Evidence for position effects as a variant ETV6-mediated leukemogenic mechanism in myeloid leukemias with a t(4;12)(q11-q12;p13) or t(5;12)(q31;p13)

Jan Cools, Nicole Mentens, Maria D. Odero, Pieter Peeters, Iwona Wlodarska, Michel Delforge, Anne Hagemeijer, and Peter Marynen

The ETV6 gene (first identified as TEL) is a frequent target of chromosomal translocations in both myeloid and lymphoid leukemias. At present, more than 40 distinct translocations have been cytogenetically described, of which 13 have now also been characterized at the molecular level. These studies revealed the generation of in-frame fusion genes between different domains of ETV6 and partner genes encoding either kinases or transcription factors. However, in a number of cases—including a t(6;12)(q23;p13), the recurrent t(5;12)(q31;p13), and some cases of the t(4;12)(q11-q12;p13) described in this work—functionally significant fusions could not be identified, raising the question as to what leukemogenic mechanism is implicated in these cases. To investigate this, we have evaluated the genomic regions at 4q11-q12 and 5q31, telomeric to the breakpoints of the t(4;12)(q11-q12;p13) and t(5;12)(q31;p13). The homeobox gene GSH2 at 4q11-q12 and the IL-3/CSF2 locus at 5q31 were found to be located close to the respective breakpoints. In addition, GSH2 and IL-3 were found to be ectopically expressed in the leukemic cells, suggesting that expression of GSH2 and IL-3 was deregulated by the translocation. Our results indicate that, besides the generation of fusion transcripts, deregulation of the expression of oncogenes could be a variant leukemogenic mechanism for translocations involving the 5’ end of ETV6, especially for those translocations lacking functionally significant fusion transcripts. (Blood. 2002;99:1776-1784) © 2002 by The American Society of Hematology

Introduction

Cytogenetic and molecular characterization has demonstrated that nonrandom chromosomal rearrangements in hematologic malignancies are associated with specific oncogenic mechanisms.1,2 Translocations and inversions can result in the generation of oncogenic fusion genes or deregulated expression of protooncogenes.1,2 The latter mechanism has been mainly described for translocations and inversions involving the immunoglobulin heavy and light chain (IgH/L) loci or T-cell receptor (TCR) loci in lymphoid leukemias and lymphomas. In these cases, expression of protooncogenes is deregulated by their juxtaposition to, respectively, B-cell- or T-cell–specific regulatory sequences of the IgH/L or TCR loci.2 This mechanism is less common in myeloid leukemias, although some examples, such as the inv (3)(q21q26) with RPN1-mediated ectopic expression of EVII, have been described.3,4 The other mechanism, the generation of oncogenic fusion proteins, is observed in both lymphoid and myeloid malignancies and frequently involves the CBFA2 (21q22), MLL (11q23), and ETV6 (12p13) loci, which are known to be fused to a wide variety of partner genes.5

Translocations involving the ETV6 gene are among the most intriguing aberrations found in hematologic malignancies, not because so many different translocation partners—more than 405—have been cytogenetically described, but also because this gene contributes to leukemogenesis through at least 3 different mechanisms. One mechanism is the modulation of the activity of the transcription factors such as CBFA2 and ARNT or the activation of the kinases PDGFRB, ABL1, ABL2, JAK2, NTRK3, and SYK by the generation of a fusion protein including the DNA binding domain of ETV6.7,16 A second mechanism is found in the MN1-ETV6 fusion protein that contains the homodimerization (pointed) domain of ETV6.7,16 The molecular analysis of a t(6;12)(q23;p13) case of TEL-ACTC1 leukemia and 3 cases of acute myeloid leukemia (AML)-M0 with a t(4;12)(q11–q12;p13) suggested that, besides the generation of fusion proteins, either chimeric kinases or chimeric transcription factors. However, the molecular analysis of a t(6;12)(q23;p13) case and the recurrent t(5;12)(q31;p13),22 with the 5q31 breakpoints ranging from the 5’ end to the 3’ end of the long fatty acyl-CoA synthetase 2 (ACS2) gene (official gene symbol KIAA0837), has indicated a lack of functionally significant fusion proteins, suggesting that another mechanism is involved in these translocations.

We previously described the in-frame fusion of a novel gene, CHIC2 (BTL), to ETV6 in 1 case of myeloid/natural killer (NK) cell leukemia and 3 cases of acute myeloid leukemia (AML)-M0 with a t(4;12)(q11–q12;p13).23 Although we have characterized CHIC2 as a palmitoylated membrane-associated protein,24 its function

From the Human Genome Laboratory, Center for Human Genetics—Flanders Interuniversity Institute for Biotechnology (VIB), and the Department of Hematology, University of Leuven, Belgium; and the Department of Genetics, University of Navarra, Pamplona, Spain.

Supported by grant G.0121.00 of the Fonds voor Wetenschappelijk Onderzoek-Vlaanderen. J.C. is an aspirant of the Fonds voor Wetenschappelijk Onderzoek-Vlaanderen.
remains unknown and the oncogenic features of CHIC2-ETV6 are not yet understood. We describe here a further molecular analysis of the recurrent t(4;12)(q11-q12;p13) based on the characterization of 2 additional cases. Surprisingly, although in these cases were located within the region spanned by the P1-based artificial chromosome (PAC) contig at the CHIC2 locus,23 the CHIC2-ETV6 fusion was absent and no other meaningful fusions could be detected. Based on this observation and taking into account that the recurrent t(5;12)(q31;p13) is also lacking a functionally significant fusion transcript,22 we hypothesized that the recurrent translocations may involve the ectopic expression of protooncogenes at the partner chromosomes. To evaluate this possibility, the genomic regions telomic to the breakpoints at 4q11-q12 and 5q31 were investigated for the presence of candidate protooncogenes, and the expression of the identified candidate genes was investigated. Our results suggest that ectopic expression of the homeobox gene GSH2 and the growth factor gene interleukin-3 (IL-3), located close to the breakpoints at 4q11-q12 and 5q31, respectively, may be the underlying leukemogenic factors explaining the pathogenesis of the t(4;12)(q11-q12;p13) and t(5;12)(q31;p13).

### Materials, patients, and methods

#### Case reports

Case 1 is a myeloid/NK cell leukemia, and cases 2, 3, and 4 are AML-M0 cases with a t(4;12)(q11-q12;p13), which were previously described.23 Two additional cases with a t(4;12)(q11-q12;p13) were collected at the University Hospital of Navarra, Spain, (case 5) and the University Hospitals Leuven, Belgium (case 6). Case 5 (male, age 47) was diagnosed as AML-M2 and was previously described.6 The material was obtained at relapse. Case 6 (male, age 25) presented with a de novo acute lymphocytic leukemia (ALL) at 11 years of age. He received chemotherapy over a period of 2 years, with complete remission. Remission was obtained but, the patient relapsed and died of leukemia 1 year after the diagnosis of AML. Case 7 (male, age 66) was diagnosed as AML-M2 and was previously described.6 The material was obtained a peripheral blood sample containing less than 10% blasts.

#### RNA isolation and cDNA synthesis

Total RNA was isolated using Trizol (Life Technologies, Paisley, United Kingdom). Poly(A)+ RNA from control bone marrow, peripheral blood, and fluorescence-activated cell sorted (FACS) cells was obtained using the mRNA isolation kit for blood/bone marrow (Roche, Mannheim, Germany). Poly(A)+ RNA from fetal brain was purchased from Clontech, Palo Alto, CA. First-strand complementary DNA (cDNA) was prepared from total RNA or poly(A)+ RNA using Moloney murine leukemia virus reverse transcriptase (RT) (Life Technologies) and random hexamer primers.

#### Rapid amplification of cDNA ends

Rapid amplification of cDNA ends (RACE) was performed as previously described.23 For 5′-RACE on case 5, nested polymerase chain reaction (PCR) was performed using the primers ETF6R3a (5′-GAATGGAGGATCTGAGTACGGC-3′) and ETF6R3b (5′-TCTCTGCTCACTGAATATTGGCCTGAAG-3′). For 3′-RACE on case 5, nested PCR was performed using the primers ETF6F1a (5′-TGAGACATTGCTGAGACTCTCTGT-3′) and ETF6F2 (5′-CCTCTGAGGCGTCCAGTGCAG-3′). For 3′-RACE on case 7, nested PCR was performed using the primers ETF6F1a and ETF6F1b (5′-ACTCTCTGCTGAGTACCATAG-3′).

#### Reverse transcriptase–polymerase chain reaction

Standard RT-PCR was performed to assess the expression of the CHIC2-ETV6 fusion gene and the expression of IL-3, CSF2, IL-5, and IL-9. The primers used were IL-3F (5′-AACTGGGAGGGTCAACAGA-3′) and IL-3R (5′-ACGTCAGTTTCCTCCGGAAAF-3′), CSF2F (5′-AGGGATGGTGCTGAGAACG-3′) and CSF2R (5′-AGGGCGATGCTGGTGAGT-3′), IL-5F (5′-CAAACGGACAGCTTTCAGA-3′) and IL-5R (5′-CAGTACCCCTCTCGACAGTT-3′), and IL-9F (5′-CCTCTGGAACGAGACGAG-3′) and IL-9R (5′-CTTGCTCCTCATCCTCTCCA-3′). TaqMan PCR was performed to analyze the expression of GSH2 using the TaqMan PCR mix (Perkin Elmer, Wellesley, MA) and 10 pm of the primers GSH2TMF (5′-AGATTCCCTGCTCACCATTG-3′), GSH2TMR (5′-TCCTCTCCAGTCCAGAGTT-3′), GAPDTP (5′-AGCCCTAACGATACGCAAGATG-3′), and GAPDTPR (5′-ATGGACTGTGGTCACGAGCTT-3′) together with 10 pm of the probes GSH2TMP (5′-FAM-CTCTGAGCCGACCGGATACCCA-TAMRA-3′) and GAPDTPM (5′-JOE-CCAAGCTTGTGAACCCCTCGGCC-TAMRA-3′). The detection of GAPD served as internal control. The Ct values given in the text represent the number of the PCR cycle in which the fluorescent signal significantly exceeds background levels.

### Generation of sequence tagged site markers

Sequence tagged sites (STSs) were designed based on available genomic sequences. The primers used were as follows: for STS-G, 5′-GCACCGGCACACCTTACAC-3′ and 5′−GAGGATGTGAAGGAGGAGGAC-3′, for STS-K, 5′−AAGGCCAGACTGTTGAGGTTG-3′ and 5′−GGTGCCAGATACGCAAGATG-3′, for STS-5A, 5′−TGGGTTGAGCTCTTGACAC-3′ and 5′−GAACCTGAGCAAAGGATTG-3′, for STS-3A, 5′−AAATGTCCTGCCAGGACTT-3′ and 5′−CCATCATTTGTAAGGTTG-3′, for STS-1L, 5′−ILF-3 and 5′−ILR, and for STS-C, CSF2F and 5′−ACAGGCCACATCTCTC-3′. PCR was performed with an annealing temperature of 60°C and 30 cycles.

### Constructs

The construct murine stem cell virus (MSCV)-GSH2, containing the full open reading frame of GSH2, was obtained by ligation a genomic fragment containing exon 1 to a PCR product containing exon 2 using the NcoI restriction site at the border of the 2 exons into the retroviral vector MSCV-puromycin (kindly provided by Dr D.G. Gilliland, Harvard, Boston, MA). The MSCV-GSH2ΔE-N construct, containing only the homeobox domain, and the MSCV-GSH2ΔE construct, containing the open reading frame of GSH2 with a deletion in the homeobox domain, were generated by PCR. The empty MSCV-puromycin vector (MSCV) served as control.

### Cell culture, viral vector production, and transduction

The HEK293T and NIH3T3 cell lines were cultured in Dulbecco modified Eagle medium containing 10% fetal calf serum. NIH3T3 cells were obtained from the American Tissue Culture Collection (ATCC, Manassas, VA). Viral vectors were produced by cotransfection of HEK293T cells with a mixture of the respective MSCV plasmids and an ecotropic packaging plasmid (Cell Genesys, Foster City, CA) using Fugene-6 (Roche). Transduction of NIH3T3 cells was performed in 6-well plates in the presence of polybrene (8 μg/mL) (Sigma, St Louis, MO). After 8 hours, the cells were washed with phosphate-buffered saline and fresh medium was added. After 24 hours, selection was started using puromycin (3 μg/mL). All experiments were started within 2 weeks after the transduction and were performed at least in double using independent transductions.

### FACS of bone marrow cells

Bone marrow cells were obtained from healthy volunteers. Mononuclear cells, obtained by using a Ficoll density gradient, were separated in a mixture of the respective MSCV plasmids and an ecotropic packaging plasmid (Cell Genesys, Foster City, CA) using Fugene-6 (Roche). Transduction of NIH3T3 cells was performed in 6-well plates in the presence of polybrene (8 μg/mL) (Sigma, St Louis, MO). After 8 hours, the cells were washed with phosphate-buffered saline and fresh medium was added. After 24 hours, selection was started using puromycin (3 μg/mL). All experiments were started within 2 weeks after the transduction and were performed at least in double using independent transductions.

Genomic clones

PAC clones for the CHIC2 locus were previously described.23 The bacterial artificial chromosome (BAC) clone 434C1, containing more than 100 kilobases (kb) upstream and 40 kb downstream of exon 1 of ETV6, was identified by database searches and obtained from the RPCI11 library (Roswell Park Cancer Institute, Buffalo, NY). Other clones from the ETV6 locus were previously described.26,27 The PAC clones P3701 and P3704 were identified by screening of the RPC15 library (Roswell Park Cancer Institute) using a probe derived from the 5′ end of the ACS2 gene. The localization of cosmids c33g8 (kindly provided by Dr M. Lovett) at the IL-3 locus was previously described.28

Results

Molecular cloning of additional t(4;12)(q11:p13) cases

We previously described the molecular analysis of a myeloid/NK cell leukemia and 3 AML-M0 cases with a t(4;12)(q11-q12;p13).23 All cases were characterized by the presence of a fusion of the first 3 exons of CHIC2 (BLT) to exons 2 to 8 of ETV6 (TEL). For 3 of these cases, the 4q breakpoint was determined by fluorescence in situ hybridization (FISH) and mapped to the region spanned by PAC 1146G14 (Figure 1). For this study, 2 additional AML cases with a t(4;12)(q11-q12;p13) were collected. Investigation of these cases by FISH revealed the breakpoints at 12p13 to be located in the CHIC2 locus at 4q11-q12. 23 Database searches identified several partially sequenced BAC clones containing the CHIC2 and GSH2 genes, which were aligned to the PAC contig (drawing not to scale). Using PCR and database searches, GSH2 and the 2 exons of HSG2 were located on this contig. The localization of the breakpoints of 6 t(4;12) cases is indicated by an arrowhead and numbered according to the cases as described in the text. Exons are represented by open or filled squares, and the transcriptional orientation is indicated by an arrow. The sequenced BAC clones are represented by their accession numbers. Tel indicates telomeric side; cen, centromeric side; open circle, negative by hybridization or PCR; filled square, positive by hybridization or PCR; k, positive by electronic PCR/database searches.

Table 1. Summary of the FISH results obtained for the 2 novel t(4;12)(q11-q12;p13) cases

<table>
<thead>
<tr>
<th>Probes (Localization)</th>
<th>Case 5</th>
<th>Case 6</th>
</tr>
</thead>
<tbody>
<tr>
<td>4q11</td>
<td>CHIC2</td>
<td></td>
</tr>
<tr>
<td>200D9</td>
<td>Exon 1</td>
<td>der(12)</td>
</tr>
<tr>
<td>238H24</td>
<td>Exons 1-6</td>
<td>split</td>
</tr>
<tr>
<td>1146G14</td>
<td>Exons 4-6</td>
<td>der(4)</td>
</tr>
<tr>
<td>12p13</td>
<td>ETV6</td>
<td></td>
</tr>
<tr>
<td>434C1</td>
<td>Telomic to exon 1/exon 1</td>
<td>der(4)</td>
</tr>
<tr>
<td>209J1</td>
<td>Exon 1/intron 1</td>
<td>der(12)</td>
</tr>
<tr>
<td>179A6</td>
<td>Exon 1/intron 1</td>
<td>der(4)</td>
</tr>
<tr>
<td>55F4</td>
<td>Intron 1/exon 2</td>
<td>der(4)</td>
</tr>
<tr>
<td>TEL11</td>
<td>Intron 2/exon 3</td>
<td>split</td>
</tr>
<tr>
<td>54D5</td>
<td>Exons 5-8</td>
<td>der(12)</td>
</tr>
<tr>
<td>148B6</td>
<td>Exon 8</td>
<td>der(12)</td>
</tr>
</tbody>
</table>

*Only part of the metaphases showed split signals. This is indicative for a breakpoint close to the centromeric side of the region covered by PAC 200D9.

238H24, while the breakpoint in case 6 was located in the region spanned by PAC 200D9 and telomeric to PAC 238H24 and thus upstream of CHIC2 (Table 1, Figure 1). This directly excluded the presence of a CHIC2-ETV6 fusion in case 6. Attempts to amplify a CHIC2-ETV6 fusion transcript in case 5 using primers in exon 1 of CHIC2 and in exon 3 of ETV6 yielded no product, whereas the control (case 1, previously described23) was positive, indicating that also in case 5 no CHIC2-ETV6 fusion transcript was present (results not shown).

We next investigated the presence of other fusion transcripts in case 5. Using 5′-RACE starting from exon 3 of ETV6, transcripts were detected fusing a novel sequence to exon 2 or directly to exon 3 of ETV6 (Figure 2). This translocation seemed to be complex, because 3′-RACE experiments showed the presence of a transcript fusing exon 2 of ETV6 to the same sequence as identified by 5′-RACE (Figure 2), indicative of a duplication of both 12p and 4q sequences associated with the translocation. No DNA from this case was available to confirm this by Southern blotting. Unfortunately, due to lack of sufficient material, a molecular analysis could not be performed for case 6.

Cloning of human GSH2/HSG2

Database searches revealed 100% identity of part of the sequence obtained by 5′-RACE from case 5 to several human expressed sequence tags. These represent the GSH2 gene; they are highly similar to the murine Gsh2 cDNA sequence. However, the sequence fused to ETV6 was derived from the complementary noncoding strand of GSH2. Part of this novel sequence overlapped with exon 2 of GSH2 and was found to be colinear with the sequence of BAC AC009614, while the other part of the novel sequence was found to be colinear with BAC AC069068. Consensus splice donor and acceptor sequences were present at the genomic sequences, indicating that these 2 parts are 2 exons of a novel gene that is located at the GSH2 locus but is transcribed in the opposite direction as GSH2 and hence was named HSG2 (Figure 1). Stop codons were found to be present in the 3 different reading frames, indicating that these exons are derived from the 5′- or 3′-untranslated region of HSG2 or that HSG2 is a noncoding transcript. The exact identity and complete cDNA sequence of HSG2 was not further analyzed. Because the complete HSG2 transcript is currently unknown, the 2 exons of HSG2 identified in this work could not be numbered and were named exon j and exon k.

Next, the 2 exons of GSH2 were cloned from a HindIII fragment of PAC 200D9. Similar to murine Gsh2, the complete 915 bp open
To investigate this, we analyzed the expression of GSH2 in reading frame of GSH2 was found to be present in 2 exons (sequence submitted to GenBank), separated by an intron of 2 kb. Several expressed sequence tags were found to match 100% with the exon sequences determined by us and confirmed the splice sites. In addition, RT-PCR on cDNA from human fetal brain using primers in the different exons confirmed the expression and splicing of GSH2 (results not shown).

An STS derived for exon 2 of GSH2 (STS-G) and for exon k of HSG2 (STS-K), respectively, was used to assess their localization in our PAC contig spanning the CHIC2 locus. GSH2, and thus also exon j of HSG2, was found to be present only in PAC 200D9, whereas exon k of HSG2 was located in the region shared between the PACs 200D9 and 238H24 (Figure 1).

Taken together, these data indicated that GSH2 was located in front of CHIC2 and that not GSH2, but its antisense gene, HSG2, was fused to ETV6 in case 5. Because the HSG2 transcript contained stop codons in all reading frames, no HSG2-ETV6 fusion protein can be generated. Thus, in case 5, no functionally significant fusion transcript was generated by the translocation. However, knowing that the GSH2 gene was in close proximity to the breakpoint in the t(4;12)(q11-q12;p13) cases and taking into account that ectopic expression of homeobox genes is a known mechanism involved in leukemia, we hypothesized that ectopic expression of GSH2 could represent an alternative leukemogenic event. To investigate this, we analyzed the expression of GSH2 in these cases and investigated its transforming properties using a focus formation assay in the NIH3T3 cell line.

Expression of GSH2 in the t(4;12)

We analyzed the expression of GSH2 by use of TaqMan real-time quantitative RT-PCR using GAPD as an internal control. As expected, expression of GSH2 was detected in control CD34+CD13+CD38+HL60 cell line 16.0 26.0 10.5

Table 2. TaqMan real-time quantitative RT-PCR analysis of the expression of GSH2

<table>
<thead>
<tr>
<th>Sample</th>
<th>C\textsubscript{T} value\textsuperscript{a}</th>
<th>(\Delta C_{T})</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fetal brain</td>
<td>15.5</td>
<td>26.0</td>
</tr>
<tr>
<td>Bone marrow</td>
<td>19.6 0.2</td>
<td>Greater than 40</td>
</tr>
<tr>
<td>Peripheral blood</td>
<td>23.4 0.2</td>
<td>Greater than 40</td>
</tr>
<tr>
<td>Spleen</td>
<td>22.6 0.3</td>
<td>Greater than 40</td>
</tr>
<tr>
<td>HL60 cell line</td>
<td>16.0 0.4</td>
<td>Greater than 40</td>
</tr>
<tr>
<td>CD34\textsuperscript{+}</td>
<td>21.1 0.9</td>
<td>Greater than 40</td>
</tr>
<tr>
<td>CD34\textsuperscript{+} CD38\textsuperscript{+}</td>
<td>33.3 0.6</td>
<td>Greater than 40</td>
</tr>
<tr>
<td>CD34\textsuperscript{+} CD38\textsuperscript{+} CD7\textsuperscript{+}</td>
<td>24.5 1.3</td>
<td>Greater than 40</td>
</tr>
<tr>
<td>CD34\textsuperscript{+} CD34\textsuperscript{+} CD38\textsuperscript{+} CD7\textsuperscript{+}</td>
<td>26.4 0.2</td>
<td>Greater than 40</td>
</tr>
<tr>
<td>CD34\textsuperscript{+} CD34\textsuperscript{+} CD38\textsuperscript{+} CD34\textsuperscript{+} CD7\textsuperscript{+}</td>
<td>23.2 0.5</td>
<td>Greater than 40</td>
</tr>
<tr>
<td>CD34\textsuperscript{+} CD34\textsuperscript{+} CD38\textsuperscript{+} CD34\textsuperscript{+} CD34\textsuperscript{+} CD7\textsuperscript{+}</td>
<td>18.0 1.0</td>
<td>Greater than 40</td>
</tr>
<tr>
<td>CD34\textsuperscript{+} CD34\textsuperscript{+} CD38\textsuperscript{+} CD34\textsuperscript{+} CD34\textsuperscript{+} CD34\textsuperscript{+} CD7\textsuperscript{+}</td>
<td>27.3 0.6</td>
<td>Greater than 40</td>
</tr>
<tr>
<td>CD34\textsuperscript{+} CD34\textsuperscript{+} CD38\textsuperscript{+} CD34\textsuperscript{+} CD34\textsuperscript{+} CD34\textsuperscript{+} CD34\textsuperscript{+} CD7\textsuperscript{+}</td>
<td>30.3 0.6</td>
<td>Greater than 40</td>
</tr>
<tr>
<td>CD34\textsuperscript{+} CD34\textsuperscript{+} CD38\textsuperscript{+} CD34\textsuperscript{+} CD34\textsuperscript{+} CD34\textsuperscript{+} CD34\textsuperscript{+} CD34\textsuperscript{+} CD7\textsuperscript{+}</td>
<td>27.7 1.2</td>
<td>Greater than 40</td>
</tr>
<tr>
<td>BM from case 1</td>
<td>22.7</td>
<td>25.4</td>
</tr>
<tr>
<td>BM from case 2</td>
<td>24.9</td>
<td>30.3</td>
</tr>
<tr>
<td>PB from case 3</td>
<td>25.0</td>
<td>33.8</td>
</tr>
<tr>
<td>BM from case 4</td>
<td>22.2</td>
<td>26.9</td>
</tr>
<tr>
<td>BM from case 5</td>
<td>28.3</td>
<td>37.1</td>
</tr>
<tr>
<td>BM from case 6</td>
<td>Nd</td>
<td>Nd</td>
</tr>
</tbody>
</table>

\textsuperscript{a}The C\textsubscript{T} value was determined in at least 3 independent experiments. It is indicated as mean \(\pm\) SD when no expression of GSH2 was detected. When expression of GSH2 was detected, individual C\textsubscript{T} values are given from 1 representative experiment and the mean \(\pm\) SD is determined from the \(\Delta C_{T}\) values obtained from the independent experiments.

\textsuperscript{†}RNA was extracted from fixed cells.

\textsuperscript{‡}Water was used as “no template control” in each experiment.
homeobox gene *NKK2-3* (data not shown) were easily detected, indicating that the negative result for *GSH2* was not due to a technical problem and that the RNA extraction from these FACS-sorted cells was adequate.

**Overexpression of *GSH2* transforms NIH3T3 cells**

To investigate the transforming properties of *GSH2* in vitro, we retrovirally transduced the NIH3T3 cell line with the empty MSCV vector or with the MSCV vector containing wild-type *GSH2* (MSCV-GSH2), *GSH2* with a deletion of the N-terminus (MSCV-GSH2delN), or *GSH2* with a deletion of the homeobox domain (MSCV-GSH2delH), respectively. Focus formation was determined as the ability to induce a significant larger number of foci than observed after transduction of the empty vector. Experiments with the different constructs were performed at the same time, with the same batch of cells. The number of spontaneous foci observed for the empty vector was 10 at maximum. In contrast, overexpression of wild-type *GSH2* resulted in the presence of more than 200 small foci per 10 cm dish (Figure 3A,B, Table 2). The deletion mutants of *GSH2* induced only a small number of foci, comparable to the empty vector (Table 3). Typically, these foci were larger in size than the foci induced by overexpression of *GSH2* (Figure 3D), although microscopically the foci induced by *GSH2* were clearly composed of transformed cells but apparently grew more slowly (Figure 3GH). We also tested whether expression of the *CHIC2-ETV6* fusion (MSCV-CHIC2ETV6) had transforming properties, but expression of *CHIC2-ETV6* did not induce focus formation in the NIH3T3 cell line (results not shown).

**Molecular analysis of an atypical CML case with a t(5;12)(q31:p13)**

The t(5;12)(q31;p13) has been described as a recurrent translocation occurring in various myeloid malignancies (at least 18 cases were previously described), often associated with eosinophilia. We collected an atypical CML (aCML) case with a t(5;12)(q31;p13) as the sole chromosomal abnormality and investigated the presence of a possible fusion transcript. This case was described previously, and the breakpoint in *ETV6* was mapped to intron 1. RACE experiments revealed the presence of an *ETV6-ACS2* fusion, confirming that this case was similar to the cases reported by Yagasaki et al. Further analysis by RT-PCR revealed the presence of 3 different *ETV6-ACS2* fusion transcripts, among which was an in-frame fusion fusing the first 11 amino acids of *ETV6* to almost the complete ACS2 protein (Figure 4). According to the RT-PCR results, the breakpoint in *ACS2* was at the 5′ end of this gene, similar to case 3 described by Yagasaki and coworkers.

However, in the 3 cases studied by Yagasaki and colleagues, only out-of-frame *ETV6-ACS2* fusion transcripts were identified, showing that the t(5;12)(q31;p13) does not generate a common in-frame fusion gene. The exact leukemogenic factor responsible for the pathogenesis of the t(5;12)(q31;p13) thus remained unresolved. According to these observations, the only invariant result of this translocation is the juxtaposition of the region directly telomeric to *ACS2* with the *ETV6* locus. We therefore investigated whether this could lead to the juxtaposition of a protooncogene on 5q31 to *ETV6* and if this could better explain the leukemogenic mechanism of this translocation.

Analysis of the region telomeric to the breakpoint at 5q31 revealed the presence of the *IL-3/CSF2* locus in the genomic sequence of BAC CTD-2198K16 (accession no. AC034228) that also contained the 5′ end of *ACS2*. This was confirmed by further analysis of this region using the available genomic sequences (accession numbers are listed in Figure 5) and by mapping of 2 PAC clones (P3701 and P3704) and a cosmide (c33g8) by PCR using STS markers specific for the 5′ end of *ACS2* (STS-5A) and the 3′ end of *ACS2* (STS-3A), *IL-3* (STS-I), and *CSF2* (STS-C) (Figure 5). A complete 1 Mb contig encompassing the *ACS2*, *IL-3*, *CSF2*, *IRF1*, *IL-5*, and *IL-4* loci was constructed previously, although the exact orientation of this contig remained unclear. This was due mainly to the small physical distance between the *IL-3* and *IL-5* loci, which made FISH-based mapping difficult and resulted in contradictions in the literature. The cloning of 4 t(5;12) cases with an *ETV6-ACS2* fusion and the known telomere to centromere orientation of *ETV6* are strongly suggestive for a telomere to centromere orientation of *ACS2*, enabling orientation of the contig described by Frazer et al. as centromere-ACS2–IL3–CSF2–IRF1–IL5–IL4–telomere. In agreement with this, Yagasaki et al. mapped the *IRF1* gene telomeric to the 5q31 breakpoint by FISH, and Grimaldi et al. cloned the t(5;14)(q31;q32) breakpoint, also confirming that *CSF2* is located telomeric to *IL-3*. Taken together, all these data are consistent with a localization of *IL-3* immediately telomeric to *ACS2*, as shown in Figure 5. Because *ACS2* and *IL-3* were both present in PAC P3704 and according to the contig

![Figure 3. Focus formation assay illustrating the oncogenic potential of *GSH2*.](image_url)
The expression of IL-3, CSF2, IL-5, and IL-9 was analyzed in a bone marrow sample obtained from the aCML case (case 7). Only expression of IL-3 could be clearly detected (Figure 6), indicating a good correlation between the position of the breakpoint and the expression of IL-3 in this myeloid sample. Although CSF2 is only 10 kb distal to IL-3, no expression of this gene was observed. The expression of IL-5 and IL-9, which are not located in the immediate surroundings of the t(5;12) breakpoint, was investigated as a control. In this way a “contamination” of the sample with normal T cells expressing IL-3 was excluded because expression of IL-5 and IL-9 would then also be detected.

Discussion

We report here a further molecular study of the recurrent translocations t(4;12)(q11-q12;p13) and t(5;12)(q31;p13). Our results described in this work together with our previous study23 and the study of Yagasaki and colleagues22 clearly indicate heterogeneity in the localization of the breakpoints at both 4q11-q12 (upstream, at the 5’ end or in intron 3 of CHIC2) and 5q31 (at the 5’ end, in the middle or at the 3’ end of ACS2). In correspondence with this, no common fusion transcript that could explain the pathogenic character of these translocations is generated by either of these translocations. In addition, the HSG2-ETV6 transcript detected in case 5 with the t(4;12) and the ETV6-ACS2 transcripts described by Yagasaki et al22 in cases with a t(5;12) do not represent in-frame fusion transcripts, raising the question as to how these transcripts contribute to the pathogenesis of the respective translocations. Two other AML cases with a t(4;12)(q11-q12;p13) and an HSG2-ETV6 fusion have been reported,36 further confirming the heterogeneity of the 4q11-q12 breakpoints. In fact, the only invariant result of these translocations is the juxtaposition of the chromosomal regions telomeric, respectively, to CHIC2 on 4q11-q12 and ACS2 on 5q31 to the ETV6 locus on 5q31 and localization of the different breakpoints of the t(5;12)(q31;p13). Schematic illustration of the small PAC/cosmid contig linking ACS2 with IL-3/CSF2 (not to scale). Database searches identified several partially or completely sequenced genomic clones containing the ACS2, IL-3, CSF2, and P4HA2 genes, confirming the PAC contig. These clones are represented by their accession numbers. Yagasaki et al mapped IRF1 telomeric to the breakpoints of the t(5;12). A 1 Mb contig was previously described33 that links ACS2 to IL-5 and is in agreement with the map shown here. The 3 t(5;12)(q31;p13) breakpoints described by Yagasaki et al are numbered as Y1, Y2, and Y3. The breakpoint described in this work is indicated with case 7, as described in the text. Genes are represented by boxes, and their transcriptional orientation is indicated by an arrow. The sequenced BAC clones are represented by their accession numbers. tel indicates telomeric side; cen, centromeric side; open circle, negative by hybridization or PCR; filled square, positive by hybridization or PCR; x, positive by electronic PCR/database searches.

Figure 5. PAC/BAC/cosmid contig covering the ACS2 and IL-3/CSF2 loci at 5q31 and localization of the different breakpoints of the t(5;12)(q31;p13). Schematic illustration of the small PAC/cosmid contig linking ACS2 with IL-3/CSF2 (not to scale). Database searches identified several partially or completely sequenced genomic clones containing the ACS2, IL-3, CSF2, and P4HA2 genes, confirming the PAC contig. These clones are represented by their accession numbers. Yagasaki et al mapped IRF1 telomeric to the breakpoints of the t(5;12). A 1 Mb contig was previously described33 that links ACS2 to IL-5 and is in agreement with the map shown here. The 3 t(5;12)(q31;p13) breakpoints described by Yagasaki et al are numbered as Y1, Y2, and Y3. The breakpoint described in this work is indicated with case 7, as described in the text. Genes are represented by boxes, and their transcriptional orientation is indicated by an arrow. The sequenced BAC clones are represented by their accession numbers. tel indicates telomeric side; cen, centromeric side; open circle, negative by hybridization or PCR; filled square, positive by hybridization or PCR; x, positive by electronic PCR/database searches.

Figure 6. Expression of IL-3 in the aCML case with a t(5;12)(q31;p13) (case 7). Expression of IL-3 but not of CSF2, IL-5, or IL-9 was detected in a bone marrow sample from the patient with aCML (case 7). By contrast, high-level expression of IL-3 could not be detected in control bone marrow cells (BM), indicating ectopic expression of IL-3 in the t(5;12) case. The faint band observed at 394 bp in control bone marrow is due to a genomic contamination of the RNA sample.
1p13. The juxtaposition of protooncogenes to the RPN1 locus, the IgH/L loci, or the TCR loci has been well illustrated and results in the deregulated expression of these protooncogenes.\textsuperscript{2,3,37} We identified the homeobox gene GSH2 and the growth factor genes IL-3 and CSF2 in the regions adjacent to the breakpoints at, respectively, q411-q12 and q531 of the t(4;12)(q11-q12;p13) and the t(5;12)(q31; p13). Ectopic expression of GSH2 was detected in all t(4;12)(q11; p13) cases, with or without the CHIC2-ETV6 fusion, and expression of IL-3 was observed in the aCML case with a t(5;12)(q31; p13), also characterized by an ETV6-ACS2 fusion. These results suggest that ectopic expression of GSH2 and IL-3 could be the underlying leukemogenic mechanisms of the t(4;12)(q11-q12;p13) and t(5;12)(q31;p13) cases, independent of the presence of fusion transcripts.

The Gsh2 gene was previously characterized as a brain-specific homeobox gene important in forebrain and hindbrain formation of the mouse.\textsuperscript{38,39} We did not find evidence that GSH2 is also expressed during hematopoiesis. Interestingly, GSH2, belonging to the ParaHox genes,\textsuperscript{40} contains a homeobox domain that is very similar to the homeobox of the clustered HOX genes, which are involved in both normal and abnormal hematopoiesis.\textsuperscript{41} In addition, overexpression of GSH2 was found to induce focus formation in the NIH3T3 cell line (Figure 3), illustrating its oncogenic potential.

Evidence for the involvement of IL-3 in leukemia was described previously by the analysis of the t(5;14)(q31;q32) in B-ALL, involving ectopic expression of IL-3 as a consequence of its translocation to the IgH locus.\textsuperscript{35,42} and by analysis of mutations found in the murine WEHI-3B cell line, which illustrated that ectopic expression of HoxB8 and IL-3 together results in AML in a mouse model.\textsuperscript{43} Although expression of IL-3 alone produced only a nonneoplastic myeloproliferative disease in this mouse model, it is clear from these studies that ectopic expression of IL-3 leads to a proliferative defect and may be involved in the multistep process of leukemogenesis.\textsuperscript{44} In addition, because IL-3 is a growth factor regulating eosinophil proliferation and differentiation,\textsuperscript{52} ectopic expression of IL-3 could also explain the eosinophilia associated with the t(5;12) cases.\textsuperscript{22,25,29,32} as was also observed for the t(5;14)(q31;q32).\textsuperscript{35,42} This is supported by the observation that continuous administration of IL-3 leads to a dose- and time-dependent eosinophilia.\textsuperscript{46}

Although we clearly detected expression of GSH2 and IL-3 in the t(4;12) and t(5;12) cases, several remarks must be taken under consideration when describing this as an abnormal expression. Bone marrow is a heterogeneous mixture of different cell types of which in leukemia only one particular cell was originally mutated and expanded. It is thus possible that the expression of a gene, which is expressed only in a rare bone marrow cell type, would become detectable by RT-PCR in leukemic bone marrow as a consequence of the clonal expansion of this cell type. Second, translocations are known to occur in open chromatin structures, suggesting that the genes found at the breakpoints are transcriptionally active, thus arguing that the detected expression of GSH2 and IL-3 in the leukemias might not be abnormal. However, the breakpoints of the t(4;12) and t(5;12) are not very close to GSH2 or IL-3 but occur in or at least close to, respectively, the CHIC2 and ACS2 genes, which are expressed in the hematopoietic system\textsuperscript{22,23} and thus most probably are located in an open chromatin region, providing space for the translocations to occur. The fact that GSH2 expression could not be detected in any of the analyzed subpopulations highly enriched for CD34+ cells further indicates that GSH2 is not expressed in the hematopoietic system, although we cannot exclude its expression in a particularly rare bone marrow cell type. However, even if GSH2 would be expressed in a rare bone marrow cell type that escaped our detection, this would not directly exclude GSH2 from its involvement in the t(4;12). Transformation might then be the result of the inability to down-regulate GSH2, rather than its induced expression. In this regard, the ectopic expression of Hoxa9 by retroviral insertion in BXH2 mice is a well-characterized example illustrating that the ectopic expression of a gene that is normally expressed in bone marrow progenitor cells can nevertheless be leukemogenic.\textsuperscript{47,48}

It has been well established that IL-3 is expressed by activated T cells, NK cells, and mast cells.\textsuperscript{49} Detection of its expression in a CML sample, containing mainly myeloid cells, is suggestive of its abnormal expression in this cell type. If expression of IL-3 would be the result of the presence of other IL-3–producing cells, then expression of other growth factors such as CSF2, IL-5, and IL-9 also would be expected, which was not the case (Figure 6). Although IL-3 and CSF2 are located very close to each other, no expression of CSF2 was detected in the aCML case (Figure 6). Interestingly, this was also observed in B-cell ALL cases with a t(5;14)(q31;q32) involving the IgH locus.\textsuperscript{42} Although IL-3 and CSF2 have an overlapping expression pattern, CSF2 expression can also be regulated separately.\textsuperscript{50,51} Elucidation of the mechanism by which expression of IL-3 becomes activated in these leukemias could provide further explanation for this.

The exact mechanism by which expression of GSH2 and IL-3 is activated is currently not known. It has become clear that the final transcription level of a gene is determined by the combined action of regulatory sequences that might be either positive or negative and the accessibility of these sequences determined by the chromatin structure and the nuclear architecture.\textsuperscript{51,52} Therefore, it might be expected that regulatory sequences as enhancers in ETV6 and silencers at the partner chromosome can be major contributors, with additional effects from the change in chromatin structure around the breakpoint. Well-known examples involving enhancer-mediated ectopic expression are the translocations involving the RPN1 locus, the IgH/L loci, or the TCR loci.\textsuperscript{2,3,17,53} These studies have also clearly illustrated that the breakpoints can be located up to 200 kb away from the deregulated protooncogene.\textsuperscript{1} Other examples of how translocations can influence the expression of genes have been described for the translocations associated with Beckwith-Wiedemann syndrome, which involve aberrant expression of the IGF2 gene by loss of imprinting,\textsuperscript{24,25} and for chromosome rearrangements at the Kit/Pdgfα loci in mutant mice, affecting the expression of Kit.\textsuperscript{54,55} In addition to the effects of enhancers and silencers, all translocations are likely to change local chromatin structure and nuclear architecture in its whole, which may be associated with changes in gene expression.

Taken together, our results indicate that the common mechanism of the t(4;12)(q11-q12;p13) cases, with or without the CHIC2-ETV6 fusion, may be the ectopic expression of GSH2. Further analysis of additional cases with breakpoints at the 5′ end of CHIC2 and analysis of the in vivo leukemogenic properties of GSH2 will be necessary to fully understand the importance of its expression in the t(4;12)(q11-q12;p13) cases. In addition, our results suggest that the AML-M0 cases represent a subgroup in the t(4;12)(q11-q12;p13) characterized by both the expression of CSF2 and the CHIC2-ETV6 fusion. Although expression of CHIC2-ETV6 does not transform the NIH3T3 cell line, it will be of interest to investigate whether the presence of the CHIC2-ETV6 fusion plays an additional role in determining the immature phenotype of the AML-M0 cases. With respect to the reported
t(5;12)(q31;p13) cases, both our data and the known leukemia-related properties of IL-3 are supportive for the ectopic expression of IL-3 as an important transforming mechanism in these cases. Although we have only been able to investigate a single case, the 3 cases described by Yagasaki et al. also involve a translocation of IL-3 to the der, supporting the possibility of IL-3-induced expression by ETV6 regulatory sequences.

Although many translocations involving ETV6 have been shown to result in the generation of an in-frame fusion gene, encoding a fusion protein with oncogenic properties, our results indicated a variable importance of a variant mechanism: the ectopic expression of a protooncogene. Other translocations, such as the reported t(6;12)(q23;p13)21 and t(7;12)(p15;p13)57 and the t(10;12)(q24;p13) and t(12;13)(p13;q12) (P. M., unpublished results, 2001), seem not to result in the generation of functionally significant fusion genes. It will be of further interest to investigate whether these cases involve ectopic expression of a protooncogene. Interestingly, several cases of the t(3;12)(q26;p13) have been described, with the 3q26 breakpoints either 5’ or 3’ to ETVI, a gene that is ectopically expressed in myeloid malignancies with the inv(3)(q21q26) or the t(3;3)(q21;q26). One of the t(3;12)(q26;p13) cases was cloned and revealed the fusion of ETV6 to MDS1/ETVII. However, at least one of the other cases described by Raymond et al.58 had breakpoints centromeric to EVII and thus could involve ectopic expression of EVII by a similar mechanism as proposed in this work. Identification of candidate protooncogenes for other translocations involving breakpoints at the 5’ end of ETV6 will be greatly facilitated by the availability of the complete human genome sequence.

Acknowledgments

We thank Vic Van Duppen for the FACS analysis and Betty Emanuel for performing FISH experiments. We are very grateful to Drs G. Verhoef, C. Cabrol, P. Talman, and C. Bilhou-Nabera for providing valuable patient material.

References


Evidence for position effects as a variant ETV6-mediated leukemogenic mechanism in myeloid leukemias with a t(4;12)(q11-q12;p13) or t(5;12)(q31;p13)

Jan Cools, Nicole Mentens, Maria D. Odero, Pieter Peeters, Iwona Wlodarska, Michel Delforge, Anne Hagemeijer and Peter Marynen