Chromosomal Translocations Involving the E2A Gene in Acute Lymphoblastic Leukemia: Clinical Features and Molecular Pathogenesis

By Stephen P. Hunger

Over the past decade, significant advances have been made in understanding the molecular details of lymphopoiesis and the pathogenesis of lymphoid neoplasia. Major contributions to these advances have been made through study of Ig/T-cell receptor (TCR) genes and the factors that control their ordered recombination and by investigation of nonrandom chromosomal translocations that occur in acute lymphoblastic leukemia (ALL). In addition, the advent of technology that allows investigators to genetically inactivate or "knockout" individual genes by creating homozygous null mice has facilitated efforts to identify proteins whose function is essential for normal development. These disparate lines of investigation have shown that the protein products of the E2A gene are essential for normal lymphopoiesis and can be converted into oncogenic transcription factors by chromosomal translocations. In this review, I will summarize briefly our current understanding of the function of E2A proteins within the lymphoid system, particularly genetic evidence showing that they are absolutely required for normal B-cell development. Attention will then be turned to discussing two translocations, t(1;19)(q23;p13) and t(17;19)(q21-22;p13), that disrupt E2A. The clinical features of t(1;19)+ and t(17;19)+ ALLs will be reviewed, particularly those of the former that illustrate the clinical importance of molecular diagnosis. I will also discuss, in detail, the structural features and functional properties of E2A-PBX1 and E2A-HLF, the fusion proteins created by the t(1;19) and t(17;19), respectively, and will review the experimental evidence implicating these chimeras in leukemogenesis.

E2A Proteins: Identification and Function

E2A encodes several transcription factors via alternative splicing of precursor mRNA. Rearrangement of the Ig and TCR genes and expression of their protein products on the surface of lymphocytes plays a critical role in lymphopoiesis. The contribution of various regulatory elements present in the enhancers of these genes to transcription of the re-arranged alleles has been extensively investigated. Ig and TCR gene enhancers can drive transcription from a wide variety of promoters in cells of lymphoid origin but not in other cell types. By mutation analysis, several short DNA sequences have been defined that are critical to this process. One class of these regulatory elements, referred to as E-boxes (after Ephrussi), were initially identified as protein binding sites by in vivo footprinting studies. One such E-box site, αE2, is important for transcription of the Ig α light chain gene. Murre et al used this site to screen a B-cell-derived cDNA expression library and identified a clone encoding a protein designated E12 that bound specifically to a αE2 probe, but not to related E-box sites.

E12 and E47, encoded by a cDNA identified because of its homology to E12, are protein products of a single gene termed E2A; they differ only in a region called the basic helix-loop-helix (bHLH) domain, and this difference results from alternative usage of E12- and E47-specific bHLH-encoding exons. One additional E2A gene product, E2-5, which differs from E47 only at the extreme amino terminus, has been identified. The bHLH motif is an ~60 amino acid domain present in a number of proteins including products of the Drosophila gene daughterless and myogenic proteins of the MyoD family. Structural modeling predicted that the HLH domain consisted of two amphipathic helices separated by a loop region of variable length (HLH), preceded by a region of basic amino acids (bHLH). More detailed studies have shown that these domains were essential for protein dimerization and DNA binding. More detailed studies have shown that the HLH motif mediates dimerization of identical or related monomers and that protein dimers, but not monomers, can bind DNA in a sequence-specific manner. DNA-binding is mediated by the region of basic amino acids located at the amino terminus of the first helix. The preferred DNA-binding sequence of a bHLH:bHLH dimer is determined by the identity of its component monomers. Recent x-ray crystallographic studies have confirmed and extended these predictions and shed additional light on the structural factors that determine whether homodimerization or heterodimerization is energetically favored.

E12, E47, E2-5 (which will subsequently be referred to collectively as E2A proteins or simply E2A), and other bHLH proteins function as transcription factors that bind to specific regulatory elements in target-gene promoter or enhancer regions and interact (directly or indirectly) with protein components of the basal transcription machinery to control the rate of transcription by RNA polymerase II. Transcription factors contain specific modular protein domains that function in transcriptional activation or repression. E2A functions as a transcriptional activator and this function was initially mapped to the amino-terminal two-thirds of the protein. More recently, two discrete transcriptional activation domains (TADs), named activation domains (AD) I and II (the latter is also referred to as the loop helix, but is distinct from the similarly named HLH domain), have been identified within this portion of E2A proteins.

E2A proteins are widely expressed and have been implicated in regulation of a diverse array of cellular processes including myogenic differentiation and transcription of Ig and pancreatic genes. In most cases, E2A heterodimer-
izes with tissue-specific bHLH proteins to regulate expression of specific target genes.\textsuperscript{2,20} A family of HLH proteins that lack a basic region, the Id proteins, interferes with transcriptional activation by heterodimerizing with E2A and other bHLH proteins and sequestering them in complexes that are incapable of binding DNA.\textsuperscript{22-25} Despite widespread expression of E2A mRNA and protein, DNA-binding E2A homodimers appear to be present only in B-lineage lymphocytes.\textsuperscript{26,28} Recent evidence indicates that E2A homodimers are stabilized by an intermolecular disulfide bond, that this association occurs only in B cells, and that it can be disrupted by an activity present in extracts of non-B cells.\textsuperscript{27}

**Genetic or functional inactivation of E2A abrogates B-cell development in the mouse.** Several recent reports have highlighted the essential role of E2A proteins in B-cell development. Two groups eliminated E2A function genetically by creating knockout mice in which either the E12-specific bHLH exon\textsuperscript{29} or both the E12- and E47-specific bHLH exons\textsuperscript{30} were disrupted by gene targeting. Heterozygotes were bred to generate homozygous mutant mice that did not produce any significant amount of full-length E2A mRNA or protein and will be considered, for the purposes of this discussion, to be E2A null. Sun\textsuperscript{30} used a B-cell–specific promoter to constitutively overexpress the dominant-negative Id1 gene in the developing B cells of transgenic mice. Of course, the latter experimental approach is not discretely targeted at E2A, but rather should interfere with the function of any bHLH proteins present in developing B cells that dimerize with Id1. Despite these different experimental approaches, the phenotype was the same in each case. Mutant mice developed normally to term, but did not form mature B cells, whereas maturation of other hematopoietic lineages appeared intact. B-cell progenitors are formed in the mutant mice, but maturation is arrested at an early stage. The initial event in IgH gene rearrangement, formation of a D\textsubscript{JH} Heavy joint, does not occur and expression of the B-lineage–specific marker B220 is not observed.\textsuperscript{28-30} The mechanisms that underlie this gross perturbation of B lymphopoiesis remain to be defined.

**Proposed mechanisms for E2A function.** Because E2A behaves as a transcriptional activator in experimental assays, it is reasonable to hypothesize that normal B-cell development requires that E2A homodimers or heterodimers induce transcription of one or more critical downstream target genes. The phenotype of E2A null mice indicates that other bHLH proteins present in B-cell precursors cannot compensate for the absence of E2A, which suggests at least two possibilities: either E2A homodimers or heterodimers of E2A with some other bHLH protein are essential for B-cell development. Whereas DNA-binding E2A homodimers have been detected only in B cells, no B-cell–specific E2A heterodimerization partner has been identified,\textsuperscript{26,28} suggesting that the crucial complex is an E2A homodimer.

In addition to their important role in transcriptional regulation, E2A proteins have been shown to have growth-inhibitory properties and, when overexpressed in fibroblasts, can block cell cycle progression near the G\textsubscript{1}-S border.\textsuperscript{31} The observation that this growth-inhibitory function is not completely dependent on the integrity of the bHLH domain\textsuperscript{31} raises the intriguing possibility that E2A has functions that are separable from its role in transcriptional regulation. Thus, critical functions of E2A proteins in B-cell ontogeny might not be limited to transcriptional regulation. Further studies directed at identifying genes induced by E2A and better defining the role of E2A proteins in cell cycle progression should provide additional insights into the regulation of lymphopoiesis.

**TWO DIFFERENT E2A FUSION PROTEINS ARE CREATED BY CHROMOSOMAL TRANSLocations in ALL**

The study of chromosomal translocations in human cancer has led to identification of a large number of proto-oncogenes implicated in malignant transformation.\textsuperscript{32} Molecular studies have shown that translocations lead to oncogenic conversion of cellular proto-oncogenes by one of two general mechanisms. The first is juxtaposition with one of the transcriptionally active Ig or TCR gene loci resulting in dysregulated expression of a structurally intact protein whose cellular presence is normally under tight homeostatic control. In other cases, translocations create a fusion gene that produces a fusion mRNA that encodes a chimeric protein with novel structural and functional properties not possessed by either of the constituent wild-type proteins.

Shortly after the E2A gene was discovered, it was found to be located within chromosome band 19p13.3 at the breakpoint of the t(1;19)(q23;p13) that is observed in childhood ALL.\textsuperscript{33} Further studies showed that this translocation fused E2A to a previously unidentified chromosome 1 gene originally named PRL and subsequently named PBX1, creating an E2A-PBX1 fusion gene on the der(19) chromosome.\textsuperscript{34,35} Analysis of a large number of cases showed that identical E2A-PBX1 fusion mRNAs were present in greater than 95% of t(1;19)\textsuperscript{+} ALLs.\textsuperscript{36} Subsequently, a second ALL translocation, the t(17;19)(q21-22;p13), was found also to interrupt E2A, in this case fusing it to a gene named HLF.\textsuperscript{37,38} Each of these translocations creates a fusion protein with structural and functional properties of a chimeric transcription factor in which the E2A bHLH domains are replaced with functionally distinct DNA binding and protein dimerization domains from heterologous proteins. Substantial progress has been made in understanding the mechanisms by which these chimeras contribute to leukemogenesis.

\textit{t(1;19)(q23;p13)}

**Clinical features.** The t(1;19) was recognized to be a recurring translocation in childhood ALL in the early 1980s.\textsuperscript{39,42} Subsequent studies have shown it to be present in about 5% of ALLs, including 20% to 25% of pre-B cases (leukemias that express cytoplasmic Ig [cIg] but not surface Ig), making it the most common chromosomal rearrangement in children with leukemia (Table 1).\textsuperscript{43,44} It is also found in a small percentage (<1%) of early pre-B (cIg+, sIg+) cases and in some ALLs with a transitional pre-B (cIg+, sIg+) phenotype.\textsuperscript{45-47} In addition, this translocation is detected in adults with ALL and occasionally in acute myelogenous leukemia (AML), T-cell ALL, and lymphomas, but the prevalence has not been well defined in these clinical set-
tions. Two different cytogenetic varieties of the t(1;19) are observed. The balanced t(1;19) is a simple reciprocal translocation that does not result in any net loss or gain of genetic material. The more frequently observed unbalanced type, 19q-+der(19)t(1;19), presumably arises by nondisjunction leading to loss of the der(1) and replacement with a second copy of the remaining normal chromosome 1 resulting in trisomy of chromosome 1q telomeric to PBX1 and monosomy of chromosome 19 telomeric to E2A. The unbalanced form can arise during clonal evolution, because the two types are observed simultaneously or metachronously in some patients.

Prognostic import. After recognition of the t(1;19) as a distinct clinicopathologic entity, it was noted that patients with this translocation appeared to have a high risk of relapse. The t(1;19) is also associated with known high-risk clinical features, including an elevated leukocyte count at diagnosis, non-white race, and central nervous system leukemia. In addition, several groups found that children with pre-B ALL fared worse than those with an early pre-B phenotype. For these reasons, it was not certain whether the t(1;19) itself was an independent adverse prognostic factor or if it was simply associated with other high-risk features. In a prospective trial performed in the early to late 1980s, the Pediatric Oncology Group (POG) found that children with pre-B ALL fared worse than those with early pre-B ALL; this difference appeared to be restricted to those pre-B patients with chromosomal translocations, particularly the t(1;19). The adverse outcome of t(1;19)+ patients remained significant even after adjustment for recognized adverse clinical features, indicating that it was indeed an independent risk factor. The St Jude group also observed that patients with the t(1;19) and other translocations enrolled in the early 1980s Total Therapy X trial had an inferior outcome, but found that this difference was no longer present for patients treated more intensively in the mid-late 80s on Total Therapy XI, consistent with observations that adverse prognostic indicators can be overcome with more effective therapy. These results await confirmation in larger trials currently underway, but emphasize the importance of identifying high-risk biologic features such as the t(1;19).

As discussed above, the t(1;19) occurs in both a balanced and an unbalanced form, raising the question of whether these differences might have prognostic import. One retrospective study that analyzed a heterogeneous group of patients from several different centers found that those with the unbalanced der(19)t(1;19) had a significantly better treatment outcome than patients with the balanced t(1;19). In contrast, the St Jude group analyzed a more homogeneous group of pre-B ALL patients treated on four consecutive clinical trials and found that the outcome of patients with a der(19)t(1;19) was not significantly different from those with a balanced t(1;19). These latter results are consistent with the fact that there is no difference in the E2A-PBX1 fusion proteins generated by balanced and unbalanced t(1;19). However, a trend toward improved outcome for those with unbalanced translocations in the St Jude study suggests that this question merits further investigation in a larger cohort of patients.

As described above, the t(1;19) occurs in both a balanced and an unbalanced form, raising the question of whether these differences might have prognostic import. One retrospective study that analyzed a heterogeneous group of patients from several different centers found that those with the unbalanced der(19)t(1;19) had a significantly better treatment outcome than patients with the balanced t(1;19). In contrast, the St Jude group analyzed a more homogeneous group of pre-B ALL patients treated on four consecutive clinical trials and found that the outcome of patients with a der(19)t(1;19) was not significantly different from those with a balanced t(1;19). These latter results are consistent with the fact that there is no difference in the E2A-PBX1 fusion proteins generated by balanced and unbalanced t(1;19). However, a trend toward improved outcome for those with unbalanced translocations in the St Jude study suggests that this question merits further investigation in a larger cohort of patients.

Table 1. Clinical Features of t(1;19)(q23;p13)+ ALL

<table>
<thead>
<tr>
<th>Frequency</th>
<th>~5% of childhood ALLs. Also occurs in adult ALLs, but frequency not well defined.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Phenotype</td>
<td>&gt;90% pre-B (clg+, slg-). Also observed in early pre-B (clg+, slg+) and transitional pre-B (clg+, slg+).</td>
</tr>
<tr>
<td>Cell surface markers</td>
<td>&gt;90% CD9+, CD19+, CD20+, CD21*, CD22*, CD34*.</td>
</tr>
<tr>
<td>Associated clinical features at presentation</td>
<td>Elevated leukocyte count, non-white race, and central nervous system leukemia.</td>
</tr>
<tr>
<td>Outcome</td>
<td>Unfavorable prognosis with antimitabolite-based therapy.</td>
</tr>
<tr>
<td>Molecular abnormality</td>
<td>&gt;65% have E2A-PBX1 fusion. Rare cases lack E2A-PBX1 fusion (usually early pre-B, CD34+).</td>
</tr>
</tbody>
</table>
sion on t(1;19) ALLs. They found that 20 of 21 cases shared an identical, complex immunophenotype defined by homogeneous high-level expression of CD19, CD10, and CD9; complete absence of CD34; and at least partial absence of CD20. This immunophenotype was present in only 8% of childhood B-precursor ALLs, more than half of which had cytogenetic and/or molecular evidence of a t(1;19). The sole t(1;19) that did not conform to this phenotype also lacked E2A gene rearrangements, E2A-PBX1 fusion mRNAs, and was hyperdiploid. The St Jude group has confirmed the association of this phenotype (and extended it to also include CD22+/CD21−) with E2A-PBX1+ t(1;19)s and found that its presence or absence correlates closely with cIg expression.

The preponderance of evidence suggests that a large majority of t(1;19)+ ALLs are cIg+, express a characteristic profile of cell surface antigens, and contain E2A-PBX1 fusion, whereas a small minority of cases, despite appearing cytogenetically identical, are cIg−, do not conform to the phenotype described above, and lack E2A or PBX1 abnormalities. Although data are limited due to infrequent occurrence, these variant t(1;19)s appear to have a good prognosis when treated with less intensive chemotherapy. Accumulating evidence indicates that, despite this tight linkage with cIg status, some t(1;19)+ ALLs that express typical E2A-PBX1 fusion mRNAs are cIg−. In addition, a small percentage of E2A-PBX1+ cases have an L3, Burkitt’s-like morphology and/or express slg. As mentioned earlier, cytogenetically indistinguishable t(1;19)s have also been observed rarely in T-ALLs, AMLs, lymphomas, and nonhematopoietic tumors. E2A gene rearrangements and/or E2A-PBX1 fusion mRNAs have been detected in 2 cases of AML, arguing that this molecular abnormality is not confined solely to lymphoid leukemias. This phenotypic heterogeneity among E2A-PBX1+ cases, although infrequent, is not particularly surprising for, as is discussed below, E2A-PBX1 is transforming when expressed in a variety of cell types of both hematopoietic and nonhematopoietic origin. Thus, based on currently available evidence, it is probably more useful to analyze t(1;19)+ ALLs for the presence of E2A-PBX1 fusion rather than cIg expression, because the former is more reliably determined, is biologically relevant, and appears to correlate best with prognosis.

Molecular features of the t(1;19) and their implications for molecular diagnosis. In the vast majority of t(1;19)+ ALLs, E2A is fused to PBX1, leading to production of at least two species of E2A-PBX1 chimeric proteins that differ at their extreme carboxy termini due to differential mRNA splicing (Fig 1). The biologic relevance of this alternative splicing is currently unknown. Breakpoints in the E2A gene almost invariably occur in the 3.5 kb intron that lies between exons 13 and 14. The genomic organization of PBX1 has not been fully determined, but it is assumed that the breakpoints occur in a single intron as the identical portion of PBX1 is joined consistently to E2A in fusion mRNAs. One variant t(1;19)+ ALL has been described in which the site of E2A-PBX1 fusion differs from other reported cases; this case provides information about portions of PBX1 that are dispensable for transformation. Several groups have also detected t(1;19)+ ALLs that contain 27 additional nucleotides at the junction between E2A and PBX1 coding sequences in fusion mRNAs. These additional nucleotides, which are identical in each case, are found in ~5% to 10% of t(1;19)+ ALLs and appear to arise from a differentially spliced exon of either E2A or PBX1, but their exact derivation is unknown.

The consistent nature of E2A-PBX1 rearrangements has important implications for molecular diagnosis. Using E2A cDNA probes, gene rearrangements can be detected in almost all E2A-PBX1+ t(1;19) ALLs (the most informative restriction enzymes are Bgl II and Xba I). In contrast, probing standard Southern blots with PBX1 probes will not consistently show gene rearrangements because the breakpoints are dispersed within an intron that is at least 50 kb in size. Reverse transcriptase-polymerase chain reaction (RT-PCR) can be performed on RNA isolated from leukemic cells and will detect a fusion product in greater than 95% of patients with t(1;19)+ ALLs. This technique can also detect subclinical levels of minimal residual disease in patients in remission and may prove useful for monitoring response to therapy. Molecular analysis frequently detects translocations that are not observed karyotypically; in fact, several studies suggest that 25% to 50% of leukemias with E2A-PBX1 fusion are missed by standard cytogenetics. In light of observations that E2A-PBX1—expressing t(1;19)s fare poorly with less aggressive antimetabolite based therapy, but respond relatively well to more intensive therapy, a strong justification exists for performing molecular
E2A TRANSLOCATIONS IN ALL

The PBX proteins are the mammalian homologues of the fly protein extradenticle (exd)\(^9\)\(^9\) and the product of the *C. elegans* ceh-20 gene\(^6\)\(^6\) and are distantly related to the yeast proteins MATA1 and MATA2.\(^9\) In contrast to most fly homeodomain-containing proteins and their vertebrate counterparts, exd and PBX proteins share significant homology within the entire protein, suggesting that there may be conservation of their functional properties as well.\(^9\) Mutations in *exd*, like mutations in HOM genes, cause homeotic transformations; in other words, one body segment is transformed to resemble that of another.\(^9\)\(^9\) Genetic evidence in *Drosophila* suggests that *exd* may function as a cofactor necessary for selector gene function.\(^9\) Recent data support this theory. *exd* interacts directly with a subset of HOM proteins and this heterocomplex binds cooperatively to DNA.\(^9\)\(^9\) These observations partially explain a long-standing enigma in developmental biology. Given that most HOM proteins bind similar DNA sequences in vitro with comparable affinities, how is biologic specificity achieved in vivo?\(^9\)\(^9\) The new data suggest that biologic specificity may be achieved in part through cooperative DNA-binding by selector and nonselector homeodomain-containing proteins, thereby determining the spectrum of target sites recognized and/or the regulatory effects of the Hox complex on subordinate genes.

In light of the remarkable sequence conservation between *exd* and PBX proteins, it is likely that the latter may serve a similar function as cofactors that interact with mammalian Hox proteins. Supporting this hypothesis, Chang et al\(^6\) found that PBX proteins (and E2A-PBX1) can bind cooperatively with certain Hox proteins to a DNA probe that contains both PBX and Hox recognition sequences. This interaction was dependent on the integrity of discrete domains in each protein. For PBX proteins, this consisted of the homeodomain and 25 carboxy-terminal amino acids, whereas Hox proteins required an evolutionarily conserved hexapeptide motif that is located immediately N-terminal to the homeodomain in a subset of Hox (and homologous HOM) proteins. These observations suggest that E2A-PBX1 may not regulate transcription of subordinate target genes by itself, but rather might function in a heterocomplex with specific Hox proteins.

**E2A-PBX1 functions experimentally as a potent oncogene.** Experimentally, E2A-PBX1 is capable of transforming cells of several different histologic origins; however, a model that recapitulates pre-B ALL as seen in children with the t(1;19) has not yet been identified. E2A-PBX1 induces transformation and loss of contact inhibition in NIH 3T3 mouse fibroblasts,\(^10\)\(^10\) causes AML when lethally irradiated mice are reconstituted with marrow progenitors infected with retroviruses that express E2A-PBX1,\(^10\)\(^6\) and causes T-cell lymphomas after expression in transgenic mice.\(^10\) Several important insights into E2A-PBX1 function have been gained from these studies.

Dedera et al\(^10\) expressed E2A-PBX1 as a transgene under the control of regulatory elements derived from the IgH promoter/enhancer, a strategy designed to direct protein expression to lymphoid cells, particularly those of B lineage. All transgenic mice died of malignant lymphomas by 5 months of age. However, unexpectedly, the tumors were
not of B lineage but displayed a phenotype (CD4+/CD8+/CD3-/+), consistent with transitional intermediate thymocytes, presumably related to unanticipated higher level expression of the transgene in T-cell, as compared with B-cell, progenitors. Animals evaluated before tumor development exhibited an interesting phenotype with reduced numbers of lymphoid cells. Paradoxically, this reduction occurred despite the presence of an increased number of thymocytes that were actively traversing the cell cycle and was explained by a concomitant increase in cells undergoing programmed cell death or apoptosis.

The mechanisms of, and structural requirements for, transformation have been analyzed in several systems. Kamps and Baltimore observed that lethally irradiated mice reconstituted with marrow infected with E2A-PBX1–encoding retroviruses developed AML, and that in some cases malignant cells exhibited factor-dependent growth in culture, whereas others could survive in the absence of exogenous growth factors. Based on these results, they hypothesized that a primary effect of E2A-PBX1 might be to block differentiation of hematopoietic cells. This postulate was investigated by infecting marrow progenitors with E2A-PBX1–encoding retroviruses and assessing their ability to grow in medium that contained specific myeloid growth factors. They found that progenitors infected with E2A-PBX1 were able to proliferate for long periods of time in culture without differentiating into mature myeloid cells as long as exogenous granulocyte-macrophage colony-stimulating factor (GM-CSF) was provided, but the cells died rapidly in the absence of this cytokine. A mechanism involving blocked differentiation is consistent with the clinical observation that in most cases of ALL associated with expression of E2A-PBX1, leukemic cells are arrested at a discrete stage of lymphoid maturation and with the experimentally observed differentiation blockade in E2A-PBX1 transgenic mice. Furthermore, although the mechanism of death after growth factor withdrawal was not directly assessed by Kamps and Wright, one would anticipate that it would be apoptotic death, similar to that observed in the “premalignant” stage of tumorigenesis in transgenic mice. Taken together, these observations suggest that a major primary effect of E2A-PBX1 is to arrest differentiation at a discrete stage, but that further genetic events are necessary for cells to exhibit a fully transformed phenotype and overcome growth factor dependence and/or escape from apoptotic signals. In future studies, investigators are expected to focus on defining the nature of these cooperating genetic events.

The 3T3 transformation system has been used to perform a detailed structure-function study to define which portions of E2A-PBX1 are necessary for oncogenesis. Retrovirus-mediated overexpression of PBX1, 2, or 3 did not lead to transformation indicating that, unlike some of the HOX proteins, wild-type PBX proteins are not oncogenic. E2A and/or PBX proteins do not have latent transforming properties that can be exposed by removal of specific protein domains, because a transformed phenotype was not observed when truncated E2A or PBX proteins that included only those portions present in the (1;19)-derived chimera were overexpressed. Rather, oncogenesis requires unique domains from each constituent of the chimera, because E2A-PBX1 readily induced transformation in this system. The essential contribution of E2A mapped to its two TADs; when either of these was removed transformation was substantially decreased or abolished. Chimeric constructs in which E2A residues were fused to PBX2 or PBX3 transformed as effectively as did E2A-PBX1, indicating that the essential PBX contribution is present in each of the three family members. The minimum portion of PBX1 required for oncogenesis in this study was the homeodomain and 25 amino acids immediately C-terminal, the same portion necessary for interaction with select HOX proteins. Surprisingly, a construct that encoded an E2A-PBX1 chimeric protein from which the homeodomain was discretely deleted (while retaining flanking residues essential for HOX protein interactions) also transformed 3T3 cells and behaved indistinguishably from the full-length chimera in the transgenic mouse system.

Potential models for transformation induced by E2A-PBX1. These data, coupled with observations on the transcriptional regulatory properties of PBX and E2A-PBX1 proteins, suggest several potential models to explain the leukemogenic properties of E2A-PBX1 (Fig 2). The currently available evidence suggests that PBX proteins normally bind DNA cooperatively with select HOX proteins, and that this heterocomplex then regulates transcription of subordinate target genes in a positive or negative manner. If we assume that the PBX/HOX complex normally activates target gene transcription, fusion to E2A might bypass regulatory processes that control the presence and/or activity of PBX and HOX proteins, and E2A-PBX1 could directly activate target gene transcription by itself. In a second possible scenario, E2A-PBX1 might still bind DNA cooperatively with a specific HOX protein, but the potent transactivation domains of E2A would enable this heterocomplex to directly activate target gene transcription without the need for other events required by PBX/HOX complexes. Alternatively, PBX/HOX complexes might normally inhibit transcription and E2A-PBX1/HOX complexes would induce transcription of genes normally repressed in lymphoid cells.

In the context of these models, the experimental evidence that the PBX1 homeodomain is not required by E2A-PBX1 for transformation is explained by the fact that the homeodomain-deleted chimera retains the residues necessary for interaction with specific (and so far undefined) HOX proteins. Presumably, the relevant HOX protein binds DNA and, via protein-protein interaction, “tethers” E2A-PBX1 to the promoter allowing crucial E2A effector domains to activate target gene transcription. These assumptions predict that an E2A-PBX1 chimera that lacks both the homeodomain and flanking residues necessary for HOX protein interaction would not be transformation competent, a prediction that can be tested experimentally.

In considering these models, one must remember that the functional differences (if any) between the highly similar PBX proteins are unknown. It is not clear whether their functions are completely redundant or potentially unique to each family member. The fact that E2A-PBX2 and E2A-PBX3 chimeric proteins transform NIH 3T3 cells as efficiently as E2A-PBX1 indicates that any differences that
might exist do not play an important role in transformation. These observations also suggest that the redirected lymphoid expression of portions of PBX1 that occurs consequent to the t(1;19) is not a crucial leukemogenic event. It is important to emphasize that protein-DNA and protein-protein interactions that occur in t(1;19) lymphoblasts are potentially quite complex and that E2A-PBX1 might have to compete with PBX2 and/or PBX3 for DNA binding sites and/or protein dimerization partners. Further progress in understanding the cascade of events initiated by E2A-PBX1 will require identification of the relevant HOX proteins and leukemogenic target genes.

t(17;19)(q21-22; p13)

Clinical features. The t(17;19)(q21-q22; p13) is a recently recognized nonrandom translocation present in approximately 1% of childhood B-precursor ALLs. Because of the infrequent occurrence of the t(17;19), no association with distinct clinical features and/or treatment outcome has been definitively identified. All cases described to date have been B-lineage leukemias, but there is no consistent linkage with cIg status. A potential association with acquired coagulation abnormalities has been noted in several reports. Raimondi et al. described 4 patients with t(17;19)-ALL, 2 of whom had disseminated intravascular coagulation (DIC) at initial diagnosis. Molecular analysis was subsequently performed on 3 of these patients, and the 2 with DIC, but not a third who lacked this clinical feature, had E2A-HLF fusion. Clinical and laboratory evidence of DIC has been observed in at least 1 other case of t(17;19)-ALL with E2A-HLF fusion; in others, clinical DIC was absent but a complete coagulation profile was not performed. Although the prognostic import of the t(17;19) has not been studied prospectively, it appears to be associated with a poor outcome. Seven patients have been described with molecularly documented E2A-HLF fusion and each has died as a direct consequence of their leukemia (A.T. Look, personal communication, June 1995).
Different E2A-HLF fusion proteins created after type I and II t(17;19)s. Two distinct types of genomic rearrangements resulting in E2A-HLF fusion have been described in t(17;19)-ALLs (Fig 3). Type I translocations result from crossovers between E2A intron 13 and HLF intron 3.37,38 In contrast to t(1;19) breakpoints, which are dispersed throughout E2A intron 13, breakpoints occur at the identical nucleotide in each type I t(17;19) examined to date. One would anticipate that splicing of precursor fusion mRNA would join sequences encoded by E2A exon 13 (e13) to those from HLF exon 4 (e4); however, this is not the case. Because these exons are in different translational reading frames, direct e13-e4 splicing would cause a frameshift and produce a truncated E2A protein that excludes all HLF residues. Instead, genomic DNA spanning the 17;19 junction is spliced into the processed mRNA to maintain an open reading frame that encodes a functional E2A-HLF fusion protein (e13-Ins-e4 splicing). The cryptic exon spanning the 17;19 breakpoint contains E2A intronic sequences at its 5′ end, HLF intronic sequences at its 3′ end, and nontemplated nucleotides in the middle. More recently, a second type of E2A-HLF chimeric protein has been described. These result from type II translocations in which crossovers occur between E2A intron 12 and HLF intron 3. E2A exon 12 and HLF exon 4 are in the same reading frame, so direct e12-e4 splicing yields a functional E2A-HLF chimeric protein. Thus, type I and II fusion proteins contain the same portions of HLF but differ on the basis of whether or not residues specified by E2A exon 13 and a cryptic exon, unique to each chimera, are present. These structural differences do not appear to have any major functional consequences, because type I and II E2A-HLF chimeric proteins behave identically in all DNA-binding and transcriptional regulatory assays in which they have been tested.117

FUNCTIONAL PROPERTIES OF E2A-HLF AND HLF

HLF is basic leucine zipper (bZIP) transcription factor. HLF is a transcription factor of the bZIP superfamily. The leucine zipper domain forms an amphipathic α-helix that allows protein monomers to homodimerize and/or heterodimerize via a parallel coiled-coil interaction and the adjacent highly charged basic region contacts the major groove of DNA and mediates site-specific DNA binding by protein dimers.112-114 The bZIP family includes several dozen proteins that can be classified into distinct subfamilies on the basis of dimerization and DNA-binding specificities. One of the more recently recognized subfamilies consists of DBP,115 TEF/VBP,116,117 and HLF.37,38 These have been collectively termed PAR proteins because, in addition to possessing highly related bZIP domains, each also includes an adjacent proline and acidic amino acid rich domain.116

The normal role of the PAR proteins is not well defined. Each has been suggested to be involved in developmental stage- or tissue-specific gene regulation, in large part based on restricted patterns of mRNA and/or protein expression. The properties of DBP, the first PAR protein identified, have been best characterized, and evidence suggests that it is important for a function that is unique to mature, nonproliferat-
ing hepatocytes. Proposed functions include regulating hepatocyte-specific expression of the albumin, cytochrome P450 CYP2C6, and cholesterol 7α hydroxylase genes. Based on its pattern of mRNA expression during embryogenesis and its ability to activate transcription of reporter genes under the control of the thyroid-stimulating hormone β gene promoter, it has been hypothesized that TEF is involved in regulation of cell differentiation in the anterior pituitary, but confirmatory functional evidence is lacking. VBP, the chicken homologue of TEF, has been suggested to play a pivotal role in estrogen-dependent regulation of vitellogenin II (VTGII) gene transcription. The normal function of HLF is unknown. HLF mRNA is expressed in a tissue-specific manner, being transcribed at relatively high levels in liver and liver-derived cell lines, at lower levels in lung and kidney, and not at all in hematolymphoid cells. Consistent with these observations, HLF protein is detected in hepatocyte-derived, but not lymphoid, cell lines.

Transcriptional regulatory properties. Each of the PAR proteins can activate transcription of reporter genes containing appropriate binding sites in transient transfection assays. DBP functions in a cell type-specific manner, activating transcription exclusively in cells of hepatic origin. In contrast, TEF/VBP and HLF activate transcription in several cell types. The PAR proteins bind DNA as homodimers or when heterodimerized with other PAR proteins, but dimerization with non-PAR bZIP proteins has not been detected and structural constraints predict that it is unlikely to occur. The PAR proteins bind to an overlapping set of target sites, eg, HLF and TEF bind optimally to the identical consensus sequence, and display an indistinguishable pattern of binding site preferences. The DNA-binding and transcriptional regulatory properties of HLF and E2A-HLF have been compared directly. Unlike the case with E2A-PBX1, fusion of HLF to E2A does not convert a nonactivator to an activator because wild-type HLF is clearly capable of activating transcription of artificial reporter genes containing either a positive or negative transcriptional effect being essential. However, loss of transforming properties after deletion of either of the E2A TADs argues that transcriptional activation is essential.

In summary, the weight of current evidence suggests that neoplastic transformation by E2A-HLF is likely mediated via activation of transcription of crucial downstream target genes, the identities of which are currently unknown. More detailed structure/function analyses are warranted to define better the domains necessary for transformation. Ultimately, identification of the relevant target genes will be required to understand the mechanisms of transformation.

ECTOPIC EXPRESSION OF bHLH PROTEINS AFTER CHROMOSOMAL REARRANGEMENTS IN T-ALL: POTENTIAL INTERACTIONS WITH E2A

Ectopic expression of bHLH proteins in T-ALL. Modification of E2A function mediated by chromosomal translocations has also been implicated in the pathogenesis of T-cell ALL. Dysregulated expression of one of these related bHLH proteins occurs after several genetic rearrangements that collectively occur in ~30% of T-ALLs. The most commonly affected is TAL1 (SCL, TCL1), expression of which is deregulated by juxtaposition with a TCR locus in 3% to 5% of T-ALL patients by the t(1;7)(p34;q35) and t(1;14)(p32;q11) or rarely by the t(1;3)(p34;q21) that fuses TAL1 with a novel chromosome 3 locus. In addition, a submicroscopic deletion that occurs in ~25% to 30% of T-ALLs also leads
to dysregulated expression by fusing TALI to regulatory elements of a gene located ~90 kb upstream. Two other infrequently observed translocations, the (t(7;19)(q34;p13) and (t(7;9)(q34;q32), result in dysregulated expression of LYL1 and TAL2, respectively, in a small percentage of T-ALLs.

**Function of TAL1, TAL2, and LYL1 and their interactions with E2A.** TAL1, TAL2, and LYL1 share extensive sequence homology with one another that is restricted to the bHLH domains. TAL1 is the most extensively studied member of this bHLH protein subfamily. TAL1-null knockout mice die in utero because they are unable to form embryonic red blood cells, arguing for a critical role for this protein in embryonic hematopoiesis. It is likely that TAL1 also plays an important role in early myelopoiesis because the knockout mice display a profound decrease in the number of myeloid cells that can be cultured from the yolk sac. The normal function of LYL1 and TAL2 is currently uncertain.

Each of these three related bHLH proteins is capable of heterodimerizing with E2A and such heterodimers bind preferentially to DNA targets distinct from those bound by E2A homodimers. (A. Miyamoto and M.L. Cleary, personal communication, July 1995). The manner by which ectopic expression of TAL1, TAL2, and LYL1 contributes to leukemogenesis has not yet been defined. In heterodimeric complexes with E2A proteins, they may regulate unique target genes or may modify transcription by altering the balance between E2A homodimers and E2A-Id dimers. The fact that they appear to act in concert with E2A provides another example of the importance of E2A proteins in normal and malignant lymphopoiesis.

**SUMMARY AND FUTURE PERSPECTIVES**

Since their identification 6 years ago, E2A proteins have been found to play an indispensable role in B-cell lymphopoiesis and to be involved in the pathogenesis of a subset of B-precursor ALLs via replacement of their bHLH regions with heterologous DNA binding domains. Furthermore, E2A proteins are also thought to play a role in the pathogenesis of some T-ALLs via interaction with ectopically expressed bHLH proteins. Given these facts, it is apparent that continued investigations into the functional properties of wild-type and chimeric E2A proteins are indicated. It will be important to define what essential properties E2A contributes to the leukemic chimeras. Portions of E2A present in E2A-PBX1 and E2A-HLF clearly function as transcriptional activation domains, but it is uncertain whether this is indeed the sole essential E2A contribution. Evidence implicating the amino terminal portion of E2A proteins in control of cell cycle progression raises the possibility that there may be essential, nontranscriptional regulatory functions for wild-type E2A proteins. In the future, it can be expected that these properties will be better defined and their potential contribution to the leukemogenic properties of E2A chimeras will be addressed experimentally.

**ACKNOWLEDGMENT**

I thank Thomas Lion, Wolf-Dieter Ludwig, Elizabeth Machntyre, and Lorna Secker-Walker for helpful discussions regarding clinical aspects of the t(1;19). I am particularly indebted to Mike Cleary, Toshiya Inaba, David LeBrun, Tom Look, Lorrie Odom, and Rob Wilson for reviewing the manuscript and making constructive suggestions.

**REFERENCES**

17. Quong MW, Massari ME, Zwant R, Murre C: A new transcriptional-activation motif restricted to a class of helix-loop-helix proteins is functionally conserved in both yeast and mammalian cells. Mol Cell Biol 13:792, 1993
19. Cordle SR, Henderson E, Masuoka H, Weil PA, Stein R: Pancreatic beta-cell-type-specific transcription of the insulin gene is


34. Kamps MP, Murre CM, Sun X-H, Baltimore D: A new homeobox gene contributes the DNA binding domain of the t(1;19) translocation protein in pre-B ALL. Cell 60:535, 1990


37. Hunger SP, Ohyashiki K, Toyama K, Cleary ML: HLF, a novel hepatic bZIP protein, shows altered DNA-binding properties following fusion to E2A in t(1;19) acute lymphoblastic leukemia. Genes Dev 6:1608, 1992


47. Yamada T, Craig JM, Hawkins JM, Jonassy G, Secker-Walker LM: Molecular investigation of 19p13 in standard and variant translocations: The E12 probe recognizes the 19p13 breakpoint in cases with t(1;19) and acute leukemia other than pre-B immunophenotype. Leukemia 5:56, 1991


63. Devaraj PE, Foroni L, Janossy G, Hoffbrand AV, Secker-Walker LM: Expression of the E2A-PBX1 fusion transcripts in t(1;19)(q23;p13) and der(19)t(1;19) at diagnosis and in remission of acute lymphoblastic leukemia with different B lineage immunophenotypes. Leukemia 9:821, 1995


65. Troussard X, Valensi F, Salamon-Nguyen F, Debert C, Flandrin G, Maclntyre E: Correlation of cytoplasmic Ig μ(Cμ) and E2A-PBX1 fusion transcripts in t(1;19) B lineage ALL: Discrepancy in Cμ detection by slide immunofluorescence and flow cytometry. Leukemia 9:518, 1995 (letter)


68. Mellentin JD, Nourse J, Hunger SP, Smith SD, Cleary ML: Molecular analysis of the t(1;19) breakpoint cluster region in pre-B-cell ALL. Genes Chromosom Cancer 2:229, 1990


76. Van Dijk MA, Voorhoeve M, Murre C: Pbx1 is converted into a transcriptional activator upon acquiring the N-terminal region of E2A in pre-B-cell acute lymphoblastic leukemia. Proc Natl Acad Sci USA 90:6061, 1993

77. LeBrun DP, Cleary ML: Fusion with E2A alters the transcriptional properties of the homeodomain protein PBX1 in t(1;19) leukemias. Oncogene 9:1641, 1994

78. Lu Q, Wright DD, Kamps MP: Fusion with E2A converts the Pbx1 homeodomain protein into a constitutive transcriptional activator in human leukemias carrying the t(1;19). Mol Cell Biol 14:3938, 1994


98. van Dijk MA, Murre C: extradenticle raises the DNA binding specificity of homeotic selector gene products. Cell 78:617, 1994
104. Kamps MP, Baltimore D: E2A-Pbx1, the t(1;19) translocation protein of human pre-B-cell acute lymphocytic leukemia, causes acute myeloid leukemia in mice. Mol Cell Biol 13:351, 1993
111. Hunger SP, Brown R, Cleary ML: DNA-binding and transcriptional regulatory properties of hepatic leukemia factor (HLF) and the t(17;19) acute lymphoblastic leukemia chimera E2A-HLF. Mol Cell Biol 14:5986, 1994
117. Iyer SV, Davis DL, Se SN, Burch JB: Chicken vitellogenin gene-binding protein, a leucine zipper transcription factor that binds to an important control element in the chicken vitellogenin II promoter, is related to rat DBP. Mol Cell Biol 11:4863, 1991
120. Lavery DJ, Schibler U: Circadian transcription of the cholesterol 7a hydroxylase gene may involve the liver-enriched bZIP protein DBP. Genes Dev 7:1871, 1993
Chromosomal translocations involving the E2A gene in acute lymphoblastic leukemia: clinical features and molecular pathogenesis

SP Hunger