

Biochemical and Biophysical Research Communications 294 (2002) 23-34



www.academicpress.com

A complex containing PBX2 contributes to activation of the proto-oncogene *HOX11*

R.L. Brake,* U.R. Kees, and P.M. Watt

Division of Children's Leukaemia and Cancer Research, Telethon Institute for Child Health Research and Center for Child Health Research, The University of Western Australia, PO Box 855, West Perth, WA 6872, Australia

Received 18 April 2002

Abstract

Ectopic expression of the homeobox gene *HOX11* is associated with a significant proportion of childhood T-cell acute lymphoblastic leukaemias (T-ALLs). We hypothesise that one mechanism of gene deregulation involves overcoming the silencing mechanism(s) of gene expression present in normal cells. Here, we describe a search for trans-acting factors that control transcriptional activity from a distal 5′ region of the *HOX11* promoter. We have identified a region of this promoter which contributes significantly to *HOX11* activation and two distinct regulatory elements are involved. First, a PBX2 Regulatory Element *PRE-1048* has been identified which contains a novel DNA-binding sequence and mediates significant activation of the *HOX11* gene in K562 cells. This is the first report of a homeobox gene being specifically regulated by PBX2 and the second report of a vertebrate homeobox target gene of a PBX protein. The PREP1 protein was also shown to be part of the *PRE-1048*-binding complex. The other regulatory element we describe here *RE-1019* contains little sequence conservation to known transcription control elements. It appears that this element is a novel sequence that binds an as yet unidentified factor, mediating significant activation of the *HOX11* gene in K562 cells. This is the first detailed report of elements that mediate regulation of the proto-oncogene *HOX11*. © 2002 Elsevier Science (USA). All rights reserved.

HOX11 is a proto-oncogene deregulated in childhood T-cell acute lymphoblastic leukaemias (T-ALLs). It is located on chromosome 10, distant from any of the HOX gene clusters [1]. HOX11 homologues exist in mice [2], Drosophila [3], and Xenopus [4]. HOX11 expression occurs in various tissues during embryogenesis, including developing nervous tissue of the hind brain in both mice and *Xenopus*. Mice express *HOX11* in the splenic anlage where it appears to direct the genesis of the spleen. Interestingly, HOX11 expression appears to decline in these tissues, such that at birth expression is no longer detectable. HOX11 null mice are born asplenic [5,6], suggesting that the major role of *HOX11* in mice is the development of the spleen during embryogenesis [6,7]. Past this stage in development, the gene becomes silenced via an unknown mechanism.

The processes by which *HOX11* becomes deregulated in T-ALL remain an area of current research. It is known

that a common mechanism involves the juxtaposition of the *HOX11* gene to one of the two T-cell receptor (TCR) loci [8–10]. These balanced chromosomal translocations account for at least 5% of all T-ALLs and have revealed the presence of a breakpoint cluster in the 5′ regulatory region of *HOX11*. Recent evidence suggests that there is an alternative mechanism of *HOX11* deregulation that accounts for a much larger proportion of T-ALL patient specimens which express *HOX11* [11,12]. The detailed mechanism by which *HOX11* is deregulated in these cases is, however, not known. To unravel the mechanism of deregulation of *HOX11* expression in cancer, it is important to determine the role of transcriptional elements involved in its normal activation.

The PBX family of proteins belongs to the superfamily of homeodomain-containing proteins referred to as TALE for (Three Amino Loop Extension). The PBX family consists of four major isoforms [13,14]. Homologues also exist in the zebra fish [15,16], Caenorhabditis elegans (ceh-20), and Drosophila melanogaster (exd) [17]. The Drosophila homologue extradenticle is responsible for regulating the specific activity of homeotic selector

^{*} Corresponding author. Fax: +618-9489-7700. E-mail address: rachaelb@ichr.uwa.edu.au (R.L. Brake).

proteins (HOM C). The DNA sequence conservation in critical domains of these proteins is very high, suggesting some redundancy in protein function.

PBX proteins act as HOX co-factors by forming complexes bound to DNA and activating various target genes. The co-binding of these proteins increases the stability of the protein complex (for review see [18]). PBX has also been shown to form both dimeric and trimeric complexes, with homeobox proteins and proteins such as *Meis (Drosophila: Homothorax)* and *Prep 1 (Drosophila: Knox1)*, which can act as non-DNA-binding members of these regulatory complexes [19]. While PBX proteins were originally identified as transcriptional activators, there is some evidence to suggest that they also contain a repressive domain upstream of the homeodomain that appears to function independent of DNA binding [20]. Therefore the functional role of this family of proteins is complex and remains to be fully elucidated.

The original PBX member, PBX1a, was identified at a chromosomal breakpoint which occurs in 30% of pre B-cell ALLs [21,22]. PBX1a and its truncated isoform PBX1b are lineage restricted in their expression. Notably, cells of the lymphoid lineage are negative for PBX1a expression [13]. The common chromosomal translocation present in ALL juxtaposes the DNA-binding domain of PBX1a to the transcriptional activator E2A. The resultant fusion protein is a novel transcriptional activator that is abnormally expressed in lymphoid tumours.

PBX2 and PBX3 appear to have related functions and reveal only subtle differences in their tissue specificity as determined by Northern blot analyses. PBX2 is ubiquitously expressed in adult tissues, with the exception of heart tissue, while PBX3 is predominantly expressed in the adult ovary but is also expressed in a wide range of tissues [13]. Differential levels of expression are also observed during embryonic development where PBX1 and PBX2 appear to be co-expressed at equivalent levels, whereas PBX3 appears to be restricted to certain tissues [13]. In contrast, PBX4 gene expression is restricted specifically to tissues of the testis [14].

Very recently, a PBX mutant named *Lazarus* was identified in zebra fish and shown to be identical to *pbx4* [15]. *Lazarus/pbx4* interacts with a *HOX* gene product as expected. Intriguingly, *Lazarus* mutants mimic homoebox gene mutants in the mouse, suggesting that the Pbx4/Hox1b/Meis3/ complex appears to regulate *HOX* target genes in a manner analogous to *Drosophila extradenticle* [15]. This provides indirect evidence of a PBX protein regulating a *HOX* target gene in vertebrates.

Previously, we reported a comprehensive mapping study of the *HOX11* promoter, which used transient transfection to demonstrate that the DNA sequence between -1000 and -1059 of the *HOX11* promoter contained significant transcriptional elements [23]. This study focused on the detection of negative regulatory elements and as such failed to identify any positive

regulatory elements (PRE). We believe that the PREs described in the current report were masked by adjacent negative elements which functioned in the transient transfection assays. We now describe the fine mapping of the 5' regulatory region of the *HOX11* promoter distal to the first 1000 bp. We have identified two transcription control elements that mediate significant activation of the *HOX11* gene in a *HOX11*-expressing cell line, K562. To our knowledge, this is the first detailed characterisation of promoter elements driving transcription of *HOX11*.

Materials and methods

RE-1059 MUT 3'D

GA -3'

Oligonucleotides and DNA probes

Probes used for EMSA analysis

RE-1059 G/S 5'	sense 5'-CTAGCTGGTTGAATTCATTCAA ACACTGTAGC-3'.
	antisense 5'-CCGGGCTACAGTGTTTGAA
	TGAATTCAACCAG-3'.
Non-specific	sense 5'-CTAGCCTTTCTAAAAAGCTGGA
	CTTGAAATTTC-3'.
	antisense 5'-CCGGGAAATTTCAAGTCCA
	GCTTTTAGAAAGG-3'.
PRS	sense 5'-CGATACATCAATCAAATGGT-3'
	antisense 5'-ACCATTTGATTGATGTACT
	G-3'
1048 core PRE	sense 5'-GTTGAATTCATTCAAACACT-3'
	antisense 5'-AGTGTTTGAATGAATTCA
	AC-3'
1048 PRE MUT	sense 5'-GTTGAATGACTGACAACACT-3'
	antisense 5'-AGTGTTCTGAGTCATTCA
	AC-3'
RE-1059 G/S 5'A	sense 5'-TGGTTGAATTCATTCA-3'
	antisense 5'-TGAATGAATTCAACCA-3'.
RE-1059 G/S 5'B	sense 5'-TGAATTCATTCAAACA-3'
	antisense 5'-TGTTTGAATGAATTCA-3'.
RE-1059 G/S 5'C	sense 5'-TCATTCAAACACTGTA-3'
	antisense 5'-TACAGTGTTTGAATGA-3'.
RE-1059 G/S 3'	sense 5'-TGCCTCTTTCGAACCCTGTAG
	GA-3'.
	antisense 5'-TCCTACAGGGTTCGAAAG
	TGGCA -3'.
RE-1059 G/S 3'A	sense 5'-TGCCTCTTTCGAACC-3'
	antisense 5'-GGTTCGAAAGAGGCA-3'.
RE-1059 G/S 3'B	sense5'-TCTTTCGAACCCTGT-3'
	antisense 5'-ACAGGGTTCGAAAGA-3'.
RE-1059 G/S 3'C	sense 5'-TTCGAACCCTGTAGGA-3'
	antisense 5'-TCCTACAGGGTTCGAA-3'.
RE-1059 MUT 3'A	sense 5'-CGGCAAGAACCCTGTAGGA-3'
	antisense 5'-TCCTACAGGGTTCTTGCC
	G-3'.
RE-1059 MUT 3'B	sense5'-TCTTCGTGCTCCTGTAGGA-3'
	antisense 5'-TCCTACAGGAGCACGAAG
	A -3'.
RE-1059 MUT 3'C	sense 5'-TCTTTCGACGGCCATAGGA-3'
	antisense 5'-TCCTATGGCCGTCGAAAG

sense 5'-TCTTTCGAACCCTCGGCCA-3'

antisense 5'-TGGCCGAGGGTTCGAAA

PCR primers for DNaseI footprinting

FOR 5'-TGTCTGGAATAGCCAACCCTGG-3' REV 5'-TGCCCATGCCTCAACATACC-3'

Plasmids and their construction

The plasmid construct P-1059 used for stable transfection was generated by the digestion of a 4.8-kb (*BamHI/NotI*) restriction fragment from the 5′ regulatory region of the *HOX11* gene with *EcoRI* and *NotI*. The fragment was subcloned into a modified form of pGL3-basic, pGL3-basicΔNotI which has been previously described [23]. P-1000 was created by Exonuclease III digestion of the subclone P-1059. All subsequent oligonucleotides and DNA fragments used in this investigation were subcloned into the *Ecl*136II site of the pGL3-polylinker (Promega, USA). Unmodified vectors used in this investigation include pEGFP (Clonetech) and pPGK–Puromycin, supplied by Dr. W. Greene (Murdoch University, Western Australia).

Preparation of nuclear protein extracts

Approximately 5×10^7 cells were collected, washed in phosphate-buffered saline, divided into five tubes, and placed on ice, where they were resuspended in 400 µl ice-cold buffer A (10 mM Hepes, pH 7.9, 10 mM KCl, 0.1 mM EDTA, 0.1 mM EGTA) and allowed to swell for 20 min. Then, $25\,\mu$ l of 10% Nonident-40 was added and the cells vortexed and centrifuged. The supernatant was removed and the cells were resuspended in 100 µl cold buffer C (20 mM Hepes, 0.42 M NaCl, 1.0 mM EDTA, 1.0 mM EGTA, 20% Glycerol) and rocked vigorously for 30 min at 4 °C. The cells were then centrifuged at 4 °C and the supernatant was removed and aliquoted on ice. Each buffer contained the same cocktail of proteinase inhibitors including 0.5 mM DTT, 1 µg/ml aprotinin, 1 µg/ml leupeptin, 0.5 mM PMSF, and 20 mM sodium orthovanadate.

Electrophoretic Mobility Shift Assays (EMSAs)

Ten pmol of single-stranded oligonucleotides was end-labelled with 50 μCi of $[\gamma^{-32}P]$ dATP and annealed to 2× molar concentration of the unlabelled complementary oligonucleotide. 0.1 pmol of probe was used per reaction with 1 µg poly dIdC and 5 µg nuclear extract. The binding buffer contained 20 mM Hepes, pH 7.9, 20% glycerol, 0.1 M KCl, and 0.2 mM EDTA. All components were incubated at 4°C for 20 min before loading onto a 38× 50-cm 6% native polyacrylamide (PAGE) gel. Electrophoresis was performed at 21 V/cm at 2-4 °C in 0.5× TBE. After the gel was dried, complexes were visualised using autoradiography and Phosphoimaging (Molecular dynamics). Unlabelled competitor oligonucleotides or supershift antibodies (Santa-Cruz) were added 10 min, prior to the addition of the labelled probe. 400 ng specific (antiPBX) and non-specific rabbit polyclonal antibodies were used for supershift assays. Assignment of complexes between experiments was determined by interpolation of the RF values of each of the bands on the autoradiographs to a log/linear curve of RF values of DNA standards.

DNA footprinting

Ten pmol of the forward PCR primer were end-labelled with $50\,\mu\text{Ci}$ of $[\gamma^{-32}\text{P}]\text{dATP}$. This primer was then purified and incorporated in a hot PCR ((95 °C 8 min), 95 °C 40 s, 52 °C 40 s, 72 °C 40 s, × 30 cycles). The radiolabelled double-stranded DNA product was purified and quantitated. The DNaseI protection assay was performed by incubating $10\,\mu\text{I}$ of $5\times$ binding buffer (20 mM Hepes, $100\,\text{mM}$ KCl, $0.2\,\text{mM}$ EDTA, $1.0\,\text{mM}$ DTT, $8\,\text{mM}$ MgCl₂, 10% glycerol) with $1\,\mu\text{g}$ of pdIdC (Sigma) and various concentrations of nuclear extract. This mix was incubated at room temperature for $30\,\text{min}$. Subsequently, $20\,\text{pmol}$ PCR probe was added and further incubated for $10\,\text{min}$ at $22\,^{\circ}\text{C}$. DNaseI digestion was performed using $0.01\,\text{U}$ of DNaseI (Gibco-

BRL) in a $1\times$ DNaseI buffer freshly prepared from a $5\times$ stock (3.12 mM NaCl₂, 2.5 mM Hepes Buffer, 0.625 mM MgCl₂, 0.25 mM CaCl₂, pH 7.8). Upon completion of the digestion, each reaction was extracted using phenol/chloroform and ethanol precipitated. Each sample was then resuspended in 7 µl formamide loading buffer and heated prior to loading. The samples were run on a 8% denaturing 38×50 or 16×18 cm PAGE gel. Electrophoresis was performed under denaturing conditions at 51 V/cm for approximately 5 h. After drying, the gel was visualised using autoradiography or Phosphoimaging (Molecular dynamics).

Nucleotide sequence searches

Searches for potential transcription factor-binding sites were performed with the TESS algorithm against the TRANSFAC database and by using MacVector (Oxford Molecular Group) subsequence searching software.

Transient transfection assays

Transient transfection assays were performed as previously published [23] and the results shown represent at least three independent experiments. Analysis of the reporter gene activity using the Dual Light Assay (Tropix, Perkin–Elmer) has also been previously described [23], with the variation that analysis was conducted using the Ascent Luminometer (Lab systems).

Stable transfection

Constructs P(-1059) and P(-1000) were linearised using BamHI. Twenty micrograms of test plasmid was introduced with 5 and 3 μ g of pPGK–puromycin and pGFP, respectively. pPGK–puromycin was linearised using ScaI and pGFP linearised using HindIII. Conditions for transfection have been previously described [23]. Following transfection, cells were allowed to outgrow for a period of 3 days, prior to the addition of 1.0 μ g/ml puromycin. Cells were maintained in RPMI+glutamine with the addition of 20% foetal calf serum and puromycin. At day 14, the cells were separated on Ficoll gradient and seeded in new media at a density of 1×10^5 /ml. Cells were then outgrown and harvested at days 18 and 20. At each time point, the cells were checked for viability (trypan blue) and GFP expression (via flow cytometry) and lysed for luciferase analysis.

Flow cytometry

Flow cytometry was used to determine the percentage of pooled clones that had stable integration of the co-transfection control pGFP into open chromatin. A gated population of cells was identified above the autofluorescence background level, determined from transfectants containing the pPGK-puromycin construct alone. Then, a 485-nm band-pass filter was used to analyse 10 000 gated cells. GFP emissions were determined on days 18 and 20 posttransfection. There was a Gaussian distribution of GFP emission. The mean level of fluorescence for P(-1059) ranged between 56 and 69 units, for P(-1000) 47 and 60 units and the vector only control ranged between 40 and 49 units. The percentage of cells expressing GFP varied by less than 4% at each time point. These results suggested that a complex population of pooled cells was present and a variety of integration sites and copy numbers of integration existed within the pool of cells. Thus, these results enabled us to make a valid comparison between the transcriptional activity mediated by each construct.

Cell culture and transfections

The K562 [24] cell line has been previously described. Growth conditions and electroporation procedures were followed as previously published [23].

Results

Significant transcription control elements found 1kb upstream of the HOX11 gene

To search for the presence of functionally significant transcription control elements outside of the first 1000 bp of the HOX11 promoter, a series of constructs was introduced into the erythroleukaemic cell line, K562. These cells express HOX11 and as such contain all the necessary elements required for its transcriptional activation. Our previous study had identified potent transcriptional elements in the distal HOX11 promoter [23]. Because of concerns about chromatin-dependent effects, we repeated the promoter analysis using stable integration of reporter constructs into the genome. Here, we describe the detailed analysis of two positive elements identified in this analysis. Fig. 1 demonstrates the location of all probes relative to the HOX11 locus 10q24. The two defining constructs integrated into the K562 cells were deletion mutants terminating at -1000 and -1059 bp, respectively, and linked to a downstream luciferase reporter gene. To approximate a normal chromatin configuration, stable clones were generated. Analysis was performed on large pools of stable transfectants to minimise the influence of position effect variegation.

Stable transfection of reporter gene constructs generated complex pools containing thousands of independent clones integrated at distinct chromosomal loci

Stable maintenance of reporter constructs was achieved by co-transfection with the plasmid pPGK–puromycin and selection for puromycin resistance. Co-transfection of limiting quantities of another plasmid expressing the GFP gene enabled us to determine the percentage of cells expressing GFP via stable integration into transcriptionally permissive open chromatin, as measured by flow cytometry. The percentage of cells in the stable transfectant pool that expressed detectable levels of GFP is shown in Fig. 2. Cells (10⁷) were initially transfected and more than 10⁵ cells survived the transfection process. The percentage of cells expressing GFP remained relatively constant during the subsequent culture period and was similar between constructs,

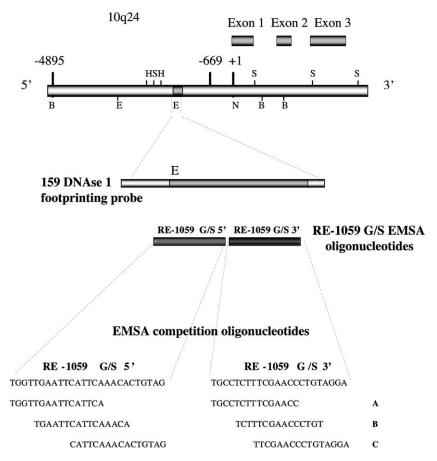


Fig. 1. A map of the *HOX11* locus demonstrating the location of double-stranded oligonucleotides and PCR products used in this investigation. The restriction enzymes are (S) *SacI*; (B) *BamHI*; (E) *EcoRI*; (N) *NotI*; (H) *HindIII*. Also annotated are the coding exons of *HOX11*. All numeric labelling refers to the start methionine as +1.

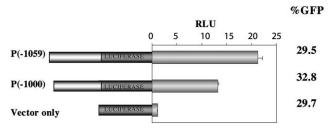


Fig. 2. Luciferase expression of stable pools of clones in K562 cells containing reporter gene constructs relevant to RE-1059. The transcriptional activity in relative light units (RLU) is compared to a vector only control (pGL3-Basic). Analysis was performed at two-time points and values were normalised relative to the vector only control. Error bars represent the range of normalised values obtained. The percentage of transfectants containing a co-transfection plasmid expressing GFP was determined by flow cytometry (% GFP) and is indicated in the right-hand column.

indicating that each pool contained more than 10,000 independent clones. Because approximately 10-fold less GFP plasmid than reporter plasmid was co-transfected, we can be confident that the complexity of the reporter construct pools also exceeded 10,000. These findings of high clonal complexity enabled us to make a reliable comparison between constructs without necessitating further correction for position effects.

Detection of luciferase was performed at two independent time points approximately 20 days post-transfection and these results are presented in Fig. 2. The construct P(-1000) demonstrated significant levels of reporter gene activation and this activation was expected and is consistent with proximal promoter elements [23]. However, the presence of only an additional 59 base pairs in the construct P(-1059) led to an approximately twofold increase in transcription over the activity of construct P(-1000) (Fig. 2). This finding using stable integration of reporter gene constructs confirmed the presence of positive regulatory elements between P(-1059) and P(-1000). These elements activated HOX11 expression to a level above that mediated by proximal promoter elements alone and were the focus of this investigation.

Defining the protein-binding sites around the distal positive element at -1059 of the HOX11 promoter

DNaseI footprinting was performed to define the transcription control elements of this distal region of the *HOX11* promoter. A 159-bp PCR product was designed to cover the region of DNA between P(-1059) and P(-1000). The location of this PCR product is represented in Fig. 1. Three prominent footprint regions were identified which coincided with this area. These protected regions (I–III) are annotated on the left of Fig. 3, as are the boundaries of the promoter fragment of interest.

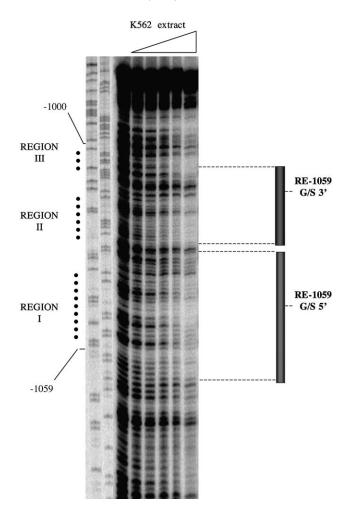


Fig. 3. DNaseI footprinting analysis of the *HOX11* promoter. The boundaries of the region of focus for this investigation (-1000 to -1059) are presented to the right-hand side of the figure. Also indicated are the boundaries of each of the oligonucleotide fragments used for subsequent analysis. To the left of the figure the protected regions of DNA and are indicated by dotted lines (I–III). Lanes 1 and 2 contain the A and T DNA sequencing reactions of the 159-bp probe while lanes 2–8 contain digests using 0.01 U of DNaseI. Lane 3 contains no nuclear extract while lanes 4–8 contained increasing amounts of nuclear extract from the K562 cell line, ranging between 2.8 and 22.0 μg.

As a result of the DNaseI footprinting data showing evidence that transcription factor complex/es may bind to these regions, oligonucleotides were synthesised which completely encompassed the protected DNA sequences of regions I and II. The boundaries of these oligonucleotides are designated RE-1059 5' and 3' and are also presented in Figs. 1 and 3.

The RE-1059 5' oligonucleotide was then used in electrophoretic mobility shift assays (EMSA) in conjunction with nuclear protein extract derived from the K562 cell line (Figs. 4 and 5). Fig. 4 reveals the formation of at least five protein complexes shown in the uppermost portion of the autoradiograph. Of these only complexes 2 and 4 appeared to be specific (compare lane 10 containing no competitor with lane 2 containing

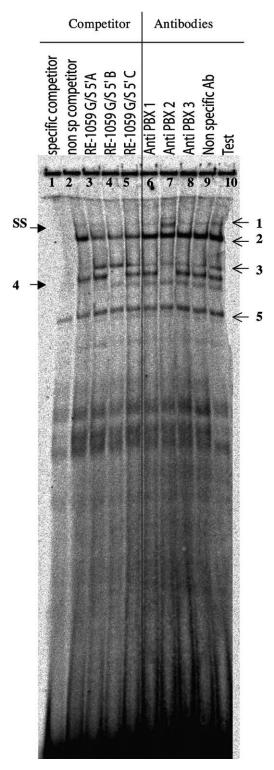


Fig. 4. EMSA analysis using probe RE-1059 G/S 5' and K562 nuclear extracts. Lane 1 contains 50-fold excess of unlabelled probe as a specific competitor, while lane 2 contains 50-fold excess unlabelled non-specific competitor. Lanes 3–5 contain 100-fold excess of the fragment competitors RE-1059 G/S 5'A–C. Supershift experiments using specified antibodies were run in lanes 6–9. Lane 10 contained the test reaction for comparison. Complexes of unknown specificity are indicated on the right margin while the PBX-containing complexes are on the left.

non-specific competitor). To rule out the influence of linker sequences at the 5' and 3' ends of the RE-1059 5' oligonucleotide, competitors were designed containing only the natural core sequence. The competitors were 16-bp long and contained between 6 and 9 bp of overlapping DNA sequence. The locations of these oligonucleotides (RE-1059 G/S 5'A, B and C) are depicted in Fig. 1. The competitions with these oligonucleotides are also presented in Fig. 4 (lanes 3–5). They reveal the absence of specific competition for complex 2, which may be an artefact due to the presence of flanking linker sequences. In contrast, complex 4 was specifically competed away by the addition of the oligonucleotide RE-1059 G/S 5'B, (compare lanes 10 and 4 in Fig. 4) while not being affected by the other two oligonucleotides.

Potential transcription factor-binding sites present in the RE-1059 G/S 5'B sequence were determined by submitting the sequence to multiple subsequence databases. Several candidate transcription factor-binding sequences were identified. Of these, no cognate DNA-binding sequence corresponded perfectly to the sequence of 1059 G/S 5'B. The field of candidate proteins was narrowed based upon statistical likelihood and the critical core sequence was identified from the fragment competitors. One such candidate was the homeodomain-containing protein PBX.

PBX2 binds to a distal positive element in the HOX11 promoter

To test the hypothesis that the HOX11 promoter may contain an element responsible for PBX protein binding, supershift assays were performed using specific antibodies directed against the family of PBX proteins, PBX1, 2, and 3. These results are shown in Fig. 4. The addition of a PBX2-specific antibody to the EMSA assay led complex 4 to supershift, (compare lanes 10 and 7). This supershifted complex was specific for PBX2 binding as no other antibody raised against the PBX family of proteins caused a supershift of complex 4. A control antibody of unrelated specificity failed to alter the mobility of any of the complexes on the gel (see lane 9). Thus, PBX2 was identified as specifically binding to the HOX11 promoter. The PBX-binding sequence identified, HOX11 PBX Regulatory Element (PRE)-1048 (5'-TTCATTCAA-3'), varies from the canonical PBX regulatory sequence (PRS) (5'-ATCAATCAA-3') but has conserved the related (5'-NTCANTCA-3') motif.

To further confirm the specificity of *PRE-1048*, competition EMSA analysis was performed using the 1048 core PRE oligonucleotide as a probe. This probe (18-bp in length) was designed so that competition analysis could be performed using the canonical PRS identified by Lu and colleagues [28] and a mutant oligonucleotide which alters essential bases of the consensus. This EMSA analysis confirmed that the PRS could

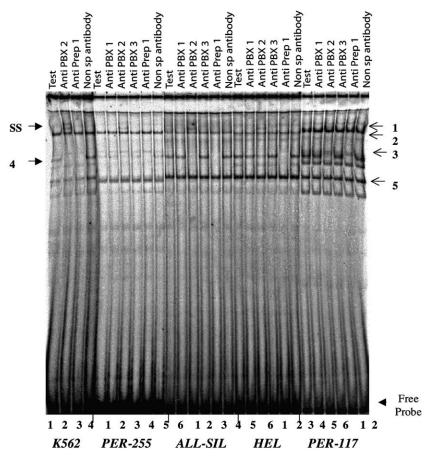


Fig. 5. EMSA supershift analysis using probe RE-60 G/S 5' and nuclear extracts from five cell lines. HOX11-positive lines included K562, PER-255, and ALL-SIL shown beneath the panel. HOX11-negative lines included HEL-900 and PER-117. The relevant antibodies added are annotated above the lanes.

specifically compete with the PBX2-containing complex, whereas the mutant oligonucleotide 1048 PRE MUT could not (data not shown). The mutant oligonucleotide contained six transversion mutations of nine essential bases of the PRS. These findings confirm that the *PRE-1048* element alone can function as a PBX-binding site.

Functional significance of PBX binding to the HOX11 promoter

To determine whether there was a relationship between PBX2 binding to the *HOX11* promoter and the role of *HOX11* as an oncogene, supershift EMSA experiments were performed using nuclear extracts, derived from various leukaemia cell lines (Fig. 5). These cell lines include T-ALL cell lines which express *HOX11* due to known translocations; PER-255 and ALL-SIL [25,26] and a T-ALL line which does not express *HOX11*; PER-117. The overall profile of the supershift analysis appeared similar between cell lines, with variation in complex formation only occurring in complexes determined to be non-specific.

The *HOX11*-expressing erythroleukaemia cell line K562 [24] was also analysed, together with a second

erythroleukaemic cell line that does not express *HOX11*, HEL-900 [23,27]. The results of these supershift experiments are also presented in Fig. 5 and reveal specific binding of the PBX2 protein to the *HOX11* promoter in all lines tested (compare complex 4 between lanes 1 and 2 for K562, lanes 1 and 3 for PER-255, lanes 1 and 3 for ALL-SIL, lanes 1 and 3 for HEL-900, lanes 1 and 3 for PER-117).

Prep 1 is part of a complex including PBX2 and the HOX11 distal positive element

The above results indicated that the PBX2 protein is widely expressed, as expected, and is a component of a functional complex that binds to the *PRE-1048* element in the cell lines tested. The large size of the complex suggested that the protein complex containing PBX2 may include additional binding partners. To investigate this further, we used an antibody directed against PREP 1. This protein, the PBX Regulatory Protein 1, has been identified as a heterodimeric partner of PBX2 on target genes of PBX2 characterised thus far. The results of this supershift experiment are also presented in Fig. 5. They indicate that the PREP 1 protein was present in all

complexes tested, since the inclusion of antibody against this protein abolished complex 4 (compare complex 4 between lanes 1 and 3 for K562, lanes 1 and 5 for PER-255, lanes 1 and 5 for ALL-SIL, lanes 1 and 5 HEL-900, and lanes 1 and 5 PER-117).

Identification of RE-1019, a second regulatory element

To assess the specificity of the other protein/DNA complexes identified by DNAseI footprinting presented in Fig. 3, an oligonucleotide corresponding to complex II was used for EMSA analysis. The results of these experiments are presented in Fig. 6 and have resulted in the identification of a second novel regulatory element RE-1019. Initial EMSA assays using K562 nuclear extract were designed to determine the specificity and potentially the identity of the protein species binding the DNA sequence contained in RE-1059 G/S 3'. In Fig. 6, the test reaction (lane 1) reveals the presence of five different complexes within this DNA region. Upon addition of increasing amounts of non-specific unlabelled DNA (of the same sequence length), complex 4 was the only unambiguously specific complex. These findings confirmed that complex 4 required the DNA within this oligonucleotide specifically to form a protein/DNA complex.

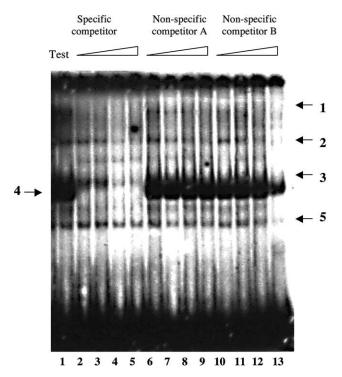


Fig. 6. EMSA analysis of *RE-1019* using the probe RE-1059 G/S 3' and nuclear extract from the K562 cells. This EMSA demonstrates specificity of complex 4 for the oligonucleotide RE-1059 G/S 3'. Increasing amounts $(4\times, 10\times, 100\times, 200\times)$ of unlabelled oligonucleotide were added of either RE-1059 G/S 3' (specific competitor) or one of the two unrelated oligonucleotides of the same DNA length as non-specific controls.

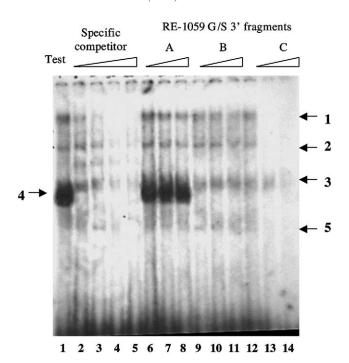


Fig. 7. EMSA analysis of *RE-1019*. Unlabelled fragment competition of RE-1059 G/S 3' with overlapping fragment competitors. Each competitor was added in increasing molar concentration $(10\times, 50\times, 100\times)$. Fragment competitors B and C both competed for the binding of complex 4.

To refine the exact boundaries required for the formation of this complex, overlapping oligonucleotides (containing only the endogenous DNA sequence) were used as unlabelled fragment competitors. The DNA sequences of these competitors are presented in Fig. 1. According to the phasing of these competitors, approximately 10 bp DNA sequence was common to two oligonucleotides. The results of this EMSA are shown in Fig. 7. With increasing amounts of the unlabelled fragment competitors, it was clearly demonstrated that the DNA sequence within RE-1059 G/S 3'A failed to compete for the formation of complex 4, confirming that the DNA sequence required for the formation of this complex was not contained within this oligonucleotide. As a result of this EMSA, boundaries of the elements functioning in complex 4 were refined to a 19-bp length of DNA, incorporating the DNA sequences common to both RE-1059 G/S 3'B and C.

Mutation analysis confirms the location of essential bases of RE-1019

The DNA sequence common to the overlapping region of the oligonucleotides B and C and 4-bp flanking on either side was then compared to the various transcription factor databases to look for candidate proteins binding this sequence. The analysis revealed multiple candidate proteins. Based on this information, a muta-

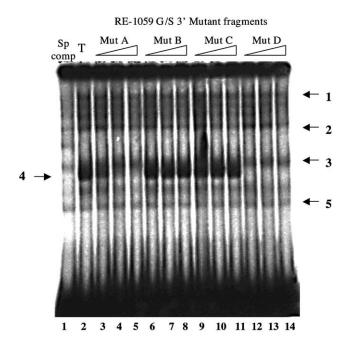


Fig. 8. EMSA analysis of *RE-1019*. Mutant competition of complex 4. Four overlapping oligonucleotides containing core mutations were added in increasing molar concentrations ($10\times$, $50\times$, $100\times$). These competitions demonstrated that Mut B and Mut C failed to compete for complex 4.

tional analysis was designed to refine the required DNA sequence for protein binding. Four 20-bp oligonucleotides containing the core DNA sequence of RE-1059 G/S 3'B and C were synthesised. These oligonucleotides contained a 6-bp mutation phased through the DNA sequence, each overlapping the next by 2 bp. These four oligonucleotides labelled Mut A–Mut D were included in EMSA analysis as unlabelled competitors (Fig. 8). They demonstrated that the DNA sequences contained within Mut B and Mut C failed to compete for the formation of complex 4, suggesting that the mutations within these oligonucleotides destroyed the DNA motif required for the formation of complex 4.

This experiment refined the DNA sequence required for the formation of complex 4–10 bp. Bioinformatic analysis of this core consensus reduced the field of likely candidate proteins further. These protein candidates were then organised according to homology with previously identified consensus sequences and statistical likelihood. This analysis generated a list of seven known factors and supershift analysis was performed on the most likely candidates. These experiments failed to demonstrate supershifting of any of the complexes formed using RE-1059 G/S 3' as a probe (data not shown).

Function of PRE-1048 and RE-1019

To determine if *PRE-1048* and *RE-1019* could re-capitulate the activation mediated by the deletion

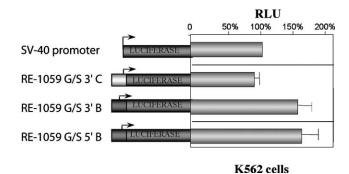


Fig. 9. K562 cells transfected with the *RE-1059* element using a heterologous promoter. Constructs introduced are indicated on the left of the figure, while the normalised luciferase activity is graphically presented on the right. The results are the means of three independent transfections. Error bars represent standard errors of those means.

mutant P(-1059) (shown in Fig. 2) and assess if there was any correlation with HOX11 expression, the DNA probes RE-1059 G/S 5'B (representing PRE-1048) and RE-1059 G/S 3'B and C (representing *RE-1019*) were linked upstream of a heterologous promoter (SV-40) and were transiently introduced into the HOX11-expressing K562 cells. These constructs were co-transfected with a constitutively expressing β -galactosidase construct to enable normalisation for transfection efficiency. The results of these experiments are presented in Fig. 9. The transcriptional activity mediated by these regions was variable and in K562 cells RE-1059 G/S 5'B and RE-1059 G/S 3'B both led to a similar fold activation as mediated by the construct P(-1059). Thus, these 16-bp fragments each stimulate significant transcription from a heterologous promoter. Therefore together these elements contribute strongly to the transcriptional activity of the *HOX11* promoter mediated by this 60-bp region outside of the first 1000 bp of the HOX11 promoter. In contrast, the fragment RE-1059 G/S 3'C appeared not to alter transcription of the heterologous promoter, suggesting that the element RE-1019 was functionally contained within the fragment RE-1059 G/S 3'B.

Discussion

This study has identified two novel regulatory elements of the *HOX11* proto-oncogene which mediate significant activation in the *HOX11*-expressing cell line, K562. It is likely that the combined effect of these elements is responsible for the activity detected in stable transfectants (Fig. 2). One of these elements, *PRE-1048*, provides the first evidence of a homeobox gene specifically regulated by a complex containing PBX2 and its co-factor PREP1. The target sequence of this binding is novel, but retains critical features with the PRS identified by in vitro binding assays [28]. The other element *RE-1019* also mediates significant activation of the

HOX11 gene, however, the factor/s binding to it remains to be elucidated. These two elements appear to contribute significantly to activation of the HOX11 proto-oncogene. We propose that these elements could play a significant role during oncogenesis to increase the level of ectopically expressed HOX11 protein in a subset of T-ALLs.

HOX11 plays a crucial role in spleen development during embryogenesis. While we have demonstrated that PBX2 specifically binds to PRE-1048 in leukaemic cells, its role in the normal regulation of HOX11 is unknown. We speculate that PBX2 may have a role in facilitating the normal development of the mammalian spleen via an effect on HOX11 expression. It is also possible that the effect of PBX2 binding (in a HOX11-positive cell line, such as K562) is indirectly related to HOX11 expression itself. According to this scenario, the products of downstream target genes of HOX11 could co-complex with PBX2 and PREP1 on PRE-1048 to form an activator complex. This hypothesis is consistent with the known feedback regulation of several other HOX genes [29–31].

To our knowledge, this is the first report of an element controlling HOX gene expression which only binds PBX2. HOXBI is one of the few identified HOX target genes that bind a PBX protein, however, this regulation involves PBX1, not PBX2 as described here. The HOXBI gene provides a precedent for a HOX gene autoregulating its function by heterodimerisation with its PBX partner on its own promoter [30]. Interestingly, Allen and colleagues [32] have demonstrated that HOX11 and PBX2 can cooperatively bind the PBX regulatory sequence (PRS) in in vitro binding assays. We investigated this possibility in the regulation of HOXII expression by PRE-1048. However, we were unable to demonstrate the presence of HOXII in a complex with PBX2 (data not shown).

Several studies have examined the in vitro binding preferences of the PBX protein family to particular DNA sequences. However, in contrast, there is little evidence for in vivo DNA targets. PBX1-specific target sequences include the TSEII element of the Somatostatin gene [33] and the CRS1 region of the bovine CYP 17 gene [34]. More recently, the pancreatic homeodomain protein gene PDX1 was identified as a specific target of PBX1b [35]. Redundancy in the binding of PBX family members to a particular target sequence has also been reported. The UEF 3 domain of the Urokinase Plasminogen Activator (UPA) gene binds a PBX/PREP1 hetrodimer which can either involve PBX1 or PBX2 [36]. Similarly, PBX1 or PBX2 can bind the FBP element on the COL5A2 gene promoter. PBX1 or PBX2 binds this element in conjunction with PREP1 and HOXB1 [37] and can mediate cell type-specific regulation, depending on the PBX species involved. This possibility was addressed in the current investigation, however, PBX2

was the only PBX protein identified as capable of binding *PRE-1048*.

Previous reports have suggested that distinct PBX/HOX protein-protein interactions appear to have different DNA-binding specificities [36,38]. To identify potential co-factors associated with PBX2 in the cell lines tested, we compared consensus sequences of PBX-binding targets. The consensus sequence identified in this investigation PRE-1048 TTCATTCAA (complement: TTGAATGAA) contained some similarity to previously identified PBX-PREP1 DNA targets, including the *urokinase* gene enhancer -TGACAG, the PRS -TGATTGAT, and an element in the HOXB1 promoter -GTGATGGAT. Our supershift experiments using an antibody directed against the PREP1 protein (Fig. 6) confirmed the presence of PREP1 in the protein complex in all cell lines tested.

The PREP1 protein has been reported to form a strong complex with PBX either with or without complexing to DNA. Since it appears to utilise structurally different dimerisation motifs than does a PBX-HOX interaction [19], it is possible for all three of these protein partners to form one complex. Berthelsen and colleagues also demonstrated that a PBX1-PREP1 complex, which bound an element in the HOXB1 promoter, had no significant effect on transcription however, if PBX1-HOXB1 dimers were preferentially formed, a moderate level of activation was observed. The level of activation was increased when all three proteins were allowed to form a trimeric complex. These findings led to the proposal that differential regulation of the element in the HOXB1 promoter may be mediated by competition for DNA binding by different PBXcontaining complexes. Similar findings were reported regarding the regulation of the COL5A2 gene promoter [37]. We propose that such a mechanism of differential regulation may be operating at the HOX11 PRE-1048 element, where components of the PBX-containing complex may differ between cell lines.

Unlike PRE-1048, RE-1019 lacks any significant sequence similarity to any previously published regulatory elements. Likewise, the factor which binds this sequence is also at present unknown. The DNA consensus sequence of RE-1019 is (5'-TCGAACCC-3') and it lacks any palindromic sequences, a common feature of transcription control elements. By supershift analysis it was confirmed that neither the glucocorticoid receptor nor the krupple-like factor EKLF or BKLF binds to this element (data not shown). Other candidate proteins contained some sequence conservation, including the Ets, Z-box, and AP-3 families of TFs. The factors which bind RE-1019 mediated significant activation of the HOX11 promoter (Fig. 9). It appears that the functional activity of this element is contained only within the fragment RE-1059 G/S 3'B, while EMSA analysis suggested that both RE-1059 G/S 3'B and C contained the

DNA sequence responsible for the formation of complex 4. Interestingly, the mutant competition assays (Fig. 8) revealed that the mutated fragment Mut A (RE-1059 G/S 3' Mut A) required an increased concentration of unlabelled oligonucleotide to compete for the binding of complex 4. We determined if the cloning of RE-1059 G/S 3'B or C led to the fortuitous recapitulation of flanking DNA sequences endogenous to HOX11, however, this was not the case. Therefore it appears likely that the essential bases for the formation of a functional complex 4 involve the 6 bases mutated in RE-1069 G/S 3' Mut B and the flanking bases on either side mutated in Mut A and Mut C, respectively. To identify the factor(s) responsible for this activation either affinity chromatography could be used or expression libraries could be probed with the consensus sequence of RE-1019.

The role of *HOX11* in leukaemogenesis remains to be defined. It has been well documented that deregulation of HOX11 appears to be strictly limited to the T-ALL phenotype and that normal T-cells are devoid of HOX11 expression [11,25,39]. Therefore there are several key questions which remain to be answered regarding HOX11 and T-ALLs. Is the deregulation of HOX11 in T-cells sufficient to produce overt leukaemia? Moreover, why is ectopic HOX11 expression specifically associated with T-cell phenotype? There have been several investigations undertaken to answer these questions: Hawley and colleagues [40] demonstrated the transforming potential of HOX11 by transducing murine bone marrow cells with recombinant retrovirus MSCV-HOX11. Very few of the transplant recipient mice developed malignancies and the tumours had a long latency period of up to 12 months [41]. This finding suggests that secondary events were required for the full oncogenic phenotype to develop. Interestingly, it has been noted that overexpression of both HOX11 and PBX2 leads to a reduction of contact-dependent growth inhibition when introduced into NIH3T3 cells [32]. This observation supports the view that TALE homeoproteins may have a role in HOX11-mediated leukaemogenesis. Furthermore, there is evidence to suggest that PBX1 is also required for normal spleen development. DiMartino and colleagues [43] report that PBX1 (-/-) embryos while being embryonically lethal show spleen agenesis, which is consistent with a HOX11 (-/-) phenotype [42]. This group has also generated double heterozygote mice (HOX11 and PBX1) and confirmed their role in normal spleen development.

Several issues need further examination, including identification of the postulated additional proteins involved in the PBX2-containing complex. Second, to determine if *PRE-1048* can be used by other binding proteins to disrupt the normal regulation of *HOX11* and induce leukaemogenesis. Lastly, to ascertain if abnormal expression of PBX2 in leukaemic cells can be specifically

associated with HOX11 expression in T-ALL patient samples. Or alternatively, if down-regulation of PBX2 by mutation disrupts the normal expression levels of HOX11. This report focuses on the significant activation of the HOX11 promoter mediated by a 60-bp region of DNA located immediately upstream of the first 1 kb of DNA. We propose that this region of DNA could play a significant role in leukaemogenesis by increasing the level of *HOX11* protein in the event of gene deregulation. Furthermore, this report strengthens the case made previously for TALE homeoproteins having a role in HOX11-mediated leukaemogenesis. Answering these questions will ultimately aid in our understanding of both the developmental regulation of HOX11 and the molecular mechanism by which HOX11-specific leukaemogenesis may occur. While this study provides the first evidence of a HOX gene specifically regulated by PBX2, it is likely that there are numerous such target genes. As these genes are identified, our knowledge of the complex regulation mediated by the family of PBX proteins will be expanded.

Acknowledgments

We thank Dr. Katrin Hoffman and Mr. Stewart Cattach for the technical assistance and Dr. Wayne Greene and Dr. Peter Dallas for helpful criticism of the manuscript. This study was supported by NHMRC of Australia, Child Health Research Foundation of WA, Children's Leukaemia and Cancer Research Foundation, and Crown Scientific, Australia.

References

- F. Apiou, D. Flagiello, C. Cillo, B. Malfoy, M.F. Poupon, B. Dutrillaux, Cytogenet. Cell Genet. 73 (1996) 114–115.
- [2] T.N. Dear, I. Sanchez-Garcia, T.H. Rabbitts, R.A. DePinho, N. Schreiber-Agus, F.W. Alt, Proc. Natl. Acad. Sci. USA 90 (1993) 4431–4435.
- [3] T.N. Dear, T.H. Rabbitts, Gene 141 (1994).
- [4] K.D. Patterson, P.A. Krieg, Dev. Dyn. 214 (1999) 34-43.
- [5] B. Kanzler, T.N. Dear, Dev. Biol. 234 (2001) 231-243.
- [6] T.N. Dear, W.H. Colledge, M.B. Carlton, I. Lavenir, T. Larson, A.J. Smith, A.J. Warren, M. Evans, M. Sofroniew, T.H. Rabbitts, Development 121 (1995) 2909–2915.
- [7] C.W.M. Roberts, J.R. Shutter, S.J. Korsmeyer, Nature 368 (1994) 747–749.
- [8] I.E. Dubé, S. Kamel-Reid, C.C. Yuan, M. Lu, X. Wu, G. Corpus, S. Raimondi, W.M. Crist, A.J. Carroll, J. Minowada, J.B. Baker, Blood 78 (1991) 2996–3003.
- [9] M.A. Kennedy, S.R. Gonzalez, U.R. Kees, F. Lampert, N. Dear, T. Boehm, T.H. Rabbitts, Proc. Natl. Acad. Sci. USA 88 (1991) 8900–8904.
- [10] M. Zutter, R.D. Hockett, C.W. Roberts, E.A. McGuire, J. Bloomstone, C.C. Morton, L.L. Deaven, W.M. Crist, A.J. Carroll, S.J. Korsmeyer, Proc. Natl. Acad. Sci. USA 87 (1990) 3161–3165.
- [11] P.D. Salvati, P.R. Ranford, J. Ford, U.R. Kees, Oncogene 11 (1995) 133–138.

- [12] A.A. Ferrando, D.S. Neuberg, J. Staunton, M.L. Loh, C. Huard, S.C. Raimondi, F.G. Behm, C.-H. Pui, J.R. Downing, D.G. Gilliland, E.S. Lander, T.R. Golub, A.T. Look, Cancer Cell 1 (2002) 75.
- [13] K. Monica, N. Galili, J. Nourse, D. Saltman, M. Cleary, Mol. Cell. Biol. 11 (1991) 6149–6157.
- [14] K. Wagner, A. Mincheva, B. Korn, P. Lichter, H. Popperl, Mech. Dev. 103 (2001) 127–131.
- [15] H. Popperl, H. Rikhof, H. Chang, P. Haffter, C. Kimmel, C. Moens, Mol. Cell. 6 (2000) 255–267.
- [16] N. Vlachakis, D.R. Ellstrom, C.G. Sagerstrom, Dev. Dyn. 217 (2000) 109–119.
- [17] T.R. Burglin, G. Ruvkun, Nat. Genet. 1 (1992) 319-320.
- [18] R.S. Mann, S.K. Chan, TIG 12 (1996) 258-262.
- [19] J. Berthelsen, V. Zappavigna, F. Mavilio, F. Blasi, EMBO J. 17 (1998A) 1423–1433.
- [20] Q. Lu, M.P. Kamps, Mol. Cell. Biol. 16 (1996) 1632–1640.
- [21] M.P. Kamps, C. Murra, X. Sun, D. Baltimore, Cell 60 (1990) 547– 555.
- [22] J. Nourse, J.D. Mellentin, N. Galili, J. Wilkinson, E. Stanbridge, S.D. Smith, M.L. Cleary, Cell 60 (1990) 535–545.
- [23] R.L. Brake, U.R. Kees, P.M. Watt, Oncogene 17 (1998) 1787-1796.
- [24] C. Lozzio, B. Lozzio, Blood 45 (1975) 321-334.
- [25] M. Hatano, W.M. Roberts, M. Minden, W.M. Crist, S.J. Korsmeyer, Science 253 (1991) 79–82.
- [26] U.R. Kees, R. Lukeis, J. Ford, O.M. Garson, Blood 74 (1989) 369–373.
- [27] P. Martin, T. Papayannopoulou, Science 216 (1982) 1233-1235.
- [28] Q. Lu, D.D. Wright, M.P. Kamps, Mol. Cell. Biol. 14 (1994) 3938–3948.
- [29] E. Ferretti, H. Marshall, H. Popperl, M. Maconochie, R. Krumlauf, F. Blasi, Development 127 (2000) 155–166.

- [30] H. Popperl, M. Bienz, M. Studer, S.K. Chan, S. Aparicio, S. Brenner, R.S. Mann, R. Krumlauf, Cell 81 (1995) 1031–1042.
- [31] C.A. Schnabel, Y. Jacobs, M.L. Cleary, Oncogene 19 (2000) 608–616
- [32] T.D. Allen, Y.X. Zhu, T.S. Hawley, R.G. Hawley, Leuk. Lymphoma 39 (2000) 241–256.
- [33] B. Peers, T. Sharma, M. Johnson, M. Kamps, M. Montminy, Mol. Cell. Biol. 15 (1995) 3385–3393.
- [34] L.J. Bischof, N. Kagawa, J.J. Moskow, Y. Takahashi, A. Iwamatsu, A.M. Buchberg, M.R. Waterman, J. Biol. Chem. 273 (1998) 7941–7948.
- [35] G.H. Swift, Y. Liu, S.D. Rose, L.J. Bischof, S. Steelman, A.M. Buchberg, C.V.E. Wright, R.J. MacDonald, Mol. Cell. Biol. 18 (1998) 5109–5120.
- [36] J. Berthelsen, V. Zappavigna, E. Ferretti, F. Mavilio, F. Blasi, EMBO J. 17 (1998B) 1434–1445.
- [37] D. Penkov, S. Tanaka, G. Rocco, J. Berthelsen, F. Blasi, F. Rameriez, J. Biol. Chem. 275 (2000) 16681–16689.
- [38] P.S. Knoepfler, M.P. Kamps, Mol. Cell. Biol. 15 (1995) 5811– 5819.
- [39] H. Yamamoto, M. Hatano, Y. Iitsuka, N.S. Mahyar, M. Yamamoto, T. Tokuhisa, Mol. Immunol. 32 (1995) 1177–1182.
- [40] R.G. Hawley, A.Z.C. Fong, M. Lu, T.S. Hawley, Oncogene 9 (1994) 1–12.
- [41] R.G. Hawley, A.Z.C. Fong, M.D. Reis, N. Zhang, M. Lu, T.S. Hawley, Cancer Res. 57 (1997) 337–345.
- [42] J.F. DiMartino, L. Selleri, M. Firpo, K. Smith, J. Rhee, J. Van Deursen, G. Grosveld, M.L. Cleary, in: American Society of Haematology 40th Annual Meeting Pharmacia and Upjohn, Miami Beach, FL, 1998.
- [43] L. Selleri, M.L. Cleary, in: American Society of Haematology, 41st Annual Meeting, 1999.