The t(8;21) translocation in acute myeloid leukemia results in production of an AML1 - MTG8 fusion transcript

Hiroyuki Miyoshi¹, Tomoko Kozu, Kimiko Shimizu, Keiichiro Enomoto, Nobuo Maseki², Yasuhiko Kaneko³, Nanao Kamada⁴ and Misao Ohki

Department of Immunology and Virology, Saitama Cancer Center Research Institute, ²Hematology Clinic and ³Department of Laboratory Medicine, Saitama Cancer Center Hospital, 818 Komuro, Ina, Saitama 362, and ⁴Department of Hematology, Research Institute for Nuclear Medicine and Biology, Hiroshima University, Kasumi-cho, Hiroshima 734, Japan

¹Corresponding author

Communicated by T.H.Rabbitts

The t(8;21) translocation is one of the most frequent chromosome abnormalities in acute myeloid leukemia. It has been shown that the t(8;21) breakpoints on chromosome 21 cluster within a single specific intron of the AML1 gene, which is highly homologous to the Drosophila segmentation gene runt. Here we report that this translocation juxtaposes the AML1 gene with a novel gene, named MTG8, on chromosome 8, resulting in the synthesis of an AML1-MTG8 fusion transcript. The fusion protein predicted by the AML1-MTG8 transcript consists of the runt homology region of AML1 and the most part of MTG8, which contains putative zinc finger DNA binding motifs and proline-rich regions constituting a characteristic feature of transcription factors. The MTG8 gene is not expressed in normal hematopoietic cells, whereas AML1 is expressed at high levels. Our results indicate that the production of chimeric AML1-MTG8 protein, probably a chimeric transcription factor, may contribute to myeloid leukemogenesis.

Key words: acute myeloid leukemia/*AML1* gene/fusion gene/ *MTG8* gene/t(8;21) translocation

Introduction

Molecular analyses of chromosomal translocations in leukemias have revealed rearrangements of genes that seem to be involved in the control of cellular growth and/or differentiation, providing important clues to the etiologies of leukemogenesis. In myeloid leukemia, three translocations, the t(9;22) translocation in chronic myeloid leukemia (CML) (Heisterkamp *et al.*, 1983; Konopka *et al.*, 1984; Shtivelman *et al.*, 1985; for review see Kurzrock *et al.*, 1988), the t(15;17) translocation in acute promyelocytic leukemia (APL) (de Thé *et al.*, 1991; Goddard *et al.*, 1991; Kakizuka *et al.*, 1991; Pandolfi *et al.*, 1991) and the t(6;9) translocation in acute myeloid leukemia (AML) (von Lindern *et al.*, 1992), have been characterized and it has been shown that these translocations lead to the formation of fusion genes and chimeric proteins which presumably contribute to the process of neoplastic transformation and tumor progression.

The t(8;21)(q22;q22) translocation is one of the most frequent chromosome abnormalities in AML and is morphologically associated with the FAB-M2 subtype of AML (Rowley, 1984). Leukemic cells with the t(8;21) translocation are uniquely characterized by a high frequency of Auer rods and maturation of the granulocytic line (Berger *et al.*, 1982). Cytogenetically, this translocation is often accompanied by a loss of sex chromosome which is rarely observed in acute leukemias without t(8;21).

We have previously shown that the t(8;21) rearrangement was detected with a NotI linking clone specific for chromosome 21 by pulsed-field gel electrophoresis analysis (Shimizu et al., 1991). This rearrangement could also be detected with yeast artificial chromosome (YAC) clones spanning the breakpoint on chromosome 21 (Gao et al., 1991; Kearney et al., 1991). Our subsequent chromosome walking and cDNA screening have resulted in the isolation of the AML1 gene that was rearranged by the t(8;21)translocation (Miyoshi et al., 1991). It has been shown subsequently that the t(8;21) breakpoints on chromosome 21 clustered within a single specific intron of the AML1 gene (Shimizu et al., 1992). The AML1 gene is highly homologous to the Drosophila segmentation gene runt (Daga et al., 1992), which encodes a nuclear protein and regulates the expression of other pair-rule genes (Kania et al., 1990). The t(8;21) translocation occurs in variant forms in a limited number of patients whose clinical course is the same as that of patients with the standard t(8;21) translocation. The derivative chromosome 8 [der(8)] is constant in the variant forms of t(8;21), similar to the Philadelphia (Ph¹) chromosome in the t(9;22) translocation in CML (Rowley, 1982; Maseki et al., 1993). Therefore, the der(8) chromosome is closely related to malignant transformation in AML. Analysis of three complex translocation breakpoints revealed that all of them occurred in the same intron between two coding exons of the AML1 gene, like the standard t(8;21) breakpoints (Shimizu et al., 1992). Thus in the breakpoint region of the der(8) chromosome, this translocation juxtaposed a common 5' portion of the AML1 gene with a presumable counterpart gene on chromosome 8, suggesting the production of a fusion transcript.

We report here the isolation and characterization of a fusion cDNA from a t(8;21)-carrying cell line, demonstrating that the *AML1* gene is juxtaposed with a novel gene, *MTG8*, on chromosome 8 following the t(8;21) translocation. The wild-type *MTG8* cDNA was also isolated and characterized. On the basis of the structural features, both AML1 and MTG8 are inferred to be transcription factors. The expression of the *AML1*-*MTG8* fusion transcript is a consistent feature of the t(8;21) translocation, and the production of the chimeric protein is suggested to be directly involved in the pathogenesis of AML.

Results

Cloning of fusion cDNAs from an AML cell line with t(8;21)

We used the polymerase chain reaction (PCR) to amplify the fused cDNA sequences from an AML cell line with t(8;21), Kasumi-1 (Asou *et al.*, 1991), which had the same rearrangement of the AML1 gene as that in the t(8;21) AML patients (Shimizu *et al.*, 1992). The PCR was performed with an oligo(dT) primer and AML1-specific primer, resulting in the isolation of a part of the fusion gene (Kozu *et al.*, 1993). A 0.4 kb HincII fragment (CH15H2S, identical to nucleotides 2247-2701 in Figure 1A) of the PCR product was shown to be derived from chromosome 8. This

| A | CATAGAGCCAGCGGGCGGGGGGGGGGGGGGGGGCCCCGGGGCGGGGCGGGG | 120 240 360 480 720 840 960 1080 1200 1320 1440 1560 1680 34 |
|---|---|---|
| | GACGCCGGGCGCTGGCCCGGCAAGCTGAGGAGCGGCGGACCGCAGCATGGTGGAGGTGCTGGGCCGACCCGGCGAGCTGGTGCGCCACCGACAGCCCCAACTTCCTCCTGCTCCGTG D A G A A L A G K L R S G D R S M V E V L A D H P G E L V R T D S P N F L C S V | 1800 74 |
| | CTECCTACECACTEGEOCTECAACAAGACCCTECCCATCECTTTCAAGETGEGEGECCTAGEGEATGTTCCAGEATGECACTCTGGTCACTGTGATGECAGEGEATGATGAAAACTACTCG L P T H W R C N K T L P I A F K V V A L G D V P D G T L V T V M A G N D E N Y S | 1920 114 |
| | GCTGAGCTGAGAAATGCTACCGCAGCCATGAAGAACCAGGTTGCAAGATTTAATGACCTCAGGTTTGTCGGTCG | 2040 154 |
| | AACCCACCCCACACCTACCACACAGAGCCATCAAAATCACAGTGGATGGGCCCCCGAGAACCTCGAAATCGTACTGAGAAGCACTCCACAATGCCAGACTCACCTGTGGATGTGGAAG N P Q V A T Y H R A I K I T V D G P R E P R N R T E K H S T M P D S P V D V K AM 1 + | 2160 194 |
| | ACGCAATCTAGGCTGACTCCTCCAACAATGCCACCTCCCCCAACTACTCAAGGAGCTCCAAGAACCAGTTCATTTACACCGACAACGTTAACTAATGGCACGACGACGACTTCCCTACAGGC T Q S R L T P P T M P P P P T T Q G A P R T S S F T P T T L T N G T S H S P T A | 2280 234 |
| | TTGAATGGCGCCCCCCAACCCCAATGGCTTCAGCAATGGGCCTTCCTCTTCTTCTCCTCCTCTCTGGCTAATCAACAGCTGCCCCCAGCCTGTGGTGCCAGGCAACTCAGCAAGCTG L N G A P S P P N G F S N G P S S S S S S S L A N Q Q L P P A C G A R Q L S K L | 2400 274 |
| | AAAAGGTTCCTTACTACCCTGCAGCAGTTTGGCAATGACATTTCACCCGAGATAGGAGAAAGAGTTCGCACCTCGTTCTGGGACTAGTGAACTCCACTTTGACAATTGAAGAATTTCAT K R F L T T L Q Q F G N D I S P E I G E R V R T L V L G L V N S T L T I E E F H | 2520 314 |
| | TCCAAACTGCAAGAAGCTACTAACTTCCCACTGAGACCTTTTGTGATCCCATTTTTGAAGGCCAACTGCCCCACTGCGCGGGGGGGG | 2640 354 |
| | CCTGCCCAGTACCTCGCCCAGCATGAACAGCTGCTTCTGGATGCCAGCAGCCACCACCACCTGTTGACTCCTCAGAGCTGCTTCTCGATGTGAACGAAAACGGGAAACGGGAACTCCAGAC P A Q Y L A Q H E Q L L L D A S T T S P V D S S E L L L D V N E N G K R R T P D | 2760 394 |
| | AGAACCAAAGAAAATGGCTTTGACAGAGAGCCTTTGCACTCAGAACATCCAAGCAAG | 2880 434 |
| | AATGGCCTGCCTCACCCCACCTCCACCTCCACCATTACCGTTTGGATGATATGGCCATTGCCCCCCACCACGGGACCTCAGGGACCCCAGGGACCTCAGGGACCAGA N G L P H P T P P P Q H Y R L D D M A I A H H Y R D S Y R H P S H R D L R D R | 3000 474 |
| | AACAGAACCTATGGGGTTGCATGGCACACGTCAAGAAGAAATGATTGAT | 3120 514 |
| | ATGGTAGAAAAAACAAGGCGATCTCTCACCGTACTAAGGCGGTGTCAAGAAGCAGAACCGGGAAGAATTGAATTACTGGATCCGGCGGTACAAGTGACGCCGAGGACTTAAAAAAAGGTGGC M V E K T R R S L T V L R R C Q E A D R E E L N Y W I R R Y S D A E D L K K G G | 3240 554 |
| | GGCAGTAGCAGCAGCCACTCTAGGCAGCAGAGTCCCCGTCAACCCAGACCCAGTTGCACTAGACGCGCATCGGGAATTCCTTCACAGGCCTGCGTCTGGATACGTGCCAGAGGAGATCTGG G S S S S H S R Q Q S P V N P D P V A L D A H R E F L H R P A S G Y V P E E I W | 3360 594 |
| | AAGAAAGCTGAGGAGGCCGTCAATGAGGTGAAGCGCCAGGCGATGACGGAGCTGCAGAAGGCCGTGCTGAGGCGGAGCGGAAGGCCCACGACATGATCACAACAGAGAGGGCCAAGATG K K A E E A V N E V K R Q A M T E L Q K A V S E A E R K A H D N I T T E R A K M | 3480 634 |
| | GAGCGCACGGTCGCCGAGGCCAAACGGCAGGCGGCGGGGGGGG | 3600 674 |
| | AGTGGCTGTAACACAGCCCGATACTGTGGCTCATTTTGCCAGCACAAAGACTGGGAGAAGCACCATCACATCTGTGGACAGACCCTGCAGGCCCAGCAGGAGAACACCTGCAGTC S G C N T A R Y C G S F C Q H K D W E K H H H I C G Q T L Q A Q Q Q G D T P A V | 3720 714 |
| | AGCTCCTCTGTCACGCCCAACAGCGGGGCTGGGAGCCCGATGGACACCACCACCAGCAGCCACTCCGAGGTCAACCACCCGGGAACCCCCTTCCACCATAGAGACAACCCCTCGGCTAGAG S S S V T P N S G A G S P M D T P P A A T P R S T T P G T P S T I E T T P R * | 3840 752 |
| | TGAACTCAGAACTGTCGGAGGAAAGACAACACAACCCAACGCGAAACCAATTCCTCATCCTCAGCTCAAAGTTGTTTTTTTGTTTG | 3960 4080 4200 4287 |
| | | |

659 SSESCWNCGRKASETCSGCNTARYCGSFCQHKDWE-KHHHICGQT 702 MTG8 :: | || |: ||| | |:||: | | : |: |:|: RP8 74 GAHLCRVCGCLAPMTCSRCKQAHYCSKEHQTLDTQLGHKQACTQS 118

Fig. 1. (A) Nucleotide and deduced amino acid sequences of the AML1-MTG8 fusion cDNA. The fusion point is indicated by a vertical line. The sequence data reported here have been deposited in the DDBJ, EMBL and GenBank sequence databases under the accession number D13979. (B) Comparison of amino acid sequence of MTG8 with cell death-associated protein RP8. Vertical lines indicate identical amino acids. Double dots indicate conservative amino acid substitutions. Cysteines or histidines that are implicated in formation of zinc fingers are marked by asterisks.

was confirmed by hybridization of this fragment to Southern blots containing DNA from a panel of human – hamster cell hybrids and a mouse – human somatic cell hybrid line containing human chromosome 8 (data not shown). A cDNA library was constructed from the Kasumi-1 cell line and screened with the *AML1*-specific probe and the CH15H2S probe. Numerous positive clones were obtained and five of these cDNA clones hybridizing with both the probes were finally isolated (see Materials and methods). Since the library was not amplified, each isolate represented a unique cDNA. Restriction enzyme mapping and partial sequence analyses revealed that the overlapping regions of these cDNA clones were identical and that four of them contained a poly(A) tract. The longest cDNA clone (K1) was selected for subsequent sequence analysis.

The fusion cDNA sequence shows an in-frame joining of the 5' portion of the AML1 gene and a counterpart gene on chromosome 8, as could be expected (Figure 1A). The 5' portion of this sequence was 810 nucleotides longer than the AML1 sequence reported previously (Miyoshi et al., 1991) and was identical to the AML1 up to nucleotide 2110. This nucleotide position indicates the exon boundary in the AML1 gene interrupted by the t(8;21) translocation (Miyoshi et al., 1991). The deduced amino acid sequence of the fusion cDNA consists of 752 amino acids with a relative molecular mass of 83 174. The predicted amino acid sequence derived from chromosome 8 (amino acids 179-752) showed no significant homology to any known protein sequences in the SWISS-PROT (August, 1992) and Protein Identification Resource (June, 1992) databases, and hence we named this novel gene MTG8 (myeloid translocation gene on chromosome 8).

The C-terminal region (amino acids 663-699) of the predicted MTG8 protein contains two putative zinc finger DNA binding motifs: one resembling a Cys/Cys-Cys/Cys class of zinc finger typified by the glucocorticoid receptor family, and the other resembling a Cys/Cys-His/Cys class found in the retroviral nucleic acid binding proteins (for reviews see Evans and Hollenberg, 1988; Berg, 1990). These two potential zinc fingers are most similar to that of a cell death-associated protein RP8 (Owens *et al.*, 1991) with 41% amino acid identity (Figure 1B), although the amino acid similarity does not extend beyond this region. In addition, the C-terminal region (amino acids 738-751) carries a PEST region (PEST score, 16.2 according to the PC/GENE PEST-

FIND program) that confers rapid intracellular degradation of protein (Rogers *et al.*, 1986). Another notable feature of the MTG8 protein is its unusually high proline content in three regions: amino acids 187-264, 405-445 and 712-751 containing 22, 29 and 23% proline, respectively. A proline-rich region of CTF/NF-I has been identified as a transcriptional activation domain (Mermod *et al.*, 1989), and such regions have also been noted in many other transcription factors. Two of the three proline-rich regions (amino acids 187-264 and 712-751) are also rich in serine and threonine. The serine/threonine-rich segments might be sites of phosphorylation, which would potentially play a role in regulation of the MTG8 protein.

Cloning and characterization of the wild-type MTG8 cDNAs

The wild-type MTG8 cDNA clones were isolated from a human fetal brain cDNA library using the CH15H2S probe since the MTG8 gene was expressed in mouse brain (see Figure 6). Nucleotide sequence analysis of eight overlapping clones identified two types of composite cDNA sequence, named MTG8a and MTG8b. The sequence of the MTG8 portion of the AML1-MTG8 fusion cDNA was conserved in both types. As shown in Figure 2, their sequences upstream of the fusion point were divergent from each other. Each of the types was determined by sequencing at least two independent clones. We could also isolate the MTG8a type cDNA from Raji (Burkitt's lymphoma) cell line cDNA library and the MTG8b type from HEL (erythroleukemia) cell line using anchored PCR, noting that the MTG8 gene was expressed in both cell lines (see Figure 5). Accordingly these two types of MTG8 cDNAs are unlikely cloning artifacts but alternatively spliced forms. In fact, several alternative 5' cDNA clones were isolated. These results suggest that the fusion site with AML1 is an alternative splice acceptor site in the MTG8 transcript. The predicted open reading frames of MTG8a and MTG8b cDNAs code for 577 and 604 amino acid proteins, respectively. An upstream in-frame termination codon was identified in both types. However, the second common ATG codon located immediately downstream of the fusion point may in fact act as a translation initiation site, because the sequence surrounding the second ATG is more favorable for the Kozak consensus sequence (Kozak, 1987). In either case, the

MTG8a

| TGAATAAATTATTCGGCATTTAGCTTATCATTC <u>TGA</u> ATTTCACTTTTTGCTTTTTGGTGCTCTGAAACTTGCAGAGAGAG | 90 |
|---|-----|
| AAGGGGAGGGTAGGGGTTGTGATACTTTGCACACACACCCCCGTCATTGTTCTGCCTAAAGAGACAGGGCTGGGTTCAAGGCCACATGTG | 180 |
| CTCCTGTCATCCTCCACATTTCTGCTCCAAGTGCAATCCGGAGTGTCAGCTCTCCATCTGTCTCTGCCTGGCAGGCGCACGCGCCCAGCA | 270 |
| CCCTGCCTCCGGCGATGCCGCCCCAGCCCCTCTGATGGCCCTCCTCTGCTGCCACTCATTCCAGAACAGGAGGCATGAGCCCGGAACG | 360 |
| CGCTTGCTTTTAGGAGAGAGGCACCTTTCTGTGTGGTACGCTGGATTCAAGGATGCCTGATCGTACTGAGAAGCACTCCACAATGCCAGAC | 450 |
| м р Ї р в т е к н з т м р р | 13 |

MTG8b

| | AAAGAT <u>TGA</u> TCTCTGGGCTGG <u>TGA</u> 24 |
|--|--|
| ACATAATCTCTGTCCCAGTCAGAAAAGGAGAGAGAGAAATTAGCAGAGCGATTGGTGGAGAATG | TATCTGTCAAAAGAAACACTTGGAGA 114 |
| M | ISVKRNTWR 10 |
| GCACTGAGTTTAGTAATAGGTGACTGCCGGAAAAAAGGGAACTTTGAATATTGTCAAGATCGT | ACTGAGAAGCACTCCACAATGCCAGAC 194 |
| A L S L V I G D C R K K G N F E Y C Q D R | TEKHSTMPD 40 |

Fig. 2. Partial nucleotide and deduced amino acid sequences of the MTG8a and MTG8b cDNAs. The in-frame termination codon preceding the putative initiation codon are underlined. Arrows indicate the point of fusion with AML1. The sequence downstream of the fusion point is identical to that of the AML1-MTG8 fusion cDNA shown in Figure 1A. The sequence data reported here have been deposited in the DDBJ, EMBL and GenBank sequence databases under the accession numbers D14820 for MTG8a and D14821 for MTG8b.



Fig. 3. Schematic representation of the AML1, MTG8 and AML1-MTG8 proteins. *runt, runt* homology region; ATP, ATP binding site motif; Pro, proline-rich region; Zn, zinc finger DNA binding motifs. Amino acid numbers are shown above the box.

AML1-MTG8 fusion gene encodes a chimeric protein which retains the *runt* homology region of AML1 protein and most of the MTG8 protein. The characteristics of AML1, MTG8 and AML1-MTG8 proteins are schematically shown in Figure 3.

To confirm the rearrangements in the *MTG8* gene, we performed Southern blot analysis of leukemic cell DNAs from 17 AML patients with t(8;21) using the *MTG8* cDNA probes. An *MTG8a*-specific probe (R4H2RN) detected the rearranged *Eco*RI bands in nine cases including one complex t(8;4;21) translocation (Figure 4). However, the CH15H2S probe was not able to detect any rearrangements with *Eco*RI or *Bam*HI digestion. These results suggest that the breakpoints on chromosome 8 are clustered within a limited region, probably at the 5' end of the *MTG8* gene.

Expression of the AML1 – MTG8 fusion gene in t(8;21) AML

The expression of the AML1, MTG8 and AML1-MTG8 genes was examined by Northern blot analysis of RNA isolated from blood samples from the t(8;21) AML patients, other AML patients and normal individuals, and several hematopoietic cell lines. An AML1-specific probe (C6E3SS6) identified four major transcripts in all samples examined except for the Raji cell line, although abnormal bands were not clearly observed in the t(8;21) AML patients and the Kasumi-1 cell line because of multiple transcripts expressed at high levels (Figure 5). The expression of AMLI seems to be constitutive at various stages of hematopoietic differentiation, because it was detected in all hematopoietic cell lines examined, especially in myeloid lineages (Figure 5 and unpublished data). On the other hand, using an MTG8-specific portion of the AML1-MTG8 fusion cDNA (CH15H2S) as a probe, two major transcripts of 6.2 and 7.8 kb were detected in Kasumi-1 and all four t(8;21) AML samples examined (Figure 5 shows two samples), but not in normal peripheral blood samples, several cell lines without t(8;21) and the t(8;21) AML sample in remission. A 5.7 kb transcript was detected only in the Raji and HEL cell lines, possibly corresponding to a normal MTG8 transcript. Expression of the AML1-MTG8 fusion transcript in the



Fig. 4. Rearrangements of the *MTG8* gene. Leukemic cell DNAs from nine AML patients with t(8;21) were digested with *Eco*RI and hybridized to the R4H2RN probe (nucleotides 51-396 of *MTG8a* in Figure 2). Rearranged bands and germline bands (11 kb) are indicated by arrowheads and arrow, respectively. Namalwa (Burkitt's lymphoma cell line) cells were used as a normal control. Breakpoints of all patients in the *AML1* locus were previously reported (Shimizu *et al.*, 1992). Patient KK has complex t(8;4;21) translocation.

t(8;21) AML patients was confirmed by a reverse PCR analysis with primers flanking the fusion point (Kozu *et al.*, 1993). The cDNAs representative of the full-length 6.2 and 7.8 kb fusion transcripts were not cloned. However, all five independent fusion cDNA clones included the same coding frame with different lengths in the 5' untranslated region. Thus the major fusion transcripts probably encode the same fusion protein as that encoded by the cloned portion of cDNAs.

Expression of the AML1 and MTG8 genes in mouse tissues

The tissue distributions of the AML1 and MTG8 transcripts in various mouse tissues were examined by Northern blot analysis (Figure 6), since the AML1- and MTG8-specific





Fig. 5. Northern blot analysis of the AML1, MTG8 and AML1-MTG8 expressions. Poly(A)⁺ RNA was isolated from the following cell lines and blood samples: HL60, promyelocytic leukemia cell line; KU812, chronic myelogenous leukemia cell line; HEL, erythroleukemia cell line; Raji and Namalwa, Burkitt's lymphoma cell lines; Kazumi-1, AML cell line with t(8;21); t(8;21) Pt. 1 and Pt. 2, bone marrow samples from the t(8;21) AML patients, t(8;21) Pt. 2 is identical to patient HS in Figure 4; t(8;21) Pt. 2(R), peripheral blood sample from Pt. 2 in remission; AML Pt. 1 and Pt. 2, bone marrow samples from the AML patients without t(8;21); Normal 1 and 2, peripheral blood samples from normal individuals. Northern blot was sequentially hybridized with the indicated probes: AML1-specific probe (C6E3SS6, a 0.4 kb SmaI fragment of AML1 cDNA corresponding to nucleotides 945-1335) (Miyoshi et al., 1991), MTG8-specific probe (CH15H2S) and β -actin probe. Size markers (in kb) are shown at the right of the figure.

probes both hybridized strongly with mouse DNA. Various sizes of *AML1* transcripts were detected at different expression levels in lung, heart, spleen, thymus and ovary, possibly due to tissue-specific alternative splicing, but were not detected in brain, liver, kidney and testis. In contrast, *MTG8* expression was detected at a high level only in brain and at low levels in lung, heart, testis and ovary. Interestingly, no detectable levels of *MTG8* expression were observed in hematopoietic organs such as spleen and thymus.

Discussion

Cytogenetic studies on complex translocations have indicated that the critical recombinant chromosome is the der(8) chromosome (Rowley, 1982; Maseki *et al.*, 1993). It has recently been shown that complex translocation breakpoints as well as standard t(8;21) breakpoints on chromosome 21 consistently occurred in the same intron between two coding exons of the *AML1* gene (Shimizu *et al.*, 1992). Our results show that the t(8;21) translocation juxtaposes the *AML1* gene

Fig. 6. Tissue distributions of the *AML1* and *MTG8* transcripts. Poly(A)⁺ RNA was isolated from various mouse tissues. Northern blot was prepared and sequentially hybridized with the indicated probes as described in Figure 5. The autoradiographic exposure time was six times that for Figure 5.

with a novel gene, MTG8, on chromosome 8, resulting in the expression of the AML1 - MTG8 fusion transcripts from the breakpoint region of the der(8) chromosome. Thus, it is suggested that the resultant AML1-MTG8 fusion protein may be involved in the pathogenesis of AML.

The AML1 gene is highly homologous to the Drosophila segmentation gene runt (Daga et al., 1992), which encodes a nuclear protein containing an ATP binding site and regulates the expression of other pair-rule genes, although it does not contain any identifiable transcription factor motifs (Kania et al., 1990). The region of homology may be an evolutionarily conserved functional domain. Interestingly, the AML1 protein is disrupted at the C-terminal end of this homologous region by the t(8;21) translocation. In addition, the AML1 protein, as well as runt, is highly related to the α subunit of polyomavirus enhancer binding protein 2 (PEBP2) (Ogawa et al., 1993). PEBP2 is a transcription factor and a heterodimer of α and β subunits. The α subunit of PEBP2 is capable of binding to the core sequences of polyomavirus enhancer. The consensus sequence for binding, PuACCPuCA (Kamachi et al., 1990) is also found in enhancer regions of murine leukemia virus and several T cell specific genes such as T cell receptor genes (Satake et al., 1992). Based on the structural homology, AML1 is inferred to be a transcription factor with sequence-specific DNA binding properties and potential to dimerize with other factors.

The predicted structure of MTG8 protein is characterized by two putative zinc fingers and proline-rich domains, which have been shown to have transcriptional activation properties. The MTG8 protein may act as a transcription factor by binding to DNA and regulating the transcriptional activity of specific target genes. Hence, the AML1-MTG8 fusion protein is suggested to be a chimeric transcription factor, containing two putative DNA binding domains derived from AML1 and MTG8, though it is uncertain whether these domains recognize and bind to specific DNA sequences. Several chimeric transcription factors involved in leukemogenesis have been identified in recent molecular studies on non-random chromosomal translocations in hematopoietic malignancies (for reviews see Cleary, 1991; Rabbitts, 1991; Solomon et al., 1991). For example, the t(1;19) translocation in pre-B cell acute lymphoblastic leukemia results in the fusion of the transcriptional activation domain of E2A and the DNA binding homeodomain of PBX1 (Kamps et al., 1990; Nourse et al., 1990). In acute promyelocytic leukemia, the t(15;17) translocation fuses the retinoic acid receptor α gene with the *PML* gene, which encodes a novel zinc finger protein (de Thé et al., 1991; Goddard et al., 1991; Kakizuka et al., 1991; Pandolfi et al., 1991).

Northern blot analysis of AML1 and MTG8 expression may help to elucidate the role of the AML1-MTG8 fusion gene in myeloid leukemogenesis. Expression of the AML1 gene is probably constitutive in multiple hematopoietic lineages, suggesting that AML1 may play an important role in hematopoietic cell growth and/or differentiation. If so, the AML1-MTG8 fusion protein might interfere with the function of wild-type AML1 as a dominant negative mutant. Alternatively, aberrant expression of truncated MTG8 following juxtaposition with the 5' portion of AML1 as a promoter may be involved in leukemogenic transformation as in the case of c-myc in Burkitt's lymphoma, since the MTG8 gene is not expressed in normal hematopoietic cells. However, it is most likely that the resultant AML1-MTG8 fusion protein contributes directly to leukemogenesis, because the breakpoints on chromosome 21 consistently occurred in the single specific intron of AML1, and AML1-MTG8 fusion protein retains presumable functional domains of AML1 protein, such as ATP binding and DNA binding domains, which would act as functional parts of a chimeric oncoprotein. The AML1-MTG8 protein presumably acts as a chimeric transcription factor and alters the transcriptional activity of AML1- or MTG8-responsive genes which contribute to neoplastic process of hematopoietic origin. It is also possible that AML1-MTG8 would have different biological properties from those of the respective wild-type proteins. Functional analysis of the AML1-MTG8 fusion protein is required to clarify its role in AML as well as the characterization of wild-type AML1 and MTG8 proteins.

Materials and methods

Cloning of AML1 – MTG8 cDNAs

Poly(A)⁺ RNA (5 µg) from the Kasumi-1 cell line was converted to double-stranded cDNA. Size-selected cDNA library (>2.0 kb) was constructed in the λ ZAPII vector (Stratagene) with *Eco*RI–*Not*I adaptors (Pharmacia). The library was screened with both the *AML1*-specific probe (C6E6H2) (Miyoshi *et al.*, 1991) and the CH15H2S probe (a 0.4 kb *Hinc*II fragment identical to nucleotides 2247–2701 in Figure 1A). Approximately 1 × 10⁶ recombinant phages were screened, and 20 of >100 positive plaques were initially isolated. The cDNA inserts from each positive clone were checked by *Eco*RI and *Not*I digestion of the purified phage DNA and agarose gel electrophoresis. Subsequently, the inserts of five clones were subcloned into the *Not*I site of pBluescript II (KS+) (Stratagene) for further analysis.

Cloning of MTG8 cDNAs

Wild-type MTG8 cDNAs were isolated from the human fetal brain (17–18 weeks gestation) cDNA library (Stratagene) using the CH15H2S probe. Twenty positive clones were initially isolated and characterized. Eight overlapping clones were then selected for sequence analysis, and two types of cDNA, MTG8a and MTG8b, were identified. MTG8a type cDNA could also be isolated from the Raji cell line cDNA library using the CH15H2S probe. MTG8b type cDNA was also isolated from the HEL cell line using anchored PCR with antisense primer downstream of the fusion point with AML1.

DNA sequencing

The inserts of cDNA clones were subcloned into the pBluescript II (KS+) vector (Strategene), and a nested series of deletions was generated by using Exo/Mung Kit (Stratagene). Both strands of cDNA clones were sequenced by the dideoxy nucleotide method using an ALF DNA sequencer (Pharmacia).

Southern and Northern blot analyses

Leukemic cell DNAs (5 μ g) from AML patients with t(8;21) were digested with the appropriate restriction enzymes and separated on 0.6% agarose gel. Total RNA was isolated from blood samples, cell lines and mouse tissues. Poly(A)⁺ RNA was selected by using Oligotex-dT30 (Roche). Samples of poly(A)⁺ RNA (1.5 μ g) were separated on 1% agarose – formaldehyde gel. Blotting, hybridization, washing and autoradiography were performed as previously described (Miyoshi *et al.*, 1991).

Acknowledgements

We thank Dr K.Nakachi for critical reading of the manuscript, Dr O.Gotoh for computer analysis, M.Mori for technical assistance and N.Kobayashi for help in the preparation of the manuscript. We also thank Dr M.Oshimura for providing the mouse – human somatic cell hybrid line. This work was supported in part by a Grant-in-Aid for a Creative Basic Research (Human Genome Program) from the Ministry of Education, Science and Culture, a grant from Special Coordination Funds for Promoting Science and Technology of the Science and Technology Agency, and a Grant-in-Aid for a Comprehensive 10-Year Strategy for Cancer Control from the Ministry of Health and Welfare of Japan.

References

- Asou, H., Tashiro, S., Hamamoto, K., Otsuji, A., Kita, K. and Kamada, N. (1991) *Blood*, **77**, 2031–2036.
- Berg, J.M. (1990) J. Biol. Chem., 265, 6513-6516.
- Berger, R., Bernheim, A., Daniel, M.-T., Valensi, F., Sigaux, F. and Flandrin, G. (1982) *Blood*, **59**, 171-178.
- Cleary, M.L. (1991) Cell, 66, 619-622.
- Daga, A., Tighe, J.E. and Calabi, F. (1992) Nature, 356, 484.
- de Thé,H., Lavau,C., Marchio,A., Chomienne,C., Degos,L. and Dejean,A. (1991) Cell, 66, 675-684.
- Evans, R.M. and Hollenberg, S.M. (1988) Cell, 52, 1-3.
- Gao, J., Erickson, P., Gardiner, K., Le Beau, M.M., Diaz, M.O., Patterson, D., Rowley, J.D. and Drabkin, H.A. (1991) Proc. Natl Acad. Sci. USA, 88, 4882-4886.
- Goddard, A.D., Borrow, J., Freemont, P.S. and Solomon, E. (1991) Science, 254, 1371-1374.
- Heisterkamp, N., Stephenson, J.R., Groffen, J., Hansen, P.F., de Klein, A., Bartram, C.R. and Grosveld, G. (1983) Nature, **306**, 239-242.
- Kakizuka, A., Miller, W.H., Jr, Umesono, K., Warrell, R.P., Jr, Frankel, S.R., Murty, V.V.V.S., Dmitrovsky, E. and Evans, R.M. (1991) Cell, 66, 663-674.
- Kamachi, Y., Ogawa, E., Asano, M., Ishida, S., Murakami, Y., Satake, M., Ito, Y. and Shigesada, K. (1990) J. Virol., 64, 4808-4819.
- Kamps, M.P., Murre, C., Sun, X.-h. and Baltimore, D. (1990) Cell, 60, 547-555.
- Kania, M.A., Bonner, A.S., Duffy, J.B. and Gergen, J.P. (1990) *Genes Dev.*, 4, 1701–1713.
- Kearney, L., Watkins, P.C., Young, B.D. and Sacchi, N. (1991) Cancer Genet. Cytogenet., 57, 109-119.
- Konopka, J.B., Watanabe, S.M. and Witte, O.N. (1984) Cell, 37, 1035-1042.
- Kozak, M. (1987) Nucleic Acids Res., 15, 8125-8148.
- Kozu, T., Miyoshi, H., Shimizu, K., Maseki, N., Kaneko, Y., Asou, H., Kamada, N. and Ohki, M. (1993) Blood, in press.

- Kurzrock, R., Gutterman, J.U. and Talpaz, M. (1988) New Engl. J. Med., 319, 990-998.
- Maseki, N., Miyoshi, H., Shimizu, K., Homma, C., Ohki, M., Sakurai, M. and Kaneko, Y. (1993) Blood, 81, 1573-1579.
- Mermod, N., O'Neill, E.A., Kelly, T.J. and Tjian, R. (1989) Cell, 58, 741-753.
- Miyoshi, H., Shimizu, K., Kozu, T., Maseki, N., Kaneko, Y. and Ohki, M. (1991) Proc. Natl Acad. Sci. USA, 88, 10431-10434. Nourse, J., Mellentin, J.D., Galili, N., Wilkinson, J., Stanbridge, E.,
- Smith, S.D. and Cleary, M.L. (1990) Cell, 60, 535-545.
- Ogawa, E., Maruyama, M., Kagoshima, H., Inuzuka, M., Lu, J., Satake, M., Shigesada, K. and Ito, Y. (1993) Proc. Natl Acad. Sci. USA, in press.
- Owens, G.P., Hahn, W.E. and Cohen, J.J. (1991) Mol. Cell. Biol., 11, 4177-4188.
- Pandolfi, P.P., Grignani, F., Alcalay, M., Mencarelli, A., Biondi, A., LoCoco, F., Grignani, F. and Pelicci, P.G. (1991) Oncogene, 6, 1285 - 1292.
- Rabbitts, T.H. (1991) Cell, 67, 641-644.
- Rogers, S., Wells, R. and Rechsteiner, M. (1986) Science, 234, 364-368.
- Rowley, J.D. (1982) Science, 216, 749-751.
- Rowley, J.D. (1984) Cancer Res., 44, 3159-3168.
- Satake, M., Inuzuka, M., Shigesada, K., Oikawa, T. and Ito, Y. (1992) Jpn J. Cancer Res., 83, 714-722.
- Shimizu, K., Ichikawa, H., Miyoshi, H., Ohki, M., Kobayashi, H., Maseki, N. and Kaneko, Y. (1991) Genes Chrom. Cancer, 3, 163-167.
- Shimizu, K., Miyoshi, H., Kozu, T., Nagata, J., Enomoto, K., Maseki, N., Kaneko, Y. and Ohki, M. (1992) Cancer Res., 52, 6945-6948.
- Shtivelman, E., Lifshitz, B., Gale, R.P. and Canaani, E. (1985) Nature, 315, 550-554.
- Solomon, E., Borrow, J. and Goddard, A.D. (1991) Science, 254, 1153-1160.
- von Lindern, M., Fornerod, M., van Baal, S., Jaegle, M., de Wit, T., Buijs, A. and Grosveld, G. (1992) Mol. Cell. Biol., 12, 1687-1697.

Received on December 2, 1992; revised on April 8, 1993

Note added in proof

Recently, Erickson et al. [Blood, 80, 1825-1831 (1992)] and Nisson et al. [Cancer Genet. Cytogenet., 63, 81-88 (1992)] have reported the partial nucleotide sequences at the junction of the fusion transcript.