



Published in final edited form as:

J Mol Biol. 2010 July 9; 400(2): 145–154. doi:10.1016/j.jmb.2010.04.067.

Molecular mechanism of MLL PHD3 and RNA recognition by the Cyp33 RRM domain

Robert A. Hom¹, Pei-Yun Chang², Siddhartha Roy^{1,5}, Catherine A. Musselman^{1,5}, Karen C. Glass^{1,5}, Anna I. Selezneva³, Or Gozani⁴, Rustem F. Ismagilov³, Michael L. Cleary², and Tatiana G. Kutateladze¹

¹Department of Pharmacology, University of Colorado Denver School of Medicine, Aurora, Colorado 80045, USA

²Department of Pathology, University of Chicago, Chicago, Illinois 60637, USA

³Department of Chemistry, University of Chicago, Chicago, Illinois 60637, USA

⁴Department of Biological Sciences, Stanford University, Stanford, California 94305, USA

Summary

The nuclear protein Cyclophilin 33 (Cyp33) is a peptidyl-prolyl *cis-trans* isomerase that catalyzes *cis-trans* isomerization of the peptide bond preceding a proline and promotes folding and conformational changes in folded and unfolded proteins. The N-terminal RRM domain of Cyp33 has been found to associate with the third plant homeodomain (PHD3) finger of the Mixed Lineage Leukemia (MLL) proto-oncoprotein and a poly-A RNA sequence. Here, we report a 1.9 Å resolution crystal structure of the RRM domain of Cyp33 and describe the molecular mechanism of PHD3 and RNA recognition. The Cyp33 RRM domain folds into a five-stranded antiparallel β -sheet and two α -helices. The RRM domain but not the catalytic module of Cyp33 binds strongly to PHD3, exhibiting a 2 μ M affinity as measured by Isothermal Titration Calorimetry (ITC). NMR chemical shift perturbation (CSP) analysis and dynamics data reveal that the β strands and the β 2– β 3 loop of the RRM domain are involved in the interaction with PHD3. Mutations in the PHD3-binding site or deletions in the β 2/ β 3 loop lead to a significantly reduced affinity or abrogation of the interaction. The RNA-binding pocket of the Cyp33 RRM domain, mapped based on NMR CSP and mutagenesis, partially overlaps with the PHD3-binding site, and RNA association is abolished in the presence of MLL PHD3. Full-length Cyp33 acts as a negative regulator of MLL-induced transcription and reduces the expression levels of MLL target genes *MEIS1* and *HOXA9*. Together, these *in vitro* and *in vivo* data provide insight into the multiple functions of Cyp33 RRM and suggest a Cyp33-dependent mechanism for regulating the transcriptional activity of MLL.

Keywords

RRM domain; PHD finger; Cyp33; MLL; mechanism; structure

Address correspondence to: Tatiana G. Kutateladze, Department of Pharmacology, University of Colorado Denver, 12801 East 17th Avenue, Aurora, Colorado 80045-0511, Tel. 303 724-3593; Fax. 303 724-3663; Tatiana.Kutateladze@UCDenver.edu.

⁵Equal contribution

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

Introduction

The nuclear Cyclophilin 33 (Cyp33) was initially identified in human Jurkat T cells in 1996 1. It belongs to a highly conserved family of PPIases that catalyze *cis-trans* isomerization of the peptide bond preceding a proline, accelerating folding and stimulating conformational changes in folded and unfolded proteins 2. The PPIase activity is also essential for intracellular protein transport and transient protein interactions involving CyPs and can be effectively inhibited by cyclosporine A 1³. The catalytic CYP/PPIase domain of Cyp33 is located in the carboxy-terminal region of the protein and is preceded by an amino-terminal RNA-recognition motif (RRM), also known as RNA-binding domain (RBD) or ribonucleoprotein (RNP). Although it remains unclear whether the Cyp33 RRM domain directly interacts with RNA, the full length Cyp33 protein has been shown to preferentially associate with mRNAs containing an AAUAAA sequence or a poly-A tail, and this association appears to activate the PPIase activity of Cyp33 1⁴. Recent reports have also implicated the RRM domain in binding to the third plant homeodomain (PHD3) finger of MLL (myeloid/lymphoid or mixed lineage leukemia) 5⁶. This interaction has been proposed to switch the MLL function from transactivation to repression 6, however how the RRM domain recognizes the MLL PHD3 finger remains unclear.

MLL is a member of the trithorax protein family that regulates gene expression, particularly *HOX* genes, during embryonic development. MLL is translocated or mutated in a variety of aggressive human blood cancers including acute lymphoblastic and acute myelogenous leukemias 7⁸. This large, ~4,000-residue protein contains numerous functional modules including three amino-terminal DNA-binding AT hook domains, specific for AT-rich regions of the DNA minor groove, two speckled nuclear localization signals, and a transcriptional repression region containing a CXXC zinc-finger homologous to the CpG-binding domain of DNA methyltransferase 1 (DNMT1) 9¹⁰ 11. Three sequential PHD modules, an acetyl-lysine binding bromodomain and another atypical PHD finger precede a transactivation (TA) domain. The TA region has a docking site for CREB-binding protein (CBP), a histone acetyltransferase (HAT) able to acetylate histone H3 and H4 at the *HOX* area 12. The carboxy-terminal Su(var)3-9, Enhancer of Zeste, Trithorax (SET) domain shows histone methyltransferase (HMTase) activity with a high specificity for lysine 4 of H3 13¹⁴ 15. Named after the first yeast H3K4 HMTase Set1, it is highly conserved throughout the SET domain-containing proteins and is capable of producing mono-, di- and tri-methylated H3K4 marks 16. A short sequence preceding the SET domain is recognized by WDR5 17¹⁸ 19. As many other HMTases, MLL is a component of a larger nuclear complex that also contains WDR5, RbPB5 and ASH2, all of which are required for the functional assembly, chromatin targeting and enzymatic activity of the MLL complex (also referred to as the human COMPASS) 20.

The three sequential PHD fingers in MLL, which are deleted in oncogenic translocation chimeras, comprise one of the most conserved regions of MLL. Although the biological role of this region remains elusive, it has recently been shown that the second PHD finger (PHD2) is involved in dimerization, whereas PHD3 binds the Cyp33 RRM domain 5⁶ 21. The PHD finger region appears to play a regulatory role in MLL function and suppresses MLL-mediated leukaemogenesis. Inclusion of the PHD2-PHD3 fingers in the chimeric MLL-AF9 protein inhibits transformation of mouse bone marrow and leads to hematopoietic cell differentiation 22. Insertion of PHD3 into the MLL-ENL chimera suppresses MLL-ENL-induced immortalization of murine bone marrow progenitor cells 6. Clearly PHD3 is essential for the proper function of MLL, yet the molecular basis underlying its biological activities including the association with Cyp33 RRM is unknown.

In this study, we characterize binding of the RRM domain of Cyp33 to the PHD3 finger of MLL and demonstrate that this interaction abolishes the association of RRM with RNA. The crystal structure of the Cyp33 RRM domain determined at 1.9 Å resolution, a combination of NMR binding and dynamics data, mutagenesis, isothermal titration calorimetry (ITC) measurements and *in vivo* quantitative RT-PCR assays were used to elucidate the molecular mechanism of the Cyp33-MLL association. Our findings suggest a negative regulatory role of this interaction in transcription of MLL target genes.

Results and Discussion

Overall structure of the Cyp33 RRM domain

The structure of the RRM domain of human Cyp33 (residues 1–83) was determined at 1.9 Å resolution by X-ray crystallography (Fig. 1). The RRM domain folds into a five-stranded antiparallel β -sheet and two α -helices. The $\alpha 1$ helix is positioned between the $\beta 1$ and $\beta 2$ strands whereas helix $\alpha 2$ connects strands $\beta 3$ and $\beta 4$. The $\beta 1$ strand (residues V7-G11), $\alpha 1$ helix (residues D19-F26), $\beta 2$ strand (residues T33-Q36), $\beta 3$ strand (residues F49-F54), $\alpha 2$ helix (residues A57-M67) and $\beta 5$ strand (residues R75-L81) adopt a canonical $\beta \alpha \beta \alpha \beta$ RRM fold with an additional strand ($\beta 4$, residues E69-L72) pairing with the $\beta 5$ strand and forming the edge of the β sheet. Notably, residues located in the long $\beta 2$ – $\beta 3$ loop and $\beta 2$ and $\beta 3$ strands, that is, D34, I35, I37, L39, D40, E42, T43, E44, H46, R47, F49, V52, and F54 exhibited the largest chemical shift changes upon binding to PHD3 and residues of the $\beta 1$ and $\beta 5$ strands were perturbed to a lesser degree (Fig. 3a, b, and d). These data suggest that the $\beta 2$ – $\beta 3$ loop and the β strands of the Cyp33 RRM domain are directly or indirectly involved in the interaction with the PHD3 finger of MLL.

The Cyp33 RRM domain binds strongly to the PHD3 finger of MLL

To determine the role of the most perturbed residues of the Cyp33 RRM domain, we replaced them with alanine and, additionally, generated mutants with partially truncated $\beta 2$ – $\beta 3$ loop. Binding of the mutant proteins to PHD3 was examined by ITC and NMR. Deletion of the $\beta 2$ – $\beta 3$ loop in RRM $\Delta 40$ – 45 completely abolished the interaction, whereas mutation of F49 and L39 reduced the binding affinity of the RRM domain for the PHD3 finger by 35- and 10-fold, respectively, suggesting a significant contribution of hydrophobic contacts (Fig. 3c). The E42A and R47A mutants bound only slightly weaker than the wild-type protein. Together, these results demonstrate that the residues located in the $\beta 2$ – $\beta 3$ loop and the β strands constitute the binding interface of the Cyp33 RRM domain (Fig. 3b and d).

To determine the strength of the RRM-PHD3 interaction, the dissociation constant and thermodynamic parameters were measured by ITC (Fig. 2d). The K_d value was found to be $1.9 \pm 0.2 \mu\text{M}$, and is in agreement with the pattern of resonance perturbations in the NMR titration experiments. The interaction was enthalpy driven ($\Delta H = -10 \pm 0.5 \text{ kcal}$), and a negative change in entropy ($\Delta S = -9 \pm 2 \text{ cal mol}^{-1} \text{ K}^{-1}$) indicated that the proteins lost some conformational freedom upon formation of the complex.

Identification of the PHD3-binding site of the Cyp33 RRM domain

To identify the active-site residues of the RRM domain, we assigned the ^1H , ^{13}C and ^{15}N resonances of the ligand-free and PHD3-bound protein using a set of triple-resonance NMR experiments. The spectra were collected on $^{15}\text{N}/^{13}\text{C}$ -labeled RRM in the apo state and in complex with unlabeled PHD3. Shown in Fig. 3a is a histogram plot of the differences in chemical shifts for backbone amides of the RRM domain in the unbound and PHD3-bound states. Notably, residues located in the long $\beta 2$ – $\beta 3$ loop and $\beta 2$ and $\beta 3$ strands, that is, D34, I35, I37, L39, D40, E42, T43, E44, H46, R47, F49, V52, and F54 exhibited the largest chemical shift changes upon binding to PHD3 and residues of the $\beta 1$ and $\beta 5$ strands were

perturbed to a lesser degree (Fig. 3a, b, and d). These data suggest that the $\beta 2$ - $\beta 3$ loop and the β strands of the Cyp33 RRM domain are directly or indirectly involved in the interaction with the PHD3 finger of MLL.

To determine the role of the most perturbed residues of the Cyp33RRM domain, we replaced them with alanine and, additionally, generated mutants with partially truncated $\beta 2$ - $\beta 3$ loop. Binding of the mutant proteins to PHD3 was examined by ITC and NMR. Deletion of the $\beta 2$ - $\beta 3$ loop in RRM $\Delta 40$ -45 completely abolished the interaction, whereas mutation of F49 and L39 reduced the binding affinity of the RRM domain for the PHD3 finger by 35- and 10-fold, respectively, suggesting a significant contribution of hydrophobic contacts (Fig. 3c). The E42A and R47A mutants bound only slightly weaker than the wild-type protein. Together, these results demonstrate that the residues located in the $\beta 2$ - $\beta 3$ loop and the β strands constitute the binding interface of the Cyp33 RRM domain (Fig. 3b and d).

Binding of PHD3 induces changes in the dynamics of the $\beta 2/\beta 3$ loop, the $\beta 1$, $\beta 2$ and $\beta 3$ strands and the $\alpha 2$ helix of the Cyp33 RRM domain

The effect of the association with PHD3 on the dynamics of the RRM domain was investigated using ^{15}N NMR relaxation experiments carried out in the absence and presence of the MLL PHD3 finger (Fig. 4). Comparison of the average relaxation rates in the ligand-free and PHD3-bound RRM domain suggested a global stabilization of the RRM structure in the complex. Binding to PHD3 caused a decrease of R_1 values and an increase in R_2 values for the majority of the RRM residues. The most evident changes in the R_1 and R_2 relaxation rates were observed for residues that are located in the $\beta 2/\beta 3$ loop, the $\beta 1$, $\beta 2$ and $\beta 3$ strands and the $\alpha 2$ helix, particularly those that are involved in the interaction with the PHD3 finger. The lack of changes in the heteronuclear NOE values indicated that the local internal backbone mobility on a sub-nanosecond timescale does not change due to the complex formation. The significantly larger R_2 relaxation rates observed for some residues in the $\beta 1$, $\beta 2$ and $\beta 3$ strands pointed to a contribution from μs - ms dynamics in these regions of the complex. In agreement with the negative change in entropy, the dynamics data suggest that overall the RRM domain of Cyp33 becomes more rigid upon binding to the PHD3 finger.

The human Cyp33 RRM domain is specific for MLL and does not interact with other PHD fingers

To determine whether Cyp33 RRM is able to recognize any PHD-finger fold, ^1H , ^{15}N HSQC spectra of the ^{15}N -labeled RRM domain were recorded as the unlabeled PHD module of tumor suppressor ING1 was titrated into the NMR sample (Supplementary Fig. 2). The amino acid sequence of the ING1 PHD finger contains a number of conserved (within the PHD family, including MLL PHD3) residues 24; 25. Addition of a five-fold excess of ING1 PHD did not induce chemical shift changes in the NMR spectrum of RRM, implying that there is no interaction between these two proteins. Thus, binding of the Cyp33 RRM domain to the MLL PHD3 finger is specific.

We next tested the ability of other RRM domains to recognize the MLL PHD3 finger. Addition of a 5-fold excess of unlabeled *Drosophila* Cyp33 RRM to the ^{15}N -labeled MLL PHD3 finger or conversely, addition of a 5-fold excess of unlabeled MLL PHD3 to the ^{15}N -labeled *Drosophila* Cyp33 RRM caused negligible changes in ^1H , ^{15}N HSQC spectra of the proteins (Supplementary Fig. 3 and data not shown), revealing high specificity of human Cyp33 RRM toward human MLL PHD3.

The Cyp33 RRM domain binds RNA

It has recently been shown that the full length Cyp33 protein associates with polyribonucleotide poly-A and poly-U but not with poly-G or poly-C, particularly preferring mRNA that contains the AAUAAA sequence 1' 4. To test whether the RRM domain of Cyp33 is responsible for this association, the AAUAAA RNA was synthesized and used in ^1H , ^{15}N HSQC titration experiments (Fig. 5). Large chemical shift changes in the NMR spectra of ^{15}N -labeled RRM, caused by the gradual addition of RNA, indicated that the Cyp33 RRM domain directly binds the RNA sequence. The residues of RRM located in the β 1 strand (Y9, V10, G11), β 3 strand (F49, A50, F51, V52), β 5 strand (N80, L81, A82) and in the C-terminus (M85-K88) were perturbed most significantly (Fig. 5b). These residues form an extended RNA-binding site that spreads across the β -sheet surface (Fig. 5c, d). The β -sheet is commonly used by other RRM modules in the interaction with single-stranded RNA 26' 27. In fact, the perturbed residues of the Cyp33 RRM domain comprise the two conserved ribonucleoprotein (RNP) motifs RNP2 (YVGGL-13) and RNP1 (RGFADFVEF-54) (Fig. 5g), that are required for RNA recognition by a typical RRM domain 26' 27. The aromatic side chains of Y9 in RNP2 and of F49 and F51 in RNP1 that protrude orthogonally to the protein surface are ideally positioned to form stacking interactions with RNA bases (Fig. 5d).

The binding affinity of the Cyp33 RRM domain for AAUAAA ($K_d = 198 \pm 11 \mu\text{M}$) was obtained by plotting normalized chemical shift changes in the amide groups of the protein *versus* the RNA concentration (Fig. 5e). Although in general RRM domains exhibit a broad range of affinities for RNAs down to the low μM range 26' 28' 29, we point out that future studies are necessary to establish the significance of the RNA association by the Cyp33 RRM domain.

Interaction between Cyp33 RRM and MLL PHD3 disrupts the association of RRM with RNA

The binding sites of the Cyp33 RRM domain for PHD3 and RNA partially overlap (Fig. 6a), suggesting competitive binding. Indeed, when the MLL PHD3 finger was added to the Cyp33 RRM-RNA complex, the NMR resonances of RNA-bound RRM disappeared and resonances of the PHD3-bound protein appeared (Fig. 6b). The resulting ^1H , ^{15}N HSQC spectrum was almost identical to that of the RRM domain obtained upon addition of PHD3 alone, implying that the presence of RNA does not alter the PHD3-binding mode of the Cyp33 RRM domain. Because the binding affinity of the RRM domain for the MLL PHD3 finger is ~ 100 fold higher than for the AAUAAA sequence, the PHD3 finger readily displaces the RNA.

Cyp33 decreases expression levels of MLL target genes

MLL regulates expression of *HOX* and *MEIS1* genes. We therefore examined the effect of Cyp33 overexpression on endogenous MLL-mediated gene transcription. 293T cells were transfected with different doses of FLAG-HIS6-CYP33 (3–12 μg), and after two days RNA was prepared and reverse-transcribed. *HOXA9*, *MEIS1* and *MLL* expression levels were analyzed by quantitative RT-PCR. As shown in Figure 7a, Cyp33 transfection led to a decrease in the *HOXA9* and *MEIS1* expression levels, and this decrease was consistently greater with increasing doses of transfected Cyp33. Thus, these data suggest that Cyp33 negatively regulates the transcriptional function of MLL.

In conclusion, our results reveal a pivotal role of the Cyp33 RRM-MLL PHD3 interaction in the function of MLL and Cyp33. The strong binding of the Cyp33 RRM domain to the MLL PHD3 finger disrupts association of RRM with RNA and *in vivo* data indicate that Cyp33 acts as a negative regulator of transcriptional activity of MLL, reducing the expression levels of MLL target genes (Fig. 7b). Although the mechanistic details of the negative

regulation by Cyp33 remain to be determined, recent studies suggest several possible mechanisms. As we report in the accompanying paper, the MLL PHD3 finger also recognizes histone H3 trimethylated at Lys4 (H3K4me3), and this interaction is essential for MLL-dependent gene transcription. Binding of the Cyp33 RRM domain to MLL PHD3 could reduce the association of PHD3 with H3K4me3 and lead to the decrease of the target gene expression. Alternatively, the RRM-PHD3 interaction may bridge the catalytic PPIase domain of Cyp33 to MLL for the subsequent action on nearby regions of MLL or MLL-effectors, such as HDAC1, binding of which to the MLL repression region is known to be enhanced by Cyp33 5' 21. Other modules of MLL surrounding PHD3, including the adjacent PHD1 and PHD2 fingers and a bromodomain could further influence the binding activity of PHD3 and fine-tune transcriptional function of MLL. In summary, the molecular and structural details of the PHD3 and RNA recognition by the Cyp33 RRM domain described in this study provide new insights into the regulation of a cancer-critical protein by a cyclophilin.

Materials and Methods

Protein expression and purification

The pET-28 LIC vector containing DNA encoding residues 1–90 of human Cyp33 was modified to express His-tagged RRM. A shorter construct of Cyp33 RRM (residues 1–83) was generated by PCR. The unlabeled, ¹⁵N-labeled and ¹⁵N/¹³C-labeled proteins were expressed in *Escherichia coli* Rosetta 2 (DE3) in LB or minimal media supplemented with ¹⁵NH₄Cl or ¹⁵NH₄Cl/¹³C₆-glucose (Isotec). The bacterial cells were grown at 37 °C to an OD₆₀₀ of 0.8 and protein expression was induced with 1.0 mM isopropyl-β-D-thiogalactopyranoside (IPTG) at 37 °C for 5 hours. The cells were collected by centrifugation at 5,000 g, resuspended in lysis buffer (20 mM Tris-HCl, pH 7.0, 150 mM NaCl, 0.05% NP-40 and Protease Inhibitor Cocktail Tablets (Roche)) and lysed by sonication. The proteins were purified on a TALON affinity resin using a wash buffer (20 mM Tris-HCl, pH 8.0, 150 mM NaCl and 2 mM β-mercaptoethanol (ME)) and eluted with 20 mM Tris-HCl pH 8.0 buffer containing 150 mM NaCl, 2 mM ME and 150 mM imidazole. The His tag was cleaved with thrombin. The cleaved proteins were concentrated in Millipore concentrators (Millipore) and further purified by FPLC on a superdex 75 HR16/60 column in 20 mM Tris-HCl pH 6.8 buffer containing 150 mM NaCl and 2 mM dithiothreitol (DTT). The same protocol was used for expression and purification of the mutant proteins.

The PHD3 finger of MLL (residues 1565–1627) was subcloned into a pGEX-2T vector (Amersham). The unlabeled and ¹⁵N-labeled PHD3 finger was expressed in *Escherichia coli* BL21(DE3) pLysS cells and purified as described in the accompanying manuscript. The unlabeled catalytic domain of human Cyp33 and unlabeled and ¹⁵N-labeled *Drosophila* Cyp33 RRM were expressed and purified using the same procedure described above for the human Cyp33 RRM domain. The unlabeled human ING1 PHD finger was purified as in 25.

X-ray crystallography

Crystallization of the human Cyp33 RRM domain (residues 1–83) was performed using a microcapillary technique 30. The crystals were obtained at 18 °C in 0.1 M HEPES and 1.0 M tri-sodium dehydrate citrate at pH 7.6. All crystals grew in a monoclinic space group (C2) with unit cell parameters of a = 84.84 Å, b = 40.52 Å, c = 65.66 Å, α = γ = 90°, β = 127.06° with two molecules per A.U. Crystals were flash cooled in liquid nitrogen, and X-ray data were collected at 100 K on a “NOIR-1” MBC system detector at beam line 4.2.2 at the Advanced Light Source in Berkeley. A native data set was collected to a resolution of 1.85 Å. Data was processed with D*TREK 31. The molecular replacement solution was

generated using the program BALBES 32 and the structure of RRM (PDB 1CVJ) as a search model. The protein structure was further refined with CNS 33 and COOT 34 and verified with PROCHECK 35. Statistics are shown in Supplementary Table 1.

PCR Mutagenesis

Mutants of the RRM domain (L39A, E42A, R47A, G48A, F49A, F54A, RRM Δ 41–42, RRM Δ 41–43, RRM Δ 40–43 and RRM Δ 40–45) were generated using a QuickChange Site-Directed Mutagenesis Kit (Stratagene).

NMR spectroscopy and sequence specific resonance assignments

Multidimensional heteronuclear NMR spectra were recorded at 298 K on Varian INOVA 800 and 600 MHz spectrometers using pulse field gradients to suppress artifacts and eliminate water signal. Because of the slow exchange regime, all spectra were collected on 1–2 mM uniformly ^{15}N - and $^{15}\text{N}/^{13}\text{C}$ -labeled Cyp33 RRM domain (residues 1–90) first in the free form and then in complex with unlabeled MLL PHD3 finger (at a 1:2 ratio of proteins). The amino acid spin system and sequential assignments were made using $^1\text{H},^{15}\text{N}$ heteronuclear single quantum coherence (HSQC) and triple-resonance HNCACB 36 and CBCA(CO)NH 37. Spectra were processed with NMRPipe 38 and analyzed using CCPN 39, nmrDraw and in-house software programs on Sun and Silicon Graphics workstations.

NMR titrations

The ligand binding to the wild type and mutant human Cyp33 RRM domain was characterized by monitoring chemical shift changes in $^1\text{H},^{15}\text{N}$ HSQC spectra of 0.1–0.2 mM ^{15}N -labeled RRM as either unlabeled MLL PHD3 (up to 0.4 mM), the AAUAAA RNA sequence (up to 0.6 mM), or unlabeled ING1 PHD (up to 1 mM) were added stepwise. The ligand binding to the wild type MLL PHD3 finger was characterized by monitoring chemical shift changes in $^1\text{H},^{15}\text{N}$ HSQC spectra of 0.2 mM ^{15}N -labeled PHD3 as unlabeled wild type or mutant human Cyp33 RRM (up to 0.4 mM), or the catalytic domain of human Cyp33 (up to 1 mM) were added gradually. Interaction between ^{15}N -labeled MLL PHD3 (0.1 mM) and unlabeled *Drosophila* Cyp33 RRM (up to 0.5 mM), and between ^{15}N -labeled *Drosophila* Cyp33 RRM (0.1 mM) and unlabeled MLL PHD3 (up to 0.5 mM) was tested similarly.

Relaxation experiments

Changes in dynamics of the Cyp33 RRM domain upon binding to the MLL PHD3 finger was investigated by backbone amide ^{15}N relaxation experiments. The ^{15}N R_1 , R_2 and $^1\text{H},^{15}\text{N}$ steady state NOE experiments were acquired on an 800 MHz spectrometer at 298 K using the ligand-free and PHD3 (1.5 mM)-bound ^{15}N -labeled RRM 1–90 (1 mM) and analyzed as described previously 40. The ^{15}N R_1 and ^{15}N R_2 values for the unbound state were determined from the spectra collected with variable T1 delay times (20, 60, 140, 240, 360, 460, 660, 860, and 1110 ms, with times 60 and 860 ms repeated for curve-fitting error) and T2 delay times (10, 30, 50, 70, 90, and 110 ms, with times 30 and 70 ms repeated for curve-fitting error), respectively. The ^{15}N R_1 and ^{15}N R_2 values for the RRM-PHD3 complex were determined from the spectra collected with variable T1 delay times (20, 60, 140, 240, 360, 460, 660, 860, 1100, and 1500, with times 60 and 1100 ms repeated for curve-fitting error) and T2 delay times (10, 30, 50, 70, and 90 ms, with times 10 and 70 ms repeated for curve-fitting error), respectively. Recovery delays of 1.2 s were used in the measurement of both R_1 and R_2 values. NOE values were determined from spectra collected either with a 5 s relaxation delay alone or with a proton presaturation period of 3 s preceded by a 2 s relaxation delay. The R_1 , R_2 and NOE values were analyzed using the program Origin.

Isothermal Titration Calorimetry

The ITC experiments were carried out at 25 C on a VP-ITC calorimeter (MicroCal). The samples (wild type and mutant Cyp33 RRM and MLL PHD3) were dialyzed for 2 days against an assay buffer (20 mM Tris-HCl, 150 mM NaCl, 2 mM DTT and 1 mM NaN₃). The heat of the reactions was measured by making 30 sequential injections of 10 µl of PHD3 (0.65 mM) into a 1.41 ml of RRM solution (0.05 mM) (and vice versa) with spacing intervals of 60 seconds. The heat of dilution was measured by injecting the ligand protein into control buffer and subtracted from the raw data before the fitting process. Binding isotherms were analyzed by non-linear least-squares fitting of the data using Microcal ORIGIN software (Microcal).

RT-PCR assays

293T cells were transfected with FLAG-HIS6-CYP33 using the Fugene6 method. After two days RNA was prepared using Trizol (Invitrogen) and reverse-transcribed using First Strand Synthesis kit (Invitrogen). Quantitative RT-PCR was performed in duplicate on the ABI PRISM 7900 Sequence Detection System. *HOXA9* and *MEIS1* expression was calculated following normalization to *GAPDH* levels by the comparative Ct (Cycle threshold) method. Taqman probes and primers used in the study are available upon request.

Accession numbers

Coordinates and structure factors have been deposited to the Protein Data Bank with accession number 3MDF.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We thank Liang Li, Sigrid Nachtergaele and Jennifer Schlegel for discussions and help with the experiments, Jay Nix at beam line 4.2.2 of the ALS in Berkeley for help with data collection and Tara Davis for providing the initial constructs of the RRM and catalytic domains of Cyp33. This research is supported by grants from the National Institutes of Health, GM074961 and GM075827 (R.F.I.), CA55029 and CA116606 (M.L.C.) and CA113472 and GM071424 (T.G.K.).

Abbreviations

Cyp33	Cyclophilin 33
PHD	plant homeodomain
MLL	Myeloid/lymphoid or Mixed Lineage Leukemia
ITC	Isothermal Titration Calorimetry
CSP	chemical shift perturbation
RRM	RNA-recognition motif
HSQC	Heteronuclear Single Quantum Coherence

References

1. Mi H, Kops O, Zimmermann E, Jaschke A, Tropschug M. A nuclear RNA-binding cyclophilin in human T cells. *FEBS Lett.* 1996; 398:201–5. [PubMed: 8977107]

2. Wang XJ, Etkorn FA. Peptidyl-prolyl isomerase inhibitors. *Biopolymers*. 2006; 84:125–46. [PubMed: 16302169]
3. Min L, Fulton DB, Andreotti AH. A case study of proline isomerization in cell signaling. *Front Biosci*. 2005; 10:385–97. [PubMed: 15574377]
4. Wang Y, Han R, Zhang W, Yuan Y, Zhang X, Long Y, Mi H. Human Cyp33 binds specifically to mRNA and binding stimulates PPIase activity of hCyp33. *FEBS Lett*. 2008; 582:835–9. [PubMed: 18258190]
5. Fair K, Anderson M, Bulanova E, Mi H, Tropschug M, Diaz MO. Protein interactions of the MLL PHD fingers modulate MLL target gene regulation in human cells. *Mol Cell Biol*. 2001; 21:3589–97. [PubMed: 11313484]
6. Chen J, Santillan DA, Koonce M, Wei W, Luo R, Thirman MJ, Zeleznik-Le NJ, Diaz MO. Loss of MLL PHD finger 3 is necessary for MLL-ENL-induced hematopoietic stem cell immortalization. *Cancer Res*. 2008; 68:6199–207. [PubMed: 18676843]
7. Ayton PM, Cleary ML. Molecular mechanisms of leukemogenesis mediated by MLL fusion proteins. *Oncogene*. 2001; 20:5695–707. [PubMed: 11607819]
8. Hess JL. MLL: a histone methyltransferase disrupted in leukemia. *Trends Mol Med*. 2004; 10:500–7. [PubMed: 15464450]
9. Yano T, Nakamura T, Blechman J, Sorio C, Dang CV, Geiger B, Canaani E. Nuclear punctate distribution of ALL-1 is conferred by distinct elements at the N terminus of the protein. *Proc Natl Acad Sci U S A*. 1997; 94:7286–91. [PubMed: 9207083]
10. Birke M, Schreiner S, Garcia-Cuellar MP, Mahr K, Titgemeyer F, Slany RK. The MT domain of the proto-oncoprotein MLL binds to CpG-containing DNA and discriminates against methylation. *Nucleic Acids Res*. 2002; 30:958–65. [PubMed: 11842107]
11. Allen MD, Grummitt CG, Hilcenko C, Min SY, Tonkin LM, Johnson CM, Freund SM, Bycroft M, Warren AJ. Solution structure of the nonmethyl-CpG-binding CXXC domain of the leukaemia-associated MLL histone methyltransferase. *Embo J*. 2006; 25:4503–12. [PubMed: 16990798]
12. Ernst P, Wang J, Huang M, Goodman RH, Korsmeyer SJ. MLL and CREB bind cooperatively to the nuclear coactivator CREB-binding protein. *Mol Cell Biol*. 2001; 21:2249–58. [PubMed: 11259575]
13. Nakamura T, Mori T, Tada S, Krajewski W, Rozovskaia T, Wassell R, Dubois G, Mazo A, Croce CM, Canaani E. ALL-1 is a histone methyltransferase that assembles a supercomplex of proteins involved in transcriptional regulation. *Mol Cell*. 2002; 10:1119–28. [PubMed: 12453419]
14. Milne TA, Briggs SD, Brock HW, Martin ME, Gibbs D, Allis CD, Hess JL. MLL targets SET domain methyltransferase activity to Hox gene promoters. *Mol Cell*. 2002; 10:1107–17. [PubMed: 12453418]
15. Southall SM, Wong PS, Odho Z, Roe SM, Wilson JR. Structural basis for the requirement of additional factors for MLL1 SET domain activity and recognition of epigenetic marks. *Mol Cell*. 2009; 33:181–91. [PubMed: 19187761]
16. Schneider J, Wood A, Lee JS, Schuster R, Dueker J, Maguire C, Swanson SK, Florens L, Washburn MP, Shilatifard A. Molecular regulation of histone H3 trimethylation by COMPASS and the regulation of gene expression. *Mol Cell*. 2005; 19:849–56. [PubMed: 16168379]
17. Patel A, Dharmarajan V, Cosgrove MS. Structure of WDR5 bound to mixed lineage leukemia protein-1 peptide. *J Biol Chem*. 2008; 283:32158–61. [PubMed: 18829459]
18. Patel A, Vought VE, Dharmarajan V, Cosgrove MS. A conserved arginine-containing motif crucial for the assembly and enzymatic activity of the mixed lineage leukemia protein-1 core complex. *J Biol Chem*. 2008; 283:32162–75. [PubMed: 18829457]
19. Song JJ, Kingston RE. WDR5 interacts with mixed lineage leukemia (MLL) protein via the histone H3-binding pocket. *J Biol Chem*. 2008; 283:35258–64. [PubMed: 18840606]
20. Steward MM, Lee JS, O'Donovan A, Wyatt M, Bernstein BE, Shilatifard A. Molecular regulation of H3K4 trimethylation by ASH2L, a shared subunit of MLL complexes. *Nat Struct Mol Biol*. 2006; 13:852–4. [PubMed: 16892064]
21. Xia ZB, Anderson M, Diaz MO, Zeleznik-Le NJ. MLL repression domain interacts with histone deacetylases, the polycomb group proteins HPC2 and BMI-1, and the corepressor C-terminal-binding protein. *Proc Natl Acad Sci U S A*. 2003; 100:8342–7. [PubMed: 12829790]

22. Muntean AG, Giannola D, Udager AM, Hess JL. The PHD fingers of MLL block MLL fusion protein-mediated transformation. *Blood*. 2008; 112:4690–3. [PubMed: 18796627]
23. Vitali J, Ding J, Jiang J, Zhang Y, Krainer AR, Xu RM. Correlated alternative side chain conformations in the RNA-recognition motif of heterogeneous nuclear ribonucleoprotein A1. *Nucleic Acids Res*. 2002; 30:1531–8. [PubMed: 11917013]
24. Peña PV, Davrazou F, Shi X, Walter KL, Verkhusha VV, Gozani O, Zhao R, Kutateladze TG. Molecular mechanism of histone H3K4me3 recognition by plant homeodomain of ING2. *Nature*. 2006; 442:100–3. [PubMed: 16728977]
25. Peña PV, Hom RA, Hung T, Lin H, Kuo AJ, Wong RP, Subach OM, Champagne KS, Zhao R, Verkhusha VV, Li G, Gozani O, Kutateladze TG. Histone H3K4me3 binding is required for the DNA repair and apoptotic activities of ING1 tumor suppressor. *J Mol Biol*. 2008; 380:303–12. [PubMed: 18533182]
26. Clery A, Blatter M, Allain FH. RNA recognition motifs: boring? Not quite. *Curr Opin Struct Biol*. 2008; 18:290–8. [PubMed: 18515081]
27. Hargous Y, Hautbergue GM, Tintaru AM, Skrisovska L, Golovanov AP, Stevenin J, Lian LY, Wilson SA, Allain FH. Molecular basis of RNA recognition and TAP binding by the SR proteins SRp20 and 9G8. *Embo J*. 2006; 25:5126–37. [PubMed: 17036044]
28. Abdul-Manan N, O'Malley SM, Williams KR. Origins of binding specificity of the A1 heterogeneous nuclear ribonucleoprotein. *Biochemistry*. 1996; 35:3545–54. [PubMed: 8639505]
29. Nadler SG, Merrill BM, Roberts WJ, Keating KM, Lisbin MJ, Barnett SF, Wilson SH, Williams KR. Interactions of the A1 heterogeneous nuclear ribonucleoprotein and its proteolytic derivative, UP1, with RNA and DNA: evidence for multiple RNA binding domains and salt-dependent binding mode transitions. *Biochemistry*. 1991; 30:2968–76. [PubMed: 1848781]
30. Li L, Mustafi D, Fu Q, Tereshko V, Chen DL, Tice JD, Ismagilov RF. Nanoliter microfluidic hybrid method for simultaneous screening and optimization validated with crystallization of membrane proteins. *Proc Natl Acad Sci U S A*. 2006; 103:19243–8. [PubMed: 17159147]
31. Pflugrath JW. The finer things in X-ray diffraction data collection. *Acta Crystallogr D Biol Crystallogr*. 1999; 55:1718–25. [PubMed: 10531521]
32. Long F, Vagin AA, Young P, Murshudov GN. BALBES: a molecular-replacement pipeline. *Acta Crystallogr D Biol Crystallogr*. 2008; 64:125–32. [PubMed: 18094476]
33. Brunger AT, et al. Crystallography & NMR system: A new software suite for macromolecular structure determination. *Acta Crystallogr D Biol Crystallogr*. 1998; 54:905–21.
34. Emsley P, Cowtan K. Coot: model-building tools for molecular graphics. *Acta Crystallogr D Biol Crystallogr*. 2004; 60:2126–32. [PubMed: 15572765]
35. Laskowski RA, MacArthur MW, Moss DS, Thornton JM. PROCHECK: a program to check the stereochemical quality of protein structures. *J Appl Crystallogr*. 1993; 26:283–291.
36. Wittekind M, Mueller L. HNCACB, A high-sensitivity 3D NMR experiment to correlate amide-proton and nitrogen resonances with the alpha-carbon and beta-carbon resonances in proteins. *J Magn Reson*. 1993; 101:201–5.
37. Grzesiek S, Bax A. Improved 3D triple-resonance NMR techniques applied to a 31-kDa protein. *J Magn Reson*. 1992; 96:432–40.
38. Delaglio F, Grzesiek S, Vuister GW, Zhu G, Pfeifer J, Bax A. NMRPipe: a multidimensional spectral processing system based on UNIX pipes. *J Biomol NMR*. 1995; 6:277–93. [PubMed: 8520220]
39. Vranken WF, Boucher W, Stevens TJ, Fogh RH, Pajon A, Llinas M, Ulrich EL, Markley JL, Ionides J, Laue ED. The CCPN data model for NMR spectroscopy: development of a software pipeline. *Proteins*. 2005; 59:687–96. [PubMed: 15815974]
40. Cheever ML, Kutateladze TG, Overduin M. Increased mobility in the membrane targeting PX domain induced by phosphatidylinositol 3-phosphate. *Protein Sci*. 2006; 15:1873–82. [PubMed: 16877709]

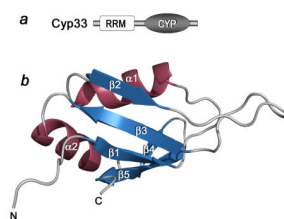


Figure 1. The crystal structure of the RRM domain of Cyp33 determined at 1.85 Å resolution. (a) Architecture of Cyp33: the amino-terminal RRM domain and the catalytic CYP domain. (b) Ribbon diagram of the RRM structure.

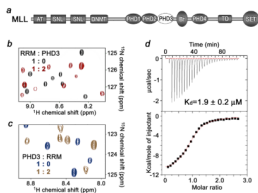
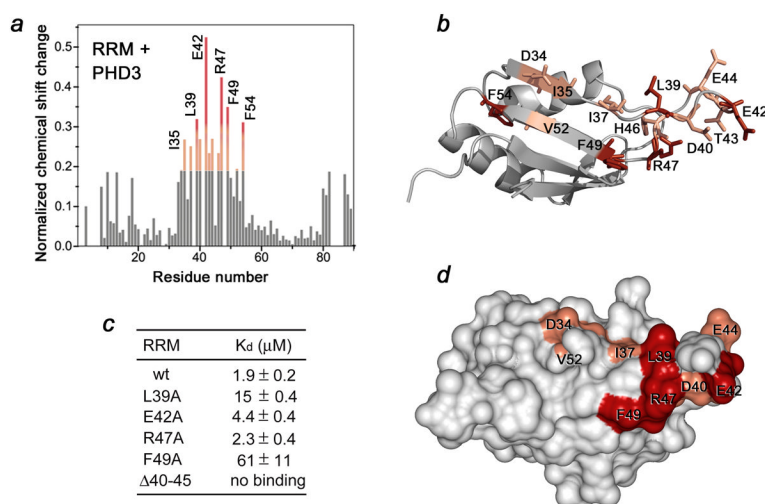


Figure 2.

Binding of the RRM domain of Cyp33 to the PHD3 finger of MLL. (a) Schematic of MLL. The PHD3 finger is shown as a white oval. (b) Superimposed ^1H , ^{15}N HSQC spectra of the ^{15}N -labeled Cyp33 RRM domain, collected in the absence and presence of a two-fold excess of unlabeled MLL PHD3. (c) Superimposed ^1H , ^{15}N HSQC spectra of the ^{15}N -labeled MLL PHD3 finger, collected in the absence and presence of a two-fold excess of unlabeled Cyp33 RRM. (d) Representative ITC curves used to calculate the binding affinity for the interaction between the Cyp33 RRM domain and the MLL PHD3 finger.

**Figure 3.**

The MLL PHD3-binding site of the Cyp33 RRM domain. (a) A histogram shows normalized ^1H , ^{15}N chemical shift changes in backbone amides of ^{15}N -labeled RRM upon addition of unlabeled PHD3 at a protein ratio of 1:1. (b, d) Residues that exhibit significant PHD3-induced resonance perturbations in (a) are mapped on the ribbon diagram (b) and the surface (d) of the Cyp33 RRM domain. Colored bars indicate significant change being greater than an average plus one standard deviation. (c) The binding affinities of wild type Cyp33 RRM and mutants for MLL PHD3, as measured by ITC. Interaction of the RRM Δ 40–45 mutant was examined by NMR. Other mutants generated, including G48A, F54A, RRM Δ 41–42, RRM Δ 41–43 and RRM Δ 40–43, precipitated during dialysis for ITC experiments.

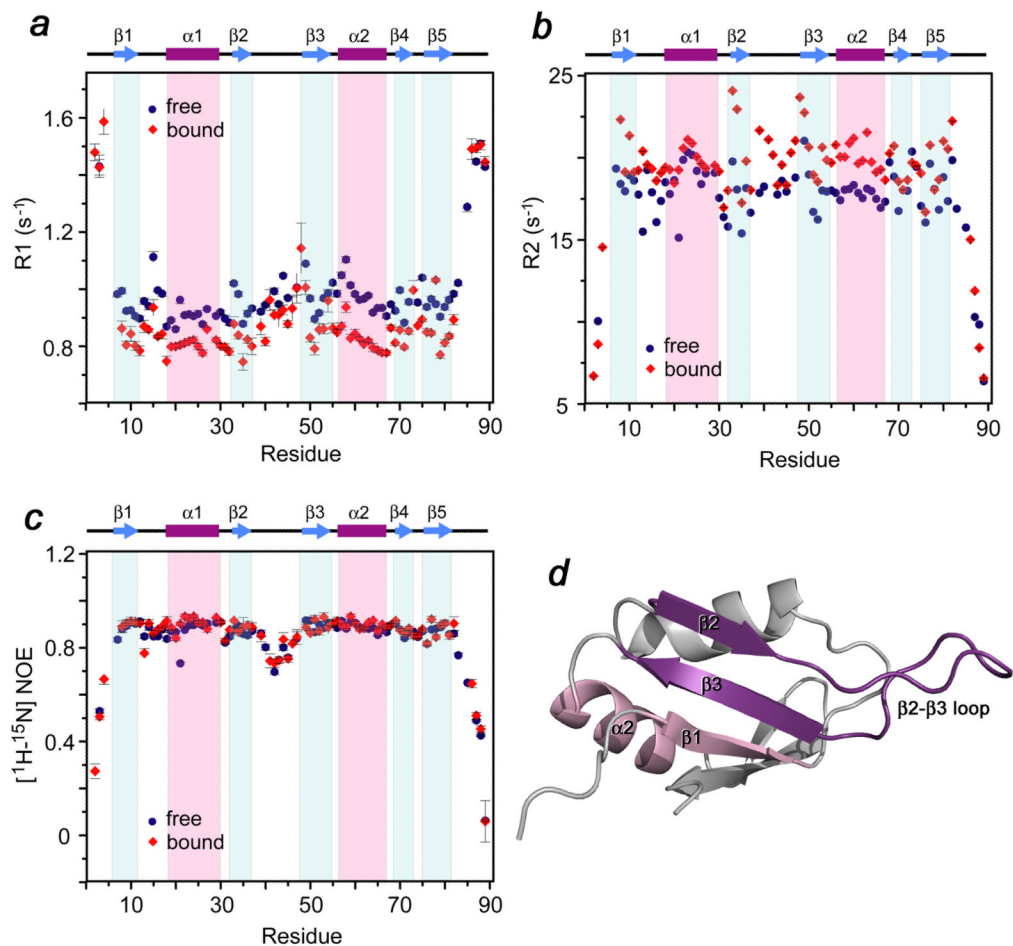
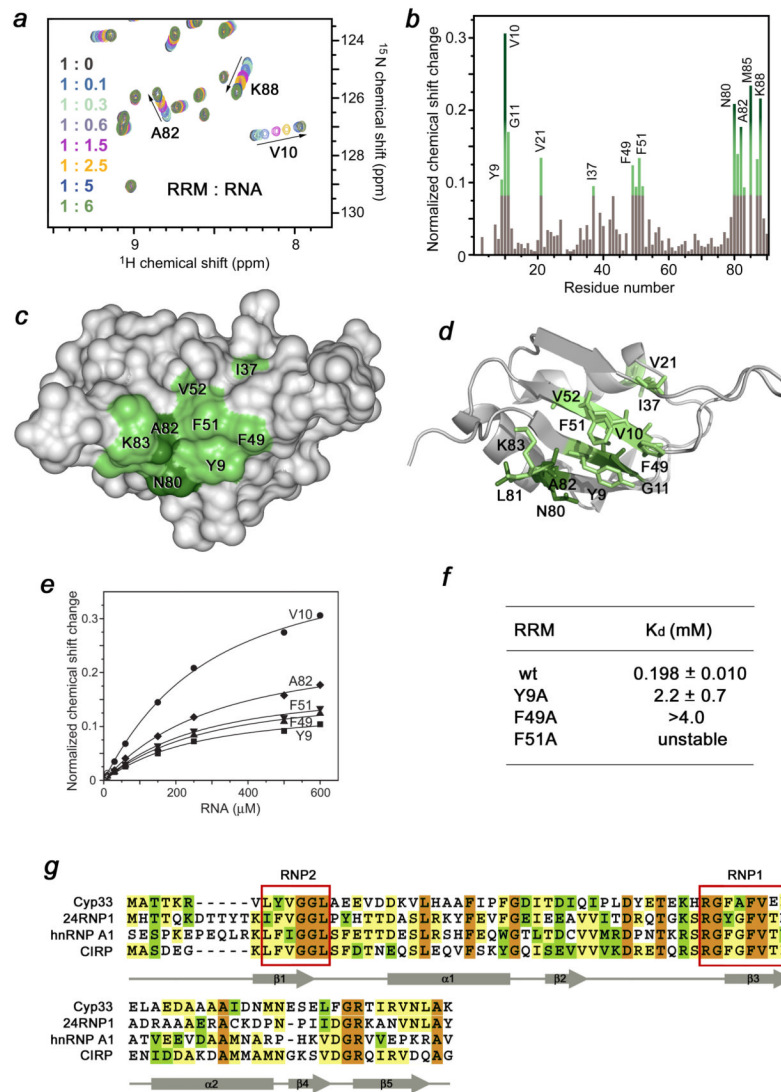


Figure 4.

Dynamics of the Cyp33 RRM domain bound and unbound. (a–c) Relaxation parameters of the RRM domain in the absence (blue) and the presence (red) of a 1.5-fold excess of MLL PHD3. R_1 , R_2 and NOE values were determined for backbone amide groups and are plotted for each residue of the Cyp33 RRM domain. The RRM secondary structure is shown above the graphs. (d) The most affected (due to the interaction with PHD3) residues of Cyp33 RRM are colored in shades of purple in the ribbon diagram of the RRM structure.

**Figure 5.**

The RNA-binding site of the Cyp33 RRM domain. (a) Eight superimposed ¹H,¹⁵N HSQC spectra of 0.1 mM Cyp33 RRM, collected as the AAUAAA RNA construct was gradually added. The spectra are color-coded according to the molar protein-RNA ratio (inset). (b) A histogram shows normalized ¹H,¹⁵N chemical shift changes in backbone amides of ¹⁵N-labeled RRM (aa 1–90) upon addition of a six-fold excess of AAUAAA. (c, d) Residues that exhibit significant RNA-induced resonance perturbations in (b) are labeled and colored in shades of green on the ribbon diagram of the structure (c) and on the surface (d) of the Cyp33 RRM domain (aa 1–83). Colored bars indicate significant change being greater than an average plus one and a half standard deviation. (e) Representative binding curves used to determine the K_d values of the Cyp33 RRM-RNA interaction by NMR spectroscopy. (f) The RNA binding affinities of the wild type and mutant RRM domain measured by NMR. (g) Alignment of the RRM domain sequences of human proteins: absolutely, moderately and weakly conserved residues are colored brown, green and yellow respectively. The RNP2 and RNP1 motifs required for the interaction with RNA are outlined by red rectangles. The secondary structure of the Cyp33 RRM domain is shown below the sequences.

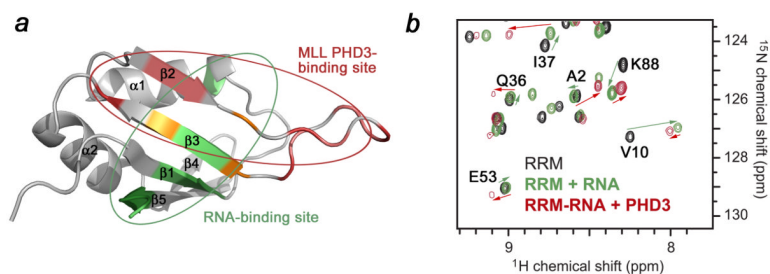


Figure 6. The MLL PHD3-binding site and the RNA-binding site of the Cyp33 RRM domain partially overlap. (a) The Cyp33 RRM domain is shown as a ribbon diagram. Residues of RRM that are perturbed by either PHD3, RNA or both ligands are colored red, green and yellow, respectively. The MLL PHD3-binding site and the RNA-binding site of the Cyp33 RRM domain are indicated by red and green circles, respectively. (b) Superimposed ^1H , ^{15}N HSQC spectra of the Cyp33 RRM domain (0.1 mM) in the ligand-free form (black), after addition of 0.6 mM RNA (green), and after subsequent addition of 0.25 mM MLL PHD3 (red).

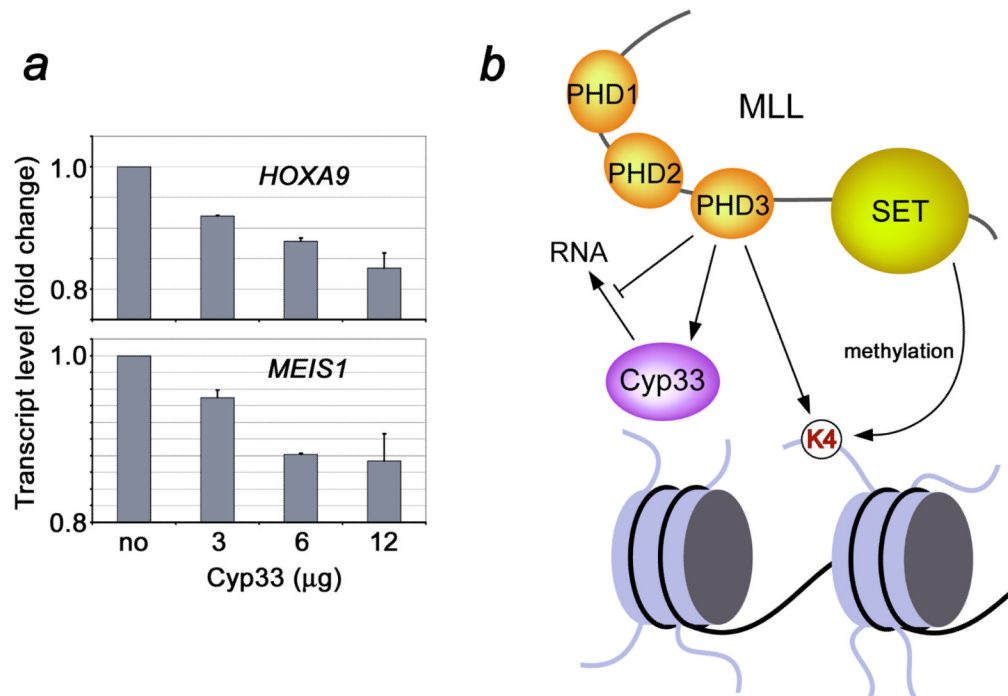


Figure 7. Cyp33 decreases MLL-dependent gene transcription. Gene expression levels were quantified in transfected 293T cells using quantitative real-time RT-PCR (a). (b) A model of the Cyp33-MLL association.