Chromatin immunoselection defines a TAL-1 target gene

Sylvia Cohen-Kaminsky,
Leïla Maouche-Chrétien, Luigi Vitelli¹,
Marie-Antoinette Vinit, Isabelle Blanchard,
Masayugi Yamamoto², Cesare Peschle^{1,3}
and Paul-Henri Roméo⁴

INSERM, U474, Hématologie Moléculaire, Hôpital Henri Mondor,
51 avenue du Maréchal de Lattre de Tassigny, 94010 Créteil, France,
¹Istituto Superiore di Sanita, Department of Hematology and
Oncology, Viale Regina Elena, 299, 00161 Rome, Italy, ²Centre for
Tsukuba Advanced Research and Institute of Basic Medical Sciences,
University of Tsukuba, Tsukuba 305, Japan and ³T.Jefferson
University, Kimmel Cancer Center, BLSB, 233 South 10th Street,
Philadelphia, PA 19107-5541, USA

⁴Corresponding author
e-mail: romeo@im3.inserm.fr

S.Cohen-Kaminsky and L.Maouche-Chrétien contributed equally to
this work

Despite the major functions of the basic helix–loop–helix transcription factor TAL-1 in hematopoiesis and T-cell leukemogenesis, no TAL-1 target gene has been identified. Using immunoprecipitation of genomic fragments bound to TAL-1 in the chromatin of murine erythro-leukemia (MEL) cells, we found that 10% of the immunoselected fragments contained a CAGATG or a CAGGTG E-box, followed by a GATA site. We studied one of these fragments containing two E-boxes, CAGATG and CAGGTC, followed by a GATA motif, and showed that TAL-1 binds to the CAGGTG E-box with an affinity modulated by the CAGATG or the GATA site, and that the CAGGTG–GATA motif exhibits positive transcriptional activity in MEL but not in HeLa cells. This immunoselected sequence is located within an intron of a new gene co-expressed with TAL-1 in endothelial and erythroid cells, but not expressed in fibroblasts or adult liver where no TAL-1 mRNA was detected. Finally, *in vitro* differentiation of embryonic stem cells towards the erythro/megakaryocytic pathways showed that the TAL-1 target gene expression followed TAL-1 and GATA-1 expression. These results establish that TAL-1 is likely to activate its target genes through a complex that binds an E-box–GATA motif and define the first gene regulated by TAL-1.

Keywords: chromatin immunoprecipitation/GATA motif/hematopoiesis/MEL cells/TAL-1 target genes

Introduction

The *tal-1* gene (also known as *scl* or *TCL-5*), was identified via a chromosomal translocation involving human chromosome 1p32 in T-cell acute lymphocytic leukemia (T-ALL) (Begley *et al.*, 1989; Finger *et al.*, 1989; Bernard *et al.*,

1990; Chen *et al.*, 1990). Further studies have shown that aberrant activation of the *tal-1* gene occurs in most patients with T-ALL (Bash *et al.*, 1995) with diverse molecular mechanisms, including chromosomal translocations, interstitial deletions and transactivation without detectable chromosomal alteration. Finally transgenic mice have demonstrated a direct role of TAL-1 in T-cell leukemogenesis (Condorelli *et al.*, 1996; Kelliher *et al.*, 1996; Larson *et al.*, 1996; Aplan *et al.*, 1997). The normal function of TAL-1 is the regulation of hematopoiesis (Green, 1996; Shivdasani and Orkin, 1996) and embryonic angiogenesis (Visvader *et al.*, 1998). Mice homozygous for deletion of the *tal-1* gene die *in utero* by embryonic day 10 as a result of the absence of blood formation (Robb *et al.*, 1995; Shivdasani *et al.*, 1995), indicating that TAL-1 has a critical function very early during hematopoietic differentiation. Furthermore, *tal-1* null embryonic stem (ES) cells failed to produce any hematopoietic lineage and to contribute to hematopoiesis in chimeric mice, implying that TAL-1 is also critical for definitive hematopoiesis (Porcher *et al.*, 1996; Robb *et al.*, 1996). Finally, transgenic rescue of hematopoietic defects of *tal-1*^{−/−} embryos has established a role for TAL-1 in angiogenesis (Visvader *et al.*, 1998).

TAL-1 belongs to the bHLH class of transcription factors (Baer, 1993). The bHLH domain allows both the formation of protein dimers and sequence-specific DNA recognition, and bHLH dimers bind to E-box motifs which have the general sequence CANNTG (Murre *et al.*, 1989). TAL-1 proteins cannot bind to DNA as homodimers but are able to bind DNA after dimerization with the ubiquitously expressed E2A gene products, E47 and E12 (Hsu *et al.*, 1991, 1994b). CASTing experiments have shown that the TAL-1/E2A heterodimers bind to the preferred sequence 5′-AACAGATGGT-3′ (Hsu *et al.*, 1994a). Using an artificial reporter gene containing multiple copies of the TAL-1/E2A binding sequence, transcriptional activity of the TAL-1/E47 heterodimer was examined in transiently transfected murine fibroblasts. While the E47 homodimer strongly activated the reporter gene, the TAL-1/E47 heterodimer was much less active, suggesting either a negative or a weak positive regulatory role of TAL-1 (Hsu *et al.*, 1994c).

The pivotal functions of TAL-1 may be accomplished by DNA binding and tight co-operation with at least two partners, the LIM-only protein LMO2 and GATA-1. TAL-1 and LMO2 genes are co-expressed during primitive and definitive hematopoiesis (Green *et al.*, 1992; Kallianpur *et al.*, 1994; Robb *et al.*, 1996), are activated in T-ALL (Begley *et al.*, 1989; Boehm *et al.*, 1991; Rabbitts, 1994), and the two proteins are functionally related. Furthermore, like TAL-1, LMO2 is also a T-cell oncogene when expressed as a transgene in mice (Boehm *et al.*, 1991; Royer-Pokora *et al.*, 1991), and LMO2^{−/−} and TAL-1^{−/−}

mice display the same phenotype (Warren *et al.*, 1994). The other potential partner of TAL-1 is GATA-1, a transcription factor present in all stages of vertebrate erythropoiesis. GATA-1 binds to the WGATAR motif present in the promoters and enhancers of several erythroid-specific genes (reviewed in Weiss and Orkin, 1995), and exerts positive or negative transactivation effects (Aird *et al.*, 1994; Raich *et al.*, 1995; Weiss and Orkin, 1995; Briegel *et al.*, 1996). An essential role for GATA-1 in erythropoiesis has been established through gene disruption. Indeed GATA-1^{-/-} ES cells fail to contribute to the mature erythroid compartment in chimeric mice (Pevny *et al.*, 1991) while expression of a normal GATA-1 transgene in the mutant ES cells rescues erythroid development (Simon *et al.*, 1992). The absence of erythropoiesis obtained in the TAL-1, LMO2 and GATA-1 null mutations suggests that these three proteins have closely related roles in erythroid differentiation. As LMO2 can bind to both GATA-1 and TAL-1 in erythroid cell lines, it has been postulated that it may act as a physical bridge between TAL-1 and GATA-1 proteins to settle a multimolecular complex (Osada *et al.*, 1995). Recent CASTing experiments have defined a CAGGTG-GATA bipartite DNA motif (Wadman *et al.*, 1997) that binds a multimolecular complex including TAL-1, E47, LMO2, GATA-1 and the LIM-domain binding protein Ldb1 (Agulnick *et al.*, 1996; Jurata *et al.*, 1996), and generates a positive transcriptional activity (Wadman *et al.*, 1997). Therefore, co-operation between TAL-1, LMO2 and a member of the GATA family may be involved in the regulation of TAL-1 target genes.

Despite its major role in hematopoiesis (Green, 1996; Shivdasani and Orkin, 1996) and leukemogenesis (Aplan *et al.*, 1992a, 1997; Condorelli *et al.*, 1996; Kelliher *et al.*, 1996; Larson *et al.*, 1996), as well as the identification of interacting partners (Valge-Archer *et al.*, 1994; Wadman *et al.*, 1994; Osada *et al.*, 1995), no TAL-1 target gene has been identified so far. Furthermore, the sequence initially defined to bind TAL-1 *in vitro* (Hsu *et al.*, 1994a) and the newly identified *in vitro* target (Wadman *et al.*, 1997) have not been found in regulatory regions of genes involved in hematopoiesis regulation. In order to identify *in vivo* TAL-1 binding sites as well as target genes regulated by TAL-1 protein, we set up a strategy based on immunoselection of DNA-protein complexes in chromatin, and we describe here one of the isolated target sequence that binds TAL-1 and is located within an intron of a new gene whose expression is consistent with TAL-1 regulation.

Results

Isolation of *in vivo* TAL-1 target genomic fragments by chromatin immunopurification

To identify TAL-1 target sequences, we developed an *in vivo* DNA-protein immunopurification strategy that enriches for short fragments bound to TAL-1 in chromatin (Figure 1). We chose the murine erythro-leukemia (MEL) cell line for this immunopurification method for two reasons. First, MEL cells have been widely used to study the molecular and cellular mechanisms that regulate terminal erythroid differentiation (Visvader *et al.*, 1991; Aplan *et al.*, 1992b; Green *et al.*, 1992; Murrell *et al.*,

Fig. 1. The *in vivo* immunoselection strategy used to isolate TAL-1 target sequences. Nuclei were isolated, treated with paraformaldehyde to stabilize *in situ* DNA-protein interactions, and then partially digested with micrococcal nuclease. The soluble chromatin fraction was loaded on a discontinuous caesium chloride gradient to select DNA-protein complexes and subjected to immunoprecipitation using monoclonal anti-TAL-1 antibodies to enrich for TAL-1-bound chromatin. After dissociation of the nucleoprotein complexes, proteins were analyzed by Western blot, and the DNA fragments cloned and sequenced to select potential TAL-1 *in vivo* binding sites for further investigations.

Table I. Sequences containing an E-box closely associated to a GATA motif immunoselected in MEL cells

CAGATG box	CAGGTG box
GATA-(4 bp)-CAGATG	CAGGTG-(9 bp)-GATA
CAGATG-(9 bp)-GATA	CAGGTG-(12 bp)-GATA
	CAGGTG-(13 bp)-GATA

1995) and secondly, TAL-1 has major functions in this cell line (Aplan *et al.*, 1992b). We obtained 450 ng of immunoprecipitated DNA, starting from 3×10^8 cells, i.e. 0.00025% of the initial DNA, and the average insert size was 500 bp. One fifth of the immunoselected material was cloned and 1000 recombinant clones from the library were ordered, among which 60 were sequenced. As expected from the absence of known TAL-1 target genes, none of the cloned fragments sequenced was recorded in sequence databases but interestingly, a high proportion (5/60, i.e. 8.3%) of potential TAL-1 target sequences contained a GATA motif close to an E box of the CAGATG or the CAGGTG type (Table I). As TAL-1 and GATA-1 are considered to be functionally related in erythroid differentiation (for review see Shivdasani and Orkin, 1996), we reasoned that these fragments would be of the utmost interest as potential TAL-1 targets in MEL cells. We therefore selected one 474 bp fragment, whose sequence is

5' CCCATTTTTCAGATGAACAGGTGGAGGCATTACATTGATAGGAAATGTGGCTAAATAGAA
GTAAAGGTCAGAGGGTGTTCGTGAGTGGCGTTTCACAGTCTGGTACAGGCAGGTTCAGC
CCTAACTGTATGACCCCTTGAGCTTCCCTTACTCACTACGCTCTGAACTGTGAGAGAACTGT
TACTTTTGTATACAAAGATCAGTTAAGACAACAGTGCCTGTTTATGGCTCCTTTCCAAAGGA
GTAGTGGGTAGGCACAAGCTGAGTGGGCATAGAGGACAAACGCCTTGTGTAAAGTAGCCAAA
GAGGAAAAAGAGCGGCTTAAAGCAGCTCTATTACCATCACCCATGGGAACCTTTCTAC
TTCCCGAGACGCTCTCGATCATGGGAAGATACATTAAAGTCAGTACTTCATAATTCTATATT
ATAAACGCTTTCCAAATTTCTCCATTTTG 3'

Fig. 2. Sequence of the genomic fragment immunoprecipitated. The sequence of the immunoprecipitated 474 bp DNA fragment is shown. The E-box1 (CAGATG), E-box2 (CAGGTG) and the GATA binding sites are shown in bold and underlined.

shown in Figure 2, for further studies for the following reasons. (i) It contained two E-boxes in tandem spaced by two nucleotides, one of which is of the CAGATG type and the other of the CAGGTG type. Furthermore, the second E box (CAGGTG) lies 13 bp 5' from a consensus GATA binding site. (ii) This fragment was located in a genomic region highly sensitive to micrococcal nuclease in MEL cells, indicating an open chromatin configuration around the immunoprecipitated region. The high micrococcal nuclease sensitivity of the immunoprecipitated fragment was assessed by comparison with the CD2 locus which is not transcriptionally activated in MEL cells, and needed 5- to 10-fold more nuclease to be digested (data not shown). Therefore, this fragment was a good candidate as a TAL-1 binding site, and we attempted first to determine if the region containing the two potential TAL-1 binding sites and the GATA motif could be a regulatory region.

The immunoprecipitated fragment binds TAL-1 *in vitro*

We studied the ability of the two E-boxes to bind TAL-1 using MEL nuclear extracts. A complex of similar mobility as the TAL-1/E2A-containing complex was observed when an oligonucleotide containing the two E-boxes was used in electrophoretic shift assays (EMSAs) (Figure 3, lanes 1 and 3) and supershift experiments using an anti-TAL-1 monoclonal antibody showed that this complex actually contains TAL-1 (Figure 3, lanes 2 and 4). To identify which E box binds TAL-1, we first used oligonucleotides that contained only the CAGATG E-box (Figure 3, lane 5) or the CAGATG E-box followed by a mutated CAGGTG E-box (Figure 3, lane 6). No TAL-1-containing complex could be obtained with these probes, suggesting that TAL-1 binds to the CAGGTG E-box. However, a 20 bp oligonucleotide (the length of the TAL-1/E2A oligonucleotide) that contained the CAGGTG E-box could not efficiently bind TAL-1 (Figure 3, lane 7). As the E-box1-E-box2 oligonucleotide contained two more nucleotides at its 3'-end, we used a 22 bp oligonucleotide that contained only the CAGGTG E-box and the same 3'-end nucleotides, and showed that this oligonucleotide binds TAL-1 (Figure 3, lanes 8 and 9). Finally, we showed that an oligonucleotide containing the CAGGTG E-box and a mutated CAGATG E-box also binds TAL-1 (Figure 3, lanes 10 and 11), demonstrating that, *in vitro*, TAL-1 can bind the CAGGTG E-box located in the immunoprecipitated fragment.

The GATA-1 site increases TAL-1 binding affinity

We then explored the role of the GATA motif located 13 bp downstream from the CAGGTG E-box. Gel-shift

Fig. 3. Gel retardation assay using MEL nuclear extracts and wild-type or mutated E-box1-E-box2 oligonucleotides. EMSAs with MEL nuclear extracts were performed using oligonucleotide probes containing the previously described TAL-1/E2A high-affinity binding site (TAL-1/E2A oligonucleotide, lanes 1 and 2), the E-box1-E-box2 sequence (lanes 3 and 4), the E-box1 sequence alone (lane 5), the E-box1 linked to a mutated E-box2 (mutant a, lane 6), the E-box2 oligonucleotide α (lane 7), the E-box2 oligonucleotide β (lanes 8 and 9) or the E-box2 linked to a mutated E-box1 (mutant b, lanes 10 and 11). Supershift experiments performed using anti-TAL-1 monoclonal antibody are shown on lanes 2, 4, 9 and 11. The band corresponding to the TAL-1/E2A heterodimer is indicated by an arrow.

experiments showed that the probe containing the two E-boxes and the GATA site generated numerous complexes (Figure 4A, lane 4). One complex contained TAL-1, and TAL-1 affinity for this composite sequence was higher than for the E-box1-E-box2 probe, and similar to the one observed for the TAL-1 consensus probe (Figure 4A, lanes 2, 3 and 4). We determined whether the GATA binding site can modulate the TAL-1 binding affinity to the CAGGTG E-box. Competition experiments were performed using the TAL-1/E2A oligonucleotide as a probe and increasing amounts of wild-type or mutant E-box1-E-box2-GATA oligonucleotides as competitors. A typical result of such an experiment is shown in Figure 4B, and indicated that the E-box1-E-box2-GATA oligonucleotide (Figure 4B, lane 5) was more efficient in competition than the E-box1-E-box2-mutantGATA oligonucleotide (Figure 4B, lane 3) and less efficient than the mutant E-box1-E-box2-GATA oligonucleotide (Figure 4B, lane 8). Quantification of the TAL-1-containing complex indicated a ratio of 2 between the E-box1-E-box2-GATA and the E-box1-E-box2-mutantGATA oligonucleotides, and a ratio of 1.5 between the mutant E-box1-E-box2-GATA and the E-box1-E-box2-GATA oligonucleotides. Altogether these results showed that TAL-1 binds to the CAGGTG E-box and that the TAL-1 binding affinity can be modulated positively (by the GATA binding site) or negatively (by the CAGATG E-box).

TAL-1 binds to the *in vivo* target as a part of a multicomponent complex present only in normal erythroid cells

As our observations of MEL cells might not reflect the TAL-1 binding during hematopoiesis, we investigated the

Fig. 4. The GATA binding site potentiates the DNA binding activity of TAL-1. **(A)** EMSAs with MEL nuclear extracts were performed using oligonucleotide probes containing the GATA sequence derived from the immunoprecipitated fragment (GATA oligonucleotide, lane 1), the two E-boxes without a downstream GATA site (E-box1–E-box2 oligonucleotide, lane 2), the TAL-1/E2A high-affinity binding site (TAL-1/E2A oligonucleotide, lane 3), and the two E-boxes followed by the GATA site (E-box1–E-box2–GATA oligonucleotide, lane 4). Bands corresponding to GATA-1- and TAL-1-containing complexes are indicated by arrows. **(B)** EMSAs with MEL nuclear extracts were performed using the TAL-1/E2A oligonucleotide as radioactive probe. A supershift experiment, performed using anti-TAL-1 antibody is shown on lane 2 and competition experiments using E-box1–E-box2–mutantGATA oligonucleotide, E-box2 oligonucleotide α , E-box1–E-box2–GATA, E box1, E-box1–mutantE-box2–GATA and mutantE-box1–E-box2–GATA oligonucleotides are shown respectively on lanes 3–8. The band corresponding to the TAL-1/E2A heterodimer is indicated by an arrow.

binding of nuclear proteins obtained from cultures of purified human hematopoietic progenitor cells (HPCs) in liquid suspension (Condorelli *et al.*, 1995). In this system HPCs are induced to undergo erythroid or granulopoietic determination in the first culture week, and terminal differentiation in the second week. In all experiments, nuclear extracts prepared from cultures at day 9 were used. The E-box1–E-box2–GATA oligonucleotide generated multiple complexes that were detected with nuclear extracts obtained from erythroid cells, and not with nuclear extracts obtained from granulocytic cells, indicating specific erythroid complexes (Figure 5, compare lanes 1 and 2, and lanes 3 and 4). One of the complexes (indicated by an arrow in Figure 5) has a mobility similar to the complex previously described (Wadman *et al.*, 1997) using a CAGGTG type E-box followed by a GATA-site oligonucleotide, and shown to contain LMO2, GATA-1, TAL-1, E2A and Ldb1 (Figure 5, lanes 1 and 3). Interestingly, no TAL-1/E2A complex was observed, indicating that TAL-1 is essentially tied up in the multicomponent complex present in erythroid cells and not in granulocytic cells.

The *in vivo* immunoprecipitated TAL-1 target sequence behaves as a *cis* regulatory acting sequence with positive transcriptional activity

To determine if the immunoprecipitated TAL-1 target sequence is a transcriptional regulatory region we used transient transfection experiments. Transcriptional activity of the E-box1–E-box2–GATA sequence cloned 5' to the minimal promoter from the glycophorin B gene (Rahuel

Fig. 5. A TAL-1-containing multicomponent complex binds the E-box1–E-box2–GATA composite motif in erythroid cells and is unilineage-specific. EMSAs using nuclear extracts from normal erythroid (lanes 1 and 3) or granulopoietic (lanes 2 and 4) cells, were performed with the oligonucleotide probe containing the *in vitro* consensus TAL-1 binding site followed by a GATA site 9 bp downstream (lanes 1 and 2), or the E-box1–E-box2–GATA oligonucleotide derived from the immunoselected sequence (lanes 3 and 4). The band corresponding to the TAL-1/E2A/LMO2/GATA-1 high-mobility complex is indicated by an arrow and the GATA-1 containing complex is also shown.

et al., 1992) was studied using the Dual Luciferase™ Reporter Assay System (Figure 6). Contrary to other studies that have used tandems of identical TAL-1/E2A binding sites of the CAGATG type (Hsu *et al.*, 1994c; Nielsen *et al.*, 1996) or of the CAGGTG type followed by a GATA site (Wadman *et al.*, 1997), we chose to work with only one copy of each oligonucleotide, which is closer to the *in vivo* context. Transcriptional activities of the wild-type or mutated E-box1–E-box2–GATA–glycophorin B (GpB) luciferase reporter constructs were compared in erythrocytic (MEL), erythro/megakaryocytic (HEL) and non-hematopoietic (HeLa) cells. No transcriptional activity was detected in HeLa cells that do not express TAL-1, while similar transcriptional activities were obtained in MEL and HEL cells that express TAL-1 (Figure 6) showing a correlation between the transcriptional activity of this sequence and TAL-1 expression. The wild-type E-box1–E-box2–GATA–GpB construct produced a 2-fold increase of luciferase activity as compared with the GpB minimal construct, and mutation of E-box1 produced a 4-fold increase in luciferase activity indicating that the CAGGTG–GATA motif acts as a positive *cis*-acting sequence whose function is negatively modulated by E-box1 (Figure 6). Double mutations that abolished both the E-box1 and E-box2, or the E-box1 and GATA binding sites resulted in background transcriptional activity demonstrating a co-operativity between TAL-1 and GATA-1 (Figure 6). Similar results were obtained when oligonucleotides were inserted in opposite orientation indicating an orientation-independent transcriptional activity (data not shown). These data showed that the region containing the two E-boxes and the GATA site from the immunoprecipitated fragment behaves as a *cis* regulatory acting sequence with a positive transcriptional activity that can be tuned by the repressive activity mediated by E-box1 (CAGATG).

The immunoprecipitated sequence is located within an intron of a new gene

Using the immunoprecipitated sequence as a probe, we screened a murine genomic library and got several pages

Fig. 6. The oligomeric TAL-1 containing complex is functional and generates positive transcriptional activity. MEL (open bar) or HeLa (hatched bar) cells were transiently transfected with the Firefly luciferase reporter plasmid constructs indicated together with a Renilla luciferase control reporter (pRL-TK). After 24 h, the cells were harvested, extracts made, and aliquots used for quantification of the luminescent signals from each of the two luciferase reporter enzymes. The values represent the average of three independent experiments using different plasmid preparations for each construct. The values are shown in arbitrary units and (H) indicates the standard error to the mean.

that encompassed the probe. One of them, containing a 18 kb insert, was cut with various restriction enzymes and the DNA fragments that overlapped or lay around the immunoprecipitated sequence were used in an exon-trapping assay. We found that a 5 kb *NcoI*-*NcoI* fragment (Figure 7A) that started within the immunoprecipitated sequence contained three exons. Neither the nucleotide sequence nor the polypeptide encoded by these exons (Figure 7A) were found in databases, indicating that these exons belong to a new transcription unit. However, the encoded polypeptide displayed a significant homology to the N-terminus of the recently cloned otogelin (Cohen-Salmon *et al.*, 1997) i.e. it contained a cysteine-rich domain that precedes the beginning of a von Willebrand Factor (vWF) D-type domain (Figure 7A). Fine mapping of the three exons defined their genomic organization and located the immunoprecipitated sequence 56 bp 5' from the first exon identified (Figure 7B). Finally, the sequence of the DNA that flanked the three exons showed intron-exon boundaries that are consistent with the GT-AG rule applied to the splicing point (Figure 7B). Therefore, we conclude that the immunoprecipitated sequence is located in an intron of a new gene.

The expression of the identified gene is consistent with TAL-1 regulation

The transcriptional status of TAL-1 and the target gene we have identified was evaluated by reverse transcription-coupled PCR (RT-PCR). We found a co-expression of these two genes in MEL cells, fetal liver and in a murine stromal cell line, M2-10 B4, that contains endothelial cells (Lemoine *et al.*, 1988) (Figure 8A). On the other hand, the putative TAL-1 target gene was not expressed in NIH 3T3 fibroblastic cell line or in adult liver where indeed no TAL-1 mRNA could be detected (Figure 8A). Finally, we used an *in vitro* differentiation assay of ES cells towards the erythro/megakaryocytic pathways to study the

expression of TAL-1, LMO2, GATA-1 and of the TAL-1-associated transcription unit. Gene expression in these cultures was also assayed by RT-PCR analysis at days 0, 3, 5, 6 and 9. As shown in Figure 8B, the TAL-1 target gene started to be detected at day 6, i.e. 1 day after the appearance of TAL-1, GATA-1 and LMO2 mRNAs. Taken together, these data show that the transcription unit we have identified is expressed in a manner consistent with TAL-1 regulation.

Discussion

Immunoselection of TAL-1 genomic binding sites

To identify TAL-1 targets we set up an immunoprecipitation assay of TAL-1 bound DNA in chromatin which may reflect the *in vivo* TAL-1 DNA-binding activity. As we and others have already discussed (Gould *et al.*, 1990; Deveau *et al.*, 1997; Yamamoto *et al.*, 1997), this method has several advantages as it does not select against low-affinity binding sites, and may take into account the complexity provided by adjacent co-operating binding sites. Immunoprecipitation of DNA-protein complexes has successfully identified *in vivo* target genes for several transcription factors in *Drosophila* (Gould and White, 1992; Graba *et al.*, 1992) as well as in mammals (Bigler and Eisenman, 1995; Grandori *et al.*, 1996; Deveau *et al.*, 1997), and it seems that the use of these *in vivo* strategies is of highest interest when addressing the question of relevant targets for transcription factors. Target sequences identified by immunoprecipitation were often different from those identified *in vitro*, with changes in core adjacent bases of the motif or even insertions in the core site. This was described for targets of UBX, the product of the *Drosophila* homeotic gene *Ultrabithorax* (Gould *et al.*, 1990), thyroid hormone receptor (Bigler and Eisenman, 1994, 1995), Pax-2 (Phelps and Dressler, 1996), MYC-MAX (Grandori *et al.*, 1996) and NF-E2 (Deveau *et al.*,

Fig. 7. The immunoselected sequence is located in an intron of a new gene. **(A)** An 18 kb murine genomic DNA that contained the immunoselected sequence (filled bar) was cut with various restriction enzymes and the *NcoI*–*NcoI* DNA fragment was cloned into the pSPL3 vector and used in an exon-trapping assay. The sequence of the three exons located in the 5 kb *NcoI*–*NcoI* DNA fragment is shown below the partial restriction map of the phage used. Down arrow indicates the end of each exon, right-angled arrow indicates the beginning of a von Willebrand D-type domain. The amino acids conserved between the polypeptide encoded by these three exons and the N-terminus of otogelin are shown in bold. The oligonucleotides used for RT-PCR analysis of the TAL-1 target gene are indicated (right-facing arrow and left-facing arrow). **(B)** A fine mapping of the three exons found in the 5 kb fragment is shown together with the location of the immunoprecipitated fragment and the sequences of the intron–exon boundaries.

1997) targets. Our work indicated that an E-box–GATA association might be important for TAL-1 function *in vivo* as ~10% of the immunoprecipitated fragments showed a GATA binding site associated to an E-box. This bipartite DNA motif, previously defined *in vitro* (Wadman *et al.*, 1997), might thus represent a hallmark for genes regulated by both TAL-1 and GATA proteins. Previous studies have shown that the association of GATA and Sp1, or GATA and Ets binding sites are respectively crucial for erythrocytic- (Frampton *et al.*, 1990) or megakaryocytic- (Lemarchandel *et al.*, 1993) specific expression, and the close association of TAL-1 and GATA binding sites might now define a particular set of hematopoietic specific genes.

Identification of a functional *in vivo* TAL-1 binding site

In order to define whether one of these immunoselected fragments represents a transcriptional regulatory sequence, we focused our study on a sequence that contains two E-boxes followed 13 bp downstream by a GATA motif. The TAL-1/E2A heterodimer could bind the CAGGTG E-box, which is not the preferred *in vitro* TAL-1/E2A

consensus binding site initially defined (Hsu *et al.*, 1994a), and its binding affinity could be modulated positively or negatively by adjacent DNA binding sites. Thus, CAGATG and CAGGTG are TAL-1 binding sites depending on adjacent bases or on closely associated DNA-binding site(s), and TAL-1 binding affinity may be tuned by other transcriptional factors to achieve fine regulation of the corresponding target gene(s). Finally, and in agreement with a previous study (Wadman *et al.*, 1997), we found that an oligomeric complex bound the CAGGTG motif followed 13 bp downstream by a GATA binding site.

Transfection experiments indicated that this complex behaved as a transcriptional transactivating complex in erythroid but not in HeLa cells, and that GATA-1 and TAL-1 are required simultaneously for efficient transcriptional activity. We also showed that the upstream E-box acted as a negative *cis*-acting sequence and thus could tune the TAL-1-containing complex transcriptional activity. One potential candidate for binding to this E-box is the newly identified and broadly expressed MIST-1 bHLH protein (Lemercier *et al.*, 1997) which is able to bind the CAGATG sequence, is present in hematopoietic cells

Fig. 8. The TAL-1 target gene mRNA levels correlate with TAL-1 expression pattern. (A) RT-PCR analysis of total RNA extracted from a stromal cell line that contains endothelial cells (M2-10B4), a fibroblastic cell line (NIH 3T3), the MEL cell line, fetal liver and adult liver were performed. The amount of cDNA used for PCR was normalized by the HPRT PCR and the identity of the TAL-1 and TAL-1 target gene fragment was assessed by hybridization with a specific oligonucleotide (data not shown). The sizes of the PCR fragments are indicated. (B) ES cells were differentiated as described in Materials and methods, and TAL-1, GATA-1, LMO2, and TAL-1 target gene PCR amplifications were performed with cDNA synthesized from total RNA extracted from embryonic bodies at days 0, 3, 5, 6 and 9 post induction using gene-specific primers. The amount of cDNA used for PCR was normalized by the HPRT PCR. The sizes of the PCR fragments are indicated.

(C.Lemercier, personal communication) and can act as a repressor (Lemercier *et al.*, 1998). We tested MIST-1 binding to E-box1 and preliminary data indicated that both E-box1 and E-box1-E-box2 oligonucleotides were able to bind MIST-1 (data not shown), suggesting that MIST-1 could be involved in the repressive activity mediated by E-box1. A similar regulation has been described in the rat prothymosin- α intron enhancer, where a second E-box element, adjacent to the E-box recognized by MYC, binds a putative repressor protein and is a critical determinant of specificity (Desbarats *et al.*, 1996).

Characterization of a TAL-1 target gene

The fragment we have immunopurified in MEL cells contains a composite E-box1-E-box2-GATA motif which can bind a TAL-1-containing complex present in normal erythroid cells and can activate the transcription of an indicator gene erythroid cell line. Therefore, this motif confers to the *in vivo* immunoselected fragment a potential regulatory transcriptional role mediated by TAL-1 at least in the erythroid lineage. Using an exon-trapping assay, we have isolated three consecutive exons, one of which was located 56 bp downstream of the immunoprecipitated region. The close association of the identified exons with

the TAL-1 binding region, together with the co-expression of TAL-1 and the identified gene in the tissues or cell lines tested, showed that this immunoselected region belongs to an intron of a novel gene potentially regulated by TAL-1. The 107 amino acid polypeptide encoded by the three exons showed a striking similarity (51% homology on a stretch of 74 amino acids) with the N-terminus of otogelin, a newly cloned secreted glycoprotein (2910 amino acid length) of the mucin-related family, specifically expressed in the inner ear and of yet unknown function (Cohen-Salmon *et al.*, 1997). The homology concerns amino acid stretches specific to otogelin and the beginning of a vWF D-type domain which is involved in multimerization of the von Willebrand factor (Voorberg *et al.*, 1990). Cloning of the full-length cDNA and characterization of the corresponding protein is now needed to establish the type of protein encoded by this gene. The precise role of this TAL-1 target gene as a mediator of TAL-1 biological function in erythroid lineage also remains to be established. Our preliminary data indicate that this gene is downregulated during terminal differentiation of MEL cells, although TAL-1 mRNA increased. This apparent discrepancy might be explained by the function(s) of LMO2 and Ldb1 which are part of the multimeric complex that was shown to bind to the bipartite E-box-GATA motif (Wadman *et al.*, 1997) and whose expression also decreased during MEL cells differentiation (Visvader *et al.*, 1997).

Finally, TAL-1 target gene expression in a cell line expressing endothelial markers is in accordance with the major role of TAL-1 in embryonic angiogenesis (Visvader *et al.*, 1998) and therefore suggests a possible role of this target gene in vascular development. Whether GATA-2 may substitute for GATA-1 activity in regulating this TAL-1 target gene in endothelial cells remains to be established.

Materials and methods

Cell cultures

The MEL cell line (strain 745) and the HeLa cell line were cultured in Dulbecco's modified Eagle's medium (DMEM) supplemented with 100 U/ml penicillin, 100 μ g/ml streptomycin and 10% fetal calf serum (FCS). The human erythro-megakaryocytic cell line (HEL) was cultured in RPMI 1640 supplemented with antibiotics and 10% FCS. In all experiments, exponentially growing cells were used.

For the *in vitro* differentiation of ES cells (clone CJ7) towards the erythrocytic and megakaryocytic lineages, ES cells were grown in 0.9% α -methylcellulose containing monothioglycerol and the following growth factors: interleukin 3, GM-CSF, interleukin 6, interleukin 11, murine stem-cell factor, thrombopoietin and erythropoietin. The cells were cultured in an humidified incubator at 37°C with 5% CO₂ for 9 days.

Hematopoietic progenitor cell (HPC) purification and liquid differentiation culture

HPC purification and the induction of their differentiation using growth factors in liquid cultures were performed exactly as previously described (Condorelli *et al.*, 1995).

In vivo immunoselection of genomic targets for TAL-1 in MEL nuclei

To isolate functionally relevant *in vivo* TAL-1 binding sites we used an immunopurification method of DNA-protein complexes in chromatin. A schematic diagram of the strategy is depicted in Figure 1. Exponentially growing MEL cells (3×10^8) were lysed in buffer A (10 mM Tris-HCl pH 7.9, 10 mM NaCl, 1.5 mM MgCl₂, 1 mM DTT) supplemented with 0.1% Triton X-100 and several protease inhibitors (2 μ g/ml aprotinin, leupeptin and pepstatin, 0.1 mM PMSF), and washed twice in buffer A.

Table II. Oligonucleotides used in the DNA binding and transcriptional studies

1. E-box1–E-box2–GATA oligonucleotide ^a	5'-ATTTTACAGATGAACAGGTGGAGGCATTACATTGATAGGAAATGT-3'
2. Mutant E-box1–E-box2–GATA oligonucleotides	
a. Mutant E-box1	5'-ATTTTACAGACGAACAGGTGGAGGCATTACATTGATAGGAAATGT-3'
b. Mutant E-box2	5'-ATTTTACAGATGAAAAGGTGGAGGCATTACATTGATAGGAAATGT-3'
c. Mutant GATA	5'-ATTTTACAGATGAACAGGTGGAGGCATTACATTCTTAGGAAATGT-3'
d. Double mutant E-box1, E-box2	5'-ATTTTACAGACGAACAGGTGGAGGCATTACATTGATAGGAAATGT-3'
e. Double mutant E-box1, GATA	5'-ATTTTACAGATGAACAGGT-3'
3. E-box1 oligonucleotide	5'-ATTTTACAGATGAACAGGT-3'
4. E-box2 oligonucleotide α	5'-AGATGAACAGGTGGAGGCAT-3'
5. E-box2 oligonucleotide β	5'-AGATGAACAGGTGGAGGCATTA-3'
6. E-box1–E-box2 oligonucleotide	5'-ATTTTACAGATGAACAGGTGGAGGCATTA-3'
7. Mutant E-box1–E-box2 oligonucleotides	
a. Mutant E-box2	5'-ATTTTACAGATGAAAAGGTGGAGGCATTA-3'
b. Mutant E-box1	5'-ATTTTACAGACGAACAGGTGGAGGCATTA-3'
8. GATA oligonucleotide	5'-ATTACATTGATAGGAAATGT-3'
9. TAL-1/E2A oligonucleotide	5'-ACCTGAACAGATGGTCCGGCT-3'
10. <i>In vitro</i> consensus TAL-1 binding sequence with a GATA site 9 bp downstream	5'-TCGGCGCCAGGTGCTGCGTCCCAGATAGGGGCCG-3'

^aE-boxes and GATA binding sites are underlined. Mutations are shown in bold.

To cross-link *in vivo* DNA–protein complexes, the intact nuclei were resuspended in 10 mM Na-phosphate buffer pH 7.4 containing 2% paraformaldehyde, incubated for 30 min at 4°C with gentle rocking, and then washed three times in buffer A. Nuclei were resuspended in buffer A supplemented with 2 mM CaCl₂ and incubated with 20 U/ml Micrococcal nuclease (Pharmacia) at 37°C for 5 min. The reaction was stopped with 5 mM EDTA. 1% NP-40 and 2% sarcosine (*N*-Methylaminoacetic acid) were added and soluble chromatin was harvested by centrifugation at 15 000 r.p.m. for 15 min at 4°C, then loaded onto a CsCl step-gradient (1.75, 1.5, 1.35 g/cm³). The gradient was centrifuged for 40 h at 20°C at 27 000 r.p.m. in a SW41 rotor (Beckman). Fractions were collected from the bottom of the tube and OD₂₆₀/OD₂₈₀ ratio was measured. Fractions containing nucleoprotein complexes were selected and dialyzed against dialysis buffer (10 mM Tris, 1 mM EDTA, 0.5 mM PMSF, pH 7.5). The pooled selected chromatin fractions were subjected to immunoprecipitation, using the anti-TAL-1 antibody (2TL136 Mab, 10 µg/ml, a gift from D.Mathieu) (Pulford *et al.*, 1995) in the dialysis buffer supplemented with 1% NP-40 for 4 h at 4°C. After addition of protein G Plus–Agarose (Santa Cruz Biotechnology), incubation was continued for an additional 4 h at 4°C. The protein G–agarose beads were washed with 10 mM Na-phosphate buffer pH 7.0 containing 1 M NaCl and 0.1% NP-40 and washed again with the above buffer without NaCl. Washing with high- and low-salt solutions was repeated for two more cycles. The bound nucleoprotein complexes were eluted with 3 M NaSCN containing 10 mM Na-phosphate buffer pH 7.0 for 1 h at 4°C with gentle rocking. Immunoprecipitate was shown to be enriched in TAL-1 protein by Western blotting. Eluate was dialyzed overnight against 0.1 M Na-phosphate buffer pH 7.0, 1 mM EDTA. The dialyzed sample was incubated with 0.1 mg/ml proteinase K (Boehringer Mannheim), 50 mM Tris–HCl buffer pH 8, 100 mM EDTA, 0.3% sodium dodecyl sulfate (SDS) at 55°C for 3 h, extracted twice with phenol–chloroform–isoamylalcohol and ethanol-precipitated in presence of dextran T40 as a carrier. The purified DNA was blunt-ended by T4 DNA polymerase and one fifth of the material was cloned into *EcoRV* digested pBluescript SK-vector (Stratagene) using library-efficiency *Escherichia coli* XL1-Blue competent cells (Stratagene).

DNA sequencing and sequence analysis

DNA sequencing was carried out using the PRISM Ready Reaction Dye Deoxy Terminator Cycle Sequencing Kit (Perkin-Elmer) on the ABI Model 377 DNA Sequencer (Applied Biosystems).

Construction of reporter plasmids and transient transfection experiments

The glycophorin B minimal promoter (GpB-37) (Rahuel *et al.*, 1992) was ligated between the *MluI* and *BglII* sites of the firefly luciferase reporter vector pGL3-Basic (Promega) to generate the plasmid pGL3-GpB. One copy of pairs of complementary oligonucleotides (45 bp) corresponding to the wild-type or mutated E-box1–E-box2–GATA sequence was then inserted into the *Acc65I* site of pGL3-GpB to create the luciferase reporter E-box1–E-box2–GATA–GpB or its mutated forms.

Transcriptional activity was tested by transient transfection of MEL, HEL or HeLa cells, using the Dual-Luciferase™ reporter Assay System (Promega). For each transfection, 10⁷ log-phase cells were washed in phosphate-buffered saline (PBS), pelleted and resuspended in 0.2 ml of ice cold PBS with 50 mM HEPES pH 7.4, containing 10 µg of supercoiled plasmid construct and 100 ng of pRL-TK control vector for normalization. Electroporations (250 V, 960 µF) were performed using the Bio-Rad Gene Pulser (Bio-Rad). The cells were resuspended in fresh medium, incubated for 24 h at 37°C, then harvested and lysed in 200 µl of Passive Lysis Buffer (Promega). Quantification of the luminescent signals from each of the two luciferase reporter enzymes was performed on 2–10 µl aliquots of supernatant using an automated luminometer equipped with two reagent injectors (LUMAT LB 9507, Berthold).

Cloning of the genomic DNA fragment that encompassed the immunoselected sequence and detection of the transcription unit linked to this sequence

The 474 bp immunoselected DNA fragment was ³²P-labelled and used as a probe to isolate phages from a λGEM12 mouse genomic library (a kind gift from Généthron, Evry, France) using standard methods (Sambrook *et al.*, 1989). A 18 kb genomic DNA fragment that contained the immunoprecipitated sequence was selected and cut with various restriction enzymes to isolate subfragments (4–6 kb) that were used in an exon-trapping assay (Exon Trapping System, Gibco-BRL). Data base searches for nucleotides and amino acids sequence alignment were achieved using the N.C.B.I. Blast program (Altschul *et al.*, 1997).

RT-PCR analysis

Total RNAs were isolated from cell lines or tissues using Trizol (Gibco-BRL), and random priming cDNAs were synthesized from 1 µg of total RNA. A fraction (2%) of the cDNA was used for the HPRT PCR and the amount of cDNA used for each of the specific amplification was normalized by the amount of the HPRT PCR product. Products were visualized after separation on a 1.5% agarose gel by ethidium bromide staining and hybridized with a ³²P-labelled internal oligonucleotide to assess the specificity of the product obtained.

Sequence of oligonucleotides

Oligonucleotides used as probes and competitors in EMSAs, and in the construction of reporter plasmids are shown in Table II. TAL-1, GATA-1, LMO2 and HPRT oligonucleotides used for RT-PCR are described elsewhere (Keller *et al.*, 1993). The location of TAL-1 target gene oligonucleotides used for expression studies is indicated in Figure 7A.

Nuclear extracts and DNA binding assays

Nuclear extracts were prepared from MEL cell line and from normal erythroid and granulocytic cells in culture (Schreiber *et al.*, 1989), and DNA binding assays were performed essentially as described (Hsu *et al.*, 1994a; Condorelli *et al.*, 1995; Wadman *et al.*, 1997). Quantification of the TAL-1-containing complex was done using a PhosphorImager (Molecular Dynamics).

Acknowledgements

We would like to thank Dr G.Uzan for the kind gift of ES cells. We are grateful to Dr S.Chretien (INTS, Paris, France) and to the members of the INSERM U.474, especially to Dr D.Beaupain, Dr M.-C.Garel, Dr V.Mignotte and Dr N.Raich for constant support and helpful discussions during this work. We thank Anne-Marie Dulac for typing the manuscript. This work was supported by the Institut National de la Santé et de la Recherche Médicale, by the Ligue Nationale contre le Cancer and by the ARC. L.Maouche-Chretien was supported by the Fondation de France.

References

- Agulnick, A.D., Taira, M., Breen, J.J., Tanaka, T., Dawid, I.B. and Westphal, H. (1996) Functional and physical interaction of Ldb1, a novel LIM domain binding factor, with the LIM homeodomain protein Lhx1/Xlim 1. *Nature*, **384**, 270–272.
- Aird, W.C., Parvin, J.D., Sharp, P.A. and Rosenberg, R.D. (1994) The interaction of GATA-binding proteins and basal transcription factors with GATA box-containing core promoters. A model of tissue-specific gene expression. *J. Biol. Chem.*, **269**, 883–889.
- Altschul, S.F., Madden, T.L., Schäffer, A.A., Zhang, J., Zhang, Z., Miller, W. and Lipman, D.J. (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res.*, **25**, 3389–3402.
- Aplan, P.D., Lombardi, D.P., Reaman, G.H., Sather, H.N., Hammond, G.D. and Kirsch, I.R. (1992a) Involvement of the putative hematopoietic transcription factor SCL in T-cell acute lymphoblastic leukemia. *Blood*, **79**, 1327–1333.
- Aplan, P.D., Nakahara, K., Orkin, S.H. and Kirsch, I.R. (1992b) The SCL gene product: a positive regulator of erythroid differentiation. *EMBO J.*, **11**, 4073–4081.
- Aplan, P.D., Jones, C.A., Chervinsky, D.S., Zhao, X., Ellsworth, M.K., Wu, C., McGuire, E.A. and Gross, K.W. (1997) An scl gene product lacking the transactivation domain induces bony abnormalities and cooperates with LMO1 to generate T-cell malignancies in transgenic mice. *EMBO J.*, **16**, 2408–2419.
- Baer, R. (1993) TAL1, TAL2 and LYL1: a family of basic helix–loop–helix proteins implicated in T cell acute leukaemia. *Semin. Cancer Biol.*, **4**, 341–347.
- Bash, R.O., Hall, S., Timmons, C.F., Crist, W.M., Amylon, M., Smith, R.G. and Baer, R. (1995) Does activation of the TAL1 gene occur in a majority of patients with T-cell acute lymphoblastic leukemia? A pediatric oncology group study. *Blood*, **86**, 666–676.
- Begley, C.G. *et al.* (1989) Chromosomal translocation in a human leukemic stem-cell line disrupts the T-cell antigen receptor δ -chain diversity region and results in a previously unreported fusion transcript. *Proc. Natl Acad. Sci. USA*, **86**, 2031–2035.
- Bernard, O., Guglielmi, P., Jonveaux, P., Cherif, D., Gisselbrecht, S., Mauchauffe, M., Berger, R., Larsen, C.J. and Mathieu-Mahul, D. (1990) Two distinct mechanisms for the SCL gene activation in the t (1;14) translocation of T-cell leukemias. *Genes Chromosom. Cancer*, **1**, 194–208.
- Bigler, J. and Eisenman, R.N. (1994) Isolation of a thyroid hormone-responsive gene by immunoprecipitation of thyroid hormone receptor-DNA complexes. *Mol. Cell. Biol.*, **14**, 7621–7632.
- Bigler, J. and Eisenman, R.N. (1995) Novel location and function of a thyroid hormone response element. *EMBO J.*, **14**, 5710–5723.
- Boehm, T., Foroni, L., Kaneko, Y., Perutz, M.P. and Rabbitts, T.H. (1991) The rhombotin family of cysteine-rich LIM-domain oncogenes: Distinct members are involved in T-cell translocations to human chromosomes 11p15 and 11p13. *Proc. Natl Acad. Sci. USA*, **88**, 4367–4371.
- Briegleb, K., Bartunek, P., Stengl, G., Lim, K.-C., Beug, H., Engel, J.D. and Zenke, M. (1996) Regulation and function of transcription factor GATA-1 during red blood cell differentiation. *Development*, **122**, 3839–3850.
- Chen, Q. *et al.* (1990) The tal gene undergoes chromosome translocation in T cell leukemia and potentially encodes a helix–loop–helix protein. *EMBO J.*, **9**, 415–424.
- Cohen-Salmon, M., El-Amraoui, A., Leibovici, M. and Petit, C. (1997) Otolin: A glycoprotein specific to the acellular membranes of the inner ear. *Proc. Natl Acad. Sci. USA*, **94**, 14450–14455.
- Condorelli, G., Vitelli, L., Valtieri, M., Marta, L., Montesoro, E., Lulli, V., Baer, R. and Peschle, C. (1995) Coordinate expression and developmental role of Id2 protein and TAL1/E2A heterodimer in erythroid progenitor differentiation. *Blood*, **86**, 164–175.
- Condorelli, G.L., Facchiano, F., Valtieri, M., Proietti, E., Vitelli, L., Lulli, V., Huebner, K., Peschle, C. and Croce, C.M. (1996) T-cell-directed TAL-1 expression induces T-cell malignancies in transgenic mice. *Cancer Res.*, **56**, 5113–5119.
- Desbarats, L., Gaubatz, S. and Eilers, M. (1996) Discrimination between different E-box-binding proteins at an endogenous target gene of *c-myc*. *Genes Dev.*, **10**, 447–460.
- Deveaux, S., Cohen-Kaminsky, S., Shivdasani, R.A., Andrews, N.C., Filipe, A., Kuzniak, I., Orkin, S.H., Roméo, P.-H. and Mignotte, V. (1997) p45 NF-E2 regulates expression of thromboxane synthase in megakaryocytes. *EMBO J.*, **16**, 5654–5661.
- Finger, L.R., Kagan, J., Christopher, G., Kurtzberg, J., Herschfield, M.S., Nowell, P.C. and Croce, C.M. (1989) Involvement of the TCL5 gene on human chromosome 1 in T-cell leukemia and melanoma. *Proc. Natl Acad. Sci. USA*, **86**, 5039–5043.
- Frampton, J., Walker, M., Plumb, M. and Harrison, P.R. (1990) Synergy between the NF-E1 erythroid-specific transcription factor and the CACCC factor in the erythroid-specific promoter of the human porphobilinogen deaminase gene. *Mol. Cell. Biol.*, **10**, 3838–3842.
- Gould, A.P. and White, R.A.H. (1992) *Connectin*, a target of homeotic gene control in *Drosophila*. *Development*, **116**, 1167–1174.
- Gould, A.P., Brookman, J.J., Strutt, D.I. and White, R.A.H. (1990) Targets of homeotic gene control in *Drosophila*. *Nature*, **348**, 308–312.
- Graba, Y., Aragnol, D., Laurenti, P., Garzino, V., Charnot, D., Berenger, H. and Pradel, J. (1992) Homeotic control in *Drosophila*: the *scabrous* gene is an *in vivo* target of *Ultrabithorax* proteins. *EMBO J.*, **11**, 3375–3384.
- Grandori, C., Mac, J., Siébelt, F., Ayer, D.E. and Eisenmann, R.N. (1996) Myc–Max heterodimers activate a DEAD box gene and interact with multiple E box-related sites *in vivo*. *EMBO J.*, **15**, 4344–4357.
- Green, T. (1996) Master regulator unmasked. *Nature*, **383**, 575–576.
- Green, A.R., Lints, T., Visvader, J., Harvey, R. and Begley, C.G. (1992) SCL is coexpressed with GATA-1 in hemopoietic cells but is also expressed in developing brain. *Oncogene*, **7**, 653–660.
- Hsu, H.-L., Cheng, J.-T., Chen, Q. and Baer, R. (1991) Enhancer-binding activity of the tal-1 oncoprotein in association with the E47/E12 helix–loop–helix proteins. *Mol. Cell. Biol.*, **11**, 3037–3042.
- Hsu, H.-L., Huang, L.H., Tsou Tsan, J., Funk, W., Wright, W.E., Hu, J.-S., Kingston, R.E. and Baer, R. (1994a) Preferred sequences for DNA recognition by the TAL1 helix–loop–helix proteins. *Mol. Cell. Biol.*, **14**, 1256–1265.
- Hsu, H.-L., Wadman, I. and Baer, R. (1994b) Formation of *in vivo* complexes between the TAL-1 and E2A polypeptides of leukemic T cells. *Proc. Natl Acad. Sci. USA*, **91**, 3181–3185.
- Hsu, H.-L., Wadman, I., Tsou Tsan, J. and Baer, R. (1994c) Positive and negative transcriptional control by the TAL1 helix–loop–helix protein. *Proc. Natl Acad. Sci. USA*, **91**, 5947–5951.
- Jurata, L.W., Kenny, D.A. and Gill, G.N. (1996) Nuclear LIM interactor, a rhombotin and LIM homeodomain interacting protein, is expressed early in neuronal development. *Proc. Natl Acad. Sci. USA*, **93**, 11693–11698.
- Kallianpur, A.R., Jordan, J.E. and Brandt, S.J. (1994) The SCL/TAL-1 gene is expressed in progenitors of both the hematopoietic and vascular systems during embryogenesis. *Blood*, **83**, 1200–1208.
- Keller, G., Kennedy, M., Papayannopoulou, T. and Wiles, M.V. (1993) Hematopoietic commitment during embryonic stem cell differentiation in culture. *Mol. Cell. Biol.*, **13**, 473–486.
- Kelliher, M.A., Seldin, D.C. and Leder, P. (1996) Tal-1 induces T cell acute lymphoblastic leukemia accelerated by casein kinase II α . *EMBO J.*, **15**, 5160–5166.
- Larson, R.C., Lavenir, I., Larson, T.A., Baer, R., Warren, A.J., Wadman, I., Nottage, K. and Rabbitts, T.H. (1996) Protein dimerization between Lmo2 (Rbn2) and Tal1 alters thymocyte development and potentiates T cell tumorigenesis in transgenic mice. *EMBO J.*, **15**, 1021–1027.
- Lemarchandel, V., Ghysdael, J., Mignotte, V., Rahuel, C. and Roméo, P.-H. (1993) GATA and Ets *cis*-acting sequences mediate megakaryocyte-specific expression. *Mol. Cell. Biol.*, **13**, 668–676.
- Lemerrier, C., To, R.Q., Swanson, B.J., Lyons, G.E. and Konieczny, S.F. (1997) Mist1: a novel basic helix–loop–helix transcription factor exhibits a developmentally regulated expression pattern. *Dev. Biol.*, **182**, 101–113.
- Lemerrier, C., To, R.Q., Carrasco, R.A. and Konieczny, S.F. (1998) The basic helix–loop–helix transcription factor Mist1 functions as a transcriptional repressor of MyoD. *EMBO J.*, **17**, 1412–1422.

- Lemoine, F.M., Humphries, R.K., Abraham, S.D.M. and Krystal, G. (1988) Partial characterization of a novel stromal cell-derived pre-B-cell growth factor active on normal and immortalized pre-B cells. *Exp. Hematol.*, **16**, 718–726.
- Murre, C. *et al.* (1989) Interactions between heterologous helix–loop–helix proteins generate complexes that bind specifically to a common DNA sequence. *Cell*, **58**, 537–544.
- Murrell, A.M., Bockamp, E.-O., Göttgens, B., Chan, Y.S., Cross, M.A., Heyworth, C.M. and Green, A.R. (1995) Discordant regulation of SCL/TAL-1 mRNA and protein during erythroid differentiation. *Oncogene*, **11**, 131–139.
- Nielsen, A.L., Nørby, P.L., Pedersen, F.S. and Jørgensen, P. (1996) E-box sequence and context-dependent TAL1/SCL modulator of basic helix–loop–helix protein-mediated transcriptional activation. *J. Biol. Chem.*, **271**, 31463–31469.
- Osada, H., Grutz, G., Axelson, H., Forster, A. and Rabbitts, T.H. (1995) Association of erythroid transcription factors: complexes involving the LIM protein RBTN2 and the zinc-finger protein GATA1. *Proc. Natl Acad. Sci. USA*, **92**, 9585–9589.
- Pevny, L., Simon, M.C., Robertson, E., Klein, W.H., Tsai, S.-F., D'Agati, V., Orkin, S.H. and Constantini, F. (1991) Erythroid differentiation in chimaeric mice blocked by a targeted mutation in the gene for transcription factor GATA-1. *Nature*, **349**, 257–260.
- Phelps, D.E. and Dressler, G.R. (1996) Identification of novel Pax-2 binding sites by chromatin precipitation. *J. Biol. Chem.*, **271**, 7978–7985.
- Porcher, C., Swat, W., Rockwell, K., Fujiwara, Y., Alt, F.W. and Orkin, S.H. (1996) The T cell leukemia oncoprotein SCL/tal-1 is essential for development of all hematopoietic lineages. *Cell*, **86**, 47–57.
- Pulford, K., Lecoite, N., Leroy-Viard, K., Jones, M., Mathieu-Mahul, D. and Mason, D.Y. (1995) Expression of TAL-1 proteins in human tissues. *Blood*, **85**, 675–684.
- Rabbitts, T.H. (1994) Chromosomal translocations in human cancer. *Nature*, **372**, 143–149.
- Rahuel, C., Vinit, M.-A., Lemarchandel, V., Cartron, J.-P. and Roméo, P.-H. (1992) Erythroid-specific activity of the glycophorin B promoter requires GATA-1 mediated displacement of a repressor. *EMBO J.*, **11**, 4095–4102.
- Raich, N., Clegg, C.H., Grofti, J., Roméo, P.-H. and Stamatoyannopoulos, G. (1995) GATA1 and YY1 are developmental repressors of the human epsilon-globin gene. *EMBO J.*, **14**, 801–809.
- Robb, L., Lyons, L., Li, R., Hartley, L., Köntgen, F., Harvey, R.P., Metcalf, D. and Begley, C.G. (1995) Absence of yolk sac hematopoiesis from mice with a targeted disruption of the scl gene. *Proc. Natl Acad. Sci. USA*, **92**, 7075–7079.
- Robb, L., Elwood, N.J., Elefanti, A.G., Köntgen, F., Li, R., Barnet, L.D. and Begley, C.G. (1996) The scl gene product is required for the generation of all hematopoietic lineages in the adult mouse. *EMBO J.*, **15**, 4123–4129.
- Royer-Pokora, B., Loos, U. and Ludwig, W.-D. (1991) TTG-2, a new gene encoding a cysteine-rich protein with the LIM motif, is overexpressed in acute T-cell leukaemia with the t (11;14) (p13;q11). *Oncogene*, **6**, 1887–1893.
- Sambrook, J., Fritsch, E.F. and Maniatis, T. (1989) *Molecular Cloning. A Laboratory Manual*, Second Edition. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
- Schreiber, E., Matthias, P., Müller, M.M. and Schaffner, W. (1989) Rapid detection of octamer binding proteins with 'mini extracts', prepared from a small number of cells. *Nucleic Acids Res.*, **17**, 6419.
- Shivdasani, R.A. and Orkin, S.H. (1996) The transcriptional control of hematopoiesis. *Blood*, **87**, 4025–4039.
- Shivdasani, R.A., Mayer, E.L. and Orkin, S.H. (1995) Absence of blood formation in mice lacking the T-cell leukaemia oncoprotein tal-1/SCL. *Nature*, **373**, 432–434.
- Simon, M.C., Pevny, L., Wiles, M.V., Keller, G., Constantini, F. and Orkin, S.H. (1992) Rescue of erythroid development in gene targeted GATA-1 mouse embryonic stem cells. *Nature Genet.*, **1**, 92–97.
- Valge-Archer, V.E., Osada, H., Warren, A.J., Forster, A., Li, J., Baer, R. and Rabbitts, T.H. (1994) The LIM protein RBTN2 and the basic helix–loop–helix protein TAL1 are present in a complex in erythroid cells. *Proc. Natl Acad. Sci. USA*, **91**, 8617–8621.
- Visvader, J., Begley, C.G. and Adams, J.M. (1991) Differential expression of the LYL, SCL and E2A helix–loop–helix genes within the hemopoietic system. *Oncogene*, **6**, 187–194.
- Visvader, J.E., Mao, X., Fujiwara, Y., Hahn, K. and Orkin, S.H. (1997) The LIM-domain binding protein Ldb1 and its partner LMO2 act as negative regulators of erythroid differentiation. *Proc. Natl Acad. Sci. USA*, **94**, 13707–13712.
- Visvader, J.E., Fujiwara, Y. and Orkin, S.H. (1998) Unsuspected role for the T-cell leukemia protein SCL/tal-1 in vascular development. *Genes Dev.*, **12**, 473–479.
- Voorberg, J., Fontijn, R., van Mourik, J. and Pannekoek, H. (1990) Domains involved in multimer assembly of von Willebrand factor (vWF): multimerization is independent of dimerization. *EMBO J.*, **9**, 797–803.
- Wadman, I., Li, J., Bash, R.O., Forster, A., Osada, H., Rabbitts, T.H. and Baer, R. (1994) Specific *in vivo* association between the bHLH and LIM proteins implicated in human T cell leukemia. *EMBO J.*, **13**, 4831–4839.
- Wadman, I.A., Osada, H., Grütz, G.G., Agulnick, A.D., Westphal, H., Forster, A. and Rabbitts, T.H. (1997) The LIM-only protein Lmo2 is a bridging molecule assembling an erythroid, DNA-binding complex which includes the TAL1, E47, GATA-1 and Ldb1/NLI proteins. *EMBO J.*, **16**, 3145–3157.
- Warren, A.J., Colledge, W.H., Carlton, M.B.L., Evans, M.J., Smith, A.J.H. and Rabbitts, T.H. (1994) The oncogenic cysteine-rich LIM domain protein Rbtn2 is essential for erythroid development. *Cell*, **78**, 45–58.
- Weiss, M.J. and Orkin, S.H. (1995) GATA transcription factors: key regulators of hematopoiesis. *Exp. Hematol.*, **23**, 99–107.
- Yamamoto, M., Takahashi, S., Onodera, K., Muraosa, Y. and Engel, J.D. (1997) Upstream and downstream of erythroid transcription factor GATA-1. *Genes Cells*, **2**, 107–115.

Received December 12, 1997; revised July 2, 1998;
accepted July 3, 1998