Proc. Natl. Acad. Sci. USA
Vol. 83, pp. 2619–2623, April 1986
Immunology

Cloning and sequence analysis of complementary DNA encoding an aberrantly rearranged human T-cell γ chain

(cDNA/DNA sequence)

DENO P. DIALYNAS*, CORNELIS MURRE†, THOMAS QUERTERMOUS†, JEREMY M. BOSS*, JEFFREY M. LEIDEN*, J. G. SEIDMAN†, AND JACK L. STROMINGER*

*Department of Biochemistry and Molecular Biology, Harvard University, 7 Divinity Avenue, Cambridge, MA 02138; and †Department of Genetics, Harvard Medical School, 45 Shattuck Street, Boston, MA 02115

Contributed by Jack L. Strominger, December 16, 1985

ABSTRACT **Complementary DNA (cDNA) encoding a human T-cell γ chain has been cloned and sequenced. At the junction of the variable and joining regions, there is an apparent deletion of two nucleotides in the human cDNA sequence relative to the murine γ-chain cDNA sequence, resulting simultaneously in the generation of an in-frame stop codon and in a translational frameshift. For this reason, the sequence presented here encodes an aberrantly rearranged human T-cell γ chain. There are several surprising differences between the deduced human and murine γ-chain amino acid sequences. These include poor homology in the variable region, poor homology in a discrete segment of the constant region precisely bounded by the expected junctions of exon CII, and the presence in the human sequence of five potential sites for N-linked glycosylation.**

The murine T-cell γ-chain transcription unit is constructed through T-cell-specific somatic juxtaposition of germ-line variable (V), joining (J), and constant (C) region elements (1). Both in this context and at the level of amino acid sequence homology, the γ chain is similar to the α and β chains of the antigen receptor on mature T cells, more generally referred to as the T-cell receptor. The γ-chain locus is distinguished in part by the apparent paucity of V- and J-region elements; in fact, the greatest apparent variability resides at the V–J junction, generated by unknown mechanisms (2).

The function of the γ chain, which has yet to be identified at the protein level, is unknown. To determine which properties of the γ chain are essential to its function, and thereby gain insight into that function, we have initiated a characterization of the human γ-chain locus. We report here the isolation and sequence analysis of complementary DNA (cDNA) clones encoding an aberrantly rearranged human T-cell γ chain.

MATERIALS AND METHODS

Enzymes. Restriction enzymes were from New England Biolabs, Bethesda Research Laboratories, and Boehringer Mannheim. Avian myeloblastosis virus reverse transcriptase was from Life Sciences. T4 DNA ligase, *Eco*RI methylase, and *Eco*RI linkers were from New England Biolabs. The Klenow fragment of DNA polymerase I was from Bethesda Research Laboratories. Nuclease S1 was from Boehringer Mannheim.

RNA Blot Analysis. Poly(A)$^+$ RNA was denatured by formaldehyde, electrophoresed, and transferred onto a nitrocellulose filter (3). Hybridization was carried out in 50% formamide/6× SSC (1× SSC = 0.15 M NaCl/0.015 M Na citrate) at 42°C for 16 hr, using a nick-translated *Pst* I/*Bam*HI fragment encoding part of the human T-cell γ-chain CI exon (4) as a probe. The filters were then washed twice at room temperature with 2× SSC, followed by two washes at 53°C with 0.2× SSC.

Construction of the HPB-MLT cDNA Library. Total RNA was isolated from logarithmic-phase HPB-MLT cells by the lithium chloride/urea method (5). Poly(A)$^+$ RNA was prepared by one cycle of oligo(dT)-cellulose affinity chromatography. First-strand cDNA was synthesized with avian myeloblastosis virus reverse transcriptase, using as template 30 μg of HPB-MLT poly(A)$^+$ RNA that had first been denatured by heating at 70°C for 3 min. Second-strand cDNA synthesis using the Klenow fragment of DNA polymerase I followed denaturation of the first-strand reaction product with 12.5 mM CH_3HgOH. The cDNA hairpin loop was then cleaved with nuclease S1, after which the cDNA was methylated with *Eco*RI methylase, blunt-ended using the Klenow fragment of DNA polymerase I, and ligated to *Eco*RI linkers. The cDNA was then digested with *Eco*RI, chromatographed on Sepharose CL-4B both to remove free linkers and to size-fractionate the cDNA, and finally ligated to *Eco*RI-cut λgt10 vector (6). The product of the ligation reaction was packaged, giving 3.3×10^6 independent recombinant phage, as determined on *Escherichia coli* C600, with an average insert size of 1100 base pairs. This library, designated MLT:λgt10A, was carried through one cycle of amplification on *E. coli* C600Hfl$^+$ such that 85% of the resultant phage contain inserts.

Screening the HPB-MLT cDNA Library for the Human T-Cell γ Chain. An aliquot of cDNA library MLT:λgt10A was plated on *E. coli* C600Hfl$^+$ at a density of 3.5×10^4 plaque-forming units per 150-mm plate. Duplicate filters were prepared from 10 plates and screened. Hybridization was carried out in 50% formamide/6× SSC at 42°C for 16 hr, using a nick-translated *Pst* I/*Bam*HI fragment encoding part of the human T-cell γ-chain CI exon (4) as probe. The filters were then washed twice at room temperature with 2× SSC, followed by two washes at 53°C with 0.2× SSC. Positive recombinant phage were plaque-purified. Two clones, Tγ-1 and Tγ-2, have been characterized extensively.

DNA Sequence Analysis. The inserts from Tγ-1 and Tγ-2 were subcloned into the *Eco*RI site of pUC9 (7) to give pTγ-1 and pTγ-2, respectively. The insert in pTγ-1 was sequenced by the chemical-degradation method (8). Fragments to be sequenced were labeled at the 3′ end using the Klenow fragment of DNA polymerase I (3). To sequence from the *Pst* I and *Acc* I sites, *Eco*RI/*Pst* I and *Eco*RI/*Acc* I fragments from pTγ-1 were subcloned into pUC13 (7) and then sequenced from the *Hin*dIII site in the pUC13 polylinker. The internal *Bam*HI fragment of pTγ-1 was subcloned into pUC13

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

Abbreviations: V, J, and C regions, variable, joining, and constant regions.

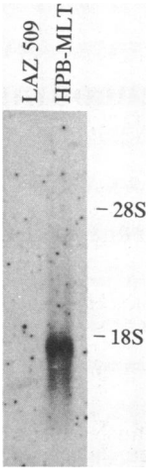

FIG. 1. RNA blot analysis of formaldehyde-denatured poly-(A)$^+$ RNA from B-cell line LAZ 509 and T-cell line HPB-MLT for expression of the human T-cell γ-chain transcript.

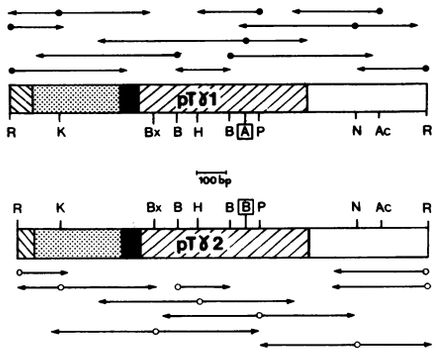

FIG. 2. Both the restriction map and sequencing strategy are given here for the inserts in pTγ-1 and pTγ-2. Restriction sites indicated are *Eco*RI (R), *Kpn* I (K), *Bst*XI (Bx), *Bam*HI (B), *Hin*cII (H), *Ava* II (A), *Pst* I (P), *Nco* I (N), and *Acc* I (Ac). The approximate junctions of the leader peptide ▨, V ▨, J ■, C ▨, and 3'-untranslated □ regions are indicated. The insert in pTγ-1 was sequenced by the chemical-degradation method (8) (●). The insert in pTγ-2 was sequenced by the dideoxy chain-termination method (9, 10) (○). bp, Base pairs.

and then sequenced both from the *Eco*RI and from the *Hin*dIII sites in the pUC13 polylinker. To sequence from the *Kpn* I site directly in pTγ-1, fragments were generated by using *Asp*718 (Boehringer Mannheim). The insert in pTγ-2 was sequenced by the dideoxy chain-termination method (9, 10). Restriction fragments as indicated were subcloned into either M13 mp8, M13 mp9, M13 mp18, or M13 mp19 (7). M13 mp8 and M13 mp9 were from Amersham, whereas M13 mp18 and M13 mp19 were from New England Biolabs. Sequencing was carried out using dATP [α-^{35}S] and a sequencing kit from Amersham.

RESULTS

RNA Blot Analysis of Human T-Cell Lines for Expression of the γ-Chain Transcript. To clone cDNA encoding the human T-cell γ chain, it was essential to identify a human cell line expressing the γ-chain transcript. To this end, a panel of human T-cell lines was screened by RNA blot analysis, using a *Pst* I/*Bam*HI fragment encoding part of the human γ-chain CI exon (4) as probe. By this criterion, the T-cell line HPB-MLT was found to express the γ-chain transcript or a transcript highly homologous to it, whereas the B-cell line LAZ 509 did not (Fig. 1). Other T-cell lines analyzed by RNA blot gave either no signal at all or an ambiguous signal (data not shown). The apparent size of the hybridizing RNA is 1.6

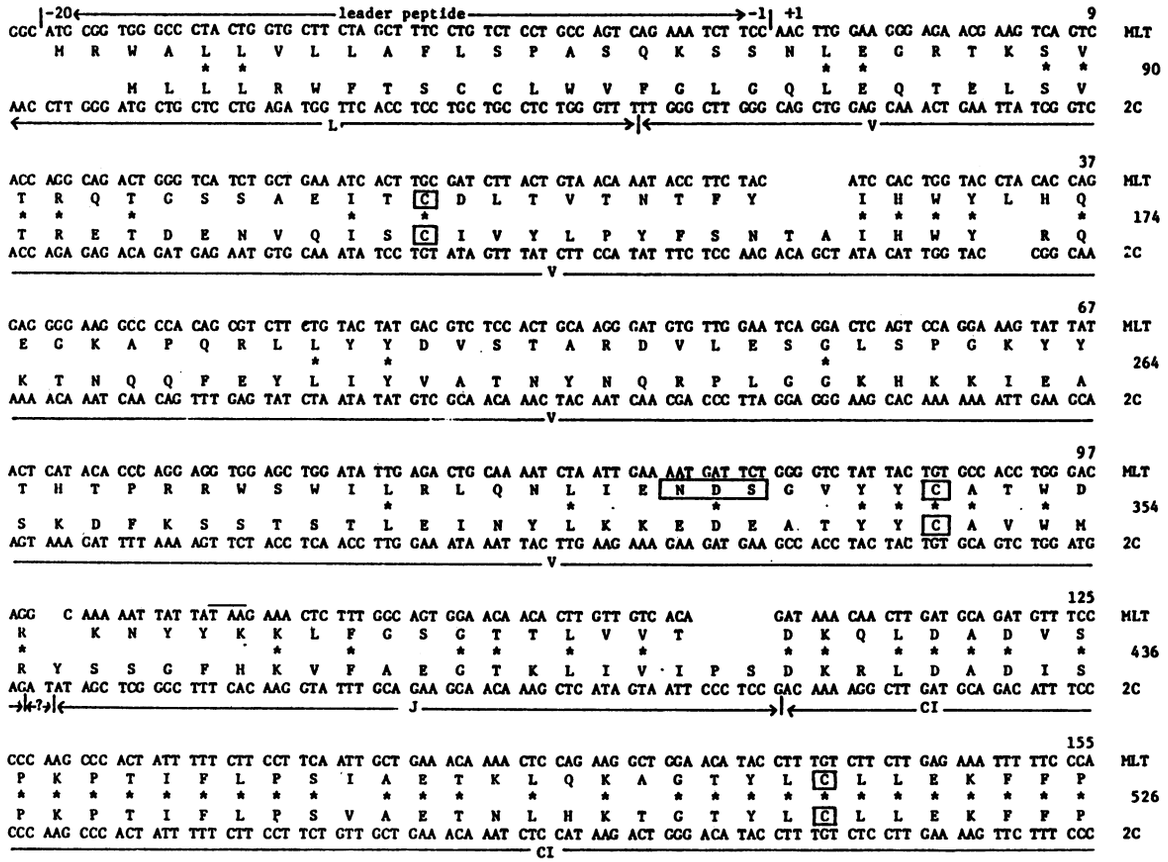

FIG. 3. *(Figure continues on the opposite page.)*

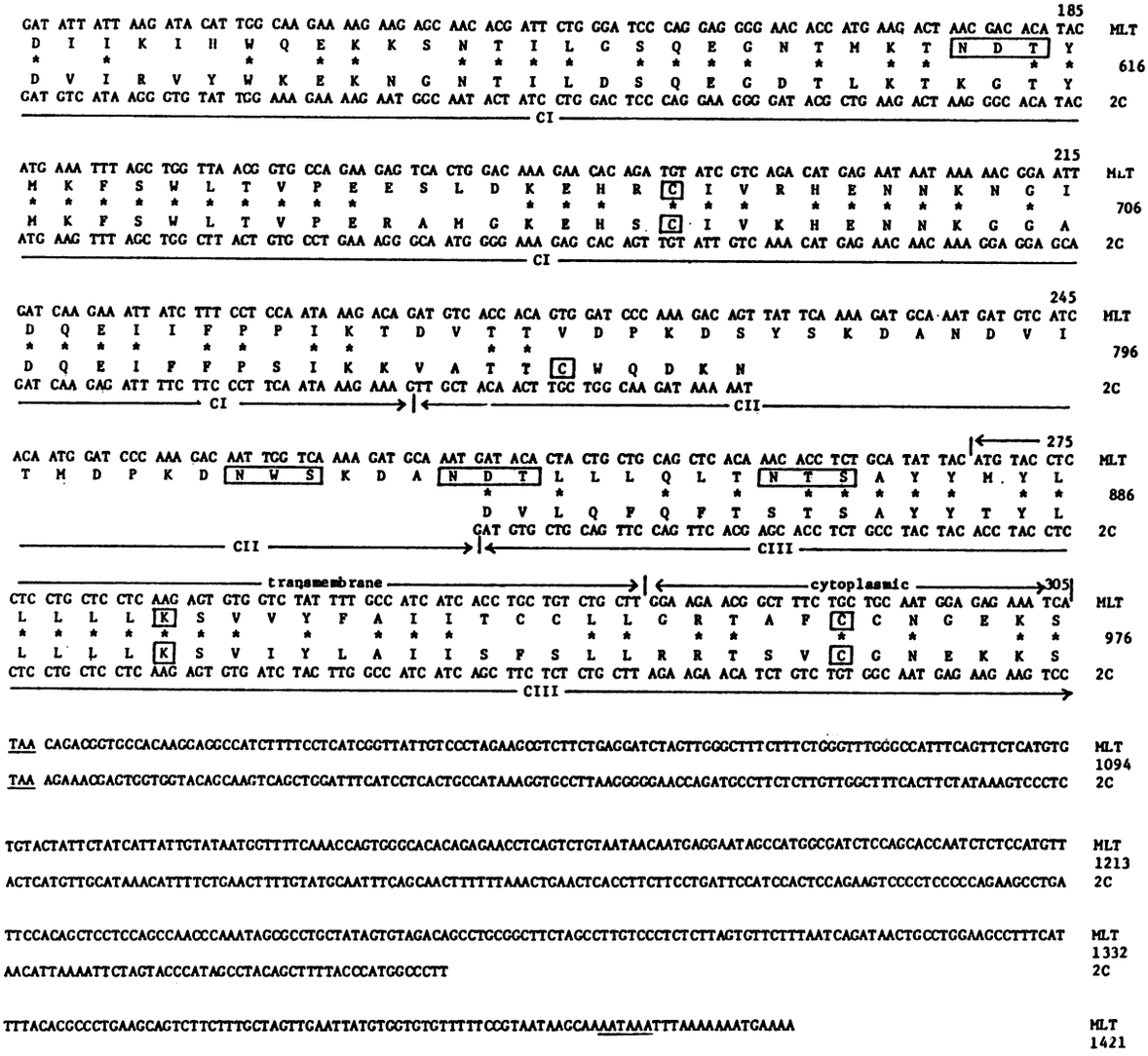

FIG. 3. Comparison of the human and murine T-cell γ-chain cDNA sequences. The nucleotide sequence and corresponding amino acid sequence of pTγ-1/2 (from the human T-cell line HPB-MLT) are compared with the nucleotide sequence and corresponding amino acid sequence of pHDS4 (from murine cytotoxic T-cell line 2C) (11). Nucleotide position for sequence pTγ-1/2 is indicated to the right of the sequence; amino acid position is indicated above the sequence. Amino acids are designated by the one-letter code. In the alignment of the amino acid sequences, several gaps have been introduced as shown. Aligned amino acids are indicated by an asterisk. The approximate junctions of the leader peptide, transmembrane, and cytoplasmic domains are indicated above the sequence, by analogy with those published for pHDS4 (11). The established (1) exon junctions for pHDS4 are indicated below the sequence. The polyadenylylation signal (AATAAA) is underlined. Cysteines believed to be involved in either intrachain or interchain disulfide linkage, potential sites for N-linked glycosylation (N-X-S or N-X-T), and the conserved lysine residue in the transmembrane region are boxed. Note that at the V–J junction there is an apparent dinucleotide deletion in the human cDNA sequence relative to the murine cDNA sequence, resulting simultaneously in the generation of an in-frame stop codon (TAA, overlined) and in a translational frameshift.

kilobases. By flow cytometric analysis, HPB-MLT has characteristics both of an immature (coexpression of T4 and T8) and of a mature (expression of T3) thymocyte (data not shown).34

Screening the HPB-MLT cDNA Library for the Human T-Cell γ Chain. The HPB-MLT cDNA library MLT:λgt10A was screened with the same probe used in the RNA blot analysis. Representation of the γ chain in this library was found to be 0.02%, 1/10th that which has been found in this library for the T-cell receptor β chain (data not shown). Positive recombinant phage were plaque-purified. Two clones, Tγ-1 and Tγ-2, have been characterized extensively.

DNA Sequence Analysis of the HPB-MLT T-Cell γ-Chain cDNA. The inserts from Tγ-1 and Tγ-2 were subcloned into the *Eco*RI site of pUC9 to give pTγ-1 and pTγ-2, respectively. Initially, the insert in pTγ-1 was sequenced in entirety. Because certain features of the pTγ-1 sequence were unexpected, as discussed below, cDNA clone pTγ-2 was also sequenced in entirety. The sequencing strategies for pTγ-1 and pTγ-2 are shown in Fig. 2.

The sequence of pTγ-1 and pTγ-2 is given in Fig. 3. The pTγ-2 sequence begins at nucleotide position 17 in the pTγ-1

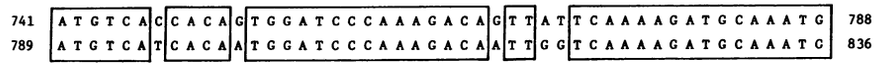

FIG. 4. Putative exon CII for the γ-chain cDNA from HPB-MLT consists precisely of a direct repeat of a 48-nucleotide sequence.

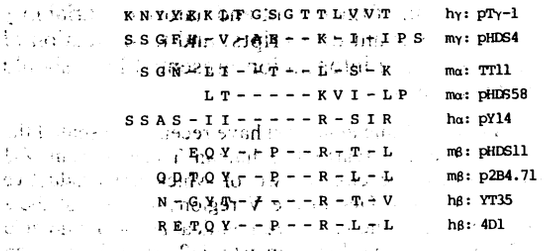

FIG. 5. Comparison of the amino acid sequence of the pTγ-1/2 J region with the amino acid sequence of human (h) or murine (m) J regions expressed in T-cell receptor α- or β-chain cDNAs. The one-letter amino acid code is used. Amino acids shared between a given sequence and pTγ-1/2 are indicated by a dash in the given sequence. The sequences presented here, except for pTγ-1/2, have been published elsewhere as indicated: pHDS4 (11), TT11 (13), pHDS58 (16), pY14 (17), pHDS11 (11), p2B4.71 (14), YT35 (18), 4D1 (15).

sequence. Apart from this, the pTγ-2 sequence is identical to that for pTγ-1, except that cytosine at position 805 in pTγ-1 is replaced by thymine in pTγ-2. This nucleotide difference, which most likely represents an error by reverse transcriptase during construction of the cDNA library, destroys a *Bam*HI site in pTγ-2 and creates an *Ava* II site in pTγ-1. The more representative sequence is pTγ-2, in that six other γ-chain cDNA clones isolated from library MLT:λgt10A also lack an *Ava* II site by restriction analysis (data not shown). The sequence given in Fig. 3 is referred to below as pTγ-1/2.

DISCUSSION

There are a number of points to be made about the pTγ-1/2 sequence. Note that at the V–J junction there is an apparent deletion of two nucleotides in the human cDNA sequence relative to the murine γ-chain cDNA sequence pHDS4 (11), resulting simultaneously in the generation of an in-frame stop codon and in a translational frameshift. The pTγ-1/2 sequence thus encodes an aberrantly rearranged γ chain. In this context, pTγ-1/2 is formally analogous to the murine γ-chain cDNA clone pHDS34 (2), which also sustains at the V–J junction a dinucleotide deletion relative to pHDS4, with similar consequences. As the origin of this pHDS4 dinucleotide is unknown (1), the basis for its absence in pTγ-1/2 remains obscure; however, it has been conjectured that this pHDS4 dinucleotide may have arisen either through terminal transferase activity or through utilization of an as yet unidentified diversity-region exon (1). Southern blot analysis of the γ-chain locus in HPB-MLT (12) and restriction analysis of three other γ-chain cDNA clones from library MLT:λgt-10A (data not shown) together suggest, but do not prove, that there exists but a single γ-chain transcript in HPB-MLT, represented by the pTγ-1/2 sequence.

Even more striking is the discontinuity in amino acid homology between pTγ-1/2 and pHDS4 C regions. Homology is high at both the 5′ and 3′ ends, but it is poor in a discrete segment precisely bounded by the expected junctions of exon CII. The putative pTγ-1/2 CII exon is 3 times the length of the pHDS4 CII exon; moreover, close analysis reveals that it consists precisely of a direct repeat of a 48-nucleotide sequence (Fig. 4). The relationship at the nucleotide level between the putative pTγ-1/2 CII exon and the pHDS4 CII exon is obscure. At least one consequence of this difference is that nowhere in the putative pTγ-1/2 CII exon is there a counterpart to the cysteine found in the pHDS4 CII exon. It is unclear what selective force(s) operate on the γ chain, such that the divergence between humans and mice observed for this segment of the constant region is tolerated.

As indicated, the pTγ-1/2 and pHDS4 constant regions are highly homologous in the segments corresponding to exons CI (74%) and CIII (64%) of pHDS4. Several of the conserved amino acids in exon CIII are of particular interest. These include the Y-X-X-L-L-X-K-X-X-X-Y sequence in the transmembrane domain and the cysteine in the cytoplasmic domain. As previously pointed out (1), the conservation of this transmembrane sequence extends to that of the β chain of the T-cell receptor. The conservation of the cytoplasmic cysteine is especially interesting in light of the apparent lack of conservation of the cysteine found in murine exon CII. It may be that, at least for the human γ chain, any interchain disulfide linkage is effected through the cytoplasmic cysteine.

We have shown elsewhere that there exist two γ-chain C-region gene segments in the human genome, designated C1

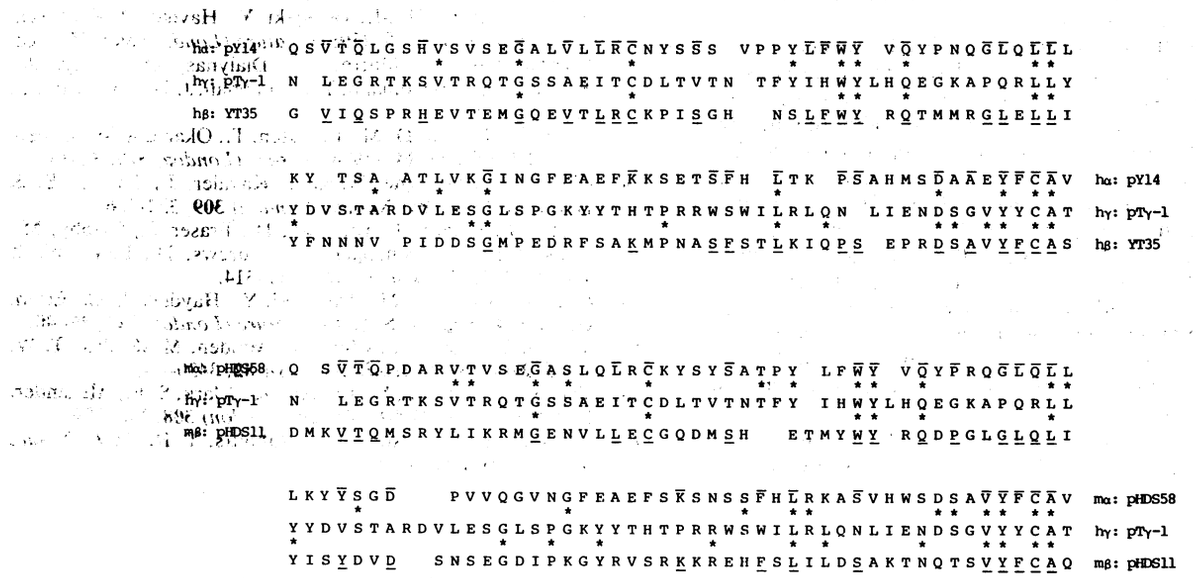

FIG. 6. Comparison of the amino acid sequence of the pTγ-1/2 V region with the amino acid sequence of human (h) or murine (m) V regions expressed in T-cell receptor α- or β-chain cDNAs. The one-letter amino acid code is used. The alignment adopted for pY14 and YT35 is that used previously (17), as is the alignment adopted for pHDS58 and pHDS11 (16). Amino acids aligned between pTγ-1/2 and either of the two companion sequences are indicated by an asterisk. Amino acids aligned between the two companion sequences are indicated by underlining or overlining.

and C2 (4, 12). The CI exon sequenced previously and the pTγ-1/2 CI exon belong to the C1 and C2 segments, respectively (4, 12). Comparison of the two deduced amino acid sequences reveals four differences, no one of which coincides with any one of the nine amino acid differences between the two sequenced murine CI exons (1). There is no obvious contribution of the CI exon to any functional difference between the two γ-chain C-region gene segments.

From Fig. 3, it is apparent that amino acid homology between human and murine γ-chain cDNA sequences extends into the J region. The J region utilized by pTγ-1/2, however, is truncated by two amino acids at its 3′ end relative to that of pHDS4. The homology between pTγ-1/2 and pHDS4 J regions for the most part extends to the J regions found in either murine or human T-cell receptor α- or β-chain cDNAs (Fig. 5). Given that to date all murine γ-chain genomic and cDNA J-region sequences have been found to be identical at the amino acid level (1, 2), one *a priori* might have expected the homology between pTγ-1/2 and pHDS4 J regions to be greater than that found (35%).

As only one of the three closely related murine γ-chain genomic V regions so far identified appears to be transcribed (1, 2), the murine γ-chain V region can thus far be operationally defined as a constant segment. One might therefore have expected the homology between human and murine γ-chain V regions to be comparable to that found for the segments corresponding to the murine CI and CIII exons. The failure to identify a human γ-chain genomic V region, using a murine γ-chain cDNA as probe, under conditions that allowed identification of a human γ-chain CI exon (4) suggested that this was not the case. That this is in fact not the case is indicated in Fig. 3, where it is apparent that the amino acid homology between human and murine γ-chain V regions is relatively poor. This is underscored by the observation that the homology of the pTγ-1/2 V region with the murine γ-chain V region (25%) is not much greater than that with the V region of the human T-cell receptor α (18%) or β (20%) chains and is in fact comparable to that with a V region from a murine T-cell receptor α-chain cDNA (24%) (Fig. 6). The basis for the poor homology between human and murine γ-chain V regions is unclear, although the possibilities include fixation of different V regions that recognize either different epitopes on a highly conserved ligand or different epitopes on a highly divergent ligand.

Finally, a striking feature of the pTγ-1/2 sequence presented in Fig. 3 is the presence of five potential sites for N-linked glycosylation (N-X-S or N-X-T), four of which reside in the C region. This contrasts with the notable absence of such sites in the murine γ-chain sequence (11). Their absence in the murine sequence argues against an absolute requirement for N-linked glycosylation in γ-chain function. The human γ-chain polypeptide has a greater predicted molecular weight (M_r 40,000) than does its murine counterpart (M_r 33,000) (11) due, in part, to the elongated putative CII exon evident in the pTγ-1/2 sequence. Utilization of the potential sites for N-linked glycosylation in the pTγ-1/2 sequence would result in an even greater difference in apparent molecular weight between the human and murine γ-chain proteins.

An understanding of human T-cell γ-chain function will require an analysis of somatic rearrangement of the γ-chain locus, a determination of the extent and nature of variability represented within γ-chain transcripts, and identification of the γ-chain protein. The information presented here should be useful to these ends.

Note Added in Proof. Lefranc *et al.* (19) have recently presented the DNA sequence of several rearranged human T-cell γ-chain V–J regions found on genomic clones, two of which are productive (in-frame) rearrangements. One of the V regions and one of the J regions presented by Lefranc *et al.* are identical at the amino acid level to the V and J regions found in pTγ-1/2, as well as to an unrearranged V region and an unrearranged J region found on genomic clones by Quertermous *et al.* (T.Q., W. Strauss, C. M., D.P.D., J.L.S. & J.G.S., unpublished observations).

D.P.D. is pleased to acknowledge the valuable advice of Drs. Paul Wong, Sam Speck, and particularly Doug Melton during construction of cDNA library MLT:λgt10A. D.P.D. was supported by a Damon Runyon–Walter Winchell postdoctoral fellowship. J.M.B. was supported by National Institutes of Health Postdoctoral Fellowship F32 AI06860. J.G.S. was supported by National Institutes of Health Grants AI-19938 and AI-18436 and by a grant from the Mallinkrodt Foundation. J.L.S. was supported by National Institutes of Health Grant AM-30241.

1. Hayday, A. C., Saito, H., Gillies, S. D., Kranz, D. M., Tanigawa, G., Eisen, H. N. & Tonegawa, S. (1985) *Cell* **40**, 259–269.
2. Kranz, D. M., Saito, H., Heller, M., Takagaki, Y., Eisen, H. N. & Tonegawa, S. (1985) *Nature (London)* **313**, 752–755.
3. Maniatis, T., Fritsch, E. F. & Sambrook, J. (1982) *Molecular Cloning: A Laboratory Manual* (Cold Spring Harbor Laboratory, Cold Spring Harbor, NY).
4. Murre, C., Waldmann, R. A., Morton, C. C., Bongiovani, K. F., Waldmann, T. A., Shows, T. B. & Seidman, J. G. (1985) *Nature (London)* **316**, 549–552.
5. Auffray, C. & Rougeon, F. (1980) *Eur. J. Biochem.* **107**, 303–314.
6. Hunyh, R. V., Young, R. A. & Davis, R. W. (1984) in *DNA Cloning Techniques: A Practical Approach*, ed. Glover, D. H. (IRL, Oxford).
7. Vieira, J. & Messing, J. (1982) *Gene* **19**, 259–268.
8. Maxam, A. M. & Gilbert, W. (1980) *Methods Enzymol.* **65**, 499–560.
9. Biggin, M. D., Gibson, T. J. & Hong, G. F. (1983) *Proc. Natl. Acad. Sci. USA* **80**, 3963–3965.
10. Sanger, R., Nicklen, S. & Coulson, A. R. (1977) *Proc. Natl. Acad. Sci. USA* **74**, 5463–5467.
11. Saito, H., Kranz, D. M., Takagaki, Y., Hayday, A. C., Eisen, H. N. & Tonegawa, S. (1984) *Nature (London)* **309**, 757–762.
12. Quertermous, T., Murre, C., Dialynas, D., Duby, A., Strominger, J., Waldmann, T. & Seidman, J. G. (1985) *Science* **231**, 252–255.
13. Chien, Y., Becker, D. M., Lindsten, T., Okamura, M., Cohen, D. I. & Davis, M. M. (1984) *Nature (London)* **312**, 31–35.
14. Chien, Y., Gascoigne, N. R. J., Kavaler, J., Lee, N. E. & Davis, M. M. (1984) *Nature (London)* **309**, 322–326.
15. Jones, N., Leiden, J., Dialynas, D., Fraser, J., Clabby, M., Kishimoto, T., Strominger, J., Andrews, D., Lane, W. & Woody, J. (1985) *Science* **227**, 311–314.
16. Saito, H., Kranz, D.M., Takagaki, Y., Hayday, A. C., Eisen, H. N. & Tonegawa, S. (1984) *Nature (London)* **312**, 36–40.
17. Yanagi, Y., Chan, A., Chin, M., Minden, M. & Mak, T. W. (1985) *Proc. Natl. Acad. Sci. USA* **82**, 3430–3434.
18. Yanagi, Y., Yoshika, Y., Leggett, K., Clark, S. P., Alexander, I. & Mak, T. W. (1984) *Nature (London)* **308**, 145–159.
19. Lefranc, M. P., Forster, A. & Rabbitts, T. H. (1986) *Nature (London)* **319**, 420–422.