Cancer Cell
Article

Structural Basis for Recognition of SMRT/N-CoR by the MYND Domain and Its Contribution to AML1/ETO's Activity

Yizhou Liu,[1] Wei Chen,[3,5] Justin Gaudet,[3,5] Matthew D. Cheney,[3,5] Liya Roudaia,[3] Tomasz Cierpicki,[1] Rachel C. Klet,[1] Kari Hartman,[4] Thomas M. Laue,[4] Nancy A. Speck,[3,*] and John H. Bushweller[1,2,*]
[1] Department of Molecular Physiology and Biological Physics
[2] Department of Chemistry
University of Virginia, Charlottesville, VA 22908, USA
[3] Department of Biochemistry, Dartmouth Medical School, Hanover, NH 03755, USA
[4] Center to Advance Molecular Interaction Science, University of New Hampshire, Durham, NH 03824, USA
[5] These authors contributed equally.
*Correspondence: nancy.speck@dartmouth.edu (N.A.S.), jhb4v@virginia.edu (J.H.B.)
DOI 10.1016/j.ccr.2007.04.010

SUMMARY

AML1/ETO results from the t(8;21) associated with 12%–15% of acute myeloid leukemia. The AML1/ETO MYND domain mediates interactions with the corepressors SMRT and N-CoR and contributes to AML1/ETO's ability to repress proliferation and differentiation of primary bone marrow cells as well as to enhance their self renewal in vitro. We solved the solution structure of the MYND domain and show it to be structurally homologous to the PHD and RING finger families of proteins. We also determined the solution structure of an MYND-SMRT peptide complex. We demonstrated that a single amino acid substitution that disrupts the interaction between the MYND domain and the SMRT peptide attenuated AML1/ETO's effects on proliferation, differentiation, and gene expression.

INTRODUCTION

AML1/ETO is produced as a result of the t(8;21) found in acute myeloid leukemia (AML) of the M2 subtype (Miyoshi et al., 1993). It consists of the N terminus of AML1, including its DNA-binding Runt domain, and almost the entire ETO protein. AML1/ETO causes leukemia in mice in cooperation with other mutated genes such as TEL/PDGFRβ, ISCBP deficiency, and the FLT3 internal tandem duplication (Grisolano et al., 2003; Schessl et al., 2005; Schwieger et al., 2002). It interacts with many other proteins including CBFβ, SMRT/N-CoR, PLZF, GFI-1, SIN3, and numerous histone deacetylases (Hug and Lazar, 2004). Determining which of these interactions contributes to AML1/ETO's activity and which pathways are affected by these interactions is necessary in order to develop targeted therapy for this leukemia.

Many of AML1/ETO's interactions with other proteins are mediated by the five domains that are conserved with the *Drosophila* Runt, Lozenge, and Nervy proteins. The most N-terminal domain is the Runt domain from the AML1 fusion partner, which is an s type immunoglobulin fold that binds both DNA and CBFβ (Bravo et al., 2001; Tahirov et al., 2001). DNA binding by the Runt domain is thought to be essential for AML1/ETO's leukemogenic activity (Grisolano et al., 2003). C-terminal to the Runt domain is the Nervy homology region 1 (NHR1), otherwise known as the ETO TBP-associated factor homology domain (eTAFH) (Zhang et al., 2004). NHR2, or the hydrophobic heptad repeat (HHR) domain, is a helical tetramer that

SIGNIFICANCE

Chromosomal rearrangements often result in the production of chimeric proteins with altered function. In many cases, aberrant recruitment of corepressor complexes has been shown to contribute substantially to the oncogenic potential of these chimeric proteins. The MYND domain (NHR4) of ETO, responsible for recruitment of the corepressors SMRT and N-CoR, becomes fused to the AML1 protein as a result of the 8;21 translocation. We have solved the structure of an MYND domain/SMRT complex. Based on the structural information, we have used a point mutation to show that SMRT/N-CoR recruitment by the MYND domain plays a key role in the block in proliferation and differentiation seen with AML1/ETO and alters the expression of genes associated with these functions.

mediates oligomerization with AML1/ETO, ETO, or the ETO homologs MTG16 and MTGR1 (Kitabayashi et al., 1998; Liu et al., 2006; Minucci et al., 2000; Zhang et al., 2004). Oligomerization through the HHR domain appears to be essential for AML1/ETO's activity in repressing cell proliferation and differentiation and in promoting the clonogenic ability of primary hematopoietic progenitors (Hug et al., 2002; Liu et al., 2006; Shimada et al., 2000; Zhang et al., 2001). NHR3 interacts with the regulatory subunit of type II cyclic AMP-dependent protein kinase (PKA RIIα) (Fukuyama et al., 2001). Its in vivo contribution to AML1/ETO function is not well characterized.

NHR4, also known as the myeloid-Nervy-DEAF-1 (MYND) domain was predicted based on its primary sequence to have two putative, non-DNA-binding zinc fingers (Gross and McGinnis, 1996) and binds the silencing mediator of retinoid and thyroid hormone receptor (SMRT) and nuclear receptor corepressor (N-CoR) complexes (Gelmetti et al., 1998; Lutterbach et al., 1998a; Wang et al., 1998). The MYND domain is defined by a C_6HC zinc-chelating motif that is found in a number of other nuclear proteins such as Nervy, DEAF-1, BS69, PDCD2, and Bop. MYND domains have frequently been implicated in transcriptional repression (Ansieau and Leutz, 2002; Gottlieb et al., 2002; Ladendorff et al., 2001; Lausen et al., 2004; Scarr and Sharp, 2002; Sims et al., 2002). Recruitment of N-CoR/SMRT and their associated histone deacetylases (HDACs) to AML1 target genes is one mechanism by which AML1/ETO is thought to aberrantly repress transcription and contribute to leukemogenesis (Gelmetti et al., 1998; Lutterbach et al., 1998b).

Here we report the solution structure of the AML1/ETO MYND domain, showing it is structurally homologous to the PHD and RING finger families of proteins. In addition, we solved the structure of a complex between the AML1/ETO MYND domain and a peptide from SMRT. Based on that structure, we introduced single amino acid substitutions into the AML1/ETO MYND domain to assess the role of the SMRT/N-CoR interaction in vivo. We show that mutations that disrupt SMRT/N-CoR binding reverse AML1/ETO's repression of cell proliferation, mildly affect its ability to repress granulocyte differentiation, and attenuate alterations in gene expression caused by AML1/ETO in lineage-negative Sca-1^+ c-kit^+ (LSK) bone marrow cells.

RESULTS

Solution Structure of the MYND Domain from AML1/ETO

The ^{15}N-^{1}H HSQC spectrum of the AML1/ETO MYND domain (aa 658–707) clearly indicated a structured entity (Figure 1A), but upon addition of 5 mM EDTA, it collapsed to a poorly dispersed spectrum characteristic of an unfolded species (data not shown). Refolding was readily achieved by dialysis against 50 μM $ZnCl_2$, confirming that Zn^{2+} is essential for maintaining the MYND structure. The putative zinc-chelating residues include cysteines 663, 666, 674, 677, 683, 687, 699, and one of the three consecutive histidines (695, 696, and 697). To unequivocally identify which histidine is chelated to Zn, we ran a long range ^{15}N-^{1}H HSQC (Pelton et al., 1993) at pH 6.0 which clearly indicated that the Nε2 of H695 is resistant to protonation and thus is the zinc-chelating atom (see Figure S1A in the Supplemental Data available with this article online). Mutation of H695 to Ala resulted in an HSQC spectrum similar to that obtained in the presence of EDTA (Figure S2D and data not shown).

We solved a high-resolution structure of the AML1/ETO MYND domain using a combination of NOEs, chemical shifts, and extensive dipolar couplings (Table 1). The overall fold of the AML1/ETO MYND domain adopts an interleaved zinc-chelating topology as observed in the PHD and RING domains. Zinc center I consists of Cys 663, 666, 683, and 687, and zinc center II is formed by Cys 674, 677, 699, and His695 (Figure 1B). The relative positioning of secondary structure elements follows the ββα motif frequently observed in PHD and RING fingers (Pascual et al., 2000; Zheng et al., 2002). These structural features greatly differ from those of the LIM zinc-finger domain, which also contains two zinc atoms but adopts a sequential zinc-chelating topology and forms two juxtaposed domains with a certain degree of structural similarity (Velyvis et al., 2001).

Spadaccini et al. (Spadaccini et al., 2006) recently reported that the DEAF-1 MYND domain (backbone RMSD = 7.8 Å to our ETO MYND domain) has a sequential zinc-chelating topology like that of the LIM domains. This is a completely different fold than we describe here, which is surprising given the very high sequence similarity between MYND domains (Figure 1C). The ETO MYND domain structure has been refined and validated with residual dipolar couplings from two independent alignments. Fitting only the conserved residues in the DEAF-1 MYND domain structure to the 19 corresponding NH dipolar couplings obtained for the ETO MYND domain yields an average Q factor of 84%, i.e., the structure is highly inconsistent with the dipolar couplings measured for the ETO MYND domain. A more extensive discussion of the differences between the two structures and the experimental data that validates the ETO MYND domain structure is provided in the Supplemental Data, including Figure S3.

The AML1/ETO MYND Domain Recognizes a "PPPLI" Motif in SMRT and N-CoR

Residues 1031–1273 in repression domain III (RDIII) of SMRT were previously shown to contain the binding site for ETO (Zhang et al., 2001). We confirmed that an aa 1031–1273 SMRT polypeptide bound the MYND domain by HSQC titration (Figure 1A). N-CoR and SMRT are predicted to be mostly unstructured in their repression domains suggesting a peptide recognition motif. We therefore screened a peptide library spanning the SMRT interacting region using fluorescence anisotropy to identify the minimal MYND-binding consensus sequence in SMRT. We identified three consecutive peptides that displayed significantly elevated anisotropy in the presence of the MYND domain (Figure 1D) and share the common sequence motif NPPPLI, corresponding to SMRT residues

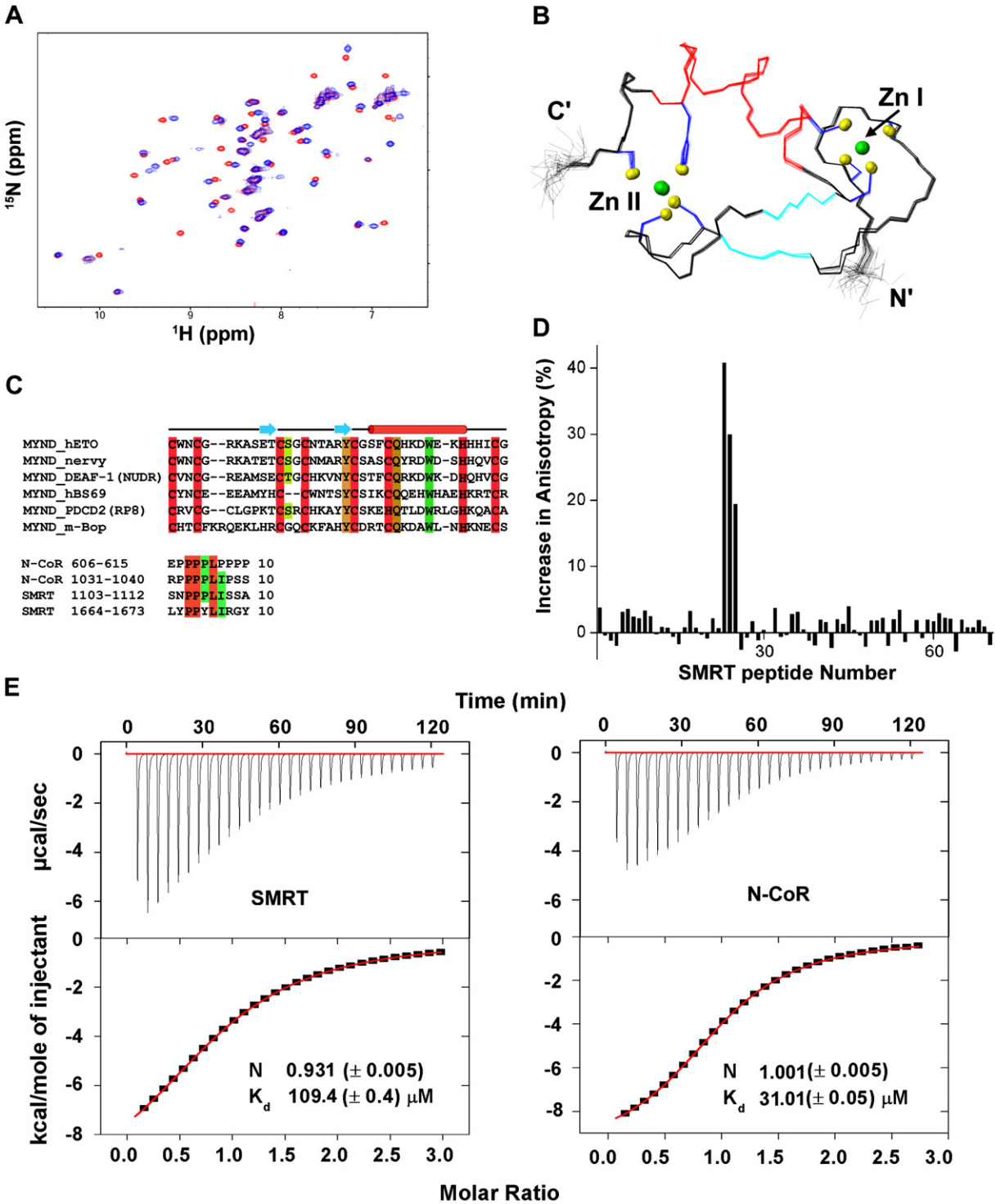

Figure 1. Structure of the ETO MYND Domain and Its Interactions with SMRT and N-CoR

(A) Overlay of ^{15}N-^{1}H HSQC spectra of the free MYND domain (red) and MYND + SMRT1031-1273 (blue) collected at pH 6.0 and 37°C at 500 MHz.

(B) Overlay of the backbone of 31 conformers representing the solution structure of the MYND domain with α helical residues (S685–H695) in red, β sheet residues (E672–T673 and R681–Y682) in cyan, zinc-chelating side-chains in blue, zinc-chelating atoms in yellow, and zinc atoms in green.

(C) Sequence alignment of MYND domains (Zn-chelating residues in red, additional conserved residues in green and brown) and interacting regions from SMRT and N-CoR (identical residues in red, conserved residues in green).

(D) Fluorescence anisotropy measurements on 70 fluorescein-labeled peptides spanning residues 1031–1273 of SMRT. The y axis corresponds to the difference between the anisotropy in the presence and absence of the MYND domain.

(E) Isothermal titration calorimetry (ITC) measurements of MYND domain binding to peptides derived from SMRT and N-CoR.

Table 1. Statistics for NMR Data Collection and Structure Calculations

	MYND	SMRT-MYND
Total NOEs observed	959	1922
Total nonredundant NOEs	505	964
Intraresidue (i = j)	169	309
Sequential (|i − j| = 1)	129	210
Medium-range (1 < |i − j| < 5)	79	126
Long-range (|i − j| > 4)	128	319
NOES between SMRT and MYND	NA	121
Total restraints for zinc chelation	25	25
Dihedral angle restraints		
Φ	15	17
Ψ	15	17
X1	17	16
Total RDCs for structural calculation		
HN	33	34
NC′	29	30
C′Cα	30	32
Total RDCs for structural validation (Q_{free})*		
HN	28	31
RMSD of distance restraints (maximum) (Å)	0.001 ± 0.000 (0.055)	0.0033 ± 0.0005 (0.099)
RMSD of dihedral angles (maximum) (°)	0.507 ± 0.033 (2.838)	0.608 ± 0.034 (4.203)
RMSD of RDC (Hz)		
HN	0.911 ± 0.012	0.828 ± 0.038
NC′	0.258 ± 0.003	0.165 ± 0.009
C′Ca	0.624 ± 0.005	0.443 ± 0.013
RMSD of covalent geometry		
Bond lengths (Å)	0.001 ± 0.000	0.001 ± 0.000
Bond angles (°)	0.383 ± 0.005	0.353 ± 0.004
Impropers (°)	0.240 ± 0.006	0.239 ± 0.005
Q_{free}*	0.192 ± 0.004	0.185 ± 0.002
Ramachandran plot statistics (%)	residues 661–700	residues 661–700 (MYND), 1105–11 (SMRT)
Most favored	78.2	79.9
Additionally allowed	20.4	19.8
Generously allowed	1.3	0.3
Disallowed	0	0
RMSD for NMR ensemble (Å)	residues 661–700	residues 661–700 (MYND), 1105–11 (SMRT)
Backbone	0.18	0.14
Heavy atoms	0.72	0.61

* Q_{free} is calculated as described previously (Cornilescu et al., 1998) except that the RDCs are used only for validation but not for structure calculation.

1104–1109. A sequence similarity search revealed a similar binding site with the sequence PPPLI in N-CoR (aa 1033–1037). Isothermal titration calorimetry (ITC) was employed to determine the stoichiometry and binding affinity between the MYND domain and the SMRT and N-CoR peptides (Figure 1E). A further sequence matching in N-CoR and SMRT using a looser standard yielded one more potential binding site in both proteins: N-CoR 606–615

and SMRT 1664–1673 (Figure 1C). Interaction of the MYND domain with SMRT 1664–1673 was confirmed by an HSQC titration that showed similar chemical shift changes as seen for SMRT 1101–1113, though the affinity appeared weaker (data not shown).

Solution Structure of the SMRT-MYND Complex

We attempted to solve the structure of a MYND domain/SMRT peptide complex; however, half-filtered NOESY spectra of a differentially labeled complex ($^{13}C/^{15}N$-MYND + unlabeled SMRT) yielded only 12 intermolecular NOEs, which were insufficient to define the structure of the complex. Complexes with modest affinity often suffer from limited intermolecular NOE information due to significant off rates. To decrease the off rate and thereby facilitate the collection of NOE data, we fused the SMRT peptide sequence (SMRT 1101–1113) to the N terminus of the MYND domain (AML1/ETO 658–707) with a 7 aa linker (ENLYFQG) between them (subsequently referred to as SMRT-MYND). This coupling served to increase the effective local concentration of the SMRT peptide and resulted in longer residence times and increased NOEs. ^{15}N-^{1}H HSQC spectra of SMRT-MYND showed that resonances from the MYND portion of SMRT-MYND overlapped with resonances from the MYND + SMRT noncovalent complex, but not with those of the free MYND domain (Figure S1B). Furthermore, the 12 NOEs previously identified in the noncovalent complex were all present in NOESY spectra of SMRT-MYND (data not shown), indicating that the fusion did not change the mode of interaction. Indeed, $\{^{1}H\}^{15}N$ NOE measurements showed that the linker region is flexible, while the MYND and SMRT portions of the fusion protein display the same dynamic behavior as one another, consistent with a well-formed complex (Figure 2A).

The structure of MYND-SMRT showed that SMRT binds in an extended conformation to a hydrophobic pocket in MYND (Figure 2B), burying a total solvent-accessible surface area of 476 Å^2. L1108, I1109, and S1110 from SMRT form a short antiparallel β sheet with E672, T673, and C674 of the MYND domain, delineated by four interstrand hydrogen bonds (Figure 2C). The side chain of L1108 protrudes deep into the hydrophobic pocket of MYND (Figure 2B). This provides a structural rationalization for the importance of Leu at this position, which has been noted by others (Lausen et al., 2004). P1105 packs its side chain on top of W692 of MYND (Figure 2C). Such a proline-tryptophan packing arrangement has been frequently observed in protein-protein interactions (Glaser et al., 2001). Another residue showing 100% conservation among MYND domains is Q688, which forms a hydrogen bond to the CO of SMRT P1106 through Hε22 (Figure 2C). S675 also has a high degree of conservation among MYND domains (Figure 1C). The S675 γ proton potentially forms a hydrogen bond with L1108 CO of SMRT. Indeed this γ proton undergoes unusually slow exchange with water and displays a downfield shifted chemical shift of 6.47 ppm in MYND-SMRT, indicative of hydrogen bonding.

Mutations in the MYND Domain Disrupt SMRT Peptide Binding

We used the structure of the SMRT-MYND complex to rationally design mutations that disrupt SMRT/N-CoR binding in order to examine the importance of this interaction for AML1/ETO's function. H695 is a zinc-chelating residue, and ^{15}N-^{1}H HSQC spectra showed the MYND domain structure was severely disrupted upon substitution of H695 with Ala (Figure S2D), resulting in a protein which cannot bind to the SMRT peptide. Substitution of H689 with Ala, on the other hand, had little or no effect on either the structure or on SMRT binding (Figure S2D and Figure 3A). The W692A mutation resulted in perturbations only of spatially proximal residues in the ^{15}N-^{1}H HSQC spectrum (Figure 3B and Figure S2A), consistent with this mutation causing no significant change in the protein structure. Substitution of W692 with Ala substantially blocked binding to the SMRT peptide, increasing the K_d more than 12-fold (Figure 3A). Ala substitutions for Q688 and S675 preserved the MYND structure (Figures S2B and S2C) and impaired SMRT peptide binding 12-fold and 2.5-fold, respectively (Figure 3A).

Despite the fact that the W692A and H695A mutations decreased binding of the isolated MYND domain to the SMRT peptide, we were unable to detect a difference in full-length AML1/ETO's ability to coimmunoprecipitate full length SMRT or N-CoR when proteins containing either of these two mutations were overexpressed in Cos7 cells (data not shown). This was not entirely unexpected since we previously found that a C-terminal truncation of AML1/ETO that removed both the Nervy and MYND domains entirely did not affect SMRT or N-CoR binding using the same approach (Liu et al., 2006). AML1/ETO can oligomerize with itself, ETO, and the ETO homologs MTGR1 and MTGR2 (MTG16) (Kitabayashi et al., 1998; Liu et al., 2006; Minucci et al., 2000; Zhang et al., 2001). Therefore, our inability to detect an effect of the MYND mutations in coimmunoprecipitation assays could be due to AML1/ETO's ability to form mixed tetramers with endogenous ETO homologs that have intact MYND domains. In addition, a second N-CoR binding site was mapped to sequences between eTAFH and HHR (Amann et al., 2001), which could also explain the retention of N-CoR and SMRT binding by the full-length, MYND-mutated AML1/ETO proteins.

Contribution of the MYND-SMRT/N-CoR Interaction to AML1/ETO Function

Previous studies showed that deletion of the MYND domain impaired AML1/ETO's ability to repress the differentiation of established hematopoietic cell lines (Gelmetti et al., 1998; Klampfer et al., 1996; Lutterbach et al., 1998a). However, in primary bone marrow cells, the MYND domain's contribution to AML1/ETO's repression of granulocyte differentiation and proliferation and to its ability to confer serial replating was only observed when coupled to an HHR domain deletion (Hug et al., 2002). We transduced primary, lineage-depleted ($CD5^-$, $B220^-$, $Mac\text{-}1^-$, $Gr\text{-}1^-$, $Ter119^-$, and Lin^-) mouse bone marrow cells with

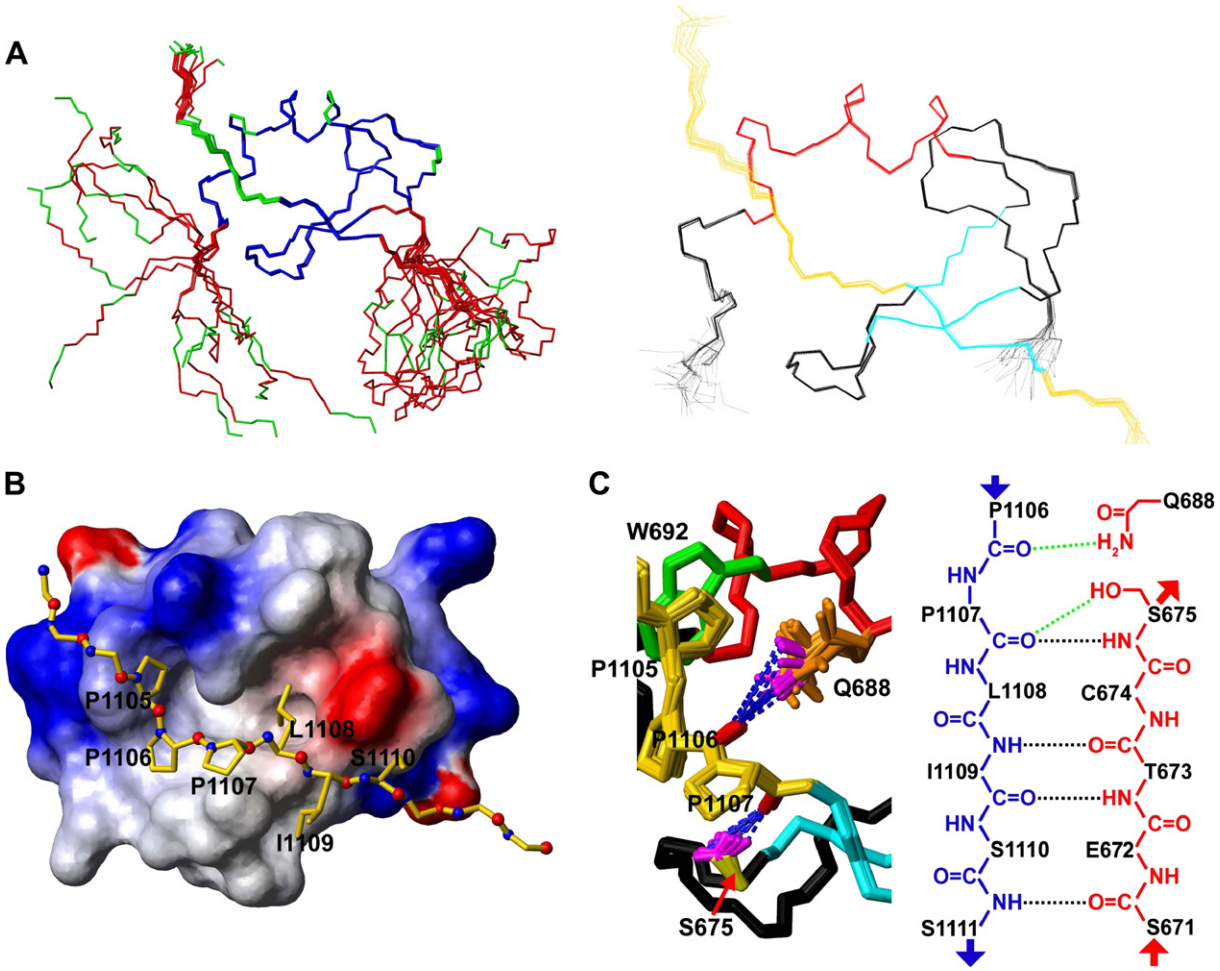

Figure 2. Structure of SMRT-MYND

(A) Left: Overlay of the backbone of 10 conformers representing the solution structure of SMRT-MYND. Residues with $\{^1H\}^{15}N$ NOE > 0.6 in blue, residues with $\{^1H\}^{15}N$ NOE < 0.6 in red, and residues with no $\{^1H\}^{15}N$ NOE information in green. Right: Overlay of the backbone of 25 conformers of SMRT-MYND, showing only the structured portion of the protein. The MYND domain is black with helical residues in red and β sheet residues in cyan. SMRT is gold with β sheet residues in cyan.

(B) Surface representation of the MYND domain with SMRT peptide bound. Electrostatic potential mapped onto surface (blue, positive electrostatic potential; red, negative electrostatic potential).

(C) Left: Sequence-specific interactions between the MYND domain and SMRT illustrating the stacking of MYND W692 and SMRT P1105, the hydrogen bond between MYND Q688 and SMRT P1106, and the hydrogen bond between MYND S675 and SMRT-P1107. Hydrogen bonds are displayed as blue dotted lines. Colors are the same as in the right figure in panel (A). Twenty-five conformers are shown. Right: Schematic illustration of the observed hydrogen bonds between MYND (red) and SMRT (blue). Backbone-backbone hydrogen bonds are black and sidechain-backbone hydrogen bonds are green.

retroviruses expressing green fluorescent protein (GFP) alone (MIGR1), AML1/ETO, or an AML1/ETO protein truncated immediately N-terminal to the MYND domain (W664X) (Figure 3C) and assessed granulocyte differentiation after 7 days of culture in the presence of IL-3, IL-6, SCF, and G-CSF in the successfully transduced (GFP^+) cells. Consistent with data reported by Hug et al. (Hug et al., 2002), we observed that deletion of the entire MYND domain and sequences C-terminal to it did not impair AML1/ETO's ability to inhibit the differentiation of $Gr\text{-}1^+$ $Mac\text{-}1^+$ cells from Lin^- bone marrow cells in vitro (Figures 3D and 3E). We therefore mutated seven Leu residues in the HHR domain at the tetramer interface that prevent its oligomerization (m7) (Liu et al., 2006) and assessed the effects of the MYND mutants in the context of the m7 HHR mutation. As we showed previously (Liu et al., 2006), the m7 mutation partially ameliorated AML1/ETO's inhibition of granulocyte differentiation, and deletion of the MYND domain in the context of this m7 mutation further weakened AML1/ETO's negative effect on granulocyte differentiation (Figures 3D and 3E). The W692A and H695A mutations in the context of the m7 mutation also significantly impaired, but did not completely reverse, AML1/ETO's repression of granulocyte differentiation (Figures 3D and 3E). All three m7 + MYND mutant AML1/ETO proteins accumulated to the same level as the m7 mutant in

retrovirally transduced NIH3T3 cells (Figure 3F). We could not detect the proteins by western blot in the transduced primary bone marrow cells (data not shown).

AML1/ETO inhibits the short-term proliferation of human and mouse primary bone marrow cells (Hug et al., 2002; Mulloy et al., 2002). It was previously reported that deletion of the MYND domain had no effect on AML1/ETO's antiproliferative activity (Hug et al., 2002). In our hands, deletion of the MYND domain (W664X) ameliorated this antiproliferative effect, although the percentage of cells that incorporated BrdU was still significantly lower than for MIGR1 transduced cells (Figure 3G). Both the H695A and W692A mutations alleviated AML1/ETO's repressive effect on proliferation to a similar extent as the W664X truncation (Figures 3G and 3H).

Deletion of the MYND domain did not affect AML1/ETO's ability to confer increased self-renewal capacity on hematopoietic progenitors in vitro, consistent with data from Hug et al. (Hug et al., 2002) (data not shown), so presumably the W692A and H595A mutations would also not affect AML1/ETO's serial replating activity, although we did not examine this directly. Since in our hands, the m7 mutation in the HHR domain destroyed AML1/ETO's ability to confer increased self-renewal (Liu et al., 2006), we could not use the m7 mutant as a sensitized background upon which to further examine the effects of the W692A and H695A mutations.

Loss of the MYND-SMRT/N-CoR Interaction Attenuates the AML1/ETO Signature in LSK Cells

To gain insights into the molecular basis by which eliminating SMRT/N-CoR binding through the MYND domain relieves the proliferation block caused by AML1/ETO in Lin^- bone marrow cells but not its ability to confer serial replating capacity, we analyzed the effect of the W692A mutation on gene expression changes caused by AML1/ETO. We prepared Lin^- $Sca\text{-}1^+$ mouse bone marrow cells, subdivided them into 11 pools, and independently transduced these cell pools with MIGR1, AML1/ETO, or the AML1/ETO W692A mutant (Figure 4A). Two days later, ~25,000 transduced (GFP^+) Lin^- $Sca1^+$ $c\text{-}kit^+$ (LSK) cells from each sample were sorted by FACS into RNA lysis buffer for RNA purification, labeling, and hybridization to Affymetrix Mouse Genome 430 2.0 arrays (Figures 4A and 4B). We drew a wide gate around the Sca-1 population for sorting purposes (Figure 4B) because cell surface Sca-1 levels varied considerably between samples (Figure 4F).

We uploaded the raw expression data to Iobion Informatic's GeneTraffic v. 3.2 server (Irizarry et al., 2003) and set expression values from the three MIGR1-transduced samples as the baseline for comparison to AML1/ETO and W692A transduced samples. Unsupervised clustering separately grouped the MIGR1, AML1/ETO, and W692A samples, indicating that the three gene profiles were significantly different (data not shown). We screened the data by significance analysis of microarrays (SAM) (Tusher et al., 2001) and found genes corresponding to 2229 probe sets were differentially expressed in AML1/ETO-transduced LSK cells as compared to LSK cells transduced with MIGR1 alone (Figure 4C). These included genes previously reported to be dysregulated in the presence of AML1/ETO in myelomonocytic cell lines (Alcalay et al., 2001; Shimada et al., 2000; Yan et al., 2004) (Figure 4E). One upregulated gene was *Ly6a*, which encodes the Sca-1 protein, and, indeed, increased levels of Sca-1 were detected on AML1/ETO-transduced LSK cells (Figure 4F). c-kit, on the other hand, was not differentially expressed either at the mRNA (data not shown, inferred from microarray data) or protein level (Figure 4F).

A smaller number of genes (823 probe sets) were differentially expressed in LSK cells transduced with the AML1/ETO W692A mutant as compared to LSK cells transduced with MIGR1 (Figure 4C). We compiled the two probe sets (AML1/ETO versus MIGR1 plus W692A versus MIGR1) into a single "master" list. There was considerable overlap between the two probe sets, thus the master list consisted of 2444 probe sets representing 1727 annotated genes (Figures 4C and 4D and Tables S1 and S2). We screened the 1727 annotated genes by SAM to identify those that were differentially regulated by AML1/ETO versus W692A, which yielded a list of 1231 genes (Figure 4C and Table S2). The expression of 1113 (90%) of these differentially expressed genes was attenuated by the W692A mutation. The remaining 118 differentially expressed genes were either dysregulated in the opposite direction by AML1/ETO and the W692A mutant, or dysregulated to a greater extent by the AML1/ETO W692A protein than by AML1/ETO. Thus, the most prominent pattern was that the W692A mutation attenuated AML1/ETO's activity (Table S2).

A trivial explanation for attenuated activity would be that the relative levels of the AML1/ETO W692A protein were lower. Unfortunately, we were unable to detect either the AML1/ETO or W692A proteins by western blot analysis of purified GFP^+ cells from retrovirally transduced primary bone marrow cells and therefore could not directly rule out this possibility. However, the levels of the bicistronic mRNAs encoding the AML1/ETO and W692A proteins were similar, based both on the intensity of GFP fluorescence (data not shown) and by qRT-PCR analysis (Figure S5). It still remains possible that the W692A protein accumulated at lower levels in LSK cells because it was no longer stabilized by the association with protein complexes normally recruited by the MYND domain.

We identified 496 genes (28.7% of the master gene list) whose expression was dysregulated to the same extent by AML1/ETO and the W692A mutant (Figure 4C and Table S2). Therefore, if the W692A mutation did affect AML1/ETO protein levels, it impacted the expression of some genes and not others.

We identified 79 cell-cycle-related genes whose expression was affected to a relatively small (ranging from 1.51- to 5.4-fold) but significantly different extent by AML1/ETO and the W692A mutant (Table 2). The differential expression of these genes presumably contributes to the major biological difference we observed, namely the repression of proliferation by AML1/ETO and the relief of that repression by the W692A mutation. The more minor effects we observed in granulocyte differentiation correlate with

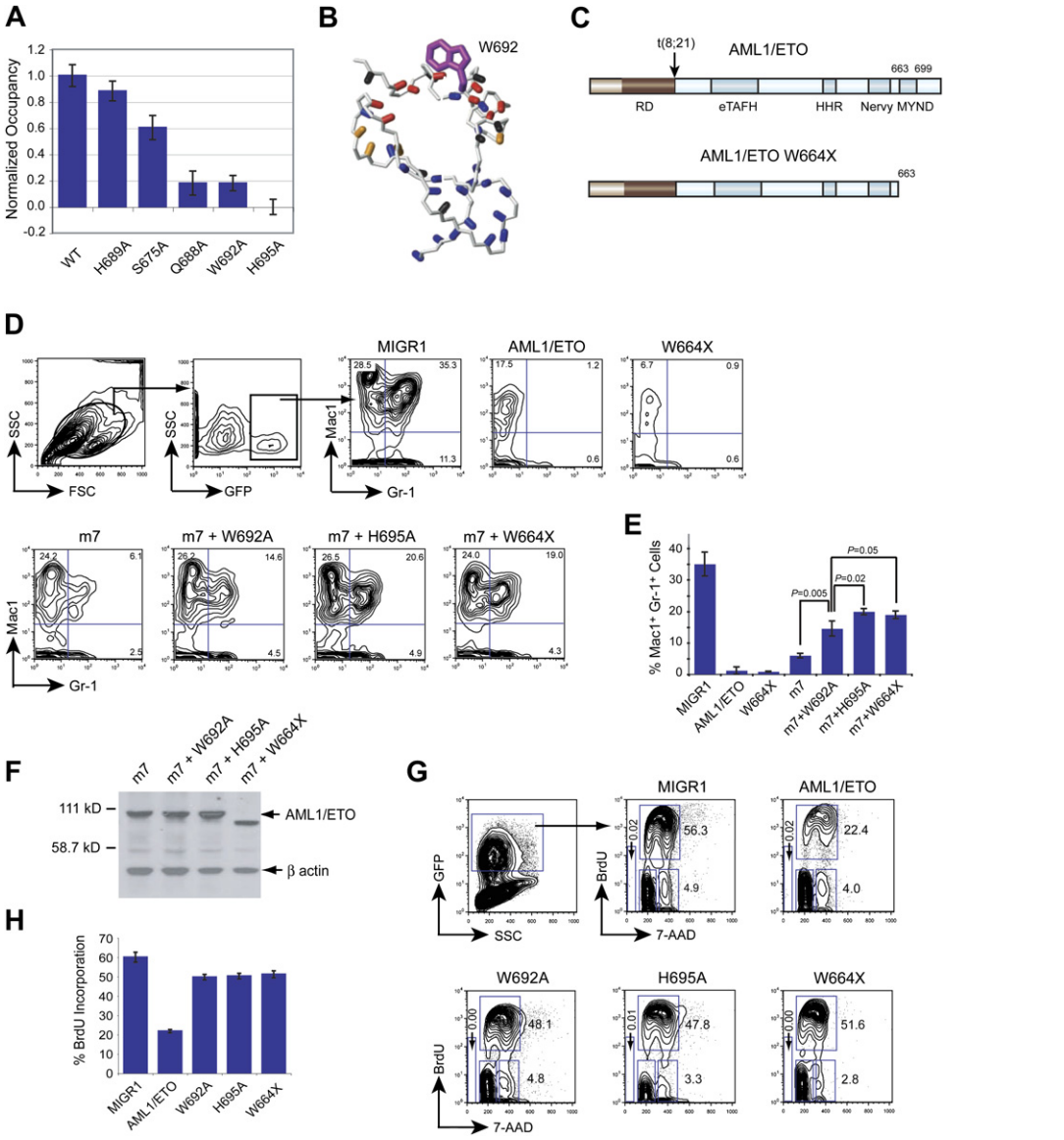

Figure 3. AML1/ETO Function Is Impaired by Mutations in the MYND Domain that Disrupt SMRT Binding

(A) Fluorescence anisotropy measurements of the SMRT1031-1273 fluorescein-labeled peptide binding with MYND mutants. Plot shows normalized change in occupancy (or normalized change in anisotropy) using the equation [FA(mutant) − FA(H695A)]/[FA(WT) − FA(H695A)], where FA(mutant), FA(H695A), and FA(WT) stand for the fluorescence anisotropy values of the peptide in the presence of the mutant MYND, of H695A MYND, and of wild-type MYND, respectively. Since H695A is unfolded, the fluorescence anisotropy in the presence of this protein was used as the zero point to correct for the effects of increased viscosity with increasing protein concentration, rather than the value for the peptide alone, on the fluorescence polarization measurement. SMRT peptide concentration was kept at 1 μM and WT and mutant MYND domains at 161 μM. Binding constants of the mutants relative to that of WT were calculated as $(([\text{mutant}]_{\text{total}} + K_D(\text{mutant}))/[\text{mutant}]_{\text{total}})/(([\text{WT}]_{\text{total}} + K_D(\text{WT}))/[\text{WT}]_{\text{total}}) = (\text{FA(WT)} - \text{FA(H695A)})/(\text{FA(mutant)} - \text{FA(H695A)})$, where $[\text{mutant}]_{\text{total}} = [\text{WT}]_{\text{total}} = 161$ μM, and $K_D(\text{WT}) = 109.4$ μM based on ITC. Error bars represent the mean ±SD for three independent measurements.

(B) Structure of the MYND domain illustrating the localized chemical shift perturbations caused by the W692A mutation. NH atoms are colored as follows: NH chemical shift change >0.1 ppm, red; NH chemical shift change between 0.04 and 0.1 ppm, orange; NH chemical shift change <0.04 ppm, blue; unassigned, black. The side chain of W692 is shown in purple.

(C) Schematic diagram of AML1/ETO and the truncated W664X mutant. Sequences derived from AML1 are brown, and those from ETO are blue.

(D) Representative flow of Lin$^-$ BM cells infected with MIGR1 retroviruses expressing AML1/ETO and mutated derivatives following seven days of culture in the presence of IL3, IL6, SCF, and G-CSF. Cells within the forward and side scatter gates were further gated for GFP expression, and GFP positive cells examined for Mac-1 and Gr-1 expression. The experiments were performed twice with triplicate samples.

(E) Average percentages of Gr-1$^+$Mac-1$^+$ cells (±SD) from a representative experiment with triplicate samples.

(F) Western blot probed with an antibody to the Runt domain, demonstrating expression of the m7 AML1/ETO HHR mutant and mutated MYND derivatives thereof in retrovirally-transduced NIH3T3 cells.

differences in the expression of genes required for that process (Table 3).

DISCUSSION

The structure of the MYND domain/SMRT peptide complex provided insights into how specificity is achieved and structure-based mutations enabled us to assess the contribution of SMRT/N-CoR binding through the MYND domain to AML1/ETO function. As is the case with other domains such as SH3 and WW that bind to Pro-rich sequences, the question arises how specificity can be achieved with a binding motif consisting exclusively of nonpolar residues that cannot form specific hydrogen bonds or electrostatic interactions. In the case of the MYND domain/SMRT (N-CoR) binding, specificity is achieved through a combination of proline-tryptophan packing interactions, other van der Waals interactions via the complementary surfaces, and by means of the conformational rigidity of the Pro-rich sequence that minimizes the entropic cost of binding and positions backbone nuclei favorably for hydrogen bond formation. The first Pro in the "PPPLIP" motif packs against W692. The carbonyls of the second and third prolines form hydrogen bonds with Q688 and S675, respectively. Although hydrogen bonds to the backbone generally cannot provide specificity, the relative geometrical positions of W692, Q688, and S675 on the MYND domain favor substrates with a certain backbone conformation for hydrogen bond formation. N-CoR binds the MYND domain with higher affinity than SMRT, and the only difference in the directly interacting region is that P1038 of N-CoR, which is one of the three consecutive residues forming a β strand with the MYND domain, is replaced by S1110 in SMRT. S1110, with its HN pointing away from the β strand, does not seem to participate in any specific interaction, and it is not obvious why a Pro at this position increases binding affinity. It is tempting to speculate that MYND's higher affinity for N-CoR is because the proline reduces the entropic penalty for binding. Indeed, ITC data show that ΔH is quite similar for the MYND/N-CoR and MYND/SMRT interactions (-9883 cal mol^{-1} and -9836 cal mol^{-1}, respectively) while it is ΔS that dominates the difference in binding affinity (–12.4 cal mol^{-1} K^{-1} and –14.5 cal mol^{-1} K^{-1}, respectively).

Using the structural information, we introduced mutations in the AML1/ETO MYND domain that specifically disrupted SMRT (and presumably N-CoR) binding without significantly affecting the MYND domain structure. Impairing SMRT/N-CoR binding to the MYND domain attenuated AML1/ETO's activity only very modestly in primary bone marrow cells, as had been shown by others using an MYND domain deletion (Hug et al., 2002). AML1/ETO's acute repression of cell proliferation was alleviated but not completely eliminated, and its negative effect on granulocyte differentiation was ameliorated, but only in the context of a mutation that impaired oligomerization. In fact, the presence of an oligomerization domain (HHR) in AML1/ETO complicates efforts to document the importance of other ETO sequences for AML/ETO's activity since the formation of mixed tetramers with endogenous ETO or the ETO homologs MTGR1 and MTGR2 could dampen a mutation's effects. Nevertheless, the effect of a single amino acid substitution that decreased SMRT/N-CoR binding by 12-fold was relatively extensive, at least with respect to the number of genes (1231) whose expression was significantly different by SAM analysis. Presumably included among these are genes whose differential expression underlies the perturbations in proliferation and granulocyte differentiation that were observed by us and others (Gelmetti et al., 1998; Hug et al., 2002; Klampfer et al., 1996).

Cell-cycle-related genes that were differentially expressed included those with established roles in cell-cycle regulation and checkpoint control, DNA repair, DNA replication, and mitosis. Many genes listed in Table 2 promote these processes and were downregulated in the presence of AML1/ETO, with this effect attenuated by the W692A mutation. Conversely, several genes that negatively regulate cell-cycle progression or proliferation (*Ccng2*, *G0s2*, *Tgfb1*) were upregulated by AML1/ETO and less so by the AML1/ETO W692A mutant. The largest absolute change in expression among this group was observed for *Vegfc* which is 75-fold upregulated with AML1-ETO and attenuated to 14-fold with AML1-ETO W692A. Elevated levels of VEGF have been observed for AML patients and correlate with reduced survival as well as reduced complete remission rates in these patients (Aguayo et al., 2000, 2002), arguing for the potential applicability of antiangiogenic therapy in t(8;21) leukemia.

AML1/ETO expression also resulted in the downregulation of genes involved in myeloid cell differentiation (Table 3). The most highly differentially expressed gene in this group is *Gfi1*, which promotes T cell proliferation, granulocyte maturation, restricts the proliferation, and thus maintains the quiescence of hematopoietic stem cells (Hock and Orkin, 2006). *Gfi1* expression was upregulated by AML1/ETO (5.3-fold), and that upregulation was attenuated by the W692A mutation. *Cebpa*, a known suppressor of leukemogenesis and positive regulator of granulopoiesis (Nerlov, 2004), was downregulated 8.3-fold in the presence of AML1/ETO, and that repression was alleviated by the W692A mutation.

Some of the more interesting genes are likely to be those that are highly dysregulated by both AML1/ETO and the W692A mutant. Deletion of the MYND domain

(G) BrdU incorporation 48 hr after transduction of Lin$^-$ bone marrow cells with MIGR1 expressing GFP, AML1/ETO, or the AML1/ETO MYND mutants. GFP$^+$ cells were analyzed for BrdU and 7-AAD incorporation following a 1 hr BrdU pulse. No differences in sub G1 (apoptotic) cells were observed. Shown is a representative of three experiments.

(H) Average percentage of gated BrdU$^+$ cells from scatter plots in panel G (±SD) (n = 3). All samples were significantly different ($p \leq 0.01$) from either MIGR1 or AML1/ETO transduced cells.

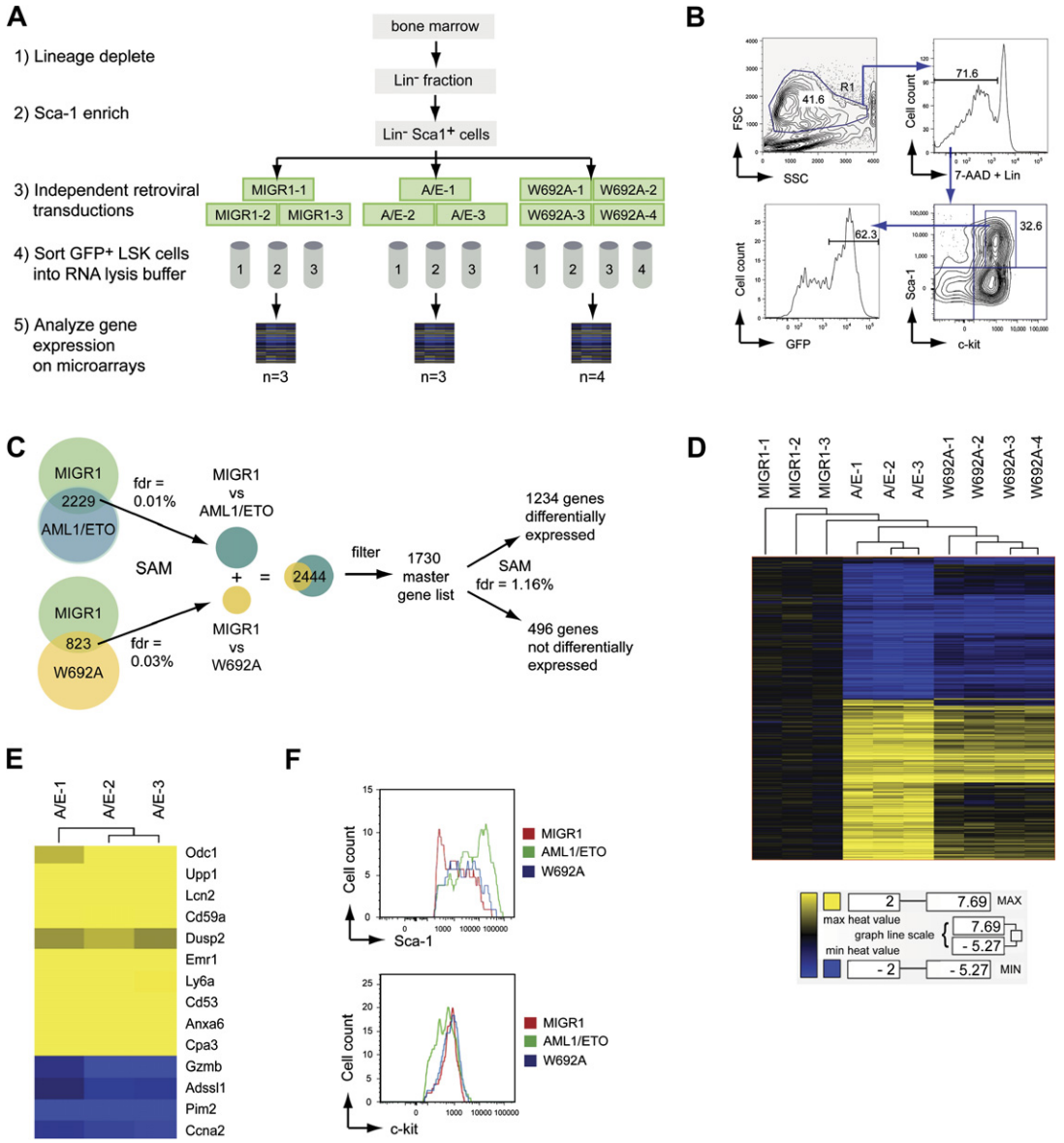

Figure 4. Gene Expression Profiles of LSK Cells Expressing AML1/ETO nd the AML1/ETO W692 Mutant

(A) Experimental scheme.

(B) Isolation of the retrovirally transduced (GFP+) LSK cells used for microarray analyses.

(C) Microarray data analysis. Expression values from the three MIGR1-transduced samples were set as the baseline for comparison to AML1/ETO- and AML1/ETO W692A-transduced samples. The AML1/ETO and AML1/ETO W692A data were filtered to remove all genes without at least 3 observations of log2 ratio = 0.58 (1.5-fold) from MIGR1 samples, yielding 6034 probe sets. SAM performed on AML1/ETO versus MIGR1 with a false discovery rate (FDR) of 0.01% identified 2229 significant probe set changes. SAM on AML1/ETO W692A versus MIGR1 with an FDR of 0.03% identified 823 significant changes. These were combined to yield a master list of 2444 probe sets, which are provided in Table S1. The 2444 probe sets were manually filtered, removing those corresponding to nonannotated sequences, "hypothetical proteins," "cDNA sequences," and "expressed sequences." In addition, for genes whose expression was reported as significant by multiple probe sets, only one probe set was included in the subsequent data analyses. The filtered list contained 1727 annotated genes (Table S2). SAM was performed on these 1727 genes to compare expression changes caused by AML1/ETO and AML1/ETO W692A (relative to MIGR1) with FDR = 1.16%, from which 1231 genes that were significantly differently expressed and 496 genes that were not significantly differently expressed were identified (Table S2).

(D) Unsupervised hierarchical clustering of the master list of 1727 genes using the average clustering of Pearson correlation coefficient, depicted as a heat map. Yellow represents genes overexpressed relative to the average level in the three MIGR1-transduced samples, and blue represents underexpressed genes.

(E) Examples of previously identified genes (Shimada et al., 2000; Yan et al., 2004) differentially regulated by AML1/ETO in LSK cells.

(F) Histograms documenting cell surface expression of Sca-1 and c-kit on the transduced LSK cells.

Table 2. Cell-Cycle Related Genes Dysregulated ≥2-Fold by AML1/ETO and to a Lesser Extent or Not at All by the AML1/ETO W692A Mutant

Gene Symbol	Fold Change AML1/ETO	Fold Change AML1/ETO W692A
Il1a	5.86	1.08
Vegfc	75.06	14.32
Cxcl1	9.51	2.00
Il1b	8.13	2.48
Dst	6.79	2.44
G0s2	3.55	1.42
Sash1	4.17	1.69
Mcm2	−3.25	−1.38
Jun	4.86	2.08
Ranbp1	−2.37	−1.02
Rprm	2.75	1.19
Fancd2	−2.69	−1.18
Mycn	−4.13	−1.81
Ccdc5	−2.70	−1.22
Ccna2	−2.83	−1.29
Tacc3	−2.89	−1.32
Exo1	−3.56	−1.64
Skp2	−2.50	−1.16
Gas2l1	2.91	1.37
Ckap2	−2.71	−1.29
Bub1b	−2.45	−1.17
Prc1	−2.35	−1.13
Ccnb1	−2.12	−1.02
Itgb1	2.23	1.08
Gsg2	−3.03	−1.47
Pdgfa	2.27	1.11
Ccng2	2.28	1.12
Mcm6	−2.28	−1.13
Bub1	−2.57	−1.29
Aurkb	−3.23	−1.63
Fbxo5	−2.10	1.07
Rad51	−2.34	−1.20
Cdca5	−2.23	−1.15
App	3.23	1.67
Ris2	−2.24	−1.17
Sesn1	2.35	1.24
Sgol1	−2.27	−1.20
Mcm7	−2.37	−1.27
Fos	13.12	7.04
Tgfb1	4.25	2.34
Aatf	−2.76	−1.53
Chek1	−2.49	−1.38
Dusp1	3.27	1.84
Racgap1	−2.37	−1.36
Mapk3	2.58	1.50
Rras	2.54	1.50
Topbp1	−2.66	−1.57
Mcm8	−2.53	−1.49
Sesn3	3.06	1.86
Cdc45l	−2.11	−1.28
Kif11	−2.39	−1.46
Dnajc2	−2.14	−1.30
Mad2l1	−2.08	−1.26
Mcm3	−2.96	−1.80
Ran	−2.25	−1.38
Mlh1	−2.12	−1.34
Spag5	−2.46	−1.55
Incenp	−2.67	−1.69
Cdc7	−2.65	−1.69
Tfdp1	−2.18	−1.39
Elk3	2.40	1.55
Ppp3ca	2.26	1.47
Atm	−2.68	−1.75
Nbn	−2.31	−1.57
Casp3	2.31	1.59
Tube1	−2.58	−1.78
Rassf5	−2.42	−1.68
Ccnd1	−3.82	−2.73
Ddx11	−2.20	−1.58
Pard6g	−2.35	−1.69
Mre11a	−2.80	−2.03
Rbl1	−2.18	−1.60
Atr	−2.45	−1.83
Myc	−2.36	−1.84
Hus1	−2.22	−1.80
Ccnt2	−2.14	−1.77

Genes are listed in the order of the magnitude difference between AML1/ETO and AML1/ETO W692A (down to a 1.2-fold difference). Fold change represents the difference in fold expression relative to MIGR1-transduced LSK cells. The shaded area indicates genes whose fold expression difference is ≥2.

Table 3. Genes Involved in Myeloid Cell Differentiation and/or Function that Were Differentially Expressed by AML1/ETO (≥2-fold) and the AML1/ETO W692A Mutant

Gene Symbol	Gene Name	Fold Change AML1/ETO	Fold Change AML1/ETO W692A
Gfi1	growth factor independent 1	5.28	1.16
Il4	interleukin 4	4.59	1.16
Cbfa2t3h	core-binding factor, runt domain, alpha subunit 2, translocated to, 3 homolog (human)	−6.05	−2.01
Cebpa	CCAAT/enhancer-binding protein (C/EBP), alpha	−8.32	−2.82
Cxcl4	chemokine (C-X-C motif) ligand 4	3.60	1.41
Nfe2	nuclear factor, erythroid derived 2	−8.07	−3.30
Pirb	paired-Ig-like receptor B	2.86	1.21
Trem3	triggering receptor expressed on myeloid cells 3	−2.88	−1.31
Myadm	myeloid-associated differentiation marker	2.66	1.23
Mitf	microphthalmia-associated transcription factor	2.60	1.27
Ncor2	nuclear receptor corepressor 2	−3.69	−1.81
Ngp	neutrophilic granule protein	−9.11	−4.49
Plscr1	phospholipid scramblase 1	4.07	2.07
Ifi203	interferon-activated gene 203	2.46	1.28
Mlf1ip	myeloid leukemia factor 1 interacting protein	−2.27	−1.20
Cd300lf	CD300 antigen like family member F	2.23	1.28
Hmgb3	high mobility group box 3	−2.72	−1.70
Spib	Spi-B transcription factor (Spi-1/PU.1 related)	−2.22	−1.47
Mcl1	myeloid cell leukemia sequence 1	2.09	1.46
Bcl11a	B cell CLL/lymphoma 11A (zinc finger protein)	−4.18	−3.07
Tirap	toll-interleukin 1 receptor (TIR) domain-containing adaptor protein	−3.97	−2.92
Gfi1b	growth factor-independent 1B	−3.55	−2.61
Runx1	runt-related transcription factor 1	2.49	1.83

Genes are listed in the order of the magnitude difference between AML1/ETO and AML1/ETO W692A. Fold change represents the difference in fold expression relative to MIGR1-transduced LSK cells. The shaded area indicates genes whose fold expression difference between AML1/ETO and W692A is ≥2.

did not affect one of AML1/ETO's most important properties, that is its ability to confer serial replating to primary bone marrow cells, which is a readout of self renewal potential (Hug et al., 2002). Furthermore, a truncated form of AML1/ETO or a naturally occurring splice variant that lacks both the Nervy and MYND domains was leukemogenic in mice (Yan et al., 2004). In fact, the truncated AML1/ETO was more highly leukemogenic than the full-length protein and could cooperate with full-length AML1/ETO to induce a rapid disease (Yan et al., 2004, 2006). One of the genes whose expression was most highly upregulated by both AML1/ETO (24-fold) and the W692 mutant (23-fold) was high mobility group protein 2 (*Hmga2*). HMGA proteins are small, architechural transcription factors involved in proliferation, apoptosis, differentiation, and senescence (Narita et al., 2006; Reeves, 2001). Upregulation of *HMGA2* expression has been found in a variety of solid tumors in humans, and chromosomal rearrangements of *HMGA2* are common in benign mesenchymal tumors (Reeves, 2001). In senescent cells, *HMGA* was shown to repress *MCMs* and *cyclin A* (Narita et al., 2006). As shown in Table 2, *cyclin A* and *MCMs* 2, 3, 6, 7, and 8 are all downregulated by AML1/ETO as well as the mutant AML1/ETO, albeit to a lesser extent, suggesting increased *HMGA2* may well contribute to the AML1/ETO phenotype.

In summary, our results demonstrate that recruitment of SMRT/N-CoR by the MYND domain contributes to the transcriptional dysregulation caused by AML1/ETO and to its repression of both proliferation and granulocyte differentiation. The two-hit hypothesis for leukemia development posits the need for both a block in differentiation and a proliferative advantage (Gilliland and Tallman, 2002). It is well established in mouse models that AML1-ETO by itself does not cause leukemia but requires additional cooperating mutations (Grisolano et al., 2003; Schessl et al., 2005; Schwieger et al., 2002). Full-length AML1-ETO can clearly

achieve the block in differentiation posited in the two-hit model but is not able to provide the necessary proliferative advantage. The recent demonstration by Zhang and coworkers (Yan et al., 2004, 2006) that a truncated form of AML1-ETO lacking the Nervy and MYND domains causes leukemia, unlike AML1-ETO itself, suggests that a truncated form of the protein can provide both the differentiation block and the proliferative advantage necessary. This is in good agreement with our results showing that loss of SMRT/N-CoR binding to the MYND domain results in increased proliferation, i.e., abrogation of MYND domain function provides or contributes significantly to the necessary proliferative advantage. The loss of this domain in the truncation in the case of truncated AML1-ETO or altered regulation, perhaps by phosphorylation, of this domain in the case of other cooperating mutations may provide a mechanism to achieve both hits through one protein. In the latter case, the signaling pathways involved in such regulation may therefore prove to be useful targets for therapeutic intervention.

EXPERIMENTAL PROCEDURES

Protein Expression and Purification

All proteins were cloned, expressed, and purified using standard procedures (details in Supplemental Data).

Peptide Binding Measurements by Fluorescence Anisotropy

All fluorescence experiments were carried out at 25°C on a FluroMax-3 fluorescence spectrophotometer (Horiba Jobin Yvon, Edison, NJ) with an external temperature control unit. The concentration of fluorescein-labeled peptides was kept at 5 μM during SMRT binding epitope screening and 1 μM during MYND mutant binding experiments. Peptides 12 aa in length with 9 aa overlaps spanning the entire SMRT sequence and with an N-terminal fluorescein were employed for epitope mapping. The excitation and emission wavelengths were set to 494 nm and 525 nm, respectively. Anisotropy values were averaged over 1 s during the epitope screening experiments and over 5 s during the MYND mutant binding experiments following a 2 min pre-equilibration at 25°C.

Peptide Binding Measurements by Isothermal Titration Calorimetry

ITC experiments were carried out on a MCS-ITC isothermal titration calorimetric system (MicroCal, Northampton MA) at 26°C. The calorimetric cell containing 170 or 257 μM MYND domain was titrated with 3.3 mM SMRT (1101–1113: YG-TISN*PPPLI*SSAK) or 2 mM N-CoR (N-CoR 1000–1011; YG-TRP*PPPLI*PSSK) peptides. Both MYND and SMRT/N-CoR were thoroughly dialyzed into the same buffer containing 25 mM Bis-Tris (pH 7.5), 0.1 M NaCl, 1 mM DTT, and 50 μM $ZnCl_2$. Data were analyzed using Origin 7.0 (OriginLab, Northampton MA).

Analytical Ultracentrifugation

Sedimentation velocity runs were done at 20°C using an AN-60 titanium rotor (Beckman) and SedVel60 centerpieces with 1.2 cm path length and quartz windows. The runs were done at 35,000 rpm, and interference scans were acquired using Beckman software. Analyses were done using Sedfit (Schuck, 2000) with (96 resolution, 0.5–10 s) radial and time independent noise at 95% confidence intervals.

NMR Spectroscopy

Samples for structure determination were prepared in 25 mM Bis-Tris (pH 6.8), 50 μM $ZnCl_2$, and 1 mM DTT. One hundred millimolar NaCl was included in the buffer for charged-gel samples to alleviate line broadening due to protein-gel interaction. Data were collected on a 600 MHz Varian Inova Spectrometer (Palo Alto, CA) equipped with a cryogenically cooled probe at 32°C. Resonance assignment and NOESY data collection were carried out using standard experiments.

For alignment, a negatively charged gel compressed to 4% concentration and consisting of 50% 2-acrylamido-2-methyl-1-propanesulfonic acid and 50% acrylamide (50 − M) and a zwitterionic charged gel compressed to 4% concentration consisting of 50% of a 1:1 mixture of (3-acrylamidopropyl)-trimethylammonium chloride/acrylic acid and 50% acrylamide (50 + MA) were employed (Cierpicki and Bushweller, 2004). Three types of dipolar couplings, $^1D_{HN}$, $^1D_{NC'}$, and $^1D_{C'C\alpha}$, were measured using IPAP (Ottiger et al., 1998) and 3D HNCO-based experiments (Yang et al., 1999). All NMR spectra were processed and analyzed using NMRPipe and Sparky (T.D. Goddard and J.M. Kneller, University of California, San Francisco).

Structure Calculations

Structure calculations were carried out using the simulated annealing protocol in CNS (Brunger et al., 1998). Details are provided in the Supplemental Data.

Retroviral Transductions

Mutated AML1/ETO proteins were transferred from pBluescript into the MIGR1 vector (Pear et al., 1998), and retroviruses were prepared as described previously (Liu et al., 2006). Primary bone marrow mononuclear cells were harvested from 5- to 6-week-old C57BL/6 mice (Jackson Labs, Bar Harbor, ME) and subjected to immunomagnetic negative selection using the Lineage cell depletion Kit (Miltenyi Biotech, Auburn, CA). For the microarray experiments, the Lin^- cells were further enriched for $Sca1^+$ cells by positive selection using the Anti-Sca-1 MicroBead Kit (Miltenyi Biotech). Cells (6×10^5 in 3 ml) were plated in ultralow adhesion 6-well plates (Costar, Corning, NY) and incubated overnight at 37°C, 5% CO_2 in transplant media (RPMI, 20% FCS, penicillin and streptomycin, 10 ng/ml IL-3, 20 ng/ml IL-6 [R&D Systems, Minneapolis, MN], 10 ng/ml SCF [Stem Cell technologies, Vancouver, Canada]). Twelve hours later, 6×10^5 cells in 2 ml of fresh transplant media were added to 6-well plates (Cellstar, GBO, NC) coated with 100 μg Retronectin (Takara, Madison, WI). Retroviral supernatants (2 ml), 4 μl of 40 mg/ml polybrene and 40 μl 1 M Hepes were added to each well, and the cells were centrifuged for 90 min at 1400 g, 37°C. Cells for microarray experiments were subjected to a second round of transduction 24 hr later as described above.

Granulocyte Differentiation

Granulocyte differentiation was assessed as described previously (Liu et al., 2006).

Cell-Cycle Kinetics

BrdU incorporation by Lin^- bone marrow cells was assessed 48 hr after retroviral transduction according to the manufacturer's instructions (APC-BrdU flow kit, BD PharMingen, San Diego, CA). Cells were incubated with 30 μM BrdU for 1 hr and stained with Allophcocyanin (APC)-conjugated anti-BrdU antibody and 7-AAD. Data were collected on a FACSCalibur (BD Biosciences, San Jose CA) and analyzed by FlowJo (version 6.1.1, Tree Star, San Carlos, CA).

Microarray Analyses

Lin^- $Sca1^+$ bone marrow cells independently transduced with MIGR1 (three separate pools), AML1/ETO (three pools), or AML1/ETO W692A (four pools) were stained 48 hr later with PE conjugated anti-CD117 (clone 2B8, BD Bioscience), PerCp-Cy5.5 conjugated Streptavidin (BD Bioscience), APC-conjugated anti-Sca1 (clone D7, Ebiosciences, SanDiego, CA), and biotin-conjugated antibody cocktail (Miltenyi Biotech). Lin^- $Sca\text{-}1^+$ $c\text{-}kit^+$ GFP^+ cells were isolated on a FACSAria (Becton Dickinson). Approximately 25,000 cells were collected for each sample. Cells were sorted into 600 μl of RNA lysis buffer (Qiagen buffer RLT with β-mercaptoethanol, Valencia CA), incubated at 42°C for 1 hr, and vortexed for 1 min to facilitate cell lysis. Total RNA was extracted

and DnaseI-treated using Rneasy Mini and DnaseI kits (Qiagen) according to the manufacturer's protocol. Total RNA quantity was measured using a ND-1000 spectrophotometer (Nano-Drop, Wilmington, DE), and RNA integrity was verified on an Agilent Bioanalyzer using a RNA 6000 Pico chip (Agilent Technologies, Santa Clara, CA). To minimize potential sample variability, 50 ng of total RNA from each sample was used as template for T7-based linear RNA amplification (RiboAmp kit, Arcturus, Mountain View, CA). Poly-inosinic poly-cytidilic acid (200 ng, Sigma, St. Louis, MO) was added as a carrier during the first round of cDNA synthesis. 1 μg of double-stranded cDNA derived from a second round of cDNA synthesis (RiboAmp kit) was used as template for generating biotinylated cRNA (BioArray High Yield Transcription kit, Enzo Life Sciences, Farmingdale NY). Labeled cRNA was then DnaseI-treated and purified on Rneasy minicolumns (Qiagen). Samples were submitted to the Norris Cotton Cancer Center and Dartmouth Medical School Microarray Shared Resource facility for fragmentation and hybridization to Mouse Whole Genome 430 2.0 arrays (Affymetrix, Santa Clara, CA).

Supplemental Data

The Supplemental Data include Supplemental Experimental Procedure, five supplemental figures, and two supplemental tables and can be found with this article online at http://www.cancercell.org/cgi/content/full/11/6/483/DC1/.

ACKNOWLEDGMENTS

We thank Carol Ringelberg and Richard Cowper for bioinformatics support. This work was supported by RO1 CA108056 (J.H.B. and N.A.S.). M.D.C was supported by T32 GM08704 and J.G. by T32 AR07576.

Received: January 4, 2007
Revised: February 23, 2007
Accepted: April 2, 2007
Published: June 11, 2007

REFERENCES

Aguayo, A., Kantarjian, H., Manshouri, T., Gidel, C., Estey, E., Thomas, D., Koller, C., Estrov, Z., O'Brien, S., Keating, M., et al. (2000). Angiogenesis in acute and chronic leukemias and myelodysplastic syndromes. Blood *96*, 2240–2245.

Aguayo, A., Kantarjian, H.M., Estey, E.H., Giles, F.J., Verstovsek, S., Manshouri, T., Gidel, C., O'Brien, S., Keating, M.J., and Albitar, M. (2002). Plasma vascular endothelial growth factor levels have prognostic significance in patients with acute myeloid leukemia but not in patients with myelodysplastic syndromes. Cancer *95*, 1923–1930.

Alcalay, M., Orleth, A., Sebastiani, C., Meani, N., Chiaradonna, F., Casciari, C., Sciurpi, M.T., Gelmetti, V., Riganelli, D., Minucci, S., et al. (2001). Common themes in the pathogenesis of acute myeloid leukemia. Oncogene *20*, 5680–5694.

Amann, J.M., Nip, J., Strom, D.K., Lutterbach, B., Harada, H., Lenny, N., Downing, J.R., Meyers, S., and Hiebert, S.W. (2001). ETO, a target of t(8;21) in acute leukemia, makes distinct contacts with multiple histone deacetylases and binds mSin3A through its oligomerization domain. Mol. Cell. Biol. *21*, 6470–6483.

Ansieau, S., and Leutz, A. (2002). The conserved Mynd domain of BS69 binds cellular and oncoviral proteins through a common PXLXP motif. J. Biol. Chem. *277*, 4906–4910.

Bravo, J., Li, Z., Speck, N.A., and Warren, A.J. (2001). The leukemia-associated AML1 (Runx1)–CBF beta complex functions as a DNA-induced molecular clamp. Nat. Struct. Biol. *8*, 371–378.

Brunger, A.T., Adams, P.D., Clore, G.M., DeLano, W.L., Gros, P., Grosse-Kunstleve, R.W., Jiang, J.S., Kuszewski, J., Nilges, M., Pannu, N.S., et al. (1998). Crystallography & NMR system: A new software suite for macromolecular structure determination. Acta Crystallogr. D Biol. Crystallogr. *54*, 905–921.

Cierpicki, T., and Bushweller, J.H. (2004). Charged gels as orienting media for measurement of residual dipolar couplings in soluble and integral membrane proteins. J. Am. Chem. Soc. *126*, 16259–16266.

Cornilescu, G., Marquardt, J.L., Ottiger, M., and Bax, A. (1998). Validation of protein structure from anisotropic carbonyl chemical shifts in a dilute liquid crystalline phase. J. Am. Chem. Soc. *120*, 6836–6837.

Fukuyama, T., Sueoka, E., Sugio, Y., Otsuka, T., Niho, Y., Akagi, K., and Kozu, T. (2001). MTG8 proto-oncoprotein interacts with the regulatory subunit of type II cyclic AMP-dependent protein kinase in lymphocytes. Oncogene *20*, 6225–6232.

Gelmetti, V., Zhang, J., Fanelli, M., Minucci, S., Pelicci, P.G., and Lazar, M.A. (1998). Aberrant recruitment of the nuclear receptor corepressor-histone deacetylase complex by the acute myeloid leukemia fusion partner ETO. Mol. Cell. Biol. *18*, 7185–7191.

Gilliland, D.G., and Tallman, M.S. (2002). Focus on acute leukemias. Cancer Cell *1*, 417–420.

Glaser, F., Steinberg, D.M., Vakser, I.A., and Ben-Tal, N. (2001). Residue frequencies and pairing preferences at protein-protein interfaces. Proteins *43*, 89–102.

Gottlieb, P.D., Pierce, S.A., Sims, R.J., Yamagishi, H., Weihe, E.K., Harriss, J.V., Maika, S.D., Kuziel, W.A., King, H.L., Olson, E.N., et al. (2002). Bop encodes a muscle-restricted protein containing MYND and SET domains and is essential for cardiac differentiation and morphogenesis. Nat. Genet. *31*, 25–32.

Grisolano, J.L., O'Neal, J., Cain, J., and Tomasson, M.H. (2003). An activated receptor tyrosine kinase, TEL/PDGFßR, cooperates with AML1/ETO to induce acute myeloid leukemia in mice. Proc. Natl. Acad. Sci. USA *100*, 9506–9511.

Gross, C.T., and McGinnis, W. (1996). DEAF-1, a novel protein that binds an essential region in a Deformed response element. EMBO J. *15*, 1961–1970.

Hock, H., and Orkin, S.H. (2006). Zinc-finger transcription factor Gfi-1: Versatile regulator of lymphocytes, neutrophils and hematopoietic stem cells. Curr. Opin. Hematol. *13*, 1–6.

Hug, B.A., and Lazar, M.A. (2004). ETO interacting proteins. Oncogene *23*, 4270–4274.

Hug, B.A., Lee, S.Y., Kinsler, E.L., Zhang, J., and Lazar, M.A. (2002). Cooperative function of Aml1-ETO corepressor recruitment domains in the expansion of primary bone marrow cells. Cancer Res. *62*, 2906–2912.

Irizarry, R.A., Hobbs, B., Collin, F., Beazer-Barclay, Y.D., Antonellis, K.J., Scherf, U., and Speed, T.P. (2003). Exploration, normalization, and summaries of high density oligonucleotide array probe level data. Biostatistics *4*, 249–264.

Kitabayashi, I., Ida, K., Morohoshi, F., Yokoyama, A., Mitsuhashi, N., Shimizu, K., Nomura, N., Hayashi, Y., and Ohki, M. (1998). The AML1–MTG8 leukemic fusion protein forms a complex with a novel member of the MTG8(ETO/CDR) family, MTGR1. Mol. Cell. Biol. *18*, 846–858.

Klampfer, L., Zhang, J., Zelenetz, A.O., Uchida, H., and Nimer, S.D. (1996). The AML1/ETO fusion protein activates transcription of BCL-2. Proc. Natl. Acad. Sci. USA *93*, 14059–14064.

Ladendorff, N.E., Wu, S., and Lipsick, J.S. (2001). BS69, an adenovirus E1A-associated protein, inhibits the transcriptional activity of c-Myb. Oncogene *20*, 125–132.

Lausen, J., Cho, S., Liu, S., and Werner, M.H. (2004). The nuclear receptor co-repressor (N-CoR) utilizes repression domains I and III for interaction and co-repression with ETO. J. Biol. Chem. *279*, 49281–49288.

Liu, Y., Cheney, M.D., Gaudet, J.J., Chruszcz, M., Lukasik, S.M., Sugiyama, D., Lary, J., Cole, J., Dauter, Z., Minor, W., et al. (2006). The

tetramer structure of the Nervy homology two domain, NHR2, is critical for AML1/ETO's activity. Cancer Cell *9*, 249–260.

Lutterbach, B., Sun, D., Schuetz, J., and Hiebert, S.W. (1998a). The MYND motif is required for repression of basal transcription from the multidrug resistance 1 promoter by the t(8;21) fusion protein. Mol. Cell. Biol. *18*, 3604–3611.

Lutterbach, B., Westendorf, J.J., Linggi, B., Patten, A., Moniwa, M., Davie, J.R., Huynh, K.D., Bardwell, V.J., Lavinsky, R.M., Rosenfeld, M.G., et al. (1998b). ETO, a target of t(8;21) in acute leukemia, interacts with the N-CoR and mSin3 corepressors. Mol. Cell. Biol. *18*, 7176–7184.

Minucci, S., Maccarana, M., Cioce, M., De Luca, P., Gelmetti, V., Segalla, S., Di Croce, L., Giavara, S., Matteucci, C., Gobbi, A., et al. (2000). Oligomerization of RAR and AML1 transcription factors as a novel mechanism of oncogenic activation. Mol. Cell *5*, 811–820.

Miyoshi, H., Kozu, T., Shimizu, K., Enomoto, K., Maseki, N., Kaneko, Y., Kamada, N., and Ohki, M. (1993). The t(8;21) translocation in acute myeloid leukemia results in production of an AML1–MTG8 fusion transcript. EMBO J. *12*, 2715–2721.

Mulloy, J.C., Cammenga, J., MacKenzie, K.L., Berguido, F.J., Moore, M.A., and Nimer, S.D. (2002). The AML1-ETO fusion protein promotes the expansion of human hematopoietic stem cells. Blood *99*, 15–23.

Narita, M., Narita, M., Krizhanovsky, V., Nunez, S., Chicas, A., Hearn, S.A., Myers, M.P., and Lowe, S.W. (2006). A novel role for high-mobility group a proteins in cellular senescence and heterochromatin formation. Cell *126*, 503–514.

Nerlov, C. (2004). C/EBPalpha mutations in acute myeloid leukaemias. Nat. Rev. Cancer *4*, 394–400.

Ottiger, M., Delaglio, F., and Bax, A. (1998). Measurement of J and dipolar couplings from simplified two-dimensional NMR spectra. J. Magn. Reson. *131*, 373–378.

Pascual, J., Martinez-Yamout, M., Dyson, H.J., and Wright, P.E. (2000). Structure of the PHD zinc finger from human Williams-Beuren syndrome transcription factor. J. Mol. Biol. *304*, 723–729.

Pear, W.S., Miller, J.P., Xu, L., Pui, J.C., Soffer, B., Quackenbush, R.C., Pendergast, A.M., Bronson, R., Aster, J.C., Scott, M.L., and Baltimore, D. (1998). Efficient and rapid induction of a chronic myelogenous leukemia-like myeloproliferative disease in mice receiving P210 bcr/abl-transduced bone marrow. Blood *92*, 3780–3792.

Pelton, J.G., Torchia, D.A., Meadow, N.D., and Roseman, S. (1993). Tautomeric states of the active-site histidines of phosphorylated and unphosphorylated IIIGlc, a signal-transducing protein from Escherichia coli, using two-dimensional heteronuclear NMR techniques. Protein Sci. *2*, 543–558.

Reeves, R. (2001). Molecular biology of HMGA proteins: Hubs of nuclear function. Gene *277*, 63–81.

Scarr, R.B., and Sharp, P.A. (2002). PDCD2 is a negative regulator of HCF-1 (C1). Oncogene *21*, 5245–5254.

Schessl, C., Rawat, V.P., Cusan, M., Deshpande, A., Kohl, T.M., Rosten, P.M., Spiekermann, K., Humphries, R.K., Schnittger, S., Kern, W., et al. (2005). The AML1-ETO fusion gene and the FLT3 length mutation collaborate in inducing acute leukemia in mice. J. Clin. Invest. *115*, 2159–2168.

Schuck, P. (2000). Size-distribution analysis of macromolecules by sedimentation velocity ultracentrifugation and lamm equation modeling. Biophys. J. *78*, 1606–1619.

Schwieger, M., Lohler, J., Friel, J., Scheller, M., Horak, I., and Stocking, C. (2002). AML1-ETO inhibits maturation of multiple lymphohematopoietic lineages and induces myeloblast transformation in synergy with ICSBP deficiency. J. Exp. Med. *196*, 1227–1240.

Shimada, H., Ichikawa, H., Nakamura, S., Katsu, R., Iwasa, M., Kitabayashi, I., and Ohki, M. (2000). Analysis of genes under the downstream control of the t(8;21) fusion protein AML1–MTG8: Overexpression of the TIS11b (ERF-1, cMG1) gene induces myeloid cell proliferation in response to G-CSF. Blood *96*, 655–663.

Sims, R.J., III, Weihe, E.K., Zhu, L., O'Malley, S., Harriss, J.V., and Gottlieb, P.D. (2002). m-Bop, a repressor protein essential for cardiogenesis, interacts with skNAC, a heart- and muscle-specific transcription factor. J. Biol. Chem. *277*, 26524–26529.

Spadaccini, R., Perrin, H., Bottomley, M.J., Ansieau, S., and Sattler, M. (2006). Structure and functional analysis of the MYND domain. J. Mol. Biol. *358*, 498–508.

Tahirov, T.H., Inoue-Bungo, T., Morii, H., Fujikawa, A., Sasaki, M., Kimura, K., Shiina, M., Sato, K., Kumasaka, T., Yamamoto, M., et al. (2001). Structural analyses of DNA recognition by the AML1/Runx-1 Runt domain and its allosteric control by CBFbeta. Cell *104*, 755–767.

Tusher, V.G., Tibshirani, R., and Chu, G. (2001). Significance analysis of microarrays applied to the ionizing radiation response. Proc. Natl. Acad. Sci. USA *98*, 5116–5121.

Velyvis, A., Yang, Y., Wu, C., and Qin, J. (2001). Solution structure of the focal adhesion adaptor PINCH LIM1 domain and characterization of its interaction with the integrin-linked kinase ankyrin repeat domain. J. Biol. Chem. *276*, 4932–4939.

Wang, J., Hoshino, T., Redner, R.L., Kajigaya, S., and Liu, J.M. (1998). ETO, fusion partner in t(8;21) acute myeloid leukemia, represses transcription by interaction with the human N-CoR/mSin3/HDAC1 complex. Proc. Natl. Acad. Sci. USA *95*, 10860–10865.

Yan, M., Burel, S.A., Peterson, L.F., Kanbe, E., Iwasaki, H., Boyapati, A., Hines, R., Akashi, K., and Zhang, D.E. (2004). Deletion of an AML1-ETO C-terminal NcoR/SMRT-interacting region strongly induces leukemia development. Proc. Natl. Acad. Sci. USA *101*, 17186–17191.

Yan, M., Kanbe, E., Peterson, L.F., Boyapati, A., Miao, Y., Wang, Y., Chen, I.M., Chen, Z., Rowley, J.D., Willman, C.L., and Zhang, D.E. (2006). A previously unidentified alternatively spliced isoform of t(8;21) transcript promotes leukemogenesis. Nat. Med. *12*, 945–949.

Yang, D.W., Venters, R.A., Mueller, G.A., Choy, W.Y., and Kay, L.E. (1999). TROSY-based HNCO pulse sequences for the measurement of (HN)-H-1-N-15, N-15-(CO)-C-13, (HN)-H-1-(CO)-C-13, (CO)-C-13-C-13(alpha) and (HN)-H-1-C-13(alpha) dipolar couplings in N-15, C-13, H-2-labeled proteins. J. Biomol. NMR *14*, 333–343.

Zhang, J., Hug, B.A., Huang, E.Y., Chen, C.W., Gelmetti, V., Maccarana, M., Minucci, S., Pelicci, P.G., and Lazar, M.A. (2001). Oligomerization of ETO is obligatory for corepressor interaction. Mol. Cell. Biol. *21*, 156–163.

Zhang, J., Kalkum, M., Yamamura, S., Chait, B.T., and Roeder, R.G. (2004). E protein silencing by the leukemogenic AML1-ETO fusion protein. Science *305*, 1286–1289.

Zheng, N., Schulman, B.A., Song, L., Miller, J.J., Jeffrey, P.D., Wang, P., Chu, C., Koepp, D.M., Elledge, S.J., Pagano, M., et al. (2002). Structure of the Cul1-Rbx1-Skp1-F boxSkp2 SCF ubiquitin ligase complex. Nature *416*, 703–709.

Accession Numbers

The coordinates for the MYND and SMRT-MYND structures have been deposited in the PDB with accession codes 2OD1and 2ODD, respectively. The microarray data discussed in this publication have been deposited in NCBIs Gene Expression Omnibus (GEO, http://www.ncbi.nlm.nih.gov/geo/) with accession number GSE7324.