Published in final edited form as:

J Mol Biol. 2017 June 30; 429(13): 2066–2074. doi:10.1016/j.jmb.2016.09.020.

Structural insight into recognition of methylated histone H3K4 by Set3

Jovylyn Gatchalian^{1,3}, Muzaffar Ali^{1,3}, Forest H. Andrews¹, Yi Zhang¹, Alexander S. Barrett², and Tatiana G. Kutateladze^{1,2,*}

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

²Program in Structural Biology and Biochemistry, University of Colorado School of Medicine, Aurora, CO 80045, USA

Abstract

The plant homeodomain (PHD) finger of Set3 binds methylated lysine 4 of histone H3 *in vitro* and *in vivo*, however precise selectivity of this domain has not been fully characterized. Here, we explore the determinants of methyllysine recognition by the PHD fingers of Set3 and its orthologs. We use X-ray crystallographic and spectroscopic approaches to show that the Set3 PHD finger binds di- and trimethylated states of H3K4 with comparable affinities and employs similar molecular mechanisms to form complexes with either mark. Composition of the methyllysine-binding pocket plays an essential role in determining the selectivity of the PHD fingers. The finding that the histone-binding activity is not conserved in the PHD finger of Set4 suggests different functions for the Set3 and Set4 paralogs.

Graphical abstract

*Corresponding author: tatiana.kutateladze@ucdenver.edu; phone (303) 724-3593; fax (303) 724-3663. Department of Pharmacology, University of Colorado School of Medicine, 12801 E. 17th Ave., Aurora, CO 80045, USA.

³Equal contribution

Accession numbers

Atomic coordinates for the structures of the Set3 PHD-H3K4me3 and Set3 PHD-H3K4me2 complexes have been deposited in Protein Data Bank under accession codes 5TDW and 5TDR.

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

Keywords

PHD finger; Set3; Set4; histone; methylation

Introduction

Methylation of histone H3 at lysine 4 is essential for transcriptional regulation^{1; 2; 3}. In yeast, the catalytic subunit of the COMPASS complex, Set1, produces mono-, di-, and tri-methylated marks (H3K4me1, H3K4me2 and H3K4me3) that concentrate in separate but overlapping genomic regions⁴. High levels of H3K4me3 are detected in the promoters and around transcription start sites (TSS) of actively transcribed genes within the yeast genome⁵. In contrast, H3K4me2 is found enriched in regions downstream of TSS and in the middle of the genes, whereas H3K4me1 is located downstream of the dimethylated species⁵. The Set1-catalyzed H3K4 dimethylation is targeted by the yeast histone deacetylase complex Set3C, whose enzymatic subunits Hos2 and Hst1 deacetylate histones H3 and H4^{6; 7; 8; 9}. Set3C has been characterized as both transcriptionally repressive and activating complex^{7; 8; 10; 11}, and its recruitment to chromatin has been found to depend in part on the interaction with phosphorylated CTD of Pol II and on the histone-binding activity of the Set3 subunit^{6; 12}.

The Set3 subunit contains a plant homeodomain (PHD) finger followed by a Su(var)3–9, Enhancer of Zeste, Trithorax (SET) domain. This two-domain composition is conserved in the Set3 paralog, yeast Set4, and in orthologs, including fly UpSET and human mixed lineage leukemia 5 (MLL5, also known as KMT2E)^{13; 14; 15}. Although the biological role of the SET domain of this subfamily of proteins remains poorly understood (beyond the finding that MLL5 possesses no intrinsic methyltransferase activity¹⁴ and no catalytic activity has been reported for Set3 or UpSET either), the PHD fingers of Set3, UpSET and MLL5 have been shown to recognize methylated H3K4^{6; 9; 16; 17; 18}. Interaction of the PHD finger with H3K4me3 stabilizes MLL5 at chromatin^{16; 19}, and binding of the Set3 PHD finger to H3K4me2 mediates deacetylation of histones in 5'-transcribed gene regions by the Set3C complex⁶. Tight association of the Set3 PHD finger with H3K4me3 has also been observed *in vitro* by histone peptide pull-down assays and fluorescence spectroscopy¹⁷. Despite the high significance of the Set3-H3K4me interaction, selectivity of the Set3 PHD finger toward methylated H3K4 species has not been fully characterized.

In this study, we explore the determinants of methyllysine recognition by the PHD finger of Set3. We show that the Set3 PHD finger binds H3K4me3 and H3K4me2 with comparable affinities and employs similar molecular mechanisms to form complexes with either mark. The finding that the histone-binding activity is not conserved in the PHD finger of Set4 suggests different functions for the Set3 and Set4 paralogs.

Results and Discussion

Set3 PHD shows selectivity for H3K4me3/me2

To determine whether the Set3 PHD finger has preference for H3K4 methylation states, we measured dissociation constants (K_d s) for the interactions of the Set3 PHD finger with histone H3K4me3 and H3K4me2 peptides using intrinsic fluorescence binding assays. We found that Set3 PHD binds H3K4me3 with affinity of 20 μ M, which is in agreement with the previously measured binding affinity of 18 μ M¹⁷ (Fig. 1a and b). A slightly weaker affinity (K_d = 30 μ M) was measured for binding of the Set3 PHD finger to H3K4me2. To validate the fluorescence data, we examined these interactions in ¹H,¹⁵N heteronuclear single quantum coherence (HSQC) NMR experiments. The unlabeled histone peptides (residues 1–12 of histone H3) were titrated into the NMR sample containing uniformly ¹⁵N-labeled Set3 PHD, and ¹H,¹⁵N HSQC spectra were recorded after each addition of the peptide (Fig. 1c). Large chemical shift perturbations (CSPs) in the intermediate exchange regime on the NMR time scale observed upon addition of H3K4me3 peptide indicated direct binding of Set3 PHD (Fig. 1c, top panel). Titration of H3K4me2 peptide to the Set3 PHD finger resulted in an overall similar pattern of CSPs, though the intermediate-to-fast exchange regime suggested that this interaction is less robust compared to the interaction with H3K4me3 (Fig. 1c). A fast exchange regime observed in ¹H,¹⁵N HSQC spectra upon titration of H3K4me1 and H3K4me0 peptides indicated a significant decrease in binding affinity of the Set3 PHD finger (K_d s = 790 μ M and 1.4 mM, respectively, as measured by NMR). Together, these data demonstrate that the Set3 PHD finger recognizes both H3K4me3 and H3K4me2, slightly favoring the trimethylated mark, but it selects against the lower methylation states of H3K4.

Structural mechanism for the recognition of H3K4me3

To gain insight into the basis for the methyllysine selectivity, we determined the crystal structure of the Set3 PHD finger in complex with H3K4me3 peptide (residues 1–12 of H3) to a resolution of 1.7 Å (Fig. 2 and Table 1). The Set3 PHD finger folds into a double-stranded antiparallel β sheet and a 3_{10} -helical turn that are stabilized by two zinc-binding clusters. In addition to the canonical PHD-like fold, a long α -helix is present in the C-terminus of Set3 PHD. The bound peptide forms the third antiparallel β -strand and pairs with β 2 via characteristic backbone-backbone hydrogen bonds involving R2, K4, and T6 of the peptide and Gly129, Thr131, and Gln133 of the protein (Fig. 2b). Further stabilization of the H3K4me3 peptide is achieved through formation of hydrogen bonds between the free amino group of A1 and the carbonyl oxygen atoms of Pro155 and Asp156.

In addition to the backbone-mediated polar contacts, intermolecular hydrogen bonds to the side chains of the peptide reinforce the interaction (Fig. 2b). Particularly, the side chain of R2 adopts two conformations: the guanidino group of one of the rotamers donates hydrogen bonds to the side chains of Gln133 and Asn138 as well as to the backbone carbonyl group of Cys134 and makes a water-mediated hydrogen bond to the side chain of Asp135. The second conformation of R2 allows for the interaction with the side chain amide of Gln133. Other water-mediated hydrogen bonds are observed between Q5 and the backbone carbonyl group of Ile151, and between the hydroxyl group of T6 and the carboxylate group of Asp127. Trimethylated K4 of the peptide is bound in the pocket composed of Ile118,

Asp127, Thr131, and Trp140, with the latter stabilizing the trimethylammonium moiety via a cation- π interaction (Fig. 2b).

Structural basis for binding of H3K4me2

To rationalize the capability of the Set3 PHD finger to bind both trimethylated and dimethylated H3K4 marks, we obtained the crystal structure of the H3K4me2-bound protein at a 1.4 Å resolution (Fig. 2c and Table 1). Overall, the structure of the PHD-H3K4me2 complex shows no major differences as compared to the structure of the PHD-H3K4me3 complex, and the two structures superimpose with an RMSD of <0.1 Å. Furthermore, the dimethylammonium group of K4 is positioned at the same (~3.6 Å) distance to Trp140 as the trimethylammonium group of K4 is positioned, implying that like K4me3, K4me2 is also restrained through the cation- π interaction with the indole ring of Trp140. Further analysis of the binding pocket reveals that no direct or water-mediated interaction is present between Asp127 and K4me2 (the distance between the nitrogen atom in K4me2 and the carboxyl group of Asp127 is 5.1 Å, Fig. 2d). Instead, the methyl groups of K4me2 are buried in the hydrophobic part of the binding cage, lined with Ile118 and Thr131.

A possible explanation for the inability of Asp127 to be engaged with the dimethylammonium moiety might arise from the rigid position of Asp127. The side chain of Asp127 is constrained due to the water-mediated hydrogen bond with T6 of the peptide, and its backbone conformation is restricted by interactions of adjacent residues (Fig. 2e). Carboxylate groups of Asp126 and Asp128 form robust contacts, i.e. a salt-bridge and hydrogen bonds with Arg180 and Gln177 residues in the C-terminal α -helix. Further, hydrophobic zip-like interactions between Leu124 and Ile181, Ile144 and Ile176, and Ile121 and Ile170 provide additional strength to this interface (indicated by orange ovals in Fig. 2e). These extensive contacts can prevent Asp127 from moving closer to the dimethylammonium group to form a hydrogen bond. Considering the lack of this hydrogen bond (that usually accounts for selectivity of some Tudors or MBT toward mono- and dimethyllysine over trimethyllysine^{20; 21}) and the position of the methyl moieties, it appears that the main driving force for binding of the K4me2 mark by Set3 is largely cation- π and hydrophobic in nature. A substantial decrease or paucity of hydrophobicity in H3K4me1 or H3K4me0 likely prevents recognition of these species by the Set3 PHD finger.

Composition of the methyllysine-binding pocket confers selectivity

An alignment of amino acid sequences of the PHD fingers from yeast Set3, human MLL5 and fly UpSET shows that two residues in the methyllysine-binding pocket are absolutely conserved: an aspartate and a tryptophan (colored orange in Fig. 3a). However the residues that create the “base” of the binding pocket differ (colored pink in Fig. 3a). Whereas Set3 PHD contains Ile118 and Thr131, in MLL5 PHD these residues are Thr119 and Met132 (Fig. 3b). Of note, a methionine in this position (Met132 in MLL5) is commonly found in other PHD fingers that recognize H3K4me3²². We therefore hypothesized that the presence of Thr in Set3 PHD may be the reason for the lack of strong selectivity toward trimethyllysine. The hydroxyl group of Thr131 in the Set3 PHD-H3K4me3 and Set3 PHD-H3K4me2 complexes points directly toward K4 of the peptide, causing steric hindrance and arching of the hydrophobic part of the K4 side chain (Fig. 2d and 3b). By contrast, the side

chain of K4me3 is bound in a fully extended conformation and lies essentially parallel to the flattened Met132 in the MLL5 PHD-H3K4me3 complex.

To test this hypothesis, we generated the T131M mutant of Set3 PHD and examined its interaction with H3K4me3 and H3K4me2 peptides by tryptophan fluorescence and NMR (Fig. 3c and d). As shown in Figure 3d, mutation of Thr131 to Met resulted in an increase in binding of the Set3 PHD finger to both peptides, however this mutant indeed became ~3-fold more selective for H3K4me3 than for H3K4me2 ($K_d = 8 \mu\text{M}$ and $2.5 \mu\text{M}$ for H3K4me2 and H3K4me3, respectively, as measured by fluorescence). An overall similar pattern of CSPs observed in the Set3 PHD finger upon titration with H3K4me3 peptide or H3K4me2 peptide suggested that the binding mode of T131M Set3 is conserved (Fig. 3c). Concomitantly, the M132T mutant of MLL5 PHD appears to have lost its preference for H3K4me3 (Fig. 3d, e). Whereas the wild type MLL5 PHD finger binds with a ~5-fold higher affinity to H3K4me3 than it binds to H3K4me2¹⁶, the M132T mutant of MLL5 recognizes tri- and dimethyllysine almost equally well, as does wild type Set3 PHD, exhibiting only 1.5-fold selectivity toward the trimethylated mark (Fig. 3d). Collectively, these findings suggest that the base residue in the methyllysine-binding pocket plays an important role in determining selectivity of PHD fingers for di- or trimethylated states of H3K4.

The aromatic cage is not sufficient for Set4

Although amino acid sequences of the PHD fingers of Set3 and Set4 are very similar (Fig. 3a), Set4 PHD has been shown to possess no H3K4me-binding activity¹⁷, and NMR titration experiments confirmed that it does not recognize H3K4me3 or H3K4me0 (Fig. 4a). Comparison of the methyllysine-binding site residues of Set3 and MLL5 with the corresponding residues of Set4 reveals that Trp and Asp are conserved in all three proteins, however Set4 contains different residues, Phe175 and Asn161, in two other positions (Fig. 4b). In the attempt to generate a gain-of-function mutant of Set4 PHD finger, we replaced Phe175 and Asn161 in the Set4 PHD finger with those found in either Set3 or MLL5 (Fig. 4c). Additionally, Lys182 of Set4 located in a would-be R2-binding pocket was substituted with an Asn, as in Set3 PHD finger (Fig. 4d). Although all generated mutants of Set4 PHD finger were folded and stable, none was able to bind H3K4me3 (Supplementary Fig. S1; the list of Set4 mutants and tested additives is shown in Fig. 4c). Together these results show that engineering the Set3-active aromatic cage within the Set4 PHD finger is not sufficient to induce binding of Set4 to H3K4me tails.

Concluding remarks

In conclusion, our study demonstrates that the Set3 PHD finger binds dimethylated and trimethylated H3K4 almost equally well but discriminates against H3K4me1 and H3K4me0. The me3/me2 selectivity of the Set3 PHD finger is reminiscent to that of the Pygo PHD finger, though the latter locks H3K4me2 through an additional hydrogen bond involving the dimethylammonium group of lysine and the carboxyl group of Asp that creates an opposite-to-Trp wall in the aromatic cage²³. Unlike Pygo, the Set3 PHD finger appears to not use this polar contact and relies on largely cation- π and hydrophobic interactions to accommodate K4me2. Interestingly, similar to Set3 and Pygo, the PHD fingers of MLL5 and Taf3 contain

an Asp residue in their methyllysine-binding aromatic cages (Fig. 3a), however in contrast to Set3 and Pygo, these PHD fingers prefer H3K4me3 over H3K4me2 by ~5- to 10-fold^{16; 24; 25}. Both MLL5 and Taf3 contain a Met as a base residue in their aromatic cages that at least in the case of MLL5 plays a role in higher selectivity for H3K4me3 compared to H3K4me2.

The base residue in the methyllysine-binding cage of Set3 is Thr131, which produces steric hindrance and causes the side chain of methyllysine to be buried not as deeply as it is buried in the binding cage of MLL5. Our data suggest that the Thr residue of Set3 contributes to the ability of Set3 to accommodate dimethylated and trimethylated H3K4 almost equally well. Substitution of the Thr131 with Met results in gaining the selectivity of Set3 toward H3K4me3. *In vivo*, PHD-mediated recruitment of the Set3C complex helps to localize this complex to the chromatin regions enriched in H3K4me2⁶. This may be due to the presence of other regions within the Set3 protein that make binding to H3K4me2 more favorable over binding to H3K4me3. It is also possible that other Set3C subunits or associating complexes preoccupy H3K4me3 or play a role in regulating the Set3 histone binding activity.

Our study further demonstrates that the inability of the Set4 PHD finger to bind H3K4me3 is not due to its impaired methyllysine-binding pocket. We have previously shown that the complete aromatic cage in the Tudor domain of *Drosophila* Pcl is necessary but insufficient for binding to H3K36me3 and that the adjacent hydrophobic patch is also required²⁶. The fact that we were unable to generate the gain-of-function mutant of Set4 by engineering the Set3-like aromatic cage suggests that other surface residues play an imperative role. Nevertheless, because the wild type Set4 PHD finger does not have histone-binding activity, Set3 and Set4 paralogs most likely have distinct functions in the cell.

Materials and Methods

Protein purification and mutagenesis

The PHD fingers of Set3 (aa 115-184 and 116-184), MLL5 (aa 117-181) and Set4 (aa 160-226) were expressed in *E. coli* Rosetta2 or BL21 (DE3) pLysS cells in Luria Broth or in M9 minimal media supplemented with ¹⁵NH₄Cl and ZnCl₂. Protein production was induced with 0.5 mM IPTG for 16–18 hours at 18 °C. After harvest, cells were lysed by sonication and the GST-fusion proteins were purified on glutathione sepharose 4B beads (GE Healthcare). Subsequently, the GST-tag was cleaved with Thrombin protease (GE healthcare). The cleaved proteins were concentrated in a final solution containing 20 mM Tris-HCl (pH 6.9), 150 mM NaCl and 5 mM dithiothreitol. For crystallization, Set3 PHD was further purified by size exclusion chromatography on a Hi-Prep Sephacryl S100 16/60 column. Point mutants were prepared using the Stratagene QuickChange XL Site Directed Mutagenesis kit according to the manufacturer's instructions.

NMR Spectroscopy

The ¹H, ¹⁵N HSQC spectra of 0.1 mM uniformly ¹⁵N-labelled wild type or mutant PHD fingers were collected at 298 K on a Varian INOVA 600 MHz spectrometer equipped with a cryogenic probe. Binding was analyzed through monitoring chemical shift perturbations

observed in ^1H , ^{15}N HSQC spectra of PHD fingers upon addition of histone peptides (synthesized by the University of Colorado Denver Peptide Core Facility) or 16mer dsDNA. The dissociation constants (K_d s) were determined by a nonlinear least-squares analysis in Kaleidagraph using the equation:

$$\Delta\delta = \Delta\delta_{\max} \frac{\left(([L] + [P] + K_d) - \sqrt{([L] + [P] + K_d)^2 - 4[P][L]} \right)}{2[P]}$$

where [L] is concentration of the peptide, [P] is concentration of the protein, δ is the observed chemical shift change, and $\delta_{\max}$ is the normalized chemical shift change at saturation. Normalized 27 chemical shift changes were calculated using the equation

$$\Delta\delta = \Delta\delta = \sqrt{(\Delta\delta_H)^2 + \left(\frac{\Delta\delta_N}{5}\right)^2}, \text{ where } \delta \text{ is the change in chemical shift in parts per million (ppm).}$$

Fluorescence spectroscopy

Spectra were recorded at 25 °C on a Fluoromax-3 spectrofluorometer (HORIBA). The samples containing 1–2 μM wild type or mutated Set3 PHD finger and progressively increasing concentrations of the histone peptide were excited at 295 nm. Emission spectra were recorded between 315 and 405 nm with a 0.5 nm step size and a 1 s integration time and averaged over 3 scans. The K_d values were determined using a nonlinear least-squares analysis and the equation:

$$\Delta I = \Delta I_{\max} \frac{\left(([L] + [P] + K_d) - \sqrt{([L] + [P] + K_d)^2 - 4[P][L]} \right)}{2[P]}$$

where [L] is the concentration of the histone peptide, [P] is the concentration of the protein,

I is the observed change of signal intensity, and $I_{\max}$ is the difference in signal intensity of the free and bound states of the protein. Each K_d value was averaged over three separate experiments, with error calculated as the standard deviation between the runs.

X-Ray crystallography

The Set3 PHD finger (aa 116-184) was crystallized in complex with H3K4me2 (1-12) and H3K4me3 (1-12) peptides. The purified PHD finger (7 mg/mL) was incubated with each histone peptide in a 1:1.5 molar ratio in 20 mM Tris-HCl buffer (pH 6.9), 150 mM NaCl and 5 mM DTT for two hours prior to crystallization. Good quality diffracting crystals were obtained by seeding at 4 °C in 0.2 M tri-sodium citrate and 20% (w/v) PEG 3350. Data collection was performed at the ALS 4.2.2 beamline, Berkeley. The structure of the Set3 PHD-H3K4me3 complex was solved by molecular replacement using the structure of MLL5 PHD as a model (PDBID 4L58). The structure of the Set3 PHD-H3K4me2 complex was solved using molecular replacement and the structure of Set3 PHD-H3K4me3 as a model. Manual model building was performed using Coot, the structure was refined using Phenix,

and the final structure was verified by Molprobity^{28; 29; 30}. The crystallographic statistics is shown in Table 1.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We thank Jay Nix at beam line 4.2.2 of the ALS in Berkeley for help with X-ray crystallographic data collection. This research was supported by grants from the NIH, GM100907 and GM106416 to T.G.K. J.G. is an NIH NRSA predoctoral fellow and F.H.A. is an American Heart Association postdoctoral fellow.

References

1. Black JC, Van Rechem C, Whetstone JR. Histone lysine methylation dynamics: establishment, regulation, and biological impact. *Mol Cell*. 2012; 48:491–507. [PubMed: 23200123]
2. Wozniak GG, Strahl BD. Hitting the ‘mark’: interpreting lysine methylation in the context of active transcription. *Biochim Biophys Acta*. 2014; 1839:1353–61. [PubMed: 24631869]
3. Strahl BD, Ohba R, Cook RG, Allis CD. Methylation of histone H3 at lysine 4 is highly conserved and correlates with transcriptionally active nuclei in *Tetrahymena*. *Proc Natl Acad Sci U S A*. 1999; 96:14967–72. [PubMed: 10611321]
4. Shilatifard A. The COMPASS family of histone H3K4 methylases: mechanisms of regulation in development and disease pathogenesis. *Annu Rev Biochem*. 2012; 81:65–95. [PubMed: 22663077]
5. Pokholok DK, Harbison CT, Levine S, Cole M, Hannett NM, Lee TI, Bell GW, Walker K, Rolfe PA, Herbolzheimer E, Zeitlinger J, Lewitter F, Gifford DK, Young RA. Genome-wide map of nucleosome acetylation and methylation in yeast. *Cell*. 2005; 122:517–27. [PubMed: 16122420]
6. Kim T, Buratowski S. Dimethylation of H3K4 by Set1 recruits the Set3 histone deacetylase complex to 5′ transcribed regions. *Cell*. 2009; 137:259–72. [PubMed: 19379692]
7. Pijnappel WW, Schaft D, Roguev A, Shevchenko A, Tekotte H, Wilm M, Rigaut G, Seraphin B, Aasland R, Stewart AF. The *S. cerevisiae* SET3 complex includes two histone deacetylases, Hos2 and Hst1, and is a meiotic-specific repressor of the sporulation gene program. *Genes Dev*. 2001; 15:2991–3004. [PubMed: 11711434]
8. Wang A, Kurdastani SK, Grunstein M. Requirement of Hos2 histone deacetylase for gene activity in yeast. *Science*. 2002; 298:1412–4. [PubMed: 12434058]
9. Kim T, Xu Z, Clauder-Munster S, Steinmetz LM, Buratowski S. Set3 HDAC mediates effects of overlapping noncoding transcription on gene induction kinetics. *Cell*. 2012; 150:1158–69. [PubMed: 22959268]
10. Yu Y, Zhou H, Deng X, Wang W, Lu H. Set3 contributes to heterochromatin integrity by promoting transcription of subunits of Clr4-Rik1-Cul4 histone methyltransferase complex in fission yeast. *Sci Rep*. 2016; 6:31752. [PubMed: 27538348]
11. Hang M, Smith MM. Genetic analysis implicates the Set3/Hos2 histone deacetylase in the deposition and remodeling of nucleosomes containing H2A.Z. *Genetics*. 2011; 187:1053–66. [PubMed: 21288874]
12. Govind CK, Qiu H, Ginsburg DS, Ruan C, Hofmeyer K, Hu C, Swaminathan V, Workman JL, Li B, Hinnebusch AG. Phosphorylated Pol II CTD recruits multiple HDACs, including Rpd3C(S), for methylation-dependent deacetylation of ORF nucleosomes. *Mol Cell*. 2010; 39:234–46. [PubMed: 20670892]
13. Emerling BM, Bonifas J, Kratz CP, Donovan S, Taylor BR, Green ED, Le Beau MM, Shannon KM. MLL5, a homolog of *Drosophila* trithorax located within a segment of chromosome band 7q22 implicated in myeloid leukemia. *Oncogene*. 2002; 21:4849–54. [PubMed: 12101424]
14. Sebastian S, Sreenivas P, Sambasivan R, Cheedipudi S, Kandalla P, Pavlath GK, Dhawan J. MLL5, a trithorax homolog, indirectly regulates H3K4 methylation, represses cyclin A2 expression, and

promotes myogenic differentiation. *Proc Natl Acad Sci U S A*. 2009; 106:4719–24. [PubMed: 19264965]

15. Rincon-Arano H, Halow J, Delrow JJ, Parkhurst SM, Groudine M. UpSET recruits HDAC complexes and restricts chromatin accessibility and acetylation at promoter regions. *Cell*. 2012; 151:1214–28. [PubMed: 23177352]
16. Ali M, Rincon-Arano H, Zhao W, Rothbart SB, Tong Q, Parkhurst SM, Strahl BD, Deng LW, Groudine M, Kutateladze TG. Molecular basis for chromatin binding and regulation of MLL5. *Proc Natl Acad Sci U S A*. 2013; 110:11296–301. [PubMed: 23798402]
17. Shi X, Kachirskaja I, Walter KL, Kuo JH, Lake A, Davrazou F, Chan SM, Martin DG, Fingerma IM, Briggs SD, Howe L, Utz PJ, Kutateladze TG, Lugovskoy AA, Bedford MT, Gozani O. Proteome-wide analysis in *Saccharomyces cerevisiae* identifies several PHD fingers as novel direct and selective binding modules of histone H3 methylated at either lysine 4 or lysine 36. *J Biol Chem*. 2007; 282:2450–5. [PubMed: 17142463]
18. Lemak A, Yee A, Wu H, Yap D, Zeng H, Dombrovski L, Houliston S, Aparicio S, Arrowsmith CH. Solution NMR Structure and Histone Binding of the PHD Domain of Human MLL5. *PLoS One*. 2013; 8:e77020. [PubMed: 24130829]
19. Ali M, Hom RA, Blakeslee W, Ikenouye L, Kutateladze TG. Diverse functions of PHD fingers of the MLL/KMT2 subfamily. *Biochim Biophys Acta*. 2014; 1843:366–71. [PubMed: 24291127]
20. Taverna SD, Li H, Ruthenburg AJ, Allis CD, Patel DJ. How chromatin-binding modules interpret histone modifications: lessons from professional pocket pickers. *Nat Struct Mol Biol*. 2007; 14:1025–40. [PubMed: 17984965]
21. Musselman CA, Lalonde ME, Cote J, Kutateladze TG. Perceiving the epigenetic landscape through histone readers. *Nat Struct Mol Biol*. 2012; 19:1218–27. [PubMed: 23211769]
22. Musselman CA, Kutateladze TG. Handpicking epigenetic marks with PHD fingers. *Nucleic acids research*. 2011; 39:9061–71. [PubMed: 21813457]
23. Fiedler M, Sanchez-Barrena MJ, Nekrasov M, Mieszczanek J, Rybin V, Muller J, Evans P, Bienz M. Decoding of methylated histone H3 tail by the Pygo-BCL9 Wnt signaling complex. *Mol Cell*. 2008; 30:507–18. [PubMed: 18498752]
24. van Ingen H, van Schaik FM, Wienk H, Ballering J, Rehmann H, Dechesne AC, Kruijzer JA, Liskamp RM, Timmers HT, Boelens R. Structural insight into the recognition of the H3K4me3 mark by the TFIID subunit TAF3. *Structure*. 2008; 16:1245–56. [PubMed: 18682226]
25. Vermeulen M, Mulder KW, Denisov S, Pijnappel WW, van Schaik FM, Varier RA, Baltissen MP, Stunnenberg HG, Mann M, Timmers HT. Selective anchoring of TFIID to nucleosomes by trimethylation of histone H3 lysine 4. *Cell*. 2007; 131:58–69. [PubMed: 17884155]
26. Gatchalian J, Kingsley MC, Moslet SD, Rosas Ospina RD, Kutateladze TG. An aromatic cage is required but not sufficient for binding of Tudor domains of the Polycomblike protein family to H3K36me3. *Epigenetics*. 2015; 10:467–73. [PubMed: 25923537]
27. Grzesiek S, Stahl SJ, Wingfield PT, Bax A. The CD4 determinant for downregulation by HIV-1 Nef directly binds to Nef. Mapping of the Nef binding surface by NMR. *Biochemistry*. 1996; 35:10256–61. [PubMed: 8756680]
28. Adams PD, Afonine PV, Bunkoczi G, Chen VB, Davis IW, Echols N, Headd JJ, Hung LW, Kapral GJ, Grosse-Kunstleve RW, McCoy AJ, Moriarty NW, Oeffner R, Read RJ, Richardson DC, Richardson JS, Terwilliger TC, Zwart PH. PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr D Biol Crystallogr*. 2010; 66:213–21. [PubMed: 20124702]
29. Emsley P, Lohkamp B, Scott WG, Cowtan K. Features and development of Coot. *Acta Crystallogr D Biol Crystallogr*. 2010; 66:486–501. [PubMed: 20383002]
30. Chen VB, Arendall WB 3rd, Headd JJ, Keedy DA, Immormino RM, Kapral GJ, Murray LW, Richardson JS, Richardson DC. MolProbity: all-atom structure validation for macromolecular crystallography. *Acta Crystallogr D Biol Crystallogr*. 2010; 66:12–21. [PubMed: 20057044]

Highlights

- Set3 PHD finger binds H3K4me3 and H3K4me2 with comparable affinities
- Composition of the methyllysine-binding pocket confers selectivity
- The Set3-like aromatic cage is not sufficient to induce binding of Set4 to H3

Figure 1. Set3 PHD selects for histone H3K4me3 and H3K4me2

(a) Binding affinities of Set3 PHD to indicated histone peptides as measured by (a) intrinsic tryptophan fluorescence or (b) NMR. (b) Representative binding curves used to determine the K_d values by fluorescence. (c) Superimposed ^1H , ^{15}N HSQC NMR spectra of Set3 PHD recorded while indicated histone H3 peptides (1-12) were titrated in. Spectra are color coded according to the Set3 PHD:histone peptide molar ratio. Arrows indicate chemical shift changes.

Figure 2. Molecular mechanism for the recognition of H3K4me3 and H3K4me2 by the Set3 PHD finger

(a) The Set3 PHD finger is depicted as a solid surface colored wheat with the H3K4me3 peptide shown as green sticks. (b) The ribbon diagram of the Set3 PHD-H3K4me3 complex. Dashed lines represent hydrogen bonds. Residues of the PHD finger are denoted by a three-letter code, whereas residues of the histone peptide are denoted using a single-letter code. (c) Structural overlays of the Set3 PHD finger bound to H3K4me3 and H3K4me2 peptides. (d) Zoom in of the aromatic cage residues of the H3K4me2-bound Set3 PHD finger. The distances are indicated by dashed lines and labeled. (e) The C-terminal helix of the PHD finger stabilizes the structure. Blue dashed lines indicate hydrogen bonds and a salt bridge and orange circles indicate zip-like hydrophobic interactions.

Figure 3. Topology of the aromatic cage defines selectivity

(a) Alignment of amino acid sequences of the PHD fingers. Zinc-coordinating residues are green, absolutely conserved residues in the aromatic cage are orange, and variable residues in the aromatic cage are dark pink and light pink. Each tenth residue of Set3 is indicated by a red box and labeled. (b) Structural overlays of the H3K4me3-bound PHD fingers of MLL5 (PDB 4L58) and Set3. The aromatic cage residues of MLL5 PHD and Set3 PHD are light blue and wheat, respectively. (c) Superimposed ^1H , ^{15}N HSQC spectra of the T131M mutant of Set3 PHD color-coded according to the protein:peptide molar ratio. (d) Binding affinities (K_d values) of Set3 PHD and MLL5 PHD to the indicated peptides. (a) Taken from ¹⁶. (e) Representative binding curves used to determine the K_d values by NMR.

Figure 4. Set4 PHD does not have DNA or histone binding activities

(a) Superimposed ^1H , ^{15}N HSQC spectra of Set4 PHD color-coded according to the Set4 PHD:peptide molar ratio. (b) Topology of the aromatic cages of Set3, Set4 and MLL5: the variant residues are colored pink. (c) The list of Set4 PHD mutants and tested additives. (d) Structural overlays of the H3K4me3-bound Set3 PHD finger (wheat) and a Phyre2-generated model of Set4 PHD finger (light grey). Histone H3 residues R2 and K4me3 are modeled as green sticks.

Table 1

Data collection and refinement statistics for the Set3 PHD complexes (molecular replacement)

	Set3:H3K4me2 ₁₋₁₂ peptide	Set3:H3K4me3 ₁₋₁₂ peptide
Data collection		
Space group	<i>P</i> 22121	<i>P</i> 22121
Cell dimensions		
<i>a</i> , <i>b</i> , <i>c</i> (Å)	25.93, 43.72, 84.70	25.87, 43.68, 85.06
α , β , γ (°)	90, 90, 90	90, 90, 90
Resolution (Å)	1.42(1.47–1.42) *	1.7(1.76–1.70) *
<i>R</i> _{sym} or <i>R</i> _{merge}	16.5(91.6)	28.3(131)
<i>I</i> / σ <i>I</i>	7.5(0.9)	14 (3.5)
Completeness (%)	98.4(87.8)	99.9 (97.1)
Redundancy	3.5(2.5)	26.5(27.1)
Refinement		
Resolution (Å)	38.85–1.42	42.54–1.70
No. reflections	18586	11089
<i>R</i> _{work} / <i>R</i> _{free}	16.3/21.3	16.5/19.9
No. atoms	867	855
Protein	662	647
Ion	4	3
Water	204	205
<i>B</i> -factors	16.20	14.90
Protein	12.70	11.60
Ligand/ion	10.70	10.40
Water	27.70	25.20
R.m.s. deviations		
Bond lengths (Å)	0.002	0.008
Bond angles (°)	0.58	1.13
Ramachandran quality		
Favored (%)	97.4	97.4
Allowed (%)	2.6	2.6
Outliers (%)	0	0

* Values in parentheses are for highest-resolution shell. Data was collected from a single crystal.