Erythroid-specific inhibition of the tal-1 intragenic promoter is due to binding of a repressor to a novel silencer

C. Courtes, N. Lecointe, L. Le Cam, Fulbert Baudoin, C. Sardet, D. Mathieu-Mahul

To cite this version:
C. Courtes, N. Lecointe, L. Le Cam, Fulbert Baudoin, C. Sardet, et al.. Erythroid-specific inhibition of the tal-1 intragenic promoter is due to binding of a repressor to a novel silencer. Journal of Biological Chemistry, American Society for Biochemistry and Molecular Biology, 2000, 275 (2), pp.949-958. 10.1074/jbc.275.2.949 . hal-02196811

HAL Id: hal-02196811
https://hal.archives-ouvertes.fr/hal-02196811
Submitted on 27 May 2021

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L’archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d’enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Distributed under a Creative Commons Attribution 4.0 International License
Erythroid-specific Inhibition of the tal-1 Intragenic Promoter Is Due to Binding of a Repressor to a Novel Silencer

(Received for publication, July 30, 1999, and in revised form, October 20, 1999)

Christine Courtes‡, Nathalie Lecointe, Laurent Le Cam§, Florence Baudoin, Claude Sardet, Danièle Mathieu-Mahul¶

From the Institut de Génétique Moléculaire, UMR 5535, IFR 24, 1919 Route de Mende, F 34293, Montpellier, France

The basic helix-loop-helix tal-1 gene plays a key role in hematopoiesis, and its expression is tightly controlled through alternative promoters and complex interactions of cis-acting regulatory elements. tal-1 is not expressed in normal T cells, but its transcription is constitutive in a large proportion of human T cell leukemias. We have previously described a downstream initiation of tal-1 transcription specifically associated with a subset of T cell leukemias that leads to the production of NH2-truncated TAL-1 proteins. In this study, we characterize the human promoter (promoter IV), embedded within a GC-rich region in exon IV, responsible for this transcriptional activity. The restriction of promoter IV usage is assured by a novel silencer element in the 3′-untranslated region of the human gene that represses its activity in erythroid but not in T cells. The silencer activity is mediated through binding of a tissue-specific nuclear factor to a novel protein recognition motif (designated tal-RE) in the silencer. Mutation of a single residue within the tal-RE abolishes both specific protein binding and silencing activity. Altogether, our results demonstrate that the tal-1 promoter IV is actively repressed in cells of the erythro-megakaryocytic lineage and that this repression is released in leukemic T cells, resulting in the expression of the tal-1 truncated transcript.

Differentiated hematopoietic cells have a limited life span and have to be continuously replenished from self-renewing pluripotent stem cells, which reside in the bone marrow and generate progenitor cells committed to proceed along one of the maturation pathways. To satisfy the variable needs of the different compartments, a tight control is required for the process of self-renewal, commitment, maturation, and survival for each differentiation stage within all of the lineages. Besides the critical role played by growth factors in regulating these processes, lineage-specific transcription factors are likely to control the expression of target genes that will confer the final cellular phenotype (for a review, see Ref. 1). Among members of the large family of the basic helix-loop-helix proteins (bHLH), which are known to control important steps in cell determination, differentiation, and growth (2, 3), tal-1 (also known as SCL) has been shown to play a pivotal role in the regulation of hematopoiesis (4–7).

Tal-1 was originally identified through its involvement in a rare chromosomal translocation specifically associated with human T cell acute leukemias (T-ALLs) (Refs. 8–12; for a review, see Ref. 13). In normal adult tissues, tal-1 expression is restricted to some hematopoietic tissues (8, 14, 15) and endothelial cells lining small blood vessels (16–18). Ectopic tal-1 expression is observed in about 50% of T-ALLs and is the most common genetic anomaly associated with this pathology (19). Indeed, in addition to chromosomal translocations, 12–28% of T-ALL patients harbor local tal-1 recombinations, termed tal+, which remove 90–100 kilobase pairs of upstream sequences from the tal-1 locus (11, 20). Finally, some T-ALLs and leukemic T cell lines display tal-1 expression without known tal-1 gene alteration (11, 21). An etiological role for tal-1 in T-ALL development is also supported by its strong similarity with two distinct bHLH genes, LYL-1 and tal-2, also involved in sporadic chromosomal translocations associated with T-ALLs (for a review, see Ref. 22). Moreover, the oncogenic potential of tal-1 in T cell lineage has been clearly demonstrated in transgenic mouse models (23, 24).

During hematopoiesis, tal-1 is expressed at low levels in early progenitor cells. Commitment to the erythroid and megakaryocytic lineages is correlated with an increase in tal-1 gene expression, whereas it is shut down in other cells, notably in the T lymphoid lineage (8, 14, 15, 17, 25). Gene targeting experiments have demonstrated that tal-1 is essential for early embryonic development (4, 5) and for the development of all hematopoietic lineages, including the T cell lineage (6, 26), indicating that tal-1 function is crucial in very early hematopoiesis. Indeed, several reports led to the conclusion that TAL-1 proteins, in concert with two other cell-specific factors, GATA-1 and LMO2, act as positive regulators of erythroid differentiation (27–29). Recent reports have shown that tal-1 also regulates the development of the vascular system (30–32). All of these observations strongly suggest that during embryonic development tal-1 activity may be required as early as the formation of hemangiblast, the common progenitor for both hematopoietic and endothelial lineages. Finally, tal-1 expression was also detected in midbrain and spinal cord of mouse embryos (17), and very recent transgenic studies in mice have identified enhancers that recapitulate physiological tal-1 expression in these tissues (13, 33).

Tal-1 encodes two major protein isoforms, full-length pp48–50 and N-terminally truncated pp24–28, both of which

EMSA, electrophoretic mobility shift assay; WT, wild type; PCR, polymerase chain reaction; LMPCR, ligation-mediated polymerase chain reaction.
possess the bHLH domain and heterodimerize with the ubiquitously expressed bHLH E-proteins, E47, E12, and HEB (11, 34, 35). Both types of TAL-1/E47 heterodimers were found to bind preferentially to the AACAAGATGGT E box (36), but only the pp48–50 TAL-1 species contains an amino-terminal transactivation domain (37). A putative tal-1 target gene, whose function is still unknown, has recently been identified (38).

Several studies have demonstrated that the expression of the tal-1 gene is itself highly regulated during adult hematopoiesis through complex interactions of cis-acting positive and negative regulatory elements. Two alternative 5′ promoters, Ia and Ib, have been identified in both human and murine genes. The activity of the promoter Ia is restricted to the erythroid and megakaryocytic lineages and is mediated by the erythroid transcription factor GATA-1 (39, 40). The promoter Ib was found to be active in primary myeloid cells through the action of PU-1, Sp1, and Sp3 transcription factors (41, 42) as well as in the leukemic T cells devoid of obvious chromosomal rearrangements (43). Finally, we have previously described a third downstream promoter, which appears to be specifically activated in a subset of human leukemic T cells (44).

In this study, we defined the third tal-1 promoter, termed promoter IV in this paper, that lies within the coding sequences of exon IV and does not by itself display cell type specificity. Moreover, we have identified a novel silencer element in the 3′-untranslated region of tal-1 (3′-UTR), which modulates the promoter IV activity in a manner indistinguishable from the regulation of the tal-1 promoter IV in hematopoietic cell lines. Promoter IV silencing is mediated through the binding of a novel tissue-specific nuclear factor to this silencer element. Altogether, our results demonstrate that the tal-1 promoter IV is actively repressed in cells of the erythro-megakaryocytic lineage and that this repression is released in leukemic T cells, resulting in expression of the tal-1 truncated transcript.

**EXPERIMENTAL PROCEDURES**

**Cell Cultures and Transfections**—All hematopoietic cells were maintained in RPMI 1640-Glutamax medium (Life Technologies, Inc.) supplemented with 10% fetal calf serum. HeLa and Swiss-3T3 cells were grown in Dulbecco’s modified Eagle’s medium (Life Technologies, Inc.) with 10% fetal calf serum. ECV304 cells were maintained in M199 (Life Technologies, Inc.) medium supplemented with 10% fetal calf serum and 10 μg Glutamax and human endothelial cells derived from umbilical cord in Medium 200 supplemented with low serum growth supplement fromCascade Biologicals, Inc. All cell lines were obtained from the ATCC, except human umbilical vein endothelial cells (from Cascade Biologicals). Frozen human primary T cells were kindly provided by Brigitte Kahn-Perles (INSERM U119, Marseille, France).

All transfections were performed as described previously (45) using DMRIE-C reagent (Life Technologies, Inc.) as per the manufacturer’s instructions. The TK-RL plasmid (Promega) encoding the Renilla luciferase gene was cotransfected with the plasmid of interest in each experiment to correct for transfection efficiency. 40 h post-transfection, the lacZ gene was cotransfected with the plasmid of interest (Promega). The two fragments (−127/+3427 and −127/+1330, in which residue +1 represents the first nucleotide following the stop codon) derived from tal-1 cDNAs were cloned into the BamHI–SalI sites of pGL3-basic promoter IV construct. The −127/+1330 fragment was also cloned in the antisense orientation into the BamHI site of pGL3 basic/promoter IV construct. The +562/+833 fragment was obtained by PCR amplification using the primers 5′-CGGGATCCCTGTCTTTCC-3′ (upper) and 5′-GGGTGGACGATCATGTAAG-3′ (lower) and subcloned into the BamHI–SalI sites of the pGL3 basic/−240/+114 promoter IV construct. Site-directed mutagenesis of residue 666 (G to A) was performed by a PCR-based method using the following primers: upper, 5′-GGTACAAAGTGCTACTTCTTCCTCCCAATG-3′; lower, 5′-CATTGAGGAGAAAATGAGTGTGCAC-3′.

All of these constructions were controlled by sequencing, carried out on double-stranded DNA with dye-terminator chemistry, and the products were resolved using an ABI Prism 377 automated sequencer.

In Vivo DMS Gemonic Footprinting—In vivo footprinting with the LMPCR procedure was performed essentially as described elsewhere (46, 47). Briefly, 107 cells were treated with the guanine methylation agent DMS at 0.2% for 5 min at room temperature in their medium (RPMI plus fetal calf serum). As a reference, genomic DNA from the same cell type was methylated in vitro with 0.5% DMS for 4 min at room temperature. Fingerprinting of methylated bases was performed in 1 M piperidine at 95 °C for 30 min. Two micrograms of cleaved genomic DNA were used for LMPCR carried out as described (46, 47). The following primers were designed to analyze the upper and lower strands of the 3′-UTR region of the tal-1 gene corresponding to the +290/+1090 fragment: SL1, 5′-GTCACTGCTTTTAGCTGG-3′; SL2, 5′-TTGGGATCCCTGTCTTTCCCTC3′; SL3, 5′-ATCTCCTGCGTTGCATTAAACGG-3′; ASG1, 5′-CAGAGAATGCAACAGAGG-3′; ASG2, 5′-ATCTCCTGCAACCAGGTG-3′; ASG3, 5′-TCAGACCAGCGTACCCGGCATCAGG-3′.

All reactions were performed in the presence of 10% MeSO. The PCR conditions were as follows: (i) first elongation, 15 min at 76 °C; (ii) PCR amplification, 4 min at 95 °C followed by 30 cycles of 1 min at 95 °C, 2 min at 58 °C, 3 min at 76 °C, ending with 7 min at 76 °C; (iii) labeling was performed by extension of 1 min at 95 °C followed by 2 min at 62 °C for SL3 or 58 °C for ASG3, 3 min at 76 °C in the presence of 32P-labeled SL3 or ASG3 primer (primers were 5′-end-labeled with T4 polynucleotide kinase (New England Biolabs)) and [γ-32P]ATP (NEN Life Science Products)). Specific activity of the labeled primer was 3–1010 cpm/pmol. Samples were run on a 5% sequencing gel at 10 kV. Dried gels were analyzed with a PhosphorImager (Molecular Dynamics, Inc., Sunnyvale, CA).

**Electrophoretic Mobility Shift Assay**—All nuclear extracts were prepared as described hereafter. 2 × 107 cells were centrifuged for 3 min at 1500 × g and incubated for 15 min on ice in four cell volumes in the hypotonic buffer A with 0.3 μM sucrose (60 μM KCl, 15 μM HEPES, pH 7.8, 15 μM NaCl, 14 μM β-mercaptoethanol, 0.15 μM spermine, 0.5 μM spermidine, protase inhibitors, and 1 mM dithiothreitol). Nonidet P-40 was added to a final concentration of 0.03% for T cells or 0.1% for other cells. Cell extracts were incubated for 1 min on ice and centrifuged for 10 min at 2000 × g through a 0.9 μM sucrose cushion in buffer A. Nuclei were resuspended in one nuclei packed volume of freezing buffer (75 μM NaCl, 0.5 mM EDTA, 20 μM HEPES, pH 7.8, 50% glycerol, protease inhibitors, and 1 mM dithiothreitol) and immediately frozen (−80 °C). For use, the nuclei were thawed on ice, and two volumes of extraction buffer were added (0.5 mM NaCl, 1.2 mM EDTA, 25 mM HEPES, pH 7.8, protease inhibitors and 1 mM dithiothreitol). Nuclear extracts were shaken gently for 15 min at 4 °C and centrifuged for 10 min at 10,000 g, and protein concentration of supernatants was quantified using the BCA protein assay reagent (Pierce).

[untitled]
The third tal-1 promoter lies within the coding sequences of exon IV and does not display tissue specificity. A, different tal-1 human genomic fragments were cloned upstream of the luciferase reporter gene and transfected in the Jurkat T cell line. The numbers refer to the position relative to the major transcription start site (+1). B, the −240/+114 promoter IV-basic construct was used for all transfections in the indicated cell lines. K562 and HEL are human erythroid-megakaryocytic cell lines; Jurkat, Rex, CEM, and HSB2 are tal-1-expressing human leukemic T cell lines; KIT 225 is a tal-1-negative human leukemic T cell line. Discrepancies in firefly luciferase activities due to transfection efficiency were normalized by comparison with Renilla luciferase activity driven by the co-transfected TK-RL reporter vector. The values were set relative to the activity of pGL3-basic plasmid devoid of promoter sequences (arbitrarily set at 1). All data represent the mean value ± S.D. of at least three independent experiments. Each experimental point was determined in triplicate. C, nucleotide sequence of the human promoter IV region. The DNA poly(dI-dC) (Roche Molecular Biochemicals). For DNA competition experiments, a 10–150 fold excess of unlabeled double-stranded oligonucleotide was added at this step. The reaction was then incubated for 10 min with the radioactive probe (150,000 cpm) and for another 15 min with 1 μg of BSA. Protein-DNA complexes were resolved in 4% acrylamide non-denaturing gels in 0.25× TBE at 4 °C, dried, and visualized by autoradiography.

UV Cross-linking—The lower strand oligonucleotide WT G666-3 (see Fig. 9C) was annealed to the 14-mer (5‘-GATCAAAGTTCTGC-3’) and filled with Klenow large fragment in the presence of 125 μT bromo-2-dUTP, dGTP, and dATP and 1 MBq of [α-32P]dCTP (111 TBoq/mol; Amersham Pharmacia Biotech). Interaction of the probe with cellular extracts in a 3-fold scaled up experiment and separation of the complex by electrophoresis were carried out as described for EMSA experiments, except that the gel was covered with Saran Wrap and autoradiographed without drying. The bands corresponding to the protein-DNA complex were isolated and exposed to a UV source (100 μJ/cm²) for 20 min. Slices were then boiled for 5 min in Laemmli buffer and run onto a 14% polyacrylamide-SDS denaturing gel, which was dried after electrophoresis and exposed for autoradiography.

RESULTS

Characterization of the Third tal-1 Promoter within the Coding Sequences of Exon IV—The 5′ truncated tal-1 transcripts observed in some leukemic T cells were presumed to result from transcriptional promotion downstream of the two major tal-1 promoters Ia and Ib (43, 48). To identify the genomic sequences responsible for the promotion of this transcription, 5′-deletion mutants of a human genomic segment encompassing the previously described transcription start site referred here as nucleotide +1 (48), were cloned upstream of the firefly luciferase reporter gene and tested by transient transfection assays in the Jurkat T cell line, which expresses both the full-length and the truncated tal-1 transcripts (43). Variations in transfection efficiency were controlled adjusting levels to the control vector. Complete deletion of intronic sequences contained in this DNA fragment (construct −240/+114) did not affect the promoter activity. Similarly, the deletion of the 5′ part of exon IV (construct −149/+114) preserved the promoter activity. Further deletion (up to nucleotide −73) slightly reduced the transcriptional potential of the gene segment (30–40% decrease), while the promoter activity was completely lost in the construct −28/+114. These results indicated that the fragment −149−28 contains elements that are essential for the promoter that lies within the coding sequences of tal-1. The −73/−28 fragment is highly G/C-rich in content and contains two potential Sp1-binding sites (Fig. 1C). Since this promoter is entirely embedded in exon IV, we refer to it as promoter IV in this paper.

Promoter IV per se Does Not Determine Tissue Specificity of Expression—We next assessed the tissue specificity of the tal-1 promoter IV in the context of the luciferase reporter gene. A panel of hematopoietic and non-hematopoietic cell lines (listed in Fig. 1B) were transiently transfected with the reporter −240/+114. Unexpectedly, this reporter construct was able to drive reporter gene transcription in all cell lines tested, although its efficiency varied from 7- to 30-fold of the activity of the control vector. Among hematopoietic cells tested, the reporter displayed promoter activity in cells belonging both to erythroid (K562, HEL), and T cell (Jurkat, Rex, CEM, KIT sequence (coding strand) of the −240/+114 segment is shown. The arrow shows the major transcription start site, indicated as nucleotide +1 in the reporter constructs. The two Sp1-like consensus sites are underlined.
225, HSB2) lineages. Thus, the activity of the promoter IV was not correlated with the expression of the endogenous truncated tal-1 transcript. Clearly, sequences identified as promoter IV are not sufficient to explain the cell type-specific expression of truncated tal-1 mRNA.

The 3′-UTR of the Human tal-1 Gene Functions as a Cell Type-specific Silencer on Promoter IV—The lack of cell type specificity of the promoter IV reporter constructs was discordant with our previous findings that endogenous promoter IV usage was restricted to leukemic T cells (43). To address this point, we checked whether other tal-1 gene sequences located outside of promoter IV could negatively regulate promoter IV activity in other hematopoietic cell lines, notably in erythroid cells.

Several reasons prompted us to first test whether a negative element might be located within the 3′-untranslated region of tal-1. Indeed, the immature hematopoietic DU528 cell line, characterized by an unusual t(1;14) chromosomal translocation resulting in the loss of most of the tal-1 3′-UTR sequences, displays high levels of tal-1 transcripts initiated exclusively from promoter IV (43, 48). Thus, this chromosomal rearrangement might have deleted a negative element controlling promoter IV activity in hematopoietic cells. Secondly, the 3′-UTR is unusually large (about 3500 bp), and sequence analysis revealed the presence of numerous potential binding sites for transcription factors, such as GATA, bHLH, Ikaros/LyF, and ETS, all regulators of hematopoietic differentiation (for a review, see Ref. 1).

To test this possibility, various restriction fragments derived from the tal-1 3′-UTR were placed downstream of the luciferase gene driven by the promoter IV (construct -240/+114 used in Fig. 1). The constructs were then tested in transient transfections in K562 and Jurkat cells. The level of luciferase activity generated by each reporter was compared with that of the promoter IV construct containing no additional sequences. The results of this functional assay (Fig. 2A) showed that portions of the 3′-UTR of tal-1, placed in cis with the promoter IV strongly inhibited luciferase gene expression in K562 (up to 70% inhibition). In contrast, in Jurkat cells, the same 3′-UTR fragments either slightly enhanced promoter IV activity (1.5-fold) or were without effect. Removal of a large part of the 3′ sequences of 3′-UTR (nucleotides 1089–3427, with nucleotide 1 designated here as the first nucleotide following the stop codon) or the 5′ sequences (nucleotides -127/+318), did not significantly alter the repression of promoter IV. Similar inhibition of promoter IV expression was observed when the 3′-UTR fragments were in the inverted orientation. Finally, the inhibitory function of the 3′-UTR was also observed when it was inserted in either orientation upstream of the promoter IV (see Fig. 2A).

Altogether, our data demonstrate that a cis-regulatory element lying in the 319–1089 gene segment beyond the stop codon of the tal-1 gene represses promoter IV in K562 independently of its orientation and of its position relative to the transcription start site. Importantly, the 3′-UTR repressive element does not block promoter IV activity in Jurkat T cells, in agreement with the observation that promoter IV is active in human leukemic T cells.

We then assessed the tissue specificity of the silencing mediated by the 3′-UTR element by testing its activity in a panel of hematopoietic and nonhematopoietic cells. As shown in Fig. 2B, the 319–1089 3′-UTR segment efficiently repressed promoter IV in the erythroid cell line HEL. In contrast, no inhibition was observed in nonhematopoietic cells and in transformed T cell lines; in fact, the construct displayed a positive effect on promoter IV activity in most of the T cell lines tested (up to 2-fold activation in the HSB2 and KIT 225 T cells. These data show that the activity of the promoter IV-Luc-3′-UTR construct recapitulates exon IV-initiated tal-1 transcription in human hematopoietic cells.

The tal-1 3′-UTR Functions as a Cell Type-specific Silencing Element on a Heterologous Promoter—We next tested whether the repression mediated by the 3′-UTR of tal-1 was specific for the promoter IV. Several fragments derived from the 3′-UTR were introduced into a vector downstream of the luciferase gene driven by the SV40 promoter and tested in different cell lines. As shown in Fig. 3, the constructs strongly inhibited the SV40 promoter in K562 cells (up to 80% inhibition), while they had no effect in Jurkat T cells. These data clearly indicate that the silencing element functions in a cell type-specific manner but is not specific for the tal-1 promoter IV.
Occupation of a Novel Factor-binding Site Correlates with the Silencing Activity of the 3'-UTR—The sequence of the 290–1090 fragment (see Fig. 4) contains several possible binding sites for nuclear factors that have been shown to control crucial steps of hematopoietic development. Given that a tissue-specific silencing element was contained in the 319–1090 fragment of the 3'-UTR, we examined binding of specific nuclear proteins to this DNA segment. To address this question, we first performed a comparative analysis of the in vivo DNA-protein contacts that occur on the 3'-UTR of tal-1, using an in vivo genomic footprinting technique based on dimethyl sulfoxide/ligation-mediated PCR. This technique was applied simultaneously to DNA prepared from K562 and Jurkat cells. Guided by the presence of these putative binding sites, we designed our LMPCR primers to cover two gene segments of interest: the first surrounding an array of three consensus GATA sites (nucleotides 775–807) and the second containing several consensus-like Ik2/Lyf-binding sites (nucleotides 320–770). DNA-protein contacts that differed in the two cell lines were studied in detail, since they could account for the cell type-specific silencing activity of the 3'-UTR. Comparison of the cleavage pattern of in vivo DMS-methylated K562 and Jurkat DNA with naked DNA revealed no significant difference in the protection/occupation of the GATA sites (nucleotides 775–807; data not shown). In contrast, our analysis clearly detected a differential protection/occupation of a G residue (nucleotides 666 indicated by an arrow in Fig. 4) preceded by a hypersensitive site in K562 (noted by two asterisks in Fig. 4). No obvious protection of G residues in the same area was observed on the opposite strand (not shown). Importantly, this DNA-protein contact was not observed on the Jurkat tal-1 3'-UTR.

The differential protection/occupation strongly supports the notion that this sequence might be involved in the binding of a repressor complex. Although computer analysis of the protected area failed to identify known binding sites, it was flanked by two Ik2/Lyf core-like sequences in opposite orientation (GGGATTC). This cell type-specific protected sequence was termed tal-RE (for tal-1 repressive element). Identification of a Specific tal-RE Binding Activity in K562 Cells—To identify a cellular tal-RE binding activity, we then carried out EMSAs using K562 nuclear extracts and a probe including the two Lyf-core motifs and the central region identified as tal-RE (WT G666-1, see Fig. 9C). Several in vitro DNA protein complexes were detected (Fig. 5A); the four major ones were termed I, IIa, IIb, and III. Competition experiments showed that the four complexes were specific for the tal-1 probe, since they were competed by a 50-fold excess of unlabeled wild type probe but were unaffected by a 100-fold excess of two unrelated oligonucleotides (Sp1 and E2F). Additional competition experiments showed that the complex III was specific for the tal-RE, whereas the large complexes I, IIa, and IIb were not. Indeed, competition with an unlabeled double-stranded oligonucleotide bearing a mutation in the tal-RE (mut A666, Fig. 9C) did not affect the formation of complex III even at a 100-fold excess but eliminated the formation of the large complexes I, IIa, and IIb (Fig. 5A). Conversely, the mut A666 oligonucleotide used as the EMSA probe generated the large complexes I, IIa, and IIb but failed to form the complex III (see Fig. 5B). These data clearly indicate that the critical residue G666, which is protected in vivo in K562 cells, is involved in the complex III formation. The pattern of the DNA-protein complexes obtained with Jurkat nuclear extracts was quite different from that of K562 (see Fig. 5B). A major complex migrated at the same position as K562 complex I, while both IIa and IIb complexes were strongly reduced and the tal-RE specific complex III was absent.

Altogether, these experiments show that K562 cells in which the tal-1 promoter IV is repressed contain a specific tal-RE binding activity that can be detected as a single DNA-protein complex (complex III). A point mutation in the tal-RE (G666 to A) abolishes the complex III formation. This is in perfect agreement with our footprint experiments, showing that the tal-RE occupancy in K562 cells requires the central G666.

Mutation of Residue 666 (G to A) within the tal-RE Abolishes the Silencing Effect of the 3'-UTR on Promoter IV—The above data strongly suggested that complex III formation on the tal-RE might account for the silencing activity of the 3' UTR in K562. To assess the functional relevance of the tal-RE in vivo, we introduced the point mutation (G666 to A) found to disrupt complex III in the above EMSA experiments, within the 319–1089 wild type reporter construct. As shown in Fig. 6, the mutated 319–1089 fragment showed no inhibition of promoter IV when introduced in K562 cells, clearly demonstrating that the in vivo protected residue G666 is critical for the silencing activity of the 3'-UTR. A shorter gene segment (561–833) encompassing the tal-RE flanked by the Lyf motifs and the array of the GATA-like sites (as indicated in Fig. 4) was also tested and found to have no effect on promoter IV activity in K562 cells (Fig. 6). These data show that the tal-RE is essential but appears to be insufficient to confer silencing activity to the 3'-UTR in K562 cells.

tal-RE Binding Activity Exhibits Cell Type Specificity—To obtain further information about the tissue specificity of the nuclear factor(s) binding to the tal-RE and leading to complex III formation, EMSA experiments were performed using the G666-1 (see Fig. 9C) as the probe with nuclear extracts from a variety of cell lines. As shown in Fig. 7, the patterns of DNA-protein complexes were quite heterogenous between the different cell lines, particularly for the slowly migrating complexes. However, a DNA-protein complex, migrating at the same position as K562 complex III, was clearly detected with nuclear extracts from cells belonging to the erythroid-megakaryocytic lineage (HEL, UT7, A745). The identity of complex III was confirmed by competition experiments with a 100-fold excess of the mut A666 (see Fig. 9C) unlabeled oligonucleotide, which did not affect its formation (data not shown). In contrast, no retarded band migrating as complex III was formed with nuclear extracts from any of the T cell lines tested, primary T cells, or nonhematopoietic cells. Thus, the cell type distribution of tal-RE-binding, as reflected by complex III formation, is in perfect agreement with the transcriptional repression activity of the 3'-UTR (Fig. 2B).

Preliminary Characterization of the Polypeptide Binding to
the tal-RE—To characterize the nuclear protein(s) that bind to the tal-RE, we performed UV cross-linking experiments, using a probe labeled with $^{32}$P- and bromodeoxyuridine (oligonucleotide G666-1; see Fig. 9C). The experiments were carried out with nuclear extracts from both K562 and A745 cells, since the latter displayed a very high tal-RE binding activity (see Fig. 7). The bands corresponding to the tal-RE complex formed with K562 and A745 nuclear extracts were cut out, exposed for 20 min to UV, and loaded onto polyacrylamide-SDS denaturing gel. As shown in Fig. 8, an identical band migrating at the position 55 kDa was observed in the lanes corresponding to both cell extracts. The addition of the nonlabeled double-mutated oligonucleotide (mut A666/667, see Fig. 9C) to the binding reaction did not modify the protein content of the complex. The effective molecular mass obtained after subtracting the mass of the double-stranded oligonucleotide was estimated at around 40 kDa. No additional bands were visible even after long exposure of the gel, suggesting that the tal-RE-binding complex most likely consists of a single protein species.

Further Characterization of the tal-RE Sequence Involved in Protein Binding—In order to more precisely identify tal-RE nucleotides involved in the in vitro binding of the nuclear factor(s), we first determined the minimum size of the double-stranded oligonucleotide required for efficient binding. We performed UV cross-linking experiments, using a probe labeled with $^{32}$P- and bromodeoxyuridine (oligonucleotide G666-1; see Fig. 9C). The experiments were carried out with nuclear extracts from both K562 and A745 cells, since the latter displayed a very high tal-RE binding activity (see Fig. 7). The bands corresponding to the tal-RE complex formed with K562 and A745 nuclear extracts were cut out, exposed for 20 min to UV, and loaded onto polyacrylamide-SDS denaturing gel. As shown in Fig. 8, an identical band migrating at the position 55 kDa was observed in the lanes corresponding to both cell extracts. The addition of the nonlabeled double-mutated oligonucleotide (mut A666/667, see Fig. 9C) to the binding reaction did not modify the protein content of the complex. The effective molecular mass obtained after subtracting the mass of the double-stranded oligonucleotide was estimated at around 40 kDa. No additional bands were visible even after long exposure of the gel, suggesting that the tal-RE-binding complex most likely consists of a single protein species.
Fig. 6. Mutation of residue 666 (G to A) within the tal-RE abolishes the silencing activity of the 3'-UTR. Wild type, mutant, or deleted segments from the 3'-UTR of the human tal-1 gene as indicated (residue +1 represents the first nucleotide following the stop codon) were cloned downstream of the polyadenylation site of the luciferase gene driven by the promoter IV and transiently transfected into K562 cell line. Discrepancies in firefly luciferase activities due to transfection efficiency were normalized by comparison with Renilla luciferase activity driven by the co-transfected TK-RL reporter vector. The values are expressed as the percentage of the activity of the promoter IV-basic construct containing only the promoter IV. All data represent the mean value ± S.D. of at least three independent experiments. Each experimental point was determined in triplicate.

formed EMSAs with nuclear extracts from both K562 and A745 cells, using a series of truncated oligonucleotides derived from the original 39-mer G666-1 (see Fig. 9C) and followed the formation of complex III. As shown in Fig. 9A, shortening the probe to 29 or 24 base pairs (G666-2 and G666-3, respectively) had no effect on the binding, whereas the band corresponding to the tal-RE-protein complex formed with the 19-mer G666-4 appeared with a reduced intensity, suggesting that the minimal size for optimal in vitro tal-RE-binding was the 24-mer G666-3 in which the G666 occupied the central position. Therefore, individual point mutations on residues flanking the G666 were introduced in the oligonucleotide G666-3, as shown in Fig. 9C. These mutated double-stranded DNAs were used in EMSA as cold competitors against the G666-3 wild type probe. Fig. 9B presents the results obtained with A745 nuclear extract, although identical results were observed with K562 nuclear extract (not shown). As shown at the top of Fig. 9B, three mutated oligonucleotides in addition to the previously used A666, displayed either no competition at all (A667) or ineffective competition (A663, A669). Consistent with these findings, neither labeled A663, A667, nor A669, formed an observable complex with both cell extracts (data not shown). To further extend the sequence required for protein binding, we tested additional point mutations of nucleotides more distant from G666, as shown in Fig. 9B (bottom), both mutations A661 and A670 strongly disrupted the competition effect of the cold oligonucleotide, as well as mutations A662 and A671, although less efficiently. On the other hand, oligonucleotides mutated at a more distant position, 657–660 and 672–676 (only some of them are shown in the figure), competed as efficiently as the wild type oligonucleotide, indicating that they are not directly involved in the factor binding. Interestingly, the quadruple mutation of those residues 673–676 deleted in the 19-mer WT G666-4 did not affect the competition effect of the cold oligonucleotide.

From all of these experiments, we deduced the tal-RE binding sequence as 5'-GTTNGCNNTC-3', which requires 5 nucleotides of flanking sequences on both sides to be efficiently recognized in vitro by the tal-RE-binding protein(s).

DISCUSSION

In this study, we investigated molecular mechanisms underlying tal-1 gene expression, focusing on the origin of an intragenic transcription initiation that occurs exclusively in a significant proportion of human leukemic T cells and leads to the production of the short TAL-1 proteins (11, 43). Transient transfections and mutational deletions allowed us to map the sequences responsible for this activity in the middle of the coding sequences of exon IV, embedded within a highly GC-rich region (80% G/C content). There is no TATA box consensus at the expected position (around ~30) from the major transcription start site; however, the sequence TTAAA between residues –24 and –20 could be a TATA-like element. In addition, two Sp1-like consensus sites are present, strongly suggesting that promoter IV initiation might be mediated through the binding of Sp1-like proteins. Several transcription start sites within exon IV are observed in the DU528 cell line, which may reflect the presence of a strong secondary structure in this region. All of these structural features are usually associated with the so-called housekeeping genes. Indeed, we found that promoter IV itself functions in a non-tissue-specific manner when tested in transient transfection experiments. Since these findings did not reflect the restricted profile of tal-1 truncated transcripts among distinct hematopoietic lineages, we searched for additional regulatory sequences located in a region distinct from the 5’ promoter IV region.

A negative regulatory element was identified in the 3'-UTR of the human tal-1 gene which inhibits promoter IV activity in erythroid but not in T cells. The 3'-UTR regulatory element can be defined as a classical silencer (49) based on its ability (i) to repress promoter IV as well as a heterologous promoter, (ii) to function in a position-independent manner, and (iii) to inhibit transcription when placed in either orientation with respect to the transcriptional start site.

Intragenic transcriptional regulatory sequences have been found in numerous genes and, in the vast majority, these elements map within DNA regions close to the transcription start site. In contrast, the promoter IV silencer is located at a rather long distance away from the 5’ promoter IV region (8 kilobase pairs). Although rare, this situation has been described for other silencers controlling the expression of genes coding for the T cell receptor a-chain (50), CD4 (51), and interleukin-4 (52). Examples of cis-regulatory sequences located in the 3'-UTR are even rarer. In eukaryotes, the 3'-UTR is usually seen as a key repository of information for regulating mRNAs in the cytoplasm, since it contains signals for controlling translation, stability, and localization (reviewed in Ref. 53). Transcriptional silencing is a rare function for a 3'-UTR that, to our knowledge, has been previously described for only two other loci, interleukin-4 (52) and a rat serine protease inhibitor (54).

Although silencers have been functionally demonstrated in numerous vertebrate genes, the question of how they act on the target promoters is still unanswered. We have identified a cell type-specific protein-DNA complex (tal-RE binding activity) that correlates with the silencer activity in vitro. Mutation of a single residue within the tal-RE (G666 to A) results in a loss of both protein binding and silencer activity, demonstrating that the tal-RE-binding protein(s) plays a central role in regulating silencer function in erythroid cells. We have defined the sequence required for tal-RE-binding (GTTNGCNNTC) that appears to be a novel protein recognition motif. UV cross-linking experiments have allowed the initial characterization of a cell-specific nuclear factor(s) of about 40 kDa that interacts with the tal-RE. Given the absence in tal-RE of a palindromic sequence usually associated with the binding of homodimeric polypeptide, the binding factor(s) probably does not bind as a homodimer. Two inverted direct repeats (5'-TGGGGA...TC-CCC-3') are flanking the tal-RE that may confer some sec-

2 C. Courtes, unpublished results.
were performed with a 32P-labeled G666-1 probe (see Fig. 9C), and nuclear extracts were prepared from the indicated cell lines. The same amount of nuclear extract (3 μg) was used in all lanes. K562, HEL, and UT7 are human erythro-megakaryocytic cell lines; A745 is a murine erythroid cell line; Jurkat, Rex, CEM, and DUS82 are tal-1-expressing human leukemic T cell lines; KIT 225 and MOLT-13 are tal-1-negative human T cell lines; quiescent (NS) or PHA-stimulated (S) human primary T lymphocytes are shown; HUVEC are primary human umbilical vein endothelial cells; and Swiss-3T3 are murine fibroblasts.

secondary structure to the surrounding proximal DNA. Interestingly, these direct inverted repeats match IK2/LyF core-like consensus. Several studies have reported that Ikaros complexes are localized to discrete heterochromatin foci associated with transcriptionally silent genes (55, 56), and more recently, Ikaros proteins have been described as being able to target chromatin remodeling and deacetylation complexes [57]. Although we failed to detect the presence of proteins on these sequences by in vivo footprinting, it will be of interest to test the possibility that some members of the Ikaros family might cooperate with the tal-RE-binding factor in the silencing activity of the tal-1 3′-UTR.

The mechanisms and factors involved in the repression of eukaryotic genes have not been studied as extensively as those involved in activation. Repressors appear to act by either modifying chromatin or inhibiting some basal processes in transcription initiation and/or elongation. The 3′-UTR DNA segment 562–833, in which the tal-RE occupies a central position, is not able to inhibit promoter IV activity in K562 cells. One possibility is that this short fragment cannot adopt a structural configuration required for specific interaction between tal-1 promoter IV and the repressor element. Another explanation could be that the tal-RE binding activity, although essential, is not sufficient to confer the repressor activity to the 3′-UTR, supporting the notion that the 3′-UTR silencing activity probably operates through a complex set of regulatory elements. As judged by its strong negative effect on SV40 promoter in erythroid cells, the 3′-UTR silencing element probably recruits cell-specific repressors distinct from the tal-RE binding factor that directly interfere with the general transcriptional machinery. GATA consensus sequences in the silencer region are required for the developmental expression of the human ε-globin gene (59, 60) as well as for the activity of the CD4 silencer (51), and more recently, the erythroid GATA-1 factor has been described as acting as an in vivo repressor of human ε-globin gene expression (61). Several potential GATA sites are also present in different domains of the 3′-UTR of the human tal-1 gene. This common occurrence raises the possibility that the erythroid-specific GATA-1 factor might be involved in the 3′-UTR silencer function. Further analysis will thus be required to identify possible additional cis-elements, distinct from tal-RE, as well as the trans-acting factors that might be crucial for the transcriptional repressor activity of the silencer.

Whatever the mechanism of transcription repression through the silencer domain, it is clear that the factor binding to the tal-RE is of central importance for regulating tal-1 promoter IV in human erythroid cells. tal-RE binding activity is detected in all cells tested of the erythro-megakaryocytic lineages that normally express tal-1 from promoter Ia (39, 40, 62). The absence of the silencer activity in all T cell lines tested, regardless of their tal-1 expression, as well as in nonhematopoietic cells coincides perfectly with the lack of tal-RE-interacting protein. In tissues that do not normally transcribe tal-1, the peculiar chromatin structure associated with the highly GC-rich content of the human tal-1 locus, notably hypermethylated (63), probably renders the locus inaccessible to Sp1-like transcription factors and thus does not allow transcription initiation at the endogenous promoter IV; there is no need to repress promoter IV activity in these cells, thus in agreement with the absence of silencing activity of the 3′-UTR in these cells.

The absence of tal-RE-binding proteins in leukemic T cells cannot be correlated to the leukemic phenotype, since human
primary T cells do not display any tal-RE binding activity either. Promoter IV has been found to operate in a subset of human T-ALLs with the exception of those associated with tal\(^4\) deletions (43), implying that truncated TAL-1 proteins participate in T cell oncogenesis. Indeed, in a transgenic mouse model, pp24–28 TAL-1 proteins were found to cooperate with LMO-1 to generate T cell malignancies (44). Constitutive expression of TAL-1 proteins in an immature T cell is likely to exert a dual effect. First, the inappropriate formation of TAL-1\(\eta\)-E47 heterodimers diverts the E47 proteins from their normal target DNA (64) and, as a consequence, inhibits their normal functions in early T cell development (Ref. 65; for a review, see Ref. 13). In that context, both large and short TAL-1 protein expression are likely to produce the same effects, since they share the same dimerization partners. Constitutive expression of pp48–50 TAL-1 in T lineage cells has been shown to prevent apoptosis (66, 67) and to regulate new target genes in association with GATA-3 (68, 69). It is unlikely that short TAL-1 proteins display the same biological function, since the transactivation \(\text{NH}_2\) domain of TAL-1 is probably involved in both activities.

This study demonstrates that one of the functions of the human tal-\(I\) 3'-UTR is to inhibit the expression of transcripts encoding the pp24–28 TAL-1 proteins in cells from erythroid and megakaryocytic lineages. pp48–50 TAL-1 proteins have been reported to act as positive regulators of erythroid differentiation (14, 28, 70), while the biological function of pp24–28 is still unknown. In close association with two other erythroid factors, LMO2 and GATA-1, pp48–50 is likely to activate a set of specific genes required for normal erythroid differentiation (38, 71). Since pp24–28 proteins lack the amino-terminal transactivation domain, it is tempting to speculate that pp24–28 heterodimers might not be efficient activators of transcription and consequently could interfere with normal erythroid development.

The 3'-UTR silencer element identified here illustrates a mechanism for excluding tal-\(I\) promoter IV usage in the erythroid lineage. Given the tissue-restricted distribution of tal-RE binding activity, it will be of importance to identify the tal-RE-binding protein and establish its role in erythroid cell differentiation.

**Acknowledgments**—We are grateful to Urszula Hibner and Naomi Taylor for critical comments on the manuscript and to Jean-Marie Blanchard and Alexandre Philips for helpful discussions. We greatly appreciate the excellent technical assistance of Christianne Dohet and Cécile Lambert. We are greatly indebted to Brigitte Kahn-Perles for providing human primary T cells.

**REFERENCES**

1. Orkin, S. H. (1995) *Curr. Opin. Cell Biol.* 7, 870–877
2. Murre, C., McCaw, P. S., Vaessin, H., Caudy, M., Jan, L. Y., Jan, Y. N., Cabrera, C. V., Buskin, J. N., Hausschka, S. D., Lasser, A. B., Weinstein, H., and Baltimore, D. (1999) *Cell* 58, 537–544
3. Murre, C., McCaw, P. S., and Baltimore, D. (1989) *Cell* 56, 777–783
4. Robb, L., Lyons, I., Li, R., Hartley, L., Kontgen, F., Harvey, R. P., Metcalf, D., and Begley, C. G. (1995) *Proc. Natl. Acad. Sci. U. S. A.* 92, 7075–7079
5. Shiodasani, R. A., Mayer, E. L., and Orkin, S. H. (1995) *Nature* 373, 432–434
6. Purche, C., Swat, W., Rockwell, K., Fujiwara, Y., Alt, F. W., and Orkin, S. H. (1996) *Cell* 86, 47–57
7. Robb, L., Elwood, N. J., Elefanty, A. G., Kontgen, F., Li, R., Barnett, L. D., and Begley, C. G. (1996) *EMBO J.* 15, 4123–4129
8. Begley, C. G., Aplan, P. D., Denning, S. M., Haynes, B. F., Waldmann, T. A., and Kirsch, I. R. (1998) *Proc. Natl. Acad. Sci. U. S. A.* 95, 10128–10132
9. Aplan, P. D., Lombardi, D. P., Ginsberg, A. M., Cosman, J., Bertness, V. L., in EMSA experiments. The small arrow denotes the position of the G\(^{666}\) found to be occupied *in vivo* in K562. The dotted lines in the sequence of the mutant oligonucleotides indicate that, with the exception of the indicated residue, all other nucleotides are conserved. The Tal-RE consensus was deduced from the EMSA competition experiments shown in B.

**Fig. 9.** Preliminary characterization of the DNA sequences required for tal-RE binding. **A**, EMSAs were carried out with either A745 or K562 nuclear extract using the indicated \(32\)P-labeled WT probe described in C. Note that the faint complex formed with the G\(^{666}-4\) (19-nucleotide) probe displays a slightly slower migration. **B**, in all lanes, EMSAs were performed using the \(32\)P-labeled G\(^{666}-3\) WT probe and A745 nuclear extracts. Increasing amounts of the indicated WT or mutated cold competitors were added to the binding reactions. The specific tal-RE-binding complex III formed with A745 nuclear extract is denoted by an arrow. **C**, sequence of the different oligonucleotides used.
