Cell, Vol. 40, 259–269, February 1985, Copyright © 1985 by MIT

0092-8674/85/020259-11 $02.00/0

Structure, Organization, and Somatic Rearrangement of T Cell Gamma Genes

Adrian C. Hayday, Haruo Saito, Stephen D. Gillies, David M. Kranz, Gary Tanigawa, Herman N. Eisen, and Susumu Tonegawa
Center for Cancer Research
and Department of Biology
Massachusetts Institute of Technology
Cambridge, Massachusetts 02139

Summary

We present the initial characterization of a novel family of genes that rearrange in T cells, but do not encode either of the defined (α/β) subunits of the clone-specific heterodimer of the T cell receptor. The family comprises at least three variable (V) gene segments, three constant (C) gene segments, and three junction (J) gene segments. In a cloned cytolytic T lymphocyte, 2C, one of each of these fragments has productively rearranged to yield an expressed VJC transcription unit, which shows no evidence for somatic mutation. Short sequences similar to those implicated in immunoglobulin gene and T cell receptor β chain gene rearrangement flank the V and J segments of this family. The linkage of two of the three V gene segments has been determined: the segments lie approximately 2.5 kb apart, and are arranged head-to-head. The inverted arrangement may cast light upon the mechanisms utilized by lymphocytes for gene rearrangement.

Introduction

The immune system of vertebrates has the capacity to cope with challenge by a potentially immense repertoire of foreign products. In coping with this challenge, the system in effect mounts a two-pronged attack: the recognition of circulating antigens by B cell immunoglobulins, and the recognition of cell-bound antigens by specific receptors on T cells. The B cell immunoglobulins and the antigen receptors on T cells both display clonal diversity. Extensive studies of immunoglobulin genes (reviewed by Tonegawa, 1983) have established that somatic rearrangement and mutation of limited germ-line genetic information plays a major role in generating the clonal diversity of the B cell response. The extent to which such a mechanism might be used to generate the diversity of the T cell response remained a central issue in immunology.

The T cell receptor might be considered to comprise, in its minimal form, a heterodimeric protein of α and β subunits, each of approximately 45 kd (Allison et al., 1982; Meuer et al., 1983a; Haskins et al., 1983). Each of the α and β subunits seems to contain constant and variable regions (Acuto et al., 1983; McIntyre and Allison, 1983; Kappler et al., 1983). Although T cells recognize antigen in the context of the products of the major histocompatibility complex (MHC), many arguments (e.g., Kappler et al., 1981; Hünig and Bevan, 1982) have been forwarded to consider the T cell receptor as a single receptor recognizing some combination of antigen and MHC product, rather than two receptors recognizing antigen and MHC product individually (reviewed by Zinkernagel and Doherty, 1979; Matzinger, 1981). According to this viewpoint, the $\alpha\beta$ heterodimer might contain the complete antigen specificity of a T cell and, indeed, evidence exists to suggest that the $\alpha\beta$ heterodimer can recognize aspects of both antigen and MHC product (Marrack et al., 1983).

Genes for the β chain (Acuto et al., 1984) of the heterodimer were recently isolated (Hedrick et al., 1984; Yanagi et al., 1984). Analysis of these genes indicated that, as for immunoglobulins, diversity in the β chain is generated, at least in part, by the somatic rearrangement of germ-line gene segments (Chien et al., 1984a; Gascoigne et al., 1984; Siu et al., 1984a; Malissen et al., 1984). Since then, we and others (Saito et al., 1984b; Chien et al., 1984b) have isolated genes that most likely encode the α chain of the heterodimeric T cell receptor (Hannum et al., 1984). These genes also apparently undergo T-cell-specific rearrangement, the possible permutations and combinations of which are again likely to generate diversity in the heterodimer.

Earlier than this, however, we cloned another distinct cDNA (clone 4/203) from the same T cell clone, which we initially thought coded for the α chain (Saito et al., 1984a). Since it is now clear that this gene does not encode a component of the heterodimeric T cell receptor (Hannum et al., 1984; Saito et al., 1984b), we call it the γ gene. Like the clones for the α and β chains, the γ gene possesses conspicuous properties. It has sequence and organizational similarities to both Ig genes and genes for the β chain of the T cell receptor. It was expressed only in T cells, and it, or a gene highly homologous to it, was expressed in all cytolytic T cells that we examined. Furthermore, the expressed gene was generated by an apparently T-cell-specific rearrangement of multiple gene segments. Hence the function of the putative γ gene product is a most interesting problem to be resolved. One approach to this is to characterize the γ gene family. This approach can indicate how similar in organization, structure, and regulation γ genes are to genes encoding immunoglobulins and the heterodimeric T cell receptors. It can indicate the potential diversity that may be generated by the rearrangement of γ genes, and thereby imply the capabilities of γ gene products. This approach has been followed here.

Results

The Isolation of Germ-Line Clones

The cDNA, pHDS4/203, cloned from an alloreactive (anti L^d) CTL clone, 2C (Kranz et al., 1984), can, by analogy with immunoglobulin (Ig) genes and T cell receptor α and β genes, be split into a variable (V) segment a constant (C) segment and, between the two, a junction (J) segment

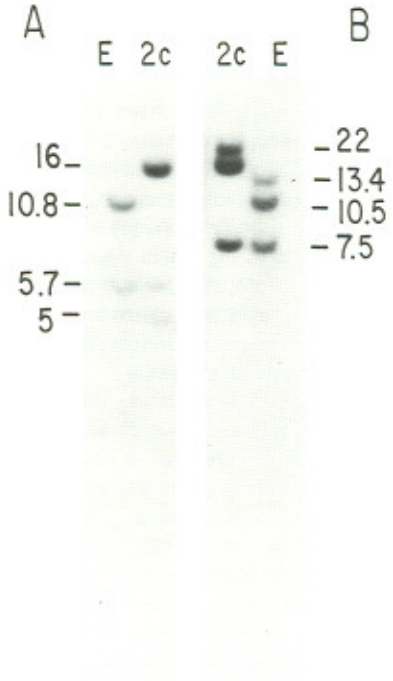

Figure 1. Eco RI-Digested Balb/c Embryo DNA (E) and 2C Cell DNA (2C) Electrophoresed in 0.8% Agarose, Transferred to Nitrocellulose, and Hybridized to a V Probe (pcV) (A), or a C Probe (pcJC) (B)

Sizes of the fragments (in kb) are indicated by the numbers shown, and were deduced from co-electrophoresis with Hind III-cut λ DNA.

(Saito et al., 1984a). The V segments and the C segments were used individually to probe Eco RI digests of Balb/c embryo and 2C DNA (Figure 1). In Balb/c embryo DNA, the V probe detects two Eco RI fragments, of 10.8 and 5.7 kb, respectively. The C probe detects three Eco RI fragments of 13.4, 10.5, and 7.5 kb, respectively. Each of these five fragments was isolated from a λ phage library of Eco RI-cleaved Balb/c embryo DNA, and was analyzed by restriction analysis coupled with hybridization to fragments of the cDNA clone (see Figure 2 for the regions of the cDNA covered by each probe), and by direct DNA sequencing. The clones, according to their sizes in kb, and to the probe to which they hybridize, were denoted: V10.8, V5.7, J-C13.4, J-C10.5, C7.5. The general organization of each clone is depicted in Figure 2.

Three C Regions, Three J Regions, Three V Regions

The clones J-C13.4, J-C10.5, and C7.5 each contain a single constant region that is split into three exons. This was determined initially by hybridization to fragments of the cDNA clone (total cDNA, pcJC, and pc3′U.T.; Figure 2), and was later confirmed, for J-C10.5 and C7.5, by DNA sequence analysis. In addition, use of probe pcJC indicated that clones J-C13.4 and J-C10.5, but not clone C7.5, each contain at least one germ-line J gene segment. Locations of these J segments were determined roughly by Southern blotting analysis using the pcJC probe and more accurately by DNA sequencing (see below). When a germline J region probe, probe gJ (Figure 2), is dissected from J-C10.5 and used to probe Eco RI-digested embryo DNA,

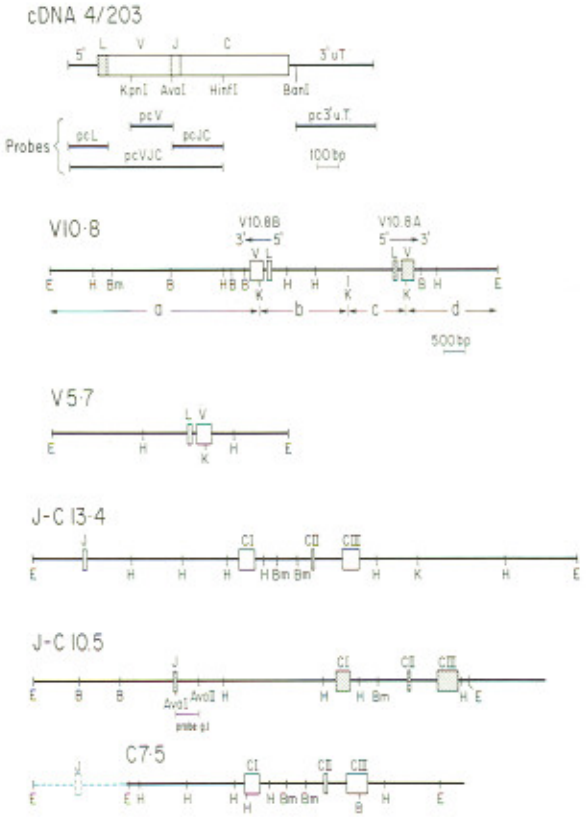

Figure 2. Recombinant DNA Clones Used in This Study

cDNA 4/203, cloned from 2C cells, and segmented into putative leader segment (L), variable segment (V), junction segment (J), constant segment (C), and 5′ and 3′ untranslated regions. Probes derived from 4/203 are indicated. Genomic clones contain coding regions (boxed). Shaded boxes denote coding regions used in the productive rearrangement in 2C cells. E, Eco RI; H, Hind III; Bm, Bam HI; K, Kpn I; B, BgL II. Restriction maps are not necessarily complete.

it detects three Eco RI fragments, of 13.4, 10.5, and approximately 2.3 kb, respectively (Figure 3). The former two fragments correspond to the cloned fragments J-C13.4 and J-C10.5. (Since probe gJ derives from C10.5 it is not surprising that hybridization is strongest to the 10.5 kb fragment). The hybridization of probe gJ to a 2.3 kb fragment indicates the presence of a third, homologous germline J segment. This has been confirmed by probing other restriction digests (Xba I, Bam HI, Hind III) of embryo DNA. The genomic location of the third J segment relative to the other coding sequences of this gene family is still unknown. However, preliminary data indicate its location to be immediately upstream of C7.5 (Figure 2). This location would be consistent with subsequent analysis of rearranged fragments in 2C DNA (see below), and also predicts an interval of between 3.5 and 4.5 kb between J and C gene segments, which is comparable to the distance of approximately 3.8 kb that separates both J13.4 and C13.4, and J10.5 and C10.5. By hybridization intensity, the 2.3 kb J segment is the least homologous of the three.

V5.7 contains one variable gene segment. Conversely, analysis of V10.8 shows that Kpn I fragments a and d each hybridize to variable region probe cV (Figure 2). Since these fragments lie approximately 3.5 kb apart, it suggests

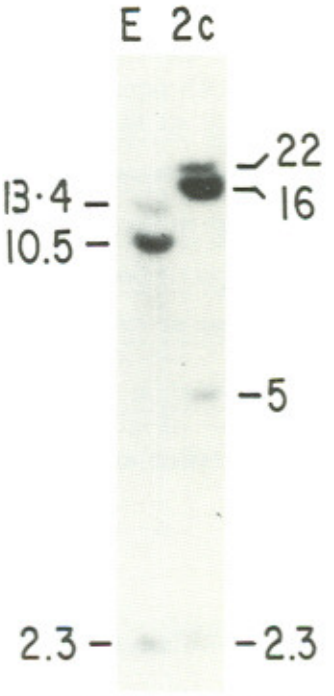

Figure 3. Eco RI-Digested Balb/c Embryo DNA (E) and 2C Cell DNA (2C) Probed with gJ (See Figure 2)

Numbers denote sizes of fragments in kb.

that V10.8 contains two variable gene segments. A probe L (Figure 2), specific for the hydrophobic, putative leader segment of V, hybridizes to Kpn I fragments b and c. A similar result is obtained using other restriction analyses, e.g., Msp I (data not shown). Since neither Kpn I nor Msp I split the leader segment (sequence data: see below), it suggests that there are two leader segments, one on Kpn Ib and one on Kpn Ic. Moreover, these two putative leader segments are flanked by the two Kpn I fragments (a and d), both of which hybridize with the V region probe cV. These results suggest that there exist two V gene segments on V10.8 (confirmed by DNA sequencing; see below) that are inverted relative to one another in a head-to-head organization. In conclusion, the homologous members of this gene family comprise three C, three V gene segments, and three clusters of J gene segments.

DNA Sequence Analysis

V Segments

The two V segments, V10.8A and V10.8B, that have been completely sequenced, are shown in Figure 4. Each contains two exons: a 5′ exon that encodes the large part of a putative hydrophobic leader segment, and a 3′ exon of approximately 300 bp, which can combine with J to encode the rest of the proposed leader and the major body of the variable region. The position of the intron between these two exons agrees with that found for all Ig heavy and light chain genes; that is, it occurs four amino acids before the finish of the putative leader peptide. The two introns differ by the deletion/insertion of 12 bp. A 12 bp deletion in the analogous intron distinguishes one of the five Balb/c germ-line IgV_x21 genes from the other members of that family (Heinrich et al., 1984). For the T cell gene presented here, as for the IgV_x21 gene (Heinrich et al., 1984), the deletion occurs between short direct repeats (GTG), which make it impossible to position it unequivocally (Figure 4).

The two V gene sequences shown here differ over the coding regions, by 16 point mutations and one triplet deletion/insertion. This degree of variation is similar to that seen between the most divergent genes of either the IgV_HT15 family or the IgV_HNP^b family that have been analyzed to date (Crews et al., 1981; Bothwell et al., 1981). However, it is considerably greater variation than that existing between the most closely related members of either of these families or of the V_x21 family (Heinrich et al., 1984). Like the variation witnessed in the studies of those Ig gene families, the changes between the two V regions shown here are widely distributed throughout the whole gene. Furthermore, there are considerably more replacements than silent mutations (14r, 2s), which, with the triplet insertion/deletion, lead to 15 amino acid differences between the potential gene products. None of these amino acid changes occur at very highly conserved positions, suggesting that in themselves none of them are enough to abrogate the function of the potential gene product. The least conservative of the changes in fact occur in the putative leader segment. The overall organizational similarity of the V region of this gene family to Ig V regions has previously been discussed (Saito et al., 1984a).

The J Regions

Short restriction fragments of J-C10.5 and J-C13.4, which hybridize to pcJC but are upstream of C, were analyzed in order to define the J region clusters for this gene family. The restriction fragments were each about 1 kb. Throughout their lengths the fragments from J-C10.5 and J-C13.4 showed high homology as judged by conservation of restriction sites. However, the only occurrence of an Ava I site was a single position in J-C10.5. Over 750 bp were sequenced from the J-C10.5 fragment, and over 250 bp were sequenced from the J-C13.4 fragment. As indicated in Figure 4, sequencing data strongly suggested the occurrence of only one J segment in J-C10.5. Sequencing data and homology to J-C10.5 suggested that J-C13.4 also contained only one J. In this respect the gene family resembles the murine Igλ light chain genes (Blomberg and Tonegawa, 1982) and differs from the T cell receptor β genes (Siu et al., 1984a; Gascoigne et al., 1984; Chien et al., 1984a; Malissen et al., 1984). The two sequenced J regions shown in Figure 4 differ by only 4 nucleotides over the 60 bp coding region; none of these changes affect the coding potential. This coding potential is quite distinct at its 5′ end from J segments of murine Ig genes or T cell receptor β chain genes. Over the 3′-most ten amino acids, the coding potential is more conserved: there are 6 out of 10 matches with the corresponding region of several T cell receptor β chain Js, 5 out of 10 matches with Igκ or Igλ light chain Js, but only 3 out of 10 matches with IgH J segments.

The J cluster contained in the germ-line 2.3 kb fragment has not been sequenced, so it is as yet unknown whether it too contains only one J.

Figure 4. Nucleotide Sequences of γ Gene Variable and Junction Segments

The intron–exon organization was determined by sequence comparison with cDNA 4/203. The potential protein coding regions are indicated in triplet form and, in themselves, show *no* alterations when compared with the sequence of 4/203 (Saito et al., 1984a). Agreement of V108B, V5.7, and J13.4 sequences to the sequences shown is indicated by dashed lines; disagreements are clearly denoted. Splice junction sequences are overlined, so that they may be compared with the consensus sequences: 5′ GTAAGT. . . . . .$(Py)_6$-CAG.3′ suggested by Breathnach and Chambon (1981). Also denoted by overlines are the heptamer and nonamer consensus sequences. An estimate of the start site for transcription (arrowed) is denoted nucleotide 1 and underlined. Underlined upstream of this are putative transcription signals (see text), and their distances from nucleotide 1. The initiator methionine, boxed, is also estimated. An ATG upstream of this, which is probably transcribed, is underlined.

C Regions

Complete sequence was obtained for the C coding regions of C10.5 and C7.5 (Figure 5). Comparison with the cDNA sequence indicates that both coding regions are divided into three exons (Figure 4). The first is an immunoglobulin-like domain of 330 bp (110 amino acids), the same length as Ig CH1 (Kabat et al., 1983). This domain features a 6 out of 8 match of the most highly conserved residues in IgCH1, and a 12 out of 17 match of the invariant residues of the constant regions of Ig light chains. In some of these matches, the domain defined here is more similar to either the Ig light chain, or the Ig heavy chain, than the Ig light and heavy chains are to each other. A similar situation has been described for the product of the murine T cell receptor β chain gene (Gascoigne et al., 1984).

The second exon of C10.5 encodes only ten amino acids. A similarly short exon encodes the hinge region of IgG heavy chains (Sakano et al., 1979a), and also exists in the structural gene for the T cell receptor β chain constant region (Gascoigne et al., 1984). The hinge regions of Ig heavy chains always include (often multiple copies of) cysteine and proline (Kabat et al., 1983). The proline residues are thought to be crucial in conferring flexibility to the hinge region. The short second exon of the constant region presented here encodes a single cysteine, but noticeably no proline. Exactly this situation occurs in the short second exon of the T cell receptor β chain constant region (Gascoigne et al., 1984). The second exon of C7.5 contains a 15 bp insertion relative to C10.5, clearly distinguishing it from the cDNA clones from 2C cells, and is almost certainly defective, for reasons discussed below and in the legend to Figure 5.

The third exon contains about 545 bp. It includes a region that we have proposed to encode the transmembrane domain (Saito et al., 1984a). Like the transmembrane domains of Ig genes, the sequence of this domain is rich in hydrophobic residues. However, it has a conspicuous property, which it shares with the T cell receptor β chain of both mouse and humans, and which distinguishes all three of these from the transmembrane domains encoded by Ig genes. This feature is a potentially charged lysine residue, flanked at a distance on both sides by tyrosines (Figure 6). The incorporation of bulky tyrosine residues into a transmembrane domain, at conserved distances from an internal, polar residue, is provocative of some specific function associated with the transmembrane domain of these proteins other than to anchor the protein. Strikingly, there is strong conservation of

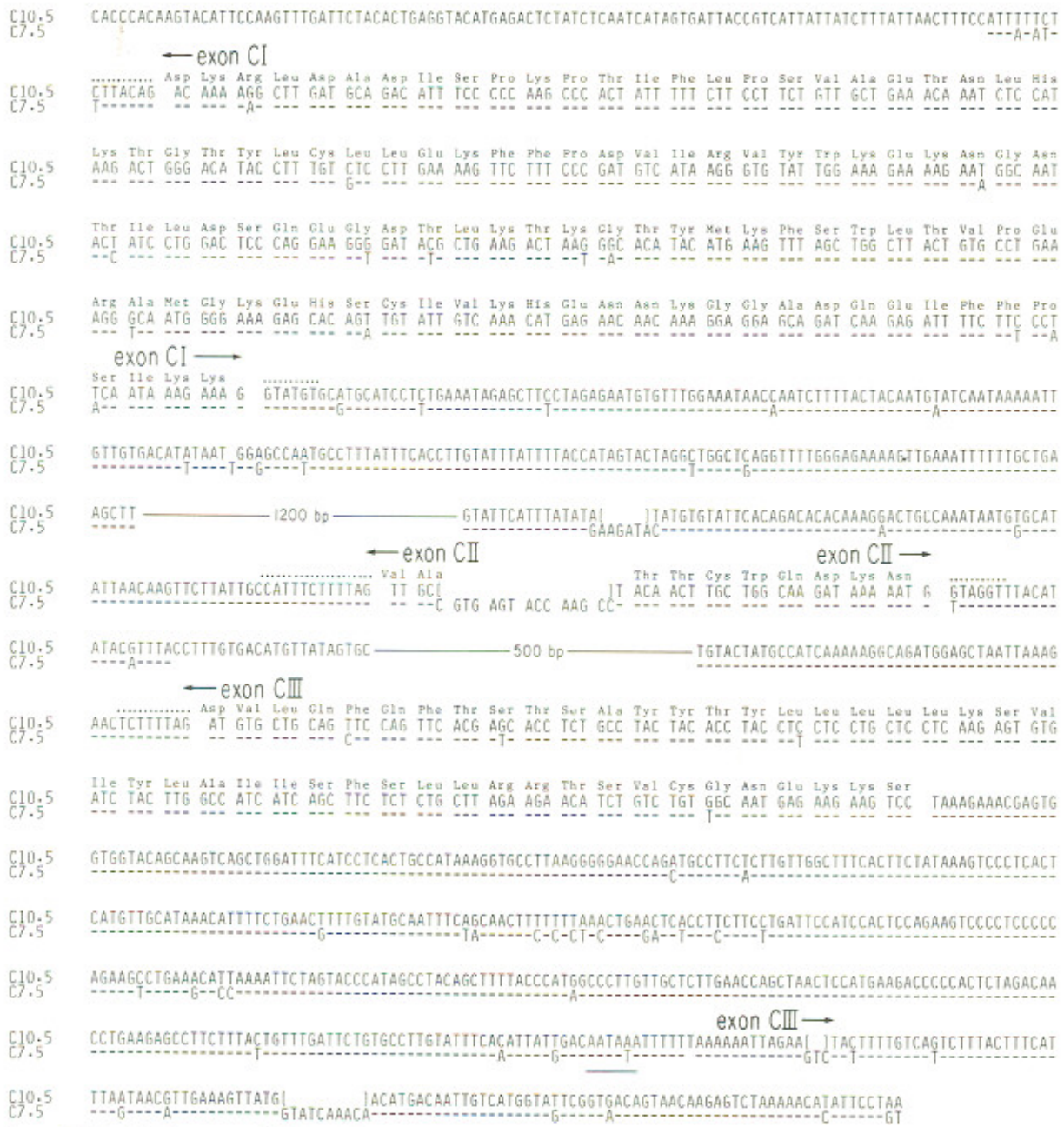

Figure 5. Nucleotide Sequences of C10.5 and C7.5 and Coding Potential of C10.5

The intron–exon organization was determined by comparison with clone 4/203 (Saito et al., 1984a), and with other unpublished cDNA clones. Splice junction sequences are overlined (dotted), and a putative polyadenylation signal (Proudfoot and Brownlee, 1976), which appears defective in C7.5, is underlined twice.

two leucine residues between the first tyrosine residue and the lysine. If an α helix is adopted by the transmembrane domain, the residues Y, L, K, Y encoded by γ genes, and by the β chain genes would all face the same surface of the helix. The third exon also encodes a small C-terminal stretch of hydrophilic residues, which we have proposed to extend into the cytoplasm (Saito et al., 1984a). Conspicuous among these residues is a cysteine. By contrast, genes for the T cell receptor β chain constant regions encode the short cytoplasmic domain in a fourth exon, separate from that encoding the putative transmembrane domain (Gascoigne et al., 1984). The third exon of the constant region described here also includes a 3′ untranslated sequence of over 400 nucleotides.

Noncoding Regions

A composite VJC gene, generated by the rearrangement of gene segments described in this paper, would contain five exons and four introns. As for immunoglobulin genes, the exon–intron boundaries occur predominantly after the first nucleotide of a coding triplet. The sequences at the intron–exon junctions are denoted by dots in Figures 4 and 5. They adhere closely to consensus sequences for

Figure 6. The Proposed Transmembrane Domains Encoded by the Gene Family Described Here and by the T Cell Receptor β Chain Genes, α Chain Genes, and Ig Genes

Sequences derive from: (a) this study and Saito et al. (1984a); (b), (c), Gascoigne et al. (1984); (d) Yanagi et al. (1984); (e) Saito et al. (1984b); (f) Tyler et al. (1982); (g) Rogers et al. (1981); (h) Rogers et al. (1980).

splicing sites (Breathnach and Chambon, 1981; Mount, 1981), with the exception of the splice donor at the end of exon 2 of C7.5. Here, the sequence TTAGGT . . . contrasts with the universal occurrence of GTNN. . . . Therefore, C7.5 is unlikely to give rise to an RNA that can be properly spliced.

The initiation point of the messenger RNA from which the cDNA cloned from 2C cells is derived has not been mapped. However, a candidate start site exists at nucleotides 169–170 on V108 (marked as nucleotide 1 on Figure 4). The sequence in this region, CCTTTAACCTAC, closely resembles an mRNA cap site consensus sequence, Py – – – PyAPyPyPyPyPy (Breathnach and Chambon, 1981). The occurrence at the start site of neighboring A residues, instead of a single A residue, is unusual but not without precedent (Breathnach and Chambon, 1981). At 28 bp upstream of the proposed cap site is the sequence CATACAAAAGATG. This can be compared with the consensus TATA box sequence (M. Goldberg, Ph.D. thesis, Stanford University, 1979; Breathnach and Chambon, 1981) $TATA^{A}/_{T}A^{A}/_{T}$–G––G. Although the C residues in the sequence described here are in unusual locations, the sequence, relative to the consensus sequence, retains ATA followed by an A-rich stretch, followed by G––G. It is therefore conceivable that this sequence acts as a TATA box for V108 transcription. Upstream of this sequence (at about −61 bp) is the sequence GGCCATTCT. This sequence is only one nucleotide removed from the consensus "CAT box" sequence (Efstratiadis et al., 1980; Breathnach and Chambon, 1981) GGCCAATCT, which occurs about 70 bp upstream of numerous structural genes. Furthermore, at about 132 bp upstream of the proposed initiation site is the sequence GAATATAGCAAAT. Almost exact copies of this sequence occur 121 bp upstream of the proposed start site for human V_H HG3 (Rechavi et al., 1983), and 112 bp upstream of the proposed start site for murine V_H108A (Givol et al., 1981). The latter two sequences are part of a "cd element," noted by Falkner and Zachau (1984) as a highly conserved sequence occurring between 85 and 117 bp upstream of the proposed start sites of murine and human heavy chain variable region genes.

Initiation of Translation

A putative initiator methionine codon has been denoted for V108A, and for V108B (Figure 4), based on analogy to Ig genes. Initiation at this point would produce a hydrophobic leader segment of 18 amino acids. However, neither for V108A nor for V108B (were it to be transcribed) is this *met* codon the first *met* codon in the presumed messenger RNA. Rather, an ATG occurs 134 bp upstream, at the start of a reading frame that remains open for eight amino acids (Figure 4). Initiation at an internal ATG has been considered previously (e.g., Mulligan and Berg, 1981). However, in the work described by Mulligan and Berg, the deletion of upstream ATG triplets led to increased protein levels from the internal ATG. Furthermore, a consensus sequence has been proposed for an efficient eukaryotic translational start site. This sequence is $CC^{A}/_{G}CCATGG$ (Kozak, 1984). In particular, the A residue 3 bp before the ATG is thought to influence translational efficiency greatly. The initiation site that we have proposed, TTGGGATGC, does not resemble the consensus sequence. The upstream initiation point has actually a slightly better match—AGATCATGA. For these reasons, the translation of the primary product of the genes described here may be a relatively inefficient process and, therefore, possibly a stage at which expression of the genes can be regulated.

The Somatic Generation of a Productively Rearranged Gene

Three rearranged Eco RI fragments can be detected with a clone 4/203 cDNA probe in 2C DNA. The sizes of these fragments are 22 kb, 16 kb, and 5 kb (Figure 1). A priori, any one of these fragments may contain the productively rearranged gene that gave rise to clone 4/203 cDNA. However, use of separate V and C probes (Figure 1) shows that the 22 kb band hybridizes only to C, and the 5 kb band only to V, whereas the 16 kb band hybridizes to both C and V. The failure of the 22 kb band to hybridize to a homologous V probe suggests that it does not represent the productive VJC rearrangement. The failure of the 5 kb band to hybridize to C may reflect movement of V to a J sequence outside of an Eco RI fragment that contains C. In this case the associated C gene segments will remain in germ-line configuration. This hypothesis is supported by the fact that the 5 kb band does react with the J region probe, gJ (Figures 2 and 3). Since the only C-reactive Eco RI fragment that remains in germ-line configuration is the 7.5 kb band, it is likely that the joined VJ is associated with C7.5. However, as discussed above, C7.5 is defective. Furthermore, it contains numerous single-base differences, by comparison with the clone 4/203cDNA sequence (Figure 5). Therefore, it is likely that this rearrangement is nonproductive and distinct from the one represented by the clone 4/203cDNA. Thus the remaining rearranged Eco RI fragment (16 kb), which reacts with V and C probes, constitutes the best candidate for the productive rearrangement.

Of the five germ-line clones (Figure 2) that make up the gene family described here it is V10.8A and J-C10.5 that, in all likelihood, recombine to form the productive gene.

Figure 7. Sequences of V108A, cDNA 4/203 and J10.5 in the Proposed Region of V-J Joining

An AT dinucleotide in 4/203, which cannot be derived from V108A or J10.5, is underlined. The heptamer–nonamer signals of V108A and J10.5 are printed in heavy type.

Figure 8. Heptamer–nonamer Sequences That Flank the V and J Segments of the Gene Family Described Here, Compared with Those That Flank T Cell Receptor β Chain Gene Segments, and Ig Light Chain Gene Segments

The numbers denote the nucleotide distance between each element. The sequences derive from the following sources: V108A, V108B, J10.5, J13.4, this study; V_β2B4, J_β1'3,4,5, Chien et al. (1984a); J_β1.1 Gascoigne et al. (1984); IgV$_\kappa$21.B, Heinrich et al. (1984); IgJ$_\kappa$.1, Sakano et al. (1979a), IgV$_\lambda$1, IgJ$_\lambda$1, Bernard et al. (1978).

The reasons are as follows: this recombination generates a rearranged Eco RI fragment of the correct size (i.e., 16 kb); the V108 sequence agrees, over the whole V segment, exactly with that of the cDNA (Saito et al., 1984a); V108B and V5.7 do not; the sequence of C10.5 agrees exactly with that of the cDNA (Figure 5); the sequence of C7.5 does not; and furthermore, the sequence of J-C10.5 agrees almost exactly with the J segment sequence of the cDNA; J-C13.4 does not (see Figure 7). The only differences are an A and a T residue that occur in the cDNA at the V–J junction, which do not occur in either V10.8 or J-C10.5 (Figure 7). Can this AT dinucleotide be explained by use of an as yet undetected J region? As stated above, the J regions have been mapped, on J-C10.5 and J-C13.4, by hybridization to small restriction fragments. Direct DNA sequencing was then performed on 750 bp of C10.5, and 250 bp of C13.4 in the appropriate regions (Figure 4). These analyses strongly suggested that no more J regions occurred on these clones. Furthermore, the J region of clone 4/203 contains an Ava I site, which is contained in J-10.5, but which is not present in or around J-13.4. The use of a J segment from the 2.3 kb J region is excluded by the result that the C-containing fragment itself rearranges to 16 kb (discussed above).

The inclusion of extra nucleotides at the VJ junctions of Ig genes is very common (reviewed by Tonegawa, 1983). The nucleotides may be inserted in a template-independent fashion, for example by terminal transferase (Alt and Baltimore, 1982). An AT dinucleotide is not the favored substrate of terminal transferase. In addition, this "unexplained" dinucleotide occurs at three out of four junctions that we have examined (D. M. K. et al., submitted; J. S. Heilig, unpublished). This situation is reminiscent of the analysis of Ig heavy chain junctions, where it was found that varying numbers of nucleotides, often of similar sequence from junction to junction, could not be explained by V or J sequences (Early et al., 1980; Sakano et al., 1980). In that case, the discovery of short D segments (Sakano et al., 1981) that link the V region to the J region explained most of the extra bases. Does the AT dinucleotide derive from a D region? Although it is unusual for a D segment to contribute consistently only 2 bp to a VJ junction (and the same 2 bp at that) (Kurosawa and Tonegawa, 1982), the discovery of D segments in the T cell receptor β chain gene families (Kavaler et al., 1984; Siu et al., 1984b) increases the possibility that this gene family also utilizes a D region. Therefore, it is quite possible that the productive rearrangement that occurs in 2C cells involves V10.8, D, and J-C10.5. The expressed product of this rearrangement is *not* somatically altered in its V region, relative to the germ-line V sequence.

Consensus Signal Sequences for Gene Rearrangement

The germ-line V, D, and J segments of Ig genes that can be rearranged are flanked by highly conserved heptamer and nonamer sequences (reviewed by Tonegawa, 1983). The consensus sequences for some of these are shown in Figure 8. The V108A and V108B genes of this family are both flanked to the 3′ side by the same, very similar heptamer–nonamer sequence, with a spacing in between of 23 bp (Figure 8). Similarly, both J segments sequenced are flanked to the 5′ side by a nonamer–heptamer with a 12 bp spacing (Figure 8). The V-flanking sequences of the Ig genes are essentially an inverted repeat of the J-flanking sequences (Tonegawa, 1983). This seems generally true for the consensus sequences of this family although the repeats are even more imperfect than, for instance, Vλ1-Jλ1 (Figure 8). The spacing of either 12 bp or 23 ± 1 bp between the heptamer and nonamer is a highly conserved feature of Ig genes. An empirical rule was drawn up (Early et al., 1980; Sakano et al., 1980) that gene segments could only be recombined if one was flanked by a 23 bp spacer and the other by a 12 bp spacer. Thus, according to this rule, the J and V segments of the gene family described here could be directly joined. In the T cell receptor β chain family, where VDJ joining can apparently occur, the arrangement of spacers is: V – 23 bp – 12 bp – D – D – 23 bp – 12 bp – J (Kavaler et al., 1984). The implication of this arrangement is that not only VDJ joining can occur, but also V-J and V-D-D-J. If the gene family described here does indeed include D elements, presumably the same potential flexibility will be apparent.

Discussion

This paper presents the initial characterization of the γ gene family, the members of which rearrange in T cells to produce a functional transcription unit. The organization of the gene family, the members of the gene family, and the transcription product that they produce, bear remarkable resemblance to the genes and products of immuno-

globulin loci. Like immunoglobulin (Ig) genes, the genes considered here are organized into constant (C) segments, variable (V) segments, and joining (J) segments. The first exon of the C-region genes encodes a domain of 110 amino acids that, by its length, and by its conservation of structurally significant residues, is clearly like an immunoglobulin-C-region domain. Similarly, the V-J segments can encode domains like those of Ig, with characteristic residues, such as *trp* and *cys*, at appropriate spacings, and highly conserved structures surrounding these. Furthermore, the coding capacity of the J segments is very similar to that of Igκ and Igλ J segments. The J and V segments are flanked by heptamer and nonamer sequences that are identical in organization, and similar in sequence to those of Ig genes. At least one of the V regions (V108A) has upstream of it a sequence that is conserved upstream of human and mouse Ig V_H genes and, furthermore, the V gene segment is split by an intron at exactly the same coding triplet as are all Ig heavy or light chain V segments analyzed to date.

The genes described here share many of these dramatic similarities to Ig genes, with the genes encoding the α and β chains of the heterodimeric T cell receptor (TCR). The C-region exons are similar to those of the TCR β chain C regions in encoding an Ig-like domain, a short *cys*-containing peptide, and a distinctive transmembrane domain. The J segments are even more similar to TCR J segments than to Ig J segments, and a computer search of the NIH protein sequence data bank (see Experimental Procedures) shows that the sequence most homologous to the V108A sequence described here is the sequence of a human TCR β chain V region (Yanagi et al., 1984). Clearly, the γ genes are very similar to those encoding the α and β chains of the TCR. Nevertheless, they are both distinct and unlinked in the murine genome (D. M. K. et al., unpublished; Lee et al., 1984; Caccia et al., 1984).

The study of Ig genes has demonstrated that diversity among immunoglobulin molecules is predominantly generated by three somatic mechanisms: the combinatorial diversity that results from joining multiple gene segments in different combinations, the junctional diversity that results from joining gene segments imprecisely and at different places, and the hypervariability introduced by somatic mutation of rearranged V regions (reviewed by Tonegawa, 1983). Studies of genes encoding the β chain of the heterodimeric TCR suggest that at least the former two of these mechanisms generate diversity among heterodimeric TCR molecules (Gascoigne et al., 1984; Chien et al., 1984a; Siu et al., 1984a). Preliminary characterization of α chain genes suggests that these too display combinatorial and junctional diversity (Saito et al., 1984b; Chien et al., 1984b). The gene family described here has little obvious capacity for combinatorial diversity. Of the C regions characterized, one (C7.5) is defective. If the gene family is organized so that V5.7 can only recombine to C7.5, then the defect in C7.5 will deplete the functional gene pool of V5.7. No defects were apparent from the sequence analysis of both V108B and the heptamer and nonamer sequences downstream of it. However, this V segment is oriented oppositely to the segment that is productively used, V108A. While some viewpoints of Ig gene rearrangement (Lewis et al., 1982) would regard this orientation as no preclusion to V108B rearrangement, it may be notable that we have not yet found this gene segment rearranged (D. M. K. et al., submitted). In fact, Southern analysis of several cloned CTLs and of heterogeneous thymocytes indicates that rearrangement of V10.8 to C10.5 may be the predominant productive rearrangement (D. M. K. et al., submitted). It is possible that γ gene rearrangements include further nonhomologous V gene segments, not characterized here. However, we have failed to detect a variety of rearranged C fragments that might be indicative of utilization of even a few such nonhomologous V segments (Saito et al., 1984a; D. M. K. et al., submitted). If, as we believe, combinatorial diversity is restricted by the common use of only one V segment and one C segment, then it is probably further restricted by the apparent availability of only one J (J-10.5) to mediate this rearrangement.

By contrast, junctional diversity may be a general phenomenon among the products of the γ genes. Of four junctions so far sequenced (admittedly from three different strains), no two are the same, in the position of breakage of the V segment and/or in the incorporated nucleotides that at present cannot be attributed to germ-line genes (D. M. K. et al., submitted; J. S. Heilig, unpublished). On the basis of an AT dinucleotide that occurs consistently among these "extra" nucleotides, the case has been made that γ-D segments exist. Utilization of these in gene rearrangement may increase junctional diversity further.

As for the TCR β chain genes, there is as yet no evidence of somatic mutation in the γ gene V regions. However, few cases have yet been studied. Furthermore, cases such as 2C cells represent primary-response T cells. Among immunoglobulins, somatic mutation is mainly apparent in IgG molecules of the "mature response," and is not so prevalent in IgM molecules. By analogy, somatic mutations may be present in the TCR molecules of T cells from hyperimmunized mice.

In summary, the primary characterization of the γ genes presented here reveals a family of genes, highly related to those encoding immunoglobulins or the heterodimeric T cell receptors. Despite the suggested small size of the gene family, a capacity to generate diversity is apparent.

The strong implication of such diversity is that the putative products can interact with a variety of substrates. Coupled with the likely location of the putative products on the cell membrane (the cDNA is derived from mRNA of membrane-bound polysomes and the predicted protein has a highly hydrophobic region reminiscent of a transmembrane peptide), the capability to interact with a variety of molecules suggests that molecules encoded by the gene family described here may participate, as well as the αβ heterodimer, in determining the clone-dependent specificity of T cells. Since no antisera directed against the αβ heterodimer reveals additional subunits beyond those of T3 (Allison et al., 1982; McIntyre and Allison, 1983; Haskins et al., 1983, 1984; Meuer et al., 1983b; Kranz et al., 1984; Staerz et al., 1984), it is likely that the putative

γ protein is either only loosely associated with the heterodimer or is all or part of a separate membrane protein. The hypothesis that T cells have two receptors, one for foreign antigen and one for MHC molecules, has previously been proposed in various forms and extensively debated, together with proposals for a single T cell receptor (von Boehmer et al., 1978; Blanden and Ada, 1978; Langman, 1978; Cohn and Epstein, 1978; Janeway et al., 1976; Zinkernagel and Doherty, 1979; Matzinger, 1981; Hünig and Bevan, 1982). Although experimental data (e.g., Kappler et al., 1981; Hünig and Bevan, 1982) have been regarded as ruling out some versions of the two-receptor hypothesis, other versions have not been excluded. Alternatively, the potential junctional diversity among γ gene products may enable them to interact with the same product, but with different affinities. Such interactions with, for example, a cellular antigen may constitute a step in development of cells in the thymus into mature T cells.

A second consequence of gene rearrangement in the IgH and Igκ loci of both humans and mice has recently been defined. The rearrangement brings a functional V segment promoter within the sphere of influence of a potent tissue-specific transcriptional enhancer element (Gillies et al., 1983; Banerji et al., 1983; Picard and Schaffner, 1984; Hayday et al., 1984), that exists in the DNA between J and C. It will be of interest to assay the 3.8 kb of sequence between J and C10.5 for analogous activity, especially considering the similarity of other putative regulatory sequences, such as those upstream of V108, to those of Ig genes. Transcriptional control signals for Ig genes show clear tissue specificity (Gillies et al., 1983; Banerji et al., 1983; Hayday et al., 1984). Will analogous signals for this gene family function only in T cells or do they also function in B cells? The closely related heptamer–nonamer sequences that probably mediate gene rearrangement can function in T or B cells, as judged by the high frequency of Ig D-J rearrangement in T cells.

Experimental Procedures

Southern Analyses

Analyses were performed on restricted DNA, electrophoresed in horizontal agarose gels, and transferred to nitrocellulose. Filters were hybridized to nick-translated probes (Rigby et al., 1977) according to Wahl et al. (1979), and were filmed at −70°C using intensifying screens.

Molecular Cloning and Analysis

Eco RI-cleaved embryonic DNA was cloned into either λ-wes (Leder, et al., 1977) or λ Charon 4A (Blattner et al., 1977), according to previously described procedures (Sakano et al., 1979b), and was screened with radioactive probes (Figure 2) (Benton and Davis, 1977). After restriction mapping (Smith and Bernstiel, 1976) fragments of the clones were subcloned into plasmid before being sequenced (Maxam and Gilbert, 1980). Sequences were analyzed using SEQ programs (Brutlag et al., 1982) and homologies of predicted amino acid sequences sought using the protein sequence database, compiled by W. C. Barker, L. T. Hunt, B. C. Orcutt, D. G. George, L. S. Yeh, H. R. Chen, M. C. Blomquist, G. C. Johnson, E. I. Seibel-Ross, and R. S. Ledley, at the National Biomedical Research Foundation, Washington, D.C. 20007.

Acknowledgments

We are indebted to Howard Siegel and to John McMaster for their enthusiastic assistance, and to Eleanor Basel for the painstaking preparation of this manuscript. We thank Y. Takagaki and Federico Robbiati for helpful discussions. This work has been supported by grants CA-28900-04 (S. T.) and CA-14051 (a core grant to S. E. Luria). D. M. K. is supported by the Arthritis Foundation.

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 November 5, 1984; revised December 3, 1984

References

Acuto, O., Hussey, R. E., Fitzgerald, K. A., Protentis, J. P., Meuer, S. C., Schlossman, S. F., and Reinherz, E. L. (1983). The human T cell receptor: appearance in ontogeny and biochemical relationship of α and β subunits on IL-2 dependent clones and T cell tumors. Cell *34*, 717–726.

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

Allison, J. P., McIntyre, B. W., and Bloch, D. (1982). Tumor specific antigen of murine T-lymphoma defined with monoclonal antibody. J. Immunol. *129*, 2293–2300.

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. Natl. Acad. Sci. USA *79*, 4118–4122.

Azuma, T., Igras, V., Reilly, E., and Eisen, H. N. (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.

Banerji, J., Olson, L., and Schaffner, W. (1983). A lymphocyte-specific cellular enhancer is located downstream of the joining region in immunoglobulin heavy chain genes. Cell *33*, 729–739.

Benton, W., and Davis, R. (1977). Screening λ gt recombinant clones by hybridisation to single plaques *in situ*. Science *196*, 180.

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

Blanden, R. V., and Ada, G. L. (1978). A dual recognition model for cytotoxic T cells based on thymic selection of precursors with low affinity for self H-2 antigens. Scand. J. Immunol. *7*, 181–190.

Blattner, F. R., Williams, B. G., Blechl, A. E., Denniston-Thompson, K., Faber, H. E., Fulong, L.-A., Grunwald, D. J., Kiefer, D. O., Moore, D. D., Sheldon, E. L., and Smithies, O. (1977). Charon phages: safer derivatives of bacteriophage λ for DNA cloning. Science *196*, 161.

Blomberg, B., and Tonegawa, S. (1982). Organisation and expression of mouse λ light chain immunoglobulin genes. In Gene Regulation. (New York: Academic Press).

Bothwell, A. L. M., Paskind, M., Reth, M., Imanishi-Kari, T., Rawesky, K., and Baltimore, D. (1981). Heavy chain variable region contribution to the NP^b family of antibodies: somatic mutation evident in a γ2a variable region. Cell *24*, 624.

Breathnach, R., and Chambon, P. (1981). Organisation and expression of eukaryotic split genes, coding for protein. Ann. Rev. Biochem. *50*, 349.

Brutlag, D. L., Clayton, J., Friedland, P., and Kedes, L. (1982). SEQ: a nucleotide sequence analysis and recombination system. Nucl. Acids. Res. *10*, 279–294.

Caccia, N., Kronenberg, M., Saxe, D., Haars, R., Bruns, G. A. P., Goverman, J., Malissen, M., Willard, H., Yoshikai, Y., Simon, M., Hood, L., and Mak, T. W. (1984). The T cell receptor β chain genes are located on chromosome 6 in mice and chromosome 7 in humans. Cell *37*, 1091–1099.

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

Chien, Y.-h., Becker, D. M., Lindsten, T., Okamura, M., Cohen, D. J.,

and Davis, M. M. (1984b). Third type of murine T cell receptor gene. Nature *312*, 31–35.

Cohn, M., and Epstein, R. (1978). T cell inhibition of humoral responsiveness. II. Theory on the role of restrictive recognition in immune regulation. Cell. Immunol. *39*, 125–153.

Crews, S., Griffin, J., Huang, H., Calame, K., and Hood, L. (1981). A single V_H gene segment encodes the immune response to phosphorylcholine: somatic mutation is correlated with the class of the antibody. Cell *25*, 59–66.

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

Efstratiadis, A., Posakony, J. W., Maniatis, T., Lawn, R., O'Connell, C., Spritz, R. A., DeRiel, J. K., Forget, B. G., Weissman, S. M., Slightom, J. L., Blechl, A. E., Smithies, O., Barelle, F. E., Shoulders, C. C., and Proudfoot, N. J. (1980). The structure and evolution of the human β-globin gene family. Cell *21*, 653.

Falkner, F. G., and Zachau, H. (1984). Correct transcription of an immunoglobulin κ gene requires an upstream fragment containing conserved sequence elements. Nature *310*, 71–74.

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

Gillies, S. D., Morrison, S. L., Oi, V. T., and Tonegawa, S. (1983). A tissue-specific transcriptional enhancer element is located in the major intron of a rearranged immunoglobulin heavy chain gene. Cell *33*, 717–728.

Givol, D., Zakut, R., Effron, K., Rechavi, G., Ram, D., and Cohen, J. B. (1981). Diversity of germ-line immunoglobulin V_H genes. Nature *292*, 426–430.

Hannum, T. H., Kappler, J. W., Trowbridge, I. S., Marrack, P., and Freed, J. H. (1984). Immunoglobulin-like nature of a α chain of a human T cell antigen/MHC receptor. Nature *312*, 65–67.

Haskins, K., Kubo, R., White, J., Pigeon, M., Kappler, J., and Marrack, P. (1983). The major histocompatibility complex-restricted antigen receptor in T cells. I. Isolation with a monoclonal antibody. J. Exp. Med. *157*, 1149–1169.

Haskins, K., Hannum, C., White, J., Roehm, N., Kubo, R., Kappler, J., and Marrack, P. (1984). The antigen specific, major histocompatibility complex restricted receptor on T cells. VI. Antibody to a receptor allotype. J. Exp. Med. *160*, 452–471.

Hayday, A. C., Gillies, S. D., Saito, H., Wood, C., Wiman, K., Hayward, W. S., and Tonegawa, S. (1984). Activation of translocated human c-*myc* by an enhancer in the immunoglobulin heavy chain locus. Nature *307*, 334–340.

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

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

Hunig, T. R., and Bevan, M. J. (1982). Antigen recognition of cloned cytotoxic T lymphocytes follows rules predicted by the altered self hypothesis. J. Exp. Med. *155*, 111–125.

Janeway, C. A., Jr., Wigzell, H., and Binz, H. (1976). Hypothesis: two different V_H gene products make up the T cell receptors. Scand. J. Immunol. *5*, 993–1001.

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

Kappler, J. W., Skidmore, B., White, J., and Marrack, P. (1981). Antigen-inducible, H2 restricted, IL-2 producing T cell hybridomas. Lack of independent antigen and H-2 recognition. J. Exp. Med. *153*, 1198–1214.

Kappler, J., Kubo, R., Haskins, K., Hannum, C., Marrack, P., Pigeon, M., McIntyre, B., Allison, J., and Trowbridge, I. (1983). The major histocompatibility complex-restricted antigen receptor on T cells in mouse and man: identification of constant and variable peptides. Cell *35*, 295–302.

Kavaler, J., Davis, M. M., and Chien, Y.-h. (1984). Localisation of a T cell receptor diversity-region element. Nature *310*, 295–302.

Kozak, M. (1984). Point mutations close to the AUG initiator codon affect the efficiency of translation of rat pre-proinsulin *in vivo*. Nature *308*, 241–246.

Kranz, D. M., Sherman, D. H., Sitkovsky, M. V., Pasternack, M. S., and Eisen, H. N. (1984). Immunoprecipitation of cell surface structures of cloned cytotoxic T lymphocytes by clone-specific antisera. Proc. Natl. Acad. Sci. USA *81*, 573–577.

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

Langman, R. E. (1978). The role of the major histocompatability complex in immunity: a new concept in the functioning of a cell-mediated immune system. Rev. Physiol. Biochem. Pharmacol. *81*, 1–37.

Leder, P., Tiemeier, D., and Enquist, L. (1977). EK2 derivatives of bacteriophage lambda useful in the cloning of DNA from higher organisms: the λ gt WES system. Science *196*, 175.

Lee, N., D'Eustachio, P., Pravtcheva, D., Ruddle, F. H., Hedrick, S. M., and Davis, M. M. (1984). Murine T cell receptor β-chain is encoded on chromosome 6. J. Exp. Med. *160*, 905–913.

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

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

Marrack, P., Shimankevitz, R., Hannum, C., Haskins, K., and Kappler, J. (1983). The major histocompatability complex-restricted antigen receptor on T cells. IV. An anti-idiotypic antibody predicts both antigen and I-specificity. J. Exp. Med. *158*, 1635–1646.

Matzinger, P. (1981). A one-receptor view of T cell behaviour. Nature *292*, 497–501.

Maxam, A., and Gilbert, W. (1980). Sequencing end-labeled DNA with base-specific chemical cleavages. Meth. Enzymol. *65*, 499.

McIntyre, B. W., and Allison, J. P. (1983). The mouse T cell receptor: structural heterogeneity of molecules of normal T cells defined by xenoantiserum. Cell *34*, 739–746.

Meuer, S. C., Fitzgerald, K. A., Hussey, R. E., Hogdon, J. C., Schlossman, S. F., and Reinherz, E. L. (1983a). Clonotypic structures involved in antigen-specific human T cell function. Relationship to the T3 molecular complex. J. Exp. Med. *157*, 705–719.

Meuer, S. C., Cooper, D. A., Hodgdon, J. C., Hussey, R. E., Fitzgerald, K. A., Schlossman, S. F., and Reinherz, E. L. (1983b). Identification of the receptor for antigen and major histocompatability complex on human inducer T lymphocytes. Science *222*, 1239–1242.

Mount, S. M. (1981). A catalogue of splice junction sequences. Nucl. Acids Res. *10*, 457–472.

Mulligan, R. C., and Berg, P. (1981). Factors governing the expression of a bacterial gene in mammalian cells. Mol. Cell. Biol. *1*, 449–459.

Picard, D., and Schaffner, W. (1984). A lymphocyte specific enhancer in the murine Igk gene. Nature *307*, 80–82.

Proudfoot, N. J., and Brownlee, G. G. (1976). 3' Non-coding region sequences in eukaryotic messenger RNA. Nature *263*, 211.

Rechavi, G., Ram, D., Glazer, L., Zakut, R., and Gival, D. (1983). Evolutionary aspects of immunoglobulin heavy chain variable region (V_H) gene subgroups. Proc. Natl. Acad. Sci. USA *80*, 855–859.

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

Rogers, J., Early, P., Carter, C., Calame, K., Bond, M., Hood, L., and Wall, R. (1980). Two mRNA with different 3' ends encode membrane-bound and secreted forms of Ig μ chains. Cell *20*, 303–312.

Rogers, J., Chai, E., Souza, L., Carter, C., Word, C., Kuehl, M., Eisenberg, D., and Wall, R. (1981). Gene segments encoding the transmembrane carboxy termini of immunoglobulin γ chains. Cell *26*, 19–27.

Saito, H., Kranz, D. M., Takagaki, Y., Hayday, A. C., Eisen, H. N., and Tonegawa, S. (1984a). Complete primary structure of a heterodimeric T cell receptor derived from cDNA sequences. Nature *309*, 757–762.

Saito, H., Kranz, D. M., Takagaki, Y., Hayday, A. C., Eisen, H., and Tonegawa, S. (1984b). A third rearranged and expressed gene in a clone of cytotoxic T lymphocytes. Nature *312*, 36–39.

Sakano, H., Rogers, J. H., Huppi, K., Brack, C., Traunecker, A., Maki, R., Wall, R., and Tonegawa, S. (1979a). Domains and the hinge region of an immunoglobulin heavy chain are encoded in separate DNA segments. Nature *277*, 627–633.

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

Sakano, H., Maki, R., Kurasawa, 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.

Sakano, H., Kurosawa, Y., Weigert, M., and Tonegawa, S. (1981). Identification and nucleotide sequence of a diversity DNA segment (D) of immunoglobulin heavy-chain genes. Nature *290*, 562.

Siu, G., Clark, S. P., Yoshikai, Y., Malissen, M., Yanagi, Y., Strauss, E., Mak, T. W., and Hood, L. (1984a). The human T cell antigen receptor is encoded by variable, diversity, and joining gene segments that rearrange to generate a complete V gene. Cell *37*, 393–401.

Siu, G., Kronenberg, M., Strauss, E., Haars, R., Mak, T. W., and Hood, L. (1984b). The structure, rearrangement and expression of Dβ gene segments of the murine T-cell antigen receptor. Nature *311*, 344–350.

Smith, H. O., and Bernstiel, M. (1976). A simple method for DNA restriction site mapping. Nucl Acids Res. *3*, 238.

Southern, E. (1975). Detection of specific sequences among DNA fragments separated by gel electrophoresis. J. Mol. Biol. *98*, 503.

Staerz, U. D., Pasternack, M. S., Klein, J. R., Benedetto, J. D., and Bevan, M. J. (1984). Monoclonal antibodies specific for a murine cytotoxic T-lymphocyte clone. Proc. Natl. Acad. Sci. USA *81*, 1799–1803.

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

Tonegawa, S., Sakano, H., Maki, R., Traunecker, A., Heinrich, G., Roeder, W., and Kurosawa, Y. (1981). Somatic reorganisation of immunoglobulin genes during lymphocyte differentiation. Cold Spring Harbor Symp. Quant. Biol. *40*, 839–857.

Tyler, B. M., Cowman, A. F., Gerondakis, S. D., Adams, J. M., and Bernard, O. (1982). mRNA for surface immunoglobulin λ chain encodes a highly conserved transmembrane sequence and a 28-residue intracellular domain. Proc. Natl. Acad. Sci. USA *79*, 2008–2112.

von Boehmer, H., Haas, W., and Jerne, N. K. (1978). Major histocompatibility complex-linked immune responsiveness is acquired by lymphocyte of low responder mice differentiating in the thymus of high responder mice. Proc. Natl. Acad. Sci. USA *75*, 2439–2442.

Wahl, G. M., Stern, M., and Stark, G. R. (1979). Efficient transfer of large DNA fragments from agarose gels to diazobenzyloxymethyl paper and rapid hybridisation using dextran sulphate. Proc. Natl. Acad. Sci. USA *76*, 3683.

Yanagi, Y., Yoshikai, Y., Leggett, K., Clark, S. P., Aleksander, I., and Mak, T. W. (1984). A human T cell-specific cDNA clone encodes a protein having extensive homology to immunoglobulin chains. Nature *308*, 145–149.

Zinkernagel, R. M., and Doherty, P. C. (1979). MHC-restricted cytotoxic T cells: studies on the biological role of polymorphic major transplantation antigens determining T-cell restriction specificity, function, and responsiveness. Adv. Immunol. *27*, 52–177.