The EMBO Journal vol.13 no.3 pp.504–510, 1994

Generation of the *AML1–EVI-1* fusion gene in the t(3;21)(q26;q22) causes blastic crisis in chronic myelocytic leukemia

Kinuko Mitani, Seishi Ogawa, Tomoyuki Tanaka, Hiroyuki Miyoshi[1], Mineo Kurokawa, Hiroyuki Mano, Yoshio Yazaki, Misao Ohki[1] and Hisamaru Hirai[2]

Third Department of Internal Medicine, Faculty of Medicine, University of Tokyo, Tokyo 113 and [1]Department of Immunology and Virology, Saitama Cancer Center Research Institute, Saitama 362, Japan

[2]Corresponding author

Communicated by K.Alitalo

The t(3;21)(q26;q22) translocation, which is one of the consistent chromosomal abnormalities found in blastic crisis of chronic myelocytic leukemia (CML), is thought to play an important role in the leukemic progression of CML to an acute blastic crisis phase. The *AML1* gene, which is located at the translocation breakpoint of the t(8;21)(q22;q22) translocation found in acute myelocytic leukemia, was also rearranged by the t(3;21)(q26;q22) translocation. Screening of a cDNA library of the t(3;21)-carrying leukemic cell line cells (SKH1) resulted in the isolation of two potentially complete AML1–EVI-1 chimeric cDNAs of 6 kb. Two species of AML1–EVI-1 fusion transcripts of 8.2 and 7.0 kb were detected in SKH1 cells. These cells expressed the 180 kDa AML1–EVI-1 fusion protein containing an N-terminal half of AML1 including a runt homology domain which is fused to the entire zinc finger EVI-1 protein. The AML1–EVI-1 fusion transcript was consistent in all three cases of the t(3;21)-carrying leukemia examined by RNA-based PCR. These findings strongly suggest that the t(3;21) translocation results in the formation of a new class of chimeric transcription factor which could contribute to the leukemic progression of CML through interference with cell growth and differentiation.

Key words: AML1/chronic myelocytic leukemia/EVI-1/transcription factor/translocation

Introduction

Defined karyotypic abnormalities are associated with specific subtypes of human leukemias. Molecular characterization of these abnormalities, especially reciprocal translocations, has shown that the chromosomal abnormalities themselves are implicated in leukemogenesis by altering the function or activities of genes located at or near the translocation breakpoint (Solomon *et al.*, 1991). The inappropriate expression of transcription factors or the creation of fusion mRNAs and fusion transcription factor proteins by chromosomal translocation are important mechanisms in leukemogenesis (Nichols and Nimer, 1992). It is well known that the t(15;17) in acute promyelocytic leukemia (M3 according to the French–American–British classification) (de The *et al.*, 1991; Kakizuka *et al.*, 1991), t(8;21) in acute myelocytic leukemia with maturation (M2) (Miyoshi *et al.*, 1991, 1993; Nucifora *et al.*, 1993), t(6;9) in acute myelocytic or myelomonocytic leukemia (M2 or M4) (von Lindern *et al.*, 1992), and t(1;19) in acute lymphocytic leukemia (pre-B cell type) (Kamps *et al.*, 1990; Nourse *et al.*, 1990) result in the formation of chimeric transcription factor genes, *PML–RARα*, *AML1–MTG8* or *AML1–ETO*, *DEK–CAN* and *E2A–PBX1*, respectively. However, little is known about the biological activities of these chimeric transcription factors (Kamps *et al.*, 1991).

Chronic myelocytic leukemia (CML) is a clonal disorder of pluripotent hematopoietic stem cells usually with a biphasic clinical course (Fialkow *et al.*, 1977; Sandberg, 1980). An initial chronic phase characterized by leukocytosis with maturation is usually followed by acceleration of the disease and ultimately by an acute blastic crisis phase characterized by cellular proliferation, maturation arrest and karyotypic clonal evolution (Alimena *et al.*, 1987). A reciprocal translocation, t(9;22)(q34;q11), is the hallmark of the chronic phase of CML (Rowley, 1973) and generates the *BCR–ABL* fusion gene (Heisterkamp *et al.*, 1983) which is translated into the BCR–ABL fusion protein with an enhanced tyrosine kinase activity (Konopka *et al.*, 1984) directly responsible for its *in vivo* tumorigenic potential (Daley *et al.*, 1990). However, the molecular events leading to blastic crisis of CML have not been well characterized although alterations of the *p53* gene, including rearrangements, deletions or point mutations, have been reported in ~30% of CML patients at blastic crisis phase (Ahuja *et al.*, 1991; Foti *et al.*, 1991; Feinstein *et al.*, 1992). It is known that additional chromosomal abnormalities to the t(9;22) translocation, such as +8, +19, +Ph or iso(17q), appear in or prior to the blastic crisis phase of CML (Alimena *et al.*, 1987). The t(3;21)(q26;q22) translocation is one of these additional chromosomal abnormalities and ~20 cases have been reported so far (Schneider *et al.*, 1991). It is not frequently observed in blastic crisis of CML and is a reciprocal translocation (Rubin *et al.*, 1987; Coyle and Najfeld, 1988) that is unique because other common additional chromosomal changes are mainly trisomy or isochromosome. Because this structural change occurs, like other additional abnormalities, prior or near to the time of blastic crisis development (Rubin *et al.*, 1987; Coyle and Najfeld, 1988), it could play a causative role in blastic crisis of CML even if the involvement of this abnormality in blastic crisis would be limited to a small percentage of CML.

The t(3;21) translocation is observed not only in blastic crisis of CML but also in myelodysplastic syndrome (MDS)-derived leukemia (Rubin *et al.*, 1990). Because it is very rare that *de novo* acute leukemia carries the t(3;21) translocation (Rubin *et al.*, 1990), the appearance of this chromosomal abnormality may be a key event when hematological disorders arising from hematopoietic stem

© Oxford University Press

cells transform to the acute leukemic phase (Chen *et al.*, 1991). Thus the molecular characterization of the t(3;21) translocation should provide useful insights into the mechanism in transformation of pluripotent stem cell disorders.

We report here the isolation of AML1-EVI-1 fusion cDNAs from a human leukemic cell line, SKH1, carrying the t(3;21) translocation. Structural analysis reveals that the AML1-EVI-1 fusion protein is a chimeric transcription factor including a runt homology domain (Daga *et al.*, 1992) from the AML1 (Miyoshi *et al.*, 1991) and two zinc finger domains and an acidic domain from EVI-1 (Morishita *et al.*, 1990a). PCR analysis with mRNAs from leukemic cells from three patients with the t(3;21) translocation showed the identical AML1-EVI-1 mRNA junction. We propose that the AML1-EVI-1 fusion protein could trigger blastic crisis in CML showing the t(3;21) translocation.

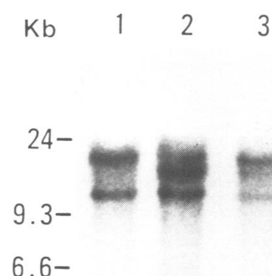

Fig. 1. Rearrangement of the *AML1* gene in leukemic cells with the t(3;21)(q26;q22) translocation. Genomic DNAs (10 μg) from leukemic cells of two patients with blastic crisis of CML (patient 1, lane 2; patient 2, lane 3) and from white blood cells of one normal control (lane 1) were digested with *Bam*HI restriction enzyme, electrophoresed on a 0.8% agarose gel and transferred to a nylon membrane filter. The blot was hybridized with the AML1 cDNA probe. Rearranged bands of 15 and 17 kb are shown in lanes 2 and 3, respectively. The DNA markers are aligned.

Results

Modification of the AML1 gene by the t(3;21) translocation

To determine the breakpoints of the t(3;21) translocation, we performed pulsed field gel electrophoresis on leukemic cells from a patient with blastic crisis of CML carrying the t(3;21) translocation in addition to the Ph translocation (patient 1), and screened for rearrangements with a number of DNA probes assigned on chromosome 3q26 or 21q22. Among them, rearranged bands were detected for *Not*I (1600 kb) and *Sfi*I (220 kb) with an AML1 cDNA probe (C6E6H2) (data not shown). The *AML1* gene is located at chromosomal band 21q22 and is known to be rearranged by the t(8;21) translocation (Miyoshi *et al.*, 1991). Because the germline band for *Sfi*I was 100 kb, the breakpoint on chromosome 21 of the t(3;21) translocation should have been within 100 kb from the cDNA probe. To identify the breakpoint precisely, Southern analysis was performed on leukemic cells from two cases (patient 1 and patient 2) of t(3;21)-carrying blastic crisis of CML. After *Bam*HI digestion, the AML1 cDNA probe hybridized to two germline fragments of 11 and 19 kb and the rearranged bands (15 kb in patient 1 and 17 kb in patient 2, respectively) were detected in both cases (Figure 1). Therefore, the rearrangement had occurred within the *Bam*HI fragment of 11 or 19 kb, as is the case with the *AML1* gene rearrangement in the t(8;21) translocation (Miyoshi *et al.*, 1991). These data suggest that the *AML1* gene of the leukemic cells was rearranged by the t(3;21) translocation.

Isolation of the AML1-EVI-1 fusion cDNA

We established a human leukemic cell line, SKH1, from patient 1 and constructed a cDNA library from poly(A)$^+$ mRNAs of SKH1 cells to investigate the probability of fusion mRNAs which cross the t(3;21) breakpoint. Subsequent screening of 1.2×10^6 recombinant λ phages of the library with the AML1 probe resulted in the isolation of 10 recombinant clones, containing normal AML1 sequences only in their 5′ regions. Two of them were potentially complete cDNAs of 6.0 kb, bearing the 5′ non-coding and following coding sequences of AML1, sequences different from AML1 and poly(A) tails. The restriction maps of these two cDNAs (clones 4 and 10) were identical (data not shown). The complete nucleotide sequence of clone 4 cDNA extends for 5937 nucleotides and contains a single long open reading frame of 4185 nucleotides, encoding a 1395 amino acid protein with a predicted molecular weight of 154 245 (Figure 2). The nucleotide sequence of the clone showed that the normal AML1 sequence was abruptly interrupted at the end of the runt homology region, as in the case of the chimeric cDNA in the t(8;21) translocation (Miyoshi *et al.*, 1991). In the following sequence of 338 bp (CHR3), there was no in-frame termination codon. Surprisingly, human EVI-1 sequence from the second exon to the poly(A) tail (Morishita *et al.*, 1990a) followed the CHR3 sequence. The CHR3 sequence was mapped to human chromosome 3 by Southern analysis with human-mouse hybrid cell DNAs (data not shown). The 200 bp fragment derived from the CHR3 detected human EVI-1 transcripts in HEC-1 cells by Northern analysis (data not shown). These observations demonstrate that the CHR3 sequence should be derived from a 5′ non-coding exon of human *EVI-1* gene. Compared with the published human EVI-1 sequence (Morishita *et al.*, 1990a), there were nine nucleotide changes: three were silent and six caused the alteration of five amino acids. The changed amino acids detected in the AML1-EVI-1 fusion protein were identical to the corresponding amino acids of mouse EVI-1 (Morishita *et al.*, 1988). The insertion of 27 bp at 2281-2307 bp found in the human EVI-1 sequence was not present in the chimeric cDNA sequence or the mouse EVI-1 sequence.

The predicted structure of the *AML1-EVI-1* gene product is shown in Figure 3. The AML1-EVI-1 fusion protein is a chimeric transcription factor including a runt homology domain from AML1 and two zinc finger domains from EVI-1; in total three DNA binding domains and an acidic domain from EVI-1 as a transcriptional activation domain.

The AML1-EVI-1 fusion transcripts

To clarify the expression of the AML1-EVI-1 fusion transcripts by the t(3;21) translocation, radiolabeled AML1 and EVI-1 probes were examined for hybridization to poly(A)$^+$ mRNAs extracted from SKH1 and KU812 cells. KU812 cells showed the normal AML1 transcripts with six major bands of 8.2, 7.9, 7.4, 6.7, 4.4 and 2.7 kb (Figure 4). Several bands hybridized with the AML1 probe were also

TTTCC AGGCA CTCTC ATTCA TAGAG CCAGC GGGCG CGGGC GGGAC GGGCG CCCCG CGGCC 60
GGACC CAGCC AGGGC ACCAC GCTGC CCGGC CCTGC GCCGC CAGGC ACTTC TTTCC GGGGC 120
TCCTA GGGAC GCCAG AAGGA AGTCA ACCTC TGCTG CTTCT CCTTG GCCTG CGTTG GACCT 180
TCCTT TTTTT GTTGT TTTTT TTTGT TTTTC CCCTT TCTTC CTTTT GAATT AACTG GCTTC 240
TTGGC TGGAT GTTTT CAACT TCTTT CCTGG CTGCG AACTT TTTCC CCAAT TGTTT TCCTT 300
TTACA ACAGG GGGAG AAAGT GCTCT GTGGT CCGAG GCGAG CCGTG AAGTT GCGTG TGCGT 360
GGCAG TGTGC GTGGC AGGAT GTGCG TGCGT GTGTA ACCCG AGCCG CCCGA TCTGT TTCGA 420
TCTGC GCCGC GGAGC CCTCC CTCAA GGCCC GCTCC ACCTG CTTGG CGGTT ACGCG GCGCT 480
CGTGG GTGTT CGTGC CTTCG GAGCA GCTAA CCGGC GGGTG CTGGG CGACG GTGGA GGAGT 540
ATCGT TCTCG CTGCT TGCCC GAGTC AGGGC TGAGT CACCC CAGCT GATGT AGACA GTGGC 600
TGCCT TCCGA AGAGT GCGTG TTTGC ATGTG TGTGA CTCTG CGGCT GCTCA ACTCC CAACA 660
AACCA GAGGA CCAGC CACAA ACTTA ACCAA CATCC CCAAA CCCGA GTTCA CAGAT GTGGG 720
AGAGC TGTAG AACCC TGAGT GTCAT CGACT GGGCC TTCTT ATGAT TGTTG TTTTA AGATT 780
AGCTG AAGAT CTCTG AAACG CTGAA TTTTC TGCAC TGAGC GTTTG ACAGA ATTCA TTGAG 840
AGAAC AGAGA ACATG ACAAG TACTT CTAGC TCAGC ACTGC TCCAA CTACT GAAGC TGATT 900
TTCAA GGCTA CTTAA AAAAA TCTGC AGCGT ACATT AATGG ATTTC TGTTG TGTTT AAATT 960
CTCCA CAGAT TGTAT TGTAA ATATT TTATG AAGTA GAGCA TATGT ATATA TTTAT ATATA 1020
CGTGC ACATA CATTA GTAGC ACTAC CTTTG GAAGT CTCAG CTCTT GCTTT TCGGG ACTGA 1080
AGCCA GTTTT GCATG ATAAA AGTGG CCTTG TTACG GGAGA TAATT GTGTT CTGTT GGGAC 1140
TTTAG ACAAA ACTCA CCTGC AAAAA ACTGA CAGGC ATTAA CTACT GCAAC TTCCA AATAA 1200
TGTGT TTGCT GATCG TTTTA CTCTT CGCAT AAATA TTTTA GGAAG TGTAT GAGAA TTTTG 1260
CCTTC AGGAA CTTTT CTAAC AGCCA AAGAC AGAAC TTAAC CTCTG CAAGC AAGAT TCGTG 1320
GAAGA TAGTC TCCAC TTTTT AATGC ACTAA GCAAT CGGTT GCTAG GAGCC CATCC TGGGT 1380
CAGAG GCCGA TCCGC AGAAC CAGAA CGTTT TCCCC TCCTG GACTG TTAGT AACTT AGTCT 1440
CCCTC CTCCC CTAAC CACCC CCGCC CCCCC CCACC CCCCG CAGTA ATAAA GGCCC CTGAA 1500
CGTGT ATGTT GGTCT CCCGG GAGCT GCTTG CTGAA GATCC GCGCC CCTGT CGCCG TCTGG 1560
TAGGA GCTGT TTGCA GGGTC CTAAC TCAAT CGGCT TGTTG TG ATG CGT ATC CCC GTA 1617
Met Arg Ile Pro Val 5

GAT GCC AGC ACG AGC CGC CGC TTC ACG CCG CCT TCC ACC GCG CTG AGC CCA GGC 1671
Asp Ala Ser Thr Ser Arg Arg Phe Thr Pro Pro Ser Thr Ala Leu Ser Pro Gly 23

AAG ATG AGC GAG GCG TTG CCG CTG GGC GCC CCG GAC GCC GGC GCT GCC CTG GCC 1725
Lys Met Ser Glu Ala Leu Pro Leu Gly Ala Pro Asp Ala Gly Ala Ala Leu Ala 41

GGC AAG CTG AGG AGC GGC GAC CGC AGC ATG GTG GAG GTG CTG GCC GAC CAC CCG 1779
Gly Lys Leu Arg Ser Gly Asp Arg Ser Met Val Glu Val Leu Ala Asp His Pro 59

GGC GAG CTG GTG CGC ACC GAC AGC CCC AAC TTC CTC TGC TCC GTG CTG CCT ACG 1833
Gly Glu Leu Val Arg Thr Asp Ser Pro Asn Phe Leu Cys Ser Val Leu Pro Thr 77

CAC TGG CGC TGC AAC AAG ACC CTG CCC ATC GCT TTC AAG GTG GTG GCC CTA GGG 1887
His Trp Arg Cys Asn Lys Thr Leu Pro Ile Ala Phe Lys Val Val Ala Leu Gly 95

GAT GTT CCA GAT GGC ACT CTG GTC ACT GTG ATG GCT GGC AAT GAT GAA AAC TAC 1941
Asp Val Pro Asp Gly Thr Leu Val Thr Val Met Ala Gly Asn Asp Glu Asn Tyr 113

TCG GCT GAG CTG AGA AAT GCT ACC GCA GCC ATG AAG AAC CAG GTT GCA AGA TTT 1995
Ser Ala Glu Leu Arg Asn Ala Thr Ala Ala Met Lys Asn Gln Val Ala Arg Phe 131

AAT GAC CTC AGG TTT GTC GGT CGA AGT GGA AGA GGG AAA AGC TTC ACT CTG ACC 2049
Asn Asp Leu Arg Phe Val Gly Arg Ser Gly Arg Gly Lys Ser Phe Thr Leu Thr 149

ATC ACT GTC TTC ACA AAC CCA CCG CAA GTC GCC ACC TAC CAC AGA GCC ATC AAA 2103
Ile Thr Val Phe Thr Asn Pro Pro Gln Val Ala Thr Tyr His Arg Ala Ile Lys 167

ATC ACA GTG GAT GGG CCT CGA GAA CCT CGA AAT AAT GAG TGT GTA TAT GGC AAC 2157
Ile Thr Val Asp Gly Pro Arg Glu Pro Arg Asn Asn Glu Cys Val Tyr Gly Asn 185

TAC CCT GAA ATA CCT TTG GAA GAA ATG CCA GAT GCA GAT GGA GTA GCC AGC ACT 2211
Tyr Pro Glu Ile Pro Leu Glu Glu Met Pro Asp Ala Asp Gly Val Ala Ser Thr 203

CCC TCC CTC AAT ATT CAA GAG CCA TGC TCT CCT GCC ACA TCC AGT GAA GCA TTC 2265
Pro Ser Leu Asn Ile Gln Glu Pro Cys Ser Pro Ala Thr Ser Ser Glu Ala Phe 221

ACT CCA AAG GAG GGT TCT CCT TAC AAA GCC CCC ATC TAC ATC CCT GAT GAT ATC 2319
Thr Pro Lys Glu Gly Ser Pro Tyr Lys Ala Pro Ile Tyr Ile Pro Asp Asp Ile 239

CCC ATT CCT GCT GAG TTT GAA CTT CGA GAG TCA AAT ATG CCT GGG GCA GGA CTA 2373
Pro Ile Pro Ala Glu Phe Glu Leu Arg Glu Ser Asn Met Pro Gly Ala Gly Leu 257

GGA ATA TGG ACC AAA AGG AAG ATC GAA GTA GGT GAA AAG TTT GGG CCT TAT GTG 2427
Gly Ile Trp Thr Lys Arg Lys Ile Glu Val Gly Glu Lys Phe Gly Pro Tyr Val 275

GGA GAG CAG AGG TCA AAC CTG AAA GAC CCC AGT TAT GGA TGG GAG ATC TTA GAC 2481
Gly Glu Gln Arg Ser Asn Leu Lys Asp Pro Ser Tyr Gly Trp Glu Ile Leu Asp 293

GAA TTT TAC AAT GTG AAG TTC TGC ATA GAT GCC AGT CAA CCA GAT GTT GGA AGC 2535
Glu Phe Tyr Asn Val Lys Phe Cys Ile Asp Ala Ser Gln Pro Asp Val Gly Ser 311

TGG CTC AAG TAC ATT AGA TTC GCT GGC TGT TAT GAT CAG CAC AAC CTT GTT GCA 2589
Trp Leu Lys Tyr Ile Arg Phe Ala Gly Cys Tyr Asp Gln His Asn Leu Val Ala 329

TGC CAG ATA AAT GAT CAG ATA TTC TAT AGA GTA GTT GCA GAC ATT GCG CCG GGA 2643
Cys Gln Ile Asn Asp Gln Ile Phe Tyr Arg Val Val Ala Asp Ile Ala Pro Gly 347

GAG GAG CTT CTG CTG TTC ATG AAG AGC GAA GAC TAT CCC CAT GAA ACT ATG GCG 2697
Glu Glu Leu Leu Leu Phe Met Lys Ser Glu Asp Tyr Pro His Glu Thr Met Ala 365

CCG GAT ATC CAC GAA GAA CGG CAA TAT CGC TGC GAA GAC TGT GAC CAG CTC TTT 2751
Pro Asp Ile His Glu Glu Arg Gln Tyr Arg Cys Glu Asp Cys Asp Gln Leu Phe 383

GAA TCT AAG GCT GAA CTA GCA GAT CAC CAA AAG TTT CCA TGC AGT ACT CCT CAC 2805
Glu Ser Lys Ala Glu Leu Ala Asp His Gln Lys Phe Pro Cys Ser Thr Pro His 401

TCA GCA TTT TCA ATG GTT GAA GAG GAC TTT CAG CAA AAA CTC GAA AGC GAG AAT 2859
Ser Ala Phe Ser Met Val Glu Glu Asp Phe Gln Gln Lys Leu Glu Ser Glu Asn 419

GAT CTC CAA GAG ATA CAC ACG ATC CAG GAG TGT AAG GAA TGT GAC CAA GTT TTT 2913
Asp Leu Gln Glu Ile His Thr Ile Gln Glu Cys Lys Glu Cys Asp Gln Val Phe 437

CCT GAT TTG CAA AGC CTG GAG AAA CAC ATG CTG TCA CAT ACT GAA GAG AGG GAA 2967
Pro Asp Leu Gln Ser Leu Glu Lys His Met Leu Ser His Thr Glu Glu Arg Glu 455

TAC AAG TGT GAT CAG TGT CCC AAG GCA TTT AAC TGG AAG TCC AAT TTA ATT CGC 3021
Tyr Lys Cys Asp Gln Cys Pro Lys Ala Phe Asn Trp Lys Ser Asn Leu Ile Arg 473

CAC CAG ATG TCA CAT GAC AGT GGA AAG CAC TAT GAA TGT GAA AAC TGT GCC AAG 3075
His Gln Met Ser His Asp Ser Gly Lys His Tyr Glu Cys Glu Asn Cys Ala Lys 491

GTT TTC ACG GAC CCT AGC AAC CTT CAG CGG CAC ATT CGC TCT CAG CAT GTC GGT 3129
Val Phe Thr Asp Pro Ser Asn Leu Gln Arg His Ile Arg Ser Gln His Val Gly 509

GCC CGG GCC CAT GCA TGC CCG GAG TGT GGC AAA ACG TTT GCC ACT TCG TCG GGC 3183
Ala Arg Ala His Ala Cys Pro Glu Cys Gly Lys Thr Phe Ala Thr Ser Ser Gly 527

CTC AAA CAA CAC AAG CAC ATC CAC AGC AGT GTG AAG CCC TTT ATC TGT GAG GTC 3237
Leu Lys Gln His Lys His Ile His Ser Ser Val Lys Pro Phe Ile Cys Glu Val 545

TGC CAT AAA TCC TAT ACT CAG TTT TCA AAC CTT TGC CGT CAT AAG CGC ATG CAT 3291
Cys His Lys Ser Tyr Thr Gln Phe Ser Asn Leu Cys Arg His Lys Arg Met His 563

GCT GAT TGC AGA ACC CAA ATC AAG TGC AAA GAC TGT GGA CAA ATG TTC AGC ACT 3345
Ala Asp Cys Arg Thr Gln Ile Lys Cys Lys Asp Cys Gly Gln Met Phe Ser Thr 581

ACG TCT TCC TTA AAT AAA CAC AGG AGG TTT TGT GAG GGC AAG AAC CAT TTT GCG 3399
Thr Ser Ser Leu Asn Lys His Arg Arg Phe Cys Glu Gly Lys Asn His Phe Ala 599

GCA GGT GGA TTT TTT GGC CAA GGC ATT TCA CTT CCT GGA ACC CCA GCT ATG GAT 3453
Ala Gly Gly Phe Phe Gly Gln Gly Ile Ser Leu Pro Gly Thr Pro Ala Met Asp 617

AAA ACG TCC ATG GTT AAT ATG AGT CAT GCC AAC CCG GGC CTT GCT GAC TAT TTT 3507
Lys Thr Ser Met Val Asn Met Ser His Ala Asn Pro Gly Leu Ala Asp Tyr Phe 635

GGC GCC AAT AGG CAT CCT GCT GGT CTT ACC TTT CCA ACA GCT CCT GGA TTT TCT 3561
Gly Ala Asn Arg His Pro Ala Gly Leu Thr Phe Pro Thr Ala Pro Gly Phe Ser 653

TTT AGC TTC CCT GGT CTG TTT CCT TCC GGC TTG TAC CAC AGG CCT CCT TTG ATA 3615
Phe Ser Phe Pro Gly Leu Phe Pro Ser Gly Leu Tyr His Arg Pro Pro Leu Ile 671

CCT GCT AGT TCT CCT GTT AAA GGA CTA TCA AGT ACT GAA CAG ACA AAC AAA AGT 3669
Pro Ala Ser Ser Pro Val Lys Gly Leu Ser Ser Thr Glu Gln Thr Asn Lys Ser 689

CAA AGT CCC CTC ATG ACA CAT CCT CAG ATA CTG CCA GCT ACA CAG GAT ATT TTG 3723
Gln Ser Pro Leu Met Thr His Pro Gln Ile Leu Pro Ala Thr Gln Asp Ile Leu 707

AAG GCA CTA TCT AAA CAC CCA TCT GTA GGG GAC AAT AAG CCA GTG GAG CTC CAG 3777
Lys Ala Leu Ser Lys His Pro Ser Val Gly Asp Asn Lys Pro Val Glu Leu Gln 725

CCC GAG AGG TCC TCT GAA GAG AGG CCC TTT GAG AAA ATC AGT GAC CAG TCA GAG 3831
Pro Glu Arg Ser Ser Glu Glu Arg Pro Phe Glu Lys Ile Ser Asp Gln Ser Glu 743

AGT AGT GAC CTT GAT GAT GTC AGT ACA CCA AGT GGC AGT GAC CTG GAA ACA ACC 3885
Ser Ser Asp Leu Asp Asp Val Ser Thr Pro Ser Gly Ser Asp Leu Glu Thr Thr 761

TCG GGC TCT GAT CTG GAA AGT GAC ATT GAA AGT GAT AAA GAG AAA TTT AAA GAA 3939
Ser Gly Ser Asp Leu Glu Ser Asp Ile Glu Ser Asp Lys Glu Lys Phe Lys Glu 779

AAT GGT AAA ATG TTC AAA GAC AAA GTA AGC CCT CTT CAG AAT CTG GCT TCA ATA 3993
Asn Gly Lys Met Phe Lys Asp Lys Val Ser Pro Leu Gln Asn Leu Ala Ser Ile 797

AAT AAT AAG AAA GAA TAC AGC AAT CAT TCC ATT TTC TCA CCA TCT TTA GAG GAG 4047
Asn Asn Lys Lys Glu Tyr Ser Asn His Ser Ile Phe Ser Pro Ser Leu Glu Glu 815

CAG ACT GCG GTG TCA GGA GCT GTG AAT GAT TCT ATA AAG GCT ATT GCT TCT ATT 4101
Gln Thr Ala Val Ser Gly Ala Val Asn Asp Ser Ile Lys Ala Ile Ala Ser Ile 833

GCT GAA AAA TAC TTT GGT TCA ACA GGA CTG GTG GGG CTG CAA GAC AAA AAA GTT 4155
Ala Glu Lys Tyr Phe Gly Ser Thr Gly Leu Val Gly Leu Gln Asp Lys Lys Val 851

GGA GCT TTA CCT TAC CCT TCC ATG TTT CCC CTC CCA TTT TTT CCA GCA TTC TCT 4209
Gly Ala Leu Pro Tyr Pro Ser Met Phe Pro Leu Pro Phe Phe Pro Ala Phe Ser 869

CAA TCA ATG TAC CCA TTT CCT GAT AGA GAC TTG AGA TCG TTA CCT TTG AAA ATG 4263
Gln Ser Met Tyr Pro Phe Pro Asp Arg Asp Leu Arg Ser Leu Pro Leu Lys Met 887

GAA CCC CAA TCA CCA GGT GAA GTA AAG AAA CTG CAG AAG GGC AGC TCT GAG TCC 4317
Glu Pro Gln Ser Pro Gly Glu Val Lys Lys Leu Gln Lys Gly Ser Ser Glu Ser 905

CCC TTT GAT CTC ACC ACT AAG CGA AAG GAT GAG AAG CCC TTG ACT CCA GTC CCC 4371
Pro Phe Asp Leu Thr Thr Lys Arg Lys Asp Glu Lys Pro Leu Thr Pro Val Pro 923

TCC AAG CCT CCA GTG ACA CCT GCC ACA AGC CAA GAC CAG CCC CTG GAT CTA AGT 4425
Ser Lys Pro Pro Val Thr Pro Ala Thr Ser Gln Asp Gln Pro Leu Asp Leu Ser 941

ATG GGC AGT AGG AGT AGA GCC AGT GGG ACA AAG CTG ACT GAG CCT CGA AAA AAC 4479
Met Gly Ser Arg Ser Arg Ala Ser Gly Thr Lys Leu Thr Glu Pro Arg Lys Asn 959

CAC GTG TTT GGG GGA AAA AAA GGA AGC AAC GTC GAA TCA AGA CCT GCT TCA GAT 4533
His Val Phe Gly Gly Lys Lys Gly Ser Asn Val Glu Ser Arg Pro Ala Ser Asp 977

GGT TCC TTG CAG CAT GCA AGA CCC ACT CCT TTC TTT ATG GAC CCT ATT TAC AGA 4587
Gly Ser Leu Gln His Ala Arg Pro Thr Pro Phe Phe Met Asp Pro Ile Tyr Arg 995

GTA GAG AAA AGA AAA CTA ACT GAC CCA CTT GAA GCT TTA AAA GAG AAA TAC TTG 4641
Val Glu Lys Arg Lys Leu Thr Asp Pro Leu Glu Ala Leu Lys Glu Lys Tyr Leu 1013

AGG CCT TCT CCA GGA TTC TTG TTT CAC CCA CAA ATG TCA GCT ATT GAA AAC ATG 4695
Arg Pro Ser Pro Gly Phe Leu Phe His Pro Gln Met Ser Ala Ile Glu Asn Met 1031

GCA GAA AAG CTA GAG AGC TTC AGT GCC CTG AAA CCT GAG GCC AGT GAG CTC TTA 4749
Ala Glu Lys Leu Glu Ser Phe Ser Ala Leu Lys Pro Glu Ala Ser Glu Leu Leu 1049

CAG TCA GTG CCC TCT ATG TTC AAC TTC AGG GCG CCT CCC AAT GCC CTG CCA GAG 4803
Gln Ser Val Pro Ser Met Phe Asn Phe Arg Ala Pro Pro Asn Ala Leu Pro Glu 1067

AAC CTT CTG CGG AAG GGA AAG GAG CGC TAT ACC TGC AGA TAC TGT GGC AAG ATT 4857
Asn Leu Leu Arg Lys Gly Lys Glu Arg Tyr Thr Cys Arg Tyr Cys Gly Lys Ile 1085

TTT CCA AGG TCT GCA AAC CTA ACA CGG CAC TTG AGA ACC CAC ACA GGA GAG CAG 4911
Phe Pro Arg Ser Ala Asn Leu Thr Arg His Leu Arg Thr His Thr Gly Glu Gln 1103

CCT TAC AGA TGC AAA TAC TGT GAC AGA TCA TTT AGC ATA TCT TCT AAC TTG CAA 4965
Pro Tyr Arg Cys Lys Tyr Cys Asp Arg Ser Phe Ser Ile Ser Ser Asn Leu Gln 1121

AGG CAT GTT CGC AAC ATC CAC AAT AAA GAG AAG CCA TTT AAG TGT CAC TTA TGT 5019
Arg His Val Arg Asn Ile His Asn Lys Glu Lys Pro Phe Lys Cys His Leu Cys 1139

GAT AGG TGT TTT GGT CAA CAA ACC AAT TTA GAC AGA CAC CTA AAG AAA CAT GAG 5073
Asp Arg Cys Phe Gly Gln Gln Thr Asn Leu Asp Arg His Leu Lys Lys His Glu 1157

AAT GGG AAC ATG TCC GGT ACA GCA ACA TCG TCG CCT CAT TCT GAA CTG GAA AGT 5127
Asn Gly Asn Met Ser Gly Thr Ala Thr Ser Ser Pro His Ser Glu Leu Glu Ser 1175

ACA GGT GCG ATT CTG GAT GAC AAA GAA GAT GCT TAC TTC ACA GAA ATT CGA AAT 5181
Thr Gly Ala Ile Leu Asp Asp Lys Glu Asp Ala Tyr Phe Thr Glu Ile Arg Asn 1193

TTC ATT GGG AAC AGC AAC CAT GGC AGC CAA TCT CCC AGG AAT GTG GAG GAG AGA 5235
Phe Ile Gly Asn Ser Asn His Gly Ser Gln Ser Pro Arg Asn Val Glu Glu Arg 1211

ATG AAT GGC AGT CAT TTT AAA GAT GAA AAG GCT TTG GTG ACC AGT CAA AAT TCA 5289
Met Asn Gly Ser His Phe Lys Asp Glu Lys Ala Leu Val Thr Ser Gln Asn Ser 1229

GAC TTG CTG GAT GAT GAA GAA GTT GAA GAT GAG GTG TTG TTA GAT GAG GAG GAT 5343
Asp Leu Leu Asp Asp Glu Glu Val Glu Asp Glu Val Leu Leu Asp Glu Glu Asp 1247

GAA GAC AAT GAT ATT ACT GGA AAA ACA GGA AAG GAA CCA GTG ACA AGT AAT TTA 5397
Glu Asp Asn Asp Ile Thr Gly Lys Thr Gly Lys Glu Pro Val Thr Ser Asn Leu 1265

CAT GAA GGA AAC CCT GAG GAT GAC TAT GAA GAA ACC AGT GCC CTG GAG ATG AGT 5451
His Glu Gly Asn Pro Glu Asp Asp Tyr Glu Glu Thr Ser Ala Leu Glu Met Ser 1283

TGC AAG ACA TCC CCA GTG AGG TAT AAA GAG GAA GAA TAT AAA AGT GGA CTT TCT 5505
Cys Lys Thr Ser Pro Val Arg Tyr Lys Glu Glu Glu Tyr Lys Ser Gly Leu Ser 1301

GCT CTA GAT CAT ATA AGG CAC TTC ACA GAT AGC CTC AAA ATG AGG AAA ATG GAA 5559
Ala Leu Asp His Ile Arg His Phe Thr Asp Ser Leu Lys Met Arg Lys Met Glu 1319

GAT AAT CAA TAT TCT GAA GCT GAG CTG TCT TCT TTT AGT ACT TCC CAT GTG CCA 5613
Asp Asn Gln Tyr Ser Glu Ala Glu Leu Ser Ser Phe Ser Thr Ser His Val Pro 1337

GAG GAA CTT AAG CAG CCG TTA CAC AGA AAG TCC AAA TCG CAG GCA TAT GCT ATG 5667
Glu Glu Leu Lys Gln Pro Leu His Arg Lys Ser Lys Ser Gln Ala Tyr Ala Met 1355

ATG CTG TCA CTG TCT GAC AAG GAG TCC CTC CAT TCT ACA TCC CAC AGT TCT TCC 5721
Met Leu Ser Leu Ser Asp Lys Glu Ser Leu His Ser Thr Ser His Ser Ser Ser 1373

AAC GTG TGG CAC AGT ATG GCC AGG GCT GCG GCG GAA TCC AGT GCT ATC CAG TCC 5775
Asn Val Trp His Ser Met Ala Arg Ala Ala Ala Glu Ser Ser Ala Ile Gln Ser 1391

ATA AGC CAC GTA TGA CGTTA TCAAG GTTGA CCAGA GTGGG ACCAA GTCCA ACAGT 5830
Ile Ser His Val 1395

AGCAT GGCTC TTTCA TATAG GACTA TTTAC AAGAC TGCTG AGCAG AATGC CTTAT AAACC 5890
TGCAG GGTCA CTCAT CTAAA GTCTA GTGAC CTTAA ACTGA ATGAT TT(A)n 5937

Fig. 2. Nucleotide sequence and deduced amino acid sequence of the AML1−EVI-1 fusion cDNA in the t(3;21) translocation. Nucleotides and amino acids are numbered at the end of each line. An arrow indicates the boundary between AML1 and EVI-1 in the t(3;21) translocation. Underlined nucleotide sequence represents the CHR3 region that is derived from the 5′ non-coding exon of human *EVI-1* gene. The initiation codon of human *EVI-1* in the third exon is double-underlined. Asterisks above the sequence show nucleotide changes in comparison with the published human EVI-1 sequence. An open triangle indicates the insertion site of 27 bp found in the published human EVI-1 sequence, but not in the mouse sequence.

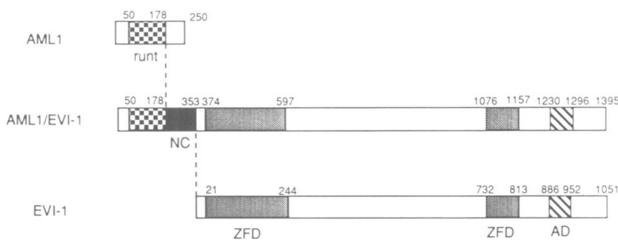

Fig. 3. Schematic structures of the AML1, AML1 – EVI-1 and EVI-1 gene products. As a consequence of the t(3;21) translocation, the 178 amino acid N-terminal portion of AML1, containing the runt homology domain, is fused to the entire of EVI-1, containing the 175 amino acids derived from the sequences of 5′ non-coding exons and the 1042 amino acids derived from the coding sequences. Checked boxes indicate a runt homology domain in AML1. The sequences derived from non-coding exons of EVI-1 (NC; non-coding) in AML1 – EVI-1 are shaded. Two zinc finger domains (ZFD) in EVI-1 are dotted. An acidic domain (AD) in EVI-1 is hatched. Amino acid positions are indicated.

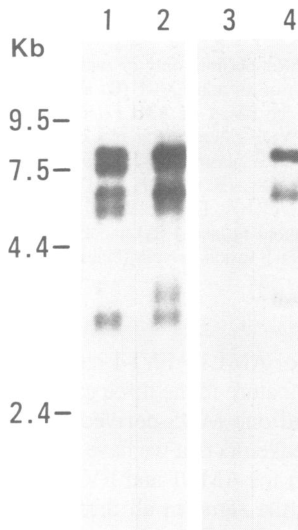

Fig. 4. Expression of *AML1* and *AML1 – EVI-1* fusion genes in KU812 and SKH1 cells. One μg of poly(A)$^+$ mRNAs isolated from KU812 (lanes 1 and 3) and SKH1 (lanes 2 and 4) cells were electrophoresed in a formaldehyde denaturing agarose gel and transferred to a nylon filter. The filter was hybridized with the AML1 (lanes 1 and 2) or EVI-1 (lanes 3 and 4) probe. The RNA markers are aligned.

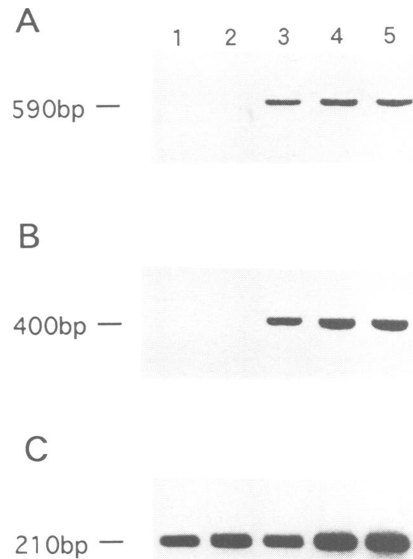

Fig. 5. Analysis of AML1 – EVI-1 fusions in t(3;21)-carrying leukemias by RNA-based PCR. RNA-based nested PCR was performed using the outer and then the inner sets of primers described in Materials and methods. cDNA synthesis was specifically primed with an oligonucleotide complementary to the AML1 (**C**) or EVI-1 (**A** and **B**) sequence just downstream of the fusion point. The PCR products with AML1 – EVI-1-specific primers (A and B) and AML1-specific primers (C) are shown with the predicted size by ethidium bromide staining. Lanes 1 and 2, normal control; lane 3, patient 3 (MDS-derived leukemia); lane 4, patient 2 (blastic crisis of CML); lane 5, patient 1 (blastic crisis of CML). The expected sizes of the products are 590, 400 and 210 bp for A, B and C, respectively.

observed in SKH1 cells. In contrast, no transcripts were detected in KU812 cells with the EVI-1 probe (described in Materials and methods). Two major transcripts of 8.2 and 7.0 kb in size, which were detected only in SKH1 cells and hybridized with both AML1 and EVI-1 probes, should have originated from the *AML1 – EVI-1* fusion gene. Thus, *AML1 – EVI-1* chimeric gene is transcribed to fusion mRNAs of at least two species.

AML1 – EVI-1 fusion is a consistent feature of the t(3;21) translocation

To determine whether the formation of AML1 – EVI-1 chimeric mRNAs is consistent in leukemic cells carrying the t(3;21) translocation, we investigated two further cases with the t(3;21) translocation (one with blastic crisis of CML and the other with MDS-derived leukemia) for the presence of the fusion mRNAs by RNA-based PCR. The primer sets used are described in Materials and methods. cDNA synthesis was specifically primed with an oligonucleotide complementary to the EVI-1 sequence just downstream of the fusion point. The PCR products amplified with mRNAs from SKH1 cells showed the size expected from the sequence of the AML1 – EVI-1 fusion cDNA. The amplified products were the same size as with mRNAs from SKH1 cells, and were observed with mRNAs from these two other cases with the t(3;21) translocation (Figure 5), indicating that mRNA sequences of the two cases spanning the AML1 – EVI-1 fusion point are identical to those of SKH1 cells. Amplified fusion cDNAs were detected in all three cases by Southern analysis using a radiolabeled probe with CHR3 sequence (data not shown), which is specific for the AML1 – EVI-1 fusion cDNA. These data suggest that the formation of AML1 – EVI-1 fusion mRNA is a common phenomenon in the t(3;21)-carrying leukemia and that the fusion points should be the same in the mRNA sequences.

An antiserum against AML1 or EVI-1 detects an AML1 – EVI-1 fusion protein in SKH1 cells

Polyclonal antibodies to AML1 and EVI-1 were raised in rabbits against synthetic peptides (106 – 125 a.a.) of AML1 and maltose-binding protein fusion of the N-terminal 255 amino acids of EVI-1 (T.Tanaka *et al.*, manuscript submitted), respectively. An anti-AML1 antibody detected the AML1 protein of 70 kDa in HL60 cells (Figure 6A), which are known to express AML1 mRNAs (Miyoshi *et al.*, 1993). Both (i) Cos1 cells transfected with the SRα expression plasmid (Toyoshima *et al.*, 1993) containing the full length of AML1 – EVI-1 cDNA and (ii) SKH1 cells

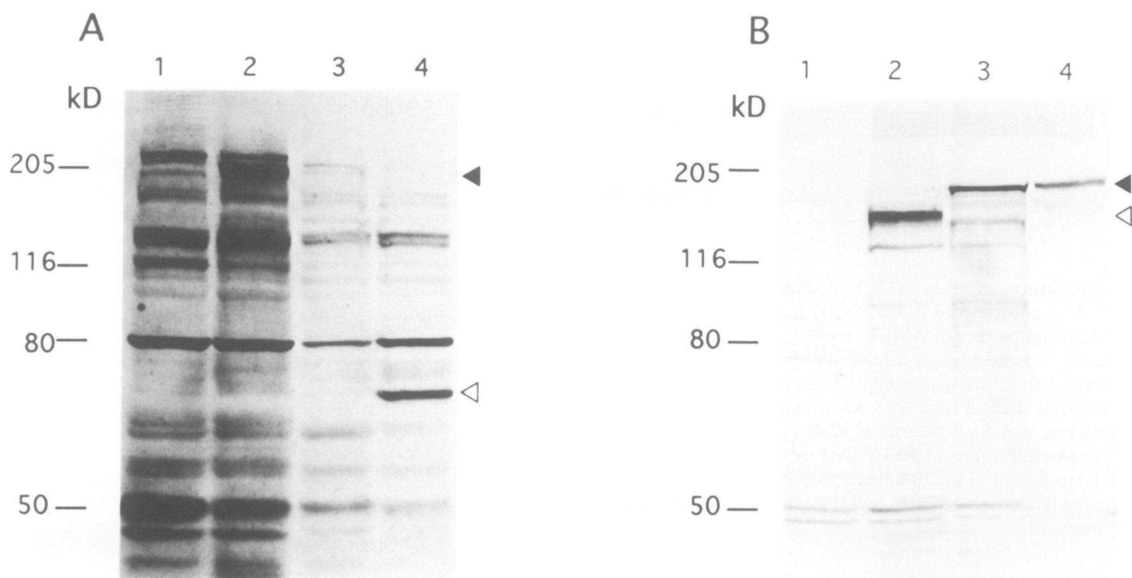

Fig. 6. Western blot analysis of whole cell extracts in Cos1 or NIH3T3 cells transfected with the SRα plasmid only or with the SRα plasmid containing the EVI-1 or AML1−EVI-1 cDNA, SKH1 cells and HL60 cells with an anti-AML1 (**A**) or an anti-EVI-1 (**B**) antibody. Proteins from Cos1 or NIH3T3 cells transfected with the SRα plasmid only, or with the SRα plasmid containing the EVI-1 or AML1−EVI-1 cDNA, SKH1 cells and HL60 cells, were fractioned on a 7.5% SDS−polyacrylamide gel and electrotransferred onto PVDF filters. After blocking, filters were incubated with an anti-AML1 (A) or an anti-EVI-1 (B) antibody. (A) Lane 1, Cos1 cells transfected with the SRα plasmid only; lane 2, Cos1 cells transfected with the SRα plasmid containing the AML1−EVI-1 cDNA; lane 3, SKH1 cells; and lane 4, HL60 cells. (B) Lane 1, NIH3T3 cells transfected with the SRα plasmid only; lane 2, NIH3T3 cells transfected with the SRα plasmid containing the EVI-1 cDNA; lane 3, NIH3T3 cells transfected with the SRα plasmid containing the AML1−EVI-1 cDNA; and lane 4, SKH1 cells. Normal AML1 protein in lane 4 (A) and EVI-1 protein expressed in NIH3T3 cells in lane 2 (B) are marked by open triangles. The molecular weight of the AML1−EVI-1 fusion protein (lanes 2 and 3 in A; lanes 3 and 4 in B) indicated by closed triangles is ~180 kDa.

showed the 180 kDa proteins, neither observed in HL60 cells. On the other hand, an anti-EVI-1 antibody detected the EVI-1 150 kDa protein in NIH3T3 cells transfected with the SRα expression plasmid containing the full length of EVI-1 cDNA (Figure 6B). The 180 kDa proteins were observed with an anti-EVI-1 antibody in SKH1 cells as well as in NIH3T3 cells transfected with the SRα expression plasmid containing the full length of AML1−EVI-1 cDNA. None of these proteins was detected using a pre-immune serum for AML1 or EVI-1 (data not shown). Thus the 180 kDa protein, detected in SKH1 cells by Western analysis both with anti-AML1 and anti-EVI-1 antibodies, should be the AML1−EVI-1 fusion protein. These observations demonstrate that the formation of the *AML1−EVI-1* fusion gene leads to the expression of the fusion protein.

Discussion

We have demonstrated in this study that the t(3;21)(q26;q22) chromosomal translocation results in synthesis of AML1−EVI-1 fusion transcripts: mainly two species of 8.2 and 7.0 kb, that cross the t(3;21) breakpoint and code for a fusion protein of 180 kDa with the potential properties of a chimeric transcription factor. The t(3;21) translocation is found usually in blastic crisis of CML (Rubin *et al.*, 1987; Coyle and Najfeld, 1988), leukemia developed from MDS (Rubin *et al.*, 1990) or secondary leukemia from hematopoietic proliferative diseases (Dastugue *et al.*, 1990), but rarely in *de novo* acute myelocytic leukemia (Rubin *et al.*, 1990). This raises the possibility that the molecular event underlying the t(3;21) translocation has a critical role in the progression from a preleukemic to a leukemic state. The fusion transcripts of AML1−EVI-1 have been demonstrated to be a consistent feature in the three cases (two blastic crisis cases of CML and one MDS-derived leukemia case) with t(3;21)-carrying leukemia that we have examined. Moreover, the point of fusion for AML1 and EVI-1 mRNA sequences was shown to be the same in all three cases with t(3;21)-carrying leukemia. Our data suggest that the translocation-induced fusion protein, AML1−EVI-1, is important in the leukemic progression of t(3;21)-carrying hematopoietic stem cell disorders.

The *AML1* gene has been isolated as a gene which is located at the translocation breakpoint of chromosome 21 in the t(8;21) translocation specifically found in acute myelocytic leukemia (M2 according to the French−American−British classification) (Miyoshi *et al.*, 1991). The AML1 product has a runt homology domain that is highly homologous to the product of *Drosophila* segmentation gene runt (Daga *et al.*, 1992) and the α subunit of polyomavirus enhancer binding protein 2 (PEBP2/PEA2) (Bae *et al.*, 1993). The *Drosophila* gene runt encodes a nuclear protein regulating the expression of other pair-rule genes, although it does not contain any identifiable transcription factor motif (Kania *et al.*, 1990). The runt homology domain is thought to include a DNA binding domain. The *AML1* gene seems to be constitutively expressed in multiple hematopoietic lineages because human leukemic cell lines derived from myeloid, B lymphoid and T lymphoid cells show normal AML1 transcripts at different expression patterns (Miyoshi *et al.*, 1993). Thus AML1 could play an important role in hematopoietic cell growth and/or differentiation through transcriptional regulation. It is interesting that the AML1 protein is disrupted at the end of the runt homology domain

by the t(3;21) translocation. Thus the AML1–EVI-1 fusion protein contains a DNA binding domain from the AML1 protein.

The truncation of the carboxy portion following the runt homology domain in AML1 itself, or replacement of it by the entire EVI-1, could change the function of AML1 as a transcription factor in several mechanisms. (i) The chimeric transcription factor would lose its ability to bind target genes for AML1. (ii) The transcription factor would lose the transactivation ability for target genes, even if it would bind the same set of genes that AML1 does. It is possible that the DNA binding activity of AML1–EVI-1 could be increased in comparison with AML1 itself. In this case, the chimeric transcription factors might work against AML1 in a dominant negative manner by competing for target DNA sites. (iii) It would change a spectrum of target genes in comparison with AML1 under the influence of EVI-1 function.

The *EVI-1* gene was originally identified as a common site of viral integration in murine myeloid leukemias (Morishita *et al.*, 1988) and is frequently expressed in murine myeloid leukemias (Morishita *et al.*, 1988; Bartholomew *et al.*, 1989; Bartholomew and Ihle, 1991). Expression of the *EVI-1* gene is also observed in human acute myelocytic leukemias, especially with 3q26 chromosomal abnormalities such as the t(3;3)(q21;q26) and ins(3;3)(q21q25q26) (Morishita *et al.*, 1992b). EVI-1 is a zinc finger protein with 10 zinc finger motifs located in two domains and an acidic domain located C-terminal to the second set of zinc fingers, indicating that EVI-1 is a site-specific DNA binding protein (Morishita *et al.*, 1988). Although the consensus sequence for EVI-1 binding has been reported (Perkins *et al.*, 1991), the target genes of EVI-1 remain unknown. Thus the exact mechanism by which expression of EVI-1 is involved in leukemogenesis is still unresolved. However, expression of EVI-1 in an IL-3-dependent murine myeloid leukemia cell line is experimentally shown to lose its ability to differentiate to granulocytes in response to granulocyte colony-stimulating factor (Morishita *et al.*, 1992a). The EVI-1 could be involved in leukemogenesis by a mechanism of differentiation block. The AML1–EVI-1 fusion protein contains the entire EVI-1 protein, including two zinc finger domains and an acidic domain, and should be expressed under the control of the *AML1* gene promoter. Normal EVI-1 transcripts have been reported as not being expressed in murine bone marrow cells (Morishita *et al.*, 1990b). The abnormal expression of EVI-1 protein using the AML1 promoter in human hematopoietic cells could contribute to the development of leukemia.

CML and MDS are hematopoietic stem cell disorders characterized by hypercellular bone marrow with myeloid maturation. An event that induces the appearance of additional chromosomal abnormalities or mutations of oncogenes or tumor suppressor genes seems to cause clonal evolution and the development of chronic leukemia to the acute phase. Our results support the hypothesis that the t(3;21) translocation, which results in fusion of the whole of EVI-1 to AML1, could have at least a role in the development of CML or MDS to an acute leukemic phase by blocking the ability of cells to differentiate. Alteration in transcriptional control has been proposed to be involved in oncogenesis (Cleary, 1991; Rabbitts, 1991). In acute leukemias, chimeric transcription factors like E2A–PBX1 in the t(1;19) (Kamps *et al.*, 1990, 1991; Nourse *et al.*, 1990), PML–RARα in the t(15;17) (de The *et al.*, 1991; Kakizuka *et al.*, 1991) and AML1–MTG8 or AML1–ETO in the t(8;21) (Miyoshi *et al.*, 1991, 1993; Nucifora *et al.*, 1993) have been demonstrated to play an important role in leukemogenesis. We have demonstrated that the same process, the formation of chimeric transcription factors by chromosomal translocation, also causes the transformation of chronic stem cell disease to acute leukemia. Although the target genes remain to be found for both AML1 and EVI-1 as transcription factors, there could be two possible mechanisms for the transformation by the AML1–EVI-1 fusion protein. The first is that the AML1–EVI-1 fusion protein could lose its transactivation ability or work on a set of genes different to AML1 or EVI-1 which are critical to the process of proliferation and differentiation of hematopoietic cells. The second relates to the abnormal expression of the EVI-1 protein under the control of the AML1 promoter. The identification of target genes for AML1 and EVI-1 should help to clarify the issue. The AML1–EVI-1 fusion protein is a potentially bifunctional transcription factor with multiple DNA binding properties. Further biological studies and functional analyses of the AML1–EVI-1 fusion protein should provide useful insights into the mechanism of oncogenesis mediated by the alteration of transcriptional control.

Materials and methods

Cell line

Human leukemic cell line, SKH1, was established from a patient with megakaryoblastic crisis of chronic myelocytic leukemia carrying the t(3;21) and t(9;22) translocations as the only cytogenetic abnormalities. The establishment and characterization of HL60, KU812 and HEC-1 cell lines have been described previously (Collins *et al.*, 1977; Morinaga *et al.*, 1983; Kishi, 1985).

Probes

The AML1 cDNA probe is the C6E6H2 probe described previously (Miyoshi *et al.*, 1991). A 2.0 kb fragment of 3′ terminal human EVI-1 cDNA sequence (Morishita *et al.*, 1990a) was used as the EVI-1 cDNA probe.

cDNA cloning and sequencing

An oligo(dT)-primed cDNA library was prepared by converting poly(A)$^+$ mRNAs of SKH1 cells into cDNAs using RNaseH$^-$ MMLV reverse transcriptase (SuperScript, Bethesda Research Laboratories) and oligo(dT) primers and ligation into the Lambda Zap II vector (Stratagene). The library was screened with random priming-labeled AML1 probe and two potentially full-length cDNA clones (clones 4 and 10) were isolated. Nested deletions of clone 4 were generated using exonuclease III according to the manufacturer's instructions (Pharmacia) and sequenced by the dideoxy chain termination method using Sequenase (USB). Gaps and ambiguities were resolved by using oligonucleotide primers.

PCR analysis of fusion transcripts

RNA-based PCR was performed by specifically priming cDNA synthesis, followed by nested PCR. PCR primers used in the analysis of the AML1 and AML1–EVI-1 fusion transcripts were as follows. For the experimental results shown in Figure 5A, outer set, 5′-AAGTCGCCACCTACCACA-GA-3′ and 5′-CCGGCGCCATAGTTTCATGG-3′; inner set, 5′-GCCATC-AAAATCACAGTGCA-3′ and 5′-GGATAGTCTTCGCTCTTCAT-3′; for Figure 5B, outer set, same as in Figure 5A; inner set, 5′-GCCATCAAA-ATCACAGTGCA-3′ and 5′-TTGTAAAATTCGTCTAAGAT-3′; and for Figure 5C, outer set, 5′-AAGTCGCCACCTACCACAGA-3′ and 5′-CCT-GCATCTGACTCTGAGGC-3′; inner set, 5′-GCCATCAAAATCACA-GTGCA-3′ and 5′-TGAGGGTTAAAGGCAGTA-3′. PCR amplification was performed using a thermal cycler. The program for each amplification was as follows: 1 min at 94°C, 2 min at 55°C and 3 min at 72°C, entrained for 25 cycles; 7 min at 72°C for final extension. The PCR product was size-fractioned by electrophoresis in a 3% agarose gel.

Western analysis of fusion proteins

Polyclonal antibodies to AML1 and EVI-1 were raised in rabbits against synthetic peptide (106–125 a.a.) of AML1 and maltose-binding protein fusion of the N-terminal 255 amino acids of EVI-1 (T.Tanaka *et al.*, manuscript submitted), respectively. A 4.1 kb *Eco*RI fragment from clone 4 containing the entire open reading frame of the AML1–EVI-1 sequence was subcloned into the SRα expression plasmid and the resulting plasmid was transiently transfected into Cos1 cells by the DEAE–dextran method (Sambrook *et al.*, 1989) or NIH3T3 cells by the calcium phosphate–DNA precipitate method (Ausubel *et al.*, 1992). The SRα expression plasmid containing a 3.3 kb *Eco*RI fragment with the entire open reading frame of the EVI-1 sequence was also transiently transfected into NIH3T3 cells by the calcium phosphate–DNA precipitate method. Cells indicated were solubilized in a lysis buffer (10 mM Tris–HCl pH 7.4, 1 mM EDTA, 500 kallikrein inhibitor U/ml aprotinin, 1 mM phenylmethylsulfonyl fluoride, 1 mM sodium orthovanadate, 50 mM sodium fluoride and 1% NP-40), fractioned on a 7.5% SDS–polyacrylamide gel and transferred to a PVDF membrane (Millipore). The filter replica was blocked with 10% Skim Milk, treated with an anti-AML1 (1:100) or an anti-EVI-1 antibody (1:700), washed and incubated with goat anti-rabbit IgG antibody coupled to alkaline phosphatase (Promega, 1:5000). Bands were visualized by incubation with nitroblue tetrazolium-bromochloroindolyl phosphate (Promega).

Acknowledgements

We thank Dr K.Morishita for providing the EVI-1 cDNA probe, Drs Y.Harada and M.Nagai for providing case samples, and Dr K.Inamori, Ms N.Yokoyama, Dr K.Ohgami and Dr T.Sudou for special technical assistance. This work was supported in part by Grants-in-Aid for Cancer Research from the Ministry of Health and Welfare and from the Ministry of Education, Science and Culture of Japan.

References

Ahuja,H., Bar-Eli,M., Arlin,Z., Advani,S., Allen,S.L., Goidman,J., Snyder,D., Foti,A. and Cline,M. (1991) *J. Clin. Invest.*, **87**, 2042–2047.
Alimena,G., De Cuia,M.R., Diverio,D., Gastaldi,R. and Nanni,M. (1987) *Cancer Genet. Cytogenet.*, **26**, 39–50.
Ausubel,F.M., Brent,R., Kingston,R.E., Moore,D.D., Seidman,J.G., Smith,J.A. and Struhl,K. (1992) *Current Protocols in Molecular Biology*. Greene Publishing Associates and Wiley Interscience, NY, pp. 9.1.1–9.1.7.
Bae,S.C., Yamaguchi-Iwai,Y., Ogawa,E., Maruyama,M., Inuzuka,M., Kagoshima,H., Shigesada,K., Satake,M. and Ito,Y. (1993) *Oncogene*, **8**, 809–814.
Bartholomew,C. and Ihle,J.N. (1991) *Mol. Cell. Biol.*, **11**, 1820–1828.
Bartholomew,C., Morishita,K., Askew,D., Buchberg,A., Jenkins,N.A., Copeland,N.G. and Ihle,J.N. (1989) *Oncogene*, **4**, 529–534.
Chen,Z., Morgan,R., Baer,M.R., Ligorsky,R. and Sandberg,A.A. (1991) *Cancer Genet. Cytogenet.*, **57**, 153–159.
Cleary,M.L. (1991) *Cell*, **66**, 619–622.
Collins,S.J., Gallo,R.C. and Gallagher,R.E. (1977) *Nature*, **270**, 347–349.
Coyle,T. and Najfeld,V. (1988) *Am. J. Hematol.*, **27**, 56–59.
Daga,A., Tighe,J.E. and Calabi,F. (1992) *Nature*, **356**, 484.
Daley,G.Q., van Etten,R. and Baltimore,D. (1990) *Science*, **247**, 824–830.
Dastugue,N., Pris,J. and Colombies,P. (1990) *Cancer Genet. Cytogenet.*, **44**, 275–276.
de The,H., Lavau,C., Marchio,A., Chomienne,C., Degos,L. and Dejean,A. (1991) *Cell*, **66**, 675–684.
Feinstein,E., Cimino,G., Gale,R.P. and Canaani,E. (1992) *Leukemia*, Suppl. **1**, 37–43.
Fialkow,P.J., Jacobsen,R.J. and Papayannopoulou,T. (1977) *Am. J. Med.*, **63**, 125–130.
Foti,A., Ahuja,H.G., Allen,S.L., Koduru,P., Schuster,M.W., Schulman,P., Bar-Eli,M. and Cline,M.J. (1991) *Blood*, **77**, 2441–2444.
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.S., Dmitrovsky,E. and Evans,R.M. (1991) *Cell*, **66**, 663–674.
Kamps,M.P., Murre,C., Sun,X.-H. and Baltimore,D. (1990) *Cell*, **60**, 547–555.
Kamps,M.P., Look,A.T. and Baltimore,D. (1991) *Genes Dev.*, **5**, 358–368.
Kania,M.A., Bonner,A.S., Duffy,J.B. and Gergen,J.P. (1990) *Genes Dev.*, **4**, 1701–1713.
Kishi,K. (1985) *Leuk. Res.*, **9**, 381–390.
Konopka,J.B., Watanabe,S.M. and Witte,O.N. (1984) *Cell*, **37**, 1035–1042.
Miyoshi,H., Shimizu,K., Kozu,T., Maseki,N., Kaneko,Y. and Ohki,M. (1991) *Proc. Natl Acad. Sci. USA*, **88**, 10431–10434.
Miyoshi,H., Kozu,T., Shimizu,K., Enomoto,K., Maseki,N., Kaneko,Y., Kamada,N. and Ohki,M. (1993) *EMBO J.*, **12**, 2715–2721.
Morinaga,M., Yonehara,S., Tomita,Y. and Kuramoto,T. (1983) *Int. J. Cancer*, **31**, 21–28.
Morishita,K., Parker,D.S., Mucenski,M.L., Jenkins,N.A., Copeland,N.G. and Ihle,J.N. (1988) *Cell*, **54**, 831–840.
Morishita,K., Parganas,E., Douglass,E.C. and Ihle,J.N. (1990a) *Oncogene*, **5**, 963–971.
Morishita,K., Parganas,E., Parham,D.M., Matsugi,T. and Ihle,J.N. (1990b) *Oncogene*, **5**, 1419–1423.
Morishita,K., Parganas,E., Matsugi,T. and Ihle,J.N. (1992a) *Mol. Cell. Biol.*, **12**, 183–189.
Morishita,K., Parganas,E., Willman,C.L., Whittaker,M.H., Drabkin,H., Oval,J., Taetle,R., Valentine,M.B. and Ihle,J.N. (1992b) *Proc. Natl Acad. Sci. USA*, **89**, 3937–3941.
Nichols,J. and Nimer,S.D. (1992) *Blood*, **80**, 2953–2963.
Nourse,J., Mellentin,J.D., Galili,N., Wilkinson,J., Stanbridge,E., Smith,S.D. and Cleary,M.L. (1990) *Cell*, **60**, 535–545.
Nucifora,G., Birn,D.J., Erickson,P., Gao,J., LeBeau,M.M., Drabkin,H.A. and Rowley,J.D. (1993) *Blood*, **81**, 883–888.
Perkins,A.S., Fishel,R., Jenkins,N.A. and Copeland,N.G. (1991) *Mol. Cell. Biol.*, **11**, 2665–2674.
Rabbitts,T.H. (1991) *Cell*, **67**, 641–644.
Rowley,J.D. (1973) *Nature*, **243**, 290–293.
Rubin,C.M., Larson,R.A., Bitter,M.A., Carrino,J.J., Le Beau,M.M., Diaz,M.O. and Rowley,J.D. (1987) *Blood*, **70**, 1338–1342.
Rubin,C.M., Larson,R.A., Anastasi,J., Winter,J.N., Thangavelu,M., Vardiman,J.W., Rowley,J.D. and Le Beau,M.M. (1990) *Blood*, **76**, 2594–2598.
Sambrook,J., Fritsch,E.F. and Maniatis,T. (1989) In Ford,N., Nolan,C. and Ferguson,N. (eds), *Molecular Cloning, A Laboratory Manual*. Cold Spring Harbor Laboratory Press, NY, pp. 16.41–16.46.
Sandberg,A.A. (1980). In Sandberg,A.A. (ed.), *The Chromosomes in Human Cancer and Leukemia*. Elsevier, New York, pp.183–261.
Schneider,N.R., Bowman,W.P. and Frenkel,E.P. (1991) *Ann. Genet.*, **34**, 256–263.
Solomon,E., Borrow,J. and Goddard,A.D. (1991) *Science*, 254, 1153–1160.
Toyoshima,H., Kozutumi,H., Maru,Y., Hagiwara,K., Furuya,A., Mioh,H., Hanai,N., Takaku,F., Yazaki,Y. and Hirai,H. (1993) *Proc. Natl Acad. Sci. USA*, 90, 5404–5408.
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 July 15, 1993; revised on November 11, 1993