npg

# Structure of the SPRY domain of human Ash2L and its interactions with RbBP5 and DPY30

Cell Research (2012) 22:598-602. doi:10.1038/cr.2012.9; published online 10 January 2012

## Dear Editor,

Histone H3 lysine 4 (H3K4) methylation is closely correlated to gene transcription activation [1]. In mammals, the MLL family (also called KMT2) histone methyltransferase, which includes MLL1 to 4, SET1A, and SET1B, are the principal enzymes that carry out the global methylation of H3K4 in vivo [1]. In contrast to other SET domain-containing histone methyltransferases, MLL family proteins exhibit very low intrinsic histone methyltransferease (HMT) activity by themselves. Their full activity is only achieved in the presence of WDR5, Ash2L, RbBP5, and DPY30, the common regulatory components of all MLL family complexes [2]. Ash2L (Absent, small or homeotic discs-like 2) is a trithorax group protein, and a critical regulator of all MLL complexes. Knockdown of Ash2L results in a global reduction of histone H3K4 trimethylation [3]. The role of Ash2L in methylation regulation is dependent on its two binding partners, RbBP5 and DPY30 [2, 4]. The interaction between Ash2L and RbBP5 is essential for MLL1 complex integrity and activity regulation [2, 5]. The Ash2L-DPY30 interaction was recently reported to be important for the regulation of nucleosomal H3K4 trimethylation and the differentiation potential of embryonic stem cells [6]. Ash2L protein adopts a modular configuration. From the N- to the Cterminus, it contains a PHD finger, a WH motif, an SPRY domain, and a DPY-30 binding-motif (DBM) (Figure 1A). The SPRY domain is the primary moiety in Ash2L that recognizes RbBP5, mediates MLL1 complex assembly, and regulates MLL1 enzymatic activity [5]. Here we report the crystal structure of SPRY domain of Ash2L, and dissect its interactions with RbBP5 and DPY30.

Bioinformatical and biochemical analyses suggested that the SPRY domain of Ash2L contains a large 44-residue unstructured loop (residues 402-445) in the middle of the domain (Supplementary information, Figure S1A, S1B and S1C). An Ash2L construct (Ash2L<sub>SPRY</sub>) with a deletion of this loop retains the same HMT activity as the full-length Ash2L (Supplementary information, Figure S1D). We determined the crystal structure of Ash2L<sub>SPRY</sub> using the single-wavelength anomalous dispersion method with mercury (MeHgAc)-derivative crystals, and refined to a resolution of 2.1 Å (Supplementary information, Table S1). The high-quality experimental electron density map enabled us to fit and refine the structure except for the C-terminal 13 residues (residues 511-523).

Ash2L<sub>SPRY</sub> adopts a tadpole-like conformation with three recognizable modules, an 11-residue N-terminal pre-SPRY motif, a central 200-residue globular fold that forms the core of Ash $2L_{SPRY}$ , followed by a 25-residue C-terminal post-SPRY motif (Figure 1B). Both the pre-SPRY and the post-SPRY motifs are short and composed of a  $3_{10}$  helix and either one or two  $\beta$  strands (Figure 1B). These two motifs form a two-stranded  $\beta$  sheet that protrudes from the globular SPRY domain, forming the tail of the tadpole-like structure (Figure 1B). The post-SPRY motif makes extensive contacts with one side of the  $\beta$ -sandwich of Ash2L<sub>SPRY</sub> and thus plays an important role in stabilizing the core of the SPRY domain (Supplementary information, Figure S2A). The central core of Ash $2L_{SPRY}$  adopts a distorted  $\beta$ -sandwich conformation, consisting of two layers of concave-shaped  $\beta$ -sheets (Figure 1B). All these  $\beta$  strands are arranged in an antiparallel configuration. The two  $\beta$ -sheet layers stack together mainly through Van der Waals contacts by a group of hydrophobic residues (Supplementary information, Figure S2B). These internal hydrophobic residues are evolutionarily conserved from yeast to humans (Supplementary information, Figure S3), suggesting that all the Ash2L homologs share the similar SPRY fold as the human Ash2L<sub>SPRy</sub>. The deleted 44-residue loop F' between strands  $\beta$ 11 and  $\beta$ 12, is away from the structural core, and thus is dispensable for the stability of the Ash $2L_{SPRY}$ domain (Figures 1B). Of note, sequence alignment shows that the yeast Bre2 protein, a distant homolog of human Ash2L, contains two large loops, a 40-residue loop between strands  $\beta 6$  and  $\beta 7$  and another 112-residue loop connecting  $\beta 11$  and  $\beta 12$  (Supplementary information, Figure S3).

The crystal structure of Ash2L<sub>SPRY</sub> closely resembles



other SPRY domains (Supplementary information, Figure S4). Ash $2L_{SPRY}$  has several unique features compared with other SPRY domains. Most notably, Ash $2L_{SPRY}$  contains both a pre- and a post-SPRY motifs and the latter is essential for the correct folding of the protein (Supplementary information, Figure S2A). In contrast, all other SPRY domains with known structures to date lack the post-SPRY motif. Secondly, the connecting loop regions have great variation in Ash $2L_{SPRY}$  compared with other SPRY domains (Supplementary information, Figure S3). These structural disparities in the loop regions are likely correlated with their different roles in specific protein recognition. Another unique feature of Ash2L is the dimerization of Ash $2L_{SPRY}$  found in the crystal struc-

ture (Supplementary information, Figure S5A and S5B). However, the affinity of the observed dimeric interface is too weak to stably hold two Ash $2L_{SPRY}$  molecules together *in vitro* (Supplementary information, Figure S5C and S5D). In addition, Ash2L dimerization appears to have no effect on the HMT activity of the MLL1 complex (Supplementary information, Figure S5E). Further investigation is needed to clarify the *in vivo* function of Ash $2L_{SPRY}$  dimerization.

Recent studies revealed that the SPRY domain of Ash2L recognizes a stretch of acidic residues of RbBP5 [5, 7]. We presumed that the interaction between Ash2L and RbBP5 is very likely mediated by electrostatic contacts. Consistent with this notion, our previous data Figure 1 Crystal structure of the SPRY domain of Ash2L. (A) Domain organization of human Ash2L, RbBP5 and DPY30. In Ash2L, the PHD domain is colored in blue, the WH domain in red, the SPRY domain in yellow, the disorder loop within the SPRY domain in pink, and the DBM (DPY30 binding motif) in tint: in RbBP5, the WD40-repeats domain in purple, the ABM (Ash2L binding motif) in orange, and the WBM (WDR5 binding motif) in magenta; in DPY30, the dimerization domain (DD) in green. Numerals indicate residue numbers at the boundaries of various subdivisions. Protein interactions are indicated with grey-shaded areas. (B) Two perpendicular views of the Ash2L<sub>SPRY</sub> structure. The pre-SPRY motif is colored in purple, the SPRY domain in yellow, and the post-SPRY motif in cyan. The N- and the C-terminus of the polypeptide are labeled. The blue-dashed line indicates the unstructured loop (Loop F') that was deleted in the Ash2L<sub>SPRY</sub> construct. Loop D is shown in orange. (C) The RbBP5-binding surface on Ash2L. Ash2L<sub>SPRY</sub> is in surface representation and colored according to its electrostatic potential (positive potential, blue; negative potential, red). Positively charged Ash2L residues that are important for RbBP5 binding are labeled. (D) Summary of ITC analysis of the interaction between Ash2L<sub>SPPY</sub> and its mutants with a RbBP5 peptide (residue 344-381) that contains the ABM of RbBP5 (nd: not detectable by ITC). (E) The RbBP5-Ash2L interaction is important for the HMT activity of MLL1. The HMT assay was performed in the presence of MLL1<sub>SFT</sub>, WDR5, RbBP5, and Ash-2L<sub>SPRY</sub> or its mutants at a concentration of 0.5 µM. The HMT activities with mutant Ash2L<sub>SPRY</sub> proteins are normalized to that of the wild-type Ash2L<sub>SPRY</sub>. Each assay was performed in triplicate and error bars represent standard deviation. (F) A short fragment of Ash2L (residues 510-523) C-terminal to the SPRY domain is important for DPY30 binding. GST-tagged DPY30 C-terminal dimerization domain was incubated with different fragments of Ash2L in the GST-pull-down assay. The left panel is 20% of the input samples and the right panel is the bound fraction. For better visualization of Ash2L<sub>510-523</sub> on SDS-PAGE, sumotagged peptides were used. As a control, sumo protein itself did dot bind to GST-DPY30<sub>DD</sub> (G) The model structure of Ash-2L<sub>SPRY-DBM</sub>-DPY30<sub>DD</sub> based on the PKA-RIIa/D-AKAP2 complex structure (left panel). Ash2L<sub>SPRY</sub> is colored in yellow, Ash2L<sub>DBM</sub> in orange, and two DPY30<sub>DD</sub> protomers in magenta and cyan, respectively. The interface between Ash2L<sub>DBM</sub> and DPY30<sub>DD</sub> is highlighted in the right panel. Ash2L residues V509, L513, V516, L517, V520 face the hydrophobic groove formed by the α1 helix from both DPY30 protomers. The hydrophobic residues important for binding are shown in stick models. (H) Single point mutations of Ash2L residues on the hydrophobic interface partially interfered with the interaction with DPY30, whereas a triple mutation 3E (L513E/L517E/L520E) completely abolished the interaction.

showed that Arg343 of Ash2L is important for RbBP5 binding [5]. Arg343 is localized at a highly positively charged surface formed by strands  $\beta$ 7,  $\beta$ 8 and loop E of Ash2L<sub>SPRY</sub> (Figure 1C). This observation led us to hypothesize that this basic concaved surface on Ash2L<sub>SPRY</sub> might be the binding site for RbBP5. To test this idea, we examined whether mutations of residues on the predicted Ash2L-RbBP5 interface could weaken or disrupt the Ash2L-RbBP5 interaction. In support of our model and consistent with our previous results, alanine substitution of the conserved Arg343 of Ash2L<sub>SPRY</sub> was sufficient to completely abolish the interaction with RbBP5 in the ITC assay (Figure 1D). Similarly, alanine mutation of another arginine residue Arg367 also impaired the interaction (Figure 1D). In addition, several other positively charged residues of Ash2L<sub>SPRY</sub> (Lys369, Lys370, and Lys476) also play roles in RbBP5 interaction, as mutations of all these residues weaken the interaction to varying degrees (Figure 1D). Moreover, these RbBP5 binding-deficient mutants showed partial or complete loss of ability to stimulate the HMT activity of MLL1, in a manner consistent with the severity of the Ash2L-RbBP5 interaction defect (Figure 1E). Taken together, we conclude that the identified basic surface on Ash2L<sub>SPRY</sub> is necessary for the binding to RbBP5. All the mutant proteins described above displayed unaltered biophysical properties, as analyzed by gel-filtration chromatography (data not shown), assuring

that the altered affinity of the mutants for RbBP5 is not attributable to a change in the structural integrity of the resulting proteins. Ash $2L_{SPRY}$  employs a novel molecular surface to interact with RbBP5, as shown by the comparison with structures of other SPRY domain-containing complexes (Supplementary information, Figure S6), reinforcing the notion that SPRY domain is a diversified protein-protein interaction module.

Beside RbBP5, another important Ash2L-binding partner is DPY30 [4]. The C-terminus of Ash2L (residues 508-534) has been proposed to mediate the interaction with DPY30 [8]. We further characterized the Ash2L-DPY30 interaction by an *in vitro* GST-pull-down assay. Both the full-length DPY30 and its C-terminal dimerization <u>domain</u> (DPY30<sub>DD</sub>; residues 45-99) could efficiently pull-down Ash2L, suggesting that  $DPY30_{DD}$  is sufficient for the interaction with Ash2L (Supplementary information, Figure S7A). Next, various constructs of Ash2L that contain the SPRY domain and end at different positions at the C-terminus of Ash2L were evaluated for their ability to interact with DPY30<sub>DD</sub>. The Ash2L<sub>SPRY</sub> construct used for the structural study (residues 276-523) retained the binding activity with DPY30, whereas the construct ending at residue 510 failed to bind to DPY30 (Figure 1F). This result suggested that a short 14-residue C-terminal fragment of Ash2L (residues 510-523) is required for the interaction with DPY30. Notably, this short frag-



ment was also found to be sufficient for DPY30 binding (Figure 1F).

Although Ash2L<sub>510-523</sub> was included in the crystallization construct of Ash2L<sub>SPRV</sub>, residues 510-523 are not visible in the electron-density map, indicating that the Ash-2L<sub>DBM</sub> peptide is not part of the compact SPRY domain. DPY30<sub>DD</sub> adopts a canonical four-helix-bundle conformation with extensive hydrophobic surface [9]. The structure of DPY30<sub>DD</sub> is similar to that of PKA RIIa, the RIIa isoform of a cAMP-dependent protein kinase [10]. The two structures can be superimposed with an r.m.s.d. value of 2.1 Å (Supplementary information, Figure S7B). In the PKA-RIIa/D-AKAP2 complex, the D-AKAP2 peptide forms a helical structure and binds to the hydrophobic cleft on the dimeric interface of PKA RIIa [10]. Notably, secondary structure prediction analysis suggested that, similar to the D-AKAP2 helix, residues 510 to 523 of Ash2L also likely adopt an  $\alpha$ -helical structure (data not shown). Therefore, we hypothesized that Ash- $2L_{DBM}$  might bind to DPY30 in a similar manner to the D-AKAP2 helix in the PKA-RIIa/D-AKAP2 complex. Sequence alignment of Ash2L<sub>DBM</sub> with the D-AKAP2 peptide revealed that the hydrophobic residues that are involved in the complex formation in the PKA-RIIa/ D-AKAP2 complex are well-conserved in Ash2L (Supplementary information, Figure S7C). Based on these analyses, we generated a model structure of  $Ash2L_{SPRY}$ complexed with DPY30<sub>DD</sub> (Figure 1G). In this model, the Ash2L-DPY30 interaction is mainly mediated by hydrophobic contacts; Ash2L residues Val509, Leu513, Val516, Leu517, and Val520 from one side of the Ash- $2L_{DBM}$  helix fit into the hydrophobic groove formed by a group of hydrophobic residues from the  $\alpha$ 1 helix of both DPY30<sub>DD</sub> protomers (Leu57, Val61, Leu65, Leu66, and Met69) (Figure 1G). This is consistent with previous results that substitution of DPY30 Leu65 and Leu66 with glutamates completely abolished the interaction with Ash2L [8]. To further corroborate our structural model, we generated a series of Ash2L mutants and examined the effects of these mutations on the Ash2L-DPY30 interaction. Despite mutations of individual Ash2L residues on the Ash2L-DPY30 interface only slightly weakened the interaction, a triple Ash2L mutation L513E/L517E/ V520E completely disrupted the interface (Figure 1H). A definitive confirmation of this interaction model between Ash $2L_{DBM}$  and DPY $30_{DD}$  has to wait until the atomic structure of the Ash2L<sub>DBM</sub>-DPY30<sub>DD</sub> complex becomes available. Nonetheless, our structural and biochemical data showed that DPY30 and RbBP5 bind to different sites on Ash2L<sub>SPRY</sub>, suggesting that they mediate independent interactions with Ash2L. Consistent with this idea, addition of DPY30 had no effect on the interaction

between Ash2L<sub>SPRY</sub> and RbBP5 (Supplementary information, Figure S7D).

In the present work, we report the crystal structure of the C-terminal SPRY domain of human Ash2L. Our structural and biochemical analyses reveal a basic surface on Ash2L as the RbBP5-binding interface, and demonstrate that this interface is crucial for both RbBP5 binding and MLL1 methyltransferase activity regulation. We also determine the minimum motif of Ash2L that is required for DPY30 interaction, and propose a structural model for the Ash2L-DPY30 interaction. Together, these results provide a framework for the further investigation of the structure of the MLL complexes.

#### Accession numbers

Atomic coordinates and structure factors have been deposited with the Protein Data Bank accession code 3TOJ.

### Acknowledgments

ML acknowledges generous financial support from NIH (RO1 GM083015) and the American Cancer Society. ML is a Howard Hughes Medical Institute Early Career Scientist. The General Medicine and Cancer Institutes Collaborative Access Team has been funded in whole or in part by federal grants from the National Cancer Institute (Y1-CO-1020) and the National Institute of General Medical Science (Y1-GM-1104). Use of the Advanced Photon Source was supported by the US Department of Energy, Office of Science, Office of Basic Energy Sciences, under contract No DE-AC02-06CH11357.

Yong Chen<sup>1, 2</sup>, Fang Cao<sup>3</sup>, Bingbing Wan<sup>1, 2</sup>, Yali Dou<sup>2, 3</sup>, Ming Lei<sup>1, 2</sup>

<sup>1</sup>Howard Hughes Medical Institute; <sup>2</sup>Department of Biological Chemistry; <sup>3</sup>Department of Pathology, University of Michigan Medical School, 1150 W. Medical Center Drive, Ann Arbor, MI 48109, USA

Correspondence: Ming Lei E-mail: leim@umich.edu

#### References

- Ruthenburg AJ, Allis CD, Wysocka J. Methylation of lysine 4 on histone H3: intricacy of writing and reading a single epigenetic mark. *Molecular cell* 2007; 25:15-30.
- 2 Dou Y, Milne TA, Ruthenburg AJ, *et al.* Regulation of MLL1 H3K4 methyltransferase activity by its core components. *Nat Struct Mol Biol* 2006; **13**:713-719.
- 3 Steward MM, Lee JS, O'Donovan A, Wyatt M, Bernstein BE, Shilatifard A. Molecular regulation of H3K4 trimethylation by ASH2L, a shared subunit of MLL complexes. *Nat Struct Mol Biol* 2006; 13:852-854.
- 4 Cho YW, Hong T, Hong S, *et al.* PTIP associates with MLL3and MLL4-containing histone H3 lysine 4 methyltransferase complex. *J Biol Chem* 2007; **282**:20395-20406.

602

- 5 Cao F, Chen Y, Cierpicki T, *et al.* An Ash2L/RbBP5 heterodimer stimulates the MLL1 methyltransferase activity through coordinated substrate interactions with the MLL1 SET domain. *PLoS One* 2010; 5:e14102.
- 6 Jiang H, Shukla A, Wang X, Chen WY, Bernstein BE, Roeder RG. Role for Dpy-30 in ES cell-fate specification by regulation of H3K4 methylation within bivalent domains. *Cell* 2011; 144:513-525.
- 7 Avdic V, Zhang P, Lanouette S, *et al.* Structural and biochemical insights into MLL1 core complex assembly. *Structure* 2011; **19**:101-108.
- 8 South PF, Fingerman IM, Mersman DP, Du HN, Briggs SD. A conserved interaction between the SDI domain of Bre2 and the Dpy-30 domain of Sdc1 is required for histone methylation and gene expression. *J Biol Chem* 2010; **285**:595-607.
- 9 Wang X, Lou Z, Dong X, et al. Crystal structure of the Cterminal domain of human DPY-30-like protein: a component of the histone methyltransferase complex. J Mol Biol 2009; 390:530-537.
- 10 Kinderman FS, Kim C, von Daake S, *et al.* A dynamic mechanism for AKAP binding to RII isoforms of cAMP-dependent protein kinase. *Mol Cell* 2006; **24**:397-408.

(**Supplementary information** is linked to the online version of the paper on the *Cell Research* website.)