Characterization of the Promoter of Human Leukocyte-specific Transcript 1

A SMALL GENE WITH A COMPLEX PATTERN OF ALTERNATIVE TRANSCRIPTS*

Received for publication, May 31, 2000, and in revised form, August 7, 2000
Published, JBC Papers in Press, August 15, 2000, DOI 10.1074/jbc.M004700200

Xiaofeng Yu and Sherman M. Weissman‡

From the Department of Genetics, Boyer Center for Molecular Medicine, Yale University School of Medicine, New Haven, Connecticut 06510

The gene for the human leukocyte-specific transcript 1 (LST1) encodes a small protein that modulates immune responses and cellular morphogenesis. The LST1 transcripts are expressed at high levels in dendritic cells. Because of the complex splicing pattern, use of alternative 5′-untranslated exons, and a biologically interesting pattern of expression of LST1 mRNA, we studied the human LST1 gene promoter and regulatory elements. We identified an additional upstream 5′-untranslated exon in U937 monocyctic cells. Transient transfection studies demonstrated that the combination of regions from −1363 to −621 with −112 to −54, relative to the translation start codon, produced the highest level of transcripts from among the various constructs tested, but the pattern of transcripts produced was only a subset of those produced from the endogenous gene. DNase I footprinting analysis and electrophoretic mobility shift assays showed that oligonucleotide probes corresponding to three regions, −1171 to −1142 (BI), −1136 to −1111 (BII), and −783 to −751 (BIV), bound proteins in U937 nuclear extracts. Competition and supershift electrophoretic mobility shift assay did not identify any known transcription factors responsible for BI probe binding. These studies suggest that a novel DNA-binding site and interaction of multiple regulatory elements may be involved in mediating the expression of the various forms of LST1 mRNA.

The leukocyte-specific transcript 1 (LST1) gene, the human homologue of the mouse B144 transcript, is located within 15 kilobases upstream of the tumor necrosis factor cluster in the major histocompatibility complex class IV region, a region that contains a number of genes that may play a role in various aspects of stress, inflammation, and immune responses (1–8).

The human LST1 mRNA is most actively transcribed in monocytes and dendritic cells, and its mRNA levels can be enhanced by interferon (IFN)-γ, suggesting a role of LST1 in the immune response (9–12). A recent study showed that this gene product had an inhibitory effect on lymphocyte proliferation (13). Transient transfection studies in our laboratory showed that the LST1 gene product induced extensive thin cytoplasmic extensions in a wide variety of cells in vitro and, in particular, was strongly expressed in dendritic cells in vivo and in vitro. Due to alternative splicing, the LST1 gene in human and mouse encodes multiple transcripts, each 800 nucleotides or less in length. Previous studies on the human LST1 gene showed that this gene consisted of at least eight exons and four introns, including four different alternative 5′-UT exons (termed 1A, 1B, 1C, and 1D, respectively) and four additional exons spanning 2.7 kilobases. The extensive alternative splicing of LST1 mRNA results in high structural diversity of the putative encoded polypeptides (9). The DNA sequences between these 5′-UT exons do not consistently show sequence features resembling promoters.

There are certain unusual features of the relationship between the human LST1 mRNA and the gene sequence. For example, one alternative intron of the LST1 gene begins with the dinucleotide GA rather than the common GT or the occasional GC. Interestingly, each of the four published 5′-UT exons is between 100 and 150 nucleotides in length and lies closely downstream of the sequence CCCAG. Furthermore, the genomic sequence in the vicinity of the 5′-UT exons contains at least two other stretches of sequences about the same length as the published 5′-UT exons in which the nucleotide stretches begin with a CCCAG and end with an apparently conventional 5′-end splice site sequence. Therefore, it is possible that one or more additional 5′-UT exons may exist in the 5′-flanking region of the human LST1 gene.

We carried out the present study to address the following questions. (a) Does the human LST1 gene have an additional 5′-UT exon, which serves as a common initial exon, or does transcription apparently begin separately upstream of each of the closely expressed alternative 5′-UT exons? (b) Where is the promoter in the 5′-flanking region of the human LST1 gene? (c) Which cis-acting elements are involved in the human LST1 gene expression? (d) Can we separate elements responsible for production of subsets of transcription initiation sites and splicing events in the LST1 gene? We have identified an additional 5′-UT exon upstream of the published exon 1A in U937 cells, a human monocytic cell line, indicating that there are at least five alternative 5′-UT first exons (termed 1a, 1b, 1c, 1d, and 1e)

* This work was supported by National Institutes of Health Grant HL63357-01 (to S. M. W.). The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked “advertisement” in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

‡ To whom correspondence should be addressed: Dept. of Genetics, Boyer Center for Molecular Medicine, Yale University School of Medicine, 295 Congress Ave., New Haven, CT 06510. Tel.: 203-737-2283; Fax: 203-737-2286; E-mail: Sherman.Weissman@yale.edu.

© 2000 by The American Society for Biochemistry and Molecular Biology, Inc.

Printed in U.S.A.
that are spliced directly to exon 2 containing the ATG translation start codon. The exon 1b-exon 2 splicing pattern was the most abundant transcript of the human LST1 gene in U937 cells and the major transcription initiation site was located upstream of exon 1b. We have also characterized the human LST1 promoter region and regulatory elements. These studies suggest that a novel DNA-binding site and interaction of multiple regulatory elements may be involved in the regulation of the human LST1 gene expression.

EXPERIMENTAL PROCEDURES

Rapid Amplification of cDNA Ends (5'-RACE)—The total RNA of U937 cells (6 × 10^6) was isolated by using TRIZOL reagent (Life Technologies, Inc.), and digested with 2 μl (10 units/μl) of RNase-free DNase I (Roche Molecular Biochemicals) at 37 °C for 30 min. 2 μg of the total RNA from U937 cells were reverse transcribed into first-strand cDNA using ThermoScript II Reverse Transcriptase (Life Technologies, Inc.) and the reverse primer GSP-RTN (see Table I) designed from the exon 2 region of the human LST1 cDNA. After degrading the original RNA template with RNase H, the first-strand cDNA was extended from its 3'-end using terminal transferase (Roche Molecular Biochemicals) to form a poly(A) tail. The poly(A)-tailed cDNA was then used as a template in 30 cycles of polymerase chain reaction (PCR) at 94 °C for 1 min, 55 °C for 1 min, and 72 °C for 1 min with the antisense primer GSP3 and the sense primer QT and QO (14). The products of PCR amplifications were subjected to a second PCR amplification with the primer GSP42N and the primer QI for another 30 cycles. The PCR products of the second amplification were separated by gel electrophoresis in a 1.2% agarose gel. Purified PCR fragments were cloned into the TA vector in the pGLO243 construct.
Characterization of Human LST1 Promoter

Schematic representation of the human LST1 gene and its multiple transcripts with the 5′-UT alternative exon 1. A, each of the 5′-UT alternative exons 1a (1a, 1b, 1c, 1d, and 1e) can be spliced to exon 2 that contains the translation start codon (ATG) of the human LST1 gene. Exons are represented as boxes and each of the 5′-UT alternative exons is filled with different patterns. Introns are represented as lines between exons. The LST1 mRNA transcribed from each alternative exon 1 is shown on the top. Numbers represent the nucleotide positions (in kilobases) relative to the ATG translation start site, and restriction enzymes used for generating the report constructs are shown on the lower part of the figure. B, RT-PCR of the additional 5′-UT exon 1a of the LST1 gene. The DNA band amplified from RNA of U937 cells by RT-PCR with the sense primer USP2 and the antisense primer GSP3 is indicated by an arrow.

1 min. PCR products were electrophoresed in a 2.0% agarose gel and stained with the TA vector (Invitrogen) for sequencing.

Nuclear Extract Preparation—The nuclear extracts were prepared according to the method of Lee et al. (18). Briefly, after U937 cells were washed with a phosphate buffer, the pellets were resuspended in one packed cell volume of buffer A (10 mM Hepes, pH 8.0, 1.5 mM MgCl2, 10 mM KCI, 1 mM DTT, 0.5 mM phenylmethylsulfonyl fluoride) and placed on ice for 15 min before disruption with a hypodermic syringe with a narrow-gauge total RNAashed. After centrifugation, the crude nuclear pellet was suspended in 0.7 volume of buffer C (20 mM Hepes pH 8.0, 1.5 mM MgCl2, 25% glycerol, 420 mM NaCl, 0.2 mM EDTA, pH 8.0, 1 mM DTT, 0.5 mM phenylmethylsulfonyl fluoride) and incubated on ice by stirring for 30 min. The nuclear debris was centrifuged and the supernatant was dialyzed for 1.5 h in 500 ml of buffer D (20 mM Hepes pH 8.0, 1.5 mM MgCl2, 25% glycerol, 420 mM NaCl, 0.2 mM EDTA, pH 8.0, 1 mM DTT, 0.5 mM phenylmethylsulfonyl fluoride). After centrifugation of the extracts, the protein concentration of the supernatant was determined with a protein assay kit (Bio-Rad) and stored at −80 °C until use.

DNase I Footprinting Analysis—The two segments spanning from −1364 to −545 and from −1001 to −621 plus −113 to −55, which were generated by digesting the plg242SA and the plg14266Ap constructs with SacI-NcoI, were incubated with nuclear extracts and subjected to DNase I footprinting analysis. The 3′-end of the segment was labeled by filling in 5′-overhangs with [γ-32P]dCTP and Klenow polymerase. The unincorporated [γ-32P]dCTP was removed by using Sephadex G-50 columns (Roche Molecular Biologicals). Hybridization of 1 ng of the labeled 24SA probe (405 bp) was used as a control for transfection efficiency. In addition, different cell lines, such as K562, Jurkat, 293T, and HeLa (from our laboratory storage) were also used in the transfection experiments with certain reporter constructs. For the IFN-γ stimulation experiment, U937 cells transfected with the plg2423Ap construct were treated with 50, 100, 200, or 500 units/ml IFN-γ (Life Technologies, Inc.) for 6, 24, and 48 h prior to performing the luciferase assay.

RT-PCR—After U937 cells were transfected with promoter report constructs (e.g. plg2423, plg54212, plg4214, plg2425S, and plg43C8S) for 48 h, the total RNA of U937 cells (6 × 105) was isolated by using TRIZOL reagent (Life Technologies, Inc.) and digested with 2 μl of total RNA containing 1 μg of linearized pCMV-NcoI (Roche Molecular Biologicals) at 37 °C for 30 min to remove traces of contaminating DNA. 2 μg of total RNA from these cells were reverse transcribed into first strand cDNA with ThermoScript II Reverse Transcriptase (Life Technologies, Inc., Life Technologies, Inc.). 2 μl of the first strand cDNA were used in 50 μl of PCR with the sense primers for different forms of the 5′-UT exon 1 (such as USP2 for exon 1a, USP14 for exon 1b, USP7 for exon 1c, USP8 for exon 1d, and USP10 for exon 1e) and the antisense primers (GSP42N) from exon 2. The DNase-treated RNA without reverse transcriptase was used as a control for the PCR reactions. The PCR reaction was performed for 30 cycles at 94 °C for 1 min, 55 °C for 1 min, 72 °C for 1 min. PCR products were electrophoresed in a 2.0% agarose gel and stained with x-ray film.
Fig. 2. Nucleotide sequence of the 5′-flanking region comprising the promoter of the human LST1 gene. Nucleotides are numbered on the left from the ATG translation start codon (indicated by a rhombus) present in exon 2. Potential cis-acting sequences determined by searching the data bases of TRANSFAC and TESS are indicated by dotted lines. The 5′-UT exons are underlined. An additional 5′-UT sequence identified by RT-PCR and the extension sequences of unreported exon 1 identified by 5′-RACE from U937 cells are in italic. The major transcription start site identified by S1 nuclease protection analysis is indicated by a star or triangle if identified by 5′-RACE. The shadowed area is the region contained in the pGL2423Ap construct (−1363 to −621 plus −112 to −54) that has the highest promoter activity. The cells. The nucleotides with and EMSA with nuclear extracts of U937 identified by DNase I footprinting assay moligonucleotides were added to 8 was removed by using Sephadex G-25 columns. 10 ng of radiolabeled polynucleotide kinase. The reactions were stopped by adding 80 KCl, 1 mM EDTA, 20% glycerol, and 1 mM DTT. The reaction mixtures were incubated at room temperature for 20 min. The DNA-protein complexes were separated from the free DNA by electrophoresis through a 6% nondenaturing polyacrylamide gel in TBE buffer.

Electrophoretic Mobility Shift Assay (EMSA)—Oligonucleotides for mobility shift analyses were end-labeled with [γ-32P]ATP and T4 polynucleotide kinase. The reactions were stopped by adding 80 μl of TE buffer and heated at 70 °C for 10 min. The unincorporated [γ-32P]ATP was removed by using Sephadex G-25 columns. 10 ng of radiolabeled oligonucleotides were added to 8 μg of the U937 cell nuclear extracts in a final volume of 25 μl containing 1 μg of nonspecific competitor poly(dI- dC) (Amersham Pharmacia Biotech), 50 mM Tris-HCl, pH 8.0, 50 mM KCl, 1 mM EDTA, 20% glycerol, and 1 mM DTT. The reaction mixtures were incubated at room temperature for 20 min. The DNA-protein complexes were separated from the free DNA by electrophoresis through a 6% nondenaturing polyacrylamide gel in 0.25 × TBE buffer.

Digesting probes were analyzed by electrophoresis on a 6% urea-polyacrylamide gel. The two probes were sequenced with the T7 sequencing kit.

RESULTS

An Additional 5′-UT Alternative Exon 1 of the Human LST1 Gene Was Identified in the U937 Cells—Using RNA from U937 cells that strongly express the LST1 gene, we performed RT-PCR with primers (see Table I) designed to be unique for upstream sequences of the potential first exons and identified an additional exon 1 (termed exon 1a) that resides upstream of the published exon 1A (equivalent to the exon 1b in present study) of the LST1 gene (Fig. 1). Therefore, there are five alternative first exons (termed 1a, 1b, 1c, 1d, and 1e) in the 5′-flanking region of the human LST1 gene (Fig. 2). To determine the 5′-ends of the various mRNA forms, we performed 5′-RACE with RNA from U937 cells. The sequencing results from 36 DNA clones of 5′-RACE products demonstrated that 23 clones (64%) represented the exon 1b-exon 2 splicing pattern (Fig. 3); two clones (5.6%) represented the exon 1c-exon 2 splicing pattern; one clone (2.8%) represented the exon 1e-exon 2 splicing pattern; and 11 clones (27.8%) represented unspliced RNA containing exon 1e, the intron between exon 1e and exon 2. These findings indicate that the transcription initiation site in exon 1b of the LST1 gene is preferentially utilized, and the exon 1b-exon 2 splice form is the most abundant transcript of the human LST1 mRNA.

The 5′-Flanking Region of the Human LST1 Gene Has Several Regions That Contribute to the Strong Promoter Activity—To characterize the LST1 gene promoter, a series of different reporter plasmid constructs that contained various 5′-flanking regions of the LST1 gene were generated in the luciferase reporter plasmid pGL3-Basic. We first prepared several deleted constructs that contained the 5′-flanking regions of each possible transcript of this gene (Fig. 4A). The levels of relative luciferase activities from extracts of transiently transfected U937 cells showed that the pGL2423 construct (−1363 to −54) that contained all 5′-UT alternative exons of the LST1 gene including exons 1a, 1b, 1c, 1d, and 1e had substantially
Characterization of Human LST1 Promoter

34601

Fig. 3. 5'-RACE analysis of the 5'-UT alternative exons of the human LST1 gene in U937 cells. The 5'-RACE experiment was performed using a pair of primers, QI and GSP42N, as described under “Experimental Procedures.” The 5'-RACE products were subcloned and sequenced. 23 out of 36 cDNA products with exon 1b-exon 2 splicing are denoted with arrows. The thick arrows indicate that more than one 5'-RACE product has the same upstream end. The nucleotide sequence of transcript of exon 1b-exon 2 splicing is shown on the top. Numbers are relative to the translation start codon within exon 2. The triangle represents the major transcription start site by the majority of 5'-RACE products (8 clones). The star represents the major transcription start site identified by an S1 nuclease protection assay. The sequence complementary to the antisense primer GSP42N is underlined. The arrow represents the junction of exon 1b and exon 2.

stronger promoter activity than those constructs with longer or shorter 5'-flanking regions of the LST1 gene. The promoter activity of the pGL14266 construct (−1001 to −54) with deletion of the exon 1a-exon 1b genomic region (−1363 to −1001) dropped by a large amount compared with that of the pGL2423 construct, which suggests that the region from exon 1a to exon 1b (−1363 to −1001) is important for the LST1 promoter activity. However, as shown in Fig. 4B, the pGL24ApSt construct that contained the region from exon 1a to exon 1b (−1363 to −545) alone did not demonstrate strong promoter activity. Based on the pGL2423 construct, we further narrowed down the region within exon 1a-exon 2 and found that (a) the apparent promoter activity of the pGL12SN (−1277 to −268) with deletion of the exon 1e-exon 2 region (−267 to −54) was reduced about 50% compared with that of the pGL12SN (−1277 to −54); (b) on comparison of the pGL24ApSt (−1363 to −545) with the pGL2S11N construct (−1363 to −579), the promoter activity of the latter construct that contained the exon 1b-exon 1c region was about 60% higher than that of the pGL24ApSt, indicating that the exon 1b-exon 1c region is also important; (c) comparison of the pGL12SN (−1277 to −268) with the pGL12SN (−1277 to −579) showed that the promoter activity of the latter construct with the deletion of the exon 1c-exon 1e region (−578 to −268) was relatively higher than that of the pGL12SN (Fig. 4B). This result may indicate that the region from exon 1c to upstream of exon 1e contains some silencer elements.

The above deleted constructs did not contain the upstream region of exon 2 that included the splice site at the beginning of exon 2, and this may be a reason why these constructs (e.g., pGL12SN, pGL12SN, pGL2S11N etc.) produced relatively lower levels of luciferase mRNA. Therefore, beginning with the pGL2423 construct, we generated the deleted pGL2423Ap construct (−1363 to −621 plus −112 to −54) that retained the sequences from the region upstream of exon 2. As we expected, the result demonstrated that the pGL2423Ap construct produced more RNA than the other constructs (Fig. 4C). Further deletion of this pGL2423Ap construct remarkably reduced its promoter activity as seen with constructs pGL12SN (−1277 to −621 plus −112 to −54), pGL14266Ap (−1001 to −54 plus −112 to −54), pGL24W13 (−1363 to −811 plus −112 to −54), and pGL24SA (−1363 to −954 plus −112 to −54). Furthermore, the promoter activities of two constructs that contained the exon 1a-exon 1b region (e.g., pGL24ApSt) and exon 1b-exon 1c region (e.g., pGL14266Ap), respectively, were significantly decreased (Fig. 4C). We found that the region from the end of exon 1b to the beginning of exon 1c (−1001 to −621) had an important role in the LST1 gene promoter activity. For instance, the promoter activity of the pGL24SA that deleted this region was reduced from about 70 to 10% compared with the pGL2423Ap (Fig. 4C). Thus, taken together, the data indicate that the combinations of these three regions (−1363 to −954, −954 to −621, and −112 to −54) are necessary for maximal LST1 transcript production.

To address the potential competitive effect of the other promoter regions, we constructed two other constructs, pGL14266Luc− (−1001 to −54) that derived from the pGL14266 construct by deleting the luciferase gene, and pGL74226Luc− (−571 to −54) that derived from the pGL74226 construct by deleting the luciferase gene, to perform co-transfection experiments with the pGL2423Ap construct. The results showed that both these constructs containing some of the alternative promoter regions could reduce the levels of relative luciferase activity of the pGL2423Ap construct from about 69 to 20% (for pGL14266Luc− construct) or to 28% (for pGL74226Luc− construct), indicating that these regions may compete partially for common transcription factors.

The 5'-UT Exon-splicing Pattern of the Human LST1 Gene in Transfected U937 Cells—We transfected U937 cells with certain constructs to identify the splicing pattern in U937 cells. We first transfected U937 cells with the pGL2423 construct that contained the regions encoding the 5'-UT alternative ex-
Fig. 4. Identification of promoter region of the LST1 gene by transient transfection in U937 cells. **Left panel,** structure of the 5'-flanking region and reporter constructs with serial deletions. **Right panel,** the levels of relative luciferase activities of different constructs. The
Characterization of Human LST1 Promoter

34603

FIG. 5. RT-PCR analysis of the splicing pattern of 5′-UT alternative exons in transfected U937 cells. RT-PCR of transfected U937 cells was performed as described under “Experimental Procedures.” A, lanes 1–5: RT-PCR results from U937 cells transfected with the pGL2423 construct by using each alternative exon 1 sense primer with antisense Luc 2 primer (7 bp downstream of GSP42N primer) from the luciferase gene (see Table I) which represent the transcripts of exon 1a-exon 2, exon 1b-exon 2, exon 1c-exon 2, 1d-exon 2, and exon 1e-exon 2 in transfected U937 cells. Lanes 6–10, RT-PCR results from untransfected U937 cells by using each of alternative exon 1 sense primer with antisense GSP42N primer from exon 2 represent the transcripts of exon 1a-exon2, exon 1b-exon 2, exon 1c-exon 2, exon 1d-exon 2, and exon 1e-exon 2 splicing patterns in endogenous U937 cells. The schematic representation of the pGL2423 construct used for transfection and relevant primers used for RT-PCR are shown on the top. B, lane 1, RT-PCR result from U937 cells transfected with the pGL3C8S construct by using the sense USP10S primer from exon 1e and the antisense Luc2 primer from the luciferase gene represents the transcript of exon 1e-exon 2 in transfected U937 cells. M represents 100-bp markers. The schematic representation of the pGL3C8S construct used for transfection and relevant primers used for RT-PCR are shown on the top.

Results represent the mean ± S.E. for three independent transfection experiments with duplicate samples. A, the promoter activities of construct containing 2.4 kilobases of the 5′-flanking region of the LST1 gene (e.g. pGL54212, −2401 to −54) and constructs with the 5′-end serial deletions. B, the promoter activities of constructs with serial deletions from both sides based on the pGL2423 construct (−1363 to −54) that shows significant promoter activity. C, the promoter activities of constructs with serial deletions from both sides but containing part of the region upstream of exon 2 (−112 to −54). The 5′-UT exons are filled with different patterns and the coding region is filled with dark. Enzyme sites used for generating certain constructs are shown on the top. Numbers are nucleotides relative to the translation start codon.
nuclear extracts of U937 cells (Fig. 7, A and B).

We performed gel shift analysis with five double-stranded DNA candidate probes (see Table I) termed BI, BII, BIII, BIV, and BV corresponding to DNase-protected regions (I, II, and III, located within exon 1a-exon 1b region; and IV and V, located within exon 1b-exon 1c region). The EMSA results demonstrated that BI, BII, and BIV bound to nuclear extracts of U937 cells and produced distinct binding patterns, and BI and BII showed stronger shift bands (Fig. 8, A and B). The competition EMSA with nonradioactive competitors (e.g., unlabelled BI, II, and IV) showed that these cold competitors could significantly inhibit the binding of radioactive probes to nuclear extracts (Fig. 8, A, B, and D). By searching the data bases of TRANSFAC and TESS for putative transcription factors, we found some putative DNA-binding sites within these probe regions, for example, C/EBPβ and GATA sites in BI probe, PU.1 and SOX5 sites in BII probe, and an SP1 site in BIV probe. However, EMSA competition assays with the cold consensus C/EBPβ or PU.1 oligonucleotides could not inhibit the BI or the BII probe binding to U937 nuclear extracts (Fig. 8, A and B). The protein-DNA complexes of the labeled PU.1 and C/EBPβ consensus probes with the U937 nuclear extracts had a different electrophoretic motility from that of the BI and the BII probes (Fig. 8C). Attempts at supershift analyses with either anti-PU.1 rabbit antibody or anti-SRY goat antibody that can cross-react with SOX5 protein provided evidence that PU.1 and SOX5 transcription factors were not involved in DNA-protein binding (Fig. 8E). EMSA with mutated probes at the transcription factor-binding sites of PU.1 and SOX5 showed that the binding patterns were similar to that of the original probes (Fig. 9A). These findings suggest that a novel DNA-binding site in the promoter region may be associated with human LST1 gene regulation. Since the two 1mBII and 2mBII probes, which separated the putative PU.1- and SOX5-binding sites, did not demonstrate significant gel shifts (Fig. 9B), we assumed that the nucleotides between these two binding sites may be related to DNA-protein binding.

In order to verify this DNA-protein binding site, we performed EMSA by introducing a “CC” deletion mutation (6mBII) in the BII probe sequence that lies between the PU.1- and SOX5-binding sites. The result showed that this mutant probe could not effectively bind to the U937 nuclear extracts (Fig. 9A). When the probe BI was divided into two parts (2mBI and 3mBU) or narrowed down (4mBI) from both ends of this probe, we found that both the divided probes and the narrowed probe demonstrated a significant gel shift as the original BI did (Fig. 9C), and the competition EMSA with these cold mutant probes (2mBI, 3mBI, and 4mBI) could inhibit these bindings (data not shown). The other mutant probes (5mBI and 6mBI) did not effect the formation of the DNA-protein complexes (Fig. 9D). These findings indicate that multiple regulatory elements may be involved in LST1 gene expression. Also, we found the cold consensus SP1 oligonucleotide could inhibit BIV probe binding to U937 nuclear extracts but the cold consensus PU.1 oligonucleotide could not (Fig. 8E).

The Novel DNA-binding Site in the LST1 Gene Promoter Region Is Necessary for Gene Expression—To determine the functional significance of the DNA-binding site found in our EMSA result, we made the mutant construct termed pGL2423mAp based on the pGL2423Ap construct which changed the nucleotides at positions −1126 through −1123 (5’-ACTTCCTCTCCTACTAAATGCGTGGG-3’) from “CCTA” between the PU.1 and SOX5 sites to “AGCT” within the BII probe sequence. The promoter activity of this mutant construct was significantly reduced compared with the pGL2423Ap control in U937, K562, and Jurkat cells although the levels of the promoter activities were apparently different in these different cell lines (Fig. 10, A-C). These data confirm that a novel DNA-binding site in the human LST1 promoter region plays an important role in regulation of this gene expression.

**DISCUSSION**

Based on the observation that each of the four known 5’-UT exons of the human LST1 gene lies closely downstream of the sequence CCCAG, we anticipated that other stretches of sequences that are preceded by CCCAG and followed by a plausible GT containing 5’-end splice site may also serve as 5’-UT exons. The template regions for the initial exons are close to one another and are not all preceded by obvious transcription factor-binding sites, so it seemed unlikely that a conventional promoter was located separately immediately upstream of each exon. We therefore also tried to identify an upstream common exon by RT-PCR with primers from exon 2 that contains the translation start codon and from regions within each of these putative exons. This uncovered an additional 5’-UT alternative exon (termed exon 1a) about 180 nucleotides upstream of the
previously published exon 1A of the human *LST1* gene (Figs. 1–2). However, our 5’-RACE analysis using RNA from U937 cells did not show a common initial exon preceding the 5’-UT alternative exons. We sequenced 36 cDNA clones of our 5’-RACE products and found that the majority of cDNA clones (64%) was transcripts from exon 1b to exon 2 (Fig. 3). The frequency of exon 1b-exon 2 splicing representation in the isolated cDNA clones indicates that the transcription initiation site in exon 1b of the *LST1* gene is preferentially used in U937 cells. We also analyzed the EST data base of *LST1* cDNAs focused on the 5’-end region and found that four cDNA from an adult brain showed exon 1b-exon 2 splicing. One cDNA from U937 cells stimulated with IFN-γ showed exon 1c-exon 2 splicing. 7 of 8 cDNA from fetal liver, spleen, heart, and placenta showed exon 1e-exon 2 splicing, but one cDNA from a postnatal brain contained the intron between exons 1e and exon 2. Combining the EST data with our 5’-RACE results, there is a suggestion that the 5’-UT splicing pattern of the human *LST1* gene may have tissue specificity, that is, the exon 1b-exon 2 splicing or exon 1e-exon 2 unspliced form appears to be abundant in U937 cells or adult tissues, while exon 1e-exon 2 splicing is more easily found in the fetal tissues, such as fetal liver, spleen, placenta and heart, etc. Also, de Baey’s (9) results of RT-PCR showed that the major *LST1* transcripts initiated from exon 1A (equivalent to our exon 1b) were detected in lymphocytes (e.g. PBMC, CD4+ T cell clones, CD8+ T cell clones, and B lymphoblastoid cell line LG2), in monocyctic cell lines (e.g. U937 and Mono Mac 6), and also in many human tissues (e.g. lung, tonsil, thymus, placenta, kidney, spleen, and liver). However, a high level of *LST1* mRNA was detected only in macrophage and monocyctic cell lines (3, 6). The previous RT-PCR data from our group showed that the expression pattern of the *LST1* gene in U937 cells was similar to that found in human monocyctic and dendritic cells derived from human blood (data not shown). Thus, we undertook the present study on the *LST1* gene promoter in U937 cells and speculated that the region from the end of exon 1a to upstream of exon 1b probably contained the major *LST1* promoter. However, our transient transfection studies did not fully support this prediction. The pGL24ApSt construct containing only this region (~1363 to ~954) did not show a high level of promoter activity (Fig. 4B).

![Fig. 7. DNase I footprinting analysis of the human LST1 promoter. The DNA fragments used for footprinting analysis were prepared by digestion of the pGL24SA construct (A) or the pGL14266Ap construct (B) with ScaI-NcoI and radiolabeled by filling in the 5’-overhangs with [γ-32P]dCTP and Klenow enzyme. DNase I footprinting analysis was performed in the absence of nuclear extracts (A, lanes 1 and 4; B, lanes 1 and 5) or in the presence of nuclear extracts from U937 cells (A, lanes 2 and 3; B, 2–4). The DNA fragments used for footprinting analysis are shown on the top, and the boxed areas represent the different potential protected regions termed as BI, BII, BIII, BIV, and BV. Numbers represent the nucleotides relative to the translation start codon.](image-url)
parent promoter activity (Fig. 4C). However, the pGL2423Ap construct combining three genomic regions (−1363 to −1001, −1001 to −621, and −112 to −54) together formed an active promoter. In addition, transfection experiments with some reporter constructs (e.g., pGL54212, pGL4214, pGL2423, pGL14266, and pGL2423Ap) in different cell lines, such as K562, Jurkat, 293T, and HeLa cells, showed much weaker promoter activities in these cells than in U937 cells (data not shown).

Holzinger's data showed that the human LST1 gene expression was enhanced by stimulation with 200 units/ml IFN-γ for
probe sequences contain a putative PU.1 and C/EBP factors showed that this region contained a number of putative flanking region by searching the data bases of transcription. The 5' were designed so that transcription started at approximately 48 h. To address the question if the IFN-γ treatment would regulate the LST1 promoter, we have performed the transient transfection experiments with the pGL2423Ap construct in U937 cells treated with IFN-γ. Our result did not show that the LST1 promoter activity was increased by IFN-γ treatment with different doses or for different times. This may be because an IFN-γ-activated site present in intron 3 of the LST1 gene is not included in the pGL2423Ap construct (6), or the regulation of LST1 gene expression by IFN-γ may be due to post-transcriptional mechanisms.

In addition to a number of alternative splice forms that occur within the coding region of the LST1 gene in human monocytes, there are at least five 5'-UT alternative exons spliced to exon 2 in U937 cells (9, 13). To understand the splicing patterns of the 5'-UT alternative exons in the transfected cells or if the 5'-UT exon splicing would influence LST1 gene expression, we used an RT-PCR approach to analyze the cDNA in U937 cells transfected with our reporter constructs. Only exon 1a-exon 2 and exon 1b-exon 2 splicing that occurred in the transcripts from the different constructs tested were found in transfected U937 cells (Fig. 5A). We did not identify any exon 1e-exon 2 splicing in the mRNA derived from constructs transiently transfected into U937 cells even when the constructs were designed so that transcription started at approximately the 5' end of exon 1e (Fig. 5B).

Analysis of the 2.4-kilobase nucleotide sequence of the 5'-flanking region by searching the data bases of transcription factors showed that this region contained a number of putative regulatory cis-acting elements (Fig. 2). For example, BII and BI probe sequences contain a putative PU.1 and C/EBPβ site, respectively. PU.1 is a myeloid-specific transcription factor that has a role in hematopoietic cell differentiation, proliferation, and apoptosis (19–26). C/EBPβ is expressed in myelomonocytic cells and its binding sites have been detected within regulatory and lipopolysaccharide-responsive elements of genes expressed in monocytes and macrophages (27–30). We focused on the BII and BI probes and tried to identify which transcription factors might be involved in the regulation of the LST1 gene expression. However, the EMSA with mutant probes at the putative PU.1, SOX5, C/EBPβ, and GATA-binding sites did not show any evidence that these DNA-binding sites were related to these factors (Figs. 8A and 9, A and D).

EMSA with radiolabeled or cold consensus C/EBPβ and PU.1 probe, or the supershift EMSA with anti-PU.1 and SRY antibodies did not show that these transcription factors were responsible for their observed gel shifts (Fig. 8, A–C). The putative C/EBPβ site is predicted only by searching TRANSFAC but not by TESS, and C/EBPβ is unlikely to be the factor binding to the BI probe. Evidence from another study (9) also showed that C/EBPβ was not involved in LST1 gene expression. The PU.1 site (GGAA) is located at the extreme 5’-end of the BII probe and it needs additional bases on its 5’ side to form an appropriate protein-binding motif (31). When we narrowed down the BII probe (5'-ACTTCCTCTCTCTAATGCTGGG-3') from both sides, we found the mutant 5mBII probe (5'-TTTCCTCTCTCAAATGCTGG-3') also significantly bound to U937 nuclear extracts (Fig. 9A). However, when we divided the BII probe into two parts sequences (such as 1mBII and 2mBII) and performed EMSA, neither of these probes bound to U937 nuclear extracts (Fig. 9B). Taken together, these data indicate that there may be a novel DNA-binding site present between the putative PU.1 and SOX5 sites within the BII probe sequence. Furthermore, EMSA with 6mBII probe and transfection study with the pGL2423Ap construct that was mutated in the putative DNA-binding site confirmed that this binding site plays an important role in DNA-protein binding and LST1 promoter activity (Figs. 9A and 10).

In summary, we have identified an additional alternative exon 1 (termed exon 1a) residing upstream of the published exon 1A. The exon 1a-exon 1c genomic region (−1363 to −621) combined with sequences upstream of exon 2 (−112 to −54) of the human LST1 gene had the highest promoter activity. The exon 1b-exon 2 splice form was the dominant transcript of the human LST1 gene in U937 cells, and the major transcription initiation site was located at 205 base pairs position upstream of exon 1b. Transient transfection studies in U937 cells did not produce transcripts with the unusual splicing pattern seen in a portion of the endogenous transcripts. A novel DNA-binding site in the human LST1 promoter region and the interaction of multiple transcription factors may play an important role in regulating the LST1 gene expression.

Acknowledgments—We thank Daniel G. Tenen, Fengzhi Li, Arati Khanna-Gupta, and all members, especially Yao-Hui Chen, in Weiss-
Characterization of Human LST1 Promoter

REFERENCES

1. Gruen, J. R., and Weissman, S. M. (1997) Blood 90, 4252–4265
2. Campbell, R. D., and Trowsdale, J. (1997) Immunol. Today 16, 43
3. Nalabolu, S. R., Shukla, H., Naillur, G., Parimoo, S., and Weissman, S. M. (1996) Genomics 31, 215–222
4. Browning, J. L., Ngam-ek, A., Lawton, P., DeMarinis, J., Tizard, R., Chow, E. P., Hession, C., O’Brine-Greco, B., Foley, S. F., and Ware, C. F. (1993) Cell 72, 847–856
5. de Baey, A., Holzinger, I., Scholz, S., Keller, E., Weiss, E. H., and Albert, E. (1995) Hum. Immunol. 42, 9–14
6. Holzinger, I., de Baey, A., Messer, G., Kack, G., Zwierzina, H., and Weiss, E. H. (1995) Immunogenetics 42, 315–322
7. Iris, F. J. M., Bougueleret, L., Prieur, S., Caterina, D., Primas, G., Perrot, V., Jurka, J., Rodriguez-Tome, P., Claverie, J. M., Dausslet, J., and Cohen, D. (1993) Nat. Genet. 3, 137–145
8. Tague, I., Shen, F. W., Steinmetz, M., and Boyse, E. A. (1987) Immunogenetics 26, 378–380
9. de Baey, A., Fellerhoff, B., Maier, S., Martinozzi, S., Weidle, U., and Weiss, E. H. (1996) Hum. Immunol. 47, 65–68
10. de Baey, A., Fellerhoff, B., Maier, S., and Weiss, E. H. (1996) Immunogenetics 42, 315–322
11. Holzinger, I., de Baey, A., Messer, G., Kack, G., Zwierzina, H., and Weiss, E. H. (1995) Immunogenetics 42, 315–322
12. Asano, H., Kobayashi, T., Uchida, K., Hayashi, S., Yokoyama, I., Inoko, H., and Takagi, H. (1997) Tissue Antigens 50, 484–487
13. Holzinger, I., Eibl, B., Pauly, M., Giesser, U., Hentges, F., Auer, B., Pall, G., Schratzberger, P., Niederwieser, D., Weiss, E. H., and Zwierzina, H. (2000) J. Immunol. 164, 3169–3176
14. Frohman, M. A., Dush, M. K., and Martin, G. R. (1988) Proc. Natl. Acad. Sci. U. S. A. 85, 8996–8902
15. Ho, S. N., Hunt, H. D., Horton, R. M., Pullen, J. K., and Pease, L. R. (1989) Gene (Amst.) 77, 51–58
16. Kusters-van Someren, M., Kishi, K., Lundell, T., and Gold, M. H. (1995) Biochemistry 34, 10620–10627
17. Bradford, M. M. (1976) Anal. Biochem. 72, 248–254
18. Lee, K. A. W., Zerovitz, K., and Akusjarvi, G. (1994) Cell Biology: A Laboratory Handbook, pp. 668–670, Academic Press, Inc., New York
19. Klemm, M. J., McRitchie, S. R., Celada, A., Van Beveren, C., and Maki, R. A. (1990) Cell 61, 113–124
20. Wasylyk, B., Hahn, S. L., and Giovane, A. (1993) Eur. J. Biochem. 211, 7–18
21. Oikawa, T., Yamada, T., Kiikara, N. F., Yamamoto, H., Kindoh, H., Hitomi, Y., and Hashimoto, Y. (1999) Cell Death Differ. 6, 599–608
22. McRitchie, S. R., Henkel, G. W., and Maki, R. A. (1999) J. Leukoc. Biol. 66, 727–732
23. Zhang, P., Behre, G., Pan, J., Iwama, A., Wara-aswapati, N., Radomska, H. S., Auro, P. E., Tenen, D. G., and Sun, Z. J. (1999) Proc. Natl. Acad. Sci. U. S. A. 96, 8765–8770
24. Nerlov, C., Querfurth, E., Kizousesa, H., and Graf, T. (2000) Blood 95, 2543–2551
25. Borras, F. E., Lliberas, J., Maki, R. A., and Celada, A. (1995) J. Biol. Chem. 270, 24385–391
26. Rekhtman, N., Radparvar, F., Evans, T., and Skoulchi, A. I. (1999) Genes Dev. 13, 1398–1411
27. Wedel, A., and Ziegler-Heitbrock, H. W. L. (1995) Immunobiology 193, 2–4
28. Akira, S., Ishihuk, H., Sugita, T., Tanabe, O., Kinoshita, S., Nashi, Y., Nakajima, T., Hirano, T., and Kishimoto, T. (1990) EMBO J. 9, 1897–1906
29. Chen, P. L., Riley, D. J., Chen-Kiang, S., and Lee, W. H. (1996) Proc. Natl. Acad. Sci. U. S. A. 93, 465–469
30. Brasier, A. R., and Kumar, A. (1994) J. Biol. Chem. 269, 10341–10351
31. Matsui, K., Xiao, S., Fine, A., and Ju, S. T. (2000) J. Immunol. 164, 3062–3068