect the gene structures of those single cells. 2) the potential to rearrange does not change over the time during which the cells are growing in culture 3) the cells will grow at the same rate whether or not they are rearranged 4) the rate of rearrangement applies to each chromosome in a cell independently 5) the error introduced by subculturing only a fraction of the population when the cells are split is not significant, and 6) rearrangement occurs before, not during cell division.

Scoring the number of recombination events on the basis of the number of rearranged chromosomes in a cell population is not an accurate reflection of the recombination frequency because rearranged alleles can rearrange a second time (see chapter two). Also the same rates of recombination may not

apply to primary and secondary recombination. Therefore it is better to calculate the frequency of rearrangment from the number of cells containing no rearrangements at all. An inductive approach to the derivation of an expression relating the fraction of unrearranged cells in an expanded population to the recombination frequency follows.

If P is the frequency of rearrangement per chromosome per cell division, then 1-P is the frequency of non-rearrangement per chromosome per cell division. $(1-P)^2$ is the frequency of non-rearrangement of two chromosomes per cell division. Since unrearranged cells come from other unrearranged cells, the number of unrearranged cells in a culture is equal to $(1-P)^2$ multiplied by the number of unrearranged cells that existed previously. If $X_{N-1}$ equals the number of unrearranged cells at generation N-1, then at generation N, $X_N = 2X_{N-1}(1-P)^2$. The fraction of cells that are unrearranged at generation N, (which is what is measured) is given by:

$$\frac{2X_{N-1}(1-P)^2}{2^N}$$

When N = 1, this is $\frac{2(1-P)^2}{2}$; when N = 2, this is $\frac{2\cdot2(1-P)^2(1-P)^2}{2\cdot2}$;

when N = 3, this is $\frac{2\cdot2\cdot2(1-P)^2(1-P)^2(1-P)^2}{2\cdot2\cdot2}$; and so on. The expression for the unrearranged fraction of cells in a population after N cell divisions of a single unrearranged cell is $(1-P)^{2N}$. Table II gives the measured values for the number of cells in "embryonic" configuration after a known number of days' expansion from a single cell for several subcloning experiments. In the last column of the table is given the probability of recombination per

1 Number of days that elapsed since the day the subclone was itself a single cell and the day it was recloned.

2 Numbers of days that the cell line was in culture since PD was derived (this is not continuous, at least one freezing and thawing occurred in each generation of subcloning. In the case of PDA-1-1, the line was under mycophenolic acid selection for some of the period during growth in culture.)

3 For the purpose of comparison, this column shows the number of days, total that elapsed between the day that the particular subline was assayed, and the day that the parent PD31 subclone was derived. As above, these periods were interrupted by freezing and thawing of the cells.

4 Number of cells that were unrearranged/total number assayed. In the case of PD 31-3 and its derivatives, the unrearranged cells were cells that had not altered either of their rearranged chromosomes relative to the parental configuration. In the case of PDA-1-1 the number likewise refers to cells in the same configuration as their progenitor, however a single-copy substrate configuration was assayed, not the endogenous genes.

5 Frequency was calculated as described in the text, based on the percentage of unrearranged cells scored.

6 This number was calculated by pooling the data for the PD31 subclones after calculating a weighted average for the number of days in culture. The calculations for each subclone separately is shown in parentheses. The data were pooled because there is no evidence that subclones become permanently stable (or conversly display elevated recombination frequencies) in this sub-lineage. For example, a line such as PD 31-3-1 descended from PD31-3 without rearrangement during approximately 68 cell divisions, but then its derivatives rearranged at some point during the next 58 cell divisions. Conversly, PD31-3-4 had rearranged during the 68 cell divisions that occurred since the derivation of PD31-3, but then didn't generate more recombinants than its siblings over the next 52 cell divisions. Note that even without pooling the data, the probability of rearrangement in each of the two subclones for which recombination was measured is still lower than that associated with primary recombinations.

7 The probability of rearrangement for a single copy substrate was calculated from the observed frequency = $(1-P)^N$.

TABLE II

| PD SUB-LINE | #DAYS ASSAY[1] | #DAYS SINCE PD[2] | #DAYS SINCE PD31[3] | #CELLS UNCHANGED[4] | FREQUENCY[5] |
|---|---|---|---|---|---|
| | | PRIMARY RECOMBINATION FREQUENCIES | | | |
| PD | ≈42 | | | 3/13 | 0.009 |
| PD31 | 22 | ≈64 | | 1/10 | 0.025 |
| PD31 | 55 | ≈97 | | 1/11 | 0.011 |
| PD31-4 | 27 | ≈91 | 49 | 3/10 | 0.011 |
| | | SECONDARY RECOMBINATION FREQUENCIES | | | |
| PD31-3 | 34 | ≈98 | 56 | 13/15 | 0.0010 |
| PD31-3-1 | 29 | ≈127 | 85 | 6/8 | 0.0008[6] (0.0025) |
| -4 | 26 | ≈124 | 82 | 13/13 | (<0.023) |
| -7 | 19 | ≈117 | 75 | 6/8 | (0.0038) |
| -9 | 22 | ≈120 | 78 | 9/9 | (<0.025) |
| -10 | 26 | ≈124 | 82 | 3/3 | (<0.010) |
| -13 | 26 | ≈120 | 78 | 9/9 | (<0.020) |
| | | FREQUENCY OF RECOMBINATION OF INTRODUCED SEQUENCES | | | |
| PDA-1-1 | ≈35 | ≈275 | ≈211 | 12/15 | 0.002[7] |

chromosome per generation as calculated using the above expression (estimating two doublings per day).

I have shown the raw data because the error in these measurements can be expected to be large due to the sample size. Nevertheless, the calculations seem surprisingly consistent. The rates of primary recombination vary, but all of the four measurements put the value near 1% per chromosome per cell division. Interestingly, when the rate of secondary recombination was calculated in the same manner as the rate of primary recombination, (by scoring how many cells derived from a cell bearing dual kappa gene rearrangements rearranged neither allele after a known number of days in culture), the frequency was about ten fold lower. This does not seem to be an effect of the length of time that a cell line was growing in culture as can be seen by comparing PD31-3 (which undergoes secondary rearrangement) to it's sibling, PD31-4 (which undergoes primary recombination). PD31-3 was unrearranged in 13 of 15 cells analyzed whereas P31-4 was unrearranged in only 3 out of 10.

Data from a second generation of PD 31-3 subclones was pooled as explained in the legend to the table, and the probability of secondary rearrangement calculated on the basis of the analysis of 50 other subclones was similar, 0.08%. (Note that the cells that had rearranged relative to their siblings, PD31-3-4 and PD31-3-13 (Figure 7) were no more likely to rearrange again, even though they contained a "rearrangeable allele": therefore pooling the data seems justified. The frequency of rearrangement of the PD31-3-1 and PD31-3-7 subclones, taken as individual values is still less than the frequency of primary rearrangement). It would be difficult to improve upon these measurements because secondary recombination rates may be different for different alleles and/or different subclone lineages. The

analysis here compares recombination rates in the same lineage among cells that have been in culture a comparable period of time, and also includes independent measurements of the secondary rearrangement rate of the same allele.

The frequency of recombination of introduced sequences (see next chapter) was measured in one cell line, PDA-1 which also is a PD31 derivative. In PDA-1, the frequency of secondary rearrangement of the introduced substrate measured under non-selective conditions was 0.2%. It should be noted that the parent line PDA was initially selected for the ability to recombine, so that this result does not provide information as to whether or not a gene at a new chromosomal position can rearrange as well as one at the proper site. It is interesting nevertheless that a cell that has been in culture for approximately 7 months still has the ability to actively rearrange, providing some assurance that the PD system may be useful for experiments that might require prolonged selection, or several consecutive manipulations. Further, this result demonstrates that it is possible to measure, under non-selective conditions, an inversional rearrangement rate that is comparable to the recombination rates of endogenous genes (this will be discussed in more detail in the next chapter).

To summarize, primary recombination frequencies are roughly 1 or 2% per chromosome per cell division in the PD cell line. Secondary recombinations occur with an approximately ten-fold reduction in rate. It is curious that secondary recombinations occur more slowly, as there is every indication that secondary coding joint recombinations, like primary recombinations, are $V_K$-to-$J_K$ joining events (see next section). Either some positive or negative regulator of rearrangement is affected when a chromosome

recombines, or the long-range topography of the locus is less favorable after becoming altered by an initial recombination event.

One conclusion that is suggested by these data is that the variants found in the PD31-3 subclone series (Figure 7) as well as those in the PD34 subclone series (Figure 10) are probably in fact derived from pre-existing rearrangements because the frequency of primary recombination is not so high that an unrearranged allele present in the original PD31-3 or PD34 cells would have escaped detection.

## GENE STRUCTURE AND PROTEIN PRODUCTION OF SECONDARY RECOMBINANTS

When a rearranging cell line has been recloned to the extent that secondary kappa gene recombinations can be scored, the patterns of rearrangements are simple enough that they can (usually) be resolved into an interpretable series of precursors and products. It is also theoretically possible to take advantage of a panel of related variants to quickly assign protein products to specific alleles. The protein production in cells of the PD34 and the PD31-3 sub-lineages were studied with this end in mind.

### Rearrangement in PD does not discontinue with the acquisition of a functionally assembled gene

Given the assumption that to a first approximation rearrangement will generate a random set of alleles with variant, random restriction sites, the assortment of coding joints found among the PD34 subclones (Figure 9B) can not all have been independently derived from a completely "germ-line" precursor cell. One of the configurations among the subclones is more

"ancestral" than at least some of the others. Although it is not possible to unambiguously define the lineage on the basis of final structures, nevertheless it is possible to argue that the "4.9kb" allele must have been present (in at least one instance) in a PD34 cell when it underwent secondary rearrangement. To be more specific, the "4.9" kb allele must have been formed either before or after the acquisition of the 6.8 kb rearrangement. If before; then one might conservatively suppose that the 6.8 kb allele, as well as the 5.5 kb allele and the deleted allele were independent primary recombinations of an embryonic gene. However, in that case, the absence of the 4.9 kb allele in PD34-14 can only be explained as the secondary deletion of a pre-existing rearrangement. If the other alternative pertained, (that is that the 4.9 kb allele were acquired after the 6.8 kb allele), then one must postulate that the 5.5 kb allele in PD34-1 must have originated as a secondary alteration of the pre-existing 6.8 kb allele. Either way, it is evident that 1) a rearranged band in the PD34 series must have been secondarily altered and 2) the secondary alteration occurred in a cell that contained the 4.9 kb allele.

At least once in the PD34 sublineage, a pre-existing rearrangement (as opposed to an embryonic allele) underwent a gross deletion resulting in the loss of Cκ hybridization. Deletions of Cκ are frequently observed in λ-producing tumors (Hieter et al 1981, Coleclough et al 1981). These deletions are very likely due to an aberrant form of essentially the same process that fuses V to J but which involves a site 3' to the constant region (Durdick et al 1984).

It has been proposed that during normal B cell differentiation, the protein templated by a productively rearranged alelle negatively regulates the recombination process itself (Rose et al 1976, Alt et al 1980). Because

it is likely that both types of secondary alterations in PD34 subclones, that is, gross deletions and further rearrangements, are manifestations of the continued activity of the V-to-J recombinase in these lines, it is interesting to ask whether or not any of the alleles in the PD34 series are productively rearranged. The "4.9 kb" allele present in PD 34-1,-4,-7 and -11 can be identified as a productively rearranged allele because all of these subclones (but not PD34-14) can be demonstrated to manufacture an immunoprecipitable κ protein after LPS induction (not shown). Thus, according to the arguments outlined in the preceeding paragraph, a functionally assembled allele, the 4.9 kb allele, was present in PD34 prior to secondary alteration events. Apparently the assembly of a "good" gene is not sufficient to shut off recombination in PD. This observation does not necessarily contradict models based on feedback-regulation mediated by the protein product of a rearranged kappa gene however, because PD will not manufacture detectable levels of light chain protein without LPS induction.

Both alleles in PD can become functionally rearranged

Allelic exclusion refers to a phenomenon defined in rabbits, the basic observation being that a B cell in an F1 rabbit will only express one each of its of heavy and light chains alleles. (Pernis et al 1965, Weiler 1965, Pernis et al 1970). The frequency with which normal murine B cells are kappa "double-producers" has not been measured, nor have myelomas been systematically screened for this property. In the rat, where allotypic markers are available, this figure is <8% (Tsukamato et al 1984). These measurements are difficult due to the problem of distinguishing the immunoglobulins manufactured by a normal B cell from cytophilic antibody.

Therefore the actual frequency of double producer cells may be much lower. If, as has been proposed, (Rose et al 1976, Alt et al 1981), allelic exclusion of light chain expression is the result of the regulated shut down of gene rearrangement after productive gene assembly, and if PD shows no evidence for this, then PD should not exhibit allelic exclusion.

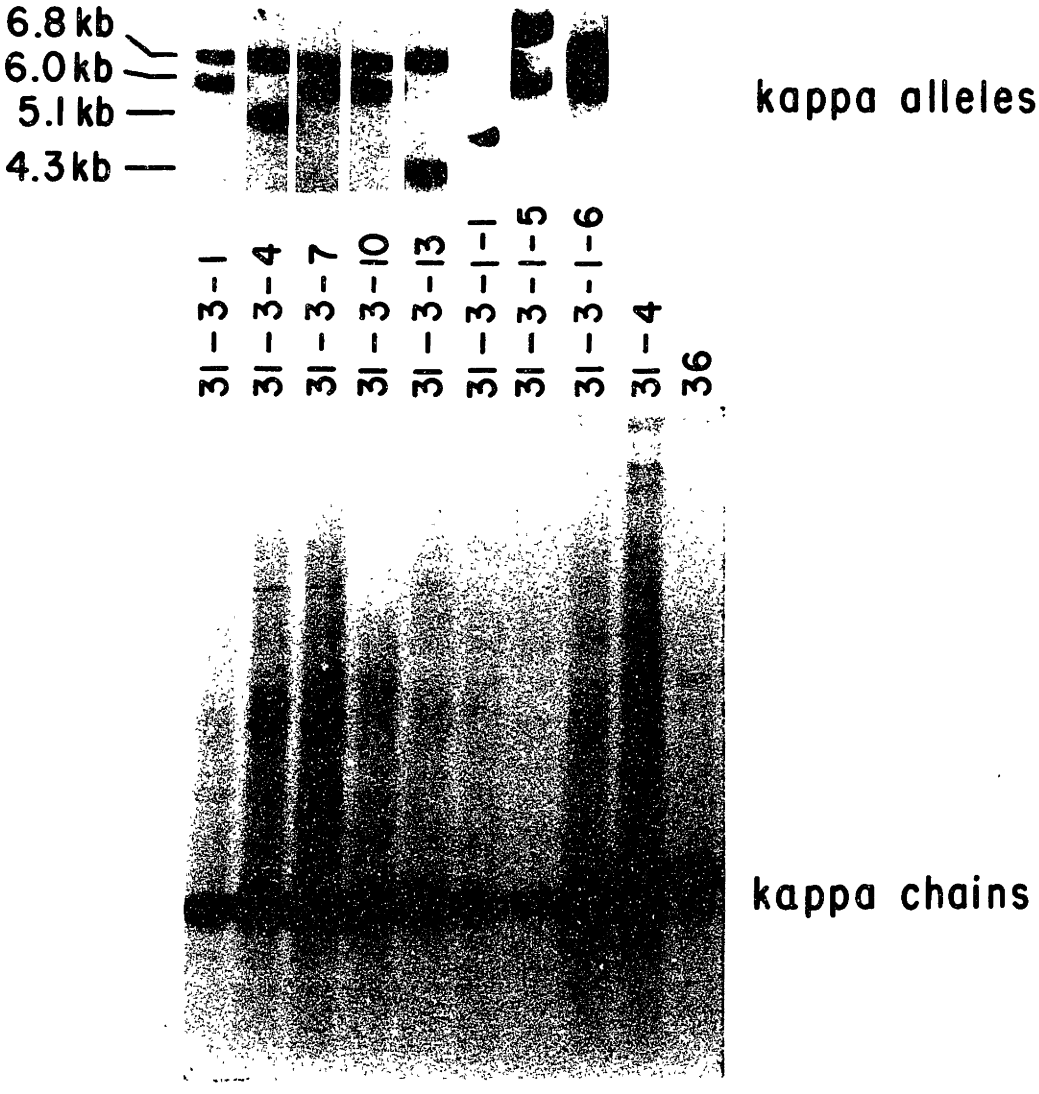
In the first subcloning of PD31-3, the PD31-3-4 and PD31-3-13 derivatives both contained variant coding joints that were not found in any other sibling isolate. The two bands in each segregated together upon subcloning (not shown). In each subclone both alleles appeared to be functional templates for kappa immunoglobulin in an apparent exception to the rule of allelic exclusion. SDS-polyacrylamide gel electrophoresis of metabolically labelled extracts of PD31-3-4 and PD31-3-13 revealed two species that were immunoprecipitable with an anti-mouse kappa antiserum (Figure 11). The other PD31-3 derivatives that were analyzed for kappa protein production (PD31-4, -7, -9, and 10) all had the same coding joints as the uncloned PD31-3 cell line and all displayed a single kappa light chain (Figure 11).

To demonstrate that the two protein species detected in PD31-3-4 and PD31-3-13 were indeed allelic products a second group of PD31-3 derivatives (the quaternary subclones shown in Figure 7) was analyzed in which some of the members had altered the other homologue from the one that was variant in PD31-3-4 and -13. The expectation was that the 6.8 kb coding joint must be the template for the common protein band found in all of the PD31-3 derivatives in the first group, and that in the second group, extracts of the lines that lacked the 6.8 kb allele (PD31-3-1-1 and PD31-3-1-5) would not contain the common protein. Unfortunately the analysis was not

Figure 11

Kappa gene structure and protein production in the PD31-3 derivatives.

The top panel is an excised portion of the autoradiogram in Figure 8 showing the stuctures of the coding-joint containing fragments in this series. The lower panel is an autoradiogram of the kappa chain species detected on a 10% SDS polyacrylamide gel. Cells were cultured in the presence of 50 g/ml bacterial lipopolysaccharide for 48 hours before they were metabolically labeled with $^{35}$S-methionine. Labeled lysates were incubated with rabbit anti-mouse kappa serum and immune complexes were precipitated with Staph A. Samples were run under reducing and denaturing conditions (see Materials and Methods).


6.8 kb
6.0 kb
5.1 kb
4.3 kb
kappa alleles
31-3-1
31-3-4
31-3-7
31-3-10
31-3-13
31-3-1-1
31-3-1-5
31-3-1-6
31-4
36
kappa chains

definitive because PD31-3-1-1 and PD31-3-1-5 still exhibited a protein in common with PD31-3-4 and PD31-3-13.

Obviously the only alternative to the conclusion that the two protein species in PD31-3-4 and PD31-3-13 are allelic products is that the two species are generated from a single allele. Pursuing this logic, either the gene structure/product correlation is coincidence (in which case the variants are produced by the 6.8 kb allele and the protein production in the second panel is unexplained), or all of the proteins detected are made by the same gene (the 6.0 kb allele) that has been altered in both PD31-3-4 and PD31-3-13. In these two lines then, a rearrangement had to have occurred that both changed the restriction site(s) several kb away from the coding sequences, and changed the coding sequences themselves. The alleles would have to have been altered in a fairly subtle way so that the original product is still produced but in addition a new species is made. This is hard to envision.

A far more simple explanation is that the precursor, PD31-3, was itself a double producer, and that both proteins that it manufactured comigrated in the gel system used. Thus the variants in the second panel retained one allele that could still produce a band in common with all the others. Co-migration of unrelated kappa proteins is not infrequent (eg. Rice and Baltimore 1982) and it can be seen as well that a labeled, immunoprecipitated extract from PD31-4 cells (which is a heterogeneous population although it appears embryonic by southern blot analysis) exhibit a faint but discernible signal in the same region of the gel as the band associated with PD31-3. The signal in the PD 31-4 lane is probably not a background band because it is not observed in the subclone PD36.

To show that two alleles can be productively rearranged and expressed in PD, both alleles from PD31-3-4 were molecularly cloned. The analysis of these clones is incomplete. However, the DNA sequence of the 5.1 kb allele has confirmed it to be an in/frame V/Jκ5 junction (S.Lewis and D. Schatz unpublished).

Independent of these considerations, the PD31-3-1 derivative unambiguously demonstrates that in PD cells, rearrangement can occur even though a productively assembled allele already exists.

Further analysis is required to resolve the issue of whether or not PD 31-3 and its derivatives are double producers. It is worth pointing out that other rearranging A-MuLV transformants have been reported to manufacture two distinct protein species that persist upon subcloning (Zeigler et al 1984), so that the PD cell line is apparently not exceptional in this regard.

Independent of these considerations, the PD31-3-1 derivatives unambiguously demonstrate that in PD cells, rearrangement can occur even though a productively assembled allele already exists.

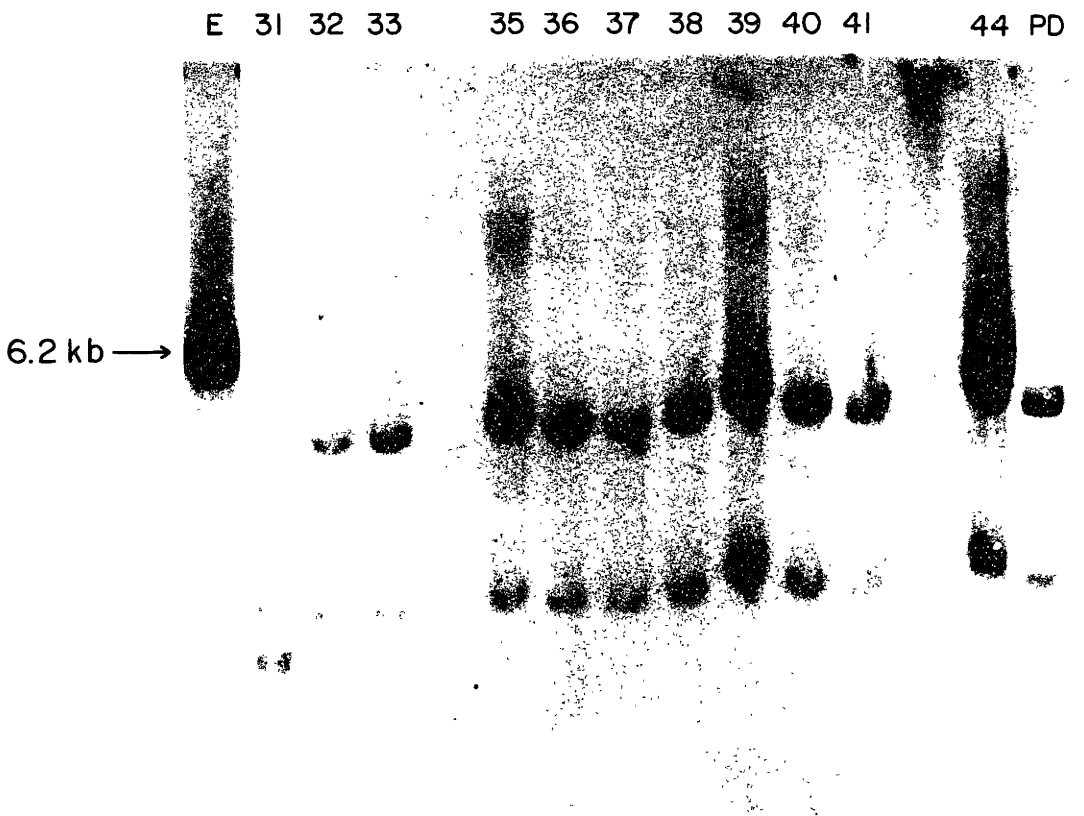
REARRANGEMENT AT μ AND λ.

A consistent finding among the PD derivatives is that cells in the process of actively rearranging their kappa genes do not display an equivalent degree of rearrangement at Jλ or Jh. The PD primary subclones were analyzed by southern blot for rearrangement at Jh by digestion of DNA samples with Eco R1. NIH/Swiss liver DNA was included as an unrearranged control and exhibited a single 6.2 kb band when hybridized to a probe specific for the Jh region (Alt et al 1981b). By this assay, all of the PD

Figure 12

Jh analysis of PD primary subclones.

Eco RI-digested DNA samples were subjected to southern blot analysis, using the pR1-Jh plasmid (Figure 4) as a probe. NIH liver DNA was included as an unrearranged control. (Subclone PD34 is not shown but using the same analysis exhibited the same Jh structure as PD32 --not shown).


E
31
32
33
35
36
37
38
39
40
41
44
PD
6.2 kb

subclones had rearranged both chromosomal copies of Jh, as had the uncloned PD cell line (Figure 12). With one exception, the subclones had identical Jh gene configurations; the variant, PD31, contained one Jh allele common to its sibling clones and one unique band. The similarity of the Jh structure among the PD subclones contrasted with the enormous variability at the Jκ locus (Figure 5), and indicated that all the variability at the κ locus must have been acquired after Jh reorganization. This was confirmed by analysis of the variant subclone PD31. All of the secondary subclones derived from PD31 had identical Jh structure (not shown), although, as indicated in Figure 6, these subclones had a variety of κ gene rearrangements. Thus even though a limited amount of heavy-chain gene reorganization occurred after the PD progenitor cell was transformed by A-MuLV, the Jh alleles did not rearrange during active κ gene rearrangement. The relative stability of the heavy chain locus is understandable in light of the recent report by De Pinho et al (1984) that PD has a V/D/J rearrangement of both heavy chain gene copies. In this circumstance there is obviously no opportunity for further recombinational assembly.

The nature of the heavy chain gene rearrangement detected in PD31 is not known. Although many of the primary subclones produced heavy chain protein, PD31 did not (Figure 3). It is possible that the rearrangement in PD could be an intron deletion of the type described by Alt et al (1982), or it could be related to the switch recombinations that have been documented to occur in some A-MuLV transformants (Akira et al 1983, De Pinho et al 1984).

The prior rearrangement of the heavy chain locus provides evidence that within a clone of cells, the onset of gene rearrangement at different loci may be asynchronous. This evidence does not resolve a more basic issue

which is whether μ genes are regulated so that they initiate and terminate recombination before κ genes or whether the asynchrony is the result of inherently different recombination rates at simultaneously induced loci.

To analyze the λ gene structures in PD and its various subclones, Eco R1 digested DNA was probed by hybridization to PABλ2. This probe will detect most rearrangements of Vλ1,, Vλ2, Cλ1, Cλ2 and Cλ3 as a result of a variety of cross-hybridizations (Miller et al 1981). There was no example of λ gene rearrangement by this assay among the primary subclones, nor in twelve PD34, ten PD31, and seven PD32 and PD39 derivatives, nor among ten PD31-3 tertiary and quaternary subclones assayed. Because deletions of the kappa constant region are frequently observed in λ-producing hymbridomas and leukemias (Heiter et al 1981, Coleclough 1981), it would have been interesting to determine whether a PD subclone such as PD34-28 (not shown) that had deleted both kappa constant regions would be more likely to rearrange its λ locus. This experiment hasn't been done, however a similar analysis was carried out on an 18-81 derivative (see below) and no λ recombinants were found. Although it is clear that variants at the λ locus do not accumulate to any detectable level, once again it is not possible to say whether this is due to ordering of the gene rearrangements at the two light chain loci (with kappa preceding lambda), or to a different inherent probability of rearrangement at the two loci, (so that kappa rearrangement is more likely than lambda), or to a precommitment of the cell to rearrange kappa in preference to lambda.

SUMMARY

To summarize, the PD cell line undergoes frequent rearrangement of its kappa genes during growth in culture. Recombination yields two qualitatively different products which on the basis of their hybridization properites appear to be roughly reciprocally related. These are termed reciprocal joints and coding joints, and either junction can be further altered by secondary recombination events. Secondary rearrangements occur regardless of whether the coding joint encodes an in frame (immunoprecipitable) kappa protein or not.

Because reciprocal joints as well as coding joints can undergo secondary recombination events, non-reciprocity among coding joints and reciprocal joints in myelomas may well be the result of secondary rearrangement of one, or both, initially-reciprocal products of kappa gene rearrangement.

There is strong circumstantial evidence that PD subclones can contain two productively recombined alleles. In contrast to this apparent lack of allelic exclusion, recombination in PD appears to be excluded with regard to isotype. The heavy chain locus is stably recombined and the λ locus is embryonic (to the limits of the assay) in all subclones tested.

## OTHER REARRANGING A-MULV TRANSFORMANTS

To extend these observations four other cell lines provided by Dr. Naomi Rosenberg and Dr. Michael Boss were examined . The four cell lines had the common property that at some time in their culture history they had each been capable of producing a light chain protein and for each, either the initial transformant or an appropriate series of subclones were available. These lines were 2M3/M ( Clark and Rosenberg, 1980; Alt et al.,

1981), ABC1 (Boss Greaves and Teich, 1979), 18-8 (which has been studied extensively with regard to spontaneous heavy chain class switching, Alt et al 1980, 1982, Burrows et al 1981, 1983, Wabl et al 1984, and has also been examined for light chain production and rearrangement after fusion to a non-producing myeloma, Riley et al. 1981), and the line 204-1-7 (which has presented a variable kappa protein phenotype-Naomi Rosenberg unpublished, and which has been included in the studies of Alt et al 1981, 1984, and Yancopulos et al 1984). There are four other independent A-MuLV transformants (designated 298-8, 196-313, 196-22, 38B9) that, by the criteria of either presenting sub-molar bands upon southern blot analysis or subclonal heterogeneity may rearrange in culture as well, but these have not been included in the study.

2M3/M

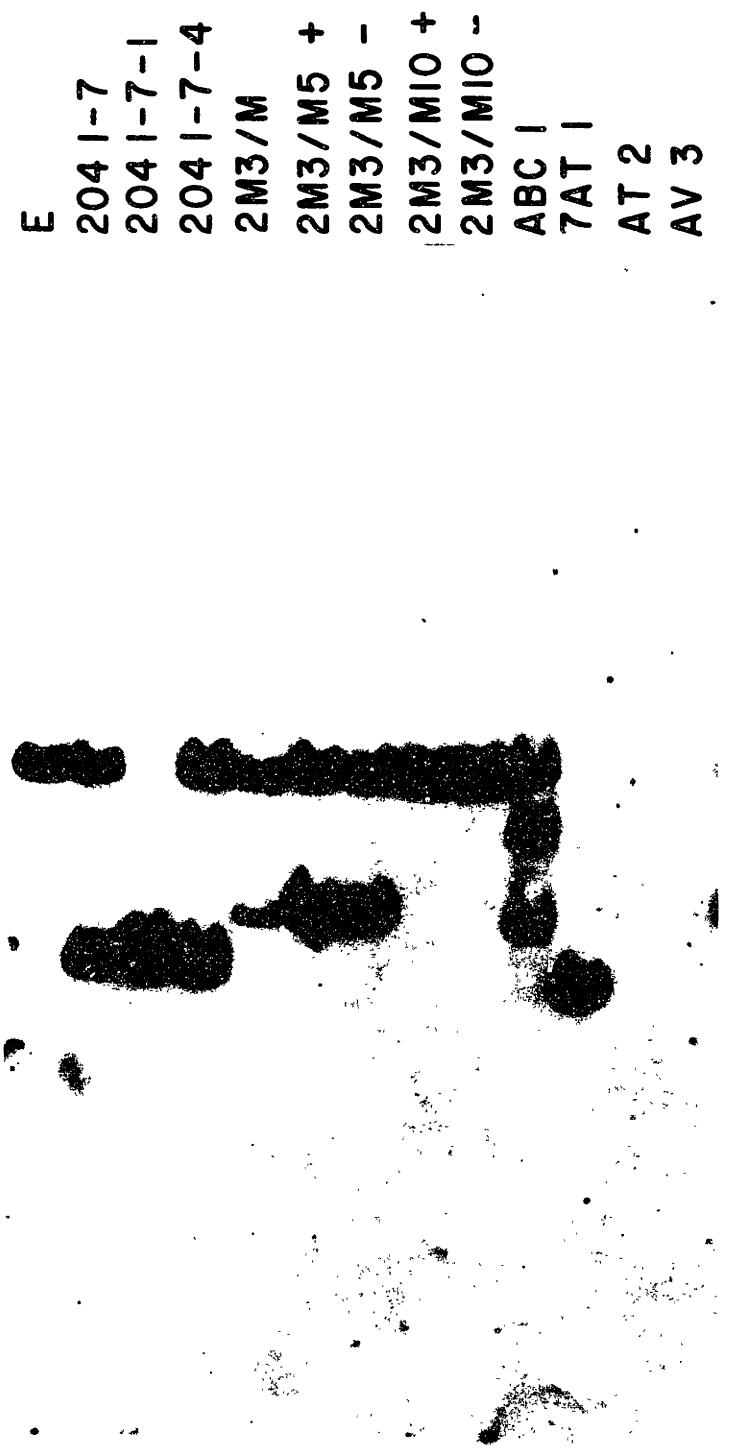
Data on 2M3/M and its subclones has been published (Clark and Rosenberg 1980, Alt et al. 1980). 2M3/M was derived by superinfection of an Abelson-transformed line, 2M3, with Moloney murined leukemia virus. Early in its passage it made detectable LPS-inducible kappa protein. 2M3/M is an exceptional type of A-MuLV transformant in that it manufactures light chain without any concommitant production of heavy chain. Fewer than 4% of all A-MuLV transformants have this phenotype (Alt et al 1981). On long term culture 2M3/M began to loose $\kappa$ expression and was at that point subcloned (Clark and Rosenberg 1980). Subclones 2M3/M5 and 2M3/M10 differ in that M5 could be induced to express kappa protein by LPS whereas M10 could not. M5 and M10 have both been proven to be subclones of 2M3/M (Alt et al. 1980). The effects of LPS on C$\kappa$ transcription in these subclones parallels its

effects on protein production. Detectable levels of stable Cκ transcripts as analyzed on Northern blots were LPS-induced only in 2M3/M5. To determine whether LPS had a detectable effect on the kappa gene structure and in general whether the potential to respond to LPS could be attributed to the genome structure of the subclones, DNA was prepared from cultures with and without LPS stimulation. Figure 13 shows that LPS had no evident effect on the kappa genes. On the basis of our studies with PD, a smear of lower molecular weight bands, and/or a decrease in the intensity of hybridization relative to un-treated cultures would have been expected, should LPS have induced heterogeneous kappa gene rearrangements in one culture and not another. While this experiment does not rule out the possibility that LPS might have some effect on gene rearrangement, there is no evidence for it here, nor is there any difference between the two subclones in this regard. On the other hand, this analysis showed that M5 contained a rearranged allele not present in M10. Since M10 displayed only a germ-line band (confirmed by Bam H1 digests and by hybridization to RH5'), the lack of transcription or kappa protein production can be simply explained as a lack of rearranged alleles. This experiment indicates that a pre-existing difference in the gene structure of two subclones can account for the differential response to LPS. Kappa protein production, when it is detected in A-MuLV transformants, has been reported to be unstable (Clark and Rosenberg 1980, Siden et al 1979, M. Boss personal communication); the protein is usually only detected upon LPS induction and often is lost gradually over time. Although no systematic examination of cultures before and after loss of LPS induction has been undertaken, as shown below, significant changes in the ratios of subpopulations of a culture can be observed after long term passage. The fact that LPS inducibility is often

Figure 13

Kappa gene structure in three independantly derived A-MuLV transformants

Southern blot of lines 204-1-7, 2M3/M, ABC 1 and their derivatives. DNA samples were digested with Eco RI and Bam HI and probed with nick-translated pHBCκ.


E
2041-7
2041-7-1
2041-7-4
2M3/M
2M3/M5 +
2M3/M5 -
2M3/M10 +
2M3/M10 -
ABC 1
7AT 1
AT 2
AV 3

lost may indicate that the subpopulation of cells in the culture that has the capacity to respond to LPS is at a slight growth disadvantage.

Similar to PD, the 2M3 subclones were variant only at the kappa locus; the two rearranged Jh hybridizing bands co-migrate (Alt et al. 1980) and the λ genes are embryonic (not shown).

18-8

18-8 is an in-vitro transformant of adult Balb/c bone marrow (Pratt et al, 1977). Figure 14 shows the kappa gene analysis of 18-8 and the primary subclones 18-81, 18-82, 18-84, 18-85 and 18-86. Included for comparison are three 18-81 populations (18-81A, 18-81B, 18-81C) that were maintained separately soon after derivation of the 18-81 cell line. 18-81A has been subcloned and two secondary subclones, 18-81A2 and 18-81A20, are also shown (these particular lines, their clonal origin and their heavy chain gene structure have been presented in Alt et al 1980, 1982; and De Pinho et al 1984).

The autoradiogram in Figure 14 has been deliberately overexposed to enhance submolar bands. It is evident that the parent 18-8 population must have undergone extensive deletion and rearrangement of its kappa genes, because no distinct bands were detected by hybridization of its DNA to probe HBCκ. Individual 18-8 subclones displayed rearranged bands some of which were productive (eg. kappa protein production by 18-84 and 18-86 is demonstrated in Rosenberg et al. 1979). The alterations in 18-81 secondary subclones, 18-81A2 and 18-81A20 must have occurred after the parent primary line (18-81) was derived.

Figure 14

Kappa gene structure in the A-MuLV transformant 18-8 and its derivatives.

Southern blot analysis was performed using pHRCκ as a hybridization probe on Eco RI plus Bam HI-digested cellular DNA.


LIVER
18-86
18-85
18-84
18-82
18-81A20
18-81A2
18-81A
18-81C
18-81B
18-8

The degree of "drift" in the cloned 18-81 cell line can be extreme as shown in Figure 14. 18-81 C for example was split from a culture of 18-81 and has, during passage, been overtaken by a subpopulation bearing a single rearranged Cκ allele and no germ line copies. 18-81 populations that look similar with probe HBCκ can be shown to be heterogeneous when the longer RBJCκ probe is used. 18-81A appears to have a germline configuration, however when probed with RBJCκ, another band, presumably a reciprocal joint,appears and is present in both subclonal derivatives, 18-81 and 18-81A20. This contrasts to reports from other laboratories with isolates of 18-81 that bear completely unrearranged kappa alleles (Burrows et al 1983).

These observations in themselves are not informative, however they emphasize the necessity for proving the identity of any two cultures of a line that has been maintained separately for any period of time, for they can differ to quite a significant extent, even without having been subcloned. In this context, applying the term "parental" to a culture of cells from which subclones were derived (eg Alt et al 1984) can be misleading, as there is no basis for the assumption that the gene structures observed in the subclone DNA bears any derivative relationship to the bands observed (as the major population) in the uncloned culture without confirming this through lineage (eg Lewis et al 1982, Sugiyama et al 1983, Zeigler et al 1984)) and/or structural (eg Desiderio 1984) analysis.

Riley et al (1981) reported induction of kappa gene expression (and presumably rearrangement) by fusion of a drug marked 18-81 subclone to a non-producing myeloma. Because the fusion process imposes a stringent selection on input cells, and because the 18-81 line spontaneously rearranges its kappa locus, the question of precursor/product relationships in this experiment made the results difficult to interpret.

The heavy chain gene structure of 18-81 has been intensively studied (Alt et al 1980, Burrows et al 1981, Alt et al 1982, Burrows et al 1983, Wabl et al 1984, De Pinho et al 1984). Burrows et al (1981) reported that their isolate of 18-81 apparently underwent class switching (mediated by a deletional rearrangement) before light chain gene rearrangement had occurred, in apparent contradiction to the view that class switch recombination is representative of an event that occurs exclusively in antigen-reactive, mature B cells and is governed by interactions with T cells (reviewed by Cebra et al 1983). A similar observation has been made by Akira et al (1983) who report that switch recombination in some of their A-MuLV transformants may even precede functional Vh exon formation. PD has been reported capable of undergoing class switching events as well (DePinho et al 1984). Based on these observations, either A-MuLV transformants do not accurately reflect processes that occur in normal developing B cells, or pre-B cells can in fact undergo switching events. Some support for the latter interpretation comes from a study of pre-B leukemias (Kubagawa et al 1983) and normal B cells (Cebra et al 1983), but the issue of whether isotype switching patterns are the result of a stochastic recombinational event that can occur at any point during B cell development or whether they represent a regulated response to antigen and cellular interactions is not resolved.

To examine λ gene structure in an 18-8 derivative, the line 18-81A2 was specifically chosen for subclonal analysis because 18-81A2 no longer has any detectable Cκ hybridization in its DNA. The high frequency of κ gene deletion in λ-producing cells, (Hieter et al 1981, Coleclough et al 1981, Korsmeyer et al 1982), indicated that there may be a correlation between

kappa gene deletion and the onset of lambda gene rearrangement. In thirteen 18-81A2 sucbclones, however, there was no example of λ gene rearrangement.

204-1-7

Lines 204-1-7, 204-1-7-1 and 204-1-7-4 are an initial transformant and two subclones, respectively. Clonality was confirmed by analysis of the restriction fragment containing Abelson-specific proviral sequences (not shown). 204-1-7 was LPS inducible for kappa protein when first tested, but like 2M3/M, lost this capability after long term culture. Subclone 7-1 is a constitutive kappa producer, 7-4 is not (nor can it be induced by LPS (N. Rosenberg personal communication).

Again, analysis of the kappa genes in these lines revealed variability at the kappa locus. 7-1 and 7-4 differ in that 7-1, the kappa protein producer, lacked a second Cκ copy found in 7-4 (figure 16). Both the 7-1-specific allele and a second, unrearranged allele were present in 7-4, the non-producer. Identity of the rearranged alleles in 7-1 and 7-4 was confirmed by Bam H1 digestion (not shown). In the case of transformant 204-1-7 then, the variability acquired during culture consisted of deletion of a Cκ allele. This provides another example where a secondary Cκ deletional event occurred in the presence of a productively rearranged allele: an allele in the common precursor to 1-7-1 and 1-7-4 must have been deleted in a cell that had already assembled one functional kappa gene. Although the "functionally assembled" allele is expressed in subclone 1-7-1, the same gene is silent in 1-7-4. The basis for the non-expression of the allele in 1-7-4 is unknown. No similar example of clonal variation in the expression of the same light chain rearrangement has been observed in any PD sublines.

ABC-1

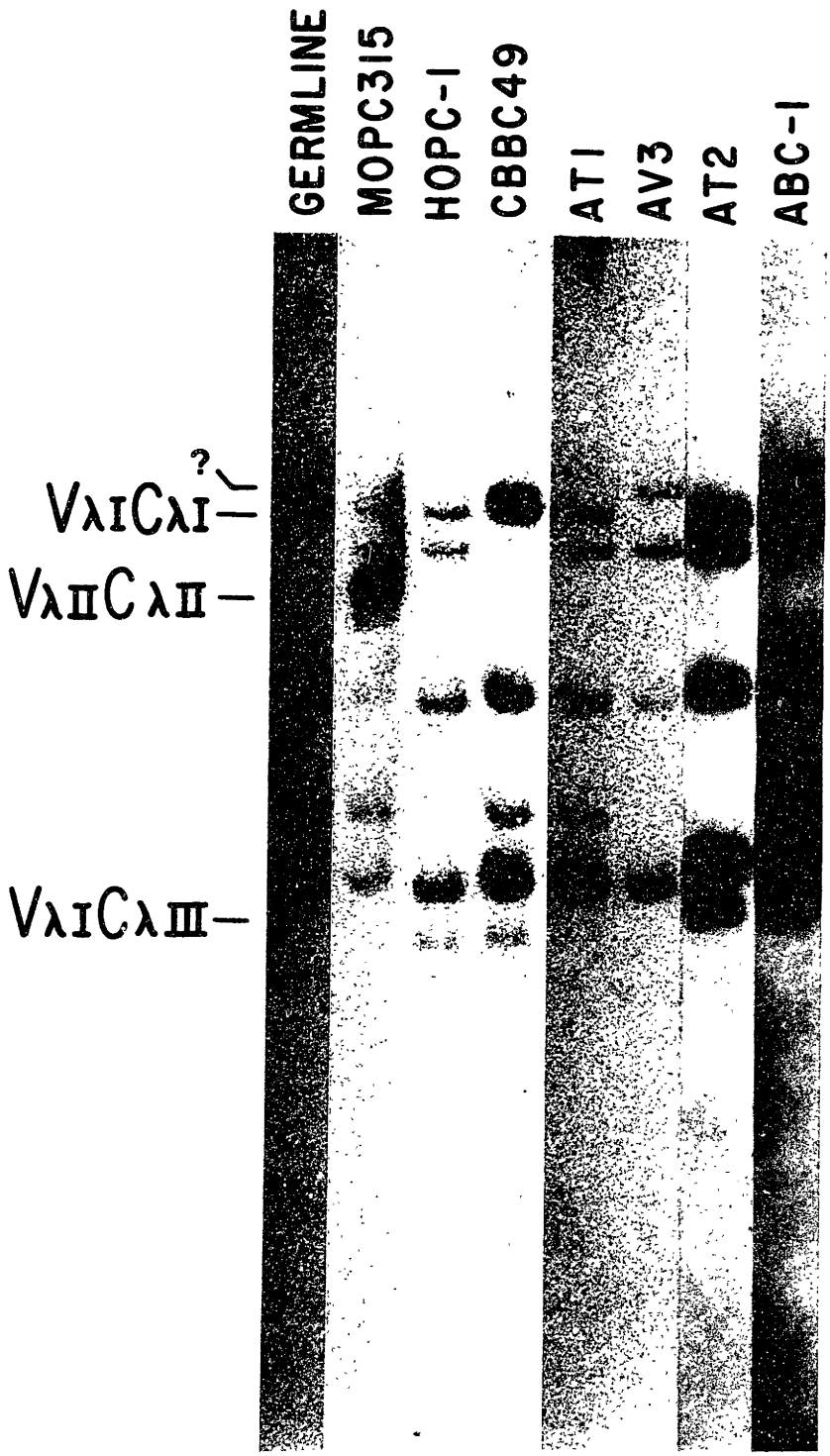
ABC-1, 7AT1, AT2 and AV3 are all derived from a single Abelson virus transformant. Clonality has been demonstrated by analysis of proviral sequences (S. Goff, unpublished). ABC-1 was isolated from in vitro transformation of a Balb/c bone marrow culture (Teich et al, 1979). ABC-1 is LPS inducible for μ heavy chain protein production, and at various times also produced a lambda light chain (M. Boss personal communication, F.Alt unpublished). AT1, AT2, and AV3 are all related, sublines of ABC-1, which were single-cell cloned after derivation.

In contrast to any other Abelson transformant, the λ locus in these lines had apparently undergone diverse rearrangements. No two lines had the same λ gene configuration, and lambda protein production was detectable in all three ABC-1 derivatives (F. Alt and M. Woo, personal communications). 7AT1, AT2, and AV3 each bore a rearranged band corresponding to the Vλ/Jλ1 rearrangements characteristically found in λ1 producing cell lines (assayed as described above using probe pABλ2), and in addition, AT2 exhibited a λ3-type rearrangement (corresponding to Vλ1/Jλ3). AV3 displayed a band that is similar in size to that representing a Vλ2-to-Jλ1 recombination as described by Reilly et al (1984). ABC-1 DNA showed a λ3-type rearrangement along with a complete set of unrearranged bands (Figure 17). This series of lines shows that it is possible for λ gene rearrangement to occur after transformation by Abelson virus. Because a λ3 type rearrangement was observed in uncloned ABC-1 as a detectable, prominent band, it is conceivable that λ3 type gene rearrangements occur readily and that lower expression of the λ3 subtype in the serum is not due to a lower frequency of λ3 recombination during differentiation (Blomberg and Tonegawa, 1982).

Figure 15

Lambda gene structure in ABC-1 sub-lines.

DNA samples were digested with Eco RI and analyzed using nick-translated pABλ2. This plasmid is a cDNA clone of the Vλ2/Jλ2 transcript in MOPC 315. By this analysis (Miller et al 1981), rearranged bands corresponding to Vλ1/Jλ1, Vλ1/Jλ3 and Vλ2/Jλ2 can be visualized. DNA samples from HOPC-1, CBPC49, and MOPC315 were included as markers for each type of rearrangement as indicated.


GERMLINE
MOPC315
HOPC-1
CBBC49
AT1
AV3
AT2
ABC-1
VλIICλII
VλICλIII

When the kappa genes in these lines were analyzed (Figure 13), they displayed variability . ABC-1 DNA contains three prominent Cκ rearrangements (and re-cloning ABC-1 revealed considerable heterogeneity of the kappa genes among subclones--not shown). AT2 has one Cκ allele, and neither 7AT1 nor AV3 contain any sequences hybridizing to probe HBCκ. No kappa protein production has been detected in ABC-1, or the three related sublines (M. Boss personal communication). Kappa protein production in the diverse ABC-1 subclones has not been tested.

SUMMARY

To summarize, these data establish that rearrangement of light chain loci can occur after transformation in five separately derived Abelson virus-infected lines. At least one clone from each of these five transformants can be shown capable of producing a light chain protein, bearing out the conclusion that the rearrangements occurring in culture are authentic V-to-J recombination events. There are differences among these lines with regard to heavy chain production, and whether or not LPS induction is required to visualize light chain protein.

It has been suggested that light chain gene rearrangement is negatively regulated by a feed back mechanism in which the regulatory signal is specifically the successful association of heavy and light chain proteins to form an IgM molecule (Alt et al 1981). No conclusion can be made as to the validity of this model based on the phenotypes of the cell lines in this survey. However it may be significant that one line (2M3/M) cannot make heavy chain, another displays variability in expression of light chains (204-1-7), and others (18-81, PD and ABC-1) do not make detectable light

chain in the absence of LPS. Studies to directly test whether, upon LPS induction, subclones of PD can assemble co-precipitable heavy and light chain products showed that relative to a control line, 70Z/3 (which manufactures a comparable amount of μ chain, is LPS inducible for light chain, and is known to assemble IgM (Paige et al 1978,1981)) there was little association of the two chains in PD (not shown).

A common property of these lines is the orderly pattern with which the light chain loci are actively rearranged. In total approximately seventy subclones of various kappa-expressing/kappa-rearranging lines were assayed for λ gene rearrangements and none was observed (in these experiments a minimum of approximately 45 independent kappa gene recombination events could be scored). In contrast, ABC-1 exhibited extensive alterations of both the kappa and the lambda locus. This pattern is consistent with the proposal that, during differentiation, kappa light chain gene rearrangement precedes lambda (Maki 1980, Alt 1981, Heiter 1981, Korsmeyer 1981, 1983). However it is just as consistent with a probabilistic model (Perry 1981, Coleclough 1981) in which the order is not a reflection of differential regulation but rather of the frequencies with which the two loci undergo rearrangement. Because the A-MuLV transformants described here do apparently exhibit "isotype exclusion" they may provide a good model system for pursuing this question further.

The frequency with which cell lines derived by in vitro bone marrow infection produce light chain protein is low, less than 10% (Alt et al 1981). Light chain positive cells are rarely found in fetal liver transformants (Alt et al 1981, Denis et al 1984). The five lines described here were derived using a variety of transformation protocols; although they were all recovered from in vitro infections of adult bone marrow, either a

P160 or P120 Abelson stock prepared with either Moloney-MuLV or Friend MuLV helper virus was used. In addition to the lines described here, Whitlock et al (1983) have described A-MuLV transformants that similarly rearrange that were isolated after infection of long term bone marrow cultures rather than bone marrow itself. Because the only alterations in the protocols that seem to affect the frequency with which light chain gene rearranging cell lines are transformed by Ableson virus all have to do with the type of tissue infected or with manipulation of the culture, the limiting factor may be the number of target cells available (Denis et al 1984). It is unknown whether these targets are unusual pre-B cells or simply a fleeting (and therefore low-level) normal intermediate in B cell differentiation.

Although the rearranging Abelson transformants appear capable of "differentiation" in that upon passage derivatives can be isolated which appear to have acquired new properties (such as de novo light chain protein production or new heavy chain isotypes), the extent to which the changes reported here represent an orderly developmental program is debatable. The lines appear to be capable of recombining their kappa or lambda light chain genes, but provide no direct evidence that this is a regulated process, or for example, that they can exhibit the progressive changes in transcription and translation of immunoglobulin genes that accompany (and define) B cell maturation (Wall and Keuhl, 1983).

## III. FORMULATION OF A MODEL

The process of joining Vκ to Jκ can be correlated with the deletion of DNA presumed to lie between these two elements. A subset of Vκ and Jκ sequences is absent from the genomes of certain kappa-producing myelomas (Seidman et al 1980). This is the result predicted if Vκ genes rearranged to Jκ elements lying in direct orientation, through a simple intramolecular recombination between the two targets (Hozumi and Tonegawa 1976).

Other studies have shown that the region 5' to a Jκ element is not always eliminated from the DNA of a rearranged cell; it can be retained in reorganized form (Steinmetz et al 1980, Van Ness et al 1982). The DNA sequence of the recombinant junctions in such reorganized fragments is consistent with that expected of the reciprocal recombination product of Vκ/Jκ joining, being a direct fusion of what appears to be the 3' flank of a Vκ gene segment to the 5' flank of a Jκ element (Steinmetz et al 1980, Hochtl et al 1982). This finding contradicts the view that Vκ-to-Jκ rearrangement is a deletional event ocurring between elements in direct orientation.

No simple alternative to a deletion model is consistent with all of the properties of these putative reciprocal joints, however. A myeloma with a coding joint on each chromosome often has only one reciprocal joint and frequently none (Steinmetz et al 1980, Selsing and Storb 1981, Van Ness et al 1982, Hochtl et al 1982). In addition, although a reciprocal joint appears to be a direct fusion of the 3' flank of a Vκ element and the 5' flank of a Jκ element, in no single case has either flank been derived from

the particular Vκ and Jκ elements found to be fused in the coding joints of the same myeloma cell (Steinmetz et al 1980, Hochtl et al 1982). In fact, all of the reciprocal joints that have been sequenced to date (Steinmetz et al 1980, Hochtl et al 1982, Alt and Baltimore 1982) as well as almost all of those analyzed by southern blot (Selsing et al 1984), have been J1 joining signal fusions. This appear to apply only to the reciprocal joints isolated from myelomas and hybridomas, however; in contrast, the reciprocal joints in PD often derive from Jκ elements other than Jκ1 (Lewis et al 1982, Chapter II). These observations have been explained to varying degrees on the basis of three proposals: that Vκ-to-Jκ joining involves deletion followed by occasional reintegration of the region between Vκ and Jκ (Steinmetz et al 1980), that it occurs by unequal sister chromatid exchange (Van Ness et al 1982), or that some of the time inversional events are involved (Lewis et al 1982)(Figure 16).

The initial suggestion that was made (Steinmetz et al 1980), was that reciprocal joints represent the reintegration of DNA that was reciprocally joined and thus looped out by the intial Vκ-to-Jκ fusion. A second recombination event reintegrates the DNA thereafter. Presumably, (although this was not incorporated into the initial proposal), secondary Vκ-to-Jκ joining events are responsible for the peculiar relationship of reciprocal joints and coding joints in myelomas.

Although there is no direct evidence that reintegration of DNA that has become recombinationally unlinked from a mammalian genome can occur, somatic cells clearly have the means by which to integrate foreign transfected DNA. It seems possible at least that a large, perhaps entangled, region of a chromosome might get reintegrated with some degree of efficiency. Even so, a deletion/reintegration proposal raises the question of why reciprocal

Figure 16

Three proposals for the origin of reciprocal joints.

Symbols for Vκ, Jκ and joining signals are defined in table 1.

A. Deletion/reintegration of the sequences intervening between Vκ and Jκ.

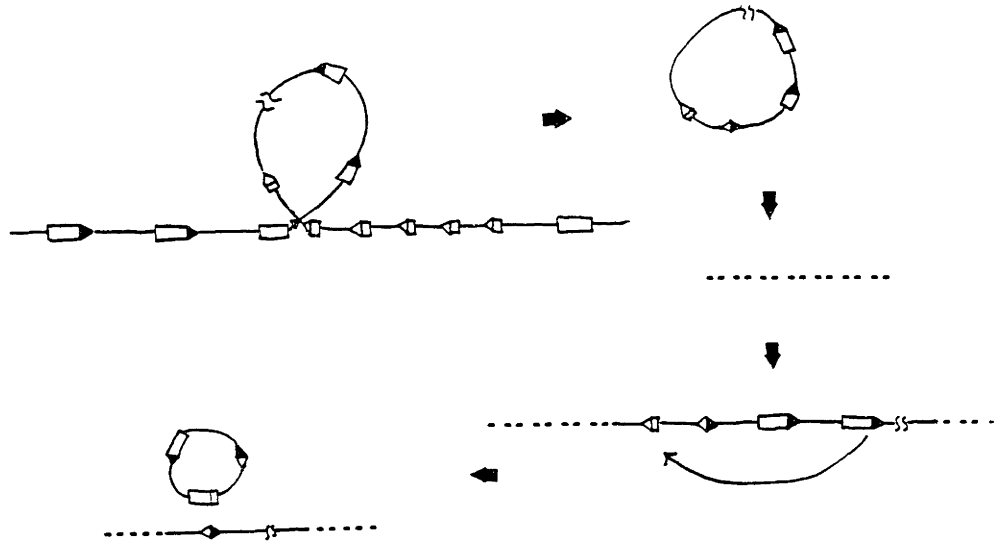
As shown this corresponds to the proposal of Steinmetz et al (1980) up to the last step depicting a secondary recombination event between Vκ and Jκ sequences contained within the translocated DNA. Equally feasible would be secondary rearrangment of Vκ and Jκ across the pre-formed coding joint; in any case some type of secondary recombination must occur in order for this model to adequately explain non-reciprocity of reciprocal joints and coding joints.

B. Unequal sister chromatid exhange.

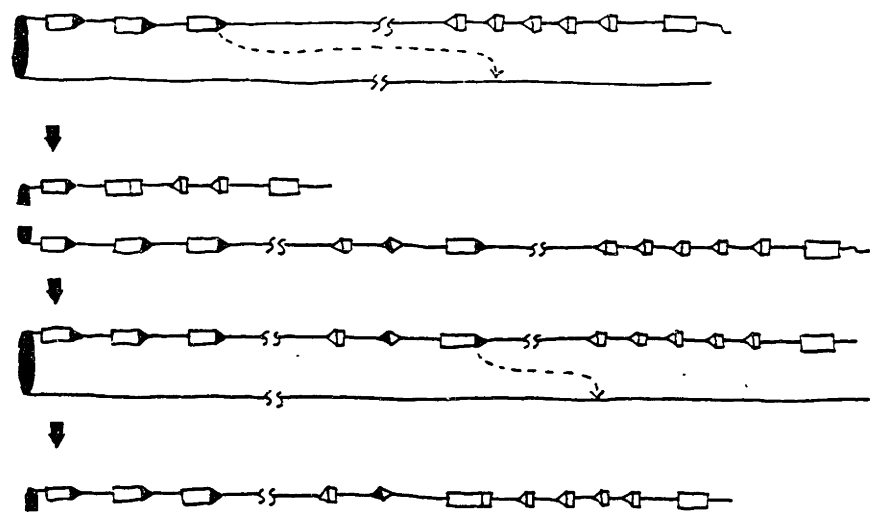
One homolog is shown. In the first step, rearrangement occurs in trans between Vκ and Jκ on sister chromatids. In the second step the centromere divides and the two products of rearrangement segregate into different daughter cells. If, as is shown in step three, the duplicated chromosome (which had not acquired a functional coding joint after the first round of recombination) undergoes replication and V/J joining again, some of the time a chromosome would be constructed in which a reciprocal joint and a coding joint would segregate together. In that circumstance, the reciprocal joint and the coding joint would not be reciprocally related.

C. Inversion/deletion.

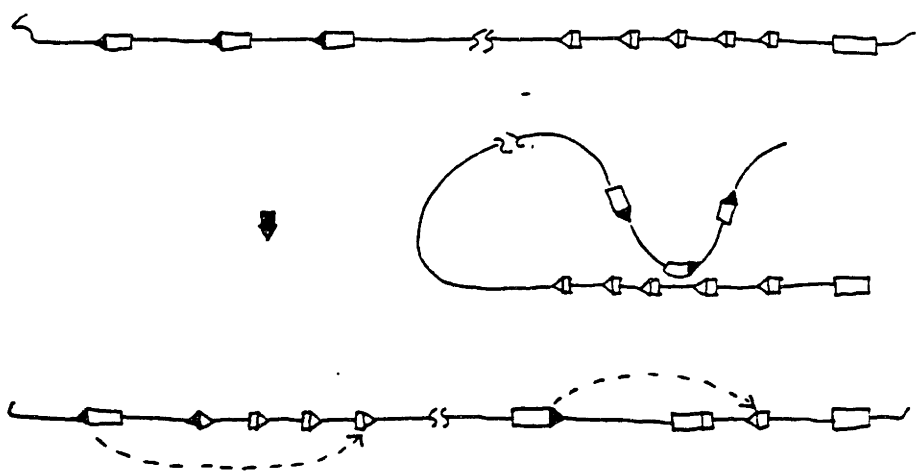
If some or all Vκ gene segments are oriented backward with respect to the Jκ gene segments, then Vκ/Jκ joining events could result in the inversion of the intervening DNA (rather than deletion as shown in panel A). As a consequence, reciprocally derived recombination products would persist in the DNA of a rearranged cell. Coding joints and reciprocal joints could be rendered non-reciprocal by secondary rearrangement events as indicated.



A



B



C

joints should be found at kappa locus but not at the lambda locus or the heavy chain locus as well. There is no obvious reason why sequences from one locus in particular should be more likely to reintegrate.

In the unequal sister chromatid exchange model (Figure 16 B), non-reciprocity of the reciprocal joints and coding joints in a particular cell is an inherent property of the joining mechanism. By this scheme, recombination would occur between sister chromatids and therefore coding joints and reciprocal joints would necessarily segregate into separate daughter cells. The two junctions would be created by an intermolecular recombination event after DNA replication; as a result, one chromatid would bear a deletion of all the sequences between Vκ and Jκ, and the other would bear a duplication of the same region. The deleted chromosome would contain the coding joint, the duplicated chromosome would contain the reciprocal joint. After cell division, the daughter cell that received the reciprocal joint, lacking a productively rearranged allele, could be imagined to undergo another round of recombination. Two products would again be formed, and again the deleted chromatid would acquire a productive rearrangement whereas the chromatid bearing a duplication would not. If in this second instance the Vκ segment that joined to the Jκ element lay closer to the Jκ element than the initial reciprocal joint junction (so that the reciprocal joint was not included in the V to J interval), then the chromatid that acquired a coding joint would retain the reciprocal junction from the previous rearrangement. The two junctions would not have been derived in the same recombination event and they would not bear a reciprocal relationship to one another. Thus, the unequal sister chromatid exchange model adequately accounts for both the non-reciprocity of coding joints and

reciprocal joints and the lower frequency with which the latter are found in myelomas (Van Ness et al 1982, Lewis et al 1982, Selsing et al 1984).

Unequal sister chromatid exchange directly implies that certain properties must characterize the V-to-J joining mechanism. First, it suggests that V-to-J joining can occur between target sequences on different DNA molecules. As mentioned in Chapter I, information bearing on this issue is scanty, however the unidirectional joining of D segments, the possibility that the proximity of V to J as mapped may have an effect upon recombination frequency, and the chromosomal organization of the rearranging loci all peripherally suggest that tracking along the DNA may be a feature of the recombination and that therefore it is an intramolecular event. If unequal sister chromatid exchange could be shown to occur, it would provide the first concrete evidence against tracking. Second, according to the unequal sister chromatid exchange model, the products of a conservative, reciprocal V/J joining event are proposed to segregate at cell division, and therefore one daughter cell should acquire a duplication of a large segment of the kappa locus. As a unique prediction of unequal sister chromatid exchange, the demonstration of such duplications would go far to substantiate the proposal. No evidence for duplications of kappa sequences occurring as a result of V region exon assembly has ever been reported, however.

None of the above considerations argues strongly against the possibility of an unequal sister chromatid exchange mechanism. If reciprocal joints are in fact the result of unequal sister chromatid exchange, however, two features of reciprocal joints require further explanation; it is curious that they should accumulate only at the kappa locus, and it is unclear why Jκ1 derived junctions, in particular, should predominate.

A third possibility which, mechanistically, is very similar to the first, suggests that the formation of reciprocal joints is a consequence of the orientation of the elements involved (Figure 16 C). Supposing that Vκ gene segments occur in either one of two orientations relative to Jκ (as might be imagined if during evolution an inversion of a portion of the Vκ region of the kappa locus occurred) then Vκ-to-Jκ recombinations could be either inversional or deletional processes. If a backward Vκ gene segment were to somatically rearrange to Jκ, a reciprocal recombination would result in a large inversion of all the sequences between Vκ and Jκ. A coding joint would lie at one border of the inversion, and a reciprocal joint would lie at the other. As documented in the PD cell line, secondary rearrangments of both coding joints and reciprocal joints can occur. If these secondary rearrangements are likewise either inversional or deletional Vκ-to-Jκ recombination events, (Figure 16 C), there would be several consequences; 1) coding joints and reciprocal joints isolated as final structures from a rearranged genome need not be reciprocally related, and 2) sequences between Vκ and Jκ could be deleted to varying extents (but never duplicated). By this scheme one might not find a 1:1 ratio of reciprocal joints and coding joints in a cell, if either reciprocal joints could be deleted as a secondary event or if they were not generated every time V-to-J joining occurred. Repeated inversion/deletion events do not directly lead to a predominance of Jκ1 reciprocal joints, but this observation can be easily accomodated if it were true that deletional rather than inversional rearrangement is the more frequent form of secondary reciprocal joint recombination. There are several reasons why this might be so (discussed in the next Chapter); as a consequence the Jκ cluster in a reciprocal junction

would tend to be efficiently "whittled down" to a junction which is the fusion of a Vκ-derived joining signal to the Jκ1 heptamer.

The three proposals outlined above; deletion/re-integration, unequal sister chromatid exchange and inversion/deletion, each demand different properties of the V/J joining mechanism. Deletion/re-integration necessitates that the circularized region bearing reciprocal junctions that is unlinked from the cell's DNA during joining become reintegrated fairly efficiently. A unique requirement of the unequal sister chromatid exchange model is that trans-recombination can occur. And an inversion/deletion model requires that the joining enzymes must be able to recombine joining signals in either of two orientations. In order for both the deletion/reintegration inversion/deletion models to adequately explain the non-reciprocity of recombinant junction, they require that secondary recombination be frequent.

One could argue that an inversional model and unequal sister chromatid exchange are mutually exclusive proposals. If inverted joining took place between two chromatids it would produce two aberrant structures, an acentric and a dicentric chromosome. To be efficient, it at least seems reasonable to expect that an inversional joining mechanism would probably have evolved in such a way as to preclude the trans-recombinations required by unequal sister chromatid exchange.

Inversional joining could be simply demonstrated if it were possible to analyze the physical linkage of Vκ and Jκ. Without any reasonable estimate of the distance between these two elements, cloning the entire region is not a tempting approach. For this reason it seemed worthwhile to approach the question of topography more indirectly, while approaching the question of

mechanism more directly by asking whether the requirements of the inversion/deletion model mentioned above are met by activities in the cell.

The analysis of the endogenous genes in PD (Chapter II) indicated that secondary recombination could be observed experimentally. To demonstrate that Vκ and Jκ gene segments could be joined in a reaction involving inversion, I introduced a recombination substrate into PD cells that was constructed so that Vκ was backward relative to Jκ.

## IV. DESIGN OF A RECOMBINATION SUBSTRATE

Although precursors and products of rearrangement can be identified by studying the clonal history of the cell in which they occur, there are basic limitations to this approach. Extensive analysis is required to unambiguously establish whether a given gene is the result of single or of multiple alterations of its precursor. In addition it is non-trivial to isolate the corresponding unrearranged V gene segment from among related reiterated copies in the germ-line. Further there is little opportunity for experimental intervention in the system.

A system in which one can test recombination of defined substrates in vivo, might eventually provide access to issues that would be impossible to study in a strictly endogenous context. Exogenously supplied substrates should facilitate experimentation directed at the understanding of the regulation of V-to-J joining, the mechanism of the process, the substrate specificity of the enzymes, the cell types in which joining can occur, etc. Different methods of introduction of the substrate would result in different configurations inside the cell. The choice of method of introduction was guided by the expectation that a stably integrated, low copy number substrate would present the best approximation of endogenous gene sequences to the recombination enzymes.

### General Considerations

Different methods of introducing an exogenous substrate should result in different configurations of the DNA inside the cell. The choice of

what general method to use was guided by the expectation that a stably integrated, low copy number substrate would present the best approximation of endogenous gene sequences to the recombination enzymes. There were several reasons for electing to use a retrovirus vector beyond this. First, introduction of the substrate as a defective retrovirus would simplify further analysis because retroviruses integrate at a specific sequence at the borders of their LTRs (Panganiban and Temin 1983). This feature ensures that the substrate sequences will not be fortuitously interrupted or rearranged in the course of stable integration into the chromosome. Second, as a general rule the copy number of retrovirally introduced sequences is low, usually single integrants can be selected without difficulty. Third, retrovirus infection is far more efficient than transfection (Mulligan 1983). All of these features provided definite advantages in an undertaking where it was expected that the frequency of recombination might be very low and where the initial analysis would would have to rely on the interpretation of genomic southern blotting experiments.

In outline, the plan of the experiment was to construct a substrate using an MuSV-based retroviral vector (Mulligan 1983), infect it into PD cells and then select for a marker gene that would be activated by a V/J joining event. In order to do this the DNA construction was built from pMSVgpt (Mulligan 1983), and designed so that V-to-J joining would occur inversionally . Expression of the Eco gpt gene (Mulligan and Berg 1980) was to be conditional on its having been reoriented by the inversion.

The packaging limit of a retroviral particle has not been directly determined (Mulligan 1983) however it seemed reasonable to design a substrate that was roughly the size of a wild-type Mo-MuLV genome. Other constraints were that the substrate had to be able to passage as virus

before recombination, and provide efficient expression of gpt afterwards. The location of polyadenylation and splice sites in the insert could be expected to affect viral titres, and the expression of gpt had to be designed so that a recombination that is known to be site specific yet "flexible" would reliably activate the marker. Additional concerns centered on which sequences from the kappa locus were likely to be absolutely required for recognition/rearrangement by the recombinase.

In response to the space constraints, the Jκ cluster without Cκ was included in the substrate. Because the endogenous reciprocal joints in PD (which by any scheme must be located far from Cκ) could still undergo secondary rearrangement it seemed possible that Cκ might be dispensed with in this experiment. The entire Jκ cluster was incorporated into the substrate instead of a single Jκ element to increase the liklihood of detecting rearrangement in the event that some one but not another Jκ element would be favored in this particular construct. The Vκ element that was used was a genomic Vκ 21-C gene segment (Heinrich et al 1984) provided by S. Tonegawa. The Vκ 21-C gene segment has been identified as the precursor to the functionally rearranged Vκ element in the myeloma MOPC 321 (Heinrich et al 1984). This was important to know, as a certain number of V gene segments are pseudo-genes, and might not be good substrates for rearrangement.

A major difference between the final configuration of the substrate and an endogenous kappa locus was that the Vκ sequences would be only a few kb away from the Jκ segments instead of 100 kb or more. However Dh-to-Jh recombination can be observed over relatively short (<1kb) distances in some B and T cells (Kurosawa et al 1981) and it was anticipated that the

proximity of Vκ to Jκ on the substrate therefore should not in itself preclude recombination.

The structure of the recombination substrate, pVJG

On the basis of these considerations, we placed Vκ and Jκ in transcriptionally opposite orientations with the gpt marker between them as represented in Figure 17. In the substrate pVJG, the gpt gene was placed into the interval between Vκ and Jκ without a linked eucaryotic promoter so that in the absence of joining it would be expected to be transcriptionally silent. Vκ-to-Jκ joining would proceed via inversion, and this would serve to re-orient the gpt gene so that it could be expressed using the retroviral 5' LTR promoter. The selection was designed around an inversion event rather than a deletion for three reasons: one was to provide a direct demonstration of inversional Vκ-to-Jκ joining, another was to take advantage of the lower background of non-specific events anticipated to occur inversionally, and a final reason was that inversional joining provides the means by which to isolate two recombinant junctions instead of just one.

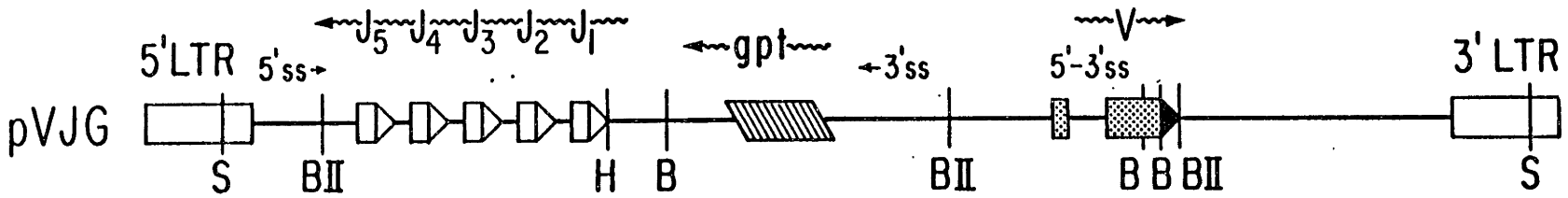
Vκ and Jκ and gpt elements were arranged as represented in Figure 17. To ensure expression of gpt after inversion, splice sites were positioned so that the region between the LTR promoter and the gpt coding sequences (which would have to include one of the two recombinant junctions) would be spliced out of the mRNA. (A construct in which the gpt gene had been linked to an envelope 3' splice site was kindly provided by Dr. S. Hartman). By this means it was hoped that the exact location and sequence of the junction site (a variable in this experiment because of the multiple Jκ elements and also because of the inherent flexibility of V/J joining) would have little effect

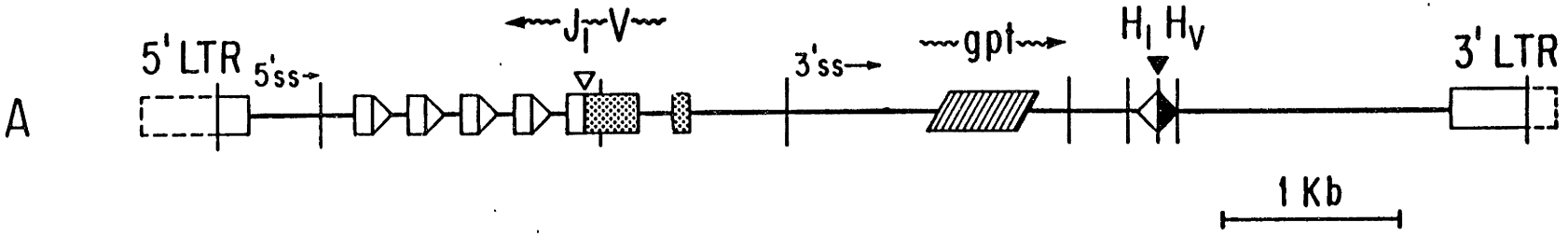
Figure 17

Structure of a substrate designed to detect $V_K$-to-$J_K$ recombination

Upper: VJG in unrecombined form. Symbols are defined in Table 1. See text for a detailed description.

Lower: VJG after recombination. Site specific $V_K$-to$J_K$1 rearrangement is depicted. Inversion would reorient the Eco gpt gene so that it could be expressed using the viral 5' LTR promoter of the vector.


pVJG
5'LTR
5'ss
J5 J4 J3 J2 J1
gpt
3'ss
V
5'-3'ss
3' LTR
S
BII
H
B
BII
B B BII
S


A
5'LTR
5'ss
J1 V
3'ss
gpt
H1 HV
3' LTR
1 Kb

on gpt expression.

Between the LTRs were located (in the 5' to 3' orientation); the 5' splice site of MuSV which is used by the wild-type virus in producing subgenomic envelope gene mRNA; the Jκ segments (in 3' to 5' transcriptional orientation); the Escherichia coli xanthine-guanine phosphoribosyl transferase (gpt) gene (in 3' to 5' orientation) and an inverted envelope gene 3' splice site. This was followed by the 5'-to-3' oriented genomic Vκ21-C segment derived from Balb/c DNA . The DNA following the Vκ21-C segment in the construct is derived from MuSV and includes the 3' LTR with its associated primer binding site. The rest of the construct is the pBR322 backbone (and other sequences) of pMSVgpt. These would not be included in the retrovirus form of the substrate, and are not indicated in the Figure.

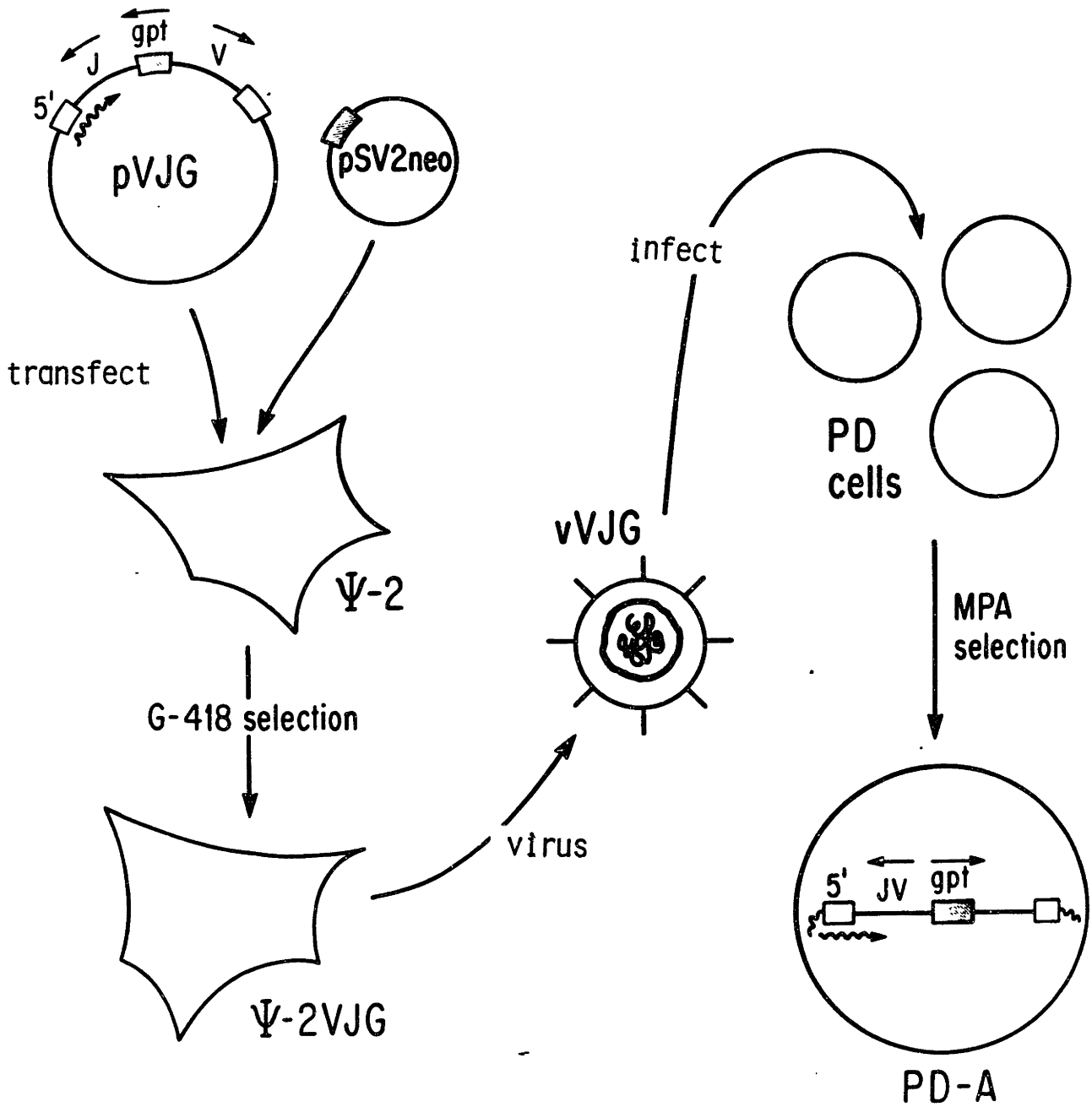
Derivation of a Recombinant

The plan of the experiment is detailed in Figure 18. The DNA construct was transfected into Ψ-2, an NIH 3T3 derivative that constitutively produces murine leukemia virus (MuLV) proteins without packaging the resident viral genome (Mann et al 1982). A co-transfected plasmid carrying a neomycin resistance gene (pSV2neo) (Southern and Berg, 1982) allowed the isolation of stably transfected Ψ-2 cells. The virions produced by the transfected "Ψ-2VJG" cells were used to infect PD cells. One to two days after the infection of $10^5$PD cells with substrate virus, the cells were placed in microtitre wells at 2-4x$10^3$ cells per well and were grown thereafter in the

Figure 18

Outline of the procedure used to introduce substrate sequences into PD.

10 µg of pVJG and 1 µg of pSV2neo (Southern and Berg, 1982) were transfected into Ψ-2 cells as described in Mann et al (1983). Four neo-positive Ψ-2 cell lines were selected in medium containing 1 mg/ml of the drug G418 (Gibco) (Colbere-Garapin et al 1981), then screend for the production of recombinant virus (vVJG). A virus producer was found among the four transfectants by testing the culture medium from each for the ability to convert PD cells to mycophenolic acid resistance. For this assay, procedures for harvesting virus and carrying out viral infections as outlined in Chapter 5 (Materials and Methods) were followed, following which PD cells were plated in mycophenolic acid-containing medium (details are given in the text). Gpt-positive clones appeared in 3-4 weeks and were maintained in selective medium.


gpt
J
V
5'
pVJG
pSV2neo
transfect
Ψ-2
G-418 selection
Ψ-2VJG
virus
vVJG
infect
PD
cells
MPA
selection
5'
JV
gpt
PD-A

presence of 0.5 μg per ml of mycophenolic acid. At three to four weeks following plating, from one to five mycophenolic acid-resistant clones appeared per experiment and nineteen were isolated for further study. All but one of the isolates expressed the introduced bacterial enzyme by direct assay (Mulligan and Berg, 1980) and the negative one was discarded.

To examine the structure of the recombined substrate, cellular DNA samples of the PD derivatives were analyzed by southern blot using a gpt probe. Several different restriction enzymes were used to generate crude maps of the structure of the integrated proviral sequences. On the basis of this analysis the cell line PD-A was selected for more detailed examination.

VJG Sequences are Site-Specifically recombined

The proviral sequences in PD-A DNA were cloned into λgt wes•λB by digesting PD-A DNA with SACI, an enzyme that cuts only in the LTRs of the substrate, and screening recombinant phage for the clone of interest with the gpt gene probe.

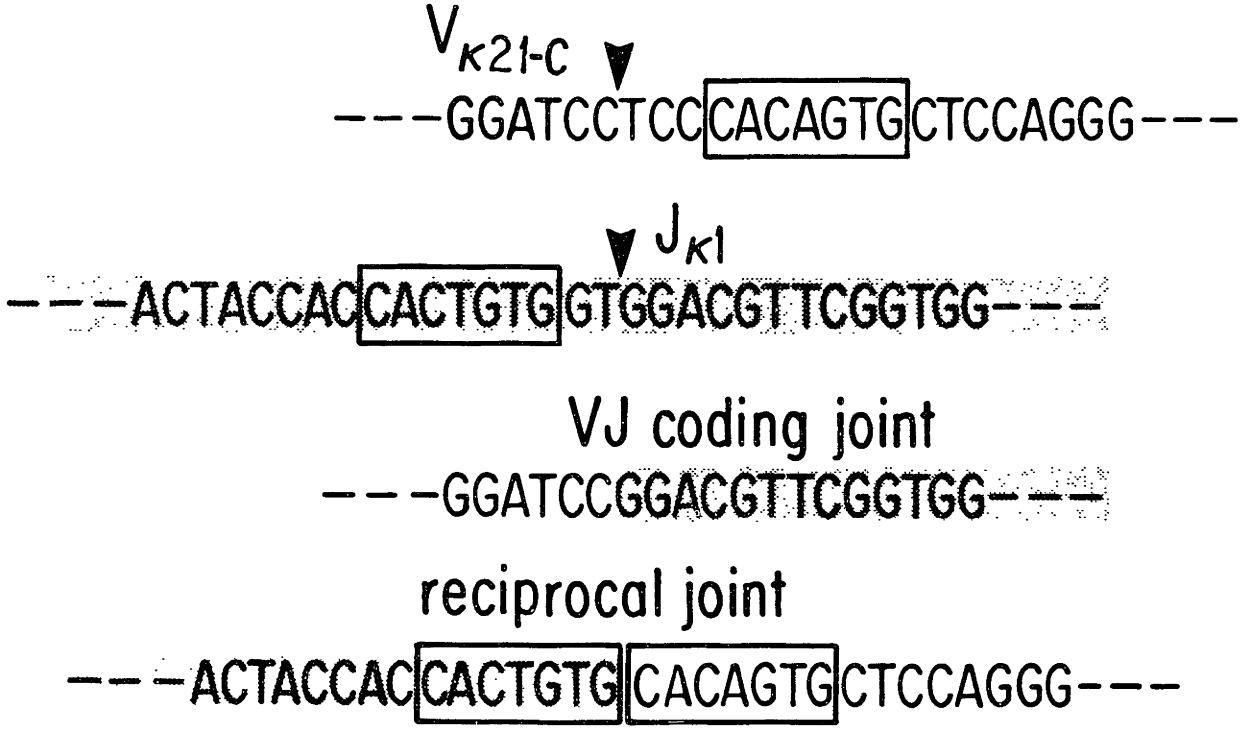
Restriction enzyme analysis of the cloned provirus from PD-A (hereafter referred to as recombinant A) indicated that it had the structure shown in Figure 19 (bottom). The only alteration relative to the original construction appeared to be a large internal inversion, with end points that were located in regions which suggested that a Vκ-to-Jκ recombination had occurred.

The DNA sequence at the inversional end points was determined using the method of Maxam and Gilbert. The precursor Vκ21-C and Jκ1 sequences are shown in Figure 19 together with the two recombinant joints. The sites at

Figure 19

Sequences of precursors and products of recombination.

The germ-line Vκ21-C and Jκ1 gene segments are shown on the top two lines. Vκ21-C was sequenced directly, Jκ1 sequence is from a published report (Sakano et al 1979), and was confirmed directly.The heptamer elements of each of the associated joining signals are boxed, nonamer elements are not included in the figure. The coding joint and the reciprocal joint resulting from recombination are shown in the last two lines. The black triangles demarcate the sites of crossing over that formed the coding joint. The crossover sites of the reciprocal joint are located at the borders of the heptamer elements. It can be seen that the bases in between these two crossover sites on both Vκ and Jκ are lost from the products of Vκ-to-Jκ recombination.


$V_{\kappa 21\text{-}C}$
---GGATCCTCCCACAGTGCTCCAGGG---
$J_{\kappa 1}$
---ACTACCACCACTGTGGTGGACGTTCGGTGG---
VJ coding joint
---GGATCCGGACGTTCGGTGG---
reciprocal joint
---ACTACCACCACTGTGCACAGTGCTCCAGGG---

which the Vκ21-C and Jκ1 elements were fused to form a coding joint can be unambiguously assigned and are indicated by arrows in the Figure. The cross-over point on Vκ 21-C is identical to that observed in the myeloma MOPC321 (where the same V gene fused with Jκ4 -Sakano et al 1979). In the present case, the reading frames of the Vκ and Jκ elements don't match, similar to the types of junctions that are often found on non-expressed alleles (Coleclogh 1983).

The structure of the reciprocally-formed junction in PD-A is typical of all of the endogenously-arising reciprocal joints that have been sequenced (Steinmetz et al 1980, Hochtl et al 1982, Alt and Baltimore 1982, Hochtl and Zachau 1983, Selsing et al 1984). The reciprocal joint was an exact apposition of the joining signals that were initially associated with Vκ21-C and Jκ1 (Figure 19).

The structures of these two recombinant products left little doubt that the enzymes in the cell can recognize the kappa gene segments on the substrate and rearrange them site-specifically.

This study demonstrated that inversional Vκ -to-Jκ joining can occur, (for estimates of the frequency relative to endogenous kappa gene recombination see Table II) and for the first time showed that a junction analogous to the reciprocal joints isolated from myelomas can be formed as a primary product of V/J recombination.

## V. THE PRODUCTS OF V/J RECOMBINATION ARE ASYMMETRICALLY FORMED

One way to learn about some of the details of the Vκ-to-Jκ joining reaction is to examine the fine structure of recombinant junctions. The sequences of these junctions are informative because the component gene segments are joined to one another in a region devoid of homology (Bernard et al 1976). The crossover sites reveal where the parental duplexes must have been interrupted before ligation created the recombinant linkage. Through the use of an inversion substrate, all the products of a single recombination event are represented--one can examine a coding joint pairwise with the corresponding reciprocal joint.

To this end, nine recombinant forms of VJG were analyzed. I have attempted to 1) examine some of the properties of the recombinase through structural analysis, and 2) characterize how successfully this system allows the retrieval of site-specifically recombined sequences.

Based on the initial southern blot mapping of the proviral substrate in fourteen mycophenolic acid resistant PD derivative, all of the PD isolates examined harbored a recombinant substrate that appeared to have been site-specifically rearranged (in particular, use of a gpt gene probe on Bgl II-digested DNA revealed a characteristic band of approximately 2 kb in each; Lewis et al 1984). No evidence for background activation of the gpt gene by any means other than site-specific Vκ-to-Jκ recombination was detected although a sizeable fraction of the PD isolates contained recombined substrates that appeared to have undergone more complex rearrangements than the inversion observed in PD-A (not shown). To examine additional examples of Vκ-to-Jκ1 recombination, some isolates that appeared identical to PD-A were chosen for analysis. To investigate the variants, several isolates

that were recombined to another Jκ segment and several of the more complex recombinants were chosen as well.

A genomic library was prepared from each of eight mycophenolic acid resistant PD derivatives, and screened as before. Recombinant substrate sequences were recovered as SacI inserts in λ gt/wes.λB. The structures of these recombinant clones is diagrammed in Figure 20.

## PROPERTIES OF CODING JOINTS

The recombinant junctions in variable region exons are not precisely specified (Sakano et al 1979, Max et al 1979); 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 and 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. (This can be accomplished if, for example, a particular idiotype is selected among hybridoma cells, see Manser and Gefter 1984). In the present system, it is possible to examine junctional diversity in a context where the precursors are known, and most importantly, without selection for expression. To this end, five independent Vκ/Jκ1 recombinants (A,I,O,N, and X) and three Vκ/Jκ4 recombinants (D,M, and P) were isolated. The sequences of these junctions are shown in Figure 21.

One consequence of junctional diversity was immediately apparent. Only three of the eight coding joints (those in recombinants O, D, and P) occur in the same translational reading frame as a functional kappa immunoglobulin gene. This ratio is close to the value expected if the recombination

Figure 20

Structures of VJG recombinants.

Structures of the SacI-to-SacI region of each recombinant after cloning into phage λ are diagrammed. Each recombinant junction is designated by one of the symbols listed in the key (see Table 1 for definitions). The sequenced regions are shown below each diagram. Restriction endonuclease sites are abbreviated as follows: S, SacI; BII, BglII; H, HindIII; B, BamHI; Ha, HaeIII; Hf, HinfI; Sa, Sau 3AI. Recombinants A, N, O and P were isolated from PD31-infected derivatives as described in Chapter 3. 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 in Chapter 2. PD34-28-1 is a derivative of PD34, similar to PD34-14 or 34-11 in Figure 8 except that it has deleted both coding joints.

# Structures of VJG Recombinants

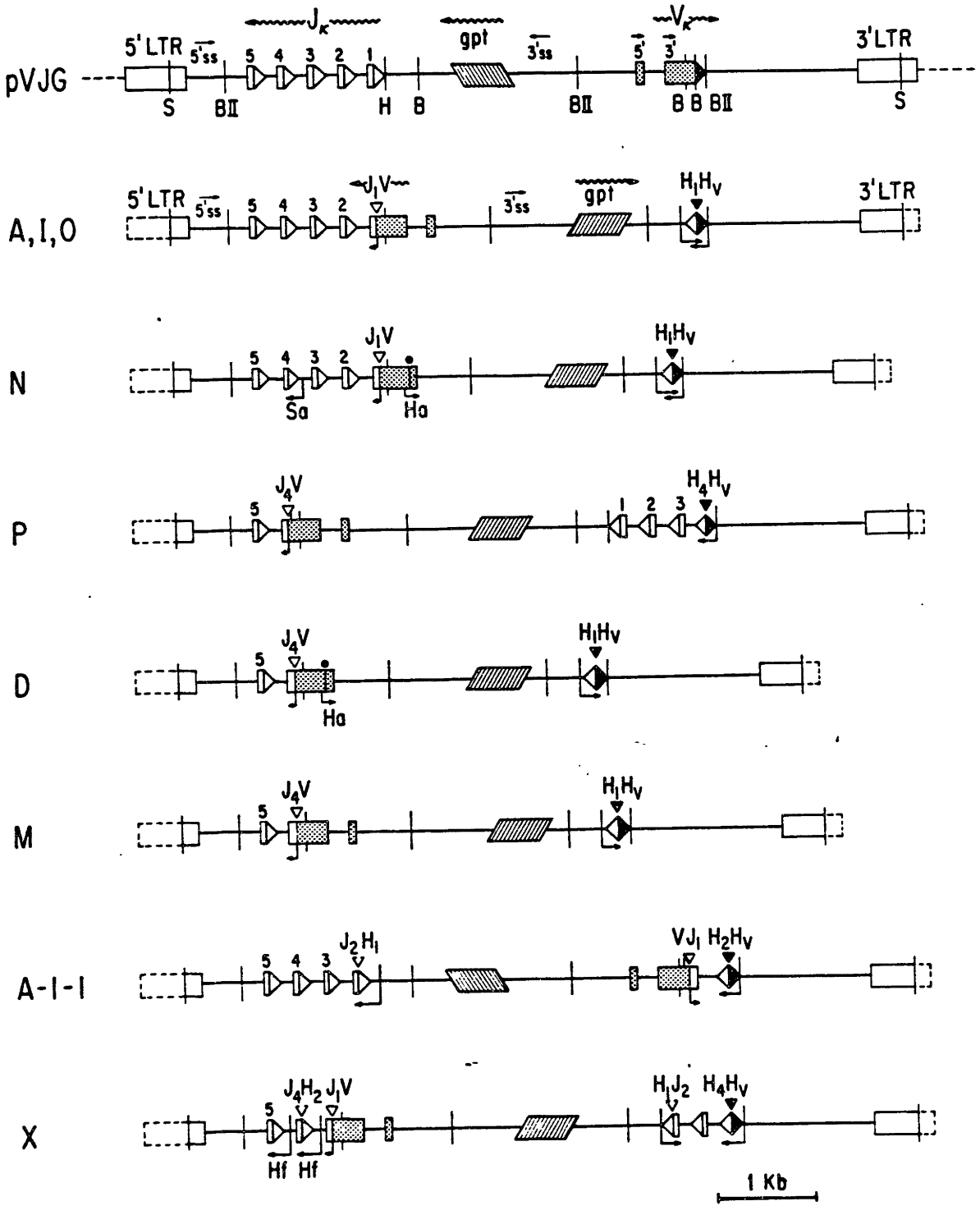


Figure 21

Sequences of recombinant junctions.

Recombinant junctions were sequenced by the method of Maxam and Gilbert (1977). The junctions are symbolized as in Figure 20. Heptameric elements of the joining signals are boxed; the spacer regions and nonamer elements of the joining signals are not included in the sequence. Sequencing gels were read for 20 bases or more on either side of the junctions in each case, except for the coding joints. These were seqenced by labeling at a Bam HI site in the $V_K$ coding sequences 8 bases 5' to the joining signal. Ambiguous crossover sites due to redundancy between the recombining sequences are indicated by multiple triangles. The translational reading frames of the V and J gene segments are indicated by the wavy lines. The lower case letters in the coding joint of recombinant D represent a 2 bp insertion. The total number of base pairs that are missing from each reciprocal pair of junctions are shown in the right hand columns. The number of base pairs lost upon formation of the reciprocal junctions in recombinant X is based upon the scheme in Figure 24, but alternative derivations would give a different value.

# Sequences of Recombinant Junctions

| Coding Joints | | Reciprocal Joints | | # Δ bp |
|---|---|---|---|---|
| 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 bp |
|---|---|---|---|---|---|---|
| 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(?) |

process were to fuse Vκ and Jκ elements randomly with respect to the coding properties of the product (Altenburger et al 1980, Max et al 1980). The frequency with which rearrangement can generate functional coding joints is an important factor in attempting to understand how the rearrangement process might be regulated (reviewed by Coleclough 1983). While the present data are limited, they represent an essentially un-biased measurement of the coding fidelity of Vκ/Jκ gene rearrangement and fully confirm earlier proposals (Altenburger et al 1980, Max et al 1980).

The variability of Vκ/Jκ coding joints can be examined more closely by comparing recombination sites. Among the eight coding joints presented in Figure 21, 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 hollow triangles in Figure 21. In Figure 22 (hollow triangles) these data are compiled relative to unrearranged Vκ and Jκ elements.

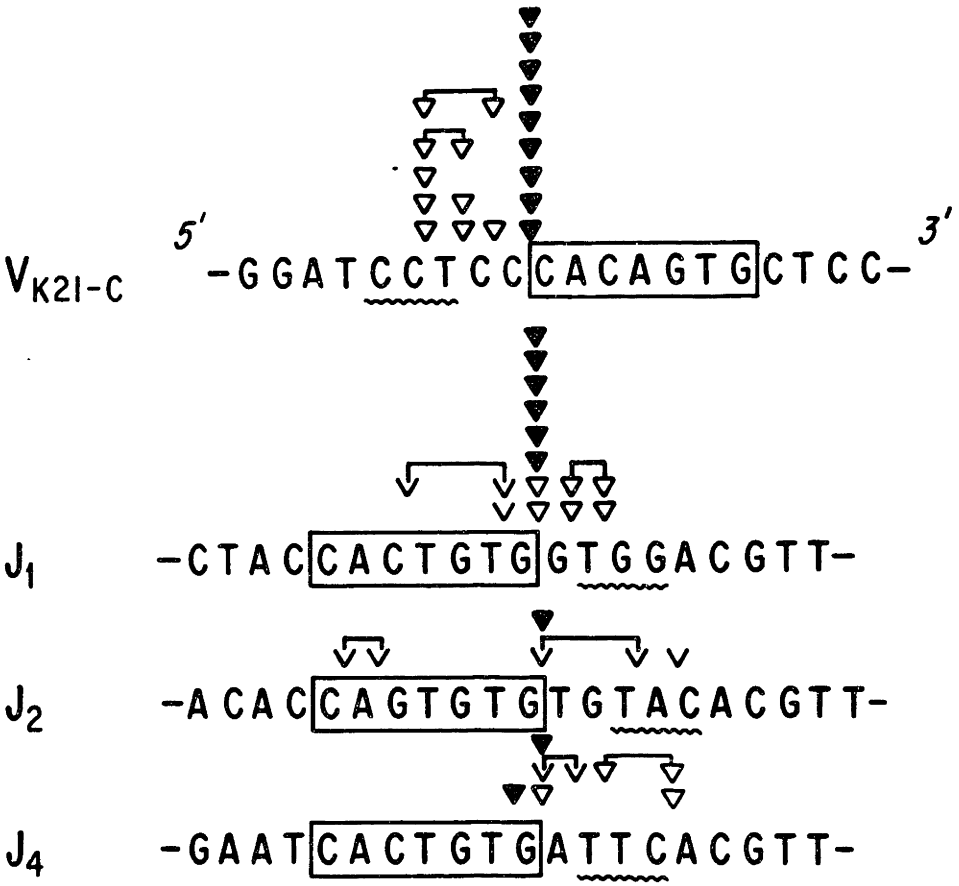
Recombination sites can be seen to cover a region on the Vκ element ranging from one to three bases away from the heptamer border. On the Jκ element they are located either directly at the heptamer border or up to four bases in toward the coding sequences. As suggested by the frequency and distribution of sites shown in Figure 22, the positioning of a coding joint crossover on either component element (Vκ or Jκ) appears to be random within a small region. Other possible patterns that would not be evident in this compilation were searched for. It was hoped 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

Figure 22

Distribution and frequency of crossover sites.

The DNA sequence of unrearranged Vκ 21-C, Jκ1, and Jκ4 elements in VJG were determined directly . The unrearranged Jκ2 sequence is from Sakano et al (1979). Symbols above the crossover sites are as in Figure 20. Ambigous junction sites are delimited by a bar with triangles at the right and left extremes. Codon 95 of Vκ, and codon 96 of each Jκ element (numbered according to Kabat et al, 1983) are underlined.

# Distribution and Frequency of X-Over Sites



sites on Vκ and Jκ as measured from either their heptamer borders or from other landmarks such as the last codon, revealed no positive or negative correlation. In addition it was clear that a variable number of base pairs are missing relative to the precursor sequences from recombinant junctions (Figure 21). This will be discussed further when the data for reciprocal joints is presented. It appears from the data in Figures 21 and 22 that if there is any restriction on the location a given Vk crossover site with respect to a given Jk crossover site, any pattern that exists must be fairly subtle because there is no hint of it in this collection. A novel property of the V-to-J joining mechanism is that some step in the recombination process generates variability in a random or near-random fashion.

## A 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 which are not derived from either of the two precursor gene segments. These two bases did not exist in the sequence of the Vκ and Jκ4 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 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. As discussed in the introduction, N regions consist of short stretches of nucleotides of fairly random base composition, which appear as apparent insertions at Vh/Dh or Dh/Jh junctions of heavy chain genes (Sakano et al,1981, Kurosawa et al 1982, Alt

and Baltimore 1982, Desiderio et al 1984) and at the junctions of assembled β chain variable region exons in T cells (Kavaler et al 1984, Siu et al 1984). N regions have not been reported to be present in joined light chain genes (with one possible exception, Max et al 1980).

Beyond that, little is known of their origin. It has been suggested that N regions are added onto free ends of DNA through the action of a template-independent polymerase such as terminal transferase (Alt and Baltimore, 1982) or that they may be recombinationally acquired from pre-existing germ line sequences (Kurosawa et al 1982).

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 (Wilson et al 1983, Anderson et al 1984, Wake et al 1984), and have been demonstrated at the breakpoints of myc gene translocations as well (Gerondikas et al, 1984). Because of this, no obvious criteria distinguish N regions from non-specific insertions. The junctional insert in recombinant D does not seem to have been caused by terminal transferase, because enzymatic assays of cellular extracts from PD-A, -X and -D showed no measurable activity above background in any of these lines (S. Desiderio, personal communication). Whatever their cause, the two extra bases found at the Vk/Jk4 junction of recombinant D demonstrate 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 Vκ/Jκ coding joints, the Hv/H1 (abbreviations defined in Table I) reciprocal joints found in the recombinants A,I,O,N and the Hv/H4 junction in P are identical types of fusions. In every case, the joining signals located at the 3' flank of Vκ and at the 5'flank of Jκ were joined to one another exactly at the coding proximal borders of their heptamers (see Figure 21).

The invariant structure of reciprocal joints was first noted in endogenously arising recombinants (Steinmetz et al 1980, Hochtl et al 1982, Selsing et al 1984). However, in these examples, all were Jκ1 derived and all were non-reciprocal to the coding joints in the same cell. The possibility therefore arose, that these endogenous Jκ1-derived reciprocal joints were not directly formed in functional Vκ-to-Jκ joining events, but represented a by-product of a related reaction (see following sections). The recombinants discussed above, which include repeated isolations of Vκ-to-Jκ1 recombination events as well as a product of a Vκ-to-Jκ4 rearrangement, definitively demonstrate that precise recombination sites are a general feature of reciprocally formed junctions.

A consequence of the qualitative difference between reciprocal joints and coding joints is that the two products of a reciprocal Vκ-to-Jκ joining event are not exact reciprocals of one another at the nucleotide level. Because coding joint crossover sites 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 results from every recombination event. In our collection, from one to six bases that existed in the precursor sequences are missing from the products (Figure 21). Whether the bases disappeared during formation of the coding joint, the reciprocal joint, or both cannot be determined. Nevertheless the microscopically non-conservative nature of immunoglobulin gene rearrangement is a provocative feature. Analogy can be made to the integration of many eucaryotic and procaryotic transposons which appear to insert into host DNA likewise through grossly reciprocal but non-conservative recombination events. 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 transpcseable elements have been suggested to result from the introduction of staggered breaks on the target duplex (reviewed by Kleckner, 1981). Although the present case is an instance where base pairs are lost, not added, it is conceivable that the non-conservative aspect of V-to-J recombination similarly reflects the way in which the participating duplexes are cut and reconnected (see below).

## SUBSTRATE SEQUENCES MAY FREQUENTLY UNDERGO SECONDARY RECOMBINATION EVENTS

The 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. Deletional joining appears to have occurred in several of the PD cell lines that were isolated. As mentioned previously, the initial mapping studies of various mycophenolic acid resistant PD derivatives indicated the integrated substrate in several of these lines may have undergone multiple recombination events. These appeared

to have been mediated by reoriented joining signals. We detected deletions in a region of the substrate that suggested that the deletions were formed by the secondary recombination of a reciprocal joint with the Jκ1 element.

To examine examples of deleted recombinants, the substrate from the lines PD-M and PD-D were molecularly cloned. In both cases, the coding joints proved to be fusions of Vκ and Jκ4 but the reciprocal joints had been derived from Vκ and Jκ1. Because M and D were missing all of the Jκ cluster lying 3' to the Jκ1 heptamer and 5' to Jκ4 coding sequences, their structures were consistent with the secondary recombination diagrammed in Figure 23. As indicated in the Figure, the recombination of a reciprocal joint with Jκ1, if it occurs in a fashion analogous to primary recombination events, will have two consequences; the formation of an H1/Hv reciprocal joint, and the deletion of a portion of the Jκ cluster. These alterations had evidently occurred in both D and M.

Mapping data indicated that deletions similar to those observed in D and M, --every one of which was accompanied by the presence of an H1/Hv reciprocal joint-- existed in five of the fourteen original PD isolates. Because the structures that were inferred from the in situ maps of the integrated substrates were confirmed in the cases of PD-M and PD-D, one could argue from these observations that 1) secondary reciprocal joint recombinations occur fairly frequently, and 2) H1/Hv reciprocal joints are over-represented as a consequence.

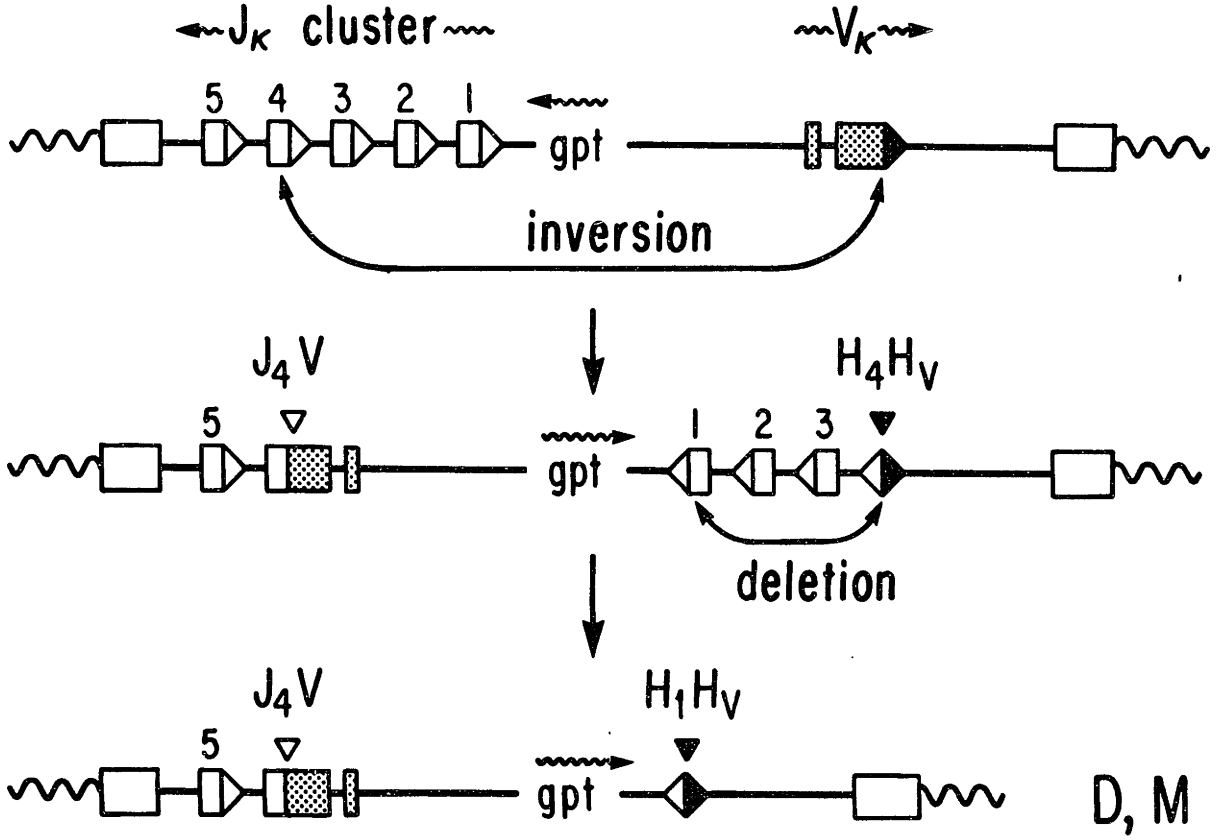
The deletive secondary rearrangement of the substrate provides a simple and consistent paradigm for both the non-reciprocal junctions and the H1/Hv bias that are generally observed in the rearranged kappa alleles of myelomas (Steinmetz et al 1980, Hochtl et al 1982, Selsing et al 1984). An obvious explanation for the H1/Hv bias found among the recombinant substrate in PD

Figure 23

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 Figure 20.

# Probable Origin of Recombinants D and M



derivatives is that H1/Hv junctions represents the limit recombination product of continued rearrangement; in the process of rearranging to Jκ1, all other similarly-oriented Jκ elements are eliminated. By analogy, supposing that endogenous gene segments are processed similarly to gene segments on the introduced substrate, repeated reciprocal joint recombination events might well underlie the parallel Hv/H1 bias and non-reciprocal junctions observed in endogenous myeloma genes. (In contrast to the model proposed in Figure 16C, secondary reciprocal joint recombinations may not involve intact Vκ gene segments, and thus are related, but not identical, to Vκ-to-Jκ joining events).

To directly demonstrate 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 is provided by the examination of both products of a secondary rearrangement, not just one. To this end it was preferable to isolate an example of a secondary inversion rather than a secondary deletion. Accordingly, the line PD-A (Figure 20) was chosen for the analysis because in order to recombine a second time the substrate in this line would have to invert due to the orientation of all the available target Jκ segments.

To isolate secondary recombinations of the substrate in PD-A, the line was subcloned in selective medium and then one subclone, PDA-1, was passaged

in nonselective medium for approximately a month. It was necessary to passage cells in non-selective 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, I subcloned PDA-1 in nonselective medium and mapped the substrate in these lines as before. The substrate appeared to have rearranged relative to the PD A-1 precursor in three of fifteen subclones analyzed. The line PDA-1-1 (which was mycophenolic acid sensitive) was chosen for further study.

Recombinant A-1-1 was cloned into λ gt wes•λ b as described, and is shown in Figures 20, 21, and 24A. The Vκ joining signal of the original reciprocal joint in PDA-1 had apparently recombined with Jκ2. The DNA sequences of the recombinant junctions in A-1-1 are presented in Figures 21 and 25. As expected, a) the coding joint in A-1-1 was identical to that in recombinant A; b) the Jκ1 heptamer that had been incorporated into the original reciprocal joint was now linked to Jκ2 coding sequences, forming a H1/Jκ2 junction; and c) a new reciprocal joint had been formed from the Jκ2 heptamer sequences and that of Vκ (resulting in an H2/Hv joint).

One can conclude from this that the joining signals in reciprocal junctions are not recombinationally inert, and can participate in secondary recombinations with unrearranged Jκ elements. The enzymes that rearrange immunoglobulin genes can evidently target not only intact gene segments, but also joining signals that have been unlinked from their corresponding coding sequences. Further, the rate with which the inversional event occurs (Table 1) can be measured. It is clear that to a first approximation, the frequency of recombination between "inverted" joining signals is of the same order of magnitude as is necessary to account for rearrangement of

Figure 24

Derivations of recombinants A-1-1 and X.

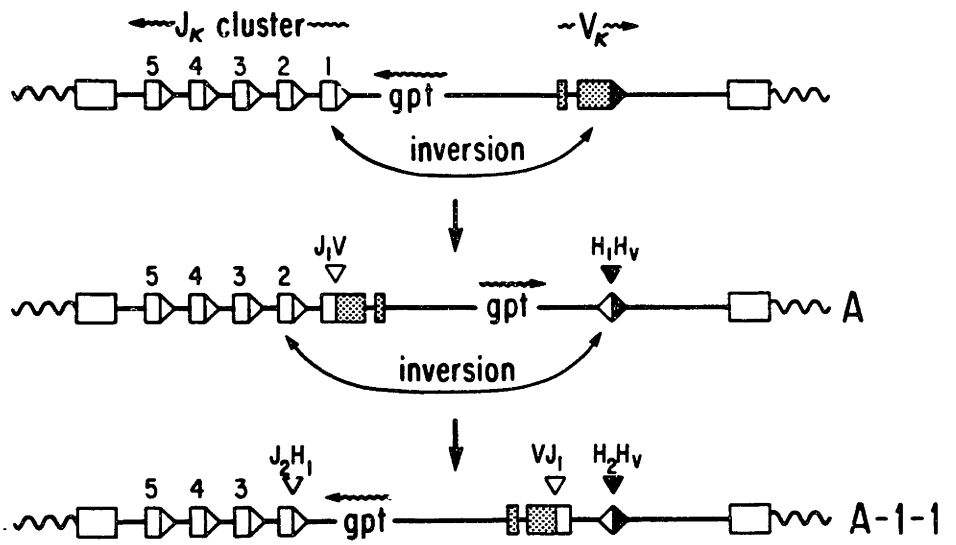
A. Origin of recombinant A-1-1.

Schematic shows the rearrangements that are known to have lead to the final structure of recombinant A-1-1 as described in the text.

B. Probable origin of recombinant X.

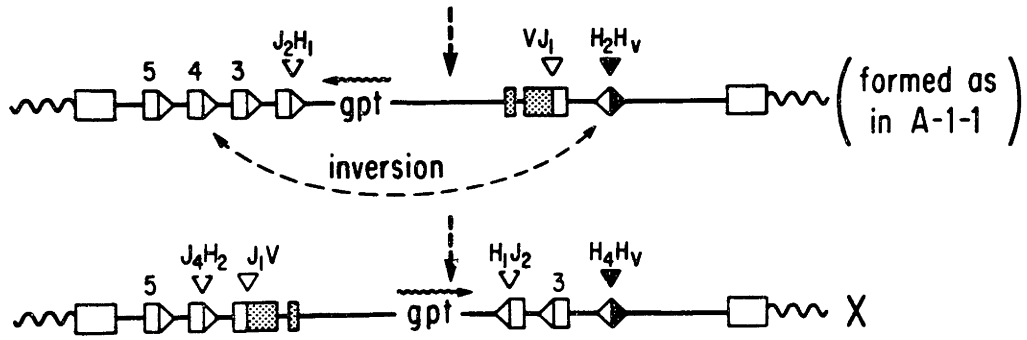
At least three stepwise recombinations must have occurred to form recombinant X. If no more than three rearrangements took place, an unambiguous order can be deduced from the final structure of X. The first two recombinations are as shown for recombinant A-1-1 (Panel A), the last rearrangement is shown in this panel,(panel B).

## Origin of Recombinant A-1-1




---

## Probable Origin of Recombinant X



endogenous genes. One final implication of this experiment is that unequal sister chromatid exchange (see discussion in Chapter III) must not be an obligate pathway for Vκ/Jκ joining. On the basis of its structure, the A-1-1 recombinant evidently arose as a simple, reciprocal, intrachromosomal inversion.

## ASYMMETRY IS AN INTEGRAL FEATURE OF THE JOINING PROCESS

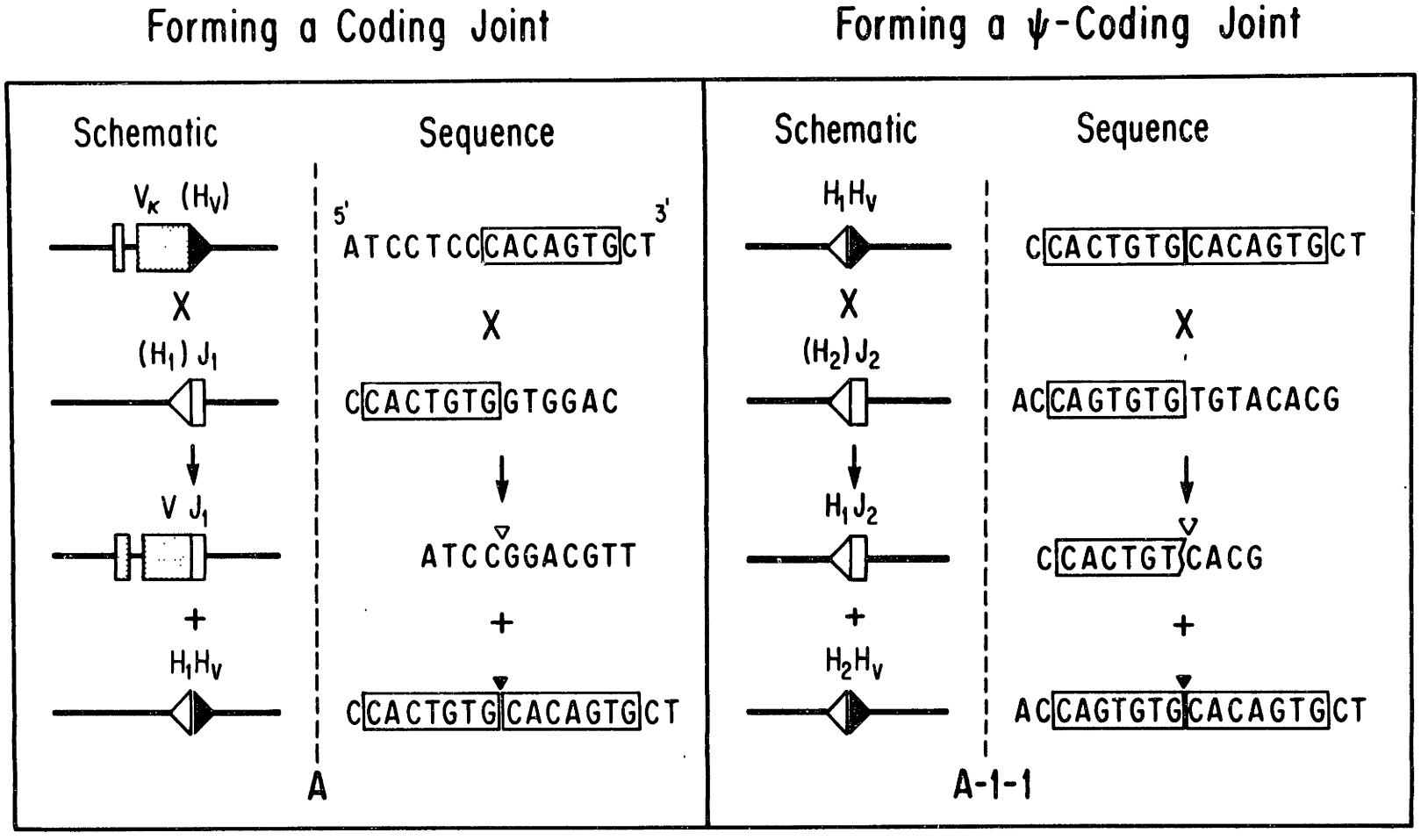
The special case of secondary rearrangement of a reciprocal joint provided the opportunity to determine whether the asymmetry of the joining process --whereby coding joints are variable but reciprocal joints are invariant-- depends upon the sequence of the joined elements or instead reflects an underlying asymmetry in the enzymology. Secondary recombination parallels primary recombination (Figure 24) with the exception that Vκ coding sequences are replaced by a Jκ1 joining signal (H1). 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, one would expect the H1/J2 "pseudo-coding joint" to contain an intact Jκ1-derived joining signal; in the second, one would expect the pseudo-coding joint to contain an H1 heptamer lacking one or more bases at its coding-proximal border.

It is evident from the sequence of the pseudo-coding joint in recombinant A-1-1 (Figures 21 and 24) that during the joining process a G-C base pair was lost from the Jκ1 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.

Figure 25

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 H1 (right) replace those designated V (left).


Forming a Coding Joint
Forming a ψ-Coding Joint
Schematic
Sequence
Schematic
Sequence
$V_K$ ($H_V$)
X
($H_1$) $J_1$
V $J_1$
+
$H_1H_V$
5'
ATCCTCCCACAGTGCT
3'
X
CCACTGTGGTGGAC
ATCCGGACGTT
+
CCACTGTGCACAGTGCT
A
$H_1H_V$
X
($H_2$) $J_2$
$H_1J_2$
+
$H_2H_V$
CCACTGTGCACAGTGCT
X
ACCAGTGTGTGTACACG
CCACTGTCACG
+
ACCAGTGTGCACAGTGCT
A-1-1

The loss of the terminal G from the Jk1 heptamer in recombinant A-1-1 was not adventitious. There are two other examples of pseudo-coding joints in junctions found in recombinant X (Figure 21). One of these represents a fusion of the same two elements that were joined together in A-1-1; the heptamer 5' to Jκ1 was joined to the coding sequences of Jκ2. The other pseudo-coding joint was formed from the 5' flank of Jκ2 and the coding sequences of Jκ4. Both of these pseudo-coding joints were like the pseudo-coding joint in A-1-1 in that the Jκ 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 (H1 and J2). Thus pseudo-coding joints are qualitatively analogous to coding joints; the presence of a Jκ joining signal in place of Vκ 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 Figure 20 and its probable derivation is diagrammed in Figure 24B. The origin of X is of interest because in addition to the V/J1 coding joint, the H1/J2 pseudo-coding joint, and the H2/J4 pseudo-coding joint mentioned in preceeding sections, it has an H4/Hv reciprocal joint. The H4/Hv reciprocal joint in X is exceptional in that it is not a direct union between the heptamer of Jκ4 and that of Vκ. Instead the Jκ4 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, and it is therefore impossible to know whether the joining signal was missing a base previous to, or as a consequence of,

joining. The simplest series of recombinations (as well as the most probable) is a scheme which proceeds identically to the derivation of A-1-1 (Figure 24A) and then in addition includes one final inversion (Figure 24B). The remaining Jκ elements on the substrate were not recombinant, as determined by sequencing the Jκ5 element and inferred from the map to be true of Jκ3 as well. All other regions of X indicated in Figures 20 and 24B, appeared intact.

The junction in recombinant X is the only imprecise reciprocal joint among either the nine examples described here (Figure 21) or the nine endogenously-derived sequences reported previously (Steinmetz et al 1980, Hochtl et al 1982, Hochtl and Zachau, 1983, Selsing et al 1984). Although the terminal G of the heptamer may have been missing from the Jκ4 heptamer prior to joining, this seems unlikely. As mentioned in Chapter I, (GTG)-(CAC) appears at the coding-proximal border of every heptamer that is part of a known-functional joining signal, and this triplet is altered by the missing base in X. A tentative conclusion then is that the G residue was eliminated during recombination, not before. By this view, the junction in recombinant X probably signifies little more than that precise reciprocal joint junctions are formed as a general rule, but this rule is not absolute (Figure 22--solid triangles).

## 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, each exhibited the internal deletion, so these recombinants were sequenced in the deleted region. As indicated in Figure 20, the deletion was found to be a precise excision of the intron between the Vκ leader peptide exon and the Vκ coding sequences proper (data not shown). 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 after inversion.

## GENERAL UTILITY OF RETROVIRALLY-INTRODUCED RECOMBINATION SUBSTRATES

The splice-induced deletions were the only structural variations revealed by the initial mapping experiments that were not due to site-specific recombination. The substrate may have acquired other alterations that are not detected on a gross level, however, the integrity of the substrate appears to be remarkably well-maintained. This system, involving an inversional selection scheme in a substrate that is retrovirally introduced into cells is well-adapted to the study of recombination because 1) no background activation of the selectable marker appears to have occurred and 2) the artifacts introduced by using a retrovirus vector are minimal. It may be that this general approach can be profitably applied to the study of other types of somatic recombination as well.

## MECHANISM OF IMMUNOGLUBULIN GENE REARRANGEMENT

While the present work represents the most extensive compilation of precursor and product sequences from Vκ-to-Jκ joining events, these data are

not readily interpretable in terms of a specific model for variable region exon formation. It is useful, nevertheless to to consider the information collected here in light of schemes that have been previously suggested.

Two features of V/J assembly imply that the mechanism is novel. One unusual feature is the flexibility of coding joint formation. This seems to be an intrinsic property of the process, revealed as a result of the fact that crossing over takes place in a region of no homology. The other unusual feature is that the reciprocal recombination products are not exact reciprocals of one another. Overall, the two products of Vκ-to-Jκ joining have lost some of the base pairs that were 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 (Selsing et al 1984, Alt and Baltimore,1982). 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 Jκ joining signals (Desiderio and Baltimore 1984, Kataoka et al 1984). In addition, duplex breaks are thought to initiate recombination in a variety of other systems (Strathern et al 1982, Szostak et al 1983). If in the present example recombination begins with 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. Instead, coding joints are held in proximity to one another while an exonuclease removes some of their sequences (Alt and Baltimore 1982). The other proposal is that an exonuclease that is specific 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 (Selsing et al 1984). These suggestions are similar in that the coding joint junction is not thought to be formed until the component termini have been trimmed.

The observed structure of pseudo-coding joints rule out the second version (Selsing et al 1984) in its simplest form. Joining signal sequences are evidently not resistant to the putative nuclease (nor are coding sequences specifically sensitive), because pseudo-coding joints can be formed in which heptamers do not remain intact (Figures 21 and 24).

In both of the studies cited above, the structure of truly reciprocal recombination products was inferred from the structures of randomly isolated, non-reciprocal junctions. I have directly demonstrated that bases are indeed missing from the products of a single recombination event (Figure 19). 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 procaryotic elements (Kikuchi and Nash 1979, Kleckner 1981), 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 (see Fig. 26) whereby the bases that are missing in the final

Figure 26 (cont.)

to the structures in Panel 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 non-equivalent 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, Vκ and Jκ, 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 panel C. One can imagine that during the non-reciprocal ligation steps the target V and J elements are designated 3'- or 5'-end donors by the spacer lengths of their joining signals.

Figure 26

Scission and ligation reactions in Vκ-to-Jκ joining could preceed 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 Vκ-to-Jκ joining need not necessarily be a removal from the coding joint alone, nor must an exonuclease necessarily act as an integral, intermediate participant in the process (Alt and Baltimore 1982, Selsing et al 1984).

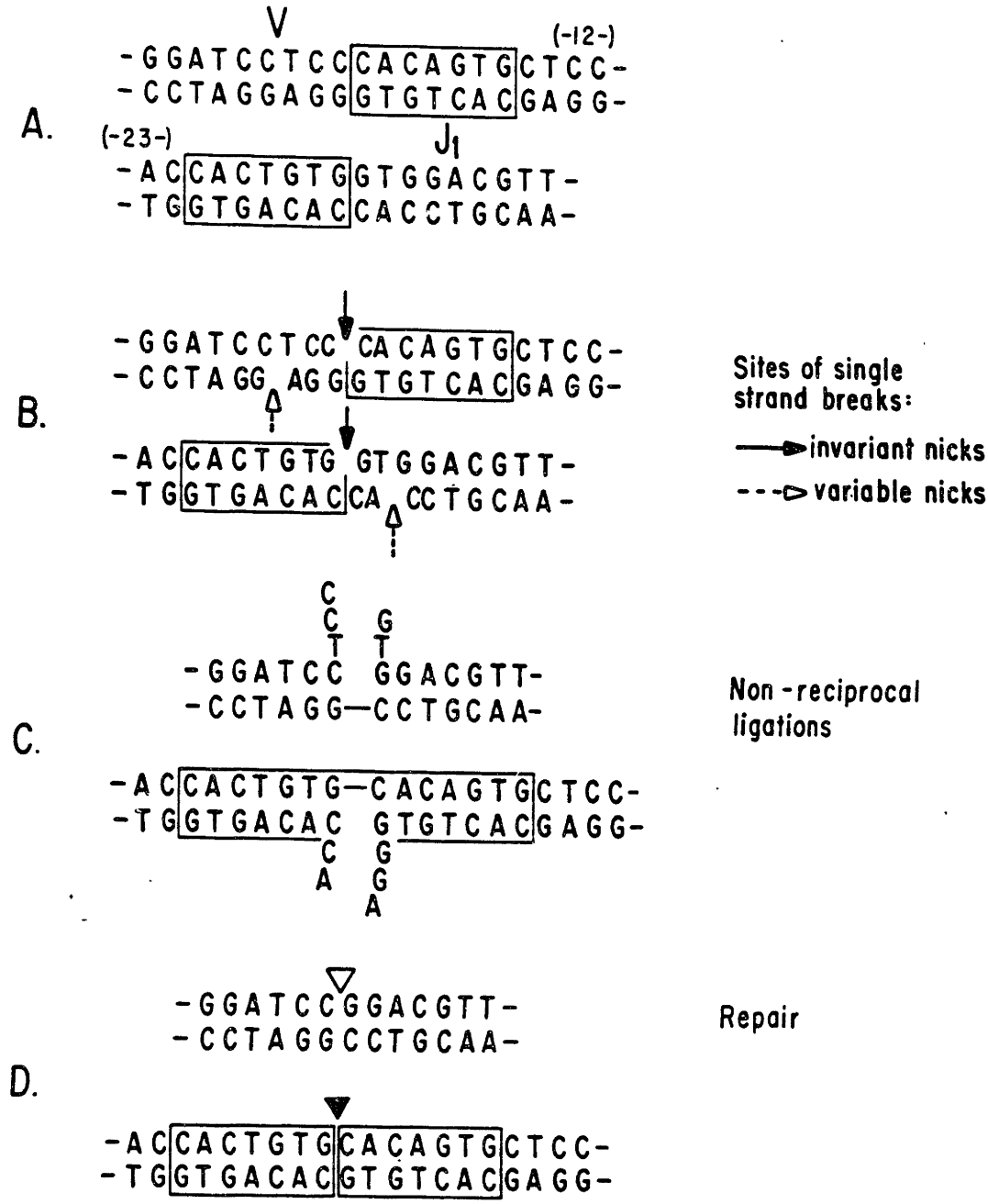
A. Vκ and Jκ precursor sequences. Heptamer sequences are boxed. Sequences to the left of the heptamer (V) or to the right of the heptamer (J1) are the coding sequences. The remainder (including the boxed heptamers) in each line partially represent the joining signals; the length of the spacer region 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 non-reciprocal ligations are indicated. The 3' end of an invariantly-nicked strand from the Vκ joining signal is joined to the 5' end of the invariantly-nicked strand of the Jκ 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 which are each ligated on only one of their two strands. Each duplex contains non-complementary tails opposite the newly formed junction.

D. Maturation of the recombinant junctions. The structures in Panel C could be trimmed and sealed by non-specific 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 and/or tightly coupled steps could be assembled that would lead


A.
V
(-12-)
-GGATCCTCCCACAGTGCTCC-
-CCTAGGAGGGTGTCACGAGG-
(-23-)
J1
-ACCACTGTGGTGGACGTT-
-TGGTGACACCACCTGCAA-
B.
-GGATCCTCCCACAGTGCTCC-
-CCTAGGAGGGTGTCACGAGG-
-ACCACTGTGGTGGACGTT-
-TGGTGACACCACCTGCAA-
Sites of single strand breaks:
invariant nicks
variable nicks
C.
-GGATCCGGACGTT-
-CCTAGG—CCTGCAA-
Non-reciprocal ligations
-ACCACTGTG—CACAGTGCTCC-
-TGGTGACAC GTGTCACGAGG-
D.
-GGATCCGGACGTT-
-CCTAGGCCTGCAA-
Repair
-ACCACTGTGCACAGTGCTCC-
-TGGTGACACGTGTCACGAGG-

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 (Alt and Baltimore 1982, Selsing et al 1984).

The scheme in Figure 26 is not, of course, the only alternative to the previous proposals. To go further, 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 (Kikuchi and Nash 1979, Been and Champoux 1981, Reed 1981, Halligan et al 1982, Krasnow and Cozzarelli 1983, Bullock et al 1984, Mizuuchi 1984). The ubiquity of enzymes that cut and re-ligate DNA by preserving the bond energy in an intermediate in which they are attached to one of the cut ends, has led to the proposal that the term "DNA strand transferases" be applied to the category in general (Mizuuchi 1984). The pattern of cuts and re-ligations shown in Figure 26 provide one conceptual route toward a mechanism that could rely principally on a combination of reciprocal and non-reciprocal reactions catalyzed by this type of enzyme.

SUMMARY

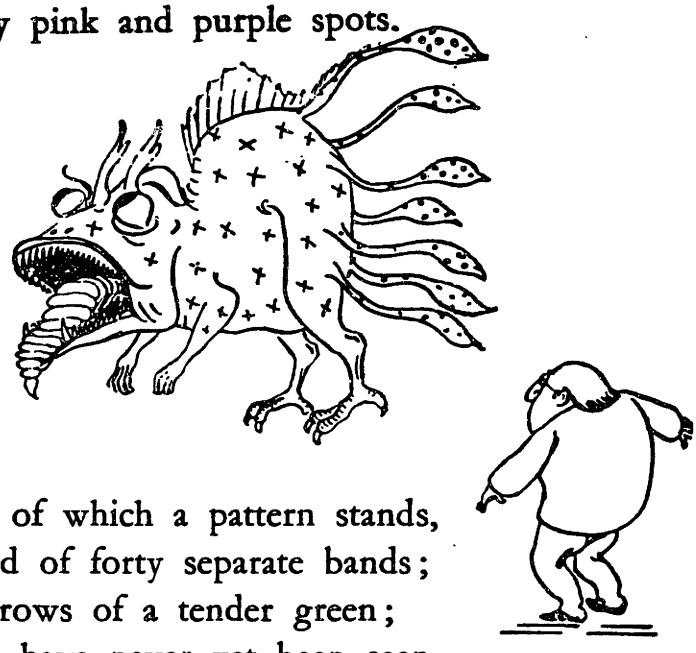
The features of Vκ-to-Jκ recombination (and by extension, recombination of the other antigen-receptor genes in the immune system) are i) coding joints display variable junctions, with no apparent restriction on the association of one particular Vκ crossover with one particular Jk crossover site; ii) 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 Vκ coding sequences; iv) joining elements can recombine in either orientation; v) a Jκ 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.

I have directly tested an inversion/deletion model for kappa gene joining by introducing a substrate that was designed to reflect,in miniature, the way the endogenous kappa locus itself may be arranged; that is with Vκ "backward" relative to Jκ. I have shown 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 Jκ gene segments , and a Jκ1 bias among reciprocal joints). While this is not a critical test, it firmly establishes that the model is feasible and can account for the observations I sought to explain. Other models such as unequal sister chromatid exchange (Van Ness, 1982) or reintegration of DNA that is deleted in the initial

joining step (Steinmetz et al 1980) have not been similarly supported.

# The Microbe

The Microbe is so very small
You cannot make him out at all,
But many sanguine people hope
To see him through a microscope.
His jointed tongue that lies beneath
A hundred curious rows of teeth;
His seven tufted tails with lots
Of lovely pink and purple spots.



On each of which a pattern stands,
Composed of forty separate bands;
His eyebrows of a tender green;
All these have never yet been seen—
But Scientists, who ought to know,
Assure us that they must be so. . . .
Oh! let us never, never doubt
What nobody is sure about! Hillaire Belloc

# IV. MATERIALS AND METHODS

## Cells

All A-MuLv transformed lymphoid cell lines were grown in RPMI medium containing 10% heat-inactivated fetal calf serum and 50μM 2-mercaptoethanol, except for the line ABC-1 which was cultured in Dulbeccoe's Modified Eagle medium supplemented with 10% calf serum. 70Z/3 was carried in RPMI with 10% calf serum and 50 m 2-mercaptoethanol. The Ψ-2 cell line was carried in DME with 10% calf serum.

PD was derived by Dr. N. Rosenberg by infection of adult NIH/Swiss bone marrow cells in vitro as described by Rosenberg and Baltimore (1976), using Friend murine leukemia virus as the helper virus. PD is referred to in some publications as 300-18 (Alt et al 1984, De Pinho et al 1984). The derivation of other cell lines is included in the references cited. For simplicity the subline IATV3H, of ABC-1 was renamed AV3.

All subclonings were performed by limiting dilution, seeding 0.2 cells per well. The number of days that elapsed between single-cell cloning and freezing sample stocks of each subclone was kept to a minimum and typically was around fifteen days. All subclonings were from recently thawed stocks, and records were kept of the number of days that elapsed in culture for each "generation" of cloning.

LPS induction

To induce kappa protein expression in PD sub-lines, $2X10^4$ cells/ml were treated for 48 hours with 50µg/ml of Salmonella typhimurium lipopolysaccharide (Sigma) in standard medium before metabolic labelling.

Transfection

Transfection of the pVJG construct into Ψ-2 was by the calcium phosphate method of Graham and Van der Eb as modified by Parker and Stark (1976) and described in Mann et al (1983). Twenty four hours after seeding approximately $5X10^5$ Ψ-2 cells (a 1:20 split from a confluent plate) onto 100mM plates, 10µg of pVJG DNA and 1 g of pSV2neo DNA were applied as a calcium phosphate precipitate. To form the precipitate, 0.5 ml Hepes buffered saline (137mM NaCl, 5mM KCl, 0.7mM $Na_2HPO_4$, 6mM dextrose and 21mM N-2-hydroxyethylpiperazine-N'-1-ethanesulfonic acid pH 7.0-7.1 adjusted with NaOH, and if necessary, readjusted before use) was mixed with DNA plus 32λ 2M $CaCl_2$ and incubated at room temperature for 30' to 45' until a precipitate visualized as a slight cloudiness had formed. After removing the culture supernatant this was left on the cells for 20', at room temperature. Then the cells were washed once gently with medium and fed with 10ml medium per plate. After 4hrs incubation in a 37°C humidified incubator, cells were osmotically shocked by replacing the medium with 2 ml Hepes buffered saline/15% glycerol, (prewarmed) for 3.5' to 4' (37°C). Following glycerol shock, the cells were washed once with medium, then fed

10 ml medium. After 2 days growth, standard medium was replaced with selective medium.

Transfectants were isolated by selection for the co-transfected plasmid pSV2neo (Southern and Berg 1982). Medium was changed for medium containing 1 mg/ml of the drug G-418 (Colbere-Garapin et al 1981), and cells were fed every three days until visible colonies appeared (about 2 weeks later). Single colonies were picked and tested for production of substrate as described in the text.

Virus

Culture supernatants were collected from subconfluent Ψ-2VJG cells 8-10 hours after medium change and then filtered through 0.45μ filters.

Infections were carried out by suspension of $1x10^5$ PD cells in 1 ml of either freshly collected or thawed virus stocks in the presence of 8μg/ml polybrene (Aldrich) for 1-2 hours (with periodic resuspension).

Mycophenolic acid selection

Mycophenolic acid resistant cells were isolated by plating infected PD cultures into medium containing 0.5μg/ml mycophenolic acid (Lilly) one to two days post-infection. Cells were seeded at a density of $2-5x10^3$ cells per well in microtitre dishes at the start of the selection. The medium contained 250 μg/ml xanthine, 15μg/ml hypoxanthine, and glutamine at 150 μg/ml (Mulligan and Berg 1981, Rice and Baltimore 1982).

GPT gel assay

E. coli XGPRT activity was tested by the in situ gel assay of Mulligan and Berg (1980). To prepare extracts for the assay, approximately $5X10^6$ cells from a subconfluent culture were harvested, washed once in PBS (phosphate buffered saline) and then resuspended in 1 ml ice-cold PBS, made 10mM DTT (dithiothreitol), and 15% glycerol. After disruption by sonication, 0.1 ml of 0.5% deoxycholate, 1% non-idet P40 was added to the lysates. Extracts were clarified by spinning 15' at 4°C in an eppendorf microfuge for 15' following which supernatants were stored at -70°C until use.

Samples were electrophoresed on a 7.5% polyacrylamide gel containing 200mM Tris pH 8.5 that had been polymerized with riboflavin as a catalyst (5Ym/ml final), and had been pre-run for about 10'. (Electrophoresis was performed in a 4°C cold room). The running buffer was 50mM Tris-glycine pH 8.5, samples were loaded in 20mM Tris pH 8.4, 50% sucrose, .001% bromophenol blue.

After electrophoresis, the gel was transferred to a 37°C warm room and the reaction cocktail (100 mMTris pH 7.4, 100μCi/ml $^3$H-Guanine -Schwarz/Mann, specific activity 20 Ci/mMol--, 10mM 5-phophorylribose-1-pyrophosphate--Sigma--, 10mM $MgCl_2$, 2.5mg/ml bovine serum albumin) was spread over its surface. This was allowed to soak in for 5' to 10' and then the gel was wrapped in Saran Wrap (Dow) for the remainder of the 1 hr incubation period. A polyethyleneimine-cellulose plate (Macherey-Nagel & Co.) pre-equilibrated with $H_2O$ (Cambridge municipal water supply) was wiped

free of excess water and laid over the gel for 15'. After this period the PEI plate was washed in de-ionized $H_2O$ and dried. The signal was visualized by fluorography exposed at -70°C on pre-flashed X ray film (Kodak).

Metabolic labeling and immunoprecipitation

$1x10^7$ cells were washed one time in phosphate buffered saline and then resuspended in 2ml of serum-free DME without methionine. After addition of 125 μCi of $^{35}S$ methionine (NEN, specific activity 1Ci/mMol), cells were incubated at 37°C for 1-2 hours. At the end of this period cells were pelleted and lysed in phospholysis buffer (10mM Na $PO_4$ pH 7.5. 10mM NaCl, 10% Triton X-100, 5% deoxycholate, 1% sodium dodecyl sulfate--Witte and Baltimore 1978) with 0.1mM phenylmethylsufonyl fluoride. Lysates were clarified for 1 hr at 100,000 Xg.

$1-3X10^6$ cpm of lysate was incubated overnight on ice with an amount of rabbit antimouse κ serum (gateway) which on the basis of previous titration was known to be in excess. Immune precipitates were collected by incubation with 20μl of 10% formalin-fixed S. aureus (The Enzyme Center, Tufts Medical Center) 1-2 hours on ice in the presence of 5% fetal calf serum. Staph A (S. aureus) pellets were washed one time in phospholysis buffer adjusted to pH 8.0 and 0.5 M NaCl, then washed a second time with phospholysis buffer after transfer to new tubes.

Staph A pellets were resuspended in a reducing sample buffer (60mM Tris-HCl pH 6.8, 5% 2-mercaptoethanol, 2% SDS), boiled for 1'-2', and subjected to SDS polyacrylamide gel electrophoresis according to Laemmli (1970).

For competition immunoprecipitation antiserum plus cold lysates were incubated for one hour at room temperature before addition of labeled extracts. After incubation overnight, samples were prepared as above. The quantity of antiserum and cold lysate was determined in pilot experiments as the minimum amount of antiserum necessary to precipitate a standard sample, and the amount of cold lysate sufficient to compete the signal.

Southern blot analysis

High molecular weight cellular DNA was prepared as described by Alt et al (1980). Approximately 2-4X$10^7$ cells (10 ml of a confluent culture) were washed in phosphate buffered saline and either frozen as cell pellets at -70°C until use or resuspended directly in 2ml of 10mM Tris (ph 7.5-8.0) 5mM EDTA (ethylenediaminetetraacetic acid) 100mM NaCl. The cell suspension was then taken up in a pasteur pipet and quickly squirted into an equal volume of the same buffer freshly made 1% SDS and 400 µg/ml proteinase K, while vortexing on low speed. The lysates were put into 15 ml capped tubes and were allowed to digest at 37°C while slowly rolling in a coffee can (Maxwell House) on a roller bottle apparatus for 4 hours or overnight. Samples were extracted 3X with distilled phenol equilibrated with Tris-HCl pH 8.0, each time slowly rocking for at least 20', then 3x with chloroform.

Following extraction, each sample was precipitated by mixing with an equal volume of isopropanol. The DNA that came out of solution was hooked out with a bent pasteur pipet and allowed to resuspend in 1 ml of 10mM Tris pH 8.0, 1mM EDTA (TE) at room temperature overnight.

DNA digestions typically were carried out with with 1 unit/µg enzyme for four hours, followed by another addition and incubation period. Eco R1 and Bam H1 digestions were in 100 mM Tris 7.4 and 10mM $MgCl_2$. Other buffers were as recommended by the vendor (New England Biolabs). After digestion, samples were adjusted to 0.5% SDS, 12.5% EDTA and 0.4 M NaCl, phenol extracted and precipitated with 2.5 volumes of ethanol.

The digested samples were resuspended in a small volume of TE and electrophoresed on 0.7% agarose gels in 100mM Tris-borate, pH 8.3.

Following electrophoresis, DNA was transferred to nitrocellulose by the method of Southern (1975) as described in Alt et al (1980). The DNA was first denatured in situ by rocking the gel in 0.5M Na OH, 0.6M NaCl for 45', which was then neutralized by two changes (rocking) of 1M Tris pH 7.4, 1.5M NaCl. Transfer was by the capillary method in a buffer of 0.15M Na Citrate, 1.5M NaCl (10X SSC) to nitrocellulose paper, following which the paper was baked in an 80°C vacuum oven for 2 hours.

Transferred filters were prehybridized in 50% formamide, 5X SSCPE (20X SSCPE = 2.4M NaCl 0.3M Na Citrate, 0.2M $KH_2PO_4$, 20mM EDTA) 5X Denhardt's solution (50X Denhardts = 1% BSA, 1% ficoll, 1% polyvinylpyrrolidone), and 500µg/ml boiled herring sperm DNA at 42°C for about eight hours. Hybridization was at 42°C in the same formamide solution adjusted to a

final 10% dextran sulfate (Alwine et al, 1980) and only 100μg/ml carrier herring sperm DNA. $1X10^7$ cpm of probe (usually prepared from whole plasmid DNA rather than gel-purified fragments) that had been labelled to a specific activity of at least $2X10^8$ cpm/μg by the method of Rigby et al (1977) was used.

After hybridization overnight, filters were washed in 2X SSCPE, 0.1% SDS at 67°C with several changes for several hours. Washing procedures were more stringent when probes made from pRI-Jh were used, and included a final wash in 0.1X SSC, 0.1%SDS for one half hour at 67°C as described by Alt et al (1981). Autoradiography was carried out at -70°C using Du Pont "lightning plus" screens.

DNA constructions

pRBJCκ is the 6.8 Eco R1 to Bam H1 fragment of the genomic clone MEMJC2 (Seidman and Leder 1978) in pBR322. pRH5', pH5', pHJκ, and pHBCκ were prepared from pRBJCκ by ligating the appropriate fragment into the corresponding site of pBR322.

pVJG was constructed in several steps. The Eco R1-to-HindIII fragment of Igκ5E (a gift of S. Tonegawa) was subcloned into pBR322 and provided by S. Desiderio. The Eco R1 site 3' to Vκ21-C in this plasmid was changed to a Sal 1 site, and the Bal 1 site in the pBR322 was changed to a Cla 1 site (both by linker addition). This construct was opened up by doubly digesting with Cla 1 and Sal 1, and the fragment containing Vκ21-C plus pBR322 sequences was gel purified. Two fragments were then inserted: the

Bgl II-to-Xho 1 partial prepared from pTRP-3'ss-MSVgpt (a gift of S. Hartman and R. Mulligan) which contains the envelope 3' splice site and the gpt gene, and the Sal 1-to-Cla 1 fragment excised from pHJκ containing the Jκ cluster. A partial BglII-to-Sal 1 digest of the resulting construct liberated the Vκ-gpt-Jκ region, which was then inserted into the Xho I-Bgl II backbone of pMSVgpt (actually pTRP-3'ss-MSVgpt).

Isolation of recombinant VJG

SacI digested cellular DNA was size-selected on salt gradients of 1.25-5.0M NaCl (TE pH 8.0) that were centrifuged for 4 hours at 260,000Xg. A cut of approximately 6-9kb was generally used for ligation into SacI digested, purified, phosphatase treated λ gtwes•λB arms.

Packaging extracts were provided by A. Gifford. Libraries of 200-400,000 phage were plated onto 24.5 X 24.5 bioassay dishes (Nunc) and screened, using standard techniques (Maniatis et al 1983), with nick-translated p-L10, which is a plasmid containing gpt sequences that was originally derived from pPT-1 (Mulligan and Berg, 1980).

Isolation of coding joints and reciprocal joints from PD31-3-4

PD31-3-4 DNA was partially digested with MboI and size selected on sucrose gradients (Maniatis et al 1983). Fragments of an average of 15kb were ligated into purified BamHI digested Charon 30 λ phage arms, packaged, and plated as above. Packaging extracts were provided by N. Andrews. Approximately $2X10^6$ recombinant phage were screened using nick-translated

pRBJCκ as a probe. Several positives were analyzed by southern blot, using Eco RI and BamHI, (with a pRBJCκ probe) and loading digested 31-3-4 DNA alongside as a marker. Clones representing the "5.1 kb"and "6.8 kb" coding joints as well as the approximately 3.0 kb unique reciprocal joint were identified. The 5.1 kb allele was sequenced for about 110 base pairs 5' to the Vκ.Jκ5 junction and about 160 base pairs 3'.

## LITERATURE CITED

Acuto, O., M. Fabbi, J. Smart, C. Poole, J. Protentis, H. Royer, S. Schlossman and E. Reinherz (1984) Purification and $NH_2$-terminal amino acid sequencing of the β subunit of a human T-cell antigen receptor (1984) Proc. Natl. Acad. Sci. USA 81:3851-3855.

Akira, S., H. Sugiyama, N. Yoshida, H. Kikutani, Y Yamamura and T. Kishimoto (1983) Isotype switching in murine pre-B cell lines, Cell 34:545-556.

Allison, J., B. MacIntyre and D. Block (1982) Tumor specific antigen and murine T-lymphoma defined with monoclonal antibody, J. Immunol. 129:2293.

Allwine, J., D. Kemp, J. Parker, J Reiser, G. Stark and G. Wahl (1980) detection of specific RNAs or specific fragments of DNA by fractionation in gels and transfer to diazobenzyloxymethyl paper in Methods in Enzymology 68. R. Wu ed New York: Academic Press pp 220-242.

Alt, F., V. Enea, A Bothwell and D. Baltimore, (1980) Activity of multiple light chain genes in murine myeloma cells producing a single functional light chain, Cell 21:1-12.

Alt, F., N. Rosenberg, S. Lewis, E. Thomas and D. Baltimore (1981) Organization and reorganization of immunoglobulin genes in A-MuLV-transformed cells: rearrangement of heavy but not light chain genes, Cell 27:381-390.

Alt, F. and Baltimore (1982) Joining of immunoglobulin heavy chain gene segments: implications from a chromosome with evidence of three D-Jh fusions. Proc. Natl. Acad. Sci USA 79:4118-4122.

Alt, F. G, Yancopoulos, T. Blackwell, C. Wood, E. Thomas, M. Boss, R. Coffman, N. Rosenberg, S. Tonegawa and D. Baltimore (1984) Ordered rearrangement of immunoglobulin heavy chain variable region segments. EMBO J. 3:1209-1219.

Altenburger, W., M. Steinmetz and H. Zachau (1980) Functional and non-functional joining in immunoglobulin light chain genes of a mouse myeloma, Nature 287:603-607.

Azuma, T., V.Igras, E. Reilly and H. Eisen (1984) Diversity at the variable-joining region boundary of λ light chains has a pronounced effect on immunoglobulin ligand-binding activity, Proc. Natl. Acad. Sci. USA 81:6139-6143.

Beach,D. (1983) Cell type switching by DNA transposition in fission yeast, Nature 305:682-688.

Bentley, D. and T. Rabbitts (1981) Human Vκ immunoglobulin gene number: implications for the origin of antibody diversity. Cell 24:613-623.

Bernard, O., N. Hozumi and S. Tonegawa (1978) Sequences of mouse immunoglobulin light chain genes before and after somatic changes. Cell 15:1133-1144.

Blackwell, T., and F. Alt (1984) Site-specific recombination between immunoglobulin D and Jh segments that were introduced into the genome of a murine pre-B cell line, Cell 34:105-112.

Blomberg, B. and S. Tonegawa (1982) DNA sequences of the joining regions of mouse λ light chain immunoglobulin genes, Proc. Natl. Acad. Sci. USA 79:530-533.

Blomberg, B. A. Traunecker, H. Eisen and S. Tonegawa (1981) Organization of four mouse λ light chain immunoglobulin genes Proc. Natl. Acad. Sci. USA 78:3765-3769.

Broach, J., V. Guarascio and M. Jayaram (1982) Recombination within the yeast plasmid 2μ circle is site-specific, Cell 29:227-234.

Burrows,P., M. LeJeune and J. Kearney (1979) Evidence that murine pre-B cells synthesize μ heavy chains but no light chains, Nature 280:838-840.

Burrows, P., G. Beck and M. Wabl (1981) Expression of μ and γ immunoglobulin heavy chains in different cells of a cloned mouswe lymphoid line Proc. Natl. Acad. Sci. USA 78:564-568.

Burrows, P., G. Beck-Engeser and M. Wabl (1983) Immunoglobulin heavy-chain class switching in a pre-B cell line is accompanied by DNA rearrangement, Nature 306:243-246.

Cebra J., J. Komisar and P. Schweitzer (1984) Ch isotype "switching" during normal B-lymphjocyte development Ann. Rev. Immunol. 2:493-598.

Cheung, S, K. Arndt and P.Lu (1984) Correlation of lac operator DNA imino proton kinetics with its function, Proc. Natl. Acad. Sci. USA 81:3665-3669.

Chien, Y.-h., N. Gascoigne, J Kavaler, N. Lee and Mark Davis (1984) Somatic recombination in a murine T-cell receptor gene, Nature 309:322-326.

Chien, Y-h, D. Becker, T. Lindsten, M. Okamura, D. Cohen and M. Davis (1984) A third type of murine T-cell receptor gene. Nature 312:31-35.

Clark, S.,Y. Yoshikai, S. Taylor, G. Siu, L. Hood and T. Mak (1984) Identification of a diversity segment of human T-cell receptor □-chain and comparison with the analogous murine element. Nature 311:387-389.

Coffman, R. (1983) Surface antigen expression and immunoglobulin gene rearrangement during mouse pre-B cell development Immunol. Rev. 69:5-23.

Coffman, R. and I Weissman (1983) Immunoglobulin gene rearrangement during pre-B cell differentiation. J Mol. Cell Immunol. 1:31.

Colebere-Garapin, F., F. Horodniceanu, P. Kourilsky and A.-C. Garapin (1981) A new dominant hybrid selective marker for higher eucaryotic cells. J. Mol. Biol. 150:1-14.

Coleclough, C (1983) Chance, necessity and antibody gene dynamics, Nature 303:23-26.

Davis,M., Y-h Chien, N. Gascoigne and S. Hedrick (1984) A murine T cell receptor gene complex: isolation, structure and rearrangement, Immunol. Rev. 81:235-258.

Denis, K., L. Treiman, J. St. Claire and O. Witte (1984) Long-term cultures of murine fetal liver retain very early B lymphoid phenotype, J. Exp. Med. 160:1087-1101.

DePinho,R., K. Kruger, N. Andrews, S. Lutzker, D. Baltimore and F. Alt (1984) Molecular basis of heavy-chain class switching and switch region deletion in an Abelson virus-transformed cell line, Mol. Cell. Biol. 4:2905-2910.

Desiderio, S., and D. Baltimore (1984) Double-stranded cleavage by cell extracts near recombinational signal sequences of immunoglobulin genes , Nature 308:860-862.

Desiderio, S. G. Yancopoulos, M. Paskind, E. Thomas, M. Boss, N. Landau, F. Alt, and D. Baltimore (1984) Insertion of N regions into heavy-chain genes is correlated with expression of terminal deoxytransferase in B cells, Nature 311:752-755.

Dreyer, W. and J. Bennett (1965) The molecular basis of antibody formation: a paradox, Proc. Natl. Acad. Sci.54 864-869.

Du Pasquier, L. (1982) Antibody diversity in lower vertebrates--why is it so restricted? Nature 290:311-313.

Durdick, J., M. Moore and E. Selsing (1984) Novel κ light chain gene rearrangements in mouse λ light chain-producing B lymphocytes, Nature 307:749-755.

Early, P. H. Huang, M. Dais, K Calame and L. Hood (1980) An immunoglobulin heavy chain variable region gene is generated from three segments of DNA: Vh, D and Jh. Cell 19:981-992.

Early, P., C. Nottenburg, I. Weissman and L. Hood (1982) Immunoglobulin gene rearrangemnts in normal mouse B cells. Mol. Cell. Biol. 2:829-836.

Emorine, L. K. Dreher, T. Kindt and E. Max (1983) Rabbit immunoglobulin κ genes: structure of a germline b4 allotype J-C locus and evidence for several b4-related sequences in the rabbit genome. Proc. Natl. Acad. Sci. USA 80:5709-5713.

Gascoigne, N. Y.-h. Chien, D. Becker, J. Kavaler and M. Davis (1984) Genomic organization and sequence of T-cell receptor β-chain constant- and joining-region genes, Nature 310:387-391.

Graham, R., and A. Van der Eb (1973) A new technique for the assay of infectivity of human adenovirus 5 DNA, Virol. 52:456-467.

Hedrick,S., E. Nielsen, J. Kavaler, D. Cohen and M.Davis (1984) Sequence relationships between putative T-cell receptor polypeptides and immunoglobulins, Nature 308:153-158.

Heidmann, O. and F. Rogeon (1983) Diversity in the rabbit immunoglobulin kappa chain variable regions is amplified by nucleotide deletions and insertions at the V-J junction. Cell 34:767-777.

Heinrich, G., A. Traunecker and S. Tonegawa (1984) Somatic mutation creates diversity in the major group of mouse immunoglobulin κ light chains, J. Exp. Med. 159:417-435

Heiter, P. S. Korsmeyer, T. Waldmann and P.Leder (1981) Human immunoglobulin κ light-chain genes are deleted or rearranged in λ-producing B cells Nature 290:368-372.

Hochtl, J., C. Muller and H. Zachau (1982) Recombined flanks of the variable and joining segments of immunoglobulin genes. Proc. Natl. Acad. Sci. USA 79:1383-1387.

Hochtl, J. and H. Zachau (1983) A novel type of aberrant recombination in immunoglobylin genes and its implications for the V-J joining mechanism Nature 302:260-263.

Honjo T.( 1983) Immunoglobulin genes. Ann. Rev. Immunol. 1:499-528.

Hozumi, N. and S. Tonegawa (1976) Evidence for somatic rearrangement of immunoglobulin genes coding for variable and constant regions, Proc. Natl. Acad. Sci 73:3628-3632.

Jeske, D., J. Jarvis, C. Milstein and J. Capra (1984) Junctional diversity is essential to antibody activity J. Immunol. 133:1090-1093.

Joho R., C. Nottenburg, R. Coffman, and I. Weissman (1983) Rearrangement and expression of immunoglobulin genes. Curr Top. Dev. Biol. 18:15-58.

Kabat, E., T. Wu, H. Bilofsky, M. Reid-Miller and H. Perry (1983) Sequences of proteins of immunological interest. U.S. Dept. of Health and Human Serivices, NIH.

Kappler, J. R. Kubo, K. Haskins, C. Hannum, P. Marrack, M. Pigeon, B. MacIntyre, J. Allison and I Trowbridge (1983) The major histocompatibility complex-restricted antigen receptor on T cells in mouse and man: Identification of constant and variable peptides. Cell 34:295.

Kavaler, J. M. Davis, and Y-h Chien (1984) Localization of a T-cell receptor diversity-region element, Nature 310:421-423.

Klein, G., S. Ohno, N. Rosenberg, F. Weiner, J. Spira and D. Baltimore (1980) Cytogenetic studies on Abelson-virus-induced mouse leukemias, Int. J. Cancer 25:805-811.

Korsmeyer, S., P. Heiter, J. Ravetch, D. Poplack, G. Waldmann and P. Leder (1981) Developmental hierarchy of immunoglobulin gene rearrangements in human leukemic pre-B-cells. Proc. Natl. Acad. Sci. USA 78:7096-7100.

Korsmeyer, S. A. Arnold, A. Bakshi, J. Ravetch, U. Siebenlist, P. Heiter, S. Sharrow, T. LeBien, J. Kersey, D. Poplack, P. Leder and T. Waldmann (1983)

Immunoglobulin gene rearrangement and cell surface antigen expression in acure lymphocytic leukemias of T cell and B cell precursor origins, J. Clin, Invest. 71:301-313.

Krasnow, M. and N. Cozzarelli (1983) Site-specific relaxation and recombination by the Tn3 resolvase: recognition of the DNA path between oriented res sites. Cell 32:1313-1324.

Kubagawa, H. M. Mayumi, W. Crist and M. Cooper (1983) Immunoglobulin heavy chain switching in pre-B leukemias, Nature 301:340-342.

Kurosawa, Y., H. von Boehmer, W Haas, H. Sakano, A. Traueneker and S. Tonegawa (1981) Identification of D segments of immunoglobulin heavy-chain genes and their rearrrangement in T lymphocytes. Nature 290:565-570.

Kurosawa, Y. and S. Tonegawa (1982) Organization, structure, and assembly of immunoglobulin heavy chain diversity DNA segments. J. Exp. Med. 155:201-218.

Korsmeyer, S. P. Hieter, S. Sharrow, C. Goldman, P. Leder and T. Waldmann (1982) Normal human B cells display ordered light chain gene rearrangements and deletions, J. Exp. Med 156:975-983.

Laemmli, U. (1970) Cleavage of structural proteins during the assembly of the head of bacteriophage T4, Nature 227:650-655.

Leder, P., E. Max, J. Seidman, S.-P Kwan, M. Scharff, M. Nau and B. Norman (1981) Recombination events that activate, diversify, and delete immunoglobulin genes Cold Spring Harb. Symp. quant. Biol. 45:859-865.

Levitt, D. and M. Cooper (1980) Mouse pre-B cells synthesize and secrete μ heavy chains but not light chains Cell 19:617-625.

Lewis, S. N. Rosenberg, F. Alt and D. Baltimore (1982) Continuing kappa-gene rearrangement in a cell line transformed by Abelson murine leukemia virus, Cell 30:807-816.

Lewis, S., A. Gifford and D. Baltimore (1984) Joining of Vκ to Jκ gene segments in a retroviral vector introduced into lymphoid cells, Nature 308:425-428.

Litman, G., L. Berger, K. Murphy, R. Litman, K. Hinds, C. Jahn and B. Erikson (1983) Complete nucleotide sequence of an immunoglobulin $V_h$ gene homologue from Caiman, a phylogenetically ancient reptile. Nature :349-352.

Mak, T. and Y. Yanagi (1984) Genes encoding the human T cell antigen receptor, Immunol. Rev. 81:235-258 (1984).

Maki, R., J. Kearney, C. Paige and S. Tonegawa (1980) Immunoglobulin gene rearrangement in immature B cells. Science 209:1366-1369.

Mann, R., R. Mulligan and D. Baltimore (1983) Construction of a retrovirus packaging mutant and its use to produce helper-free defective retrovirus, Cell 33:153-159.

Max, E., J. Seidman and P. Leder (1979) Sequences of five potential recombination sites encoded close to an immunoglobulin kappa constant region gene, Proc. Natl. Acad. Sci. 76:3450-3457.

Max, E., J. Seidman, H. Miller and P. Leder (1980) Variation in the crossover point of kappa immunoglobulin gene V-J recombination: evidence from a cryptic gene. Cell 21:793-799.

Malissen, M., K. Minard, S. Mjoisness, M. Kronenberg, J. Goverman, T. Hunkapiller, M. Prystowsky, Y. Yoshikai, F. Fitch, T. Mak and L. Hood (1984) Mouse T cell antigen receptor: structure and organization of constant and joining gene segments encoding the β polypeptide, Cell 37:1101-1110.

Manser,T., S. Huang and M. Gefter, (1984) Influence of clonal selection on the expression of immunoglobulin variable region genes, Science 226:1283-1288.

Melchers, F., J. Andersson and R. Phillips (1977) Ontogeny of murine B lymphocytes: development of Ig synthesis and of reactivities to mitogens and to anti-Ig-antibodies, Cold Spring Harb. Symp. Quant. Biol. 41. New York pp.147-158.

Miller, J., A. Bothwell and U. Storb (1981) Physical linkage of the constant region genes for immunoglobulins λ1 and λ2 Proc. Natl. Acad. Sci USA 78:3829-3833.

Mizuuchi, K. (1984) Mechanism of transposition of bacteriophage Mu: polarity of the strand transfer reaction at the initiation of transposition, Cell 39:395-404.

Mulligan R. and P. Berg (1980) Expression of a bacterial gene in mammalian cells Science 209:1422-1427.

Mulligan, R. and P. Berg (1981) Selection for animal cells that express the E. coli gpt gene coding for xanthine-guanine phosphoribosyltransferase. Proc. Nat. Acad. Sci. USA 78:2072-2076.

Mulligan R. (1983) Construction of highly transmissable mammalian cloning vehicles derived from murine retroviruses. In Experimental Manipulation of Gene Expression, M. Inouye, ed. (New York: Academic Press), pp. 155-173.

Nash, H.A (1981) Annu Rev. Genet. 15:143-167.

Paige, C., P. Kincade and Peter Ralph (1978) Murine B cell leukemia line with inducible surface immunoglobulin expression Jour. Immunol 121:641-647.

Paige, C., P. Kincade and Peter Ralph (1981) Independant control of immunoglobulin heavy and light chain expression in a murine pre-B-cellline Nature 292:631-633.

Palacios, R., G. Henson, M. Steinmetz and J. McKearn (1984) Interleukin-3 supports growth of mouse pre-B-cell clones in vitro, Nature 309:126-131.

Parker, B. and G. Stark (1979) Regulation of simian virus 40 transcription: sensitive analysis of the RNA species present early in infections by virus or viral DNA, J. Virol 31:360-369.

Patten, P., T. Yokota, J. Rothbard, Y. Chien, K.Arai and M. Davis (1984) Structure, expression and divergence of T-cell receptor β-chain variable regions, Nature 312:40-46.

Pernis, B., G. Chiappino, A. Kelus and P. Gell (1965) Cellular localization of immunoglobulin with different allotype specificities in rabbit lymphoid tissues, J. Exp. Med 122:853-876.

Pernis, B., L. Fourni and L. Amante (1970) Immunoglobulin spots on the surface of rabbit lymphocytes, J. Exp. Med 132:1001-1018.

Perry, R., D. Kelley, C. Coleclough and J. Kearney (1981) Organization and expression of immunoglobulin genes in fetal liver hybridomas, Proc. Nat. Acad. Sci. USA. 78:247-251.

Pratt, D., J. Strominger, R. Parkman, D. Kaplan, J. Schwaber, N. Rosenberg,and C. Scher (1977) Abelson Virus-transformed lymphocytes: null cells that modulate H-2 Cell 12:683-690.

Reilly, E., B. Blomberg, T. Imanishi-Kari, S. Tonegawa and H. Eisen (1984) Restricted association of V and J-C gene segments for mouse λ light chains. Proc. Natl. Acad. Sci. USA 81:2484-2488.

Rice, D. and D. Baltimore (1982) Regulated expression of an immunoglobulin κ gene introduced into a mouse lymphoid cell line, Proc. Natl. Acad. USA 79:7862-7865.

Rigby, P., M. Dieckmann, C. Rhodes and P. Berg (1977) Labelling deoxyribonucleic acid to high specific activity in vitro by nick-translation with DNA polymerase I, J. Mol. Biol. 113:237-251.

Riley, S. e. Brock and W. Kuehl (1981) Induction of light chain expression in a pre-B cell line by fusion to myeloma cells Nature 289:804-806.

Rose, S., M. Kuehl and G. Smith (1977) Cloned MPC11 myeloma cells express two kappa genes: a gene for a complete light chain and a gene for a constant region polypeptide. Cell 12:453-462.

Rosenberg N., and D. Baltimore (1976) A quantitative assay for transformation of bone marrow cells by Abelson murine leukemia virus. J. Exp. Med 147:1453-1463.

Royer, H. A Bensussam, O. Acuto and E. Reinherz (1984) Functional isotypes arre not encoded by the constant region genes of the β subunit of the T cell receptor for antigen/ major histocompatibility complex, J. Exp. Med. 160: 947-952.

Ruley, E. and M. Fried (1983) Clustered illegitimate recombination events in mammalian cells involving very short sequence homologies Nature 304:181-184.

Saito, H. D. Kranz, Y. Tagaki, A. Hayday, H. Eisen, and S Tonegawa (1984) A third rearranged and expressed gene in a clone of cytotoxic T lymphocytes. Nature 312:36-40.

Sakano, S. K. Huppi, G. Heinrich and S. Tonegawa (1979) Sequences at the somatic recombination sites of immunoglobulin light-chain genes, Nature 280:288-294.

Sakano, H. R. Maki, Y Kurosawa, W. Roeder, and S. Tonegawa (1980) Two types of somatic recombination are necessary for the generation of complete immunoglobulin heavy-chain genes. Nature 286:676-683.

Sakano, H., Y. Kurosawa, M Weigert and S. Tonegawa (1981) Identification of D segments of immunoglobulin heavy-chain genes and their rearrangement in T lymphocytes. Nature 290:562-565.

Samelson, L. R. Germaine and R. Schwartz (1983) Monoclonal antibodies against the antigen receptor on a cloned T-cell hybrid Proc. Natl. Acad. USA 80:6972.

Seidman, J. and P. Leder (1978) The arrangement and rearrangement of antibody genes Nature 276:790-795.

Seidman, J., E. Max and P. Leder (1979). A κ immunoglobulin gene is formed by site-specific recombination without further somatic mutation. Nature 280:370-375.

Seidman, J., M. Nau, B. Norman, S.-P. Kwan, M. Scharff and P. Leder (1980), Immunoglobulin V/Jκ recombination is accompanied by deletion of joining site and variable region segments. Proc. Nat. Acad. Sci. USA 77:6022-6026.

Seidman J. and P. Leder (1980) A mutant immunoglobulin light chain is formed by abberrant DNA- and RNA-splicing events Nature 286:779-783

Selsing, E., J. Voss and U. Storb (1984) Immunoglobulin gene 'remnant' DNA--implications for antibody gene recombination, Nucleic Acids Res. 12:4229-4246.

Siden, E., D. Baltimore, D. Clark and N. Rosenberg (1979) Immunoglobulin synthesis by lymphoid cells transformed in vitro by Abelson murine leukemia virus, Cell 16:389-396.

Siden, E., F. Alt, L Shinefeld, V Sato and D. Baltimore (1981) Synthesis of immunoglobulin μ chain gene products precedes synthesis of light chians during B-lymphocyte development, Proc. Natl. Acad. Sci. USA 78:1823-1827.

Simon,M. and M.Silverman (1983) Recombinational regulation of gene expression in bacteria, p.211-227 in Gene Function in Procaryotes, ed. J. Beckwith, Cold Spring Harbor, N.Y.

Siu, G. M. Kronenberg, E. Strauss, R. Haars, T. Mak and L. Hood (1984) The structure, rearrangement and expression of $D_\beta$ gene segments of the murine T-cell antigen receptor. Nature 311:344-350.

Southern, J. (1975) Detection of specific sequences among DNA fragments separated by gel electrophoresis, J. Mol. Biol. 38:503-517.

Southern, P. and P. Berg (1982) Transformation of mammalian cells to antibiotic resistance with a bacterial gene under control of the SV40 early region promoter J. Mol. and App. Genet. 1:327-341.

Strathern, J., A. Klar, J. Hicks, J. Abraham, J. Ivy, K. Nasmyth and C. McGill (1982) Homothallic switching of yeast mating type cassettes is initiated by a double stranded cut in the MAT locus, Cell 31:183-192.

Sugiyama, H., S Akiro, N. Yoshida, S. Kishimoto, Y. Yamamura, P. Kincade, T. Honjo and T. Kishimoto (1982) Relationship between the rearrangement of immunoglobulin genes, the appearance of a B lymphocyte antigen, and immunoglobulin synthesis in murine pre-B cell lines, J. Immunol. 126:2793-2797.

Sugiyama, H., S. Akira, H. Kikutani, S. Kishimoto, Y. Yamamura and T. Kishimoto (1983) Functional V region formation during in vitro culture of a murine immature B precursor cell line, Nature 303:812-815.

Teich, N. M. Boss and T. Dexter (1978) Infection of mouse bone marrow cells with Abelson Murine Leukemia virus and establishment of producer cell lines in Moder Trands in Human Leukemia III, ed R. Neth and R. Gallo. Springer-Verlag, Berlin, 1979 pp.487-490.

Tonegawa, S., C. Brack, N. Hozumi and V. Pirotta (1977) Organization of immunoglobulin genes, Cold spring Harbor Symp. Quant. Biol. 42, pp.921-931.

Tonegawa, S.(1983) Somatic generation of antibody diversity Nature 302:575-581.

Tsukamoto, A., I. Weissman and S. Hunt (1984) Allelic exclusion in rat kappa immunoglobulin chains: extent of Jκ rearrangement in normal B lymphocytes, EMBO J. 3:975-981.

Valbuena, O., K. Marcu, M. Weigert and R Perry (1978) Multiplicity of germline genes specfying a group of related mouse κ chains with implications for the generation of immunoglobulin diversity. Nature 276:280-283.

Vogler, L., J. Preud'homme, M. Seligmann, W. Gathings, W. Christ, M. Cooper and F. Bollum (1981) Diversity of immunoglobulin expression in leukaemic cells resembling B-lymphoyte precursors, Nature 290:339-341.

Wabl, M., G. Beck-Engeser and P. Burrows (1984) Allelic inclusion in the pre-B-cell line 18-81, Proc. Nat., Acad, Sci. USA 81:876-870.

Wall, R., and M. Kuehl Biosythesis and regulation of immunoglobulins (1983) Ann. Rev. Immunol. 1:393-422.

Weigert, M. R. Perry , D. Kelly, T. Hunkapiller, J. Schilling and L. Hood (1980) The joining of V and J gene segments creates antibody diversity Nature 283:497-499.

Whitlock, C., and O. Witte (1982) Long term culture of B lymphocytes and their precursors from murine bone marrow. Proc. Natl. Acad. Sci. 79:3608-3612.

Whitlock, C., S. Zeigler, L. Treiman, J. Stafford and O Witte (1983) Differentiation of cloned populations of immature B cells after transformation with Abelson murine leukemia virus, Cell 32:903-911.

Witte, O. and D. Baltimore (1978) Relationship of retrovirus polyprotein cleavages to virion maturation studied with temperature-sensitive murine leukemia virus mutants J. Virol, 26:750-761.

Wood, D. and C. Coleclough (1984) Different joining region J elements of the murine κ immunoglobulin light chain locus are used at markedly different frequencies, Proc. Natl. Acad. Sci. USA (1984)

Wood, C., and S. Tonegawa (1983) Diversity and joining segments of mouse immunoglobulin heavy chain genes are closely linked and in the same orientation: implications for the joining mechanism, Proc. Natl. Acad. Sci. USA 80:3030-3034.

Zeigler, S., L. Treiman and O. Witte (1984) κ gene diversity among the clonal progeny of pre-B lymphocytes, Proc. Natl. Acad., Sci, USA 81:1529-1533.

Cell, Vol. 30, 807–816, October 1982, Copyright © 1982 by MIT



# Continuing Kappa-Gene Rearrangement in a Cell Line Transformed by Abelson Murine Leukemia Virus

**Susanna Lewis, Naomi Rosenberg,***
**Frederick Alt† and David Baltimore**
Whitehead Institute for Biomedical Research,
and Center for Cancer Research
and Department of Biology
Massachusetts Institute of Technology
77 Massachusetts Avenue
Cambridge, Massachusetts 02139
* and Cancer Research Center
and Department of Pathology
Tufts University
55 Kneeland Street
Boston, Massachusetts 02111

## Summary

**A cell line transformed by Abelson murine leukemia virus, called PD, is capable of carrying out κ-gene rearrangement while growing in culture. Subclones of PD have diverse κ-gene structures, and some derivatives show evidence of continued joining activity after as many as three subclonings. Analysis of PD sublineages has shown that a rearranged chromosome can undergo secondary κ-gene rearrangements, producing either a new rearrangement or a deletion of $C_\kappa$. Although the PD line actively rearranges its κ genes, its rearranged heavy-chain genes show little variation, and there is no rearrangement of λ genes. In PD subclones, DNA fragments representing the reciprocal product of κ-gene rearrangement are often evident, and they may undergo either further rearrangement or deletion. The implications of multiple rearrangements on a single chromosome and of the maintenance of reciprocal fragments are considered in the context of a model that postulates that the $V_\kappa$ and $J_\kappa$ segments are not all organized in the DNA in the same transcriptional direction, leading to inversions rather than deletions during joining.**

## Introduction

Rearrangement of the DNA encoding immunoglobulin κ chains during B-cell differentiation has been inferred from comparisons of germline DNA with myeloma cell DNA (Lenhard-Schuller et al., 1978; Seidman and Leder, 1978; Sakano et al., 1979; Seidman et al., 1979). Sequence analysis of myeloma immunoglobulin genes has allowed a fairly precise specification of the consequences of rearrangement, but many aspects of the rearrangement process itself have not been systematically investigated because recombination of myeloma immunoglobulin genes after establishment of a transplantable tumor occurs rarely, if at all.

From the previous analyses, it is known that the κ locus on chromosome 6 of the mouse (Swan et al., 1979) has a large number (probably hundreds) of variable ($V_\kappa$) regions, four functional and one nonfunctional joining ($J_\kappa$) regions and a single constant ($C_\kappa$) region. By bringing together a $V_\kappa$ segment with a $J_\kappa$ segment, the cell constructs the complete κ variable region several thousand base pairs upstream from $C_\kappa$ (Seidman and Leder, 1978; Lenhard-Schuller et al., 1978; Sakano et al. 1979; Seidman et al., 1979). The process of joining $V_\kappa$ to $J_\kappa$ can sometimes be correlated with the deletion of DNA sequence by the demonstrated absence of a subset of $J_\kappa$ segments and $V_\kappa$ genes in the genomes of certain myelomas (Seidman et al., 1980). Other evidence, however, indicates that as often DNA lying 5′ to a joined $J_\kappa$ is retained in a reorganized form (Steinmetz et al., 1980). The structure of the reorganized 5′ sequences suggests that they may be generated during the process of $V_\kappa$–$J_\kappa$ joining, perhaps representing reciprocal recombination products (Steinmetz et al., 1980; Hochtl et al., 1982). A given cell line, however, often has a smaller number of these putative reciprocal fragments than rearranged κ constant regions (Steinmetz et al., 1980; Selsing and Storb, 1981; Van Ness et al., 1982; Hochtl et al., 1982). In addition, sequence analysis of several such apparent reciprocal fragments showed that they were not the reciprocal product of an assembled κ gene in the same cell (Steinmetz et al., 1980; Hochtl et al., 1982). These observations have led to the proposal that $V_\kappa$–$J_\kappa$ joining involves unequal sister-chromatid exchange (Van Ness et al., 1982).

A more detailed examination of κ-gene rearrangement requires a biologic system in which $V_\kappa$–$J_\kappa$ rearrangement can be studied as an ongoing event. Most Abelson murine leukemia virus (A-MuLV) transformants derived by in vitro infection of adult bone marrow cells bear stably rearranged heavy-chain genes along with germline κ and λ constant regions (Alt et al., 1981b). A few lines, however, are capable of rearranging their κ loci in culture (Alt et al., 1981b; Riley et al., 1981). We report here the properties of one such cell line, called PD.

Study of subclones derived from PD has demonstrated that once a chromosome has undergone one rearrangement, it can still undergo further rearrangements. Furthermore, the presence of reciprocal fragments from the rearrangement process is evident in the PD lineage, and secondary loss of such fragments has been shown. We suggest that the formation of reciprocal fragments could be a consequence of the relative orientation of the $V_\kappa$ and the $J_\kappa$ segments.

## Results

### Light-Chain Gene Rearrangement during In Vitro Culture

The A-MuLV-transformed line PD was derived by in vitro infection of bone marrow cells from an adult NIH/

† Present address: Department of Biochemistry, College of Physicians and Surgeons, Columbia University, New York, New York 10027.

Swiss mouse (Rosenberg and Baltimore, 1976). The transformant was isolated as a single focus growing in soft agar, adapted to liquid culture and carried for about 6 weeks before being frozen. When PD cells from an early freezing were thawed and subcloned, about half of the derivatives displayed light-chain proteins that had not been detected in the uncloned PD culture (N. Rosenberg, unpublished observations). The light-chain proteins produced by different subclones migrated differently during electrophoresis through SDS-polyacrylamide gels (Figure 1). The species detected in metabolically labeled lysates by immunoprecipitation with a rabbit antimouse κ serum could be competed by an MPC11 lysate but not by a MOPC104E lysate, confirming the identity of these proteins as the κ light-chain isotype (MPC11 contains κ protein but no λ; MOPC104E produces λ light chain but no κ, data not shown).

To examine the structure of the κ light-chain genes in the subclones of PD (which will be referred to as primary subclones), we extracted DNA from each subclone and doubly digested it with Bam HI and Eco RI. The $J_κ$- and $C_κ$-coding sequences are entirely contained within a 6.8 kb fragment when unrearranged DNA is digested with these enzymes (Alt et al., 1980). Therefore, with an appropriate probe, $V_κ$–$J_κ$ joining is usually detectable as an alteration in the size of the band generated by Bam HI and Eco RI codigestion. Rearrangement was assayed by hybridization to nick-translated DNA from either plasmid pRBJC$_κ$, which contains $C_κ$ and $J_κ$, or with pHBC$_κ$, which contains $C_κ$ but not $J_κ$ (Figure 2).

When DNA from the uncloned PD line was analyzed with either probe, a single 6.8 kb band was observed (Figure 3A, lane PD; other data not shown) that migrated coincidentally with the fragment from BALB/c liver DNA (Figures 3A and 3B, lane L). Unlike liver DNA, the PD line also displayed a smear of indistinct lower molecular weight bands. When 13 primary subclones from PD were analyzed with one or both probes, 10 had strongly hybridizing bands at positions different from the 6.8 kb germline band and different from one another (Figures 3A and 3B; the extra bands detected specifically by pRBJC$_κ$ will be discussed below). It is evident from this result that the PD line accumulated heterogeneity at its κ locus sometime after it was isolated. Because those lines with apparently unrearranged κ genes (PD31, PD35 and PD40) contained no detectable κ protein, while the lines with rearranged κ genes had κ proteins (Figures 1 and 3), many of the observed rearrangements appear to be the result of in-phase $V_κ$–$J_κ$ joining. Not all of the κ rearrangements we have observed have led to κ-protein production, and the fraction of productive rearrangements is under investigation.

Because PD had been in culture for several weeks before its first subcloning, the diversity displayed among the primary subclones could have resulted from a single burst of recombination activity. To determine whether the line could continue to rearrange during in vitro culture, we cloned seven of the primary subclones again; 7 to 20 secondary subclones were derived from each. DNA from these secondary subclones was analyzed for κ-gene rearrangement as described above. Striking examples of continued rearrangement were found in subclones of PD31 and of PD40 (Figure 4 shows the PD31 series). Although uncloned PD31 or PD40 had a prominent embryonic band (Figure 3A and 3B), nearly all of the secondary subclones had undergone diverse rearrangements at the κ loci. Other primary subclones analyzed in this manner (PD34, PD36 and PD42) similarly showed evidence of continued rearrangment in culture. Based on the extent of the variability recovered among the

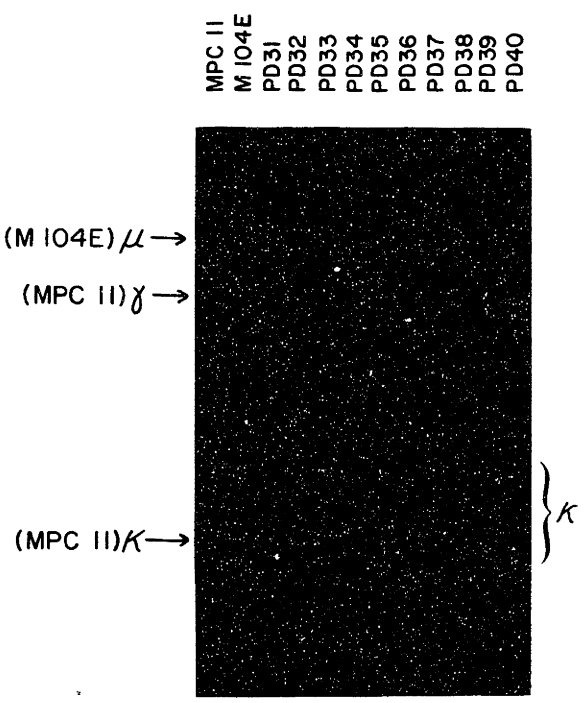



Figure 1. Protein Production in PD Primary Subclones

$^{35}$S-methionine-labeled lysates from metabolically labeled cells were immunoprecipitated with antimouse immunoglobulin serum as described by Rosenberg and Witte (1980). The MPC11 γ and κ proteins and MOPC104E μ protein are indicated. Bracket various κ proteins produced by the indicated primary PD subclones.

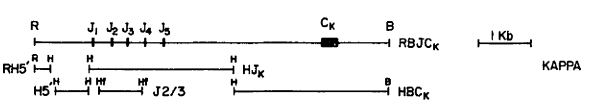



Figure 2. Probes Used for the Analysis of $J_κ$ Rearrangements in PD Derivatives

Shown are κ subclones in pBR322. The $J_κ$ and $C_κ$ regions in embryonic BALB/c DNA are entirely contained between Eco RI and Bam HI sites, defining the region present in plasmid pRBJC$_κ$. The fragments represented in the subcloned plasmids pRH5′, pH5′, pHJ$_κ$, pJ2/3 and pHBC$_κ$ are indicated. R: Eco RI. B: Bam HI. H: Hind III. Hf: Hinf II.

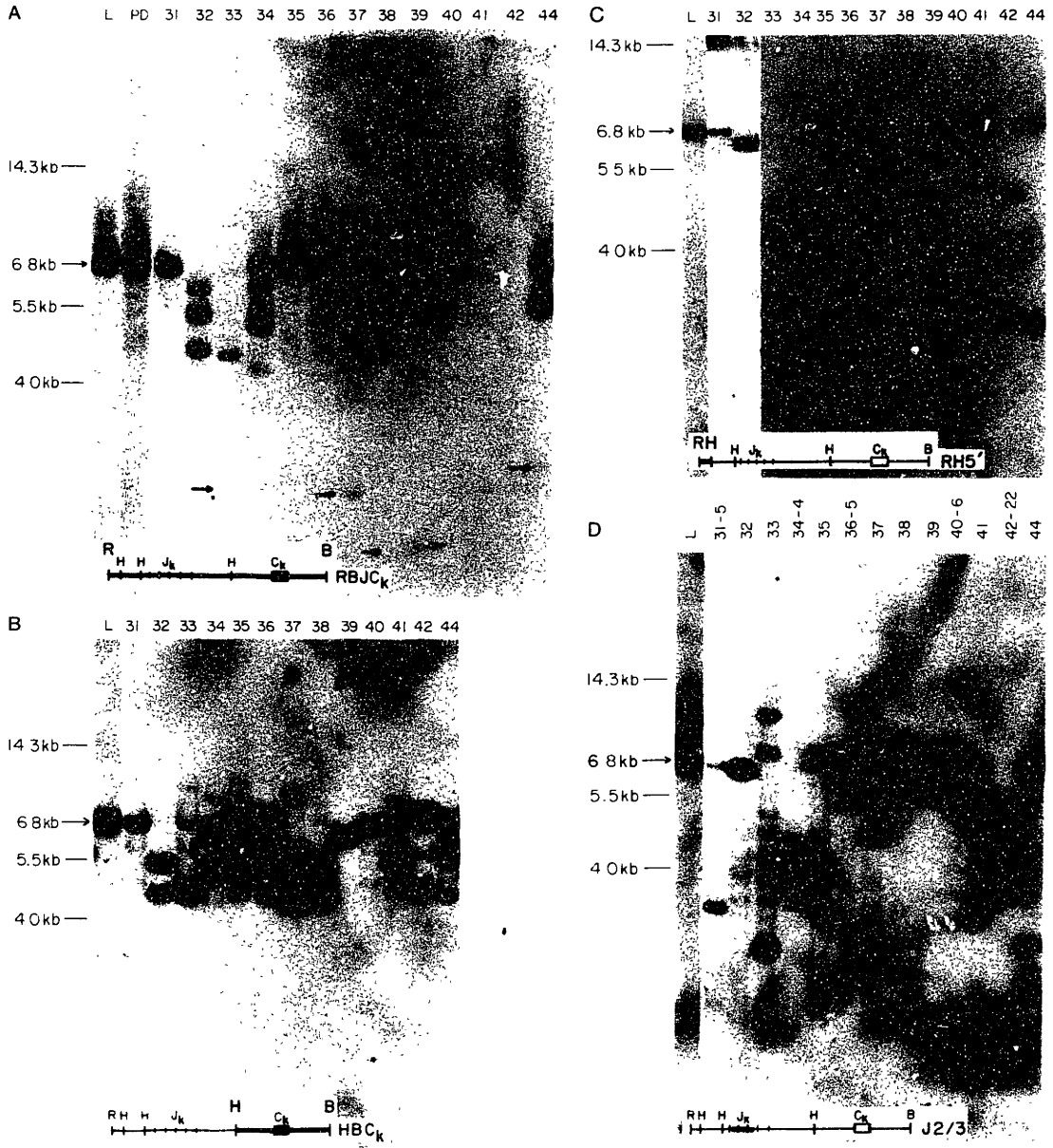



Figure 3. Kappa-Gene Structure in PD Primary Subclones

Approximately 12 μg genomic DNA from each subclone was digested with Eco RI and Bam HI, subjected to electrophoresis through 0.7% agarose gels, transferred to nitrocellulose paper and assayed for hybridization to the indicated probes. The plasmids pRBJC$_\kappa$, pHBC$_\kappa$, pRH5′ and pJ2/3 were labeled by nick translation to a specific activity of approximately 4 × 10$^8$ cpm/μg. All procedures were performed as described by Alt et al. (1980).

(A) Primary subclones analyzed with the pRBJC$_\kappa$ probe. Arrows: several bands that did not reproduce well.

(B) PD primary subclones analyzed with the pHBC$_\kappa$ probe. The autoradiograph was deliberately overexposed to accentuate submolar bands.

(C) Primary subclones analyzed with the pRH5′ probe. Lanes 31 and 32 (containing PD31 and PD32 DNA) are from a separate experiment.

(D) PD subclones analyzed with the pJ2/3 probe. In some cases secondary subclone DNA was substituted for the corresponding primary subclone DNA. This was done to maximize the detection of bands corresponding to those in (C). Because rearranged fragments would be expected to have a variable degree of homology to pJ2/3, nonhomogeneous primary subclone DNA could lead to lack of detection of a reciprocal fragment. Most secondary subclones in this figure bore the same reciprocal fragments detected in the primary subclone from which they were derived, and thus the data are comparable to those in (C). The exception is PD31-5, which is different from PD31 (see Figure 4).

secondary subclones of these lines, however, they were not undergoing rearrangement as actively as PD31 or PD40. Still other primary subclones (PD32 and PD39) showed no variation upon subcloning and therefore appeared to be stable. As seen in Figure 3B, of 13 primary subclones, only PD32 and PD39 lacked a background of submolar bands detected by pHBC$_\kappa$.

Thus the PD line, at the time when primary subclones were derived, was predominantly composed of cells capable of undergoing continuing κ-gene rearrangement. In only two primary subclones had rearrangement activity slowed or perhaps stopped.

**A Rearranged Kappa Allele Can Undergo Further Recombination**

PD31-3 is a secondary subclone of PD31 that bears two rearranged κ alleles. These alleles were detected as 6.8 kb and 6.0 kb Eco RI–Bam HI fragments when probed with pHBC$_κ$ (Figure 5). The 6.8 kb allele happened to migrate indistinguishably from an embryonic band in this analysis, but when the DNA was cut with Bam HI and probed with pHBC$_κ$, no embryonic-sized bands could be detected in PD31 DNA. Furthermore, a probe specific for the region 5′ to J$_κ$, (Figure 2), failed to hybridize to the 6.8 kb Eco RI–Bam HI fragment detected with pHBC$_κ$ (Figure 5C). Because no embryonic gene copies could be detected in PD31-3, subclones of it were examined for variants. With pHBC$_κ$ as a probe, most of the PD31-3 tertiary subclones ha[illegible]e κ-gene structure evident in the uncloned PD[illegible]3 line (Figure 5C). Out of a total of 11 subclones e[illegible]mined, two lacked a 6.0 kb band and had a novel band in its place (PD31-3-4 and PD31-3-13 in Figure 5C). Thus either these two variants arose through alteration of the 6.0 kb allele itself, or they were formed from a germline allele that was present in PD31-3 at the time of its derivation but that rearranged thereafter.

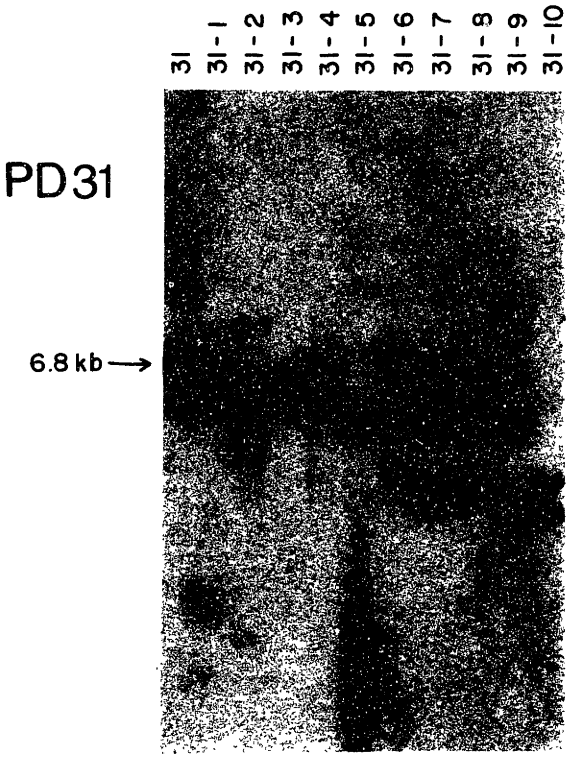



Figure 4. C$_κ$ Rearrangement in PD31 Secondary Subclones

DNA samples from PD31 secondary subclones were digested with Eco RI and Bam HI and were probed with pHBC$_κ$ as described in Figure 3.

To test critically whether a rearranged chromosome could again undergo rearrangement, we cloned PD31-3-1, a subclone with the same apparent gene structure as PD31-3 (two rearranged C$_κ$ alleles on 6.0 and 6.8 kb Eco RI–Bam HI fragments; Figure 5C) again to produce a set of quaternary clones. By this procedure, PD31-3-1 should have been purified of any hypothetical "embryonic" alleles present in PD31-3 (Figure 6A shows, as expected, that in PD31-3-1, the 6.8 kb allele lacks homology to pRH5′). Among the quaternary subclones of PD31-3-1 were again found variants, this time lacking the 6.8 kb band (Figure 5D). The recombination giving rise to these variants took place within or 3′ to the J$_κ$ region, because the 6.8 kb band hybridized to pJ2/3 (Figure 2), whereas the derivative 5.5 kb and 7.6 kb bands did not (data not shown).

The results of analysis of subclones of PD31-3 and PD31-3-1 showed unambiguously that a chromosome that has undergone one rearrangement of its C$_κ$-containing DNA could undergo another rearrangement. This was evident from the serial subcloning and also from comparison of, for example, PD31-3-1, PD31-3-4 and PD31-3-1-6. Both alleles present in one cell line (PD31-3-1) were found in conjunction with a novel rearrangement in one of the other cell lines. Because the alleles appear together in PD31-3-1, random rearrangement of a hypothetical embryonic gene can account for, at most, one of the two other variants (PD31-3-4 or PD31-3-1-6). The other variant must have arisen by rearrangement of a rearranged gene. (Secondary recombination has also been observed among two PD31-3-10 derivatives; data not shown.)

The PD34 secondary subclone series demonstrated that secondary reorganization could also produce deletions of a previously rearranged κ constant region. Both a 6.8 kb and a 4.9 kb rearranged allele occurred in some PD34 variants (PD34-4, PD34-7; Figure 6B), whereas each allele was represented singly in others (PD34-11, PD34-14; Figure 6B). By similar reasoning to that presented above, at least one previously rearranged allele must have undergone a subsequent deletion to generate a line lacking that allele. These deletions may represent either recombination events or loss of chromosome 6.

(By the Eco RI–Bam HI codigestion procedure both PD31-3 and PD34 showed 6.8 kb bands representing rearranged alleles. This is not coincidental—these lines were initially chosen for further study on the mistaken assumption that they contained a 6.8 kb germline gene copy. The 6.8 kb band in PD34 represents a different nongermline allele than that in PD31-3 [it does not hybridize to pJ2/3] and was shown to

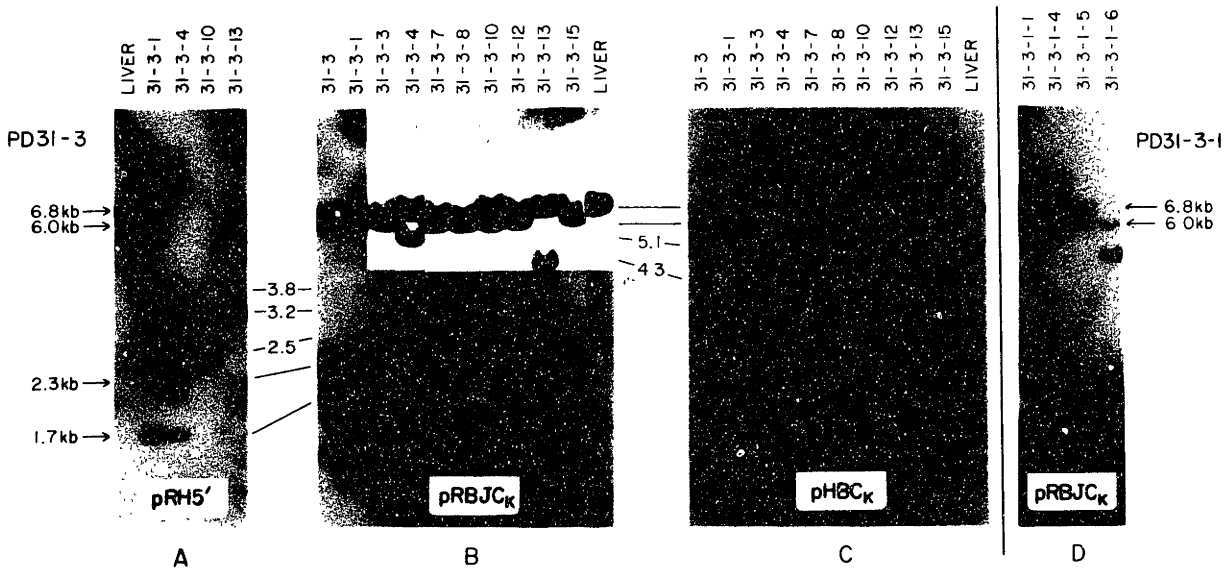



Figure 5. Kappa-Gene Structure of PD31-3 Tertiary Subclones, and of PD31-3-1 Quaternary Subclones

DNA from representative PD31-3 tertiary subclones and representative PD31-3-1 quaternary subclones was digested with Eco RI and Bam HI and probed with the indicated nick-translated plasmids. Arrow (B): a band in PD31-3-4 DNA that is detected with the pRBJC$_\kappa$ probe but not with pRH5' or pHBC$_\kappa$. A region of the print in (C) was masked to retain some of the detail, while allowing for the reproduction of the fainter bands.

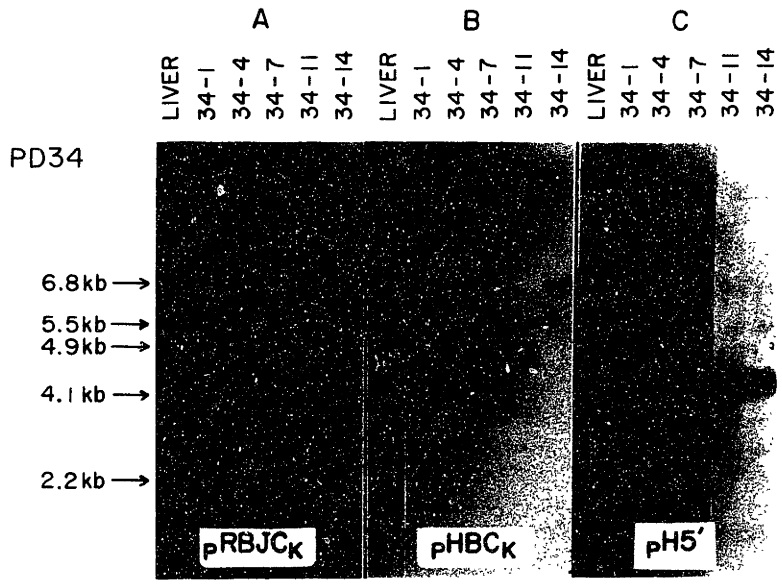



Figure 6. Kappa-Gene Structure of PD34 Secondary Subclones

DNA samples from representative PD34 secondary subclones were digested with Eco RI and Bam HI and probed with the indicated nick-translated plasmids as described in Figure 3.

be rearranged by the procedures outlined above; data not shown.)

**Rearrangement at $J_H$ and $J_\lambda$**

The PD primary subclones were analyzed for rearrangement at $J_H$ by digestion of DNA samples with Eco RI. NIH/Swiss liver DNA contained a single 6.2 kb band when hybridized to probes specific for the $J_H$-containing Eco RI fragment (Alt et al., 1981b). By this assay, the $J_H$-hybridizing bands in various separately derived, in vitro adult bone marrow A-MuLV transformants appeared unique to each line (Alt et al., 1981b), indicating that reorganization at the heavy-chain locus is random. All of the PD subclones had rearranged both chromosomal copies of $J_H$, as had the uncloned PD line (data not shown). With one exception, those subclones had identical $J_H$ gene configurations; the variant, PD31, contained one $J_H$ allele common to its sibling clones and one unique band. The similarity of the $J_H$ structure among the PD subclones contrasted with the enormous variability at the $J_\kappa$ locus, and indicated that the variability that had accumulated at the κ locus (Figure 3) must have been acquired after $J_H$ reorganization. Evidence for a sequential reorganization of first $J_H$ and then $J_\kappa$ was confirmed by analysis of the variant subclone, PD31.

All of the secondary subclones derived from PD31 had identical $J_H$ structures, although, as shown in Figure 4, they had various κ-gene rearrangements. Thus, even though a limited amount of heavy-chain-gene reorganization may have occurred after the PD progenitor cell was transformed by A-MuLV, the $J_H$ alleles were stabilized prior to the onset of κ-gene rearrangement.

Eco RI–digested DNA from the primary subclones was analyzed for λ-gene rearrangements by hybridization to nick-translated $pAB_{\lambda II}$. This probe will detect most rearrangements of $V_{\lambda I}$, $V_{\lambda II}$, $C_{\lambda I}$, $C_{\lambda II}$ or $C_{\lambda III}$ as a result of a variety of cross-hybridizations (Miller et al., 1981). No primary subclone had λ-gene rearrangements, nor have we detected λ-gene rearrangement among secondary subclones of PD31, PD32, PD34 or PD39 (a total of 50 lines were assayed, among which existed 37 different $C_\kappa$ rearrangements; data not shown).

**Reorganization of the Kappa Locus Does Not Always Occur by Deletion**

The restriction enzymes used in our analysis of PD κ genes did not cut within the region of homology to the $pRBJC_\kappa$ probe (Figure 2). The deletion model of $V_\kappa$–$J_\kappa$ joining (Sakano et al., 1979; Max et al., 1979; Seidman et al., 1980) therefore predicts that by our assay, a stable PD derivative should have displayed no more than one new band for each chromosome it had rearranged. Contrary to this expectation, more than two bands were observed in primary subclones PD32 and PD39. Because the "extra" bands persisted upon subcloning (data not shown), the trivial explanation of their contribution by variant subpopulations in each culture was ruled out.

By subdividing probe $pRBJC_\kappa$ into smaller regions, we determined the structure of the supernumerary bands among the PD primary subclones more precisely. The probes used are diagrammed in Figure 2. They subdivided $pRBJC_\kappa$ as follows: $pHBC_\kappa$, as described earlier, contained $C_\kappa$ and represented approximately the 3′ half of $pRBJC_\kappa$; pRH5′ represented the 5′ end of $pRBJC_\kappa$ and lacked all $J_\kappa$ sequence; pJ2/3 included only the sequence around $J_\kappa 2$ and $J_\kappa 3$.

Identical sets of Eco RI–Bam HI–digested primary subclone DNA (with minor variations as noted) were hybridized to $pRBJC_\kappa$ (Figure 3A), $pHBC_\kappa$ (Figure 3B), pRH5′ (Figure 3C) and pJ2/3 (Figure 3D). $pHBC_\kappa$ and pRH5′ (probes representing the extreme 3′ and 5′ regions of $pRBJC_\kappa$, respectively) detected nonoverlapping sets of rearranged bands (with the possible exception of one of the fragments detected in PD37). Probe pJ2/3 hybridized to some of the bands that were recognized by $pHBC_\kappa$ (as would be expected), and all of those that hybridized to pRH5′ as well as to one unique fragment (the 7.1 kb band in PD39; Figure 3A and 3D). Together these three probes hybridized to all of the bands detected by $pRBJC_\kappa$ itself. (Migration rates of low molecular weight fragments are distorted by variation in the absolute amount of DNA analyzed from experiment to experiment; thus their relative positions in Figure 3 vary slightly.)

The gene structure of the lines comprising the PD34 secondary subclone series (shown in Figure 6) was analyzed in more detail by use of various digestion procedures and probes. Again, bands reactive with either a 3′ or a 5′ probe (but not both) were detected. Because each rearranged structure in the PD34 subclones series has a unique distribution among the subclones shown in Figure 6, we could unambiguously map the recombination site or sites for all bands detected by $pRBJC_\kappa$ in this series to within the Hind III–Xba I fragment containing the $J_\kappa$ region (data not shown). Hind III cuts approximately 80 bases 5′ to $J_\kappa 1$, and Xba cuts about 300 bases 3′ to $J_\kappa 5$ (Max et al., 1981).

These data imply that during the rearrangement process, the region homologous to $pRBJC_\kappa$ is often divided by recombinations within (or very near) the $J_\kappa$ region. Such recombinations unlink a portion of the κ locus 5′ to the $J_\kappa$ region from that 3′ to $J_\kappa$. Since the bands that were not detected by $pHBC_\kappa$ had a roughly reciprocal pattern of hybridization to those that were, we will refer to the former as reciprocal fragments, and to the latter as $C_\kappa$ rearrangements. Almost certainly, the reciprocal fragments we have observed are similar to the fragments detected in various myelomas by Steinmetz et al. (1980), Selsing and Storb (1981), Hochtl et al. (1982) and Van Ness et al. (1982). Sequence analysis of several such fragments has led to the suggestion that they represent a byproduct of the $V_\kappa$–$J_\kappa$ joining process (Steinmetz et al., 1980; Hochtl et al., 1982; see Discussion).

Steinmetz et al. (1980), Hochtl et al. (1982) and Selsing and Storb (1981) have emphasized that the individual reciprocal fragments they analyzed do not appear to be reciprocally related to the specific $V_\kappa$–$J_\kappa$ joins of a given line. A dynamic picture leading to the same conclusion may be had by examination of κ genes in the PD34 subclone series. The reciprocal fragments observed in the PD34 series (Figure 6C) could not be correlated uniquely to a given $C_\kappa$ rearrangement (Figure 6B), as might be predicted for the reciprocal products of a single $V_\kappa$–$J_\kappa$ recombination event. Four PD34 subclones had the same reciprocal fragments, but did not have the same pair of $C_\kappa$ rearrangements. Additionally, one line (PD34-7) had a unique reciprocal fragment, while bearing the common $C_\kappa$ structure of the PD34 subclones. That the same reciprocal fragment was present in both PD34-11 and PD34-14 (Figure 6C) implied that deletion of either chromosomal copy of the $C_\kappa$-coding region could occur without a concurrent loss of the reciprocal fragment. This suggests that $C_\kappa$ deletions are recom-

bination events rather than losses of an entire chromosome, although other explanations are possible.

We have shown that $C_\kappa$ rearrangements can undergo further reorganization, resulting in either their deletion or their conversion into new structures. Likewise, reciprocal fragments, once formed, are not inert in the PD line. Figure 5 shows that PD31-3-1, -7, -10 and -15 have identical $C_\kappa$ rearrangements (Figure 5C), while they vary as to the accompanying reciprocal fragments they display (Figures 5A and 5B). Because each of these lines has two identical rearranged $C_\kappa$ alleles, the differences in the pRH5′-hybridizing fragments in this series were acquired *after* the reciprocal fragments were unlinked from the rearranged constant regions. Most of the PD31-3 subclones possessed a 2.3 kb reciprocal fragment; its absence in PD31-3-15 suggests that reciprocal fragments may become deleted as a secondary event (compare PD31-3-1, -3, -8 and -12). A more direct argument comes from the observation that deletion of this 2.3 kb band is evident in subclones of PD31-3-1-1 (data not shown). In this lineage, both a 2.3 kb and a 1.7 kb reciprocal fragment were detected through three subclonings (that is, in DNA extracted from PD31-3, PD31-3-1 and PD31-3-1-1) before the 2.3 kb band was deleted from PD31-3-1-1-3. Thus, occasionally, a preformed reciprocal fragment may be cleared from the genome of a cell in a secondary deletional event. This can occur in the absence of secondary $C_\kappa$ rearrangement (Figure 5C; other data not shown). The pattern of pRH5′-hybridizing bands in PD31-3-1 versus PD31-3-10 (Figure 5A) or in PD34-4 versus PD34-7 (Figure 6C) also suggests the possibility (discussed below) that reciprocal fragments may be converted into new forms as a secondary event.

Whereas most of the reciprocal fragments we detected are hybridizable to pRH5′, an alternative form of reciprocal fragment in PD31-3-4 (arrow in Figure 5B) hybridized to pRBJC$_\kappa$ but not to the subregions represented by pRH5′ or pHBC$_\kappa$ (Figures 6A–6C). Fragments with similar properties have been detected in other PD derivatives (PD39; Figure 3; and PD31-6; data not shown) and in the DNA of some myelomas (Van Ness et al., 1982). Regardless of the specific manner in which this type of reciprocal fragment is generated (one possibility will be diagrammed in Figure 7), a minimum of two recombination events must have occurred in order to unlink a piece of the $J_\kappa$ region from germline sequences on both sides. This may be taken as a further indication of secondary rearrangement of previously rearranged structures (Van Ness et al., 1982).

## Discussion

The A-MuLV-transformed cell line PD is an unusual transformant in that it undergoes extensive DNA rearrangement at its κ locus during culture, resulting in productive κ-gene assembly. Most in vitro–derived A-MuLV transformants from adult bone marrow are stably rearranged at their heavy-chain loci and show no rearrangements at either κ or λ (Alt et al., 1981b). The fact that many PD sublines produce κ protein suggests that many recombinations we observe are the result of $V_\kappa$–$J_\kappa$ joining activity. It is possible that some of the recombinations are of a different nature. It is not understood why PD, and a few other such lines, have a propensity to rearrange their light-chain genes at $C_\kappa$, while most other A-MuLV lines remain embryonic (Alt et al., 1981a, 1981b; S. Lewis, N. Rosenberg, F. Alt and D. Baltimore, manuscript in preparation), but the existence of such exceptional transformants allows certain aspects of immunodifferentiation to be investigated.

### Order of Light-Chain Rearrangements

One striking characteristic of PD and of similar lines actively recombining at $J_\kappa$ is that rearrangement is occuring only at the κ locus, not at $J_H$ or λ. $J_H$ alleles have been stably maintained in a rearranged configuration; the λ genes have remained embryonic. Pooling our data from PD and several other rearranging lines, among 66 subclones we can assay the occurrence of a minimum of 43 different recombination events at the κ locus without detecting a single recombination event at λ. Thus the onset of active rearrangement at the various immunoglobulin loci in the clonal PD line parallels what is known about the process in pre-B cells, as inferred from hybridoma studies (Maki et al., 1980; Perry et al., 1981), from comparisons among A-MuLV transformants derived at various stages of ontogeny (Alt et al., 1981b) and from the properties of some human acute lymphocytic leukemias (Korsmeyer et al., 1981). Comparison of the extent of κ-gene rearrangement with that of λ-gene rearrangement in the PD line is consistent with an ordered model of light-chain-gene rearrangement in which κ is followed by λ, rather than a probabilistic mechanism (Alt et al., 1980, 1981b; Hieter et al., 1981; Colclough et al., 1981).

### Regulation of Kappa-Gene Rearrangement

We have been able to demonstrate unambiguously that a chromosome rearranged at $J_\kappa$ can undergo further alterations; sometimes a new rearrangement is produced, other times the constant region of the κ locus is deleted. The demonstration of multiple rearrangements of a single chromosome for both heavy-chain (Alt and Baltimore, 1982) and light-chain immunoglobulin genes suggests that this capability should be considered in any model of gene rearrangement during normal B-cell differentiation. For light chains, multiple recombination could be a normal process, representing a mechanism whereby an aber-

rantly rearranged allele is corrected by the joining of a second variable region to a $J_\kappa$ segment lying 3′ to the aberrant joint. In this case, regulated termination of gene rearrangement would appear necessary both to maintain allelic exclusion and to prevent inactivation of correctly assembled $\kappa$ genes. One possibility is that the termination of $J_\kappa$ rearrangement depends on protein production (see arguments by Alt et al., 1980, 1981b). Expression of both heavy- and light-chain genes is low and variable among PD subclones for reasons not yet fully understood, and may be the basis for the ability of the cell lines to continue $\kappa$-gene rearrangement in culture.

**Reciprocal Fragments and the Mechanism of $V_\kappa$–$J_\kappa$ Joining**

We assume that the reciprocal fragments we have found are similar to the ones found by Steinmetz et al. (1980), Hochtl et al. (1982), Selsing and Storb (1981) and Van Ness et al. (1982). Those that have been sequenced (Steinmetz et al., 1980; Hochtl et al., 1982) are all joins of the recombination signal sequence 5′ to $J_\kappa 1$ to a recombination signal sequence apparently derived from a $V_\kappa$ region; they are the reciprocal products of a $V_\kappa$–$J_\kappa$ joining reaction. Their presence indicates that joining is a reciprocal recombination event, and that either product of a joining reaction may remain as part of chromosomal DNA.

Our studies show that reciprocal fragments can be detected in the DNA of cells actively undergoing $\kappa$-gene rearrangements. As is the case with some myelomas (Seidman et al., 1980; Steinmetz et al., 1980; Van Ness et al., 1982), however, reciprocal fragments are not always found. Occasionally a line such as PD36 (Figure 3C) appears to be free of such structures. The majority of PD derivatives (we estimate 80%) that have dual $C_\kappa$ rearrangements bear only a single reciprocal fragment.

One possible explanation for the lower number of reciprocal fragments compared with the number of $C_\kappa$ rearrangements comes from the demonstration that reciprocal fragments are not necessarily stable as formed (Figures 5 and 6). It is probably misleading to attach too much significance to the ratio of reciprocal fragments to $C_\kappa$ rearrangements in a terminally reorganized cell line. It is possible that reciprocal fragments are formed directly with every $V_\kappa$–$J_\kappa$ joining event, but that they then frequently undergo secondary deletion. Alternatively, reciprocal fragments may be formed only a fraction of the times that a $V_\kappa$ gene and a $J_\kappa$ segment recombine. A model that accommodates both possibilities is presented below.

The sequenced reciprocal fragments from myelomas (Steinmetz et al., 1980; Hochtl et al., 1982) show one systematic difference from the reciprocal fragments in PD subclones: the myeloma fragments are all joins involving $J_\kappa 1$. PD fragments generally hybridize to pJ2/3 (Figure 3D) and therefore do not occur at $J_\kappa 1$. Furthermore, the sequenced myeloma reciprocal fragments from a given cell are not the expected reciprocal product of the cell's $C_\kappa$ rearrangements (Steinmetz et al., 1980; Hochtl et al., 1981). A possible explanation of this finding is offered by our observation that secondary rearrangements of reciprocal fragments can occur. Because initial rearrangements not at $J_\kappa 1$ leave at least one recognition sequence in the reciprocal fragment (that 5′ to $J_\kappa 1$), reciprocal fragments may serve as substrates for further $V_\kappa$ recombination reactions even though they are no longer linked to $C_\kappa$. Such secondary recombinations could continue until no recognition sequences remain, generating reciprocal fragments joined at the $J_\kappa 1$ recognition sequence.

Several models of $V_\kappa$–$J_\kappa$ recombination have been proposed to explain the origin of reciprocal fragments and will not be reviewed here (Steinmetz et al., 1980; Selsing and Storb, 1981; Van Ness et al., 1982; Hochtl et al., 1982). Of these, the model of unequal sister-chromatid exchange most adequately explains both the frequency with which reciprocal fragments are detected and the isolation of lines that bear reciprocal fragments apparently unrelated to their accompanying assembled $\kappa$ genes (Van Ness et al., 1982). We would like to suggest that another model for $V_\kappa$–$J_\kappa$ joining, originally mentioned by Hozumi and Tonegawa (1976), can, with some modification, account for all the published data as well as our present findings.

If a $V_\kappa$ gene joins to a $J_\kappa$ segment by inversion, rather than deletion, the reciprocal recombination product would be formed and would remain in the genome of a rearranged cell. A simple inversion model was originally rejected because it predicts that sequences 5′ to $J_\kappa$ should always be retained in a rearranged chromosome. Seidman et al. (1980) demonstrated the deletion of $V_\kappa$ and $J_\kappa$ sequences in certain myelomas, presumably occurring as a consequence of the joining of $V_\kappa$ to $J_\kappa$ in those lines. If only some $V_\kappa$ genes are oriented "backwards," relative to the J segments, however, then both inversion and deletion could occur during $\kappa$-gene assembly. Furthermore, because we have shown that secondary rearrangements can occur, these could serve to rearrange further or delete recombined sequences (discussed above). The nonreciprocal structure (Steinmetz et al., 1980) and the ratio of reciprocal fragments to $C_\kappa$ rearrangements (Steinmetz et al., 1980; Van Ness et al., 1982; see above) then could be accounted for fully by the occurrence of secondary recombination events and by the postulate that at some point in the assembly process, $V_\kappa$ genes with the potential to recombine with $J_\kappa$ segments occur in both polarities. This is outlined in Figure 7. To simplify the figure, rearrangements of those segments linked to $C_\kappa$ only are shown; $V_\kappa$ to $J_\kappa$ recombinations involving segments of reciprocal fragments might also occur, however, with similar consequences (deletion or inversion). In addition, germline $V_\kappa$ genes are illustrated in a mixture

of orientations, but it is possible that *all* $V_\kappa$ genes in the germ line are in inverted orientation relative to the $J_\kappa$ segments. In this case, the first $V_\kappa$–$J_\kappa$ recombination at a chromosome must occur via inversion; subsequent $V_\kappa$–$J_\kappa$ joinings could then occur via either inversion or deletion essentially as illustrated. Details of the model are presented in the legend to Figure 7.

We have shown that the joining signals of $D_H$ and $J_H$ may recombine to invert regions of the heavy-chain locus (Alt and Baltimore, 1982). This could not have occurred by recombination between sister chromatids, because a dicentric and an acentric chromosome would result from such an event. Because the similarity of the structures of the joining regions of heavy- and light-chain genes suggests that a similar, if not identical, enzyme system catalyzes their recombination (Early et al., 1980; Sakano et al., 1980), inversion at the heavy-chain locus strengthens the possibility that $V_\kappa$–$J_\kappa$ joining may involve inversions as well. Furthermore, there has been no report to date that a reciprocal fragment represents a duplication of $J_\kappa$ segments, or any other portion of the κ locus; because such duplication is a specific prediction of the model of unequal sister-chromatid exchange (Van Ness et al., 1982), evidence for it should be sought.

By a combination of lineage analysis and sequencing studies we hope to distinguish between the various proposals. The A-MuLV transformant PD and other similar lines should prove valuable in addressing questions about the control and mechanism of light-chain-gene rearrangement during immunodifferentiation that have not been accessible with the use of other systems.

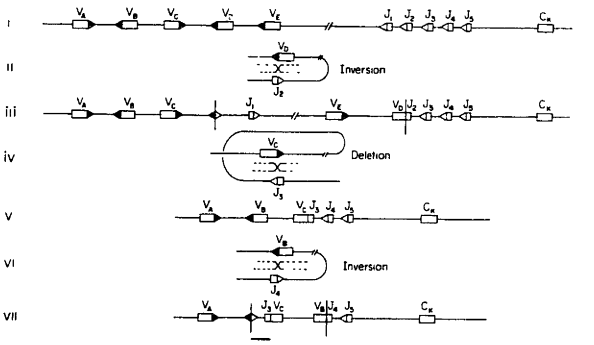



Figure 7. Retention or Deletion of Reciprocal Recombination Products May Be Dictated by the Orientation of the V Genes Relative to the J Segments

The figure is designed to account for the following observations, as discussed in the text: the formation of reciprocal fragments accompanying some $V_\kappa$–$J_\kappa$ recombinations but not others; the deletion of reciprocal fragments as a secondary event; the generation of $C_\kappa$ rearrangements independent of an associated reciprocal fragment; and the formation of reciprocal fragments that are recombinant at 3′ and 5′ sites. (i) Possible germlike organization of the κ-gene segments. It is postulated that some $V_\kappa$ genes ($V_B$, $V_D$, $V_E$) are oriented in the opposite transcriptional direction from the $J_\kappa$ segments, while others ($V_A$, $V_C$) are oriented in the same transcriptional direction as the $J_\kappa$s. (▶—) The recognition sequence associated with the $V_\kappa$ genes in the left-to-right orientation:

$$(\text{V gene})\ \frac{\text{CACAGTG}}{\text{GTGTCAC}}\text{–12 bp–}\frac{\text{ACAAAAACC}}{\text{TGTTTTTGG}}$$

(Max et al., 1979; Sakano et al., 1979, 1980). (—◀) The recognition sequence associated with the $J_\kappa$ segments in the left-to-right- orientation:

$$\frac{\text{GGTTTTTGT}}{\text{CCAAAAACA}}\text{–23 bp–}\frac{\text{CACTGTG}}{\text{GTGACAC}}\ (J_\kappa \text{ segment}).$$

(ii) If a backwards $V_\kappa$ segment (one in the opposite transcriptional orientation from the $J_\kappa$s; in this case $V_D$) is recombined with a $J_\kappa$ segment ($J_2$ here) then an inverted join would occur much as described by Alt and Baltimore (1982). (iii) The resulting product would have all of the DNA between the recombination points in inverted configuration. An appropriate $V_D$–$J_2$ join would be formed with both segments in the same transcriptional direction. The reciprocal fragment (—◀▶—) is also formed. (iv) A forward-oriented $V_\kappa$ segment ($V_C$) joins to a $J_\kappa$ segment ($J_3$). (v) The consequence is a deletion of the previously joined fragment and the production of a new $V_C$–$J_3$ κ gene. (vi) A further join by inversion can occur. (vii) The consequence is an appropriately joined κ gene ($V_B$–$J_4$), and a reciprocal join in which the DNA between the $J_2$ recognition sequence and $J_3$ is unlinked from the DNA segments that flanked it on both its 3′ and 5′ sides in the germline configuration (bar). This type of structure results any time an inversion occurs as a secondary event.

## Experimental Procedures

### Cells

The A-MuLV(P120) transformant PD (also called 300-18) was produced by infection of adult NIH/Swiss bone marrow cells in vitro, as described by Rosenberg and Baltimore (1976). The helper virus was Friend murine leukemia virus. The PD line was isolated as a single colony in soft agar, adapted to liquid culture as described by Rosenberg and Baltimore (1978) and Alt et al. (1981a, 1981b) and first frozen approximately 6 weeks after its derivation. PD was carried in RPMI medium containing 10% heat-inactivated fetal calf serum and 50 μM 2-mercaptoethanol.

All subclones were derived by thawing frozen stocks and cloning from these. The time a given PD derivative was in culture between its isolation as a single cell and its subsequent cloning was generally not more than 3–4 weeks. Cloning was carried out by limiting dilution, seeding 0.2 cells per well.

### Recombinant DNA Plasmids

The plasmid pRBJC$_\kappa$ was derived by subcloning the 6.8 kb $C_\kappa$-bearing Eco RI–Bam HI fragment of a longer genomic clone, MEMJC2 (Seidman and Leder, 1978; a gift from P. Leder), into the corresponding restriction sites of pBR322. The plasmids pRH5′, pH5′ and pHJ$_\kappa$ were subcloned from pRBJC$_\kappa$ by ligating the appropriate Hind III fragment into the Hind III site of pBR322. pHBC$_\kappa$ was derived from pRBJC$_\kappa$ by subcloning the 2.8 kb Hind III–Bam HI fragment into Hind III–Bam HI-digested pBR322. pJ2/3 (a gift from P. Leder) contains $J_\kappa$2 and $J_\kappa$3 sequences on a 860 base Hinf II fragment cloned into pBR322. (We refer to this plasmid as pJ2/3 rather than pJ4/3 [Seidmen et al., 1980] in keeping with accepted terminology; Max et al., 1981). pAB$_{\lambda II}$ contains the full-length (V and C) $\lambda_{II}$ cDNA sequence of MOPC315, as described by Bothwell et al. (1981). pRI-J$_H$ is the $J_H$-containing genomic Eco RI fragment in pBR322, as described by Alt et al. (1981b).

### Analysis of Heavy- and Light-Chain-Gene Rearrangement in Cellular DNA

DNA preparation, restriction enzyme digestions, agarose gel electrophoresis, DNA blotting procedures, probe preparations and hybridization procedures were performed as described by Alt et al. (1980). Specific details or modifications are included in the figure legends.

### Metabolic Labeling and Immunoprecipitation

All procedures for labeling of cells with $^{35}$S-methionine, preparation of extracts, immunoprecipitation and analyses by SDS-polyacryl-

amide gel electrophoresis have been described by Siden et al. (1979), Alt et al. (1980) and Rosenberg and Witte (1980).

**Acknowledgments**

This work was supported by grants from the National Cancer Institute. D. B. was supported also by the American Cancer Society and by a contribution from the Whitehead Charitable Foundation. D. B. is an American Cancer Society Research Professor. N. R. has a Research Career Development Award from the National Cancer Institute. S. L. was supported by a National Science Foundation Fellowship.

The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

Received May 20, 1982; revised June 12, 1982

**References**

Alt, F. and Baltimore, D. (1982). Joining of immunoglobulin heavy chain gene segments: implications from a chromosome with evidence of three D-$J_H$ fusions. Proc. Nat. Acad. Sci. USA *79*, 4118–4122.

Alt, F. W., Enea, V., Bothwell, A. L. M. and Baltimore, D. (1980). Activity of multiple light chain genes in murine myeloma cells producing a single, functional light chain. Cell *21*, 1–12.

Alt, F. W., Rosenberg, N. E., Lewis, S., Casanova, R. J. and Baltimore, D. (1981a). Variations in immunoglobulin heavy and light chain gene expression in a cloned Abelson murine leukemia virus-transformed cell line. In B Lymphocytes in the Immune Response, N. Klinman, D. Mosier, C. Scher and E. Vitetta, eds. (New York: Elsevier/North-Holland), pp. 33–41.

Alt, F., Rosenberg, N., Lewis, S., Thomas, E. and Baltimore, D. (1981b). Organization and reorganization of immunoglobulin genes in A-MuLV-transformed cells: rearrangement of heavy but not light chain genes. Cell *27*, 381–390.

Bothwell, A. L. M., Paskind, M., Schwartz, R. C., Sonenshein, G. E., Gefter, M. L. and Baltimore, D. (1981). Dual expression of λ genes in the MOPC-315 plasmacytoma. Nature *290*, 65–67.

Coleclough, C., Perry, R., Karjalainen, K. and Weigert, M. (1981). Aberrant rearrangements contribute significantly to the allelic exclusion of immunoglobulin gene expression. Nature *290*, 372–378.

Early, P., Huang, H., Davis, M., Calame, K. and Hood, L. (1980). An immunoglobulin heavy chain variable region gene is generated from three segments of DNA: $V_H$, D and $J_H$. Cell *19*, 981–992.

Hieter, P. A., Korsmeyer, S. J., Waldmann, T. A. and Leder, P. (1981). Human immunoglobulin κ light chain genes are deleted or rearranged in λ-producing B cells. Nature *290*, 368–372.

Hochtl, J., Muller, C. R., and Zachau, H. (1982). Recombined flanks of the variable and joining segments of immunoglobulin genes. Proc. Nat. Acad. Sci. USA *79*, 1383–1387.

Hozumi, N. and Tonegawa, S. (1976). Evidence for somatic rearrangement of immunoglobulin genes coding for variable and constant regions. Proc. Nat. Acad. Sci. USA *73*, 3628–3632.

Korsmeyer, S. I., Hieter, P. A., Ravetch, J. V., Poplack, D. G., Waldmann, T. A. and Leder, P. (1981). Developmental hierarchy of immunoglobulin gene rearrangements in human leukemic pre-B-cells. Proc. Nat. Acad. Sci. USA *78*. 7096–7100.

Lenhard-Schuller, R., Hohn, B., Brack, C., Hirama, M. and Tonegawa, S. (1978) DNA clones containing mouse immunoglobulin κ chain genes isolated by *in vitro* packaging into phage λ coats. Proc. Nat. Acad. Sci. USA. *75*, 4709–4713.

Maki, R., Kearney, J. K., Paige, C. and Tonegawa, S. (1980). Immunoglobulin gene rearrangement in immature B cells. Science *209*, 1366–1369.

Max, E., Seidman, J. and Leder, P. (1979). Sequences of five potential recombination sites encoded close to an immunoglobulin κ constant region gene. Proc. Nat. Acad. Sci. USA *76*, 3450–3454.

Max, E., Maizel, J. and Leder, P. (1981). The nucleotide sequence of a 5.5 kilo-base DNA segment containing the mouse κ immunoglobulin $J_\kappa$ and C region genes. J. Biol. Chem. *236*, 5116–5120.

Miller, J., Bothwell, A. and Storb, U. (1981). Physical linkage of the constant region genes for immunoglobulins $\lambda_I$ and $\lambda_{III}$. Proc. Nat. Acad. Sci. USA *78*, 3829–3833.

Perry, R. P., Kelley, D. E., Coleclough, C. and Kearney, J. F. (1981). Organization and expression of immunoglobulin genes in fetal liver hybridomas. Proc. Nat. Acad. Sci. USA *78*, 247–251.

Riley, S. C., Brock, E. J. and Keuhl, W. M. (1981). Induction of light chain expression in a pre-B cell line by fusion to myeloma cells. Nature *298*, 804–806.

Rosenberg, N. and Baltimore, D. (1976). A quantitative assay for transformation of bone marrow cells by Abelson murine leukemia virus. J. Exp. Med. *143*, 1453–1463.

Rosenberg, N. and Baltimore, D. (1978). The effect of helper virus on Abelson virus-induced transformation of lymphoid cells. J. Exp. Med. *147*, 1126–1141.

Rosenberg, N. and Witte, O. (1980). Abelson murine leukemia virus mutants with alterations in the virus-specific P120 molecule. J. Virol. *33*, 340–348.

Sakano, H., Huppi, K., Heinrich, G. and Tonegawa, S. (1979). Sequences at the somatic recombination sites of immunoglobulin light chain genes. Nature *280*, 388–294.

Sakano, H., Maki, R., Furasaw, Y., Roeder, W. and Tonegawa, S. (1980). Two types of somatic recombination are necessary for the generation of complete immunoglobulin heavy chain genes. Nature *286*, 676–683.

Seidman, J. G. and Leder, P. (1978). The arrangement and rearrangement of antibody genes. Nature *276*, 790–795.

Seidman, J. G., Max, E. E. and Leder, P. (1979). A κ-immunoglobulin gene is formed by site-specific recombination without further somatic mutation. Nature *280*, 370–375.

Seidman, J. G., Nau, M. M., Norman, B. Kwan, S.-P., Scharff, M. and Leder, P. (1980). Immunoglobulin $V/J_\kappa$ recombination is accompanied by deletion of joining site and variable region segments. Proc. Nat. Acad. Sci. USA *77*, 6022–6026.

Selsing, E. and Storb, U. (1981). Mapping of immunoglobulin variable region genes: relationship to the "deletion" model of immunoglobulin gene rearrangement. Nucl. Acids Res. *9*, 5725–5735.

Siden, E. J., Baltimore, D., Clark, D. and Rosenberg, N. E. (1979). Immunoglobulin synthesis by lymphoid cells transformed in vitro by Abelson murine leukemia virus. Cell *16*, 389–396.

Steinmetz, M., Altenburger, W. and Zachau, H. G. (1980). A rearranged DNA sequence possibly related to the translocation of immunoglobulin gene segments. Nucl. Acids Res. *8*, 1709–1720.

Swan, D., D'Eustachio, P., Leinwand, L., Seidman, J., Keithley, D. and Ruddle, F. H. (1979). Chromosomal assignment of mouse κ light chain genes. Proc. Nat. Acad. Sci. USA *76*, 2735–2739.

Van Ness, B., Coleclough, C., Perry R. and Weigert, M. (1982). DNA between variable and joining gene segments of immunoglobulin κ light chain is frequently retained in cells that rearrange the κ locus. Proc. Nat. Acad. Sci. USA *79*, 262–266.

Reprinted from Nature, Vol. 308, No. 5958, pp 425-428, 29 March 1984
Macmillan Journals Ltd., 1984



# Joining of $V_\kappa$ to $J_\kappa$ gene segments in a retroviral vector introduced into lymphoid cells

**Susanna Lewis, Ann Gifford & David Baltimore**

Whitehead Institute for Biomedical Research; Department of Biology and Center for Cancer Research, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA

*$V_\kappa$ to $J_\kappa$ recombination occurred within a DNA construct introduced into cells in the form of a defective retrovirus. Analysis of one recombinant demonstrated that $V_\kappa$–$J_\kappa$ joining can occur via inversion with a net loss of a few base pairs.*

IMMUNOGLOBULIN genes are somatically assembled from widely separated DNA segments (see ref. 1 for a review). For κ light chains, recombination brings together two coding units to form the complete variable region portion of the gene: one of perhaps several hundred inherited $V_\kappa$ gene segments joins to one of five $J_\kappa$ segments in a reaction that occurs during B-cell development. The signals that allow recombination enzymes in the cell to recognize immunoglobulin gene segments and to fuse them are not defined, but in all cases rearranged DNA is found previously linked to a heptamer–spacer–nonamer of characteristic structure, suggesting a critical role for this DNA element in the joining reaction. A system which allows introduction into cells of defined DNA constructions that will then undergo the joining process would greatly facilitate the study of the sequence requirements and, more generally, the organizational and mechanistic features of this site-specific DNA recombination process. We report here the development of such a system.

In outline, we approached the problem by infecting a cell line that undergoes immunoglobulin gene rearrangement with a genetically engineered retrovirus containing a germ-line $V_\kappa$ segment (linked to its heptamer/nonamer joining signal) and the five germ-line $J_\kappa$ segments. We used a selectable marker to recover cells that had putatively rearranged the introduced substrate; we then cloned and sequenced one resident provirus to prove that $V_\kappa$–$J_\kappa$ joining had occurred.

## Substrate design

The selection procedure was suggested by a model we presented previously. Studies of κ gene rearrangement have shown that in addition to the $V_\kappa J_\kappa$ coding joint that is formed, it is often possible to detect additional DNA fragments with hybridization properties and sequence characteristics suggesting that they originate as reciprocal recombination products of $V_\kappa$–$J_\kappa$ joining[2–6]. To explain the presence of such reciprocal joints, we proposed that some (and possibly all) $V_\kappa$ elements are organized in germ-line DNA such that their transcriptional orientation is opposite to that of the $J_\kappa$ elements and that inversions joining $V_\kappa$ to $J_\kappa$ would concomitantly generate the reciprocal joint[4]. On the basis of this idea, in our retrovirus construction we placed $V_\kappa$ and $J_\kappa$ in transcriptionally opposite orientations with a selectable marker between them. The marker was positioned in noncoding orientation relative to the promoter in the 5′ long terminal repeat (LTR) and therefore was not expected to be expressed unless activated for expression by an inversional $V_\kappa$–$J_\kappa$ joining event. The selection was designed to depend on an inversional rather than a deletional recombination for three reasons: to minimize background from rearrangements unrelated to $V_\kappa$–$J_\kappa$ joining, to allow recovery of the coding joint as well as the noncoding reciprocal joint, and to test directly whether it is possible for $V_\kappa$–$J_\kappa$ joining to occur via inversion.

A representation of the relevant region of the pMSVgpt-based plasmid[7] constructed for these experiments is shown in Fig. 1 (top). The sequences designed to serve as the joining substrate were flanked by retroviral LTRs derived originally from Moloney sarcoma virus (MSV). Between the LTRs were located (in the 5′ to 3′ orientation of an RNA transcript): the 5′ splice site of MSV which is used by the wild-type virus to allow envelope gene expression; the packaging sequence of MSV[8]; the $J_\kappa$ segments (in 3′ to 5′ transcriptional orientation); the *Escherichia coli* xanthine–guanine phosphoribosyl transferase (*gpt*) gene (in 3′ to 5′ orientation) and an inverted envelope gene 3′ splice site. This was followed by a 5′ to 3′ oriented genomic $V_\kappa 21$–$C$ segment derived from BALB/c DNA (provided by Dr S. Tonegawa). This particular $V_\kappa$ segment has been identified as having rearranged into the active allele in the MOPC 321 myeloma[9,10]. The DNA following the $V_\kappa 21$–$C$ segment in the construct is irrelevant except for the 3′ LTR and its associated primer binding site.

## Retrovirus infection

The plan of the experiment is detailed in Fig. 2. The DNA construct was transfected into ψ-2, a fibroblast cell line that constitutively produces all of the murine leukaemia virus (MuLV) proteins but is unable to package its resident helper virus genome[8]. A co-transfected plasmid carrying a neomycin resistance gene allowed the isolation of stably transfected ψ-2 cells. The virions produced by the transfected cells were used to infect PD cells, a lymphoid cell line that undergoes continuous rearrangement of its κ gene segments[4].

There were several advantages to using a retrovirus vector as a means of introducing the $J_\kappa$ and $V_\kappa$ segments into lymphoid cells. First, introduction of the sequences to be joined as a defective retrovirus simplified further analysis because retroviruses are generally integrated at the edges of their LTRs. This feature ensures that the substrate sequences will not be fortuitously interrupted or rearranged by their insertion into chromosomal DNA. Second, retrovirus infection is a more efficient and controlled process than DNA transfection. Third, the efficiency of introducing substrate could be maximized by using the ψ-2 packaging cell line because the virus stocks it produces are free of helper virus[8].

The PD cell line which we chose to infect with the substrate virus is an Abelson murine leukaemia virus-transformed lymphoid cell line that can carry out $V_\kappa$–$J_\kappa$ joining in culture[4] and is superinfectible by Moloney MuLV. One to two days after the infection of $10^5$ PD cells with substrate virus, the cells were placed in microtitre wells at $2–4 \times 10^3$ cells per well and were grown thereafter in the presence of 0.5 μg ml$^{-1}$ mycophenolic acid to select clones expressing the *E. coli gpt* gene[7,11]. At 3–4 weeks following plating, from one to five mycophenolic acid-resistant clones appeared per experiment and a number were isolated for further study. All but one of the isolates expressed the introduced bacterial enzyme by direct assay (ref. 11; data not shown) and the negative one was discarded.

## Inversional joining

Mycophenolic acid-resistant PD clones in which the *gpt* gene was being expressed could have arisen either by inversional

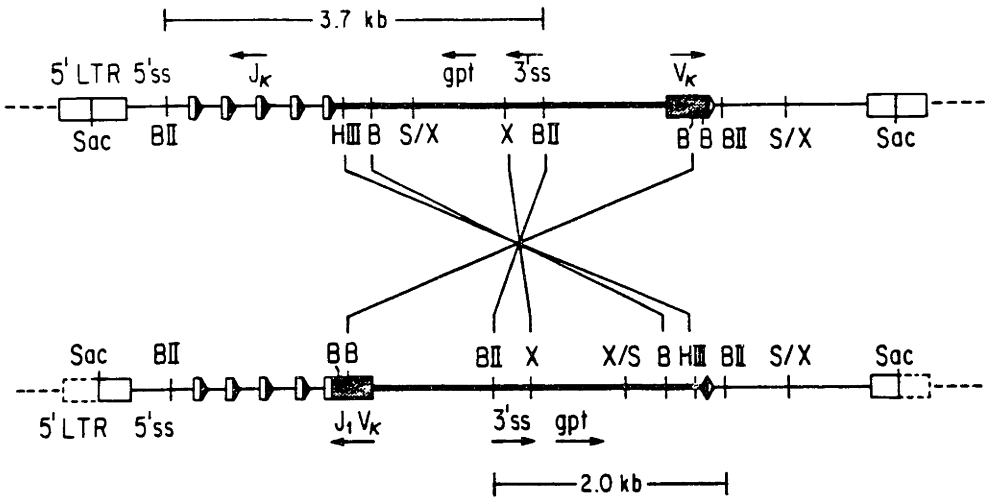



**Fig. 1** Maps of the substrate (top) and one predicted product of $V_\kappa$–$J_\kappa$ joining (bottom). Top, partial map of the substrate, pVJG. The sources of the component fragments were as follows: the *Bgl*II–*Xho*I fragment bearing a 5′ LTR, 5′ splice site, pBR322 sequences, polyoma origin of replication and 3′ LTR (of which only the ends of the fragment bearing the 5′ LTR, 5′ splice site and 3′ LTR are represented in the figure) was derived from pMSVgpt[7]; the *Bgl*II–*Sal*I fragment including the $J_\kappa$ cluster was derived from pHJ$_\kappa$ (ref. 4); the *Xho*I–*Bgl*II fragment containing the *gpt* gene and the 3′ splice site was derived from a construct denoted pTRP-3′ss-MSVgpt (a gift of S. Hartman and R. Mulligan, unpublished) and the *Bgl*II–*Sal*I fragment containing the $V_\kappa$ sequences was derived from Ig$_\kappa$5E (a gift from S. Tonegawa). The *Eco*RI–*Hin*dIII $V_\kappa$-containing fragment of Ig$_\kappa$5E was first subcloned into pBR322 (designated pV$_\kappa$14 and constructed by S. Desiderio) and then the unique *Eco*RI site was changed to a *Sal*I site. The rest of the construction was built in four stages, the details of which can be provided on request. Represented restriction endonuclease sites are abbreviated: Sac, *Sac*I; X, *Xho*I; BII, *Bgl*II; B, *Bam*HI. All the cleavage sites for these enzymes occurring between the *Sac*I sites in the LTRs of pVJG are indicated. Bottom, predicted structure of the region between the LTRs in pVJG if $V_\kappa$ joined to $J_\kappa 1$ by inversion. This map also shows restriction sites found in a λ phage clone designated 2-1 and derived from cell line PD-A as described in the text. Clone 2-1 was isolated from a library of $5\times10^5$ recombinant λ phage bearing size-selected *Sac*I-cut DNA fragments from PD-A cell DNA.

$V_\kappa$–$J_\kappa$ joining, so that the gene was expressed from the 5′-LTR promoter, or by the *gpt* gene acquiring an adventitious promoter from cellular sequences at the insertion site. (The retroviral construct was designed so that the *gpt* gene could be expressed via a spliced mRNA but we have not determined whether the expected spliced mRNA is produced.) The structure that would arise after inversional joining of $V_\kappa$ to $J_\kappa 1$ is shown in Fig. 1 (bottom); it is evident that a new series of fragments should be detected by restriction enzyme cleavage at sites around the joints of an inverted segment. As one test of site-specific joining, DNA was prepared from a number of mycophenolic acid-resistant PD cell isolates, cut with *Bgl*II, and probed by hybridization to *gpt* gene sequences. In the original construct, the *gpt* gene occurred on a 3.7-kb *Bgl*II fragment (as indicated in Fig. 1, top). This fragment was detected in the virus-producing ψ-2 derivative (Fig. 3, lane ψ-2VJG), indicating that no major rearrangement of the substrate had occurred in the virus-producing cell line; the parental ψ-2 line had only the high molecular weight *gpt*-containing fragment introduced during its derivation (Fig. 3, lane ψ-2; see ref. 8. Rearrangement of the substrate in recipient PD cells should produce novel *Bgl*II fragments. The DNA from all nine mycophenolic acid-resistant PD clones assayed (representing at least seven independent isolates) showed an altered *Bgl*II fragment of 2 kb (Fig. 3). This *gpt*-containing fragment had the size that would be expected if $V_\kappa 21$–*C* joined to $J_\kappa 1$ by inversion (Fig. 1, bottom). Thus, it appeared likely that the *gpt* gene was being activated by the predicted inversional join and not by acquisition of adventitious promoters in the various mycophenolic acid-resistant PD cell lines.

## Joint regions

An inversion forms two new DNA joints: in the present case it was expected that one would be a $V_\kappa J_\kappa$ coding joint and the other would be its reciprocal joint. In all lines assayed, the *Bgl*II digest implied that the reciprocal joint was derived from a recombination to $J_\kappa 1$, and not to any of the other $J_\kappa$ genes. It seemed odd, however, that other DNA digests should indicate that the substrate in the nine PD isolates had formed a variety of coding joints, presumably via recombination to *J* genes other than $J_\kappa 1$ (the data will be presented in detail elsewhere). This apparent lack of reciprocity between coding joints and reciprocal joints, and further, this apparent over-representation of $J_\kappa 1$-derived reciprocal joints, is reminiscent of the situation in myelomas[2,6]. We believe that secondary recombination, at high

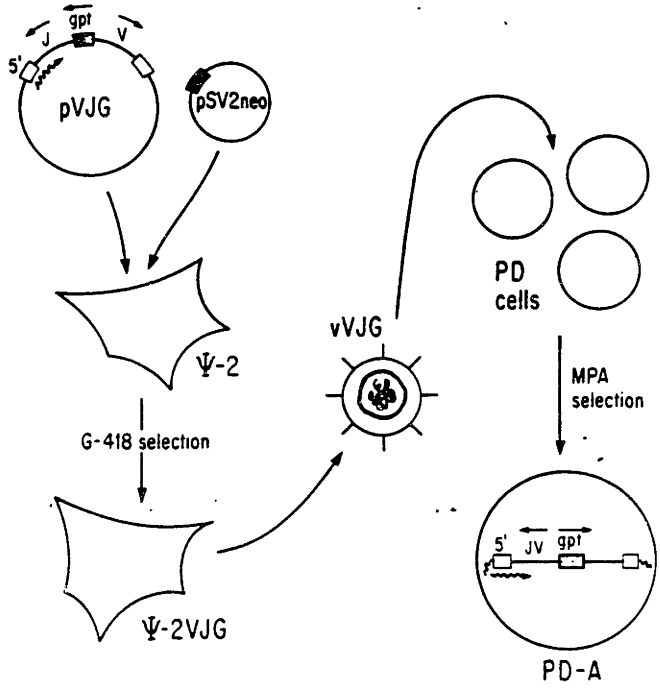



**Fig. 2** Outline of the procedure used to introduce substrate sequences into the target cell line (PD): 10 μg of pVJG and 1 μg pSV2neo (ref. 19) were transfected into ψ-2 cells as described in ref. 8. Four *neo*-positive transfected ψ-2 cell lines were selected in medium containing 1 mg ml$^{-1}$ of the drug G-418 (Gibco)[20], then screened for the production of recombinant virus (vVJG). A virus-producing line was found among these four lines by demonstrating that medium from those cells could convert PD cells to mycophenolic acid resistance. For this assay, 1 ml of medium from the transfectants was applied to PD cells in the presence of 8 μg ml$^{-1}$ polybrene for 1–2 h. One to two days post-infection, PD cells were plated directly in selective medium (RPMI 1640, 10% heat-inactivated fetal calf serum, 50 μM 2-mercaptoethanol, 0.5 μg ml$^{-1}$ mycophenolic acid (Lilly), 250 μg ml$^{-1}$ xanthine, 15 μg ml$^{-1}$ hypoxanthine and 150 μg ml$^{-1}$ glutamine[21]) as described in the text. *gpt*-Positive clones appeared in 3–4 weeks and were maintained in selective conditions.

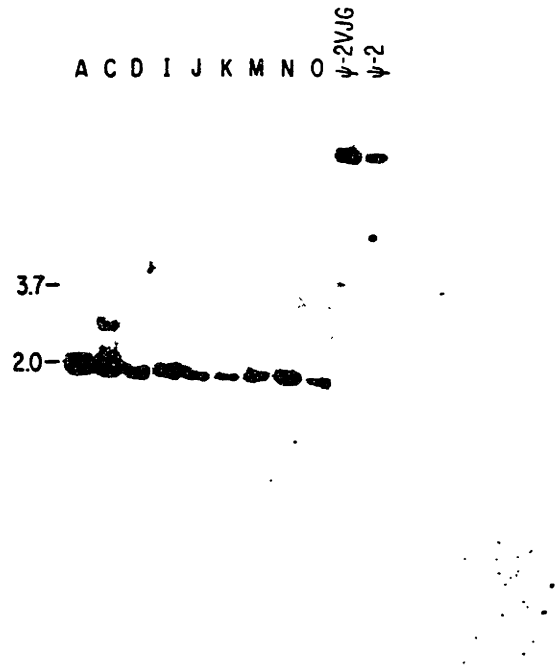



**Fig. 3** Analysis of *Bgl*II-digested cellular DNA samples from mycophenolic acid-resistant PD cell lines. PD isolates A, C, D, I, J, K, M, N and O are shown, as is the fibroblast virus producer, $\psi$-2VJG, and its parental line, $\psi$-2. The filter was probed with *gpt* gene sequences, gel-purified from a *Hin*dIII, *Bam*HI digest of plasmid pL10 (a gift of R. Mulligan) which is a subclone of pPT-1[11].

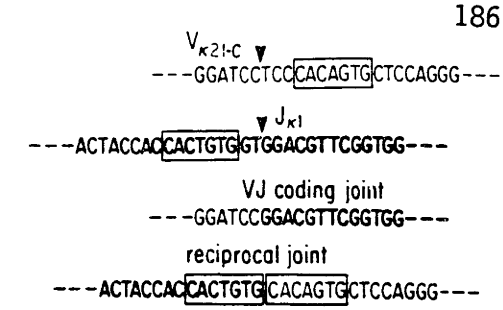



**Fig. 4** Sequence of parental and recombinant fragments. $V_{\kappa}$21–*C*: $^{32}$P end-labelled, *Bam*HI-digested DNA was prepared from pV$_{\kappa}$14 (see Fig. 1 legend) and recut with *Bgl*II. The fragment extending from the 3′-most *Bam*HI site in the coding portion of $V_{\kappa}$21–*C* to the 3′ proximal *Bgl*II site (see Fig. 1, upper) was electrophoretically purified for sequencing[12]. The sequence could be read for 65 bases past the *Bam*HI site (not shown). $J_{\kappa}$1: The sequence shown is from a published report[13]. Coding joint: Clone 2-1 DNA was cut with *Bam*HI, 5′-$^{32}$P-labelled with polynucleotide kinase and recut with *Bgl*II. The 3′-most *Bam*HI site in the coding portion of $V_{\kappa}$21–*C* was labelled by this procedure, and the fragment extending from this *Bam*HI site across the coding joint to the *Bgl*II site 3′ to the $J_{\kappa}$ cluster was purified (see Fig. 1, bottom). The sequence was read for 63 bases past the recombinant joint and agrees completely with the published $J_{\kappa}$1 sequence[13]. Reciprocal joint: The *Hin*dIII–*Bgl*II fragment spanning the reciprocal joint of clone 2·1 (Fig. 1, bottom) was sequenced in both directions. The 50 bases read 3′ to the joint were identical to the sequence obtained for the 3′ flank of $V_{\kappa}$21–*C* (not shown). The 70 bases read 5′ to the joint were identical to the published sequence of the 5′ flank of $J_{\kappa}$1 (ref. 13). Arrowheads indicate the positions of the recombination site used in coding joint formation relative to unrearranged $V_{\kappa}$21–*C* and $J_{\kappa}$1 DNA sequences; the inside borders of the boxed sequences in $V_{\kappa}$21–*C* and $J_{\kappa}$1 represent the position of the recombination site used in reciprocal joint formation.

frequency, deletes reciprocal joint regions down to fusions between the signal elements of $V_{\kappa}$ and $J_{\kappa}$1 regardless of their original structure. Therefore, to simplify the present analysis, we selected a cell line, PD-A, for further study, on the basis that its DNA appeared to have a single proviral integrant that had undergone a single inversional join without secondary deletion. It was expected that in this case we might recover substrate sequences with both a coding joint and a reciprocal joint derived from $J_{\kappa}$1, and thus recover a pair of reciprocal recombination products.

## Provirus sequence

To analyse the structure of the recombined substrate in more detail, the proviral sequences in PD-A were cloned into λgt wes·λB by digesting PD-A DNA with *Sac*I, an enzyme that cuts only in the LTRs of the substrate, and screening recombinant phage for the clone of interest with the *gpt* gene probe.

Restriction enzyme analysis of the clone we obtained indicated that it had the structure shown in Fig. 1 (bottom). The only alteration relative to the original construction appeared to be a large internal inversion, the end points of which suggested that a $V_{\kappa}$–$J_{\kappa}$ recombination had occurred.

The DNA sequence at the inversional end points was determined using the method of Maxam and Gilbert[12]. The precursor $V_{\kappa}$21–*C* and $J_{\kappa}$1 sequences are shown in Fig. 4 together with the two recombinant joints. The sites at which the $V_{\kappa}$21–*C* and $J_{\kappa}$1 elements were fused to form the coding joint can be unambiguously assigned and are indicated by arrows in the figure. The cross-over point on $V_{\kappa}$21–*C* shown here is identical to that observed in the myeloma MOPC 321 (where the same *V* gene has fused with $J_{\kappa}$2; ref. 13). Similarly, the cross-over point we observed on *J*κ1 is consistent with the site where $J_{\kappa}$1 recombined in another myeloma, MOPC 41[5]. In the present case, the reading frames of the fused $V_{\kappa}$ and $J_{\kappa}$ elements do not match, but out-of-frame joining of immunoglobulin gene components is well documented (for a review see ref. 14). Thus, the coding joint we observe has features that in all respects are consistent with its formation via the immunoglobulin gene recombination system.

Although the sequence of the coding joint was consistent with one of several possible expected structures, an even more compelling indication that immunoglobulin-specific joining processes caused the inversion in PD-A was provided by the reciprocal joint. The reciprocal joint was an exact apposition of the heptamer recognition signals that were initially associated with $V_{\kappa}$21–*C* and $J_{\kappa}$1 (Fig. 4). The analogous structure, a precise junction between heptamer elements, has been found in those reciprocal joints that have been cloned from lymphoid cells[2,5,6,15]. Thus, these data clearly show that enzymes in the cell can recognize κ gene-derived segments on the substrate and rearrange them site-specifically. By extension, the identical size of the *Bgl*II fragment in the various PD derivatives expressing *gpt* (Fig. 3) implies that similar events activated the gene in all of the cells.

## Mechanism of joining

Because both coding joints and reciprocal joints are found in many B-lymphoid tumour cells and because, as we show here, both of these joints can be generated by an inversion, the results described here support our previous proposal that in pre-B lymphocytes immunoglobulin gene joining may be accomplished through an inversional process[4]. The evidence that reciprocal joints in tumour cells are often not the direct reciprocals of the accompanying coding joints[2,6,15] does not invalidate this suggestion because there is evidence from rearranging cells that the κ locus can undergo multiple recombinations[4]. Thus, reciprocal joints and coding joints, once formed, may be subjected to subsequent recombination which would serve to obscure their original relationship. Other proposals for the origin of reciprocal joints have been made[2,3] but positive evidence for them is lacking. Because the distinction between our proposal and alternative schemes is basically topological, the question will probably remain open until more is known about the relative orientation of $V_{\kappa}$ and $J_{\kappa}$ in the murine germ line.

A direct conclusion obtained from the comparison of reciprocal recombination products and their precursor sequen-

ces (Fig. 4) is that joining can be accompanied by a net loss of a small number of nucleotides. Three nucleotides from the germ-line 3′ flank of $V_\kappa 21$–$C$ and two from the 5′ side of $J_\kappa 1$ were not found in either the coding joint or the reciprocal joint after recombination. Missing nucleotides have also been noted in a case where a non-$V_\kappa$-derived DNA segment recombined with $J_\kappa 1$ (ref. 5) and was suggested to be a general feature of immunoglobulin gene joining processes[6,15]. Nucleotides are lost because the cross-over site for the reciprocal joint is not positioned at the same base pair as the cross-over site for the coding joint (see Fig. 4). As mentioned earlier, although the cross-over site for reciprocal joints is invariant in all available examples (refs 2, 5, 6, 15 and this paper) there is evidence for considerable junctional diversity at coding joints[16–18] which indirectly implies that nucleotides must be lost during a $V_\kappa$–$J_\kappa$ joining event.

A limitation of the present study is that the recombination substrate before rearrangement had no active selectable marker. Thus, it was impossible either to titre the substrate virus stocks directly or to determine the frequency with which recombination occurred in infected cells. It is also difficult to rule out the unlikely possibility that the observed rearrangements took place in the fibroblast $\psi$-2 producer line before the virus was used to infect PD, although mycophenolic acid resistance could not be passed from the $\psi$-2 cells to NIH 3T3 cells. We are presently constructing substrates that should allow quantitation of the frequency of this sort of recombination in a variety of cell types.

## Conclusion

With the ability to examine the sequence of both products of a single joining event, we have been able to present evidence for the first time that $V_\kappa$–$J_\kappa$ joining can occur via inversion, that it is a reciprocal exchange and that the joining enzymes act asymmetrically to unite heptamer elements at a fixed position while apparently eliminating bases from the coding joint. In addition, the effective rearrangement of the substrate implies that the joining reaction is not dependent on long-range chromatin or DNA structure, nor even on the presence of the constant-region portion of the $\kappa$ gene. The ability to insert designed constructions into cells that will then undergo recombination should be very useful in further elucidating the mechanism, sequence requirements, and perhaps eventually the control of immunoglobulin gene joining.

In independent studies, K. Blackwell and F. Alt (personal communication) have observed rearrangement of heavy-chain $D$ and $J$ regions transfected into Abelson murine leukaemia virus-transformed fetal liver cells.

We thank R. Mulligan and members of his laboratory for advice in constructing pVJG and for the generous supply of component DNA; also D. Rice, R. Mann, C. Tabin and N. Andrews for advice, reagents or both. This work was supported by grants from the NCI and the American Cancer Society. D.B. is an American Cancer Society Research Professor. S.L. was supported by a NSF fellowship.

*Note added in proof*: Cloning and DNA sequence analysis of additional independently-derived proviral recombinants confirm that $V_\kappa 21$–$C$ in the substrate can recombine to $J_\kappa$ segments other than $J_\kappa 1$ and that in several examples of recombination to $J_\kappa 1$ the resulting coding joint is variable. In contrast, all reciprocal joints analysed (including that found on a clone with a $V_\kappa$–$J_\kappa 4$ coding joint) have a sequence identical to that in Fig. 4.

Received 12 December 1983; accepted 10 January 1984.

1. Tonegawa, S. *Nature* **302**, 575–581 (1983).
2. Steinmetz, M., Altenburger, W. & Zachau, H. *Nucleic Acids Res.* **8**, 1709–1720 (1980).
3. Van Ness, B., Coleclough, C., Perry, R. & Weigert, M. *Proc. natn. Acad. Sci. U.S.A.* **79**, 262–266 (1982).
4. Lewis, S., Rosenberg, N., Alt, F. & Baltimore, D. *Cell* **30**, 807–816 (1982).
5. Hochtl, J. & Zachau, H. *Nature* **302**, 260–263 (1983).
6. Hochtl, J., Muller, C. R. & Zachau, H. *Proc. natn. Acad. Sci. U.S.A.* **79**, 1383–1387 (1982).
7. Mulligan, R. in *Experimental Manipulation of Gene Expression* (ed. Inouye, M.) 155–173 (Academic, New York, 1983).
8. Mann, R., Mulligan, R. & Baltimore, D. *Cell* **33**, 153–158 (1983).
9. Lenhard-Schuller, Hohn, B., Brack, C., Hirama, M. & Tonegawa, S. *Proc. natn. Acad. Sci. U.S.A.* **75**, 4709–4713 (1978).
10. Heinrich, G., Traueneker, A. & Tonegawa, S. *J. exp. Med.* (in the press).
11. Mulligan, R. C. & Berg, P. *Science* **209**, 1422–1427 (1980).
12. Maxam, A. M. & Gilbert, W. *Proc. natn. Acad. Sci. U.S.A.* **74**, 560–564 (1977).
13. Sakano, H., Huppi, K., Heinrich, G. & Tonegawa, S. *Nature* **280**, 288–294 (1979).
14. Coleclough, C. *Nature* **303**, 23–26 (1983).
15. Alt, F. & Baltimore, D. *Proc. natn. Acad. Sci. U.S.A.* **79**, 4118–4122 (1982).
16. Weigert, M. *et al. Nature* **283**, 497–499 (1980).
17. Max, E., Seidman, J., Miller, H. & Leder, P. *Cell* **21**, 793–799 (1980).
18. Kurosawa, Y. *et al. Nature* **290**, 565–570 (1981).
19. Southern, P. & Berg, P. *J. molec. appl. Genet.* **1**, 327–341 (1982).
20. Colbere-Garapin, F., Horodniceanu, F., Kourilsky, P. & Garapin, A.-C. *J. molec. Biol.* **150**, 1–14 (1981).
21. Rice, D. & Baltimore, D. *Proc. natn. Acad. Sci. U.S.A.* **79**, 7862–7865 (1982).

Printed in Great Britain by Macmillan Production Ltd., Basingstoke, Hampshire