MOLECULAR AND CELLULAR BIOLOGY, Dec. 1987, p. 4407–4413
0270-7306/87/124407-07$02.00/0

Vol. 7, No. 12

Two *lck* Transcripts Containing Different 5′ Untranslated Regions Are Present in T Cells

ANNA F. VORONOVA,[1,2]† HASKELL T. ADLER,[1,3] AND BARTHOLOMEW M. SEFTON[1]*

Molecular Biology and Virology Laboratory, The Salk Institute, San Diego, California 92138,[1] and Departments of Biology[2] and Chemistry,[3] University of California, San Diego, La Jolla, California 92093

Received 29 June 1987/Accepted 1 September 1987

p56^{lck} is a new member of the *src* family of cellular tyrosine protein kinases. It is expressed constitutively at a low level in normal T cells and at an elevated level in the LSTRA and Thy19 Moloney murine leukemia virus-induced thymoma cell lines. It is possible that the expression of p56^{lck} at an elevated level contributes to the transformation of these thymoma cells. The structure of the mRNAs encoding p56^{lck} was examined by using an RNase protection assay. Both a chimeric *lck* mRNA containing the 5′ untranslated region of Moloney virus mRNA and a normal *lck* mRNA were found in Thy19 and LSTRA cells. The chimeric *lck* transcript was 4- to 10-fold more abundant than the normal transcript. Transcription arising from a viral promoter is therefore responsible for the elevated levels of *lck* mRNA in these two cell lines. Surprisingly, uninfected murine T cells were also found to contain *lck* transcripts with differing 5′ untranslated regions. One species of mRNA was colinear with the region of the chromosome just upstream of the initiation codon for p56^{lck}. The other appeared to arise through splicing of an unidentified 5′ untranslated exon to a sometimes cryptic splice acceptor just upstream of the region encoding p56^{lck}. These data suggest that *lck* is expressed through the use of at least two different promoters. The promoters could be subject to different forms of regulation.

p56^{lck} is a new member of the family of nonreceptor, cytoplasmic tyrosine protein kinases, the prototype of which is p60$^{c\text{-}src}$ (12, 22). The gene encoding p56^{lck}, which was previously called either *lsk*^{T} (12) or *tck* (22), is now designated *lck*. It is expressed in all T cells (11, 12, 19, 21, 22) and in some B cells (12; T. Hurley, personal communication). Neither *lck* mRNA nor p56^{lck} has been detected in hematopoietic cells of other lineages, in fibroblasts, or in nonlymphoid organs of the mouse (12, 21, 22). It appears that the expression of the *lck* gene is restricted to lymphoid cells and occurs predominantly, but not exclusively, in T cells. The presence of p56^{lck} in all T cells suggests that it may have some normal function in T lymphocytes.

p56^{lck} is expressed at an elevated level in two Moloney murine leukemia virus (Mo-MuLV)-induced murine thymomas, LSTRA and Thy19 (3, 21). Its expression at an unusually high level may contribute to the transformed phenotype of these cells. p56^{lck} is encoded in LSTRA cells by a hybrid mRNA which contains a 5′ untranslated region derived from a Mo-MuLV provirus (22). Transcription of this mRNA originates at a viral promoter and is probably produced by splicing between the normal viral splice donor for the subgenomic envelope mRNA (nucleotide 205 of the Mo-MuLV genome) and a splice acceptor site located five nucleotides upstream from the initiation codon for p56^{lck}.

The Thy19 thymoma cell line, which was derived from a BALB/c(Mo) mouse thymic tumor, also contains an elevated level of both p56^{lck} and *lck* mRNA. The reason for the overexpression of the *lck* gene in these cells remains to be determined. However, DNA from the Thy19 cell line contains 6.6-kilobase *Bam*HI and 9.0-kb *Hin*dIII restriction fragments which can be detected by hybridization with *lck* cDNA which are not found in normal murine DNA (our unpublished results). The new restriction fragments suggest that a region of the chromosome upstream of one of the alleles of the *lck* gene in Thy19 cells has undergone rearrangement. It is likely that the rearrangement of the region of the chromosome upstream of the *lck* gene contributes to the elevated levels of *lck* mRNA in these cells. One possibility is that a Mo-MuLV provirus is integrated at a site similar but not identical to that occupied by a Mo-MuLV provirus in LSTRA cells. Southern blot hybridization of DNA from the Thy19 cell line with a probe specific for the long terminal repeat of Mo-MuLV, which does not anneal significantly to the long terminal repeat of endogenous viruses (14), detects *Hin*dIII and *Bam*HI fragments of DNA from the Thy19 cell line that are the same size as those containing the rearranged 5′ end of the *lck* gene (our unpublished results). Since there are many Mo-MuLV proviruses present in Thy19 DNA (our unpublished results), this apparent similarity could be simply coincidental. Therefore the RNase protection technique (13) was used to examine the structure of the *lck* mRNA in the Thy19 cell line directly. Unexpectedly, we found that even uninfected murine T cells contained at least two species of *lck* mRNA.

MATERIALS AND METHODS

Cells. The LSTRA line of Mo-MuLV-induced BALB/c thymoma cells (7) and the BW5147 line of spontaneously transformed AKR/J mouse T lymphoma cells (10) were grown in Dulbecco-Vogt modified Eagle medium supplemented with 20% fetal calf serum and 5×10^{-5} M 2-mercaptoethanol. The Thy16 and Thy19 cell lines, which were isolated from spontaneous thymomas of BALB/c(Mo) mice (20), were a gift from M. Vogt and were all grown in the above medium. The E5 and L2 T helper cell lines, the LTT.3.4 and LTT.2.14.11 T cell hybridomas (generated by fusion of the E5 and L2 cell lines with BW5147 cells), and the RG17.16.ST cell line (derived from the fusion of AD5 T-helper cells with BW 5147 cells [6]) were all from A. Glasebrook. The WEHI 7 X-ray-induced BALB/c mouse thymoma cell line (9) was provided by D. Gruol.

* Corresponding author.
† Present address: Whitehead Institute, Cambridge, MA 02142.

RNA isolation. RNA was prepared by using guanidine thiocyanate (4). Briefly, cells were lysed in 4 M guanidine thiocyanate–0.1% antifoam A–25 mM sodium citrate (pH 7.0)–0.5% *N*-lauroyl sarcosine–0.1 M 2-mercaptoethanol. Lysates were then layered over a 3.5-ml cushion of 5.7 M CsCl in 30 mM sodium acetate (pH 5.0)–5 mM EDTA and subjected to centrifugation in an SW41 rotor for 16 h at 38,000 rpm at 22°C. The pellets containing the RNA were dissolved in diethyl pyrocarbonate-treated water and precipitated twice with ethanol in the presence of 0.3 M sodium acetate.

Isolation of *lck* cDNA clones. The normal thymic *lck* cDNA was isolated from an oligo(dT)-primed C57BL/ka thymus cDNA library in λgt10 (1), obtained from S. Hedrick. Bacteriophage containing *lck* cDNA were identified by hybridization with a *Pst*I fragment consisting of nucleotides 210 to 1081 of the hybrid *lck* cDNA described previously (22). Nucleotide sequence analysis showed the *lck* cDNA to be identical to that isolated by Marth et al. (12) from the same library. The isolation of the hybrid *lck* cDNA from LSTRA cells was described previously (22).

Isolation of λ clones containing chromosomal *lck*. A mouse genomic library from the BALB/c plasmocytoma cell line MOPC41 (17) in bacteriophage λ Charon 4A, obtained from G. Evans, was screened by using the *Pst*I fragment of *lck* cDNA described above as a probe. A positive recombinant clone containing a 17-kilobase DNA insert was isolated. The *Eco*RI restriction fragments of the insert DNA were subcloned into the plasmid pEMBL8$^+$ (5). A 1-kilobase *Eco*RI fragment containing sequences corresponding to the 5′ end of the *lck* mRNA was identified by hybridization with a synthetic oligonucleotide probe complementary to nucleotides 235 through 252 of the hybrid *lck* cDNA (22). The nucleotide sequence of the clone was determined by the dideoxyoligonucleotide chain termination method (15).

Construction of recombinant plasmids. The *Asp*718-*Eco*RI fragment from the hybrid *lck* cDNA isolated from LSTRA cells, containing nucleotides 17 through 1778 (22), was made blunt ended by filling in with the Klenow fragment of DNA polymerase I and inserted into the *Sal*I site of pIBI30 (International Biotechnologies, Inc.) by using *Xho*I linkers. The resultant plasmid was then digested with *Bam*HI and *Bgl*II restriction endonucleases and religated to generate a template plasmid designated pIBI30*lck*^{L}BB. This plasmid contains nucleotides 17 through 295 of the *lck* cDNA isolated from LSTRA cells (22).

The *Eco*RI restriction fragment from *lck* cDNA isolated from normal thymus, containing nucleotides 2 through 1748 of the published sequence (12), was inserted into the *Eco*RI site of pIBI76 (International Biotechnologies). The resultant plasmid was then digested with *Bam*HI and *Bgl*II restriction endonucleases and religated to generate a template plasmid, pIBI76 *lck*BB. This plasmid contains nucleotides 1 to 265 of the *lck* cDNA (12).

Synthesis of complementary-strand T7 RNA probes. The pIBI30*lck*^{L}BB and pIBI76*lck*BB plasmids described above were digested with restriction endonucleases *Eco*RI and *Hin*dIII, respectively, both of which cut the plasmids once in the polylinker. Linear template DNA was then purified by centrifugation through a 5 to 20% sucrose gradient in 20 mM Tris hydrochloride (pH 7.9)–2 mM EDTA–0.2 M NaCl–0.1% sarcosyl–5 μg of ethidium bromide per ml for 3 h at 48 K in an SW60 rotor (Beckman Instruments, Inc.). The linear DNA was identified visually, removed with a syringe, and precipitated with ethanol in the presence of 0.3 M sodium acetate. Transcription with T7 RNA polymerase was carried out according to the protocol of the manufacturer (Promega Biotec). Briefly, 10-μl reaction mixtures contained 40 mM Tris hydrochloride (pH 7.5), 6 mM $MgCl_2$, 2 mM spermidine, 10 mM NaCl, 10 mM dithiothreitol, 1 U of RNasin RNase inhibitor (Promega Biotec) per ml, 0.5 mM (each) rATP, rGTP, and rCTP, 50 μM rUTP, 100 μCi of [^{32}P]rUTP (600 Ci/mmol; New England Nuclear Corp.), 0.5 to 1 μg of template DNA, and 5 to 8 U of T7 RNA polymerase (Promega Biotec). Synthesis was allowed to proceed for 45 min at 37°C. The DNA template was then degraded by incubation of the above mixture with 1 U of RNase-free RQ1 DNase (Promega Biotec) at 37°C for 15 min. The RNA was extracted with an equal volume of phenol-chloroform and then precipitated with ethanol in the presence of 3.7 M ammonium acetate and 20 μg of yeast tRNA. The pellet was suspended in 80% formamide–dye mix sample buffer, and the RNA was purified by electrophoresis on a 6% acrylamide–7 M urea gel in Tris-borate buffer. The RNA was detected by autoradiography, excised from the gel, and electroeluted with an Isco 1750 electrophoretic sample collector. The RNA was then precipitated with ethanol in the presence of 0.3 M sodium acetate and 20 μg of yeast tRNA.

Solution hybridization and RNase digestion (RNase protection). Dried RNA samples (30 μg of total cellular RNA or 1 μg of the polyadenylated RNA and 30 μg of tRNA) were dissolved in 30 μl of hybridization buffer [80% formamide, 400 mM NaCl, 1 mM EDTA, and 40 mM piperazine-*N*,*N*′-bis(2-ethanesulfonic acid), pH 6.5] containing 5×10^5 cpm of RNA probe prepared as described above. The mixture was incubated at 85°C for 10 min and then allowed to anneal at 46°C overnight. After hybridization, 300 μl of RNase digestion buffer (10 mM Tris hydrochloride [pH 7.5], 5 mM EDTA, 300 mM NaCl) containing 40 μg of RNase A per ml and 2 μg of RNase T1 per ml was added. RNase digestion was allowed to proceed at 37°C for 45 min in the case of the pIBI76*lck*BB-derived probe or for 2 h in the case of the pIBI30*lck*^{L}BB-derived probe. The reaction was stopped by the addition of 20 μl of 10% sodium dodecyl sulfate and 50 μg of proteinase K (Merck & Co.) and incubation at 37°C for 15 min. The sample was then extracted once with water-saturated ether and then with phenol-chloroform (1:1), and the resultant ^{32}P-labeled RNA-RNA hybrids were precipitated by ethanol in the presence of 0.3 M sodium acetate and 20 μg of yeast tRNA. The pellets were dried, dissolved in the 80% formamide sample buffer, and analyzed on a 6% acrylamide–7 M urea gel.

RESULTS

Expression of *lck* in the Thy19 cell line. We used the RNase protection technique to analyze the structure of *lck* mRNA in Thy19 cells. Total cellular RNA or poly(A)$^+$ RNA (where indicated) from the LSTRA, Thy19, Thy16, and BW5147 cell lines were analyzed. The plasmid pIBI30*lck*^{L}BB includes an *Asp*718-*Bgl*II restriction fragment, containing nucleotides 17 through 295, from the hybrid *lck* cDNA isolated from LSTRA cells. This fragment consists of 107 nucleotides encoded by the *lck* gene, including the first 102 nucleotides of the coding region, and 172 nucleotides identical to the 5′ leader of Mo-MuLV genomic RNA. Cleavage of this DNA template with restriction endonuclease *Hin*dIII and transcription with T7 RNA polymerase yields a 330-nucleotide RNA containing 279 nucleotides encoded by the cDNA, flanked by a total of 51 nucleotides from the polylinker at its 5′ and 3′ ends (Fig. 1A). With this probe it is possible to distinguish between *lck* mRNA species initiated at a cellular

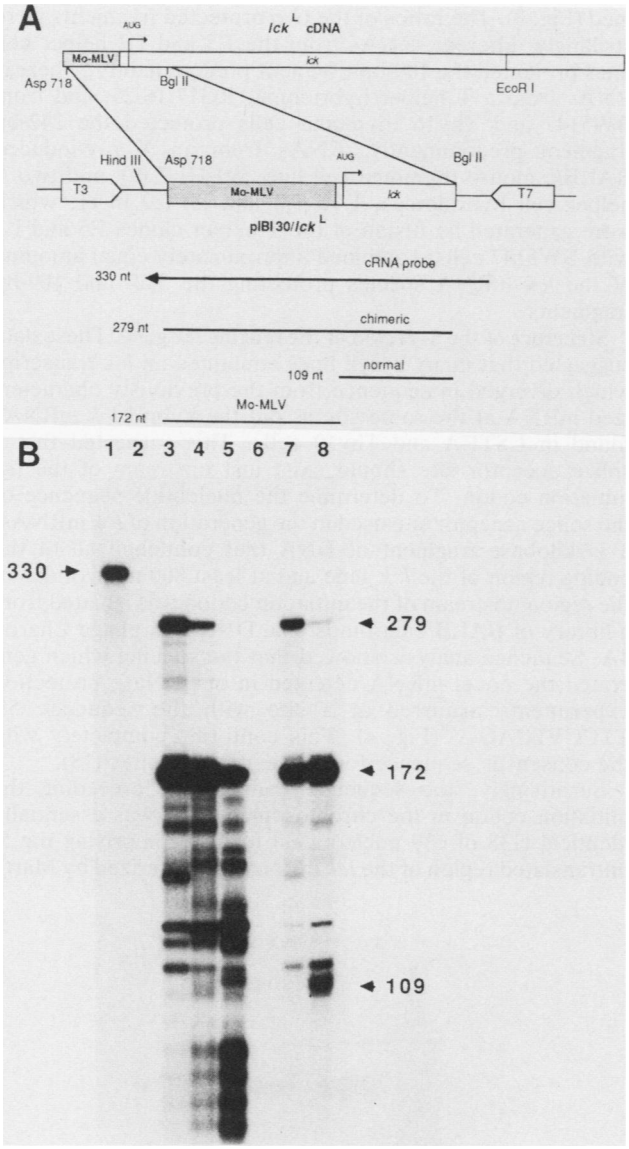

FIG. 1. *lck* mRNA containing a viral 5' untranslated region present in Thy19 cells. (A) Structure of pIBI30/*lck*^{L}BB DNA and its relationship to the chimeric *lck* cDNA from LSTRA cells. The details of the plasmid construction are described in Materials and Methods. The probe was synthesized by using T7 RNA polymerase. The probe and the fragments that would be protected by the chimeric Mo-MuLV *lck* transcript, the normal cellular *lck* transcript, and either Mo-MuLV genomic or subgenomic mRNA are shown. (B) Gel analysis of fragments of the probe protected from RNase digestion by cellular RNA as follows. Lanes: 1, undigested RNA probe; 2, 30 μg of tRNA; 3, 30 μg of total RNA from LSTRA cells; 4, 30 μg of total RNA from Thy19 cells; 5, 30 μg of total RNA from Thy16 cells; 6, 30 μg of total RNA from BW5147 cells; 7, 1 μg of poly(A)$^+$ RNA from LSTRA cells; 8, 1 μg of poly(A)$^+$ RNA from Thy19 cells.

promoter and those initiated from the Mo-MuLV promoter. Hybridization of this probe with the chimeric mRNA from LSTRA cells initiated at the Mo-MuLV promoter will protect 279 nucleotides of the probe from subsequent RNase digestion. Hybridization with mRNA initiated from a cellular promoter and therefore lacking the Mo-MuLV leader sequences will protect a 109-base-pair (bp) fragment. Finally, hybridization with Mo-MuLV genomic RNA or subgenomic mRNA will protect 172 bp of the probe.

RNA from both the LSTRA and Thy19 cell lines protected a 279-bp fragment of the probe, whereas RNA from the Thy16 and BW5147 cell lines did not (Fig. 1B). This suggests strongly that the Thy19 cell line also contains hybrid *lck* transcripts. A 172-bp protected fragment was prominent in the samples prepared with RNA from the LSTRA, Thy19, and Thy16 cell lines but not in those prepared with RNA from BW5147 cells (Fig. 1B). Such a fragment would be protected by hybridization with either Mo-MuLV genomic or subgenomic mRNA. Although it is difficult to see in Fig. 1B, a trace of a 109-bp fragment was protected by RNA from BW5147 cells. Such a species was not apparent when LSTRA and Thy16 cell RNAs were analyzed, probably because it was obscured by a multiplicity of other bands arising from hybridization with viral RNA. A 109-bp fragment was protected by poly(A)$^+$ RNA from Thy19 cells (Fig. 1B, lane 8).

Expression of *lck* in lymphoid cells. To determine whether both the normal and the rearranged alleles of the *lck* gene are transcribed in LSTRA and Thy19 cells, an RNA probe derived from a pIBI76*lck*BB plasmid was utilized (Fig. 2A). This plasmid contains nucleotides 1 to 265 of the *lck* cDNA isolated from normal thymic cells. This fragment includes 163 nucleotides of the 5' untranslated region of the *lck* mRNA and 102 nucleotides of the coding region (12). Transcription initiated at the T7 promoter and terminated at the *Eco*RI site yields a 314-nucleotide RNA probe, which includes 54 nucleotides from the pIBI76 polylinker at its 5' end. The predicted size of the fragment which would be protected by the normal cellular transcript is 242 bp. Likewise, the chimeric transcript which originates at the proviral promoter and lacks the 5' untranslated region of the cellular mRNA should protect a 109-nucleotide fragment.

RNA from LSTRA cells protected two fragments of 242 and 109 nucleotides (Fig. 2B). These correspond exactly to the predicted fragments resulting from protection by the normal cellular and chimeric transcripts, respectively. The presence of these fragments indicates that both the rearranged and the normal alleles of the *lck* gene are transcriptionally active in LSTRA cells. Quantification of these bands by densitometry revealed that the hybrid *lck* mRNA was approximately 10-fold more abundant than the normal cellular transcript.

This finding demonstrated that the predominant *lck* transcript in LSTRA cells is initiated from the Mo-MuLV promoter. Similar analysis with Thy19 RNA also yielded both a 242-bp and an approximately 109-bp protected fragment (Fig. 2B). Here the smaller fragment was only fourfold more abundant than the larger fragment. We cannot explain the small apparent difference in size between the smaller protected fragments observed with RNA from LSTRA and Thy19 cells.

RNA from the Thy16 and BW5147 cell lines also yielded small amounts of the 109-bp protected fragment. This was surprising. Although Thy16 cells are infected with Mo-MuLV, Southern blot analysis suggests that the *lck* gene has not undergone rearrangement near its 5' end (our unpublished data). Additionally, the BW5147 cell line, which was derived from a mineral oil-induced thymoma, does not contain a Mo-MuLV provirus which could give rise to a hybrid transcript.

To examine whether RNA species protecting a 109-bp fragment of this probe are peculiar to these cell lines, RNA from several other well-characterized T-cell lines which did

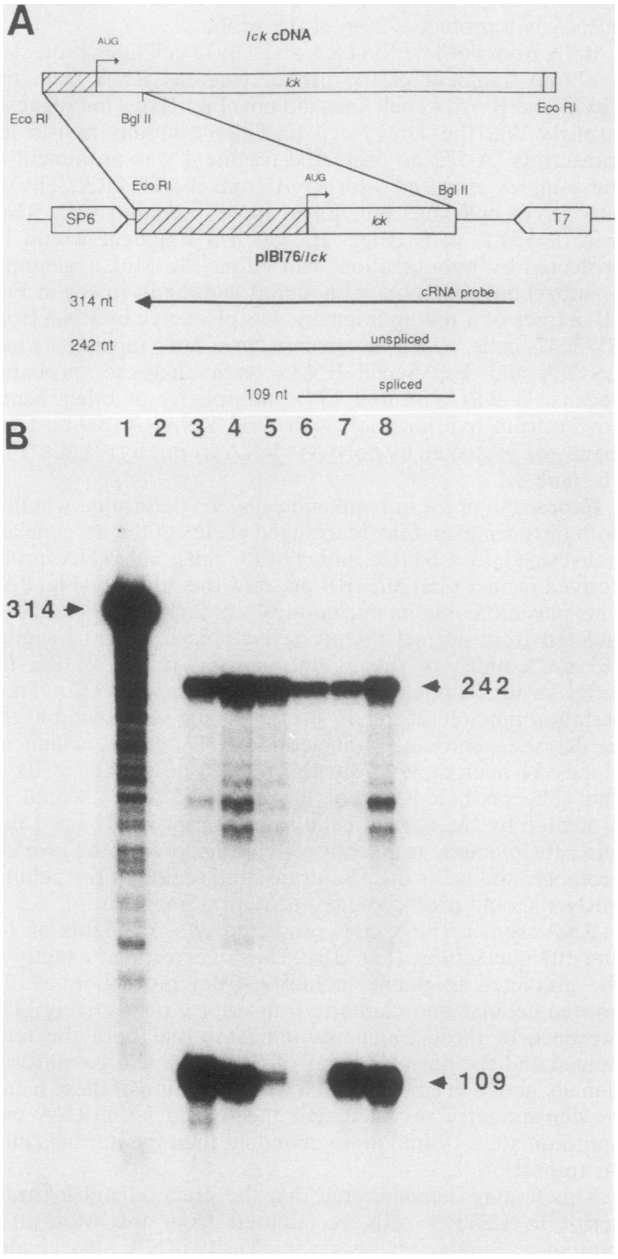

FIG. 2. LSTRA and Thy19 cells contain *lck* transcripts arising from both a viral and a cellular promoter. (A) Structure of pIBI76*lck*BB and its relationship to the *lck* cDNA described by Marth et al. (12). The details of the plasmid construction are described in Materials and Methods. T7 RNA polymerase was used to synthesize the probe. The probe and the fragments of the probe which would be protected by spliced and unspliced transcripts are diagrammed. (B) Gel analysis of fragments of the probe protected from RNase digestion by cellular RNA as follows. Lanes: 1, undigested RNA probe; 2, 30 μg of tRNA; 3, 30 μg of total RNA from LSTRA cells; 4, 30 μg of total RNA from Thy19 cells; 5, 30 μg of total RNA from Thy16 cells; 6, 30 μg of total RNA from BW5147 cells; 7, 1 μg of poly(A)$^+$ RNA from LSTRA cells; 8, 1 μg of poly(A)$^+$ RNA from Thy19 cells.

not originate from the infection of mice with Mo-MuLV was isolated and subjected to RNase protection analysis with the pIBI76*lck*-derived RNA probe. Both 242- and 109-bp fragments were protected by all of the RNA preparations examined (Fig. 3). The ratios of the two protected fragments were strikingly different. RNAs from the E5 and L2 helper cell lines protected the 109-bp fragment predominantly, whereas RNAs from a T helper hybridoma, RG17.16.25, and from BW5147 and Thy16 thymoma cells protected the 242-bp fragment predominantly. RNAs from an X ray-induced BALB/c mouse thymoma cell line, WEHI 7 (9), and two T helper cell hybridomas, LTT.3.4 and LTT.2.14.11, which were generated by fusion of the T helper clones E5 and L2 with BW5147 cells, contained approximately equal amounts of the *lck* mRNA species protecting the 242- and 109-bp fragments.

Structure of the 5′ region of the murine *lck* gene. These data suggested that many T-cell lines contained an *lck* transcript which diverged in sequence from the previously characterized mRNA at the same site as did the hybrid *lck* mRNAs found in LSTRA and Thy19 cells. This suggested that a splice acceptor site should exist just upstream of the *lck* initiation codon. To determine the nucleotide sequence of the splice acceptor site used in the generation of *lck* mRNAs, a 17-kilobase fragment of DNA that contained all of the coding region of the *lck* gene and at least 800 nucleotides of the region upstream of the initiation codon was isolated from a library of BALB/c chromosomal DNA in λ phage Charon 4A. Sequence analysis showed that the splicing which generated the novel mRNA detected in our RNase protection experiments occurred at a site with the sequence 5′-TTCCTTCAG-3′ (Fig. 4). This conforms completely with the consensus sequence for splice acceptor sites (18).

Surprisingly, the sequence immediately preceding the initiation codon in the chromosomal DNA was essentially identical (138 of 139 nucleotides) to that comprising the 5′ untranslated region of the *lck* mRNA characterized by Marth

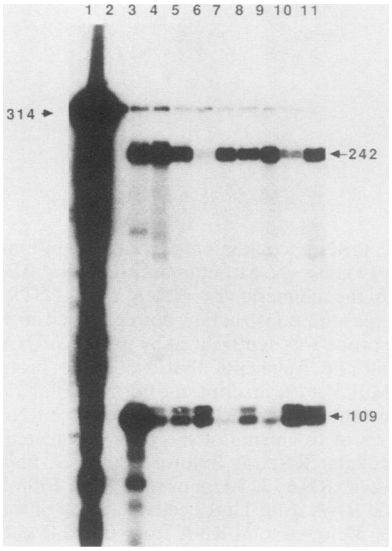

FIG. 3. *lck* transcripts containing multiple 5′ untranslated regions are present in uninfected T cells. The structure of the 5′ end of *lck* mRNA was analyzed by using the RNase protection technique with the same probe described in the legend to Fig. 2. The protected fragments were then analyzed by gel electrophoresis. Cells were protected with 30 μg of RNA as follows. Lanes: 1, undigested RNA probe; 2, tRNA; 3, total RNA from LSTRA cells; 4, total RNA from Thy16 cells; 5, total RNA from LTT2.14.11 cells; 6, total RNA from L2 cells; 7, total RNA from BW5147 cells; 8, total RNA from LTT 3.4 cells; 9, total RNA from RG17.16.25 cells; 10, total RNA from E5 cells; 11, total RNA from WEHI 7 cells.

FIG. 4. Comparison of the nucleotide sequences of the hybrid *lck* cDNA (*lck*^{L} cDNA) isolated from the LSTRA cell line, *lck* cDNA from a normal C57BL/ka mouse thymus (*lck* cDNA), and the *lck* gene (genomic *lck*) from the MOPC41 cell line. The nucleotide sequence of the *lck*^{L} cDNA is taken from Voronova and Sefton (22), and that of the *lck* cDNA is from a clone identical to that isolated by Marth et al. (12). The numbering systems for both sequences are retained here. The *lck* cDNA sequence presented here differs from that reported by Marth et al. (12) in that the sequence of the synthetic *Eco*RI linker at the 5' end of the published sequence is not included and nucleotides 61, 65, and 71 differ from those reported previously. Identical nucleotides are indicated by vertical bars. S.A. designates the consensus splice acceptor site. Eighteen nucleotides of the *lck* cDNA that have imperfect sequence homology with the corresponding region of genomic *lck* are underlined. The reverse complement of these 18 nucleotides present in genomic *lck* is boxed.

et al. (12). Present unused in the *lck* cDNA is the splice acceptor site identified in the chromosomal DNA.

DISCUSSION

Our data suggest that murine thymocytes contain more than one mRNA encoding p56^{lck}. One is that described by Marth et al. (12). The 5' untranslated region of this mRNA is colinear with the region of the chromosome immediately upstream of the sequences encoding p56^{lck}. This mRNA is present in all of the murine T cell lines we have examined and in normal thymus (12).

Most murine thymocytes also contain at least one other form of *lck* mRNA. This other mRNA contains a different 5' untranslated region. Only the five nucleotides that adjoin the initiation codon in this mRNA are the same as those in the 5' untranslated region of the mRNA characterized by Marth et al. In the chromosome, these five nucleotides comprise the right side of the sequence 5'-TTCCTTCAGGGATC-3'. The sequence 5'-TTCCTTCAG-3' found here conforms completely to the consensus sequence for a splice acceptor site (18), with the AG comprising the 3' end of an intron. It seems very likely, therefore, that the novel mRNA(s) we have identified arises through the splicing of an as yet unidentified 5' untranslated exon to this site.

We have identified this novel mRNA only because it lacks the 5' end present in the cDNA clone described by Marth et al. (12). We cannot therefore determine whether there exists only one additional *lck* mRNA or multiple *lck* mRNAs, containing a diversity of 5' ends, joined to the same splice acceptor site.

The relative abundance of mRNAs containing spliced 5' leaders was variable. Such transcripts comprised more than half of the *lck* mRNA in two T helper cell lines. In contrast, they represented only 1 to 10% of the *lck* mRNAs in the BW5147 and Thy16 thymoma cell lines. We have not yet analyzed mRNA from either primary thymus tissue or cytotoxic T lymphocytes.

lck mRNA is present at a significantly elevated level in both the LSTRA and Thy19 thymoma cell lines. RNase protection analysis shows that the most abundant *lck* mRNA in both cell lines is a chimeric molecule made up of 205 nucleotides encoded by the Mo-MuLV provirus at its 5' end and a body of 1,979 nucleotides encoded by the *lck* gene. In both cell lines, this mRNA arises as the result of integration of a Mo-MuLV provirus upstream of the *lck* coding region. The chimeric transcript is apparently initiated in the 5' long terminal repeat of the provirus and diverges in sequence from that of Mo-MuLV RNA at nucleotide 206, the splice donor used to generate the subgenomic viral mRNA. The viral leader is joined to the *lck* sequences at the same splice acceptor used to produce the novel transcript described here. It seems reasonable to conclude, therefore, that the chimeric *lck* mRNA in LSTRA and Thy19 cells is produced by splicing. The alternative possibility, that a deleted Mo-MuLV provirus is integrated exactly five nucleotides upstream of the initiation codon for p56^{lck}, is extremely unlikely. Indeed, the same chimeric *lck* mRNA is seen in both LSTRA and Thy19 cells, yet the sites of proviral integration are clearly different (data not shown).

Both the rearranged and the normal *lck* genes are transcribed in LSTRA and Thy19 cells. In LSTRA cells, *lck* transcripts arising from the viral promoter may be as much as 10-fold more abundant than transcripts arising from cellular promoters. Since we have only examined steady-state RNA levels, we do not know whether the viral promoter is more active than the cellular promoters or whether the *lck* mRNA with the viral 5' end is more stable than the mRNA with a cellular 5' end.

Discrepancy between the sequence of *lck* cDNA and the *lck* gene. Except for one residue, the sequence of *lck* chromosomal DNA is identical to that of residues 24 to 163 of the 5' untranslated region of the cDNA isolated by Marth et al.

(12). We suspect that the one-nucleotide difference at nucleotide 155 of the genomic *lck* sequence is real but is peculiar to the MOPC41 cell line from which the genomic clone was isolated. Were the sequence of BALB/c *lck* different from that of C57BL/ka *lck*, it is probable that we would not have observed a 242-bp protected fragment in Fig. 2 and 3.

In contrast, only 7 of the first 15 nucleotides of the *lck* cDNA clone are identical to those in the chromosomal DNA (Fig. 4). We have determined the nucleotide sequence of the 5′ untranslated region of an *lck* cDNA clone isolated from the same cDNA library as that used by Marth et al. (12). It is likely that we reisolated the same cDNA clone. The *lck* cDNA clone which we isolated differs from the chromosomal DNA at the same eight positions at the 5′ end of the clone.

Why does the sequence of the cDNA diverge partially from that of the chromosomal DNA in this region? It could do so because a partially homologous sequence was linked by splicing to a splice acceptor in this region. There is, however, no consensus splice acceptor site in this region. Alternatively, the fact that there is partial homology suggests that the sequence differences could have arisen as a result of the removal of a hairpin structure that primed the synthesis of the second strand of cDNA.

Residues 16 to 33 of the sequence of chromosomal *lck* DNA presented here (boxed in Fig. 4) are the exact reverse complement of the first 18 residues of the cDNA (underlined in Fig. 4). We suspect that the synthesis of the first strand of cDNA, which would be the complement of that shown here, terminated after copying the residue we have numbered 16 in the genomic *lck* sequence (Fig. 4). Were this the case, the 18 nucleotides at the end of the first strand (the complement of the boxed nucleotides) would be able to fold back and form an imperfect hybrid with residues complementary to the nucleotides numbered 39 to 56 in the sequence of genomic *lck* (10 of 18 nucleotides are complementary). The resultant hairpin structure could then have primed the synthesis of the second strand of cDNA, and the concerted action of DNA polymerase I and S1 nuclease could have generated the 5′-terminal sequence of the clone isolated by Marth et al. This would have occurred if a short portion of the strand containing the complement of the residues numbered 39 to 56 in the genomic sequence were degraded and the strand containing the nucleotides complementary to residues 16 to 33 was then copied during the fill-in reaction. We would argue, therefore, that the sequence of the first 23 nucleotides of the *lck* cDNA sequence reported by Marth et al. does not correspond to that of *lck* mRNA and that the 5′ end of the mRNA has a sequence identical to that of the *lck* gene reported here.

How many promoters give rise to *lck* transcripts? The several *lck* mRNAs present in uninfected T cells most probably arise from more than one promoter. The *lck* transcripts in thymus are approximately 2,200 nucleotides in size (12). Because the *lck* cDNA clone isolated by Marth et al. from thymus cDNA contains 2,086 nucleotides while lacking the entire poly(A) tail, almost all of the mRNA must be represented in this clone. The 5′ untranslated region of this mRNA is essentially colinear with that of the chromosome, save for the 15 nucleotides at the extreme 5′ end. We have just argued that this region of imperfect sequence homology is probably artifactual and arose as a result of hairpin formation during cDNA synthesis. It is likely therefore that almost all of the 5′ untranslated region of the *lck* mRNA is present in this cDNA clone. If this is the case, the promoter giving rise to this transcript should be located approximately 200 nucleotides upstream of the *lck* coding sequences. Inspection of the sequence of this region of the chromosome reveals no obvious TATA or CAAT elements (data not shown).

The novel *lck* mRNA described here is produced by splicing and probably arises from a second promoter located farther upstream. The sites of integration of the Mo-MuLV proviruses in LSTRA and Thy19 cells, relative to the two presumptive *lck* promoters, is not yet known.

Is there any physiological significance to the apparent existence of more than one *lck* promoter and of multiple *lck* mRNAs containing different 5′ untranslated regions? One possibility is that the use of multiple promoters increases the diversity of means through which the expression of this gene can be regulated. The multiple *lck* promoters could be sensitive to different means of induction or repression. This would increase the number of agents able to regulate the transcription of the gene. Such a mechanism has been observed previously. Both the α-amylase gene of mice (8, 16) and the alcohol dehydrogenase gene of *Drosophila melanogaster* (2) are expressed via two mRNAs which have different 5′ untranslated regions and are transcribed from separate promoters active in different cell types.

Alternatively, the nature of the 5′ end of the mRNA could affect either mRNA stability or the efficiency of translation. In this context, it is notable that the 5′ untranslated region of the *lck* mRNA which gave rise to the clone characterized by Marth et al. (12) contains three AUG codons upstream of the initiation codon for $p56^{lck}$. Interference between these several initiation codons could conceivably render the translation of this mRNA somewhat inefficient.

ACKNOWLEDGMENTS

We thank Tamara Hurley for help with the isolation and sequence analysis of the *lck* cDNA clone from thymus RNA, Vincent Raymond for advice about RNase protection analysis, Andy Glasebrook for cell lines, Glen Evans for the library of BALB/c genomic DNA, Steve Hedrick for the library of thymus cDNA, and Mark Kamps and Sheila Podell for reading the manuscript critically.

This work was supported by Public Health Service grants CA 14195 and CA 42350 from the National Institutes of Health.

LITERATURE CITED

1. **Barth, R. K., B. S. Kim, N. C. Lan, T. Hunkapiller, N. Sobic, A. Winoto, H. Gershenfeld, C. Ocada, I. L. Weissman, and L. Hood.** 1985. The murine T cell receptor employs a limited repertoire of expressed V gene segments. Nature (London) **316:** 517–523.
2. **Benyajati, C., N. Spoerel, H. Haymerle, and M. Ashburner.** 1983. The messenger RNA for alcohol dehydrogenase in *Drosophila melanogaster* differs in its 5′ end in different developmental stages. Cell **33:**125–133.
3. **Casnellie, J. E., M. L. Harrison, K. E. Hellstrom, and E. G. Krebs.** 1983. A lymphoma cell line expressing elevated levels of tyrosine protein kinase activity. J. Biol. Chem. **258:**10738–10742.
4. **Chirgwin, J. M., A. E. Pryzybala, R. J. MacDonald, and W. J. Rutter.** 1979. Isolation of biologically active ribonucleic acid from sources enriched in ribonuclease. Biochemistry **18:**5294–5299.
5. **Dente, L., G. Cesareni, and R. Cortese.** 1983. pEMBL: a new family of single stranded plasmids. Nucleic Acids Res. **11:**1645–1656.
6. **Glasebrook, A. L., A. Kelso, R. H. Zubler, J. M. Ely, M. B. Prystowsky, and F. M. Fitch.** 1982. Lymphokine products of cytolytic and noncytolytic alloreactive T cell clones, p. 342–353. *In* C. G. Fathman and F. W. Fitch (ed.), Isolation, characterization, and utilization of T lymphocyte clones. Academic Press,

Inc., New York.
7. **Glynn, J. P., A. R. Bianco, and A. Goldin.** 1964. Studies on induced resistance against isotransplants of virus-induced leukemia. Cancer Res. **24**:502–508.
8. **Hagenbuchle, O., M. Tosi, U. Schibler, R. Bovey, P. K. Wellauer, and R. A. Young.** 1981. Mouse liver and salivary gland α-amylase mRNA differ only in 5′ non-translated sequences. Nature (London) **289**:643–646.
9. **Harris, A. W., A. D. Bankhurst, S. Mason, and N. L. Warner.** 1973. Differentiated function expressed by cultured mouse lymphoma cells. II. Antigen, surface immunoglobulin and a receptor for antibody on cells of a thymoma cell line. J. Immunol. **110**:431–438.
10. **Hyman, R., and V. Stallings.** 1974. Complementation pattern of THY-1 variants and evidence that antigen loss variants "preexist" in the parental population. J. Natl. Cancer Int. **52**: 429–436.
11. **Koga, Y., N. Caccia, B. Toyonaga, R. Spolski, Y. Yanagi, Y. Yoshikai, and T. W. Mak.** 1986. A human T cell-specific cDNA clone (YT16) encodes a protein with extensive homology to a family of protein-tyrosine kinases. Eur. J. Immunol. **16**:1643–1646.
12. **Marth, J. D., R. Peet, E. G. Krebs, and R. M. Perlmutter.** 1985. A lymphocyte-specific protein-tyrosine kinase gene is rearranged and overexpressed in the murine T cell lymphoma LSTRA. Cell **43**:393–404.
13. **Melton, D. A., M. R. Rebagliati, T. Maniatis, K. Zinn, and M. R. Green.** 1984. Efficient in vitro synthesis of biologically active RNA and RNA hybrydization probes from plasmids containing a bacteriophage SP6 promoter. Nucleic Acids Res. **12**:7035–7056.
14. **Müller, R., and D. Müller.** 1984. Co-transfection of normal NIH/3T3 DNA and retroviral LTR sequences: a novel strategy for the detection of potential c-*onc* genes. EMBO J. **3**:1121–1127.
15. **Sanger, F., S. Nicklen, and A. R. Coulson.** 1977. DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. USA **74**:5463–5467.
16. **Schibler, U., O. Hagenbuchle, P. K. Wellaueer, and A. C. Pittet.** 1983. Two promoters of different strenth control the transcription of the mouse alpha-amylase gene Amy-1^a in the parotid gland and the liver. Cell **33**:501–508.
17. **Seidman, J. G., and P. Leder.** 1978. The arrangement and rearrangement of antibody genes. Nature (London) **276**:790–795.
18. **Sharp, P. A.** 1981. Speculations on RNA splicing. Cell **23**: 643–646.
19. **Trevillyan, J. M., Y. Lin, S. J. Chen, C. A. Phillips, C. Canna, and T. J. Linna.** 1986. Human T lymphocytes express a protein-tyrosine kinase homologous to p56^{LSTRA}. Biochim. Biophys. Acta **888**:286–295.
20. **Vogt, M.** 1979. Properties of "mink cell focus-inducing" (MCF) virus isolated from spontaneous lymphoma lines of BALB/c mice carrying Moloney leukemia virus as an endogenous virus. Virology **93**:226–236.
21. **Voronova, A. F., J. E. Buss, T. Patschinsky, T. Hunter, and B. M. Sefton.** 1985. Characterization of the protein apparently responsible for the elevated tyrosine protein kinase activity in LSTRA cells. Mol. Cell. Biol. **4**:2705–2713.
22. **Voronova, A. F., and B. M. Sefton.** 1986. Expression of a new tyrosine protein kinase is stimulated by retroviral promoter insertion. Nature (London) **319**:682–685.