doi:10.1016/j.jmb.2010.04.067 J. Mol. Biol. (2010) 400, 145–154

JMB

Available online at www.sciencedirect.com

ScienceDirect

ELSEVIER

Molecular Mechanism of MLL PHD3 and RNA Recognition by the Cyp33 RRM Domain

Robert A. Hom[1], Pei-Yun Chang[2], Siddhartha Roy[1]†, Catherine A. Musselman[1]†, Karen C. Glass[1]†, Anna I. Selezneva[3], Or Gozani[4], Rustem F. Ismagilov[3], Michael L. Cleary[2] and Tatiana G. Kutateladze[1]*

[1]*Department of Pharmacology, University of Colorado Denver School of Medicine, 12801 East 17th Avenue, Aurora, CO 80045, USA*

[2]*Department of Pathology, Stanford University, Stanford, CA 94305, USA*

[3]*Department of Chemistry, University of Chicago, Chicago, IL 60637, USA*

[4]*Department of Biological Sciences, Stanford University, Stanford, CA 94305, USA*

Received 1 April 2010; received in revised form 30 April 2010; accepted 30 April 2010
Available online 8 May 2010

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 RNA-recognition motif (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 μM affinity as measured by isothermal titration calorimetry. NMR chemical shift perturbation (CSP) analysis and dynamics data reveal that the β strands and the β2–β3 loop of the RRM domain are involved in the interaction with PHD3. Mutations in the PHD3-binding site or deletions in the β2–β3 loop lead to a significantly reduced affinity or abrogation of the interaction. The RNA-binding pocket of the Cyp33 RRM domain, mapped on the basis of 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.

Edited by M. F. Summers

Keywords: RRM domain; PHD finger; Cyp33; MLL; mechanism

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 peptidyl-prolyl isomerases (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.[1,3] The catalytic CYP/PPIase domain of Cyp33 is located in the carboxy-terminal region of

*Corresponding author. E-mail address: Tatiana.Kutateladze@UCDenver.edu.

† S.R., C.A.M., and K.C.G. contributed equally to this work.

Abbreviations used: Cyp33, cyclophilin 33; RRM, RNA-recognition motif; PHD, plant homeodomain; MLL, myeloid/lymphoid or mixed lineage leukemia; CSP, chemical shift perturbation; PPIase, peptidyl-prolyl isomerase; RNP, ribonucleoprotein; HMTase, histone methyltransferase; SET, Su(var)3–9, Enhancer of Zeste, Trithorax; ITC, isothermal titration calorimetry; NOE, nuclear Overhauser enhancement; HSQC, heteronuclear single-quantum coherence.

0022-2836/$ - see front matter

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.[1,4] 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).[5,6] 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.[7,8] This large, ~4000-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).[9–11] Three sequential PHD modules, an acetyl-lysine binding bromodomain, and another atypical PHD finger precede a transactivation domain. The transactivation 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.[13–15] 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 trimethylated H3K4 marks.[16] A short sequence preceding the SET domain is recognized by WDR5.[17–19] 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, constitute 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.[5,6,21] The PHD finger region appears to play a regulatory role in MLL function and suppresses MLL-mediated leukemogenesis. 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 the 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 α1 helix is positioned between the β1 and β2 strands, whereas helix α2 connects strands β3 and β4. The β1 strand (residues V7–G11), α1 helix (residues D19–F26), β2 strand (residues T33–Q36), β3 strand (residues F49–F54), α2 helix (residues A57–M67), and β5 strand (residues R75–L81) adopt a canonical βαββαβ RRM fold with an additional strand (β4, residues E69–L72) pairing with the β5 strand and forming the edge of the β sheet. The overall structure of the Cyp33 RRM domain superimposes with the structure of a typical

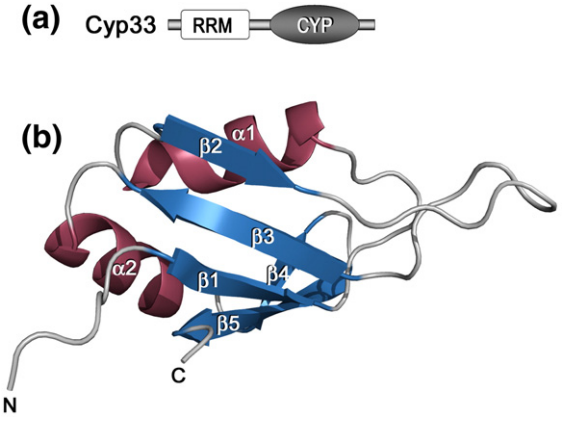


Fig. 1. The crystal structure of the RRM domain of Cyp33 determined at 1.85 Å resolution. (a) Architecture of Cyp33: the amino-terminal RRM domain and the catalytic CYP domain. (b) Ribbon diagram of the RRM structure.

RRM module, such as human hnRNP A1 [Protein Data Bank (PDB) 1L3K][23] with a root-mean-square deviation of 1.4 Å over C^{α} atoms. The diffraction data and refinement statistics for the structure are shown in Supplementary Table 1.

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

The interaction of Cyp33 RRM with PHD3 of MLL was initially tested by NMR and ITC (Fig. 2). $^{1}H,^{15}N$ heteronuclear single quantum coherence (HSQC) spectra of the ^{15}N-labeled RRM domain were recorded while unlabeled PHD3 was added stepwise to the RRM sample (Fig. 2b and data not shown). As PHD3 was titrated in, resonances corresponding to the unbound state of the RRM domain decreased in intensity, disappearing completely at an RRM-to-PHD3 ratio of 1:1. Concomitantly, another set of resonances corresponding to the RRM–PHD3 complex gradually appeared, replacing the initial set of resonances. This pattern of chemical shift changes is characteristic of slow exchange on the NMR time scale and is indicative of a robust interaction between RRM and PHD3. Likewise, the reverse titration of unlabeled Cyp33 RRM into ^{15}N-labeled MLL PHD3 caused a gradual decrease of peak intensities for the ligand-free PHD3 finger and an increase of peak intensities of the RRM-bound state of PHD3 (Fig. 2c). The slow exchange regime observed in the reverse titration confirmed the strong association between the two proteins. A lack of significant chemical shift perturbations in the ^{15}N-labeled MLL PHD3 finger upon addition of the unlabeled catalytic domain of Cyp33 indicated that the catalytic domain is not involved in the recognition of PHD3 (Supplementary Fig. 1).

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± 0.2 μM, and is in agreement with the pattern of resonance perturbations in the NMR titration experiments. The interaction was enthalpy driven ($\Delta H = -10 \pm 0.5$ kcal), and a negative change in entropy ($\Delta S = -9 \pm 2$ cal mol^{-1} 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 ^{1}H, ^{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 $^{15}N/^{13}C$-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 β2–β3 loop and β2 and β3 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 β1 and β5

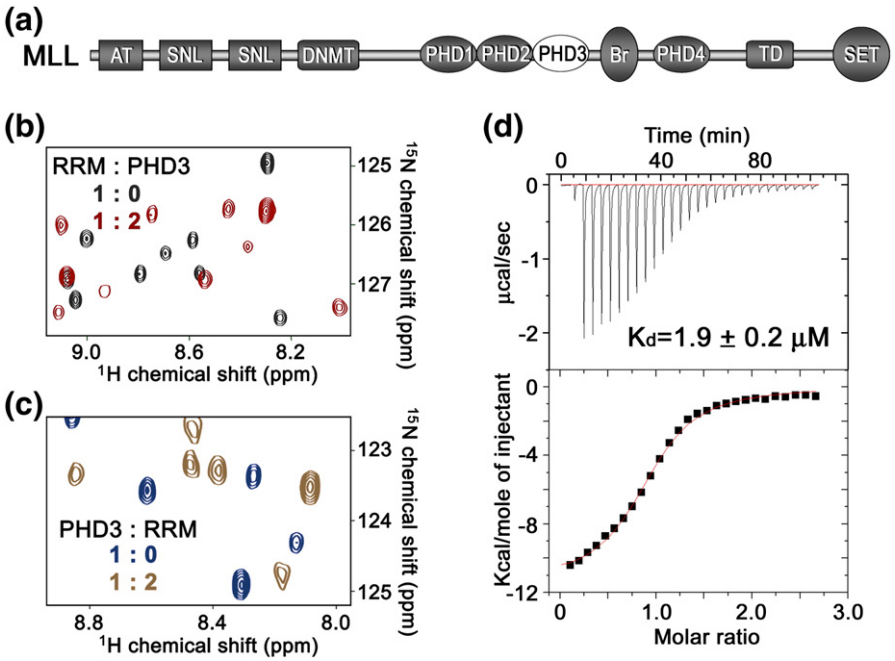
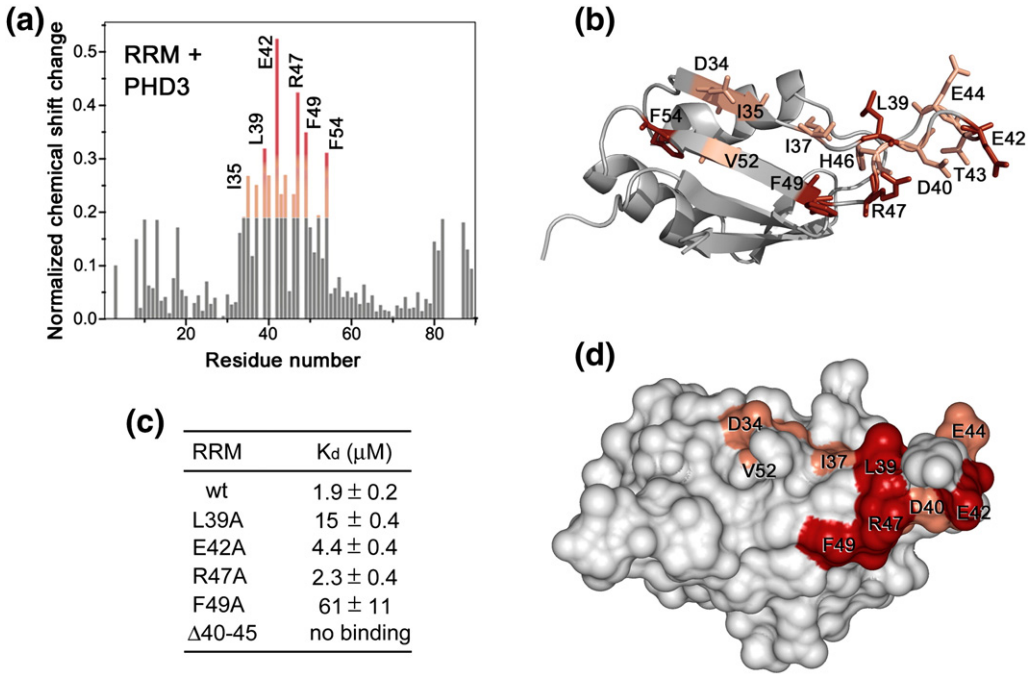


Fig. 2. Binding of the RRM domain of Cyp33 to the PHD3 finger of MLL. (a) Schematic of MLL. The PHD3 finger is shown as a white oval. (b) Superimposed $^{1}H,^{15}N$ HSQC spectra of the ^{15}N-labeled Cyp33 RRM domain, collected in the absence and presence of a twofold excess of unlabeled MLL PHD3. (c) Superimposed $^{1}H,^{15}N$ HSQC spectra of the ^{15}N-labeled MLL PHD3 finger, collected in the absence and presence of a twofold 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.



(c)

RRM	Kd (μM)
wt	1.9 ± 0.2
L39A	15 ± 0.4
E42A	4.4 ± 0.4
R47A	2.3 ± 0.4
F49A	61 ± 11
Δ40-45	no binding

Fig. 3. The MLL PHD3-binding site of the Cyp33 RRM domain. (a) A histogram shows normalized $^{1}H,^{15}N$ chemical shift changes in backbone amides of ^{15}N-labeled RRM upon addition of unlabeled PHD3 at a protein ratio of 1:1. (b and 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 RRMΔ40–45 mutant was examined by NMR. Other mutants generated, including G48A, F54A, RRMΔ41–42, RRMΔ41–43 and RRMΔ40–43, precipitated during dialysis for ITC experiments.

strands were perturbed to a lesser degree (Fig. 3a, b, and d). These data suggest that the β2–β3 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 Cyp33 RRM domain, we replaced them with alanine and, additionally, generated mutants with partially truncated β2–β3 loop. Binding of the mutant proteins to PHD3 was examined by ITC and NMR. Deletion of the β2–β3 loop in RRMΔ40–45 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 β2–β3 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 β2–β3 loop, the β1, β2, and β3 strands, and the α2 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 in 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 β2–β3 loop, the β1, β2, and β3 strands, and the α2 helix, particularly those that are involved in the interaction with the PHD3 finger. The lack of changes in the heteronuclear NOE (nuclear Overhauser enhancement) values indicated that the local internal backbone mobility on a subnanosecond time scale does not change due to the complex formation. The significantly larger R_2 relaxation rates observed for some residues in the β1, β2, and β3 strands pointed to a contribution from microsecond–millisecond 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, $^{1}H,^{15}N$ HSQC spectra of the ^{15}N-labeled RRM domain were recorded as the unlabeled PHD module of tumor

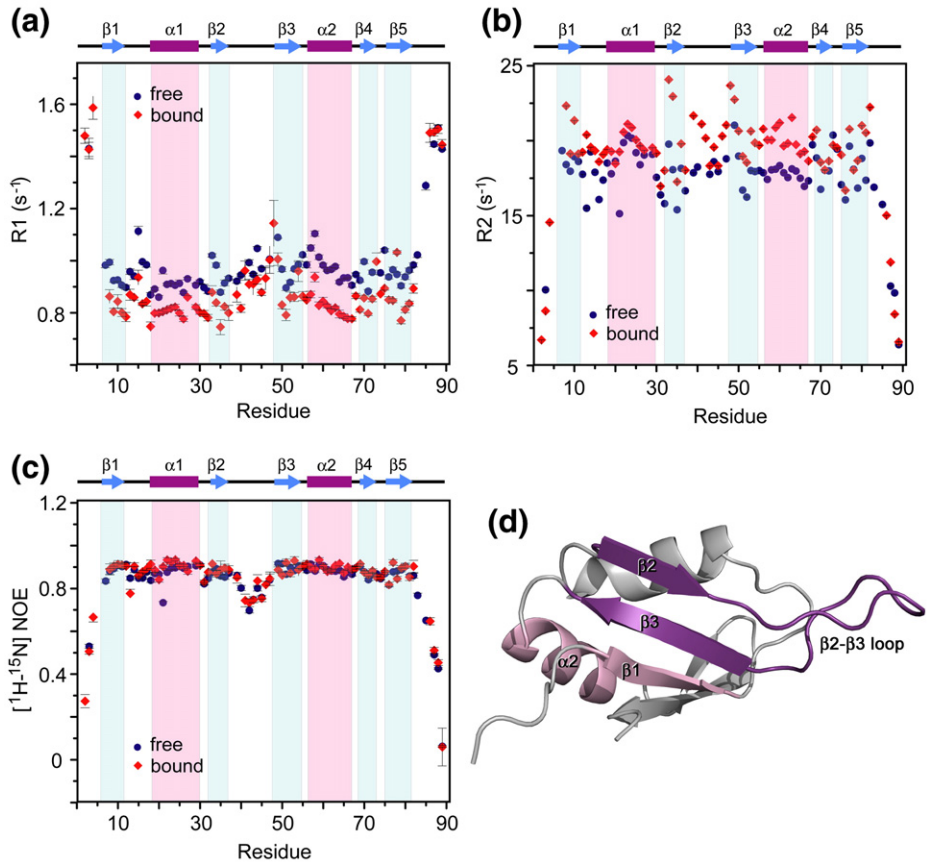


Fig. 4. Dynamics of the Cyp33 RRM domain bound and unbound. (a–c) Relaxation parameters of the RRM domain in the absence (blue) and the presence (red) of a 1.5-fold excess of MLL PHD3. R_1, R_2, and NOE values were determined for backbone amide groups and are plotted for each residue of the Cyp33 RRM 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.

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 fivefold 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 fivefold excess of unlabeled *Drosophila* Cyp33 RRM to the ^{15}N-labeled MLL PHD3 finger or, conversely, addition of a fivefold excess of unlabeled MLL PHD3 to the ^{15}N-labeled *Drosophila* Cyp33 RRM caused negligible changes in ^{1}H,^{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 ^{1}H,^{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 β1 strand (Y9, V10, G11), β3 strand (F49, A50, F51, V52), β5 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 and 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 constitute the two conserved RNP motifs RNP2 (YVGGL-13) and RNP1 (RGFAF-VEF-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±11 μM) 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

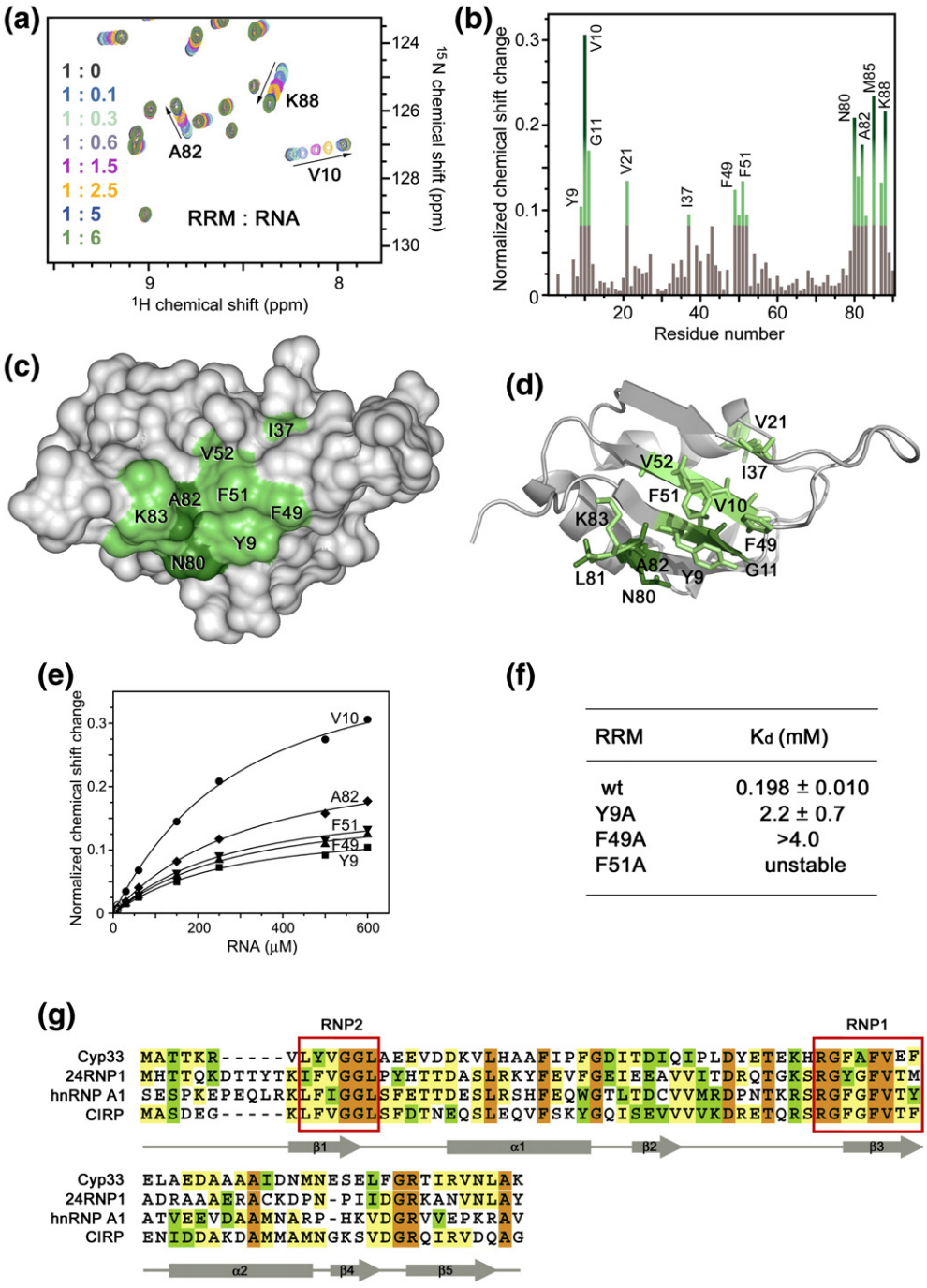


Fig. 5. The RNA-binding site of the Cyp33 RRM domain. (a) Eight superimposed ^{1}H,^{15}N HSQC spectra of 0.1 mM Cyp33 RRM, collected as the AAUAAA RNA construct was gradually added. The spectra are color-coded according to the molar protein–RNA ratio (inset). (b) A histogram shows normalized ^{1}H,^{15}N chemical shift changes in backbone amides of ^{15}N-labeled RRM (amino acids 1–90) upon addition of a sixfold excess of AAUAAA. (c and 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 (amino acids 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.

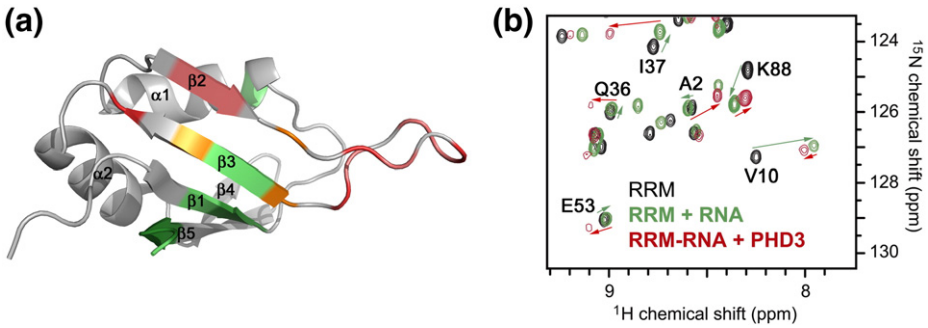


Fig. 6. The MLL PHD3-binding site and the RNA-binding site of the Cyp33 RRM domain partially overlap. (a) The Cyp33 RRM domain is shown as a ribbon diagram. Residues of RRM that are perturbed by either PHD3, RNA, or both ligands are colored red, green, and yellow, respectively. (b) Superimposed $^{1}H,^{15}N$ 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).

micromolar 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 $^{1}H,^{15}N$ 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 2 days RNA was prepared and reverse-transcribed. *HOXA9*, *MEIS1*, and *MLL* expression levels were analyzed by quantitative RT-PCR. As shown in Fig. 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*

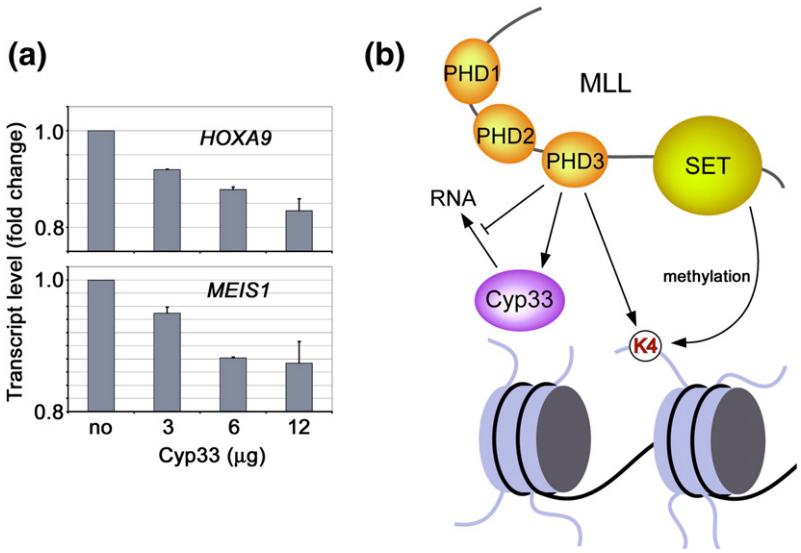


Fig. 7. Cyp33 decreases MLL-dependent gene transcription. (a) Gene expression levels were quantified in transfected 293T cells using quantitative real-time RT-PCR. (b) A model of the Cyp33–MLL association.

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 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 the 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, ^{15}N-labeled, and ^{15}N/^{13}C-labeled proteins were expressed in *Escherichia coli* Rosetta 2 (DE3) in LB or minimal media supplemented with $^{15}NH_4Cl$ or $^{15}NH_4Cl/^{13}C_6$-glucose (Isotec). The bacterial cells were grown at 37 °C to an OD_{600} of 0.8 and protein expression was induced with 1.0 mM isopropyl-β-D-thiogalactopyranoside (IPTG) at 37 °C for 5 h. The cells were collected by centrifugation at 5000*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] and eluted with 20 mM Tris–HCl (pH 8.0) buffer containing 150 mM NaCl, 2 mM β-mercaptoethanol, 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 ^{15}N-labeled PHD3 fingers were 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 ^{15}N-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 Ref. 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 trisodium dihydrate citrate at pH 7.6. All crystals grew in a monoclinic space group (*C*2) with unit cell parameters of $a = 84.84$ Å, $b = 40.52$ Å, $c = 65.66$ Å, $\alpha = \gamma = 90°$, $\beta = 127.06°$ with two molecules per asymmetric unit. 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 were processed with D*TREK.[31] The molecular replacement solution was generated with 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, RRMΔ41–42, RRMΔ41–43, RRMΔ40–43, and RRMΔ40–45) were generated with 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}N/^{13}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}H,^{15}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}H,^{15}N HSQC spectra of 0.1–0.2 mM ^{15}N-labeled RRM while 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) was added stepwise. The ligand binding to the wild-type MLL PHD3 finger was characterized by monitoring chemical shift changes in ^{1}H,^{15}N HSQC spectra of 0.2 mM ^{15}N-labeled PHD3 while unlabeled wild-type or mutant human Cyp33 RRM (up to 0.4 mM), or the catalytic domain of human Cyp33 (up to 1 mM) was added gradually. Interactions 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) were tested similarly.

Relaxation experiments

Changes in dynamics of the Cyp33 RRM domain upon binding to the MLL PHD3 finger were investigated by backbone amide ^{15}N relaxation experiments. The ^{15}N R_1, R_2 and ^{1}H,^{15}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 with 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 NaN_3). The heat of the reactions was measured by making 30 sequential injections of 10 μl of PHD3 (0.65 mM) into 1.4 ml of RRM solution (0.05 mM) (and *vice versa*) with spacing intervals of 60 s. 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 nonlinear least-squares fitting of the data with MicroCal Origin software (MicroCal).

RT-PCR assays

293T cells were transfected with FLAG-HIS6-CYP33 using the Fugene6 method. After 2 days, RNA was prepared with Trizol (Invitrogen) and reverse-transcribed with 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.

PDB accession numbers

Coordinates and structure factors have been deposited to the PDB with accession number 3MDF.

Acknowledgements

We thank Liang Li, Sigrid Nachtergaele, and Jennifer Schlegel for discussions and help with the experiments, A.I.S. and R.F.I. for the identification of crystallization conditions and providing the crystals of RRM, 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 was supported by National Institutes of Health grants GM074961 and GM075827 (R.F.I.), CA55029 and CA116606 (M.L.C.), and CA113472 and GM071424 (T.G.K.).

Supplementary Data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.jmb.2010.04.067

References

1. Mi, H., Kops, O., Zimmermann, E., Jaschke, A. & Tropschug, M. (1996). A nuclear RNA-binding cyclophilin in human T cells. *FEBS Lett.* **398**, 201–205.
2. Wang, X. J. & Etzkorn, F. A. (2006). Peptidyl-prolyl isomerase inhibitors. *Biopolymers*, **84**, 125–146.
3. Min, L., Fulton, D. B. & Andreotti, A. H. (2005). A case study of proline isomerization in cell signaling. *Front. Biosci.* **10**, 385–397.
4. Wang, Y., Han, R., Zhang, W., Yuan, Y., Zhang, X., Long, Y. & Mi, H. (2008). Human CyP33 binds specifically to mRNA and binding stimulates PPIase activity of hCyP33. *FEBS Lett.* **582**, 835–839.
5. Fair, K., Anderson, M., Bulanova, E., Mi, H., Tropschug, M. & Diaz, M. O. (2001). Protein interactions of the MLL PHD fingers modulate MLL target gene regulation in human cells. *Mol. Cell. Biol.* **21**, 3589–3597.
6. Chen, J., Santillan, D. A., Koonce, M., Wei, W., Luo, R., Thirman, M. J. *et al.* (2008). Loss of MLL PHD finger 3 is necessary for MLL-ENL-induced hematopoietic stem cell immortalization. *Cancer Res.* **68**, 6199–6207.
7. Ayton, P. M. & Cleary, M. L. (2001). Molecular mechanisms of leukemogenesis mediated by MLL fusion proteins. *Oncogene*, **20**, 5695–5707.
8. Hess, J. L. (2004). MLL: a histone methyltransferase disrupted in leukemia. *Trends Mol. Med.* **10**, 500–507.
9. Yano, T., Nakamura, T., Blechman, J., Sorio, C., Dang, C. V., Geiger, B. & Canaani, E. (1997). Nuclear punctate distribution of ALL-1 is conferred by distinct elements at the N terminus of the protein. *Proc. Natl Acad. Sci. USA*, **94**, 7286–7291.
10. Birke, M., Schreiner, S., Garcia-Cuellar, M. P., Mahr, K., Titgemeyer, F. & Slany, R. K. (2002). The MT domain of the proto-oncoprotein MLL binds to CpG-containing DNA and discriminates against methylation. *Nucleic Acids Res.* **30**, 958–965.
11. Allen, M. D., Grummitt, C. G., Hilcenko, C., Min, S. Y., Tonkin, L. M., Johnson, C. M. *et al.* (2006). Solution structure of the nonmethyl-CpG-binding CXXC domain of the leukaemia-associated MLL histone methyltrans. *EMBO J.* **25**, 4503–4512.

12. Ernst, P., Wang, J., Huang, M., Goodman, R. H. & Korsmeyer, S. J. (2001). MLL and CREB bind cooperatively to the nuclear coactivator CREB-binding protein. *Mol. Cell Biol.* **21**, 2249–2258.
13. Nakamura, T., Mori, T., Tada, S., Krajewski, W., Rozovskaia, T., Wassell, R. *et al.* (2002). ALL-1 is a histone methyltransferase that assembles a supercomplex of proteins involved in transcriptional regulation. *Mol. Cell*, **10**, 1119–1128.
14. Milne, T. A., Briggs, S. D., Brock, H. W., Martin, M. E., Gibbs, D., Allis, C. D. & Hess, J. L. (2002). MLL targets SET domain methyltransferase activity to *Hox* gene promoters. *Mol. Cell*, **10**, 1107–1117.
15. Southall, S. M., Wong, P. S., Odho, Z., Roe, S. M. & Wilson, J. R. (2009). Structural basis for the requirement of additional factors for MLL1 SET domain activity and recognition of epigenetic marks. *Mol. Cell*, **33**, 181–191.
16. Schneider, J., Wood, A., Lee, J. S., Schuster, R., Dueker, J., Maguire, C. *et al.* (2005). Molecular regulation of histone H3 trimethylation by COMPASS and the regulation of gene expression. *Mol. Cell*, **19**, 849–856.
17. Patel, A., Dharmarajan, V. & Cosgrove, M. S. (2008). Structure of WDR5 bound to mixed lineage leukemia protein-1 peptide. *J. Biol. Chem.* **283**, 32158–32161.
18. Patel, A., Vought, V. E., Dharmarajan, V. & Cosgrove, M. S. (2008). A conserved arginine-containing motif crucial for the assembly and enzymatic activity of the mixed lineage leukemia protein-1 core complex. *J. Biol. Chem.* **283**, 32162–32175.
19. Song, J. J. & Kingston, R. E. (2008). WDR5 interacts with mixed lineage leukemia (MLL) protein via the histone H3-binding pocket. *J. Biol. Chem.* **283**, 35258–35264.
20. Steward, M. M., Lee, J. S., O'Donovan, A., Wyatt, M., Bernstein, B. E. & Shilatifard, A. (2006). Molecular regulation of H3K4 trimethylation by ASH2L, a shared subunit of MLL complexes. *Nat. Struct. Mol. Biol.* **13**, 852–854.
21. Xia, Z. B., Anderson, M., Diaz, M. O. & Zeleznik-Le, N. J. (2003). 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. USA*, **100**, 8342–8347.
22. Muntean, A. G., Giannola, D., Udager, A. M. & Hess, J. L. (2008). The PHD fingers of MLL block MLL fusion protein-mediated transformation. *Blood*, **112**, 4690–4693.
23. Vitali, J., Ding, J., Jiang, J., Zhang, Y., Krainer, A. R. & Xu, R. M. (2002). Correlated alternative side chain conformations in the RNA-recognition motif of heterogeneous nuclear ribonucleoprotein A1. *Nucleic Acids Res.* **30**, 1531–1538.
24. Peña, P. V., Davrazou, F., Shi, X., Walter, K. L., Verkhusha, V. V., Gozani, O. *et al.* (2006). Molecular mechanism of histone H3K4me3 recognition by plant homeodomain of ING2. *Nature*, **442**, 100–103.
25. Peña, P. V., Hom, R. A., Hung, T., Lin, H., Kuo, A. J., Wong, R. P. *et al.* (2008). Histone H3K4me3 binding is required for the DNA repair and apoptotic activities of ING1 tumor suppressor. *J. Mol. Biol.* **380**, 303–312.
26. Clery, A., Blatter, M. & Allain, F. H. (2008). RNA recognition motifs: boring? Not quite. *Curr. Opin. Struct. Biol.* **18**, 290–298.
27. Hargous, Y., Hautbergue, G. M., Tintaru, A. M., Skrisovska, L., Golovanov, A. P., Stevenin, J. *et al.* (2006). Molecular basis of RNA recognition and TAP binding by the SR proteins SRp20 and 9G8. *EMBO J.* **25**, 5126–5137.
28. Abdul-Manan, N., O'Malley, S. M. & Williams, K. R. (1996). Origins of binding specificity of the A1 heterogeneous nuclear ribonucleoprotein. *Biochemistry*, **35**, 3545–3554.
29. Nadler, S. G., Merrill, B. M., Roberts, W. J., Keating, K. M., Lisbin, M. J., Barnett, S. F. *et al.* (1991). 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*, **30**, 2968–2976.
30. Li, L., Mustafi, D., Fu, Q., Tereshko, V., Chen, D. L., Tice, J. D. & Ismagilov, R. F. (2006). Nanoliter microfluidic hybrid method for simultaneous screening and optimization validated with crystallization of membrane proteins. *Proc. Natl Acad. Sci. USA*, **103**, 19243–19248.
31. Pflugrath, J. W. (1999). The finer things in X-ray diffraction data collection. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **55**, 1718–1725.
32. Long, F., Vagin, A. A., Young, P. & Murshudov, G. N. (2008). BALBES: a molecular-replacement pipeline. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **64**, 125–132.
33. Brunger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P. & Grosse-Kunstleve, R. W. (1998). Crystallography & NMR system: a new software suite for macromolecular structure determination. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **54**, 905–921.
34. Emsley, P. & Cowtan, K. (2004). Coot: model-building tools for molecular graphics. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **60**, 2126–2132.
35. Laskowski, R. A., MacArthur, M. W., Moss, D. S. & Thornton, J. M. (1993). PROCHECK: a program to check the stereochemical quality of protein structures. *J. Appl. Crystallogr.* **26**, 283–291.
36. Wittekind, M. & Mueller, L. (1993). 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.* **101**, 201–205.
37. Grzesiek, S. & Bax, A. (1992). Improved 3D triple-resonance NMR techniques applied to a 31-kDa protein. *J. Magn. Reson.* **96**, 432–440.
38. Delaglio, F., Grzesiek, S., Vuister, G. W., Zhu, G., Pfeifer, J. & Bax, A. (1995). NMRPipe: a multidimensional spectral processing system based on UNIX pipes. *J. Biomol. NMR*, **6**, 277–293.
39. Vranken, W. F., Boucher, W., Stevens, T. J., Fogh, R. H., Pajon, A., Llinas, M. *et al.* (2005). The CCPN data model for NMR spectroscopy: development of a software pipeline. *Proteins*, **59**, 687–696.
40. Cheever, M. L., Kutateladze, T. G. & Overduin, M. (2006). Increased mobility in the membrane targeting PX domain induced by phosphatidylinositol 3-phosphate. *Protein Sci.* **15**, 1873–1882.