The structure of the RbBP5 β-propeller domain reveals a surface with potential nucleic acid binding sites

Anshumali Mittal[1], Fruzsina Hobor[2], Ying Zhang[1], Stephen R. Martin[3], Steven J. Gamblin[1], Andres Ramos[2] and Jon R. Wilson[1],*

[1]The Francis Crick Institute, 1 Midland Road, London NW1 1AT, UK, [2]Institute of Structural and Molecular Biology, University College London, London WC1E 6XA, UK and [3]Structural Biology Science Technology Platform, The Francis Crick Institute, 1 Midland Road, London, NW1 1AT, UK

ABSTRACT

The multi-protein complex WRAD, formed by WDR5, RbBP5, Ash2L and Dpy30, binds to the MLL SET domain to stabilize the catalytically active conformation required for histone H3K4 methylation. In addition, the WRAD complex contributes to the targeting of the activated complex to specific sites on chromatin. RbBP5 is central to MLL catalytic activation, by making critical contacts with the other members of the complex. Interestingly its only major structural domain, a canonical WD40 repeat β-propeller, is not implicated in this function. Here, we present the structure of the RbBP5 β-propeller domain revealing a distinct, feature rich surface, dominated by clusters of Arginine residues. Our nuclear magnetic resonance binding data supports the hypothesis that in addition to the role of RbBP5 in catalytic activation, its β-propeller domain is a platform for the recruitment of the MLL complexes to chromatin targets through its direct interaction with nucleic acids.

INTRODUCTION

The novel motifs generated by methylation of histone lysine residues recruit effector proteins through selective recognition domains and therefore drive the cells gene expression program (1,2). This process must be tightly controlled, which is reflected in the complexity of the trithorax and polycomb methyltransferase complexes that regulate gene activation and repression through histone H3 lysine-4 and lysine-27 methylation respectively (3,4). In addition to their catalytic domains, these complexes contain components that allosterically regulate enzyme activity and others that precisely target the enzyme activity to specific genes (5,6). In higher organisms, histone H3 lysine-4 (H3K4) methylation is carried out by a family of six large multi-domain enzymes, the Mixed Lineage Leukemia (MLL) family. These enzymes share a distinctive C-terminal catalytic SET domain, which is activated by binding to a conserved four-member complex, termed WRAD (7,8). WRAD consists of two β-propeller domain proteins, WDR5 and RbBP5, the SPRY/DNA binding domain protein Ash2L and Dpy30 (9). Of these, WDR5, RbBP5 and Ash2L have been shown to be essential for catalytic activation and are thought to stabilize the active conformation of the SET domain (10,11). Dpy30 may further modulate this activity although the molecular details remain to be elucidated (12).

The MLL family are large proteins, with the major part of the sequence composed of many so-called reader domains, linked by regions of predicted low structural complexity. The function of the reader domains is to target the enzyme by recognition of chromatin or other factors (13). For example, the PHD3-Bromo cassette of MLL1, binds to its own product—methylated H3K4, and facilitates the spread of the activating mark (14). Members of the WRAD complex have also been shown to bind directly to factors that target MLL. For example, the WDR5 β-propeller can bind to unmodified histone H3 (15), Ash2L binds directly to the transcription factors Mef2 and Sox2, and Dpy30 to Oct4 (16,17). In addition to these protein: protein interactions there is mounting evidence of lncRNA-mediated MLL recruitment to specific loci in chromosomes for gene activation (18–20). The details of these interactions are uncharacterized, but one RNA binding interface has been identified on the WDR5 β-propeller and confirmed by mutation of key residues that disrupted normal H3K4 methylation pathways in cells (21). In this manuscript we explore the po-

tential of the RbBP5 β-propeller as another site of RNA interaction.

The amino terminus of RbBP5, residues 1–325, is composed of a series of canonical WD40 repeats that are predicted to form a seven bladed β-propeller domain, which is the major structural feature of the protein (Figure 1A). However, it is the region immediately following this domain, residues 340–380, that stimulates methyltransferase activity through cooperative interactions with the other members of WRAD complex and the MLL SET domain. Specifically, the region RbBP5$_{344–360}$ binds to Ash2L ([22]), RbBP5$_{372–381}$ to WDR5 ([23,24]) and RbBP5$_{330–344}$ to the MLL SET domain ([11]). Thus, only a small 40 amino acid region has an ascribed function, and the role of the β-propeller domain and the remaining 150 C-terminal residues of RbBP5 is currently unknown. Here we focus on the RbBP5 β-propeller domain and show that although it is not essential for WRAD-mediated MLL methyltransferase stimulation, it potentially has a role in targeting the MLL complex. We present the first crystal structure of the β-propeller domain and, supported by nuclear magnetic resonance (NMR) spectroscopy, propose a targeting function through binding of nucleic acid.

MATERIALS AND METHODS

Protein expression and purification

The mouse *RbBP5* gene (residues 1–380) was cloned into the pOPINS vector for protein production in *Escherichia coli* BL21 (DE3) cells ([24]). The 5′ end of *RbBP5 gene* was extended with sequences encoding a 6×His-SUMO tag followed by a Precission 3C protease site (L-E-V-L-F-Q-G-P). RbBP5 (1–380) expression was performed in *E. coli* BL21 (DE3) cells, grown in TB media. *RbBP5* (*1–322*) and *RbBP5* (*1–340*) constructs were generated by introducing a stop codon at the respective sites in *RbBP5* (*1–380*) by site-directed mutagenesis. Site directed arginine to glutamic acid mutants were made in the RbBP5 (1–340) construct. Protein production was induced at OD$_{600}$ of 0.8 with 0.2 mM Isopropyl β-D-1-thiogalactopyranoside (IPTG) for 16–20 h at 20°C. Harvested cell pellets were resuspended in 50 mM Tris–HCl (pH 7.4), 150 mM NaCl, 10% glycerol, 0.5 mM tris(2-carboxyethyl)phosphine (TCEP), 25 μg/ml DNaseI, protease inhibitor cocktail (Sigma) and cells were lysed by sonication on ice. Unbroken cells and cell debris were removed by centrifugation at 20 000 rpm (Sorvall SS-34 rotor) for 60 min at 4°C and cleared lysate was loaded onto a Ni^{2+}-NTA gravity flow column containing 3 ml resin (Qiagen). The resin was washed with 10 column volumes of buffer containing 20 mM imidazole and incubated with rhinovirus 3C protease overnight at 4°C. Untagged RbBP5 was separated from the 6×His-SUMO tag bound to Ni^{2+}-NTA resin by gravity flow. RbBP5 was finally purified by gel filtration (Superdex 75, GE Healthcare) in 50 mM Tris/HCl (pH 7.4), 150 mM NaCl, 0.5 mM TCEP and 10% glycerol and the fractions of the homogenous main peak corresponding to monomeric RbBP5 were pooled. Purified RbBP5 was snap-frozen in liquid nitrogen and stored at −80°C. The RbBP5 (1–340) construct for NMR was expressed in *E. coli* BL21 (DE3) cells in minimal media with $^{15}NH_4Cl_2$ sole as nitro-

gen source and purified by affinity and size-exclusion chromatography as described above.

Crystallization

Crystals of RBBP5 (1–380) were obtained using the vapor diffusion method. RbBP5 (1–380) protein was concentrated to 15–20 mg/ml and used for crystallization by sitting-drop method. Initial crystals appeared in 0.1M MES (pH 6.5), 40% polyethylene glycol (PEG) 200 (V/V) (A7 condition of PEGs suite from Qiagen) at 20°C. The optimized condition included a mixture of additives (H3 of Silver Bullet from Hampton Research) to the original condition. Crystals were harvested directly from the drop and flash-frozen in liquid nitrogen.

Structure determination

Data were collected at the Diamond Light Source (Oxfordshire, UK) on station I02. The reflections were indexed using XDS and reduced/scaled with programs from the CCP4i suite ([25]). The structure was solved by molecular replacement using the PHASER package using edited coordinates of WDR5. Difference maps were used to rebuild and extend the initial model using the Coot molecular graphics package ([26]). Iterative cycles of refinement were carried out using REFMAC ([27]). Structure visualization was carried out in the PyMol Molecular Graphics System (Schrodinger LLC). Electrostatic surfaces were generated using the APBS plugin ([28]), protein prepared using pdb2pqr ([29]), map calculated with grid spacing 0.5 and a solvent excluded surface generated. Conservation was potted onto the surface using the AL2CO sequence conservation analysis server ([30]), using the multiple sequence analysis presented in Supplementary Figure S2. The coordinates and structure factors for the structure have been deposited in the PDB under accession code 5OV3.

RNA preparation

All RNAs pools were purchased from Thermo Scientific (Dharmacon). Pools were de-protected following the manufacturer instructions and re-suspended in H$_2$O.

NMR spectroscopy

NMR data were acquired using a Bruker NMR spectrometer operating at 800 and 950 MHz. NMR samples consisted of 40–50 μM ^{15}N-labeled RbBP5(1–340) alone and in complex with unlabeled RNA at a 1:1 or 1:4 protein:RNA molar ratio, and were prepared in 350 μl of NMR buffer (10 mM Tris–HCl, 75 mM NaCl, pH 7.2 10% D$_2$O). RbBP5–RNA interaction measurements by HSQC NMR were performed at 37°C. To measure Arginine-side chains experiments were performed at lower pH (pH 6.9 at room temperature). The scaffold independent analysis (SIA) strategy and analysis was as previously described ([31]), using a pool of 6mer nucleotides.
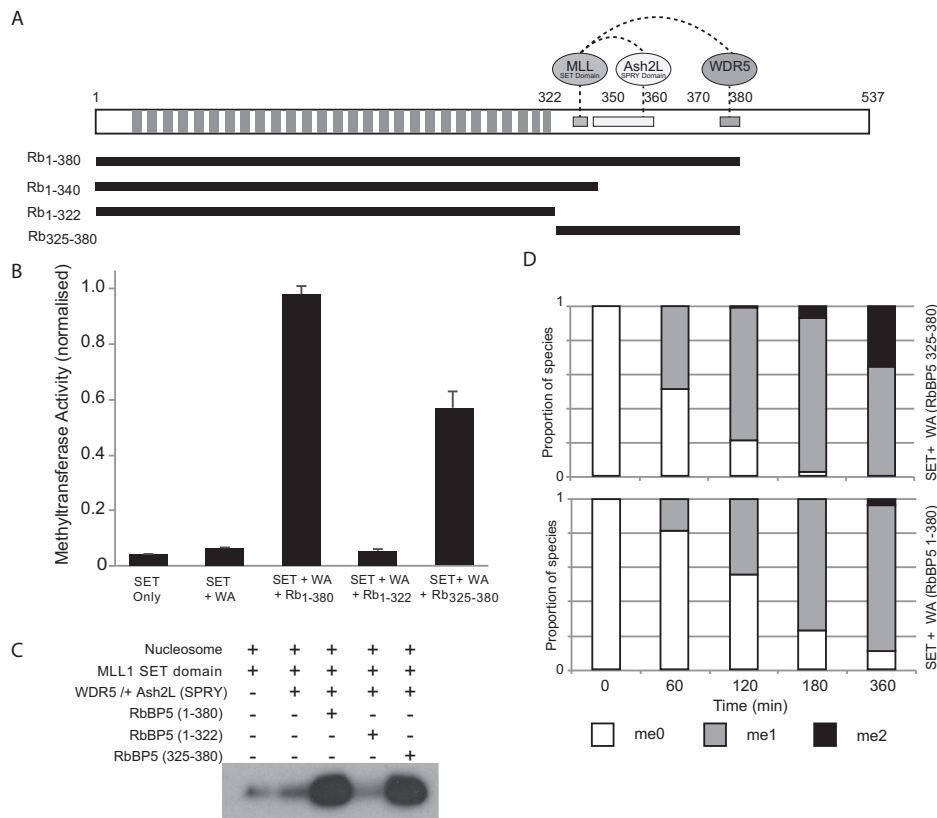
Figure 1. RbBP5 architecture and activity. (**A**) Schematic representation of the RbBP5 protein. The WD40 repeats are indicated along with the three regions shown to interact with the MLL1 SET domain, Ash2L SPRY domain and WDR5. The bars beneath represent the constructs used in the current study. (**B**) Normalized *in vitro* methyltransferase assay with H3 peptide substrate for the MLL1 (SET domain) + WDR5 + Ash2L (SPRY domain) with different RbBP5 constructs showing that only the RbBP5 340–380 region is essential for activity. (**C**) *In vitro* methyltransferase assay with recombinant mononucleosome substrate (147 bp DNA), using a histone H3K4 monomethyl antibody to follow activity. The loading control gel, stained with Coomassie Blue, is presented in Supplementary Figure S1. (**D**) Methyltransferase assay with MLL1 (SET domain) + WDR5 + Ash2L (SPRY domain) with RbBP5 construct plus and minus β propeller analysed by MALDI TOF mass spectrometry.

Methyltransferase assays

Methyltransferase assays were performed using peptide substrates based on the histone H3 amino terminal sequence (ARTKQTARKSTGGKAPR-Y). For end-point assays, H3 peptide concentration was 1 mM peptide, 0.5 mM SAM (including 0.625 mM 3H SAM [PerkinElmer]) in an assay buffer of 50 mM HEPES (pH 8), 200 mM NaCl and 0.5 mM TCEP. Following separation of the peptide from cofactor by C18 cartridge purification, the incorporation of 3H-labeled SAM into the peptide was estimated by scintillation counting, as previously described (32). Methyltransferase assays were carried out at 30°C for 60 min with a final enzyme concentration of 10 μM.

Methyltransferase assays with nucleosome substrate were performed in a buffer containing 20 mM HEPES, pH 7.8, 200 mM NaCl, 2 mM dithiothreitol (DTT). Final reagent concentrations were 5 μM MLL1 SET domain construct (containing the WDR5 interacting motif), WDR5, Ash2L and RbBP5 constructs, 100 μM SAM and recombinant mononucleosome substrate with 147 bp containing the 601 positioning sequence, prepared by the salt dialysis method (33). Reactions were incubated at 30°C for 2.5 h and stopped by addition of sodium dodecyl sulphate sample buffer. Following SDS-polyacrylamide gel electrophoresis protein was transferred to Immobilon-P[SQ] transfer membrane (Merck Millipore Ltd.) and blocked with 5% milk solution. Membrane was probed with histone H3K4 monomethyl-specific antibody (ab176877), (Abcam, Cambridge, UK) at 1:10000 dilution in phosphate buffered saline buffer with 0.05% Tween. The secondary antibody was goat anti-rabbit conjugated to HRP at 1:5000 dilution (Promega) and signal detected by autoradiography following incubation with ECL reagent (Merck Millipore Ltd).

MALDI-TOF methyltransferase assays were performed to examine the reaction products of the methyltransferase reaction with the H3 peptide at specific time points as described previously (34). The reaction mixture contained 10 μM enzyme (MLL1 SET domain containing Win motif, WDR5, Ash2L_SPRY and RbBP5 1–380 or 325–380), 200 μM SAM and 50 μM unmodified H3 peptide at 30°C for up to 6 h. At various time points, aliquots of the reaction were quenched by the addition of an equal volume of 1% trifluoroacetic acid. Samples were diluted in a 1:5 ratio with α-cyano-4-hydroxycinnamic acid and analyzed on a Bruker AutoFlex mass spectrometer (Bruker) in reflectron mode. The reactions were performed in triplicate, and the proportion of methyl species at each time point calculated by combining these multiple measurements.

Binding measurements

HPLC purified 14-mer DNA/RNA (GCATAGGTTCGA TC-6FAM/GCAUAGGUUCGAUC-6FAM) labeled with 6-FAM fluorescein molecules were purchased from Sigma. Equilibrium dissociation constants for the interaction of RbBP5 (P) with FAM–RNA (L) to form the complex RbBP5:RNA (PL) were determined using anisotropy titrations. For any mixture of RbBP5 and FAM–RNA the observed anisotropy (r_{OBS}) is given by:

$$r_{OBS} = \frac{\alpha r_{PL}[PL] + r_L[L]}{\alpha[PL] + [L]} \qquad (1)$$

where r_{PL} and r_L are the anisotropies of the complex and the free nucleic acid, and α (= 1 in this case) is the fluorescence intensity of the complex divided by that of the free nucleic acid. The data were analyzed using non-linear least-squares fits to equation (1) with r_{PL}, r_L and the dissociation constant ($K_{d,L}$) as variables, and PL and L calculated in the usual way:

$$[PL] = \frac{(K_{d,L}+[P_0]+[L_0])-\sqrt{(K_{d,L}+[P_0]+[L_0])^2-4[P_0][L_0]}}{2} \;\&\; [L] = [L_0] - [PL]$$

where the subscript 0 indicates the total concentration of the species.

All anisotropy titrations were performed at 20°C using a Jasco FP-8500 fluorimeter. The buffer (50 mM Tris pH 7.4, 150 mM NaCl, 0.5 mM DTT). Measurements with dsRNA and dsDNA were performed in identical conditions. Measurements with RbBP5 mutants were performed as above with dsRNA.

RESULTS

The RbBP5 β-propeller domain is not essential for catalysis

The RbBP5 region covering residues 340–380, immediately C-terminal to the β-propeller domain, has been shown to bind to other components of the catalytically active MLL complex and to be essential for methylation (11). We were first interested to determine if the β-propeller domain contributes to the activity of the assembled complex. We therefore compared the *in vitro* methyltransferase activity with H3 peptide substrate for the MLL1 SET domain reconstituted with Ash2L/WDR5 and different RbBP5 constructs (Figure 1). As expected the MLL1 SET domain alone exhibited negligible methyltransferase activity, which only modestly increased with addition of WDR5/Ash2L (SPRY) (Figure 1B). However, the further addition of RbBP5$_{1-380}$ strongly stimulated methylation confirming the integral role that this protein has in assembly of the catalytic core. In line with the observation of Li *et al.* (11), a RbBP5$_{325-380}$ construct also significantly stimulated activity, confirming that the elements responsible for assembling the complex and stabilizing the active SET domain conformation, are located in this region. However, interestingly, we observed that the RbBP5$_{1-322}$ construct, in combination with WDR5/Ash2L (SPRY), did not stimulate the overall activity of MLL1, indicating that the β-propeller domain does not contribute significantly to the catalytic activation mechanism. The slightly higher activity observed for

RbBP5$_{1-380}$ compared to RbBP5$_{325-380}$ may arise due to better construct stability or improved complex integrity. The equivalent methyltransferase assay, but using a recombinant mononucleosome (147 bp DNA) substrate, showed a similar pattern of MLL activation, i.e. activity was dependent on the presence of the RbBP5$_{325-380}$ region, but the β-propeller was not required (Figure 1C). This suggests that the β-propeller is neither essential for assembly of the catalytically active complex or for recognition of core nucleosome.

The potency of histone methylation as an epigenetic signaling mark lies in the high degree of specificity that recognition domains exhibit for the number of methyl groups on the methylated lysine. We were therefore interested to determine if the presence of the β-propeller domain had an effect on the final reaction product. We compared the product of the methyltransferase reaction with peptide substrate using MALDI TOF analysis for the MLL1 (SET)/WDR5/Ash2L (SPRY) complex with RbBP5$_{1-380}$ or RbBP5$_{325-380}$ constructs. In both cases the H3K4$_{me1}$ and to a lesser extent H3K4$_{me2}$ product were readily detected, consistent with the previously reported activity of the reconstituted complex with peptide substrate (10,34) (Figure 1D). The amount of me$_2$ product produced by the RbBP5$_{325-380}$ construct was slightly lower than the longer construct, which may be attributed to a modest effect on the overall stability of the complex, but these data indicate that the RbBP5 β-propeller does not significantly contribute to catalytic activity and suggests that it has a different role.

Structure of the RbBP5 β-propeller

To obtain a better understanding of the characteristics of the WD40-repeat rich region, we determined the crystal structure of mouse RbBP5. After screening a wide range of constructs, diffracting crystals of the RbBP5$_{1-380}$ construct were obtained, which contains both the β-propeller and the regions known to interact with other members of the MLL complex. The structure was determined by molecular replacement, for details see 'Materials and Methods' section, and refined to a resolution of 2.5Å. The asymmetric unit contains two copies of the construct (Figure 2A), and data collection and refinement statistics are presented in Table 1. The β-propeller domain structure is basically consistent between both copies, but the 325–380 region is largely disordered and varies significantly. Differences arise both in the course it takes with respect to the propeller, and to the number of residues which could be built into interpretable electron density. This can be interpreted as indicative of an unstructured region, but one that adopts a conformation induced by the interaction with its binding partners when assembled into the complex (11,22,24). The two copies of the β-propeller are linked in the crystal structure due to a strand swap which disrupts the seventh blade of the copy B propeller. The copy A residues 332–340 pack as a parallel β-strand against the copy B strand, displacing the N-terminal residues 18–22, which completes the propeller blade through the canonical interaction. This induces the third strand in the blade to extend by four residues to Val 326. This organisation is most likely a crystallographic artefact, notably we observed no dimerization in solution either
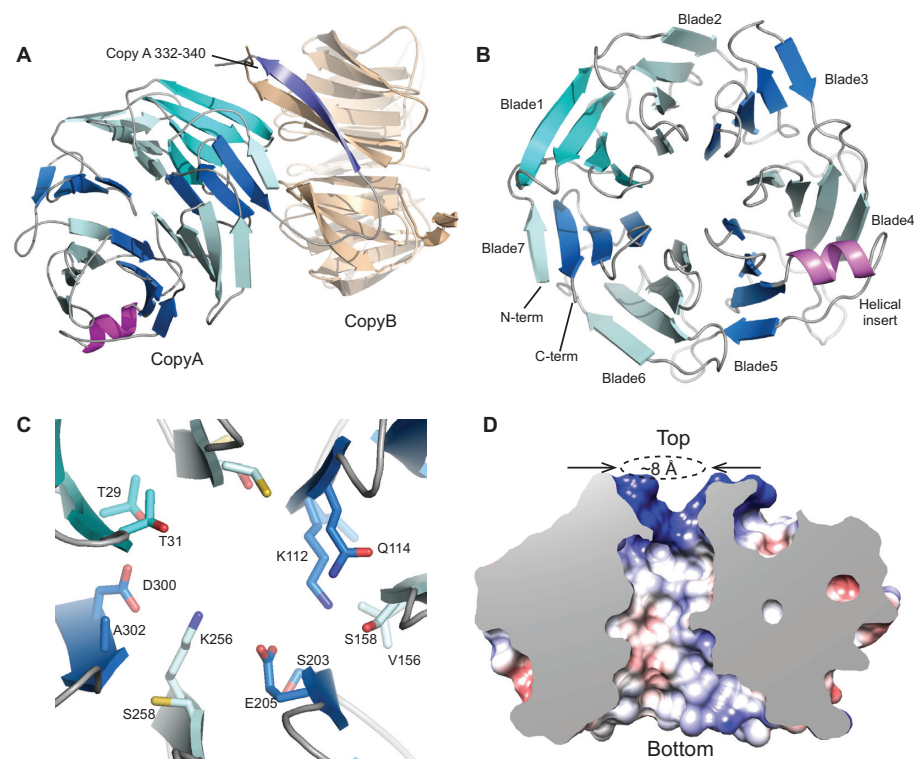
Figure 2. Structure of the RbBP5 β-propeller domain. (**A**) Cartoon representation of the RbBP5 asymmetric unit. The blades of the Copy A β-propeller are colored in alternating shades of blue. Copy B is colored in beige. The Copy A residues 332–340 displace the N-terminal strand in Copy B to form a crystallographic dimer. (**B**) Top view of the RbBP5 β-propeller showing a canonical arrangement of seven 4-stranded blades. A helical insert (pink) is located at the interface of blades 4 and 5. (**C**) The RbBP5 central channel is formed from largely polar residues. (**D**) Cut-through of RbBP5 (electrostatic surface representation) revealing the wide polar channel.

during purification or by Multi-Angle Laser Light Scattering (Supplementary Figure S3) analysis. In the analysis below we focus on the β-propeller of copy A, which represents the physiological form of the propeller.

RbBP5 forms a canonical seven bladed WD40 repeat β-propeller in which the seventh blade is formed from three anti-parallel strands from the C-terminal region of the domain and is completed with a strand originating from the N-terminus of the domain (35) (Figure 2B). Between the third and fourth strands of the 5th propeller blade there is an insert which includes a short α-helix, which lies on one face of the propeller toward the rim (Figure 2B). Helical inserts are common in WD40 repeat proteins, and have been observed in other chromatin associated β-propellers, such as RbAp48 and embryonic ectoderm development (EED), where they are associated with protein interactions. For example, in RbAp48 the helical insert on the side of the propeller interacts with Histone H4 (36,37).

A conspicuous feature of the RbBP5 β-propeller, which appears to be quite distinct from other WD40 proteins in chromatin modification complexes such as WDR5 in MLL and EED and RbAp48 in PRC2, is that the central axis is not formed from bulky aromatic side chains, but instead polar residues dominate the propeller axis (Figure 2C). This creates a much more open solvent filled channel of about 8Å diameter that links the two faces (Figure 2D). The top

Table 1. Data collection and refinement statistics

	5OV3
Data collection	
Wavelength (Å)	0.9795
Space group	$P2_12_12_1$
Cell dimensions	
a, b, c (Å)	56.81, 71.9, 178.27
α, β, γ (°)	90.00, 90.00, 90.00
Resolution (Å)	71.90–2.45 (2.58–2.45)
R_{merge}	0.10 (1.26)
R_{pim}	0.043 (0.59)
$I / \sigma I$	12.7 (1.6)
$CC_{1/2}$ (%)	99.8 (64.9)
Completeness (%)	99.8(99.2)
Redundancy	6.3 (5.4)
Refinement	
Resolution (Å)	66.7–2.45
No. reflections	27559
R_{work} / R_{free}	0.196/0.271
No. atoms	
Protein	5093
Ligand (PEGs)	23
B-factors	
Protein	59.3
R.m.s. deviations	
Bond lengths (Å)	0.009
Bond angles (°)	1.041
Ramachandran Outliers	1.39 %
Rotamer Outliers	2.3 %

channel edge is strongly basic, and at the 'bottom' the channel flares out creating a concave feature on that face of the propeller. Potentially these patches could mediate targeting interactions critical to the function of the MLL complex.

Both faces of the RbBP5 β-propeller feature prominent basic patches (Figure 3A and B). On the 'top' face, defined as the face containing the α-helical insert, the propeller axis has a complete ring of basic side-chains, which extends into a basic groove between blades six and seven, and so reaches toward the edge of the propeller. These features arise from Arginine and Lysine side-chains, with the central ring feature consisting of residues Arg34, 76, 118, 161, 162, 208 and Lys209 located on six of the seven blades (Figure 3C). These residues are distributed across the primary sequence of the domain and do not form a discrete motif; the residues are conserved in the RbBP5 homologs from higher organisms, only partially in *Neurospora*, but not at all in yeast species (Figure 3C and E; Supplementary Figure S2). This suggests that this feature may have evolved in multi-cellular organisms, which is suggestive of a potential role in cell lineage targeting, perhaps during developmental processes. The bottom face of the β-propeller is also characterized by a basic patch (Figure 3B), which extends from one side of the large cavity that surrounds the central axis to the edge of the face (Figure 3D). These residues, Arg220, 251 272, and 294, and Lys 255, which form the basic batch on the bottom surface are more broadly conserved across species, (Figure 3F and Supplemental Figure S2). In general, the bottom face of the propeller is better conserved than the top (Figure 3E and F), which may indicate that it could mediate an interaction that is fundamental to basic function.

Overall, the principal feature of the RbBP5 β-propeller surface, is that it is predominantly polar, both at the open central axis and the basic patches that characterize both top and bottom faces of the propeller. We were therefore interested to identify binding partners for this domain.

RbBP5 binds nucleic acids using an Arginine-rich surface

Given that the RbBP5 β-propeller does not have a role in modulating the activity of MLL1 we hypothesized that this domain is most likely a platform for binding to unidentified partners involved in targeting the complex to chromatin features or to specific genes. We explored whether the domain binds to histone tails using a biotin tagged RbBP5 β-propeller construct to probe a histone peptide array containing duplicates of 384 of the common epigenetic modifications (Active Motif), but no hits were identified using this strategy (data not shown). Next, we reasoned that given the basic characteristic of the features identified on the surface of the β-propeller, and the increasing evidence of involvement of long non-coding RNA in targeting MLL1 (21,38), it was plausible that the domain may recognize nucleic acid rather than a protein partner.

To test RbBP5′s capability to bind nucleic acids we used an NMR spectroscopy approach. NMR can deliver information on protein-ligand interactions covering a broad range of affinities and is ideal for such a *de novo* exploration. The $^1H^{15}N$ HSQC spectrum of the RbBP5$_{1–340}$ construct is of a good quality considering the protein's size. The resonances are well dispersed and the peaks of relatively uni-

form intensity confirm that the protein is folded, soluble and monomeric (Figure 4A). As no nucleic acid target of the protein is known, we tested RbBP5 binding to both ss-RNA, dsRNA and to dsDNA. We first titrated the protein with a pool of short 7-nucleobase RNA oligos with random sequence (7N-RNA). Titration of 7N-RNA shifted a subset of resonances, indicating an interaction is taking place with a discrete selection of RbBP5 residues (Figure 4B). The resonances are in fast exchange on a chemical shift timescale, which is consistent with a dissociation constant in the micromolar range. Further, the small number of peaks changing chemical shift indicates that the contacts with backbone moieties are not extensive and that it is likely that no major structural rearrangement is taking place in the RbBP5 upon associating with RNA. This is consistent with a surface feature such as those formed by the arginine side-chains.

Next, we examined whether the interaction with ssRNA is sequence specific using 'SIA' (31). This method provides the nucleobase preference of a protein domain for the different positions of the bound RNA, and performs best at intermediate to low affinity and with transient protein–RNA interactions. SIA was used to test the preference for each nucleobase at five consecutive positions of a bound oligonucleotide (Figure 4C). The SIA results indicate that RbBP5 has relatively little sequence specificity, with a very modest preference for G and U. Consistently, an SIA optimized RNA oligo (UAGGUUC) based on this preference bound only marginally stronger to the RbBP5 β-propeller than the randomized oligo pool (Figure 4B and D).

Given that only a small number of RbBP5 resonance shifts were observed upon addition of RNA and that the nucleobase preference exhibited was very limited, we wondered whether the RbBP5–RNA interaction was mediated with low specificity by a subset of the Arginine side chains on the β-propeller surface. In order to test this, we recorded $^1H^{15}N$ HSQC experiments optimized to capture changes in the resonances of the Arginine side chain resonances (Figure 5A). RNA titration results in the shift of a few of the existing resonances and in the appearance of a number of new resonances. The strong effect of the RNA on the linewidth of selected Arginine side chains is consistent with the RNA interacting with the charged side chains of a group of Arginines on the RbBP5 surface and slowing the exchange of the ^{15}N attached protons with the bulk water.

Next, we tested whether RbBP5 preferred DNA to RNA or if a structural preference exists for ssRNA versus dsRNA. We recorded Arginine-optimized NMR experiments using 14-mer ssRNA, dsRNA and dsDNA molecules of the same sequence (except uracil were changed for thymine in DNA) and length to titrate against RbBP5$_{1–340}$ (Figure 5B). The effect was strongest when titrating dsRNA and weakest when titrating dsDNA, with the ssRNA having an intermediate effect. To determine whether this effect can be correlated to a difference in binding affinity, we performed a fluorescence anisotropy assay to measure equilibrium dissociation constants for the binding of 6-carboxy fluorescein labeled 14-mer oligos of ssRNA, dsRNA or dsDNA of identical sequence (Figure 5C). The anisotropy measurements confirmed binding of all three nucleic acid species, consistent with the NMR titrations. Further, also consistent with the NMR titration RbBP5 showed a mod-
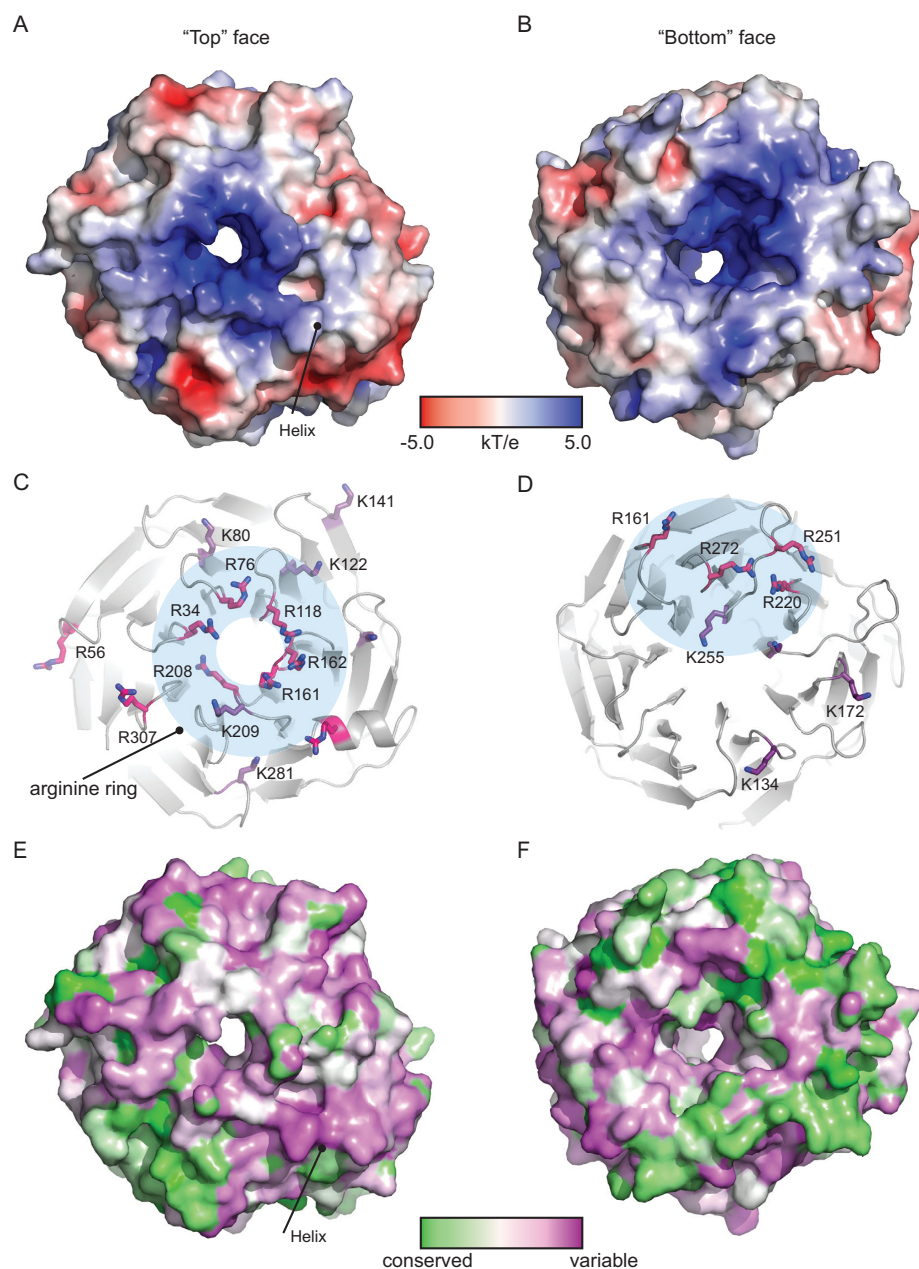
Figure 3. Surface features of RbBP5. (**A**, **C** and **E**) Views of the 'top' of the RbBP5 β-propeller, showing (**A**) Electrostatic surface of the 'top' face of the β-propeller. The position of the helical insert is indicated. (**C**) Cartoon representation of the 'top' face of the β-propeller. The Arginine (magenta) and Lysine (purple) side chains are shown, with the Arginine ring feature indicated by a blue ring. (**E**) Sequence conservation mapped onto the top face of the β-propeller (Gradient: green (conserved) to magenta (variable)). (**B**, **D** and **F**) Views of the 'bottom' surface of the propeller; (**B**) Electrostatic surface of the 'bottom' face of the β-propeller. (**D**) Bottom view of the propeller showing the Arginine (magenta) and lysine (purple) side chains. (**F**) Sequence conservation mapped onto the 'bottom' face of the β-propeller.

est preference for double stranded RNA (13.2 μM), over single stranded RNA (22.5 μM) and double stranded DNA (29.7 μM). Overall our data indicate that RbBP5 recognize dsRNA with moderate affinity and structure specificity and confirm that the interaction is mediated by Arginine side chains.

Given that there are basic patches on both faces of the β-propeller, we were interested if nucleic acid binding was limited to either face. Two candidate side chains were selected from each patch and mutated to acidic residues. The

fluorescence anisotropy binding measurement was repeated with dsRNA for wild-type β-propeller and the four mutant proteins (Figure 5D). Only modest changes in affinity were observed for the single point mutants, which is consistent with a broad interface. Both upper face mutants (R34E and R162E) exhibited a greater than 2-fold loss of affinity, whereas only one of the bottom face mutants (R220E) differed significantly from the wild-type. This limited evidence favors a model in which the upper Arginine ring promotes the interaction, but does not preclude the possibility that
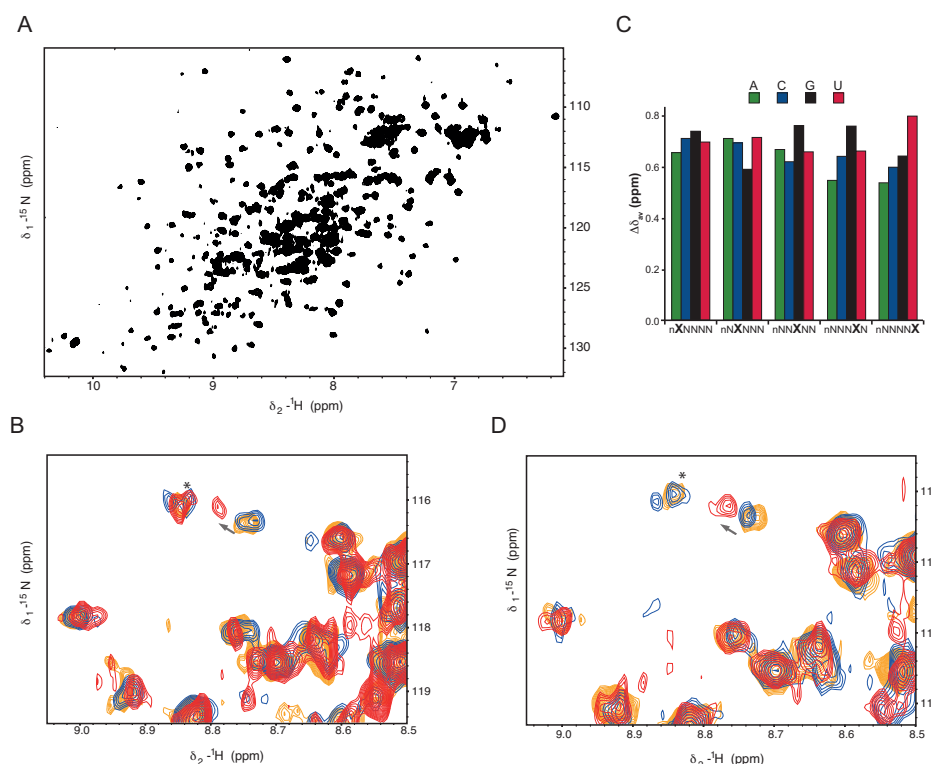
Figure 4. Analysis of RNA binding to the RbBP5 β-propeller. (**A**) $^1H^{15}N$ HSQC spectrum of the RbBP5$_{1–340}$ construct showing well-dispersed resonances. (**B**) Titration of a randomized RNA 7-mer oligo nucleotide induces selective resonance shifts in the $^1H^{15}N$ HSQC spectrum. Blue apo RbBP5, orange 1:1, red 4:1 RNA:protein. (**C**) Scaffold Independent Analysis showing the preference of RbBP5 for each nucleobase at five consecutive positions in the oligonucleotide. (**D**) Titration of a sequence optimised RNA 7-mer induces selective resonance shifts in the $^1H^{15}N$ HSQC NMR spectrum of RbBP5.

DISCUSSION

The MLL histone methylation complexes have two distinct functions; one is the targeting function, and this is the role of the major part of the MLL proteins themselves, as they are essentially a series of recognition domains linked by regions of predicted unstructured sequence. The other is the catalytic function, comprising the complex formed by the extreme C-terminal region of the MLLs and the four WRAD complex proteins (9). The role of the RbBP5 region (residues 330–380) in organizing the activated complex structure by linking Ash2L, WDR5 and the SET domain is well described (11,22,24,39). Our *in vitro* methylation data show that the main structural element of RbBP5, the β-propeller domain, does not play an integral role in MLL complex assembly or catalytic activation. This suggests that the RbBP5 β-propeller may have a role in targeting the MLL/WRAD activity or even a function outside of the MLL complex.

The other β-propeller protein in the WRAD complex, WDR5, has been shown to be a component of multi-protein complexes independently of the other WRAD proteins or MLL (40,41). For example, WDR5 is an integral component of the histone acetyltransferase complexes NSL/MOF where it binds to the NSL1 subunit (42) and the ATAC

complex (43). WDR5 is ubiquitously expressed and its high abundance in the cytoplasm also suggests that it has roles beyond chromatin modification complexes (44). There are fewer examples suggesting that RbBP5 has such a varied function. However, recent proteomic studies point to several chromatin associated proteins that may potentially participate in RbBP5 interactions in an MLL context, including Chromo Helicase Domain 8 (45), Cfp1 (46), OGT (47) and Cul4A-DDB1 (48), but further studies will be required to determine if these are direct contacts.

Given that it does not appear to have a direct role in formation of the catalytic core, the most likely function of the RbBP5 β-propeller could be in recruitment of MLL to specific sites via modified chromatin. However, using a histone modification peptide array consisting of a comprehensive range of typical posttranslational modifications associated with epigenetic signaling, we were unable to identify any hits. It should be noted that this does not preclude that RbBP5 could bind to histone tails in the context of an extended nucleosome array with linker DNA or to an untested combination of modifications. However, the prominent features on the RbBP5 surface are basic patches, and it is unlikely that the generally basic histone tails are interaction partners.

Arginine ring or Arginine patch features occur in a range of proteins, and in a number of cases have been shown to interact with nucleic acid rather than proteins (49,50). For example, the DDB2 β-propeller has a surface patch of basic residues that bind to DNA through contacts with the
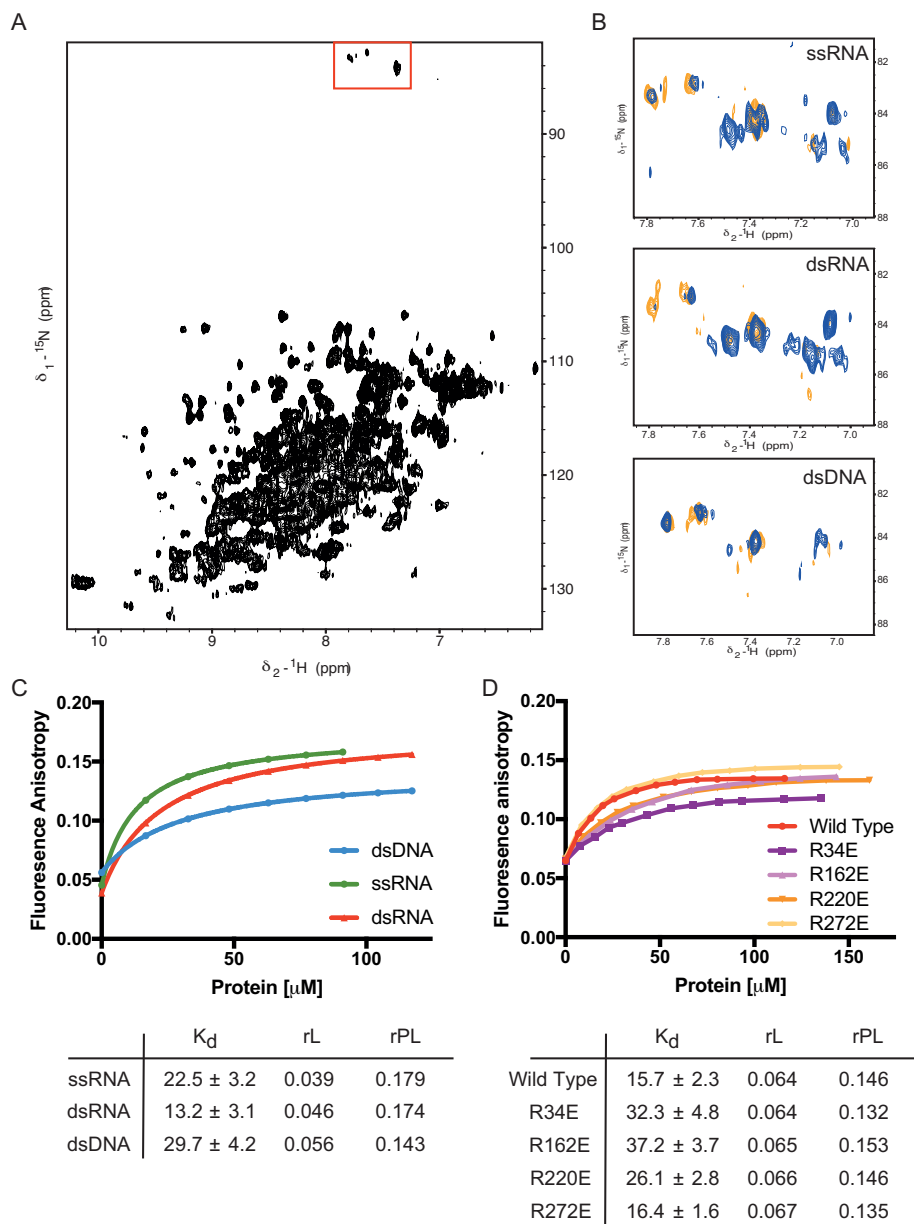
Figure 5. RbBP5 binding to nucleic acids. (**A**) ^{1}H^{15}N HSQC spectrum of RbBP5$_{1-340}$ construct optimized for the detection of arginine side chain moieties, indicated by red box at top. (**B**) Spectrum of RbBP5 free and titrated with ssRNA, dsRNA and dsDNA (orange apo RbBP5, blue 4:1 NA:protein). (**C**) Fluorescence anisotropy measurements of RbBP5$_{1-340}$ binding to fluorescein labeled nucleic acid. (**D**) Fluorescence anisotropy measurements of acidic mutants of RbBP5$_{1-340}$ binding to fluorescein labeled dsRNA.

phosphate backbone (51). This encouraged us to investigate if we could detect direct binding of RNA to the RbBP5 β-propeller. Our NMR and fluorescence anisotropy measurements show that RbBP5 binds to nucleic acid via selected Arginine side chains, consistent with the binding taking place in the large positive patch that surrounds one end of the domains channel. Further, the experiments show a modest preference of the protein for double stranded RNA over single stranded RNA. This would be consistent with a lncRNA, which are often structured, but as only modest variations in affinity were observed with the different types of nucleic acid, more experiments will be required to identify the physiological binding partner or partners.

There is growing evidence that at least part of the targeting function of chromatin regulatory protein complexes such as trithorax and polycomb group during development is mediated by lncRNAs (52,53). The list of lncRNAs with a putative role in MLL mediated H3K4 methylation is expanding, and examples include HOTTIP, Mistral, HoxBlinc and SnoRNA/116HG (20,54–57). Although the site of interaction of many of these lncRNAs is unknown, WDR5 has been implicated in some interactions (21). However, there is some evidence supporting direct interaction between RbBP5 and RNA. For example, depletion of SnoRNA/116HG, a lncRNA, is linked to Prader-Willi syndrome. In a co-precipitation experiment with RbBP5 used

as bait protein *116HG* RNA was precipitated in conditions that favored formation of RNA secondary structure (18). RbBP5 has been identified as an RNA interacting protein in two recent large-scale screens for novel RNA binding proteins (58,59). In one experiment RNA co-precipitated with a peptide that corresponded to RbBP5$_{61-76}$. Notably, this region incorporates one of the Arginines associated with one of the basic patches identified in our structural analysis.

Our structural analysis has revealed that the RbBP5 β-propeller has a feature rich surface that is likely to mediate as yet unknown binding interactions. It is also likely that, like its WRAD complex partner WDR5, the RbBP5 β-propeller may mediate interactions with multiple partners in different contexts. However, our binding analysis leaves open the possibility that lncRNA is one potential partner.

DATA AVAILABILITY

The coordinates and structure factors for the structure have been deposited in the PDB under accession code 5OV3.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

ACKNOWLEDGEMENTS

REFERENCES

1. Berger,S.L. (2007) The complex language of chromatin regulation during transcription. *Nature*, **447**, 407–412.
2. Kouzarides,T. (2007) Chromatin modifications and their function. *Cell*, **128**, 693–705.
3. Schuettengruber,B., Chourrout,D., Vervoort,M., Leblanc,B. and Cavalli,G. (2007) Genome regulation by polycomb and trithorax proteins. *Cell*, **128**, 735–745.
4. Bannister,A.J. and Kouzarides,T. (2011) Regulation of chromatin by histone modifications. *Cell Res.*, **21**, 381–395.
5. Ringrose,L. and Paro,R. (2004) Epigenetic regulation of cellular memory by the Polycomb and Trithorax group proteins. *Annu. Rev. Genet.*, **38**, 413–443.
6. Schwartz,Y.B. and Pirrotta,V. (2008) Polycomb complexes and epigenetic states. *Curr. Opin. Cell Biol.*, **20**, 266–273.
7. Miller,T., Krogan,N.J., Dover,J., Erdjument-Bromage,H., Tempst,P., Johnston,M., Greenblatt,J.F. and Shilatifard,A. (2001) COMPASS: a complex of proteins associated with a trithorax-related SET domain protein. *Proc. Natl. Acad. Sci. U.S.A.*, **98**, 12902–12907.
8. Dou,Y., Milne,T.A., Ruthenburg,A.J., Lee,S., Lee,J.W., Verdine,G.L., Allis,C.D. and Roeder,R.G. (2006) Regulation of MLL1 H3K4 methyltransferase activity by its core components. *Nat. Struct. Mol. Biol.*, **13**, 713–719.
9. Ernst,P. and Vakoc,C.R. (2012) WRAD: enabler of the SET1-family of H3K4 methyltransferases. *Brief. Funct. Genomics*, **11**, 217–226.
10. Southall,S.M., Wong,P.S., Odho,Z., Roe,S.M. and 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.
11. Li,Y., Han,J., Zhang,Y., Cao,F., Liu,Z., Li,S., Wu,J., Hu,C., Wang,Y., Shuai,J. *et al.* (2016) Structural basis for activity regulation of MLL family methyltransferases. *Nature*, **530**, 447–452.
12. Jiang,H., Shukla,A., Wang,X., Chen,W.Y., Bernstein,B.E. and Roeder,R.G. (2011) Role for Dpy-30 in ES cell-fate specification by regulation of H3K4 methylation within bivalent domains. *Cell*, **144**, 513–525.
13. Swygert,S.G. and Peterson,C.L. (2014) Chromatin dynamics: interplay between remodeling enzymes and histone modifications. *Biochim. Biophys. Acta*, **1839**, 728–736.
14. Wang,Z., Song,J., Milne,T.A., Wang,G.G., Li,H., Allis,C.D. and Patel,D.J. (2010) Pro isomerization in MLL1 PHD3-bromo cassette connects H3K4me readout to CyP33 and HDAC-mediated repression. *Cell*, **141**, 1183–1194.
15. Zhang,P., Lee,H., Brunzelle,J.S. and Couture,J.F. (2012) The plasticity of WDR5 peptide-binding cleft enables the binding of the SET1 family of histone methyltransferases. *Nucleic Acids Res.*, **40**, 4237–4246.
16. Rampalli,S., Li,L., Mak,E., Ge,K., Brand,M., Tapscott,S.J. and Dilworth,F.J. (2007) p38 MAPK signaling regulates recruitment of Ash2L-containing methyltransferase complexes to specific genes during differentiation. *Nat. Struct. Mol. Biol.*, **14**, 1150–1156.
17. Yang,Z., Augustin,J., Hu,J. and Jiang,H. (2015) Physical interactions and functional coordination between the core subunits of Set1/Mll complexes and the reprogramming factors. *PLoS One*, **10**, e0145336.
18. Powell,W.T., Coulson,R.L., Crary,F.K., Wong,S.S., Ach,R.A., Tsang,P., Alice Yamada,N., Yasui,D.H. and Lasalle,J.M. (2013) A Prader-Willi locus lncRNA cloud modulates diurnal genes and energy expenditure. *Hum. Mol. Genet.*, **22**, 4318–4328.
19. He,X., Chen,X., Zhang,X., Duan,X., Pan,T., Hu,Q., Zhang,Y., Zhong,F., Liu,J., Zhang,H. *et al.* (2015) An Lnc RNA (GAS5)/SnoRNA-derived piRNA induces activation of TRAIL gene by site-specifically recruiting MLL/COMPASS-like complexes. *Nucleic Acids Res.*, **43**, 3712–3725.
20. Deng,C., Li,Y., Zhou,L., Cho,J., Patel,B., Terada,N., Li,Y., Bungert,J., Qiu,Y. and Huang,S. (2016) HoxBlinc RNA recruits Set1/MLL complexes to activate Hox gene expression patterns and mesoderm lineage development. *Cell Rep.*, **14**, 103–114.
21. Yang,Y.W., Flynn,R.A., Chen,Y., Qu,K., Wan,B., Wang,K.C., Lei,M. and Chang,H.Y. (2014) Essential role of lncRNA binding for WDR5 maintenance of active chromatin and embryonic stem cell pluripotency. *Elife*, **3**, e02046.
22. Zhang,P., Chaturvedi,C.P., Tremblay,V., Cramet,M., Brunzelle,J.S., Skiniotis,G., Brand,M., Shilatifard,A. and Couture,J.F. (2015) A phosphorylation switch on RbBP5 regulates histone H3 Lys4 methylation. *Genes Dev.*, **29**, 123–128.
23. Patel,A., Dharmarajan,V. and Cosgrove,M.S. (2008) Structure of WDR5 bound to mixed lineage leukemia protein-1 peptide. *J. Biol. Chem.*, **283**, 32158–32161.
24. Odho,Z., Southall,S.M. and Wilson,J.R. (2010) Characterization of a novel WDR5-binding site that recruits RbBP5 through a conserved motif to enhance methylation of histone H3 lysine 4 by mixed lineage leukemia protein-1. *J. Biol. Chem.*, **285**, 32967–32976.
25. Bailey,S. (1994) The Ccp4 suite—programs for protein crystallography. *Acta Crystallogr. D Biol. Crystallogr.*, **50**, 760–763.
26. Emsley,P. and Cowtan,K. (2004) Coot: model-building tools for molecular graphics. *Acta Crystallogr. D Biol. Crystallogr.*, **60**, 2126–2132.
27. Murshudov,G.N., Skubak,P., Lebedev,A.A., Pannu,N.S., Steiner,R.A., Nicholls,R.A., Winn,M.D., Long,F. and Vagin,A.A.

(2011) REFMAC5 for the refinement of macromolecular crystal structures. *Acta Crystallogr. D. Biol. Crystallogr.*, **67**, 355–367.

28. Jurrus,E., Engel,D., Star,K., Monson,K., Brandi,J., Felberg,L.E., Brookes,D.H., Wilson,L., Chen,J., Liles,K. *et al.* (2018) Improvements to the APBS biomolecular solvation software suite. *Protein Sci.*, **27**, 112–128.

29. Dolinsky,T.J., Czodrowski,P., Li,H., Nielsen,J.E., Jensen,J.H., Klebe,G. and Baker,N.A. (2007) PDB2PQR: expanding and upgrading automated preparation of biomolecular structures for molecular simulations. *Nucleic Acids Res.*, **35**, W522–W525.

30. Pei,J. and Grishin,N.V. (2001) AL2CO: calculation of positional conservation in a protein sequence alignment. *Bioinformatics*, **17**, 700–712.

31. Beuth,B., Garcia-Mayoral,M.F., Taylor,I.A. and Ramos,A. (2007) Scaffold-independent analysis of RNA-protein interactions: the Nova-1 KH3-RNA complex. *J. Am. Chem. Soc.*, **129**, 10205–10210.

32. Xiao,B., Jing,C., Wilson,J.R., Walker,P.A., Vasisht,N., Kelly,G., Howell,S., Taylor,I.A., Blackburn,G.M. and Gamblin,S.J. (2003) Structure and catalytic mechanism of the human histone methyltransferase SET7/9. *Nature*, **421**, 652–656.

33. Luger,K., Rechsteiner,T.J. and Richmond,T.J. (1999) Preparation of nucleosome core particle from recombinant histones. *Methods Enzymol.*, **304**, 3–19.

34. Zhang,Y., Mittal,A., Reid,J., Reich,S., Gamblin,S.J. and Wilson,J.R. (2015) Evolving catalytic properties of the MLL family SET domain. *Structure*, **23**, 1921–1933.

35. Xu,C. and Min,J. (2011) Structure and function of WD40 domain proteins. *Protein Cell*, **2**, 202–214.

36. Murzina,N.V., Pei,X.Y., Zhang,W., Sparkes,M., Vicente-Garcia,J., Pratap,J.V., McLaughlin,S.H., Ben-Shahar,T.R., Verreault,A., Luisi,B.F. *et al.* (2008) Structural basis for the recognition of histone H4 by the histone-chaperone RbAp46. *Structure*, **16**, 1077–1085.

37. Song,J.J., Garlick,J.D. and Kingston,R.E. (2008) Structural basis of histone H4 recognition by p55. *Genes Dev.*, **22**, 1313–1318.

38. Wang,K.C., Yang,Y.W., Liu,B., Sanyal,A., Corces-Zimmerman,R., Chen,Y., Lajoie,B.R., Protacio,A., Flynn,R.A., Gupta,R.A. *et al.* (2011) A long noncoding RNA maintains active chromatin to coordinate homeotic gene expression. *Nature*, **472**, 120–124.

39. Avdic,V., Zhang,P., Lanouette,S., Groulx,A., Tremblay,V., Brunzelle,J. and Couture,J.F. (2011) Structural and biochemical insights into MLL1 core complex assembly. *Structure*, **19**, 101–108.

40. Trievel,R.C. and Shilatifard,A. (2009) WDR5, a complexed protein. *Nat. Struct. Mol. Biol.*, **16**, 678–680.

41. van Nuland,R., Smits,A.H., Pallaki,P., Jansen,P.W., Vermeulen,M. and Timmers,H.T. (2013) Quantitative dissection and stoichiometry determination of the human SET1/MLL histone methyltransferase complexes. *Mol. Cell. Biol.*, **33**, 2067–2077.

42. Dias,J., Van Nguyen,N., Georgiev,P., Gaub,A., Brettschneider,J., Cusack,S., Kadlec,J. and Akhtar,A. (2014) Structural analysis of the KANSL1/WDR5/KANSL2 complex reveals that WDR5 is required for efficient assembly and chromatin targeting of the NSL complex. *Genes Dev.*, **28**, 929–942.

43. Spedale,G., Timmers,H.T. and Pijnappel,W.W. (2012) ATAC-king the complexity of SAGA during evolution. *Genes Dev.*, **26**, 527–541.

44. Wang,Y.Y., Liu,L.J., Zhong,B., Liu,T.T., Li,Y., Yang,Y., Ran,Y., Li,S., Tien,P. and Shu,H.B. (2010) WDR5 is essential for assembly of the VISA-associated signaling complex and virus-triggered IRF3 and NF-kappaB activation. *Proc. Natl. Acad. Sci. U.S.A.*, **107**, 815–820.

45. Yates,J.A., Menon,T., Thompson,B.A. and Bochar,D.A. (2010) Regulation of HOXA2 gene expression by the ATP-dependent chromatin remodeling enzyme CHD8. *FEBS Lett.*, **584**, 689–693.

46. Tang,Z., Chen,W.Y., Shimada,M., Nguyen,U.T., Kim,J., Sun,X.J., Sengoku,T., McGinty,R.K., Fernandez,J.P., Muir,T.W. *et al.* (2013) SET1 and p300 act synergistically, through coupled histone modifications, in transcriptional activation by p53. *Cell*, **154**, 297–310.

47. Deplus,R., Delatte,B., Schwinn,M.K., Defrance,M., Mendez,J., Murphy,N., Dawson,M.A., Volkmar,M., Putmans,P., Calonne,E. *et al.* (2013) TET2 and TET3 regulate GlcNAcylation and H3K4 methylation through OGT and SET1/COMPASS. *EMBO J.*, **32**, 645–655.

48. Higa,L.A., Wu,M., Ye,T., Kobayashi,R., Sun,H. and Zhang,H. (2006) CUL4-DDB1 ubiquitin ligase interacts with multiple WD40-repeat proteins and regulates histone methylation. *Nat. Cell Biol.*, **8**, 1277–1283.

49. Neves,M.A., Yeager,M. and Abagyan,R. (2012) Unusual arginine formations in protein function and assembly: rings, strings, and stacks. *J. Phys. Chem. B*, **116**, 7006–7013.

50. Jarvelin,A.I., Noerenberg,M., Davis,I. and Castello,A. (2016) The new (dis)order in RNA regulation. *Cell Commun. Signal.*, **14**, 9.

51. Scrima,A., Konickova,R., Czyzewski,B.K., Kawasaki,Y., Jeffrey,P.D., Groisman,R., Nakatani,Y., Iwai,S., Pavletich,N.P. and Thoma,N.H. (2008) Structural basis of UV DNA-damage recognition by the DDB1-DDB2 complex. *Cell*, **135**, 1213–1223.

52. Vance,K.W. and Ponting,C.P. (2014) Transcriptional regulatory functions of nuclear long noncoding RNAs. *Trends Genet.*, **30**, 348–355.

53. Hiragami-Hamada,K. and Fischle,W. (2014) RNAs - physical and functional modulators of chromatin reader proteins. *Biochim. Biophys. Acta*, **1839**, 737–742.

54. Cruvinel,E., Budinetz,T., Germain,N., Chamberlain,S., Lalande,M. and Martins-Taylor,K. (2014) Reactivation of maternal SNORD116 cluster via SETDB1 knockdown in Prader-Willi syndrome iPSCs. *Hum. Mol. Genet.*, **23**, 4674–4685.

55. Bertani,S., Sauer,S., Bolotin,E. and Sauer,F. (2011) The noncoding RNA Mistral activates Hoxa6 and Hoxa7 expression and stem cell differentiation by recruiting MLL1 to chromatin. *Mol. Cell*, **43**, 1040–1046.

56. Cheng,Y., Jutooru,I., Chadalapaka,G., Corton,J.C. and Safe,S. (2015) The long non-coding RNA HOTTIP enhances pancreatic cancer cell proliferation, survival and migration. *Oncotarget*, **6**, 10840–10852.

57. Fang,K., Han,B.W., Chen,Z.H., Lin,K.Y., Zeng,C.W., Li,X.J., Li,J.H., Luo,X.Q. and Chen,Y.Q. (2014) A distinct set of long non-coding RNAs in childhood MLL-rearranged acute lymphoblastic leukemia: biology and epigenetic target. *Hum. Mol. Genet.*, **23**, 3278–3288.

58. He,C., Sidoli,S., Warneford-Thomson,R., Tatomer,D.C., Wilusz,J.E., Garcia,B.A. and Bonasio,R. (2016) High-resolution mapping of RNA-binding regions in the nuclear proteome of embryonic stem cells. *Mol. Cell*, **64**, 416–430.

59. Brannan,K.W., Jin,W., Huelga,S.C., Banks,C.A., Gilmore,J.M., Florens,L., Washburn,M.P., Van Nostrand,E.L., Pratt,G.A., Schwinn,M.K. *et al.* (2016) SONAR discovers RNA-binding proteins from analysis of large-scale protein-protein interactomes. *Mol. Cell*, **64**, 282–293.