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 characterized 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 characterized 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 \textit{in vitro} response to interleukin 5 (4, 5). 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 \textit{in vitro} response to interleukin 5 (4, 5).

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 HL1062b); placenta (Clonetech HL1008b); and recovering bone marrow (gift of I. Hirsch). Template cDNA for PCR ampli-
prolines, serines, or threonines, and to the transactivating regions of a terminal region of PEPB2aB (18), 37% of whose amino acids are conserved between the two species for 574 residues compared. There is also a high degree of similarity in the nucleic acid sequences of the respective mETOs, including its polyadenylated tract. The sequence of such a clone, mETO-1, from a 5-day mouse cerebellum library is given in Fig. 1, with a comparison with the human ETO portion of PF4-9b and mETO, starting at residue 72 (SSSSSSSLANQQLP) and at 378 (SSSSH-EQFQ), which represent the most concentrated regions of these amino acids in Fig. 2A. Although the number of other transcription factors (28). We have indicated the three criteria (C-X-2,4-C-X-2-|5-α-X-2-4-α, where "α" represents either cysteine or proline) as occurs in ATF2 transcription factor (30). One, or possibly two, of the cysteine-proline bonds might be critical for such protein-protein interactions, which are a hallmark of transcription factors, providing more complex control of gene regulation. Such protein-protein interactions, which are a hallmark of transcription factors, provide more complex control of gene regulation. Such protein-protein interactions, which are a hallmark of transcription factors, provide more complex control of gene regulation.

**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 AML1/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 reiteration of sequence involving the two longest runs of serines; starting at residue 72 (SSSSSSSLANQQLP) and at 378 (SSSSH-EQFQ).

**Translation of ETO**

The amino acid sequence of ETO portion of AML1/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 reiteration of sequence involving the two longest runs of serines; starting at residue 72 (SSSSSSSLANQQLP) and at 378 (SSSSH-EQFQ).
TATA-binding protein that form a complex, TFIID, essential for brain library is consistent with our experience in screening this motif (32).

the rat programmed cell death gene RP-8 that also has a zinc-finger of GenBank, the ETO zinc-finger motif is most similar to a region in the possibility they may be involved in protein-protein interactions the sequence of classical DNA-binding zinc fingers (29), leaving open potential metal binding domains. However, these do not closely match the sequence or histidine) used by Berg (31) to carry out computer search for homology found by BLASTX between ETO and the TATA-binding protein accessory 1784
The product was found in Fig. 3. Lung transcripts appear to lack some of the 5' sequence present in transcripts from other tissues.

REFERENCES


Opment (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 our data the amino acid sequence: gly-ala-pro-arg-

What aspects of the biological features of 8;21 AML relate directly or indirectly to the AML1/ETO fusion protein remain to be determined. For example, whether the exaggerated eosinophilic response of 8;21 AML cells to interleukin 5, or the tendency to form tumor masses outside the bone marrow, is a direct consequence of specific genes altered by this abnormal transcription factor will require an understanding of the involved target genes and their regulation.

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

E-mail alerts
Sign up to receive free email-alerts related to this article or journal.

Reprints and Subscriptions
To order reprints of this article or to subscribe to the journal, contact the AACR Publications Department at pubs@aacr.org.

Permissions
To request permission to re-use all or part of this article, contact the AACR Publications Department at permissions@aacr.org.