MOLECULAR AND CELLULAR BIOLOGY, Jan. 2003, p. 186–194 Vol. 23, No. 1
0270-7306/03/$08.00+0 DOI: 10.1128/MCB.23.1.186–194.2003

Proteolytic Cleavage of MLL Generates a Complex of N- and C-Terminal Fragments That Confers Protein Stability and Subnuclear Localization

James J.-D. Hsieh,[1] Patricia Ernst,[1] Hediye Erdjument-Bromage,[2] Paul Tempst,[2] and Stanley J. Korsmeyer[1]*

Howard Hughes Medical Institute, Departments of Pathology and Medicine, Harvard Medical School, Dana-Farber Cancer Institute, Boston, Massachusetts 02115,[1] and Molecular Biology Program, Memorial Sloan Kettering Cancer Center, New York, New York 10021[2]

Received 2 August 2002/Returned for modification 12 September 2002/Accepted 1 October 2002

The mixed-lineage leukemia gene (*MLL*, *ALL1*, *HRX*) encodes a 3,969-amino-acid nuclear protein homologous to *Drosophila* trithorax and is required to maintain proper *Hox* gene expression. Chromosome translocations in human leukemia disrupt *MLL* (11q23), generating chimeric proteins between the N terminus of MLL and multiple translocation partners. Here we report that MLL is normally cleaved at two conserved sites (D/GADD and D/GVDD) and that mutation of these sites abolishes the proteolysis. MLL cleavage generates N-terminal p320 (N320) and C-terminal p180 (C180) fragments, which form a stable complex that localizes to a subnuclear compartment. The FYRN domain of N320 directly interacts with the FYRC and SET domains of C180. Disrupting the interaction between N320 and C180 leads to a marked decrease in the level of N320 and a redistribution of C180 to a diffuse nuclear pattern. These data suggest a model in which a dynamic post-cleavage association confers stability to N320 and correct nuclear sublocalization of the complex, to control the availability of N320 for target genes. This predicts that MLL fusion proteins of leukemia which would lose the ability to complex with C180 have their stability conferred instead by the fusion partners, thus providing one mechanism for altered target gene expression.

homeotic genes of the invertebrates and *Hox* genes of the vertebrates encode a family of related transcription-regulatory proteins that play critical roles in implementing developmental patterns and conferring positional information (18). Development of a normal body plan requires continued expression of each *homeotic*/*Hox* gene throughout development, and the combination of their expression patterns determines the segmental identity of cells throughout the body. Deregulation of these expression patterns causes transformation of segmental identities during development and can contribute to a variety of human disorders, including malignancy. *Hox* gene expression is initiated by transiently expressed segmentation gene products. However, the maintenance of appropriate *Hox* expression is conferred by two classes of regulators, the positive regulatory *trithorax* group (*trx-G*) and the negative regulatory *polycomb* group (*Pc-G*), which apparently function through epigenetic mechanisms (21, 26).

Mll is a mammalian member of the *trx-G*, and loss-of-function studies indicate that it is required for midgestational development. An *Mll-exon3*^{lacZ} homozygous disruption results in embryonic death at E10.5 with multiple patterning defects including abnormal ganglia, absence of the maxillary branch of the first branchial arch, and pooling of erythroid precursors in the coelomic cavity (35). This *Mll* knockout revealed that *Mll* is required for the maintenance rather than the initiation of *Hox* gene expression in early embryogenesis (34). Although *Mll*-null embryos demonstrated correct initiation of *Hox* gene expression, including *Hoxa7* and *Hoxc8*, the subsequent maintenance of the expression of these *Hox* genes was lost. Together, these studies indicate that *Mll* is critical for the proper maintenance of *Hox* gene expression during development.

Translocations of human chromosome segment 11q23 disrupt the *MLL* gene, resulting in a common chromosome abnormality encountered in acute leukemia, which is found in 80% of infant leukemias and the majority of chemotherapy-related leukemias following treatment with topoisomerase II inhibitors (12, 30). The *MLL* gene is composed of 37 exons and encodes a 3,969-amino-acid (aa) nuclear protein with multiple distinct domains including an N terminus with three AT-hook motifs, a DNA methyltransferase homology domain (DNMT), four PHD fingers, a transactivation domain, and a highly conserved SET domain at the C terminus (11, 15, 31) (see Fig. 1A). *MLL* is fused with more than 20 different chromosomal partners on translocation, resulting in multiple unique MLL fusion proteins. The *MLL* breakpoints are tightly clustered in an 8.3-kb genomic region containing exons 8 to 14, and the derivative 11 chromosome [der(11)] encodes the MLL fusion protein, which possesses ~1,300 aa of N-terminal MLL (12). The C-terminal fusion partners of MLL are very diverse, ranging from putative transcription factors to cytoplasmic structural proteins. This has raised the question whether there is any common property conferred by the fusion proteins (2).

In this study, we demonstrate that native MLL is proteolytically processed into N-terminal p320 (N320) and C-terminal p180 (C180) fragments. We identify two protease recognition motifs that are highly conserved in MLL homologues from *Drosophila* to mammals. Processed MLL fragments form a

* Corresponding author. Mailing address: Dana-Farber Cancer Institute, One Jimmy Fund Way, Boston, MA 02115. Phone: (617) 632-6402. Fax: (617) 632-6401. E-mail: stanley_korsmeyer@dfci.harvard.edu.

stable complex based on interactions between the FYRN domain of N320 and the FYRC and SET domains of C180. Mutational analysis indicates that interaction between C180 and N320 is required for steady-state levels of N320 and correct subnuclear localization. These observations indicate that leukemogenic MLL fusion proteins would lose association with C180. Furthermore, they suggest a model in which the varied MLL fusion partners all play a common role of stabilizing the fusion product.

MATERIALS AND METHODS

Plasmid construction. Flag-tagged full-length human *MLL* cDNA inserted in the pCXN2 vector has been described (16). Deletion between two *Spe*I sites generates MLLΔ2253–2727. Oligonucleotides that encode the Flag or Myc epitope were inserted into the unique *Bst*EII site of *MLL* to generate C-terminally Flag- or Myc-tagged MLL, respectively. Flag-tagged truncated MLL constructs were generated by inserting a stop codon into a *Sal*I or *Pfl*MI site. A PCR fragment encoding aa 2720 to 3969 was inserted into the pCMV-Myc vector (Clontech) to generate Myc-tagged C180. A stop codon was added at the *Bst*EII site to generate Myc-C150ΔF/S. Green fluorescent protein (GFP) was inserted in frame at the *Bst*EII site of *MLL*. A PCR fragment encoding red fluorescent protein (RFP) was cloned into pCDNA3, and the resultant vector was used to generate the RFP-MLL-GFP construct. Truncation mutants with a deletion of FYRC and SET domains (ΔF/S) were constructed by adding a stop codon after either the Flag epitope or GFP. A *Not*I site was created in wild-type (wt) MLL to generate DG→AA and DGV→AAA MLL mutants.

MLL Ab preparation. MLL fragments consisting of aa 5 to 955 and aa 2829 to 2883 were expressed as glutathione *S*-transferase GST fusions in *Escherichia coli* BL21/DE3. GST-MLL-fusion proteins were purified as previously described (13) and injected into rabbits (Covance). Abs were purified from immunized rabbit serum by using a HiTrap protein A column (Amersham Pharmacia Biotech) as specified by the manufacturer.

Transient-transfection, immunoblotting, in vitro transcription and translation, and immunoprecipitation assays. Human embryonic kidney 293T cells were transiently transfected with the indicated MLL-expressing constructs by using Fugene 6 (Roche Molecular Biochemicals) as specified by the manufacturer. Cells were solubilized in RIPA buffer 48 h after transfection. Immunoprecipitation and immunoblot assays were performed as described previously (6) using indicated anti-Flag (Sigma), anti-N-terminal (anti-NT), or anti-CT Ab.[35]S-labeled MLL polypeptides were generated by in vitro transcription and translation using the TNT-coupled reticulocyte lysate system (Promega). Immunoprecipitation assays of[35]S-labeled polypeptides using anti-Flag antibody were described previously (6).

Indirect immunofluorescence and confocal microscopy. Fluorescence-tagged MLL was expressed in 293T cells. Cells were fixed in phosphate-buffered saline with 3% paraformaldehyde, stained with YOPRO-3 (Molecular Probes), and detected by confocal microscopy.

Purification of Flag-C180 and N-terminal sequence analysis. An MLL construct carrying a Flag tag at the C terminus was transfected into 293T cells. Cellular extracts purified from 30 10-cm dishes were incubated with anti-Flag Ab-coated beads. Immunoprecipitates were separated by sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) and transferred to a polyvinylidene difluoride (PVDF) membrane. Coomassie blue staining demonstrated one predominant band, which was excised and subjected to N-terminal Edman degradation analysis.

RESULTS

MLL is proteolytically cleaved into N-terminal p320 and C-terminal p180 fragments. To further explore the function of MLL and MLL fusion proteins, we generated several Abs against distinct regions of MLL, all of which failed to detect a protein appropriate for full-length MLL (predicted molecular mass, 430 kDa; actually migrates at ~500 kDa). However, an Ab to an N-terminal region (aa 5 to 955; anti-NT) recognized instead an ~320-kDa fragment in cells expressing a full-length MLL construct (Fig. 1C). Reciprocally, an Ab to a C-terminal region (aa 2829 to 2883; anti-CT) also failed to detect full-length MLL but did identify an ~180-kDa fragment (Fig. 1D). This suggested that MLL may undergo a posttranslational cleavage. To test this hypothesis, we tagged the N-terminal or C-terminal portion of MLL with a Flag epitope (Fig. 1B). An anti-Flag Ab detected the same 320- and 180-kDa fragments (Fig. 1E), as did the anti-NT (Fig. 1C) and anti-CT (Fig. 1D) Abs. This confirmed that full-length MLL is cleaved into N-terminal p320 (N320) and C-terminal p180 (C180) fragments. A series of internal deletion mutants were subsequently generated to localize the cleavage site. The Flag-tagged mutant (MLLΔ2253–2727) was not cleaved and was detected as a ~450-kDa protein by all three Abs (Fig. 1B to E). In total, these data indicate that MLL is cleaved and that the cleavage site(s) resides between aa 2253 and 2727.

N320 and C180 fragments form a post-cleavage complex which displays a distinct subnuclear colocalization. Numerous studies have demonstrated a punctate subnuclear localization for MLL utilizing Abs raised against the N-terminal portion of the protein (4, 5, 24, 32). Since MLL is proteolytically processed to two distinct fragments, one critical issue is whether C180 demonstrates the same or a unique subcellular localization compared to N320. To address this, RFP was fused to the N terminus of MLL and GFP was introduced into the C terminus. When this doubly tagged protein was expressed, it was cleaved correctly into RFP-N320 and GFP-C180 fragments (Fig. 2D). Expression of this dually fluorescent MLL protein in 293T cells resulted in the colocalization of RFP-N320 and GFP-C180 in a punctate nuclear pattern, as determined by confocal microscopy (Fig. 2A). This observation raised the possibility of a physical association between N320 and C180 fragments prompting the creation of a dually tagged construct bearing a Flag-tagged N terminus and a Myc-tagged C-terminus (Fig. 2B). Myc-C180 was coimmunoprecipitated with Flag-N320 by using the anti-Flag Ab (Fig. 2B). A reciprocal immunoprecipitation using an anti-Myc Ab coprecipitated the Flag-N320 as well as C180 fragments (data not shown). To examine whether endogenous MLL protein was proteolytically cleaved and assembled, cellular extracts from HeLa and 293T cells were immunoprecipitated with anti-CT Ab and subjected to Western blot analysis using either anti-CT or anti-NT Abs. The endogenous MLL protein of both cells was also processed into N320 and C180 fragments with evidence through their coimmunoprecipitation of a post-cleavage complex (Fig. 2C).

Identification of two consensus cleavage sites in MLL. To precisely determine the cleavage site, the Flag-tagged C180 MLL fragment was purified from immunoprecipitates separated by SDS-PAGE and transferred to a PVDF membrane. A Coomassie blue staining step identified one predominant 180-kDa peptide (Fig. 3A). N-terminal Edman degradation analysis of the excised polypeptide yielded a 12-aa sequence corresponding to aa 2719 to 2730 of MLL (Fig. 3A). Sequence alignment of MLL from different species indicated that this cleavage site is highly conserved and resides within the internally deleted region of the *Drosophila trx*^{E3} mutant (20). However, mutation of the cleavage site from D/GVDD to A/AADD (Fig. 3B) reduced but did not completely abolish the N320 band; however, it did reveal an increase in the amount of apparently full-length p500 MLL. Of note, this mutant generated a C-terminal fragment somewhat larger than wild-type

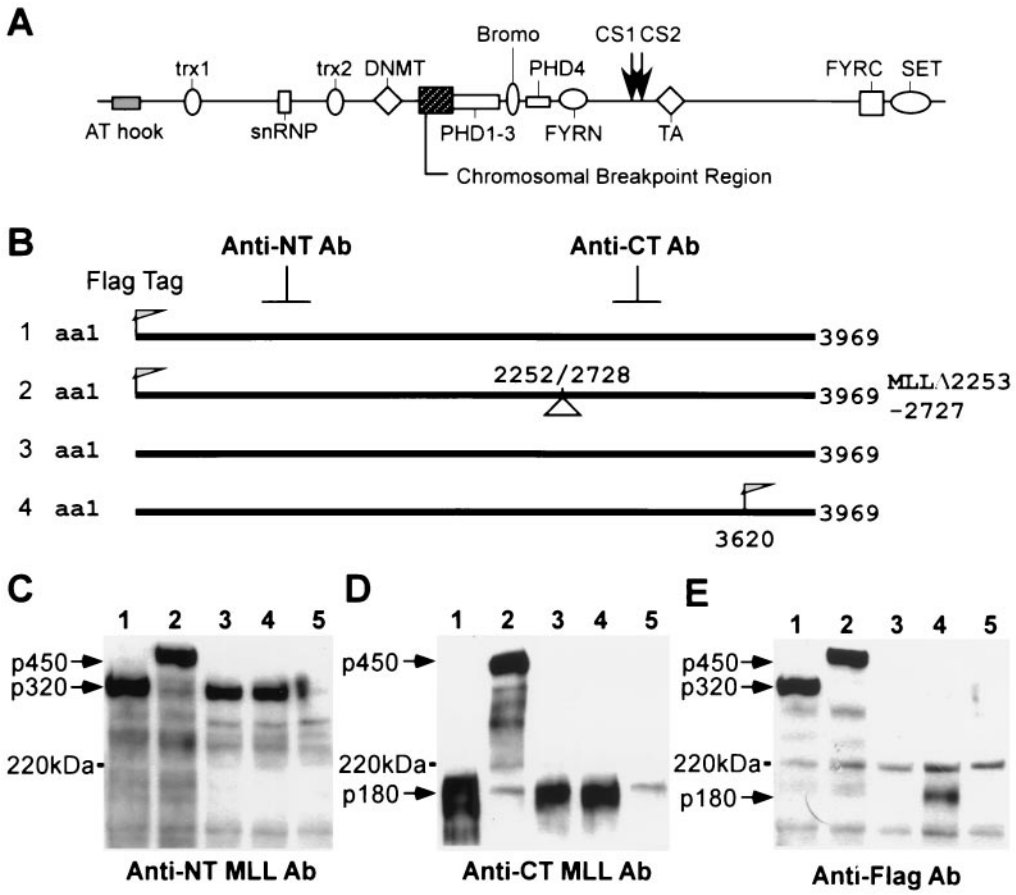

FIG. 1. MLL is proteolytically cleaved into an N-terminal 320-kDa fragment, N320, and a C-terminal 180-kDa fragment, C180. (A) Conserved domain structure of MLL in relation to cleavage sites (CS1 and CS2) and the chromosomal breakpoint region. (B) MLL expression constructs used in the experiments below are numbered 1 to 4, corresponding to the lanes where they are expressed. (C) Western blot analysis using an Ab directed to an N-terminal region of MLL (anti-NT MLL) detects p320 in cells expressing Flag-tagged MLL (lane 1), untagged MLL (lane 3), and Myc-tagged MLL (lane 4); lane 2 contains the internal deletion mutant MLLΔ2253-2727 at ~450 kDa; lane 5 contains whole-cell extracts purified from untransfected 293T cells as a control. (D) Anti-CT Ab detected a 180-kDa polypeptide in cells expressing Flag-tagged MLL (lane 1), untagged MLL (lane 3), and Myc-tagged MLL (lane 4), but the MLL deletion mutant was ~450 kDa (lane 2). (E) Anti-Flag Ab Western blot analysis detected the p320 fragment (lane 1), the ~450-kDa MLL deletion mutant (lane 2), and the C-terminal p180 fragment (lane 4).

C180 (Fig. 3B). This compilation of findings suggested the existence of an additional cleavage site upstream of the residue 2718-2719 site. A search for a similar cleavage motif identified another potential proteolytic site, D/GADD, located 53 residues proximal to the initially identified D/GVDD site. An MLL construct possessing doubly mutated motifs (Fig. 3B) was no longer processed, as evidenced by the disappearance of both N320 and C180 fragments as well as an abundance of full-length p500 (Fig. 3B). Thus, MLL appears to be proteolytically processed at two independent sites, which we have called CS1 (D/GADD, aa 2666-2667) and CS2 (D/GVDD, aa 2718-2719).

The FYRN domain of N320 directly interacts with the FYRC and SET domains of C180. The interaction sites between N320 and C180 were further refined by coexpressing truncations of the N terminus along with C180 (Fig. 4A). The region between aa 1394 and 2160 within the N320 fragment proved necessary for its interaction with C180 since residues 1 to 2160 but not residues 1 to 1393 would coimmunoprecipitate Myc-C180 (Fig. 4E). Several conserved domains lie within this 767-aa region, including PHD fingers 1 to 3, a Bromo domain, PHD finger 4, and a FYRN domain (Fig. 1A). Three conserved regions have been recognized within the C180 fragment of MLL. They include the transactivation domain that mediates CBP interaction, the SET domain at the C-terminal end, and the FYRC domain immediately N-terminal to the SET domain (Fig. 1A). We reasoned that interaction between N320 and C180 is likely to be mediated through conserved domains. To test this hypothesis, we performed a series of coimmunoprecipitation assays with the indicated MLL domains (Fig. 5). These analyses identified a direct interaction between the FYRN domain of N320 MLL and the region containing both the FYRC and SET domains of C180 MLL (Fig. 5G).

Interaction with C180 controls the stability of N320. To investigate the functional significance of the N320-C180 interaction mediated by FYRC plus SET, we deleted C-terminal residues 3621 to 3969 of MLL, which contain both domains (MLLΔF/S, Fig. 6). This C-terminally truncated mutant, MLLΔF/S, was proteolytically processed to the expected N320 fragment and the predictably smaller C150ΔF/S fragment (Fig.

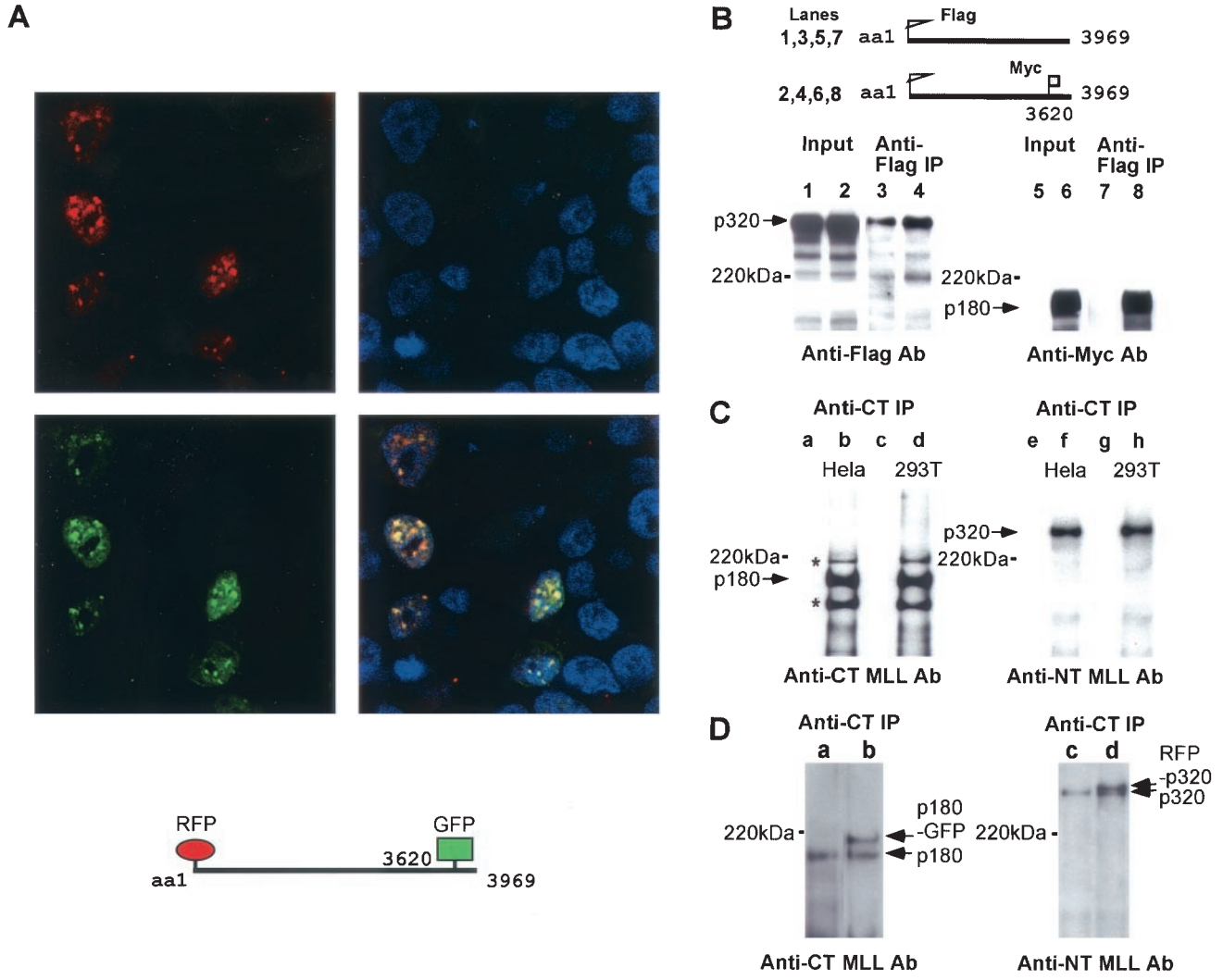

FIG. 2. Colocalization and association between the processed p320 and p180 fragments of MLL. (A) RFP-p320 and GFP-p180 colocalize to the punctate subnuclear compartment. Cells expressing dually fluorescent RFP-MLL-GFP were assessed by confocal microscopy for localization of the expected RFP-p320 (red, upper left panel) and GFP-p180 (green, lower left panel) fragments. The nuclei were stained with YO-PRO (blue, upper right panel). Colocalization of RFP and GFP resulted in yellow punctate subcellular fluorescence in the merged images (yellow, lower right panel). (B) Coimmunoprecipitation between Flag-p320 and Myc-p180. The N-terminally Flag-tagged MLL (Flag-MLL) or the N-terminally Flag-plus C-terminally Myc-tagged MLL (Flag-MLL-Myc) constructs were expressed in cells, solubilized, size fractionated in the corresponding lanes, and then subjected to immunoprecipitation (IP) using anti-Flag Ab. Western blots of input lysate or anti-Flag immunoprecipitates were developed with anti-Flag Ab or anti-Myc Ab. Flag-MLL-Myc was processed to Flag-p320 and Myc-p180 (lanes 2 and 6). Myc-p180 is efficiently coprecipitated with Flag-p320 (lane 8). (C) Endogenous MLL is cleaved to p320 and p180, which form a stable complex in cells. Cellular extracts from HeLa or 293T cells were incubated with (lane b, d, f, and h) or without (lane a, c, e, and g) anti-CT Ab and then precipitated with protein-A beads. Immunoprecipitates were subjected to SDS-PAGE and analyzed by Western blotting using either anti-NT or anti-CT Ab. Anti-CT Ab precipitated a 180-kDa protein from both cell lines tested (lanes b and d). Anti-NT Ab detected a 320-kDa protein that was coprecipitated with endogenous p180 (lanes f and h). The asterisks denote cross-reactive bands recognized by anti-CT Ab. (D) Dually fluorescent RFP-MLL-GFP is proteolytically processed as wt MLL. Cellular extracts from RFP-MLL-GFP-transfected 293T cells were immunoprecipitated with anti-CT Ab (lanes b and d). Anti-CT Ab Western blot analysis recognized C-180-GFP in addition to the endogenous C180 (lane b). RFP-N320 was detected as a slower-migrating band above the endogenous N320 (lane d). Lanes a and c contain anti-CT immunoprecipitates from untransfected 293T cells.

6). The C150ΔF/S fragment did not coprecipitate N320 (Fig. 6). This confirms the in vitro interaction between the FYRC-plus-SET region of C180 and N320. Of note, although the C150ΔF/S fragment was of comparable abundance to wt C180, we observed a 10-fold reduction in N320 levels in cells with mutant C150ΔF/S (Fig. 6). This suggests that the interaction between N320 and C180 protects N320 from degradation. To further demonstrate the effect of the N320-C180 interaction on N320 stability, we coexpressed a Flag-tagged N-terminal MLL with either Myc-tagged C180 or C150ΔF/S in 293T cells (Fig. 7A). In the presence of C180, the N-terminal MLL was stabilized and detected at high abundance (Fig. 7B, lane 1) compared to the markedly diminished level detected when expressed with C150ΔF/S (lane 2). To estimate the half-life of the N-terminal MLL fragment, cells expressing these constructs had their protein synthesis inhibited by cycloheximide. Levels

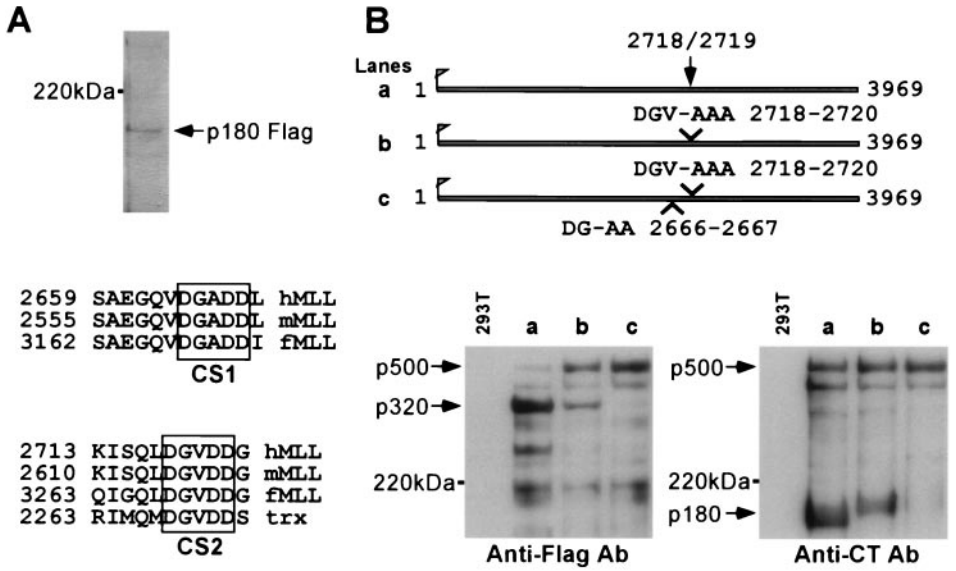

FIG. 3. Identification and verification of two MLL cleavage sites. (A) C-terminally Flag-tagged MLL was expressed in 293T cells, and the processed Flag-C180 was purified using anti-Flag Ab. Captured protein was analyzed by SDS-PAGE, transferred to PVDF membranes, visualized with Coomassie blue (arrow), excised, and subjected to Edman degradation. This yielded a 12-aa sequence corresponding to residues 2719 to 2730 of MLL. Sequence alignment among hMLL (human), mMLL (murine), fMLL (*Fugu rubripes*), and *Drosophila trx* indicates conservation of D/GVDD cleavage site 2 (CS2). A second cleavage site, D/GADD (CS1), proximal to CS2 was subsequently identified but exists only in the vertebrate MLLs, not *Drosophila trx*. (B) MLL mutants possessing the indicated amino acid substitutions are shown in the upper panel and used in the corresponding lanes below. Expression of the indicated wt or mutant MLL proteins in 293T cells was assessed by either anti-Flag Ab (left panel) or anti-CT Ab (right panel) Western blotting. Alanine substitution of CS2 (2718DGV-AAA) failed to completely abolish the proteolytic cleavage of MLL (lane b, both panels). Mutation of both cleavage sites (2666DG-AA plus 2718DG-AAA) is required for complete abolition of MLL processing, as evidenced by the disappearance of both N320 and C180 and an increase in the amount of full-length p500 (lane c, both panels).

of N-terminal MLL were determined over a serial time course by anti-Flag Western blot analyses. In the presence of C180, the estimated half-life of N-terminal MLL was longer than 8 h (Fig. 7C). In contrast, when C150ΔF/S, which fails to interact, was expressed, the estimated half-life of N-terminal MLL was shortened to ~4 h (Fig. 7D). This confirms that the processed MLL N-terminal fragment must interact with the C-terminal fragment to avoid degradation and confer stability. Thus, one rationale for a dynamic interaction between cleaved MLL fragments is to regulate the levels of the N-terminal fragment of MLL.

Intact FYRC and SET domains are required for the punctate subnuclear localization of C180. To assess the effects of interaction on subcellular protein distribution, the MLLΔF/S construct lacking the FYRC and SET domains was labeled with RFP at its N terminus and GFP at its C terminus (Fig. 8). Confocal microscopy of cells expressing this doubly fluorescent construct revealed that there was no detectable expression of the RFP-N320 fragment (Fig. 8, left upper panel). Thus, because of the low level of RFP-N320, we could not accurately assess whether N320, which possesses previously defined subnuclear localization sequences (32), was itself capable of correct localization in the absence of interaction with C180. In contrast, GFP-C150ΔF/S was easily detected in the nucleus but had lost its normal punctate subnuclear pattern and was now distributed diffusely throughout the nucleus (Fig. 8). This argues that the localization of C180 to the subnuclear compartment is governed by its association with N320, although we cannot exclude the possibility that localization is mediated directly through the FYRC and SET domains. Thus, the processing of MLL regulates both the abundance and the localization of MLL fragments, providing mechanisms to control target gene expression.

DISCUSSION

Specific proteolytic cleavage plays a critical role in many biological pathways, including the liberation of SREBP from the endoplasmic reticulum, the release of Notch from the plasma membrane, the maturation of hedgehog signaling molecules, and the activation of caspases and their subsequent cleavage of death substrates during apoptosis. Here, we demonstrate that mammalian MLL is proteolytically processed at cleavage sites which follow P1 aspartic acid residues located within two highly conserved D/GVDD and D/GADD motifs. The precise protease which mediates this cleavage remains to be determined but is likely to be the same for the similar CS1 and CS2 motifs. Of note, we have tested inhibitors of calpain, the 26S proteasome, and caspases and demonstrated no substantial effect on the cleavage. We also noted a D/GVDD motif deleted in a *Drosophila* TRX that is located within a 281-aa region deleted in a described mutant *trx* allele, trx^{E3} (32). Unlike null mutations of *trx* that fail to maintain the expression of *homeotic* genes of the *bithorax (BX-C)* and *antennapedia (ANT-C)* complexes (3), trx^{E3} displays selectively reduced expression of *ANT-C* but not *BX-C* genes (28). Subsequent studies using antibodies recognizing different regions of trx are consistent with the notion that this recognized motif serves as

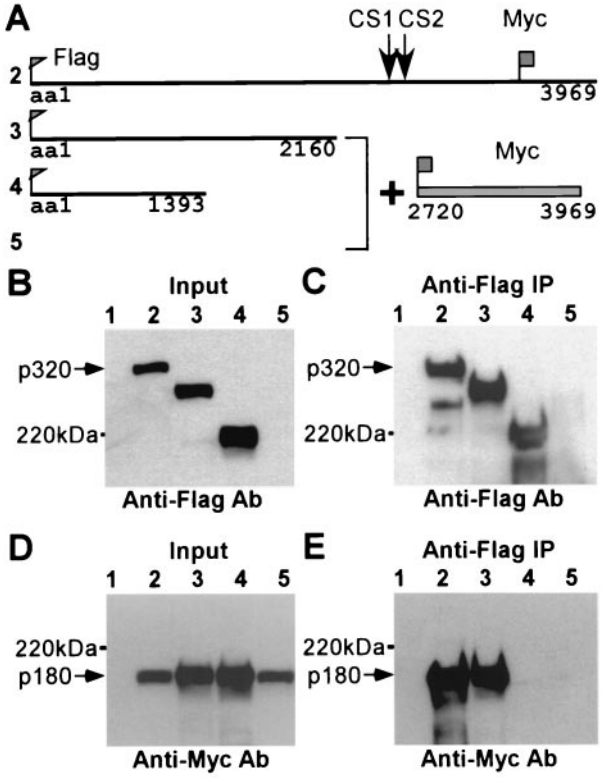

FIG. 4. aa 1394 to 2160 of N320 mediate the interaction with C180. (A) Flag-tagged wt or truncated N-terminal MLL constructs were expressed in 293T cells either alone or in combination with Myc-tagged C180. (B to E) Lane 1 of each experiment contained extracts prepared from untransfected cells. Doubly tagged MLL (Flag-MLL-Myc) was in lane 2 throughout. (E) Anti-Flag immunoprecipitation (IP) followed by anti-Myc Western blot analysis indicated that MLL aa 1 to 2160 but not 1 to 1393 efficiently interact with Myc-p180 (lanes 3 and 4).

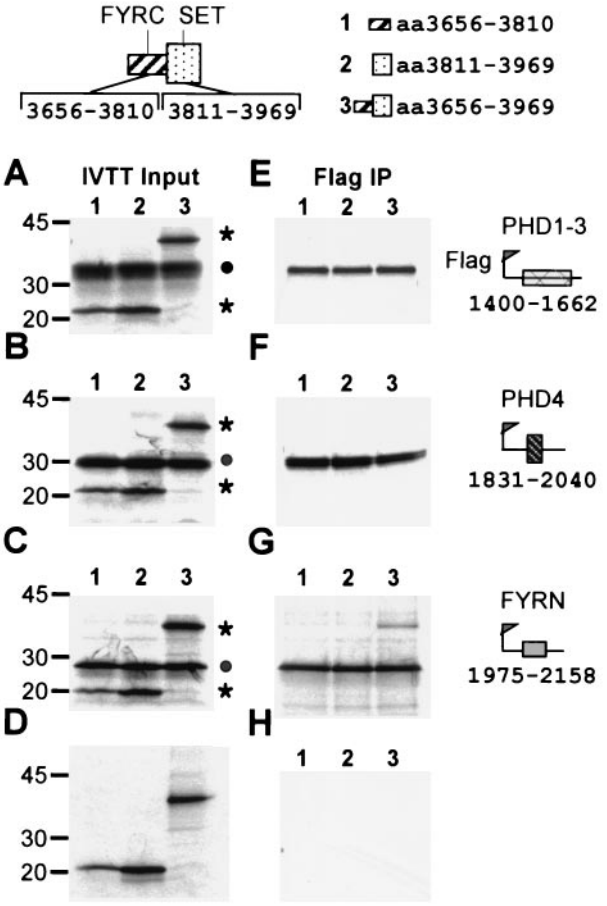

FIG. 5. Direct physical interaction between the FYRN domain of N320 and the FYRC plus SET domains of C180. The top schematic depicts the conserved C-terminal domains of MLL that were in vitro translated in the presence of [^{35}S]methionine and used in lanes 1 to 3, respectively. The side schematics depict Flag-tagged N-terminal domains that were ^{35}S labeled and used in panels A and E (PHD1 to PHD3), B and F (PHD4), and C and G (FYRN). The left four panels (A to D) show the input samples, while the right four panels (E to H) show the anti-Flag immunoprecipitation assays. Solid circles denote the positions of Flag-tagged N-terminal domains, while asterisks denote the positions of C-terminal domains. (G) The FYRN domain of N320 and the FYRC plus SET domains of C180 are the minimal interaction domains.

a proteolytic site (19). The *homeotic* defects in trx^{E3} argue that cleavage of trx has a developmental consequence.

The precise mechanisms whereby MLL exerts its principal function to maintain the expression of target genes, including *Hox* genes, remain uncertain. However, the presence of conserved motifs with MLL provides some insight. MLL possesses domains with ascribed biochemical properties: the three AT hook motifs provide minor-groove DNA binding (31, 36), the DNMT homology domain binds methylated DNA differentially and can repress transcription (8, 11, 22, 36), a conserved subnuclear localization signal directs punctate nuclear distribution (32), the PHD fingers mediate protein-protein interaction (14), the transactivation domain interacts with CREB-binding protein (13), and the highly conserved SET domain binds to histone H3, Sbf1, and ASH1, as well as itself (2, 9, 17, 25). Here we identify the D/GVDD and D/GADD motifs and assign a role for them in MLL processing as well as identifying roles for the FYRN and FYRC domains. The FYRN and the FYRC plus SET domains mediate the interaction between the N320 and C180 subunits of MLL. Of note, *Drosophila* TRX also possesses FYRN and FYRC motifs, suggesting that it may also form a complex of processed fragments. Two distant MLL family members, ALR and MLL3, are present in the human genome (23, 29) but lack the D/GVDD cleavage motifs. Interestingly, they have a varied organization of domains compared to MLL, possessing a PHD domain, a FYRN domain, a FYRC domain, and a SET domain all clustered within the C-terminal 450 aa of ALR and MLL3. This highly conserved architecture is also present in *Caenorhabditis elegans* and *Drosophila* ALR homologues (27). The interaction between the FYRN and the FYRC plus SET domains noted here for MLL raises a testable possibility that ALR and MLL3 may have clustered these domains to create a similar functional moiety. Perhaps one role for the cleavage of MLL is to physically juxtapose these previously separated domains.

Chromosomal translocations in human leukemia which disrupt MLL produce a chimeric protein in which the N-terminal portion of MLL is fused to 1 of more than 20 partners. These fusion partners do not have any recognizable motif in common,

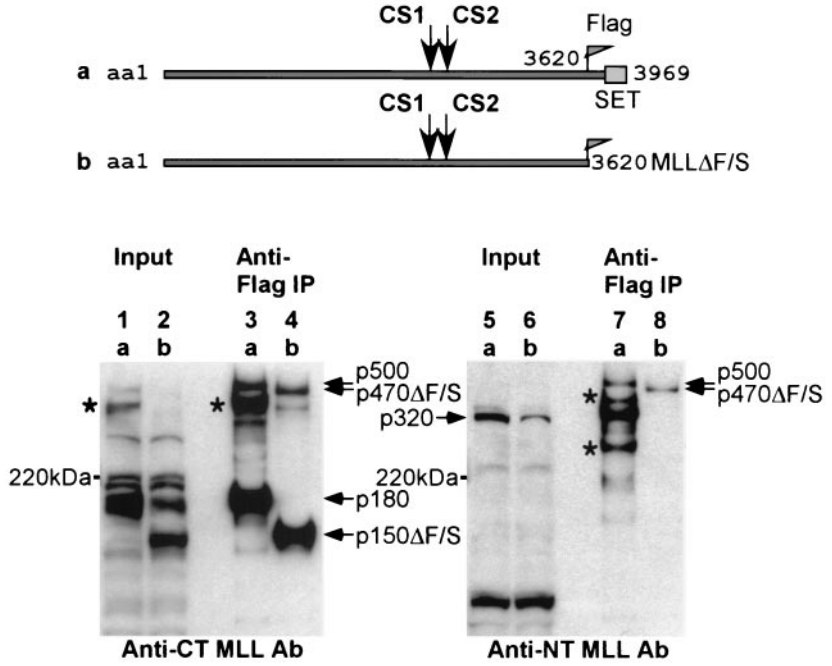

FIG. 6. The interaction between N320 and C180 of MLL confers stability to the N320. Flag-tagged wt MLL and a C-terminal deletion lacking the FYRC and SET domains (MLLΔF/S) were expressed in 293 T cells used in lanes a and b, respectively. Input cellular extracts and anti-Flag immunoprecipitate (IP) were analyzed by Western blotting with the indicated Abs. MLL-FlagΔF/S was processed to N320 and C150ΔF/S (lanes 2 and 6). In the coimmunoprecipitation assays, C150ΔF/S failed to precipitate N320 (lane 8). The protein level of C150ΔF/S was comparable to that of C180 (lanes 1 to 4). In contrast, the level of N320 was markedly reduced when it could no longer associate with the C-terminal fragment (lanes 5 and 6). Small amounts of the newly synthesized, uncleaved MLL (~500 kDa) and a few smaller migrating proteins, perhaps representing splice variants (asterisk), are noted.

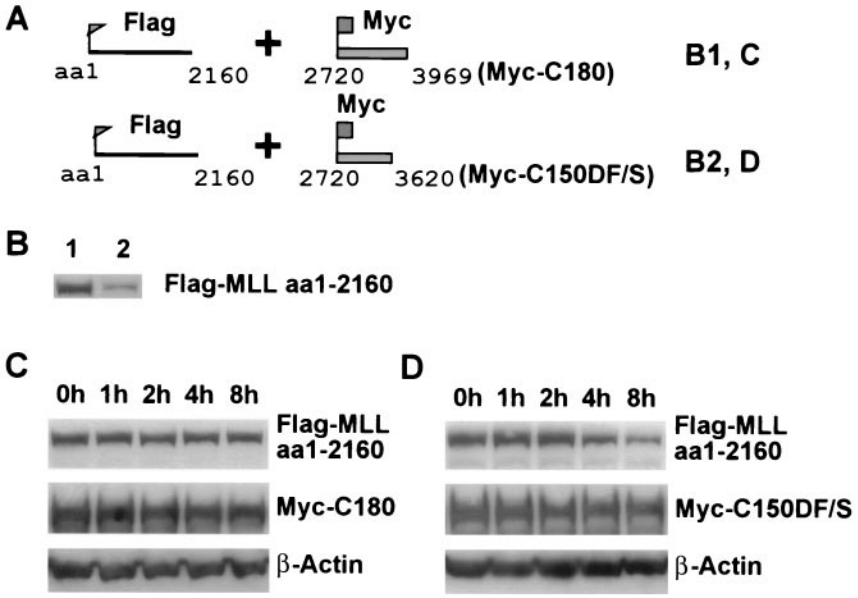

FIG. 7. The FYRC-plus-SET region of C180 protects N-terminal MLL from degradation. (A) Flag-tagged MLL aa 1 to 2140 was cotransfected with Myc-tagged C180 or C150ΔF/S in 293T cells. Cellular extracts coexpressing the MLL N terminus and C180 were used in panel B, lane 1, and panel C. Samples from extracts expressing the MLL N terminus and C150ΔF/S were analyzed in panel B, lane 2, and panel D. (B) Myc-tagged wt C terminus, C180 (lane 1) but not the truncated C150 (lane 2), increases the abundance of coexpressed N-terminal MLL, determined by anti-Flag Western blot analysis. (C and D) The degradation of N-terminal MLL in the presence of C180 (C, top panel) or C150ΔF/S (D, top panel) was compared. Cellular extracts obtained at the indicated time points after addition of 10 μg of cycloheximide per ml at 36 h following transfection were analyzed using anti-Flag Ab. The stable expression of C180 (C, middle panel) and C150ΔF/S (D, middle panel) was detected by anti-Myc Ab. Equal protein loading was determined by using anti-β actin Ab in both lower panels. The gel in panel D has been exposed five times longer than that in panel C to equilibrate the steady-state levels of Flag-MLL aa 1 to 2160.

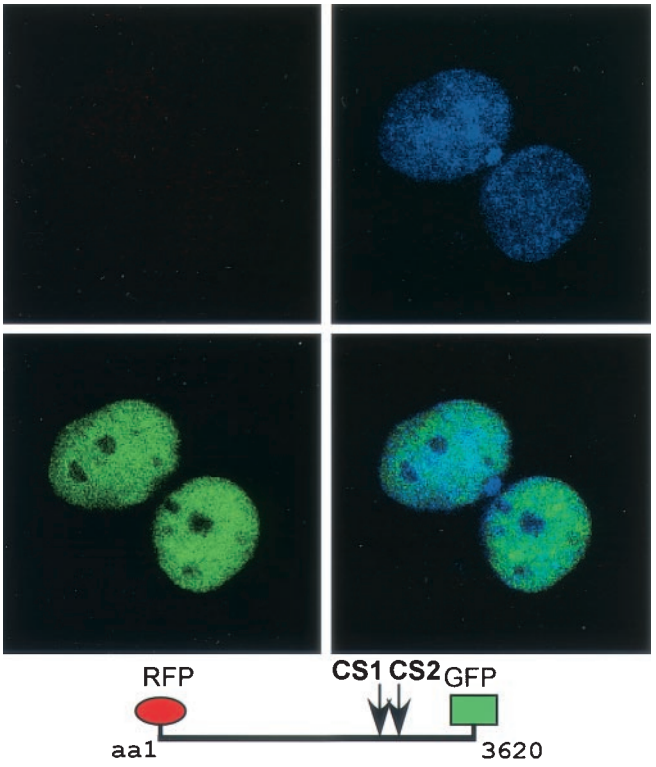

FIG. 8. Intact FYRC and SET domains are required for the punctate subnuclear localization of C180. Loss of the C180 interaction results in loss of RFP-N320 fluorescence. A dually fluorescent MLL mutant, with the FYRC and SET domains deleted (RFP-MLL-GFPΔF/S), was expressed, and the resultant RFP-N320 and GFP-C150ΔF/S fragments were assessed by confocal microscopy. GFP-C150ΔF/S was easily detected but had redistributed from the punctate nuclear pattern seen with wt C180 to a diffuse nuclear pattern (lower left panel). The same cells that expressed GFP-C180ΔF/S failed to exhibit detectable RFP-N320, as judged by the consistent lack of red fluorescence (upper left panel). The upper right panel shows the YO-PRO nuclear stain, while the lower right panel shows the merged image of RFP-N320 and GFP-C150ΔF/S.

and they range widely from nuclear factors to cytoplasmic proteins and cell junction proteins. Gene expression profiles of human leukemias bearing an *MLL* translocation identified a pattern of upregulated genes. Among these genes were some of the well-recognized targets of wt MLL, including *HOXA4*, *HOXA5*, and *HOXA9* (1, 33). This argues that the N-terminal 1,300-aa portion of MLL is sufficient to confer at least some target gene specificity to MLL fusion proteins. Mice expressing an MLL-AF9 fusion product progressed to acute myeloid leukemia, establishing the oncogenic potential of MLL chimeric proteins (7). In contrast, *Mll-exon8*^{Myc} mice in which a small myc epitope tag was fused with MLL failed to develop leukemia, arguing that a truncated MLL fragment was not in itself sufficient to cause transformation. However, *Mll-exon8*^{LacZ} mice in which MLL was fused with the β-galactosidase protein did progress to acute leukemia (10). β-Galactosidase has been extensively used in transgenic mouse models and plays no independent, demonstrated role in tumor formation. The compilation of these findings indicates that individual fusion partners play an indispensable role but that they need not be directly oncogenic. The observations here concerning protein instability may provide an explanation. The FYRN domain, required for interaction of the processed MLL fragments, is located downstream of the chromosomal breakpoint, indicating that MLL fusion proteins would lose interaction with the stabilizing C180 component. Consequently, one testable characteristic that might prove to be shared by all the partner proteins is the capacity to confer protein stability to the MLL N terminus.

ACKNOWLEDGMENTS

This work was supported in part by NIH award P01 CA68484. P.E. is a recipient of the Beckenstein Fellowship and a Special Fellow of the Leukemia and Lymphoma Society.

We thank E. Smith for editorial assistance and B. Malynn for helpful discussions and suggestions.

REFERENCES

1. **Armstrong, S. A., J. E. Staunton, L. B. Silverman, R. Pieters, M. L. den Boer, M. D. Minden, S. E. Sallan, E. S. Lander, T. R. Golub, and S. J. Korsmeyer.** 2002. MLL translocations specify a distinct gene expression profile that distinguishes a unique leukemia. Nat. Genet. **30:**41–47.
2. **Ayton, P. M., and M. L. Cleary.** 2001. Molecular mechanisms of leukemogenesis mediated by MLL fusion proteins. Oncogene **20:**5695–5707.
3. **Breen, T. R., and P. J. Harte.** 1993. Trithorax regulates multiple homeotic genes in the bithorax and Antennapedia complexes and exerts different tissue-specific, parasegment-specific and promoter-specific effects on each. Development **117:**119–134.
4. **Butler, L. H., R. Slany, X. Cui, M. L. Cleary, and D. Y. Mason.** 1997. The HRX proto-oncogene product is widely expressed in human tissues and localizes to nuclear structures. Blood **89:**3361–3370.
5. **Caslini, C., A. Shilatifard, L. Yang, and J. L. Hess.** 2000. The amino terminus of the mixed lineage leukemia protein (MLL) promotes cell cycle arrest and monocytic differentiation. Proc. Natl. Acad. Sci. USA **97:**2797–2802.
6. **Cheng, E. H., M. C. Wei, S. Weiler, R. A. Flavell, T. W. Mak, T. Lindsten, and S. J. Korsmeyer.** 2001. BCL-2, BCL-X(L) sequester BH3 domain-only molecules preventing BAX- and BAK-mediated mitochondrial apoptosis. Mol. Cell **8:**705–711.
7. **Corral, J., I. Lavenir, H. Impey, A. J. Warren, A. Forster, T. A. Larson, S. Bell, A. N. McKenzie, G. King, and T. H. Rabbitts.** 1996. An Mll-AF9 fusion gene made by homologous recombination causes acute leukemia in chimeric mice: a method to create fusion oncogenes. Cell **85:**853–861.
8. **Cross, S. H., R. R. Meehan, X. Nan, and A. Bird.** 1997. A component of the transcriptional repressor MeCP1 shares a motif with DNA methyltransferase and HRX proteins. Nat. Genet. **16:**256–259.
9. **Cui, X., I. De Vivo, R. Slany, A. Miyamoto, R. Firestein, and M. L. Cleary.** 1998. Association of SET domain and myotubularin-related proteins modulates growth control. Nat. Genet. **18:**331–337.
10. **Dobson, C. L., A. J. Warren, R. Pannell, A. Forster, and T. H. Rabbitts.** 2000. Tumorigenesis in mice with a fusion of the leukaemia oncogene *Mll* and the bacterial *lacZ* gene. EMBO J. **19:**843–851.
11. **Domer, P. H., S. S. Fakharzadeh, C. S. Chen, J. Jockel, L. Johansen, G. A. Silverman, J. H. Kersey, and S. J. Korsmeyer.** 1993. Acute mixed-lineage leukemia t(4;11)(q21;q23) generates an MLL-AF4 fusion product. Proc. Natl. Acad. Sci. USA **90:**7884–7888.
12. **Downing, J. R., and A. T. Look.** 1996. MLL fusion genes in the 11q23 acute leukemias. Cancer Treat. Res. **84:**73–92.
13. **Ernst, P., J. Wang, M. Huang, R. H. Goodman, and S. J. Korsmeyer.** 2001. MLL and CREB bind cooperatively to the nuclear coactivator CREB-binding protein. Mol. Cell. Biol. **21:**2249–2258.
14. **Fair, K., M. Anderson, E. Bulanova, H. Mi, M. Tropschug, and M. O. Diaz.** 2001. Protein interactions of the MLL PHD fingers modulate MLL target gene regulation in human cells. Mol. Cell. Biol. **21:**3589–3597.
15. **Gu, Y., T. Nakamura, H. Alder, R. Prasad, O. Canaani, G. Cimino, C. M. Croce, and E. Canaani.** 1992. The t(4;11) chromosome translocation of human acute leukemias fuses the ALL-1 gene, related to *Drosophila* trithorax, to the AF-4 gene. Cell **71:**701–708.
16. **Hanson, R. D., J. L. Hess, B. D. Yu, P. Ernst, M. van Lohuizen, A. Berns, N. M. van der Lugt, C. S. Shashikant, F. H. Ruddle, M. Seto, and S. J. Korsmeyer.** 1999. Mammalian Trithorax and polycomb-group homologues are antagonistic regulators of homeotic development. Proc. Natl. Acad. Sci. USA **96:**14372–14377.
17. **Katsani, K. R., J. J. Arredondo, A. J. Kal, and C. P. Verrijzer.** 2001. A homeotic mutation in the trithorax SET domain impedes histone binding. Genes Dev. **15:**2197–2202.
18. **Krumlauf, R.** 1994. Hox genes in vertebrate development. Cell **78:**191–201.

19. **Kuzin, B., S. Tillib, Y. Sedkov, L. Mizrokhi, and A. Mazo.** 1994. The *Drosophila* trithorax gene encodes a chromosomal protein and directly regulates the region-specific homeotic gene fork head. Genes Dev. **8:**2478–2490.
20. **Mazo, A. M., D. H. Huang, B. A. Mozer, and I. B. Dawid.** 1990. The trithorax gene, a trans-acting regulator of the bithorax complex in *Drosophila*, encodes a protein with zinc-binding domains. Proc. Natl. Acad. Sci. USA **87:**2112–2116.
21. **Pirrotta, V.** 1998. Polycombing the genome: PcG, trxG, and chromatin silencing. Cell **93:**333–336.
22. **Prasad, R., T. Yano, C. Sorio, T. Nakamura, R. Rallapalli, Y. Gu, D. Leshkowitz, C. M. Croce, and E. Canaani.** 1995. Domains with transcriptional regulatory activity within the ALL1 and AF4 proteins involved in acute leukemia. Proc. Natl. Acad. Sci. USA **92:**12160–12164.
23. **Prasad, R., A. B. Zhadanov, Y. Sedkov, F. Bullrich, T. Druck, R. Rallapalli, T. Yano, H. Alder, C. M. Croce, K. Huebner, A. Mazo, and E. Canaani.** 1997. Structure and expression pattern of human ALR, a novel gene with strong homology to ALL-1 involved in acute leukemia and to *Drosophila* trithorax. Oncogene **15:**549–560.
24. **Rogaia, D., F. Grignani, R. Carbone, D. Riganelli, F. LoCoco, T. Nakamura, C. M. Croce, P. P. Di Fiore, and P. G. Pelicci.** 1997. The localization of the HRX/ALL1 protein to specific nuclear subdomains is altered by fusion with its eps15 translocation partner. Cancer Res. **57:**799–802.
25. **Rozovskaia, T., O. Rozenblatt-Rosen, Y. Sedkov, D. Burakov, T. Yano, T. Nakamura, S. Petruck, L. Ben-Simchon, C. M. Croce, A. Mazo, and E. Canaani.** 2000. Self-association of the SET domains of human ALL-1 and of *Drosophila* TRITHORAX and ASH1 proteins. Oncogene **19:**351–357.
26. **Schumacher, A., and T. Magnuson.** 1997. Murine Polycomb- and trithorax-group genes regulate homeotic pathways and beyond. Trends Genet. **13:**167–170.
27. **Sedkov, Y., J. J. Benes, J. R. Berger, K. M. Riker, S. Tillib, R. S. Jones, and A. Mazo.** 1999. Molecular genetic analysis of the *Drosophila* trithorax-related gene which encodes a novel SET domain protein. Mech. Dev. **82:**171–179.
28. **Sedkov, Y., S. Tillib, L. Mizrokhi, and A. Mazo.** 1994. The bithorax complex is regulated by trithorax earlier during *Drosophila* embryogenesis than is the Antennapedia complex, correlating with a bithorax-like expression pattern of distinct early trithorax transcripts. Development **120:**1907–1917.
29. **Tan, Y. C., and V. T. Chow.** 2001. Novel human HALR (MLL3) gene encodes a protein homologous to ALR and to ALL-1 involved in leukemia, and maps to chromosome 7q36 associated with leukemia and developmental defects. Cancer Detect. Prev. **25:**454–469.
30. **Thirman, M. J., H. J. Gill, R. C. Burnett, D. Mbangkollo, N. R. McCabe, H. Kobayashi, S. Ziemin-van der Poel, Y. Kaneko, R. Morgan, A. A. Sandberg, et al.** 1993. Rearrangement of the MLL gene in acute lymphoblastic and acute myeloid leukemias with 11q23 chromosomal translocations. N. Engl. J. Med. **329:**909–914.
31. **Tkachuk, D. C., S. Kohler, and M. L. Cleary.** 1992. Involvement of a homolog of *Drosophila* trithorax by 11q23 chromosomal translocations in acute leukemias. Cell **71:**691–700.
32. **Yano, T., T. Nakamura, J. Blechman, C. Sorio, C. V. Dang, B. Geiger, and E. Canaani.** 1997. Nuclear punctate distribution of ALL-1 is conferred by distinct elements at the N terminus of the protein. Proc. Natl. Acad. Sci. USA **94:**7286–7291.
33. **Yeoh, E. J., M. E. Ross, S. A. Shurtleff, W. K. Williams, D. Patel, R. Mahfouz, F. G. Behm, S. C. Raimondi, M. V. Relling, A. Patel, C. Cheng, D. Campana, D. Wilkins, X. Zhou, J. Li, H. Liu, C. H. Pui, W. E. Evans, C. Naeve, L. Wong, and J. R. Downing.** 2002. Classification, subtype discovery, and prediction of outcome in pediatric acute lymphoblastic leukemia by gene expression profiling. Cancer Cell **1:**133–143.
34. **Yu, B. D., R. D. Hanson, J. L. Hess, S. E. Horning, and S. J. Korsmeyer.** 1998. MLL, a mammalian trithorax-group gene, functions as a transcriptional maintenance factor in morphogenesis. Proc. Natl. Acad. Sci. USA **95:**10632–10636.
35. **Yu, B. D., J. L. Hess, S. E. Horning, G. A. Brown, and S. J. Korsmeyer.** 1995. Altered Hox expression and segmental identity in Mll-mutant mice. Nature **378:**505–508.
36. **Zeleznik-Le, N. J., A. M. Harden, and J. D. Rowley.** 1994. 11q23 translocations split the "AT-hook" cruciform DNA-binding region and the transcriptional repression domain from the activation domain of the mixed-lineage leukemia (MLL) gene. Proc. Natl. Acad. Sci. USA **91:**10610–10614.