MOLECULAR CHARACTERIZATION OF TRANSLOCATION (6;9) IN ACUTE NONLYMPHOCYTIC LEUKEMIA

MOLEKULAIRE KARAKTERISATIE VAN TRANSLOCATIE (6;9) IN ACUTE NIETLYMFOCYTAIRE LEUKEMIE

PROEFSCHRIFT

ter verkrijging van de graad van doctor
aan de Erasmus Universiteit Rotterdam
op gezag van de Rector Magnificus
Prof. Dr. C.J. Rijnvos
en volgens besluit van het College van Dekanen.

De openbare verdediging zal plaats vinden op
woensdag 10 juni 1992 om 15.45 uur

door

Maartje Maria von Lindern
geboren te Krimpen aan de Lek

PROMOTIECOMMISSIE

Promotor:	Prof. Dr. D. Bootsma
Overige leden:	Prof. Dr. A.J. van der EB Prof. Dr. A. Hagemeijer Prof. Dr. B. Löwenberg
Co-promotor:	Dr. G.C. Grosveld

Printed by: Haveka B.V., Alblasserdam, The Netherlands.

Dit proefschrift is bewerkt binnen de vakgroep Celbiologie en Genetica van de Faculteit der geneeskunde en gezondheidswetenschappen van de Erasmus Universiteit Rotterdam. De Nederlandse Kankerbestrijding (Koningin Wilhelmina Fonds) heeft het onderzoek financieel gesteund en heeft ook bijgedragen aan de drukkosten van dit proefschrift.

..... ours are " the two experiences of adult life" of which Pavese spoke, success and failure, to kill the white whale or wreck the ship; one should not surrender to incomprehensible matter, one must not just sit down. We are here for this - to make mistakes and to correct ourselves, to stand the blows and hand them out. We must never feel disarmed; nature is immense and complex, but it is not impermeable to the intelligence; we must circle around it, pierce and probe it, look for the opening or make it.

Primo Levi,
The Periodic Table

aan allen die hun oprechte belangstelling
voor het hierbeschreven werk hebben getoond

Omslag : Reliëf op de voorgevel van de dom van Modena voorstellende het offer van Kaïn en Abel en de doodslag van Abel. Beeldhouwer: Wiligelmo, begin 12e eeuw.

INTRODUCTION

Specific chromosomal translocations are one of the defects associated with leukemia. Isolation and characterization of genes affected by these translocations may give insight into the processes of both leukemogenesis and normal hematopoiesis. When the experiments described in this thesis were started, several genes involved in translocations in lymphoid leukemia were isolated. These genes were all translocated into T-cell receptor and Immunoglobulin loci, which deregulated their expression. In myeloid leukemia only translocation (9;22) was characterized molecularly and the resulting *bcr-abl* gene was the only fusion gene known. Chapter 1 gives an overview of what is known to date about genes involved in leukemogenesis. To extend the research on the molecular characterization of translocations in myeloid leukemia, we decided to clone and characterize the translocation breakpoints of t(6;9) that characterizes a subtype of acute myeloid leukemia. Chapter 2 gives an introduction to t(6;9) AML and reports the results of our investigations. Chapter 3 discusses these results in relation to our current understanding of leukemogenesis.

CHAPTER 1

HEMATOPOIESIS AND LEUKEMIA

1.1 Normal hematopoiesis

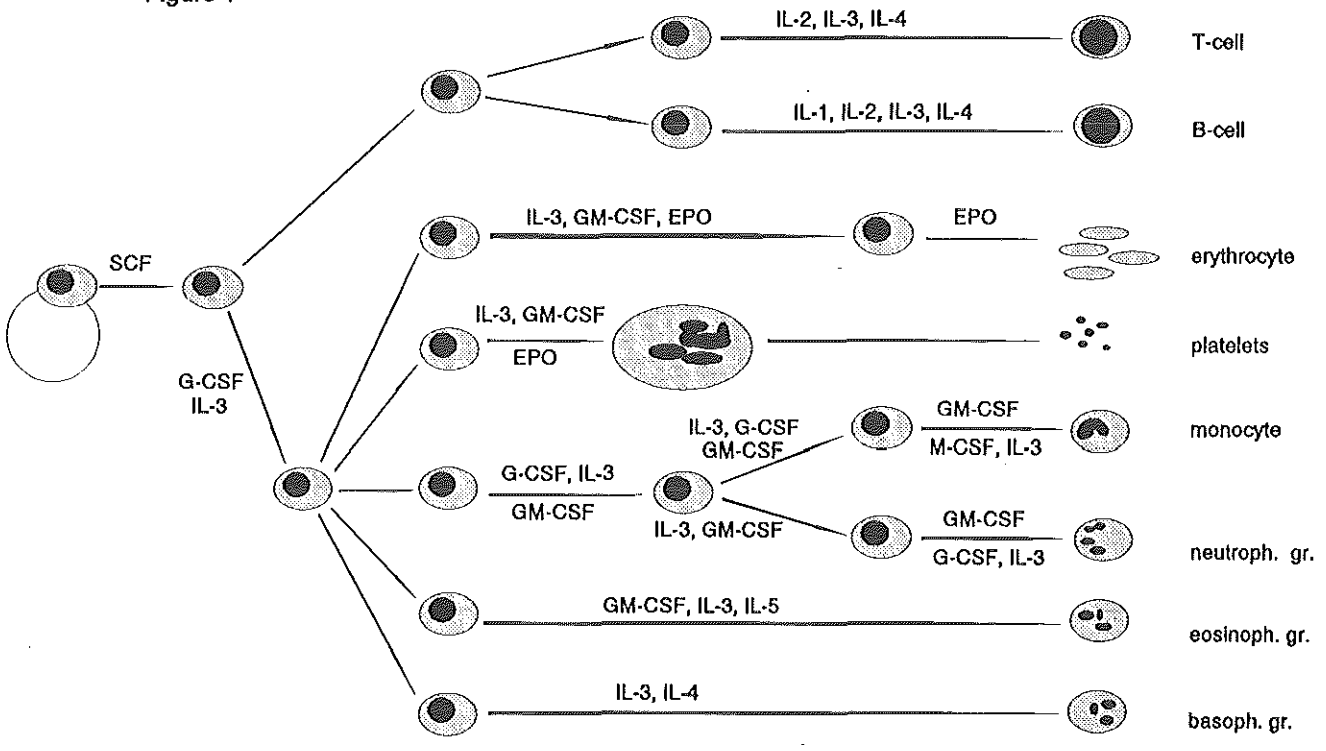
Hematopoiesis is a tightly regulated process of proliferation and differentiation that results in the presence of at least eight different types of mature and functional cells in the peripheral blood (reviewed by (152)). These cells are derived from a single hematopoietic progenitor cell in the bone marrow. In cycling progenitor cells, stochastic processes determine wether the cell will self-renew or commit itself to differentiation ((101) and refs therein). Much has been learned about progenitor cells by reconstitution experiments using lethally irradiated mice and bone marrow cells marked by retroviral integrations (38, 56, 101, 107). Following reconstitution, mature cells are derived from few progenitors that become extinct in some months. Although these progenitors are totipotent, most likely, they are already committed to differentiation. After 4-6 months mature cells are derived from a different set of progenitors that may remain stable for periods over 16 months. When stochastic processes are initially in favour of selfrenewal, a stock of totipotent noncommitted stemcells will be generated. If bone marrow is transplanted into secondary or tertiary irradiated recipients, mature cells continue to be derived from the same small number of progenitors even if several recipients are reconstituted with bone marrow of one donor. This indicates that stem cells have an apparently unlimited capacity of selfrenewal and proliferation (101, 107).

Even if large numbers of precursor cells are used in reconstitution experiments, few progenitors will produce mature cells, suggesting that few niches for early hematopoietic proliferation and differentiation are available in the stromal environment (101). Despite the fact that few clones contribute to stable long term hematopoiesis, fluctuations in progenitor activity can occur, that result in a temporal shift in the clonal upmake of peripheral blood cells (101, 107).

The control of cell numbers occurs via hematopoietic growth factors and cytokines able to influence each others functioning via a complex network of interactions (reviewed by (79, 152)). A detailed description of these networks is beyond the scope of this thesis, but to illustrate the fact, some (by no means complete examples) are mentioned. A simplified overview of targets of some growth factors and cytokines is given in figure 1.

Several growth factors have been characterized, many have pleiotropic effects and act at different stages of differentiation (126). Pleiotropic effects can be the result of the observed synergism that exists between many hematopoietic growth factors. Growth factors promoting proliferation and differentiation in the myeloid lineage are macrophage- and granulocyte colony stimulating factor (M-CSF, G-CSF), Granulocyte-macrophage-CSF (GM-CSF) and multi-CSF (IL-3) (reviewed by (43)).

Figure 1



Schematic overview of normal hematopoiesis Mature cells that occur in the blood are shown at the righthand side (gr. - granulocyte), the pluripotent stem cell is shown at the lefthand side. Some growth factors involved in proliferation and differentiation are indicated.
Il - Interleukine, CSF - colony stimulating factor, EPO - erythropoietin, SCF - stem cell factor.

These factors are produced by lymphocytes, macrophages, stromal cells, fibroblasts and endothelial cells. Multi-CSF or IL-3 mainly acts at the level of the multi-potent and early-committed stem cells of all lineages; lymphoid (B- and T-cells) as well as myeloid. GM-CSF is more potent than IL-3 in inducing terminal differentiation, but its capacity to synergize effectively with IL-3 makes GM-CSF a potent stimulator of proliferation in the myeloid lineage. The function of G-CSF and M-CSF is more restricted. G-CSF is involved in the differentiation of neutrophilic granulocytes, M-CSF in the differentiation of monocytes. Recently, a growth factor and its receptor were characterized specific for stem cells. SCF (stem cell factor or Steel factor) binds to a receptor encoded by the c-*kit* proto-oncogene (reviewed by (17)). Although SCF itself is not very potent, it synergizes dramatically with the previously mentioned growth factors in stimulating proliferation and differentiation in early totipotent stem cells.

The signal transduction of different growth factors via their cognate membrane receptor molecules seems to employ a common intracellular route. Transfection of the gene encoding the epidermal growth factor receptor (EGF-R) into a CSF dependent cell line allows that cell line to be stimulated by EGF (171, 204). This implies that expression of a growth factor receptor is the major determinant for the cells susceptibility for a certain growth factor.

Next to growth factors, cytokines have a clear influence on hematopoietic cell proliferation and differentiation. Their role in normal hematopoiesis is most likely limited but they play a crucial role in the respons of the hematopoietic system on external signals (126). Il-2 stimulates T-cells after mitogenic triggering while tumor necrosis factor (TNF) and Il-1 stimulate fibroblasts and endothelial cells to produce CSFs after bacterial infection. Some cytokines, like TGF-ß and interferons, have pleiotropic effects.

1.2 Leukemogenesis

Leukemia may occur when the tightly regulated process of proliferation and differentiation in hematopoiesis is disrupted. Depending on the differentiation potential of the leukemic cells, chronic leukemia is distinguished from acute leukemia. The latter is divided into subtypes depending on the lineage and differentiation stage of the leukemic cells. Classification of human leukemias has been standardized based on cell morphology by a group of french, american and british hematologists (14) and is known as the FAB classification.

Karyotype analysis of leukemic cells showed that distinct chromosomal aberrations are associated with specific leukemia subtypes. Therefore, it can be argued that genes affected by these aberrations play a role in

Table 1

disease	translocation	genes involved	
CML, AML, ALL	t(9;22)(q34;q11)	*abl* (#9)	*bcr* (#22)
APL	t(15;17)(q22;q12-q21)	*pml* (#15)	*RARα* (#17)
pre-B ALL	t(1;19)(q23;p13.3)	*pbx* (#1)	*E2A* (#19)
B-ALL	t(5;14)(q31;q32)	*IgH* (#14)	*Il-3* (#5)
B-CLL	t(14;19)(q32;q13.1)	*IgH* (#14)	*bcl-3* (#19)
B-lymphoma	t(11;14)(q13;q32)	*IgH* (#14)	*bcl-1* (#11)
B-lymphoma	t(14;18)(q32;q21)	*IgH* (#14)	*bcl-2* (#18)
B-lymphoma	ins(2;2)(p13;p11,2-14)	*rel* (#2)	
B-lymphoma	t(8;14)(q24;q32) t(2;8)(p12;q24) t(8;22)(q24;q11)	*IgH* (#14) *Igk* (#2) *Igλ* (#22)	c-*myc* (#8)
T-ALL	t(8;14)(q24;q11)	*TCRαδ* (#14)	c-*myc* (#8)
T-ALL	t(7;9)(q35;q34.3)	*TCRßγ* (#7)	*tan*-1 (#9)
T-ALL	t(1;14)(p32;q11)	*TCRαδ* (#14)	*scl/tal* (#1)
T-ALL	t(7;19)(q35;p13)	*TCRßγ* (#7)	*lyl* (#19)
T-ALL	t(10;14)(q24;q11) t(7;10)(q35;q24)	*TCRαδ* (#14) *TCRßγ* (#7)	*hox11* (#10)
T-ALL	t(11;14)(p15;q11) t(11;14)(p13;q11) t(7;11)(q35;p13)	*TCRαδ* (#14) *TCRßγ* (#7)	*Rhombotin* (#11p13) *Rhom-2* (#11p15)

Summary of characterized genes involved in chromosomal aberrations in leukemia.

leukemogenesis. These aberrations encompass deletions, inversions, duplications and translocations, of which, to date, the translocations are best studied at the molecular level. Localization of the cellular homologue of the viral oncogene v-*abl* on chromosome 9q34 led to the characterization of t(9;22) in CML and it was found that the translocation of c-*myc* into the immunoglobuline (*Ig*) locus results in high level expression of c-*myc* in Burkitt's lymphoma (BL).

Chromosomal regions containing the T-cell receptor (*TCR*) or *Ig* genes (chromosomes 14q11, 7q35 and 2p11, 14q32, 22q11, respectively) are frequently involved in translocations found in lymphoid malignancies other than BL (reviewed by (48, 65)). Molecular characterization of these translocations led to the isolation of some known and a number of novel genes, activated by translocation into the *TCR* or *Ig* loci (chapter 1.3).

Table 1 gives a summary of the genes that have been characterized to date as being involved in chromosomal aberrations in human leukemia.

In addition, translocation breakpoints not involving *Ig* or *TCR* genes have been cloned and characterized and these appeared to give rise to specific fusion genes, mechanistically analogous to the *bcr-abl* fusion gene in CML (Chapter 1.3).

Analysis of the genes involved in chromosomal aberrations of lymphoid and myeloid neoplasia, may eventually give more insight, not only into the process of leukemogenesis, but also into normal regulation of proliferation and differentiation in the hematopoietic system. In addition, specific chromosomal aberrations can be used as leukemia markers. T(9;22) in CML can be found in cells of all lineages, indicating that the translocation occurred in an early totipotent progenitor cell. Other leukemias have chromosomal aberrations only in restricted lineages. This suggests that the aberrations occurred in cells already committed to that lineage but it can not be ruled out that the nature of the mutation in a multipotent progenitor drives differentiation of its derivatives into one lineage.

In a large number of leukemia patients no chromosomal aberrations are detected. This is not surprising as a gross chromosomal rearrangement is but one out of many possibilities to mutate a gene. The emphasis on genes activated by chromosomal rearrangements is solely due to the fact that they are clearly marked by the cytogenetic aberration and they are relatively easily accessible through improved cloning technology .

1.3 Characterization of genes involved in specific chromosomal aberrations in human leukemia

Control of cell proliferation and differentiation is regulated at several levels: among the factors involved are growth factors and growth factor receptors, mediators of the signalling pathway, specific transcription factors as well as cell cycle regulatory proteins. One can speculate that interference with any of these levels may lead to leukemia (see also (20)). The leukemogenic effect of several oncogenes has been established in mice or chicken animal models. This chapter will focus on genes implicated in chromosomal aberrations in human leukemia.

1.3.1 Ligands and receptors

A large number of growth factors and inhibitors are involved in the regulation of proliferation and differentiation of hematopoiesis (Chapter 1.1, (49, 152)). However, disregulation of these factors or their receptors has not been implicated as a direct and major cause of leukemia.

Il-3

Nevertheless, elevated expression of IL-3 characterizes a subtype of B-ALL with eosinophilia and t(5;14)(q31;q32) (91). The translocation fuses the *IgH* gene on chromosome 14 to the *Il-3* gene on chromosome 5 (77). Breakpoints occur 5' of the transcription start site of *Il-3* and leave the encoded protein intact. As a result, IL-3 levels in serum had markedly increased due to the juxtaposition of *Il-3* to the IgH enhancer (149). The observed increase in eosinophilic granulocytes may be caused by high levels of IL-3 in serum but it can not be ruled out that the eosinophils are leukemic cells as well and carry the translocation. Experimental data suggest that precursor B-cells can give rise to myeloid cell types ((115) and refs therein).

High level IL-3 production by otherwise normal cells stimulates both proliferation and differentiation. Therefore, experimentally induced high levels of IL-3 in mice fail to produce leukemia but rather effectuates a myeloproliferative disorder with large numbers of mature myeloid cells in peripheral blood (40, 210). Assuming that mouse model systems can be extrapolated to the human situation, it is likely that another mutation will accompany the deregulation of IL-3 in this subtype of B-ALL.

tan-1

The only hematopoietic growth factor receptor that was found to be involved in leukemia progression and MDS is the M-CSF receptor or c-FMS protein (chapter 1.5.2, (176)). To date, rearrangement or mutation of growth factor receptors was not found to be associated with a specific subtype of leukemia. However, a receptor with a so far unknown ligand is truncated by t(7;9)(q34;q34.3) in human T lymphoblastic leukemia (61). The breakpoint on chromosome 7 is located in the TCRß gene; the breakpoint on chromosome 9 is located in the middle of a gene that was initially named *tcl*-3 (175), but was renamed *TAN*-1 (translocation associated *Notch* homolog) as it is highly homologous to the *Notch*-1 gene of *D.melanogaster*. The predicted TAN-1 protein has a hydrophobic putative transmembrane domain and, at the N-terminal side, 39 cysteine rich repeats: 36 epidermal growth factor (EGF) repeats and 3 *Notch/lin*-12 repeats. It has been shown for NOTCH that this domain is extracellular and binds to the EGF repeats of another transmembrane protein, DELTA (63). The presumed intracellular part of TAN-1 contains 6 repeats that also occur in the yeast cell cycle regulator genes *cdc10* and *SWI6*. These repeats most likely are involved in protein-protein interactions (137); i.e. in *Drosophila* the intracellular protein ENHANCER OF SPLIT interacts with NOTCH possibly via the *cdc10/SWI6* repeats. ENHANCER OF SPLIT shows homology to the ß subunit of mammalian G proteins. The extensive homology between TAN-1 and NOTCH argues that TAN-1 has a role in transmitting an extracellular signal, possibly via G proteins. The signal may be a diffusible ligand or proteins that play a role in cell-cell

contact homologous to NOTCH. Due to the translocation the extracellular domain is lost and the truncated receptor is expressed at relatively high level. Loss of the extracellular domain can have two possible effects: (*i*) It releases the cells from an inhibitory cell-cell interaction. Loss of cell-cell contact has been implicated before as one of the steps in tumorigenesis. (*ii*) It renders TAN ligand independent, resulting in a constant activation of the signalling pathway by the intracellular domain.

Tan-1 expression is not lymphoid specific, it is expressed in many fetal and adult tissues, relatively high in fetal spleen, adult thymus, brain and lung (61).

1.3.2 Signalling pathway

Several distinct signalling pathways exist in eukaryotic cells that may or may not interact with each other. Membrane coupled GTP-binding proteins (like RAS) and kinases (like SRC) as well as cytoplasmic and nuclear kinases (RAF, ABL) are part of the signalling pathway and have a role in tumorigenesis (reviewed by (37)). However, only $p56^{lck}$ and the *bcr-abl* fusion gene were shown to be specifically involved in human leukemogenesis.

bcr-abl

Due to the t(9;22)(q34;q11), a BCR-ABL fusion protein is formed that has enhanced kinase activity compared to the *wt* ABL protein (119, 205). The t(9;22) occurs in 95% of CML, 6-25% of ALL and 3% of AML patients (121). Almost all chromosome 22 breakpoints in CML and half of these in ALL are clustered in a 5.8 kb DNA fragment, the major breakpoint cluster region (M-bcr), in the middle of the *bcr*-gene (78, 81, 86). The remaining chromosome 22 breakpoints in ALL and few breakpoints in CML occur in the large (68 kb) first intron of the *bcr*-gene, which is named minor breakpoint cluster region (m-bcr) (88, 185). On chromosome 9, breakpoints occur in a large stretch of DNA (~200 kb), which appeared to be the first intron of the c-*abl* oncogene (16, 52). The resulting *bcr-abl* fusion genes encode chimeric *bcr-abl* transcripts with an in frame fusion of the two reading frames (88, 188). Breakpoints in the M-bcr give rise to a $p210^{bcr\text{-}abl}$, breakpoints in the m-bcr to a $p190^{bcr\text{-}abl}$.

The *bcr* and *abl* genes are expressed in all cell types, which suggests a general role of these proteins in cell metabolism rather than tissue specific functions.

The ABL protein is a nonreceptor tyrosine kinase, containing a SH3-SH2 (SRC homology regions 2 and 3) domain N-terminal of the kinase domain (reviewed by (116)). The SH2 domain is a protein-protein interaction domain as it binds to phosphorylated tyrosine residues. In general, it is postulated that the SH3 domain can interact with the cytoskeleton, but in

the ABL SH3 domain probably binds to a cellular factor that inhibits its tyrosine kinase activity: specific deletion of the SH3 sequences results in enhanced tyrosine kinase activity of the mutant protein and overexpression (>500x) of the normal ABL protein also leads to enhanced kinase activity most likely by depletion of the cellular factor binding to SH3 (166). Breakpoints in *abl* occur mostly 5' of the second exon encoding the SH3 domain. In few cases however, part of the SH3 domain was removed without any notable effects on the enhanced kinase activity of the chimeric BCR-ABL protein or on the leukemic cells (191, 199). This suggests that the cellular factor can not bind SH3 in the BCR-ABL fusion protein.

A C-terminal part of BCR has strong homology to the catalytic domain of GAP proteins (57), which suggests that BCR may have a role in a signal transduction pathway via a G-protein. This putative function of BCR is not relevant for the BCR-ABL fusion protein, since it is not present in the BCR-ABL translocation product. The first *bcr* exon, present in BCR-ABL fusion proteins, contains a novel type of serine/threonine kinase domain (141). However, the capacity of BCR sequences encoded by the first exon to bind to the ABL SH2 domain appeared to be essential for the transforming potential/enhanced tyrosine kinase activity of the BCR-ABL protein (167). Fusion of BCR to ABL most likely influences the conformation of ABL, thereby preventing the as yet unidentified cellular factor to bind to SH3, which leads to enhanced tyrosine kinase activity.

The leukemogenic potential of *bcr-abl* was tested in two ways. (*i*) Transgenic mice were generated containing a *bcr-abl* fusion gene encoding $p190^{bcr\text{-}abl}$, under control of the metallothionein-1 promoter for non-targeted expression (85). 8/10 transgenic mice died 10-58 days after birth from leukemia (myeloid and B-lymphoid). Just like in CML and blast crises of CML, the morphology of the cells showed variable degrees of differentiation. (*ii*) Consistent results were obtained by transplantation of bone marrow cells infected *in vitro* with retrovirus harbouring a *bcr-abl* gene encoding $p210^{bcr\text{-}abl}$ (50, 60, 108). Some mice developed myelomonocytic and granulocytic leukemia, that can be compared with CML, others developed B-ALL.

lck

Another nonreceptor tyrosine kinase, $p56^{lck}$, appears to be involved in t(1;7)(p34;q34), described in patients with T-ALL. The *lck* gene is translocated into the TCRß locus, which results in enhanced expression from the gene's own promoter (36). $p56^{lck}$ is a SRC-like kinase, expressed almost exclusively in lymphoid cells (reviewed by (184)). It complexes with CD4 and CD8 and most likely functions in the signalling pathway activating T-cell proliferation upon antigen recognition. $p56^{lck}$ kinase activity is stimulated by addition of IL-2 (93). Constitutively activated $p56^{lck}$ enhances T-cell responsiveness (2) and induces thymic

tumorigenesis (1). In addition, enhanced levels of $p56^{lck}$ have been observed in a number of Burkitt's lymphoma cell lines (102), suggesting cooperation between *lck* and c-*myc* in tumor progression (chapter 1.5).

1.3.3 Cell cycle regulation

Although many proteins involved in tumorigenesis, like kinases and transcription factors, may influence the regulation of the cell cycle indirectly, only very few cancer associated proteins are directly involved in the decision to start cell division. The best known example is *Rb*, a recessive oncogene involved in retinoblastoma, osteosarcoma and an increasing number of other tumors (reviewed by (140)).

Cell cycle control has been investigated in detail in yeast. Only recently, the isolation of mammalian homologues of these yeast genes disclosed some aspects of human cell cycle control (131). The kinase $p34^{cdc2}$, Rb and cyclins appear to be key factors. $p34^{cdc2}$ and Rb are expressed throughout the cell cycle, while the induction of several short lived cyclins is crucial for induction of cell cycle entry.

bcl-1/PRAD1

The t(11;14)(q13;q32) has been found frequently in centrocytic lymphoma and rarely in other types of B-cell neoplasia (208). On chromosome 14q32 the Ig heavy chain locus is involved and on chromosome 11q13 a locus has been cloned that was designated *bcl-1*. However, no transcriptionally active region was found near the translocation breakpoint (197).

The same chromosomal region is rearranged in a subset of benign parathyroid tumors. An 11q13-q15 inversion places a locus on chromosome 11q13 under control of the 5' regulatory region of the parathyroid hormone gene (178). Characterization of the 11q13 locus shows that it contains a gene encoding a protein of 295 amino acids that has 44% to 59% similarity to the family of cyclins (155). The gene was named *PRAD1* (*parathyroid adenomatosis*) and probably represents a new and different class of the cyclin gene family. *PRAD1* is highly conserved and is expressed in many tissues. Expression varies during the cell cycle and is elevated in tumor tissue. The PRAD1 protein associates with $p34^{cdc2}$ thereby inducing $p34^{cdc2}$ kinase activity.

Although the activation mechanism of *PRAD1* may be specific for parathyroid tumors, enhanced expression of the gene may be involved in several types of cancer. A region of chromosome 11q13, delimited by the oncogene *int*-2 and the *bcl*-1 locus, is amplified in a variety of epithelial tumors, notably breast and esophageal carcinoma, squamous carcinoma of the head and neck and at least in one case of hepatocellular carcinoma (122). *PRAD1* is consistently coamplified and expressed at elevated levels

in all tumor samples tested. Long range mapping analysis showed that *PRAD1* is located between *int*-2 and the *bcl*-1 locus at a distance less than 130 kb from *bcl*-1 (122, 179). Moreover, *PRAD1* is abundantly expressed in lymphomas carrying a t(11;14) and not in similar lymphomas without this translocation. This makes *PRAD1* a good candidate for the oncogene involved both in epithelial tumors and in t(11;14) B-cell lymphoma (179).

Withers et al. (209) hypothesized that a gene activated over a relatively long distance by translocation into the *IgH* locus, will have a GC-rich promoter region. Chromosome walking from the translocation breakpoint to the first CpG-island that functions as the promoter of a gene abundantly expressed in t(11;14) cells, led to the isolation of the *PRAD1* gene. The translocation breakpoint is located at a distance of 120 kb from *PRAD1*. Withers et al. (209) analyzed *bcl-1/PRAD1* transcripts in three t(11;14) cell lines and found truncations in the 3' UTR in two of them. However, Rosenberg et al. (179) found abundant expression of normally sized transcripts in 9 different patient samples.

The data strongly suggest that enhanced expression of a cyclin-related protein is involved in tumors of various origin among which hematopoietic cells.

bcl-3

The *bcl-3* gene implicated in B-cell chronic lymphocytic leukemia may also be involved in regulation of the cell cycle, although its function is not clear. The translocation breakpoints of t(14;19)(q32;q13.1) in CLL occur in the Ig heavy chain gene on chromosome 14q32 and in the *bcl-3* locus on chromosome 19q13.1 (147, 148). The translocation breakpoints cluster in the GC-rich 5' regulatory region of *bcl-3* and do not interrupt the transcriptional unit (164). *Bcl*-3 expression is clearly enhanced by the translocation. The gene encodes a 47 kD basic protein with a proline rich (25%) N-terminus followed by seven copies of a 33-37 amino acid repeat, that occurs in cell cycle associated proteins like *S. pombe* cdc10 and *S. cerevisiae* SWI4 and SWI6 and in the Notch and TAN transmembrane proteins of *D. melanogaster* and man, respectively. The C-terminus of BCL-3 contains several SPXX motifs that may have an ancillary DNA binding function. These domains are well conserved between man and mouse (18).

BCL-3 expression is cell cycle dependent as stimulation of T cells with mitogens induced *bcl-3* expression within 15 minutes. Expression was maximal between 2 and 8 hours after stimulation (164). The presence of a proline rich region that may function as a transcription activation region and the homology of BCL-3 with CDC10, SWI4 and SWI6 suggests that BCL-3 may function as a transcription factor and could induce expression of proteins responsible for cell division.

1.3.4 Transcription regulation

Several classes of transcription factors have been characterized to date. Many transcription factors are active as dimers and some dimerization domains are well defined like amphipathic helices (leucine zipper and helix-loop-helix motif) and a Zn-finger like motif (LIM motif) (reviewed in (100)). Domains responsible for transcription activation may be either acidic or proline/glycine rich (153). Well defined DNA binding domains encompass the homeobox, a stretch of basic residues combined with a leucine zipper or a helix-loop-helix domain and Zn-fingers (82). Experimental data show that DNA binding domains and transcription activation domains can be combined at random and function more or less independently (32, 92).

Deregulation of transcription factors can induce and/or repress a whole array of genes which may reprogram the cell and conceivably many oncogenes encode this type of nuclear proteins (132).

basic stretch-helix-loop-helix proteins.

In an increasing number of transcription factors the basic stretch-helix-loop-helix (bHLH) motif is recognized as a domain involved in both DNA binding (by the basic residues) and dimerization (by the helices) (13, 22, 156). All bHLH proteins bind DNA as dimers. Small variations in the distance between basic stretch and the first helix, and in the length of the loop, may determine specificity for heterodimer partners or homodimer formation. In addition to the bHLH domain, a leucine zipper (124) can be present. The E-box (CACGTG), seems to be a common DNA binding site for members of the bHLH proteins, among which the ubiquitous transcription factors E12/E47 encoded by the *E2A* gene ((128, 157) and refs therein). Transcription factors with a bHLH domain implied in leukemogenesis encompass c-MYC, SCL/TAL, LYL and E12/E47.

c-*myc*

The most elaborately studied association between an oncogene and a specific type of leukemia is the invariant translocation of c-*myc*, located on chromosome 8, into one of the Immunoglobulin (*Ig*) genes in Burkitt's lymphoma (BL) (reviewed by (139)). The Ig heavy chain locus on chromosome 14 and the Ig κ- and λ-light chain loci on chromosome 2 and 22 respectively, are involved in t(8;14), t(2;8) and t(8;22). Breakpoints are scattered around c-*myc* both 3' and 5' of the gene and often occur in its first exon and first intron; all breakpoints will leave the protein coding region intact. Breakpoints far upstream of the *myc* transcription regulatory domains are often accompanied by mutations in the first exon, which cause abrogation of a transcription elongation block, present in this region. However, these mutations are not sufficient for transformation. MYC expression in the cell is tightly regulated, the short half life of the transcript and the protein are not altered in BL, but transcription is

markedly upregulated. Expression of the Ig-gene involved in the translocation is essential for enhanced *myc* transcription even if the translocation breakpoint occurred at a large distance from the *myc* gene. Thus, translocation into the *Ig* loci is the essential oncogenic event.

The c-*myc* gene is not only involved in translocations with Ig-genes, translocation into the *TCRα/δ* locus has also been reported as the molecular result of t(8;14)(q24;q11) associated with T-ALL (26).

MYC is a 56 kD nuclear, DNA binding protein. It has a long N-terminus that may function in transcription activation, the C-terminus contains a bHLH motif and a leucine zipper (LZ). Recently, it was shown that a specific protein, designated MAX, dimerizes via the bHLH and LZ motifs (22). MAX is a small protein of 160 amino acids containing a bHLH domain and a LZ at its N-terminus and an acidic region at its C-terminus. MYC exclusively dimerizes with MAX, on the other hand, MAX dimerizes with MYC but also forms homodimers. Only MYC-MAX heterodimers or MAX homodimers show high affinity for the E-box (CACGTG), which is also the MYC binding site (21, 111, 173). What genes can be transcriptionally regulated by MYC is not yet known. It can not be excluded that MAX is an important transcription factor regulating cell proliferation, while MYC regulates the function of MAX through the formation of heterodimers.

MYC is expressed in all cycling cells. It can be induced to relatively high levels in quiescent cells by serum stimulation after which a basic expression level is maintained throughout the cell cycle (reviewed by (47)). It is speculated that high levels of MYC prevent the leukemic cells from resting. Although besides BL no other tumor is exclusively associated with high levels of c-*myc* expression, aberrant expression of this gene occurs in many tumor types. Transgenic mouse models indicate that high level expression of *myc* predisposes for cancer but is in itself not sufficient (84, 125, 201).

tal/scl and *lyl*

Other bHLH proteins are involved in T-cell leukemia and show a large degree of homology with each other. The *scl* or *tal* gene is involved in t(1;14)(p34;q11) in T-ALL, the *tal*-2 gene is involved in t(7;9)(q34;q32) T-ALL and the *lyl* gene is deregulated by t(7;19)(q34;p13) in a T-cell leukemia cell line.

T(1;14)(p32;q11) occurs in 1-2% of T-ALL and was also found in a stem cell leukemia (SCL) with lymphoid and myeloid manifestations (89, 123). In both cases the chromosome 14 breakpoint was located in the TCR*δ* gene (11, 41, 66). Cloning and characterization of sequences on chromosome 1 revealed that the same gene is involved in T-ALL as well as SCL. The gene was reported as *tal* by Chen et al., as *scl* by Begley et al. and as *tcl5* by Finger et al. Rearrangement of the *scl* gene was also reported in a case of T-ALL with a t(1;7)(p32;q35) (67).

The *tal/scl* gene has a complicated genomic organization. (5). In T-ALL, translocation breakpoints are located near a cryptic heptamer sequence in exon III (41), upstream of the ATG startcodon in exon IV. In SCL the translocation occurred in exon III as well as in the 3' untranslated region of the gene (15). Surprisingly, 'breakpoints' in exon III were also detected in 13/50 T-ALLs with a normal karyotype and in some T cell lines (6, 33). In these cases a remarkably precise deletion had occurred that measures 90 kb. This deletion mimics the t(1;14), only now *tal/scl* is under the transcriptional control of a novel gene, designated *sil* (*scl* interrupting locus) (6). The *sil* gene is exclusively expressed in hematopoietic tissues and cell lines and the putative SIL protein shows no homology with any other protein sequences (7).

Both aberrations result in a normal TAL/SCL protein, expressed at elevated levels, because *tal/scl* expression in normal T-cells is low or absent. Normal *tal/scl* expression seems to be restricted to early myeloid progenitor cells. Highest levels of expression were detected in fetal liver and in bone marrow recovering from chemotherapy, which indicates that TAL/SCL may be involved in proliferation of early hematopoietic cells (11). In addition, it has been suggested that TAL/SCL is involved in erythroid differentiation (75, 203).

The *TCRß* gene on chromosome 7 and a novel gene named *lyl*, located on chromosome 19 are involved in t(7;19)(q35;p13) (151). This translocation was cloned and analyzed in a single T-cell line. The *lyl* transcript gives rise to a 33 and a 29 kD protein. These proteins are initiated on two in frame start codons present in exon 1 and 2, respectively. The translocation occurs in the first intron, the transcription regulatory sequences are replaced by those of *TCRß* resulting in the expression of a truncated *lyl* transcript encoding only the 29 kD protein (151). Analysis of *lyl* expression in mouse tissues and cell lines revealed that the transcription is restricted to the hematopoietic system (120, 203). Highest expression levels are found in immature B-cell lines, low level expression is found in myeloid cells. Markedly, no expression can be detected in thymus or in T-cell lines.

The C-terminus of TAL/SCL contains the bHLH motif. Specific formation of heterodimers between TAL/SCL and the ubiquitous bHLH transcription factor E12/E47 was demonstrated (95) resulting in binding of the heterodimers to the E-box. Dimerization and DNA binding is dependent on the bHLH motif. The LYL protein contains a bHLH motif which is highly homologous to TAL/SCL. It is not known whether LYL can form heterodimers with E12/E47 as well. As a group of homologous muscle specific bHLH proteins is involved in the regulation of myogenesis through dimerization with E12/E47 (128), it is an attractive idea that a distinct group of hematopoietic bHLH proteins, among which TAL/SCL and LYL, dimerizes in a similar way with the ubiquitous factors E12/E47 to regulate

hematopoietic differentiation. In T-ALL, LYL and TAL/SCL may act on the same set of target genes.

The *tal*-2 gene is located at a distance of 30 kb from the t(7;9)(q34;q32) breakpoints (211). In the bHLH domain the TAL2 protein is highly homologous to TAL1/SCL and LYL but outside this domain the proteins diverge. Contrary to *tal-1/scl* and *lyl*, *tal-2* expression was not detected in the hematopoietic system, only in testis. In a cell line harbouring t(7;9)(q34;q32) *tal-2* was actively transcribed, which suggests that this gene is activated through translocation. Possibly, the TAL2 protein can act on the same target genes as TAL1/SCL and LYL.

E2A-pbx

Interestingly, the *E2A* gene, encoding the bHLH transcription factors E12/E47 that dimerize with SCL, appeared to be involved in lymphoid malignancy itself (150). *E2A* is located on chromosome 19p13.3 and is disrupted by t(1;19)(q23;p13.3) in 30% of pediatric pre-B ALL patients. Although the E12 and E47 transcription factors were first identified as proteins binding the κE2 site of the Ig-κ gene in B-cells, they are expressed ubiquitously and complex with several other bHLH proteins (as mentioned before)(156, 157). The N-terminal part of the proteins is rich in proline and glycine and may function as the transcription activating domain, the C-terminal part of the proteins contains the leucine zipper and bHLH domain. The translocation breakpoints consistently occur in the 3' part of the *E2A* open reading frame, between the leucine zipper and bHLH domain (96, 104, 160). The 1q- chromosome carrying the translocation product containing the *E2A* bHLH domain is lost in the majority of pre-B ALL patients (150). On chromosome 19p+, the 5' part of *E2A* is fused to a novel gene designated *pbx-1* (104, 160). *Pbx-1*, located on chromosome 1q23, was shown to contain a homeodomain, a DNA binding entity first identified in genes involved in *Drosophila* and mammalian development (see below)(28, 177). In the E2A-PBX fusion protein the homeodomain replaces the bHLH DNA binding domain of E2A. The molecular consequence of t(1;19) is highly consistent as an identical chimeric transcript of *E2A* and *pbx-1* could be detected in 37/38 t(1;19) bone marrow samples (96).

It can be argued that t(1;19) results in a chimeric transcription factor with the transcription activation domain of E12/E47 that will bind to the recognition sequence of the homeotic protein. As *pbx-1* expression could not be detected in non leukemic hematopoietic cells, the recombinant transcription factor may activate genes normally silent during hematopoiesis, whose activation contributes to the leukemic process (104, 160). Alternatively, the E47-PBX transcription factor may interfere with expression regulation of a gene under normal control of a PBX related homeobox factor. It will be interesting to know what genes are (in)activated, since these may be important for regulation of

hematopoiesis. As the homeobox of PBX is highly homologous to the yeast mating type protein MATa1 (158), PBX may bind a related recognition site.

Homeobox genes.

The homeobox is a 61 amino acid domain, adopting a helix-turn-helix structure, that functions in DNA binding (183). Homeobox genes were first identified in *D. melanogaster* where a large number of different homeobox genes has an important role in the regulation of *Drosophila* development. Homeobox genes are well conserved during evolution and play an important role in tissue specific transcription regulation in mammals clearly reminiscent of developmental control in *Drosophila* (28, 113). Increasing evidence shows that this family of genes is also involved in regulation of differentiation in the hematopoietic system, as several homeobox genes are expressed in specific hematopoietic cell lineages and differentiation stages (4, 118, 138, 186).

HOX11

In addition to *pbx-1*, the homeobox gene *HOX11* was found to be involved in lymphoid neoplasia.

Approximately 7% of T-ALLs possess the t(10;14)(q24;q11), which involves the *TCRα/δ* locus on chromosome 14 (217). The locus on chromosome 10 was initially named *tcl3* but isolation and characterization of the gene demonstrated that this translocation affects a homeobox gene which was subsequently designated *HOX11* (58, 83, 136). *HOX11* is also involved in t(7;10)(q35;q24) T-ALL, where the gene is translocated into the *TCRß/γ* gene (110). Moreover, rearrangements in the 5' part of *HOX11* were found in 2/56 pre-B ALLs and in 2/50 T ALLs with no apparent chromosomal aberrations (58). This indicates that *HOX11* may be involved in T- and B-cell neoplasia and that an alternative mechanism of activation exists. This may be analogous to the activation of the *tal/scl* gene by deletion of 5' regulatory sequences and juxtaposition to the *sil* gene. In all 21 cases analyzed, the breakpoints occur in a 15 kb region immediately upstream of the sequences coding for the *HOX11* 2.1 kb transcript (58, 83, 110, 136).

Expression of *HOX11* was only detected in human liver and not in thymus or T-cells (83). The leukemogenic effect may be due to ectopic expression as was argued for the *pbx-1* homeobox gene. Genes may be activated by HOX11 that should remain silent.

HOX11 can not be classified into any of the homeobox gene clusters that are homologous to *Drosophila* genes. The protein has a glycine rich N-terminus that may function in transcription activation and the homeodomain at its C-terminus is most homologous to the homeobox of the mouse protein Hlx, although this is clearly a different homeobox gene (4). HOX11 has little homology with PBX (104, 160). Possibly, HOX11

could interfere with the function of human HLX by binding to the same recognition site. The mouse *Hlx* gene is expressed in several tissues both fetal and adult but in the hematopoietic system the gene has a pronounced expression in the myeloid-macrophage lineage and in early B cells, suggesting a role for this protein in hematopoietic differentiation. *Hlx* is well conserved in several species and may have two human homologues.

Rhombotins

The members of the *rhombotin* family do not contain DNA binding motifs or transactivation domains, instead they seem to be merely dimerization domains. However, it is hypothesized that their protein products can modulate certain transcription factors, perhaps homeobox proteins, through the formation of heterodimers.

Translocations in T-ALL often occur on chromosome 11p. Separate loci are involved, 11p13 in t(11;14)(p13;q11) / t(7;11)(q35;p13) and 11p15 in t(11;14)(p15;q11), of which t(11;14)(p13;q11) is most common. The loci are juxtaposed to the *TCRδ* gene on chromosome 14 or *TCRß* on chromosome 7 (23, 182). The *11p15* or *Tgt*-1 gene was cloned and characterized first (25, 146). The translocation breakpoint is located 5' of the gene and transcription starts from the gene's own promoters. Although the gene was well expressed in T cell lines carrying the t(11;14), normal expression of this gene in the mouse appeared to be restricted to the developing nervous system with a pronounced segmental expression pattern in the rhombomeres (76, 145). Therefore the *11p15* gene was renamed the *rhombotin* gene (*Rhom-1*). Expression of this gene can also be found in other cells of the CNS and in neural crest derived cells. Careful examination revealed very low expression levels in other tissues, including thymus (27).

As high level expression was found in postmitotic neurons, it seems likely that Rhom-1 is involved in differentiation regulation rather than proliferation (76). Sequence analysis revealed that all the *Rhom-1* transcript encodes for, are two clusters of cysteins. These represent two LIM motifs (25, 146, 174), a protein-protein interaction domain also present in the *C. elegans* cell lineage specific homeobox proteins lin-11 and mec-5 and the vertebrate transcription factor Isl-1 (68, 106). The function of Rhom-1 may be to modulate the function of homeobox proteins that contain a LIM motif by forming heterodimers. Such a function would be analogous to the function of Id in the regulation of myogenic bHLH proteins (193).

In order to isolate potential factors interacting with its LIM motif, low stringency hybridizations with *Rhom-1* sequences representing the LIM motif were performed. This led to the isolation of two *Rhom-1* homologues (24). One of these, named *Rhom-2*, was 50% homologous and contained a double LIM motif without a possible DNA binding domain

as well. *Rhom-2* appeared to be located at chromosome 11p13. Comparison of the 11p13 locus involved in t(11;14)(p13;q11) and the 5' part of *Rhom-2* revealed that all these translocation breakpoints occur in a region of 25 kb upstream of *Rhom-2*. In addition, characterization of the 11p13 locus led directly to the isolation of the same gene, which is also named *Tgt-2* (181).

Like *Rhom-1, Rhom-2* is extremely well conserved during evolution but its expression pattern is different (24). Transcription of *Rhom-2* starts in early development and continues throughout adult life in almost all tissues. Remarkably, the expression in thymus is exceptionally low. Although the general function of Rhom-1 and Rhom-2 in regulation of differentiation may differ, it seems likely that they interfere with the same regulatory factors during T cell leukemogenesis. It can be hypothesized that a LIM motif containing transcription factor plays a crucial role in T cell differentiation and it will be interesting to identify this factor. It may be part of the growing number of homeobox proteins, expressed in lineage and differentiation stage dependent fashion in hematopoietic cells.

Steroid hormone receptor family.

Steroid hormone receptors are a specific class of ligand regulated transcription factors designed for an effective response of the transcription machinery of the cell to external signals (see review by (62)). They also include the thyroid hormone and retinoic acid receptors. Evidence is accumulating that some members of this receptor family may have a function in the control of hematopoietic differentiation. V-erbA the retroviral, oncogenic counterpart of the thyroid hormone receptor (TR) is able to block differentiation in chicken erythroblasts ((215) and refs therein) and the retinoic acid receptor type α (RARα) is impaired in acute promyelocytic leukemia.

pml-RARα

Acute promyelocytic leukemia (APL) is specifically associated with t(15;17)(q22;q12-q21) (127). The breakpoint on chromosome 17 maps in the first intron of the gene encoding the retinoic acid receptor type α (*RARα*) (29, 54, 134). Although DNA encoding some N-terminal sequences of RARα is separated from the rest of the gene by the t(15;17), exons encoding the DNA binding domain, the dimerization and ligand binding domains, remain intact and are fused to a novel gene on chromosome 15. This novel gene was called *pml* (promyelocytic leukemia) and analysis of its cDNA sequence predicts that the PML product is a transcription factor as well (55, 103, 165). The N-terminus contains a proline rich region, followed by a cystein rich region, a potentially α-helical region and a C-terminus with several SPXX motifs. The cystein rich region contains three cystein clusters that have homology to several DNA binding proteins and represent a new type of Zn-finger motif. The *pml*

cDNAs reported by de Thé et al. and by Kakizuka et al. differ in their 3' end, most likely due to alternative splicing. Alternative splicing of several *pml* exons has been found (73). Likewise, different *pml-RARα* fusion cDNAs have been found, though in both cases the breakpoint in *pml* occurred between the α-helical region and the SPXX motifs.

Leukemic APL cells can be forced to differentiate by administration of high levels of retinoic acid (RA) both *in vivo* and *in vitro* (42). Therefore, the effect of RARα and PML-RARα on transcription activation was assayed in the presence and absence of RA, using several reporter genes (55, 103). PML clearly alters the function of RARα. Glass et al. (72) showed that several cell type specific proteins associate with RARα and in that way regulate RARα transactivation. Recent evidence indicates that at least one of these factors is the RAR related factor RXR (114, 216). PML may affect this association. However, to evaluate the effect of the modulation of RARα function by PML, it will be essential to know the RARα and PML-RARα targets in APL cells.

Alternatively, the effect of RA on APL cells may be due to the abnormal RA dependence of PML. It has been shown that transcription factors (c-MYB, c-MYC, E1A, c-FOS and C/EBP) that do not belong to the steroid hormone receptor superfamily, can be made hormone responsive when they are linked to the ligand binding domain of the estrogen receptor (35, 59, 169, 194, 198). The resulting hybrid transcription factors bind to their cognate response elements in response to relatively high doses of estrogen. The same situation may have been created for the PML transcription factor by t(15;17). The normal function of PML may have been blocked by its fusion to the ligand binding domain of RARα and can only be restored by high doses of RA. Once the target sequences for PML are found, RA dependence of PML-RARα can be assayed. Normal expression of *pml* can be detected in all adult and fetal tissues of the mouse (73).

Both hypothesis, impairment of RARα function by PML and RA dependence of PML function, assume a dominant negative effect of the fusion product, which suggests that RARα or PML may function in complexes. The situation could be comparable to the dominant negative oncogene p53 (130).

NFκB complex.

Nuclear factor κB (NFκB) was originally described as a factor binding to the enhancer of the Ig κ light chain gene but now it is clear that the NFκB complex has a more general function (8, 129). NFκB consists of two different proteins, p50 and p65, that share a homologous DNA binding and dimerization domain. This domain is also present in proteins of the *rel* oncogene family and in the *dorsal* gene of *D. melanogaster* (reviewed by (71)). The DNA binding and dimerization domains of NFκB are distinct from those of other transcription factors. Among the proteins that

complex with NFκB is IκB, a factor that inhibits NFκB function by retaining the protein in the cytoplasm. IκB can also bind the REL proteins (112) and a gene with a function similar to IκB, *cactus*, is present in *D. melanogaster* (51). Several lines of evidence suggest that interference with NFκB function can lead to B-cell malignancies.

rel

In human B cell follicular and diffuse large cell lymphomas aberrations often occur at chromosome 2p11.2-14, a region that harbours the c-*rel* gene (34, 213). In birds, it has been shown that the v-*rel* oncogene is involved in the generation of lymphomas. V-*rel* is a truncated version of c-*rel* that lacks the N-terminal part, but contains the NFκB homologous C-terminal part with its dimerization and DNA binding domains (reviewed by (71)).

In a human cell line of a pre-T diffuse large cell lymphoma an ins(2;2)(p13;p11.2-14) was analyzed (135). Next to the normal 12 kb c-*rel* transcript an abundant transcript of ~2 kb was detected with 5' c-*rel* probes. A breakpoint in intron 5 was cloned and it was shown that a fusion gene had been formed containing the 5' part of c-*rel* and the 3' part of a novel gene named *nrg* (non-*rel* gene). Analysis of cDNA clones predicted the presence of a hybrid protein in which the C-terminal part of REL was replaced with NRG sequences. NRG has no homology to any known protein sequences and it is not known wether NRG contributes to the presumed tumorigenic effect of the REL-NRG protein. Possibly truncation of c-REL, comparable to v-REL, is the critical step. It has been shown that v-REL is able to inhibit NFκB function (10, 97).

Eight patients with non-Hodgkin's lymphomas and aberrations in chromosome region 2p11.2-14 were analyzed for alterations in c-*rel* (135). Two cases showed amplification of c-*rel*, 10- and 28-fold, respectively. Only one additional case showed a breakpoint in c-*rel* intron 8. In none of the cases rearrangement of *nrg* could be detected. These data suggest that the c-*rel* gene is nonrandomly involved in some non-Hodgkin's lymphomas but the activation mechanism of c-*rel* seems to vary. It is possible that minor mutations exist, that truncate the c-REL product, but which are not detected by chromosome analysis or Southern and Northern blot hybridizations. Another plausible explanation is that overexpression of c-*rel* competes for IκB and thereby interferes with the function of the NFκB complex of which the REL proteins may be part.

lyt-10.

Analysis of t(10;14)(q24;q32) in low grade B-cell non-Hodgkin lymphoma revealed that the translocation juxtaposes a novel gene named *lyt-10* to the *IgH* locus. (159). Characterization of *lyt-10* showed that its putative protein product has extensive homology to the C-terminal part of the NFκB p50 precursor, including the ankyrin repeats, the dimerization

and DNA binding domains. The translocation results in a truncation of *lyt-10* resulting in loss of the ankyrin repeats that inhibit its function. Whether LYT-10 is associated with the NFκB complex or functions independently awaits to be solved. Considering the multiple interactions of NFκB with other cellular factors, interference of *lyt-10* activation with NFκB function can be expected.

1.3.5 Apoptosis

Although tumor growth is mainly seen as enhanced cell proliferation, evidence increases that reduced apoptosis plays a role as well. Apoptosis or programmed cell death is an active cellular process that requires *de novo* protein synthesis (154). It is characterized by plasma membrane blebbing, loss of volume through export of water and salts and the induction of a specific endonuclease that generates DNA fragments of oligonucleosomal length. Neighbouring cells are induced to phagocytosis and cells are removed without induction of an inflammatory reaction. Apoptosis is responsible for elimination of T-cells expressing self antigens or an inappropriate set of surface markers (187, 189), it plays a role in embryo morphogenesis (172) and is induced in cells deprived of essential factors (207). Induction of proliferation and induction of the apoptosis program may be equally important to control cell numbers in normal peripheral blood. BCL-2, implicated in follicular lymphoma, appeared to be an important regulator of apoptosis in B-cells.

bcl-2

The t(14;18)(q32;q21) is specifically and exclusively associated with follicular B-cell lymphoma (214). This rather benign disease is characterized by large numbers of otherwise normal, polyclonal, small resting B-cells. The translocation juxtaposes the *bcl-2* gene to the IgH gene such that the BCL-2 protein is not altered but highly overexpressed (9, 44, 196). Transgenic mice were generated carrying a *bcl-2-Ig* minigene. High level expression in mouse B-cells appeared to be directly responsible for survival of these cells and enlargement of the B-cell compartment, but did not interfere with proliferation or differentiation of the cells (143). Overexpression of *bcl-2* extends the survival of certain hematopoietic cell lines after factor deprivation (162, 202). The effect of BCL-2 is intrinsic to the cells as transgenic B-cells also show long term survival once transplanted in normal animals (161). Moreover, the length of the immune response is seven fold increased.

Clearly, BCL-2 prevents cells from entering the process of apoptosis. How BCL-2 exerts this function is not known. The 25 kD protein has no homology to any other protein sequence, but it contains a 19 amino-acid hyrophobic stretch at its C-terminus which causes the protein to be

anchored to the inner mitochondrial membrane (90). Possibly, the protein interferes with the cell's energy status.

1.4 Leukemogenesis is a multistep process

In general, the generation of tumors is seen as a multistep process, which holds true for leukemia as well. In some cases it is evident. CML can be treated successfully until blast crises occurs (39). Follicular lymphoma associated with t(14;18) is a relatively benign disease but about 20% of the cases will develop to a malignant form (94). Patients with BL are often infected with EBV (139) and the occurrence of HTLV-I/HTLV-II is correlated with T-ALL (180).

Clear evidence for genes cooperating in leukemogenesis is derived from animal models.

The murine erythroleukemia Friend virus causes leukemia via a number of well defined steps (reviewed by (12)). Unlike other acutely transforming retroviruses, Friend virus (SFFV-P and SFFV-A) does not contain activated oncogenes. Instead, the *env* gene product can interact with the EPO-receptor (133), which leads to factor independent polyclonal proliferation of nonleukemic erythroid progenitor cells. Mice with mutations in the *scf/steel* gene or in the c-*kit* gene (17) are resistant to Friend virus which indicates that viral products may also interact with this growth factor system or that EPO has to synergize with SCF to generate erythroid proliferation. In almost all subsequent clonal expansions of malignant cells, the virus has activated *ets* related transcription factors (80) by viral integration (*spi*-1 by SFFV-P and SFFV-A; *fli*-1 by F-MuLV) . As v-*ets* is essential for erythroid transformation by the avian leukemia virus E26 (163), these data strongly suggest a role for the *ets*-gene family in erythroid development. Another event specific for late stages of Friend virus induced leukemia is inactivation of the recessive oncogene p53 (p53 reviewed by (130)). 30% of the clonal expansions have deletions or virus integrations in p53, often accompanied by loss of the normal allele.

Retroviral tagging in mice has identified a number of genes involved in tumorigenesis. It led to the characterization of *pim*-1 a serine/threonine kinase involved in T cell lymphomas. However, transgenic mice harbouring an immunoglobulin enhancer (Eμ)-linked *pim*-1 gene only developed tumors after a long latency period of more than a year (200). Retroviral infections with Molony murine leukemia virus (MoMLV) of newborn transgenic mice gave rise to tumors with very short latency periods and surprisingly all retroviral insertions were found to activate c-*myc* or N-*myc* genes. In the reverse experiment, Eμ c-*myc* transgenic mice develop lymphomas after a variable latency period (125). A preleukemic state can be discerned and tumors are clonal, indicating secondary events. Infection of Eμ-*myc* transgenic mice with Mo-MLV led to a rapid development of

lymphomas (84, 201). Characterization of the retroviral integration sites revealed a limited set of activated genes: *pim*-1 (35%), *bmi*-1 (35%) a putative transcription factor, and the as yet uncharacterized genes *pal*-1 and *bla*-1. These results point out that at least two independent genes need to be deregulated to generate leukemia. Moreover, these experiments reveal that there are specific sets of cooperating oncogenes in this model system. This may indicate that separate signal transduction pathways need to be deregulated, or deregulation of a certain pathway needs to be enhanced to overcome interactive regulation by other control mechanisms.

Comparable results were obtained with *bcl-2* transgenic mice. A number of high-grade lymphomas developing from indolent follicular hyperplasia had acquired an activated c-*myc* gene (144). In addition, mating transgenic Eμ-*myc* and Eμ-*bcl-2* mice resulted in offspring that developed early lymphoid tumors at high frequency (192). This indicates that *myc* and *bcl-2* are efficiently cooperating oncogenes.

The avian erythroblastosis virus (AEV) contains two oncogenes, v-*erbB* (mutated epidermal growth factor receptor) and v-*erbA* (mutated thyroid hormone receptor) (74). Analysis of the single oncogenes demonstrated that v-*erbB* is responsible for uncontrolled growth, while v-*erbA* inhibits differentiation. Both are needed for effective induction of erythroleukemia ((215) and refs therein).

Similarly, the murine myelomonocytic cell line WEHI-3B contains both activated *Il-3* and *Hox-2.4* genes. IL-3 leads to enhanced proliferation, Hox-2.4 inhibits differentiation. Again both are necessary to induce leukemia in a mouse model system (168).

1.4.1 Enlargement of a susceptible compartment

As is demonstrated above, a first step towards leukemogenesis can be enhanced proliferation or reduced differentiation. These steps seem to be conditional since they generate an enlarged compartment of early hematopoietic precursor cells that are susceptible to subsequent mutations, which increases the chance to generate a leukemic subclone. Both preleukemic conditions are actually encountered in human hematological disorders.

Myeloproliferative syndrome (MPS) is characterized by an increased cell mass but cells are fully capable to differentiate. Chronic leukemias, like CML, might be regarded as a form of MPS. Noteworthy, BCR-ABL can abrogate factor dependence (IL-3, IL-2, IL-7) in cell lines (45, 46, 170, 212), while MPS can be elicited in mice by overexpression of growth factors like IL-3 or GM-CSF (40, 99, 210).

Myelodysplastic syndrome (MDS) is characterized by reduced numbers of mature cells in peripheral blood. Bone marrow is often hyperplastic

which may be caused by normal feedback regulation due to the low numbers of mature cells. Probably, cells die by increased apoptosis upon entering differentiation. As MDS is mostly clonal and affects cells in all lineages, a stem cell disorder is assumed (70, 98, 195). However, a defect in stromal cells could account for the effect as well. Although little is known about MDS, growth factors and their receptors could be involved (31, 176). Most growth factors are produced by stromal cells, which could explain their putative contribution. Many growth factor receptors are located on chromosome 5q a region frequently deleted in MDS (70). Abstinence from growth factors will activate the apoptotic program in differentiating cells. Eventually, approximately 20% of MDS patients will develop acute leukemia.

Protection from apoptosis was only widely recognized as a way to enlarge a compartment and enhance the chance for secondary events, upon characterization of BCL-2 in follicular lymphoma (chapter 1.3.5). It is possible that additional genes exist that could protect against apoptosis. Enhanced expression of these genes could contribute to leukemogenesis.

Infection with Epstein Barr virus (EBV) is detected in some 96% of endemic Burkitt's Lymphoma (139). Proteins encoded by the virus are able to immortalize B-cells, although the efficiency is relatively low. Recently it was shown that latent membrane protein 1 of EBV can upregulate *bcl*-2 and block apoptosis in EBV infected cells (87). Thus EBV infection leads to an enlarged B-cell compartment.

Human T-cell leukemia virus type I or II (HTLV-I/HTLV-II) can immortalize T-cells, though the mechanism is as yet unknown. The resulting enlarged T-cell compartment predisposes for adult T-cell leukemia (reviewed by (180)).

1.4.2 Mutations involved in progression

The previously discussed translocations, associated with specific subtypes of leukemia, may be involved in early leukemogenesis. Still, little is known about the nature of mutations involved in progression.

Follicular lymphoma progresses frequently into a much more aggressive form. In some patients that develop high grade B cell malignancy, an additional t(8;14)(q24;q32) activates the c-*myc* gene (53, 69). This indicates that cooperation between MYC and BCL-2 occurs in man as well as in mouse.

Most of the previously described rearrangements of the c-*rel* gene were found in malignant B-cells that carried a t(14;18) as well (135). Therefore *bcl-2* and *rel* must also be considered cooperating oncogenes.

Although the frequencies differ largely, mutations in the *ras* gene at codons 12, 13 and 61 occur in many types of tumors. In AML and MDS the incidence of *ras* mutations is about 30% (103/412 and 38/138, respectively) (reviewed by (30)). There seems to be a preference for

monocytic or myelomonocytic diseases. A *ras* mutation was found in 3/22 blast crises of CML and only 1/41 chronic phases. This may imply a minor role for *ras* in progression of CML. In lymphoid leukemias, however, *ras* mutations are rare.

The dominant negative oncogene p53 is the most widely affected gene in human tumors (reviewed by (130)). Evaluation of p53 mutations is complicated by the fact that two ways of activation exist. (*i*) Mutations in p53 lead to a much more stable protein and thus to higher steady state levels. Mutated p53 complexes with *wt* p53, which results in a dominant effect though p53 is a recessive oncogene. Therefore, tumor samples were screened for enhanced expression. While p53 expression in normal blast cells is low or undetectable, increased levels were found in 19/34 and 8/33 cases of leukemia (117, 190). (*ii*) Once *p53*, located on chromosome 17p, was recognized as a recessive oncogene, tumor samples were screened for loss of function mutations. In the evolution of CML to blast crises 10-20 % of the patients acquires an isochromosome 17q (105), which suggests that loss of *p53* may be involved. Analysis of CML patients for *p53* rearrangements or homozygosity showed that p53 rearrangements occurred in 19/134 (14%) patients in blastcrisis or accelerated phase and in 5/128 (4%) patients in chronic phase (combined results of (3, 64, 109, 142)). Leukemic cells of these patients did not contain an i17q. This suggests that mutation of *p53* can be one of the steps that lead to progression of CML. The incidence may be over 30% or even higher as point mutations in p53 can occur at many sites in the gene, which can not be detected by Southern and Northern blotting. The association of p53 mutation and leukemia progression was also found in erythroid leukemogenesis by Friend virus in mice.

Progression of leukemia is often accompanied by the accumulation of chromosomal abnormalities, some of which are frequently recurring. In addition to the above mentioned deletion of chromosome 17p, deletion of chromosome 5q or 7q as well as trisomy of chromosome 8 is relatively common (105). Chromosome 5q harbours the genes for many growth factors and growth factor receptors (19, 206). Deletion will affect both proliferation and differentiation of the cells. The gene encoding the M-CSF (CSF-1) receptor, c-*fms*, has been studied most extensively. Mutations that upregulate the activity of the intracellular kinase domain of the receptor, have been detected in 8/48 AML patients (176). The molecular consequences of any of the other chromosomal aberrations is unknown.

Although different steps in leukemogenesis can be studied relatively easy in animal model systems, it is often difficult to categorise stages in human leukemia due to uncontrolled selection mechanisms and genetic heterogeneity. Borderlines between preleukemic stages like MDS and MPS, and overt leukemia may be unclear. Nevertheless, our understanding of leukemogenesis will increase through the study of genes aberrantly

expressed in human leukemia, the listing of cooperating oncogenes and the analysis of defined mutations in animal model systems.

1.5 References

1. Abraham, K.M., S.D. Levin, J.D. Marth, K.A. Forbush and R.M. Perlmutter. 1991. Thymic tumorigenesis induced by overexpression of p56lck. Proc Natl Acad Sci USA 88:3977-81.
2. Abraham, N., M.C. Miceli, J.R. Parnes and A. Veillette. 1991. Enhancement of T-cell responsiveness by the lymphocyte-specific tyrosine protein kinase p56lck. Nature 350:62-6.
3. Ahuja, H., M. Bar-Eli, S.H. Advani, S. Benchimol and M.J. Cline. 1989. Alterations in the p53 gene and the clonal evolution of the blast crisis of chronic myelocytic leukemia. Proc Natl Acad Sci USA 86:6783-7.
4. Allen, J.D., T. Lints, N.A. Jenkins, N.G. Copeland, A. Strasser, R.P. Harvey and J.M. Adams. 1991. Novel murine homeo box gene on chromosome 1 expressed in specific hematopoietic lineages and during embryogenesis. Genes Dev 5:509-20.
5. Aplan, P.D., C.G. Begley, V. Bertness, M. Nussmeier, A. Ezquerra, J. Coligan and I.R. Kirsch. 1990. The SCL gene is formed from a transcriptionally complex locus. Mol Cell Biol 10:6426-35.
6. Aplan, P.D., D.P. Lombardi, A.M. Ginsberg, J. Cossman, V.L. Bertness and I.R. Kirsch. 1990. Disruption of the human SCL locus by "illegitimate" V-(D)-J recombinase activity. Science 250:1426-9.
7. Aplan, P.D., D.P. Lombardi and I.R. Kirsch. 1991. Structural characterization if SIL, a gene frequently disrupted in T-cell acute lymphoblastic leukemia. Mol Cell Biol 11:5462-5469.
8. Baeuerle, P.A. 1991. The inducible transcription activator NF-kappa B: regulation by distinct protein subunits. Biochim Biophys Acta 1072:63-80.
9. Bakhshi, A., J.J. Wright, W. Graninger, M. Seto, J. Owens, J. Cossman, J.P. Jensen, P. Goldman and S.J. Korsmeyer. 1987. Mechanism of the t(14;18) chromosomal translocation: structural analysis of both derivative 14 and 18 reciprocal partners. Proc Natl Acad Sci USA 84:2396-400.
10. Ballard, D.W., W.H. Walker, S. Doerre, P. Sista, J.A. Molitor, E.P. Dixon, N.J. Peffer, M. Hannink and W.C. Greene. 1990. The v-rel oncogene encodes a kappa B enhancer binding protein that inhibits NF-kappa B function. Cell 63:803-14.
11. Begley, C.G., P.D. Aplan, S.M. Denning, B.F. Haynes, T.A. Waldmann and I.R. Kirsch. 1989. The gene SCL is expressed during early hematopoiesis and encodes a differentiation-related DNA-binding motif. Proc Natl Acad Sci USA 86:10128-32.
12. Ben-David, Y. and A. Bernstein. 1991. Friend virus-induced erythroleukemia and the multistage nature of cancer. Cell 56:831-834.
13. Benezra, R., R.L. Davis, D. Lockshon, D.L. Turner and H. Weintraub. 1990. The protein Id: a negative regulator of helix-loop-helix DNA binding proteins.37. Cell 61:49-59.
14. Bennett, J.M., D. Catovsky, M.T. Daniel, G. Flandrin, D.A. Galton, H.R. Gralnick and C. Sultan. 1985. Proposed revised criteria for the classification of acute myeloid leukemia. A report of the French-American-British Cooperative Group. Ann Intern Med 103:620-5.
15. Bernard, O., P. Guglielmi, D. Cherif, S. Gisselbrecht, M. Mauchauffe, R. Berger, C.-J. Larsen and D. Mathieu-Mahul. 1990. Two distinct mechanisms for the SCL gene activation in the t(1;14) translocation of T-cell leukemias. Genes Chrom Cancer 1:94-208.
16. Bernards, A., C.M. Rubin, C.A. Westbrook, M. Paskind and D. Baltimore. 1987. The first intron in the human c-abl gene is at least 200 kilobases long and is a target for translocations in chronic myelogenous leukemia. Mol Cell Biol 7:3231-6.

17. Bernstein, A., L. Forrester, A.D. Reith, P. Dubreuil and R. Rottapel. 1991. The murine W/c-kit and Steel loci and the control of hematopoiesis. Semin Hematol 28:138-42.
18. Bhatia, K., K. Huppi, T. McKeithan, D. Siwarski, J.F. Mushinski and I. Magrath. 1991. Mouse bcl-3: cDNA structure, mapping and stage-dependent expression in B lymphocytes. Oncogene 6:1569-1573.
19. Bishop, D.T. and C. Westbrook. 1990. Report of the committee on the genetic constitution of chromosome 5. Cytogenet cell genet 55:111-117.
20. Bishop, J.M. 1991. Molecular themes in oncogenesis. Cell 64:235-48.
21. Blackwell, T.K. and H. Weintraub. 1990. Differences and similarities in DNA-binding preferences of MyoD and E2A protein complexes revealed by binding site selection. Science 250:1104-10.
22. Blackwood, E.M. and R.N. Eisenman. 1991. Max: a helix-loop-helix zipper protein that forms a sequence-specific DNA-binding complex with Myc. Science 251:1211-7.
23. Boehm, T., R. Baer, I. Lavenir, A. Forster, J.J. Waters, E. Nacheva and T.H. Rabbitts. 1988. The mechanism of chromosomal translocation t(11;14) involving the T-cell receptor C delta locus on human chromosome 14q11 and a transcribed region of chromosome 11p15. Embo J 7:385-94.
24. Boehm, T., L. Foroni, Y. Kaneko, M.F. Perutz and T.H. Rabbitts. 1991. The rhombotin family of cysteine-rich LIM-domain oncogenes: distinct members are involved in T-cell translocations to human chromosomes 11p15 and 11p13. Proc Natl Acad Sci USA 88:4367-71.
25. Boehm, T., J.M. Greenberg, L. Buluwela, I. Lavenir, A. Forster and T.H. Rabbitts. 1990. An unusual structure of a putative T cell oncogene which allows production of similar proteins from distinct mRNAs. Embo J 9:857-68.
26. Boehm, T. and T.H. Rabbitts. 1989. The human T cell receptor genes are targets for chromosomal abnormalities in T cell tumors. FASEB J. 3:2344-2359.
27. Boehm, T., M.G. Spillantini, M.V. Sofroniew, M.A. Surani and T.H. Rabbitts. 1991. Developmentally regulated and tissue specific expression of mRNAs encoding the two alternative forms of the LIM domain oncogene rhombotin: evidence for thymus expression. Oncogene 6:695-703.
28. Boncinelli, E., R. Somma, D. Acampora, M. Pannese, M. D'Esposito, A. Faiella and A. Simeone. 1988. Organization of human homeobox genes. Hum Reprod 3:880-6.
29. Borrow, J., A.D. Goddard, D. Sheer and E. Solomon. 1990. Molecular analysis of acute promyelocytic leukemia breakpoint cluster region on chromosome 17. Science 249:1577-80.
30. Bos, J.L. 1989. ras oncogenes in human cancer: a review . Cancer Res 49:4682-9.
31. Boultwood, J., K. Rack, S. Kelly, J. Madden, A.Y. Sakaguchi, L.M. Wang, D.G. Oscier, V.J. Buckle and J.S. Wainscoat. 1991. Loss of both CSF1R (FMS) alleles in patients with myelodysplasia and a chromosome 5 deletion. Proc Natl Acad Sci USA 88:6176-80.
32. Brent, R. and M. Ptashne. 1985. A eukaryotic transcriptional activator bearing the DNA specificity of a prokaryotic repressor. Cell 43:729-736.
33. Brown, L., J.T. Cheng, Q. Chen, M.J. Siciliano, W. Crist, G. Buchanan and R. Baer. 1990. Site-specific recombination of the tal-1 gene is a common occurrence in human T cell leukemia. Embo J 9:3343-51.
34. Brownell, E., H.P. Fell, P.W. Tucker, A.H. Geurts van Kessel, A. Hagemeijer and N.R. Rice. 1988. Regional localization of the human c-rel locus using translocation chromosome analysis. Oncogene 2:527-9.
35. Burk, O. and K.-H. Klempnauer. 1991. Estrogen-dependent alterations in differentiation state of myeloid cells caused by a v-myb/estrogen receptor fusion protein. EMBO J 10:3713-3719.
36. Burnett, R.C., J.C. David, A.M. Harden, M.M. Le Beau, J.D. Rowley and M.O. Diaz. The LCK gene is involved in the t(1;7)(p34;q34) in the T-cell acute lymphoblastic leukemia-derived cell line HSB-2. Genes Chromosomes and Cancer, in the press:
37. Cantley, L.C., K.R. Auger, C. Carpenter, B. Duckworth, A. Graziani, R. Kapeller and S. Soltoff. 1991. Oncogenes and signal transduction. Cell 64:281-302.

38. Capel, B., R. Hawley, L. Covarrubias, T. Hawley and B. Mintz. 1989. Clonal contributions of small numbers of retrovirally marked hematopoietic stem cells engrafted in unirradiated neonatal W/Wv mice. Proc Natl Acad Sci USA 86:4564-8.
39. Champlin, R.E. and D.W. Golde. 1985. Chronic myelogenous leukemia: recent advances. Blood 65:1039-47.
40. Chang, J.M., D. Metcalf, R.A. Lang, T.J. Gonda and G.R. Johnson. 1989. Nonneoplastic hematopoietic myeloproliferative syndrome induced by dysregulated multi-CSF (IL-3) expression. Blood 73:1487-97.
41. Chen, Q., J.T. Cheng, L.H. Tasi, N. Schneider, G. Buchanan, A. Carroll, W. Crist, B. Ozanne, M.J. Siciliano and R. Baer. 1990. The tal gene undergoes chromosome translocation in T cell leukemia and potentially encodes a helix-loop-helix protein. Embo J 9:415-24.
42. Chomienne, C., P. Ballerini, N. Balitrand, M.E. Huang, I. Krawice, S. Castaigne, P. Fenaux, P. Tiollais, A. Dejean, L. Degos and H. De Thé. 1990. The retinoic acid receptor alpha gene is rearranged in retinoic acid-sensitive promyelocytic leukemias. Leukemia 4:802-7.
43. Clark, S.C. and R. Kamen. 1987. The human hematopoietic colony-stimulating factors. Science 236:1229-37.
44. Cleary, M.L., S.D. Smith and J. Sklar. 1986. Cloning and structural analysis of cDNAs for bcl-2 and a hybrid bcl-2/immunoglobulin transcript resulting from the t(14;18) translocation. Cell 47:19-28.
45. Cook, W.D., B.f. de St Groth, J.F. Miller, H.R. MacDonald and R. Gabathuler. 1987. Abelson virus transformation of an interleukin 2-dependent antigen-specific T-cell line. Mol Cell Biol 7:2631-5.
46. Cook, W.D., D. Metcalf, N.A. Nicola, A.W. Burgess and F. Walker. 1985. Malignant transformation of a growth factor-dependent myeloid cell line by Abelson virus without evidence of an autocrine mechanism. Cell 41:677-83.
47. Cory, S. 1986. Activation of cellular oncogenes in hemopoietic cells by chromosome translocation. Adv Cancer Res 47:189-234.
48. Croce, C.M. 1987. Role of chromosome translocations in human neoplasia. Cell 49:155-6.
49. Cross, M. and T.M. Dexter. 1991. Growth factors in development, transformation, and tumorigenesis. Cell 64:271-280.
50. Daley, G.Q., R. Van Etten and D. Baltimore. 1990. Induction of chronic myelogenous leukemia in mice by the P210bcr/abl gene of the Philadelphia chromosome. Science 247:824-30.
51. Davis, N., W. Bargmann, M.Y. Lim and H.J. Bose. 1990. Avian reticuloendotheliosis virus-transformed lymphoid cells contain multiple pp59v-rel complexes. J Virol 64:584-91.
52. de Klein, A., A. Geurts van Kessel, G. Grosveld, C.R. Bartram, A. Hagemeijer, D. Bootsma, N.K. Spurr, N. Heisterkamp, J. Groffen and J.R. Stephenson. 1982. A cellular oncogene is translocated to the Philadelphia chromosome in chronic myelocytic leukemia. Nature 300:765-767.
53. de Jong, D., B.M. Voetdijk, G.C. Beverstock, G.J. van Ommen, R. Willemze and P.M. Kluin. 1988. Activation of the c-myc oncogene in a precursor-B-cell blast crisis of follicular lymphoma, presenting as composite lymphoma. N Engl J Med 318:1373-8.
54. de Thé, H., C. Chomienne, M. Lanotte, L. Degos and A. Dejean. 1990. The t(15;17) translocation of acute promyelocytic leukaemia fuses the retinoic acid receptor alpha gene to a novel transcribed locus. Nature 347:558-61.
55. de Thé, H., C. Lavau, A. Marchio, C. Chomienne, L. Degos and A. Dejean. 1991. The PML-RARA Fusion mRNA Generated by the t(15;17) Translocation in Acute Promyelocytic Leukemia Encodes a Functionally Altered RAR. Cell 66:675 - 684.
56. Dick, J.E., M.C. Magli, D. Huszar, R.A. Phillips and A. Bernstein. 1985. Introduction of a selectable gene into primitive stem cells capable of long-term reconstitution of the hemopoietic system of W/Wv mice. Cell 42:71-9.

57. Diekmann, D., S. Brill, M.D. Garrett, N. Totty, J. Hsuan, C. Monfries, C. Hall, L. Lim and A. Hall. 1991. Bcr encodes a GTPase-activating protein for p21rac. Nature 351:400-2.
58. Dube, I.A., S. Kamel-Reid, C.C. Yuan, C. Lu, X. Wu, G. Corpus, S.C. Raimondi, W.M. Crist, A. Carroll and J.B. Baker. 1991. A novel homeobox gene is deregulated in lymphoid neoplasias with chromosomal translocation t(10;14). blood in press.
59. Eilers, M., D. Picard, K.R. Yamamoto and J.M. Bishop. 1989. Chimaeras of myc oncoprotein and steroid receptors cause hormone-dependent transformation of cells. Nature 340:66-8.
60. Elefanty, A.G., I.K. Hariharan and S. Cory. 1990. bcr-abl, the hallmark of chronic myeloid leukaemia in man, induces multiple haemopoietic neoplasms in mice. Embo J 9:1069-78.
61. Ellisen, L.W., J. Bird, D.C. West, A.L. Soreng, T.C. Reynolds, S.D. Smith and J. Sklar. 1991. TAN-1, the human homolog of the Drosophila Notch gene, is broken by chromosomal translocations in T lymphoblastic neoplasms. Cell 66:649-61.
62. Evans, R.M. 1988. The steroid and thyroid hormone receptor superfamily. Science 240:889-95.
63. Fehon, R.G., P.J. Kooh, I. Rebay, C.L. Regan, T. Xu, M.A. Muskavitch and S. Artavanis-Tsakonas. 1990. Molecular interactions between the protein products of the neurogenic loci Notch and Delta, two EGF-homologous genes in Drosophila. Cell 61:523-34.
64. Feinstein, E., G. Cimino, R.P. Gale, G. Alimena, R. Berthier, K. Kishi, J. Goldman, A. Zaccaria, A. Berrebi and E. Canaani. 1991. p53 in chronic myelogenous leukemia in acute phase. Proc Natl Acad Sci USA 88:6293-7.
65. Finger, L.R., R.C. Harvey, R.C. Moore, L.C. Showe and C.M. Croce. 1986. A common mechanism of chromosomal translocation in T- and B-cell neoplasia. Science 234:982-5.
66. Finger, L.R., J. Kagan, G. Christopher, J. Kurtzberg, M.S. Hershfield, P.C. Nowell and C.M. Croce. 1989. Involvement of the TCL5 gene on human chromosome 1 in T-cell leukemia and melanoma. Proc Natl Acad Sci USA 86:5039-43.
67. Fitzgerald, T.J., G.A.M. Neale, S.C. Raimondi and R.M. Goorha. 1991. c-tal, a helix-loop-helix protein, is juxtaposed to the T-cell receptor B chain gene by a reciprocal chromosomal translocation: t(1;7)(p32;q35). Blood 78:2686-2695.
68. Freyd, G., S.K. Kim and H.R. Horvitz. 1990. Novel cysteine-rich motif and homeodomain in the product of the Caenorhabditis elegans cell lineage gene lin-11. Nature 344:876-9.
69. Gauwerky, C.E., F.G. Haluska, Y. Tsujimoto, P.C. Nowell and C.M. Croce. 1988. Evolution of B-cell malignancy: pre-B-cell leukemia resulting from MYC activation in a B-cell neoplasm with a rearranged BCL2 gene. Proc Natl Acad Sci USA 85:8548-52.
70. Geddes, A.D., D.T. Bowen and A. Jacobs. 1990. Clonal karyotype abnormalities and clinical progress in the myelodysplastic syndrome. British journal of haematology 76:194-202.
71. Gilmore, T.D. 1990. NF-kappa B, KBF1, dorsal, and related matters. Cell 62:841-3.
72. Glass, C.K., O.V. Devary and M.G. Rosenfeld. 1990. Multiple cell type-specific proteins differentially regulate target sequence recognition by the alpha retinoic acid receptor. Cell 63:729-38.
73. Goddard, A.D., J. Borrow, P.S. Freemont and E. Solomon. 1991. Characterization of a zinc finger gene disrupted by the t(15;17) in acute promyelocytic leukemia. Science 254:1371-1374.
74. Graf, T. and H. Beug. 1983. Role of the v-erbA and v-erbB oncogenes of avian erythroblastosis virus in erythroid transformation. Cell 34:7-9.
75. Green, A.R., E. Salvaris and C.G. Begley. 1991. Erythroid expression of the 'helix-loop-helix' gene, SCL. Oncogene 6:475-9.
76. Greenberg, J.M., T. Boehm, M.V. Sofroniew, R.J. Keynes, S.C. Barton, M.L. Norris, M.A. Surani, M.G. Spillantini and T.H. Rabbits. 1990. Segmental and developmental regulation of a presumptive T-cell oncogene in the central nervous system. Nature 344:158-60.
77. Grimaldi, J.C. and T.C. Meeker. 1989. The t(5;14) chromosomal translocation in a case of acute lymphocytic leukemia joins the interleukin-3 gene to the immunoglobulin heavy chain gene. Blood 73:2081-5.

78. Groffen, J., J.R. Stephenson, N. Heisterkamp, A. de Klein, C.R. Bartram and G. Grosveld. 1984. Philadelphia chromosomal breakpoints are clustered within a limited region, bcr, on chromosome 22. Cell 36:93-9.
79. Groopman, J.E., J.M. Molina and D.T. Scadden. 1989. Hematopoietic growth factors. Biology and clinical applications. N Engl J Med 321:1449-59.
80. Gutman, A. and B. Wasylyk. 1991. Nuclear targets for transcription regulation by oncogenes. Trends Genet 7:49-54.
81. Hariharan, I.K. and J.M. Adams. 1987. cDNA sequence for human bcr, the gene that translocates to the abl oncogene in chronic myeloid leukaemia. Embo J 6:115-9.
82. Harrison, S.C. 1991. A structural taxonomy of DNA-binding domains. Nature 353:715-719.
83. Hatano, M., C.W. Roberts, M. Minden, W.M. Crist and S.J. Korsmeyer. 1991. Deregulation of a homeobox gene, HOX11, by the t(10;14) in T cell leukemia. Science 253:79-82.
84. Haupt, Y., W.S. Alexander, G. Barri, S.P. Klinken and J.M. Adams. 1991. Novel zinc finger gene implicated as myc collaborator by retrovirally accelerated lymphomagenesis in E mu-myc transgenic mice. Cell 65:753-63.
85. Heisterkamp, N., G. Jenster, H.J. ten, D. Zovich, P.K. Pattengale and J. Groffen. 1990. Acute leukaemia in bcr/abl transgenic mice. Nature 344:251-3.
86. Heisterkamp, N., K. Stam, J. Groffen, A. de Klein and G. Grosveld. 1985. Structural organization of the bcr gene and its role in the Ph' translocation. Nature 315:758-61.
87. Henderson, S., M. Rowe, C. Gregory, D. Croom-Carter, F. Wang, R. Longnecker, E. Kieff and A. Rickinson. 1991. Induction of bcl-2 expression by Epstein-Barr virus latent membrane protein 1 protects infected B cells from programmed cell death. Cell 65:1107-15.
88. Hermans, H., N. Heisterkamp, M. von Lindern, S. van Baal, D. Meijer, D. van der Plas, L.M. Wiedeman, J. Groffen, D. Bootsma and G. Grosveld. 1987. Unique fusion of BCR and c-ABL genes in Philadelphia chromosome positive acute lymphoblastic leukemia. Cell 51:33-40.
89. Hershfield, M.S., J. Kurtzberg, E. Harden, J.O. Moore, J. Whang-Peng and B.F. Haynes. 1984. Conversion of a stem cell leukemia from a T-lymphoid to a myeloid phenotype induced by the adenosine deaminase inhibitor 2'-deoxycoformycin. Proc Natl Acad Sci USA 81:253-7.
90. Hockenbery, D., G. Nunez, C. Milliman, R.D. Schreiber and S.J. Korsmeyer. 1990. Bcl-2 is an inner mitochondrial membrane protein that blocks programmed cell death. Nature 348:334-6.
91. Hogan, T.F., W. Koss, A.J. Murgo, R.S. Amato, J.A. Fontana and F.L. VanScoy. 1987. Acute lymphoblastic leukemia with chromosomal 5;14 translocation and hypereosinophilia: case report and literature review. J Clin Oncol 5:382-90.
92. Hope, I.A. and K. Struhl. 1986. Functional dissection of a eukaryotic transcriptional activator protein, GCN4 of yeast. Cell 46:885-94.
93. Horak, I.D., R.E. Gress, P.J. Lucas, E.M. Horak, T.A. Waldmann and J.B. Bolen. 1991. T-lymphocyte interleukin 2-dependent tyrosine protein kinase signal transduction involves the activation of p56lck. Proc Natl Acad Sci USA 88:1996-2000.
94. Horning, S.J. and S.A. Rosenberg. 1984. The natural history of initially untreated low-grade non-Hodgkin's lymphomas. N Engl J Med 311:1471-5.
95. Hsu, H.L., J.T. Cheng, Q. Chen and R. Baer. 1991. Enhancer-binding activity of the tal-1 oncoprotein in association with the E47/E12 helix-loop-helix proteins. Mol Cell Biol 11:3037-42.
96. Hunger, S.P., N. Galili, A.J. Carroll, W.M. Crist, M.P. Link and M.L. Cleary. 1991. The t(1;19)(q23;p13) results in consistent fusion of E2A and PBX1 coding sequences in acute lymphoblastic leukemias. Blood 77:687-93.
97. Inoue, J., L.D. Kerr, L.J. Ransone, E. Bengal, T. Hunter and I.M. Verma. 1991. c-rel activates but v-rel suppresses transcription from kappa B sites. Proc Natl Acad Sci USA 88:3715-9.

98. Janssen, J.W., M. Buschle, M. Layton, H.G. Drexler, J. Lyons, H. van den Berghe, H. Heimpel, B. Kubanek, E. Kleihauer, G.J. Mufti and C.R. Bartram. 1989. Clonal analysis of myelodysplastic syndromes: evidence of multipotent stem cell origin. Blood 73:248-54.
99. Johnson, G.R., T.J. Gonda, D. Metcalf, I.K. Hariharan and S. Cory. 1989. A lethal myeloproliferative syndrome in mice transplanted with bone marrow cells infected with a retrovirus expressing granulocyte-macrophage colony stimulating factor. Embo J 8:441-8.
100. Jones, N. 1990. Transcriptional regulation by dimerization: two sides to an incestuous relationship. Cell 61:9-11.
101. Jordan, C.T. and I.R. Lemischka. 1990. Clonal and systemic analysis of long-term hematopoiesis in the mouse. Genes Dev 4:220-32.
102. Jucker, M., H. Abts, D. Eick, G.M. Lenoir and H. Tesch. 1991. Over-expression of lck in Burkitt's lymphoma cell lines. Leukemia 5:528-30.
103. Kakizuka, A., W.H. Miller, K. Umesono, R.P. Warrell, S.R. Frankel, V.V.V.S. Murty, E. Dmitrovsky and R.M. Evans. 1991. Chromosomal translocation t(15;17) in Human Acute Promyelocytic Leukemia Fuses RARA with a Novel Putative Transcription Factor, PML. Cell 66:663 - 674.
104. Kamps, M.P., C. Murre, X.H. Sun and D. Baltimore. 1990. A new homeobox gene contributes the DNA binding domain of the t(1;19) translocation protein in pre-B ALL. Cell 60:547-55.
105. Kaneko, Y., J.D. Rowley, H.S. Maurer, D. Variakojis and J.W. Moohr. 1982. Chromosome pattern in childhood acute nonlymphocytic leukemia (ANLL). Blood 60:389-99.
106. Karlsson, O., S. Thor, T. Norberg, H. Ohlsson and T. Edlund. 1990. Insulin gene enhancer binding protein Isl-1 is a member of a novel class of proteins containing both a homeo- and a Cys-His domain. Nature 344:879-82.
107. Keller, G. and R. Snodgrass. 1990. Life span of multipotential hematopoietic stem cells in vivo. J Exp Med 171:1407-18.
108. Kelliher, M.A., J. McLaughlin, O.N. Witte and N. Rosenberg. 1990. Induction of a chronic myelogenous leukemia-like syndrome in mice with v-abl and BCR/ABL. Proc Natl Acad Sci USA 87:6649-53.
109. Kelman, Z., M. Prokocimer, S. Peller, Y. Kahn, G. Rechavi, Y. Manor, A. Cohen and V. Rotter. 1989. Rearrangements in the p53 gene in Philadelphia chromosome positive chronic myelogenous leukemia. Blood 74:2318-24.
110. Kennedy, M.A., R. Gonzales-Sarmiento, U.R. Kees, F. Lampert, N. Dear, T. Boehm and T. Rabbitts. 1991. HOX11, a homeobox-containing T-cell oncogene on human chromosome 10q24. Proc Natl Acad Sci USA 88:8900-8904.
111. Kerkhoff, E., K. Bister and K.H. Klempnauer. 1991. Sequence-specific DNA binding by Myc proteins. Proc Natl Acad Sci USA 88:4323-7.
112. Kerr, L.D., J. Inoue, N. Davis, E. Link, P.A. Baeuerle, H.J. Bose and I.M. Verma. 1991. The rel-associated pp40 protein prevents DNA binding of Rel and NF-kappa B: relationship with I kappa B beta and regulation by phosphorylation. Genes Dev 5:1464-76.
113. Kessel, M. and P. Gruss. 1990. Murine developmental control genes. Science 249:374-9.
114. Kliewer, S.A., K. Umesono, D.J. Mangelsdorf and R.M. Evans. 1992. Retinoid X receptor interacts with nuclear receptors in retinoic acid, thyroid hormone and vitamin D3 signalling. Nature 335:446-449.
115. Klinken, S.P., W.S. Alexander and J.M. Adams. 1988. Hemopoietic lineage switch: v-raf oncogene converts Emu-myc transgenic B cells into macrophages. Cell 53:857-67.
116. Koch, C.A., D. Anderson, M.F. Moran, C. Ellis and T. Pawson. 1991. SH2 and SH3 domains: elements that control interactions of cytoplasmic signaling proteins. Science 252:668-74.

117. Koeffler, H.P., C. Miller, M.A. Nicolson, J. Ranyard and R.A. Bosselman. 1986. Increased expression of p53 protein in human leukemia cells. Proc Natl Acad Sci USA 83:4035-9.
118. Kongsuwan, K., E. Webb, P. Housiaux and J.M. Adams. 1988. Expression of multiple homeobox genes within diverse hematopoietic lineages. EMBO J. 7:2131-2138.
119. Konopka, J.B., S.M. Watanabe and O.N. Witte. 1984. An alteration of the human c-abl protein in K562 leukemia cells unmasks associated tyrosine kinase activity. Cell 37:1035-42.
120. Kuo, S.S., J.D. Mellentin, N.G. Copeland, D.J. Gilbert, N.A. Jenkins and M.L. Cleary. 1991. Structure, chromosome mapping, and expression of the mouse Lyl-1 gene. Oncogene 6:961-8.
121. Kurzrock, R., J.U. Gutterman and M. Talpaz. 1988. The molecular genetics of Philadelphia chromosome-positive leukemias. N Engl J Med 319:990-8.
122. Lammie, G.A., V. Fantl, R. Smith, E. Schuuring, S. Brookes, R. Michalides, C. Dickson, A. Arnold and G. Peters. 1991. D11S287, a putative oncogene on chromosome 11q13, is amplified and expressed in squamous cell and mammary carcinomas and linked to BCL-1. Oncogene 6:439-44.
123. Lampert, F., J. Harbott, J. Ritterbach, W.D. Ludwig, C. Fonatsch, D. Schwamborn, B. Stier, A. Gnekow, V. Gerein and B. Stollmann. 1988. T-cell acute childhood lymphoblastic leukemia with chromosome 14q11 anomaly: a morphologic, immunologic, and cytogenetic analysis of 10 patients. Blut 56:117-23.
124. Landschulz, W.H., P.F. Johnson and S.L. McKnight. 1988. The leucine zipper: a hypothetical structure common to a new class of DNA binding proteins. Science 240:1759-64.
125. Langdon, W.Y., A.W. Harris, S. Cory and J.M. Adams. 1986. The c-myc oncogene perturbs B lymphocyte development in E-mu-myc transgenic mice. Cell 47:11-8.
126. Lange, W., W. Brugger, F.M. Rosenthal, L. Kanz and A. Lindemann. 1991. The role of cytokines in oncology. Int J Cell Cloning 9:252-273.
127. Larson, R.A., K. Kondo, J.W. Vardiman, A.E. Butler, H.M. Golomb and J.D. Rowley. 1984. Evidence for a 15;17 translocation in every patient with acute promyelocytic leukemia. Am J Med 76:827-41.
128. Lassar, A.B., R.L. Davis, W.E. Wright, T. Kadesch, C. Murre, A. Voronova, D. Baltimore and H. Weintraub. 1991. Functional activity of myogenic HLH proteins requires hetero-oligomerization with E12/E47-like proteins in vivo. Cell 66:305-15.
129. Lenardo, M.J., A. Kuang, A. Gifford and D. Baltimore. 1988. NF-kappa B protein purification from bovine spleen: nucleotide stimulation and binding site specificity. Proc Natl Acad Sci USA 85:8825-9.
130. Levine, A.J., J. Momand and C.A. Finlay. 1991. The p53 tumour suppressor gene. Nature 351:453-6.
131. Lewin, B. 1990. Driving the cell cycle: M phase kinase, its partners, and substrates. Cell 61:743-52.
132. Lewin, B. 1991. Oncogenic conversion by regulatory changes in transcription factors. Cell 64:303-12.
133. Li, J.P., A.D. D'Andrea, H.F. Lodish and D. Baltimore. 1990. Activation of cell growth by binding of Friend spleen focus-forming virus gp55 glycoprotein to the erythropoietin receptor. Nature 343:762-4.
134. Longo, L., P.P. Pandolfi, A. Biondi, A. Rambaldi, A. Mencarelli, C.F. Lo, D. Diverio, L. Pegoraro, G. Avanzi, A. Tabilio, D. Zangrilli, M. Alcalay, E. Donti, F. Grignani and P. Pelicci. 1990. Rearrangements and aberrant expression of the retinoic acid receptor alpha gene in acute promyelocytic leukemias. J Exp Med 172:1571-5.
135. Lu, D., J.D. Thompson, G.K. Gorski, N.R. Rice, M.G. Mayer and J.J. Yunis. 1991. Alterations at the rel locus in human lymphoma. Oncogene 6:1235-41.
136. Lu, M., Z. Gong, W. Shen and A.D. Ho. 1991. The TCL-3 proto-oncogene altered by chromosomal translocation in T-cell leukemia codes for a homeobox protein. EMBO J. 10:2905-2910.

137. Lux, S.E., K.M. John and V. Bennett. 1990. Analysis of cDNA for human erythrocyte ankyrin indicates a repeated structure with homology to tissue-differentiation and cell-cycle control proteins. Nature 344:36-42.
138. Magli, M.C., P. Barba, A. Celetti, G. de Vita, C. Cillo and E. Boncinelli. 1991. Coordinate regulation of HOX genes in human hematopoietic cells. Proc Natl Acad Sci USA 88:6348-6352.
139. Magrath, I. 1990. The pathogenesis of Burkitt's lymphoma. Adv Cancer Res 55:133-271.
140. Marshall, C.J. 1991. Tumor suppressor genes. Cell 64:313-26.
141. Maru, Y. and O.N. Witte. 1991. The BCR gene encodes a novel serine/threonine kinase activity within a single exon. Cell 67:459-468.
142. Mashal, R., M. Shtalrid, M. Talpaz, H. Kantarjian, L. Smith, M. Beran, A. Cork, J. Trujillo, J. Gutterman and A. Deisseroth. 1990. Rearrangement and expression of p53 in the chronic phase and blast crisis of chronic myelogenous leukemia. Blood 75:180-9.
143. McDonnell, T.J., N. Deane, F.M. Platt, G. Nunez, U. Jaeger, J.P. McKearn and S.J. Korsmeyer. 1989. bcl-2-immunoglobulin transgenic mice demonstrate extended B cell survival and follicular lymphoproliferation. Cell 57:79-88.
144. McDonnell, T.J. and S.J. Korsmeyer. 1991. Progression from lymphoid hyperplasia to high-grade malignent lymphoma in mice transgenic for the t(14;18). Nature 349:254-256.
145. McGuire, E.A., A.R. Davis and S.J. Korsmeyer. 1991. T-cell translocation gene 1 (Ttg-1) encodes a nuclear protein normally expressed in neural lineage cells. Blood 77:599-606.
146. McGuire, E.A., R.D. Hockett, K.M. Pollock, M.F. Bartholdi, S.J. O'Brien and S.J. Korsmeyer. 1989. The t(11;14)(p15;q11) in a T-cell acute lymphoblastic leukemia cell line activates multiple transcripts, including Ttg-1, a gene encoding a potential zinc finger protein. Mol Cel Biol 9:2124-2132.
147. McKeithan, T.W. 1990. Molecular biology of non-Hodgkin's lymphomas. Semin Oncol 17:30-42.
148. McKeithan, T.W., J.D. Rowley, T.B. Shows and M.O. Diaz. 1987. Cloning of the chromosome translocation breakpoint junction of the t(14;19) in chronic lymphocytic leukemia. Proc Natl Acad Sci USA 84:9257-60.
149. Meeker, T.C., D. Hardy, C. Willman, T. Hogan and J. Abrams. 1990. Activation of the interleukin-3 gene by chromosome translocation in acute lymphocytic leukemia with eosinophilia. Blood 76:285-9.
150. Mellentin, J.D., C. Murre, T.A. Donlon, P.S. McCaw, S.D. Smith, A.J. Carroll, M.E. McDonald, D. Baltimore and M.L. Cleary. 1989. The gene for enhancer binding proteins E12/E47 lies at the t(1;19) breakpoint in acute leukemias. Science 246:379-82.
151. Mellentin, J.D., S.D. Smith and M.L. Cleary. 1989. LYL-1, a novel gene altered by chromosomal translocation in T-cell leukemia, codes for a protein with a helix-loop-helix DNA binding motif. Cell 58:77-83.
152. Metcalf, D. 1989. The molecular control of cell division, differentiation commitment and maturation in haemopoietic cells. Nature 339:27-30.
153. Mitchell, P.J. and R. Tjian. 1989. Transcriptional regulation in mammalian cells by sequence-specific DNA binding proteins. Science 245:371-8.
154. Morris, R.G., A.D. Hargreaves, E. Duvall and A.H. Wyllie. 1984. Hormone-induced cell death. 2. Surface changes in thymocytes undergoing apoptosis. Am J Pathol 115:426-36.
155. Motokura, T., T. Bloom, H.G. Kim, H. Juppner, J.V. Ruderman, H.M. Kronenberg and A. Arnold. 1991. A novel cyclin encoded by a bcl1-linked candidate oncogene. Nature 350:512-5.
156. Murre, C., P.S. McCaw and D. Baltimore. 1989. A new DNA binding and dimerization motif in immunoglobulin enhancer binding, daughterless, MyoD, and myc proteins. Cell 56:777-83.
157. Murre, C., P.S. McCaw, H. Vaessin, M. Caudy, L.Y. Jan, Y.N. Jan, C.V. Cabrera, J.N. Buskin, S.D. Hauschka, A.B. Lassar, H. Weintraub and D. Baltomore. 1989. Interactions

between heterologous helix-loop-helix proteins generate complexes that bind specifically to a common DNA sequence. Cell 58:537-44.
158. Nasmyth, K. and D. Shore. 1987. Transcriptional regulation in the yeast life cycle. Science 237:1162-70.
159. Neri, A., C.-C. Chang, L. Lombardi, M. Salina, P. Corradini, A.T. Maiolo, R.S.K. Chaganti and R. Dalla-Favera. 1991. B cell lymphoma-associated chromosomal translocation involves candidate oncogene lyt-10, homologous to NF-kB p50. Cell 67:1075-1088.
160. Nourse, J., J.D. Mellentin, N. Galili, J. Wilkinson, E. Stanbridge, S.D. Smith and M.L. Cleary. 1990. Chromosomal translocation t(1;19) results in synthesis of a homeobox fusion mRNA that codes for a potential chimeric transcription factor. Cell 60:535-45.
161. Nunez, G., D. Hockenbery, T.J. McDonnell, C.M. Sorensen and S.J. Korsmeyer. 1991. Bcl-2 maintains B cell memory. Nature 353:71-3.
162. Nunez, G., L. London, D. Hockenbery, M. Alexander, J.P. McKearn and S.J. Korsmeyer. 1990. Deregulated Bcl-2 gene expression selectively prolongs survival of growth factor-deprived hemopoietic cell lines. J Immunol 144:3602-10.
163. Nunn, M.F. and T. Hunter. 1989. The ets sequence is required for induction of erythroblastosis in chickens by avian retrovirus E26. J Virol 63:398-402.
164. Ohno, H., G. Takimoto and T.W. McKeithan. 1990. The candidate proto-oncogene bcl-3 is related to genes implicated in cell lineage determination and cell cycle control. Cell 60:991-7.
165. Pandolfi, P.P., F. Grignani, M. Alcalay, A. Mencarelli, A. Biondi, F. LoCoco, F. Grignani and P.G. Pelicci. 1991. Structure and origin of the acute promyelocytic leukemia myl/RAR alpha cDNA and characterization of its retinoid-binding and transactivation properties. Oncogene 6:1285-92.
166. Pendergast, A.M., A.J. Muller, M.H. Havlik, R. Clark, F. McCormick and O.N. Witte. 1991. Evidence for regulation of the human ABL turosine kinase by a cellular inhibitor. Proc Natl Acad Sci USA 88:5927-5931.
167. Pendergast, A.M., A.J. Muller, M.H. Havlik, Y. Maru and O.N. Witte. 1991. BCR sequences essential for transformation by the BCR-ABL oncogene bind to the ABL SH2 regulatory domain in a non-phosphotyrosine-dependent manner. Cell 66:161-71.
168. Perkins, A., K. Kongsuwan, J. Visvader, J.M. Adams and S. Cory. 1990. Homeobox gene expression plus autocrine growth factor production elicits myeloid leukemia. Proc Natl Acad Sci USA 87:8398-402.
169. Picard, D., S.J. Salser and K.R. Yamamoto. 1988. A movable and regulable inactivation function within the steroid binding domain of the glucocorticoid receptor. Cell 54:1073-80.
170. Pierce, J.H., P.P. Di Fiore, S.A. Aaronson, M. Potter, J. Pumphrey, A. Scott and J.N. Ihle. 1985. Neoplastic transformation of mast cells by Abelson-MuLV: abrogation of IL-3 dependence by a nonautocrine mechanism. Cell 41:685-93.
171. Pierce, J.H., M. Ruggiero, T.P. Fleming, P.P. Di Fiore, J.S. Greenberger, L. Varticovski, J. Schlessinger, G. Rovera and S.A. Aaronson. 1988. Signal transduction through the EGF receptor transfected in IL-3-dependent hematopoietic cells. Science 239:628-31.
172. Pratt, R.B. and R.M. Greene. 1976. Devel. Biol. 54:135-145.
173. Prendergast, G.C., D. Lawe and E.B. Ziff. 1991. Association of Myn, the murine homolog of max, with c-Myc stimulates methylation-sensitive DNA binding and ras cotransformation. Cell 65:395-407.
174. Rabbitts, T.H. and T. Boehm. 1990. LIM domains. Nature 346:418.
175. Reynolds, T.C., S.D. Smith and J. Sklar. 1987. Analysis of DNA surrounding the breakpoints of chromosomal translocations involving the beta T cell receptor gene in human lymphoblastic neoplasms. Cell 50:107-17.
176. Ridge, S.A., M. Worwood, D. Oscier, A. Jacobs and R.A. Padua. 1990. FMS mutations in myelodysplastic, leukemic, and normal subjects. Proc Natl Acad Sci USA 87:1377-80.
177. Robertson, M. 1988. Homoeo boxes, POU proteins and the limits to promiscuity. Nature 336:522-4.

178. Rosenberg, C.L., H.G. Kim, T.B. Shows, H.M. Kronenberg and A. Arnold. 1991. Rearrangement and overexpression of D11S287E, a candidate oncogene on chromosome 11q13 in benign parathyroid tumors. Oncogene 6:449-53.
179. Rosenberg, C.L., E. Wong, E.M. Petty, A.E. Bale, Y. Tsujimoto, N.L. Harris and A. Arnold. 1991. PRAD1, a candidate BCL1 oncogene: mapping and expression in centrocytic lymphoma. Proc Natl Acad Sci USA 88:9638-9642.
180. Rosenblatt, J.D., I.S. Chen and W. Wachsman. 1988. Infection with HTLV-I and HTLV-II: evolving concepts. Semin Hematol 25:230-46.
181. Royer-Pokora, B., U. Loos and W.-D. Ludwig. 1991. TTG-2, a new gene encoding a cysteine-rich protein with the LIM motif, is overexpressed in acute T-cell leukemia with the t(11;14)(p13;q11). Oncogene 6:1887-1893.
182. Sanches Garcia, I., Y. Kaneko, R. Gonzalez-Sarmiento, K. Campbell, L. White, T. Boehm and T.H. Rabbitts. 1991. A study of chromosome 11p13 translocations involving TCR beta and TCR delta in human T cell leukaemia. Oncogene 6:577-82.
183. Scott, M.P., J.W. Tamkun and G. Hartzell. 1989. The structure and function of the homeodomain. Biochim Biophys Acta 989:25-48.
184. Sefton, B.M. 1991. The lck tyrosine protein kinase. Oncogene 6:683-6.
185. Selleri, L., M. von Lindern, A. Hermans, D. Meijer, G. Torelli and G. Grosveld. 1990. Chronic myeloid leukemia may be associated with several bcr-abl transcripts including the acute lymphoid leukemia-type 7 kb transcript. Blood 75:1146-53.
186. Shen, W.F., K. Detmer, T.A. Simonitch-Eason, H.J. Lawrence and C. Largman. 1991. Alternative splicing of the HOX 2.2 homeobox gene in human hematopoietic cells and murine embryonic and adult tissues. Nucleic Acids Res 19:539-545.
187. Shi, Y.F., B.M. Sahai and D.R. Green. 1989. Cyclosporin A inhibits activation-induced cell death in T-cell hybridomas and thymocytes. Nature 339:625-6.
188. Shtivelman, E., B. Lifshitz, R.P. Gale and E. Canaani. 1985. Fused transcript of abl and bcr genes in chronic myelogenous leukaemia. Nature 315:550-4.
189. Smith, C.A., G.T. Williams, R. Kingston, E.J. Jenkinson and J.J. Owen. 1989. Apoptosis. Nature 338:10.
190. Smith, L.J., E.A. McCulloch and S. Benchimol. 1986. Expression of the p53 oncogene in acute myeloblastic leukemia. J Exp Med 164:751-61.
191. Soekarman, D., J. van Denderen, L. Hoefsloot, M. Moret, T. Meeuwsen, S. van Baal, A. Hagemeijer and G. Grosveld. 1990. A novel variant of the bcr-abl fusion product in Philadelphia chromosome-positive acute lymphoblastic leukemia. Leukemia 4:397-403.
192. Strasser, A., A.W. Harris, M.L. Bath and S. Cory. 1990. Novel primitive lymphoid tumours induced in transgenic mice by cooperation between myc and bcl-2. Nature 348:331-3.
193. Sun, X.-H., N.G. Copeland, N.A. Jenkins and D. Baltimore. 1991. Id proteins Id1 and Id2 selectively inhibit DNA binding by one class of helix-loop-helix proteins. Mol Cell Biol 11:5603-5611.
194. Superti-Furga, G., G. Bergers, D. Picard and M. Busslinger. 1991. Hormone-dependent transcriptional regulation and cellular transformation by Fos-steroid receptor fusion proteins. Proc Natl Acad Sci USA 88:5114-8.
195. Tefferi, A., S.N. Thibodeau and L.J. Solberg. 1990. Clonal studies in the myelodysplastic syndrome using X-linked restriction fragment length polymorphisms. Blood 75:1770-3.
196. Tsujimoto, Y., M.M. Bashir, I. Givol, J. Cossman, E. Jaffe and C.M. Croce. 1987. DNA rearrangements in human follicular lymphoma can involve the 5' or the 3' region of the bcl-2 gene. Proc Natl Acad Sci USA 84:1329-31.
197. Tsujimoto, Y., E. Jaffe, J. Cossman, J. Gorham, P.C. Nowell and C.M. Croce. 1985. Clustering of breakpoints on chromosome 11 in human B-cell neoplasms with the t(11;14) chromosome translocation. Nature 315:340-3.
198. Umek, R.M., A.D. Friedman and S.L. McKnight. 1991. CCAAT-enhancer binding protein: a component of a differentiation switch. Science 251:288-92.

199. van der Plas, D.C., D. Soekarman, A.M. van Gent, G. Grosveld and A. Hagemeijer. 1991. bcr-abl mRNA lacking abl exon a2 detected by polymerase chain reaction in a chronic myelogeneous leukemia patient. Leukemia 5:457-61.
200. van Lohuizen, M., S. Verbeek, P. Krimpenfort, J. Domen, C. Saris, T. Radaszkiewicz and A. Berns. 1989. Predisposition to lymphomagenesis in pim-1 transgenic mice: cooperation with c-myc and N-myc in murine leukemia virus-induced tumors. Cell 56:673-82.
201. van Lohuizen, M., S. Verbeek, B. Scheijen, E. Wientjens, H. van der Gulden and A. Berns. 1991. Identification of cooperating oncogenes in E mu-myc transgenic mice by provirus tagging. Cell 65:737-52.
202. Vaux, D.L., S. Cory and J.M. Adams. 1988. Bcl-2 gene promotes haemopoietic cell survival and cooperates with c-myc to immortalize pre-B cells. Nature 335:440-2.
203. Visvader, J., C.G. Begley and J.M. Adams. 1991. Differential expression of the LYL, SCL and E2A helix-loop-helix genes within the hemopoietic system. Oncogene 6:187-94.
204. von Ruden, T. and E.F. Wagner. 1988. Expression of functional human EGF receptor on murine bone marrow cells. Embo J 7:2749-56.
205. Walker, L.C., T.S. Ganesan, S. Dhut, B. Gibbons, T.A. Lister, J. Rothbard and B.D. Young. 1987. Novel chimaeric protein expressed in Philadelphia positive acute lymphoblastic leukaemia. Nature 329:851-3.
206. Wasmuth, J.J., C. Park and R.E. Ferrell. 1989. Report of the committee on the genetic constitution of chromosome 5. Cytogenet Cell Genet 51:137-48.
207. Williams, G.T., C.A. Smith, E. Spooncer, T.M. Dexter and D.R. Taylor. 1990. Haemopoietic colony stimulating factors promote cell survival by suppressing apoptosis. Nature 343:76-9.
208. Williams, M.E., C.D. Westermann and S.H. Swerdlow. 1990. Genotypic characterization of centrocytic lymphoma: frequent rearrangement of the chromosome 11 bcl-1 locus. Blood 76:1387-91.
209. Withers, D.A., R.C. Harvey, J.B. Faust, O. Melnyk, K. Carey and T.C. Meeker. 1991. Characterization of a candidate bcl-1 gene. Mol. Cell. Biol. 11:4846-4853.
210. Wong, P.M., S.W. Chung, C.E. Dunbar, D.M. Bodine, S. Ruscetti and A.W. Nienhuis. 1989. Retrovirus-mediated transfer and expression of the interleukin-3 gene in mouse hematopoietic cells result in a myeloproliferative disorder. Mol Cell Biol 9:798-808.
211. Xia, Y., L. Brown, C. Ying-Chuan Yang, J. Tsou Tsan, M.J. Siciliano, R. Espinosa III, M.M. LeBeau and R.J. Baer. 1991. TAL2, a helix-loop-helix gene activated by the (7;9)(q34;q32) translocation in human T-cell leukemia. Proc Natl Acad Sci USA 88:11416-11420.
212. Young, J.C., M.L. Gishizky and O.N. Witte. 1991. Hyperexpression of interleukin-7 is not necessary or sufficient for transformation of a pre-B lymphoid cell line. Mol Cell Biol 11:854-63.
213. Yunis, J.J. Manuscript in preparation.
214. Yunis, J.J., G. Frizzera, M.M. Oken, J. McKenna, A. Theologides and M. Arnesen. 1987. Multiple recurrent genomic defects in follicular lymphoma. A possible model for cancer. N Engl J Med 316:79-84.
215. Zenke, M., A. Munoz, J. Sap, B. Vennstrom and H. Beug. 1990. v-erbA oncogene activation entails the loss of hormone-dependent regulator activity of c-erbA. Cell 61:1035-49.
216. Zhang, X.-K., B. Hoffmann, P.B.-V. Tran, G. Graupner and M. Pfahl. 1992. Retinoid X receptor is an auxiliary protein for thyroid hormone and retinoic acid receptors. Nature 335:441-446.
217. Zutter, M., R.D. Hockett, C.W. Roberts, E.A. McGuire, J. Bloomstone, C.C. Morton, L.L. Deaven, W.M. Crist, A.J. Carroll and S.J. Korsmeyer. 1990. The t(10;14)(q24;q11) of T-cell acute lymphoblastic leukemia juxtaposes the delta T-cell receptor with TCL3, a conserved and activated locus at 10q24. Proc Natl Acad Sci USA 87:3161-5.

CHAPTER 2

MOLECULAR ANALYSIS OF T(6;9) IN ANLL

CHAPTER 2.1

Clinical, cytogenetic and morphological aspects of t(6;9) acute nonlymphocytic leukemia

The translocation (6;9)(p23;q34) was first described by Rowley and Potter (20) in two patients with acute non-lymphocytic leukemia (ANLL). Since then 51 patients with this translocation have been reported and t(6;9) ANLL is now considered to be a specific leukemia subtype (Table I). At the time of diagnosis t(6;9) is often the sole chromosomal aberration present in the leukemic cells, which argues for a causal role of the translocation in leukemogenesis. Clinical characteristics comprise a poor response to therapy leading to short survival, and a young age of onset of the leukemia. In contrast to a median age of onset of 50 years for ANLL as a whole, these patients have a median age of 30 (3-69) at diagnosis.

T(6;9) ANLL has been FAB classified as M4 (18/45), M2 (20/45) or M1 (7/45) and the translocation was also found in RAEB, a subtype of MDS. Pearson et al. (18) have reported association of t(6;9) ANLL with bone marrow basophilia but this feature is variable (16/40). It has been found in t(6;9) ANLL M4, M2 and M1, and is most consistent in M2 (11/17 tested). The most likely explanation for the apparent pluriform morphology of the leukemic cells is, that the event initiating the leukemia occurred in an early hematopoietic progenitor cell, which is able to generate cells of multiple myeloid lineages. The effect of the translocation may be inhibition of differentiation, but certainly not a total block. For example, the presence of cells in all differentiation stages in bone marrow of patient 6 in Table I, led to an initial diagnosis of Ph' negative CML. Patient 50 could have been diagnosed as Ph' negative CML for the same reasons.

Adriaansen et al. (1) reported coexpression of myeloid markers and terminal deoxynucleotidyl transferase (TdT) in t(6;9) leukemic cells. In the two cases of t(6;9) ANLL analyzed, coexpression occurred in 12.5 and 25.5% of the bone marrow cells, respectively. TdT was supposed to be a typical lymphoid marker, but low levels of coexpression of this enzyme and myeloid markers appeared to be a common feature of ANLL cells (2). More samples need to be examined to know whether relatively high levels of coexpression of TdT and myeloid markers is a consistent feature of t(6;9) ANLL.

In at least 13 cases, a myelodysplastic phase preceded the onset of overt AML, in two cases sarcoidosis (17) was preceding t(6;9) ANLL, once acute myelofibrosis (AMF) (6) and once Sweet's syndrome (23). Cuneo et al. (6) indicated that two of their patients presented with ANLL, but showed MDS features in the other lineages (Trilineage MDS: TMDS

TABLE I

case no.	age	sex	additional information[1]	FAB type	marrow basophilia	additional chromosomal abnormalities[2]	therapy response[3]	survival (months)	references
1	37	F	TMDS	M4	-	t(6;9), + 8, + 13	CR	8	20, 25
2	47	F	TMDS	M1	+		NR	< 1	20
3	5	F	TMDS	M2	+	t(6;9),t(4;10)(q31;q22)	CR	9.5 +	15
4	29	M		M1/M2	?		?	?	9
5	44	F	TMDS	M2	+		PR	7	10
6	14	M	TMDS	M2	+		NR	3.5	10, 19
7	?	M	MDS → ANLL	?	?	t(6;9),-Y	?	?	12
8	12	M		M2	?		CR	6 +	21
9	51	F	TMDS	M4	+		PR	5.5	22
10	27	M	TMDS	M4	+		PR	12	22
11	49	M		M2	+		PR	10	6, 26
12	50	M	TMDS	M2	-		NR	< 1	6, 26
13	15	M	RAEB → ANLL	?	-		NR	2	6, 26
14	23	F	MDS → ANLL	M2	+	t(6;6;9)(p23;q23q25;q34)	CR	8	27, 28
15	43	F		M2	+		CR	10.5	28
16	39	M		M1	-		CR	12	5
17	24	M		M4	-		CR	19	5
18	8	F		M4	-		CR	11 +	5
19	28	M	TMDS	M2	+		?	?	18
20	42	M	MDS → ANLL	M2	+		?	?	18
21	13	F	MDS → ANLL	M4	-	t(6;9), t(2;?),t(3;?),-5, + mars.	PR	< 2	13, 23
22	25	M		?	?	t(6;9),12p-,22q-, + other	?	?	16
23	69	M	MDS →	-	?		-	6 +	8
24	3	F	MDS → ANLL	M1	-	t(6;9), + 13	PR	28	11
25	26	M		M2	-		PR	6	11
26	37	M	MDS → ANLL	M4	+	t(6;9) + mar1, + mar2	PR	10	14

27	20	M	MDS → ANLL	M4	-		NR	4	14
28	54	F		M2	-	t(6;9), + 8	CR	11	7
29	42	F		M4	-		CR	2	7
30	34	M		U/M1	-	t(6;9),-Y	CR	3 +	3
31	13	F	MDS → ANLL	M4	-	t(6;9), + 4,-9q + , + 8,i(9p), + 13	CR	25	1, 23
32	17	F	MDS → ANLL	M4	-		NR	< 1	1, 23
33	37	F	sarcoidosis → ANLL	M4	?		CR	12	17
34	33	M	sarcoidosis → ANLL	M2	?		CR	6 +	17
35	38	M	AMF → ANLL	M4	-		PR	8	6, 23
36	28	M		M1	-	t(6;9),16q +	CR	7	6
37	56	F	TMDS	M2	+		NR	3	6
38	13	F		M2	-	t(6;9), + 13	CR	> 24	24
39	63	F		M4	?		PR?	?	23
40	35	M		M2	+	t(6;9), + 8,-9q +	PR	18	23
41	28	F		M4	-		CR	17	23
42	18	M	Sweet's syndrome → ANLL	M2	-	t(6;9),inv.1	CR	> 24	23
43	19	F		M2	+		CR	14	23
44	14	F		M4	?		CR	> 36	23
45	6	F		M4	-		NR	2	23
46	28	M		M1	?		?	?	23
47	24	F		M1	?	t(6;9),del7q	?	?	23
48	10	M		M2	-		PR	20	23
49	53	M		M4	+		CR	> 4	23
50	54	F	Ph' neg. CML		-		?	11	23
51	53	F	RAEBt		-		?	15	23

(4)). TMDS can also be recognized in eight patients reported by Pearson et al. (18) and it may be present in more of the cases reported previously (6, 14). On aggregate, at least 23 out of 51 patients present myelodysplastic features. There are no indications whether the MDS phase that precedes overt t(6;9) leukemia, is a first symptom caused by the translocation or whether the translocation occurs during the MDS phase, thereby initiating leukemia. The early stages of MDS preceding t(6;9) leukemia have never been analyzed cytogenetically and a (6;9) translocation has only been found in RAEB, which is considered as MDS in transformation to overt leukemia. Although little is known about the biology of MDS, most likely, large numbers of hematopoietic cells die during differentiation. The shortage of mature cells stimulates precursor cells to proliferate, which is regulated by normal cytokine interactions and leads to a hyperplastic bone marrow. Therefore, the compartment of cycling early precursor cells is enlarged. These cycling cells are susceptible to subsequent genetic changes, which may sometimes be leukemogenic. In this respect, it is interesting that Fonatsch et al. and Cuneo et al. (6, 11) suggest that in several cases, exposure to toxic agents or radiation may be related to t(6;9) ANLL. It will be interesting to screen MDS in young people in order to find out whether t(6;9) may initiate MDS or, alternatively, occurs later during MDS, possibly induced by carcinogenic agents. The young age of onset indicates that t(6;9) is a primary defect.

References

1. Adriaansen, H.J., J.J. van Dongen, H. Hooijkaas, K. Hählen, M.B. van 't Veer, B. Löwenberg and A. Hagemeijer. 1988. Translocation (6;9) may be associated with a specific TdT-positive immunological phenotype in ANLL. Leukemia 2:136-40.
2. Adriaansen, H.J., J.J. van Dongen, M.C. Kappers-Klunne, K. Hählen, M.B. van 't Veer, J.H. Wijdenes-de Bresser, A.C. Holdrinet, E.J. Harthoorn-Lasthuizen, J. Abels and H. Hooijkaas. 1990. Terminal deoxynucleotidyl transferase positive subpopulations occur in the majority of ANLL: implications for the detection of minimal disease. Leukemia 4:404-10.
3. Bernstein, R., M.J. Cain and J. Cho. 1989. Loss of the Y chromosome associated with translocation t(6;9)(p23;q34) in a patient with acute nonlymphocytic leukemia. Cancer Genet Cytogenet 39:81-7.
4. Brito-Babapulle, F., D. Catovsky and D.A.G. Galton. 1987. Clinical and laboratory features of de novo acute myeloid leukemia with trilineage myelodysplasia. Br J Haematol 66:445-450.
5. Carroll, A.J., R.P. Castleberry, J.T. Prchal and W.H. Finley. 1985. Translocation (6;9)(p23;q34) in acute nonlymphocytic leukemia: three new cases. Cancer Genet Cytogenet 18:303-6.
6. Cuneo, A., S. Kerim, E. Vandenberghe, A. Van Orshoven, J. Rodhain, A. Bosly, P. Zachee, A. Louwagie, J.L. Michaux, P. Dal Cin and H. Van den Berghe. 1989. Translocation t(6;9) occurring in acute myelofibrosis, myelodysplastic syndrome, and acute nonlymphocytic leukemia suggests multipotent stem cell involvement. Cancer Genet Cytogenet 42:209-19.

7. Fan, Y.S., S.N. Jani Sait, A. Raza, J.M. Rowe, P. D'Arrigo and A.A. Sandberg. 1988. Translocation t(6;9)(p23;q34) in acute nonlymphocytic leukemia. Two new patients without increased bone marrow basophils. Cancer Genet Cytogenet 32:153-5.
8. Fan, Y.S., A. Raza, J. Schumer, S.N. Sait, A.W. Block, M. Snyderman and A.A. Sandberg. 1987. Translocation t(6;9)(p22.3;q34) in myelodysplastic syndrome--refractory anemia with excess blasts. Cancer Genet Cytogenet 29:135-8.
9. Fitzgerald, P.H., C.M. Morris, G.J. Fraser, L.M. Giles, J.W. Hamer, D.C. Heaton and M.E. Beard. 1983. Nonrandom cytogenetic changes in New Zealand patients with acute myeloid leukemia. Cancer Genet Cytogenet 8:51-66.
10. Fleischman, E.W., E.L. Prigogina, G.W. Iljinskaja, L.N. Konstantinova, G.P. Puchkova, M.A. Volkova, M.A. Frenkel and S.A. Balakirev. 1983. Chromosomal rearrangements with a common breakpoint at 6p23 in five cases of myeloid leukemia. Hum Genet 64:254-6.
11. Fonatsch, C., B. Stollmann, J. Holldack and A. Engert. 1987. Translocation (6;9)(p23;q34) in smoldering leukemia and acute nonlymphocytic leukemia. Cancer Genet Cytogenet 26:363-8.
12. Gold, E.J., M. Conjalka, L.M. Pelus, S.C. Jhanwar, H. Broxmeyer, A.B. Middleton, B.D. Clarkson and M.A. Moore. 1983. Marrow cytogenetic and cell-culture analyses of the myelodysplastic syndromes: insights to pathophysiology and prognosis. J Clin Oncol 1:627-34.
13. Heim, S., U. Kristoffersson, N. Mandahl, F. Mitelman, A.N. Bekassy, S. Garwicz and T. Wiebe. 1986. High resolution banding analysis of the reciprocal translocation t(6;9) in acute nonlymphocytic leukemia. Cancer Genet Cytogenet 22:195-201.
14. Horsman, D.E. and D.K. Kalousek. 1987. Acute Myelomonocytic Leukemia (AML-M4) and translocation t(6;9): two additional patients with prominent myelodysplasia. Am J Haematol 26:77-82.
15. Kaneko, Y., J.D. Rowley, H.S. Maurer, D. Variakojis and J.W. Moohr. 1982. Chromosome pattern in childhood acute nonlymphocytic leukemia (ANLL). Blood 60:389-99.
16. Levin, M., M. Le Coniat, A. Bernheim and R. Berger. 1986. Complex chromosomal abnormalities in acute nonlymphocytic leukemia. Cancer Genet Cytogenet 22:113-119.
17. Nordenson, I., L. Bjermer, G. Holmgren, P. Hornsten and A. Wahlin. 1989. t(6;9) in bone marrow cells in two patients with sarcoidosis and acute myeloid leukemia. Cancer Genet Cytogenet 38:297-300.
18. Pearson, M.G., J.W. Vardiman, M.M. Le Beau, J.D. Rowley, S. Schwartz, S.L. Kerman, M.M. Cohen, E.W. Fleischman and E.L. Prigogina. 1985. Increased numbers of marrow basophils may be associated with a t(6;9) in ANLL. Am J Hematol 18:393-403.
19. Puchkova, G.P., E.L. Prigogina, E.W. Fleischmann, T.S. Drosdova, S.A. Mayakova and I.S. Peterson. 1983. Chromosome abnormalities in chronic myeloid leukemia in children. Hum Genet 64:257-62.
20. Rowley, J.D. and D. Potter. 1976. Chromosomal banding patterns in acute nonlymphocytic leukemia. Blood 47:705-722.
21. Sandberg, A.A., R. Morgan, J.A. McCallister, B. Kaiser-McCaw and F. Hecht. 1983. Acute myeloblastic leukemia (AML) with t(6;9) (p23;q34): a specific subgroup of AML? Cancer Genet Cytogenet 10:139-42.
22. Schwartz, S., R. Jiji, S. Kerman, J. Meekins and M.M. Cohen. 1983. Translocation (6;9)(p23;q34) in acute nonlymphocytic leukemia. Cancer Genet Cytogenet 10:133-8.
23. Soekarman, D., M. von Lindern, S. Daenen, B. de Jong, C. Fonatsch, B. Heinze, C.R. Bartram, A. Hagemeijer and G. Grosveld. The translocation (6;9)(p23;q34) shows consistent rearrangement of two genes and defines a myeloproliferative disorder with specific clinical features. Submitted for publication
24. Stejskalova, E., P. Goetz and J. Stary. 1990. A translocation (6;9)(p23;q34) and trisomie 13 in a case of childhood acute nonlymphocytic leukemia. Cancer Genet Cytogenet 45:139-141.
25. Testa, J.R., U. Mintz, J. Rowley, J.W. Vardiman and H.M. Golomb. 1979. Evolution of karyotypes in acute nonlymphocytic leukemia. Cancer Res 39:3619-3627.

26. Vermaelen, K., J.L. Michaux, A. Louwagie and H. Van den Berghe. 1983. Reciprocal translocation t(6;9)(p21;q33): a new characteristic chromosome anomaly in myeloid leukemias. Cancer Genet Cytogenet 10:125-31.
27. Yunis, J.J., R.D. Brunning, R.B. Howe and M. Lobell. 1984. High-resolution chromosomes as an independent prognostic indicator in adult acute nonlymphocytic leukemia. N Engl J Med 311:812-8.
28. Yunis, J.J., M.M. Oken, A. Theologides, R.B. Howe and M.E. Kaplan. 1984. Recurrent chromosomal defects are found in most patients with non-Hodgkin's-lymphoma. Cancer Genet Cytogenet 13:17-28.

CHAPTER 2.2

The human *pim*-1 gene is not directly activated by the translocation (6;9) in acute nonlymphocytic leukemia

Marieke von Lindern[1], Ton van Agthoven[1,3], Anne Hagemeijer[1], Henk Adriaansen[2] & Gerard Grosveld[1]

[1] *Department of Cell Biology and Genetics, Erasmus University, Rotterdam, PO Box 1738, 3000 DR Rotterdam;* [2] *Department of Immunology, Academic Hospital Dijkzigt, PO Box 1738, 3000 DR Rotterdam, The Netherlands*

In Acute Nonlymphocytic Leukemia (ANLL) specified by a reciprocal translocation (6;9), defined clinical and morphological features are observed. This suggests that genes located near the breakpoints of the translocation chromosomes are involved in the generation of this subtype of leukemia. The human *pim*-1 gene has been mapped near the t(6;9) breakpoint on chromosome 6. Using somatic cell hybrids we demonstrated that the *pim*-1 gene remains on chromosome 6. We investigated whether *pim*-1 plays a role in t(6;9) ANLL. The expression of *pim*-1 is elevated in two out of three patients with t(6;9) ANLL. However, the *pim*-1 gene has a size of only 6 kb and using field inversion gel electrophoresis, no chromosomal breakpoint can be detected within a distance of 165 kb from the *pim*-1 locus. Therefore it seems more likely that the elevated expression is due to the differentiation state of the cells rather than transcriptional activation by the translocation.

Introduction

In neoplastic cells of patients with acute nonlymphocytic leukemia (ANLL) many different chromosomal aberrations can be found (Larson *et al.*, 1983; Yunis *et al.*, 1984). Among them, specific aberrations have been found to be associated with distinct morphological subtypes of ANLL. For example the reciprocal translocation (8;21)(q22;q22) is associated with acute myeloid leukemia type M2 (AML-M2) (LeBeau & Rowley, 1984) and the t(15;17)(q22;q21) is almost always and exclusively found in patients with acute promyelocytic leukemia (APL) (Rowley *et al.*, 1977, Larson *et al.*, 1984). A relatively rare subtype of ANLL (0,5%) is characterized by the t(6;9)(p23;q34) and was first described by Rowley & Potter (1976). Patients with t(6;9) ANLL are relatively young, they present with ANLL of FAB-M2 (60%), M4 (30%) or even M1 and all respond poorly to therapy (Heim *et al.*, 1986). In many cases, myelodysplastic features preceded the onset of acute leukaemia. Increased marrow basophilia has been reported particularly in M2 patients (Pearson *et al.*, 1985) and the blast cells have been found to express an immature immunologic phenotype (Adriaansen *et al.*, 1988). The specific occurrence of the t(6;9) suggests that the disruption or activation of one or more genes, located near the breakpoints on chromosome 6 or 9, is involved in the generation of this type of leukemia. Consistent chromosomal translocations, of which the breakpoints have been cloned and characterized, include the t(8;14) found in Burkitt Lymphoma and the t(9;22) in chronic myelogenous leukemia (rev. by Cory 1986). Involvement of the oncogenes c-*myc* and c-*abl* has been shown, respectively.

Oncogenes known to be located near the breakpoints of the t(6;9) are c-*abl* on chromosome 9q34 (Heisterkamp *et al.*, 1982) and human *pim*-1 on chromosome 6p21 (Nagarajan *et al.*, 1986).

Westbrook *et al.*, (1985) showed by *in situ* hybridization, that in two patients with t(6;9) ANLL the c-*abl* gene is not translocated to the 6 chromosome. More over, using Southern blots, they found no rearrangement of the 3′ end of the c-*abl* gene, which is oriented towards the breakpoint.

The mouse *pim*-1 gene plays a role in MuLV induced T-cell lymphomas (Cuypers *et al.*, 1984). Integration of proviruses, mainly in the 3′ untranslated region of the *pim*-1 gene, results in high steady state levels of the *pim*-1 transcript in these cells (Selten *et al.*, 1985). The *pim*-1 gene codes for a 33 kD protein located in the cytoplasm of the cell (Selten *et al.*, 1986; Telerman *et al.*, 1988). The mouse and the human *pim*-1 gene shows a high degree of homology (Meeker *et al.*, 1987, Domen *et al.*, 1987). Until now there is no evidence for the involvement of the *pim*-1 gene in any type of human cancer.

We investigated the expression of *pim*-1 in three patients with t(6;9) ANLL and compared it to the expression of *pim*-1 in other patients with acute myeloid leukemia. Furthermore we determined whether *pim*-1 is translocated in t(6;9) ANLL by Southern blot analysis of somatic cell hybrids containing the segregated translocation chromosomes. Subsequently we looked for rearrangements in a large area surrounding both *pim*-1 and c-*abl*, using field inversion gel electrophoresis (FIGE) (Carle *et al.*, 1986).

Results

Expression of pim-*1 in t(6;9) ANLL*

From a cDNA library of K562 a full length *pim*-1 cDNA (hp63) was isolated (Domen *et al.*, 1987).

Total RNA was extracted from bone marrow cells of 10 ANLL patients, with or without t(6;9). Patient identification, cytogenetic and morphological data are given in Table 1. All but one bone marrow sample contained more than 60% blast cells. The exception was the bone marrow of patient BA, which contained only 30% blast cells. Northern blots of the RNAs were probed with the *pim*-1 cDNA clone hp63. The results are shown in Figure 1. In the majority of the patients, *pim*-1 is

[3] Present address: Dr Daniel den Hoed Cancer Center, Groene Hilledijk 301, 3078 KG Rotterdam, The Netherlands

Received 9 August 1988; accepted 21 September 1988

Table 1 Cytogenetic aberrations and FAB classification of the blast cells present in the bone marrow of the examined patients

Patient	*Sex Age (years)*	*Cytogenetic aberration*	*FAB*	*Percentage blast cells*
DK	F, 16	t(6;9)(p23;q34)	AML M4	60
BL	F, 13	t(6;9)(p23;q34)	AML M4	75
BA	F, 3	t(6;9)(p23;q34)	AML	30
BN	M, 19	der(6)t(6;9)(p24;q32), −9, complex	AUL	>80
JR	F, 36	dir dup(6)(p21.1p22.3)	AML M4	60
WA	M, 43	normal	AUL	75
SE	M, 19	normal	AUL	95
KO	F, 68	ins(3;3), t(7;15;17), −7	AML M1	85
BV	F, 55	normal	AML M4	80
DV	M, 39	−Y, t(8;21)	AML M2	80

expressed at a low to moderate level in bone marrow cells. *Pim*-1 is expressed at a relatively high level in the bone marrow of two patients with a typical t(6;9) ANLL, i.e. DK and BA, after correction for the lower percentage of blasts in the latter case. In a third patient (BL), elevated *pim*-1 expression is not observed. As a control, the Northern blot contains RNA from bone marrow of several other ANLL patients with immature blast cell phenotype, including two patients with an aberration on chromosome 6 (BN, JR). These samples did not contain elevated *pim*-1 mRNA levels, except for patient WA. In bone marrow cells of patient WA no karyotypic aberrations were detected. Immunologically this patient presented with undifferentiated myelocytic leukemia with blast cells positive for the same set of markers as the blast cells of t(6;9) patients DK and BL (Adriaansen *et al.*, 1988).

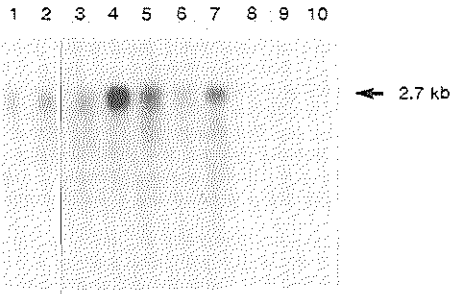


Figure 1 Expression of *pim*-1 in bone marrow of leukemic patients with or without t(6;9) ANLL. Total RNA (20 μg), isolated from bone marrow cells of patients KO, BV, SE, DK, BA, BL, WA, JR, BN, DV (lane 1 to 10) was fractionated in agarose, transferred to nitrocellulose and hybridized to the [^{32}P]-labeled hp63 probe. Before loading the blotting gel, 1/20 of each sample was tested on a 1% agarose gel containing ethidium bromide to be sure that an equal amount of RNA was present in each lane. Patients BL and DK are described in more detail by Adriaansen *et al.* (1988)

The size of the *pim*-1 transcript is 2.7 kb and therefore does not appear to be altered in patients with t(6;9) ANLL, nor in any of the other patients.

Translocation of c-abl *and* pim-*1 in t(6;9)*

In order to determine whether *pim*-1 and c-*abl* are translocated to respectively the 9q$^+$ or the 6p$^-$ chromosome in t(6;9) ANLL, bone marrow cells of patient DK were fused to hamster a23 cells and 60 hybrid cell lines were isolated. After karyotyping, 14 of these hybrid cell lines showed segregation of the chromosome of interest: 6, 6p$^-$, 9 and 9q$^+$. All informative cell lines were tested on Southern blots and gave concordant results. Only 5 representative lines are demonstrated.

DNA extracted from hybrid cell lines containing chromosome 6, 6p$^-$, 9 or 9q$^+$ were hybridized to *pim*-1 (hp63) and c-abl (K38) sequences (Table 2). The hybrid cell line MA3B contains chromosome 6p$^-$ and still hybridizes to *pim*-1 sequences, indicating that *pim*-1 is not translocated to chromosome 9q$^+$ in this t(6;9) ANLL patient. The hybrid cell line MA4C with chromosome 6 and 9q$^+$ contains human c-*abl* sequences, while in cell lines MA3B and MA6C (#6p$^-$) they are absent. Therefore c-*abl* is not translocated to chromosome 6p$^-$ in this patient either.

As a control, the human chromosome 6 markers D6S8 and D6S7 were hybridized to the same filter. D6S8 is located 4 cM telomeric of the HLA locus, while D6S7 is also located telomeric of HLA, but not genetically linked to it (Leach *et al.*, 1986). The hybrid cell line MA3B (#6p$^-$) contains D6S8 sequences. Thus the breakpoint on chromosome 6 in this t(6;9) is situated telomeric of D6S8. DNA isolated from MA3B does not hybridize to the D6S7 probe, indicating that this marker is translocated to chromosome 9q$^+$.

Table 2 hybrid cell lines* containing chromosomes 6, 6p$^-$, 9 or 9q$^+$, isolated after fusion of bone marrow cells from patient DK with a23 Chinese hamster cells

	Chromosomes						*Markers*	
Cell line	*6*	*6p$^-$*	*9*	*9q$^+$*	c-abl	pim-*1*	*D6S8*	*D6S7*
MA 13A	+				−	+	+	+
MA 3B		+			−	+	+	−
MA 6C	+	+			−	+	+	+
MA 4C	+			+	+	+	+	+
MA 5C			+		+	−	−	−

Chromosomes: + chromosome seen in at least 10/16 metaphases
Markers: + hybridizing, − no hybridizing band observed
*DNA from these hybrid lines was digested with EcoRI, fractionated in 0.7% agarose, blotted on GeneScreenPlus and hybridized to the [^{32}P]-labeled probes K38 (c-*abl*), hp63 (*pim*-1), D6S8 and D6S7

Table 3 DNA fragments*, generated by digestion with rare cutting restriction enzymes, hybridizing to *pim*-1 sequences

	MluI	*NotI*	*PvuI*	*SalI*	*SfiI*
MluI	525				
NotI	225	225			
PvuI	360	225	440		
SalI	425	225	440	600	
SfiI	250	225	250	250	250
	300		300	300	300

*DNA was isolated from Hela cells, single and double digests were performed. Bacteriophage lambda multimers were used as molecular weight markers. Sizes are in kilobases

FIGE mapping of the pim-*1 locus*

The *pim*-1 locus has been mapped with rare cutting restriction endonucleases, to determine which enzymes will screen a large area around the locus for a possible breakpoint on chromosome 6. Single and double digests were performed with SfiI, PvuI, MluI, SalI and NotI on DNA derived from Hela cells. Hybridization of the digested DNA to *pim*-1 sequences resulted in the fragments shown in Table 3.

In a single SfiI digest, the *pim*-1 probe detects two fragments of 250 and 300 kb. Both bands remain visible after hybridization with a probe containing only the large 3′ exon of *pim*-1 in which no SfiI site is present (Domen *et al.*, 1987). In a SfiI-NotI double digest only one band of 225 kb is detected. As both the 250 and the 300 kb band are present in equimolar amounts in DNA derived from Hela cells as well as in DNA derived from K562 cells, a polymorphic SfiI site flanks the *pim*-1 locus.

With the data of the MluI, PvuI and SalI single and double digests a long range restriction map can be constructed as presented in Figure 2. In double digests with NotI and any of the other enzymes used, no fragments smaller than the 225 kb NotI fragment were generated. Therefore, the NotI sites can not be mapped relative to the other restriction enzyme sites. For the same reason the SfiI sites cannot be mapped either. The NotI and SfiI sites, which are located within the 360 kb PvuI-MluI fragment are not depicted in Figure 2.

The *pim*-1 gene is located in a 360 kb PvuI-MluI fragment. At one side this fragment is flanked by a 165 kb MluI-PvuI fragment, at the other side by a 175 kb MluI-SalI fragment. DNA of the hybrid lines MA13A (#6), MA6C (#6, #6p⁻) and MA3B (#6p⁻) was digested with MluI and SalI. Hybridization to a *pim*-1 cDNA probe revealed hamster fragments of 400 and 550 kb and a human fragment of 800 kb in the SalI digest. In theMluI digest there are hamster bands of 40 and 150 kb and a human band of 800 kb (Figure 3). The fragments containing the human *pim*-1 sequences are of the same size in DNA from both chromosome 6 and chromosome 6p⁻ containing hybrid lines. However,

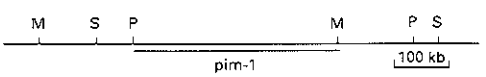


Figure 2 Physical map of a region around the *pim*-1 locus. Restriction enzyme sites are MluI (M), PvuI (P) and SalI (S). The localization of the *pim*-1 gene within the 360 kb PvuI-MluI fragment can not be determined. Double and single digests have been performed on DNA isolated from Hela cells; single digests have been confirmed on DNA isolated from K562 cells

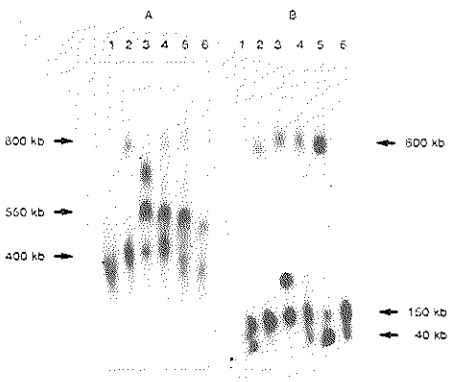


Figure 3 DNA fragments of chromosome 6 and 6p⁻ hybridizing to *pim*-1 sequences. DNA in agarose blocks, from a23 cells (lane 1) and from the hybrid lines MA13A (#6), MA6C (#6, #6p⁻), MA3B (#6p⁻), MA4C (#6, #9q⁺) and MA5C (#9) (lanes 2 to 6) was digested with SalI (a) and MluI (b) separated by FIGE, transferred to GeneScreenPlus and hybridized to *pim*-1 probes. Lambda multimers were used as molecular weight markers

these fragments are larger than the fragments seen in DNA isolated from Hela cells. This is most likely due to complete CpG methylation of at least one of the MluI/SalI recognition sites shown in Figure 2. The pattern of the hamster fragments hybridizing to *pim*-1 sequences, also varies due to CpG methylation. When the experiment is repeated with DNA isolated from KG-1 or K562 cells, fragments of 525 and 800 kb are detected in the MluI digest and fragments of 600 and 800 kb in the SalI digest. Since the fragments are present in variable proportions, it indicates that in these cell lines the MluI and SalI recognition sites are partially CpG methylated. Digesting the DNA over longer periods and with a larger amount of enzyme did not change the relative proportions of the bands. However, as the human *pim*-1 fragment has the same size in both the chromosome 6 and the chromosome 6p⁻ containing hybrid cell line, the conclusion must be that the minimal distance of *pim*-1 to the breakpoint measures 165 kb.

FIGE analysis of c-abl, *D6S8 and D6S7*

The orientation of c-*abl* on chromosome 9 is 5′ centromeric, 3′ telomeric (Heisterkamp *et al.*, 1983). A map of the NotI, SfiI and SacII sites in c-*abl* has been published by Westbrook *et al.* (1987). We could confirm the position of the SfiI and NotI sites in DNA extracted from different sources. NotI sites are present in the regions around the IB and the IA exons of c-*abl*. SfiI sites are present near the IB exon and in the most 3′ exon of c-*abl*. As c-*abl* is not translocated, a breakpoint is expected to be located 3′ of the gene. DNA of the hybrid lines MA5C (#9) and MA4C (#6 and #9q⁺) was digested with SfiI and NotI and hybridized with the c-*abl* probe K38. This probe is located 3′ of the SfiI site located in the last exon of c-*abl*. Both DNAs contain SfiI and NotI fragments of 300 kb (SfiI digest: Figure 4, NotI digest not shown). Also in MluI and SalI digests bands of the same size showed up after hybridization with K38. As the SfiI fragment is 300 kb in size, the distance between c-*abl* and the t(6;9) breakpoint on chromosome 9 is at minimum 300 kb in patient DK.

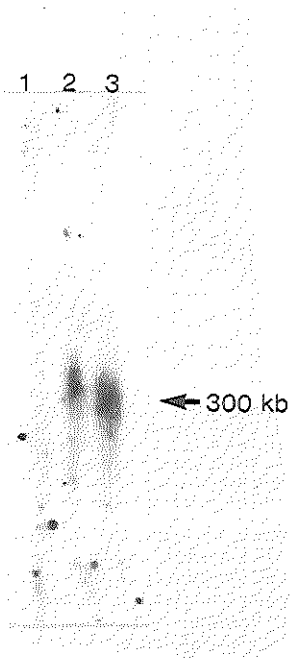


Figure 4 DNA fragments of chromosomes 9 and $9q^+$ hybridizing to c-*abl* probes. DNA, isolated in agarose blocks, from a23 cells (lane 1) and from the hybrid lines MA4C (#6, $\#9q^+$) and MA5C (#9) (lane 2 and 3) was digested with SfiI, separated by FIGE, transferred to GeneScreenPlus and hybridized to the c-*abl* probe K38. Lambda multimers were used as molecular weight markers

Filters with DNA of hybrid lines containing chromosomes 6, $6p^-$, 9 and $9q^+$ digested with SalI and MluI, were also hybridized with the chromosome 6 markers D6S8 and D6S7 (data not shown). Human specific fragments of more than 800 kb could be detected. No smaller bands were present in the $6p^-$ (D6S8) or $9q^+$ (D6S7) lanes. As the fragments migrated in the compression zone of the gel, we formally can not exclude the possibility, that aberrant fragments containing D6S8 or D6S7 are present in SalI or MluI digests of DNA from hybrid cell lines containing chromosome $6p^-$ or chromosome $9q^+$.

Discussion

We have isolated somatic cell hybrids containing the $6p^-$ and the $9q^+$ chromosome of a patient with t(6;9) ANLL. Using these cell hybrids, we could show that the *pim*-1 gene, located on chromosome 6p21 (Nagarajan *et al.*, 1986), is not translocated to the $9q^+$ chromosome. The c-*abl* gene, located on chromosome 9q34 (Heisterkamp *et al.*, 1982), is not translocated either, which is in accordance with a previous report by Westbrook *et al.* (1985).

Using field inversion gel electrophoresis, we have demonstrated that the minimal distance of the *pim*-1 locus to the breakpoint on chromosome 6 is 165 kb in a patient with t(6;9) ANLL. In man as well as in mouse the *pim*-1 locus encompasses no more than 6 kb (Selten *et al.*, 1986, Meeker *et al.*, 1987). A distance of more than 165 kb between the breakpoint and the *pim*-1 locus is probably too large to assume activation of the *pim*-1 locus by the t(6;9) in ANLL.

Northern blots with total RNA of patients with myeloid leukemia were hybridized to *pim*-1 sequences.

The control patients were selected on morphological or immunological grounds to home a bone marrow population, closely resembling the bone marrow of t(6;9) patients. One of the features of the blast cells of t(6;9) ANLL patients is that they express both myeloid surface antigens (My7, My9) and the T-cell determinant TdT (Adriaansen *et al.*, 1988). A similar set of immunological markers was present on blast cells of AUL (Acute undifferentiated leukemia) patients WA and SE as well as on the AML M1 cells of patient KO. A relatively high expression level of *pim*-1 is found in only one of these patients (WA). Out of three patients with t(6;9) ANLL, two expressed *pim*-1 at a relatively high level.

On morphological criteria t(6;9) ANLL is not diagnosed as a single FAB subtype. Most t(6;9) ANLL's are classified as FAB-M2, some as FAB-M4 and few as FAB-M1. Though it is an acute leukemia, some maturation can be found in t(6;9) ANLL. Variation in the extent of maturation is probably responsible for the observed phenotypic heterogeneity. When the expression of the *pim*-1 gene is restricted to a limited stage of hematopoietic differentiation, this phenotypic heterogeneity may account for the observed variation of *pim*-1 expression in t(6;9) ANLL.

We favor the idea that expression of *pim*-1 in t(6;9) ANLL merely reflects the differentiation stage of the blast cells. Nevertheless the expression level of *pim*-1 in t(6;9) ANLL may play a role in this leukemia through a synergistic effect. Generally, tumorigeneses is believed to be a multistep process (rev. by Klein & Klein, 1985). The expression of a kinase like *pim*-1 during the proliferation of haematopoietic cells could set the stage for the transformation of these cells by the activation (or inactivation) of a gene, which is necessary, but on its own not sufficient for full transformation of these cells into tumor cells. The elevated expression of *pim*-1 may be needed in the signaling pathway of a mitogenic stimulus in much the same way as the transformation of NIH3T3 cells by p21 *ras* seems to be dependent on protein kinase C (PKC) expression (Kamata *et al.*, 1987).

Materials and methods

RNA analysis

Total RNA was isolated, using the guanidine isothiocyanate method (Chirgwin *et al.*, 1979). 20 μg of total RNA was electrophoresed on a 0.8% agarose gel in the presence of formaldehyde, blotted onto nitrocellulose and hybridized in 10% dextran, 50% formamide (Maniatis *et al.*, 1982). Final washing was performed in 0.3 × SSC, 0.1% SDS at 65°C. DNA probes were labeled according to Feinberg & Vogelstein (1982).

Cell hybrids

Production, isolation and propagation of human-hamster cell hybrids was performed as described by Geurts van Kessel *et al.* (1981). Bone marrow cells of patient DK and a23 Chinese hamster cells (TK^-) were fused, using inactivated Sendai virus as fusogen. Independent hybrid clones were screened for the

presence of chromosomes 6, $6p^-$, 9 and $9q^+$. Cytogenetic analysis and DNA extractions were always carried out on the same batches of cells.

Southern blotting

High molecular weight DNAs, isolated as described by Jeffreys *et al.* (1977), were digested with restriction enzymes and electrophoresed on 0.7% agarose gels. The gel was transferred to GeneScreenPlus in 0.4 M NaOH. After blotting the filters were rinsed in 0.5 M Tris (pH 7), 3 M NaCl and distilled water before drying. Hybridization and washing of the filters was performed as described for Northern blots, except that SDS was added to 1% in the hybridization mix and in the first washing steps. To remove hybridized probe, filters were rinsed with 0.4 M NaOH at room temperature for 20–30 min and then neutralized with 3 × SSC.

Field inversion gel electrophoresis (FIGE)

Hybrid hamster-human cells were washed twice with 0.9% NaCl after trypsinization, suspended in 0.5% low melting agarose (BRL) in 0.9% NaCl at a concentration of 6.2×10^6/ml^{-1} and poured into precooled plexiglass casting molds, containing 80 µl slots. After cooling, the agarose blocks were incubated in lysis mix (0.5 M EDTA, pH 8.5, 1% Na-lauroyl sarcosine), containing 1 mg ml^{-1} Proteinase K at 50°C overnight. New lysis mix and Proteinase K were added and incubation at 50°C continued for 24 h. The blocks were stored in lysis mix at 4°C. Before restriction enzyme digestion, the blocks were washed twice in 10 mM Tris (pH 8.0), 1 mM EDTA (T10E1) with 1 mM phenylmethylsulphonylfluoride (PMSF) for 30 min at 50°C and once in T10E1 for 30 min at room temperature. The blocks were digested twice with 20 U restriction enzyme for 3 h in a total volume of 160 µl, under conditions recommended by the manufacturer. The blocks with digested DNA were stored in 0.5 M EDTA. Half a block was placed in a 0.6% agarose gel, along with molecular weight markers of bacteriophage Lambda multimers. The gel buffer was 0.2 × TBE (18 mM Tris, 18 mM borate, 0.4 mM EDTA) and the gel was cooled to 12°C during the run, by recirculating cooled buffer. A programmable power inverter (MJ Research, Inc.) was placed between the power supply and the geltank. The inversion program used was: initial reverse time: 0.5; reverse increment: 0.1; initial forward time: 1.0; forward increment: 0.4; 100 steps. This program gave a good separation between 50 and 1000 kb. After staining the gel with ethidium bromide to visualize the DNA, the gel was blotted on GeneScreenPlus as described above for Southern blots.

Acknowledgements

We gratefully acknowledge Drs Peter O'Connell and Ray White for providing the probes D6S7 and D6S8.

We want to thank Hans Morreau for initiating the FIGE technique, Drs K. Hählen, B. Löwenberg and A.J.P. Veerman for providing the patient material, Prof. D. Bootsma, Dies Meijer and André Hermans for helpful discussion and encouragement, Rita Boucke for typing the manuscript and Joop Fengler and Cor van Dijk *et al.* for preparing the figures.

This work was supported by the Netherlands Cancer Society (Koningin Wilhelmina Fonds).

References

Adriaansen, H.J., Van Dongen, J.J.M., Hooijkaas, H., Hählen, K., Van't Veer, M.B., Löwenberg, B. & Hagemeijer, A. (1988). *Leukemia*, **2**, 136–140.

Carle, G.F., Frank, M., Olson, M.V. (1986). *Science*, **232**, 65–68.

Chirgwin, J.M., Przybyla, A.E., MacDonald, J. & Rutter, W.J. (1979). *Biochemistry*, **18**, 5294–5299.

Cory, S. (1986). *Avd. in Cancer Res.*, **47**, 189–234.

Cuypers, H.T., Selten, G., Quint, W., Zylstra, M., Robanus Maandag, E.R., Boelens, W., Van Wezenbeek, P., Melief, C. & Berns, A. (1984). *Cell*, **37**, 141–150.

Domen, J., Von Lindern, M., Hermans, A., Breuer, M., Grosveld, G. & Berns, A. (1987). *Oncogene Res.*, **1**, 103–112.

Feinberg, A.P. & Vogelstein, B. (1982). *An. Bioch.*, **132**, 6–13.

Geurts van Kessel, A.H.M., Ten Brinke, H., Boere, W.A.M., Den Boer, W.C., De Groot, P.G., Hagemeijer, A., Meera Khan, P. & Pearson, P.L. (1981). *Cytogenet. Cell. Genet.*, **30**, 83–91.

Heim, S., Kristoffersson, U., Mandahl, N., Mitelman, F., Bekassy, A.N., Garwicz, S. & Wiebe, T. (1986). *Cancer Genet. Cytogenet.*, **22**, 195–201.

Heisterkamp, N., Groffen, J., Stephenson, J.R., Spurr, N.K., Goodfellow, P.N., Solomon, E., Carrit, B. & Bodmer, W.F. (1982). *Nature*, **299**, 747–750.

Heisterkamp, N., Stephenson, J.R., Groffen, J., Hansen, P.F., de Klein, A., Bartram, C.R. & Grosveld, G. (1983). *Nature*, **306**, 239–242.

Jeffreys, A., Flavell, R.A. (1977). *Cell*, **12**, 429–439.

Kamata, T., Sullivan, N.F., Wooten, M.W. (1987). *Oncogene*, **1**, 19–27.

Klein, G. & Klein, E. (1985). *Nature*, **315**, 190–195.

Larson, R., LeBeau, M., Vardiman, J., Testa, J., Golomb, H. & Rowley, J. (1983). *Cancer Genet. Cytogenet.*, **10**, 219–236.

Larson, R.A., Kondo, K., Vardiman, J.W., Butler, A.E., Golomb, H.M. & Rowley, J.D. (1984). *Am. J. Med.*, **76**, 827–841.

Leach, R., DeMars, R., Hastedt, S. & White, R. (1986). *Proc. Natl. Acad. Sci.*, **83**, 3909–3913.

LeBeau, M.M. & Rowley, J.R. (1984). *Cancer Surveys*, **3**, 371–394.

Maniatis, T., Fritsch, E.F. & Sambrook, J. (eds). (1982). *Molecular Cloning: A Laboratory Manual*. Cold Spring Harbor Laboratory Press: New York.

Meeker, T.C., Nagarajan, L., Ar-Rushdi, A., Rovera, G., Huebner, K. & Croce, C.M. (1987). *Oncogene Res.*, **1**, 87–101.

Nagarajan, L., Louie, E., Tsujimoto, Y., Ar-Rushdi, A., Huebner, K. & Croce, C.M. (1986). *Proc. Natl. Acad. Sci.*, **83**, 2556–2560.

Pearson, M.G., Vardiman, J.W., LeBeau, M.M., Kowley, J.D., Schwartz, S., Kerman, S.L., Cohen, M.M., Fleischman, E.W. & Prigogina, E.L. (1985). *Am. J. Hem.*, **18**, 393–403.

Rowley, J.D. & Potter, D. (1976). *Blood*, **47**, 705–721.

Rowley, J.D., Golomb, H. & Dougherty, C. (1977). *Lancet*, **i**, 549–550.

Selten, G., Cuypers, H.T. & Berns, A. (1985). *EMBO J.*, **4**, 1793–1798.

Selten, G., Cuypers, H.T., Boelens, W., Robanus-Maandag, E., Verbeek, J., Domen, J., Van Beveren, C. & Berns, A. (1986). *Cell*, **46**, 603–611.

Telerman, A., Amson, R., Zakut-Houri, R. & Givol, D. (1988). *Mol. Cell. Biol.*, **8**, 1498–1503.

Westbrook, C.A., LeBeau, M.M., Diaz, M.O., Groffen, J. & Rowley, J.D. (1985). *Proc. Natl. Acad. Sci. USA*, **82**, 8742–8746.

Westbrook, C.A., Rubin, C.M., LeBeau, M.M., Kaminer, L.S., Smith, S.D., Rowley, J.D. & Diaz, M.O. (1987). *Proc. Natl. Acad. Sci. USA*, **84**, 251–255.

Yunis, J.J. (1984). *Cancer Genet. Cytogenet.*, **11**, 125–137.

CHAPTER 2.3

The (6;9) Chromosome Translocation, Associated with a Specific Subtype of Acute Nonlymphocytic Leukemia, Leads to Aberrant Transcription of a Target Gene on 9q34

MARIEKE VON LINDERN,[1] ANNEMARIE POUSTKA,[2]† HANS LERACH,[2]‡ AND GERARD GROSVELD[1]*

Department of Cell Biology and Genetics, Erasmus University, P.O. Box 1738, 3000 DR Rotterdam, The Netherlands,[1] and European Molecular Biology Laboratory, 6900 Heidelberg 1, Federal Republic of Germany[2]

Received 25 January 1990/Accepted 1 May 1990

The specific (6;9)(p23;q34) chromosomal translocation is associated with a defined subtype of acute nonlymphocytic leukemia (ANLL). The 9q34 breakpoint is located at the telomeric side of the c-*abl* gene. Through a combination of chromosome jumping, long-range mapping, and chromosome walking, the chromosome 9 breakpoints of several t(6;9) ANLL patients were localized within a defined region of 8 kilobases (kb), 360 kb telomeric of c-*abl*. Subsequent cDNA cloning revealed that this region represented an intron in the middle of a gene, called Cain (*can*), encoding a 7.5-kb transcript. Disruption of the *can* gene by the translocation resulted in the expression of a new 5.5-kb *can* mRNA from the 6p$^-$ chromosome. Isolation of chromosome 6 sequences showed that breakpoints on 6p23 also clustered within a limited stretch of DNA. These data strongly suggest a direct involvement of the translocation in the leukemic process of t(6;9) ANLL.

The occurrence of defined chromosomal translocations in specific subtypes of leukemia strongly suggests that these translocations play an important role in the process of leukemogenesis. Translocations may activate nearby cellular genes involved in the control of proliferation and/or differentiation.

Translocations in B-cell neoplasms often involve chromosome 14q32, 2p12, or 22q11, where the immunoglobulin heavy-chain, κ light-chain and λ light-chain loci are located, respectively (10, 13, 29, 32). Translocations in T-cell neoplasms often involve chromosome 14q11 or 7q35, where the T-cell receptor (TCR) α/δ and β/γ chain loci are located (9, 25, 41). Cloning and molecular characterization of breakpoints involving these chromosomes showed that breakpoints occur within the immunoglobulin and TCR genes, e.g., t(8;14) in Burkitt's lymphoma (45), t(11;14) in B-cell lymphoma and leukemia (48), t(14;18) in follicular lymphoma (47), t(7;9) in T-cell lymphoma (42), and t(7;19) in T-cell leukemia (8). Sequence analysis of chromosomal breakpoints suggested that mechanisms involved in immunoglobulin and TCR gene rearrangement are responsible for the translocation of proto-oncogenes to the immunoglobulin or TCR gene environment. Target genes activated by a translocation have been cloned and characterized, e.g., c-*myc* (38), *bcl-2* (46), *tcl-3* (42), and *lyl-1* (36).

In contrast, similar analyses of chromosome translocations in myeloid leukemias have been less successful. To date, the only chromosomal translocation characterized at the molecular level is t(9;22), occurring in 95% of chronic myeloid leukemia, 25% of acute lymphoid leukemia, and 2% of acute myeloid leukemia (AML) cases (31). This Philadelphia translocation involves the *bcr* gene on chromosome 22 and c-*abl* on chromosome 9. As a result, chimeric *bcr-abl* genes are generated (17, 23, 44) encoding hybrid Bcr-Abl proteins (3, 30, 51). These proteins have enhanced tyrosine kinase activity and tumorigenic properties (18, 34, 35), strongly suggesting an active role of the translocation products in the generation of these leukemias. In contrast to the immunoglobulin and TCR genes, *bcr* is not a rearranging gene, and there are no indications that the Philadelphia translocation is the product of faulty gene rearrangements (5, 11, 22, 49).

To better understand the role and the mechanism of chromosome translocations in myeloid leukemogenesis, the t(6;9)(p23;q34) translocation in acute nonlymphocytic leukemia (ANLL) was analyzed. ANLL defined by t(6;9) is a relatively rare disease, associated with specific clinical and morphological features. A myeloid dysplastic phase can precede the acute leukemia, which has a poor prognosis and mostly affects young adults. Leukemic cells are classified as FAB M2 (60%), M4 (30%), or M1 (1, 20, 39). The t(6;9) is often the sole abnormality present in the neoplastic cells, which argues for a direct involvement of the translocation in the leukemic process.

As defined by cytogenetic analysis, the breakpoint on chromosome 9 maps in the same chromosomal band (9q34) as c-*abl*. However, molecular analysis showed that the t(6;9) breakpoint is located downstream of c-*abl* (52) at a distance of at least 300 kilobases (kb) (50). On chromosome 6, the human *pim-1* gene was thought to be located near the t(6;9) breakpoint (37), but long-range mapping experiments showed that the distance between the t(6;9) breakpoint and *pim-1* measures at minimum 165 kb (50). Therefore, it seems likely that neither c-*abl* nor *pim-1* is the target for translocation events in t(6;9) ANLL.

In order to characterize the gene or genes involved in the leukemic process in t(6;9) ANLL, the aim was to first clone the breakpoints on chromosomes 9q34 and 6p23, followed by identification of possible target genes in these regions. As the chromosomal breakpoints are located at a considerable distance from the available probes, a *Not*I jumping library was used (40) to generate probes that allow detection of the breakpoints by long-range mapping analysis. Subsequent

* Corresponding author.

† Present address: Department of Cell and Tumor Biology, Deutsches Krebsforschungszentrum, Im Neuenheimer Feld 280, 6900 Heidelberg 1, Federal Republic of Germany.

‡ Present address: Imperial Cancer Research Fund, Lincoln Inns Field, London WC2A 3PX, United Kingdom.

TABLE 1. Cytogenetic aberrations and FAB classification of blast cells in bone marrow of patients examined

Patient (patient no.)	Karyotypic abnormality	FAB	Reference
DK (2)	t(6:9)(p23:q34)	AML M4	1
PM	t(6:9)(p23:q34)	AML M4	20
JK	t(6:9)(p23:q34)	AML M4	
PL	t(6:9)(p23:q34)	AML M7	
SE	None	AUL[a]	1a

[a] AUL, Acute undifferentiated leukemia.

chromosome-walking experiments eventually led to the isolation of the breakpoints. Following this protocol and starting from c-*abl*, the breakpoint on 9q34 was mapped and cloned. Detailed Southern blot analysis showed that on chromosome 9 as well as on chromosome 6, the breakpoints in four t(6:9) ANLL patients mapped within remarkably small regions of DNA. On chromosome 9, this region represents an intron of a gene with an as yet unknown function, coding for a 7.5-kb transcript. Data presented in this paper argue for a defined role of the translocation in t(6:9) ANLL. The use of jumping libraries in combination with long-range restriction mapping proves to be a powerful approach in cloning chromosomal breakpoints.

MATERIALS AND METHODS

Patient material and cell lines. The cytogenetic aberrations and FAB classifications of the blast cells present in bone marrow of the patients examined are given in Table 1. Patient SE had been included as a control in a study of the expression of *pim-1* in t(6:9) ANLL (50). When an aberrant transcript was found by using *can* probes, it was decided to include this patient in the present study.

Isolation and characterization of the somatic cell hybrids established after fusion of leukemic cells from patient DK with a23 hamster cells have been described previously (50).

Southern and Northern (RNA) blotting. High-molecular-weight DNA was prepared as described by Jeffreys and Flavell (26) or in agarose blocks (50). Field inversion gel electrophoresis (FIGE) was performed as described previously (50) with a programmable power inverter (MJ Research, Inc., Boston, Mass.).

RNA was isolated by either the guanidinium isothiocyanate (6) or the LiCl-Ureum method (2). Total RNA was electrophoresed and blotted as described by Fourney et al. (15).

Filters containing RNA or genomic DNA were hybridized in 10% dextran–50% formamide (33). Final washing was performed in 0.1× SSC–0.1% sodium dodecyl sulfate at 65°C, unless stated otherwise (1× SSC is 0.15 M NaCl, 0.015 M sodium citrate). Probes were labeled by the method of Feinberg and Vogelstein (14).

Chromosome jumping and walking. Procedures for construction and screening of the *Not*I jumping library have been described by Poustka et al. (40).

To generate repeat-free probes, subcloned fragments of phage or cosmid clones were digested with frequently cutting restriction endonucleases (*Rsa*I, *Dde*I, *Hin*fI, and *Hae*III). Fragments were separated on 1 to 2% agarose gels, blotted on Zetaprobe in 0.4 N NaOH, and hybridized to sonicated total human DNA labeled with ^{32}P. Fragments that failed to hybridize were tested on Southern blots containing *Eco*RI-digested human DNA. Probes recognizing single fragments were used for hybridizations.

The CML0 cosmid library was described by Hermans et al. (24), and the CML0 λEMBL3 library was reported by de Klein et al. (11). The K562 λEMBL3 library was constructed by the method of Frischauf et al. (16). The total complexity of the library was 3×10^5 PFU. Both the CML0 (2×10^6 plaques) and the K562 (3×10^5 plaques) λEMBL3 libraries yielded on average 10 to 20 clones with every screening. For subsequent chromosome walking, DNA of the isolated phages was digested with *Eco*RI and *Eco*RI plus *Sal*I to select for the clone extending most telomerically. These clones were mapped with *Eco*RI, *Bam*HI, and *Hin*dIII and are shown in Fig. 3 (A11C2 and A11F3 to A11F9). Phage clones A11–4, A11F3, A11F4, and A11F6 were isolated from the CML0 library by using probes Ajl.R3, A11C2HSH3, A11F3BS, and A11F4EP, respectively. Cosmid clone A11C2 was isolated by using probe A11BB. Phage clones A11F7, A11F8, and A11F9 were isolated from the K562 library by using probes A11F6E.3, A11F7H, and A11F8ES, respectively (see Fig. 3).

cDNA cloning. The K562 cDNA library in λgt10 was constructed by A. Hermans (12), and the human testis cDNA library in λgt11 was purchased from Clonetech (Palo Alto, Calif.). Every screening (2×10^6 plaques of each library) gave rise to approximately 15 new cDNA clones. The K562 cDNA library was screened with a 300-base-pair (bp) genomic *Eco*RI-*Hin*dIII fragment (A11F3EH, Fig. 3) and yielded clone hX8 (see Fig. 5). hX8 was used to screen the testis cDNA library, producing clone hXT23. Further cDNA walking was performed on the testis cDNA library. cDNA inserts were subcloned into pTZ19 (Pharmacia), and each time the cDNA clone extending most telomerically was used to rescreen the same set of filters of the testis cDNA library. In this way, newly detected plaques could be distinguished from clones hybridizing to the previous probe; cDNA clones hXT37, hXT54, and hXT65 were isolated subsequently.

Cloning of chromosome 6 sequences. DNA of the hybrid cell line MA4C, containing the $9q^+$ chromosome of patient DK, was digested with *Eco*RI and loaded on a 0.7% low-melting-point agarose (Bio-Rad Laboratories) gel. After electrophoretic seperation, gel slices containing DNA fragments of approximately 4 kb were isolated and tested on a slot blot for the presence of the 4-kb *Eco*RI fragment that hybridizes to probe A11F4EP (see Fig. 3). The positive fraction was purified over a Qiagen Tip5 (Diagen) and cloned in λgt10. A total of 1.5×10^5 PFU was screened with probe A11F4EP. Two independent clones were isolated, containing the chimeric 4-kb *Eco*RI fragment.

A λEMBL3 library containing bone marrow DNA from patient PM was constructed by the method of Frischauf et al. (16). A total of 1.6×10^6 PFU was screened with probes A11F6E.3 and A11F7EBD4 (see Fig. 3 and 7). A single chimeric phage clone, M1-8 (see Fig. 7A and B), was isolated.

RESULTS

Generating probes 3′ of c-*abl*. In t(6;9) ANLL, the breakpoint on chromosome 9 is located 3′ (telomeric) of a 300-kb *Not*I fragment that contains the c-*abl* IA and common exons (Fig. 1) (51). Since c-*abl* is located at the centromeric end of this fragment, a 450-bp *Bgl*II-*Bam*HI c-*abl* probe (IA-BB, Fig. 1) was used to screen a *Not*I-*Bam*HI jumping library. Eight independent phage clones were isolated and analyzed. Six phages contained the 650-bp c-*abl* *Not*I-*Bam*HI fragment, the *supF* tag plasmid (1.9 kb), and a novel *Not*I-*Bam*HI fragment of 1.6 kb (Fig. 1).

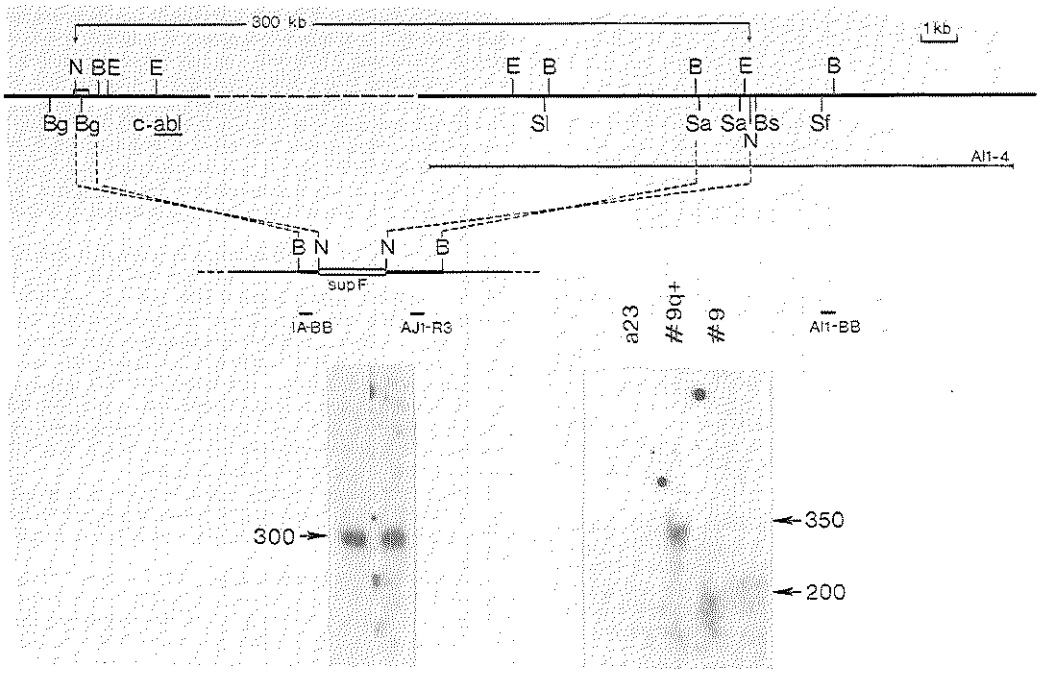


FIG. 1. Analysis of clones derived from a *Not*I jumping library. The top line represents the 300-kb *Not*I fragment containing the human c-*abl* gene at the left end (centromeric). The *Not*I site in the c-*abl* 1A exon is indicated. The line beneath the map represents one of six λ jumping clones isolated from a *Not*I-*Bam*HI jumping library. In addition to the 650-bp *Not*I-*Bam*HI c-*abl* fragment and the *supF* tag plasmid, these clones contained a 1.6-kb *Bam*HI-*Not*I fragment representing the telomeric end of the 300-kb *Not*I fragment. Probes IA-BB and Aj1R3 both detected a 300-kb *Not*I fragment in a FIGE blot containing *Not*I-digested HeLa cell DNA (bottom left). The probes are indicated by a horizontal bar beneath the jumping clones. With probe Aj1R3, a λ linking clone (All-4) was produced from which probe All-BB was isolated. This probe detected an adjacent telomeric 200-kb fragment in *Not*I-digested DNA from a somatic cell hybrid containing normal chromosome 9 (#9) (bottom right). A 350-kb *Not*I fragment was detected in a hybrid cell line containing the $9q^+$ chromosome of t(6;9) ANLL patient DK (#9q+). In hamster DNA (a23), no hybridizing fragments were detected. Abbreviations: B, *Bam*HI; Bs, *Bss*HII; E, *Eco*RI; N, *Not*I; Sa, *Sac*II; Sf, *Sfi*I.

To confirm that the six identical phages contained the other end of the 300-kb *Not*I fragment, a single-copy *Rsa*I probe of 280 bp (Aj1R3) was isolated from the novel 1.6-kb *Not*I-*Bam*HI insert. On a Southern blot of HeLa cell DNA, probe Aj1R3 recognized a 300-kb *Not*I fragment comigrating with the *Not*I fragment hybridizing to the *Bgl*II-*Bam*HI c-*abl* probe (Fig. 1). On a panel of somatic cell hybrids, probe Aj1R3 proved to be located on chromosome 9 (data not shown).

Subsequently, probe Aj1R3 was used to screen a genomic λEMBL3 library to generate a probe in the 3′-adjacent *Not*I fragment. Phage All-4 contained an insert of 16 kb, of which 7.5 kb were located downstream of the *Not*I site (Fig. 1). From the latter sequences, a 600-bp *Bgl*II-*Bam*HI single-copy probe (All*BB*, Fig. 1) was used to screen *Not*I-digested DNA from somatic cell hybrids containing the segregated chromosomes of t(6;9) ANLL patient DK (50). A 200-kb fragment was present in the cell line containing normal chromosome 9 (MA5C), while in the cell line containing chromosome $9q^+$ (MA4C), a 350-kb fragment was detected (Fig. 1). This observation suggested that the t(6;9) ANLL chromosome 9 breakpoint was located in the adjacent 200-kb *Not*I fragment. Also in the DNA of t(6;9) ANLL patient JK, an aberrant 350-kb *Not*I fragment was found in addition to the normal 200-kb *Not*I fragment (Fig. 2A).

DNAs of other t(6;9) ANLL patients were analyzed on long-range mapping gels, with several infrequently cutting restriction endonucleases and the same probe. As most infrequently cutting restriction endonucleases have the CpG dinucleotide in their recognition sequence, sites are often clustered in the G+C-rich HTF (*Hpa*II tiny fragments) islands (4). Mapping of phage All-4 showed that a *Bss*HII and an *Sfi*I site were located within 2 kb of the *Not*I site, 300 kb downstream of c-*abl* (Fig. 1 and 3). The downstream *Sfi*I and *Bss*HII fragments measured 6 and 30 kb, respectively (Fig. 3) and did not contain the t(6;9) breakpoint (data not shown). As a consequence, the breakpoint maps 3′ of the 30-kb *Bss*HII fragment.

Chromosome walking was used to generate probes at the telomeric side of the 30-kb *Bss*HII site (see Materials and Methods). Subsequent screening of a cosmid library with probe All*BB* and of a λEMBL3 library with probe All*C2HSH3* produced cosmid All*C2* and λ-All*F3*, respectively (Fig. 3). From the phage clone, a single-copy probe

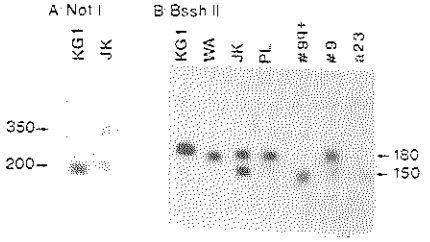


FIG. 2. Detection of the t(6;9) breakpoints on chromosome 9 by FIGE. High-molecular-weight DNA was prepared in agarose blocks, digested with infrequently cutting restriction endonucleases, and separated by FIGE. (A) *Not*I-digested DNA from t(6;9) ANLL patient JK and a normal control cell line, KG-1. The Southern blot of this gel was hybridized with All8B (Fig. 1). (B) *Bss*HII-digested DNA from cell line KG-1 and patient WA (normal control DNAs), JK and PL [t(6;9) ANLL patients], the hybrid cell lines MA-4C and MA-5C [containing chromosome $9q^+$ (#9q+) and normal chromosome 9 (#9) of t(6;9) ANLL patient DK] and the hamster fusion partner a23. The Southern blot of this gel was hybridized with All3EH (Fig. 3). Numbers at the sides of the figure indicate the size of the fragments (in kilobases).

was generated mapping 3′ of the 30-kb *Bss*HII fragment (All3EH). This probe was used to screen *Bss*HII digests of DNA derived from bone marrow cells of three different t(6;9) ANLL patients, DK, JK, and PL (Fig. 2). In control DNA (KG-1 cells and normal subject WA), a 180-kb *Bss*HII fragment was present; in DNA from all three patients, an additional aberrant fragment of 150 kb could be seen. The bone marrow sample of patient PL contained only 30% blast cells, which accounts for the low intensity of the 150-kb fragment on the autoradiogram. For patient DK, the DNAs of the somatic cell hybrids were used together with the DNA of the hamster cell fusion partner a23 as a control. It is noteworthy that probe All3EH cross-hybridized with a hamster fragment of 150 kb after the blot was washed under stringent conditions (0.3× SSC, 65°C). Detection of the 150-kb abnormal *Bss*HII band in all three patients indicated that the t(6;9) 9q34 breakpoints are located within the 180-kb *Bss*HII fragment. The fact that the aberrant fragments were identical in size in three different patients suggested that the chromosomal breakpoints are clustered on both chromosomes 9 and 6.

Precise localization of the t(6;9) breakpoint on chromosome 9q34. The next two steps of chromosome walking produced phages All4 and All6 (Fig. 3). As tested on Southern blots containing DNA from the somatic cell hybrids of patient DK, the single-copy probe All4EP (a 200-bp *Eco*RI-*Pst*I fragment, Fig. 3) was still present on the $9q^+$ chromosome, while probe All6E.3 (a 300-bp *Eco*RI fragment 3′ of All4EP, Fig. 3) was translocated to chromosome $6p^-$ (Fig. 4). As a consequence, the breakpoint on chromosome 9 in patient DK is located between All4EP and All6E.3.

The two probes were subsequently hybridized to *Bam*HI and *Bgl*II digests of DNA isolated from the bone marrow cells of four t(6;9) ANLL patients (DK, JK, PL, and PM) and acute undifferentiated leukemia (AUL) patient SE (Fig. 4). In all five patients an abnormal band was detected with either probe All4EP or All6E.3. In patient PL, two aberrant bands were detected with probe All6E.3, suggesting that the breakpoint is located within this 300-bp *Eco*RI fragment. Mapping data indicated that the breakpoints on chromosome 9 in these five patients were located within a region of 8 kb, 360 kb downstream of c-*abl*.

Further chromosome walking produced three more phage clones (All7, All8, and All9) encompassing 35 kb of DNA downstream of the 8-kb region containing the breakpoints (Fig. 3).

Chromosome 9 breakpoint maps within a gene. Probe All3EH possibly contained exon sequences, as it cross-hybridized to hamster DNA (Fig. 3). On Northern blots, the probe recognized a 7.5-kb transcript in total RNA isolated from the cell line K562 (data not shown). A K562 cDNA library was screened with All3EH, and 10 cDNA clones were isolated. The largest cDNA, hX8, was mapped with several restriction endonucleases (Fig. 5). In order to obtain full-length cDNA of the 7.5-kb transcript, the K562 and a human testis cDNA library were screened with the insert of hX8. Only the testis library produced clones containing new sequences. Further walking on the human testis cDNA library produced the overlapping cDNA clones shown in Fig. 5. The telomeric side of the cDNA clone hX8, derived from K562, was not present in any of the overlapping clones from the testis library. A probe from this telomeric part failed to detect the 7.5-kb mRNA on Northern blots containing K562 RNA (data not shown). Most likely this part of hX8

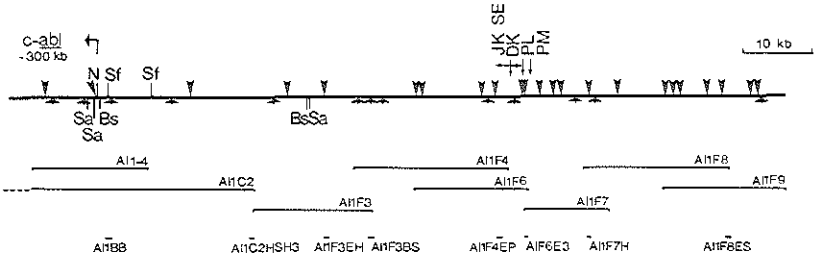


FIG. 3. Restriction map of the region on chromosome 9q34 containing the t(6;9) breakpoints. The top line shows a restriction map of the centromeric 95 kb of the 200-kb *Not*I fragment containing the t(6;9) breakpoints. This map was deduced from cosmid clone All C2 and phage clones All F3 to All F9, indicated below the map. Phage All-4 is also shown in Fig. 1. Small horizontal bars beneath the clones represent single-copy probes (All BB, All C2HSH3, All F3EH, All F3BS, All F4EP, All F6E3, All F7H, and All F8ES). As a reference point, the localization of the breakpoints in DNAs from patients DK, PM, PL, JK, and SE is indicated by small arrows above the restriction map. Large arrowheads above the line indicate *Eco*RI sites; large arrows below the line indicate *Bam*HI sites. Other sites: Bs, *Bss*HII; N, *Not*I; Sa, *Sac*II; Sf, *Sfi*I.

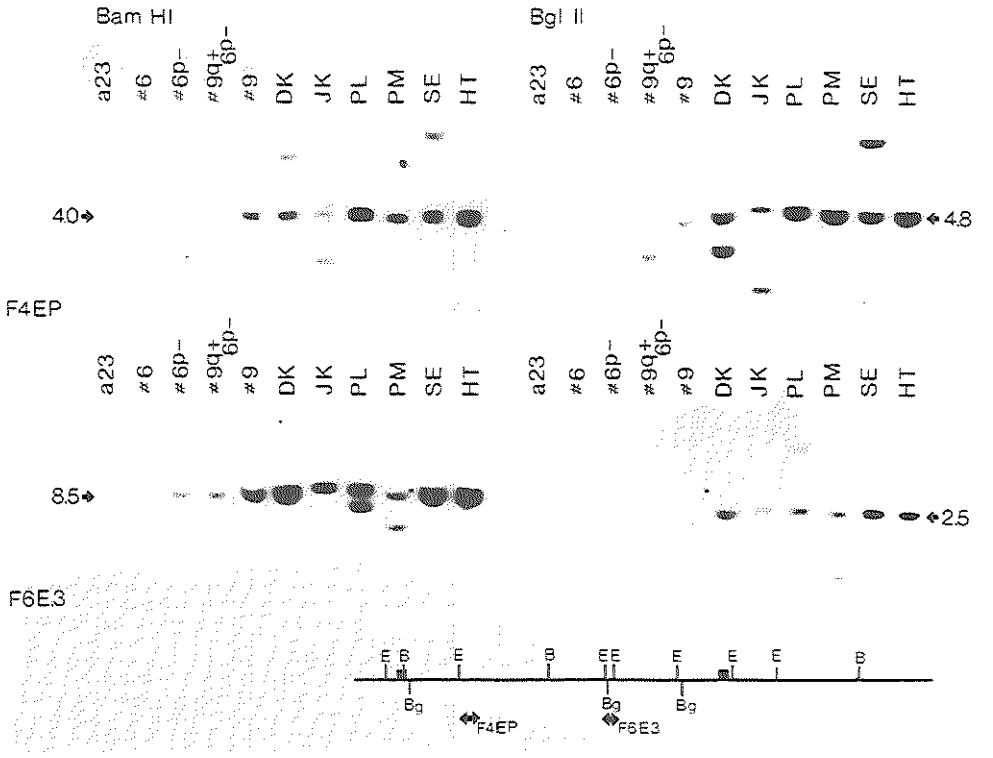


FIG. 4. Breakpoints in four t(6;9) ANLL patients and AUL patient SE map in a region of 8 kb on chromosome 9, icb-9. Probes All F4EP and All F6E.3 were hybridized to Southern blots containing DNAs digested with *Bam*HI and *Bgl*II. Lane 1, Hamster DNA (a23). Lanes 2 to 6, Hybrid cell line DNA containing chromosomes 6, $6p^-$, $9q^+/6p^-$, and 9, respectively, of t(6;9) ANLL patient DK. Lanes 6 to 10, Bone marrow DNA from patients DK, JK, PL, PM, and SE, respectively. Lane 11, Normal DNA from human thymus (HT). The position of the probes is indicated in Fig. 3 and in a more detailed map at the bottom of this figure. The sizes of the normal *Bam*HI and *Bgl*II fragments are indicated (in kilobases). Ba, *Bam*HI; Bg, *Bgl*II; E, *Eco*RI.

represents intron sequences, indicated in Fig. 5 by a dashed line.

The inserts of the five overlapping cDNA clones were used to probe Southern blots containing different digests (*Bam*HI, *Hin*dIII, *Eco*RI, and *Xba*I) of the genomic phage clones (Fig. 5). It is interesting that hXT37 appeared to hybridize to genomic fragments only centromeric and telomeric of the region encompassing the t(6;9) breakpoints and not to fragments within this 8-kb region (Fig. 5). Therefore, we can conclude that the breakpoints are located within the gene coding for the 7.5-kb transcript and seem to be clustered in one intron of this gene. We propose to call this gene Cain, *can*, and to call the intron containing the breakpoints on chromosome 9 icb-9.

The overlapping cDNA clones spanned 7.5 kb of cDNA sequences and can represent the entire mRNA (Fig. 5). Approximately 3.3 kb of the transcript was encoded at the centromeric side of icb-9 and 4.2 kb at the telomeric side. As shown in Fig. 5, the cDNA clones hybridized over an area of 65 kb of genomic DNA. Genomic sequences hybridizing to the most telomeric 1.8 kb of cDNA hXT65 have not been isolated. Therefore, the *can* gene measures more than 65 kb.

To determine the direction of transcription, the centromeric 500-bp *Eco*RI-*Pst*I fragment of cDNA clone hXT65 was inserted into phage M13 in both orientations. A ^{32}P-labeled second strand was synthesized by using the universal M13 primer. Only the ^{32}P-labeled strand 3′-*Eco*RI–5′-*Pst*I hybridized to *can* mRNA in K562 total RNA (Fig. 5B). Thus, the *can* gene is transcribed from 5′ centromeric to 3′ telomeric on the chromosome.

cDNA probes mapping both 5′ (hXT23) and 3′ (hXT54) of icb-9 were hybridized to total RNA isolated from the epithelial cell line HeLa, several hematopoietic cell lines (Daudi, Hl60, KG1, and K562), and bone marrow cells from t(6;9) ANLL patients DK and JK, AUL patient SE, and a normal control (BM) (Fig. 6). In all samples, a 7.5-kb transcript was detected in variable amounts; less abundant larger transcripts of unknown origin were also present. In addition to the 7.5-kb *can* mRNA, a new 5.5-kb transcript was detected in the patient samples with the 3′ cDNA probe only. The 5′ cDNA probe hXT23 did not detect the 5.5-kb mRNA (data not shown). This indicated that the 3′ part of the *can* gene is expressed in leukemic cells of t(6;9) ANLL patients as part of an aberrantly sized transcript. The expression of both the

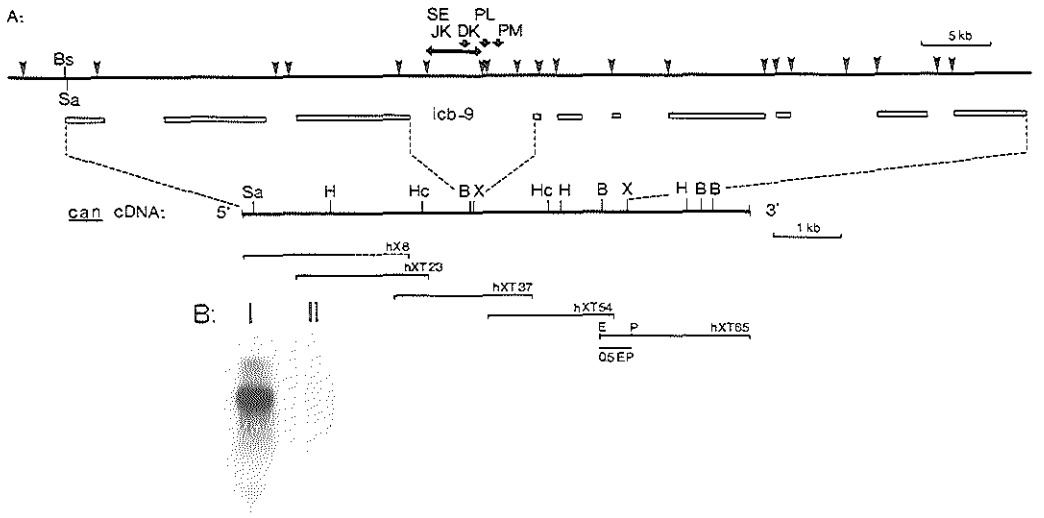


FIG. 5. Restriction map of the *can* cDNA. (A) Restriction map of the cDNA was deduced from several cDNA clones isolated from a K562 and a human testis cDNA library. Overlapping cDNA clones are shown below the restriction map. The transcriptional orientation is indicated. Clone hX8 contains intron sequences at the 3' part, which are indicated by a dashed line. The top line represents the genomic restriction map of the *can* gene. Open bars indicate genomic fragments that hybridized to cDNA sequences. Hybridizing fragments are delimited by restriction sites for *Bam*HI, *Hind*III, *Eco*RI, and *Xba*I (not shown in the genomic restriction map). The position of a number of "key" restriction sites (*Sac*II, *Bam*HI, and *Xba*I) in the cDNA is indicated on the genomic map by dashed lines. Genomic sequences corresponding to the 3' 1.8 kb of the *can* cDNA have not yet been isolated. The exact position of the breakpoints in DNA from patients DK, PL, and PM was determined (◆); the breakpoints in DNA from patients SE and JK were positioned in a 3.4-kb *Eco*RI fragment (⟷). The stretch of DNA containing the breakpoints did not hybridize to *can* cDNA and was named icb-9 (for intron containing the breakpoints on chromosome 9). The scales of the genomic and the cDNA map are indicated in kilobases. ▼ , *Eco*RI sites. Other restriction sites: E, *Eco*RI; B, *Bam*HI; H, *Hind*III; Hc, *Hinc*II; P, *Pst*I; Sa, *Sac*II; X, *Xba*I. (B) Northern blots of K562 RNA were hybridized with strand-specific probes (a 0.5-kb *Eco*RI-*Pst*I fragment [0.5EP in panel A]), derived from cDNA clone hXT65. Lane I, ^{32}P-labeled 3'-*Eco*RI-5'-*Pst*I strand. Lane II, ^{32}P-labeled 5'-*Eco*RI-3'-*Pst*I strand.

7.5- and the 5.5-kb mRNA was much more abundant in AUL patient SE than in the two t(6;9) ANLL patients.

Molecular cloning of the t(6;9) chromosomal breakpoints. To characterize sequences from chromosome 6p23 that are fused to chromosome 9, *can* probes were used to identify and clone a chimeric 6:9 genomic DNA fragment. Probe A11F4EP detected a 3.4-kb normal *Eco*RI fragment in the hybrid cell line MA-5C (normal chromosome 9 of patient DK) and a 4.0-kb *Eco*RI fragment in the hybrid cell line MA-4C (9q$^+$ chromosome of DK; data not shown). The chimeric 4.0-kb fragment was gel purified and cloned into λgt10 (Fig. 7C). Mapping analysis showed that the fragment contained 2 kb of sequences not derived from chromosome 9. These sequences were expected to be derived from chromosome 6p23. Unfortunately, it was impossible to generate a probe from the new DNA because it consisted of repetitive sequences. Even hybridization in the presence of excess amounts of single-stranded, sonicated human competitor DNA failed to generate a specific signal on Southern blots.

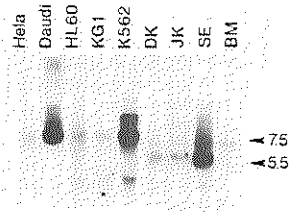


FIG. 6. *can* expression in cell lines and bone marrow cells of leukemia patients. Northern blot containing total RNA from cell lines HeLa (15 μg), Daudi (25 μg), HL60 (15 μg), KG-1 (15 μg), and K562 (15 μg) and from bone marrow cells of t(6;9) ANLL patients DK (15 μg) and JK (2 μg), AUL patient SE (15 μg), and normal control subject BM (25 μg). The filter was hybridized with cDNA clone hXT54, located 3' of icb-9. The lane containing JK RNA was exposed six times longer than the other lanes. The normal 7.5-kb *can* transcript and the 5.5-kb aberrant transcript are indicated by arrows.

To circumvent this problem, DNA from patient PM was used to construct a λEMBL3 phage library, which was screened with probes on either side of the translocation breakpoint of DNA from patient PM on chromosome 9 (A11F6E.3 and A11F7EBD4). One phage clone, M1-8, was detected that consisted in part of known chromosome 9 sequences linked to 3.7 kb of DNA with an unknown restriction pattern (Fig. 7B). A 200-bp *Dde*I fragment (M1D3), derived from the new sequences, was almost free of

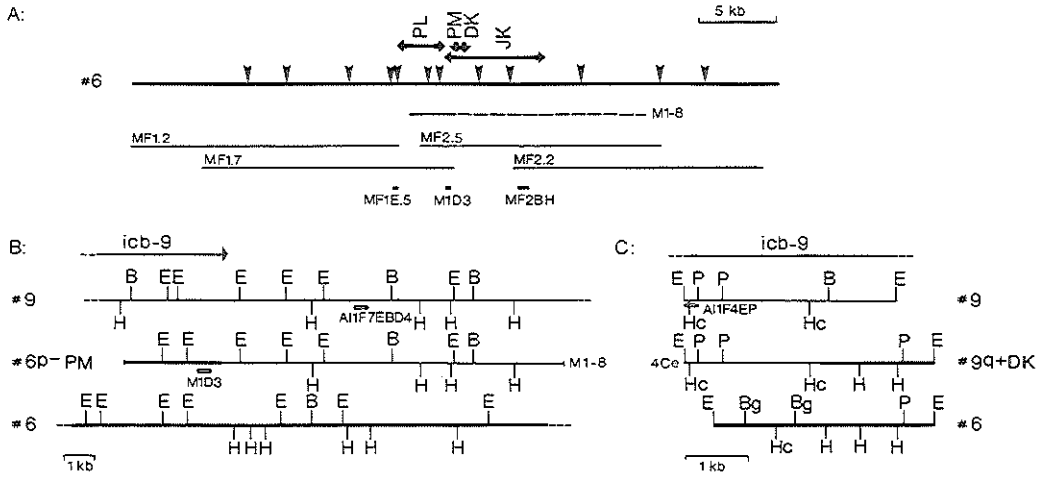


FIG. 7. Restriction map of the region on chromosome 6 in which the t(6;9) breakpoints occur. (A) The top line represents an *Eco*RI map of the region of chromosome 6 in which the t(6;9) breakpoints are located. The genomic clones are indicated below the map (MF1.2, MF1.7, MF2.2, MF2.5). The position is indicated of chromosomal breakpoints of four t(6;9) ANLL patients: PM and DK (◆) and PL and JK (⟷). M1-8 is a chimeric genomic phage clone derived from patient PM, consisting of 3.7 kb of chromosome 6 and 12 kb of chromosome 9 sequences (dashed line). Beneath the phage clones are shown the positions of three genomic probes, MF1E.5, M1D3, and MF2BH. ▼ . *Eco*RI sites. (B) The thin top line represents the genomic map of chromosome 9 sequences involved in the (6;9) translocation of patient PM. The arrow above delineates icb-9. Probe Al1F7EBD4 was used to isolate the chimeric phage clone M1-8 (the middle line), which is derived from the $6p^-$ chromosome of PM. From the chromosome 6 sequences (thick line), an almost-single-copy probe, M1D3, was isolated. The bottom line represents the chromosome 6 sequences involved in the t(6;9) of patient PM. B, *Bam*HI; E, *Eco*RI; H, *Hin*dIII. (C) The second line from the top represents a 3.4-kb *Eco*RI fragment from icb-9 involved in the (6;9) translocation of patient DK. The line below represents the 4-kb chimeric $9q^+$ fragment of patient DK, which was cloned into λgt10 and identified with probe Al1F4EP. The right-hand side of this fragment (thick line) represents chromosome 6 sequences. The bottom line shows the 3.5-kb *Eco*RI fragment of chromosome 6p23 that contains the chromosomal breakpoints from patients PM and DK. B, *Bam*HI; Bg, *Bgl*II; E, *Eco*RI; H, *Hin*dIII; Hc, *Hin*cII; P, *Pst*I.

repetitive DNA and could be used to screen the same genomic λEMBL3 library in the presence of human competitor DNA. Two phages were isolated and analyzed, MF1.7 and MF2.5, spanning 30 kb of DNA. The *Eco*RI restriction map of the putative chromosome 6p23 region is shown in Fig. 7A. Alignment of this restriction map with that of the chimeric phage M1-8 revealed that the 3.5-kb *Eco*RI fragment of phage clone MF2.5 was the normal counterpart of the chimeric 1.7-kb *Eco*RI fragment in the DNA from patient PM (Fig. 7B).

Further restriction analysis of the chimeric *Eco*RI fragments of DNAs from patients DK and PM provided (indirect) evidence that the new sequences were derived from chromosome 6. A 180-bp *Hin*fI fragment, isolated from the new sequences of the 4-kb breakpoint fragment of patient DK (4Ce, Fig. 7C), contained low-copy-number repetitive sequences. This fragment specifically hybridized to the 3.5-kb *Eco*RI fragment of phage MF2.5, in which the breakpoint of the DNA from patient PM was located. Fine mapping of clone 4Ce, the 3.5-kb *Eco*RI fragment of clone MF2.5, and the 1.7-kb chimeric *Eco*RI fragment of clone M1-8 showed that the breakpoints in the DNA from two t(6;9) patients mapped only 0.6 kb apart on chromosome 6.

On either side of probe M1D3 (chromosome 6; Fig. 7A), single-copy sequences were isolated; MF1E.5 (a 500-bp *Eco*RI fragment) and MF2BH (an 800-bp *Bgl*II-*Hin*dIII fragment) (Fig. 7A). These probes were used to screen the DNA of t(6;9) ANLL patients for breakpoints on chromosome 6 and to isolate phage clones MF1.2 and MF2.2 (Fig. 7A).

DNA from bone marrow cells of t(6;9) ANLL patients DK, JK, PL, and PM, AUL patient SE, and the somatic cell hybrids of patient DK was digested with *Eco*RV, *Hin*dIII, and *Bam*HI. Only *Eco*RV and *Hin*dIII digests are shown in Fig. 8. Both probes MF1E.5 and MF2BH hybridized to the chromosome 6 hybrid cell line MA-13C and not to the chromosome 9 hybrid MA-5C, again indicating their chromosome 6 origin. Moreover, MF1E.5 did hybridize to the cell line containing chromosome $6p^-$, while MF2BH did not hybridize to this cell line. The chromosome 6 breakpoint of patient DK must be situated between these two probes, located 8 kb apart. MF1E.5 recognized an aberrant fragment of 7 kb in *Eco*RV-digested DNA from patient PL and aberrant fragments of 8 and 9.5 kb in *Hin*dIII-digested DNA from patients PL and PM. MF2BH hybridized to large aberrant fragments of more than 15 kb in *Eco*RV-digested DNA from patients PM, DK, and JK. Only in DNA from AUL patient SE did the probes fail to detect a breakpoint in either *Eco*RV (Fig. 8), *Hin*dIII, or *Bam*HI digests (results not shown). Therefore, the breakpoints in the DNA from four out of five patients with a breakpoint within icb-9 were clustered in a relatively small region of 12 kb on chromosome 6.

DISCUSSION

Through a combination of long-range mapping, chromosome jumping, and chromosome walking techniques, it was possible to clone the chromosomal breakpoints of the (6;9)

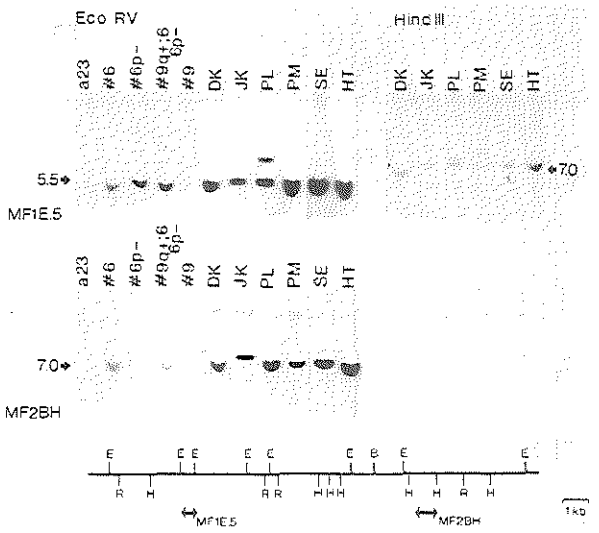


FIG. 8. Mapping of chromosome 6 breakpoints of four t(6;9) ANLL patients. Probes MF1E.5 and MF2BH were hybridized to Southern blots containing DNAs digested with *Eco*RV and *Hin*dIII. Lane 1, Hamster DNA (a23). Lanes 2 to 6, Hybrid cell line DNA containing chromosomes 6, $6p^-$, $9q^+/6p^-$, 6, and 9, respectively, of t(6;9) ANLL patient DK. Lanes 6 to 10, Bone marrow DNA from patients DK, JK, PL, PM, and SE, respectively. Lane 11, Normal DNA from human thymus (HT). Of the *Hin*dIII digest, only lanes 6 to 11 are shown, hybridized with probe MF1E.5. The position of the probes is indicated in Fig. 7A and in a more detailed map at the bottom of this figure. B, *Bam*HI; E, *Eco*RI; H, *Hin*dIII; R, *Eco*RV.

translocation in ANLL patients. Conditions for successful isolation of the t(6;9) breakpoints were favorable: Westbrook et al. (52) showed that the t(6;9) breakpoint on chromosome 9q34 is situated at the telomeric side of c-*abl*. The orientation of the c-*abl* gene on chromosome 9 is 5′ centromeric, 3′ telomeric (21). Therefore, c-*abl* exon 1A could immediately be used to screen a *Not*I jumping library, since it is situated at the 5′ end of a 300-kb *Not*I fragment. Although these conditions speeded up the procedure, the tactics are applicable to the identification of any other chromosome translocation.

***can* gene on chromosome 9.** In all four cases of t(6;9) ANLL and in an AUL patient with an apparently normal karyotype, the breakpoint on chromosome 9 maps in a limited stretch of 8 kb of DNA, 360 kb downstream of c-*abl*. This 8-kb region appeared to be an intron of a large gene (>65 kb), *can*. In control RNA, a 7.5-kb *can* transcript was present. However, in DNA from two t(6;9) ANLL patients (DK and JK), the 3′ part of the gene, translocated to the $6p^-$ chromosome, expressed an aberrant 5.5-kb mRNA.

Two different observations point to the $6p^-$ chromosome as the translocation product carrying an activated oncogene: (i) a t(6;9) ANLL patient was described who lost the $9q^+$ chromosome during the course of the disease (1), and (ii) only cDNA probes from the *can* gene that translocate to the $6p^-$ chromosome recognized an aberrant transcript in patient RNA. Alignment of the cDNA clones with the genomic restriction map showed that 3.3 kb of the 7.5-kb transcript was encoded 5′ of icb-9 and 4.2 kb was encoded 3′. There are several possibilities to generate the 5.5-kb abnormal transcript on the $6p^-$ chromosome. (i) A cryptic promoter on either chromosome 9 or chromosome 6 can be activated by the translocation. As a result, the mRNA will consist of 5′ intron sequences fused to the 3′ 4.2 kb of the *can* mRNA. (ii) Alternatively, the translocation may fuse the 3′ part of *can* to the 5′ part of a gene on chromosome 6. The promoter of the gene on chromosome 6 will drive transcription of this chimera. Cloning and sequencing of the full-length 5.5-kb cDNA will distinguish between these alternatives. It will be interesting to see whether the presumptive protein-coding capacity of the *can* mRNA is affected by the translocation.

At present, nothing is known about the function of the *can* gene. Transcription is probably not restricted to myeloid or even hematopoietic cells, because it is expressed in K562, KG-1, and HL60 cells (myeloid), Daudi cells (lymphoid), HeLa cells (epithelial), and normal adult testis. Expression in K562 cells is relatively high. This may be because the gene is located on a large amplicon that derives from the Philadelphia translocation in K562 cells. The amplicon stretches from upstream of the λ light-chain locus to at least 500 kb downstream of c-*abl*, past the *can* gene (43; M. von Lindern, unpublished results).

Sequence analysis of the cDNA clones may reveal putative functional domains of the protein, indicating its possible function in the cell. cDNA probes of the *can* gene hybridized to hamster DNA under stringent conditions (0.3× SSC, 65°C). This indicated that the gene is evolutionarily conserved between humans and hamsters.

The 5′ end of the *can* cDNA is located in an HTF island, between a *Bss*HII and a *Sac*II site. These G+C-rich areas are strongly associated with promoter regions (4). We assume that this HTF island contains the *can* promoter region. The direction of transcription supports this hypothesis. Expression studies, such as chloramphenicol acetyltransfer-

ase assays, will be needed to prove the assumption and to determine which sequences are necessary for *can* transcription.

Chromosome region 9q34 was reported to be involved in several translocations. They include t(9;14)(q34;q11) (7), t(9;17)(q34;q23) (28), t(2;9)(q33;q34) (27), and t(7;9)(q36;q34) (19), all occurring in lymphoblastic lymphoma. Most likely, the immunoglobulin κ light chain on chromosome 2q33, the β TCR on chromosome 7q36, and the α TCR on chromosome 14q11 are involved in the translocations. Reynolds et al. (42) cloned the breakpoint of the t(7;9)(q36;q34), starting from the β TCR gene. Part of a gene (*tcl-3*) was isolated, flanking the translocation breakpoint on chromosome 9. However, the restriction map and the size of *tcl-3* mRNA are different from those of the *can* gene. At present, involvement of the *can* gene in other types of leukemia, such as t(2;9), t(7;9), t(9;14), and t(9;17) in lymphoblastic lymphoma, has not been excluded. Investigations into this possibility are in progress.

Breakpoint area on chromosome 6p23. In a library representing the genome of leukemic cells of patient PM, only a single phage clone was found to contain chromosome 9 sequences fused to putative chromosome 6 sequences. Although there is no direct proof that chromosome 6 sequences were cloned, indirect evidence seems sufficient. (i) The putative chromosome 6 sequences of the $9q^+$ fragment of DNA from patient DK (clone 4Ce) were present in the same genomic 3.5-kb *Eco*RI fragment as the sequences of the $6p^-$ chimeric phage clone M1-8, derived from patient PM. In fact, the breakpoints of DNAs from patients DK and PM mapped only 0.6 kb apart on chromosome 6. (ii) Although the panel of hybrids used was not fully informative for the chromosomal localization of the isolated sequences, probes from the putative chromosome 6 region hybridized to somatic cell lines containing chromosome 6 (MA-13A) and not to hybrids containing chromosome 9 (MA-5C). Probe MF1E.5 hybridized to the cell line containing the $6p^-$ chromosome (MA-3B); probe MF2BH did not. This implies that the chromosome 6 breakpoint is located between these two probes, located 8 kb apart. (iii) Probes MF1E.5 and MF2BH (chromosome 6; Fig. 7A) recognized a breakpoint in genomic DNA not only from patients PM and DK, but also from the two other t(6;9) ANLL patients; PL and JK.

Mapping of the breakpoints on chromosome 6 showed that they were scattered over a stretch of 12 kb. Therefore, the localization of the breakpoints of different t(6;9) ANLL patients is amazingly precise on both chromosome 9 and 6. The clustering of breakpoints on chromosome 6 strongly suggests a functional necessity for this configuration.

In AUL patient SE, no breakpoint was detected. Therefore, this patient may be an exception, but the absence of a breakpoint on chromosome 6 cannot be excluded. The breakpoint area on chromosome 6 is extremely rich in repetitive sequences, which prevents the isolation of other informative probes. This inconvenience, in combination with the localization of the mapped restriction sites in this area, makes *Eco*RV the only suitable restriction enzyme for detection of chromosome 6 breakpoints at present. Aberrant fragments can comigrate with the normal *Eco*RV fragment on Southern blots. Alternatively, the breakpoint of patient SE may map outside the 12-kb region limited by the *Eco*RV sites. Long-range mapping analysis of the DNA from patient SE should detect a breakpoint on chromosome 6, unless the *can* gene of this patient is fused to sequences located elsewhere in the genome. On the transcriptional level, patient SE was different from t(6;9) ANLL patients DK and JK. A similarly sized abnormal *can* transcript (5.5 kb) was detected, but it was expressed at a considerably higher level (Fig. 6). In fact there are no data available specifying patient SE as a t(6;9) ANLL patient. Cytogenetic analysis failed to show a t(6;9), and the leukemic cells were phenotypically different (AUL instead of AML M2 or M4).

The segments of chromosome 6 and 9 that are translocated in t(6;9) ANLL are rather small. The $6p^-$ and $9q^+$ chromosomes are not easily distinguishable from their normal counterparts in a cytogenetic analysis. Depending on the quality of chromosome preparations, the translocation can be missed. Probes described in this paper may contribute to a relatively easy and reliable diagnosis of t(6;9) ANLL, comparable to the molecular detection of *bcr* breakpoints in chronic myelogenous leukemia. However, more patients have to be analyzed to define the general applicability of these probes.

Clustering of the breakpoints on chromosomes 9 and 6 in combination with disruption of the *can* transcription unit argue for a causative role of the (6;9) translocation in this type of leukemia. Further analysis is needed to define the exact nature of the *can* disruption and its effect on the biological activity of the putative CAN protein.

ACKNOWLEDGMENTS

We thank D. Bootsma for continuous support and discussion, J. Abels, H. Adriaansen, A. Hagemeijer, B. Löwenberg, C. Mecucci, F. Mitelman, H. van den Berghe, and P. W. Wijermans for patient material, M. Fornerod, M. Jaegle, D. Meijer, and S. van Baal for support and critical reading of the manuscript, and R. Boucke and M. Kuit for preparation of the manuscript and photographic work.

This work is supported by the Dutch Cancer Foundation (Koningin Wilhelmina Fonds).

LITERATURE CITED

1. **Adriaansen, H. J., J. J. M. Van Dongen, H. Hooijkaas, K. Hählen, M. B. Van't Veer, B. Löwenberg, and A. Hagemeijer.** 1988. Translocation (6:9) may be associated with a specific TdT-positive immunological phenotype in ANLL. Leukemia 2:136–140.

1a. **Adriaansen, H. J., A. H. M. Geurts van Kessel, J. H. F. M. Wydenes-de Bresser, E. van Drunen-Schoenmaker, and J. J. M. van Dongen.** 1990. Expression of the myeloid differentiation antigen CD33 depends on the presence of human chromosome 19 in human-mouse hybrids. Ann. Hum. Genet. 54:115–119.

2. **Auffray, C., and F. Rougeon.** 1980. Purification of mouse immunoglobin heavy chain messenger RNA from total myeloma tumor RNA. Eur. J. Biochem. 107:303–314.

3. **Ben-Neriah, Y., G. Q. Daley, A. M. Mes-Masson, O. N. Witte, and D. Baltimore.** 1986. The chronic myelogenous leukemia-specific p210 protein is the product of the *bcr/abl* hybrid gene. Science 232:212–214.

4. **Bird, A. P.** 1986. CpG-rich islands and the function of DNA methylation. Nature (London) 321:209–213.

5. **Chen, S. J., Z. Chen, J. D. Grausz, J. Hillion, L. d'Auriol, G. Flandrin, C. J. Larsen, and R. Berger.** 1988. Molecular cloning of a 5′ segment of the genomic *phl* gene defines a new breakpoint cluster region (bcr2) in Philadelphia-positive acute leukemias. Leukemia 2:634–641.

6. **Chirgwin, J. M., A. E. Przybyla, J. MacDonald, and W. J. Rutter.** 1979. Isolation of biologically active ribonucleic acid from sources enriched in ribonuclease. Biochemistry 18:5294–5299.

7. **Clare, N., D. Boldt, G. Messerschmidt, P. Zeltzer, K. Hansen, and L. Manhoff.** 1986. Lymphocyte malignancy and chromosome 14: structural aberrations involving band q11. Blood 67:704–709.

8. **Cleary, M. L., J. D. Mellentin, J. Spies, and S. D. Smith.** 1988. Chromosomal translocation involving the beta T-cell receptor gene in acute leukemia. J. Exp. Med. 167:682–687.

9. **Croce, C. M., M. Isobe, A. Palumbo, J. Puck, J. Ming, D. Tweardy, J. Erikson, M. Davis, and G. Rivera.** 1985. Gene for alpha chain of human T-cell receptor: location on chromosome 14 region involved in T-cell neoplasms. Science **227**:1044–1047.
10. **Croce, C. M., M. Shander, J. Martinis, L. Cicurel, G. C. D'Ancona, T. W. Dolby, and H. Koprowski.** 1979. Chromosomal location of the genes for human immunoglobin heavy chains. Proc. Natl. Acad. Sci. USA **76**:3416–3419.
11. **de Klein, A., T. Van Agthoven, C. Groffen, N. Heisterkamp, J. Groffen, and G. Grosveld.** 1986. Molecular analysis of both translocation products of a Philadelphia-positive CML patient. Nucleic Acids Res. **14**:7071–7081.
12. **Domen, J., M. von Lindern, A. Hermans, M. Breuer, G. Grosveld, and A. Berns.** 1987. Comparison of the human and mouse *pim*-1 cDNA's: nucleotide sequence and immunological identification of the in vitro synthesized *pim*-1 protein. Oncogene Res. **1**:103–112.
13. **Erikson, J., J. Martinis, and C. M. Croce.** 1981. Assignment of the genes for human λ immunoglobulin chains to chromosome 22. Nature (London) **294**:173–175.
14. **Feinberg, A. P., and B. Vogelstein.** 1982. A technique for radiolabeling DNA restriction endonuclease fragments to high specific acitivity. Anal. Biochem. **132**:6–13.
15. **Fourney, R. M., J. Miyakoshi, R. S. Day III, and M. C. Paterson.** 1988. Northern blotting: efficient RNA staining and transfer. Focus **10**:5–6.
16. **Frischauf, A., H. Lehrach, A. Poustka, and N. Murray.** 1983. Lambda replacement vectors carrying polylinker sequences. J. Mol. Biol. **170**:827–842.
17. **Grosveld, G., T. Verwoerd, T. Van Agthoven, A. de Klein, K. Ramachandran, N. Heisterkamp, K. Stam, and J. Groffen.** 1986. The chronic myelocytic cell line K562 contains a breakpoint in *bcr* and produces a chimeric *bcr/c-abl* transcript. Mol. Cell. Biol. **6**:607–616.
18. **Hariharan, I. K., A. W. Harris, M. Crawford, H. Abud, E. Webb, S. Cory, and J. M. Adams.** 1989. A *bcr*-v-*abl* oncogene induces lymphomas in transgenic mice. Mol. Cell. Biol. **9**: 2798–2805.
19. **Hecht, F., R. Morgan, R. M. Gemmill, B. Hecht, and S. D. Smith.** 1985. Translocations in T-cell leukemia and lymphoma. N. Engl. J. Med. **313**:758–759.
20. **Heim, S., U. Kristoffersson, N. Mandahl, F. Mitelman, A. N. Bekassy, S. Garwiesz, and T. Wiebe.** 1986. High resolution banding analysis of the reciprocal translocation t(6:9) in acute nonlymphocytic leukemia. Cancer Genet. Cytogenet. **22**:195–201.
21. **Heisterkamp, N., J. R. Stephenson, J. Groffen, P. F. Hansen, A. de Klein, C. R. Bartram, and G. Grosveld.** 1983. Localisation of the c-*abl* oncogene adjacent to a translocation breakpoint in chronic myelocytic leukaemia. Nature (London) **306**:239–242.
22. **Heisterkamp, N., K. Stam, J. Groffen, A. de Klein, and G. Grosveld.** 1985. Structural organisation of the *bcr* gene and its role in the Ph′ translocation. Nature (London) **315**:758–761.
23. **Hermans, A., N. Heisterkamp, M. von Lindern, S. van Baal, D. Meijer, D. van der Plas, L. M. Wiedeman, J. Groffen, D. Bootsma, and G. Grosveld.** 1987. Unique fusion of *bcr* and c-*abl* genes in Philadelphia chromosome-positive acute lymphoblastic leukemia. Cell **51**:33–40.
24. **Hermans, A., L. Selleri, J. Gow, L. Wiedemann, and G. Grosveld.** 1989. Molecular analysis of the Philadelphia translocation in myelogenous and acute lymphoblastic leukemia. Cancer Cells **7**:21–26.
25. **Isobe, M., J. Erikson, B. S. Emanuel, P. C. Nowell, and C. M. Croce.** 1985. Location of gene for β subunit of human T-cell receptor at band 7q35, a region prone to rearrangements in T-cells. Science **228**:580–582.
26. **Jeffreys, A., and R. A. Flavell.** 1977. A physical map of the DNA regions flanking the rabbit β-globin gene. Cell **12**:429–439.
27. **Kaneko, Y., G. Frizzera, N. Maseki, M. Sakurai, Y. Komada, M. Sakurai, Y. Hiyoshi, H. Nakadate, and T. Takeda.** 1988. A novel translocation, t(9:17)(q34:q23), in aggressive childhood lymphoblastic lymphoma. Leukemia **2**:745–748.
28. **Kaneko, Y., D. Variakojis, L. Kluskens, and J. D. Rowley.** 1982. Lymphoblastic lymphoma: cytogenetic, pathologic and immunologic studies. Int. J. Cancer **30**:273–279.
29. **Kirsch, I. R., C. C. Morton, K. Nakahara, and P. Leder.** 1982. Human immunoglobulin heavy-chain genes map to a region of translocation in malignant B lymphocytes. Science **216**:301–303.
30. **Konopka, J. B., S. M. Watanabe, and O. N. Witte.** 1984. An alteration of the human c-*abl* protein in K562 leukemia cells unmasks associated tyrosine kinase activity. Cell **37**:1035–1042.
31. **Kurzrock, R., J. K. Gutterman, and M. Talpaz.** 1988. The molecular genetics of Philadelphia chromosome-positive leukemias. N. Engl. J. Med. **13**:990–998.
32. **Malcolm, S., P. Barton, C. Murphy, M. A. Ferguson-Smith, D. L. Bentley, and T. H. Rabbitts.** 1982. Localization of human immunoglobulin K light-chain variable-region genes to the short arm of chromosome 2 by *in situ* hybridization. Proc. Natl. Acad. Sci. USA **79**:4957–4961.
33. **Maniatis, T., E. F. Fritsch, and J. Sambrook.** 1982. Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
34. **McLaughlin, J., E. Chianese, and O. N. Witte.** 1987. *In vitro* transformation of immature hematopoietic cells by the P210 *bcr/abl* oncogene product of the Philadelphia chromosome. Proc. Natl. Acad. Sci. USA **84**:6558–6562.
35. **McLaughlin, J., E. Chianese, and O. N. Witte.** 1989. Alternative forms of the *bcr-abl* oncogene have quantitatively different potencies for stimulation of immature lymphoid cells. Mol. Cell. Biol. **9**:1866–1874.
36. **Mellentin, J. D., S. D. Smith, and M. L. Cleary.** 1989. *lyl-1*, a novel gene altered by chromosomal translocation in T-cell leukemia, codes for a protein with a helix-loop-helix DNA binding motif. Cell **58**:77–83.
37. **Nagarajan, L., E. Louie, Y. Tsujimoto, A. Ar-Rushdi, K. Huebner, and C. M. Croce.** 1986. Localization of the human *pim* oncogene (PIM) to a region of chromosome 6 involved in translocations in acute leukemias. Proc. Natl. Acad. Sci. USA **83**:2556–2560.
38. **Nishikura, K., A. Ar-Rushdi, J. Erikson, R. Watt, G. Rovera, and C. M. Croce.** 1983. Differential expression of the normal and of the translocated human c-*myc* oncogene in B cells. Proc. Natl. Acad. Sci. USA **80**:4822–4826.
39. **Pearson, M. G., J. W. Vardiman, M. M. LeBeau, J. D. Rowley, S. Schwartz, S. L. Kerman, M. M. Cohen, E. W. Fleischman, and E. L. Prigogina.** 1985. Increased numbers of marrow basophils may be associated with a t(6:9) in ANLL. Am. J. Hematol. **18**:393–403.
40. **Poustka, A., T. M. Pohl, D. P. Barlow, A. M. Frischauf, and H. Lehrach.** 1987. Construction and use of human chromosome jumping libraries from *Not*I-digested DNA. Nature (London) **325**:353–355.
41. **Rabbitts, T. H., M. P. Lefranc, M. A. Stinson, J. E. Sims, J. Schroder, M. Steimetz, N. L. Spurr, E. Solomon, and P. N. Goodfellow.** 1985. The chromosomal location of T-cell receptor genes and a T-cell rearranging gene: possible correlation with specific translocation in human T-cell leukaemia. EMBO J. **4**:1461–1465.
42. **Reynolds, T. C., S. D. Smith, and J. Sklar.** 1987. Analysis of DNA surrounding the breakpoints of chromosomal translocations involving the β T-cell receptor gene in human lymphoblastic neoplasms. Cell **50**:107–117.
43. **Selden, J. R., B. S. Emanuel, E. Wang, L. Cannizzaro, A. Palumbo, J. Erikson, P. C. Nowell, G. Rovera, and C. M. Croce.** 1983. Amplified Cλ and c-*abl* genes are on the same marker chromosome in K562 leukemia cells. Proc. Natl. Acad. Sci. USA **80**:7289–7292.
44. **Shtivelman, E., B. Lifshitz, R. P. Gale, and E. Canaani.** 1985. Fused transcript of *abl* and *bcr* genes in chronic myelogenous leukaemia. Nature (London) **315**:550–554.
45. **Taub, R., I. Kirsch, C. Morton, G. Lenoir, D. Swan, S. Tronick, S. Aaronson, and P. Leder.** 1982. Translocation of the c-*myc* gene into immunoglobulin heavy-chain locus in human Burkitt lymphoma and murine plasmacytoma cells. Proc. Natl. Acad. Sci. USA **79**:7837–7841.
46. **Tsujimoto, Y., and C. M. Croce.** 1986. Analysis of the structure,

transcripts, and protein products of *bcl-2*, the gene involved in human follicular lymphoma. Proc. Natl. Acad. Sci. USA 83: 5214–5218.

47. **Tsujimoto, Y., L. R. Finger, J. Yunis, P. C. Nowell, and C. M. Croce.** 1984. Cloning of the chromosomal breakpoint of neoplastic B-cells with the t(14;18) chromosome translocation. Science 226:1097–1099.
48. **Tsujimoto, Y., J. Yunis, L. Onorato-Showe, J. Erikson, P. C. Nowell, and C. M. Croce.** 1984. Molecular cloning of the chromosomal breakpoint of B-cell lymphomas and leukemias with the t(11:14) chromosome translocation. Science 224:1403–1406.
49. **Van der Feltz, M. J. M., M. K. K. Shivji, G. Grosveld, and L. M. Wiedemann.** 1988. Characterization of the translocation breakpoint in a patient with Philadelphia-positive, *bcr*-negative acute lymphoblastic leukaemia. Oncogene 3:215–219.
50. **von Lindern, M., T. Van Agthoven, A. Hagemeijer, H. Adriaansen, and G. Grosveld.** 1989. The human *pim-1* gene is not directly activated by the translocation (6:9) in acute nonlymphocytic leukemia. Oncogene 4:75–79.
51. **Walker, L. C., T. S. Ganesan, S. Dhut, B. Gibbons, T. A. Lisker, J. Rothbard, and B. D. Young.** 1987. Novel chimaeric protein expressed in Philadelphia-positive acute lymphoblastic leukaemia. Nature (London) 329:8511–853.
52. **Westbrook, C. A., M. M. LeBeau, M. O. Diaz, J. Groffen, and J. D. Rowley.** 1985. Chromosomal localization and characterization of c-*abl* in the t(6;9) of acute nonlymphocytic leukemia. Proc. Natl. Acad. Sci. USA 82:8742–8746.

CHAPTER 2.4

The translocation (6;9), associated with a specific subtype of acute myeloid leukemia, results in the fusion of two genes, *dek* and *can*, and the expression of a chimeric, leukemia specific *dek-can* mRNA

Marieke von Lindern, Maarten Fornerod, Sjozèf van Baal, Martine Jaegle, Ton de Wit, Arjan Buijs and Gerard Grosveld

dept. Cell Biology and Genetics, Erasmus University, Rotterdam, The Netherlands.

Molecular and Cellular Biology,12,*in press.*

ABSTRACT

The translocation (6;9) is associated with a specific subtype of acute myeloid leukemia (AML). Previously, it was found that breakpoints on chromosome 9 are clustered in one of the introns of a large gene named *Cain (can)*. cDNA probes derived from the 3' part of *can* detect an aberrant, leukemia specific 5.5 kb transcript in bone marrow cells from t(6;9) AML patients. cDNA cloning of this mRNA revealed that it is a fusion of sequences encoded on chromosome 6 and 3' *can*. A novel gene on chromosome 6 was isolated which was named *dek*. In *dek* the t(6;9) breakpoints also occur in one intron. As a result the *dek-can* fusion gene, present in t(6;9) AML, encodes an invariable *dek-can* transcript. Sequence analysis of the *dek-can* cDNA showed that *dek* and *can* are merged without disruption of the original open reading frames and therefore the fusion mRNA encodes a chimeric DEK-CAN protein of 165 kDa. The predicted DEK and CAN proteins have molecular masses of 43 and 220 kDa respectively. Sequence comparison with the EMBL database failed to show consistent homology with any known protein sequences.

INTRODUCTION

Defined karyotypic aberrations are associated with specific subtypes of leukemia. Detailed molecular characterization of these aberrations may identify genes involved in leukemogenesis and in the precise regulation of proliferation and differentiation in the hematopoietic system. Translocations are the best studied chromosomal abnormalities. As the result of a translocation the function or activity of oncogenes located at or near the translocation breakpoint is altered. In myeloid leukemia three translocation breakpoints have been cloned and analyzed at the molecular level.

The two best studied, t(9;22) in chronic myeloid leukemia (CML) (27, 43), and t(15;17) in acute promyelocytic leukemia (APL) (2, 8, 12), result in the formation of chimeric genes that encode fusion proteins. In CML this is a BCR-ABL protein, that has an enhanced tyrosine kinase activity (34, 49), directly responsible for its *in vivo* tumorigenic potential (14, 25). In APL a PML-RARα fusion protein is found, that represents an altered transcription factor (16, 33).

The third translocation is the t(6;9)(p23;q34) found in a specific subtype of AML (1, 39, 41). This leukemia is characterized by a poor prognosis, affects young adults and is FAB classified mostly as M2 or M4 and rarely as M1. A region on chromosome 9 was cloned and analyzed, situated 360 kb downstream of the *c-abl* gene. It was found that breakpoints are clustered in a region of 8 kb in 5 patients, 4 with t(6;9) AML and 1 with acute undifferentiated leukemia (AUL) (47). Through cDNA cloning this region could be identified as one of the introns of a large gene (>100 kb) encoding a 7 kb transcript. This intron was named *icb*-9: intron containing the breakpoints on chromosome 9, and is situated in the middle of a gene named *Cain*, (*can*). The 3′ part of *can* is translocated to the 6p- chromosome and only 3′ *can* probes detect an additional, leukemia specific 5.5 kb transcript in bone marrow cells from t(6;9) AML patients. No additional transcripts were detected with 5′ *can* probes. The breakpoint region on chromosome 6p23 was isolated from a genomic λEMBL3 library constructed of bone marrow DNA of one of the t(6;9) patients. An area of 40 kb of chromosome 6 DNA was cloned in overlapping phages. Southern blot analysis showed that breakpoints of t(6;9) AML patients are clustered in a relatively small region of 12 kb.

This article reports the cloning of a cDNA representing the 5.5 kb aberrant transcript specific for t(6;9) AML; the isolation of a novel gene, *dek*, on chromosome 6p23; and the sequence analysis of both *can* and *dek* cDNAs.

MATERIALS AND METHODS

Northern blotting. Patient material and cell lines used were described previously (47, 48). RNA of mouse tissue was isolated from BCBA mice. RNA was isolated by either the guanidinium isothiocyanate (11) or LiCl-method (5). Total RNA was electrophoresed and blotted as described by Fourney et al. (20). Equal amounts of rRNA were loaded: before loading the samples on a denaturing gel, 5% of each sample was loaded on a nondenaturing agarose gel to estimate the amount of rRNA and to adapt the sample quantity if necessary. Northern blots were hybridized in 10% dextran (40). Northern blots of mouse tissues were hybridized with human probes at 3xSSC 65°C, and filters were washed with 1xSSC at 65°C for *dek* probes and with 0.3xSSC at 65°C for *can* probes. Probes were labelled by the method of Feinberg and Vogelstein (19).

cDNA cloning. 100 μg total RNA from patient DK was heat-denatured and annealed to 10 μg of a 21-mer, 5′GAAGGACTAGGTGCACCATGT3′, at 55°C. First strand synthesis was done with Avian RT (26). Second strand synthesis was done according to the RNaseH method (24). The DNA was blunt ended with T4 polymerase and treated with EcoRI

methylase (Sigma). EcoRI linkers were ligated onto the cDNA with T4 ligase and RNA ligase (40) and after EcoRI digestion, the cDNA was size selected on a Sephacryl S-1000 column. cDNA larger than 1 kb was ligated into the EcoRI site of λgt10 (31). Phage DNA was packaged using packaging extracts (GIGA gold;Stratagene). $19x10^6$ pfu were generated of which only 10% contained an insert, estimated by analysis of randomly picked phages. Most likely the other 90% contained linker sequences. The human testis cDNA library in λgt11 was purchased from Clonetech (Palo Alto, Calif.). The CML0 λEMBL3 library was described by Hermans et al. (28)

Sequence determination and analysis. Restriction fragments of cDNA clones were subcloned in M13. Overlapping cDNA sequences were determined on both strands by dideoxy sequencing (42). Initially M13 primers were used; when no suitable restriction sites were present a primer was generated based on the already available cDNA sequence. To establish intron-exon borders, genomic fragments containing the exon of interest were subcloned into M13 and a primer near the putative intron-exon border was generated to prime the sequence reaction. Sequences were analyzed with the computerprogram Microgenie and the EMBL database was used to search for homologous sequences at both the nucleotide and amino acid levels.

The nucleotide sequence data reported in this paper will appear in the EMBL, Genbank and DDBJ Nucleotide Sequence Databases under the accession numbers X64228 (*can*) and X64229 (*dek*).

RACE cloning 3' end *can*. 30 μg total RNA of bone marrow cells from AUL patient SE was heat-denatured and first strand cDNA was synthesized with Avian RT, using 100 pmol of the 35-mer 5'GTCGCGAATTCGTCGACGCGTTTTTTTTTTTTTTT3' as a primer (21, 28). Excess of primer was removed by isopropanol precipitation. 1/100 of the cDNA reaction was amplified using Taq polymerase (Perkin Elmer Cetus) and the primers 5'GTCGCGAATTCGTCGACGCG3' and 5'GCCTTTGGATCCCTGGGACCAACCGC3'. The latter primer is located 180 bp upstream the polyA signal in *can* cDNA. The amplified fragment of 230 bp was sequenced using a protocol for direct sequencing of fragments produced by an asymmetric PCR (32).

RESULTS

Analysis of the *can* gene and transcript. As reported previously, a nearly full length *can* cDNA was isolated in the overlapping cDNA clones hXT23, hXT37, hXT54 and hXT65 (47). Originally cDNA clone hX8 was thought to represent the 5' part of the *can* mRNA. However, a more detailed mapping analysis showed that the 5' part of hX8 does not belong to the *can* gene and is in fact not even located on chromosome 9 (data not shown). Therefore hX8 must be considered a cloning artifact. As several (11) independent cDNA clones appeared to have 5' ends mapping close to the 5' end of hXT23, we assumed that the 5' end of the latter clone maps in the vicinity of the *can* mRNA cap site.

The genomic map of *can*, reported previously, extended over 70 kb but did not include the 3' part of the *can* gene. Therefore, cDNA clone hXT65 was used to screen a genomic λEMBL3 library and many hybridizing phages were isolated. Clones Al1F10.6, Al1F10.2, Al1F10.8 and Al1F10.12 were selected since they covered the largest stretch of DNA, and were analyzed in more detail. As indicated in Fig. 1A, a gap is still present between

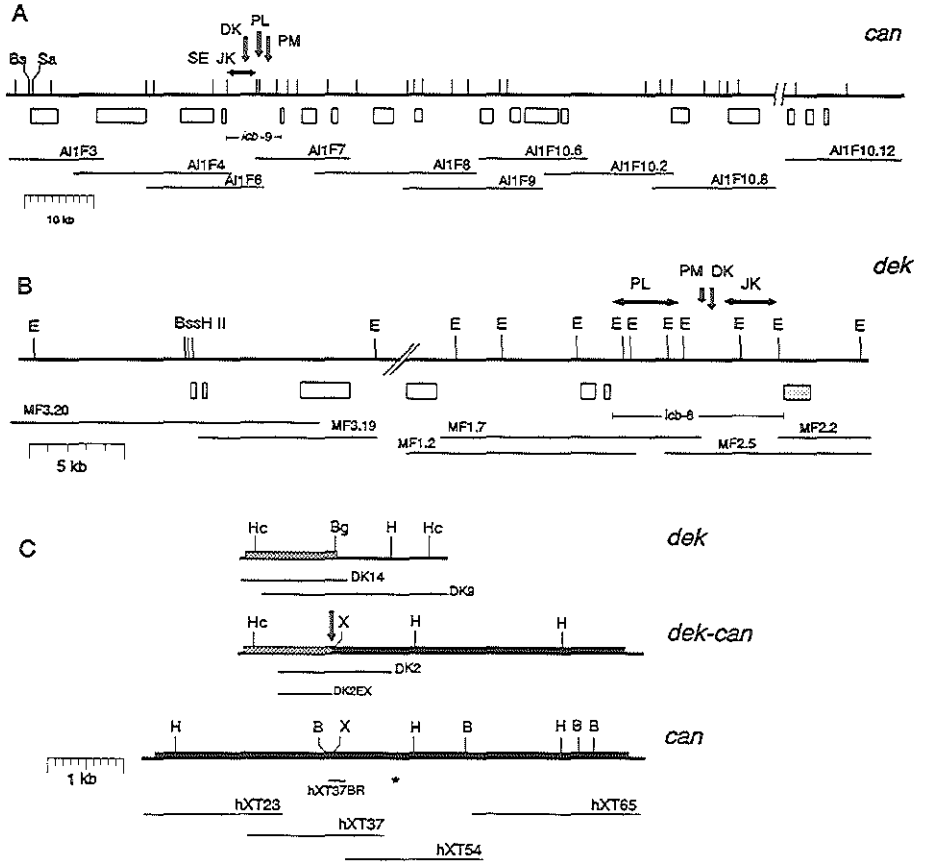


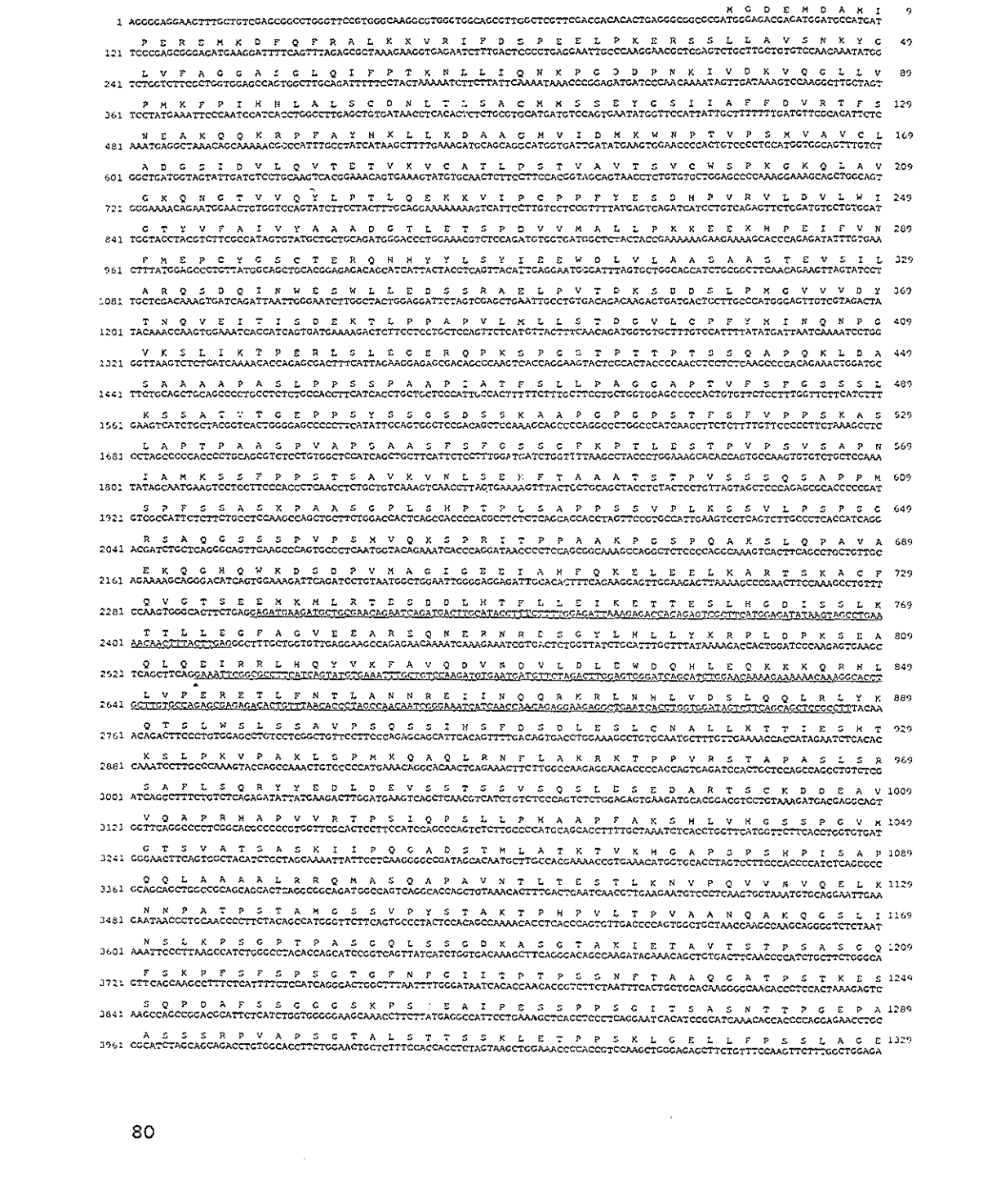
Figure 1. Restriction maps of the *can* and *dek* gene, *dek*, *dek-can* and *can* cDNAs. (A) The top line represents a genomic map of the *can* gene. Vertical lines represent *Eco*RI sites. Open boxes represent restriction enzyme fragments hybridizing to *can* cDNA probes. The position of the breakpoints of 4 t(6;9) AML patients, DK, PM, JK, PL and AUL patient SE is indicated with arrows. They are all located in *icb*-9. Below the map isolated genomic phages are depicted. AI1F3, AI1F4, AI1F6, AI1F7, AI1F8 and AI1F9 were reported previously. AI1F10.6, AI1F10.2, AI1F10.8 and AL1F10.12 were isolated with cDNA clone hXT65 as probe. The gap between AI1F10.8 and AI1F10.12 is at maximum 40 kb. The scale is indicated in kb. (B) The second line represents a genomic map of *dek*. Open boxes indicate restriction fragments hybridizing to cDNA probes, these fragments were delimited by various restriction enzyme sites not shown in this map. Stippled boxes are mapped exons. The positions of the breakpoints of t(6;9) AML patients DK, PM, PL and JK are indicated with

AI1F10.8 and AI1F10.12. The total amount of *can* sequences cloned in phages is 130 kb. Since the gene is located on a *Bss*HII fragment of 170 kb (47) and no *Bss*HII site is present in AI1F10.12, it was deduced from fig. 1A that the gap between AI1F10.8 and AI1F10.12 can range between 1 and 40 kb. Restriction enzyme fragments that contain exons were determined by hybridization of Southern blots containing *Eco*RI, *Bam*HI and *Hin*dIII digests of the phages with *can* cDNA clone hXT65 (Fig. 1C).

The overlapping *can* cDNA clones were sequenced and appeared to contain a large open reading frame (ORF) of 6270 nt, encoding a putative protein of 220 kDa (Fig. 2). This ORF starts in clone hXT23 and ends in clone hXT65. A 700 bp *Hin*dIII-*Pst*I fragment of phage AI1F3, in which the *Bss*HII site is located (AI1F3E4HP) was also sequenced. Fig. 3A shows that the sequence of AI1F3E4HP is colinear with hXT23 up to its 5′ end. As other cDNA clones have 5′ ends mapping near the 5′ end of hXT23, it has to be tested whether this region contains *can* promoter sequences. At the 5′ end the *can* cDNA contains ATG start codons at positions 95-97, 107-109 and 115-117. The sequence around the codon at position 95 is concordant with the consensus sequence postulated by Kozak (35), which suggests that this methionine is probably the start of the CAN protein. The first stopcodon in this frame is at position 6365. The sequence of cDNA clone hXT65 ends immediately 3′ of what appeared to be a variant polyadenylation signal: ATTAAA (nt 6562-6567). As no polyA tail was present in this clone, the 3′ end of the *can* transcript was amplified with the RACE protocol (21) from a position 180 bp 5′ of the polyA signal to the polyA tail. The sequence of this amplified fragment showed that the polyA tail starts 16 nt downstream of the ATTAAA signal. The 3′ end of hXT65 hybridized to genomic λEMBL3 phage AI1F10.12. Sequence analysis showed that the 3′ exon of *can* is present in this phage. Its sequence is colinear with the cDNA sequence down to the polyA tail (Fig. 3B).

Since previous mapping data localized the t(6;9) breakpoints in the middle of cDNA clone hXT37 (Fig. 1C), the breakpoints must occur within the *can* ORF. To exactly localize the position of *icb*-9 within the ORF, genomic clones

arrows, they are all located in *icb*-6. MF3.20, MF3.19, MF1.2, MF1.7, MF2.5 and MF2.2 are λEMBL3 phages from which the map has been deduced. The gap between MF3.19 and MF1.20 is estimated to be only a few kb. The scale is indicated in kb. (C) The bottom part shows restriction maps of the cDNAs of *dek, dek-can* and *can* are depicted. A scale for the cDNA maps is indicated. Arrows indicate the position of the breakpoints. The open reading frame of *dek* is indicated by a hatched bar, the open reading frame of *can* by a solid bar on top of the lines that indicate the cDNAs. The chimaeric cDNA DK2 has been isolated from a primed cDNA library that was made with a primer indicated by an asterix. This library was screened with probe hXT37BR. DK9 and DK14 are *dek* cDNAs isolated with probe DK2EX from a λgt11 cDNA library derived from human testis RNA. B:*Bam*HI, Bg:*Bgl*II, E:*Eco*RI, H:*Hin*dIII, Hc:*Hin*cII, X:*Xba*I.



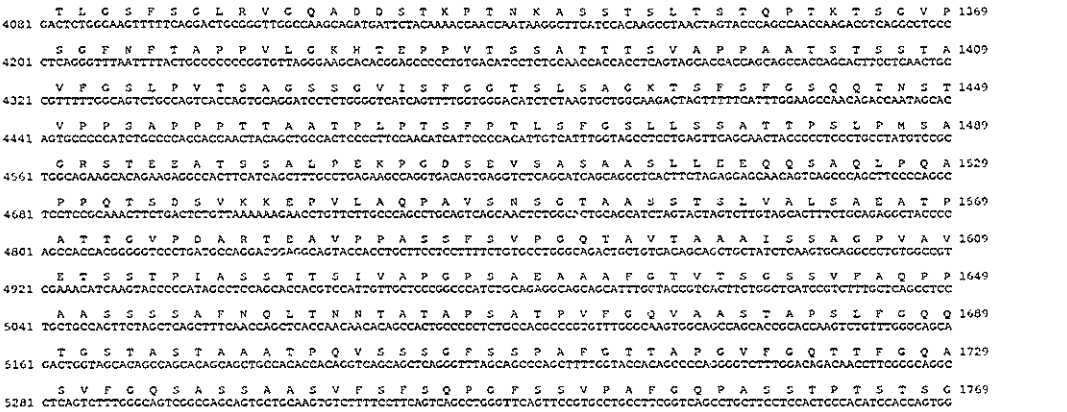


Figure 2. cDNA sequence and putative amino acid sequence of the CAN protein. The position of the t(6;9) breakpoint is indicated with a black triangle (nt 2530-31). The putative leucine zipper (boxA; Figure 6A) and amphiphatic helices (boxB; Figure 6A) are underlined.

were used to sequence the intron-exon borders delineating this intron. This showed that the translocation breakpoints occur between codons 812 and 813 (nt 2530-2531) in the ORF of the *can* mRNA (Fig. 2, 3C). Because of the translocation, 4053 nucleotides of the *can* cDNA are encoded on the 6p- chromosome. As a consequence, cDNA probes located within these 4053 nt recognize a specific 5.5 kb transcript in bone marrow cells from t(6;9) AML patients (47).

Cloning the *dek-can* hybrid cDNA. To resolve the identity of the t(6;9) AML, specific 5.5 kb mRNA a primed cDNA library was constructed using

total RNA of bone marrow from t(6;9) patient DK and a 21 nt primer mapping 800 bp downstream of the translocation breakpoint in the *can* cDNA (Fig. 1C). Part of the library ($2x10^6$ pfu) was screened with a 360 bp *Bam*HI-*Rsa*I (hXT37BR) fragment, indicated in Fig. 1C. Two clones (DK1 and DK2; 1.3 and 1.5 kb respectively) were isolated and characterized. They appeared to be colinear with *can* cDNA from the primer at the 3' end of the cDNA clones, exactly up to the 5' end of the exon flanking *icb*-9 at its 3' side. Upstream of this point both clones are identical but deviate completely from the *can* sequence. To determine the chromosomal origin of these sequences, a 5' DK2 fragment (probe DK2EX, a 700 bp *Eco*RI-*Xba*I fragment, Fig. 1C) was hybridized to a Southern blot containing DNA of a hybrid cell panel with the segregated translocation chromosomes involved in the t(6;9) (48). The probe hybridized to DNA of cell lines containing chromosome 6 and 6p- (results not shown). The same probe was hybridized to a Northern blot containing RNA of Hela cells, hematopoietic cell lines (Daudi, HL60, KG1, K562) and bone marrow cells of t(6;9) AML patient DK

A: 5'*can*	
atgcgaggtcaactgcgcgccgctcgggctcAGGGGAGGAAGTTTGCTGTCGAGCGGCCT	A11F3E4HP
AGGGGAGGAAGTTTGCTGTCGAGCGGCCT	hXT23
GGGTTCCGTGGGCAAGGCGTGGGAGGCAGCGTTGGCTCGTTC	A11F3E4HP
GGGTTCCGTGGGCAAGGCGTGGGAGGCAGCGTTGGCTCGTTC	hXT23
B: 3'*can*	
ATATTTCATGTTGGGTTTTCCCTCCCACTATTAAACAGTCTGTTTCCGTACagaacgtat	A11F10.12St
ATATTTCATGTTGGGTTTTCCCTCCCACTATTAAACA	hXT65
ATATTTCATGTTGGGTTTTCCCTCCCACTATTAAACAGTCTGTTTCCGTACAAAAAAAAA	hXE230
C: *icb*-9	
GGATCCCAAGAGTGAAGCTCAGCTTCAGgtaggagatctatgtaaatctgtttaaaagatt	A11F4BE
GGATCCCAAGAGTGAAGCTCAGCTTCAGGAAATTCGGCGCCTTCATCAGTATGTGAAATTT	hXT37
tatttccctttaatttttttcttatagGAAATTCGGCGCCTTCATCAGTATGTGAAATTT	A11F7E2
D: 5'*dek*	
cgccgcctcccagaacctcttcgtgccctcggcgtgccaGGCCCGCGGCGGCCGAAATCC	MF3.20Hc1.4
GGCCCGCGGCGGCCGAAATCC	DK14
GCGgtaagggctgcttcgggacgcggagggcccggtcgcccctgccgcggccgacgggg	MF3.20Hc1.4
GCGGTTCACAGCATGTCCGCCTCGGCCCCTGCTGCGGAGGGGGAGGGAACCCCCACCCAG	DK14
E: 3'*dek*	
TTATAAACCTTGTCAATAAAAATAAATCTAAATCactggtgttttaagtcacttgcattt	MF2.5E4.5
TTATAAACCTTGTCAATAAAAATAAATCTAAATCAAAAAAAAAAA	DK9
gatatcttataggtgtatatagc	MF2.5E4.5
F: *icb*-6	
CACAATGAAACAGATTTGCAAAAAGgtaattagacaaatgtttagattatttgctttgct	MF1.7E2.7
CACAATGAAACAGATTTGCAAAAAGGTCTATGAAAATTATCCTACTTATGATTTAACTGA	DK9
tatttttccttttcactatacatagGTCTATGAAAATTATCCTACTTATGATTTAACTGA	MF2.5E4.5

and AUL patient SE. This revealed the presence of a 2.7 kb transcript in all lanes and an additional 5.5 kb transcript in the t(6;9) AML bone marrow sample (Fig. 4). This 5.5 kb transcript is identical in size to the aberrant transcript detected with 3'*can* probes in this patient (47). These results proved that sequences encoded by a gene on chromosome 6 are present in the t(6;9) AML specific 5.5 kb transcript, which is thus identified as a chimeric mRNA.

It is noteworthy that in AUL patient SE no aberrant transcript was detected by the chromosome 6 probe, while hybridization with 3' *can* probes clearly detected an aberrant mRNA of 5.5 kb in the Northern blot of this patient (47). This result is in agreement with the observation that bone marrow cells from AUL patient SE contained a breakpoint in *icb*-9, but failed to show a breakpoint in band p23 on chromosome 6 that contains breakpoints of t(6;9) patients.

Cloning *dek* cDNA. To isolate a cDNA of the normal 2.7 kb transcript encoded on chromosome 6, a λgt11 cDNA library, derived from human testis RNA, was screened with probe DK2EX. In total 24 clones were isolated and analyzed by restriction enzyme mapping. Two overlapping cDNA clones were identified (DK9 and DK14) that contained 2.7 kb of contiguous sequences, probably representing the full length transcript. The full length *dek* cDNA clone of 2699 nt was sequenced (Fig. 5). It contains an ORF of 1125 nt encoding a putative protein of 375 amino acids with an estimated molecular mass of 43 kDa, followed by a large 3' untranslated region (UTR) of 1541

Figure 3. Comparison of genomic and cDNA sequences of *dek* and *can*. (A) A genomic 700 bp *Hin*dIII-*Pst*I fragment isolated from phage AI1F3 (AI1F3E4HP) contains the most 5' sequences of cDNA clone hXT23. Presumably, these sequences belong to the first exon of *can*. (B) A genomic 900 bp *Stu*I fragment isolated from phage AI1F10.12 (AI1F10.12St) contains the most 3'exon of *can*. cDNA clone hXT65 ends immediately 3' of the polyA signal, therefore a 230 bp cDNA fragment was generated with the RACE protocol (hXE230) that contains the polyA signal and polyA tail. The polyA signal is underlined. (C) The intron-exon borders flanking *icb*-9 were determined to prove that the t(6;9) breakpoints are located in a single intron. A 1.2 kb *Bam*HI-*Eco*RI fragment from phage AI1F4 (AI1F4BE) contains part of the exon flanking *icb*-9 upstream. In fact, the *Bam*HI site is located in the exon. A 2 kb *Eco*RI fragment from phage AI1F7 (AI1F7E2) contains the exon flanking *icb*-9 downstream. The sequence of cDNA hXT37 shows both exons joined together. (D) A 1.4 kb *Hin*cII fragment from the genomic *dek* phage MF3.20 (MF3.20Hc1.4) contains the most 5' *dek* cDNA sequences of clone DK14 in a 240 bp *Bss*HII fragment. (E) *Dek* has a large 3' exon located in a 4.5 kb *Eco*RI fragment of phage MF2.5 (MF2.5E4.5). The sequence encompassing the 3' end of this exon is shown together with the 3' end of cDNA clone DK9. (F) The intron-exon borders flanking *icb*-6 were determined. Downstream of *icb*-6 only one large 3' exon is present in a 4.5 kb *Eco*RI fragment of phage MF2.5, the 5' border of which is shown. The exon flanking *icb*-6 upstream is located in a 2.7 kb *Eco*RI fragment of phage MF1.7 (MF1.7E2.7). The sequence of cDNA clone DK9 shows both exons joined together.

nt. The ATG codon is located at position 34-36 which matches the Kozak consensus sequence (35). The predicted amino acid sequence of the ORF is shown in Fig. 5. The 3′ end of the cDNA sequence contains two AATAAA polyA addition sites next to each other, at position 2682 and 2688, followed by a polyA stretch at position 2702.

As reported previously (47), the genomic area of chromosome 6p23, containing the breakpoints of 4 t(6;9) AML patients, was isolated in 4 overlapping λEMBL3 clones. The t(6;9) breakpoints appeared to map in a stretch of 12 kb in the middle of this region. The cDNA clones DK9 and DK14 were hybridized to Southern blots containing DNA of these genomic phages, digested with several restriction enzymes. Genomic fragments hybridizing to cDNA probes are present at either side of, but not within, the region that contains the chromosome 6 translocation breakpoints, but not within (Fig. 1B). Initially, the 12 kb breakpoint region was mapped by Southern blot analysis of genomic DNA and was delineated by two *Eco*RV sites. More precise mapping of the cloned chromosomal DNA reduced the size of the breakpoint region on chromosome 6p23 to an intron of 9 kb. Similar to what was done for the *can* gene, the intron-exon borders of the *dek* exons were sequenced, that flank *icb*-6. This showed that the translocation breakpoints occur between codons 349 and 350, almost at the C-terminus of the *dek* ORF (Fig. 5, 3F). The intron containing the breakpoints on chromosome 6 was termed *icb*-6. Fusion of *dek* and *can* via the introns *icb*-9 and *icb*-6 results in transcription of a chimeric mRNA in which the ORFs of *dek* and *can* are merged without disruption of their original reading frames. As a result, the DEK-CAN protein has a predicted molecular mass of 165 kDa.

Analysis of the *dek* gene. Hybridization of *dek* cDNA probes to the previously isolated λEMBL3 phages occurred 3′ of *icb*-6 to only one apparently

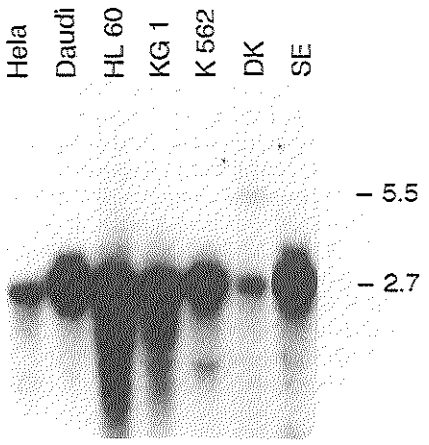


Figure 4. Northern blot containing 20 μg total RNA extracted from the cell lines Hela, Daudi, HL60, KG1 and K562 and from bone marrow cells of t(6;9) AML patient DK and AUL patient SE. The size of the mRNAs hybridizing to the *dek* cDNA probe DK14 is indicated in kilobases.

continuous stretch of genomic DNA. As the restriction maps of this genomic DNA and the 3′ *dek* cDNA are colinear, it is likely that only a single 1.6 kb 3′ exon is present downstream of *icb*-6. To further substantiate this point, the 3′ end of this exon was sequenced and was found to be colinear with *dek* cDNA down to the polyA tail (Fig. 3E). However, a tiny intron in this exon can only be excluded by complete sequencing of the exon.

Hybridization of 5′ *dek* cDNA probes to genomic phages indicated, that the 5′ end of the *dek* gene was not contained within them. Therefore a genomic λEMBL3 library was screened with a 270 bp *Eco*RI-*Hin*cII fragment derived from the 5′ end of cDNA clone DK14. Phages (24) were isolated and characterized. Two overlapping phages that spanned the largest stretch of DNA were phage MF3.19 and MF3.20 (Fig. 1B). MF3.20 contains five *Bssh*II sites mapping close together in a 1.4 kb *Hin*cII fragment (MF3.20Hc1.4). MF3.20Hc1.4 is the most 5′ fragment hybridizing to the 270 bp *Eco*RI-*Hin*cII cDNA probe. Fine mapping and subsequent sequence analysis showed that

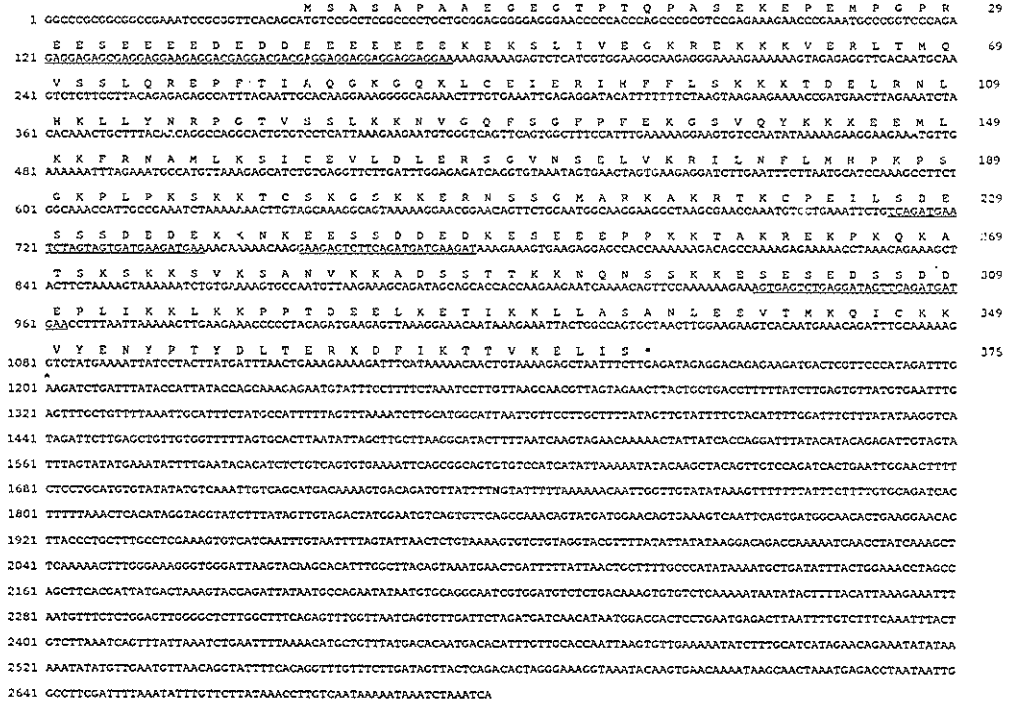


Figure 5. cDNA sequence and putative amino acid sequence of the DEK protein. The position of the t(6;9) breakpoint is indicated with a black triangle (nt 1080-81). Acidic regions are underlined.

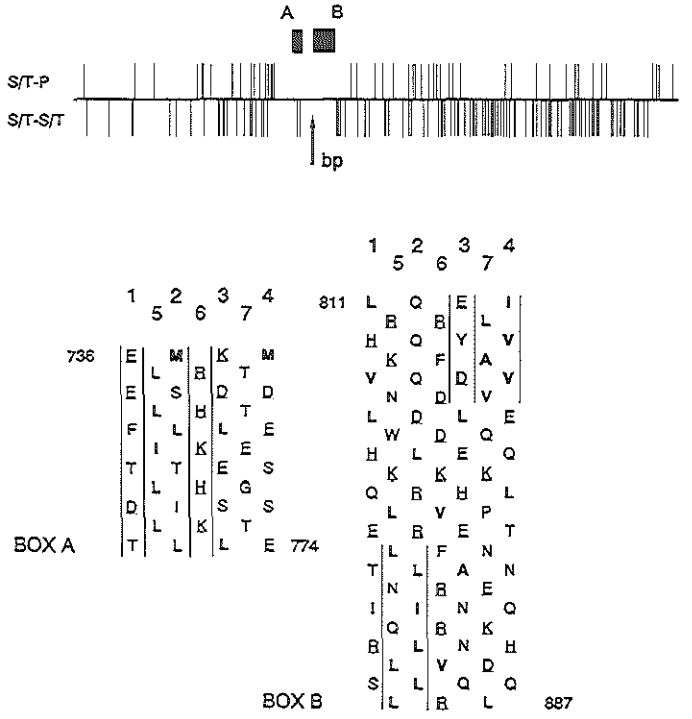


Figure 6. (A) Domains of the putative CAN protein that may have functional significance. The top line represents the 2192 amino acids putative CAN protein. The position of S/T-P and S/T-S/T dimers is indicated with vertical lines. Box A (amino acid 736-775) represents a putative leucine zipper. Box B (amino acid 811-887) represents two amphipathic helices separated by a region of charged amino acids. An arrow indicates the position of the translocation breakpoint between boxes A and B. The amino acid sequence of box A and B is given underneath the CAN protein. The first 4 amino acid residues are written in a horizontal row and the next three are placed below them and in between. In this way the sequence can be read as an helical wheel, cut open at one side. Charged amino acids are underlined, hydrophobic residues are in bold. Vertical lines indicate hydrophobic or charged sides of the predicted helical structure. (B) A part of the predicted CAN protein sequence present in boxB (shown above) is homologous to the human estrogen receptor. The homologous sequences are aligned. Identical amino acids are indicated with vertical lines, similar amino acids are indicated with dots. The C-terminal 22 amino acids of the ER (underlined) are essential for ER homodimerization.

the most 5' cDNA sequences are present in a 240 bp *Bssh*II fragment, preceded by a region rich in G/C (Fig. 3D).

*Bssh*II sites are mostly clustered in CpG-islands which often appear to represent promoter sequences (7). Therefore, this region in *dek* may well encompass the promoter area. Moreover, linking of a 520 bp genomic fragment, mapping immediately 5' of the cDNA homologous sequence, to a CAT reporter gene showed that this DNA region contains a strong promoter activity (Results not shown).

The 3' end of MF3.19 did not overlap with the most 5' end of the previously isolated phages (MF1.2); a gap is still present in the map. Long range mapping analysis (unpublished results) indicated that the distance between the translocation breakpoint of patient DK in *icb*-6 and the *Bssh*II sites is approximately 30 kb. As 27 kb of DNA between the *Bssh*II sites and the breakpoint of DK is present in the genomic phages, the gap is estimated to measure only a few kb.

Analysis of the *can* and *dek* cDNA sequence. Comparison of the predicted amino acid sequences of both *can* and *dek* with the EMBL data base failed to reveal any substantial homology to known protein sequences. However, detailed analysis identified some structures that may have functional significance. In CAN, from aa 736 to 774, N-terminal of *icb*-9, a leucine zipper motif is located. The leucine repeat consists of L-740, L-747, I-754, L-761 and L-768, and may represent a protein-protein dimerization domain. Projected on a helical wheel (Fig. 6A) hydrophobic residues at position +1 (relative to the leucines), acidic residues at position -1 and basic residues at position +2 are present. These residues may stabilize the formation of protein dimers through additional electrostatic interactions between the leucine repeats of proteins.

Immediately C-terminal of the breakpoint in CAN from aa 811 to 887 two putative amphipathic helices are present separated by a loop of 25 residues of which 13 are charged either positively (7) or negatively (6). The hydrophobic backbone of the first amphipathic helix is formed by I-814, L-817, V-821, A-824, V-828 and V-831 (Fig. 6A). The second amphipathic helix contains a heptad Leucine repeat consisting of L-861, I-868, L-875 and L-882 (Fig. 6A). A region encompassing the C-terminal part of the loop and the C-terminal amphipathic helix (aa 840 to 887) shows homology to the human estrogen receptor (ER) dimerization domain: 30% of the residues is identical, 57% similar (Fig. 6B). The C-terminal 22 aa of this homology region in the ER were shown to be essential for the formation of ER homodimers (38).

Many SP, TP and SS, ST, TS or TT dimers are present both N-terminal and C-terminal of the amino acid stretch containing the putative leucine zipper and amphipathic helices (Fig. 6A). This sequence motif has been proposed to have an ancillary role in DNA binding. At the C-terminus there is a

recurrence of phenylalanine residues often in combination with S/T-P or S/T-S/T dimers.

In the predicted DEK protein no specific structures could be recognized apart from a continuous stretch of acidic residues at the N-terminus, three acidic regions interspersed with serines and a very high overall percentage (42%) of charged amino acids (H, R, K, E, D).

Expression of *dek* and *can*. The expression pattern of *dek* and *can* in different mouse tissues may give a clue to the possible function of these genes. 20 μg of total RNA of bone marrow, spleen, thymus, brain, liver, kidney, testes, ovary, placenta and whole embryos of 10, 13, 16 and 19 days after conception, was loaded on a denaturing agarose gel. Hybridization of *dek* and *can* cDNA probes to hamster and mouse derived hybrid cell lines showed that both genes are conserved between species (unpublished results). Thus, the human *dek* cDNA clone DK14 and *can* cDNA clones

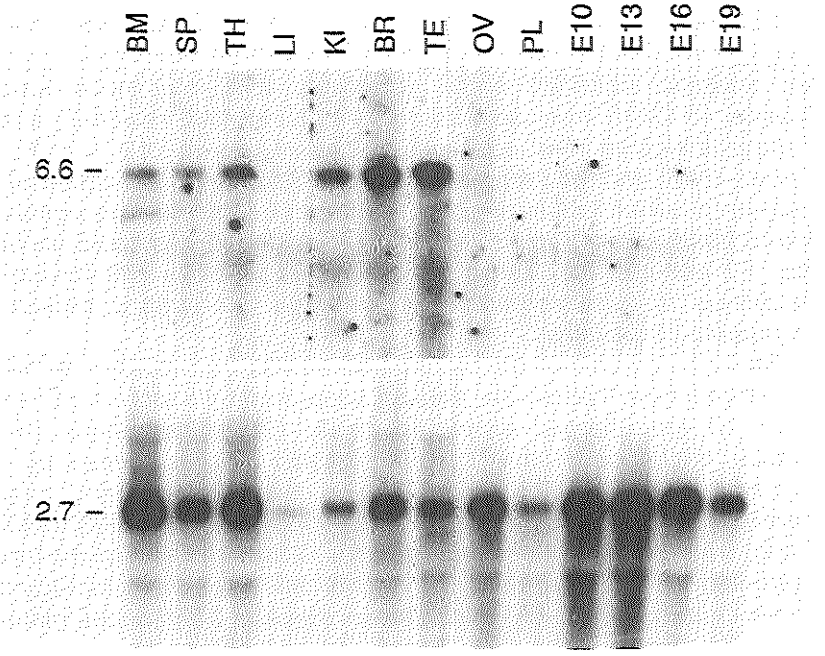


Figure 7. Northern blot containing total RNA of various mouse tissues, hybridized to *can* cDNA probe hXT37 and hXT56 (A) and to *dek* cDNA probe DK14 (B). The size of the transcript is indicated in kilobases. BM:bone marrow, SP:spleen, TH:thymus, LI:liver; KI:kidney, BR:brain, TE:testes, OV:ovary, PL:placenta(13 days after conception), E10, E13, E16, E19:embryos aged 10, 13, 16 and 19 days after conception respectively.

hXT37 and hXT56 were used to screen for mouse *dek* and *can* transcripts. As shown in Fig. 7 *dek* and *can* are expressed in all tissues. *Dek* is expressed at a relatively high level, while *can* seems to have a more restricted expression pattern. *Can* expression was easily detected in RNA of thymus, spleen, bone marrow, kidney, brain and testes but hardly visible in all other tissues or in whole embryos during development.

DISCUSSION

A novel fusion gene is present in leukemic cells carrying a (6;9)(p23;q34) translocation. The translocation breakpoints on chromosome 9 occur in one intron of the *can* gene: *icb*-9. Translocation breakpoints on chromosome 6 occur in one intron of the *dek* gene: *icb*-6. As a result of the translocation a *dek-can* fusion gene is generated encoding a chimeric *dek-can* transcript. The sequence of this chimeric cDNA predicts it to encode a 165 kDa DEK-CAN protein.

Although the precise position of the breakpoints in *icb-9* and *icb-6* may vary, the same exons of *dek* and *can* are joined by splicing of the primary transcript of the fusion gene. The invariable *dek-can* transcript can be used as a marker of t(6;9) AML that can be sensitively monitored by PCR (44). This may be a great advantage for diagnosis, monitoring response to chemotherapy and detection of minimal residual disease after bone marrow transplantation.

If steady state levels of *dek-can* and *dek* transcripts are compared in bone marrow from patient DK, it appears that *dek* mRNA is much more abundant than *dek-can* mRNA. The bone marrow from patient DK contains >90% leukemic cells, of which every cell contains one chromosome 6 and one chromosome 6p-. In this cell population the overall number of alleles of the normal *dek* gene and the fusion gene are about equal and both are driven by the *dek* promoter. Higher steady state levels of *dek* mRNA could be due to a longer half life of *dek* transcripts compared with that of *dek-can* transcripts. Alternatively, enhancer sequences could be present at the 3' side of the *dek* gene which are involved in transcription activation. The enhancer would be removed from the fusion gene by the translocation.

The cellular function of DEK and CAN and the way DEK-CAN may interfere with normal hematopoiesis, is still obscure. Neither of the two genes shows expression that is confined to the hematopoietic system. In fact, screening a Northern blot containing RNA samples of different mouse tissues showed that *dek* is expressed ubiquitously. *Can* is also expressed in all tissues, though at much lower and more variable levels. The tissues expressing *can* at a relatively high level include spleen and bone marrow. Since *can* mRNA is also found in human hematopoietic cell lines, it is unlikely that, because of the translocation, ectopic expression of *can* in hematopoietic cells would directly be involved in leukemogenic transformation. More likely, replacement

of the N-terminal part of CAN by DEK sequences generates a protein, that has different properties and is involved in transformation.

A breakpoint in *can* (*icb*-9) was also demonstrated in bone marrow cells from an AUL patient (SE) with an apparently normal karyotype (47). However, no breakpoint could be found in *dek*. In concordance with this observation, an aberrant transcript of 5.5 kb, detected by 3' *can* probes in bone marrow RNA from this AUL patient, failed to hybridize to 5' *dek* probes. cDNA cloning results strongly suggest that in this patient, *can* forms a fusion gene in which the 5' sequences are derived from another as yet unknown gene (Chapter 2.7). Therefore, it is possible that the C-terminal part of CAN contains domains involved in the leukemogenic process that may be activated by different N-terminal moieties.

Preliminary protein localization data were obtained by immunocytochemistry with antibodies directed against DEK and CAN, in COS cells transiently expressing high levels of CAN, DEK or DEK-CAN protein. CAN appears to be mainly cytoplasmic, while DEK has a strictly nuclear localization. The fusion of DEK to CAN results in a protein with a nuclear localization (M.F., M.v.L., G.G. manuscript in preparation).

In view of these data, the analysis of the *can* cDNA sequence revealed some structures that may be indicative for its function.

(*i*) An amphipathic helix with a heptad leucine repeat is predicted by the sequence just 5' of *icb*-9. This leucine zipper motif has been detected in many proteins like FOS, JUN, GCN4 and CCAAT/enhancer binding protein (10, 50), in which it mediates the formation of either homo- or heterodimers. A basic stretch of amino acids, juxtaposed to the leucine zipper, can function as a DNA binding element. In CAN no basic region is present adjoining the leucine zipper and hence this helix most likely functions as a dimerization domain. Not only the addition of novel sequences to the 3' part of *can*, but also the removal of the original 5' part of the gene may contribute to the putatively tumorigenic properties of the *dek-can* fusion gene. As the leucine zipper is detached from the C-terminal CAN sequences by the translocation it is tempting to speculate that this structure may be the interaction site for a factor that could regulate CAN activity.

(*ii*) The protein sequence just 3' of *icb*-9 predicts two amphipathic helices separated by a stretch of 25 amino acids, containing many charged residues. Several arguments suggest that this domain may function in protein dimerization. (a) The C-terminal putative helix and part of the preceding charged amino acids show homology to the hormone binding region of the human and mouse ER (23). It has been shown that the mouse ER contains a strong dimerization domain adjoining the hormone binding domain (18, 38). The entire sequence containing both domains is conserved within the steroid receptor family. N-terminal, the homology between the ER and CAN extends beyond the homology between the ER and other steroid hormone receptors. However, CAN has no homology to the hormone binding domain immediately C-terminal of the dimerization domain. It is interesting that the homologous

protein domain in another member of the steroid hormone receptor family, the retinoic acid receptor type α (RARα), was shown to dimerize with multiple cell type-specific proteins, which have not yet been characterized. Dimerization increased the affinity of the receptor for its cognate binding sequence (22). In addition, homology of CAN to the ER is noteworthy with regard to the finding that the RARα is involved in t(15;17) in acute promyelocytic leukemia (APL) (2, 8, 15). It will be interesting to analyze whether CAN can form heterodimers with the ER, or other members of the steroid hormone receptor family. (b) Although no direct homology is present, the putative structure of CAN just C-terminal of *icb*-9 (aa 811-887), architecturally resembles amino acid 82-162 of transcription factor AP-4, a basic stretch-helix-loop-helix protein (30). This part of AP-4 contains an additional dimerization domain, which, like this region in CAN, consists of two amphipathic helices separated by a stretch of 28 aa, containing many charged residues (30).

(*iii*) Many SP and TP dimers are present both N-terminal and C-terminal of the region containing the putative leucine zipper and amphipathic helices. A proline preceded by a serine or a threonine forms a ßturnI, stabilized by formation of hydrogen-bonds between the serine/threonine and the backbone of two amino acids following the proline (45). A ßturnI conformation can also be assumed by serine or threonine dimers. S/T-P dimers are clustered around DNA binding domains of many proteins that associate with DNA in a sequence specific manner. Suzuki (45) proposes that the S/T-P-X-X (X for any amino acid) motif will bind in the minor groove of DNA in a sequences, independent manner. This may stabilize a specific interaction of the DNA binding motif in the major groove.

In the C-terminal cluster of S/T-P and S/T-S/T dimers in CAN, the aromatic residue phenylalanine is often recurring. The C-terminal part of RNA polymerase II of both yeast and mammals contains a SPTSPSY repeat (3, 13), which is essential for its function (4). Suzuki (46) argues that the ßturnI-X-Y motif may be essential for DNA binding and shows that the aromatic ring of the tyrosine residue in this repeat can intercalate into the DNA. In *Drosophila* RNA polymeraseII, tyrosine is replaced by another aromatic residue: phenylalanine (6). A structure of ßturns combined with aromatic residues is therefore postulated to be a novel type of DNA binding domain. In the 3' part of CAN a S/T-S/T/P-X-F sequence occurs 14 times. We will study whether this region has DNA binding capacity, either by itself or by stabilizing DNA binding domains of transcription factor complexes that contain the CAN protein.

The predicted protein sequence of DEK contains a remarkably high percentage of charged amino acids. At the N-terminus (aa 30-47) DEK contains a continuous stretch of acidic residues. Three other acidic stretches are present from aa 227-236, 241-248 and 301-310. They contain acidic residues only interspersed by serine residues. Acidic regions were mainly found in two types of nuclear proteins (17). (*i*) Chromatin associated proteins

such as nucleolin and high mobility group (HMG) proteins contain acidic regions that can interact with the basic domains of histones (36, 37). These proteins also contain a conserved DNA binding domain, the HMG-box, a sequence motif that is not present in DEK. (*ii*) A class of transcriptional activators among which are Herpes simplex virus VP16 protein and the yeast transcription factor GCN4 contain an acidic patch that can interact with the RNA polymeraseII complex (9, 29).

Many basic amino acids are present in the DEK protein next to the acidic regions. The calculated pI of DEK is 8.9. Because of these basic stretches, several putative nuclear localization signals can be recognized. DEK is completely devoid of hydrophobic stretches.

We speculate that replacement of N-terminal CAN sequences by almost the entire DEK protein, may activate the transforming potential of CAN. However, the mechanism of this putative activation remains to be determined. Analysis of the primary structure of DEK and CAN combined with the preliminary localization data, suggests that these proteins may have a function in the cell nucleus.

Up to now breakpoints of three different translocations in myeloid leukemia have been cloned and molecularly analyzed. Thus far the formation of fusion genes seems to be the predominant effect of translocations in myeloid leukemia.

ACKNOWLEDGMENTS

We thank D. Bootsma and D. Meijer for continuous support and discussion, J. Abels, H. Adriaansen, A. Hagemeijer and B. Löwenberg for patient material, M. Kuit and T. de Vries Lentsch for photographic work. We are indebted to Dr. A. Coulson (MRC, Edinburgh) for his help with computer homology searches. This work was supported by the Dutch Cancer Foundation.

LITERATURE CITED

1. Adriaansen, H.J., J.J.M. van Dongen, H. Hooijkaas, K. Hählen, M.B. van 't Veer, B. Löwenberg and A. Hagemeijer. 1988. Translocation (6;9) may be associated with a specific TdT-positive immunological phenotype in ANLL. Leukemia 2:136-40.
2. Alcalay, M., D. Zangrilli, P.P. Pandolfi, L. Longo, A. Mencarelli, A. Giacomucci, M. Rocchi, A. Biondi, A. Rambaldi, F. Lo Coco, D. Diverio, E. Donti, F. Griniani and P.G. Pelicci. 1991. Translocation breakpoint of acute promyelocytic leukemia lies within the retinoic acid receptor alpha locus. Proc. Natl. Acad. Sci. USA 88:1977-81.
3. Allison, L.A., M. Moyle, M. Shales and C.J. Ingles. 1985. Extensive homology among the largest subunits of eukaryotic and prokaryotic RNA polymerases. Cell 42:599-610.
4. Allison, L.A., J.K. Wong, V.D. Fitzpatrick, M. Moyle and C.J. Ingles. 1988. The C-terminal domain of the largest subunit of RNA polymerase II of Saccharomyces cerevisiae, Drosophila melanogaster, and mammals: a conserved structure with an essential function. Mol. Cell. Biol. 8:321-9.
5. Auffray, C. and F. Rougeon. 1980. Purification of mouse immunoglobulin heavy-chain messenger RNAs from total myeloma tumor RNA. Eur. J. Biochem. 107:303-14.

6. Biggs, J., L.L. Searles and A.L. Greenleaf. 1985. Structure of the eukaryotic transcription apparatus: features of the gene for the largest subunit of Drosophila RNA polymerase II. Cell 42:611-21.
7. Bird, A.P. 1986. CpG islands and the function of DNA methylation. Nature 321:209-13.
8. Borrow, J., A.D. Goddard, D. Sheer and E. Solomon. 1990. Molecular analysis of acute promyelocytic leukemia breakpoint cluster region on chromosome 17. Science 249:1577-80.
9. Brent, R. and M. Ptashne. 1985. A eukaryotic transcriptional activator bearing the DNA specificity of a prokaryotic repressor. Cell 43:729-36.
10. Busch, S.J. and P. Sassone-Corsi. 1990. Dimers, leucine zippers and DNA-binding domains. Trends Genet. 6:36-40.
11. Chirgwin, J.M., A.E. Przybyla, J. MacDonald and W.J. Rutter. 1979. Isolation of Biologically Active Ribonucleic Acid from Sources Enriched in Ribonuclease. Biochemistry 18:5294-9.
12. Chomienne, C., P. Ballerini, N. Balitrand, M.E. Huang, I. Krawice, S. Castaigne, P. Fenaux, P. Tiollais, A. Dejean, L. Degos and H. de Thé. 1990. The retinoic acid receptor alpha gene is rearranged in retinoic acid-sensitive promyelocytic leukemias. Leukemia 4:802-7.
13. Corden, J.L., D.L. Cadena, J.J. Ahearn and M.E. Dahmus. 1985. A unique structure at the carboxyl terminus of the largest subunit of eukaryotic RNA polymerase II. Proc. Natl. Acad. Sci. USA 82:7934-8.
14. Daley, G.Q., R. Van Etten and D. Baltimore. 1990. Induction of chronic myelogenous leukemia in mice by the P210bcr/abl gene of the Philadelphia chromosome. Science 247:824-30.
15. de Thé, H., C. Chomienne, M. Lanotte, L. Degos and A. Dejean. 1990. The t(15;17) translocation of acute promyelocytic leukaemia fuses the retinoic acid receptor alpha gene to a novel transcribed locus. Nature 347:558-61.
16. de Thé, H., C. Lavau, A. Marchio, C. Chomienne, L. Degos and A. Dejean. 1991. The PML-RARA Fusion mRNA Generated by the t(15;17) Translocation in Acute Promyelocytic Leukemia Encodes a Functionally Altered RARα. Cell 66:675-84.
17. Earnshaw, W.C. 1987. Anionic regions in nuclear proteins. J. Cell. Biol. 105:1479-82.
18. Fawell, S.E., J.A. Lees, R. White and M.G. Parker. 1990. Characterization and colocalization of steroid binding and dimerization activities in the mouse estrogen receptor. Cell 60:953-62.
19. Feinberg, A.P. and B. Vogelstein. 1983. A technique for radiolabeling DNA restriction endonuclease fragments to high specific activity. Anal. Biochem. 132:6-13.
20. Fourney, R.M., J. Miyakoshi, R.S. Day III and M.C. Paterson. 1988. Northern Blotting: efficient RNA Staining and Transfer. Focus 10:5-6.
21. Frohman, M.A. 1990. Rapid Amplification of cDNA Ends (RACE): User-Friendly cDNA Cloning. Amplifications 5:11-15.
22. Glass, C.K., O.V. Devary and M.G. Rosenfeld. 1990. Multiple Cell Type-Specific Proteins Differentially Regulate Target Sequence Recognition by the alpha Retinoic Acid Receptor. Cell 63:729-38.
23. Green, S., P. Walter, V. Kumar, A. Krust, J.M. Bornert, P. Argos and P. Chambon. 1986. Human oestrogen receptor cDNA: sequence, expression and homology to v-erb-A. Nature 320:134-9.
24. Gubler, U. and B.J. Hoffman. 1983. A simple and very efficient method for generating cDNA libraries. Gene 25:263-9.
25. Heisterkamp, N., G. Jenster, H.J. ten, D. Zovich, P.K. Pattengale and J. Groffen. 1990. Acute leukaemia in bcr/abl transgenic mice. Nature 344:251-3.
26. Hermans, A., J. Gow, L. Selleri, M. von Lindern, A. Hagemeijer, L.M. Wiedemann and G. Grosveld. 1988. bcr-abl oncogene activation in Philadelphia chromosome-positive acute lymphoblastic leukemia. Leukemia 2:628-33.
27. Hermans, A., N. Heisterkamp, M. von Linden, S. van Baal, D. Meijer, D. van der Plas, L.M. Wiedemann, J. Groffen, D. Bootsma and G. Grosveld. 1987. Unique fusion of bcr and c-abl genes in Philadelphia chromosome positive acute lymphoblastic leukemia. Cell 51:33-40.

28. Hermans, A., L. Selleri, J. Gow, L. Wiedeman and G. Grosveld. 1989. Molecular Analysis of the Philadelphia Translocation in Myelogenous and Acute Lymphoblastic Leukemia. Cancer Cells 7:21-26.
29. Hope, I.A. and K. Struhl. 1986. Functional dissection of a eukaryotic transcriptional activator protein, GCN4 of yeast. Cell 46:885-94.
30. Hu, Y.F., B. Luscher, A. Admon, N. Mermod and R. Tjian. 1990. Transcription factor AP-4 contains multiple dimerization domains that regulate dimer specificity. Genes Dev. 4:1741-52.
31. Huynh, T.V., R.A. Young and R.W. Davis. 1984. Construction and Screening cDNA Libraries in λgt10 and λgt11. p. 49-78. *In* Glover, D. (ed.), DNA Cloning Techniques: A Practical Approach. IRL Press, Oxford.
32. Kadowaki, T., H. Kadowaki and S.I. Tayler. 1990. A nonsense mutation causing decreased levels of insulin receptor mRNA: Detection by a simplified technique for direct sequencing of genomic DNA amplified by the polymerase chain reaction. Proc. Natl. Acad. Sci. USA 87:658-62.
33. Kakizuka, A., W.H. Miller, K. Umesono, R.P. Warrell, S.R. Frankel, V.V.V.S. Murty, E. Dmitrovsky and R.M. Evans. 1991. Chromosomal translocation t(15;17) in human acute promyelocytic leukemia fuses RARα with a novel putative transcription factor, PML. Cell 66:663-74.
34. Konopka, J.B., S.M. Watanabe and O.N. Witte. 1984. An alteration of the human c-abl protein in K562 leukemia cells unmasks associated tyrosine kinase activity. Cell 37:1035-42.
35. Kozak, M. 1988. A profusion of controls. J. Cell. Biol. 107:1-7.
36. Krohne, G. and W.W. Franke. 1980. A major soluble acidic protein located in nuclei of diverse vertebrate species. Exp. Cell. Res. 129:167-89.
37. Lapeyre, B., H. Bourbon and F. Amalric. 1987. Nucleolin, the major nucleolar protein of growing eukaryotic cells: an unusual protein structure revealed by the nucleotide sequence. Proc Natl Acad Sci U S A 84:1472-6.
38. Lees, J.A., S.E. Fawell, R. White and M.G. Parker. 1990. A 22-amino-acid peptide restores DNA-binding activity to dimerization-defective mutants of the estrogen receptor. Mol. Cell. Biol. 10:5529-31.
39. Pearson, M.G., J.W. Vardiman, B.M. Le, J.D. Rowley, S. Schwartz, S.L. Kerman, M.M. Cohen, E.W. Fleischman and E.L. Prigogina. 1985. Increased numbers of marrow basophils may be associated with a t(6;9) in ANLL. Am. J. Hematol. 18:393-403.
40. Sambrook, J., E.F. Fritsch and T. Maniatis. 1989. Molecular Cloning: a Laboratory Manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, New York.
41. Sandberg, A.A., R. Morgan, J.A. McCallister, M.B. Kaiser and F. Hecht. 1983. Acute myeloblastic leukemia (AML) with t(6;9) (p23;q34): a specific subgroup of AML? Cancer Genet. Cytogenet. 10:139-42.
42. Sanger, F., S. Nicklen and A.R. Coulson. 1977. DNA Sequencing with Chain-terminating Inhibitors. Proc. Natl. Acad. Sci. USA 74:5463-67.
43. Shtivelman, E., B. Lifshitz, R.P. Gale and E. Canaani. 1985. Fused transcript of abl and bcr genes in chronic myelogenous leukaemia. Nature 315:550-4.
44. Soekarman, D., M. von Lindern, S. Daenen, B. de Jong, C. Fonatsch, B. Heinze, C.R. Bartram, A. Hagemeijer and G. Grosveld. The translocation (6;9)(p23;q34) shows consistent rearrangement of two genes and defines a myeloproliferative disorder with specific clinical features. Blood, in press.
45. Suzuki, M. 1989. SPXX, a Frequent Sequence Motif in Gene Regulatory Proteins. J. Mol. Biol. 207:61-84.
46. Suzuki, M. 1990. The heptad repeat in the largest subunit of RNA polymerase II binds by intercalating into DNA. Nature 344:562-5.
47. von Lindern, M., A. Poustka, H. Lerach and G. Grosveld. 1990. The (6;9) chromosome translocation, associated with a specific subtype of acute nonlymphocytic leukemia, leads to aberrant transcription of a target gene on 9q34. Mol. Cell. Biol. 10:4016-26.

48. von Lindern, M., T. van Agthoven, A. Hagemeijer, H. Adriaansen and G. Grosveld. 1989. The human pim-1 gene is not directly activated by the translocation (6;9) in acute nonlymphocytic leukemia. Oncogene 4:75-9.
49. Walker, L.C., T.S. Ganesan, S. Dhut, B. Gibbons, T.A. Lister, J. Rothbard and B.D. Young. 1987. Novel chimeric protein expressed in Philadelphia positive acute lymphoblastic leukaemia. Nature 329:851-3.
50. Ziff, E.B. 1990. Transcription factors: a new family gathers at the cAMP response site. Trends Genet. 6:69-72.

CHAPTER 2.5

The translocation (6;9)(p23;q34) shows consistent rearrangement of two genes and defines a myeloproliferative disorder with specific clinical features

D. Soekarman[1], M. von Lindern[1], S. Daenen[2], B. de Jong[2], C. Fonatsch[3], B. Heinze[4], C.R. Bartram[5], A. Hagemeijer[1] and G. Grosveld[1]

[1]dept. Cell Biology and Genetics, Erasmus University, Rotterdam, The Netherlands.
[2]dept. Hematology and dept. Genetics, University of Groningen, Groningen, The Netherlands.
[3]dept. Human Genetics, University Medical school of Lübeck, Lübeck, Germany.
[4]Inst. Occupational & Social Medicin, University of Ulm, Ulm, Germany.
[5]dept. Pediatrics, Section of Molecular Biology, University of Ulm, Ulm, Germany.

Blood, in press

SUMMARY

Translocation (6;9)(p23;q34) is a cytogenetic aberration which can be found in specific subtypes of both acute myeloid leukemia (AML) and myelodysplastic syndrome (MDS). This translocation is associated with an unfavorable prognosis. Recently, the genes involved in the t(6;9) were isolated and characterized. Breakpoints in both the *dek*-gene on chromosome 6 and the *can*-gene on chromosome 9 appear to occur in defined regions which allows us to diagnose this type of leukemia at the molecular level. Moreover, due to the translocation a chimeric *dek-can* mRNA is formed which, as we show here, is an additional target for diagnosis via cDNA-preparation and the polymerase chain reaction (PCR). We studied 17 patients whose blood- and/or bone marrow cells showed a t(6;9) with karyotypic analysis (Table 1). Fourteen patients suffered from AML, one patient had a refractory anemia with excess of blasts in transformation (RAEBt), one patient had an acute myelofibrosis (AMF) and one patient a chronic myeloid leukemia (CML). In 9 cases studies at the DNA- and RNA-level were possible while in 7 cases only the DNA could be analyzed. In one case only RNA was available. Conventional Southern blot analysis showed the presence of rearrangements of both the *dek*-gene and the *can*-gene. In both genes, breakpoints cluster in one intron in the patients investigated. The presence of a consistent chimeric *dek-can* product after cDNA-preparation followed by the PCR was demonstrated. We conclude from our data that the t(6;9) is found in myeloproliferative disorders with typical clinical characteristics. This translocation results in highly consistent abnormalities at the molecular level.

INTRODUCTION

Since the discovery in 1960 of the Philadelphia chromosome in cases of chronic myeloid leukemia (CML) by Nowell and Hungerford, a large number of leukemias has been associated with specific chromosomal translocations (1,2). The development of new techniques enabled molecular biologists to isolate and characterize a number of genes involved in reciprocal chromosome translocations. Well-known examples are the t(8;14) in Burkitt's lymphoma (BL) in which the *myc*-gene on chromosome 8 is linked to the IgH-chain locus on chromosome 14 and t(9;22) in CML in which breakpoints occur in the *abl*-gene on chromosome 9 and the *bcr*-gene on chromosome 22 (3,4,5).

In a specific subgroup of AML a t(6;9)(p23;q34) can be found (6-10). Patients with this type of leukemia are usually quite young and their prognosis is poor. Blast-cells are mostly classified as FAB-M2 or M4 (90%) and in a minority as M1(10%). At the time of diagnosis the t(6;9) is usually the sole cytogenetic aberration. Additional karyotypic abnormalities are rare but may occur during progression of the disease (9-18). Recently, the genes located at the chromosomal breakpoints of this translocation were isolated and characterized (19). The gene on chromosome 6 which participates in the reciprocal exchange is called *dek* and encompasses 40 kb (von Lindern et al, manuscript in preparation). Southern blot analysis of 4 patients with t(6;9) indicated that breakpoints are located in one intron of 9 kb which is called 'intron containing breakpoints on chromosome 6' or icb-6. The *can*-gene on chromosome 9 is more than 130 kb in length. Here, breakpoints occur in one intron of 7.5 kb (icb-9) which is located in the middle of the gene. The *can*-gene is transcribed into a 6.6 kb mRNA. Due to the translocation the 3' part of the *can*-gene is fused to the 5' part of *dek*, resulting in a chimeric *dek-can* gene on the 6p- derivative (Chapter 2.4) This chimeric gene is transcribed into an aberrant 5.5 kb mRNA. The functions of the normal *dek* and *can* gene products are as yet unknown and it is equally unclear in which way the hybrid product may be involved in leukemogenesis.

In this study we analyzed 14 patients with AML, one patient with RAEBt, one patient with AMF and one patient with CML whose blood or bone marrow cells carried a t(6;9). Investigation of the leukemic cells at the DNA- and RNA-level confirms the highly consistent involvement of both the *dek*- and *can*-genes in this translocation. The myeloproliferative disorder marked by a t(6;9) appears as a distinct clinical entity which, as we show here, can now be diagnosed and monitored at the molecular level.

PATIENTS

Clinical and hematologic data of the patients are given in table 1. In the case of

patients 1 to 5 various data were published previously (13,19,20,21). Patients 6 to 17 were newly admitted cases and fresh or frozen samples were sent to us for molecular investigations by the following centers: Regional Cancer Center Marseille, France (patients 6 and 17), Universitätsklinik Ulm, Germany (patients 7 and 16), University Hospital Groningen, The Netherlands (patient 8), Centre Regional de Transfusion Sanguine et de Génétique Humaine Bois-Guillaume, France (patient 9), Free University Hospital Amsterdam, The Netherlands and Stichting Nederlandse Werkgroep Leukemie bij Kinderen The Hague, The Netherlands (patient 10), Medical Center of the University of Amsterdam, The Netherlands (patients 11 and 15), Imperial Cancer Research Fund, Saint Bartholomew's Hospital, London, United Kingdom (patients 12 and 13), Children's Cancer Research Institute Vienna, Austria (patient 14).

MATERIALS AND METHODS

Samples. Bone marrow aspirates and blood samples were collected in heparinized tubes. After isolation of the white fraction by Dextran or a Ficoll-Hypaque gradient, cells were frozen and stored in liquid Nitrogen until used.

Conventional Southern Blot Analysis. DNA was isolated from blood or bone marrow cells according to standard procedures (22) or high molecular weight DNA was prepared in agarose plugs as described previously (23). The following restriction enzymes were used for digestion of DNA: EcoRV, BamHI, HindIII and/or BglII. DNA-fragments were separated on a 0.7% agarose gel and blotted onto nylon filters (Zeta-probe, Biorad Lab., Richmond, CA) according to the manufacturers instructions. Probes used for hybridization, hybridization- and washing-conditions were described previously (19). In short, rearrangement of the *dek*-gene is detectable using radio-labeled probes MF1E.5 (a 500 bp EcoRI-EcoRI fragment) and MF2BH (an 800 bp BglII-HindIII fragment). Probes AL1F4EP (a 200 bp EcoRI-PstI fragment) and AL1F6E.3 (a 300 bp EcoRI-EcoRI fragment) are used for the detection of breakpoints in the *can*-gene. In figures 1A and 1B a schema-tic representation is given of a simplified restriction-map of both icb-6 and icb-9 and the localization of the probes used is indicated.

Polymerase Chain Reaction (PCR). RNA-isolation, cDNA-preparation and PCR-conditions were described previously (24,25). As a control for the cDNA-synthesis and PCR-reaction not only the chimeric *dek-can* cDNA was amplified from the patient RNA but also the normal *can* cDNA. 1 unit Taq polymerase was added per reaction (Cetus Corp., Emeryville, CA, USA or BRL, Gibco Lab., Life Technol. Inc., NY, USA). Primers used for cDNA-synthesis and PCR-amplification consisted of either of the two following sets of sequences:

Primer-set I
3'primer in *can* = 5' ACCAGGTGATTCAGCCT 3'
5'primer in *can* = 5' CTGAAAACAACTTTACTTGA 3'
5'primer in *dek* = 5' CCTACAGATGAAGAGTTAA 3'
or:
Primer-set II
3'primer in *can* = 5' GTGTCTCTCGCTCTGG 3'
5'primer in *can* = 5' AAGAGACCACAGAGTCG 3'
5'primer in *dek* = 5' GGCCAGTGCTAACTTGG 3'

The anneal-temperature for these primers was chosen at 45°C. The PCR was performed using 24 cycles of denaturation (1 min. 30 sec. at 92°C), annealing (1 min. at 45°C) and extension (5 min. at 72°C).Analysis of residual disease was performed by

Table 1 CLINICAL AND HEMATOLOGIC DATA OF 17 PATIENTS WITH A TRANSLOCATION (6;9)(p23;q34)

Patient	Age/ Sex[c]	FAB	Clinical phase[d]	Karyotypic abnorm.	% Abn. metaph.	Baso-philia[e]	WBC 10^9/l	% Blasts BM	% Blasts PB	Southern blot[f]	PCR[g]	Response to therapy[h]	Survival after diagnosis
1[a]	13/F	RAEB M4	diagn. relapse	t(6;9) t(6;9) + add.abn.	100	no	1,66			+	RNA n.a.	PR	± 18 months
2[a]	63/F	M4	diagn.	t(6;9)	93		87	47	30	+	no ampl.	PR	
3[a]	38/M	AMF AMF	diagn. relapse	NM t(6;9)	30	no	5,2	9	5	+	no ampl.	PR	29 months
4[a]	17/F	M4	diagn.	t(6;9)	97	no	21,2	59	18	+	+	no therapy	3/4 months
5[b]	13/F	M4	diagn. relapse	t(6;9) 6p- +add.abn.	87 91	no	29,4	73	58	DNA n.a.	+	CR	30 months
6	35/M	M2	diagn. relapse	t(6;9) t(6;9), +8 6p-	80 20 100	yes	12,6 92	58	35 95	+	no ampl.	PR	18 months
7	28/F	M4	relapse	t(6;9)	91	no	5,1	92	20	+	+	CR	17 months
8	18/M	M2	diagn. after ther.	t(6;9),inv.1 ND	35	no	10	25 10	< 1	+ DNA n.a.	RNA n.a. BM: + PB: -	CR	> 24 months
9	19/F	M2	relapse	t(6;9)	100	yes	100	87	94	+	+	CR	14 months
10	14/F	M4	diagn.	t(6;9)				30		+	+	CR	> 36 months

11	6/F	M4	diagn.	t(6;9)		no	262	80	45	+	+	NR	2 months
12	28/M	M1	diagn.	t(6;9)	100		65,9		79	+	+		
13	24/F	M1	diagn.	t(6;9), del.7q t(6;9)	40 60		54,8		95	+	RNA n.a.		
14	10/M	M2	diagn.	t(6;9)	100	no	47,5	52	37	+	no ampl.	PR	20 months
15	53/M	M4	relapse	t(6;9)	91	yes	100	41	31	+	+	PR	> 4 months
16	54/F	CML Ph', bcr'	diagn.	t(6;9)	100	no	92		9	+	RNA n.a.		11 months
17	53/F	RAEBt	diagn. after ther.	t(6;9) ND	100	no	2,5	22	3	+	+	no chemoth.	15 months

a) Published previously by Von Lindern et al., ref. 19
b) Published previously by Adriaansen et al., ref. 21
c) Age in years at diagnosis Sex: M = male, F = female
d) diagn. = at diagnosis
after ther. = after therapy: patient 8 - chemotherapy, patient 17 - IL2 therapy
e) > 0.25% basophils in the bone marrow
f) + = Southern blot analysis shows rearrangement of the dek- and the can- gene
DNA n.a = DNA was not available.
g) + = PCR analysis shows amplification of a chimeric dek-can fragment
- = no amplification of dek-can while can is amplified.
RNA n.a. = RNA was not available
no ampl. = despite intact RNA no amplification by PCR
h) CR = complete remission, PR = partial remission, NR = no remission

NM = no mitosis
ND = not done
WBC = white blood cell count
BM = bone marrow
PB = peripheral blood

serial 10-fold dilutions of 1 μg of total patient RNA into 10 μg of yeast RNA. Procedures for cDNA-synthesis and PCR were as described above. Although we do not know yet the exact copy number of the *can*- and *dek-can* mRNA per leukemic cell, we assume that both show comparable levels of expression based on data obtained by Northern blot analysis (2 to 3 fold more *dek-can* mRNA than *can* mRNA, ref.19). Amplification by PCR of the *can*- and *dek-can* fragment occurs with the same efficiency. We compare PCR-results for *can* with *dek-can*. If 100% of cells with *can* mRNA also express *dek-can* mRNA all dilution samples will be positive for both *can* and *dek-can*. If 10% of cells carry a t(6;9) no *dek-can* signal is present in the highest dilution while this still contains *can* signal.

For the sequence-specific detection of amplified fragments two oligomers were designed:

detection of normal *can*-cDNA: 5' GTTATCTGCATTTGCT 3'

detection of *dek-can* fusion-cDNA: 5' GCAAAAAGGAAATTCG 3'

Both ^{32}P end-labeled probes were hybridized at 39°C for 3 hours and filters were washed in 1xSSPE for 1 hour at the same temperature. A schematic representation of the PCR-procedure is given in figure 2.

RESULTS

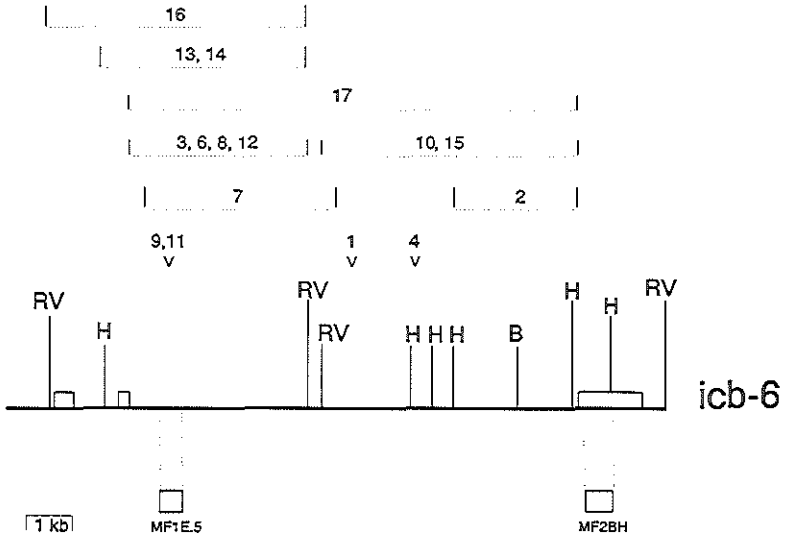


Figure 1A. Restriction map of the icb-6. Exons are depicted as open boxes. The localization of probes MF1E.5 and MF2BH are indicated as open boxes beneath the map. Stippled lines or arrowheads indicate localization of the breakpoint in the *dek*-gene of the patients. The patients are numbered according to table 1. The breakpoint-containing fragments of patients 1 and 4 were cloned and the precise breakpoints were localized (Ref. 19). The breakpoints of patients 9 and 11 fall into probe MF1E.5. RV = *Eco*RV, H = *Hin*dIII, B = *Bam*HI.

Karyotyping of blood or bone marrow cells of patients was performed in the centers that sent samples for molecular analysis. In metaphases of all patients a t(6;9)(p23;q34) had been found. Patients 1 to 4 were analyzed previously with Southern blot (table 1, ref. 19). Additionally, we obtained DNA from 12 newly admitted cases (patients 6-17, table 1). DNA of each patient was digested using a minimum of three different restriction enzymes. In figures 3 and 4 some results are shown of conventional Southern blot analysis in various patients. A combination of three different restriction enzymes and four probes enabled us to localize breakpoints in the *dek*- and the *can*-gene in all 16 patients. Moreover, digestion with an additional restriction enzyme (BglII or EcoRI) (or results of PCR-experiments, see further) narrowed down the breakpoint localization in most patients to the icb-6 and icb-9.

An example is patient 10: exact localization of the breakpoint in the *can*-gene was impossible using Southern blot alone since only the *Eco*RV-digest generated an aberrant fragment with probes AL1F4EP and AL1F6E.3 (figure 4), while the *Hin*dIII- and *Bam*HI-digests showed germ-line bands (data not shown). However, a PCR-experiment generated a chimeric *dek-can* fragment (see below), indicating that in frame joining of

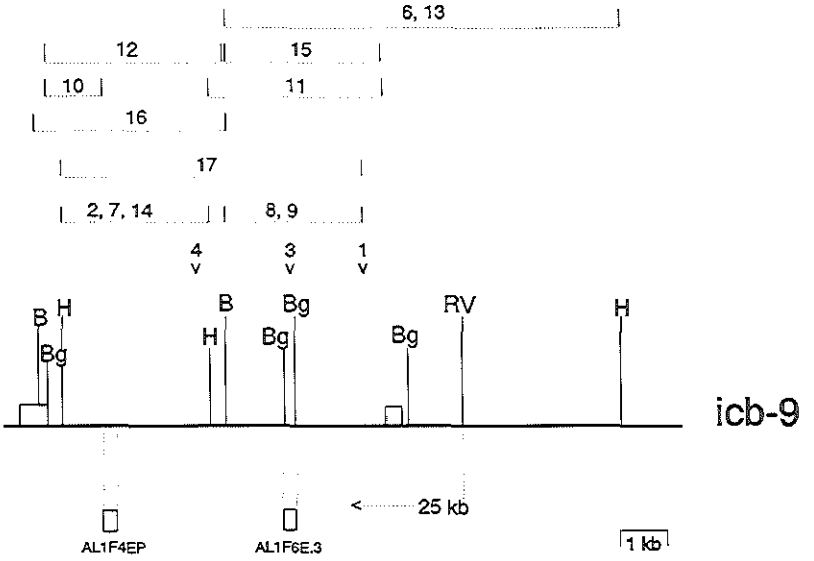


Figure 1B. Restriction map of icb-9. Exons are indicated as open boxes. Probes AL1F4EP and AL1F6E.3 are indicated beneath the map as open boxes. Stippled lines or arrowheads indicate localization of the breakpoint in the *can*-gene of the patients which are numbered according to table 1. The breakpoint-containing fragments of patients 1 and 4 were cloned and the precise breakpoints were localized (Ref. 19). The breakpoint of patient 3 falls into probe AL1F6E.3. RV = *Eco*RV, H = *Hin*dIII, B = *Bam*HI, Bg = *Bgl*II.

dek and *can* had occurred.

In the *Eco*RV-digest of this patient, 3' *dek* probe MF2BH detects an aberrant band of 7.5 kb (figure 3). However, a much larger fragment should be expected since the first *Eco*RV site 5' of icb-9 in *can* is located 16 kb upstream of this intron (Figure 1B). Therefore, we assume that patient 10 has a large deletion of the 5' *can*-gene which has not been mapped more precisely. The most accurate prediction for the position of the icb-9 breakpoint in this patient is between the 5' end of icb-9 and probe AL1F4EP (results not shown).

Patient 8 was diagnosed in a preleukemic phase preceding overt AML. His bone marrow contained 25% blasts while in the periferal blood only 1% blasts was observed (table 1). He suffered from paraneoplastic neutrophilic dermatosis also called variant Sweet's syndrome (26). Blood differential counts indicated that more than 95% of his leucocytes were granulocytes. The marrow also contained variably high numbers of granulocytes which diluted the blasts and therefore it was difficult to diagnose AML on purely morphological grounds. We determined whether the periferal blood granulocytes carried the t(6;9). Indeed, Southern blot analysis of both blood and bone marrow generated germ line and aberrant fragments of equal intensity indicating the presence of the t(6;9) in the vast majority of the marrow and blood cells and implying that the

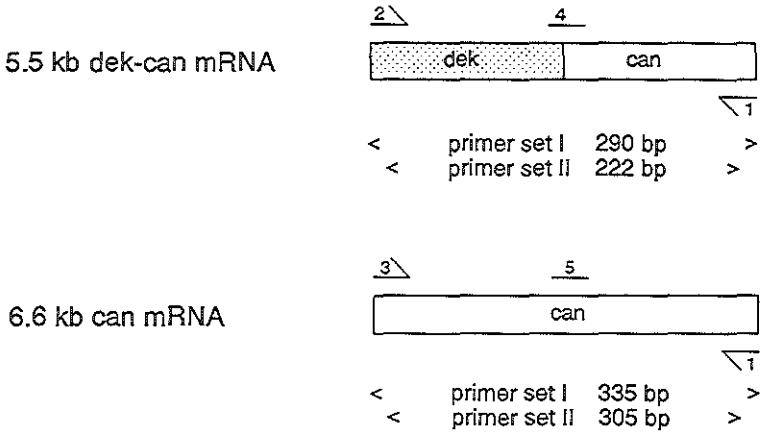


Figure 2. Schematic representation of the reversed PCR of *dek-can* and *can* mRNA. *dek* is represented as a stippled bar while *can* is shown as an open bar. A 3'primer in *can* (1) is used to make cDNA while the PCR is performed with this primer in combination with a 5'primer either in *dek* (2) or in *can* (3). This results in amplification of a *dek-can* or *can* fragment which can be detected by the specific oligo-probes 4 and 5. The size of the fragments generated by the various primer-sets are indicated beneath each drawing.

granulocytes carried the t(6;9) as well (figure 3).

In 14 cases (patients 2-12, 14, 15, 17, table 1) sufficient material was available for RNA-isolation and subsequent PCR-amplification of chimeric *dek-can* cDNA. As shown in figure 2, primer-set I generates a fragment of 290 bp when a chimeric *dek-can* mRNA is present and a 335 bp fragment from the *can* mRNA-template. However, these primers were not optimal and gave rise to aspecific priming on the *can*-mRNA (see figure 6). Therefore, primer set II was made. The *dek-can* product of primer set II is 222 bp while the *can*-fragment is 305 bp (figure 2). Results of the analysis of 4 patients using primer-set II are shown in figure 5A. Via PCR a hybrid *dek-can* mRNA was detected in 10 patients with a t(6;9), while RNA from leukemia patients without a t(6;9) or healthy controls yielded only the fragment derived from the *can* mRNA. No conclusions can be drawn for patients 2, 3, 6 and 14. Although intact RNA of these patients was available, neither a *dek-can* or a *can*-fragment was detected after cDNA-preparation followed by PCR.

Since PCR is a sensitive diagnostic method we used this technique for analysis of material containing a minority of cells carrying the t(6;9). Comparison of amplification of *can* with amplification of *dek-can* in a

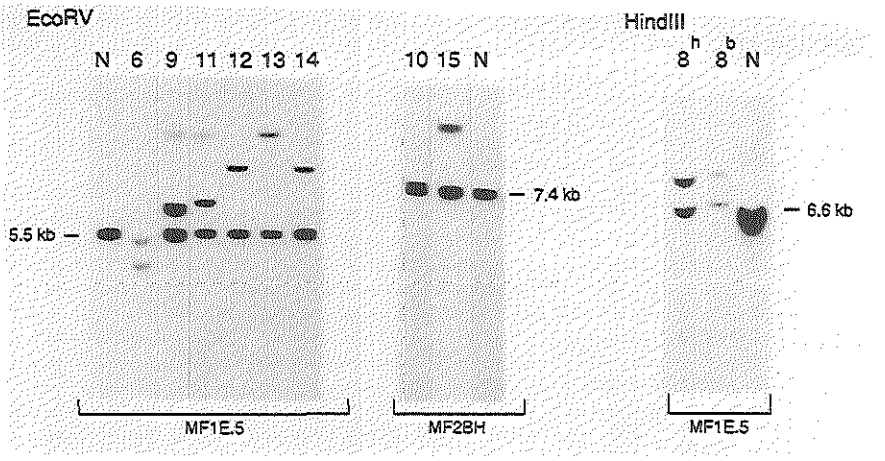


Figure 3. Detection of breakpoints in the *dek*-gene using *Eco*RV (patients 6,9,11,12,13,14,10,15) or *Hin*dIII (patient 8) as restriction enzymes in combination with probes MF1E.5 and MF2BH. The patients are numbered according to table 1. The size of the germline bands is indicated in kilobases (kb's). Aberrant fragments have the following sizes: patient 6: 4.4kb, patient 9: 7.5kb and >23kb, patient 11: 8kb and >23kb, patient 12: 16kb, patient 13: >23kb, patient 14: 16kb, patient 10: 7.5kb, patient 15: >23kb, patient 8: 9.5kb. 8h=DNA of periferal blood cells from patient 8, 8b=DNA of bone marrow cells from patient 8. N=DNA of thymus or white blood cells from a non-leukemic individual.

dilution series gives an estimation of the fraction of cells with t(6;9) in a sample from a patient. Both blood and bone marrow cells from patient 8 were investigated after chemotherapy. Hematomorphologic studies indicated that still 10% blasts were present in the bone marrow (table 1). With PCR the presence of residual disease in bone marrow cells was seen in dilutions 10^{-1} and 10^{0}.

The signal of *dek-can* in the 10^{0} dilution is comparable with the *can* signal in the 10^{-2} dilution, indicating that roughly 1/100 cells carry the t(6;9) (see Materials and Methods). In the blood no t(6;9) carrying cells could be detected (figure 6).

Patient 17 had been diagnosed as an RAEBt and in diagnostic phase all metaphases showed a t(6;9). Blood cells were analyzed by us from a period after 4 days of IL-2 therapy which was given for activation of the T-cells. The karyotype of this sample was unknown but 3% blasts had been observed in the periferal blood. Since this low amount of blasts may be difficult to detect by Southern blotting we decided to use PCR. Clearly, a chimeric *dek-can* fragment was generated, indicating persistence of the translocation (figure 5B). Despite the low number of blast cells, Southern blot analysis confirmed this result (data not shown).

DISCUSSION

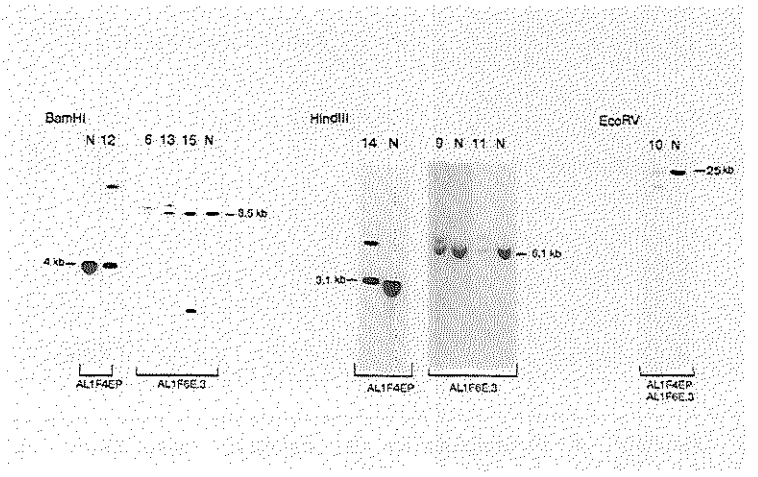


Figure 4. Detection of breakpoints in the *can*-gene of 8 patients using restriction-enzymes *Bam*HI, *Hin*dIII or *Eco*RV in combination with probes AL1F4EP and AL1F6E.3. The patients are numbered according to table 1. The size of the germline fragments is indicated in kilobases (kb's). Aberrant fragments have the following sizes: patient 12: >23kb, patient 6: 9.5kb, patient 13: 10kb, patient 15: 2.4kb, patient 14: 5kb, patient 9: 6.5kb, patient 11: 8kb, patient 10: 11kb. N=DNA of thymus or white blood cells from a non-leukemic individual.

The entity of acute myeloid leukemia (AML) consists of a heterogeneous group of diseases. Subclassification according to the French-American-British Cooperative Group (FAB) facilitates diagnosis and enables physicians from different hematologic centers to exchange and compare data (27). Moreover, the prognosis for a patient depends on the FAB-subtype which is found. Additional independent prognostic factors are the various chromosomal abnormalities which are linked to specific FAB-subtypes (28,29). A t(6;9)(p23;q34) can be found in 0.5 to 4% of patients with AML (6-8,11,14,30-32). To date, 34 patients have been reported in the literature whose karyotype showed a t(6;9) (6-12,14-18,30-36). When the 17 patients are added who were analyzed by us at the molecular level, some general features emerge for this group of 51 patients. Diagnosis of this disease is usually made in the second or third decade of life in contrast to AML as a whole group in which the median age is above 60 years. No striking male or female preponderance is found. FAB-classification of the AML was frequently reported as M2 or M4 with a minority of M1. However, the translocation is not restricted to these subtypes of AML. Cuneo et al (20) pointed out that a number of patients with t(6;9) was reported who were diagnosed as a refractory anemia with

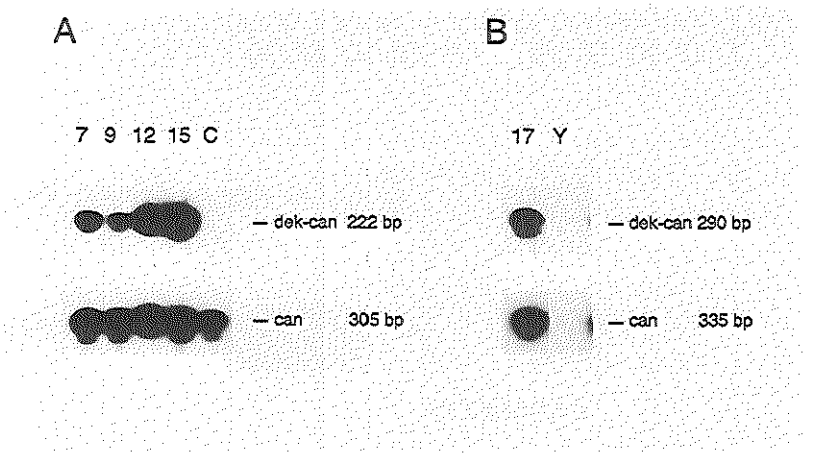


Figure 5. PCR detection of *can* and *dek-can* transcripts.
A. Results of a PCR-experiment of four patients using primer set II. C = control material from an ALL patient with a normal karyotype. Primers in *can* generate a 305 basepair *can* fragment in all five individuals while a primer in *can* in combination with a primer in *dek* generates a 222 basepair *dek-can* fragment only in the four patients with a t(6;9).
B. Results of the PCR-experiment of patient 17 using primer set I. Y = yeast RNA, used as a negative control. Primers in *can* generate a 335 bp fragment while a primer in *can* in combination with a primer in *dek* generate a 290 bp *dek-can* fragment in the patient. Y = Yeast stays negative for both *can* and *dek-can*.

excess of blasts (RAEB) which is a subtype of myelodysplastic syndrome (MDS). Close scrutiny of the literature confirmed this observation and we estimated RAEB with t(6;9) to be third in rank after AML-M2 and M4 with this translocation (11,13,14,16,30,35). Patient 3 presented with acute myelofibrosis (AMF) which eventually evolved into AML-M4. One of our 17 patients was diagnosed as a Ph-, BCR- CML. CML with t(6;9) was reported once before by Fleischman (31). However, in some cases differentiation between CML and MDS/AML may be difficult. Whether t(6;9) truly may occur in CML or whether the hematologic subclassification is incorrect, is not clear (37). The molecular data are strongly in favour of the second possibility.

Pearson et al proposed that a correlation might exist between the t(6;9) and basophilia, though this has been contradicted by others (10,11,15). In our patient-group this phenomenon was reported in two patients with AML-M2 and in one with AML-M4. However, to establish the real incidence of basophilia in this specific patient-group, standardized prospective studies are needed since detection of this feature involves careful

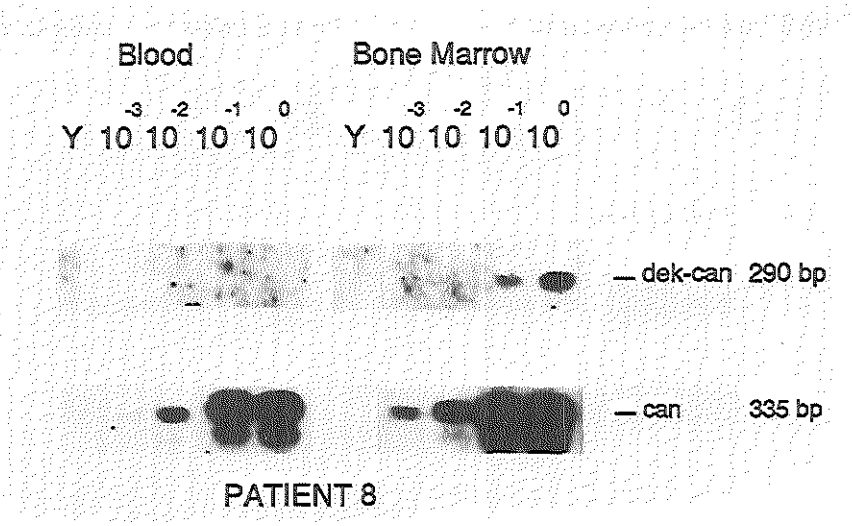


Figure 6. Results of the PCR in patient 8. RNA was isolated from blood and bone marrow after a second chemotherapy course. 1 μg total RNA of patient 8 was serially 10-fold diluted in 10 μg yeast RNA. $10^0 = 1$ μg patient RNA in 10 μg yeast RNA, $10^{-1} = 0.1$ μg patient RNA in 10 μg yeast RNA, etc. cDNA-preparation and PCR were performed according to the protocol using primer set I. The 335 basepair *can* fragment which is generated by the PCR is detectable both in the blood and the bone marrow of this patient. A 290 bp chimeric fragment is visible in the first two dilution steps of the bone marrow sample indicating that 1/10x1/10 = 1/100 cells contain the t(6;9). Y = 10 μg of yeast RNA which is used as a negative control.

examination of marrow morphology.

To date, the immunophenotype of blast cells has only been studied in two patients and this showed HLA-DR/TdT/CD13-positivity in both cases (21). It would be interesting to investigate a larger group of patients with t(6;9) for TdT-positivity since its presence in leukemias of myeloid origin indicates an unfavorable prognosis.

One of our patients (patient 8) presented with a variant of Sweet's syndrome, a paraneoplastic syndrome characterized by high fever, cutaneous lesions and granulocytosis (26). At diagnosis, his blood contained mainly mature granulocytes. Southern blot analysis of the blood and bone marrow cells of this patient showed germ line and aberrant fragments of similar intensity (Figure 3, patient 8). This indicated that the *dek-can* rearrangement was present in the majority of cells investigated. Therefore, we assume that cells which appear as "mature" granulocytes in the blood of this patient at diagnosis carry the *dek-can* rearrangement. This observation gives support to the finding of others that leukemic cells can differentiate into polymorphonuclear leucocytes (38,39).

Karyotyping usually shows a simple t(6;9)(p23;q34) at the time of diagnosis. It is noteworthy that in patients 5 and 6 a 6p- derivative was seen in relapse while the 9q+ was lost. This is a cytogenetic indication that the 6p- derivative that carries the *dek-can* fusion gene is important in this type of leukemia. Additional aberrations are rare but may occur, especially during progression of the leukemia (9-18). The most frequently observed extra cytogenetic abnormalities are trisomy 8 and trisomy 13. Trisomy 8 is often seen during progression of myeloid leukemias and is found in many cases of AML and CML in accelerated or blastic phase (40). Trisomy 13 is a rare event in leukemias but has been reported to occur as a sole aberration in AUL, AML and RAEB (41-44). The finding of a trisomy 8 or a trisomy 13 in any leukemia is an ominous sign and heralds an unfavorable outcome of the disease. One of the patients studied here (patient 13) showed deletion of the long arm of chromosome 7 in addition to a t(6;9). This abnormality is associated with the so called secondary AML or MDS and also predicts an unfavorable prognosis (45). To what extent all these secondary cytogenetic changes contribute to the overall prognosis of patients with a t(6;9) is unknown at present.

The discovery of a translocation (6;9) in metaphases of a patient with leukemia is alarming since such patients respond poorly to therapy: in our group of patients only half of the recorded cases achieved a complete remission after therapy which is in concordance with the data in the literature. This forms a sharp contrast to a comparable age-group of patients with AML in whom a 77% complete remission rate can be achieved with chemotherapy (46,47). Mostly, survival does not exceed 3 years and in the group we analyzed only one case, patient 10, survived 3 years after a bone marrow transplantation and is still in remission. Correct diagnosis of t(6;9) is of utmost importance in trying to improve the

prognosis for these patients in future. Since the reciprocal translocation involves small chromosomal fragments of similar morphology, cytogenetic diagnosis of this disease can be difficult. Initially, von Lindern et al (19) showed that four out of four t(6;9) patients contained breakpoints in both icb-6 and icb-9. Since icb-6 and icb-9 represent introns in the *dek*- and *can*-genes respectively, the translocation apparently fuses the same *dek* 'donor'-exon to the same *can* 'acceptor'-exon, resulting in the formation of uniform *dek-can* fusion genes on the 6p- chromosome of these patients. The data presented in this paper corroborate and further extend the initial observation: the translocation is amazingly precise and highly consistent in 17 t(6;9) patients analyzed at the molecular level. Since the total target size for the translocation per haploid genome amounts to less than 20 kb of DNA (icb-6 and icb-9), this may well explain the low incidence of the translocation in AML (0.5%-4%). Standardized Southern blot analysis, using restriction enzymes EcoRV, HindIII and BamHI in combination with the limited number of four probes, is a reliable method for diagnosis of t(6;9) and is therefore clinically applicable.

Due to the limited amount of material, intact RNA could only be analyzed in 14 cases (table 1). In 10 out of 14 cases the PCR generated a *dek-can* chimeric fragment using one primer set and confirmed the results obtained by Southern blot analysis. Intact RNA of 4 patients neither yielded *can*- nor *dek-can* fragments after amplification. Probably this is due to technical insufficiency, since inhibition of the Taq polymerase by substances in blood has been described (48). (Moreover, Northern blot analysis of patient 2 showed a *dek-can* mRNA, see ref. 19). We infer from our results that the chimeric gene is expressed in the leukemic cells of all the t(6;9)-patients that could be analyzed. Moreover, the chimeric fragment is detected in all these patients using the same *dek-can* oligomer, indicating that it originates from the same exon fusion product in these 10 cases.

The uniform findings at the DNA- and RNA-level in 17 patients with t(6;9) indicate that this translocation is highly suitable for molecular detection by Southern blotting and reversed PCR. The latter technique allows sensitive detection of residual leukemic cells after chemotherapy and after bone marrow transplantation through monitoring the presence of the *dek-can* mRNA. The consistent finding of the chimeric product in patients with a t(6;9) also strongly argues for a distinct causative role of the *dek-can* fusion gene in this myeloproliferative disorder. Occurrence of this translocation in various subtypes of AML (M1,2,4,AMF) and the observations in patient 8 indicate that the *dek-can* gene product does not cause maturation block. Whether once taken place the translocation invariably leads to leukemia remains to be established.

Acknowledgements
The authors are grateful to the following physicians and cytogeneticists who so

generously sent material and provided clinical data: Dr M. Lafage (patients 6 and 17), Dr C. Bastard (patient 9), Dr E. van Wering, Dr A. van der Does, Stichting Nederlandse Werkgroep Leukemie bij Kinderen (SNWLK) and Dr A. Veerman (patient 10), Dr R. Slater and Dr H. Behrendt (patient 11), Dr R. Slater and Dr M. van Oers (patient 15), Dr B. Young and Dr A. Lister (patients 12 and 13), Dr P. Ambros (patient 14). The secretarial assistance of R. Boucke and the photographic work of T. de Vries Lentsch are gratefully acknowledged. We thank Prof.Dr. D. Bootsma for continuous support.

REFERENCES

1. Nowell PC, Hungerford DA: A minute chromosome in human granulocytic leukemia. Science 132:1497, 1960
2. Trent JM, Kaneko Y, Mitelman F: Report of the committee on structural chromosome changes in neoplasia. Cytogenet Cell Genet 51:533, 1989
3. Dalla-Favera R, Martinotti S, Gallo RC, Erikson J, Croce CM: Translocation and rearrangements of the c-*myc* oncogene locus in human undifferentiated B-cell lymphomas. Science 219:963, 1983
4. Heisterkamp N, Stephenson R, Groffen J, Hansen PF, de Klein A, Bartram CR, Grosveld G: Localization of the c-*abl* oncogene adjacent to a translocation breakpoint in chronic myelocytic leukaemia. Nature 306:239, 1983
5. Groffen J, Stephenson JR, Heisterkamp N, de Klein A, Bartram CR, Grosveld G: Philadelphia chromosomal breakpoints are clustered within a limited region, bcr, on chromosome 22. Cell 36:93, 1984
6. Rowley JD, Potter D: Chromosomal banding patterns in acute nonlymphocytic leukemia. Blood 47:705, 1976
7. Schwartz S, Jiji R, Kerman S, Meekins J, Cohen MM: Translocation (6;9) (p23;q34) in acute nonlymphocytic leukemia. Cancer Genet Cytogenet 10:133, 1983
8. Sandberg AA, Morgan R, McCallister JA, Kaiser-McCaw B, Hecht F: Acute myeloblastic leukemia (AML) with t(6;9)(p23;q34): A specific subgroup of AML? Cancer Genet Cytogenet 10:139, 1983
9. Carroll AJ, Castleberry RP, Prchal JT, Finley WH: Translocation (6;9)(p23;q34) in acute nonlymphocytic leukemia: three new cases. Cancer Genet Cytogenet 18:303, 1985
10. Bernstein R, Cain M, Cho J: Loss of the Y chromosome associated with translocation t(6;9)(p23;q34) in a patient with acute nonlymphocytic leukemia. Cancer Genet Cytogenet 39:81, 1989
11. Pearson MG, Vardiman JW, Le Beau MM,Rowley JD, Schwartz S, Kerman SL, Cohen MM, Fleischman EW, Prigogina EL: Increased numbers of marrow basophils may be associated with a t(6;9) in ANLL. American Journal of Hematology 18:393, 1985
12. Horsman DE, Kalousek DK: Acute myelomonocytic leukemia (AML-M4) and translocation t(6;9)(p23;q34): Two additional patients with prominent myelodysplasia. Am J Hemat 26:77, 1987
13. Heim S, Kristoffersson U, Mandahl N, Mitelman F, Bekassy AN, Garwicz S, Wiebe T: High resolution banding analysis of the reciprocal translocation t(6;9) in acute nonlymphocytic leukemia. Cancer Genet Cytogenet 22:195, 1986
14. Gold EJ, Conjalka M, Pelus LM, Jhanwar SC, Broxmeyer H, Middleton AB, Clarkson BD, Moore MAS: Marrow cytogenetic and cell-culture analyses of the myelodysplastic syndromes: Insights to pathophysiology and prognosis. J Clin Oncol 1:627, 1983
15. Fan Y, Sait SNJ, Raza A, Rowe JM, D'Arrigo P, Sandberg AA: Translocation t(6;9)(p23;q34) in acute nonlymphocytic leukemia. Two new patients without increased bone marrow basophils. Cancer Genet Cytogenet 32:153, 1988
16. Fonatsch C, Stollmann B, Holldack J, Engert A: Translocation (6;9)(p23;q34) in smoldering leukemia and acute nonlymphocytic leukemia. Cancer Genet Cytogenet 26: 363, 1987
17. Levin M, Le Coniat M, Bernheim A, Berger R: Complex chromosomal abnormalities in

acute nonlymphocytic leukemia. Cancer Genet Cytogenet 22:113, 1986
18. Stejskalova E, Goetz P, Stary J: A translocation (6;9)(p32;q34) and trisomy 13 in a case of childhood acute nonlymphocytic leukemia. Cancer Genet Cytogenet 45:139, 1990
19. von Lindern M, Poustka A, Lerach H, Grosveld G: The (6;9) chromosome translocation, associated with a specific subtype of acute nonlymphocytic leukemia, leads to aberrant transcription of a target gene on 9q34. Mol Cell Biol 10:4016, 1990
20. Cuneo A, Kerim S, Vandenberghe E, Van Orshoven A, Rodhain J, Bosly A, Zachee P, Louwagie A, Michaux J, Dal Cin P, Van Den Berghe H: Translocation t(6;9) occurring in acute myelofibrosis, myelodysplastic syndrome, and acute nonlymphocytic leukemia suggests multipotent stem cell involvement. Cancer Genet Cytogenet 42:209, 1989
21. Adriaansen HJ, van Dongen JJM, Hooijkaas H, Hahlen K, van 't Veer MB, Löwenberg B, Hagemeijer A: Translocation (6;9) may be associated with a specific TdT-positive immunological phenotype in ANLL. Leukemia 2:136, 1988
22. Maniatis TE, Fritsch EF, Sambrook J: Molecular cloning: a laboratory manual, Cold Spring Harbor, NY: Cold Spring Harbor Laboratory, 1982
23. von Lindern M, van Agthoven T, Hagemeijer A, Adriaansen H, Grosveld G: The human pim-1 gene is not directly activated by the translocation (6;9) in acute nonlymphocytic leukemia. Oncogene 4:75, 1989
24. Hermans A, Gow J, Selleri L, von Lindern M, Hagemeijer A, Wiedemann LM, Grosveld G: bcr-abl oncogene activation in Philadelphia chromosome-positive acute lymphoblastic leukemia. Leukemia 2:628, 1988
25. Soekarman D, van Denderen J, Hoefsloot L, Moret M, Meeuwsen T, van Baal J, Hagemeijer A, Grosveld G: A novel variant of the *bcr-abl* fusion product in Philadelphia chromosome positive acute lymphoblastic leukemia. Leukemia 4:397, 1990
26. Cooper PH, Innes DJ, Greer KE: Acute febrile neutrophilic dermatosis (Sweet's syndrome) and myeloproliferative disorders. Cancer 51:1518, 1983
27. Bennett JM, Catovsky D, Daniel MT, Flandrin G, Galton DAG, Gralnick HR, Sultan C: Proposed revised criteria for the classification of acute myeloid leukemia. Ann Int Med 103:620, 1985
28. Bitter MA, Le Beau MM, Rowley JD, Larson RA, Golomb HM, Vardiman JW: Associations between morphology, karyotype, and clinical features in myeloid leukemias. Human Pathology 18:211, 1987
29. Arthur DC, Berger R, Golomb HM, Swansbury GJ, Reeves BR, Alimena G, Van Den Berghe H, Bloomfield CD, de la Chapelle A, Dewald GW, Garson OM, Hagemeijer A, Kaneko Y, Mitelman F, Pierre RV, Ruutu T, Sakurai M, Lawler SD, Rowley JD: The clinical significance of karyotype in acute myelogenous leukemia. Cancer Genet Cytogenet 40:203, 1989
30. Vermaelen K, Michaux J, Louwagie A, Van Den Berghe H: Reciprocal translocation t(6;9)(p21;q33): A new characteristic chromosome anomaly in myeloid leukemias. Cancer Genet Cytogenet 10:125, 1983
31. Fleischman EW, Prigogina EL, Iljinskaja GW, Konstantinova LN, Puchkova GP, Volkova MA, Frenkel MA, Balakirev SA: Chromosomal rearrangements with a common breakpoint at 6p23 in five cases of myeloid leukemia. Hum Genet 64:254, 1983
32. Fitzgerald PH, Morris CM, Fraser GJ, Giles LM, Hamer JW, Heaton DC, Beard MEJ: Nonrandom cytogenetic changes in New Zealand patients with acute myeloid leukemia. Cancer Genet Cytogenet 8:51, 1983
33. Yunis JJ: Recurrent chromosomal defects are found in most patients with acute nonlymphocytic leukemia. Cancer Genet Cytogenet 11:125, 1984
34. Yunis JJ, Brunning RD, Howe RB, Lobell M: High-resolution chromosomes as an independent prognostic indicator in adult acute nonlymphocytic leukemia. The New Engl J Med 311:812, 1984
35. Fan Y, Raza A, Schumer J, Sait SNJ, Block AW, Snyderman M, Sandberg AA: Translocation t(6;9)(p22.3;q34) in myelodysplastic syndrome-refractory anemia with excess blasts. Cancer Genet Cytogenet 29:135, 1987
36. Nordenson I, Bjermer L, Holmgren G, Hornsten P, Wahlin A: t(6;9) in bone marrow cells in two patients with sarcoidosis and acute myeloid leukemia. Cancer Genet Cytogenet

38:297, 1989
37. Martiat P, Michaux JL, Rodhain J: Philadelphia-negative (Ph-) chronic myeloid leukemia (CML): Comparison with Ph+ CML and chronic myelomonocytic leukemia. Blood 78:205, 1991
38. Fearon ER, Burke PJ, Schiffer CA, Zehnbauer BA, Vogelstein B: Differentiation of leukemia cells to polymorphonuclear leukocytes in patients with acute nonlymphocytic leukemia. N Engl J Med 315:15, 1986
39. Fialkow PJ, Singer JW, Raskind WH, Adamson JW, Jacobson RJ, Bernstein ID, Dow LW, Najfeld V, Veith R: Clonal development, stem-cell differentiation, and clinical remissions in acute nonlymphocytic leukemia. N Engl J Med 317:468, 1987
40. Heim S, Mitelman F: Numerical chromosome aberrations in human neoplasia. Cancer Genet Cytogenet 22:99, 1986
41. Dohner H, Arthur DC, Ball ED, Sobol RE, Davey FR, Lawrence D, Gordon L, Patil SR, Surana RB, Testa JR, Verma RS, Schiffer CA, Wurster-Hill DH, Bloomfield CD: Trisomy 13: A new recurring chromosome abnormality in acute leukemia. Blood 76:1614, 1990
42. Sreekantaiah C, Baer MR, Morgan S, Isaacs JD, Miller KB, Sandberg AA: Trisomy/tetrasomy 13 in seven cases of acute leukemia. Leukemia 4:781, 1990
43. Limon J, Baran W, Babinska M, Hellmann A: Trisomy 13 in a case of myelodysplastic syndrome.
44. Beverstock GC, Ploem JE, Wessels H, v.d. Keur D, Mollevanger P: Trisomy 13 and myelodysplastic syndrome. Cancer Genet Cytogenet 48:179, 1990
45. Golomb HM, Alimena G, Rowley JD, Vardiman JW, Testa JR, Sovik C: Correlation of occupation and karyotype in adults with acute nonlymphocytic leukemia. Blood 60:404, 1982
46. Schiffer CA, Lee EJ, Tomiyasu T, Wiernik PH, Testa JR: Prognostic impact of cytogenetic abnormalities in patients with de novo acute nonlymphocytic leukemia. Blood 73:263, 1989
47. Gale RP: Acute myelogenous leukemia: Progress and controversies. Leukemia 4:529, 1990
48. Panaccio M, Lew A: PCR based diagnosis in the presence of 8% (v/v) blood. Nucl Acid Res 19:

CHAPTER 2.6

Can, a putative oncogene associated with myeloid leukemogenesis, may be activated by fusion of its 3' half to different genes: characterization of the *set* gene

Marieke von Lindern[1], Sjozèf van Baal[1], Joop Wiegant[2], Anton Raap[2], Anne Hagemeijer[1] and Gerard Grosveld[1].

[1]Dept. of Cell Biology and Genetics, Erasmus University, Rotterdam, The Netherlands.
[2]Dept. of Cytochemistry and Cytometry, University of Leiden, Leiden, The Netherlands.

Submitted for publication

ABSTRACT

The translocation (6;9)(p23;q34) in acute nonlymphocytic leukemia results in the formation of a highly consistent *dek-can* fusion gene. Translocation breakpoints invariably occur in single introns of *dek* and *can*, that were named *icb*-6 and *icb*-9, respectively. In a case of acute undifferentiated leukemia, a breakpoint was detected in *icb*-9 of *can* while no breakpoint could be detected in *dek*. Genomic and cDNA cloning showed that instead of *dek*, a different gene was fused to *can*, which was named *set*. *Set* encodes transcripts of 2.0 and 2.7 kb that result from the use of alternative polyadenylation sites. Both transcripts contain the ORF for a putative SET protein with a predicted molecular mass of 32 kD. The *set-can* fusion gene is transcribed into a 5 kb transcript that contains a single open reading frame predicting a 155 kD chimeric SET-CAN protein. The SET sequence shows no consistent homology with any known proteins. The only common sequence motif of SET and DEK proteins is an acidic region. SET has a long acidic tail, of which a large part is present in the predicted SET-CAN fusion protein. The *set* gene is located on chromosome 9q34, centromeric of c-*abl*. Since a *dek-can* fusion-gene is present in t(6;9) AML and a *set-can* fusion-gene was found in a case of AUL, we assume that *can* may function as an oncogene activated by fusion of its 3' part to either *dek*, *set* or perhaps even other genes.

INTRODUCTION

Translocations are the best studied non-random chromosomal aberrations associated with specific subtypes of leukemia. Due to a translocation, an oncogene can be activated through alterations in regulatory DNA sequences that leave the encoded protein intact (e.g. *myc*) or through formation of a fusion gene, encoding a chimeric protein (e.g. *bcr-abl*). The t(9;22) associated with chronic myeloid leukemia

(CML), acute myeloid leukemia (AML) and acute lymphoblastic leukemia (ALL) (27) results in the expression of a chimeric BCR-ABL protein with enhanced tyrosine kinase activity (15, 18, 25, 36, 42). Pendergast et al. showed that defined sequences encoded by the first exon of *bcr* interact with the SH2 domain of ABL (31). This interaction is essential for the activation of the ABL tyrosine kinase activity and for transforming capacity of BCR-ABL. More recently, other fusion genes have been isolated. The t(1;19), occurring in childhood pre-B cell acute leukemia, fuses the *E2a* gene, encoding transcription factors E12/E47, to a novel homeobox gene *PBX1* (24, 30). The t(15;17), strongly associated with acute promyelocytic leukemia (APL), fuses part of the retinoic acid receptor type α (*RARα*) to a novel gene on chromosome 15 named *PML*, which is predicted to be a transcription factor (8, 23). *Bcr-abl*, *E2A-pbx* and *pml-RARα* seem to be highly consistent partners.

Previously we reported the cloning of t(6;9) breakpoints (40). T(6;9) is the hallmark of a specific subtype of AML characterized by a poor prognosis and a young age of onset. It is FAB classified mostly as M2/M4 and rarely as M1 or RAEB (1, 34, 37). On chromosome 9 breakpoints take place in a specific intron, *icb*-9, of a large gene (>140 kb) named *Cain (can)* (40). On chromosome 6 breakpoints also occur in a single intron, *icb*-6, of a gene named *dek* (39). The result of t(6;9) is the formation of a *dek-can* fusion gene on chromosome 6p-, which is transcribed into an invariable, 5.5 kb, leukemia specific *dek-can* mRNA (37). The fusion transcript encodes a 165 kD chimeric protein, which derives from the in frame fusion of *dek* and *can* open reading frames. Sequence comparison of DEK and CAN with the EMBL data base shows no homology to any known protein sequences. CAN contains several putative dimerization motives and the C-terminal part may function as an ancillary DNA binding domain. The DEK protein contains 43% of charged amino acids and several acidic domains (39).

Surprisingly, a breakpoint in *icb*-9 of *can* was also detected in the bone marrow sample from a patient with acute undifferentiated leukemia (AUL) and an apparently normal karyotype. No breakpoint in *dek* could be detected in this case (39). In this paper we report the isolation and characterization of a novel gene named *set*, that was fused to *can* in the leukemic cells of this patient. A chimeric *set-can* transcript was detected whose sequence predicts a SET-CAN protein of 155 kD.

MATERIALS AND METHODS

Southern and Northern blotting. Patient material and human cell lines used were described previously (40). T(9;22) hybrid cell lines used were described (7, 14). DNAs used in the Zoo-blot were derived from human thymus, NIH3T3 cells (mouse), a *Potorous tridactylis* cell line (marsupial), chicken liver, *Aulonocara stuartgranti* liver (fish), *Xenopus laevis* liver, DM2 cell line (*D. melanogaster*) and *S. cerevisiae* strain

GRF18. High molecular weight DNA was prepared as described by Jeffreys et al. (21). KG1 DNA was isolated and digested in agarose blocks as described previously (41). Restriction enzyme reactions were performed as recommended by the suppliers. Southern blots were done as described by Sambrook et al, (33). RNA of mouse tissue was isolated from BCBA mice. RNA was isolated by either the guanidinium isothiocyanate (6) or LiCl-method (3). Total RNA was electrophoresed and blotted as described by Fourney et al. (11). Equal amounts of ribosomal RNA were loaded: before loading the samples on a denaturing gel, 5% of each sample was loaded on a nondenaturing agarose gel to estimate the amount of ribosomal RNA and to adapt the sample quantity if necessary. Southern and Northern blots were hybridized in 10% dextran (33). Northern blots of mouse tissues were hybridized with human *set* probes at 65°C, and filters were washed with 0.3xSSC at 65°C. Probes were labelled by the method of Feinberg and Vogelstein (10).

Genomic and cDNA libraries. A λEMBL3 library of bone marrow DNA of patient SE was constructed according to Frischauf et al. (12) and $2x10^6$ pfu were screened with the *can* probes Al1F6E.3 and AL1F4EP (Figure 1). A human testis cDNA library in λgt11 was purchased from Clonetech (Palo Alto, Calif.).

Cloning chimeric *set-can* cDNA. 50 μg total RNA of bone marrow cells from AUL patient SE was heat-denatured and first strand cDNA was synthesized with Avian RT and 100 pmol of a 20-mer primer: 5'CCACCAGGTGATTCAGCCTC3', located 200 bp downstream of the translocation breakpoint in the *can* cDNA (13, 19). cDNA was size selected on a Sephacryl S-1000 column (≥800 bp). Terminal deoxynucleotidyl transferase (TdT) was used to tail first strand cDNA with deoxyadenosine. Following heat denaturation, second strand cDNA was synthesized with Klenov-polymerase and the 35-mer 5'GTCGCGAATTCGTCGACGCGTTTTTTTTTTTTTTT3'. Part of the double stranded cDNA (1/20) was amplified with Taq polymerase (Perkin Elmer Cetus) and the primers 5'GTCGCGAATTCGTCGACGCG3' and 5'TTTGAATTCGTCGACCAGATGCTGA-TCCCACTCC3'. The latter primer contains *Sal*I and *Eco*RI recognition sites fused to a 20-mer sequence located 86 bp downstream of the translocation breakpoint in the *can* cDNA sequence. No PCR product of a specific size was generated but a smear of PCR fragments hybridized to *can* cDNA probe hXT37BR derived from *can* cDNA, covering sequences 3' of the translocation breakpoint. Although this probe overlaps with the amplification primers, conditions were such that no hybridization to the primers was allowed. Blots were washed at 65°C with 0.1x SSC. DNAs larger than 800 bp were isolated from an agarose gel, reamplified, cut with *Eco*RI and cloned into λgt10. The resulting library was screened with *can* probe hXT37BR.

Cloning the 3' end of the 2 kb *set* mRNA. 30 μg total RNA of bone marrow cells from AUL patient SE was heat-denatured and first strand cDNA was synthesized with Avian RT, using 100 pmol of the 35-mer 5'GTCGCGAATTCGTCGACGCGTTTTTTTTTTTTT-TTT3' as a primer (13, 19). Excess of primer was removed by isopropanol precipitation. 1/100 of the cDNA reaction was amplified using Taq polymerase (Perkin Elmer Cetus) and the primers 5'GTCGCGAATTCGTCGACGCG3' and either 5'GTTTGGGTGGGTTAGTGGGC3' or 5'CCACTCAATGGGAGAATCAGC3'. The latter primers cover nt 1372-1391 and nt 1519-1539 of the *set* cDNA respectively. Amplified fragments of 350 and 220 bp were sequenced using a protocol for direct sequencing of fragments produced by an asymmetric PCR (22).

Sequence determination and analysis. Restriction fragments of cDNA clones were subcloned in M13. Overlapping cDNA sequences were determined on both strands by dideoxy sequencing (35). Initially M13 primers were used; when no suitable restriction sites were present a primer was generated based on the already available cDNA sequence. To establish intron-exon borders, genomic fragments containing the exon of interest were subcloned into M13 and a cDNA primer near the putative intron-exon

border was generated to prime the sequence reaction. Sequences were analyzed with the computerprogram Microgenie (Beckman) and the EMBL database was used to search for homologous sequences both at the nucleotide and amino acid level.

In situ **hybridization** was performed as described by Arnoldus et al. (2).

RESULTS

Previously it was shown that *can* probe AI1F4EP, located at the 5' end of *icb*-9 (Figure 1), hybridized to an aberrant fragment in bone marrow DNA of AUL patient SE (40). On a Northern blot, 3' *can* cDNA probes hybridized to an aberrant 5 kb transcript in total RNA of bone marrow cells of patient SE (39). These data proved that a breakpoint is present in

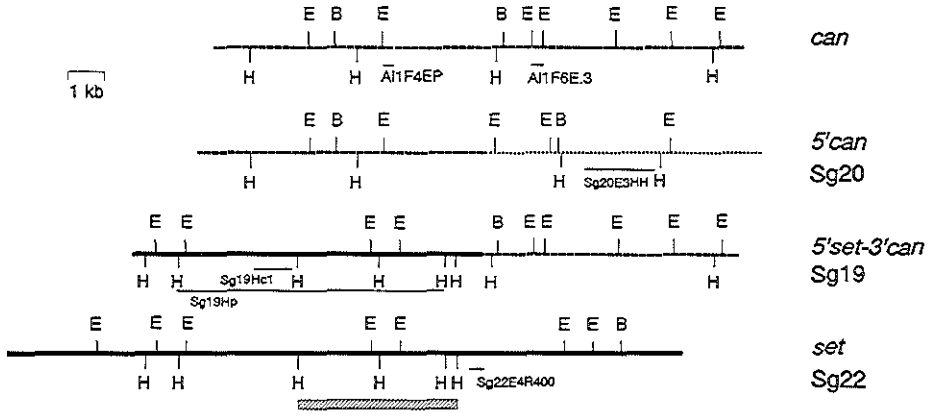


Figure 1. Restriction maps of *can, set-can* and *set*. A bold dotted line indicates *can* sequences, a solid line *set* sequences. A thin dotted line indicates sequences fused to *can,* whose normal localization is unknown. The top line shows a region of *can* around *icb*-9. The single copy probes AI1F4EP and AI1F6E.3 are located at either side of the translocation breakpoint of patient SE. The second line represents phage Sg20, isolated with probe AI1F4EP. Sg20 contains 5' *can* sequences and novel sequences 3' of the translocation breakpoint. A 1.9 kb *Hin*cII-*Hin*dIII fragment Sg20E3HH could be used as single copy probe. The third line represents phage Sg19, isolated with probe AI1F6E.3. Sg19 contains 3' *can* sequences fused to novel sequences 5' of the translocation breakpoint, that were shown to represent the *set* gene. A 1 kb *Hin*cII fragment (Sg19Hc1) was isolated as single copy probe. The genomic fragment Sg19Hp was used for *in situ* hybridization of chromosome preparations. The fourth line represents phage Sg22, which contains *wt set* and was isolated with probe Sg19Hc1. The striped bar indicates the region of *set* homologous to multiple copies in the genome. Sg22E4R400 is a 400 bp *Rsa*I fragment isolated as an additional single copy probe just 5' of the translocation breakpoint. The scale is indicated by a bar. E: *Eco*RI, B: *Bam*HI, H: *Hin*dIII.

icb-9 of the *can* gene. However, the 5 kb aberrant transcript failed to hybridize to 5' *dek* cDNA probes that invariably detected the leukemia specific *dek-can* transcript of similar size in t(6;9) AML cells. This suggested to us that instead of *dek* another gene may be fused to *can* in leukemic cells of patient SE.

Isolation of the *set* gene. To isolate and characterize DNA sequences fused to *can* in AUL patient SE, a genomic λEMBL3 phage library was constructed of high molecular weight DNA of leukemic cells of this patient. In total $2x10^6$ pfu were screened with *can* probes AI1F4EP and AI1F6E.3, located at either side of the translocation breakpoint (Figure 1). Hybridizing clones (12) were analyzed by restriction enzyme mapping. Most clones contained DNA from the normal *can* allele, but clones Sg19 and Sg20 (hybridizing to AI1F6E.3 and AI1F4EP respectively) contained *can* sequences fused to DNA with an unknown restriction enzyme pattern. Subclones of phages Sg19 and Sg20 were tested for DNA fragments not containing any repetitive sequences (not shown). A 1 kb *Hin*cII fragment from Sg19 (Sg19Hc1) and a 1.9 kb *Hin*dIII-*Hin*cII fragment from Sg20 (Sg20E3HH)(Figure 1) could be used as single copy probes to screen the genomic library again in order to obtain non-rearranged clones overlapping with Sg19 and Sg20. A total of 10 phages was isolated and analyzed, 2 hybridizing to Sg19Hc1 and 8 hybridizing to Sg20E3HH. Phage Sg22 hybridized to probe Sg19Hc1 located at the 5' side of the breakpoint and extended beyond the breakpoint at the 3' end (Figure 1). However, hybridization experiments failed to detect overlap between Sg22 and Sg20 , which contains DNA juxtaposed to *can* at the 3' end of the breakpoint. Similarly, no overlap could be detected between Sg19 and phages hybridizing to probe Sg20E3HH that extend at the 5'end beyond the breakpoint present in Sg20 (results not shown). Therefore it was concluded that a stretch of DNA flanking the

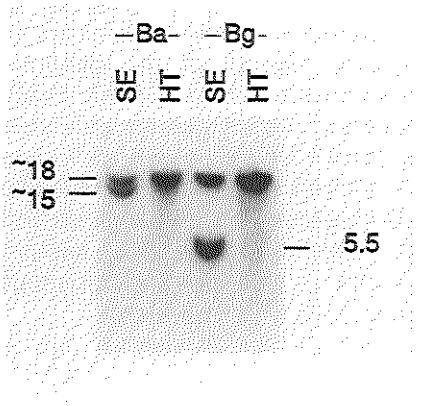


Figure 2. Southern blot containing DNA isolated of bone marrow cells of patient SE (lanes 1 and 3) and of human thymus (lanes 2 and 4), digested with the restriction enzymes *Bam*HI (lanes 1 and 2) and *Bgl*I (lanes 3 and 4). The blot was hybridized to probe Sg22E4R400 (indicated in Figure 1).

translocation breakpoint was deleted. The size of this deletion measures at minimum 6 kb and may be much larger. As 3', but not 5', *can* cDNA probes detected an aberrant transcript in RNA of leukemic cells of patient SE, attention was focused on the newly isolated DNA fragments 5' of the translocation breakpoint fused to 3' *can*. We reasoned that these sequences could be part of a novel fusion gene associated with this case of AUL.

To prove that Sg19 indeed represents the translocation breakpoint, probe Sg19Hc1 was hybridized to a Southern blot containing DNA isolated of bone marrow cells from patient SE. In addition to a large ~18 kb *Bam*HI fragment, a fragment of ~15 kb was detected specific for the AUL sample (not shown). A second single copy probe could be isolated which maps just 5' of the translocation breakpoint (Sg22E4R400; Figure 1). This probe hybridized to the same *Bam*HI fragments, and to a (normal) large ~18 kb *Bgl*II fragment and an aberrant 5.5 kb *Bgl*II fragment in DNA of leukemic cells of patient SE (Figure 2). Although the large fragments in the Southern blot are difficult to size, comparison of the restriction maps of Sg19 and Sg22 shows that the size difference of the *Bam*HI fragments should be 3 kb and the size of the *set Bam*HI fragment at minimum 18 kb.

Isolation of *set* cDNA. To isolate cDNA sequences representing the aberrant transcript detected in RNA of patient SE, we used an anchored PCR reaction according to the protocol for rapid amplification of cDNA ends (RACE) (13). The PCR reaction was primed with a *can* oligonucleotide located 86 nt downstream of the translocation breakpoint in *dek-can* cDNA. Hybridization of the PCR products to a *can* cDNA probe just 3' of the breakpoint, hXT37BR (figure 3), showed a smear of fragments. Therefore size selected DNA (≥800 bp) was isolated from a preparative agarose gel and cloned into λgt10 to generate a small selective cDNA library. Many phages hybridized to *can* cDNA probe hXT37BR and six were grown and analyzed. One clone was only weakly hybridizing, three clones contained *wt can* sequences and two clones (SE3 and SE4, 300 and 700 bp in size respectively) contained 86 bp of *can* sequences linked to unknown DNA sequences. The 5' 500 bp of clone SE4 (an *Eco*RI-*Rsa*I fragment) were isolated and were hybridized to a Southern blot containing restriction enzyme digests of the genomic phages Sg19, Sg20 and Sg22. Strong hybridization to a 5.5 kb *Eco*RI fragment in phages Sg19 and Sg22 was detected (schematically indicated in figure 3). On a Northern blot containing total RNA of several cell lines (K562, Hela, Daudi, HL60, KG1) and AUL patient SE, the SE4ER probe hybridized to two transcripts of 2.7 and 2.0 kb present in all cell lines and to additional transcripts of 5 and 6.5 kb, specific for the AUL sample (Figure 4). The 5 kb transcript is identical in size to the aberrant transcript detected by 3' *can* probes in RNA of leukemic cells of this patients.

Together the data prove that in AUL patient SE the 3' part of *can* is

fused to a novel gene, distinct from the previously isolated *dek* gene. This novel gene was named *set*.

To isolate full length *set* cDNA, probe SE4ER was used to screen a λgt11 testis library. Six hybridizing clones were isolated and analyzed. The overlapping cDNAs SE9 and SE10 are shown in figure 3, SE10 extends most 5', SE9 most 3'. Different subfragments of clones SE9 and SE10 were hybridized to a Southern blot containing restriction enzyme digests of the genomic phages Sg19, Sg20 and Sg22. The 5' part of SE10 hybridized to the same 5.5 kb *Eco*RI fragment that hybridized to the fusion cDNA clone SE4. Surprisingly, the 3' part of SE10 and almost the entire clone SE9 hybridized to a 1.1 kb *Eco*RI-*Hin*dIII fragment in Sg19 that is situated 3' of the 5.5 kb *Eco*RI fragment, but 5' of the translocation breakpoint (Figure 3). Fine mapping of cDNA clone SE9 and

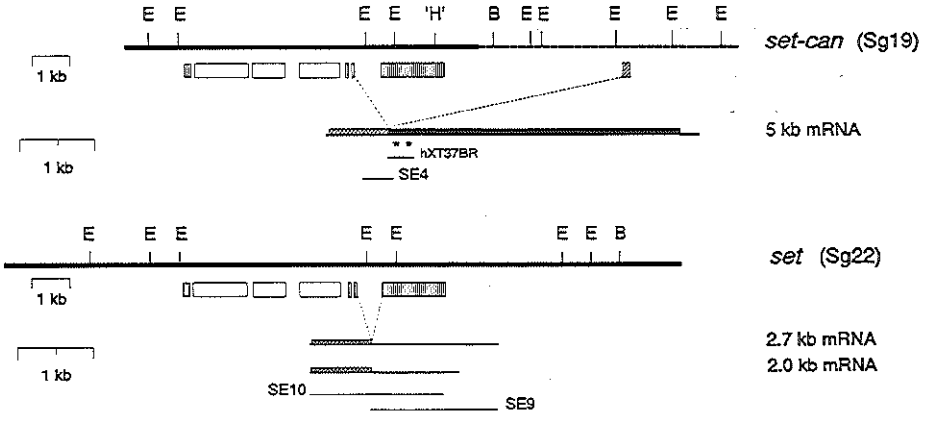


Figure 3. The *set* gene. The top line represents part of the *set-can* gene. The second restriction map represents the *wt set* gene. A striped line indicates *can* sequences, a solid line *set* sequences. Open boxes below the maps indicate restriction enzyme fragments hybridizing to cDNA sequences. Stippled boxes are defined *set* exons, the striped box is a defined *can* exon. Below the *set-can* map the 5 kb fusion transcript is indicated. The *can* ORF is indicated by a solid bar, the *set* ORF by a hatched bar. SE4 is the fusion cDNA clone, isolated via the RACE protocol and hybridizing to *can* cDNA probe hXT37BR. Asterixes indicate the positions of the oligonucleotides used to amplify the fusion cDNAs. A stippled line indicates which *set* exon is spliced to the *can* exon 3' of *icb*-9. The last 3' *set* exon is alternatively spliced and not present in the 5 kb fusion transcript. Below the *set* map the 2.0 and 2.7 kb transcripts are indicated. They differ only in the polyadenylation signal used. The stippled line indicates the exon spliced to a *can* exon in the *set-can* gene and to the 3' last exon in the *wt set* gene. SE9 and SE10 are cDNA clones isolated from a λgt11 human testis library. Scales of genomic and cDNA representations are indicated in kb. All *Eco*RI (E) and *Bam*HI sites are indicated. 'H' is only one of the *Hin*dIII sites indicated for clarity of the text.

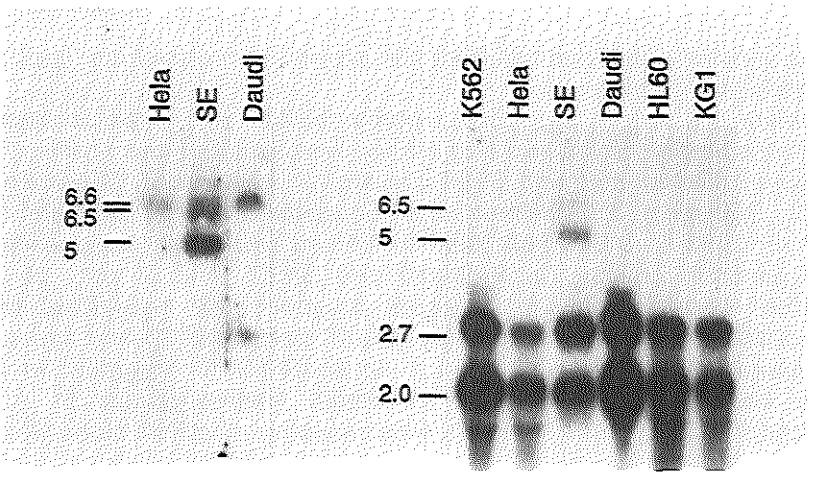


Figure 4. Northern blots. Total RNA (20 μg) isolated from cell lines HeLa and Daudi and from bone marrow of AUL patient SE was hybridized to *can* cDNA probes 3' of the translocation breakpoint (first panel). Total RNA (20 μg) isolated from cell lines K562, Hela, Daudi, HI60 and KG1 and from bone marrow of AUL patient SE was hybridized to the *set* cDNA probe SE10 (second panel). The size of the transcripts is indicated in kb.

the genomic fragment to which it hybridized showed that restriction maps of the two fragments are colinear. These data suggest that the *set* gene contains a large exon at its 3' end, which is situated 5' of the *set-can* translocation breakpoint. Since these exon sequences are not present in the fusion cDNAs SE3 and SE4, they must be excluded from the 5 kb *set-can* transcript by alternative splicing.

Markedly, next to the aberrant 5 kb transcript, a weak 6.5 kb transcript was detected in total RNA of leukemic cells of patient SE after hybridization with the *set* cDNA probe SE4 (Figure 4A). A transcript of the same size was also detected by the 3' *can* cDNA probe hXT54, because the transcript migrates just a little faster than the normal 6.6 kb *can* transcript (Figure 4B). This suggests that, in addition to the 5 kb transcript, a second 6.5 kb *set-can* transcript is present in the leukemic cells. Possibly, this transcript contains part of the 3' *set* exon, spliced to *can* via a cryptic splice donor site in the 3' *set* exon.

Sequence analysis of *set* cDNA. The nucleotide sequence of SE9 and SE10 was determined. Together these clones contained 2570 bp of cDNA sequences. An open reading frame of 831 nt is present, encoding a protein of 277 amino acids with a predicted molecular weight of 32 kD. A long 3' untranslated region of 1733 nt follows the ORF (Figure 5). This 3' UTR contains many stopcodons in all reading frames. The ORF starts immediately at the 5' end of SE10 at position 4 of the cDNA, and the first

stopcodon in this reading frame is found at position 835. Comparison of the sequence of the fusion cDNA SE4 and cDNA clone SE10 shows that nt 813 (amino acid 269) is fused in frame to the *can* ORF. Sequence determination of part of the homologous genomic clone (Sg19E5.5) showed that this position 813 in the *set* cDNA corresponds to an exon-intron border (figure 6B). The 3' border of this intron was determined by sequencing a 300 bp *Hin*dIII fragment of genomic phage Sg19 (figure 6D). A large part of the exon 3' of this intron was sequenced from the Sg19 subclones Sg19E.8 and Sg19E3. This confirmed the suggestion that in the normal *set* gene, the exon spliced to *can* in the fusion gene is followed by a single large 3' exon of 1756 nt. This exon still contains 7 amino acids of the ORF and the entire 3'UTR. The polyadenylation signal for the 2.7 kb transcript is found at position 2542, followed by the polyA tail at position 2562. The genomic sequence overlapping the 3' end of the cDNA is shown in figure 6D.

As the ATG start codon is present at the very 5' end of SE10, an

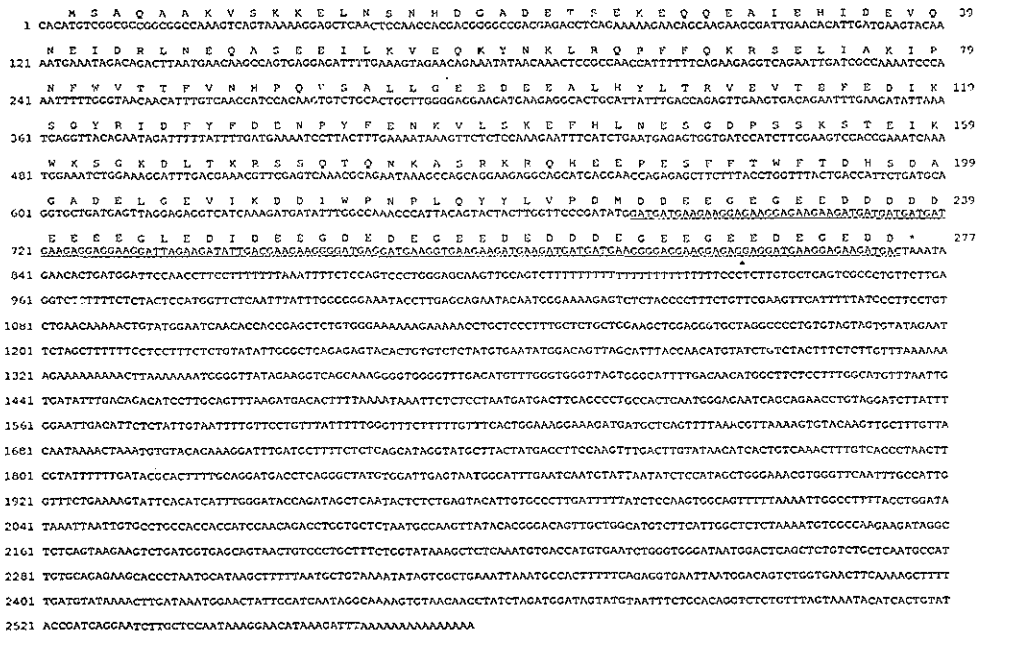


Figure 5. cDNA and putative amino acid sequence of the SET protein. The position where *set* is fused to *can* in the *set-can* fusion protein is indicated with a black triangle (nt 813). The acidic tail is underlined.

anchored PCR was employed to clone the 5′ end of the *set* transcript. Unfortunately, due to the high G/C content of the DNA, we were not successful. The 5′ cDNA sequences are contained in the 5′ end of the 5.5 kb *Eco*RI fragment of the genomic phage Sg19. The genomic nucleotide sequence preceding the ATG codon was determined (Figure 6A), which appeared to be rich in CpG. CpG rich areas are found to be associated with certain promoter regions (4). Therefore it seems likely that this ATG encodes the first methionine of the putative SET protein and that the

```
A: 5'set

ccctttctctccccctccccgctcccccccccgaccgcgggagcagCACATG    Sg19E5S.8
                                                CACATG    SE10

TCGGCGCCGGCGGCCAAAGTCAGTAAAAAGGAGCTCAACTCCAACCACGA      Sg19E5S.8
TCGGCGCCGGCGGCCAAAGTCAGTAAAAAGGAGCTCAACTCCAACCACGA      SE10

CGGGGCCGACGAGACCTCAGgtgagagcagcgacccgggggccggccccg      Sg19E5S.8
CGGGGCCGACGAGACCTCAGAAAAAGAACAGCAAGAAGCGATTGAACACA      SE10


B: 3' end 2.7 kb set mRNA

GGAATCTTGCTCCAATAAAGGAACATAAAGATTTtttttggactggggtc      Sg19H300
GGAATCTTGCTCCAATAAAGGAACATAAAGATTTAAAAAAAAAAAAAAAAA     SE9

gattctccttgttttataagagaaatgttaccttgcctattgatt           Sg19H300


C: 3' end 2 kb set mRNA

TTTTAAACGTTAAAGTGTACAAGTTGCTTTGTTACAATAAAACTAAATGT      SE9
TTTTAAACGTTAAAGTGTACAAGTTGCTTTGTTACAATAAAACTAAATGT      SE220

GTACACAAAGGATTTGATGCTTTTCTCTCAGCATAG                    SE9
GTACACAAAAAAAAAAAAAAA                                   SE220


D: last set intron

ATGATGATGAAGGGGAGGAAGGAGAGgtaaaa.....                   Sg19E5
ATGATGATGAAGGGGAGGAAGGAGAGGAGGATGAAGGAGAAGATGACTA       SE10
gattaactgcccactttttctttcagGAGGATGAAGGAGAAGATGACTA       Sg19E800
```

Figure 6. Comparison of genomic and cDNA sequences of *set*. (A) A 5.5 kb *Eco*RI fragment in Sg19 contains the most 5′ sequences of *set* cDNA clone SE10 in a 800 bp *Sst*I fragment (Sg19E5S.8). (B) The end of the 3′ exon of *set* is present in cDNA clone SE9 and in a genomic 300 bp *Hind*III fragment derived from phage Sg19. The polyA signal is underlined. (C) A 2 kb mRNA is generated by alternative polyadenylation on nt 1690 of the *set* cDNA sequence. SE9 represents part of the 2.7 kb mRNA, SE220 is the amplified 3′ end of the 2 kb *set* mRNA. The polyA signal is underlined. (D) The intron-exon borders of the last *set* intron were sequenced to determine the position of the *set-can* fusion in the *set* cDNA sequence. The exon that can be spliced to *can* is present in a genomic 5.5 *Eco*RI fragment derived of phage Sg19 (Sg19E5). The border of the 3′ exon is present in a 800 bp *Eco*RI fragment of the same phage clone (Sg19E800). Genomic sequences are aligned with cDNA clone SE10.

genomic region 5' of the cloned cDNA contains the promoter sequences.

The putative SET protein contains an extremely high percentage of acidic residues: 32% (98 aa). Half of those (43 aa) are present at the C-terminus, forming a long acidic tail. Most of the acidic residues (37) are present in the chimeric SET-CAN protein. If the amino acid sequence of SET is compared to DEK, only the acidic stretches show homology. Comparison of the SET sequence to the EMBL database revealed

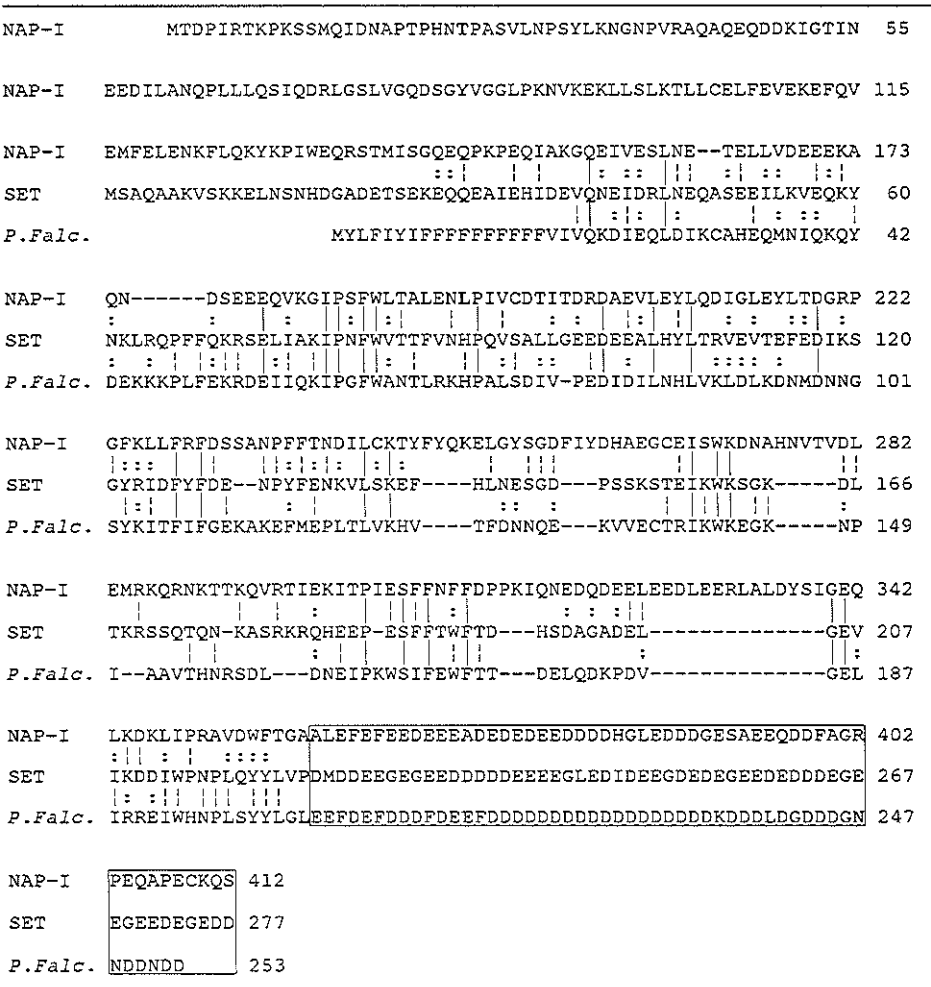


Figure 7. Alignment of the SET protein with a putative protein of *P. falciparum (P. falc.)* and the NAP-I protein of *S. cerevisiae*. Identical aa in all three proteins are indicated with a vertical line. Identical aa in SET and NAP-I or in SET and *P.falc.* are indicated with an interupted vertical line. Similar aa are indicated with colons.

homology to a putative *Plasmodium falciparum* protein (29) and to a nucleosome assembly protein, NAP-I, of *Saccharomyces cerevisiae* (20a) (Figure 7). Both proteins have a large stretch of acidic residues at their C-terminus. In addition, comparison of aa 38 to 221 of SET to aa 20 to 201 of the *P. falciparum* protein showed 33% (61/187) identity and 50% (94/187) similarity. Comparison of aa 1 to 220 of SET to aa 116 to 362 of *S. cerevisiae* NAP-I showed an identity of 24% and 36% similarity. However, some regions are more homologous than others, i.e. SET aa 69 to 143 showed 34% identity and 54% similarity to NAP-I.

The 3' part of the *set* gene is not single copy. Upon screening genomic phage Sg19 for fragments free of repetitive sequences, it was found that a 1 kb *Eco*RI-*Hin*dIII fragment just upstream of the *can* sequences (figure 3) detected a number of bands after hybridization to a Southern blot containing human DNA (Figure 8A). This suggested that at least 10 cross hybridizing copies of this sequence are present in the genome. This *Eco*RI-*Hin*dIII fragment appeared to be part of the 3' exon of *set*. Also cDNA probe SE4ER appeared to be multi-copy. This probe is located upstream of the G/A rich stretch of DNA encoding the acidic tail of the SET protein. On long range mapping Southern blots with *Bssh*II digested DNA, a comparable number of fragments, ranging from 30 - 800 kb in size, could be detected (Figure 8B). This indicates that the sequences hybridizing to *set* are located at large distances from each other and may well be scattered over the genome. Only the very 5' end of the *set* gene is single copy, which includes probe Sg19Hc1 (Figure 8A,D). Probe Sg22E4R400 located immediately 3' of the 3' exon, is also single copy.

To test whether both the 2.7 and 2.0 kb transcript are encoded by the cloned *set* gene and not by any of the other *set* homologous alleles, total RNA of the cell line KG1 was hybridized to (1) the single copy probe Sg19Hc1, (2) the entire cDNA clone SE10 and (3) the 3' terminal 500 bp of cDNA clone SE9 (SE9R520) (Figure 8C). Sg19Hc1 as well as SE10 detected the 2.7 and the 2.0 kb mRNAs, while the 3' probe SE9R520 detected only the larger of the two transcripts. These results indicate that both transcripts are encoded by the *set* gene and that the difference in size is most likely due to alternative polyadenylation. Two putative polyA signals for the 2 kb mRNA are present in the 3' UTR, at position 1483 and 1682 respectively. To test whether these signals are used, the RACE protocol (13) was employed. cDNA generated with an adaptor-oligodT primer was amplified with an adaptor oligonucleotide sequence and a primer located at either position 1372-1391 or a primer located at position 1519-1539. This resulted in fragments of 350 and 220 nt respectively. Direct sequencing of these fragments showed that the putative polyA signal at position 1483 is not used, while the polyA signal at position 1682 results in polyadenylation of the mRNA at position 1702 (Fig. 6C)

A 'zoo-blot' containing DNA derived from man, mouse, marsupial, bird,

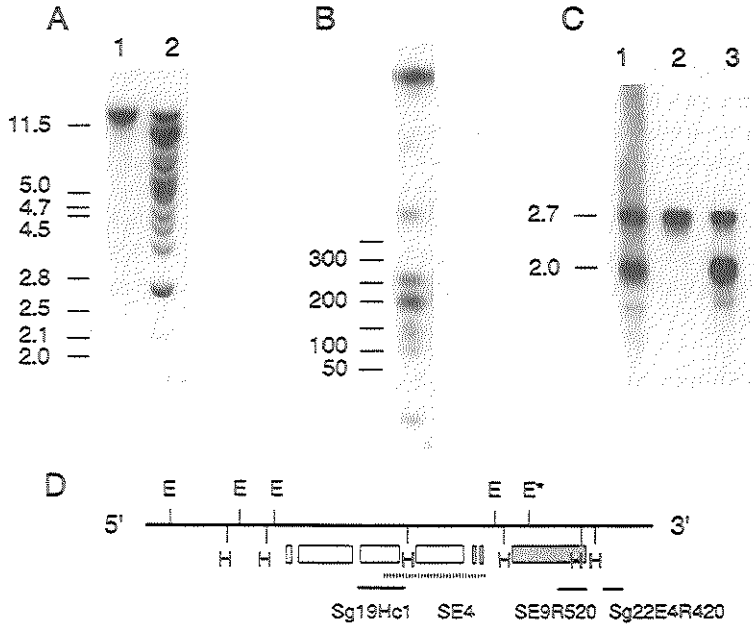


Figure 8. (A/B) Multiple copies of *set* are present in the genome. Southern blots containing DNA derived of human thymus, digested with *Bgl*II (A), or DNA derived of the cell line KG1, digested with *Bssh*II (B), were hybridized to probes Sg19Hc1 (A:lane 1) and SE9R520 (A:lane 2, B). A size marker (in kb) is shown next to the blots. (C) Both *set* transcripts are derived from the cloned *set* gene. Northern blots with RNA derived of the cell line KG1 were hybridized to probes Sg19HcII (lanes 1), SE9R520 (lanes 2) and SE10 (lanes 3). The size of the transcripts is indicated in kb. (D) The position of the probes is indicated in the genomic map below the blots. Open boxes underneath the *set* restriction map indicate restriction fragments hybridizing to *set* cDNA probes. Probes Sg19Hc1, SE9R520 and Sg22E4R420 are indicated by bars, cDNA probe SE10 contained cDNA sequences encoded upstream of the *Eco*RI site marked with an asterix. The estimated position of SE4 is indicated with an interrupted line.

amphibian, fish, fly and yeast was hybridized to the cDNA probe SE10 to see whether the *set* gene and the 3' *set* repeated sequences are conserved between species (Figure 9). Hybridization of SE10 to chicken, reptile and fish DNA could still be detected under relatively stringent conditions (1xSSC,65°C). Markedly, the repetitive nature of the 3' part of *set* was not only seen in man but also in mouse.

Expression of *set*. The expression pattern of *set* in different mouse tissues was analyzed by Northern blotting. Total RNA (20 μg) of bone

marrow, spleen, thymus, brain, liver, kidney, testes, ovarium, placenta and whole embryos 10, 13, 16 and 19 sacrificed days post coitum was loaded on a denaturing agarose gel. Hybridization of *set* cDNA probe SE10 to mouse DNA showed that the gene is evolutionary well conserved (Figure 9). Thus, the human *set* cDNA clone SE10 was used to screen for mouse *set* transcripts. As shown in Figure 10 *set* is expressed in all adult tissues analyzed. The expression during embryogenesis is remarkable. *Set* expression is relatively high in the youngest embryos tested (10 days old) and decreases during development of the embryos. In mouse not two but three or four transcripts were detected. The nature of the smaller mRNAs is not known. The band migrating just under the 18S rRNA band may be background due to compression of the background smear by the bulk of rRNA.

Chromosomal localization of the *set* gene. The karyotype of leukemic cells of patient SE appeared to be normal and gave no clue as to where in the genome *set* is located. A biotinylated genomic fragment from phage Sg19, encompassing three *Hin*dIII fragments of 3.3, 2.3 and 1.9 kb (Sg19Hp, figure 1) was used to determine the chromosomal localization by *in situ* hybridization. Although some background fluorescence was present, a clear signal was detected at the tip of the long arm of chromosome 9 (figure 11A/B). The oncogenes c-*abl* involved in t(9;22)

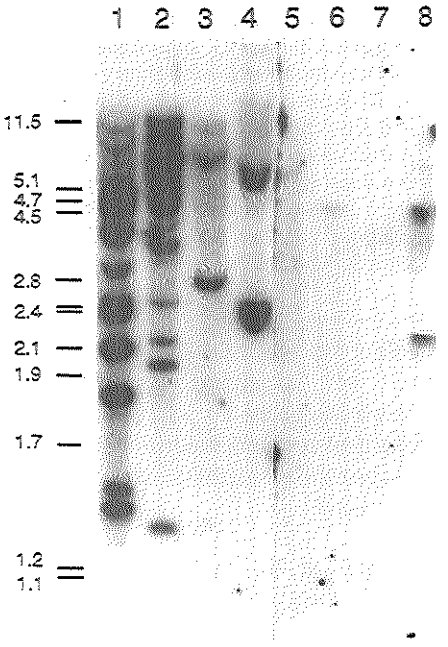


Figure 9. Zoo blot. DNA derived from man (human thymus), mouse (NIH3T3 cells), marsupial (*Potorous tridactylis* cell line), bird (chicken liver), amphibian (*Xenopus laevis* liver), fish (*Aulonocara stuartgranti* liver), fly (*D. melanogaster* cell line) and yeast (*S. cerevisiae*) (lanes 1 to 8) was digested with *Eco*RI, blotted and hybridized to the *set* cDNA probe SE10. The blot was washed at 65°C in 1xSSC. A size marker is indicated in kb.

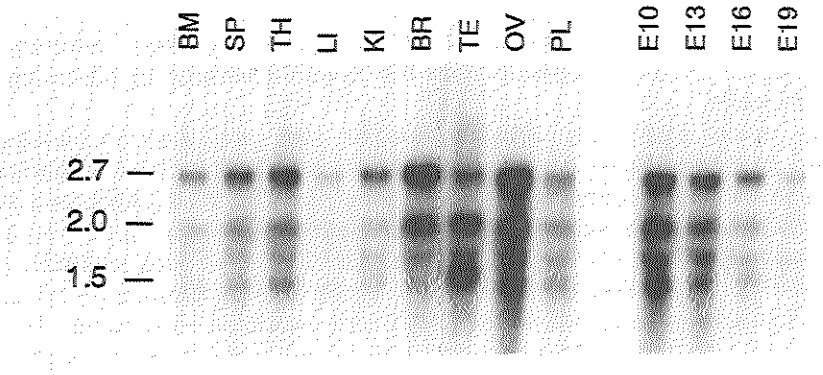


Figure 10. Expression of *set* in mouse tissues. A Northern blot containing total RNA derived of various tissues of BCBA mice was hybridized to the human *set* cDNA probe SE10. The filter was washed at 65°C, 0.3xSSC. BM:bone marrow, SP:spleen, TH:thymus, LI:liver, KI:kidney, BR:brain, TE:testes, OV:ovarium, PL:placenta, E10, E13, E16, E19: embryos aged 10, 13, 16 and 19 days post coitum. Lanes E10-E19 are exposed shorter than the other lanes. In the original exposure, the signal of lane E10 is comparable to the signal of lane SP.

and *can* involved in t(6;9) are also located on chromosome 9q34. Somatic cell hybrids containing the 9q+ or 22q- chromosomes of t(9;22) were hybridized to probe Sg19Hc1, to determine the localization of *set* relative to c-*abl* and *can*. As shown in figure 11C, *set* sequences hybridized to cell lines carrying the normal chromosome 9 (17CB-10) or the 9q+ chromosome (8CB-7B, 15CB-7D) but not to cell lines carrying the 22q- chromosome (Wedy9, WESP-2A). This confirmed the results of the *in situ* hybridization and locates the *set* gene on chromosome 9, centromeric of c-*abl*. The physical distance between *set* and *abl* is unknown.

DISCUSSION

In t(6;9) AML a specific fusion of *dek* to *can* is found. All t(6;9) leukemic cells analyzed to date, invariably contained a translocation breakpoint in *icb*-6 of *dek* and in *icb*-9 of *can* (37). We now provided evidence that *can* can be involved in at least two different translocation events. In leukemic cells of a patient with AUL *can* is fused to a novel gene named *set*. The *set-can* fusion gene encodes a 5 kb chimeric transcript whose nucleotide sequence predicts a 150 kD SET-CAN fusion protein. The finding that the same part of CAN is linked by translocation to two different protein moieties at its N-terminal side, suggests that the C-terminal part of CAN contains domains involved in leukemogenesis that can be activated in more than one way. Markedly, the phenotype of the leukemic cells

carrying the *set-can* fusion gene was very immature in this patient (AUL with rearrangement of T cell receptor ß, δ and γ genes), whereas a variable degree of differentiation into the myeloid lineage is observed in t(6;9) AML cells.

A large scale study for involvement of *dek* and *can* in leukemia confirmed the specificity of the association of t(6;9) with the *dek-can* fusion gene. However, among the 320 cases of MDS, AUL, AML and ALL studied, two leukemia samples contained a breakpoint in *icb*-9 of *can* while no breakpoint in *dek* could be detected. One case was a RAEB, the other a c-ALL (38). These cases were tested for a breakpoint in *set*, which so far gave a negative answer (Bartram unpublished results). These results suggest that in addition to *dek* and *set* even more genes might be able to activate *can*. Although the different types of CAN fusion-proteins are all associated with leukemia, the differentiation potential of the leukemic cells may be influenced by the N-terminal moiety of the fusion

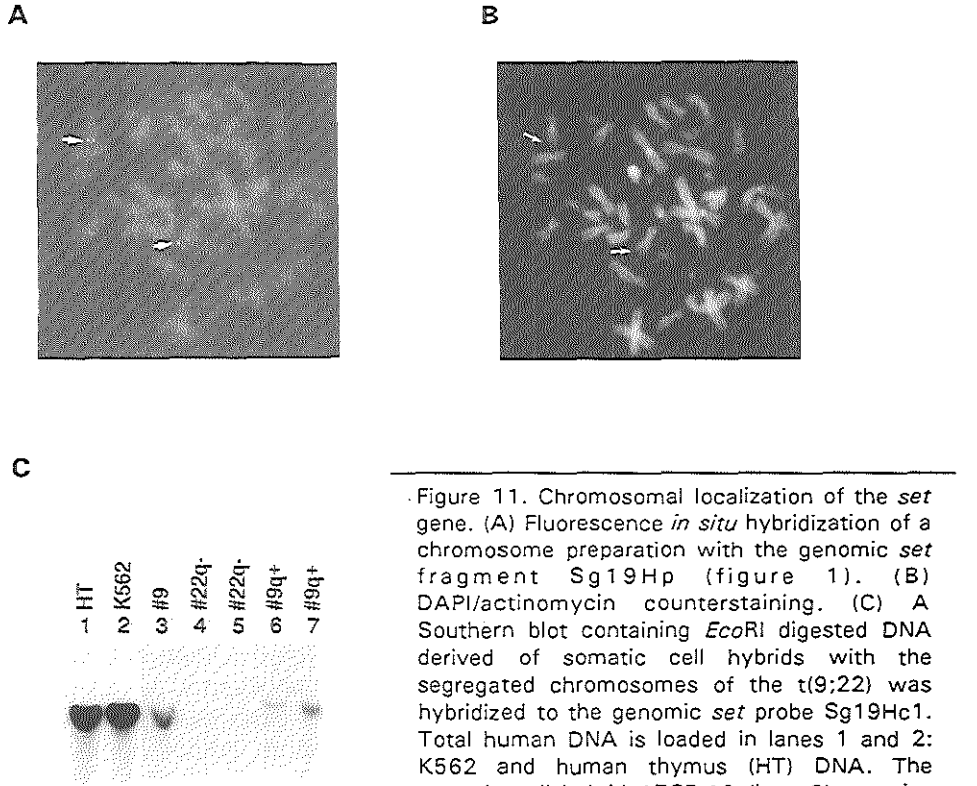


Figure 11. Chromosomal localization of the *set* gene. (A) Fluorescence *in situ* hybridization of a chromosome preparation with the genomic *set* fragment Sg19Hp (figure 1). (B) DAPI/actinomycin counterstaining. (C) A Southern blot containing *Eco*RI digested DNA derived of somatic cell hybrids with the segregated chromosomes of the t(9;22) was hybridized to the genomic *set* probe Sg19Hc1. Total human DNA is loaded in lanes 1 and 2: K562 and human thymus (HT) DNA. The somatic cell hybrid 17CB-10 (lane 3) contains entire chromosome 9, Wedy9 and WESP-2A (lanes 4, 5) contain the 22q- chromosome and 8CB-7B and 15CB-7D (lanes 7, 8) contain the 9q+ chromosome.

protein.

The predicted protein sequence of SET shows no homology with DEK apart from the fact that both proteins contain acidic regions. Many proteins containing acidic regions are located in the nucleus and may have different functions (reviewed by (9)). Although the function of CAN and DEK is not known yet, domains present in CAN and DEK suggest a role in transcription regulation (39). Analogous to acidic domains in VP16 and GAL4 (5, 20) the acidic motifs of DEK and SET could serve as transcription activation domains. On the other hand, acidic domains are also present in proteins associated with chromatin, like nucleolins (28) and high mobility group proteins (HMG-family) (26). Functional assays are needed to determine whether the acidic domains of SET and DEK are essential for the putative transforming capacity of the DEK-CAN and SET-CAN fusion proteins.

The homology of SET with the *S. cerevisiae* nucleosome assembly protein NAP-I (20a) argues that SET may be a nuclear protein. This is interesting since fusion of CAN to DEK results in a nuclear localization of the fusion protein, while CAN itself is mainly cytoplasmic. Fusion of SET to CAN could have the same effect and result in a nuclear localization of the SET-CAN fusion protein. A nuclear localization of the C-terminal part of CAN may be essential for the putative leukemogenic effect of the fusion proteins.

In addition to the acidic stretch at the N-terminus, the entire SET protein and the NAP-I protein are 24% identical. SET also shows homology to a putative protein of *P. falciparum*, whose function is unknown (29). Possibly, the functions of SET and the putative *P. falciparum* protein in the cell are related to the function of the nucleosome assembly protein. SET and the *P.falciparum* protein are similar in size and their overall identity is 33%. The NAP-I protein is larger and extends at the N-terminal side.

Like *dek*, the *set* gene is expressed in all tissues of the mouse, suggesting that SET has a rather general function in the organism. Still the expression pattern is not entirely identical to *dek*. Most remarkable is the high expression level in early embryos. It will be interesting to analyze by *in situ* hybridization whether high expression of *set* in early embryos is round all over the embryo or whether it is restricted to specific structures.

Set is a relatively small gene of 8 kb. The most 5' sequences of the cloned *set* cDNA are located in a CpG island that measures at least 1000 bp in the genomic DNA. The size of the mRNAs on Northern blots is estimated to be 2.7 and 2 kb. The cloned cDNA sequences represent 2.5 and 1.7 kb. This difference in size may be due in part to the fact that the polyA tail is not included in the cDNA. However, it is more likely that sequences at the 5' side of the transcript are missing from the cDNA clones. Since CpG islands are strongly associated with promoter regions (4), we assume that the 5' part of phage Sg19 contains the most 5' *set*

sequences and the *set* promoter.

In comparison with other fusion genes, the position of the breakpoint in the *set-can* fusion gene is peculiar as it is located 800 bp 3' of the *set* gene and not in an intron. Apparently, the primary transcript is not terminated in the 4.8 kb that separate the 3' end of *set* from the most 5' exon of the translocated *can* gene, but proceeds to the end of the *can* gene. Subsequently, the 3' *set* exon must be skipped by splicing to produce an in frame *set-can* fusion transcript. This alternative splicing may be rather efficient. Large exons seldomly occur within a gene and it has been reported that large exons are preferentially spliced out of primary transcripts (16, 32). Markedly, next to an aberrant 5 kb transcript, a weaker 6.5 kb transcript can be detected in total RNA of leukemic cells of patient SE which hybridized to cDNA clone SE10. An additional transcript of similar size does also hybridize to 3' *can* cDNA probes and is just smaller than the 6.6 kb normal *can* transcript. This suggests that the 6.5 kb mRNA may be a *set-can* fusion transcript containing part of the 3' *set* exon, spliced in via a cryptic splice donor site. Considering the intensities of the signals on the Northern blot, the use of the cryptic splice site is not very efficient. The presence of transcripts containing part of the 3' *set* exon may be of no functional importance as the natural *set* stopcodon is present in this mRNA. The longer fusion transcripts would encode only *wt* SET protein.

Hybridization of total RNA of leukemic cells of patient SE with *set* cDNA probes shows that the steady state levels of *set* are much higher than those of *set-can*, despite the presence of >90% leukemic cells in the bone marrow sample from which RNA was isolated. As each cell contains one *set* promoter driving *wt set* expression and one *set* promoter driving transcription of the *set-can* fusion gene, additional elements regulate the relative numbers of *set* and *set-can* transcripts. (*i*) Transcription of the *set* gene may be upregulated by an enhancer element located 3' of the gene, which is removed by the translocation. (*ii*) The presence of the last 3' exon on the rearranged chromosome may give rise to production of *wt set* transcripts from the *set-can* fusion gene. Transcription termination and polyadenylation of normal *set* transcripts will decrease the level of *set-can* transcripts in the cell. (*iii*) The difference in *set* and *set-can* steady state levels may also be due to a reduced halflife of *set-can* compared to *set*.

The karyotype of the leukemic cells of patient SE appeared to be normal. We show here that *set* is localized on chromosome 9q34. Thus, the chromosomal aberration is not detected by cytogenetic means since the distance between *set* and *can* is relatively small. The chromosomal rearrangement that took place to fuse *set* to *can* may be either an insertion, an inversion or a translocation. A deletion can be excluded because genomic phages were cloned that contained novel sequences fused to the 5' part of *can* as well as to its 3' part. Fluorescent *in situ* hybridization techniques (2) may distinguish between the other

possibilities: a translocation, an insertion or an inversion.

Only the 5' end of *set* appeared to be single copy. The 3' part of the *set* gene, including exons coding for the SET protein, is present at least 10 times in the human genome. The *set* gene may be part of a gene family whose members have highly homologous 3' ends and specific 5'leaders. Alternatively, nonexpressed pseudogenes lacking the 5' end of *set*, may be present in the genome. Similarly, several (4) duplications of the 3' *bcr* gene are present on chromosome 22 at large distances from each other (17). These *bcr* copies are not expressed and contain conserved exons as well as conserved introns. The *bcr*-related genes are conserved in gorilla and chimpanzee but not in mouse (17).

Northern blot and PCR analysis suggest that both the 2.0 and 2.7 kb *set* transcripts are encoded by the *set* gene identified in this paper. However, it can not be excluded that *set* homologous genes are expressed in specific cell types. The Zoo-blot showed that the *set* gene is evolutionary well conserved and that a multiplication of *set* occurred after the separation of marsupials and mammals but before separation of primates and rodents. It is surprising that multiple copies generated so long ago remain so well conserved not only in the protein coding part but also in the large 3' UTR. It is not known whether introns are conserved as well.

ACKNOWLEDGMENTS

We thank D. Bootsma, D. Meijer and M. fornerod for continuous support and discussion, J. Abels and H. Adriaansen for patient material, M. Kuit and T. de Vries Lentsch for photographic work. This work is supported by the Dutch Cancer Foundation.

LITERATURE CITED

1. Adriaansen, H.J., J.J.M. van Dongen, H. Hooijkaas, K. Hählen, M.B. van 't Veer, B. Löwenberg and A. Hagemeijer. 1988. Translocation (6;9) may be associated with a specific TdT-positive immunological phenotype in ANLL. Leukemia 2:136-40.
2. Arnoldus, E.P.J., J. Wiegant, I.A. Noordermeer, J.W. Wessels, G.C. Beverstock, G.C. Grosveld, M. van der Ploeg and A.K. Raap. 1990. Detection of the Philadelphia chromosome in interphase nuclei. Cytogenet Cell Genet 54:108-111.
3. Auffray, C. and F. Rougeon. 1980. Purification of mouse immunoglobulin heavy-chain messenger RNAs from total myeloma tumor RNA. Eur J Biochem 107:303-14.
4. Bird, A.P. 1986. CpG islands and the function of DNA methylation. Nature 321:209-213.
5. Brent, R. and M. Ptashne. 1985. A eukaryotic transcriptional activator bearing the DNA specificity of a prokaryotic repressor. Cell 43:729-736.
6. Chirgwin, J.M., A.E. Przybyla, J. MacDonald and W.J. Rutter. 1979. Isolation of Biologically Active Ribonucleic Acid from Sources Enriched in Ribonuclease. Biochemistry 18:5294-5299.
7. de Klein, A., A. Geurts van Kessel, G. Grosveld, C.R. Bartram, A. Hagemeijer, D. Bootsma, N.K. Spurr, N. Heisterkamp, J. Groffen and J.R. Stephenson. 1982. Nature 300:765-767.

8. De Thé, H., C. Lavau, A. Marchio, C. Chomienne, L. Degos and A. Dejean. 1991. The PML-RARA Fusion mRNA Generated by the t(15;17) Translocation in Acute Promyelocytic Leukemia Encodes a Functionally Altered RAR. Cell 66:675 - 684.
9. Earnshaw, W.C. 1987. Anionic regions in nuclear proteins. J Cell Biol 105:1479-82.
10. Feinberg, A.P. and B. Vogelstein. 1983. A technique for radiolabeling DNA restriction endonuclease fragments to high specific activity. Anal Biochem 132:6-13.
11. Fourney, R.M., J. Miyakoshi, R.S. Day III and M.C. Paterson. 1988. Northern Blotting: efficient RNA Staining and Transfer. Focus 10:5-6.
12. Frischauf, A., H. Lehrach, A. Poustka and N. Murray. 1983. Lambda replacement vectors carrying polylinker sequences. J. Mol. Biol. 170:827-842.
13. Frohman, M.A. 1990. Rapid Amplification of cDNA Ends (RACE): User-Friendly cDNA Cloning. Amplifications 5:11 - 15.
14. Geurts van Kessel, A.H.M., P.A.T. Tetteroo, A.E.G.K. von dem Borne, A. Hagemeijer and D. Bootsma. 1983. Expression of human myeloid-associated surface antigens in human-mouse myeloid cell hybrids. Proc. Natl. Acad. Sci. USA 80:3748-3752.
15. Grosveld, G., T. Verwoerd, T. van Agthoven, A. de Klein, K.L. Ramachandran, N. Heisterkamp, K. Stam and J. Groffen. 1986. The chronic myelocytic cell line K562 contains a breakpoint in bcr and produces a chimeric bcr/c-abl transcript. Mol Cell Biol 6:607-16.
16. Hawkins, J.D. 1988. A survey on intron and exon lengths. Nucleic Acids Res 16:9893-908.
17. Heisterkamp, N. and J. Groffen. 1988. Duplication of the *bcr* and gamma-glutamyl transpeptidase genes. Nucleic Acids Research 16:8045-8056.
18. Hermans, A., N. Heisterkamp, M. von Linden, S. van Baal, D. Meijer, D. van der Plas, L.M. Wiedemann, J. Groffen, D. Bootsma and G. Grosveld. 1987. Unique fusion of bcr and c-abl genes in Philadelphia chromosome positive acute lymphoblastic leukemia. Cell 51:33-40.
19. Hermans, A., L. Selleri, J. Gow, L. Wiedeman and G. Grosveld. 1989. Molecular Analysis of the Philadelphia Translocation in Myelogenous and Acute Lymphoblastic Leukemia. Cancer Cells 7:21-26.
20. Hope, I.A. and K. Struhl. 1986. Functional dissection of a eukaryotic transcriptional activator protein, GCN4 of yeast. Cell 46:885-94.
20a. Ishimi, Y. and Kikuchi, A. 1991. Identification and molecular cloning of yeast homolog of nucleosome assembly protein I which facilitates nucleosome assembly in vitro. J. Biol. Chem. 266:7025-7029.
21. Jeffreys, A. and R.A. Flavell. 1977. A physical map of the DNA regions flanking the Rabbit B-globin gene. Cell 12:429-439.
22. Kadowaki, T., H. Kadowaki and S.I. Tayler. 1990. A nonsense mutation causing decreased levels of insulin receptor mRNA: Detection by a simplified technique for direct sequencing of genomic DNA amplified by the polymerase chain reaction. Proc. Natl. Acad. Sci. USA 87:658-662.
23. Kakizuka, A., W.H. Miller, K. Umesono, R.P. Warrell, S.R. Frankel, V.V.V.S. Murty, E. Dmitrovsky and R.M. Evans. 1991. Chromosomal translocation t(15;17) in human acute promyelocytic leukemia fuses RARA with a novel putative transcription factor, PML. Cell 66:663 - 674.
24. Kamps, M.P., C. Murre, X.H. Sun and D. Baltimore. 1990. A new homeobox gene contributes the DNA binding domain of the t(1;19) translocation protein in pre-B ALL. Cell 60:547-55.
25. Konopka, J.B., S.M. Watanabe and O.N. Witte. 1984. An alteration of the human c-abl protein in K562 leukemia cells unmasks associated tyrosine kinase activity. Cell 37:1035-42.
26. Krohne, G. and W.W. Franke. 1980. A major soluble acidic protein located in nuclei of diverse vertebrate species. Exp Cell Res 129:167 - 189.
27. Kurzrock, R., J.U. Gutterman and M. Talpaz. 1988. The molecular genetics of Philadelphia chromosome-positive leukemias. N Engl J Med 319:990-8.
28. Lapeyre, B., H. Bourbon and F. Amalric. 1987. Nucleolin, the major nucleolar protein of

growing eukaryotic cells: an unusual protein structure revealed by the nucleotide sequence. Proc Natl Acad Sci USA 84:1472-6.

29. Lenstra, R., L. d'Auriol, B. Andrieu, J. Le Bras and F. Galibert. 1987. Cloning and sequencing of Plasmodium falciparum DNA fragments containing repetitive regions potentially coding for histidine-rich proteins: identification of two overlapping reading frames. Biochem Biophys Res Commun 146:368-77.
30. Nourse, J., J.D. Mellentin, N. Galili, J. Wilkinson, E. Stanbridge, S.D. Smith and M.L. Cleary. 1990. Chromosomal translocation t(1;19) results in synthesis of a homeobox fusion mRNA that codes for a potential chimeric transcription factor. Cell 60:535-45.
31. Pendergast, A.M., A.J. Muller, M.H. Havlik, Y. Maru and O.N. Witte. 1991. BCR sequences essential for transformation by the BCR-ABL oncogene bind to the ABL SH2 regulatory domain in a non-phosphotyrosine-dependent manner. Cell 66:161-171.
32. Robberson, B.L., G.J. Cote and S.M. Berget. 1990. Exon definition may facilitate splice site selection in RNAs with multiple exons. Mol Cell Biol 10:84-94.
33. Sambrook, J., E.F. Fritsch and T. Maniatis. 1989. Molecular Cloning: a Laboratory Manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, New York.
34. Sandberg, A.A., R. Morgan, J.A. McCallister, M.B. Kaiser and F. Hecht. 1983. Acute myeloblastic leukemia (AML) with t(6;9) (p23;q34): a specific subgroup of AML? Cancer Genet Cytogenet 10:139-42.
35. Sanger, F., S. Nicklen and A.R. Coulson. 1977. DNA Sequencing with Chain-terminating Inhibitors. Proc Natl Acad Sci USA 74:5463 - 5467.
36. Shtivelman, E., B. Lifshitz, R.P. Gale and E. Canaani. 1985. Fused transcript of abl and bcr genes in chronic myelogenous leukaemia. Nature 315:550-4.
37. Soekarman, D., M. von Lindern, S. Daenen, B. de Jong, C. Fonatsch, B. Heinze, C.R. Bartram, A. Hagemeijer and G. Grosveld. The translocation (6;9)(p23;q34) shows consistent rearrangement of two genes and defines a myeloproliferative disorder with specific clinical features. Blood, in press
38. Soekarman, D., M. von Lindern, D. van der Plas, L. Selleri, C.R. Bartram, P. Martiat, D. Culligan, R.A. Padua, K.P. Hasper-Voogt, A. Hagemeijer and G. Grosveld. DEK-CAN rearrangement is restricted to translocation (6;9)9p23;q34). Leukemia, in press
39. von Lindern, M., M. Fornerod, S. van Baal, M. Jaegle, T. de Wit, A. Buijs and G. Grosveld. The t(6;9), associated with a specific subtype of AML, results in the fusion of two genes, dek and can, and the expression of a chimeric, leukemia specific dek-can mRNA. Moll. Cell. Biol. in press:
40. von Lindern, M., A. Poustka, H. Lerach and G. Grosveld. 1990. The (6;9) chromosome translocation, associated with a specific subtype of acute nonlymphocytic leukemia, leads to aberrant transcription of a target gene on 9q34. Mol Cell Biol 10:4016-26.
41. von Lindern, M., T. van Agthoven, A. Hagemeijer, H. Adriaansen and G. Grosveld. 1989. The human pim-1 gene is not directly activated by the translocation (6;9) in acute nonlymphocytic leukemia. Oncogene 4:75-9.
42. Walker, L.C., T.S. Ganesan, S. Dhut, B. Gibbons, T.A. Lister, J. Rothbard and B.D. Young. 1987. Novel chimeric protein expressed in Philadelphia positive acute lymphoblastic leukaemia. Nature 329:851-3.

CHAPTER 2.7

Characterization of the translocation breakpoint sequences of two *dek-can* fusion genes present in t(6;9) AML, and a *set-can* fusion gene found in a case of AUL.

Marieke von Lindern[1], Dimitri Breems[1,2], Sjozèf van Baal[1], Henk Adriaansen[2] and Gerard Grosveld[1].

[1]Department of Cell Biology and Genetics, [2]Department of Immunology, Erasmus University, Rotterdam, The Netherlands.

Submitted for publication

ABSTRACT

The t(6;9) associated with a subtype of AML was shown to generate a fusion between the 3' part of the *can* gene on chromosome 9 and the 5' part of the *dek* gene on chromosome 6. The same part of the *can* gene appeared to be involved in a case of AUL as well, where it was fused to the *set* gene. Genomic sequences around the translocation breakpoint were determined in two t(6;9) samples and in the case of the *set-can* fusion. Although coexpression of myeloid markers and terminal deoxynucleotidyl transferase was shown to be one of the characteristics of t(6;9) AML, no addition of random nucleotides at the translocation breakpoint could be found. In addition, the breakpoint regions did not reveal heptamer-nonamer sequences, purine-pyrimidine tracts, a chi-octamer motif or *Alu*-repeats. The sequence in which the translocation breakpoints occurred was enriched in A/T. Markedly, the specific introns in which clustering of breakpoints occurs in *dek* and *can*, both contain a LINE-1 element. As LINE-1 elements occur with a moderate frequency in the human genome, the presence of such an element in both breakpoint regions may be more than coincidental and may play a role in the translocation process.

INTRODUCTION

An ever increasing number of leukemia subtypes are shown to be associated with specific translocations and several translocation breakpoints have recently been cloned and analyzed. Due to translocation, oncogenes are activated either through interference with expression regulation or through the formation of fusion genes (reviewed by (11)).

Translocations in lymphoid neoplasias often take place in the

immunoglobuline (Ig) and T cell receptor (TCR) genes (12, 16). Cloning and molecular characterization of breakpoints involving these genes, strongly suggests that mechanisms that effectuate Ig and TCR gene rearrangement are also involved in the translocation of proto-oncogenes to the Ig or TCR gene loci. The recombinases responsible for Ig or TCR rearrangement recognize the heptamer-nonamer sequence CACAGTG-spacer-ACAAAAACC (33). In a large number of lymphoid malignancies this sequence, or part of it, can be recognized in oncogenes translocated to the Ig or TCR loci (7). Analysis of the DNA sequence near such translocation breakpoints suggests that recombinases can act on sequences with a minimal homology to the nonamer-heptamer sequence.

The enzyme terminal deoxynucleotidyl transferase (TdT) adds at random nucleotides to the recombination site of segments of rearranging Ig or TCR genes (22). A similar addition of nucleotides has been described for the translocation breakpoint of some lymphoid malignancies even when only a minimal homology to a heptamer sequence can be decerned (7) This implicates that TdT is active during the translocation event.

In addition to the heptamer-nonamer sequence, alternating purine-pyrimidine tracts near the translocation breakpoints may promote normal and aberrant rearrangements of Ig/TCR genes (6). Purine-pyrimidine tracts have the potential to form Z-DNA, thereby rendering the chromatin structure accessible to enzymes mediating translocation events.

Krowczynska et al. (21) reported the presence of the minisatellite VTR1.1, a *chi*-like octamer sequence (GC(A/T)GG(A/T)GG), near or at recombination sites of Ig/TCR genes as well as genes involved in chromosomal translocations.

In contrast, genes or enzymes involved in rearrangements have never been detected in myeloid cells. Analysis of translocation breakpoints suggested an involvement of repetitive DNA elements either through recombination or through the formation of an accessible chromatin structure (10). Repeated DNA elements in the mammalian genome are divided in LINES (long interspersed elements) and SINES (short interspersed elements). The LINE or L1 element is 6 kb in size and contains two open reading frames that are well conserved during evolution. The complete element occurs approximately $4x10^3$ times in the human genome, but 5' truncated versions are repeated roughly $2x10^4$ times (15, 29). The SINES include the primate *Alu* family; sequences are about 360 bp in size and less conserved during evolution. They occur with a frequency of $3x10^5$ copies in the primate genome (28). *Alu* repeats are often present at or near t(9;22) breakpoints in chronic myeloid leukemia (CML) and acute lymphoid leukemia (ALL) (10, 14, 34). L1 sequences were detected near rearranged sequences in some lymphoid malignancies (25, 26).

Analysis of t(9;22) breakpoints also revealed the presence of A/T rich

regions (10, 24), which may be a target for the nucleosome phasing α-protein, a HMG-like protein that can bind three distinct sites in α-satellite DNA, which brings them together in order to form nucleosomes (31). Its DNA binding is not very specific; in addition to consensus α-sites, the protein is able to bind any A/T rich sequence.

Previously it has been shown that translocation breakpoints in t(6;9) acute myeloid leukemia (AML) occur in one specific intron of the *Cain (can)* gene on chromosome 9 and the *dek* gene on chromosome 6. These introns were named *icb*-9 (intron containing the breakpoint on chromosome 9) and *icb*-6, respectively. Within the introns the breakpoints are scattered over a region of 7.5 (*icb*-9) and 9 kb (*icb*-6) of DNA. Translocation breakpoints were isolated from two t(6;9) AML cases (DK and PM) (36) and from one acute undifferentiated leukemia patient (AUL; SE) (35). This patient had an apparently normal karyotype but molecular analysis showed that *can* was fused to another gene present on chromosome 9, band q34, which we designated *set*. In two t(6;9) AML patients analyzed (one of these was patient DK), a relatively large percentage (25.5%, DK) of leukemic cells expressed myeloid markers (CD13, CD33) together with TdT (2). Leukemic cells of patient SE coexpressed TdT and myeloid markers as well (1). Therefore, we set out to investigate whether addition of random nucleotides by TdT had occurred during the translocation event in these myeloid leukemias. This paper reports the sequence analysis of the translocation breakpoints of these three patients. The involvement of specific sequences in the translocation event are discussed.

MATERIALS AND METHODS

Immunological characterization of leukemic cells of patients DK and SE has been reported by Adriaansen et al. (1, 2) (patient 2 and R.S., respectively). Molecular characterization of the translocation present in leukemic cells of patients DK, PM and SE has been reported by Von Lindern et al.(35, 36). These papers also report the cloning and preliminary characterization of the DNA fragments used in this study.

Sequence determination and analysis. A 600 bp *Hin*cII fragment of the genomic fragment Sg19E3 was subcloned in M13mp18, a 500 bp *Hin*cII-*Bam*HI of Al1F5E4 was subcloned in pTZ18, a 700 bp *Hin*cII-*Hin*dIII fragment of MmF2.5E3, a 800 bp *Hin*cII-*Eco*RI fragment of Mm1.7, a 850 bp *Hin*cII-*Eco*RI fragment of Al1F7E2, a 1.65 kb *Hin*cII-*Pst*I fragment of 4CE and a 1.5 kb *Pst*I-*Hin*dIII fragment of Al1F5E4 were subcloned in pBluescript (figure 1). Standard vector primers were used to sequence inserted fragments, only Sg22E4 was sequenced directly with a specific primer. Sequences were determined by double strand (ds) dideoxy sequencing (27), mostly on only one of the strands. Sequences were analyzed with the computerprogram Microgenie and the EMBL database was used to search for homologous sequences at the nucleotide level. The nucleotide sequence data will appear in the EMBL, GenBank and DDBJ Nucleotide Sequence Databases under the accession numbers X63687 (*set*), X63688 (*dek, icb*-6), X63689 and X63690 (*can, icb*-9).

To investigate the length and position of the LINE-1 repeat, the ends of all available

subclones (pTZ or pBluescript plasmids) of *icb*-9 and *icb*-6 were sequenced (ds) with the universal and M13 primers. The DNA stretches of which sequences were analyzed for homology with the LINE-1 repeat are indicated in figure 4.

Polymerase chain reaction (PCR). The genomic fragment containing the translocation breakpoint of patient DK on the 6p- chromosome, was amplified from the somatic hybrid cell line MA3B (37) using as a 5' primer 5'TTGAATTCTAATATCCAGAATCTAT-AAGG3' and as 3' primer 5'TTGAATTCACTAATGTTGCAGTGTGGTG3'. Both primers contain an *Eco*RI site at the 5' side to facilitate cloning of the amplified fragment in pBluescript. Annealing was performed at 50°C, amplification at 72°C, denaturation at 94°C.

RESULTS

Previously we reported the cloning of the translocation breakpoints of patients DK, PM (t(6;9) AML) and SE (AUL) and the immunological characterization of leukemic cells of patients DK and SE (1, 2, 36). Comparison of the restriction maps of chromosomes 9 and 6 with sequences of the derivative chromosomes roughly indicated the location of the breakpoints in *set, dek* and *can* within a few hundred basepairs. The translocation breakpoints of patient PM occur in a 3.2 kb *Eco*RI fragment in *dek* (MmF2.5E3) and in a 2.1 kb *Eco*RI fragment (AI1F7E2) in *can*, resulting in a 1.7 kb *Eco*RI fragment (Mm1.7) on the 6p- chromosome (Figure 1A). The translocation breakpoints of patient DK occur in the same fragment in *dek* (MmF2.5E3) and in a 3.1 kb *Eco*RI fragment (AI1F5E4) in *can*, generating a 4 kb *Eco*RI fragment (4CE) on the 9q+ chromosome (Figure 1B). In patient SE translocation breakpoints occur in a 4.3 kb *Eco*RI fragment in *set* (Sg22E4) and in a 3.1 kb *Eco*RI fragment in *can* (AI1F5E4), generating a 3.6 kb *Eco*RI fragment (Sg19E3) containing 5'*set*, 3'*can* sequences (Figure 1C). A detailed restriction map of these fragments was constructed to localize the position of the translocation breakpoints more precisely. Small fragments around the translocation breakpoints (underlined in figure 1) were subcloned in M13 for ss DNA sequencing or in plasmid vectors for ds DNA sequencing. Sequences of the *wt* chromosome 9 and 6 fragments are shown in Figure 2, in which arrows indicate the position of the breakpoints.

In leukemic cells an aberrant fusion transcript can be detected by 3' *can* probes and not by 5' *can* probes (35, 38). This suggests that the *5'dek-3'can* and *5'set-3'can* fusion genes are important for the leukemogenic process but not their reciprocal counterparts. In concordance with these observations, two t(6;9) AML patients were reported whose leukemic cells contained the 6p- translocation chromosome, but had lost the reciprocal 9q+ chromosome (2, 30). During Ig/TCR rearrangement, addition of nucleotides takes place at the "coding joint". (4). The cloned breakpoints of patients PM and SE represent the expressed fusion genes on the 6p- and der.9 chromosome

respectively, but of patient DK only the fragment derived of the 9q+ chromosome was cloned. Therefore the translocation breakpoint of patient DK on the "coding chromosome" was isolated via PCR. From the sequence around the breakpoint, two primers were chosen that would generate a fragment of 270 bp representing the 6p- chromosome. The amplified fragment was subcloned and sequenced.

In figure 3, the sequences are shown of *wt* and fusion fragments encompassing the translocation breakpoints. The fusion fragments of

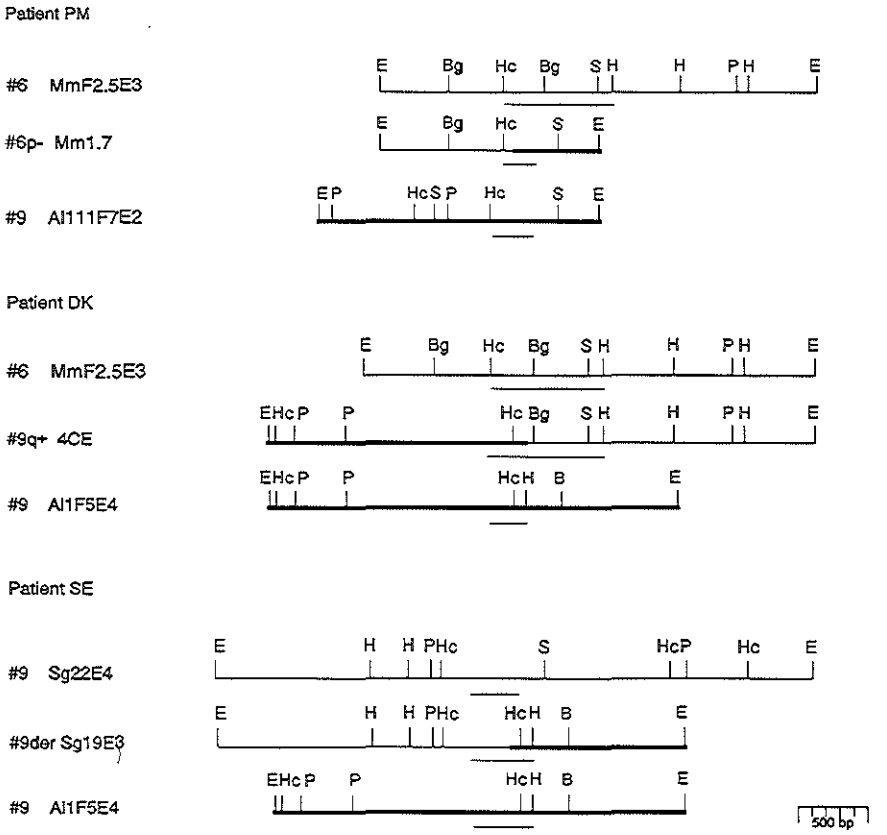


Figure 1. Localization of the translocation breakpoints of patients PM and DK (t(6;9) AML) and patient SE (AUL) by restriction enzyme mapping. Bold lines indicate *can* sequences, thin lines indicate *dek* or *set* sequences. The scale is indicated by a bar of 500 bp. The chromosomal origin of the mapped fragments is indicated (#), further details can be found in the text. Underlined subfragments were sequenced. B:*Bam*HI, Bg:*Bgl*I, E:*Eco*RI, H:*Hin*dIII, Hc:*Hin*cII, P:*Pst*I, S:*Sst*I.

patients PM and SE show a neat recombination of sequences derived from *can* with sequences from *dek* and *set*, respectively. Only of patient DK, fusion fragments from both translocation chromosomes were sequenced, and comparison of both fragments with *wt can* and *dek* sequences revealed a more complex situation. Of chromosome 9, 46 nucleotides (nt) are present on both the 6p- and the 9q+ chromosome and 31 nt of chromosome 6 are missing from both translocation chromosomes.

icb-9 (*can*) breakpoints patients DK, SE

```
  1 TCATATTTTGCAATGTGAATTGTGCTGCTGTAAACATGCGTGTGCAAGTGTCTTTTCATGTAATGACTGTTTTCCTCTGCGTAGATACAC
 91 AGTAGTAGGATTGCTGGATCAAATGGTAGGTACTTTTAGTTCTTTAAGAAATCTCCATATTGTTTTCTGTAGTGGTTGTACTAGTTTACA
181 TTCCCAACAGCAGTGTAAAAGTGTTCCCTTTTCACCACATCCACACCAACATCTATTATTTTTTGATTTTTAAATTACGGCCATTCTTGC
                                                          ^bp SE
271 AGAAGTAAGACGGTATTGCATTGTGGTTTTGATTTGCATTTGCCTGATAATTAGTGATGTTAAGCATTTTTTTGGATGTTTGTCAACCAT
                                                                    ^bp 6p- DK
361 TTGTATATCTTTTAAGAATTGTCTGTTCATGTCCTTAGCCCACTTTTTGATGGGATTGTTTGTTTTGTTCTTGCTGATATGTTTGAGTTC
                    ^bp 9q+ DK
451 CTTATAGATTCTGGATATTAGTCCATTATCGAATGCATAGTTTGTGAAAATTTTCTCCCACTCTGTGAGTTGTCTGTTTACTCTGCTGAT
541 TATTTCTTTTGCTGTGCAGAAGCTT
```

set breakpoint patient SE

```
  1 TAACAGTGAAACTTAGCACAACTAGTGGTTTTACATGCTGGTGAGCGTCTAGGGCTGTCAGGTTATAGGTCACAATAACGTTTATTTTCG
 91 GGAATGTAAAAAATCGGGAGCCCTTTAAGGTAGTAGTAGTTCACTCAACTATTGTGATCTCCTGAGGTGCCACTAATATCTGCAAATCTT
181 GTAAGAGAACTAGGCTGTGAGAAAACAAAATGGCAACCTTACTGTACCAAAAACTTCCATACACAGTAGCCAGCATGCCTTGT
                  ^bp SE
```

icb-6 (*dek*) breakpoints patients DK, PM

```
  1 GTCGACTTTCTGTCTCACTGATCTCACTAATGTTGCAGTGTGGTGTTAAAGTCTACCATTATTATTGTGTGGGAGTCTACGTCTCTTTGT
 91 AGGTCTCTAAGGACTTGCTTTATGAATCTGGGTGCTCCTGTATTGGGTGCATATATATTCAGGATAGTTAGCTCTCCTTGTTGAATTGAT
                ^bp PM                                                          ^bp 6p- DK
181 CCCTTTACCATTATGTAATGCCCTTCTTTGTCTCTTTTGATCTTTGTTGGTTTAAAGTCTGTTTTATCAGAGACTAGGATTGCAACCCCT
                    ^bp 9q+ DK
271 GTTTTTTTTGTTTTCCATTTGCTTGGTAGATCTTCCTCCATCCCTTCATTTTGAGCCTATGTGTGTCTCTGCACATGAGATGGGTCTCCT
361 GAATACAGCACACTGATGGGTCTTGACTCTTTATCCAGTTTGTCAGTCTGTGTCTTTAATTGGAGCATTTAGCCCATTTACATTTAAGGT
451 TAATGTTGTTATGTGTGAATTTGATCCTGTCATTATGATGTTAGCTGGTTATTTTGCTCGTTAGTTGATGCAGTTTCTTCCTAGCATCGA
541 TGGTCTTTACAATTTGGCATGTTTTTGCAGTGGCTGGTACCGGTTGTTCCTTTCCATGTTTAGTGCTTCCTTCAGGAGCTCTTGTAAGGC
631 AGGCCTGGTCGTGACAAAATCTCTCAGCATTTGCTTGTCTGTAAAGGATTTTATTTCTCCTTCACTTATGAAGCTT
```

icb-9 (*can*) breakpoint patient PM

```
  1 GTCAACAGTCGCCCAAAATTTAAATAAAATTATTGCAGGCCTATAATAAGTTAAATAGCTAAAATTTTAAATAATCACAGATTCAGTTTT
 91 TAGTGCTGACAGTGTTCTTTGATTTTGCAAACAAATGAGTATTTCTCAAATGGGAAGACGTCTTATATGTTCTATGCTGTGAATAGATAG
                                                                    ^bp PM
181 GTTTAGAATTACTTTCAGCACCGTTTTGTCTCCATTACAGTTAATTTTATGGGTGGGAGAGCAAAATCTAAATGGATGCACTGTCTGAGT
271 ACCAGAATGAATGGAAAA
```

Figure 2. Nucleotide sequence of the genomic *wt dek, set* and *can* DNA fragments containing the translocation breakpoints of patients DK, PM and SE. Fragments that were sequenced are underlined in figure 1. Translocation breakpoints are indicated by arrows.

Despite the presence of TdT in the nucleus of leukemic cells of at least two out of three patients, no addition of random nucleotides had occurred at the translocation breakpoints of any of these patients. Close inspection of the sequences presented in Figure 2 failed to reveal heptamer-nonamer sequences, purine-pyrimidine tracts, a *chi*-octamer motif or *Alu*-repeats. However, sequence comparison with the EMBL database showed that the *can* sequence of the reverse strand of fragment AI1F5E4 (containing breakpoints of DK and SE) is 76.8% homologous to nt 5009- 5579 of the human LINE-1 repeat (L1Hs). The reverse strand *dek* sequence of fragment MmF2.5E3 (containing breakpoints of DK and PM) is 94.1% homologous to nt 1734-2442 of the L1Hs repeat. Sequences from fragments AI1F7E2 and SE22E4 showed no homology to the LINE-1 element.

To determine the length and position of the LINE-1 element in *icb*-9 and *icb*-6, small stretches of DNA were sequenced all over the introns (figure 4). The homology of sequences overlapping the LINE-1 ranged from 70 to 98%. The repeat present in *icb*-9 of *can* consists of the 3' 1.7 kb of a LINE-1 element, the repeat in *icb*-6 consists of a near full length, 6 kb element. The sparse sequences available of the LINE-1 elements in both introns do not show gross rearrangements of the repeats. However, the consensus restriction pattern, e.g. the 1.9 kb *Hin*dIII repeat, is not present in these elements.

In *icb*-9 the LINE-1 repeat is located in the middle of the intron around the *Bam*HI and *Hin*dIII sites. The repeat covers ~23% of the intron and

```
Patient PM                              ↓
#6 (dek)   cgtctctttgtaggtctctaaggacttgctttatgaatctgggtgctcct
#6p-       cgtctctttgtaggtctctaatgacAGATAGGTTTAGAATTACTTTCAGC
#9 (can)   GTCTTATATGTTCTATGCTGTGAATAGATAGGTTTAGAATTACTTTCAGC
                                        ↑
Patient DK              ↓      deletion 31 bp
#6 (dek)   gttagctctccttgttgaattg-//-attatgtaatgcccttctttgtct
#6p-       gttagctctcctGTCAACCATT-//-CATGTCCTTAGCCCACTTTTTGAT
#9q+       TTTTGGATGTTTGTCAACCATT-//-CATGTCCTTAgccctttctttgtct
#9 (can)   TTTTGGATGTTTGTCAACCATT-//-CATGTCCTTAGCCCACTTTTTGAT
                              duplication 46 bp   ↑

Patient SE                              ↓
#9 (set)   atcttgtaagagaactaggctgtgagaaaacaaaatggcaaccttactgt
#9der.     atcttgtaagagaactaggctgtgaATTTTTTGATTTTTAAATTACGGCC
#9 (can)   CACCACATCCACACCAACATCTATTATTTTTTGATTTTTAAATTACGGCC
                                        ↑
```

Figure 3. Alignment of the nucleotide sequences of *wt* and derivative translocation chromosomes immediately next to the translocation breakpoint. The *can* sequence is in uppercase, *dek* and *set* sequences are in lowercase letters.

contains the translocation breakpoints of patients DK and SE but not the one of patient PM. In *icb*-6 the LINE-1 repeat covers 66% of the 9 kb intron. The translocation breakpoints of patients DK and PM are located in the middle of the repeat in a sequence that is deleted by truncation of the repeat in *icb*-9.

All sequences around the translocation breakpoints, located in a L1 element or not, are clearly enriched in A/T and may be a target for the nucleosome phasing α-protein (31).

DISCUSSION

Although the function of TdT seems to be restricted to lymphoid cells, its presence can be detected in the majority of acute myeloid leukemias (3). Coexpression of myeloid markers and TdT is correlated with an immature phenotype, though this association is not absolute. It has been shown that TdT can be active in rearrangements that do not involve Ig or TCR genes. The t(1;14), occurring in 3% of T cell ALL, translocates the *tal* oncogene into the TCR α/δ locus (5, 9, 17). A near consensus heptamer sequence is present in the *tal* gene near the translocation breakpoint and addition of random nucleotides was found at the breakpoint. Surprisingly, a precise 90 kb deletion involving the *tal* gene was found in 25% of the T-ALL patients with an apparently normal karyotype (8). Heptamer sequences and addition of random nucleotides are found at the site of rearrangement, implicating active participation of recombinase and TdT in this process, without participation of the TCR genes. Adriaansen et al. suggested that t(6;9) AML may be a leukemic subtype that can be characterized by relatively high numbers of leukemic cells expressing myeloid markers as well as TdT (2). Also leukemic cells of AUL patient SE, carrying a *set-can* fusion gene, expressed both TdT and myeloid markers (1). In contrast to t(6;9) AML, rearrangement of TCRß, TCRγ and TCRδ had occurred in the leukemic cells of patient SE (1). However, sequence comparison of the translocation breakpoints with the *wt dek, set* and *can* counterparts shows that no addition of random nucleotides had taken place. The variety of phenotypes that can be seen in t(6;9) AML suggests that the translocation takes place in an early precursor cell (13, 30). The most plausible explanation is that no TdT was present in the cell at the time the translocation event occurred.

In Ph' acute leukemia (ALL and ANLL), rearrangement of *IgH* and *TCR* genes is a common finding (10b). However, analysis of the t(9;22) breakpoints never showed involvement of nonamer-heptamer sequences or addition of random nucleotides at the translocation breakpoints. Since (*i*) rearrangements of *Ig* genes are polyclonal and (*ii*) the t(9;22) is supposed to occur in a very early cell type, also here the translocation probably takes place before TdT is expressed.

Although *Alu*-repeats were found to adjoin t(9;22) breakpoints (10, 14), they were not detected near the breakpoints of *dek-can* and *set-can* fusions. Instead, homology was found with the L1 repetitive DNA element (reviewed in (15, 29)). Sequencing short fragments all over the two introns showed that *icb*-9 contains a truncated 3' part of a LINE-1 repeat, whereas *icb*-6 contains an almost full length element. A sequence at the 3' end of the L1 element is present in the *icb*-6 repeat, but not in the *icb*-9 repeat, which may lack the last few hundred nucleotides. Both elements have an antisense orientation relative to *dek* and *can*. The homology of intron sequences with L1Hs varies between 75 and 98%. As the incidence of L1 elements is one in 150 kb for truncated elements and one in 750 kb for intact L1 elements the presence of a L1 element in both introns is a remarkable coincidence (29). The chance of finding two elements in ~17 kb of DNA is around 0.15%. Inversions and deletions are a common finding in L1 repeats, but the random sequences obtained of the L1 repeats in these two introns (Figure 4) show no major rearrangements of the elements.

Involvement of L1 repeats has been reported for translocations mediated by the Ig/TCR recombinase (25, 26, 39). In these cases sequences resembling the nonamer-heptamer sequence could be detected in the L1 element next to the recombination site as well as in the juxtaposed DNA sequence. However, near the translocation breakpoints of PM, DK and SE

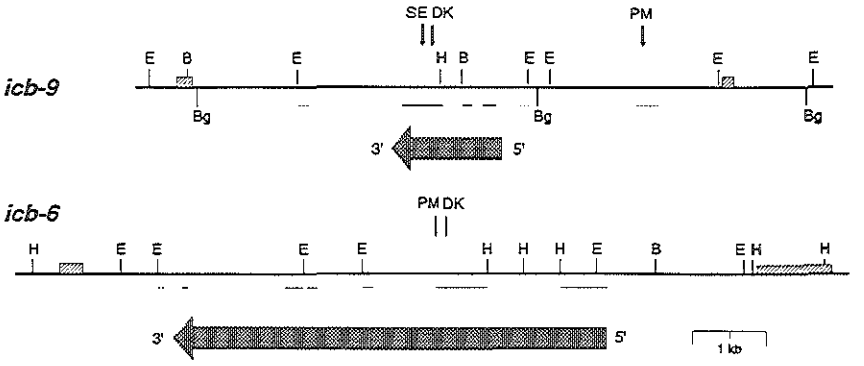


Figure 4. Representation of the introns containing the breakpoints in *can* and *dek*, *icb*-9 and *icb*-6, respectively. The position of translocation breakpoints of patients DK, PM and SE is indicated with arrows. Striped boxes indicate the position of exons. The position and orientation of the LINE-1 element is indicated with a large arrow below the restriction maps. DNA stretches that were sequenced to determine the position and length of the L1 repeat are indicated with short lines underneath the restriction map. Continuous lines indicate 75-95% homology with L1Hs, dotted lines indicate no homology with L1Hs. The scale is indicated with a bar of 1 kb. B:*Bam*HI, Bg:*Bgl*II, E:*Eco*RI, H:*Hin*dIII.

no nonamer-heptamer consensus can be detected. Moreover, t(6;9) breakpoints are scattered all over the repeat and adjoining sequences (30).

It can not be settled whether the L1 elements might have promoted the translocation. As *icb*-9 contains a truncated element, in which the L1 sequences containing the translocation breakpoints on chromosome 6 are absent, there was no homologous recombination between the repeat sequences. A relevant observation may be the high incidence with which aberrations seem to occur within L1 sequences (15). Since L1 elements were found to be involved in translocations, it has been suggested that the chromatin structure of the repeats may be important. The same factors that enhance the occurrence of rearrangements within the elements, may be involved in the translocation event associated with the acute leukemias. In addition, pairing of *icb*-9 and *icb*-6 via the L1 repeats could play a role in the translocation event.

No L1 element was detected at the translocation breakpoint in the *set* gene of SE. However, the presence of a L1 element a little more downstream of the *set* gene has not been excluded.

A/T rich regions were reported by Chen et al. and Papadopoulos et al. (10, 24). They are very abundant in the sequences of all translocation breakpoints studied in this report. The combined sequences of all three breakpoints contain more than 65% A/T. This may influence the chromatin structure, making it more accessible to presumed recombinases. The A/T stretches may also be a target for the nucleosome phasing α-protein (31). This protein can bind three A/T rich sites and perhaps could bring together two (broken) chromosomes containing such sites resulting in the observed recombination.

It is interesting that at the translocation breakpoints in the *set* and *can* genes of patient SE, the A/T rich regions are composed of dAdT stretches. These sequences could have formed a triple helix structure, which can be stabilized by a recombinase protein that binds to triple helices (18, 20). This might have served as the contact point between the two genes during the translocation event. Subsequently, the chromosomes must have broken right next to the triple helix structure and were ligated together.

Only of patient DK both reciprocal fusion fragments were sequenced. Analysis of genomic phages containing the reciprocal breakpoints of patient SE revealed a large deletion of unknown size, adjacent to the translocation breakpoint and 3' of the *set* gene (38). Therefore, analysis of the breakpoint fragment on the reciprocal translocation chromosome was judged to be irrelevant. Although it was attempted, we were not able to isolate the breakpoint fragment of patient PM on the 9q+ chromosome by PCR.

Analysis of the reciprocal fusion fragments of patient DK reveals a particular structure. 31 bp of chromosome 6 are lost. Loss of sequences is a feature regularly encountered in other translocations as well. Howe-

ver, we were surprised to find a duplication of 46 bp of chromosome 9. Although the mechanism of the translocation remains elusive, a possible explanation for the duplication may be the formation of a staggered break with sticky ends of at least 46 nt that were repaired by DNA polymerase before the chromosome 6 ends were fused.

ACKNOWLEDGMENTS

We thank D. Bootsma for continuous support and discussion, J. Abels, A. Hagemeijer, B. Löwenberg and F. Mitelman for patient material. This work is supported by the Dutch Cancer Foundation.

LITERATURE CITED

1. Adriaansen, H.J., P.W.C. Soeting, I.L.M. Wolvers-Tettero and J.J.M. Van Dongen. 1991. Immunoglobulin and T cell receptor gene rearrangements in acute non-lymphocytic leukemia. Leukemia 5:744-751.
2. Adriaansen, H.J., J.J.M. van Dongen, H. Hooijkaas, K. Hählen, M.B. van 't Veer, B. Löwenberg and A. Hagemeijer. 1988. Translocation (6;9) may be associated with a specific TdT-positive immunological phenotype in ANLL. Leukemia 2:136-40.
3. Adriaansen, H.J., J.J.M. van Dongen, M.C. Kappers-Klunne, K. Hählen, M.B. van 't Veer, J.H. Wijdenes-de Bresser, A.C. Holdrinet, E.J. Harthoorn-Lasthuizen, J. Abels and H. Hooijkaas. 1990. Terminal deoxynucleotidyl transferase positive subpopulations occur in the majority of ANLL: implications for the detection of minimal disease. Leukemia 4:404-10.
4. Alt, F.W. and D. Baltimore. 1982. Joining of immunoglobulin heavy chain gene segments: implications from a chromosome with evidence of three D-JH fusions. Proc. Natl. Acad. Sci. USA 79:4118-22.
5. Begley, C.G., P.D. Aplan, M.P. Davey, K. Nakahara, K. Tchorz, J. Kurtzberg, M.S. Hershfield, B.F. Haynes, D.I. Cohen, T.A. Waldmann and I.R. Kirsch. 1989. Chromosomal translocation in a human leukemic stem-cell line disrupts the T-cell antigen receptor delta-chain diversity region and results in a previously unreported fusion transcript. Proc. Natl. Acad. Sci. USA 86:2031-5.
6. Boehm, T., L. Mengle-Gaw, U.R. Kees, N. Spurr, I. Lavenir, A. Forster and T.H. Rabbitts. 1989. Alternating purine-pyrimidine tracts may promote chromosomal translocations seen in a variety of human lymphoid tumours. Embo J 8:2621-31.
7. Boehm, T. and T.H. Rabbitts. 1989. The human T cell receptor genes are targets for chromosomal abnormalities in T cell tumors. FASEB J. 3:2344-2359.
8. Brown, L., J.T. Cheng, Q. Chen, M.J. Siciliano, W. Crist, G. Buchanan and R. Baer. 1990. Site-specific recombination of the tal-1 gene is a common occurrence in human T cell leukemia. Embo J 9:3343-51.
9. Chen, Q., J.T. Cheng, L.H. Tasi, N. Schneider, G. Buchanan, A. Carroll, W. Crist, B. Ozanne, M.J. Siciliano and R. Baer. 1990. The tal gene undergoes chromosome translocation in T cell leukemia and potentially encodes a helix-loop-helix protein. Embo J 9:415-24.
10. Chen, S.J., Z. Chen, L. d'Auriol, M. Le Coniat, D. Grausz and R. Berger. 1989. Ph1 + bcr- acute leukemias: implication of Alu sequences in a chromosomal translocation occurring in the new cluster region within the BCR gene. Oncogene 4:195-202.
10b. Chen, S.J., G. Flandrin, M.-T. Daniel, F. Valensi, L. Baranger, D. Grausz, A. Bernheim, Z. Chen, F. Sigaux and R. Berger. 1988. Philadelphia-positive acute leukemia: lineage promiscuity and inconsistently rearranged breakpont cluster region. Leukemia 2:261-273.

11. Cleary, M.L. 1991. Oncogenic conversion of transcription factors by chromosomal translocations. Cell 66:619-22.
12. Croce, C.M. 1987. Role of chromosome translocations in human neoplasia. Cell 49:155-6.
13. Cuneo, A., S. Kerim, E. Vandenberghe, A. Van Orshoven, J. Rodhain, A. Bosly, P. Zachee, A. Louwagie, J.-L. Michaux, P. Dal Cin and H. Van den Berghe. 1989. Translocation t(6;9) occuring in acute myelofibrosis, myelodysplastic syndrome, and acute nonlymphocytic Leukemia suggests multipotent stem cell involvement. Cancer Genet. Cytogenet. 42:209-219.
14. De Klein, A., T. Van Agthoven, C. Groffen, N. Heisterkamp, J. Groffen and G. Grosveld. 1986. Molecular analysis of both translocation products of a Philadelphia-positive CML patient. Nucl. Acids Res. 14:7071-7082.
15. Fanning, T.G. and M.F. Singer. 1987. LINE-1: a mammalian transposable element. Biochim. Biophys. Acta 910:203-212.
16. Finger, L.R., R.C. Harvey, R.C. Moore, L.C. Showe and C.M. Croce. 1986. A common mechanism of chromosomal translocation in T- and B-cell neoplasia. Science 234:982-985.
17. Finger, L.R., J. Kagan, G. Christopher, J. Kurtzberg, M.S. Hershfield, P.C. Nowell and C.M. Croce. 1989. Involvement of the TCL5 gene on human chromosome 1 in T-cell leukemia and melanoma. Proc Natl Acad Sci U S A 86:5039-43.
18. Hsieh, P., C.S. Camerini-Otero and R.D. Camerini-Otero. 1990. Pairing of homologous DNA sequences by proteins: evidence for three-stranded DNA. Genes & Development 4:1951-1963.
19. Kazazian HH, J.r., C. Wong, H. Youssoufian, A.F. Scott, D.G. Phillips and S.E. Antonarakis. 1988. Haemophilia A resulting from de novo insertion of L1 sequences represents a novel mechanism for mutation in man. Nature 332:164-6.
20. Kiyama, R. and D. Camerini-Otero. A triplex DNA-binding protein from human cells: purification and characterization. Science, in press.
21. Krowczynska, A.M., R.A. Rudders and T.G. Krontiris. 1990. The human minisatellite consensus at breakpoints of oncogene translocations. Nucleic Acids Res 18:1121-7.
22. Landau, N.R., D.G. Schatz, M. Rosa and D. Baltimore. 1987. Increased frequency of N-region insertion in a murine pre-B-cell line infected with a terminal deoxynucleotidyl transferase retroviral expression vector. Mol. Cell. Biol. 7:3237-43.
23. Morse, B., P.G. Rotherg, V.J. South, J.M. Spandorfer and S.M. Astrin. 1988. Insertional mutagenesis of the myc locus by a LINE-1 sequence in a human breast carcinoma. Nature 333:87-90.
24. Papadopoulos, P.C., A.M. Greenstein, R.A. Gaffney, C.A. Westbrook and L.M. Wiedemann. 1990. Characterization of the translocation breakpoint sequences in Philadelphia-positive acute lymphoblastic leukemia. Genes Chromosom. Cancer 1:233-9.
25. Pear, W.S., S.F. Nelson, H. Axelson, G. Wahlstrom, H. Bazin, G. Klein and J. Sumegi. 1988. Aberrant class switching juxtaposes c-myc with a middle repetitive element (LINE) and an IgH intron in two spontaneously arising rat immunocytomas. Oncogene 2:499-507.
26. Rabbitts, P.H., J. Douglas, P. Fischer, E. Nacheva, A. Karpas, D. Catovsky, J.V. Melo, R. Baer, M.A. Stinson and T.H. Rabbitts. 1988. Chromosome abnormalities at 11q13 in B cell tumours. Oncogene 3:99-103.
27. Sanger, F., S. Nicklen and A.R. Coulson. 1977. DNA Sequencing with Chain-terminating Inhibitors. Proc. Natl. Acad. Sci. USA 74:5463-5467.
28. Schmid, C.W. and W.R. Jelinek. 1982. The Alu family of dispersed repetitive sequences. Science 216:1065-70.
29. Singer, M.F. and J. Skowronski. 1985. Making sense out of LINES: long interspersed repeat sequences in mammalian genomes. Trends Biochem. Sci. 10:119-122.
30. Soekarman, D., M. von Lindern, S. Daenen, B. de Jong, C. Fonatsch, B. Heinze, C.R. Bartram, A. Hagemeijer and G. Grosveld. The translocation (6;9)(p23;q34) shows consistent rearrangement of two genes and defines a myeloproliferative disorder with specific clinical features. Blood, in press.

31. Strauss, F. and A. Varshavsky. 1984. A protein binds to a satellite DNA repeat at three specific sites that would be brought into mutual proximity by DNA folding in the nucleosome. Cell 37:889-901.
32. Swergold, G.D. 1990. Identification, characterization, and cell specificity of a human LINE-1 promoter. Mol. Cell. Biol. 10:6718-29.
33. Tonegawa, S. 1983. Somatic generation of antibody diversity. Nature 302:575-81.
34. van der Feltz, M.J., M.K. Shivji, P.B. Allen, N. Heisterkamp, J. Groffen and L.M. Wiedemann. 1989. Nucleotide sequence of both reciprocal translocation junction regions in a patient with Ph positive acute lymphoblastic leukaemia, with a breakpoint within the first intron of the BCR gene. Nucleic Acids Res. 17:1-10.
35. von Lindern, M., M. Fornerod, S. van Baal, M. Jaegle, T. de Wit, A. Buijs and G. Grosveld. The t(6;9), associated with a specific subtype of AML, results in the fusion of two genes, dek and can, and the expression of a chimaeric, leukemia specific dek-can mRNA. Mol. Cell. Biol., in press.
36. von Lindern, M., A. Poustka, H. Lerach and G. Grosveld. 1990. The (6;9) chromosome translocation, associated with a specific subtype of acute nonlymphocytic leukemia, leads to aberrant transcription of a target gene on 9q34. Mol. Cell. Biol. 10:4016-26.
37. von Lindern, M., T. van Agthoven, A. Hagemeijer, H. Adriaansen and G. Grosveld. 1989. The human pim-1 gene is not directly activated by the translocation (6;9) in acute nonlymphocytic leukemia. Oncogene 4:75-9.
38. von Lindern, M., S. van Baal, J. Wiegant, T. Raap, A. Hagemeijer and G. Grosveld. Can, a putative oncogene associated with myeloid leukemogenesis, can be activated by fusion of its 3' half to different genes: characterization of the set gene. Submitted for publication
39. Yancopoulos, G.D., R.A. DePinho, K.A. Zimmerman, S.G. Lutzker, N. Rosenberg and F.W. Alt. 1986. Secondary genomic rearrangement events in pre-B cells: VHDJH replacement by a LINE-1 sequence and directed class switching. Embo J. 5:3259-66.

CHAPTER 3

DISCUSSION

3.1 DEK-CAN and leukemogenesis

Putative function of DEK-CAN.

As is shown in the previous chapter, the generation of a *dek-can* fusion gene is a consistent feature of t(6;9) acute myeloid leukemia (AML). Interestingly, a *set-can* fusion gene was isolated from leukemic cells of a patient with acute undifferentiated leukemia (AUL). This suggests that the C-terminal part of CAN contains domains responsible for the leukemogenic activity of the fusion proteins. These domains can be activated by fusion to either DEK or SET sequences. Genomic *can* probes that detect aberrant fragments due to translocation breakpoints in *icb*-9, detected aberrant fragments in bone marrow cells without a t(6;9), of two more patients (33). Aberrant fragments could not be detected with *dek* or *set* probes in these patients. One patient suffered from RAEB, the other patient from cALL. This implies that the *can* gene may be activated by fusion to yet other genes.

The possible function of CAN, DEK and SET is still elusive. DEK and SET are non-homologous proteins, they only share a stretch of acidic aa residues. These acidic domains could therefore play a role in the activation of CAN. The C-terminal part of CAN contains a putative dimerization domain and a motif present in a number of transcription factors that may have an ancillary DNA binding capacity. A possible function of acidic domains is transcriptional activation, which prompted us to consider a function for DEK-CAN or SET-CAN in transcription complexes.

If DEK-CAN has a role in transcription regulation, its intracellular localization should be nuclear. Polyclonal antibodies against DEK and CAN were raised in rabbits that specifically recognize DEK and CAN proteins respectively (11). After transient expression of DEK and CAN at high levels in COS cells, immunocytochemical analysis indicated that the localization of CAN is mainly cytoplasmic in these cells, while DEK and DEK-CAN are localized exclusively in the nucleus. CAN was detected in large vesicles in the cytoplasm. The origin of these vesicles is unknown, but they are clearly different from lysosomes or mitochondria and may represent an artifact of overexpression. Markedly, the DEK-CAN protein is located in the nucleus. Fusion of the C-terminal part of CAN to DEK results in translocation of CAN from the cytoplasm to the nucleus. These results support the hypothesis that DEK-CAN could be involved in transcription complexes.

Recently, two other fusion genes involved in transcription regulation were isolated from leukemic cells (reviewed by (5)). As discussed in chapter 1.3.4, the E47-PBX fusion protein may target transcription activation by E47/E12 to genes regulated by the PBX homeoprotein. The PML-RARα protein may result in retinoic acid dependence of PML function or in an impaired functioning of RARα. In addition, the retroviral MYB-ETS fusion protein is much more competent in inducing erythroleukemia than cooperation of v-MYB and v-ETS as separate proteins. These examples illustrate different

mechanisms to deregulate the activity of transcription factors through the formation of fusion genes. Although DEK-CAN and SET-CAN could have a putative role in transcription regulation, none of the consensus DNA binding motifs (14) were recognized in DEK, SET or CAN. However, it has been hypothesized that in addition to several classes of DNA binding transcription factors, there exist 'adaptor' proteins that bind to transcription factors, to bring them in close contact with the basic transcription machinery (26). These putative adaptor proteins thus provide the interaction between transcription factors and the RNA polymerase complex. If CAN is such an adaptor molecule, whose specificity is modulated by fusion to the acidic domains of DEK or SET, the effect would be analogous to the fusion of E47 to PRL: dislocation of a transcription activation domain, resulting in an aberrant transcription pattern of target genes.

Transcription factors and hematopoiesis.

Molecular characterization of chromosomal aberrations in leukemia has identified a number of genes with a potential leukemogenic activity (chapter 1.3). A relatively large number of these genes encodes transcription factors. As deregulation of transcription factors can directly cause abnormal expression of genes regulating proliferation and differentiation, a potential oncogenic effect can be anticipated. Characterization of the leukemogenic effect of transcription factors may at the same time reveal normal mechanism of hematopoietic control. Little is known about transcription factors regulating hematopoiesis.

The role of homeobox genes in the regulation of differentiation in *Drosophila* development is well established. Homologous genes are also expressed in specific hematopoietic lineages and differentiation stages (22, 27). The ectopic expression of HOX11 in T-ALL (9, 15, 21) may activate normally silent genes, that inhibit T-cell differentiation. Alternatively, HOX11 could bind to the DNA binding site recognized by another homeobox protein and thereby, through competition, impair the normal function of this transcription factor in differentiation. Ectopic expression of rhombotins (2) may interfere with the same factor as HOX11, since the LIM motif is a dimerization domain found in homeobox proteins. Analysis of the function of HOX11 and rhombotins in T-cell differentiation may identify genes involved in lymphoid differentiation.

In contrast to HOX11 and Rhombotins, the basic stretch-helix-loop-helix proteins SCL and LYL are specifically expressed in hematopoietic cells, although their expression is restricted to the myeloid lineage. Due to translocation into the TCR locus, they are expressed at elevated levels in T-cells (1, 4, 10, 31). The leukemogenic effect in T-cells may be related to their normal function in myeloid cells. Markedly, SCL expression is maximal in proliferating cells. As T-ALL is a disease with somewhat uniform clinical features, HOX11, Rhombotins, SCL and LYL may act upon genes with analogous functions, or even, perhaps, on the same set of responder genes (32).

In the case of fusion genes, both partners may have a role in proliferation or differentiation of hematopoietic cells. As discussed above, fusion of two transcription factors can direct transcription activation of genes that are normally not acted upon. Further investigation into the function of PML and PBX will reveal what role these factors have in (dis)regulation of the hematopoietic system. Although aberrant expression of some genes may interfere with the function of specific hematopoietic factors, other gene products may have a more general function in the cell like c-MYC (6). Analysis of these genes will reveal general, basic principles in regulation of proliferation and differentiation. CAN as well as DEK and SET are expressed in hematopoietic cells. However, their ubiquitous expression argues against a specific role in hematopoiesis as may be the case for SCL or LYL. Nevertheless, the effect of the DEK-CAN fusion gene could be restricted to hematopoietic cells. The dimerization motifs in CAN could result in the formation of heterodimers with cell type specific factors. As the dimerization motif C-terminal of the translocation breakpoint is homologous to the dimerization motif of the estrogen receptor, it is tempting to speculate on interaction of CAN with the estrogen receptor. Whether the estrogen receptor plays a role in the regulation of hematopoiesis is not known.

Leukemia subtypes.

Chromosomal aberrations are associated with specific leukemia subtypes. Translocations involving TCR loci give rise to T-cell malignancies, which is not surprising as the enhanced expression of translocated genes is restricted to T-cells, since these genes are controlled by the TCR enhancers. Similarly, translocations involving Ig loci result in B-cell malignancies. The c-*myc* gene can be activated by translocation into the Ig- or TCR loci (28), though translocation into the IgH locus is most frequent. This may be due to the translocation mechanism, since in Burkitt's lymphoma the site of the translocation breakpoint relative to c-*myc* varies with the Ig locus involved. The fact that *Rhombotins* and *HOX11* are involved in T-ALL and not in B-ALL, may therefore be a consequence of the translocation mechanism as well. Alternatively, the encoded proteins may interfere with a factor specifically regulating T-cell, and not B-cell, differentiation. A lineage specific effect must also underlie the association of fusion genes with specific leukemia subtypes. T(15;17) is strictly associated with APL, while t(1;19) is associated with pre-B ALL. In addition to the association with a specific hematopoietic lineage, chromosomal aberrations can be associated with a specific differentiation stage within that lineage. Some of the lineage-, or differentiation stage specific associations can be explained, now that the genes involved are characterized. Factors involved in proliferation like BCL-1 (24, 37), or in cell death like BCL-2 (29), give relatively benign disease with large numbers of normally differentiating cells. In contrast, enhanced expression of genes that interfere with differentiation, like the *Rhombotin* genes, results in large numbers of immature cells and more acute neoplasia.

Consequently it can be argued that PML-RARα interferes with promyelocyte differentiation and E47-PBX with B-cell differentiation. T(9;22), generating a BCR-ABL fusion gene, is mainly found in chronic myeloid leukemia (CML), but also in acute leukemia both myeloid and lymphoid. (23). This suggests that BCR-ABL has a role in proliferation that is not lineage restricted. T(6;9), i.e. the *dek-can* fusion gene, characterizes AML, but is also detected in RAEB. Some patients were initially diagnosed as (blast crisis of) CML, because large numbers of differentiating cells are present in the bone marrow. Clearly, leukemic cells have a variable differentiation potential. Therefore, we assume that DEK-CAN has an effect on the proliferation of myeloid cells. The supposed effect of genes involved in leukemogenesis on proliferation and differentiation in the diverse hematopoietic lineages is indicated in table 1.

Multistep tumorigenesis in t(6;9) AML.

As already discussed in chapter 1.5, leukemogenesis is a multistep process. Activation of BCL-2 gives rise to low grade lymphoma, additional activation of c-MYC results in high grade lymphoma (8, 12, 30). Formation of a BCR-ABL fusion gene results in chronic myeloid leukemia, subsequent mutations may be needed for transition into blast crises. Likewise, t(6;9) AML is often preceded by a MDS phase and myelodysplastic features are regularly found in all non-leukemic lineages. Therefore it has been hypothesized that t(6;9) is a stem cell disorder underlying MDS features as well as acute leukemia (7, 19). Alternatively, mutations resulting in MDS and activation of CAN could be cooperating to generate acute leukemia. MDS is a hematological disorder in which differentiation is hampered, while DEK-CAN probably enhances proliferation. This could be a potent combination in leukemogenesis. Mutations causing MDS may precede the translocation event or occur later, enhancing the biological effect of the translocation in hematopoietic cells.

Nothing is known about mutations involved in progression of t(6;9) ANLL. Additional chromosomal aberrations that occur do not only comprise trisomy of chromosome 8, but also of chromosome 13. Unfortunately, it is not known what genes on these chromosomes could play a role in leukemia progression. The trisomy may produce a gene dosage effect, the additional chromosome may encode an altered gene product. Also a combination of these possibilities could effectuate progression.

Diagnosis and therapy

Since specific translocations characterize defined leukemia subtypes, detection of translocation breakpoints by Southern blotting or polymerase chain reaction (PCR) is a useful extension of karyotypic diagnosis. Moreover, the sensitivity of the PCR technique makes it extremely suitable for detection of minimal residual disease (MRD) after therapy. Translocations involving TCR or Ig genes are often very precise due to the role of heptamer-nonamer sequences in many lymphoid translocations (chapter 1.4), i.e. t(14;18) can

TABLE 1

	Proliferation	Differentiation
Myeloid cells	IL-3, BCR-ABL DEK-CAN	PML-RARα
B-cells	IL-3, BCR-ABL, BCL-1, BCL-2, BCL-3, MYC	E47-PBX
T-cells	MYC	SCL, LYL, HOX11, Rhombotins

A tentative classification of oncogene products with regard to their putative function in leukemogenesis. Deregulation of the hematopoietic system can be effectuated through increased proliferation or through inhibition of differentiation. Increased proliferation may be the effect of decreased cell death (BCL-2). As a result of the complex regulation of proliferation and differentiation, interference with differentiation will affect proliferation and vice-versa. SCL and LYL may have an effect on other lineages as well, but the proteins are only classified in a lineage where their effect has been observed.

be detected by PCR on DNA of leukemic cells. T(9;22) and t(6;9) result in the formation of specific fusion transcripts, but at the genomic the translocation breakpoints are scattered over relatively long introns. Consequently, these translocations can be sensitively monitored by RNA-PCR, but not by genomic PCR. The formation of an invariant *dek-can* transcript in t(6;9) ANLL appeared to be a valuable target for both diagnosis and detection of MRD by RNA-PCR (chapter 2.4).

Specific treatment of leukemia, antagonizing the effect of a fusion gene or abnormal gene expression, is still far away. Nevertheless, improved diagnosis and characterization of leukemia subtypes will lead to improved protocols for chemotherapy. In addition, various approaches of immunotherapy are studied. Fusion proteins like DEK-CAN are extremely suitable targets for immunotherapy. Antibodies can be raised against a specific fusion peptide, which excludes side effects due to reactivity against the protein in its normal function. Although DEK-CAN is a nuclear protein, peptides of all intracellular proteins are presented in the context of MHC classI molecules at the cell surface. Kast et al. (20) raised cytotoxic T-lymphocytes (CTL) against cells transformed by adenovirus type 5 (Ad5) early region 1 (E1). These CTL clones stimulated by IL-2 eradicated very efficiently tumors of Ad5E1 transformed cells. This approach of immunotherapy may be suitable for leukemia. Once transgenic mice are generated that express the *dek-can* fusion gene, they may develop leukemia. If so, CTLs raised against *dek-can* transformed, syngeneic mouse cells, can be tested in these mice for their

antileukemic effect.

3.2 Future directions.

To provide evidence that the product of the translocation (6;9), the DEK-CAN protein, is directly involved in leukemogenesis, the presumed transforming activity must be demonstrated.

Hematopoietic cell lines will be transfected with *dek, can* and *dek-can* cDNA constructs to assay the capacity of these genes to render these cells factor independent, and to assay their tumorigenicity in nude mice and in syngeneic animals. Some cell lines still have a limited differentiation potential, therefore, these cells can be used to analyze the effect of DEK-CAN on differentiation. The phenotype of some t(6;9) ANLL patients indicates that DEK-CAN will not completely block differentiation, nevertheless, differentiation may be hampered.

Transgenic mice will be generated as well. Analysis of these mice will answer two questions:

(*i*) Does DEK-CAN have a specific hematopoietic effect? As DEK and CAN are expressed ubiquitously, a *dek-can* fusion gene may interfere with regulation of proliferation and differentiation in other tissues as well. To test the effect in transgenic mice, a *dek-can* fusion construct will be placed under control of the normal *dek* promoter. The *bcr* and *abl* genes are also expressed in all tissues. Transgenic mice, expressing *bcr-abl* from the ubiquitously expressed metallothionein-1 promoter (17), specifically developed hematopoietic disorders. This suggests that the tumorigenic effect of BCR-ABL is most pronounced in the hematopoietic system. However, the mice died 10-58 days after birth and neoplasia in other tissue may occur after longer latency periods. In contrast, a general role in tumorigenesis has clearly been demonstrated for BCL-1 (24) and MYC (6).

(*ii*) Does DEK-CAN have a direct effect or are additional mutations needed to generate leukemia? Transgenic mice harbouring a *bcr-abl* gene die within 10-58 days after birth (17) suggesting that BCR-ABL has an almost direct effect and needs no or few cooperating oncogenes to produce leukemia. However, transgenic mice harbouring an activated c-*myc* gene develop neoplasia only after a relatively long latency period. Introducing additional oncogenes reduces this latency period dramatically (16, 25, 34, 35). Dependent on the phenotype of *dek-can* transgenic mice, cooperating oncogenes can be selected by infection of the mice with Murine leukemia virus as described by van Lohuizen et al. (35, 36).

In addition, the presumed function of CAN and DEK in transcription regulation must be assayed. Synthesis and purification of large amounts of protein in the Baculovirus system or as Glutathion-S transferase (GST)-fusion proteins will provide the possibility to analyze DNA binding and protein-protein interactions in a *in vitro* system. As DNA binding and transcription

activation domains are able to function independently (3, 18) fusion proteins can be made between characterized transcription factors and CAN or DEK to screen protein domains for activity. Cross linking experiments may help to find proteins that dimerize with CAN (13). It will be of specific interest to investigate whether CAN can interfere with the function of the estrogen receptor with regard to the homology of CAN to the Estrogen receptor dimerization domain. Once transcription activation or DNA binding activity of CAN can be demonstrated, target genes may be found. These may be important genes regulating proliferation and differentiation in the myeloid lineage.

Presumably, CAN can be activated by fusion to DEK as well as SET. Therefore, it will be interesting to assay the minimal requirements for activation of CAN. A stretch of acidic amino acid residues may be sufficient. Possibly, truncation of CAN is essential as well. Deletion of N-terminal domains of CAN (e.g. the leucine zipper) may reveal mechanisms regulating the normal function of CAN. However, these experiments need a test system to assay the tumorigenic effect of DEK-CAN and SET-CAN and related constructs.

Finally, analysis of the function of *dek-can* and other genes affected by translocations, may identify genes (de)regulated by these oncogene products that are essential for the complex regulation of hematopoiesis. This may be a major contribution in our understanding of this process.

References

1. Begley, C.G., P.D. Aplan, S.M. Denning, B.F. Haynes, T.A. Waldmann and I.R. Kirsch. 1989. The gene SCL is expressed during early hematopoiesis and encodes a differentiation-related DNA-binding motif. Proc Natl Acad Sci USA 86:10128-32.
2. Boehm, T., L. Foroni, Y. Kaneko, M.F. Perutz and T.H. Rabbitts. 1991. The rhombotin family of cysteine-rich LIM-domain oncogenes: distinct members are involved in T-cell translocations to human chromosomes 11p15 and 11p13. Proc Natl Acad Sci USA 88:4367-71.
3. Brent, R. and M. Ptashne. 1985. A eukaryotic transcriptional activator bearing the DNA specificity of a prokaryotic repressor. Cell 43:729-736.
4. Chen, Q., J.T. Cheng, L.H. Tasi, N. Schneider, G. Buchanan, A. Carroll, W. Crist, B. Ozanne, M.J. Siciliano and R. Baer. 1990. The tal gene undergoes chromosome translocation in T cell leukemia and potentially encodes a helix-loop-helix protein. Embo J 9:415-24.
5. Cleary, M.L. 1991. Oncogenic conversion of transcription factors by chromosomal translocations. Cell 66:619-22.
6. Cole, M.D. 1991. Myc meets its Max. Cell 65:715-6.
7. Cuneo, A., S. Kerim, E. Vandenberghe, A. Van Orshoven, J. Rodhain, A. Bosly, P. Zachee, A. Louwagie, J.L. Michaux, P. Dal Cin and H. Van den Berghe. 1989. Translocation t(6;9) occurring in acute myelofibrosis, myelodysplastic syndrome, and acute nonlymphocytic leukemia suggests multipotent stem cell involvement. Cancer Genet Cytogenet 42:209-19.
8. De Jong, D., B.M. Voetdijk, G.C. Beverstock, G.J. van Ommen, R. Willemze and P.M. Kluin. 1988. Activation of the c-myc oncogene in a precursor-B-cell blast crisis of follicular lymphoma, presenting as composite lymphoma. N Engl J Med 318:1373-8.
9. Dube, I.A., S. Kamel-Reid, C.C. Yuan, C. Lu, X. Wu, G. Corpus, S.C. Raimondi, W.M. Crist,

A. Carroll and J.B. Baker. 1991. A novel homeobox gene is deregulated in lymphoid neoplasias with chromosomal translocation t(10;14). blood in press.
10. Finger, L.R., J. Kagan, G. Christopher, J. Kurtzberg, M.S. Hershfield, P.C. Nowell and C.M. Croce. 1989. Involvement of the TCL5 gene on human chromosome 1 in T-cell leukemia and melanoma. Proc Natl Acad Sci USA 86:5039-43.
11. Fornerod, M., M. von Lindern and G. Grosveld. manuscript in preparation.
12. Gauwerky, C.E., F.G. Haluska, Y. Tsujimoto, P.C. Nowell and C.M. Croce. 1988. Evolution of B-cell malignancy: pre-B-cell leukemia resulting from MYC activation in a B-cell neoplasm with a rearranged BCL2 gene. Proc Natl Acad Sci USA 85:8548-52.
13. Glass, C.K., O.V. Devary and M.G. Rosenfeld. 1990. Multiple cell type-specific proteins differentially regulate target sequence recognition by the alpha retinoic acid receptor. Cell 63:729-38.
14. Harrison, S.C. 1991. A structural taxonomy of DNA-binding domains. Nature 353:715-719.
15. Hatano, M., C.W. Roberts, M. Minden, W.M. Crist and S.J. Korsmeyer. 1991. Deregulation of a homeobox gene, HOX11, by the t(10;14) in T cell leukemia. Science 253:79-82.
16. Haupt, Y., W.S. Alexander, G. Barri, S.P. Klinken and J.M. Adams. 1991. Novel zinc finger gene implicated as myc collaborator by retrovirally accelerated lymphomagenesis in E mu-myc transgenic mice. Cell 65:753-63.
17. Heisterkamp, N., G. Jenster, H.J. ten, D. Zovich, P.K. Pattengale and J. Groffen. 1990. Acute leukaemia in bcr/abl transgenic mice. Nature 344:251-3.
18. Hope, I.A. and K. Struhl. 1986. Functional dissection of a eukaryotic transcriptional activator protein, GCN4 of yeast. Cell 46:885-94.
19. Horsman, D.E. and D.K. Kalousek. 1987. Acute Myelomonocytic Leukemia (AML-M4) and translocation t(6;9): two additional patients with prominent myelodysplasia. Am J Haematol 26:77-82.
20. Kast, W.M., R. Offringa, P.J. Peters, A.C. Voordouw, R.H. Meloen, A.J. van der EB and C.J.M. Melief. 1989. Eradication of adenovirus E1-induced tumors by E1A-specific cytotoxic T lymphocytes. Cell 59:603-614.
21. Kennedy, M.A., R. Gonzales-Sarmiento, U.R. Kees, F. Lampert, N. Dear, T. Boehm and T. Rabbitts. 1991. HOX11, a homeobox-containing T-cell oncogene on human chromosome 10q24. Proc. Natl. Acad. Sci. USA 88:8900-8904.
22. Kongsuwan, K., E. Webb, P. Housiaux and J.M. Adams. 1988. Expression of multiple homeobox genes within diverse hematopoietic lineages. EMBO J. 7:2131-2138.
23. Kurzrock, R., J.U. Gutterman and M. Talpaz. 1988. The molecular genetics of Philadelphia chromosome-positive leukemias. N Engl J Med 319:990-8.
24. Lammie, G.A., V. Fantl, R. Smith, E. Schuuring, S. Brookes, R. Michalides, C. Dickson, A. Arnold and G. Peters. 1991. D11S287, a putative oncogene on chromosome 11q13, is amplified and expressed in squamous cell and mammary carcinomas and linked to BCL-1. Oncogene 6:439-44.
25. Langdon, W.Y., A.W. Harris, S. Cory and J.M. Adams. 1986. The c-myc oncogene perturbs B lymphocyte development in E-mu-myc transgenic mice. Cell 47:11-8.
26. Lewin, B. 1990. Commitment and activation at pol II promoters: a tail of protein-protein interactions. Cell 61:1161-4.
27. Magli, M.C., P. Barba, A. Celetti, G. de Vita, C. Cillo and E. Boncinelli. 1991. Coordinate regulation of HOX genes in human hematopoietic cells. Proc. Natl. Acad. Sci. USA 88:6348-6352.
28. Magrath, I. 1990. The pathogenesis of Burkitt's lymphoma. Adv Cancer Res 55:133-271.
29. McDonnell, T.J., N. Deane, F.M. Platt, G. Nunez, U. Jaeger, J.P. McKearn and S.J. Korsmeyer. 1989. bcl-2-immunoglobulin transgenic mice demonstrate extended B cell survival and follicular lymphoproliferation. Cell 57:79-88.
30. McDonnell, T.J. and S.J. Korsmeyer. 1991. Progression from lymphoid hyperplasia to high-grade malignent lymphoma in mice transgenic for the t(14;18). Nature 349:254-256.
31. Mellentin, J.D., S.D. Smith and M.L. Cleary. 1989. LYL-1, a novel gene altered by chromosomal translocation in T-cell leukemia, codes for a protein with a helix-loop-helix DNA binding motif. Cell 58:77-83.
32. Rabbitts, T.H. 1991. Translocations, master genes, and differences between the origins of

acute and chronic leukemias. Cell 67:641-644.
33. Soekarman, D., M. von Lindern, D. van der Plas, L. Selleri, C.R. Bartram, P. Martiat, D. Culligan, R.A. Padua, K.P. Hasper-Voogt, A. Hagemeijer and G. Grosveld. DEK-CAN rearrangement is restricted to the translocation (6;9)(p23-q34). Leukemia, in press.
34. Strasser, A., A.W. Harris, M.L. Bath and S. Cory. 1990. Novel primitive lymphoid tumours induced in transgenic mice by cooperation between myc and bcl-2. Nature 348:331-3.
35. van Lohuizen, M., S. Verbeek, P. Krimpenfort, J. Domen, C. Saris, T. Radaszkiewicz and A. Berns. 1989. Predisposition to lymphomagenesis in pim-1 transgenic mice: cooperation with c-myc and N-myc in murine leukemia virus-induced tumors. Cell 56:673-82.
36. van Lohuizen, M., S. Verbeek, B. Scheijen, E. Wientjens, H. van der Gulden and A. Berns. 1991. Identification of cooperating oncogenes in E mu-myc transgenic mice by provirus tagging. Cell 65:737-52.
37. Withers, D.A., R.C. Harvey, J.B. Faust, O. Melnyk, K. Carey and T.C. Meeker. 1991. Characterization of a candidate bcl-1 gene. Mol. Cell. Biol. 11:4846-4853.

SUMMARY

A number of defined leukemia subtypes is associated with specific chromosomal translocations, which suggest that these translocations may have a causal role in leukemogenesis. Molecular characterization of some of these translocations led to the isolation of some known and some novel genes activated by translocation into T-cell receptor (TCR) or Immunoglobulin (Ig) loci, or by formation of a fusion gene encoding a chimaeric protein (chapter 1).

This thesis describes the molecular characterization of t(6;9), associated with a specific subtype of acute nonlymphocytic leukemia (chapter 2). T(6;9) ANLL patients are relatively young and respond poorly to therapy (chapter 2.1). The translocation breakpoints occur in chromosome 9q34 and 6p23, regions that harbour the c-*abl* and *pim*-1 genes respectively. However, long range mapping analysis showed that these genes are not involved in t(6;9) (chapter 2.2). A jumping library was used to isolate a genomic DNA probe 300 kb downstream c-*abl*, that could detect the translocation breakpoint on chromosome 9q34. Subsequent chromosome walking and cDNA cloning led to the isolation of a novel gene named *Cain (can)* (chapter 2.3). Analysis of 17 t(6;9) ANLL cases showed that the translocation breakpoints consistently occur in a single intron of 7.5 kb, designated *icb*-9, in this large (>140 kb) gene (chapter 2.4). Using t(6;9) bone marrow cells as a source for genomic and cDNA cloning, the gene on chromosome 6p23 could be isolated, that is involved in this translocation. This gene was named *dek*. In the *dek* gene translocation breakpoints cluster in a single intron as well. This intron was designated *icb*-6 and measures 9 kb (chapters 2.4, 2.5). A *dek-can* fusion gene is generated by the translocation, that encodes a chimaeric DEK-CAN protein as DEK and CAN open reading frames are fused in frame in the chimaeric transcript (chapter 2.5). The invariable *dek-can* fusion transcript can be used as a marker of t(6;9) ANLL, since it can be sensitively monitored by the polymerase chain reaction (chapter 2.4).

In a single case of acute undifferentiated leukemia (AUL) a translocation breakpoint was detected in *icb*-9 of *can*, but not in *dek*. Genomic cloning of the breakpoint region and cDNA cloning of the observed abnormal *can* transcript, led to the isolation of a third novel gene named *set* (chapter 2.6). The finding that the same C-terminal part of CAN can be linked to two different N-terminal protein moieties, DEK or SET, suggests that *can* may be the oncogene capable of generating leukemia. DEK and SET show no homology except for a stretch of acidic amino acids. The expression of both *dek, set* and *can* is ubiquitous. Since the respective proteins show no homology with any known protein sequences, their function is not known. However, the DEK-CAN and SET-CAN fusion products may have a role in transcription regulation, regarding the presence of putative

dimerization and ancillary DNA binding domains in CAN and the presence of an acidic domain, that may function in transcription activation, in both DEK and SET (chapter 2.5, 2.6).

Translocations in lymphoid cells, involving TCR or Ig loci, are most likely mediated by recombinases that normally function in TCR or Ig rearrangements. In contrast, the mechanisms by which fusion genes are generated in myeloid cells is far from being understood. Surprisingly, analysis of the genomic sequence of t(6;9) breakpoints revealed the presence of an almost complete LINE-1 element in *icb*-6 of *dek* and a truncated LINE-1 element in *icb*-9 of *can* (chapter 2.7). Since LINE-1 elements occur with a moderate frequency in the human genome, the presence of such an element in both breakpoint regions may be more than coincidental and may play a role in the translocation process.

SAMENVATTING

Een aantal gedefinieerde leukemie subtypen is geassocieerd met specifieke chromosomale translokaties, hetgeen doet vermoeden dat deze translokaties een rol spelen bij het ontstaan van de leukemieën. Molekulaire karakterisatie van enkele van deze translokaties resulteerde in de isolatie van zowel reeds bekende als nog onbekende genen. Deze genen kunnen geaktiveerd worden door translokatie in T-cel receptor (TCR) of immunoglobuline (Ig) loci, of door het vormen van fusie genen die koderen voor chimaere eiwitten (hoofdstuk 1).

Dit proefschrift beschrijft de molekulaire karakterisatie van de t(6;9), welke geassocieerd is met een specifiek subtype acute nietlymfatische leukemie (hoofdstuk 2). T(6;9) ANLL patienten zijn relatief jong en reageren slecht op therapie (hoofdstuk 2.1). De translokatie breukpunten treden op in chromosoom 9q34 en 6p23; regio's die respektievelijk het c-*abl* en *pim*-1 gen bevatten. Lange afstands DNA kartering liet echter zien dat deze genen niet betrokken zijn bij de t(6;9) (hoofdstuk 2.2). Het gebruik van een 'jumping library' leidde tot de isolatie van een genomisch DNA fragment, 300 kb 3' van c-*abl*, waarmee het translokatie breukpunt op chromosoom 9q34 kon worden waargenomen. Door het kloneren van zowel genomisch DNA als cDNA kon vervolgens een nieuw gen geïsoleerd worden dat *Cain, (can)* genoemd is (hoofdstuk 2.3). De analyze van 17 t(6;9) ANLL gevallen liet zien dat de translokatie breukpunten onveranderlijk plaats vinden in één enkel intron van 7,5 kb, *icb*-9 genoemd, in dit grote (>140 kb) gen (hoofdstuk 2.4). Door t(6;9) beenmergcellen te gebruiken voor het kloneren van genomisch en cDNA, kon een tweede gen worden geïsoleerd dat *dek* is genoemd en op chromosoom 6p23 ligt. In het *dek* gen liggen de translokatie breukpunten ook bijeen in één enkel intron, dat *icb*-6 is genoemd en 9 kb groot is (hoofdstuk 2.4 en 2.5). Door de translokatie wordt een *dek-can* fusie gen gevormd, dat voor een chimaer DEK-CAN eiwit kodeert aangezien de open leesramen van DEK en CAN zodanig gefuseerd zijn in het hybride transcript dat de leesramen intakt blijven (hoofdstuk 2.5). Omdat minimale hoeveelheden van het onveranderlijke *dek-can* fusie transcript kunnen worden gedetecteerd met de polymerase ketting reactie, kan het als specifiek kenmerk voor t(6;9) gebruikt worden (hoofdstuk 2.4).

In een geval van acute ongedifferentieerde leukemie, werd een breukpunt waargenomen in *icb*-9 van het *can* gen, maar niet in het *dek* gen. Door genomisch DNA rond het breukpunt te kloneren en cDNA te isoleren dat het afwijkend *can* transcript van deze patient vertegenwoordigt, kon een derde, tot nu toe onbekend gen worden geïdentificeerd (hoofdstuk 2.6). Het feit dat hetzelfde C-terminale deel van CAN gekoppeld kan worden aan twee verschillende N-terminale eiwit sequenties, doet vermoeden dat *can* wellicht het oncogen is dat leukemie

kan veroorzaken. DEK en SET zijn onderling niet homoloog met uitzondering van een rij van zure aminozuren die in beide eiwitten aanwezig zijn. *Dek, set* en *can* komen in alle weefsels tot expressie. Aangezien DEK, SET en CAN geen homologie vertonen met bekende eiwit sequenties, is de functie van deze eiwitten onbekend. Het is echter niet uitgesloten dat de DEK-CAN en SET-CAN fusie produkten een funktie hebben in transcriptie regulatie gezien de dimerisatie en mogelijke DNA-binding domeinen in CAN en het zure domein in zowel DEK als SET, dat dienst zou kunnen doen als transcriptie aktivatie domein (hoofdstuk 2.5, 2.6).

Translokaties in lymfatische cellen, waarbij TCR of Ig loci betrokken zijn, worden waarschijnlijk tot stand gebracht door recombinases, die normaal een funktie hebben in de rearrangering van TCR en Ig genen. Daarentegen is er nog weinig bekend van het mechanisme dat ervoor zorgt dat fusiegenen worden gevormd in myeloïde cellen. Het vaststellen van de nucleotidezuur volgorde rond de genomische t(6;9) breukpunten liet als verrassend resultaat zien dat een vrijwel compleet LINE-1 element aanwezig is in *icb*-6 van *dek*, en een verkort LINE-1 element in *icb*-9 van *can* (hoofdstuk 2.7). Omdat LINE-1 elementen slechts met een matige frequentie in het humane genoom voorkomen, lijkt de aanwezigheid ervan in beide breukpunt regio's meer dan toevallig en zouden de elementen een rol kunnen spelen in het translokatie proces.

CURRICULUM VITAE

Maartje Maria von Lindern

21 juni 1959	geboren te Krimpen aan de Lek
juni 1977	Eindexamen VWO Scholengemeenschap 'De Krimpenerwaard'
	Examen hoger beroepsonderwijs (HBO-B)
juni 1980	-klinisch chemisch analist
september 1980	-klinisch bacteriologisch analist
september 1980	Aanvang studie Biologie, Rijksuniversiteit Leiden
september 1983	Kandidaatsexamen Biologie, tweede hoofdvak Scheikunde (B4)
februari 1986	Doctoraalexamen Biologie, Hoofdvak: -Immunologie, Dr. A. Zantema, Prof. Dr. A.J. van der Eb, Medische Biochemie, RUL Bijvakken: -Microbiologie, Dr. C.A. Wijffelman, Moleculaire Plantkunde, RUL -Moleculaire Genetica, Dr. C.G. Grosveld, Prof. Dr. D. Bootsma, Celbiologie en Genetica, EUR
maart 1986	Aanvang promotieonderzoek op de afdeling Celbiologie en Genetica, Erasmus Universiteit Rotterdam, bij Dr. G.C. Grosveld en Prof. Dr. D. Bootsma
maart 1992	Aanstelling als tijdelijk wetenschappelijk medewerker, aan het 'Institute of Molecular Pathology', Wenen, Oostenrijk, bij Dr. H. Beug

PUBLICATIONS

Domen, J., M. Von Lindern, A. Hermans, M. Breuer, G. Grosveld and A. Berns. 1987. Comparison of the human and mouse pim-1 cDNA's: nucleotide sequence and immunological identification of the in vitro synthesized pim-1 protein. Oncogene Research 1:103-112.

Hermans, H., N. Heisterkamp, M. von Lindern, S. van Baal, D. Meijer, D. van der Plas, L.M. Wiedeman, J. Groffen, D. Bootsma and G. Grosveld. 1987. Unique fusion of *BCR* and c-*ABL* genes in Philadelphia chromosome positive acute lymphoblastic leukemia. Cell 51:33-40.

Meijer, D., A. Hermans, M. von Lindern, T. van Agthoven, A. de Klein, P. Mackenbach, A. Grootegoed, D. Talarico, G. Della Valle and G. Grosveld. 1987. Molecular characterization of the testis specific c-abl mRNA in mouse. Embo J. 6:4041-8.

Hermans, A., J. Gow, L. Selleri, M. von Lindern, A. Hagemeijer, L.M. Wiedemann and G. Grosveld. 1988. *bcr-abl* oncogene activation in Philadelphia chromosome-positive acute lymphoblastic leukemia. Leukemia 2:628-33.

von Lindern, M., T. van Agthoven, A. Hagemeijer, H. Adriaansen and G. Grosveld. 1989. The human *pim*-1 gene is not directly activated by the translocation (6;9) in acute nonlymphocytic leukemia. Oncogene 4:75-9.

von Lindern, M., A. Poustka, H. Lerach and G. Grosveld. 1990. The (6;9) chromosome translocation, associated with a specific subtype of acute nonlymphocytic leukemia, leads to aberrant transcription of a target gene on 9q34. Mol. Cell. Biol. 10:4016-26.

Selleri, L., M. von Lindern, A. Hermans, D. Meijer, G. Torelli and G. Grosveld. 1990. Chronic myeloid leukemia may be associated with several bcr-abl transcripts including the acute lymphoid leukemia-type 7 kb transcript. Blood 75:1146-53.

Soekarman, D., M. von Lindern, S. Daenen, B. de Jong, C. Fonatsch, B. Heinze, C.R. Bartram, A. Hagemeijer and G. Grosveld. The translocation (6;9)(p23;q34) shows consistent rearrangement of two genes and defines a myeloproliferative disorder with specific clinical features. Blood, in press.

Soekarman, D., M. von Lindern, D. van der Plas, L. Selleri, C. Bartram, P. Martiat, D. Culligan, R.A. Padua, K.P. Hasper-Voogt, A. Hagemeijer and G. Grosveld. DEK-CAN rearrangement is restricted to the translocation (6;9)(p23-q34). Leukemia, in press.

von Lindern, M., M. Fornerod, S. van Baal, M. Jaegle, T. de Wit, A. Buijs and G. Grosveld. The t(6;9), associated with a specific subtype of AML, results in the fusion of two genes, *dek* and *can*, and the expression of a chimaeric, leukemia specific dek-can mRNA. Mol. Cell. Biol. 12: in press.

von Lindern, M., S. van Baal, J. Wiegant, T. Raap, A. Hagemeijer and G. Grosveld. *Can*, a putative oncogene associated with myeloid leukemogenesis, may be activated by fusion of its 3' half to different genes: characterization of the *set* gene. Submitted for publication

von Lindern, M., D. van Beemst, S. van Baal, H. Adriaansen and G. Grosveld. Characterization of the translocation breakpoint sequences of two *dek-can* fusion genes present in t(6;9) AML, and a *set-can* fusion gene present in a case of AUL. Submitted for publication

von Lindern, M., M. Fornerod, N. Soekarman, S. van Baal, M. Jaegle, A. Hagemeijer, D. Bootsma and G. Grosveld. Translocation (6;9) in acute nonlymphocytic leukemia results in the formation of a *dek-can* fusion gene. *In*: The molecular genetics of hematological neoplasms. B. Young (ed.) Baillière Tindall, London. in press.

Shen, Y., M. von lindern, M. den Bakker and G. Grosveld. Regional mapping of 18 *Not*I linking clones on human chromosome 22. Submitted for publication.

NAWOORD

Als je er plezier in hebt om wetenschappelijk onderzoek te doen, is dat meer dan "een baan om je brood te verdienen". Gelukkig zijn de mensen met wie ik de afgelopen 6 jaar in Rotterdam samen gewerkt heb dan ook meer dan "collega's" geweest. Natuurlijk waren er zowel hoogtepunten als tegenslagen, wetenschappelijk en persoonlijk; die konden we dan met elkaar vieren of verwerken. Professor Bootsma wil ik bedanken voor de kans die hij mij gaf om bij hem te promoveren. Ondanks mijn 'flap-uit' opmerkingen, heb ik uw vertrouwen gehad. Toen ik kwam solliciteren naar een stageplaats, antwoordde u op een wellicht wat naieve opmerking van mij:'...als je niet hard wil werken meisje, zou ik zeggen gezellig dat je een kopje koffie bent komen drinken, maar je kunt weer gaan..'. Beste Dick, ik hoop dat ik je niet teleurgesteld heb, bedankt voor je steun.

Van Gerard Grosveld heb ik, zo niet alles, dan toch heel veel geleerd. Bovendien was je een goede hulpkracht toen Sjozèf me nog niet bijstond. Gerard je hebt gelijk dat het gevaarlijk is om met een techniek bezig te zijn om de techniek. Toch denk ik met plezier terug aan alle keren dat we nieuwe klonerings-strategiën bedachten voor t(6;9), t(11;22) of zomaar. Alles heeft zijn tijd, nu kan er meer biologie gedaan worden en mijn vertrek naar Wenen is hopelijk niet het eind van onze samenwerking. Bedankt voor alles.

Dies, je was niet alleen een goede vriend, ook een handig moleculair biologisch naslagwerk en opbouwend kritikus. Ik zal mijn best doen om ook zonder jouw aansporingen behoorlijke restrictiemaps te maken zonder al te sterk wisselende schaal. Ik zal je missen, maar weet je te vinden!

André, jou missen we al heel lang en helaas voor altijd. Je was een goede vriend en ik had graag nog eens met je gediscussieerd; over de klinische kant van ons moleculaire werk en over nog veel meer....

Sjozèf, je was er altijd, wist bijna overal raad op en was nergens te beroerd voor. Het enige dat ik je wellicht zou kunnen verwijten is dat ik ~14 kb sequence gepubliceerd heb en niet weet hoe ik een sequencegel moet runnen. Dank je dat je ook de laatste klus nog wilt opknappen: mijn paranymph zijn.

Arjan, Michael en Dimitri, als student hebben jullie ieder je bijdrage geleverd. Ik heb jullie misschien wel eens teveel laten vrijzwemmen en niet alles loopt altijd zoals je zou willen, maar, zoals je merkte, ook dat is 'part of the game'.

Nike, jij was er op het goede moment en wilde al het patienten screeningswerk doen met een nauwgezetheid die ik nooit op had kunnen brengen. Dank je daarvoor en voor het feit dat ik een deel van je werk als hoofdstuk in mijn boekje kon plaatsen.

Maarten, ik ben blij dat je twee jaar terug besloot bij ons te komen

werken. Onze eendrachtige samenwerking wordt wellicht gesymboliseerd door die ene foto die je in Boston maakte, al heb ik die nog nooit gezien. Judith, ik houd je aan je belofte dat je thee en taart voor me klaar zet in Rotterdam, dan zorg ik ervoor in Wenen. Martine, al kost het veel moeite, het zal lukken met die muizen. Maarten, Martine, Judith en Ronald, zet'm op de komende tijd en houdt me op de hoogte. Wordt 'v/h de CML groep' de 'transcriptie regulatie unit'?

Licia, you came for a few weeks, you stayed more than a year. We had a very good time together and I'll always remember you. Shen Yan, you taught us that italian pasta, dutch 'speculaaskruiden' and more, all came from China as you did. I'll take your tea-container with me to Vienna, like you took it from China to Holland; to remind you of home.

Zes jaar is een lange tijd en 'v/h de CML groep' heeft veel medewerkers gehad. Annelies, je hebt me overgehaald deze baan te nemen en daar heb ik geen spijt van gehad. Ton van Agthoven, de hybride cellijnen zijn uiterst waardevol geweest. Lies, Christine, Anneke, Antoinette, Robert, Ingrid, Ton de Wit, Monique, New-Ton, Marella, bedankt voor de gezelligheid, tot ziens.

De taken waren goed verdeeld binnen de afdeling en ik zou dus heel veel mensen kunnen bedanken voor reparaties, bestellingen, oplossingen en zo meer. Laat ik maar zeggen, ik heb het naar mijn zin gehad, bedankt allemaal.

Professor van der Eb, Professor Löwenberg en professor Hagemeijer wil ik bedanken voor het kritisch lezen van mijn proefschrift. Beste Anne, je hebt je niet alleen grondig door alle papers heen gegraven, je stond ook altijd klaar om materiaal voor ons te organizeren of ons wijzer te maken op punten die niet tot de standaard wetenschappelijke bagage van een moleculair bioloog behoren. Ook de mensen in jouw groep, vooral Bep, Marianne en Ellen, hadden altijd tijd voor ons. Dank jullie wel.

Tenslotte het thuisfront in Krimpen. Jullie hebben me in de afgelopen jaren steeds weer heel veel werk uit handen genomen. Jullie bijdrage aan dit proefschrift is dan ook niet gering. Nu ik naar Wenen ga om daar een nieuwe dependance op te zetten, zal het begrip 'thuis' waarschijnlijk nog verwarrender worden, maar dat geeft niet.

Tja, Freek, ik zeg maar zo...,ik zeg maar niets. We zien wel welke bergen we aan kunnen, letterlijk en figuurlijk.