The ETO Portion of Acute Myeloid Leukemia t(8;21) Fusion Transcript Encodes a Highly Evolutionarily Conserved, Putative Transcription Factor

Paul F. Erickson, Misi Robinson, Gregory Owens, and Harry A. Drabkin

Division of Medical Oncology and the Department of Neurology, University of Colorado Health Sciences and Cancer Center, Denver, Colorado 80262; and the Eleanor Roosevelt Institute for Cancer Research, Denver, Colorado 80206

ABSTRACT

The 8;21 translocation, t(8;21)(q22;q22.3), is seen only in acute myelogenous leukemia and is characteristically associated with the M2 subtype. Subsequent to our identification of the t(8;21) breakpoint region on chromosome 21, we reported that the translocation results in the fusion of the AML1 gene on chromosome 21 with a novel gene on chromosome 8 which we called ETO (for eight twenty-one). Recently, the AML1 portion of the fusion protein has been shown to correspond to the DNA-binding and dimerization domains of the mouse gene, polyoma enhancer binding protein 2aB (pebp 2aB). We report here the complete sequence of the ETO portion of the fusion transcript as compiled from complementary DNAs from a t(8;21) AML patient and compare this with the ETO sequence from a mouse brain transcript. The deduced amino acid sequences are 99% identical. ETO has several features consistent with it being a transcription factor. The ETO sequence is different from the portion of PEBP 2aB it replaces in the AML1/ETO fusion protein, except for their common high content of proline, serine, and threonine residues. Because neither the putative zinc fingers nor the TAF110 homology domain of ETO is present in PEBP2aB, one might expect functional differences in the ability of AML1/ETO protein to affect the levels of transcription of genes normally regulated to some degree by AML1 (PEBP2aB) during myeloid differentiation. The relatively high levels of ETO in developing brain suggest that it could be involved in the regulation of some aspect of neural proliferation or differentiation.

INTRODUCTION

The 8;21 translocation, t(8;21)(q22;q22.3), is seen only in acute myelogenous leukemia and is characteristically associated with the M2 subtype (1, 2). There are several notable biological features of t(8;21) AML(3) that among others include a high response rate to chemotherapy (1), a tendency to form tumor masses outside the bone marrow (3), and a marked in vitro response to interleukin 5 (4, 5). Although AML cells containing the 8;21 translocation demonstrate histological differentiation, they retain the CD34 antigen and are frequently positive for the lymphoid-associated antigen, CD19 (6, 7). Because neither the putative zinc fingers nor the TAF110 homology domain of ETO is present in PEBP2aB, one might expect functional differences in the ability of AML1/ETO protein to affect the levels of transcription of genes normally regulated to some degree by AML1 (PEBP2aB) during myeloid differentiation. The relatively high levels of ETO in developing brain suggest that it could be involved in the regulation of some aspect of neural proliferation or differentiation.

MATERIALS AND METHODS

The cDNA clone, mETO-1, was obtained from a 5-day mouse cerebellum Agt11 library by standard screening (24) of 100,000 plaques with the human ETO portion of the t(8;21) fusion transcript contained in clone PF4-9b (10). The cDNA clone, PF4-16c, was similarly obtained from the PF4 library (10). Sequencing of the ETO portion of PF4-9b and of PF4-16c was as described (10). The sequence of double-stranded mETO-1 was assembled from deletion subclones (25) and from extension of custom primers to selected portions of its previously determined sequence. Homology searches were carried out using BLASTX (26) with GenBank.

We used the following human cDNA libraries: fetal brain (Stratagene 936206); lung (Clonetech HL1158a); kidney (Clonetech HL1123a); peripheral blood leukocyte (Clonetech HL1056b); placenta (Clonetech HL1086b); and recovering bone marrow (gift of I. Hirsch). Template cDNA for PCR amplification can be identified by PCR in every patient (11-13). Where appropriately investigated over time, minimal residual disease has been detected in patients who were thought to be in complete clinical remission. About one-third of these cases subsequently relapsed (11, 14), while two-thirds of the patients have had a prolonged survival (12, 14). Accordingly, more experience will be required before clinical decisions can be made on the basis of minimal residual disease detected by the polymerase chain reaction (14).

Those chromosomal translocations in acute myeloid leukemia described to date have repeatedly resulted in the fusion of genes encoding transcription factors, the consequence of which may lead to a block in differentiation (15, 16). We and others reported that the AML1 gene had an extended and high degree of homology to a putative Drosophila transcription factor, runt (10, 17). Recent studies suggest that the full-length AML1 gene product probably is identical to a mouse transcription factor, PEBP 2aB, that binds the polyoma virus enhancer (18). They also demonstrate that runt is homologous to a second member of the PEBP2 family, 2aA, over its entire extent (19). Runt has been shown by genetic methods to be involved in the regulation of three developmental pathways: in segmentation as a pair-rule gene (20); in neurogenesis (21); and as a numerator element in sex determination (22). Therefore, AML1 seems to be the type of gene appropriate for an important role in hematopoiesis. The AML1 (pebp 2aB) transcript is present in acute myeloid leukemias that lack the 8;21 translocation, while Northern blot analysis of different subtypes of leukemia suggested that the ETO gene product is specifically seen in t(8;21) AMLs (12). Because the runt homology domain of AML1 (pebp 2aB) gene product is now known to be DNA binding (18), we would expect that the portion of the ETO gene present in the 8;21 fusion product to encode a transcriptional activation domain. Alternatively, the fusion transcript could result in a framework that would produce a truncated AML1 (PEBP 2aB) protein that could conceivably act in a dominant negative manner.

Our present results demonstrate that the fusion transcript results in an open reading frame that continues through the ETO gene. The AML1/ETO fusion transcript is predicted to encode a protein of approximately M, 85,000. The portion of the ETO gene contained in the t(8;21) fusion is highly conserved between human and mouse and has features indicative of a transactivating domain. Also of interest, ETO has some amino acid sequence homology with a recently described Drosophila TATA-binding protein accessory factor TAF110 (23).
-47 ACT TCG ACA GCA CTT ACT CTA CTA AYA ACT GNC TCC CEE AAE AAG ACG ACC TNC GAA TTA CTA TTT GAA
-45 ACT TCG ACA GCA CTT ACT CTA CTA AYA ACT GNC TCC CEE AAE AAG ACG ACC TNC GAA TTA CTA TTT GAA
-37 TCT CAA G
-3 cy e gln asp

<table>
<thead>
<tr>
<th>A</th>
<th>A</th>
<th>A</th>
<th>A</th>
<th>A</th>
<th>A</th>
<th>A</th>
<th>A</th>
<th>A</th>
<th>A</th>
<th>A</th>
<th>A</th>
<th>A</th>
<th>A</th>
<th>A</th>
<th>A</th>
<th>A</th>
</tr>
</thead>
<tbody>
<tr>
<td>T</td>
<td>G</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
</tr>
</tbody>
</table>

**RESULTS**

We used the *ETO* portion of the fusion transcript PF4-9b to screen for clones in various cDNA libraries that would contain the 3' end of *ETO*, including its polyadenylated tract. The sequence of such a clone, mETO-1, from a 5-day mouse cerebellum library is given in Fig. 1, along with a comparison with the human ETO portion of PF4-9b and PF4-16c. At the amino acid level, there are only five differences between the two species for 574 residues compared. There is also a high degree of similarity in the nucleic acid sequences of the respective 3' untranslated regions.

The amino acid sequence of ETO portion of AML 1/ETO contains a high proportion (27%) of serine, threonine, and proline residues (Fig. 1). In this respect, it is similar to the transactivating, carboxy terminal region of PEPB2aB (18), 37% of whose amino acids are prolines, serines, or threonines, and to the transactivating regions of a number of other transcription factors (28). We have indicated the three most concentrated regions of these amino acids in Fig. 2A. Although their functional significance is at present unclear, there is an interesting repetition of sequence involving the two longest runs of serines; starting at residue 72 (SSSSSSSLANQQLP) and at 378 (SSSSH-SSSURALQQLP), with their functional significance at present unclear, there is an interest.

**ETO SEQUENCE**

<table>
<thead>
<tr>
<th>A</th>
<th>A</th>
<th>A</th>
<th>A</th>
<th>A</th>
<th>A</th>
<th>A</th>
<th>A</th>
<th>A</th>
<th>A</th>
<th>A</th>
<th>A</th>
<th>A</th>
<th>A</th>
<th>A</th>
<th>A</th>
<th>A</th>
</tr>
</thead>
<tbody>
<tr>
<td>T</td>
<td>G</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
</tr>
</tbody>
</table>

Fig. 1. Sequence of *ETO* gene contained in the t(8;21) AML fusion transcripts, PF4-9b and PF4-16c, compared with and extended in the 5' direction by its mouse homologue, mETO-1. The mouse nucleic acid sequence and deduced amino acid sequence is given; the human sequence where different is given above and below. Amino acid residue number is underlined.

**REFERENCES**

1. Act TCG ACA GCA CTT ACT CTA CTA AYA ACT GNC TCC CEE AAE AAG ACG ACC TNC GAA TTA CTA TTT GAA
2. ACT TCG ACA GCA CTT ACT CTA CTA AYA ACT GNC TCC CEE AAE AAG ACG ACC TNC GAA TTA CTA TTT GAA
3. cy e gln asp
TATA-binding protein that form a complex, TFIID, essential for the transcription of ETO in brain is dramatically higher in the brain library is consistent with our experience in screening this protein. The concentration of ETO in brain is dramatically higher in the brain library is consistent with our experience in screening this protein.

The possibility they may be involved in protein-protein interactions could be investigated further. The sequence of classical DNA-binding zinc fingers (29), leaving open the question of whether these stars are involved in the coordination of zinc. The starred cysteines and histidines may not necessarily be involved in coordination of zinc.

The activated, but not basal, transcription (23) is present in AML. Finally, it is now clear from PCR studies that the aberrant transcript is involved, in a consistent fashion, in every case of t(8;21) AML. In this study, the fusion of two transcription factor genes, AML1 and ETO, producing an aberrant transcription factor, AML1/ETO, essential to the etiology of t(8;21) AML (10). Our studies, combined with those of others, strongly support this hypothesis. In particular, the deduced amino acid sequence of the ETO portion of the t(8;21) fusion transcript has multiple proline/serine/threonine-rich domains characteristic of a transcription activation domain. Also, because neither the putative zinc fingers nor TAF110 homology domain of ETO is present in PEBP2ab (18), one might expect functional differences in the ability of AML1/ETO protein to affect the levels of transcription of genes normally regulated to some degree by AML1 (PEBP2ab) during myeloid differentiation. There is the possibility, which must be experimentally verified, that AML1/ETO regulates most of the genes affected by AML1 itself because the runt homology domain of AML1 (PEBP2ab) is its DNA-binding domain (18) and is present in AML/ETO. Since AML1 (or at least its transcript) is present in AMLs in addition to t(8;21) AML (12), it appears to have an essential role in myeloid hematopoiesis at the stages of differentiation often blocked in AML. Finally, it is clear from PCR studies that the aberrant transcript is involved, in a consistent fashion, in every case of t(8;21) AML (11–13).

Human and mouse ETO are highly evolutionarily conserved: 99% identical amino acids for the 574 residues present in the fusion transcript. Recently, the sequence of two alternate transcripts of human ETO (referred to as MTG8) has independently been reported (33). The larger of these (MTG8b) has 29 amino acids coded by the gene 5' to the t(8;21) breakpoint. A comparison of these residues with those of mouse (Fig. 1) shows that the alanine at position –8 has been substituted by glutamine in the human sequence, and the valine at –17 by leucine. The degree of conservation for ETO is even higher than that for ALL (91%), the human homologue of the Drosophila gene trithorax that is involved in most leukemias having translocations at 11q23 (34). There is total conservation of sequence between mouse and human in the region of ETO homologous to Drosophila TAF110 and only one substitution in the two adjacent P/S/T-rich regions (Fig. 2B). The expression of ETO in brain (Fig. 4) correlates strongly with levels of neural proliferation and differentiation there during development.
product was found in Fig. 3, lung transcripts appear to lack some of the 5′ sequence present in transcripts from other tissues.

REFERENCES


Opinion (35) Therefore, we hypothesize that ETO is a transcription factor important to regulation of some aspect of one of these processes. Others have suggested from ETO nucleic acid sequence near the breakpoint that the fusion transcript contains a region that corresponds to one of the three highly conserved coding regions of mammalian D-type cyclins (36). However, in this region of nucleic acid homology we deduce from the data the amino acid sequence: gly-al-a-pro-arg-ser-phil-ser-thr-pro-thr-thr. Since the corresponding sequence of D-type cyclin is leu-thr-ala-glu-lys-leu-cys-ile-tyr-thr-asn (37), we conclude that ETO bears no relationship to D-type cyclin.

What aspects of the biological features of 8;21 AML cells can be detected in the 8;21 fusion mRNA in patients with (8;21) acute myeloid leukemia. Blood, 80: 1825–1831, 1992.


Note Added in Proof

Using a second set of primers, we have recently obtained a PCR product from the lung cDNA library extending from nucleotide 997 to 1724. Since no...
The *ETO* Portion of Acute Myeloid Leukemia t(8;21) Fusion Transcript Encodes a Highly Evolutionarily Conserved, Putative Transcription Factor

Paul F. Erickson, Misi Robinson, Gregory Owens, et al.


Updated version  Access the most recent version of this article at: [http://cancerres.aacrjournals.org/content/54/7/1782](http://cancerres.aacrjournals.org/content/54/7/1782)

<table>
<thead>
<tr>
<th>E-mail alerts</th>
<th>Sign up to receive free email-alerts related to this article or journal.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reprints and Subscriptions</td>
<td>To order reprints of this article or to subscribe to the journal, contact the AACR Publications Department at <a href="mailto:pubs@aacr.org">pubs@aacr.org</a>.</td>
</tr>
<tr>
<td>Permissions</td>
<td>To request permission to re-use all or part of this article, contact the AACR Publications Department at <a href="mailto:permissions@aacr.org">permissions@aacr.org</a>.</td>
</tr>
</tbody>
</table>