MOLECULAR AND CELLULAR BIOLOGY, Nov. 1994, p. 7111–7123
0270-7306/94/$04.00+0

Vol. 14, No. 11

The Lymphoid Transcription Factor LyF-1 Is Encoded by Specific, Alternatively Spliced mRNAs Derived from the *Ikaros* Gene

KYUNGMIN HAHM,[1] PATRICIA ERNST,[1] KIERSTEN LO,[1] GREGG S. KIM,[1] CHRISTOPH TURCK,[2] AND STEPHEN T. SMALE[1]*

Howard Hughes Medical Institute, Molecular Biology Institute, and Department of Microbiology and Immunology, University of California School of Medicine, Los Angeles, California 90024-1662,[1] and Howard Hughes Medical Institute and Department of Medicine, University of California School of Medicine, San Francisco, California 94143[2]

Received 24 May 1994/Accepted 2 August 1994

The lymphocyte-specific DNA-binding protein LyF-1 interacts with a critical control element in the terminal deoxynucleotidyltransferase (TdT) promoter as well as with the promoters for other genes expressed during early stages of B- and T-cell development. We have purified LyF-1 and have obtained a partial amino acid sequence from proteolytic peptides. The amino acid sequence suggests that LyF-1 is a zinc finger protein encoded by the *Ikaros* gene, which previously was implicated in T-cell development. Recombinant Ikaros expressed in *Escherichia coli* bound to the TdT promoter, and antisera directed against the recombinant protein specifically blocked the DNA-binding activity of LyF-1 in crude extracts. Further analysis revealed that at least six distinct mRNAs are derived from the *Ikaros/LyF-1* gene by alternative splicing. Only two of the isoforms possess the N-terminal zinc finger domain that is necessary and sufficient for TdT promoter binding. Although both of these isoforms bound to similar sequences in the TdT, λ_5, V_{preB}, and *lck* promoters, one isoform contains an additional zinc finger that resulted in altered recognition of some binding sites. At least four of the Ikaros/LyF-1 isoforms were detectable in extracts from B- and T-cell lines, with the relative amounts of the isoforms varying considerably. These data reveal that the LyF-1 protein is encoded by specific mRNAs derived from the alternatively-spliced *Ikaros* gene, suggesting that this gene may be important for the early stages of both B- and T-lymphocyte development.

The hematopoietic stem cell is a self-renewing progenitor of multiple cell lineages, including the cell lineages needed for an effective immune response (1). At an early stage of hematopoiesis, molecular events occur that commit a fraction of the multipotential cells to the B- and T-lymphoid lineages. To gain insight into the molecular control of the earliest stages of B- and T-cell development, we are studying the regulatory mechanisms for proteins involved in the immunoglobulin (Ig) and T-cell receptor gene rearrangements. These recombination events are thought to be initiated at very early stages of B- and T-cell development. Therefore, the proteins which carry out the rearrangements must be activated at an even earlier stage of lymphopoiesis. Three developmentally regulated proteins are known to be of central importance to the gene rearrangement process in both lymphoid pathways: RAG1 (40), RAG2 (33), and terminal deoxynucleotidyltransferase (TdT) (22). RAG1 and RAG2 are sufficient for activating recombination when transfected into fibroblasts (21, 33) and are likely to be essential components of the recombinase (30, 42). TdT is required for inserting nucleotides (N regions) at the coding joints during recombination, thereby increasing diversity of the Ig and T-cell receptor repertoires (10, 21, 23, 24). Although not essential for recombination, TdT expression is tightly restricted to early lymphocytes (5, 25). Because TdT is a component of the recombination machinery and is expressed exclusively in early B and T cells, an analysis of its regulation provides an appropriate starting point for studying the earliest stages of lymphopoiesis.

Previous studies of the murine TdT promoter have focused on two important control elements. The first element, called an initiator (Inr), functionally replaces the TATA box, in that it directs the location of the TdT transcription start site and is sufficient for basal transcription of the TdT gene (43). The Inr, which overlaps the start site, functions in all cell types and does not appear to contribute directly to lymphocyte specificity. However, the presence of an Inr instead of a TATA box may be involved indirectly in appropriate TdT regulation. For example, experiments from our laboratory and others have suggested that Inr-mediated transcription proceeds through a different rate-limiting step than does TATA-mediated transcription (6, 50). Thus, the presence of an Inr in the TdT promoter may be required for an appropriate response to specific upstream activators that stimulate transcription through a specific rate-limiting step.

A second control element, called D′, is located approximately 60 bp upstream of the Inr and contains nucleotides critical for transcription in lymphocytes (7, 28). Multiple proteins expressed in B and T cells interact with this element, but it is not known whether one of these proteins or a combination of proteins is needed for D′ function. Two Ets family members, Ets-1 (49) and Fli-1 (3), bind with high affinity to the D′ element and are expressed in TdT-expressing cell lines, but neither Ets-1 nor Fli-1 expression is restricted exclusively to lymphocytes (3, 4, 37). A distinct B- and T-cell-specific protein called LyF-1 binds to two overlapping sites in the TdT D′ element (7, 28). The proximal LyF-1-binding site closely coincides with the Ets protein-binding site and with the sequences needed for lymphocyte-specific transcription (7).

To study the role of LyF-1 in TdT expression and in early B- and T-cell development, we have purified the LyF-1 protein and have obtained a partial amino acid sequence. The amino acid sequence revealed that LyF-1 is a zinc finger protein encoded by a gene called *Ikaros*. This gene was recently isolated from a λgt11 library screened with a portion of the

* Corresponding author. Phone: (310) 206-4777. Fax: (310) 206-3800.

CD3δ enhancer and was thought to be expressed specifically in the T-cell lineage (9). The results presented below suggest that the *Ikaros/LyF-1* gene may regulate gene expression in both the B- and T-lymphocyte lineages. Furthermore, the *Ikaros/LyF-1* gene was found to generate at least six mRNA isoforms which encode proteins with variable zinc finger domains and DNA-binding properties.

MATERIALS AND METHODS

Purification and amino acid sequence analysis of LyF-1. LyF-1 was purified from RLm11 (murine TdT^+ thymoma) nuclear extracts as described by Lo et al. (28), with minor modifications. RLm11 cells were grown to a density of 1.5×10^6/ml in 2-liter roller bottles. Nuclear extracts from 9×10^{10} cells were fractionated on a 50-ml heparin-agarose column in HGED buffer (20 mM *N*-2-hydroxyethylpiperazone-*N*′-2-ethanesulfonic acid [HEPES; pH 7.9], 20% glycerol, 1 mM EDTA, 1 mM dithiothreitol, 0.1 mM phenylmethylsulfonyl fluoride). Following a 0.075 M KCl wash, 40 to 60 mg of total protein typically eluted with 0.2 M KCl. The 0.2 M KCl pool was diluted to 0.1 M KCl with an equal volume of HGED buffer containing 4 mg of poly(dI-dC) (Pharmacia) and 0.03% Nonidet P-40 (NP-40). (The final concentration of 0.015% NP-40 was lower than the 0.05% used previously [28]. This lower concentration appeared to prevent protein aggregation and resulted in the elimination of the 50-kDa band from the purified preparations.) Following a 15-min incubation on ice, precipitated proteins were removed by ultracentrifugation at 25,000 rpm for 15 min in an SW28 rotor (Beckman). Supernatants were loaded onto a 4.5-ml sequence-specific DNA affinity column (20), which contained multimers of the oligonucleotide 5′-GATCCATTTTGGGAGAAAG-3′ and its complementary sequence. The column was washed with 3 to 5 column volumes of HGED.1 (0.1 M KCl) containing 0.015% NP-40. Bound proteins were eluted sequentially with 3 to 5 column volumes each of HGED.3 (0.3 M KCl) and HGED.6 (0.6 M KCl) containing 0.015% NP-40. The second pass over the DNA affinity column was performed with the 0.6 M KCl eluate from the first pass. The first-pass 0.6 M KCl eluates were pooled and diluted to 0.1 M KCl with HGED.0 containing 0.015% NP-40. The diluents were loaded onto a 1-ml DNA affinity column, which was then washed with 3 to 5 column volumes of HGED.1. Proteins were eluted with 3 to 5 column volume of HGED.3 and HGED.6 containing 0.015% NP-40. Fractions were tested for gel shift and DNase I footprinting activity; proteins were visualized by sodium dodecyl sulfate (SDS)-polyacrylamide gel electrophoresis (PAGE) (12% gel) followed by silver staining.

To isolate the 60-kDa protein for microsequencing, first-pass 0.6 M eluates were analyzed by SDS-PAGE, and appropriate fractions were pooled. After nine runs of the 4.5-ml DNA affinity column, approximately 15 μg of 60-kDa protein was obtained. The pooled proteins were precipitated by 20% trichloroacetic acid (Fisher), run on three separate preparative SDS–10% polyacrylamide gels, and transferred to a polyvinylidene difluoride membrane (Bio-Rad), using 10 mM 3-[cyclohexylamino]-1-propane sulfonic acid (CAPs) (pH 11.0) and 10% methanol as the transfer buffer. The membrane was stained with Ponceau S, and the 60-kDa band was excised and digested with trypsin (sequencing grade; Boehringer Mannheim Biochemicals). The tryptic peptides were separated by using a reverse-phase high-pressure liquid chromatography column (C18; 2.1 mm by 25 cm; Applied Biosystems) equilibrated with 0.1% trifluoroacetic acid (Sequenal grade; Pierce Chemical) in water at a flow rate of 0.25 ml/min. The peptides were eluted with a 90-min gradient of 0.1% trifluoroacetic acid–90% acetonitrile. Amino-terminal sequence analysis of the tryptic peptides was carried out on a model 470A gas-phase sequencer equipped with a model 120A on-line phenylthiohydantoin-amino acid analyzer and a model 900A data analysis station, using version 1.20 (Applied Biosystems).

Isolation of multiple *Ikaros/LyF-1* cDNAs. To isolate *Ikaros/LyF-1* cDNAs by PCR, first-strand cDNA was generated from total RLm11 RNA, using oligo(dT) (Promega) as a primer as described by Sambrook et al. (39). Each of the PCRs described below was performed with the first-strand cDNA obtained from 1 μg of RNA. Reactions were carried out for 35 cycles with the following conditions: denaturing at 94°C for 1 min, annealing at 55°C for 1 min, and extension at 72°C for 2 min. Sense primers contained a *Bgl*II site, and antisense primers contained an *Eco*RI site.

To obtain full-length *Ikaros/LyF-1* cDNAs, the following primers corresponding to the reported *Ikaros* cDNA (9) were used: 5′-GATAGATCTATGGATGTCGATGAGGGTCAA GAC-3′ (translation start site sense primer) and 5′-GATGAA TTCTTAGCTCAGGTGGTAACGATGCTC-3′ (translation stop site antisense primer). A PCR product of approximately 1.3 kb was isolated and inserted into pSP72 (Promega) digested with *Bgl*II and *Eco*RI. The nucleotide sequence of this cDNA was determined with a Sequenase kit (U.S. Biochemical) and was found to encode isoform IV.

To isolate additional Ikaros isoforms, the full-length cDNA obtained by PCR was used to screen nitrocellulose filters (a gift of Owen Witte's laboratory, University of California, Los Angeles) containing plaques from a 70Z/3 pre-B-cell cDNA library (11). Phage DNAs from six positive plaques were isolated and converted to a Bluescript vector (Stratagene), and the cDNA sequences were determined. These cDNAs corresponded to isoforms I and V.

To eliminate possible errors introduced by PCR into the C-terminal half of isoform IV, a *Pfl*MI-*Pst*I fragment from a cDNA isolated from the 70Z/3 library (containing the C-terminal half of the cDNA and a portion of the 3′ untranslated sequence) was substituted for the same fragment in the form IV cDNA.

To isolate additional cDNAs, N-terminal fragments were generated by PCR from the RLm11 RNA and inserted into pSP72 cleaved with *Bgl*II and *Eco*RI. The primers used were the translation start site sense primer (described above) and 5′-GATGAATTCGTCCATCACGTGGGATGTCATCAT-3′, an antisense primer which hybridizes at the 5′ end of the C-terminal sequence (CTS). The various N-terminal PCR clones were sequenced, and a *Bgl*II-*Pfl*MI fragment from each of these cDNAs was fused to the C-terminal fragment isolated from the phage library. Numerous attempts to detect variations in the C-terminal half of the *Ikaros/LyF-1* cDNA revealed no evidence of heterogeneity.

Recombinant Ikaros/LyF-1 proteins. To prepare glutathione *S*-transferase (GST)–Ikaros fusion plasmids, the full-length cDNAs were cleaved with *Bgl*II and *Ssp*I and inserted into the *Bam*HI-*Sma*I site of pGEX2T (Pharmacia). To prepare plasmids expressing GST fused to either the N-terminal zinc fingers from isoform V or the common C-terminal fragment, PCRs were carried out with the following primer pairs: for the N terminus, 5′-GATAGATCTATGGGTGAACGGCCTTTC CAGTGCAAC-3′ and 5′-GATGAATTCTCAAAGAAATTT CTGAGGCATAGAGCT-3′; for the C terminus, 5′-GATAG ATCTGACAAGTGCCTGTCAGACATGCCC-3′ and 5′-GA TGAATTCTTAGCTCAGGTGGTAACGATGCTC-3′. The PCR products were inserted into pGEX2T cleaved with *Bam*HI and *Eco*RI.

pGEX fusion constructs were introduced into the SCS-1 strain of *Escherichia coli* (Stratagene). Bacterial lysates were prepared and fusion proteins were purified by standard procedures (Pharmacia), except that cells were induced with 1 mM isopropylthiogalactopyranoside (IPTG) for 2 h in the presence of 100 μM $ZnCl_2$. In addition, Triton X-100 was omitted from the lysis buffer, and all buffers contained 10 μM $ZnCl_2$. Fractions containing the purified fusion proteins were pooled, diluted 1:1 with HGED.1 containing 10 μM $ZnCl_2$, and stored at −80°C.

Ikaros antiserum and IgG purification. To generate anti-Ikaros (α-Ikaros) serum, the GST fusion protein containing the C-terminal Ikaros/LyF-1 region (amino acids 197 to 431) was purified as described above. The purified fusion protein was dialyzed against deionized water, and 100 μg of the protein was used to immunize a rabbit (Caltag, Healdsburg, Calif.). To purify IgG from the preimmune and α-Ikaros serum, protein A-Sepharose (Pharmacia) chromatography was performed as described by Harlow and Lane (16). The purified IgG from both the preimmune and immune sera contained 1 mg of protein per ml as determined with Coomassie protein assay reagent (Pierce).

Gel shift assays. To prepare ^{32}P-labeled probes containing the TdT D, TdT D′, and CD3δ elements, the following oligonucleotides and their complementary sequences were inserted into the *Bam*HI site of the pSP72 vector (Promega): TdT D, 5′-GATCCATTTTGGGAGAAAG-3′; TdT D′, 5′-GATCCAGTGAGACATTCCTTCAGCAGGAAGTTGTCTG-3′; and CD3δ, 5′-GATCCAGAAGTTTCCATGACATCATGAATGGGGGTGGCAGAGAG-3′. These plasmids were cleaved with *Bgl*II and *Xho*I and 5′ end labeled as described by Ernst et al. (7). The probe used as a negative control was the *Bgl*II-*Xho*I fragment of pSP72.

Probes containing mutations in the TdT D site were prepared as described previously (19). The mutant sequences are as follows (mutated nucleotides are underlined): m10, 5′-TGATCAGGTCCTTTCTCCCAAACTGGAGCTCTTGACACAAAGAACAGTGAACT-3′; m12, 5′-TGATCAGGTCCTTTCTCCCAAAATTGAGCTCTTGACACAAAGAACAGTGAACT-3′; m20, 5′-TGATCAGGTCCTTTCTCCCAAAAAGGAGCTCTTGACACAAAGAACAGTGAACT-3′; and m914, 5′-TGATCAGGTCCTTGAGCTCAAAATGGAGCTCTTGACACAAAGAACAGTGAACT-3′.

Binding reactions were performed as described by Lo et al. (28) except that incubation was at room temperature and the incubation buffer contained 4 μM $ZnCl_2$. Each reaction contained 10,000 cpm of ^{32}P-labeled probe. Samples were analyzed as described by Ernst et al. (7). For IgG blocking experiments, RLm11 nuclear extracts were preincubated with the purified preimmune or α-Ikaros IgG for 15 min in the binding buffer prior to addition of the probe.

DNase I footprinting assays. DNase I footprinting reactions were performed as described by Lo et al. (28) except that the binding reaction mixtures contained 2 μM $ZnCl_2$. ^{32}P-labeled probes were prepared with the TdT, λ_5, V_{preB}, and *lck* promoters as described previously (28). For the zinc chelating experiment, nuclear extracts were preincubated in 1 mM phenanthroline for 10 min and 1 to 10 μl of 1 mM zinc chloride was added. Five minutes after the addition of $ZnCl_2$, probe was added to the reaction. The IgG block experiments were identical to the standard footprinting assay except that nuclear extracts were preincubated with the purified IgG in the reaction buffer at room temperature for 15 min before addition of the probe.

Western blot (immunoblot) analysis and in vitro transcription and translation. Nuclear extracts from HAFTL-A, 38B9, 40E4, PD31, 70Z/3, S194, R1.1, and RLm11 cells were prepared as described by Lo et al. (28). For NIH 3T3 cells, a mini-nuclear extract protocol (26) was used. Nuclear extracts (15 μg) were separated on an SDS–10% polyacrylamide gel and transferred to a nitrocellulose membrane (Schleicher & Schuell) as described above. The membrane was incubated for 1 h with the blocking buffer (5% Carnation fat-free dried milk, 50 mM Tris HCl [pH 7.4], 0.5 M NaCl). Then, the membrane was incubated for 1 h each with α-Ikaros serum (1:500 dilution), with biotinylated anti-rabbit IgG (BioGenex), and with peroxidase-conjugated streptavidin (BioGenex) in an antibody dilution buffer (50 mM Tris HCl [pH 7.4], 0.5 M NaCl). Between each incubation, the membrane was washed rapidly three times with wash buffer (50 mM Tris HCl [pH 7.4], 0.5 M NaCl, 0.2% Tween 20). After the last incubation, the membrane was washed for 30 min and developed with an enhanced chemiluminescence kit (Amersham). The protein bands that reacted with α-Ikaros antibodies were visualized on Kodak XAR5 film. To compare in vitro transcription and translation products with the isoforms detected in RLm11 nuclear extracts, in vitro transcription and translation was performed (Promega) in the presence of [^{35}S]methionine.

RESULTS

Purification of LyF-1 revealed two candidates with molecular masses of 40 and 60 kDa. To isolate the gene encoding LyF-1, we purified the protein and obtained a partial amino acid sequence from proteolytic peptides. A procedure for partial purification of the protein was reported previously (28) (Fig. 1A). This procedure began with nuclear extracts from the cultured murine thymoma RLm11. The extracts were subjected to heparin-agarose chromatography and then passed twice over a sequence-specific DNA affinity column (see Materials and Methods). Coupled to the affinity column were multimers of a 19-bp sequence spanning the TdT D element, which is located at −100 in the TdT promoter and contains a high-affinity LyF-1-binding site. The TdT D element, which is not needed for promoter activity in a transient transfection assay (28), was used instead of the critical D′ element (at −60) because LyF-1 binds to a single site in the D element with high affinity and, unlike the D′ element, the D element does not contain an overlapping binding site for Ets family proteins.

The protein profiles detected in the heparin-agarose and DNA affinity column eluates were analyzed by SDS-PAGE followed by silver staining (Fig. 1B). Following the second pass over the affinity column, an abundant band of 60 kDa was detected in the 0.3 M KCl fraction (Fig. 1B, lane 4), along with a low level of gel shift activity (data not shown). The 0.6 M KCl fraction from the second pass contained no detectable DNA-binding activity, but abundant proteins of 40 and 105 kDa were observed (Fig. 1B, lane 5). (The 50-kDa protein detected previously [28] was not present in the second-pass fractions, suggesting that this protein was a contaminant [see Materials and Methods].) DNase I footprinting activity was not detected in any of the second-pass fractions, presumably because the protein was too dilute. Although the 0.6 M KCl fraction from the second affinity column lacked DNA-binding activity, we did not immediately exclude the possibility that the 40- or 105-kDa protein was LyF-1. Indeed, the proteins which elute at 0.6 M KCl must have a very high affinity for the oligonucleotide; it was possible that the high level of purity of this fraction could destabilize the protein. Of the three candidates (40, 60, and 105 kDa), the 105-kDa protein was the least likely to be LyF-1 because this protein bound tightly to an affinity column containing a different oligonucleotide that does not interact with

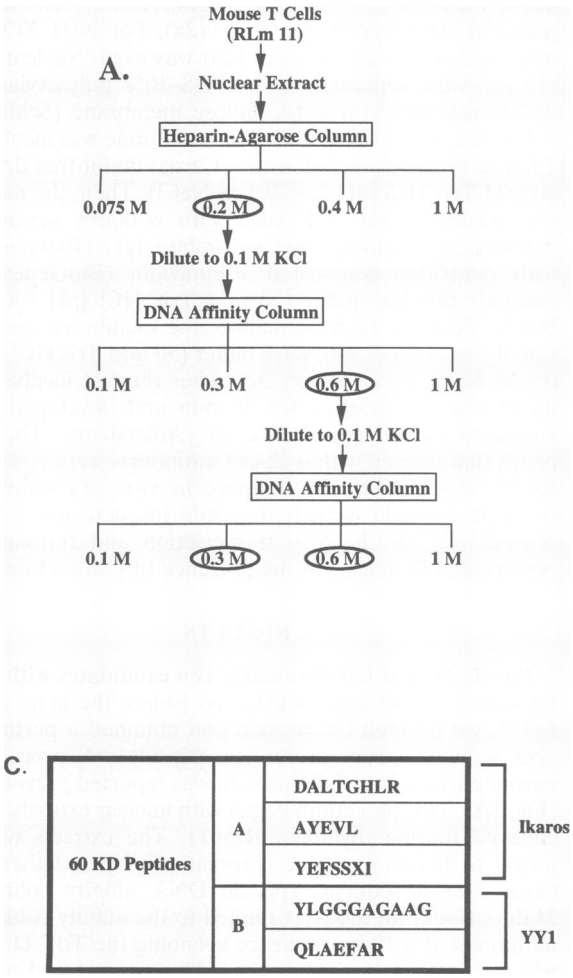

FIG. 1. Purification of LyF-1. (A) Schematic diagram of the LyF-1 purification protocol (see Materials and Methods). Column elutions were performed with buffers containing various concentrations of KCl as indicated. The protein fractions analyzed by SDS-PAGE are circled. (B) Pooled column fractions were analyzed by SDS-PAGE followed by silver staining. Included are the 0.2 M KCl eluate from the heparin-agarose column (lane 2), the 0.6 M KCl eluate from the first-pass DNA affinity column (lane 3), and the 0.3 M (lane 4) and 0.6 M (lane 5) KCl eluates from the second-pass DNA affinity column. Protein amounts in each lane were adjusted to give similar staining. Molecular weight markers (M; lanes 1 and 6) are indicated to the left in kilodaltons. Abundant proteins migrating at 105, 60, 50, and 40 kDa are indicated by the arrows to the right. Fractions containing gel shift and DNase I footprinting activities are indicated at the bottom. (C) Peptide sequences derived from tryptic fragments of proteins in the 60-kDa band. Three peptides match sequences in the Ikaros protein, and two match sequences in YY1.

LyF-1 (data not shown). The 40- and 60-kDa proteins were therefore the most likely candidates for LyF-1.

A partial amino acid sequence of the 60-kDa protein implicates Ikaros as a candidate for LyF-1. To determine which protein corresponds to LyF-1, we obtained amino acid sequences of tryptic peptides from both the 40- and 60-kDa proteins. The 40-kDa protein was found to contain domains homologous to RNA-binding proteins, suggesting that it is a single-stranded-DNA- and RNA-binding protein that bound to unannealed oligonucleotides attached to the affinity column (see reference 14). From the 60-kDa protein, we obtained five different peptide sequences (Fig. 1C), all of which were obtained in approximately equimolar amounts (data not shown). A GenBank database search revealed that these peptides were derived from two known proteins. Two peptides were derived from YY1 (also called UCRBP, δ, and NF-E1 [8, 15, 35, 41]) and three from a T-cell-specific protein called Ikaros (9). YY1 is a ubiquitously expressed zinc finger protein related to the *Drosophila* protein Krüppel (38). Ikaros was originally isolated from a λgt11 library that was screened with a radiolabeled DNA fragment from the CD3δ enhancer (9). A 48-kDa zinc finger protein related to the *Drosophila* protein hunchback (46), Ikaros contains three C_2H_2 zinc fingers at the N terminus and two at the C terminus. In the CD3δ enhancer, Ikaros bound to the sequence GAATGGGGGTGG.

It is unlikely that the five peptides were derived from a fusion protein resulting from a translocation of the *YY1* and *Ikaros* genes because the peptides came from various regions of the proteins; the molecular mass of a minimal fusion protein containing all of these peptides would be in excess of 90 kDa. Therefore, we concluded that both the YY1 and Ikaros proteins were present in the 60-kDa SDS-PAGE band that we isolated and sequenced.

LyF-1 is encoded by the *Ikaros* gene. Ikaros was the most likely candidate for LyF-1 because both proteins appear to be expressed specifically in lymphocytes. To determine if LyF-1 is Ikaros, three predictions can be tested: (i) the LyF-1-binding activity detected in nuclear extracts should depend on the presence of zinc; (ii) the LyF-1-binding activity should be altered or inhibited by α-Ikaros IgG; and (iii) recombinant Ikaros expressed in *E. coli* should specifically interact with the

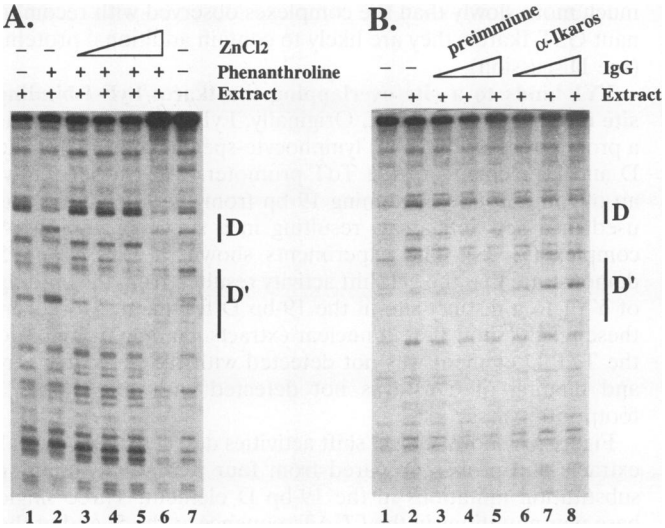

FIG. 2. Binding of LyF-1 to the TdT promoter requires zinc and is inhibited by α-Ikaros IgG. DNase I footprinting analyses were performed with the murine TdT promoter (28). (A) Footprinting reactions were performed with RLm11 nuclear extracts (15 μg) in the presence (lanes 2 to 6) or absence (lane 1) of phenanthroline. One, 2, 5, or 10 μl of 100 μM $ZnCl_2$ (lanes 3 to 6) was added following preincubation with phenanthroline (see Materials and Methods). Excess $ZnCl_2$ appeared to cause a nonspecific inhibition of footprinting (lane 6). A footprinting reaction performed in the absence of extract is shown in lane 7. The TdT D and D′ elements are indicated (vertical bars). (B) Preimmune and α-Ikaros IgG were added to DNase I footprinting reactions containing (lanes 2 to 8) or lacking (lane 1) 15 μg of RLm11 nuclear extract. Increasing amounts of preimmune (lanes 3 to 5) or α-Ikaros (lanes 6 to 8) IgG were preincubated with extract for 15 min prior to addition of the labeled probe. The TdT D and D′ elements are indicated by vertical bars.

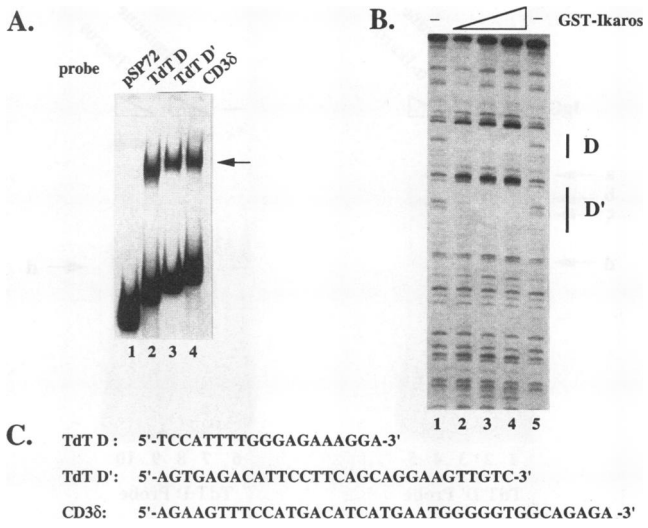

FIG. 3. Recombinant Ikaros binds to the TdT D and D′ elements. (A) Gel shift assays were performed with purified recombinant GST-Ikaros fusion protein. ^{32}P-labeled probes contained pSP72 vector sequences (lane 1), the TdT D element (lane 2), the TdT D′ element (lane 3), or the CD3δ enhancer element (lane 4). The specific protein-DNA complex is indicated (arrow). (B) DNase I footprinting assays were performed with purified recombinant GST-Ikaros fusion protein and a probe containing the TdT promoter. Binding reaction mixtures contained no protein (lanes 1 and 5) or 4 (lane 2), 20 (lane 3), or 100 (lane 4) ng of the purified GST-Ikaros. The TdT D and D′ elements are indicated by vertical bars. (C) DNA sequences of the TdT D, TdT D′, and CD3δ elements contained within the gel shift probes.

TdT D and D′ elements. These three experiments are described below, and the results shown in Fig. 2 and 3.

To test for zinc dependence, the zinc chelator phenanthroline was added to RLm11 nuclear extracts and was found to inhibit the binding of LyF-1 to the TdT D element in a DNase I footprinting assay (Fig. 2A; compare lanes 1 and 2). The binding pattern at the D′ element (which contains two overlapping LyF-1-binding sites and an Ets protein-binding site) was also altered, in that a weak band in the center of the protected region became hypersensitive to DNase digestion. We previously showed that this hypersensitive band results from the binding of Ets-related proteins to the D′ element (7). Thus, the enhancement of the hypersensitive site most likely results from an increase in the detection of Ets protein binding following the loss of LyF-1 binding. The retention of Ets protein binding in Fig. 2A, lane 2, demonstrates that the inhibition of DNA-binding activity by phenanthroline was specific for LyF-1. To confirm that depletion of zinc was responsible for the inhibition of LyF-1 binding, we added increasing concentrations of zinc chloride to DNase I footprinting reactions containing phenanthroline. Zinc chloride restored LyF-1 binding at the D element and resulted in protection of the hypersensitive site at the D′ element (Fig. 2A, lanes 3 to 5).

To determine if α-Ikaros IgG influences the LyF-1-binding activity, antiserum directed against a C-terminal fragment of the Ikaros protein was prepared and IgG from preimmune and immune sera were purified by protein A-Sepharose chromatography (see Materials and Methods). Addition of α-Ikaros IgG to a DNase I footprinting reaction with RLm11 nuclear extracts resulted in the same phenotype as was found with the phenanthroline experiment (Fig. 2B; compare lane 2 with lanes 6 to 8). Protection was lost at the TdT D element, and the hypersensitive site characteristic of Ets protein binding at the D′ element was enhanced. Addition of preimmune IgG had no effect on the DNase I footprinting pattern (Fig. 2B, lanes 3 to 5).

To test for direct binding of recombinant Ikaros to the TdT promoter, a full-length Ikaros cDNA was isolated and expressed in *E. coli* as a GST fusion protein. The cDNA isolated was identical to that reported previously (9) except for a 2-bp substitution in the N-terminal zinc finger domain and a 9-bp deletion in the C-terminal finger domain. The N-terminal difference results in two amino acid substitutions (amino acids 148 [V to M] and 149 [C to Y]). The C-terminal difference results in a three-amino-acid deletion in the zinc finger domain (amino acids 394 to 396), which improves the homology to the hunchback C-terminal zinc finger domain (see Fig. 7B). It is not clear if these changes influence the functional properties of the protein. These same differences were found when multiple cDNAs were isolated from the 70Z/3 library and when gene fragments were isolated from RLm11 mRNA by PCR (see Materials and Methods). Therefore, the differences probably result from allelic variations, since 70Z/3 and RLm11 cells are derived from BALB/c mice, whereas the original Ikaros cDNA was isolated from EL4, a C57BL-derived cell line (9).

The purified GST-Ikaros fusion protein was tested in both gel shift and DNase I footprinting assays (Fig. 3). In a gel shift assay, GST-Ikaros bound to probes derived from the TdT D and D′ elements as well as to the fragment of the CD3δ enhancer studied by Georgopoulos et al. (9) (Fig. 3A, lanes 2 to 4). In a DNase I footprinting assay, GST-Ikaros bound to

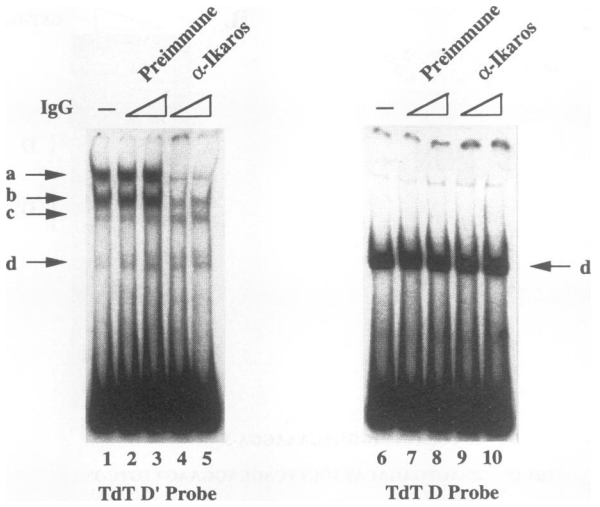

FIG. 4. Ikaros/LyF-1 is a component of two slow-mobility complexes that form on the TdT D′ element in RLm11 extracts. Gel shift assays were performed with the TdT D′ (lanes 1 to 5) or D (lanes 6 to 10) probe and RLm11 nuclear extracts (1.5 μg). Binding reaction mixtures lacked IgG (lanes 1 and 6) or contained IgG from preimmune (lanes 2, 3, 7, and 8) or α-Ikaros (lanes 4, 5, 9, and 10) serum. Complexes a through d are indicated (arrows). Complex d is likely to represent the binding of YY1 to both the D and D′ elements.

both the D and D′ elements (Fig. 3B, lanes 2 to 4), resulting in a protection pattern similar to that found with extensively purified LyF-1 (see reference 28). Taken together, the results in Fig. 2 and 3 establish that LyF-1 is encoded by the *Ikaros* gene.

Ikaros/LyF-1 is contained within a large protein-DNA complex which forms on the TdT D′ element in RLm11 nuclear extracts. Previous studies with the TdT D′ element revealed that in gel shift assay with nuclear extracts, slow-mobility complexes thought to contain multiple proteins could be detected (7). Identification of the components of these complexes may lead to an understanding of the precise protein-protein and protein-DNA interactions needed for TdT promoter function. To determine if Ikaros/LyF-1 is a component of slow-mobility complexes which form on the TdT D′ element, we added α-Ikaros IgG to binding reactions with RLm11 nuclear extracts (Fig. 4). In the absence of the antibodies, two abundant complexes were detected in the gel shift experiment (Fig. 4, lane 1, complexes a and b), both of which migrated much more slowly than the complex observed with recombinant GST-Ikaros in Fig. 3A. These slow-mobility complexes were absent when HeLa cell extracts were used, and both were observed only when zinc chloride was added to the binding reaction mixture (data not shown). These complexes are therefore distinct from the slow-mobility complexes reported previously, which were observed with several nonlymphoid cell extracts and did not require addition of zinc (6a, 7). Preincubation of α-Ikaros IgG with RLm11 nuclear extract prior to addition of the DNA probe inhibited formation of both complexes (Fig. 4, lanes 4 and 5). The α-Ikaros IgG had no effect on less abundant complexes detected in the same reaction (lanes 4 and 5, complexes c and d), and preimmune IgG had no effect on any of the bands detected with the D′ probe (Fig. 4, lanes 2 and 3).

These results demonstrate that two slow-mobility complexes which form on the TdT D′ element contain a protein encoded by the *Ikaros/LyF-1* gene. Because the complexes migrate much more slowly than the complexes observed with recombinant GST-Ikaros, they are likely to contain additional proteins (see Discussion).

YY1 binds to a site overlapping the Ikaros/LyF-1-binding site in the TdT D element. Originally, LyF-1 was identified as a protein responsible for a lymphocyte-specific footprint at the D and D′ elements of the TdT promoter (28). Subsequently, an oligonucleotide containing 19 bp from the D element was used in a gel shift assay, resulting in a single protein-DNA complex (7, 28). The experiments shown in Fig. 4 and 5 demonstrate that the gel shift activity resulted from the binding of YY1 to a distinct site in the 19-bp D fragment. Moreover, these results show that in nuclear extracts, binding of LyF-1 to the TdT D element was not detected with the gel shift assay and binding of YY1 was not detected with the DNase I footprinting assay.

Figure 5A shows the gel shift activities detected with RLm11 extracts and probes prepared from four plasmids containing substitution mutations in the 19-bp D element. Three single base pair mutations in the CCAT sequence at the 5′ end of the probe eliminated the gel shift activity (Fig. 5A, lanes 2 to 4), whereas a 4-bp mutation at the 3′ end of the element had little or no effect (Fig. 5A, lane 5). In contrast, Fig. 5C reveals that the three single base pair mutations which disrupted the gel shift activity had no effect on the LyF-1 footprinting activity (Fig. 5C, lanes 5 to 16), whereas the 4-bp mutation abolished the LyF-1 footprint (Fig. 5C, lanes 17 to 20). Consistent with these observations, the experiment in Fig. 4 reveals that α-Ikaros IgG did not influence the protein-DNA complex detected with the D element in a gel shift assay (lanes 9 and 10). These results demonstrate that the protein responsible for the gel shift activity is distinct from the Ikaros/LyF-1 protein.

Three observations suggest that YY1 is responsible for the gel shift activity detected with a probe containing the TdT D element. First, the CCAT sequence essential for formation of this protein-DNA complex resembles known high-affinity binding sites for YY1 (19, 41). Second, YY1 bound with high affinity to the sequence-specific DNA affinity column containing the TdT D element (Fig. 1). Finally, the protein-DNA complex detected with the gel shift assay migrates at the same location as a complex formed with the TdT D element and recombinant YY1 expressed in *E. coli* (Fig. 5B; recombinant YY1 was a gift of M. Gilman, Cold Spring Harbor Laboratory).

In summary, the data described above demonstrate that the gel shift activity detected in nuclear extracts with the TdT D element results from the binding of YY1 to a sequence which overlaps the Ikaros/LyF-1-binding site (Fig. 5D). Although neither the Ikaros/LyF-1-binding site nor the YY1-binding site within the D element influences TdT promoter activity in a transient transfection assay (27a, 28), this result is significant because it establishes that the gel shift assay cannot be used to measure Ikaros/LyF-1-binding activity in nuclear extracts.

Several distinct mRNAs are generated from the *Ikaros/LyF-1* gene. For the studies described above, we isolated *Ikaros/LyF-1* cDNAs by screening a phage library prepared from the 70Z/3 pre-B-cell line and also by PCR amplification of mRNA from the RLm11 T cell line. Interestingly, both of these approaches resulted in the isolation of multiple distinct cDNA products. Six different cDNAs have been isolated, and the nucleotide sequences have been determined (Fig. 6 and 7). Importantly, each form maintains the same open reading frame, and the variations occur exclusively at one location (following amino acid 53 in the original Ikaros protein sequence). All forms of the mRNA encode proteins with the same 53 amino acids at the N terminus (N-terminal sequence [NTS]) and 236 amino acids at the C terminus (CTS). Only

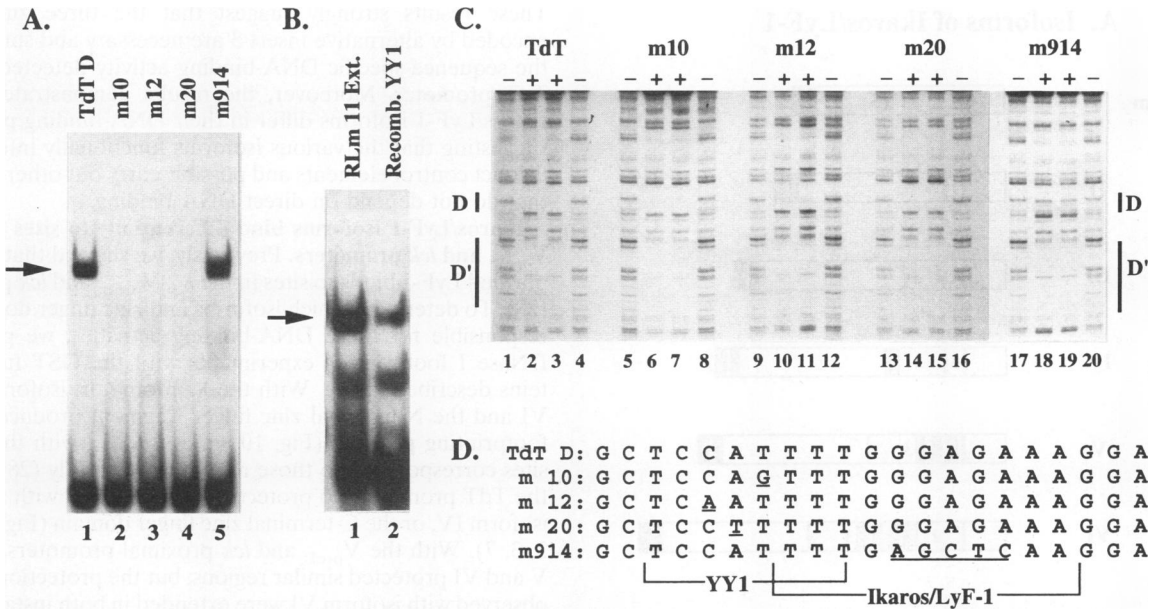

FIG. 5. YY1 is responsible for the gel shift activity detected in nuclear extracts with the TdT D element. (A) Gel shift assays were performed with RLm11 nuclear extract (1.5 μg) and probes containing the wild-type TdT D element (lane 1) or mutant m10 (lane 2), m12 (lane 3), m20 (lane 4), or m914 (lane 5). (B) Gel shift assays were performed with a probe containing the TdT D element and either RLm11 extract (lane 1) or recombinant YY1 expressed in *E. coli* (lane 2). The complex of interest is indicated with an arrow; the lower complex results from an abundant degradation product of YY1 (11a). (C) DNase I footprinting assays were performed with RLm11 extracts and probes containing either the wild-type TdT promoter (lanes 1 to 4) or the m10 (lanes 5 to 8), m12 (lanes 9 to 12), m20 (lanes 13 to 16), or m914 (lanes 17 to 20) D element mutant promoter. Binding reaction mixtures contained (lanes 2, 3, 6, 7, 10, 11, 14, 15, 18, and 19) or lacked (lanes 1, 4, 5, 8, 9, 12, 13, 16, 17, and 20) RLm11 extract. (D) DNA sequences of the wild-type and mutant TdT D elements. The approximate recognition sites for YY1 and Ikaros/LyF-1 are indicated at the bottom.

isoform I lacks an insertion between the NTS and CTS fragments. The remaining isoforms contain one or two of the three distinct insert sequences shown in Fig. 7 (inserts 1 to 3). These observations strongly suggest that the various cDNAs are derived from the *Ikaros/LyF-1* gene by alternative splicing.

Many of the protein isoforms encoded by the mRNAs differ in the number and structure of the N-terminal zinc fingers (Fig. 6). Forms I and II encode proteins which lack zinc fingers at the N terminus and therefore contain zinc fingers only in the CTS. Forms III, IV, V, and VI encode proteins which contain one, one, three, and four zinc fingers, respectively, at the N terminus, in addition to the two zinc fingers in the CTS. These observations suggest that the six isoforms may vary in their DNA binding and functional characteristics.

Multiple isoforms of Ikaros/LyF-1 are expressed in lymphoid cell lines. During the initial isolation of the *Ikaros* gene, Northern (RNA) analysis and in situ hybridization experiments suggested that it was expressed with specificity in T cells and in T-cell progenitors (9). Our previous characterization of the LyF-1 protein revealed DNA-binding activity in multiple B-cell lines (28). To address this issue, Western blot analyses were performed with antibodies directed against the CTS of Ikaros/LyF-1. In RLm11 extracts, several bands were consistently detected (Fig. 8A, lane 5). Four major bands migrate similarly to isoforms I, III, V, and VI generated by an in vitro transcription-translation protocol (Fig. 8A, lanes 1 to 4). (Isoforms III and IV differ by only 2 kDa, making it difficult to determine which of these isoforms is present in RLm11 extracts.) Additional bands which do not correspond to isoforms I through VI were also detected (Fig. 8A, lane 5); these bands could represent additional isoforms or unrelated proteins which cross-react with the antiserum. On the basis of the migration of these isoforms relative to molecular weight markers, isoform VI was likely to be contained within the 60-kDa band that was used to generate the amino acid sequence (Fig. 1).

To determine if Ikaros/LyF-1 is expressed in B-cell lines and if the relative abundance of specific isoforms varies from cell line to cell line, we performed immunoblotting with several nuclear extracts. Isoforms V and VI appear to be expressed at high levels in six cell lines representing early and late pre-B cells (Fig. 8B, lanes 1 to 6), only one of which (40E4) expresses high levels of TdT. Isoforms V and VI were also expressed at high levels in two TdT-expressing T-cell lines (lanes 8 and 9). In contrast, the bands comigrating with isoforms I and III were expressed in only a subset of the cell lines (lanes 1, 4, 6, and 9). The S194 plasma cell line and the 3T3 fibroblast line lacked expression of all Ikaros/LyF-1 isoforms (lanes 7 and 10), consistent with the expression pattern defined by Northern analysis and in situ hybridization (9). These results establish that products of the *Ikaros/LyF-1* gene are expressed in several cell lines with characteristics of immature B cells, suggesting that Ikaros/LyF-1 isoforms are expressed in normal pre-B cells.

DNA-binding properties of Ikaros/LyF-1 isoforms depend on their N-terminal zinc finger domains. The presence of multiple Ikaros/LyF-1 isoforms which differ in their zinc finger domains suggests that the DNA-binding properties of the isoforms may vary. Isoforms I, IV, V, and VI were expressed in *E. coli* and tested for binding to the TdT promoter. In the gel shift assay, isoforms V and VI formed complexes with probes containing the TdT D and D′ elements, but complexes were not detected with isoforms I and IV (Fig. 9A, lanes 7 to 10 and 13 to 16). The DNase I footprinting experiments confirm and extend these results (Fig. 9B). Isoforms V and VI, but not isoforms I and IV, protected both the D and D′ elements (Fig.

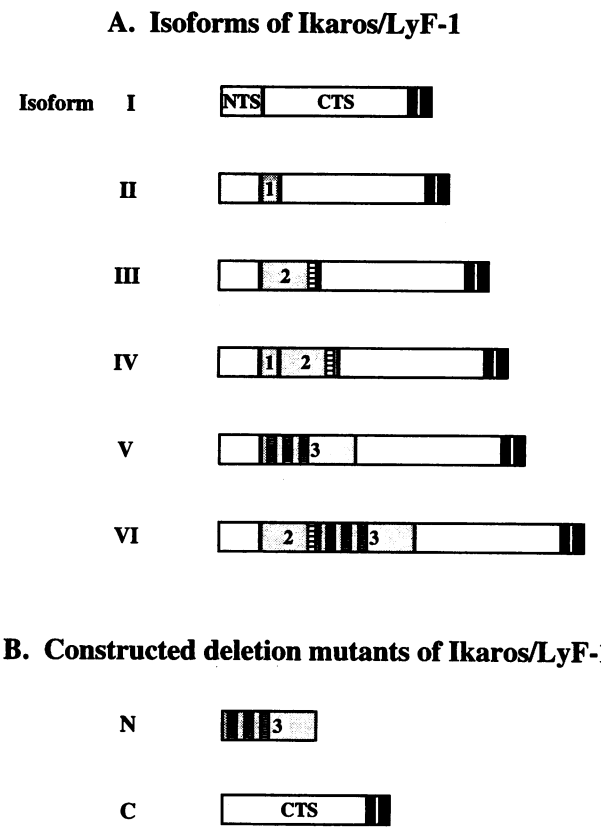

FIG. 6. Schematic diagram of Ikaros/LyF-1 isoforms. (A) Structures of the six characterized isoforms. NTS and CTS refer to the sequences found in all isoforms. Inserts 1 through 3 refer to sequences shown in Fig. 7. The zinc finger domains in inserts 2 and 3 and in the CTS are indicated with striped or solid bars. (B) Structures of the N-terminal and C-terminal zinc finger domains fused separately to the GST protein for the studies in Fig. 9 and 10.

9B, lanes 2 to 9). Interestingly, the DNase I footprinting experiment revealed that the protection patterns found with isoforms V and VI were different. Isoform V appeared to protect the D element more strongly than the D′ element (lanes 6 and 7), but isoform VI appeared to protect the D′ element more strongly than the D element (lanes 8 and 9). Furthermore, the footprint observed with isoform VI on the D′ element was somewhat longer than the footprint observed with isoform V, extending slightly farther in both the 3′ and 5′ directions. These observations suggest that the additional zinc finger present in isoform VI may have two distinct effects on the DNA-binding properties. First, it may alter the sequence specificity of the protein; second, in some instances, the additional zinc finger may contact flanking base pairs, extending the region protected from DNase digestion.

The data described above suggest that the N-terminal zinc finger domains are the primary determinants of the DNA-binding characteristics. To investigate the relative contributions of the N-terminal and C-terminal zinc finger domains, the three N-terminal zinc fingers found in isoform V and the two C-terminal zinc fingers found in all isoforms were expressed as GST fusion proteins (Fig. 7B). The C-terminal fragment was unable to bind to either the D or D′ element (Fig. 9A, lanes 12 and 18; Fig. 9B, lanes 12 and 13), but the N-terminal zinc finger domain from isoform V was sufficient for binding in both assays (Fig. 9A, lanes 11 and 17; Fig. 9B, lanes 10 and 11). These results strongly suggest that the three zinc fingers encoded by alternative insert 3 are necessary and sufficient for the sequence-specific DNA-binding activity detected with the TdT promoter. Moreover, the results demonstrate that the Ikaros/LyF-1 isoforms differ in their DNA-binding properties, suggesting that the various isoforms functionally interact with distinct control elements and possibly carry out other functions that do not depend on direct DNA binding.

Ikaros/LyF-1 isoforms bind differentially to sites in the λ_5, V_{preB}, and *lck* promoters. Previously, we showed that partially-purified LyF-1 binds to sites in the λ_5, V_{preB}, and *lck* promoters (28). To determine which isoforms and zinc finger domains are responsible for these DNA-binding activities, we performed DNase I footprinting experiments with the GST fusion proteins described above. With the λ_5 promoter, isoforms V and VI and the N-terminal zinc finger fragment produced similar footprinting patterns (Fig. 10, lanes 4 to 6), with the binding sites corresponding to those reported previously (28). As with the TdT promoter, no protection was detected with isoform I, isoform IV, or the C-terminal zinc finger domain (Fig. 10, lanes 2, 3, 7). With the V_{preB} and *lck* proximal promoters, isoforms V and VI protected similar regions, but the protection patterns observed with isoform VI were extended in both instances (Fig. 10, lanes 12, 13, 17, and 18). Hypersensitive sites were also detected at the 3′ ends of the probes with isoform VI, suggesting that weak interactions may occur at those locations. Interestingly, with both the V_{preB} and *lck* promoters, isoform I also produced specific hypersensitive bands at the 3′ ends of the radiolabeled probes. The presence of these bands suggests that isoform I may bind to specific sequences at these locations with a relatively low affinity. Taken together, the DNase I footprinting results in Fig. 9 and 10 demonstrate that, for the most part, isoforms V and VI bind to similar sequences. High-affinity binding sites for isoforms I and IV and for the C-terminal zinc finger domain were not observed, suggesting that these proteins either do not bind to DNA with high affinity or bind with high affinity to sequences not represented in the four promoters tested.

DISCUSSION

The B- and T-cell transcription factor LyF-1 is encoded by the *Ikaros* gene. The results presented here establish that the *Ikaros* gene (9) encodes the LyF-1 protein, which we previously implicated in transcriptional activation in early B and T cells (28). *Ikaros* was originally isolated as a T-cell-specific gene encoding a protein which binds to a control element in the CD3δ enhancer (9). The experiments shown here demonstrate that various isoforms of Ikaros/LyF-1 are also expressed in pre-B-cell lines representing various stages of development. Expression was detected in the early pre-B-cell line HAFTL-A, which expresses the B220 antigen and contains the Ig μ and κ genes primarily in germ line configuration (2). Ikaros isoforms were also expressed in more mature pre-B-cell lines, including the 70Z/3 cell line, which has functional rearrangements of both the μ and κ genes but does not express κ protein (29, 34). Ikaros/LyF-1 was not expressed in the S194 plasma cell line, consistent with the low level of expression observed in murine spleen (9).

Ikaros/LyF-1 was expressed in the B- and T-cell lines which express TdT (see reference 28), including 40E4, RLm11, and R1.1, consistent with a possible role in TdT regulation. However, the Ikaros/LyF-1 expression pattern does not match precisely with the TdT expression pattern during lymphopoiesis, since most of the pre-B cells tested do not express their endogenous TdT gene (28). This result implies that additional

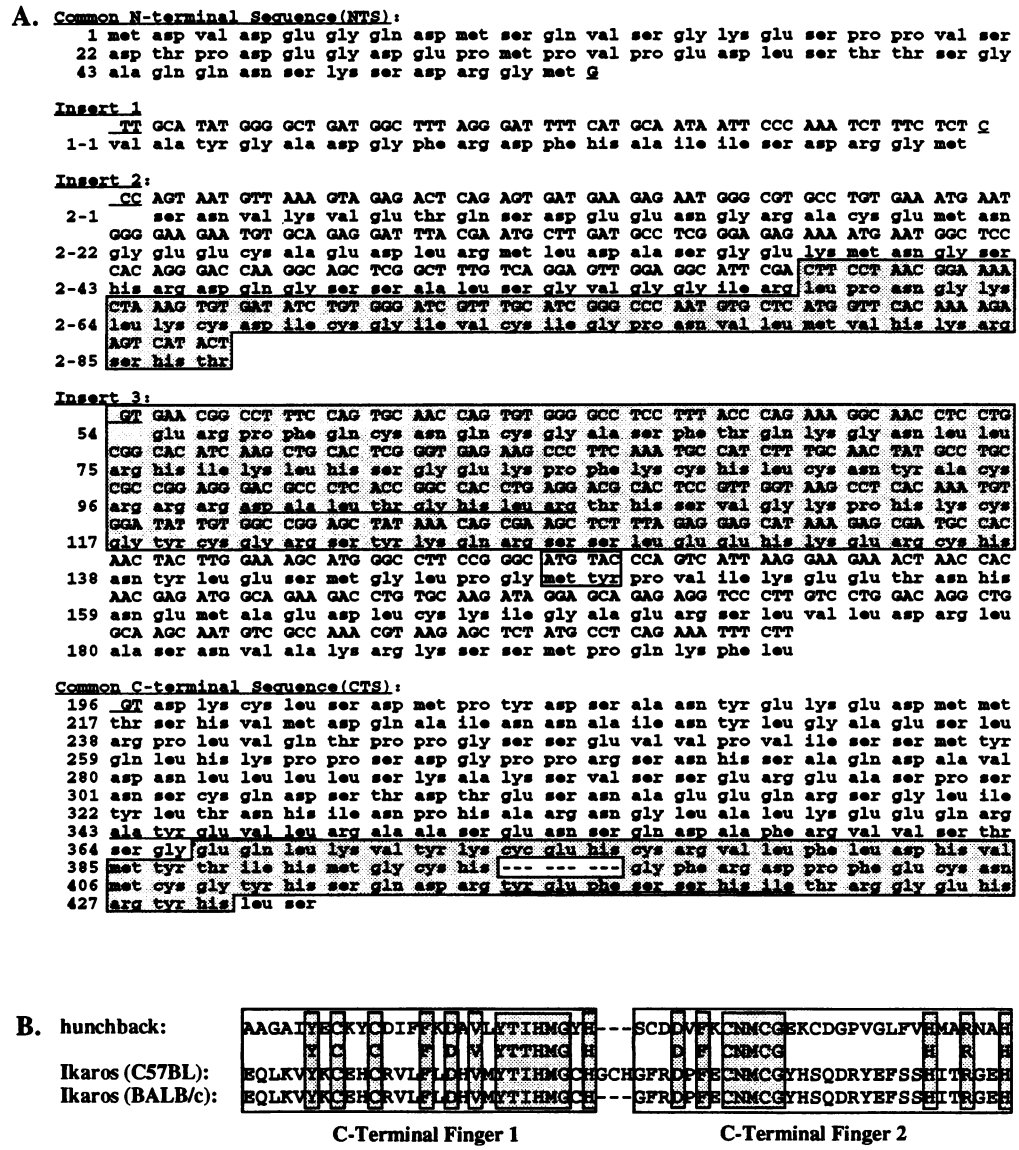

FIG. 7. DNA sequences of Ikaros/LyF-1 cDNA segments introduced into mRNA forms I through VI. (A) The common NTS and CTS were present in all isoforms. One or more of inserts 1 through 3 are found in isoforms II through VI. Isoform V, containing the NTS, CTS, and insert 3, was reported previously (9). The zinc finger domains are shaded, and the locations of the peptides obtained by amino acid sequence analysis (Fig. 1) are underlined. Open boxes indicate locations of allelic differences. The first amino acid of most segments varies depending on whether the CTS or insert 1 precedes it (see underlined nucleotides). (B) Comparison of the sequence of the C-terminal zinc finger domain found in Ikaros/LyF-1 from BALB/c-derived cell lines (RLm11 and 70Z/3) with the corresponding sequence from a C57BL-derived cell line (EL4 cells [9]) and from the *Drosophila* hunchback protein. The three-amino-acid deletion in the BALB/c sequence increases the similarity to the hunchback protein.

proteins are needed for TdT regulation, an idea that is consistent with combinatorial principles of gene regulation as well as with our observation that multiple proteins bind to the critical TdT D′ element (7).

The expression pattern and multiple isoforms of Ikaros/LyF-1 suggest that the isoforms may play roles in activating a wide range of genes during both B- and T-cell development. Several other transcription factors are expressed in both lymphoid lineages with different degrees of specificity. For example, LEF-1, which has been implicated in T-cell receptor α-chain expression, appears to be tightly restricted to T cells and immature B cells (47, 48). In contrast, Ets family members, including Ets-1 (49) and Fli-1 (3), have been implicated in the activation of a variety of B- and T-cell-specific genes (7, 13, 17, 32, 37), but these proteins are expressed in a wider range of cell types (3, 4, 37). Determination of the actual role of each of these genes during B- and T-cell development awaits the completion of gene disruption experiments.

Ikaros/LyF-1 is a component of a large protein-DNA complex that forms on the TdT D′ element. Previous studies demonstrated that the critical D′ element in the TdT promoter contains two adjacent binding sites for LyF-1, with a binding site for Ets family proteins coinciding with the proximal LyF-1 site. Although we have not determined which proteins are

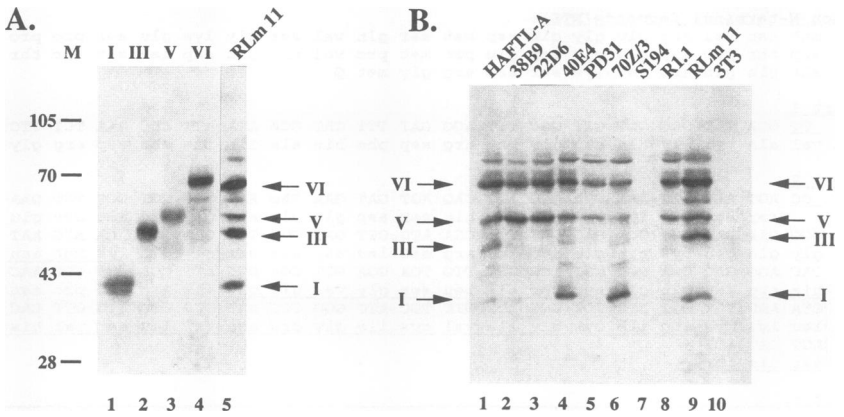

FIG. 8. Multiple isoforms of Ikaros/LyF-1 are expressed in B and T lymphoid cell lines. (A) RLm11 nuclear extracts (lane 5) and [^{35}S]Met-labeled in vitro-transcribed and translated products of isoforms I (lane 1), III (lane 2), V (lane 3), and VI (lane 4) were separated on an SDS–10% polyacrylamide gel and transferred to a nitrocellulose membrane. Western blot analysis was performed on lane 5, and the autoradiograph of the Western blot after enhanced chemiluminescence (Amersham) was compared with an autoradiograph of the [^{35}S]Met-labeled products. Sizes of molecular weight markers (m; in kilodaltons) are shown on the left. Arrows on the right indicate the mobilities of in vitro-translated products. (B) Western blot analysis was performed with nuclear extracts (15 μg) from six pre-B-cell lines, HAFTL-A (lane 1), 38B9 (lane 2), 22D6 (lane 3), 40E4 (lane 4), PD31 (lane 5), and 70Z/3 (lane 6), from the plasmacytoma S194 (lane 7), from two T-cell lines, R1.1 (lane 8) and RLm11 (lane 9), and from 3T3 fibroblasts (lane 10). An autoradiograph after enhanced chemiluminescence (Amersham) is shown. Proteins that migrate with the in vitro-translated products of isoforms I, III, V, and VI are indicated by arrows.

essential for transcriptional activation through the D′ element, the slow-mobility protein-DNA complexes detected in RLm11 extracts may provide a useful starting point for addressing this issue. These complexes contain proteins derived from the *Ikaros/LyF-1* gene, but their slow migration relative to complexes formed with the recombinant Ikaros/LyF-1 proteins suggests that they contain multiple Ikaros/LyF-1 isoforms and/or additional proteins. The abundance of these two complexes relative to the smaller complexes detected in the same experiment suggests that the proteins contained within the complexes bind cooperatively to the D′ element. It is unlikely that these complexes simply contain two copies of Ikaros/LyF-1 isoforms V or VI, since slow-mobility complexes were not detected when the recombinant proteins were tested by gel shift with the TdT D′ probe (Fig. 3A). The complexes therefore are likely to contain either a combination of Ikaros/LyF-1

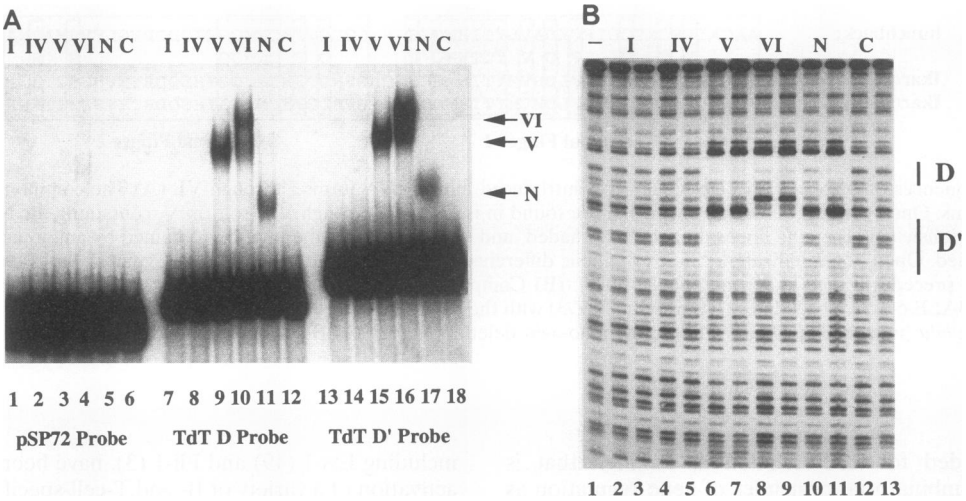

FIG. 9. Binding of Ikaros/LyF-1 isoforms to the TdT D and D′ elements. (A) Gel shift assays were performed with probes derived from the pSP72 polylinker (lanes 1 to 6), TdT D element (lanes 7 to 12), or TdT D′ element (lanes 13 to 18). Reaction mixtures contained similar amounts of recombinant GST-Ikaros fusion proteins from isoform 1 (lanes 1, 7, and 13), IV (lanes 2, 8, and 14), V (lanes 3, 9, and 15), or VI (lanes 4, 10, and 16). Experiments were also performed with GST fusions of the N-terminal zinc finger domain from isoform V (lanes 5, 11, and 17; see Fig. 6) or the common C-terminal zinc finger domain (lanes 6, 12, and 18). Arrows indicate the complexes formed with isoforms V and VI and with the N-terminal zinc finger fragment. (B) DNase I footprinting assays were performed with a TdT promoter probe and two different amounts of purified GST-Ikaros isoforms I (lanes 2 and 3), IV (lanes 4 and 5), V (lanes 6 and 7), and VI (lanes 8 and 9). The GST fusions with the N-terminal (lanes 10 and 11) and C-terminal (lanes 12 and 13) fragments were also tested (see Fig. 6). Each reaction mixture contained approximately 100 ng of the full-length recombinant protein compared with bovine serum albumin standards by SDS-PAGE. However, degradation products were present in the purified isoforms V and VI, preventing us from accurately comparing concentrations of active proteins. The TdT D and D′ sites are indicated by bars on right.

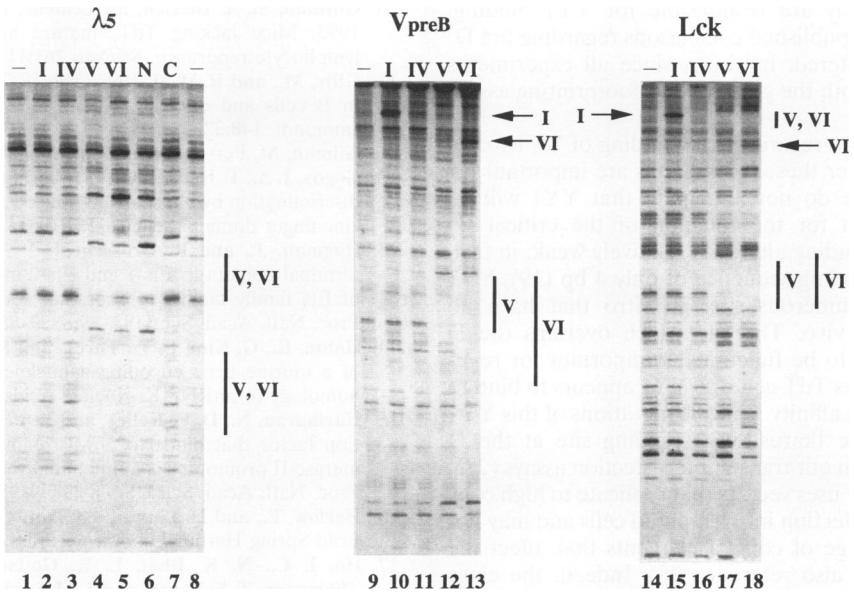

FIG. 10. Ikaros/LyF-1 isoforms bind to the λ_5, V_{preB}, and *lck* promoters. DNase I footprinting assays were performed with ^{32}P-labeled probes containing the λ_5 (lanes 1 to 8), V_{preB} (lanes 9 to 13), and *lck* (lanes 14 to 18) promoters. Probes were prepared as described previously (28). Binding reactions were carried out in the absence of protein (lanes 1, 8, 9, 14) or in the presence of approximately 100 ng of purified GST-Ikaros isoforms I (lanes 1, 9, and 15), IV (lanes 3, 11, and 16), V (lanes 4, 12, and 17), and VI (lanes 5, 13, and 18) or the GST–N-terminal zinc finger domain (lane 6) or GST–C-terminal domain (lane 8). DNA sequences protected by isoforms V and VI are indicated by bars on the right, and hypersensitive sites found with isoforms, I, V, and VI are indicated by arrows.

isoforms or an Ikaros/LyF-1 isoform in combination with other proteins. It is tempting to speculate that the complexes contain an Ikaros/LyF-1 isoform in combination with an Ets family member. Experiments are under way to purify these complexes to identify their components.

Multiple protein isoforms are generated from the *Ikaros/LyF-1* gene. An unexpected feature of the *Ikaros/LyF-1* gene is that it generates multiple mRNA and protein products, apparently through alternative splicing. Most interesting is the fact that the alternative splicing alters the domain which is critical for DNA binding. Two isoforms (V and VI) bind tightly to four different promoters, but their DNase I footprinting patterns vary. These isoforms may regulate different sets of genes by binding with different affinities to distinct promoter or enhancer elements. Isoforms I and IV (and most likely isoforms II and III) do not bind with high affinity to the promoters tested, but the hypersensitive sites detected in Fig. 10 suggest that they may bind to distinct DNA sequences. Alternatively, isoforms I through IV may carry out unique functions in developing lymphocytes or may act as dominant negative regulators of isoforms V and VI by blocking protein-protein interactions. It may be noteworthy that the two isoforms found to vary among the pre-B- and T-cell lines tested are isoforms I and III. Possibly, isoforms V and VI play a more general role in activating transcription in the B- and T-cell lineages, whereas isoforms I and III may play more specific regulatory roles.

Several other cases of altered DNA-binding properties generated through alternative splicing have been reported. For example, alternative splicing leads to two isoforms of the *Drosophila* CF2 protein, which differ in their zinc finger domains and DNA-binding specificites (12, 18). In addition, one isoform of the Wilms' tumor protein contains a specific insertion between two zinc fingers that impairs DNA binding (36). The E47 and E12 proteins represent an example of non-zinc finger proteins with different DNA-binding characteristics generated by alternative splicing (31). These proteins are derived from the same gene but contain different basic helix-loop-helix DNA-binding domains. Thus, variation of DNA-binding properties through alternative splicing appears to be a common mechanism by which related proteins might regulate the expression of a variety of different genes.

YY1 binds to a sequence overlapping the Ikaros/LyF-1 site. The finding that YY1 binds to a site overlapping the Ikaros/LyF-1 binding site in the TdT D element was surprising, especially since in nuclear extracts, YY1 binding was detected only with the gel shift assay and Ikaros/LyF-1 binding was detected only with the DNase I footprinting assay. Although the DNase I footprinting assay was used as a measure of LyF-1 binding in most of our previous experiments, the gel shift assay was sometimes used as a more sensitive alternative. We now know that YY1 and not Ikaros/LyF-1 was responsible for this activity. One issue that was addressed with the gel shift assay was the precise tissue distribution of LyF-1 (28). Gel shift activity was found in numerous lymphoid and nonlymphoid cell lines, although the amount of activity was generally increased in lymphoid extracts. Since YY1 is thought to be a ubiquitously expressed protein, experiments are under way to determine if it is indeed enriched in lymphoid cell lines. More importantly, we have never detected LyF-1 footprinting activity in extracts from a nonlymphoid cell line, and the results in Fig. 8 and those of Georgopoulos et al. (9) suggest that Ikaros/LyF-1 expression is more tightly restricted to the lymphoid lineages than we previously suspected.

We also reported binding of LyF-1 to the TdT D′ site with both the DNase I footprinting and gel shift assays (7), although the gel shift activity was much weaker than detected with the D site. We now suspect that the gel shift activity at the D′ site resulted from the binding of YY1. We previously reported that YY1 binding depends primarily on a four-nucleotide consensus sequence of CCAT (19). On the basis of this consensus sequence, two low-affinity binding sites are present in the D′

element and most likely are responsible for YY1 binding. None of our previously published conclusions regarding the D′ element need to be altered, however, since all experiments were performed with both the gel shift and footprinting assays (7).

An important question regarding the binding of YY1 to the TdT promoter is whether these interactions are important for promoter function. We do not anticipate that YY1 will be found to be important for the function of the critical D′ element because the binding affinity is relatively weak; in fact, given the consensus binding sequence of only 4 bp (19), YY1 is likely to bind to numerous sites in vitro that have no important function in vivo. The site which overlaps the D element is more likely to be functionally important for regulation of the endogenous TdT gene, as YY1 appears to bind to this site with a very high affinity. Specific mutations of this YY1 binding site and of the Ikaros/LyF-1 binding site at the D element had not effect in our transient transfection assays (27a, 28). However, this assay uses vectors that replicate to high copy number following transfection into lymphoid cells and may not depend on the full range of control elements that affect the endogenous gene (see also reference 28). Indeed, the close proximity of the YY1- and Ikaros/LyF-1-binding sites in the TdT promoter is intriguing, considering the functional antagonism between the related *Drosophila* gap proteins Krüppel and hunchback during *Drosophila* embryogenesis (44).

ACKNOWLEDGMENTS

We are extremely grateful to Owen Witte and Rudi Grosschedl for helpful discussions, to Katia Georgopoulos for communicating results prior to publication, and to Doug Black and Owen Witte for critically reading the manuscript.

This work was supported by PHS grant DK43726. K.H. and P.E. are supported by PHS National Research Service Awards GM-07104 and GM-07185, respectively.

REFERENCES

1. **Abramson, S., R. G. Miller, and R. A. Phillips.** 1977. The identification in adult bone marrow of pluripotent and restricted stem cells of the myeloid and lymphoid systems. J. Exp. Med. **145:**1567–1579.
2. **Alessandrini, A., J. H. Pierce, D. Baltimore, and S. V. Desiderio.** 1987. Continuing rearrangement of immunoglobulin and T-cell receptor genes in a Ha-*ras*-transformed lymphoid progenitor cell line. Proc. Natl. Acad. Sci. USA **84:**1799–1803.
3. **Ben-David, Y., E. B. Giddens, K. Letwin, and A. Bernstein.** 1991. Erythroleukemia induction by Friend murine leukemia virus: insertional activation of a new member of the *ets* gene family, *Fli-1*, closely linked to c-*ets*-1. Genes Dev. **5:**908–918.
4. **Bhat, N. K., R. J. Fisher, S. Fujiwara, R. Ascione, and T. S. Papas.** 1987. Temporal and tissue-specific expression of mouse *ets* genes. Proc. Natl. Acad. Sci. USA **84:**3161–3165.
5. **Bollum, F. J.** 1979. Terminal deoxynucleotidyl transferase as a hematopoietic cell marker. Blood **54:**1203–1215.
6. **Colgan, J., and J. L. Manley.** 1992. TFIID can be rate limiting in vivo for TATA-containing, but not TATA-lacking, RNA polymerase II promoters. Genes Dev. **6:**304–315.

6a. **Ernst, P.** Unpublished data.

7. **Ernst, P., K. Hahm, and S. T. Smale.** 1993. Both LyF-1 and an Ets protein interact with a critical promoter element in the murine terminal transferase gene. Mol. Cell. Biol. **13:**2982–2992.
8. **Flanagan, J. R., K. G. Becker, D. L. Ennist, S. L. Gleason, P. H. Driggers, B. Z. Levi, E. Appella, and K. Ozato.** 1992. Cloning of a negative transcription factor that binds to the upstream conserved region of Moloney murine leukemia virus. Mol. Cell. Biol. **12:**38–44.
9. **Georgopoulos, K., D. D. Moore, and B. Derfler.** 1992. *Ikaros*, an early lymphoid-specific transcription factor and a putative mediator for T cell commitment. Science **258:**808–812.
10. **Gilfillan, S., A. Dierich, M. Lemeur, C. Benoist, and D. Mathis.** 1993. Mice lacking TdT: mature animals with an immature lymphocyte repertoire. Science **261:**1175–1178.
11. **Gilly, M., and R. Wall.** 1992. The IRG-47 gene is IFN-γ induced in B cells and encodes a protein with GTP-binding motifs. J. Immunol. **148:**3275–3281.

11a. **Gilman, M.** Personal communication.

12. **Gogos, J. A., T. Hsu, J. Bolton, and F. C. Kafatos.** 1992. Sequence discrimination by alternatively spliced isoforms of a DNA binding zinc finger domain. Science **257:**1951–1955.
13. **Hagman, J., and R. Grosschedl.** 1992. An inhibitory carboxyl-terminal domain in Ets-1 and Ets-2 mediates differential binding of Ets family factors to promoter sequences of the mb-1 gene. Proc. Natl. Acad. Sci. USA **89:**8889–8893.
14. **Hahm, K., G. Kim, C. W. Turck, and S. T. Smale.** 1993. Isolation of a murine gene encoding a nucleic acid-binding protein with homology to hnRNP K. Nucleic Acids Res. **21:**3894.
15. **Hariharan, N., D. E. Kelley, and R. P. Perry.** 1991. δ, a transcription factor that binds to downstream elements in several polymerase II promoters, is a functionally versatile zinc finger protein. Proc. Natl. Acad. Sci. USA **88:**9799–9803.
16. **Harlow, E., and D. Lane.** 1988. Antibodies: a laboratory manual. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.
17. **Ho, I. C., N. K. Bhat, L. R. Gottschalk, T. Lindsten, C. B. Thompson, T. S. Papas, and J. M. Leiden.** 1990. Sequence specific binding of human Ets-1 to the T cell receptor α gene enhancer. Science **250:**814–818.
18. **Hsu, T., J. A. Gogos, S. A. Kirsh, and F. C. Kafatos.** 1992. Multiple zinc finger forms resulting from developmentally regulated alternative splicing of a transcription factor gene. Science **257:**1946–1950.
19. **Javahery, R., A. Khachi, K. Lo, B. Zenzie-Gregory, and S. T. Smale.** 1994. DNA sequence requirements for transcriptional initiator activity in mammalian cells. Mol. Cell. Biol. **14:**116–127.
20. **Kadonaga, J. R., and R. Tjian.** 1986. Affinity purification of sequence-specific DNA-binding proteins. Proc. Natl. Acad. Sci. USA **83:**5889–5893.
21. **Kallenbach, S., N. Doyen, M. Fanton d'Andon, and F. Rougeon.** 1992. Three lymphoid-specific factors account for all junctional diversity characteristic of somatic assembly of T-cell receptor and immunoglobulin genes. Proc. Natl. Acad. Sci. USA **89:**2799–2803.
22. **Kato, K., H. M. Goncalves, G. E. Gouts, and F. J. Bollum.** 1967. Deoxynucleotide-polymerizing enzymes of calf thymus gland. J. Biol. Chem. **242:**2780–2789.
23. **Komori, T., A. Okada, V. Stewart, and F. W. Alt.** 1993. Lack of N regions in antigen receptor variable region genes of TdT-deficient lymphocytes. Science **261:**1171–1175.
24. **Landau, N. R., D. G. Schatz, M. Rosa, and D. Baltimore.** 1987. Increased frequency of N-region insertion in a murine pre-B cell line infected with a terminal deoxynucleotidyl transferse retroviral expression vector. Mol. Cell. Biol. **7:**3237–3243.
25. **Landau, N. R., T. P. St. John, I. L. Weisssman, S. C. Wolf, A. E. Silverstone, and D. Baltimore.** 1984. Cloning of terminal transferase cDNA by antibody screening. Proc. Natl. Acad. Sci. USA **81:**5836–5840.
26. **Lee, K. A., A. Bindereif, and M. R. Green.** 1988. A small-scale procedure for preparation of nuclear extracts that support efficient transcription and pre-mRNA Splicing. Gene Anal. Tech. **5:**22–31.
27. **Li, Y. S., K. Hayakawa, and R. R. Hardy.** 1993. The regulated expression of B lineage associated genes during B cell differentiation in bone marrow and fetal liver. J. Exp. Med. **178:**951–960.

27a. **Lo, K.** Unpublished data.

28. **Lo, K., N. R. Landau, and S. T. Smale.** 1991. LyF-1, a transcriptional regulator that interacts with a novel class of promoters for lymphocyte-specific genes. Mol. Cell. Biol. **11:**5229–5243.
29. **Maki, R., J. Kearney, C. Paige, and S. Tonegawa.** 1980. Immunoglobulin gene rearrangement in immature B cells. Science **209:**1366–1369.
30. **Mombaerts, P., J. Iacomini, R. S. Johnson, K. Herrup, S. Tonegawa, and V. E. Papaioannou.** 1992. RAG-1-deficient mice have no mature B and T lymphocytes. Cell **68:**869–877.
31. **Murre, C., P. S. McCaw, and D. Baltimore.** 1989. A new DNA binding and dimerization motif in Ig enhancer binding, daughter-

less, Myo D, and myc proteins. Cell **56:**777–783.
32. **Nelsen, B., G. Tian, B. Erman, J. Gregoire, R. Maki, B. Graves, and R. Sen.** 1993. Regulation of lymphoid-specific immunoglobulin μ heavy chain gene enhancer by ets-domain proteins. Science **261:**82–86.
33. **Oettinger, M. A., D. G. Schatz, C. Gorka, and D. Baltimore.** 1990. *RAG-1* and *RAG-2*, adjacent genes that synergistically activate V(D)J recombination. Science **248:**1517–1523.
34. **Paige, C. J., P. W. Kincade, and P. Ralph.** 1978. Murine B cell leukemia line with inducible surface immunoglobulin expression. J. Immunol. **121:**641–647.
35. **Park, K., and M. L. Atchison.** 1991. Isolation of a candidate repressor/activator, NF-E1(YY-1,δ), that binds to the immunoglobulin κ 3′ enhancer and the immunoglobulin heavy-chain μE1 site. Proc. Natl. Acad. Sci. USA **88:**9804–9808.
36. **Rauscher, F. J., III, J. F. Morris, O. E. Tournay, D. M. Cook, and T. Curran.** 1990. Binding of the Wilms' tumor locus zinc finger protein to the EGR-1 consensus sequence. Science **250:**1259–1262.
37. **Rivera, R. R., M. H. Stuiver, R. Steenbergen, and C. Murre.** 1993. Ets proteins: new factors that regulate immunoglobulin heavy-chain gene expression. Mol. Cell. Biol. **13:**7163–7169.
38. **Rosenberg, U., C. Schroder, A. Preiss, A. Kienlin, S. Cote, I. Riede, and H. Jäckle.** 1986. Structural homology of the product of the *Drosophila Krüppel* gene with *Xenopus* transcription factor IIIA. Nature (London) **319:**336–339.
39. **Sambrook, J., E. F. Fritsch, and T. Maniatis.** 1989. Molecular cloning: a laboratory manual, 2nd ed. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.
40. **Schatz, D. G., M. A. Oettinger, and D. Baltimore.** 1989. The V(D)J recombination activating gene, *RAG-1*. Cell **59:**1035–1048.
41. **Shi, Y., E. Seto, L. S. Chang, and T. Shenk.** 1991. Transcriptional repression by YY1, a human GLI-Krüppel-related protein, and relief of repression by adenovirus E1A protein. Cell **67:**377–388.
42. **Shinkai, Y., G. Rathbun, K. P. Lam, E. M. Oltz, V. Stewart, M. Mendelsohn, J. Charron, M. Datta, F. Young, A. M. Stall, and F. W. Alt.** 1992. RAG-2-deficient mice lack mature lymphocytes owing to inability to initiate V(D)J rearrangement. Cell **68:**855–867.
43. **Smale, S. T., and D. Baltimore.** 1989. The "initiator" as a transcription control element. Cell **57:**103–113.
44. **Small, S., R. Kraut, T. Hoey, R. Warrior, and M. Levine.** 1991. Transcriptional regulation of a pair-rule stripe in *Drosophila*. Genes Dev. **5:**827–839.
45. **Smith, D. B., and K. S. Johnson.** 1988. Single-step purification of polypeptides expressed in *Escherichia coli* as fusions with glutathione S-transferase. Gene **67:**31–40.
46. **Tautz, D., R. Lehmann, H. Schnürch, R. Schuh, E. Seifert, A. Kienlin, K. Jones, and H. Jäckle.** 1987. Finger protein of novel structure encoded by *hunchback*, a second member of the gap class of *Drosophila* segmentation genes. Nature (London) **327:**383–389.
47. **Travis, A., A. Amsterdam, C. Belanger, and R. Grosschedl.** 1991. *LEF-1*, a gene encoding a lymphoid-specific protein with an HMG domain, regulates T-cell receptor α enhancer function. Genes Dev. **5:**880–894.
48. **Waterman, M. L., W. H. Fischer, and K. A. Jones.** 1991. A thymus-specific member of the HMG protein family regulates the human T cell receptor C α enhancer. Genes Dev. **5:**656–669.
49. **Watson, D. K., M. J. McWilliams, P. Lapis, J. A. Lautenberger, C. W. Schweinfest, and T. S. Papas.** 1988. Mammaian *ets-1* and *ets-2* genes encode highly conserved proteins. Proc. Natl. Acad. Sci. USA **85:**7862–7866.
50. **Zenzie-Gregory, B., A. O'Shea-Greenfield, and S. T. Smale.** 1992. Similar mechanisms for transcription initiation mediated through a TATA box or an initiator element. J. Biol. Chem. **267:**2823–2830.