

Identification and characterization of the *ARPI* gene, a target for the human acute leukemia *ALL1* gene

HIROFUMI ARAKAWA*, TATSUYA NAKAMURA*, ALEXANDER B. ZHADANOV*, VINCENZO FIDANZA*, TAKAHIRO YANO*, FLORENCIA BULLRICH*, MASAYOSHI SHIMIZU*, JANNA BLECHMAN†, ALEXANDER MAZO*, ELI CANAANI†, AND CARLO M. CROCE*‡

*Kimmel Cancer Institute, Jefferson Medical College, Philadelphia, PA 19107; and †Department of Molecular Cell Biology, Weizmann Institute of Science, Rehovot 76100, Israel

Contributed by Carlo M. Croce, February 10, 1998

ABSTRACT *ALL1*, the human homologue of *Drosophila trithorax*, is directly involved in human acute leukemias associated with abnormalities at 11q23. Using the differential display method, we isolated a gene that is down-regulated in *All1* double-knockout mouse embryonic stem (ES) cells. The gene, designated *ARPI* (also termed *RIEG*, *Ptx2*, or *Otx2*), is a member of a family of homeotic genes containing a short motif shared with several homeobox genes. Using a bacterially synthesized All1 polypeptide encompassing the AT-hook motifs, we identified a 0.5-kb *ARPI* DNA fragment that preferentially bound to the polypeptide. Within this DNA, a region of ≈100 bp was protected by the polypeptide from digestion with *ExoIII* and DNase I. Whole-mount *in situ* hybridization to early mouse embryos of 9.5–10.5 days indicated a complex pattern of *Arp1* expression spatially overlapping with the expression of *All1*. Although the *ARPI* gene is expressed strongly in bone marrow cells, no transcripts were detected in six leukemia cell lines with 11q23 translocations. These results suggest that *ARPI* is up-regulated by the All1 protein, possibly through direct interaction with an upstream DNA sequence of the former. The results are also consistent with the suggestion that *ALL1* chimeric proteins resulting from 11q23 abnormalities act in a dominant negative fashion.

Abnormalities at 11q23 are common in acute lymphocytic, acute myelomonocytic, acute monocytic, and acute myeloid leukemia, drug-induced secondary leukemias, and myelodysplastic syndrome. We and others have isolated a gene at 11q23 that is involved in these chromosome aberrations (1, 2). The gene was designated *ALL1* (also *HRX*, *HTRX*, and *MLL*). It exhibits strong homology to the *Drosophila trithorax* (*trx*) gene (3), particularly in the region encoding multiple zinc fingers (PHD fingers), as well as in the C-terminal (SET) domain. In addition, the All1 protein contains a cluster of “AT-hook” motifs involved in DNA binding (2), a region with homology to DNA-methyltransferase (4), and a transcriptional activation domain (5, 6). In *Drosophila*, *trx* maintains the expression of homeotic genes of the *bithorax* and *Antennapedia* complexes (BX-C and ANT-C), which direct body segmentation during *Drosophila* embryonic development (7, 8). Recent analysis of *All1* knockout mice suggests that *ALL1* plays a role in mammals that is similar to the role of *trx* (9).

Antibodies directed against the *trx* protein detected multiple, specific binding sites on polytene chromosomes of larvae salivary glands (10, 11). The sites included the regulatory region of a known *trx* target gene, *fork head*, and mutations of *trx* caused reduced expression of this target gene (10). These findings suggested that *trx* exerts its effects by binding directly

or indirectly to specific DNA sequences in target genes. Some of the targets for *ALL1* are likely to be the mammalian *Hox* genes. This notion is supported by the demonstration (9) of spatially altered expression of *Hoxa 7* and *Hoxc 9* in *All1* +/- knockout mice.

In acute leukemias with aberrations at 11q23, the *ALL1* gene recombines with many partner genes from a variety of chromosomes (12, 13). In addition, in a significant number of AML patients, *ALL1* undergoes partial tandem duplication (14, 15). Finally, several patients with acute lymphocyte leukemia show an identical deletion spanning *ALL1* exon 8 and flanking intronic sequences (16). The partner genes are a diverse lot. Some are transcription factors (17, 18), and others are involved in signal transduction (19). *AF9* and *ENL* (20), as well as *AF10* and *AF17* (17), are highly homologous. Because of these diverse features, the underlying mechanism for *ALL1* leukemogenicity is still unknown. It is conceivable that *ALL1* triggers leukemia through its effects on target gene(s) involved in hematopoiesis. We have proposed previously (14, 21) that *ALL1* rearrangements result in its loss of function and that the chimeric proteins exert a dominant negative effect on the normal All1 protein encoded by the intact allele present in the leukemic cells.

We recently have generated *All1* knockout embryonic stem (ES) cell lines and investigated their ability to differentiate along hematopoietic lineages after withdrawal of leukemia inhibitory factor (LIF) (22). Whereas wild-type ES cells formed lineage-restricted mature colonies, +/- and -/- knockout ES cells developed, at high frequency, immature and/or “multiphenotypic” colonies. This phenotype is similar to that of acute leukemia cells with 11q23 abnormalities. These results suggested that *ALL1* gene rearrangements in humans and *All1* knockout in mice result in similar abnormalities in hematopoietic cells. We reasoned that the *All1* -/- knockout ES cells might fail to express transcripts of genes induced by *All1*. Applying the differential display methodology (23), we isolated and characterized a gene that is underexpressed in *All1* knockout ES cells. The properties of this gene, *ARPI*, support the notion that it is a target for *ALL1*.

MATERIALS AND METHODS

Cell Lines and Differential Display. ES AB2.1 (wild type), 48^{all1+/-}, 48^{all1-/-}, and 48^{N^{all1}-/-} cells were maintained in gelatinized tissue culture dishes in DMEM supplemented with 15% heat-inactivated fetal bovine serum, 2 mM glutamine, 0.1

Abbreviations: ATH, AT-hooks domain; ES, embryonic stem; MTase, DNA methyltransferase.

Data deposition: The sequence reported in this paper has been deposited in the GenBank database (accession nos. AF048720–AF048724).

‡To whom reprint requests should be addressed at: Kimmel Cancer Institute, Jefferson Medical College, BLSB, Room 1050, 233 South, 10th Street, Philadelphia, PA 19107-5799.

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked “advertisement” in accordance with 18 U.S.C. §1734 solely to indicate this fact.

© 1998 by The National Academy of Sciences 0027-8424/98/954573-6\$2.00/0
PNAS is available online at <http://www.pnas.org>.

mM 2-mercaptoethanol, and 1,000 units/ml of recombinant leukemia inhibitory factor (GIBCO/BRL).

Using Genhunter RNA image kit (Genhunter, Nashville, TN), we performed the differential display comparing total RNAs of AB2.1, 48S^{all1-/-}, and 48N^{all1-/-} cells.

Binding-Site Selection, Exonuclease III Protection, and DNase I Footprint Analysis. Two regions of *ALL1* cDNA, encoding the AT-hooks domain (ATH) spanning nucleotides 331–948 and DNA methyltransferase (MTase) homologous region encompassing nucleotide 3346–3753, were cloned into the pET-23d vector (Novagen), which provides a stretch of six histidines. The BL21 (DE3) bacteria strain (Novagen) was used for expression. The 25-kD ATH and 18-kD MTase peptides were affinity-purified by Ni-NTA resin (Qiagen) and dialyzed against buffer D (20 mM Hepes, pH 7.9/0.1 M KCl/20% glycerol/0.2 mM EDTA/1 mM DTT). Aliquots of 1 μ g each of the ATH or MTase peptides were mixed either with 1 μ g of BSA or 1 μ g of HeLa cell nuclear extract in 10 μ l of buffer D, spotted on 1 cm² of nitrocellulose filter (BA85, Schleicher & Schuell) and air-dried for 10 min. The filter was subsequently blocked for 30 min at 4°C with 100 μ l of binding buffer (25 mM Hepes, pH 7.9/40 mM KCl/5 mM MgCl₂/1 mM DTT) containing 0.5% nonfat milk and 10 μ g of poly (di-dC) and washed with binding buffer containing 0.25% nonfat milk. Ten micrograms of ARP1 genomic cosmid clone 2 was completely digested with *Sau3AI* and ligated to a *Sau3AI* adapter formed by annealing the 24-mer 5'-GATCA-GAAGCTTGAATTTCGAGCAG-3' and 20-mer 5'-CT-GCTCGAATTCAAGCTTCT-3' oligonucleotides. After incubation for 2 hr at 4°C in 100 μ l of binding buffer, containing 500 ng of the segmented DNA preparation, the filter was washed three times with binding buffer including 0.25% milk and once with binding buffer and transferred into an Eppendorf tube. DNA was eluted by incubation with a buffer containing 20 mM Tris, pH 8.0, and 1 M KCl for 10 min on ice with occasional vortexing. The elute was diluted with water and precipitated with isopropanol in the presence of glycogen, and the DNA pellet was resuspended in 20 μ l water. One milliliter aliquot of the DNA was amplified by 25 cycles of PCR, using the 20-mer adapter oligonucleotide, and the PCR product was resolved on a 2% MetaPhor XR agarose (FMC) and cloned into pBluescript SK(-) (Stratagene).

Exonuclease III (ExoIII) protection assay was carried out by the method of Wu (24). DNase I footprint analysis was carried out by using the Sure track DNase footprint kit (Pharmacia).

Whole-Mount *In Situ* Hybridization. Embryos were obtained from natural matings of FVB/N or C57B/6J mice. The time point of vaginal plug observation was designated at 0.5 day postcoitum, and embryonic stages are listed as days postcoitum (E followed by the number of the day). Hybridization experiments were performed as described (25). Digoxigenin-labeled RNAs were prepared by using an RNA labeling kit (Boehringer Mannheim) with *ARP1* cDNA. Anti-digoxigenin antibodies (Fab fragment) conjugated with alkaline phosphatase were purchased from Boehringer Mannheim. Some E10.5 embryos were consequently incubated in isopropanol, isopropanol-Paraplast (1:1), and Paraplast (58°C) for several hours and finally embedded in Paraplast Plus (Oxford Scientific, St. Louis, MO). Sections (20–25 μ m) were cut in sagittal and transverse planes, mounted in slides, and photographed under Nomarski optics using an Olympus Vanox microscope.

RESULTS AND DISCUSSION

Identification of a Gene Down-Regulated in *All1* Knockout ES Cells. Recently we established *All1* ^{+/+} and ^{-/-} knockout ES cell lines, 48^{all1+/-}, 48S^{all1-/-}, and 48N^{all1-/-}, from parental AB2.1 ES cells and showed that these ES cells are deficient in hematopoietic differentiation in an *in vitro* colony-

formation assay (22). To identify genes down-regulated in the *All1* knockout ES cells, we examined the expression of approximately 15,000 genes in these, as well as in parental ES cells, by using a differential display kit. We carried out 240 primer-combination arbitrary reverse transcription-PCRs (RT-PCRs) and compared the bands derived from AB2.1 RNA to those derived from RNAs of 48S^{all1-/-} and 48N^{all1-/-} cells. As expected, the great majority of bands displayed similar intensity in the patterns of the three cell lines. The intensity of several bands was reduced in *All1* ^{-/-} knockout ES cells. However, these differences were proved to be false or minor by subsequent RT-PCR or Northern blot analysis. Nevertheless, one band (set 21G composed of the primer 5'-anchor dN6-TCTCTGG-3' and dTG-5'-TTTTTTTTTTTTTG-3') was reproducibly reduced in *All1* ^{-/-} knockout ES cells (indicated by arrow in Fig. 1A, 21G; lanes 48S^{all1-/-} and 48N^{all1-/-}). DNA corresponding to this band was excised from the gel, reamplified by PCR using the same primers, cloned, and sequenced. Using this DNA of 161 bp as a probe in Northern blot analysis, two mRNA species (1.4 and 2.5 kb) were observed in parental AB2.1 cells. These transcripts were considerably less abundant in 48S^{all1-/-}, 48N^{all1-/-} cells, as well as in 48^{all1+/-} cells (Fig. 1B). Sequence of this DNA indicate no relationships to known genes present in the database, but exhibited homology with the human anonymous sequences—expressed sequence tags (ESTs) gb/T64905 and

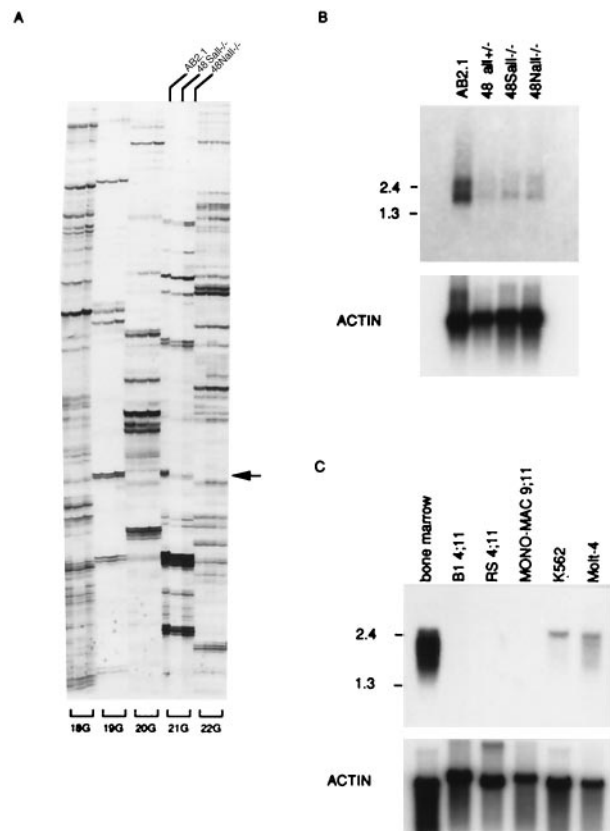


FIG. 1. Identification of a transcript down-regulated in *All1* knockout ES cells. (A) Differential display. 18G, 19G, 20G, 21G, and 22G indicate specific primer sets. The three lanes for each set correspond to the ES cell line from which the RNA was amplified: AB2.1 wild type (left lane), 48S^{all1-/-} (center lane), and 48N^{all1-/-} (right lane). Arrow points to the band apparent only in the wild-type cells. (B) Northern analysis of *Arp1* expression in ES cell lines. Two micrograms of polyadenylated RNA were electrophoretically resolved and examined for hybridization to the 161-bp probe (see text). (C) Northern analysis of *Arp1* expression in human leukemia cell lines. Analysis as in B. The 300-bp human cDNA (see text) was utilized as a probe. Numbers on the left correspond to size markers (kb).

gb/R77812. By PCR amplification, we obtained a 300-bp fragment of the human homologue. Using the 161- and 300-bp DNA fragments as probes, the expression of the gene was examined in mouse and human adult tissues. Transcripts were detected in mouse skeletal muscle, mouse kidney, human skeletal muscle, human placenta (not shown), and human bone marrow (Fig. 1C). In contrast, other tissues showed little expression. In analysis of polyadenylated RNAs from leukemic cell lines in which *ALL1* is rearranged, *ARPI* transcripts were not detected in RS4;11 [with t(4:11)], B1B [with t(4:11)], Mono-Mac6 [with t(9:11)] (Fig. 1C), or MV4;11[with t(4:11)], THP-1 [with t(9:11)], and ML-2 [with t(6:11)] (data not shown). In contrast, *ARPI* transcripts are detectable in human leukemia cell lines and bone marrow cells without 11q23 gene rearrangements (Fig. 1C).

Isolation and Characterization of Mouse and Human Full Length *ARPI* cDNAs. To isolate the full-length mouse cDNA, we prepared a cDNA library from the mouse AB2.1 wild-type ES cell line. One million independent colonies were screened, and 17 positive clones were isolated. All clones were sequenced, and 10 appeared to contain full-length cDNA. We designated the corresponding gene *ARPI* (*ALL1* responsive gene 1). Physical mapping and sequence analysis of the cDNAs indicated two species (*ARPIa* and *ARPIc*) that varied at 5' segments including the initiation codon ATG (see later in text). The predicted ORFs for the murine *Arp1a* and *Arp1c* encoded 271 and 324 aa, respectively. The human 300-bp DNA fragment was used as a probe to clone the human homologue. Two of the cDNA species (*ARPIa* and *ARPIc*) corresponded to the mouse species, and a third (*ARPIb*) contained additional sequences derived from an internal, alternatively spliced exon (see later in text). Comparison of the sequences of human and murine *ARPIa* and *ARPIc* cDNAs showed striking conservation, 90 and 89% nucleotide identity and 99 and 97% amino acid identity, respectively. The database search indicated that *ARPIa/Arp1a* and *ARPIb* are likely to be identical to *Ptx2a* (26)/*Otx2* (27)/*RIEG* (28) and *Ptx2b* (26), respectively. However, the N-terminal domain of *ARPIc* did not correspond to any known gene. Amino acid sequence and domains of *ARPIc* are shown in Fig. 2A and B, respectively. *ARPI* also has striking homology with *Ptx1/Potx*, which is thought to be a pituitary cell fate determinant (29, 30). The amino acid sequences of the two genes are 82% similar along the whole coding sequence, increasing to 86% without the N-terminal regions and increasing further to 98% when considering only homeodomains (Fig. 2C). The predicted Arp1 protein contains a previously unknown motif (Fig. 2), which we also found at the C terminus of some homeobox genes that belong to the "Phox-Aristales" group (31). We named this motif PAR (Phox-ARPI-Aristales) motif. This motif was also noticed by Semina *et al.* (28).

RIEG and *Ptx1/Potx* were described as bicoid-related homeobox transcription factors (28–30) because of lysine residue at position 9 of the third helix according to the classification of Burglin (32). We will further specify the *ARPI/RIEG/Ptx2/Otx2* and *Ptx1/Potx* proteins as members of a distinctive family within the "Ortoïd-Ortoïdentical" class of homeobox genes based on the definition of "group-specific patterns" and "intergroup-specific motifs" (31).

Chromosome Location and Genomic Structure of *ARPI*. Using a fragment of the *ARPI* gene as a probe in Southern analysis of DNAs from somatic cell hybrids and applying *in situ* hybridization, we mapped human *ARPI* to chromosome region 4q24–25. For analysis of the genomic structure, cosmid clones 2, 4, and 5 were isolated by screening 5 × 10⁵ colonies of human genomic cosmid library (CLONTECH). Physical mapping and sequencing of these cosmid clones enabled us to determine the genomic structure of the human *ARPI* gene (Fig. 3). The gene spans 30 kb. Alternative splicing gives rise to the three transcripts (*ARPIa*, *ARPIb*, and *ARPIc*). All

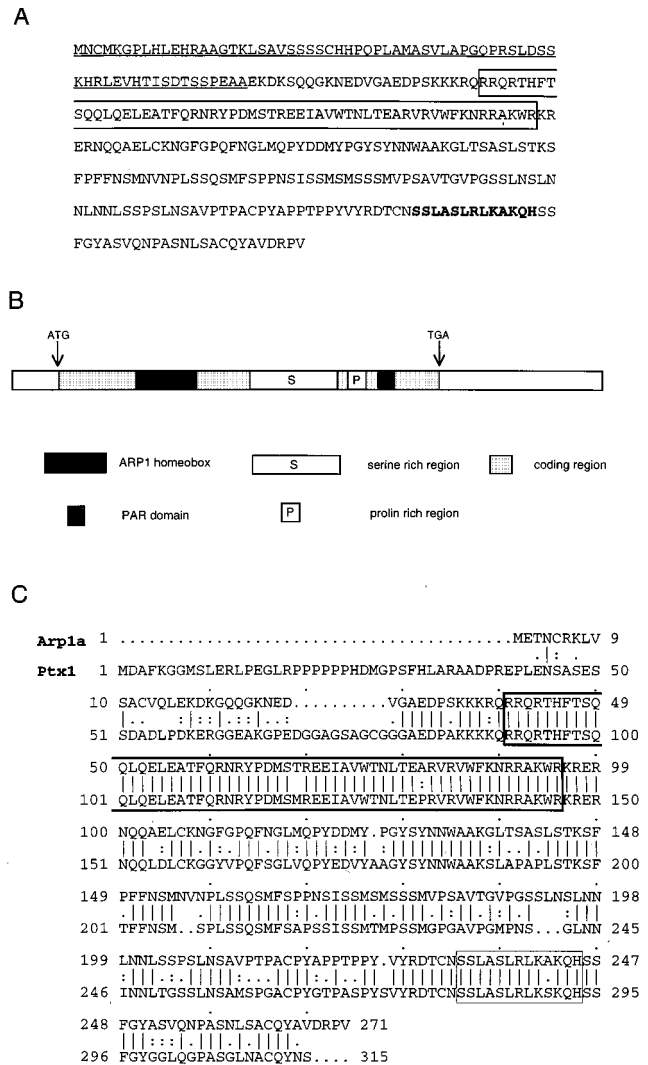


FIG. 2. Amino acids sequence, domains, and homology of *ARPI*. (A) Amino acids sequence of *ARPIc*. Underline indicates unique domain of *ARPIc*; box indicates homeodomain. Bold letters are short motifs mentioned by Semina *et al.* (28). (B) Schema of domain of *ARPIc*. (C) Similarity between the *Arp1a* and *Ptx1* proteins. The sequences were aligned by using the algorithm BESTFIT.

transcripts include sequences of exon 5, which encodes the most 5' region of the homeodomain, and of exon 6, encoding the 3' 11 aa of the homeodomain.

The AT-Hooks Domain of the All1 Protein Binds to a Sequence Positioned Upstream of the *ARPI* Transcription Start Site. To investigate the possibility that *ALL1* regulates *ARPI* expression through direct protein–DNA association, we searched for All1-binding sites within the human *ARPI* locus. "AT-hook" domains of several proteins were previously shown to be involved in direct binding to DNA (33, 34). Therefore, we expressed in *E. coli* All1 ATH, tagged at the C terminus with six histidines, and inserted it into the PET bacterial expression vector (Novagen). As a control, we expressed a similar-sized All1 polypeptide spanning the MTase region. Cosmid 2, containing the entire *ARPI* gene, was digested with *Sau3AI* enzyme, and the fragments were ligated to PCR adapters. These fragments were subjected to binding-site selection (BSS) on nitrocellulose filters spotted with each of the two bacteria-synthesized All1 polypeptides in the presence or absence of either BSA or HeLa cell nuclear extracts. The bound DNAs were eluted, PCR-amplified, and analyzed electrophoretically (Fig. 4A). A unique 0.5-kb DNA fragment was amplified after selection with the ATH polypeptide in the

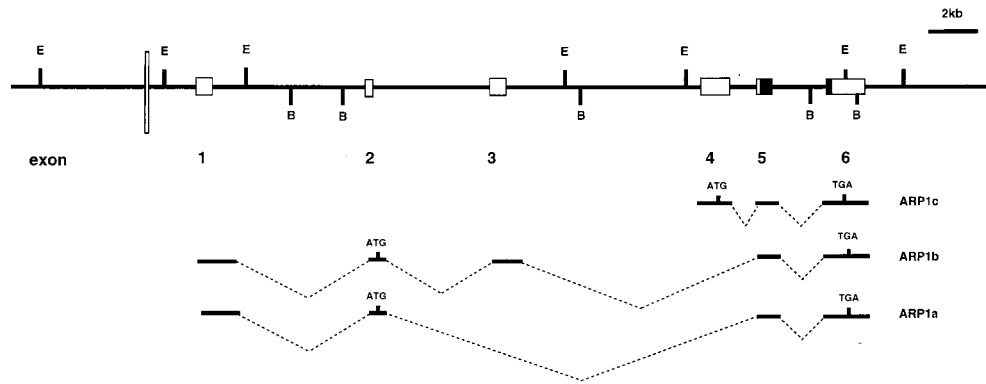


FIG. 3. Genomic structure of human *ARP1*. Exons are depicted by lightly shaded boxes. Solid box indicates *ARP1* homeodomain. Box sizes do not correspond to precise length of exons. The All1 AT-hook binding site is noted by thin, white bar. E and B correspond to *EcoRI* and *BamHI* sites, respectively. Exons included within *ARP1* transcripts are shown in the middle.

presence of BSA or HeLa cell nuclear extract. This fragment was not selected with the MTase polypeptide, but could be selected with ATH alone in a second cycle of binding and PCR amplification. A fragment of 0.19 kb that was selected by HeLa cell nuclear extract in the presence or absence of ATH or MTase polypeptides was not studied further. The selected 0.5-kb fragment containing the ATH peptide-binding site was subcloned and its sequence was determined (Fig. 4B). This DNA was found to reside within a 2-kb *EcoRI* fragment located at the 5' end of cosmid clone 2, 1.7 kb upstream of the 5' terminus of *ARP1a* and *ARP1b* cDNAs.

To narrow the binding site to ATH within the 524-bp DNA, an ExoIII protection experiment was carried out. As shown in Fig. 5A, the ATH peptide protected a 340-bp fragment of the 5' end-labeled DNA (Fig. 5A Left) and 235-bp fragment of the 3' end-labeled DNA (Fig. 5A Right) in a dose-dependent manner; no protection was found by using the MTase peptide. This indicated that the ATH-binding site spanned nucleotides 289–340 (Fig. 4B, open box). Next, a 220-bp DNA composed of nucleotides 202–421 (includes the ExoIII-protected region) was subcloned and used in a DNase I footprinting study. Nucleotides 268–321 of the lower strand (numbering starts at nucleotide 1 in Fig. 5B) and nucleotides 242–257 and 306–325 of the upper strand were consistently protected (Fig. 5B). Thus, the results of two types of assays indicated that All1 “AT hooks” polypeptide binds *ARP1* DNA in a region of around 100 bp located ≈1.7 kb upstream of the transcription start site. Utilizing this region (nucleotides 250–344) in electrophoretic mobility shift assay we observed specific binding to a protein complex in nuclear extracts from cultured cells (not shown).

Expression of *Arp1* During Mouse Embryogenesis. We examined the expression of the *Arp1* gene at different stages of mouse embryo development by Northern analysis. No expression was detected at E7.0, whereas we found strong expression at E11, E13, and E15 embryos (not shown). Ex-

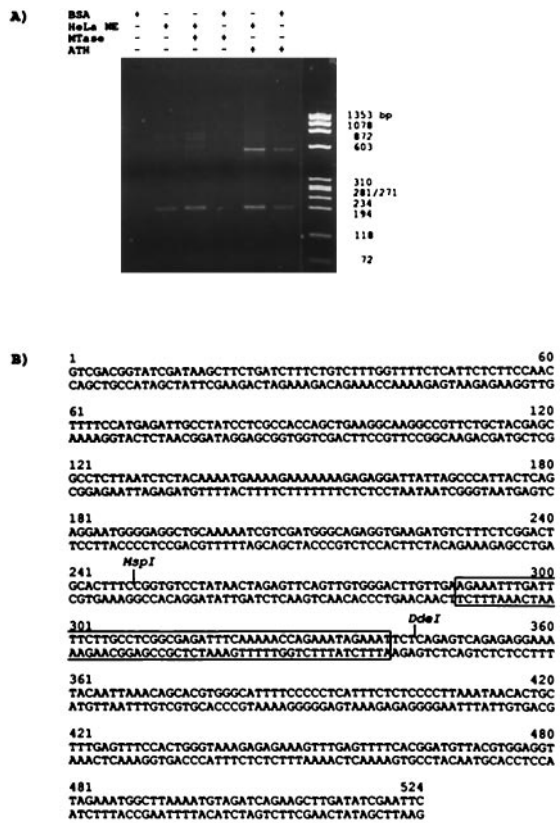


FIG. 4. Binding-site selection of *ARP1* genomic fragments. (A) Bacterially synthesized All1 polypeptides spanning the “AT-hooks” domain (ATH) or the region homologous to methyltransferases (MTase), together or without BSA or HeLa nuclear extract (NE), were reacted with the fragmented *ARP1* genomic DNA. Bound fragments were PCR-amplified and identified by agarose gel electrophoresis. (B) The sequence of the ATH-selected DNA fragment. Sequence protected by ATH from ExoIII digests is boxed.

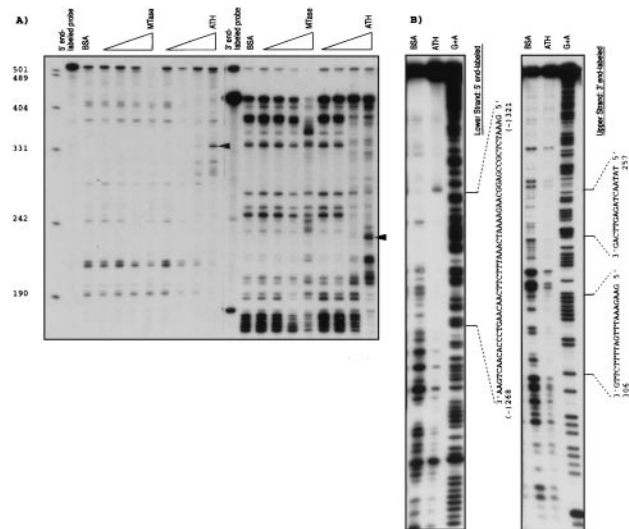


FIG. 5. Fine mapping of *ARP1* genomic sequence that binds the ATH polypeptide. (A Left and Right) Protection from ExoIII digestion of *ARP1* DNA segment end-labeled at the 5' and 3' ends, respectively. Bands protected specifically by increasing amounts of the ATH polypeptide are indicated by arrows. (B) Protection from DNaseI digestion of *ARP1* DNA segment labeled at the 5' or 3' end.

pression of the *Arp1* gene was also analyzed by whole-mount *in situ* hybridization in day E8.0–E15.5. As shown and mentioned in Fig. 6, we detect *Arp1* mRNA in a number of tissues in day E9.5–E10.5. The *All1* gene has also been shown to be widely expressed throughout embryonic development in the



FIG. 6. Spatial distribution of *Arp1* transcripts during mouse development. (A) Lateral view of E9.5 embryo. Hybridizable RNA is observed in the abdominal area (Abd), in surface ectoderm of the maxillary (Mx) and mandibular (Ma) regions of the first branchial arch. Arrowhead points out the mesenchymal component of the mandibular arch. Staining of the optic vesicle (Ov) is because of the nonspecific trapping of colored conjugates as evidenced by reaction with a sense probe. No signal was observed in spinal cord (Sc). (B) Anterior view of the head of E9.5 embryo. Facial part has been removed for better visualization of the signal. Staining is observed in the developing Rathke's pouch (Rp) and in two bilateral structures (marked by white asterisks), still to be identified. (C) Anterior view of the first branchial arches of E9.5 embryo. Arrowheads point to the sharp bands of *Arp1* expression along the surface ectoderm of the mandibular component of the first branchial arch. (D) Lateral view of E10.5 embryo. Expression is detected in a number of new structures: in the dermomyotomal compartments of the somites (dmy), in the presumable migratory myogenic precursors (of the limb muscle) localized in the ventrolateral myotome (vlmy) at the forelimb level. The signal is also seen at the periphery of the eye. Strong staining is observed in the restricted domains of the forebrain (white arrows) at the area of the cranial flexure (cf). Black triangle corresponds to the midbrain/forebrain boundary; fl, forelimb; hl, hindlimb. (E) High magnification of the anterior view of head of E10.5 embryo (facial part has been removed); expression of *Arp1* is sharply localized to the Rathke's pouch. Bilateral staining previously detected becomes stronger and more compact. (F) Dorsolateral view of E10.5 embryo. *Arp1* specific signal is seen in presumable ventral (v) and dorsal (d) myoblast cells within the forelimb (fl). (G) Anteriolateral view of E10.5 head showing bilateral distribution of *Arp1* signal in the forebrain. (H) Sagittal section of the prestained E10.5 embryo. Expression is stronger in the pars disilis (PD) than in pars intermedia (PI) within the Rathke's pouch. Expression is seen in the ectoderm and mesenchyme (arrowhead) of the mandibular branch (Ma) of the first branchial arch. (I) Staining was detected on sagittal section in epithelial cells of the bronchi (arrow) within the lung bud (Lb). (J and K) Transverse sections of the E10.5 embryo at the level of the forelimbs. Open triangle points to the myoblast precursor cells migrating into the proximal portion of the forelimb. Staining is also revealed in the mesenchyme of the dorsal mesentery (Dm); G, gut. [Bars = 200 μ m (A, D, and F–I), 100 μ m (J and K), and 50 μ m (B and C).]

neural and somitic tissues from E7.0 into adulthood (9). Interestingly, *Arp1* mRNA is detected in ectodermal cells in first branchial arch (Fig. 6A and C), ventrolateral dermomyotome of early somites (Fig. 6D and F), and neural tissues (Fig. 6D and G), overlapping the expression of *All1*.

ARPI Is a Target for ALL1. The *ARPI* gene was isolated by virtue of its down-regulation in ES cells, in which both alleles of *ALL1* were disrupted by homologous recombination (22). The disruption eliminated *ALL1* third exon, which encodes 885 aa, and contains the "AT-hook" motifs. It is conceivable that the deletion of this large and highly conserved segment of the protein disrupted its function. The elimination of the "AT-hooks" domain presumably interferes with the capacity of the protein to bind its DNA targets. In three ES cell lines in which one or two alleles of *ALL1* were disrupted, transcription of *ARPI* was reduced by more than 10-fold. In addition, in six leukemic cell lines in which *ALL1* is rearranged *ARPI* is not expressed. This is consistent with the idea (12, 14, 21) that *ALL1* rearrangements result in loss of function of the gene and, therefore, in down-regulation of its targets, such as *ARPI* (shown in this work) or *Hox* genes (22). The suggestion that *ARPI* is a primary target of *ALL1* is based on our findings that the *All1* "AT hooks" domain binds *in vitro* specifically to an *ARPI* sequence located upstream to the transcription initiation point and that the expression of *All1* coincides with *Arp1* expression pattern in several tissues including the dermomyotome. However, this suggestion would have to be supported by the demonstration *in vivo*, and possibly *in vitro*, of *All1*-dependent transcription of a reporter gene linked to the *ARPI* sequence. Our preliminary gel-shift experiments show that the sequence binds specifically to a protein complex present in mammalian nuclear extracts (data not shown). It still remains to be shown that the *All1* protein is a constituent of this complex.

We have identified a genomic *ARPI* sequence to which the *All1* AT-hook polypeptide binds *in vitro*. It was previously shown that AT hooks binding domains may bind to AT-rich sequences in a structure-specific rather than sequence-specific manner (5, 35). However, in selecting a 0.5-kb fragment out of 40 kb, the binding of *All1* polypeptide appears to be more specific in our assay. All chimeric proteins produced as a result of *ALL1* translocations, as well as *All1* proteins resulting from partial tandem duplications, retain both the *All1* AT-hook domain and the MTase domain (12, 15), suggesting that these domains are critical for the leukemogenic activity of the proteins. The DNA-binding capacity of the *All1* AT-hook polypeptide suggests that the chimeric *All1* proteins can bind to DNA targets of the normal *All1* protein. However, this binding might be nonproductive.

Potential Role of *ARPI* in Oncogenesis. *ALL1* is frequently rearranged in different kinds of acute leukemias, such as acute lymphocytic (ALL), acute myelomonocytic, acute monocytic, and acute myeloid leukemias. Furthermore, leukemic cells of the majority of patients with the t(4;11) chromosomal translocation show both precursor B cell and myeloid markers (36) and are considered to be "biphenotypic." These features suggest that alteration in expression of the target gene(s) due to *ALL1* abnormalities may affect differentiation of multipotential hematopoietic stem cell along the lymphoid and myeloid lineages. Thus, *ALL1*'s target genes involved in leukemogenesis could play an important role in the differentiation of hematopoietic stem cells to specific cell lineage. Interestingly, *ARPI* has striking homology with *Ptx1/Potx*, a pituitary cell fate determinant that plays a role in the differentiation of pituitary cells. The strong expression of *ARPI* in bone marrow cells and in skeletal muscle cells raises the possibility that it is involved in the differentiation of hematopoietic cells (and skeletal muscle cells) and therefore could be a mediator of leukemogenesis resulting from *ALL1* alterations. The lack of *ARPI* expression in all six leukemia cell lines with 11q23

translocations examined is consistent with this notion and the hypothesis (12, 14, 21) that *ALL1* rearrangements result in a loss of function of the gene and therefore in down-regulation of its targets.

While this work was in progress, *ARPIa* and *ARPIb* were isolated by others in three different contexts (26–28). The major point in those studies was that *ARPI*, termed *RIEG*, or *Otlx2*, or *Ptx2* is expressed in the pituitary gland, in the eye, and in restricted brain tissues. Moreover, the gene was found to be directly involved in Rieger syndrome (28).

Based on binding of the trx protein to many sites on polytene chromosomes (10, 11), we expect *ALL1* to regulate a variety of mammalian genes in addition to the Hox loci. *ARPI* appears to be the first example of these target genes.

We thank Dr. Victor M. Gindilis for helpful discussions and valuable suggestions. We are grateful to Dr. Kay Heubner and Dr. Jim Jaynes for critical reading of the manuscript. We particularly acknowledge the contribution of Drs. Richard R. Schmidt and Kenneth P. Chepenik to the analysis of whole-mount hybridized embryos. These studies were supported by National Cancer Institute Grants CA 39860 and CA 50507, and by grants from the Tobacco Research Council, U.S.–Israel Binational Fund, and the Minerva and Forschheimer Foundations.

- Gu, Y., Nakamura, T., Alder, H., Prasad, R., Canaani, O., Cimino, G., Croce, C. M. & Canaani, E. (1992) *Cell* **71**, 701–708.
- Tkachuk, D. C., Kohler, S. & Cleary, M. L. (1992) *Cell* **71**, 691–700.
- Mazo, A. M., Huang, D. H., Mozer, B. A. & David, I. B. (1990) *Proc. Natl. Acad. Sci. USA* **87**, 2112–2116.
- Ma, Q., Alder, H., Nelson, K. K., Chatterjeel, D., Gu, Y., Nakamura, T., Canaani, E., Croce, C. M., Siracusa, L. D. & Buchberg, A. M. (1993) *Proc. Natl. Acad. Sci. USA* **90**, 6350–6354.
- Zeleznick-Le, N. J., Harden, A. M. & Rowley, J. D. (1994) *Proc. Natl. Acad. Sci. USA* **91**, 10610–10614.
- Prasad, R., Yano, T., Sorio, C., Nakamura, T., Rallapalli, R., Gu, Y., Lehkowitz, S., Croce, C. M. & Canaani, E. (1995) *Proc. Natl. Acad. Sci. USA* **92**, 12160–12164.
- Ingham, P. W. (1983) *Nature (London)* **306**, 591–593.
- Kennison, J. A. (1995) *Annu. Rev. Genetics* **29**, 289–303.
- Yu, B. D., Hess, J. L., Horning, S. E., Brown, G. A. & Korsmeyer, S. J. (1995) *Nature (London)* **378**, 505–508.
- Kuzin, B., Tillib, S., Sedkov, Y., Mizrokhi, L. & Mazo, A. (1994) *Genes Dev.* **8**, 2478–2490.
- Chinwalla, V., Jane, E. P. & Harte, P. J. (1995) *EMBO J.* **14**, 2056–2065.
- Canaani, E., Nowell, P. C. & Croce, C. M. (1995) *Adv. Cancer Res.* **66**, 213–234.
- Bernard, O. A. & Berger, R. (1995) *Genes Chromosomes Cancer* **13**, 75–85.
- Schichman, S. A., Caligiuri, M. A., Gu, Y., Strout, M. P., Canaani, E., Bloomfield, C. D. & Croce, C. M. (1994) *Proc. Natl. Acad. Sci. USA* **91**, 6236–6239.
- Schichman, S. A., Canaani, E. & Croce, C. M. (1995) *J. Am. Med. Assoc.* **273**, 571–576.
- Lochner, K., Siegler, G., Fuhrer, M., Gereil, J., Berk, J. D., Fey, G. H. & Marschalek, R. (1996) *Cancer Res.* **56**, 2171–2177.
- Chaplin, T., Ayton, P., Bernard, O. A., Saha, V., Della, V., Hillion, J., Gregorini, A., Lillington, D., Berger, R. & Young, B. D. (1995) *Blood* **85**, 1435–1441.
- Cairns, B. R., Henry, N. L. & Kornberg, R. D. (1996) *Mol. Cell Biol.* **16**, 3308–3316.
- Kuriyama, M., Harada, N., Karoda, S., Yamamoto, T., Nakafuku, M., Iwamatsu, A., Yamamoto, D., Prasad, R., Croce, C. M., Canaani, E. & Kaibuchi, K. (1996) *J. Biol. Chem.* **271**, 607–610.
- Nakamura, T., Alder, H., Gu, Y., Prasad, R., Canaani, O., Kamada, N., Gale, R. P., Lange, B., Crist, W. N., Nowell, P. C., *et al.* (1993) *Proc. Natl. Acad. Sci. USA* **90**, 4631–4635.
- Prasad, R., Lehkowitz, D., Gu, Y., Alder, H., Nakamura, T., Saito, H., Hubner, K., Berger, R., Croce, C. M. & Canaani, E. (1994) *Proc. Natl. Acad. Sci. USA* **91**, 8107–8111.
- Fidanza, V., Melotti, P., Yano, T., Nakamura, T., Bradley, A., Canaani, E., Calabretta, B. & Croce, C. M. (1996) *Cancer Res.* **56**, 1179–1183.
- Liang, P. & Pardee, A. B. (1992) *Science* **257**, 967–971.
- Wu, C. (1985) *Nature (London)* **317**, 84–87.
- Conlon, R. A. & Rossant, J. (1992) *Development* **116**, 357–368.
- Gage, P. J. & Camper, S. A. (1997) *Hum. Mol. Genet.* **6**, 457–464.
- Muccielli, M. L., Martinez, S., Pattyn, A., Goridis, C. & Brunet, J. F. (1996) *Mol. Cell. Neurosci.* **8**, 258–271.
- Semina, E. V., Reiter, R., Leysens, N. J., Alward, W. L. M., Small, K. W., Datson, N. A., Siegel-Bartelt, J., Bierke-Nelson, D., Bitoun, P., Zabel, B. U., *et al.* (1996) *Nat. Genet.* **14**, 392–399.
- Lamonerie, T., Tremblay, J. J., Lanctot, C., Therrien, M., Gauthier, Y. & Drouin, J. (1996) *Genes Dev.* **10**, 1284–1295.
- Szeto, D. P., Ryan, A. K., O'Connell, S. M. & Rosenfeld, M. G. (1996) *Proc. Natl. Acad. Sci. USA* **93**, 7706–7710.
- Gindilis, V., Banikazemi, M., Vyasankin, A., Verlinsky, O., Matveyev, I. & Verlinsky, Y. (1994) *J. Assist. Reprod. Genet.* **11**, 244–269.
- Bürglin, T. R. (1993) in *A Guidebook for Homeobox Genes*, ed. Duboule, D. (Oxford Univ. Press, Oxford), pp. 25–71.
- Thanos, D. & Maniatis, T. (1992) *Cell* **71**, 777–789.
- Grosschedl, R., Giese, K. & Pagel, J. (1994) *Trends Genet.* **10**, 94–100.
- Solomon, M. J., Strauss, F. & Varshavsky, A. (1986) *Proc. Natl. Acad. Sci. USA* **83**, 1276–1280.
- Stong, R. C., Korsmeyer, S. J., Parkin, J. L., Arthur, D. C. & Kersey, J. H. (1986) *Blood* **67**, 391–397.