The Genomic Structure of Human Vβ6 T Cell Antigen Receptor Genes

By Yixin Li, Paul Szabo, and David N. Posnett

From the Department of Medicine, Cornell University Medical College, the Immunology Graduate Program, and the Department of Cell Biology and Genetics, Cornell University Graduate School of Medical Sciences, New York, New York 10021

Summary

Six genomic clones were characterized containing members of the human Vβ6 subfamily of T cell antigen receptor genes. There were four major findings. (*a*) New Vβ genes were discovered, including Vβ6.10, Vβ13.4, Vβ13.5, and Vβ5.5. (*b*) Members of the Vβ13, Vβ6, and Vβ5 subfamilies cluster together in the Vβ locus and may have evolved through multiple duplication events of an ancestral cassette containing Vβ13-Vβ6-Vβ5 genes. These Vβ subfamilies are used by an estimated one-third of T cells in humans and probably represent a highly useful component of the Vβ repertoire. (*c*) The promoters of Vβ13, Vβ6, and Vβ5 genes contain conserved decamer motifs, but discrete differences were observed between promoters of different Vβ subfamilies, raising the question of different transcriptional control depending on Vβ subfamily usage. (*d*) The new Vβ6.10 gene is probably a pseudogene, which may have been inactivated due to retrotransposition of Alu elements into its promoter region, a mutation affecting a highly conserved cysteine residue or mutations of the 3′ recombinase signal sequence.

The α/β TCR is responsible for recognition of the antigen/MHC complex of APC and is expressed by ~95% of human peripheral blood T cells. To create sufficient diversity for recognition of a huge number of different antigenic specificities, several mechanisms are used (1). First, two sets of genomic gene segments are available to each T cell encoding Vβ, Dβ, Jβ, Cβ, Vα, Jα, and Cα segments. The V gene loci contain approximately >50 Vα and >57 Vβ gene segments (2–9). Second, V-D-J gene rearrangement mediated by recombinase provides considerable combinatorial diversity. Third, junctions at which gene segments are joined are characterized by nucleotide loss and addition of nucleotides by terminal deoxy-transferase, thus creating what is termed "N-region" diversity.

Because most TCR gene sequences are obtained from rearranged cDNA clones, very little is known about the germline structures of V gene segments. These are of particular interest because each V gene is preceded by its own individual promoter, which will control the expression of the mature, functionally rearranged, β chain (10–12). In contrast, TCR α and β enhancers are located 3′ to the respective C regions (13–17) and can therefore be used by every T cell. Moreover, genomic sequences of V gene segments can provide useful information on the existence of possible pseudogenes, on introns, on signal sequences for recombinase, and on allelic forms of V genes. The latter question is of some importance, since allelic variations of V gene segments are apparently a common finding (18–21), and could possibly be related to disease susceptibility, although there are no conclusively established examples yet (18).

Materials and Methods

Probes. The following probes were used for screening a genomic library, for restriction mapping, and for Southern blot analyses: (*a*) a 108-bp EcoRI/BstXI fragment derived from the cDNA clone OT-1 (20), designated OT-V, which represents the 3′ end of the Vβ6.7a gene sequence; (*b*) probe Vph79 derived from the cDNA clone ph79 (4), containing 249 bp of Vβ6.7a 5′ to the internal BamHI site of Vβ6.7a (20); (*c*) an intron probe derived from clone GL-PA (20) containing a 149-bp HinfII fragment corresponding to the intron sequences of Vβ6.7b; (*d*) probe Vph5 derived from the cDNA clone ph5 (4), covering 276 bp of the coding sequence of Vβ6.3, from the ATG codon to the internal BamHI site; (*e*) two additional Vβ probes derived from the cDNA clone pCEM-1 (22) (Vβ13.2) and from the clone p12A1 (23) (Vβ5.3), kindly provided by Drs. A. Duby and J. Leiden, respectively; (*f*) the UAS-2 clone, kindly provided by A. E. Hinkkanen and J. T. Epplen (Max-Planck-Institute for Immunology, Freiburg, Germany), containing a rearranged Vβ6.7a gene segment (24).

Southern Blotting. Human genomic DNA was isolated from EBV-transformed B lymphoblastoid cell lines, digested with BamHI or other restriction enzymes, separated by 0.7% agarose gel electrophoresis, and then transferred to charge-modified nylon-66 filters (Gelman Sciences, Inc., Ann Arbor). After baking in a vacuum oven for 2 h, the blot was prewashed at 65°C 0.1× SSC in 0.5% SDS for 1 h, and prehybridized at 42°C in 5× SSPE, 5× Denhardt's, 50% formamide, 100 μg/ml denatured salmon sperm DNA,

 J. Exp. Med. © The Rockefeller University Press • 0022-1007/91/12/1537/11 $2.00
Volume 174 December 1991 1537–1547

and 0.1% SDS. Specific probes for hybridization were labeled with ^{32}P, using a random primer extension labeling kit (Boehringer Mannheim Biochemicals, Indianapolis, IN). Hybridization was performed in the prehybridization buffer with labeled probes at 10^6 cpm/ml. After hybridization, blots were washed twice with 2× SSC, 0.5% SDS for 20 min each, followed by two washes with 0.1× SSC, 0.5% SDS for 20 min each at 68°C. The filters were then exposed to Kodak XAR-5 film at −70°C with an intensifying screen. Blots were rehybridized, after removing the original signal, by incubating the blot in the prehybridization buffer at 65°C for 30 min followed by prehybridization at 42°C overnight.

Screening of the Human Genomic Library. A human genomic library consisting of placental DNA, partially digested with Sau3A, and ligated into the BamHI site of λ-EMBL3 (Clontech Laboratories, Inc., Palo Alto, CA) was obtained. The titer of the library was 3×10^9/ml. Approximately 10^7 phage were screened on 150-mm plates at a density of 5×10^5 plaques per plate. Replicate nitrocellulose filters (Schleicher & Schuell, Inc., Keene, NH) were baked under vacuum at 80°C for 2 h, prewashed with 0.1× SSC, 0.5% SDS at 65°C for 30 min, prehybridized with 5× SSPE, 5× Denhart's, 50% formamide, 100 μg/ml denatured salmon sperm DNA, 0.1% SDS at 42°C for 6 h, then hybridized overnight with ^{32}P-labeled Vβ probes (at 5×10^5 to 1×10^6 cpm/ml). The filters were washed first at room temperature with 2× SSC, 0.5% of SDS, and then twice at 65°C, followed by two washes in 0.1× SSC 0.5% SDS at 65°C, for 20 min each. After washing, filters were dried and exposed to Kodak XAR-5 film at −70°C with an intensifying screen. After the first round of screening, positive phages were selected and soaked out of the agar gel plugs into SM buffer (100 mM NaCl, 80 mM $MgSO_4$, 50 mM Tris-Cl, pH 7.5). These phages were further purified until homogeneous. λ DNA was prepared from the positive clones by liquid culture and the cesium chloride banding method (25).

Subcloning and Restriction Mapping of Genomic Clones. Purified λ DNA samples from positive clones were digested first by SalI and run on a 0.6% agarose gel. The average length of inserts in Vβ6-positive clones was ~15 kb. Each insert was subcloned into the SalI site of a plasmid vector PIBI31. After transformation into the host bacteria DH5α, the plasmid DNA was isolated from 25–50 ml bacterial culture suspension by the boiling method (25). For restriction mapping, the enzymes BamHI, EcoRI, KpnI, and SacI were used in single or double digestion reactions. The restriction maps were constructed and confirmed by Southern blot analysis using the probes as described above.

The transcriptional orientations of the Vβ genes were determined by Southern blotting and sequences analysis. For instance, specific probes corresponding to the 5′ and 3′ portions of Vβ6.7a or Vβ13.2 were used to delineate the location and orientation of the Vβ6.7a and Vβ13.2 gene segments in clone 5-2. In clone 11, a 3.9-kb KpnI fragment spans the Vβ6.10 and Vβ13.3 genes. By subcloning and sequencing this fragment the orientation of both genes could be determined. The same strategy was used for Vβ6.1 and Vβ5.5 in clone 9. For clone 4-1, a PvuII site located in the coding region of Vβ13.4 (not shown) was used to determine the orientation of Vβ13.4.

Sequencing. Relevant restriction fragments were subcloned into PIBI31. The plasmid DNA was isolated from transformed DH5α and the sequence determined by the dideoxy chain termination sequencing method of Sanger et al. (26), using the universal T3 and T7 primers, as well as specifically synthesized oligonucleotide primers. Both strands were completely sequenced. The computer program used to align sequences and compute homologies was "genalign" using the "clustered pair-wise region method" (27). Promoter sequences were analyzed by D. Ghosh (National Center for Biotechnology Information, NIH, Bethesda, MD) with "SITES, table of release 3.0 of TFD" (28). The identical match (IM) probability is the probability that the SITES data base sequence would match with 100% identity somewhere in a random sequence with equivalent length to the sequence being analyzed.

Results

Genomic Organization of Vβ6 Genes. Six genomic clones were isolated from a human placental DNA library using Vβ6.7a and Vβ6.3 cDNA probes. The restriction enzyme maps of these clones are shown in Fig. 1. From these data it is clear that three clones (clones 9, 3, and 13) represent overlapping clones of the same locus. All three clones contain the Vβ6.1 gene, which was confirmed by sequence analysis of subcloned coding and flanking regions. Moreover, the Vβ6.1 gene is adjacent to two new Vβ genes, termed Vβ13.5 and Vβ5.5, which were found ~3 kb upstream and ~4 kb downstream, respectively (Fig. 1). The other three genomic clones each represent a separate locus: clone 4-1 contains the Vβ6.3 gene, clone 5-2 contains the Vβ6.7a gene, and clone 11 contains a new Vβ gene, termed Vβ6.10. Each of these Vβ6 genes is associated with a Vβ13 gene located at variable distances upstream. Two of these Vβ13 genes are newly described genes, termed Vβ13.3 and 13.4. Analysis of large restriction enzyme fragments by pulsed field gel electrophoresis, analysis of deletion variants of the TCR β locus, and analysis of cosmid clones have indicated that Vβ13, Vβ6, and Vβ5 genes are often found in close proximity to each other (29). Fig. 1 demonstrates that members of these three subfamilies are probably contained within a cassette with a conserved order of V genes (5′ Vβ13-Vβ6-Vβ5 3′). At all four loci, Vβ6 genes are preceded by an upstream Vβ13 gene. Our data are consistent with those of Lai et al. (29), and the combined data suggest that multiple duplication events of such a cassette characterize the evolution of the TCR β locus. Clones 4-1, 5-2, and 11 (Fig. 1) lack a Vβ5 gene by Southern blotting (with a Vβ5.3 and a Vβ5.5 probe) and by PCR with primers common to all known Vβ5 sequences. However, clones 4-1 and 5-2 may not extend far enough downstream to include a putative Vβ5 gene. It is therefore not possible to absolutely exclude a Vβ5 gene associated with these loci. Moreover, the cassette alluded to above may not necessarily be complete in every instance. This would explain the variable total number of V genes in each of these subfamilies, estimated at five Vβ13 genes, 8–10 Vβ6 genes, and five Vβ5 genes (8).

Nucleotide Sequences of Vβ6, Vβ13, and Vβ5 Genes. Vβ6 is the largest human Vβ subfamily. At least 12 different sequences are known (2, 20) (Li, Y., and D. Posnett, unpublished data). Differences between some sequences are minor. For instance, there is only one silent nucleotide difference between the coding regions of Vβ6.6 and Vβ6.7a (2). Such minor differences may represent allelic forms of the same gene rather than two different genes. Two allelic forms of Vβ6.7 (Vβ6.7a and Vβ6.7b) have been described in detail and encode expressed β chains that differ at two amino acid posi-

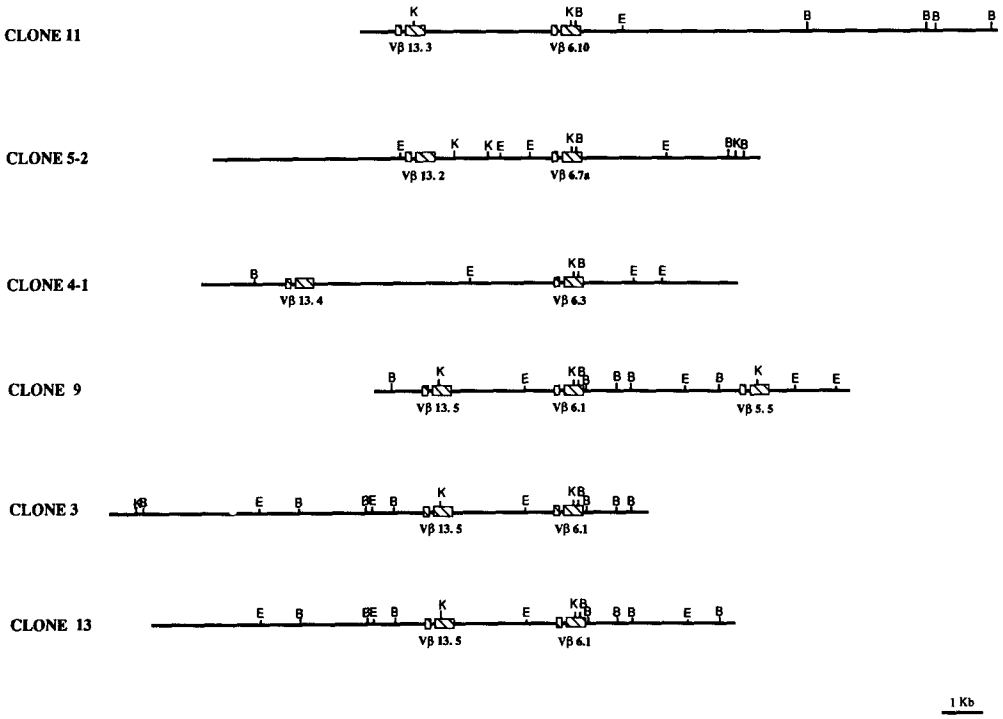

Figure 1. Restriction enzyme map of six genomic clones hybridizing with a Vβ6.7a probe. Clones 9, 3, and 13 are overlapping at the Vβ13.5-Vβ6.1-Vβ5.5 locus. The boxes represent the exons of the V genes (their size is enlarged and out of scale for clarity). B, BamHI, K, KpnI, E, EcoRI.

```
                <leader                                          > <intron
1 Vb6.3         ATGGGCACCAGGCTCCTCTGCTGGGtgGtCCTGgGttTCCTaGGGaCAG GTGAGTCCTCAGaACACCAAgtAGTtTCATT                    tTTTCTGTtTGTagG
2 Vb6.1         ATGGGCACCAGGCTCCTCTGCTGGGcaGCCCTGTGcCTCCTGGGGGCAG GTGAGTCCTCAG ACACCAAccAGTcTCATTC
3 Vb6.7a        ATGGGCACCAGGCTCCTCTTCTGGGTGGCCTTCTGTCTCCTGGGGGCAG GTGAGTCCTCAG ACACCAACCATTCTCATTgTGTGTGTGTGTGTGTGTGTGTGTGTGTGTG
4 Vb6.7a-UAS2   ATGGGCACCAGGCTCCTCTTCTGGGTGGCCTTCTGTCTCCTGGGGGCAG GTGAGTCCTCAG ACACCAACCATTCTCAT      TGTGTGTGTGTGTGTGTGTGTGTGTGTG
5 Vb6.7b        ATGGGCACCAGGCTCCTCTTCTGGGTGGCCTTCTGTCTCCTGGGGGCAt GTGAGTCCTCAG ACACCAcCCATTCTCAT           TGTGTGTGTGTGTGTGTG
6 Vb6.10        ATGGGCACaAGGCTCCTCTgCTGGGcaGCCaTaTGTCTCCTGGGGGCAG GTGAtTCCTCAG AtgCCAAgCAgTCTCc   TGTGTGTGTGTGTGTGTGTGTGTGTaTGTGTGTG

 ____(GT) repeats_____________________                                                    > <leader><V region
1 TGTGTGTGTGTGaGagAGTGGTTGTGTGTGTGTGTGTGTATGATGACTACAAaTATTTTCCTtATTCTGTTgCCAAaTTcTaTtTCCACAG ATCACACAGGTGCTGGAGTCTCCCAGtCCCCtAGg
2 TGTGTGTcTGTa        TGTGTGTGTGTGTGTGTG ATGA G  TACAAcTgTTTTCCTcATcCTGTTctCAAcTTgTgTcTCCgCAG ATCACACAGGTGCTGGAGTCTCCCAGaCCCCCAGT
3 TGTGTGTGTGTGTG TAGGGGTGGGCTGTGTGCATTCGTGTGATGACTAAAATTATTTTCCTCATTCTGTTCCCAACTT TGTCTCCACAG ATCACACAGGAGCTGGAGTCTCCCAGTCCCCCAGT
4 TGTGTGTGTGTGTG TAGGGGTGGGCTGTGTGCATTCGTGTGATGACTAAAATTATTTTCCTCATTCTGTTCCCAACTT TGTCTCCACAG ATCACACAGGAGCTGGAGTCTCCCAGTCCCCCAGT
5 TGTGTGTGTGTGTG TAGGGGTGGGCTGTGTGCATTCGTGTGATGACTAAAATTATTTTCCTCATTCTGTTCCCAACTT TGTCTCCACAG ATCACACAGGAGCTGGAGTCTCCCAGTCCCCCAGT
6 aG aTGTGTGTGTG  A   GaGaG aGaGaG A   GaGaGcTGACTAtAgTTgTTTTtCTCATTCaGTTCCCAAtTTcTGTCTCCACAG ATCACACAGGtGCTGGAGTCTCCCAGTCCCtgAGa

1 tACAAaGTCgCAaAGAgaGGAcAggATGTAGctCTCAGGTGTGATCCAATTTCgGGTCATgtatCCCTTTttTGGTACCaACAggcCCTGGGGCAGGGgCCAGAGTTTCTgAcTTAtTTCCAgaatg
2 AACAAGGTCACAGAGAAGGGAAAatATGTAGAGCTCAGGTGTGATCCAATTTCAGGTCATACTGCCCTTTACTGGTACCGACAaAGCCTGGGGCAGGGCCCAGAGTTTCTAATTTACTTCCAAGGCA
3 AACAAGGTCACAGAGAAGGGAAAGGATGTAGAGCTCAGGTGTGATCCAATTTCAGGTCATACTGCCCTTTACTGGTACCGACAGAGCCTGGGGCAGGGCCTGGAGTTTTTAATTTACTTCCAAGGCA
4 AACAAGGTCACAGAGAAGGGAAAGGATGTAGAGCTCAGGTGTGATCCAATTTCAGGTCATACTGCCCTTTACTGGTACCGACAGAGCCTGGGGCAGGGCCTGGAGTTTTTAATTTACTTCCAAGGCA
5 AACAAGGTCACAGAGAAGGGAAAGGATGTAGAGCTCAGGTGTGATCCAATTTCAGGTCATACTGCCCTTTACTGGTACCGACAGAGcCTGGGGCAGGGCCTGGAGTTTTTAATTTACTTCCAAGGCA
6 cACAAGGTagCAaAGAAGGGAAAGGATGTAGctCTCAGaTaTGATCCAATTTCAGGTCATAaTGCCCTTTAtTGGTACCGACAGAGcCTGGGGCAGGGCCTGGAGTTTccAATTTACTTCCAAGGCA

1 aaGcTcaactAGAcaAaTCgGGGCTGCCCAgtGATCGcTTCTTTGCAGaaAGGCCTGAGGGATCCGTCTCCaCTCTGAAGATCCAGCGCACAcAGaaGGaGGACTCcGCCGTGTATCTCTGTGCCAG
2 cgGGTCCggCAGAtgAcTCAGGGCTGCCCAacGATCGgTTCTTTGCAGtcAGGCCTGAGGGATCCGTCTCTcACTCTGAAGATCCAGCGCACAgAGCgGGgGGACTCaGCCGTGTATCTCTGTGCCAG
3 ACAGTGCACCAGACAAATCAGGGCTGCCCAGTGATCGCTTCTCTGCAGAGAGGACTGGGGGATCCGTCTCCACTCTGACGATCCAGCGCACACAGCAGGAGGACTCGGCCGTGTATCTCTGTGCCAG
4 ACAGTGCACCAGACAAATCAGGGCTGCCCAGTGATCGCTTCTCTGCAGAGAGGACTGGGGGATCCGTCTCCACTCTGACGATCCAGCGCACACAGCAGGAGGACTCGGCCGTGTATCTCTGTGCCAG
5 ACAGTGCACCAGACAAATCAGGGCTGCCCAGTGATCGCTTCTCTGCAGAGAGGACTGGGGaATCCGTCTCCACTCTGACGATCCAGCGCACACAGCAGGAGGACTCGGCCGTGTATCTCTGTG
6 AggaTGCAgCAGACAAATCgGGGCTtCCCcGTGATCGgTTCTCTGCAcAGAGGtCTGaGGgATCCaTCTCCACTCTGAaGtTCCAGCGCACACAGCAGGgGGACTtGGCtGTGTATCTCTGTGCCAG

          >__7mer_                                  ___9mer__
1 CAGCTTAGCCACAGCATGGCACAGTTGCCTCCTTCCTGTTCACAAACCTCA
2 CAGCTTAACCACAGCATGACACAATCGCCTCCTTCCTGCTCATAAACCTCCT
3 CAGCTTAGCCACAGCATGGCACAGTCGCCTCCTTCCTGCTCATAAACCTCAC
4 CAGC
5
6 CAGCTCAGCCACAGCACTACTGCTCCAGTGTCAGCTTGGTTCCCTAGGAAAT
```

Figure 2. Genomic sequences of Vβ6 genes. The clones are described in Fig. 1. UAS2 (24) and Vβ6.7b (20) were used for comparison. The former is a rearranged clone and the latter was obtained by PCR amplification from genomic DNA. Sequences corresponding to leader, intron, GT repeats, V region, and 3' hepatamer and nonamer (*shaded*) are indicated. Spaces were introduced in the introns to optimize alignment, and on either side of the intron for clarity. Capital letters indicate sequence identity with aligned residues above or below.

tions (18, 20). It remains possible that there are additional alleles of Vβ6.7. The Vβ6.7 gene contained in clone 5-2 has an identical coding region sequence to the prototypic UAS2 (Vβ6.7a) sequence, but differs in the intron (Fig. 2). Both sequences are also identical over ~200 bp of promoter sequence upstream of the ATG codon (see Fig. 8). Thus, the sole difference between the two sequences is the number of GT repeats found in the intron: $(GT)_{21}$ for UAS2 and $(GT)_{24}$ for the Vβ6.7a gene of clone 5-2. Previous analysis of $(GT)_n$ repeats has shown that these sequences can be very polymorphic (30, 31). Since the two alleles do not differ in peptide sequence, we have opted not to distinguish the new 5-2-derived sequence by a lower case letter and consider it another allele encoding the Vβ6.7a peptide sequence.

The genomic sequences shown in Figs. 2–4 demonstrate that the Vβ6, Vβ13, and Vβ5 genes have a similar structure. The first exon encodes NH_2-terminal leader peptide residues and the second exon encodes the rest of the leader and the V segment sequences. The introns reveal the most conspicuous differences between these three V gene subfamilies. Vβ6 subfamily members all contain $(GT)_n$ repeats in the intron, while these repeats are lacking in the Vβ13 and the Vβ5.5 genes. In each case, the $(GT)_n$ sequence appears at almost the same position in the intron, beginning at position 28-31

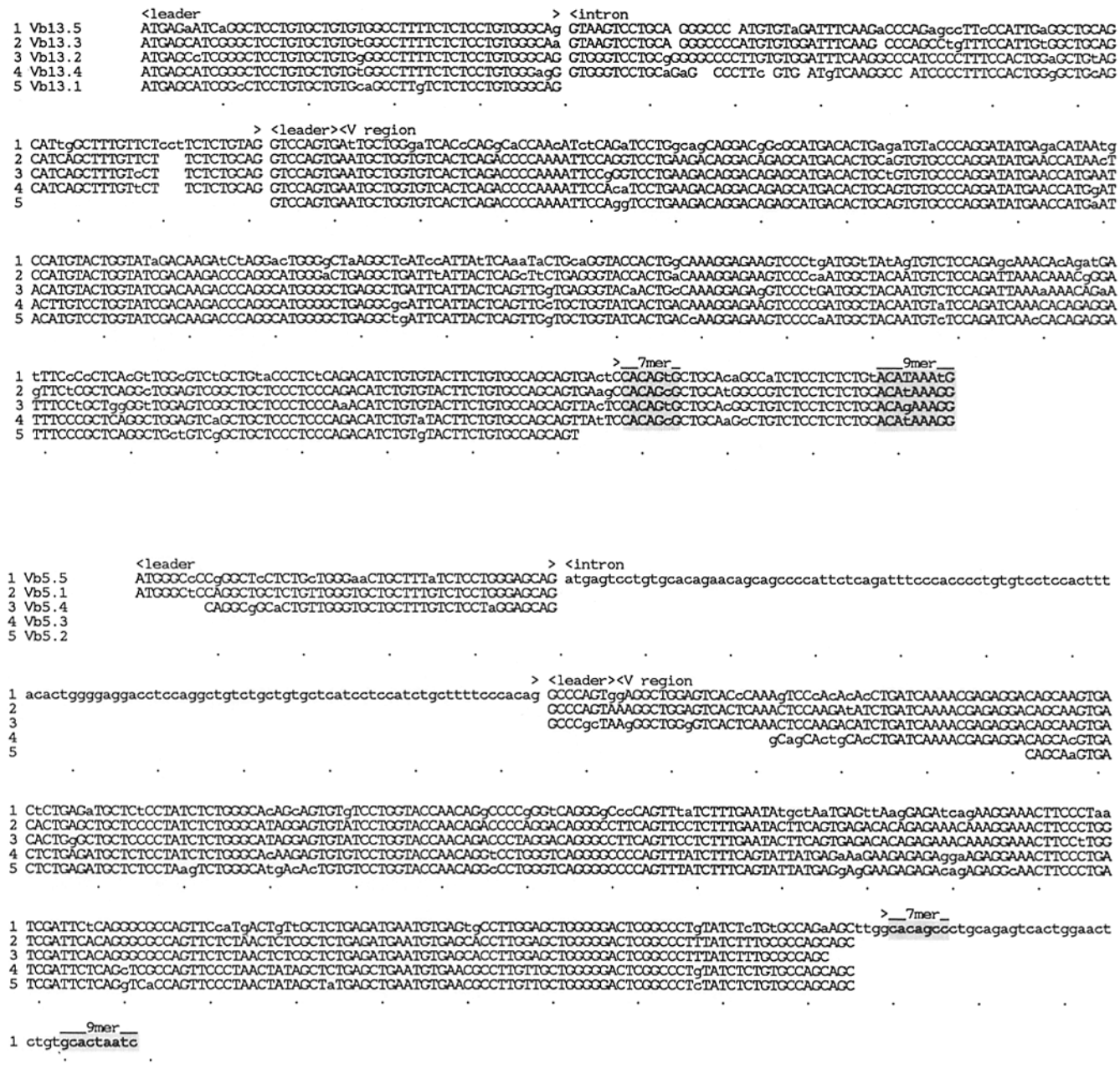

Figure 4. Genomic sequences of the Vβ5.5 gene. This sequence was obtained from clone 9 (Fig. 1) and is compared to cDNA sequences from the literature: Vβ5.1, clone HBP51 (2, 7); Vβ5.2, clone PL2.5 (6); Vβ5.3, clone 12A1 (58); Vβ5.4, clone Ph24 (4). See also legend to Fig. 2.

from the 5′ end on the intron (Fig. 2). In the three Vβ6.7 sequences shown in Fig. 2, the $(GT)_n$ repeat differs each time by a multiple of two nucleotides (n = 24 for Vβ6.7a from 5-2; n = 21 for Vβ6.7a from UAS2; n = 15 for Vβ6.7b). Similar simple repeats capable of assuming non-B secondary DNA configurations are observed in the introns of several human and murine TCR V gene segments (18). Such minisatellite repeat sequences may serve to promote recombination events such as homologous recombination and gene conversion. Alternatively, they may serve as enhancer elements and play a role in gene expression (32–35), but their exact function remains unknown.

The 3′ flanking sequences of Vβ6, Vβ13, and Vβ5.5 gene segments contain the conserved heptamer and nonamer signal sequences for recombinase-mediated VDJ recombination separated by 23 bp of nonconserved sequence (Figs. 2–4) (36). However, the nonamer sequence was absent in Vβ6.10 (Fig. 2), while the heptamer was found at the expected position. These data suggest that the Vβ6.10 gene may not be able to recombine into a functional β chain gene.

The Vβ5.5 gene (Fig. 4) represents the sole example of a genomic sequence from this Vβ subfamily. The intron begins with an AT rather than the more usual GT dinucleotide. However, the intron ends with the typical AG sequence.

All Vβ sequences from the genomic clones were compared for homology by considering exon sequences, intron sequences, and promoter sequences (see Fig. 8) separately. Homology matrices expressed in percentages are shown in Table 1. Overall, homology is highest in the coding region. Homologies were lower among the intron sequences and the promoter sequences. As expected, homologies were greatest within a Vβ subfamily. The new Vβ6.10 gene is most closely related to Vβ6.7a. The

Table 1. *Scoring Matrices for Exon, Intron, and Promoter Sequences*

		Vb6.1	Vb6.7a	Vbuas2	Vb6.10	Vb6.3	Vb5.5	Vb13.2	Vb13.3	Vb13.4	Vb13.5
		1	2	3	4	5	6	7	8	9	10
A. Exon	1		**89.3**	**89.5**	**85.6**	**81.0**	46.4	43.6	45.1	42.2	43.6
	2			**100.0**	**88.5**	**83.9**	46.1	45.6	47.1	44.2	45.3
	3				**88.5**	**83.9**	46.1	45.6	47.1	44.2	45.3
	4					**81.8**	44.9	44.2	45.1	41.6	43.3
	5						53.1	45.3	46.2	44.8	44.8
	6							52.0	52.5	49.9	49.1
	7								**91.0**	**89.8**	**79.4**
	8									**90.4**	**80.5**
	9										**80.8**
		1	2	3	4	5	6	7	8	9	10
B. Intron	1		**58.2**	**62.7**	**32.7**	**50.9**	31.8	35.2	30.2	26.5	32.6
	2			**96.1**	**42.9**	**47.4**	28.2	33.0	33.7	31.3	30.3
	3				**42.9**	**46.7**	31.5	29.5	30.2	28.9	32.6
	4					**29.2**	25.0	27.3	29.1	30.1	32.6
	5						36.3	30.7	30.3	34.9	32.6
	6							34.1	25.6	28.9	24.7
	7								**76.7**	**65.1**	**33.0**
	8									**69.9**	**69.8**
	9										**42.2**
		1	2	3	4	5	6	7	8	9	10
C. Promoter	1		**69.5**	**69.5**	**79.7**	**62.7**	25.8	31.4	33.9	30.1	26.7
	2			**100.0**	**64.0**	**61.0**	20.8	25.1	28.8	30.1	27.5
	3				**64.0**	**61.0**	20.8	25.1	28.8	30.1	27.5
	4					**57.6**	25.8	28.7	31.8	29.7	29.7
	5						25.0	24.6	19.5	26.7	24.6
	6							23.1	22.9	24.2	23.7
	7								**79.1**	**60.7**	**60.2**
	8									**59.5**	**58.9**
	9										**64.0**

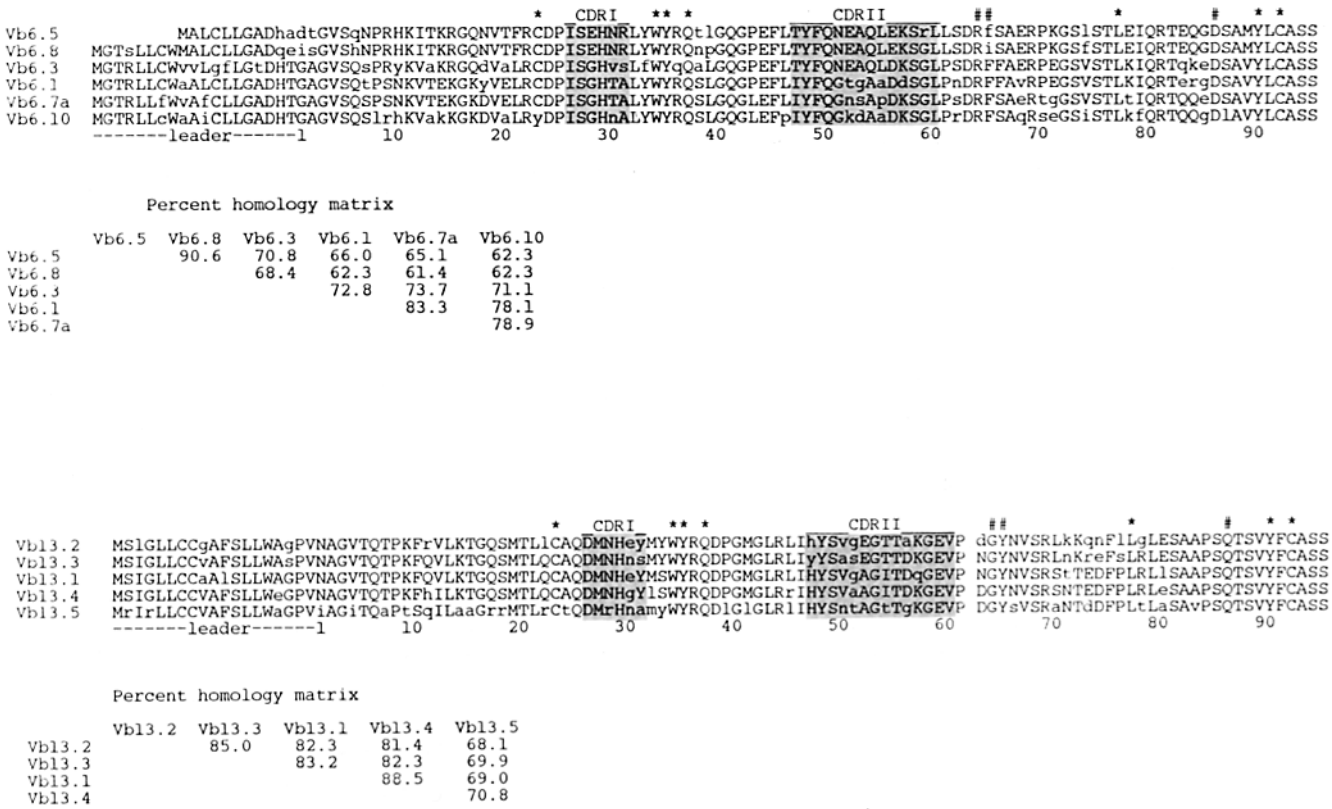

Figure 6. Vβ13 protein sequences. Sequences are derived from the genomic clones in Fig. 1 except for Vβ13.1 (22). See legend of Fig. 5 for symbols.

Vβ13 genes examined are closely related to one another, with the exception of Vβ13.5, which shows least homology with the other Vβ13 members (Table 1).

Translated Sequences of Vβ6, Vβ13, and Vβ5 Genes. In Figs. 5–7, the peptide sequences of the V genes encoded by our genomic clones are aligned with some reference sequences from the same Vβ subfamilies. The location of the CDR1 and CDR2 loops is based upon comparisons of V gene sequences with Ig V gene sequences and modeling of the TCR structure on the known Ig three-dimensional structure (37, 38). CDR1 and CDR2 are thought to interact with MHC antigen-presenting molecules (1). Vβ6 peptides are one amino acid longer than Vβ13 and Vβ5 peptides as described previously (37, 38). This is probably due to differences in the size of the CDR2 loop. Amino acid residues that are highly conserved and contribute to the Ig domain structure (37, 38) are indicated by an asterisk in Figs. 5–7, and residues that define Vβ subgroups (8, 37, 38) are identified by the pound sign. All these residues are conserved within the new Vβ sequences reported herein. There is one significant exception in the Vβ6.10 gene sequence where C^{23} is replaced by Y^{23}. Thus, the disulfide bond between C^{23} and C^{92}, which is essential for an Ig domain structure, cannot be formed by the Vβ6.10 gene product.

Peptide homology matrices (Figs. 5–7) indicate that within the Vβ6 subfamily, Vβ6.1, Vβ6.7a, and Vβ6.10 are closely related to one another. Vβ6.5 and Vβ6.8 represent a different example of closely related Vβ6 genes. Vβ6.3 is least homologous with the tested sequences. Within the Vβ13 subfamily, Vβ13.1 and Vβ13.4 are most homologous, while Vβ13.2

Figure 7. Vβ5 protein sequences. The Vβ5.5 sequence is from clone 9 (Fig. 1). Other sequences are as described in Fig. 4. See legend of Fig. 5 for symbols.

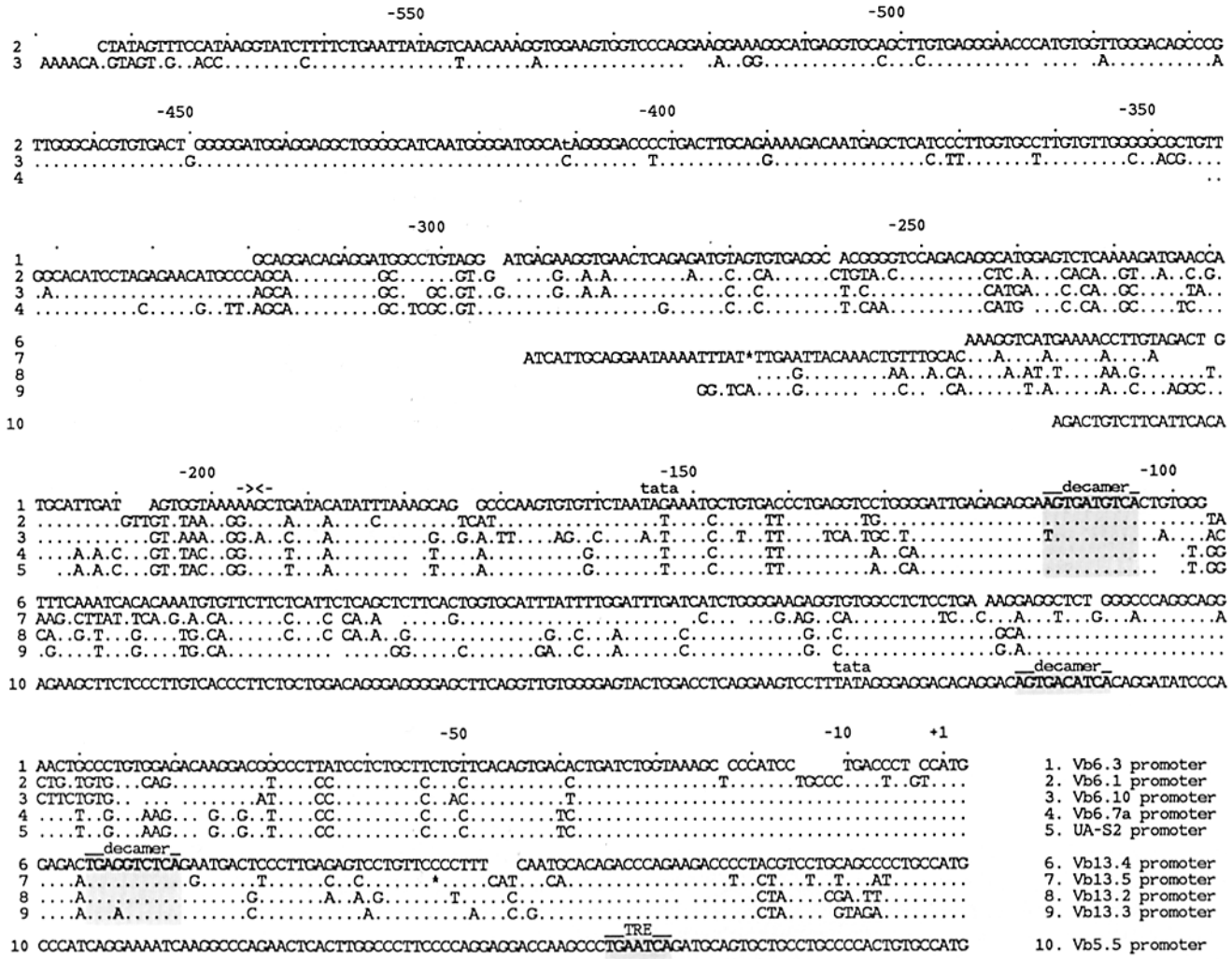

Figure 8. Vβ promoter sequences. Promoter areas were sequenced from the clones shown in Fig. 1. The promoter of the rearranged clone UAS2 (Vβ6.7a) was also sequenced. Sequences are arranged by V gene subfamily and dots indicate sequence identity with the top sequence of the respective subfamily. Within a V gene subfamily, the entire promoter region is conserved. The locations of the consensus decamer and a TRE motif in the Vβ5.5 promoter are indicated, as are possible locations of TATA sequences. The site of the transposon insertion within the Vβ6.10 promoter sequence (sequence no. 3) is indicated by two arrows. There are two variants of the Vβ13.5 promoter, which differ at marked residues (*). One is represented by genomic clones 3 and 9 and contains G at −268 and C at −53. The other is represented by clone 13 and contains A at −268 and G at −53. Thus, the genomic clones 3, 9, and 13 may represent different alleles.

and Vβ13.3 are more similar to one another. Vβ13.5 has least homology with the other Vβ13 sequences. Within the Vβ5 subfamily, Vβ5.2 and Vβ5.3 are closely related, as are Vβ5.1 and Vβ5.4. The new gene, Vβ5.5, is similarly related to both of these subgroups of Vβ genes (Fig. 8).

Promoter Sequences. Fig. 8 shows the 5′ flanking sequences for the Vβ genes described in Fig. 1. Approximately 200 bp of 5′ flanking sequence obtained in our laboratory from the clone UAS2 (Vβ6.7a) are also included for comparison. Promoter sequences were analyzed with the SITES program described by Ghosh (28). A previously described TCR decamer motif and a TRE-like motif were observed (Fig. 8) with identical match probabilities ranging between 1.74 and 7.04 × 10^{-3}. The conserved decanucleotide consensus sequence, 5′ AGTGATGTCA 3′, was found at variable positions (−80 to −106) in reference to the ATG codon. A similar decamer was observed in the 5′ flanking region of the human Vβ8.1 and several murine Vβ genes (12). This decamer sequence displays a degree of dyad symmetry reminiscent of the common features of other regulatory factor binding sites. Usually these begin with TGA and end with TCA. The numbers of nucleotides inserted between TGA and TCA are variable. For example, TGACTCA has been identified as a TRE-like motif, and may represent the binding site for Ap-1 (39). Another example is TGACGTCA, a palindromic sequence that has been associated with the cAMP response element (40, 41). Both motifs can be recognized by members of the fos/jun family (42). The decamers in Fig. 8 show a two-nucleotide insertion mode, NNTGATGTCA in the Vβ6 gene subfamily, and NNTGACATCA in the Vβ5.5 gene segment. An addi-

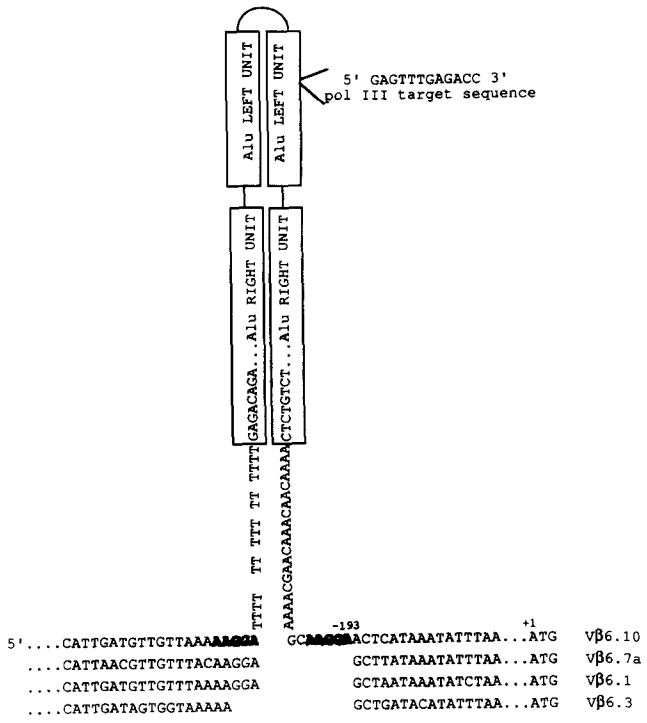

Figure 9. Transposon in the Vβ6.10 promoter. The flanking direct repeats are shaded. The 3' Alu sequence represents a sense sequence, while the 5' Alu element is a reverse complement sequence. Both are homologous to consensus Alu sequences.

tional TRE-like motif is present in the Vβ5.5 promoter with a single nucleotide insertion sequence, TGAATCA. In contrast, a novel four-nucleotide insertion decamer sequence, TGAAGTCTCA or TGAGGTCTCA, was found in all Vβ13 genes. Extensive homology surrounding these consensus decamer sequences was observed. Each decamer is preceded by a purine-rich sequence 10–20 bp long. DNAse footprint analysis in the promoter region of human Vβ8.1 (11) indicated that one of the protected areas contains the decamer motif, suggesting that this conserved motif represents a protein binding site.

A typical transposon-like insertion element was found in the 5' flanking region of the Vβ6.10 gene segment (Fig. 9). This is a 549-bp insertion at position −193 (Figs. 8 and 9). Homology between Vβ6.10 and Vβ6.1 is maintained (at 84% over 200 bp) both upstream and downstream of this insertion. The features of this insertion element include the following. A 5-bp direct repeat (AAGGA) flanks the insertion on either side. The entire insertion is characterized by two imperfect inverted repeats. The stem origin of the insertion contains a 5' run of poly(T)$_{15}$ and a 3' run of poly(A)$_{15}$ interrupted by some nucleotides (Fig. 9). Finally, the insertion element contains two Alu sequences head to tail. Although unusual, similar head-to-tail Alu sequences have previously been described (43). The 3' Alu sequence contains a perfect polIII target sequence, 5' GAGTTTGAGACC 3', in its shorter (L-handed) monomer unit (Fig. 9). Thus, this insertion element contains all the characteristics of Alu sequences (44, 45).

Discussion

Six genomic clones are described, each containing a Vβ6 gene segment. Several findings were made with these clones. (*a*) Several new human Vβ genes were discovered, including Vβ6.10, Vβ13.4, Vβ13.5, and Vβ5.5. (*b*) Members of the Vβ6, Vβ13, and Vβ5 subfamilies cluster together in the Vβ locus. These three subfamilies probably evolved by repeated duplications of a cassette containing Vβ13-Vβ6-Vβ5 genes. (*c*) The promoter regions of all the genes belonging to these three subfamilies contain conserved motifs, such as the decamer motif, probably representing binding sites for *trans*-acting regulators. The promoter regions were conserved within subfamilies but less so between subfamilies. (*d*) The Vβ6.10 gene is most likely a pseudogene, since it lacks a conserved recombinase signal sequence, it contains a mutation of the highly conserved Cys[23] necessary for formation of an Ig-like domain, and it contains a 549-bp insertion of Alu sequences in its promoter, which may interfere with promoter activity.

In a recent reevaluation of the total number of Vβ gene segments (8), minimal estimates were made of five Vβ5 genes, eight Vβ6 genes, and five Vβ13 genes. These estimates are based on counting bands on Southern blots, which is unreliable because the results depend strongly on the probe used. These estimates were also based on counts of genes on available cosmid clones (29), but they remain minimal estimates, since the entire Vβ locus has not been characterized. With this caveat in mind, it seems never the less likely that the newly described Vβ5.5 gene represents the last member of the Vβ5 subfamily. The Vβ5.5 peptide sequence is only 70.7–76.1% homologous to other Vβ5 sequences (Fig. 7). It is thus unlikely that it represents an allelic form of a previously described Vβ5 sequence, in particular since these have each been cloned from different sources and the sequences are usually identical for each Vβ5 gene (2,3).

The same argument applies to the Vβ13.3, Vβ13.4, and Vβ13.5 genes (Fig. 6). Moreover, these gene products, which are 68.1–88.5% homologous to the previously known Vβ13.1 and Vβ13.2 gene products, are clearly derived from distinct genes at different loci (Fig. 1). Thus, the Vβ13 subfamily now contains five known members as previously predicted (8). A recent publication describes a PCR-derived Vβ13 cDNA clone called IGRb14, which is 100% homologous to our Vβ13.3 coding sequence (9). It is still unclear whether other new sequences described in this paper represent new loci or allelic variants of known Vβ genes.

Vβ6 is the largest human Vβ subfamily. There are >10 previously published sequences, of which two are definitely allelic forms of the same gene and are named Vβ6.7a and Vβ6.7b (20). The Vβ6.6 sequence differs only by a single silent nucleotide change in the coding region from the Vβ6.7a sequence and probably represents another allele, although this has not been formally proven. Also, shown herein is the sequence of Vβ6.7a derived from the clone UAS2. This sequence differs from the 5-2-derived Vβ6.7a nucleotide sequence in the intron, where the former contains a (GT)$_{21}$ repeat and the latter a (GT)$_{24}$ repeat. All coding sequences and promoter sequences are alike. Thus, we chose to refer to both clones as encoding Vβ6.7a genes. These data imply

that V genes with $(GT)_n$ repeats (or similar repeats) in their introns will be characterized by numerous allelic forms (18, 30, 31). The majority of these alleles will be distinguished only by the number of dinucleotide repeat units. Occasionally, other mutations will be associated with a given allele, and a fraction of these will result in different peptides, as is the case for Vβ6.7a and Vβ6.7b (20), but not for Vβ6.6. The availability of genomic clones of Vβ6.7, such as clone 5-2, will facilitate a more extensive analysis of the allelic polymorphisms associated with this gene. It will be of interest to compare this polymorphism with that of the adjacent upstream gene, Vβ13.2, since Vβ6 introns all contain $(GT)_n$ repeats and Vβ13 introns do not.

The new Vβ6 gene, Vβ6.10, has greatest homology with Vβ6.7a and Vβ6.1 (Table 1). Yet, all three sequences represent separate genes (Fig. 1). Vβ6.10 is probably a pseudogene. The conserved nonamer in the 3′ UT region, which is considered a recombinase signal sequence, is missing. Since Vβ6.10 has not been isolated from cDNA libraries, it is probable that this V gene cannot undergo V-D-J recombination. Moreover, the highly conserved Cys^{23} residue, which is necessary for formation of an Ig-like domain and forms a disulfide bond with Cys^{92}, is changed to a Tyr^{23} in Vβ6.10. Thus, it is questionable whether a Vβ6.10-encoded β chain could fold into an Ig-like domain. Lastly, the Vβ6.10 promoter contains a 549-bp insertion element at position −193 (Fig. 9) consisting of two head-to-tail Alu sequences. Such sequences are classified as nonviral retroposons and are thought to derive from processed 7SL RNA (44–46). These elements are transcribed by RNA polymerase III initiating at the 5′ end of the left unit and transcribing through the 3′ terminal poly(A) tract. Reverse transcription is then primed on the poly(A) tract, providing a mechanism for retrotransposition (44, 45).

Little is known about the significance of mobile genetic elements in humans, and Alu sequences in particular, which frequently insert in introns or flanking sequences of genes, but generally not in exons. In certain cases, retrotransposons may have acted as insertional mutagens, resulting in inactivation (47) or activation (48) of a cellular gene. In fact, retrotransposition has been directly demonstrated in human lung carcinoma cells transfected with a target gene. These cells were screened for insertion mutations, resulting in inactivation of the gene. A newly transposed Alu sequence was identified by these means (49). In other instances, retrotransposons have inactivated genes relevant to human disease, such as the factor VIII gene in hemophilia A (50) or the c-*myc* gene in breast carcinoma cells (51). Retrotransposons have been associated with a high percentage of murine actin pseudogenes (52) and with homologous recombination in humans (53). The Alu sequences inserted in the Vβ6.10 promoter may have resulted in inactivation of this gene by interference with promoter function. The resultant Vβ6.10 pseudogene would then have undergone further mutation for lack of selective evolutionary pressures. It remains possible that the inserted Alu sequences upstream of Vβ6.10 are not present in all haplotypes and that Vβ6.10 may be a functional gene in some individuals.

The significance of a genomic cassette containing members of the Vβ13-Vβ6-Vβ5 subfamilies remains hypothetical. Clearly, these three Vβ subfamilies represent the largest Vβ subfamilies and together account for at least 18 Vβ genes, i.e., ~32% of the Vβ repertoire assuming a total of ~57 Vβ genes (8). One may estimate the usage of certain Vβ genes with mAbs specific for Vβ gene products. Thus, the mean frequencies of Vβ usage are 3.2% for Vβ5.1/5.4, 2.8% for Vβ5.2/5.3, and 4.5% for Vβ6.7 (54–56, and our unpublished results). Moreover, Vβ13 gene usage has been estimated by quantitative PCR at 9.0% for Vβ13.1, 2.7% for Vβ13.2, and 7.8% for Vβ6.1-3. Thus, a minimal estimate of expression of Vβ13, Vβ6, and Vβ5 genes is 30%, and these data are incomplete because of the absence of data for several genes in these three subfamilies. Therefore, we would propose that the Vβ genes encoded by the three Vβ subfamilies represent particularly useful Vβ genes to humans, and probably arose through multiple gene duplications of the basic cassette structure.

Lastly, the promoter sequences shown herein demonstrate strong conservation within a Vβ subfamily, but not between subfamilies (Table 1 *C*). Moreover, the consensus decamer discussed above demonstrates some discrete differences in location and sequence between Vβ subfamilies. Thus, one could postulate that control of transcription may differ in T cells depending on what Vβ subfamily is used. This could have relevance to observed differences of Vβ usage at different developmental stages of T cells (57).

This work was supported in part by National Institutes of Health grants AI-31140, CA-42046, AG-00541 and P60-AR-38520. D. N. Posnett is a recipient of the Cancer Research Institute Investigator Award, and an award from the Crohn's & Colitis Foundation of America, Inc.

Address correspondence to David N. Posnett, Department of Medicine, Box 56, Cornell University Medical College, 1300 York Avenue, New York, NY 10021.

Received for publication 13 June 1991 and in revised form 16 September 1991.

Note added in proof: The new sequences described herein can be retrieved from EMBL with the following accession numbers: X61439 (TCR Vβ5.5); X61440 (TCR Vβ6.1); X61441 (TCR Vβ6.3); X61442 (TCR Vβ6.7a); X61443 (TCR Vβ6.7b); X61444 (TCR Vβ6.10); X61445 (TCR Vβ13.2); X61446 (TCR Vβ13.3); X61447 (TCR Vβ13.4); X61653 (TCR Vβ13.5).

References

1. Davis, M.M., and P.J. Bjorkman. 1988. T-cell antigen receptor genes and T-cell recognition. *Nature (Lond.).* 334:395.
2. Toyonaga, B., and T.W. Mak. 1987. Genes of the T cell antigen receptor in normal and malignant T cells. *Annu. Rev. Immunol.* 5:585.
3. Wilson, R.K., E. Lai, P. Concannon, R.K. Barth, and L.E. Hood. 1988. Structure, organization and polymorphism of murine and human T cell receptor alpha and beta chain gene families. *Immunol. Rev.* 101:149.
4. Tillinghast, J.P., M.A. Behlke and D.Y. Loh. 1986. Structure and diversity of the human T cell receptor β chain variable region genes. *Science (Wash. DC).* 233:879.
5. Kimura, N., B. Toyonaga, Y. Yoshikai, R-P. Du, and T.W. Mak. 1987. Sequences and repertoire of the human T cell receptor α and β chain variable region genes in thymocytes. *Eur. J. Immunol.* 17:375.
6. Concannon, P., L.A. Pickering, P. Kung, and L. Hood. 1986. Diversity and structure of human T cell receptor β chain variable region genes. *Proc. Natl. Acad. Sci. USA.* 83:6598.
7. Kimura, N., B. Yoyonaga, Y. Yoshikai, M.D. Minden, and T.W. Mak. 1986. Sequence and diversity of human T cell receptor β chain genes. *J. Exp. Med.* 164:739.
8. Robinson, M.A. 1991. The human T cell receptor β chain gene complex contains at least 57 variable gene segments: identification of six Vβ genes in four new gene families. *J. Immunol.* 146:4392.
9. Ferradini, L., S. Roman-Roman, J. Azocar, H. Michalaki, F. Triebel, and T. Hercend. 1991. Studies on the human T cell receptor α/β variable region genes. II. Identification of four additional Vβ subfamilies. *Eur. J. Immunol.* 21:935.
10. Siu, G., E.C. Strauss, E. Lai, and L.E. Hood. 1986. Analysis of a human Vβ gene subfamily. *J. Exp. Med.* 164:1600.
11. Royer, H.D., and E.L. Reinherz. 1987. Multiple nuclear proteins bind upstream sequences in the promoter region of a T cell receptor β chain variable region gene: evidence for tissue specificity. *Proc. Natl. Acad. Sci. USA.* 84:232.
12. Anderson, S.J., H.S. Chou, and D.Y. Loh. 1988. A conserved sequence in the T-cell receptor β-chain promoter region. *Proc. Natl. Acad. Sci. USA.* 85:3551.
13. Winoto, A., and D. Baltimore. 1989. A novel inducible and T cell specific enhancer located at the 3' end of the T cell receptor α locus. *EMBO (Eur. Mol. Biol. Organ.) J.* 8:729.
14. Krimpenfort, P., R. deJong, Y. Uematsu, Z. Dembic, S. Ryser, H. von Boehmer, M. Steinmetz, and A. Berns. 1988. Transcription of T cell receptor β chain genes is controlled by a downstream regulatory element. *EMBO (Eur. Mol. Biol. Organ.) J.* 7:745.
15. Ho, I-C., L-H. Yang, G. Morle, and J.M. Leiden. 1989. A T cell specific transcriptional enhancer element 3' of Cα in the human T cell receptor α locus. *Proc. Natl. Acad. Sci. USA.* 86:6714.
16. Gottschalk, L.R., and J.M. Leiden. 1990. Identification and functional characterization of the human T cell receptor β gene transcriptional enhancer: common nuclear proteins interact with the transcriptional regulatory elements of the T cell receptor α and β genes. *Mol. Cell. Biol.* 10:5486.
17. Redondo, J.M., S. Hata, C. Brocklehurst, and M.S. Krangel. 1990. A T cell specific transcriptional enhancer within the human T cell receptor α locus. *Science (Wash. DC).* 247:1225.
18. Posnett, D.N. 1990. Allelic variations of human TCR V gene products. *Immunol. Today.* 11:368.
19. Robinson, M.A. 1989. Allelic sequence variations in the hypervariable region of a TCR β chain: correlation with restriction length polymorphism in human families and populations. *Proc. Natl. Acad. Sci. USA.* 86:9422.
20. Li, Y., P. Szabo, M.A. Robinson, B. Dong, and D.N. Posnett. 1990. Allelic variations in the human T cell receptor Vβ6.7 gene products. *J. Exp. Med.* 171:221.
21. Pullen, A.M., W. Potts, E.K. Wakeland, J. Kappler, and P.M. Marrack. 1990. Surprisingly uneven distribution of the T cell receptor Vβ repertoire in wild mice. *J. Exp. Med.* 171:49.
22. Duby, A.D., and J.G. Seidman. 1986. Abnormal recombination products result from aberrant DNA rearrangement of the human T cell antigen receptor beta-chain gene. *Proc. Natl. Acad. Sci. USA.* 83:4890.
23. Leiden, J.M., D.P. Dialynas, A.D. Duby, C. Murre, J. Seidman, and J.L. Strominger. 1986. Rearrangement and expression of T cell antigen receptor genes in human T lymphocyte tumor lines and normal human T cell clones: evidence for allelic exclusion of Tiβ gene expression and preferential use of a Jβ2 gene segment. *Mol. Cell. Biol.* 6:3207.
24. Hinkkanen, A.E., V. Steimle, M. Schlesier, H.H. Peter, and J.T. Epplen. 1989. The antigen receptor of an autoreactive T-cell clone from human rheumatic synovia. *Immunogenetics.* 29:131.
25. Sambrook, J., E.F. Fritsch, and T. Maniatis. 1989. Molecular Cloning: A Laboratory Manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.
26. Sanger, F., S. Nicklen, and A.R. Coulson. 1977. DNA sequencing with chain-terminating inhibitors. *Proc. Natl. Acad. Sci. USA.* 74:5463.
27. Sobel, E., and H.M. Martinez. 1985. A multiple sequence alignment program. *Nucleic Acids Res.* 14:363.
28. Ghosh, D. 1990. A relational database of transcription factors. *Nucleic Acids Res.* 18:1749.
29. Lai, E., P. Concannon, and L. Hood. 1988. Conserved organization of the human and murine T cell receptor β gene families. *Nature (Lond.).* 331:543.
30. Weber, J.L., and P.E. May. 1989. Abundant class of human DNA polymorphisms which can be typed using the polymerase chain reactions. *Am. J. Hum. Genet.* 44:388.
31. Litt, M., and J.A. Dufy. 1989. A hypervariable microsatellite revealed by in vitro amplification of a dinucleotide repeat within the cardiac muscle actin gene. *Am. J. Hum. Genet.* 44:397.
32. Nordheim, A., and A. Rich. 1983. The sequence (dC-dA)n.(dG-dT)n forms left-handed Z-DNA in negatively supercoiled plasmids. *Proc. Natl. Acad. Sci. USA.* 80:1821.
33. Hamada, H., M.G. Petrino, T. Kakunaga, M. Seidman, and B.D. Stollar. 1984. Characterization of genomic Poly(dT-dG).Poly(dC-dA) sequences: structure, organization, and conformation. *Mol. Cell. Biol.* 4:2610.
34. Hamada, H., M. Seidman, B.H. Howard, and C.M. Gorman. 1984. Enhanced gene expression by the Poly(dT-dG).Poly(dC-dA) sequence. *Mol. Cell. Biol.* 4:2622.
35. Nordheim, A., and A. Rich. 1983. Negatively supercoiled simian virus 40 DNA contains Z-DNA segments within transcriptional enhancer sequences. *Nature (Lond.).* 303:674.
36. Sakano, H., K. Huppi, G. Heinrich, and S. Tonegawa. 1979. Sequences at the somatic recombination sites of immunoglobulin light-chain genes. *Nature (Lond.).* 280:288.
37. Schiffer, M., T.T. Wu, and E.A. Kabat. 1986. Subgroups of variable region genes of β chains of T-cell receptors for antigen. *Proc. Natl. Acad. Sci. USA.* 83:4461.
38. Chothia, C., D.R. Boswell, and A.M. Lesk. 1988. The outline structure of the T cell α/β receptor. *EMBO (Eur. Mol.*

Biol. Organ.) J. 7:3745.
39. Lee, W., P. Mitchell, and R. Tjian. 1987. Purified transcription factor AP-1 interacts with TPA-inducible enhancer elements. *Cell.* 49:741.
40. Montminy, M.R., and L.M. Bilezikjian. 1987. Binding of a nuclear protein to the cyclic-AMP response element of the somatostatin gene. *Nature (Lond.).* 328:175.
41. Montminy, M.R., K.A. Sevarino, J.A. Wagner, G. Mandel, and R.H. Goodman. 1986. Identification of a cyclic-AMP-responsive element within the rat somatostatin gene. *Proc. Natl. Acad. Sci. USA.* 83:6682.
42. Vogt, P.K., and T.J. Bos. 1989. The oncogene jun and nuclear signalling. *Trends Biochem. Sci.*, 14:172.
43. Lee, M.G.-S., C. Loomis, and N.J. Cowan. 1984. Sequence of an expressed human β-tubulin gene containing ten Alu family members. *Nucleic Acids Res.* 12:5823.
44. Weiner, A.M., P.L. Deininger, and A. Efstratiadis. 1986. Nonviral retroposons: genes, pseudogenes, and transposable elements generated by the reverse flow of genetic information. *Annu. Rev. Biochem.* 55:631.
45. Jelinek, W.R., and C.W. Schmid. 1982. Repetitive sequences in eukaryotic DNA and their expression. *Annu. Rev. Biochem.* 51:813.
46. Ullu, E., and C. Tschudi. 1984. Alu sequences are processed 7SL RNA genes. *Nature (Lond.).* 312:171.
47. Canaani, E., O. Dreazen, A. Klar, G. Rechavi, D. Ram, J.B. Cohen, and D. Givol. 1983. Activation of the c-mos oncogene in a mouse plasmacytoma by insertion of an endogenous intracisternal A-particle genome. *Proc. Natl. Acad. Sci. USA.* 80:7118.
48. Hawley, R.G., M.J. Shulman, H. Murialdo, D.M. Gibson, and N. Hozumi. 1982. Mutant immunoglobulin genes have repetitive DNA elements inserted into their intervening sequences. *Proc. Natl. Acad. Sci. USA.* 79:7425.
49. Lin, C.S., D.A. Goldthwait, and D. Samols. 1988. Identification of Alu transposition in human lung carcinoma cells. *Cell.* 54:153.
50. Kazazian, H.H., Jr., C. Wong, H. Youssoufian, A.F. Scott, D.G. Phillips, and S.E. Antonarakis. 1988. Haemophilia A resulting from de novo insertion of L1 sequences represents a novel mechanism for mutation in man. *Nature (Lond.).* 332:164.
51. Morse, B., P.G. Rotherg, V.J. South, J.M. Spandorfer, and S.M. Astrin. 1988. Insertional mutagenesis of the myc locus by a LINE-1 sequence in a human breast carcinoma. *Nature (Lond.).* 333:87.
52. Min Man, Y., H. Delius, and D.P. Leader. 1987. Molecular analysis of elements inserted into mouse τ-actin processed pseudogenes. *Nucleic Acids Res.* 15:3291.
53. Mager, D.L., and N.L. Goodchild. 1989. Homologous recombination between the LTRs of a human retrovirus-like element causes a 5-kb deletion in two siblings. *Am. J. Hum. Genet.* 45:848.
54. Posnett, D.N., C.Y. Wang, and S. Friedman. 1986. Inherited polymorphism of the human T cell antigen receptor detected by a monoclonal antibody. *Proc. Natl. Acad. Sci. USA.* 83:7888.
55. Posnett, D.N., A. Gottlieb, J.B. Bussel, S.M. Friedman, N. Chiorazzi, Y. Li, P. Szabo, N.R. Farid, and M.A. Robinson. 1988. T cell antigen receptors in autoimmunity. *J. Immunol.* 141:1963.
56. Maecker, H.T., and R. Levy. 1989. Prevalence of antigen receptor variants in human T cell lines and tumors. *J. Immunol.* 142:1395.
57. Fowlkes, B.J., A.M. Kruisbeek, H. Ton-That, M.A. Weston, J.E. Coligan, R.H. Schwartz, and D.M. Pardoll. 1987. A novel population of T cells receptor $\alpha\beta$-bearing thymocytes which predominantly express a single Vβ gene family. *Nature (Lond.).* 329:251.
58. Leiden, J.M., and J. L. Strominger. 1986. Generation of diversity of the beta chain of the human T-lymphocyte receptor for antigen. *Proc. Natl. Acad. Sci. USA.* 83:4456.