REVIEW ARTICLE

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 rearranged alleles has been extensively investigated.1 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.2 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.34 One such E-box site, κE2, is important for transcription of the Ig κ light chain gene.5 Murre et al6 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 κE2 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.6-8 One additional E2A gene product, E2-5, which differs from E47 only at the extreme amino terminus, has been identified.6,9,10 The bHLH motif is an ~60 amino acid domain present in a number of proteins including products of the Drosophila gene daughterless11,12 and myogenic proteins of the MyoD family.13 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).6 Mutational studies showed that these domains were essential for protein dimerization and DNA binding.5 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.7 The preferred DNA-binding sequence of a bHLH:bHLH dimer is determined by the identity of its component monomers.14 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.15

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.9 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.16,17

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.6,16,17 In most cases, E2A heterodimer-
izes with tissue-specific bHLH proteins to regulate expression of specific target genes. 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. Despite widespread expression of E2A mRNA and protein, DNA-binding E2A homodimers appear to be present only in B-lineage lymphocytes. Recent evidence indicates that E2A homodimers are stabilized by an intramolecular 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.

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 or both the E12- and E47-specific bHLH exons 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 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{H}J\\textsubscript{H} joint, does not occur and expression of the B-lineage–specific marker B220 is not observed. 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. A basic region that 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, 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. The observation that this growth-inhibitory function is not completely dependent on the integrity of the bHLH domain 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. 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. Further studies showed that this translocation fused E2A to a previously unidentified chromosome 1 gene originally named PRL and subsequently renamed PBX1, creating an E2A-PBX1 fusion gene on the der(19) chromosome. Analysis of a large number of cases showed that identical E2A-PBX1 fusion mRNAs were present in greater than 95% of t(1;19) ALLs. 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. 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.

Clinical features. The t(1;19) was recognized to be a recurring translocation in childhood ALL in the early 1980s. 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). It is also found in a small percentage (<1%) of early pre-B (cIg+, slg−) cases and in some ALLs with a transitional pre-B (cIg−, slg+) phenotype. 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-
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.

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

<table>
<thead>
<tr>
<th>Phenotype</th>
<th>&gt;90% pre-B (clg+, slg-). Also observed in early pre-B (clg-, slg-) and transitional pre-B (clg+, slg+).</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cell surface markers</td>
<td>&gt;90% CD9+, CD10+, CD19+, CD20+, CD21+, CD22+, CD34+</td>
</tr>
<tr>
<td>Associated clinical features at presentation</td>
<td>Elevated leucocyte count, non-white race, and central nervous system leukemia.</td>
</tr>
<tr>
<td>Outcome</td>
<td>Unfavorable prognosis with antimetabolite-based therapy.</td>
</tr>
<tr>
<td>Molecular abnormality</td>
<td>&gt;85% have E2A-PBX1 fusion. Rare cases lack E2A-PBX1 fusion (usually early pre-B, CD34+).</td>
</tr>
</tbody>
</table>

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.

**Phenotype of t(1;19)+ ALLs.** Although most t(1;19)+ ALLs are classified as pre-B and express clg+, 5% to 10% of cases are clg-. In addition, clg+ cases are generally pseudodiploid and rarely contain greater than 50 chromosomes, whereas the clg- cases are generally hyperdiploid, a feature that has been associated with an excellent prognosis by a number of investigators. Privitera et al. found that t(1;19)+ ALLs were molecularly heterogeneous: detection of E2A rearrangements and expression of E2A-PBX1 fusion transcripts and chimeric proteins was limited to the clg- subset. These observations suggest that the t(1;19) observed in clg- cases does not affect E2A or PBX1, but rather involves other, yet to be identified, genes. In contrast, several groups have reported that a significant percentage of t(1;19)+ ALLs were clg- but contained E2A rearrangements and/or expressed E2A-PBX1 fusion mRNAs. Several explanations may account for at least part of these differences. First, criteria used in classifying leukemias as clg- or clg+ may differ from study to study. In support of this, groups that report a significant percentage of clg- ALLs tend also to observe a higher frequency of clg- t(1;19) than do those that have found that all clg- cases lack E2A-PBX1 fusion. Expression of cytoplasmic Igα has classically been assessed by slide immunofluorescence that can be subject to significant technical variability. Troussard et al. compared slide immunofluorescence to the newer technique of flow cytometric analysis of permeabilized cells and found that use of the latter approach reduced the number of E2A-PBX1+, clg- t(1;19) cases from 6 of 15 to 1 of 14 (1 case lacked sufficient material for flow cytometry to be performed). Secondly, it is possible that the cytogenetic classification of some leukemias may be incorrect. At least 1 case originally described as a clg- t(1;19) was subsequently analyzed using chromosome painting and fluorescence in situ hybridization (FISH) and determined to be a t(1;22). Finally, one must consider the possibility that differences in patient population may account for some of these discrepancies.

This tight association with a pre-B immunophenotype is intriguing and prompted POG investigators to perform a detailed analysis of the profile of cell surface antigen expres-
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 clg expression.

The preponderance of evidence suggests that a large majority of t(1;19)+ ALLs are clg+, express a characteristic profile of cell surface antigens, and contain E2A-PBX1 fusion, whereas a small minority of cases, despite appearing cytogenetically identical, are clg−, 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 have an L3, Burkitt's-like morphology and/or express SI−. 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 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 clg 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)+ ALL. 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 analysis to detect these translocations.
analysis for this and other high-risk chromosomal translocations in all children with ALL, particularly in those otherwise eligible for enrollment in less intensive treatment regimens.

**FUNCTIONAL PROPERTIES OF E2A-PBX1 AND PBX1: CONVERGENCE OF MOLECULAR ONCOLOGY AND DEVELOPMENTAL BIOLOGY**

**Transcriptional regulatory properties.** E2A-PBX1 contains the amino-terminal two-thirds of E2A proteins, retaining its two distinct TADs, but the E2A carboxy terminus (which includes the bHLH region) is replaced by a portion of PBX1 containing a heterologous DNA recognition motif termed the homeodomain. The potential reciprocal PBX1-E2A protein encoded by the der(1) is not produced because PBX1 is transcriptionally silent in lymphoid cells. The homeodomain is an ~60 amino acid motif originally identified in *Drosophila* homeotic selector (HOM) proteins that regulate segment identity during embryogenesis. The PBX1 homeodomain diverges significantly from that found in most *Drosophila* selectors and their mammalian homologues. Two other human proteins, PBX2 and PBX3, have been identified that share striking homology to PBX1; in fact, the homeodomains of these three proteins are nearly identical.

When the t(1;19) was cloned, the presence of a homeodomain in PBX1 suggested that it was likely to function as a transcription factor and that fusion to E2A would create a chimera that could bind to cognate PBX1 binding sites and activate target gene expression via E2A effector domains. 

Subsequent studies have confirmed these general predictions. Three groups have independently identified a consensus PBX1 binding site, 5′-ATCAATCA(A)-3′, by repetitive binding and amplification of a pool of random oligomers. The same target sequence is bound equally well by PBX1, 2, and 3 (these three proteins will subsequently be referred to collectively as PBX proteins). However, PBX and E2A-PBX1 proteins display striking differences in their transcriptional regulatory properties. PBX proteins are incapable of activating transcription of reporter genes that contain concamers of the PBX target site adjacent to a minimal promoter that drives expression of a chloramphenicol acetyl transferase (CAT) gene, whereas E2A-PBX1 functions as a potent transcriptional activator.

What then is the normal function of PBX proteins? To understand this, we must first consider data derived from studies of *Drosophila* embryogenesis. Fly HOM genes are divided into two distinct clusters, the Antennapedia and bithorax complexes. Vertebrate counterparts to these genes exist, termed the HOX genes. and they are also organized into distinct complexes within the genome. Mammalian HOX and fly HOM proteins share significant sequence homology within the homeodomain, which has been shown to mediate site-specific DNA binding, but diverge outside of this region. The function of these proteins has also been conserved throughout evolution, because knockout mice in which individual HOX genes have been genetically inactivated display discrete defects in body segment determination and the HOX genes can partially complement mutations in fly HOM genes.

The PBX proteins are the mammalian homologues of the fly protein extradenticle (exd) and the product of the *C. elegans ceh-20* gene, and are distantly related to the yeast proteins MATA1 and MATA2. In contrast to most fly homeodomain-containing proteins and their vertebrate counterparts, exd and PBX proteins share significant homology within the entire proteins, suggesting that there may be conservation of their functional properties as well. Mutations in exd, like mutations in HOM genes, cause homeotic transformations; in other words, one body segment is transformed to resemble that of another. Genetic evidence in *Drosophila* suggests that exd may function as a cofactor necessary for selector gene function. Recent data support this theory. exd interacts directly with a subset of HOM proteins and this heterocomplex binds cooperatively to DNA. 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? 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 subordinated 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. 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 murine fibroblasts, causes AML when lethally irradiated mice are reconstituted with marrow progenitors infected with retroviruses that express E2A-PBX1, and causes T-cell lymphomas after expression in transgenic mice. Several important insights into E2A-PBX1 function have been gained from these studies.

Dedera et al. 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 invoking blocked apoptosis similar to that observed in the “premalignant” stage of tumorigenesis in 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. Retroviral-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 undetermined) 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
Fig 2. Potential models for E2A-PBX1 function. (A) Based on currently available experimental evidence, PBX proteins are hypothesized to bind target gene promoters in a heterocomplex with select HOX proteins. This heterocomplex would then affect transcription positively or negatively. (B) In this scenario, PBX1 alone has no effect on target gene transcription, whereas the PBX1/HOX complex is capable of activating transcription but may be subject to a variety of regulatory influences (indicated as +*). In contrast, E2A-PBX1 might directly activate transcription by itself without requiring interaction with a specific HOX protein(s). (C) Alternatively, E2A-PBX1 might still have to bind DNA in concert with a specific HOX protein(s), but this complex could then cooperatively activate transcription without requiring other regulatory events. (D) It is also possible that PBX1 alone, or when complexed with a specific HOX protein(s), acts as a transcriptional repressor. Fusion to E2A would then result in transcription of genes that are normally inhibited in lymphoid cells.

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.107 In most cases that have been molecularly analyzed, E2A is fused to the chromosome 17 gene HLF; however, in a few cases, neither E2A nor HLF abnormalities have been detected, suggesting that this translocation, like the t(1;19), may be molecularly heterogeneous.37,38,108,109 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 al107 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.38 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.108,110 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 leukemia108-110 (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 3,7,37,38 Type I breakpoints occur within intron 13. The HLF genomic structure is based on data presented by Hunger et al.108 Type I t(17;19) breakpoints reported to date are located at the identical position within intron 13, whereas type II breakpoints occur within intron 12. The HLF genomic structure is based on data presented by Hunger et al.108 Type I 17;19 breakpoints occur in intron 3. The cryptic exon that encodes the "insertion" or "joining region" in type I E2A-HLF fusion proteins is located at the site of E2A-HLF gene fusion. Type I and II E2A-HLF fusion proteins differ from one another due to the presence or absence of residues encoded by E2A exon 13 and the insertion, which is unique to each patient. Abbreviations: AD1/AD2, transcriptional activation domains 1 and 2; bHLH, basic helix loop helix; bZIP, basic leucine zipper; PAR, proline and acidic amino acid rich domain.

**FUNCTIONAL PROPERTIES OF E2A-HLF AND HLF**

HLF is a transcription factor of the bZIP superfamily. The leucine zipper domain forms an amphipathic a-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/VP16,116,117 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. 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 and 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. In vitro, the PAR proteins bind to an overlapping set of target sites found that overexpression of E2A-HLF in fibroblasts induces malignant transformation and that this transformation is dependent on the integrity of the HLF leucine zipper domain (which is necessary for DNA binding) and the E2A TADs.

Potential models for leukemogenesis induced by E2A-HLF. Given these facts, what conclusions can be drawn about the role of E2A-HLF in leukemogenesis? Experimental evidence supports the hypothesis that E2A-HLF could function by inducing expression of downstream target genes; however, it is uncertain whether the relevant leukemogenic target genes would normally be regulated by HLF, by one of the other PAR proteins, or by other non-PAR bZIP proteins with closely related binding sites. For example, E2A-HLF can activate transcription of reporter genes containing binding sites for the C/EBP family of bZIP proteins. Another bZIP protein, E4BP4, binds to a target sequence very similar to the HLF/TEF consensus and normally functions to repress transcription. E4BP4 is present in B-lineage leukemias, including those that carry a t(17; 19), suggesting that one potential model of action of E2A-HLF could be to activate transcription of target genes normally repressed in B-cell precursors by E4BP4. Finally, one must consider possible dominant negative effects of E2A-HLF in leukemogenesis. In this scenario, E2A-HLF would heterodimerize with endogenous proteins and interfere with their function. However, no such partners have yet been identified. The fact that deletion of the HLF leucine zipper abolishes transformation by E2A-HLF in model systems is consistent with 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 three related bHLH proteins occurs after several genetic rearrangements that collectively occur in ~30% of T-ALLs. The most commonly affected is TALI (SCL, TCL5), 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 TALI 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 TAL1 to regulatory elements of a gene located ~90 kb upstream.\(^{13-15}\) Two other infrequently observed translocations, the (t7;19)(q34;p13) and (t7;9)(q34;q32), result in dysregulated expression of LYL1 and TAL2, respectively, in a small percentage of T-ALLs.\(^{16,17}\)

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.\(^{18}\) 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.\(^{19}\) 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.\(^{19,140}\) (A. Miyamoto and M.L. Cleary, personal communication, July 1995). The manner by which ektopically expressed 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.\(^{141}\) 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 Machintyre, 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 leukemias. Genes Dev 6:1608, 1992


47. Yamada T, Craig JM, Hawkins JM, Janossy 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 5:36, 1991


65. 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 90:6049, 1991


92. 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
98. Iyer SV, Davis DL, Sea 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
101. 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