JBC RESEARCH ARTICLE

A complete methyl-lysine binding aromatic cage constructed by two domains of PHF2

Received for publication, December 5, 2022, and in revised form, December 27, 2022 Published, Papers in Press, December 31, 2022,
https://doi.org/10.1016/j.jbc.2022.102862

John R. Horton, Jujun Zhou, Qin Chen, Xing Zhang, Mark T. Bedford, and Xiaodong Cheng*

From the Department of Epigenetics and Molecular Carcinogenesis, University of Texas MD Anderson Cancer Center, Houston, Texas, USA

Edited by Brian Strahl

The N-terminal half of PHF2 harbors both a plant homeodomain (PHD) and a Jumonji domain. The PHD recognizes both histone H3 trimethylated at lysine 4 and methylated nonhistone proteins including vaccinia-related kinase 1 (VRK1). The Jumonji domain erases the repressive dimethylation mark from histone H3 lysine 9 (H3K9me2) at select promoters. The N-terminal amino acid sequences of H3 (AR_2TK_4) and VRK1 (PR_2VK_4) bear an arginine at position 2 and lysine at position 4. Here, we show that the PHF2 N-terminal half binds to H3 and VRK1 peptides containing K4me3, with dissociation constants (K_D values) of 160 nM and 42 nM, respectively, which are 4 × and 21 × lower (and higher affinities) than for the isolated PHD domain of PHF2. X-ray crystallography revealed that the K4me3-containing peptide is positioned within the PHD and Jumonji interface, with the positively charged R2 residue engaging acidic residues of the PHD and Jumonji domains and with the K4me3 moiety encircled by aromatic residues from both domains. We suggest that the micromolar binding affinities commonly observed for isolated methyl-lysine reader domains could be improved *via* additional functional interactions within the same polypeptide or its binding partners.

The histone code hypothesis (1, 2) suggests that specific posttranslational modifications on histones can be translated into distinct biological outcomes through the actions of epigenetic writers, readers, and erasers. Together with DNA methylation, epigenetic marks on histone residues constitute the messengers that extend beyond the genetic code. Reader proteins bear one or more domains that recognize either specific modifications on histones or unmodified histone residues and are often associated with writer or eraser domains either within the same polypeptide or in a multiunit complex. The dual discoveries that the bromodomain of p300/CBP-associated factor binds to acetyl-lysine (3) and the chromodomain of heterochromatin protein 1 binds to histone H3 methylated at lysine 9 (4–6) functionally linked these reader domains to the regulation of gene transcription.

Shortly after the discovery of the chromodomain, the superfamily of histone methyl-lysine readers rapidly expanded to include the Tudor domain (7, 8), plant homeodomain (PHD) (9–11), malignant brain tumor domain (12, 13), ankyrin repeats (14), the proline-tryptophan-tryptophan-proline (PWWP) domain (15–17), bromo-adjacent homology domain (18–20), and the cysteine-tryptophan (CW) domain (21). In contrast to these methyl-lysine readers, the PHD domain of BHC80 and ADD domain of DNMT3A/B/L recognize unmethylated H3 lysine 4 (H3K4me0) (22, 23).

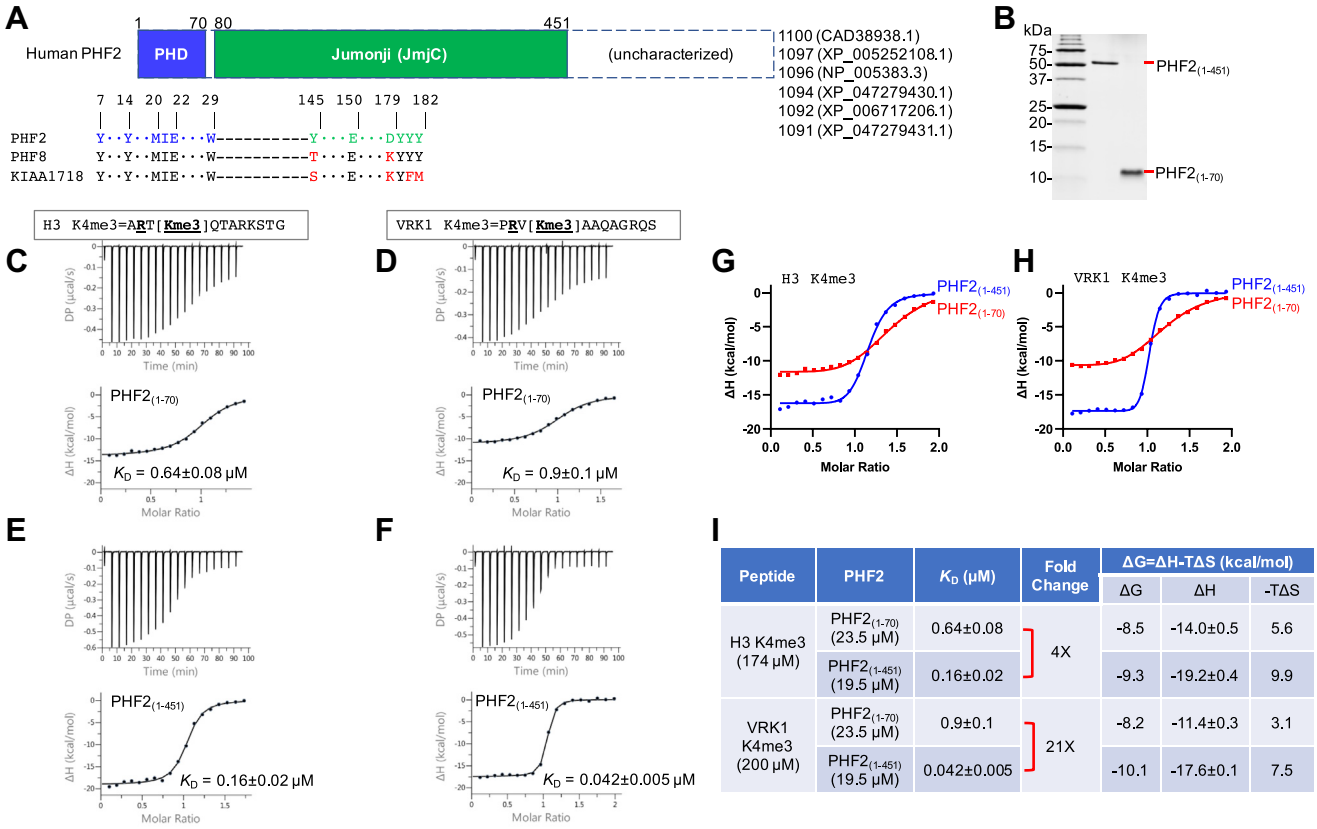
In this study, we examined the N-terminal half of PHF2 (residues 1–451) that harbors both the PHD and Jumonji domains (Fig. 1*A*). PHF2 belongs to the KDM7 family of Jumonji demethylases, which are Fe(II) and α-ketoglutarate–dependent members of a broader family of dioxygenases (24). The KDM7 family has three members in mice and humans, PHF2, PHF8, and KIAA1718 (25). PHF2 binds to histone H3 trimethylated at lysine 4 (H3K4me3) through its PHD and demethylates H3K9me2/m1 *via* its Jumonji domain (26, 27). In contrast to PHF8 and KIAA1718, PHF2 is enzymatically inactive *in vitro* (based on mass spectrometry– and fluorescence-based demethylase assays on peptide substrates) (28) and remains inactive until it is phosphorylated by protein kinase A (29). In addition to its histone substrates, PHF2 also binds to methylated nonhistone proteins that mimic the histone H3 N-terminal tail sequence (30) including vaccinia-related kinase 1 (VRK1), a serine/threonine protein kinase (31). PHF2 forms a complex with the DNA-interacting protein ARID5B, resulting in both demethylation of ARID5B by PHF2 (29) and subsequent recruitment of the PHF2–ARID5B complex to select H3K9me2-marked promoters of specific genes regulated by transcription factor Sox9 (32) or the lipogenic transcription factor ChREBP (carbohydrate-responsive element binding protein) (27). Here, we sought to better understand the mechanism of interaction between PHF2 and H3K4me3 and how that interaction influences PHF2 demethylase activity on H3K9me2.

Results

PHF2 has stronger binding affinity for K4me3 when both the PHD and Jumonji domains are present

We purified two PHF2 recombinant fragments, one containing the PHD domain (residues 1–70), hereafter $PHF2_{(1–70)}$, and another containing both the PHD and Jumonji domains (residues 1–451), hereafter $PHF2_{(1–451)}$ (Fig. 1*B*). We tested the

* For correspondence: Xiaodong Cheng, XCheng5@mdanderson.org.



Peptide	PHF2	K_D (μM)	Fold Change	ΔG=ΔH-TΔS (kcal/mol)		
				ΔG	ΔH	-TΔS
H3 K4me3 (174 μM)	$PHF2_{(1-70)}$ (23.5 μM)	0.64±0.08	4X	-8.5	-14.0±0.5	5.6
	$PHF2_{(1-451)}$ (19.5 μM)	0.16±0.02		-9.3	-19.2±0.4	9.9
VRK1 K4me3 (200 μM)	$PHF2_{(1-70)}$ (23.5 μM)	0.9±0.1	21X	-8.2	-11.4±0.3	3.1
	$PHF2_{(1-451)}$ (19.5 μM)	0.042±0.005		-10.1	-17.6±0.1	7.5

Figure 1. $PHF2_{(1–451)}$ has a stronger binding with the methylated K4me3 peptides. *A*, schematic representation of human PHF2 isoforms with variations at the C-terminus. The aromatic cage residues for K4me3 and R2-interacting acidic residues are indicated below. *B*, the two recombinant fragments of PHF2 used in the study, shown in an 18% SDS polyacrylamide gel. *C* and *D*, ITC measurements of $PHF2_{(1–70)}$ against two K4me3 peptides (N = 2). *E* and *F*, ITC measurements of $PHF2_{(1–451)}$ against two K4me3 peptides (N = 2). *G* and *H*, overlay of ITC fitting curves of the two PHF2 fragments (short and long) against H3 peptide (panel *G*) and VRK1 peptide (panel *H*). *I*, summary of PHF2 protein and peptide concentrations used in the ITC measurements and the derived fitting parameters. ITC, isothermal titration calorimetry; VRK1, vaccinia-related kinase 1.

binding of these PHF2 fragments to two synthesized 12-mer peptides, one corresponding to the N-terminal tail of histone H3 and another corresponding to the N-terminus of VRK1. The N-terminal amino acid sequences of both H3 (AR_2TK_4) and VRK1 (PR_2VK_4) bear an arginine at position 2 and lysine at position 4 within a zRxK motif (z = a small amino acid, x = any amino acid) (30), a motif that is permissive for initiator methionine cleavage by methionine aminopeptidases. To assess their binding affinities quantitatively, we used isothermal titration calorimetry (ITC) to measure the dissociation constants (K_D values) of $PHF2_{(1–70)}$ and $PHF2_{(1–451)}$, with the H3 and VRK1 peptides trimethylated at lysine 4 (K4me3). We found that $PHF2_{(1–70)}$ binds the two 12-mer peptides equally well, with similar K_D values, 0.64 μM for H3 and 0.9 μM for VRK1 (Fig. 1, *C* and *D*).

Next, we measured the binding affinities between $PHF2_{(1–451)}$ and the methylated H3 and VRK1 peptides (Fig. 1, *E* and *F*). For the H3 peptide, $PHF2_{(1–451)}$ exhibited 4 × increased binding affinity relative to that of $PHF2_{(1–70)}$ (K_D = 0.16 μM) (Fig. 1, *G* and *I*). For the VRK1 peptide, the affinity was enhanced approximately 21 × (K_D = 42 nM) (Fig. 1, *H* and *I*). These results indicate that the dual-domain fragment $PHF2_{(1–451)}$ may preferentially bind VRK1 compared to H3, as demonstrated by a 3.8× greater affinity for VRK1 relative to H3 (42 nM vs. 160 nM), whereas the isolated PHD domain $PHF2_{(1–70)}$ may slightly prefer H3 over VRK1 as shown by a 1.4× greater affinity for H3 relative to VRK1 (0.64 μM vs. 0.9 μM). This difference in binding affinity is consistent with our previously published peptide pull-down experiments (30). The relative amount of full-length endogenous PHF2 recovered from HEK293T cells was greater for VRK1-K4me3 peptide pulldowns than H3K4me3 peptide pulldowns, whereas in pulldowns performed using a GST fusion to only the PHD domain of PHF2, the H3 association was favored over the VRK1 association (30). We also note that the ∼40 nM binding affinity is one of the strongest we know for any methylated peptide characterized so far.

An aromatic K4me3-binding cage is formed by the PHF2 PHD and Jumonji domain residues

To understand the structural basis for the high-affinity interaction between $PHF2_{(1–451)}$ and K4me3 peptides, we first examined two previously characterized structures of the PHF2 PHD in complex respectively with K4me3 peptides of histone H3 and VRK1 (Fig. 2, *A* and *B*) (26, 30). These two

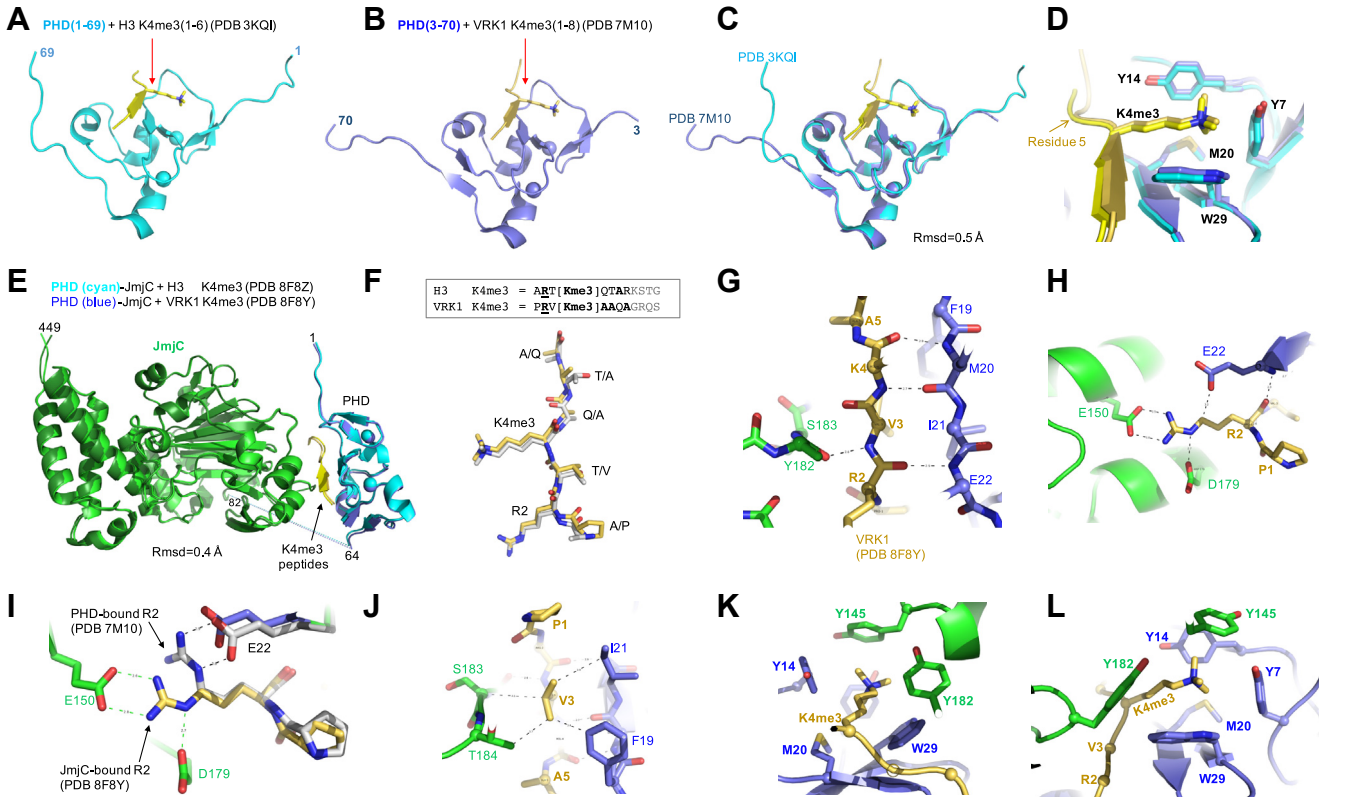


Figure 2. The PHD and Jumonji domains of PHF2 contribute to the binding of K4me3 peptides. *A*, PHD domain binds H3K4me3 peptide. *B*, PHD domain binds VRK1 K4me3 peptide. *C*, superimposition of the PHD domain structures complexed with H3 and VRK1 peptides. *D*, K4me3 sits in an open cage. The wall of the cage facing the reader is open. *E*, superimposition of PHF2$_{(1–451)}$ complexed with H3 and VRK1 peptides, which are held in between the two domains. *F*, superimposition of H3 and VRK1 peptides and their corresponding sequences (*top*). *G*, VRK1 peptide makes main-chain interactions with both PHD (*blue*) and Jumonji (*green*). *H*, VRK1 R2 residue interacts with three acidic residues (two from the Jumonji domain and one from the PHD domain). *I*, conformational change of VRK1 R2 residue from PHD-bound form (E22) to Jumonji-bound form (E150 and D179). *J*, VRK1 V3 residue interacts with F19 and I21 of PHD (*blue*) and S183 and T184 of Jumonji (*green*). *K* and *L*, two views of K4me3 encircled by five aromatic and one methionine. Y145 and Y182 (*green*) close off the open cage. Panels *L* and *D* are presented in a similar orientation. H3K4me3, histone H3 trimethylated at lysine 4; PHD, plant homeodomain; VRK1, vaccinia-related kinase 1.

structures were determined at high resolutions of 1.78 Å (PDB 3KQI) and 1.15 Å (PDB 7M10). The first six residues of H3 were observed in PDB 3KQI (Fig. 2*A*), whereas the first eight residues of VRK1 were observed in PDB 7M10 (Fig. 2*B*). Superimposition of the two structures revealed a rmsd of 0.5 Å, with conformational variations at the N- and C-termini residues (Fig. 2*C*). The bound peptides were overlaid across the first five residues, with the K4me3 sitting in an open cage formed by Y7, Y14, M20, and W29 of the PHD domain (Fig. 2*D*). Notably, these four cage residues are invariant in the other two KDM7 family members (PHF8 and KIAA1718) (Fig. 1*A*), both known to bind H3K4me3 peptides (33–36).

To better define the basis for the recognition and enhanced binding of the K4me3 peptides by PHF2$_{(1–451)}$ compared to PHF2$_{(1–70)}$, we cocrystallized PHF2$_{(1–451)}$ with each of the two H3 and VRK1 K4me3-containing peptides (Fig. 2*E*). The two complexes crystallized in space group $P2_1$, and the structures were determined to resolutions of 3.3 Å (H3) and 3.06 Å (VRK1), respectively (Table S1). For PHF2$_{(1–451)}$, we observed the complete N-terminus; however, two residues (450–451) were missing at the C-terminus, and there was a disordered linker (residues 65–81) between the PHD and Jumonji domains (Figs. 2*E* and S1). In addition, although they were used in the crystallization mixture, we did not observe Mn(II) or α-ketoglutarate in the structures, or any K4me3-containing peptide binding to the active site of the Jumonji domain. Instead, we observed that the methylated peptide bound at the PHD-Jumonji interface (Fig. 2*E*), with the first eight residues of the peptide being ordered but with residue 9 and beyond being unstructured. Unexpectedly, the side chains of Q5 and R8 of histone H3 and Q7 of VRK1 were not observed due to lack of electron density, therefore these residues of the two peptides were modeled as alanine (Fig. 2*F*).

In the following section, we describe the PHF2$_{(1–451)}$ interactions with the VRK1 peptide, with a focus on the interactions involving both the PHD and Jumonji domains. First, the VRK1 peptide displayed an extended conformation with the main-chain atoms of VRK1 residues R2-A5 forming an anti-parallel β-strand with residues F19-E22 of the PHD domain (blue in Fig. 2*G*). Further, the main-chain carbonyl oxygen of Y182 in the Jumonji domain formed a hydrogen bond with the main-chain amide nitrogen atom of VRK1 residue V3. Second, the positively charged R2 of VRK1 is surrounded by three acidic residues of PHF2$_{(1–451)}$: E22 of the PHD and E150 and D179 of the Jumonji domain (Fig. 2*H*). Compared with the structure of VRK1 bound to PHF2$_{(1–70)}$,

the VRK1 R2 side chain in complex with $PHF2_{(1–451)}$ undergoes a conformational change such that instead of interacting with E22 as in $PHD_{(1–70)}$, it interacts with E150 and D179 of the Jumonji domain in $PHF2_{(1–451)}$ (Fig. 2*I*). We note that, while E150 is conserved, D179 of PHF2 is substituted by a lysine in both PHF8 and KIAA1718 (Fig. 1*A*). This implies that the peptide-bound conformation observed between the PHD and Jumonji domains of PHF2 will not occur with either PHF8 or KIAA1718 due to electrostatic repulsion between the PHF8 and KIAA1718 lysine residues and the VRK1 peptide R2 residue. Third, the hydrophobic residue V3 of VRK1 makes van der Waals contacts with F19 and I21 of the PHD domain, the main-chain Cα atom of S183, and the side chain of T184 located within the Jumonji domain (Fig. 2*J*). Fourth, the K4me3 moiety on VRK1 is encircled by a cage of six amino acids (five aromatic residues and one methionine residue) (Fig. 2, *K* and *L*). In addition to the aforementioned cage residues (Y7, Y14, M20, and W29) of the PHD domain, two tyrosine residues (Y145 and Y182) of the Jumonji domain complete the cage by forming the final cage wall, protecting the cationic and aliphatic features of Kme3.

Finally, these observations are also true for the first four residues of histone H3 peptide (ARTK), as observed in the $PHF2_{(1–451)}$-H3 peptide complex structure. While the polar side chain of T3 of histone H3 may form a weak hydrogen bond (3.4 Å) with T184 of PHF2, the current structures of $PHF2_{(1–451)}$ do not provide a clear explanation for the difference observed in the binding affinity of the VRK1 and H3 peptides (*i.e.*, 3.8 × difference in K_D values; Fig. 2), due to the lacking electron density for the side chains of peptide residues 5 to 8 and disordered peptide beyond residue 9, where the differences between the two peptide sequences occur.

Effect of K4me3 binding on PHF2 activity on K9me2

Next, we assessed PHF2 demethylase activity on H3K9me1/2/3, in the presence and absence of added K4me3-containing peptide *in trans* (Fig. 3*A*), by Western blotting using antibodies validated for each of H3K9me1, H3K9me2, and H3K9me3 (37). We note that the published PHF2 activities were previously only detected by antibodies (26, 27, 29); and although PHF2 activity on H3K9me1 was initially reported (26) it was not confirmed by later experiments (27, 29). Using the recombinant $PHF2_{(1–451)}$ and core histones, we found that PHF8 (as a positive control) and $PHF2_{(1–451)}$ (in the presence and absence of K4me3 peptide) actively demethylate H3K9me2, that is, the decreased amount of H3K9me2 accompanies an increased amount of H3K9me1, which is the

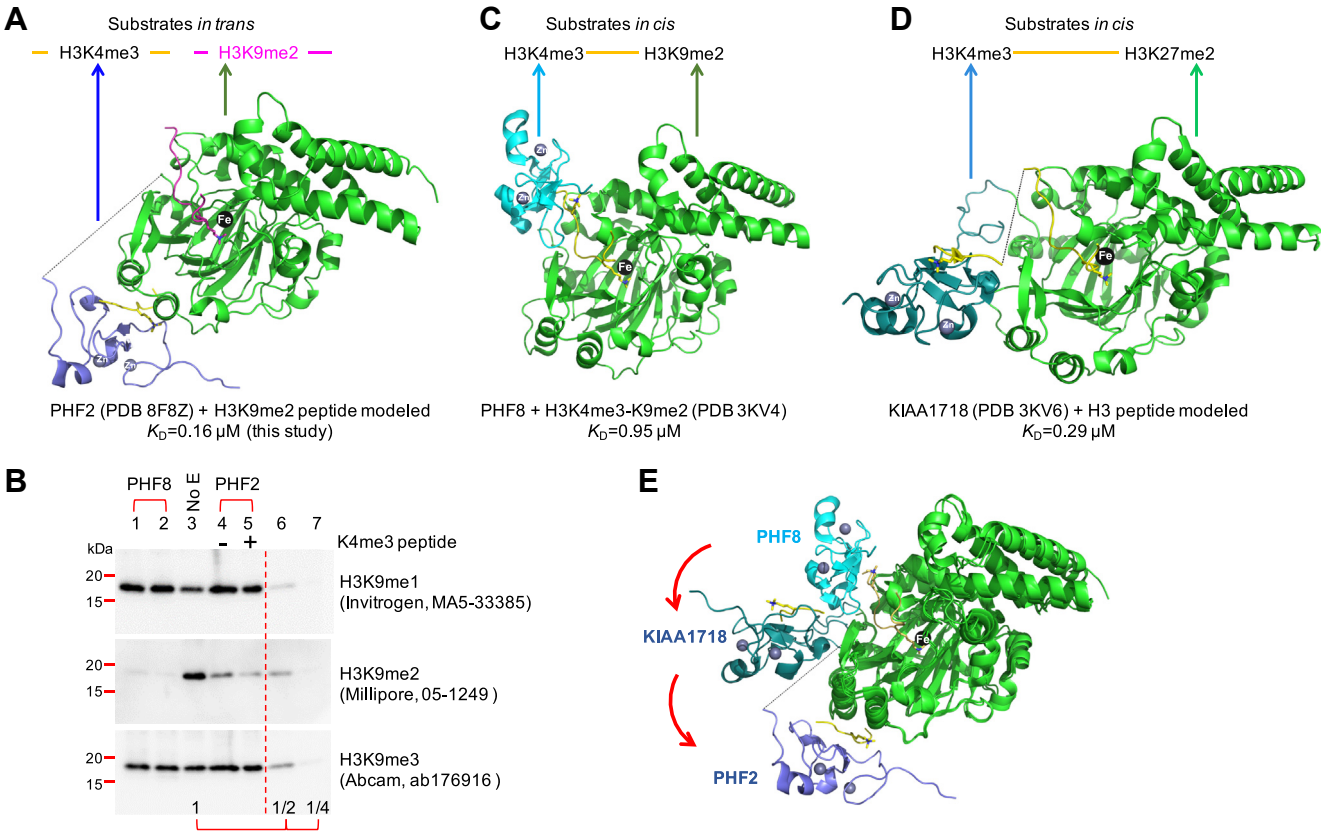


Figure 3. The effects of PHD-H3K4me3 interaction on the demethylase activity of the linked Jumonji domain. *A*, structure of $PHF2_{(1–451)}$ bound with H3K4me3 (*yellow*) and a modeled H3K9me2 (*magenta*) *in trans*. *B*, demethylase activities on core histones of PHF8 (lanes 1 and 2) and $PHF2_{(1–451)}$, in the absence (lane 4) and presence of (lane 5) of K4me3 peptide added *in trans* (N = 4 independent replicates). Lanes 3, 6, and 7 are the 2× dilution of reaction mixture without enzymes (No E). Note the decrease in H3K9me2 corresponds with an increase in H3K9me1. *C*, the structure of $PHF8_{(1–447)}$ bound with a single H3 peptide (*yellow*) containing both modifications of H3K4me3 and H3K9me2. *D*, structure of $KIAA1718_{(1–488)}$ with a modeled H3 peptide (*yellow*) of H3K4me3 and H3K27me2. *E*, superimposition of PHF8, KIAA1718, and PHF2 indicating three different locations of PHD domain in relation to Jumonji (*green*). The active sites of the Jumonji domains are labeled by metal ion Fe(II). H3K4me3, histone H3 trimethylated at lysine 4; PHD, plant homeodomain.

demethylation product of H3K9me2 (comparing lanes 1, 2, 4, and 5 to lane 3 in Fig. 3*B*). It seems that the presence of K4me3 peptide *in trans* may slightly enhance PHF2 activity on H3K9me2 (comparing lanes 4 and 5). However, due to the limitation of our Western blotting assay, we were unable to accurately quantify this difference.

Discussion

Here, we described the tight binding of K4me3 peptides derived from histone H3 and VRK1 to PHF2 *via* interactions with its PHD and Jumonji domains, which are connected by a flexible linker. Key residues contributing to the interaction of the methylated peptide with PHF2 reside within the first four residues, A$\underline{R}_2$T$\underline{K}_4$ for H3 and P$\underline{R}_2$V$\underline{K}_4$ for VRK1. The positively charged R2 residue engages with three acidic residues, one from the PHD domain and two from the Jumonji domain, and the methylated lysine (K4me3) is encircled by five aromatic residues, three from the PHD domain and two from the Jumonji domain, as well as an additional methionine. These interactions result in very high, if not the highest affinity of any methyl-lysine readers characterized to date.

Spatial limitations may impact KDM7 family demethylase activity and substrate choice

H3K4me3 is a histone modification associated with transcriptional activation, whereas H3K9me2 and H3K27me2 are associated with transcriptional repression (38). Previous studies have established that the PHD domains of the three KDM7 family members, PHF2, PHF8, and KIAA1718, bind H3K4me3-containing peptides (26, 33–36). However, the effects of PHD–H3K4me3 interaction on the enzymatic activity of the linked Jumonji domains on repressive marks H3K9me2 and/or H3K27me2 vary among the three. For PHF8, the interaction of the PHD domain with H3K4me3 and the Jumonji domain with H3K9me2 (when both modifications exist *in cis*) makes the histone peptide a better substrate for H3K9me2 demethylation (35) (Fig. 3*C*). In contrast, for KIAA1718, the enzyme is inactive on H3K9me2 when H3K4me3 is present *in cis* but becomes more selective toward H3K27me2 (35) (Fig. 3*D*). The relative configuration between the two domains renders a closed conformation for PHF8 on H3K4me3 and H3K9me2 and an open conformation of KIAA1718, enabling it to engage two methyl marks separated by more distance *in cis* on the same peptide. Thus, the structural relationship between PHD-H3K4me3 binding and the position of the catalytic Jumonji domains relative to H3K4me3 determines which repressive mark (H3K9me2 or H3K27me2) is removed by PHF8 and KIAA1718.

Superimposition of the three demethylases shows that the PHD domain of PHF2 is located at the greatest distance from the active-site of the linked Jumonji (Fig. 3*E*), and thus in contrast to PHF8, PHF2 cannot act on a single histone peptide bearing H3K4me3 and H3K9me2 *in cis* (28). We note that H3K9me2 is the only histone methyl-lysine substrate of PHF2 identified so far (27, 29). It is plausible that the PHD and Jumonji domains of PHF2 could interact *in trans* with two different histones within the same nucleosome.

Increased binding affinity is related to the number of aromatic residues in the PHD domain

Despite the vast diversity of methyl-lysine readers, the structurally characterized methyl-lysine reader domains commonly form an aromatic cage that recognizes both the cationic and hydrophobic features of methyl-lysine, as first noted in the original chromodomain structure of heterochromatin protein 1 (6) (and reviewed in (39, 40)). To better understand the interaction of the PHF2 domain with its substrates, we reviewed the existing structures of several PHD domains, including PHF2 (the focus of our study). These structures represent examples of PHD domains bound to H3K4 bearing no methylation, dimethylation, and trimethylation, as well as acetylated lysine 16 of histone H4 (H4K16).

The PHD domain of the histone-binding protein BHC80 (PHF21A) binds to unmethylated histone H3 lysine 4 (H3K4me0) *via* a negatively charged aspartate (D) side chain that bridges with the positively charged side chains of H3K4 and H3R8 and a hydrophobic methionine (M) which aligns with the aliphatic portion of H3K4 (Fig. 4*A*) (22). The PHD domain of SET3 binds H3K4me2/me3 by substituting a tryptophan (W) for the methionine (41), forming the first wall of an aromatic cage, positioned to the left of the methyl-lysine in Figure 4*B*. The PHD domain of PHF20 recognizes H3K4me2 by forming the second wall of a cage with a methionine behind the methyl-lysine and with a tryptophan to the left, as oriented in Figure 4*C* (42). The PHD domain of ING2 binds H3K4me3/me2 by constructing a left wall with a tryptophan, a back wall with a methionine, and the floor of the cage with a tyrosine residue (W-M-Y as oriented in Fig. 4*D*) (11). Interestingly, the PHD6 domain of MLL4 recognizes acetylated lysine 16 of histone H4 *via* W-L-L residues positioned in a similar configuration to that of ING2 (Fig. 4*E*) (43). Finally, the PHD domain of PHF2 recognizes H3K4me3 by adding a second tyrosine as the right-side wall of the cage, to the right of the methyl-lysine (W-M-F-F as oriented in Fig. 4*F*).

We note that these studies were carried out with the isolated reader domains (∼70 residues) extracted from often large proteins (∼1100 residues in the case of PHF2 and ∼5500 residues in the case of MLL4). Overall, we find that an increased number of aromatic residues (0–3) correlates with an increased binding affinity (*i.e.*, decreased dissociation constant K_D values varying from 20–30 μM to 1–2 μM) (Fig. 4). This is consistent with the idea that the formation of a complete aromatic cage will require additional aromatic (or hydrophobic) residues not present in the isolated PHD domains, just as in PHF2, where both the PHD and the Jumonji domains contribute to the formation of the cage.

Experimental procedures

Expression and purification of PHF2 proteins

GST-tagged human PHF2$_{(1–451)}$ from construct pXC816 (28) was expressed in BL21(DE3)CodonPlus cells using

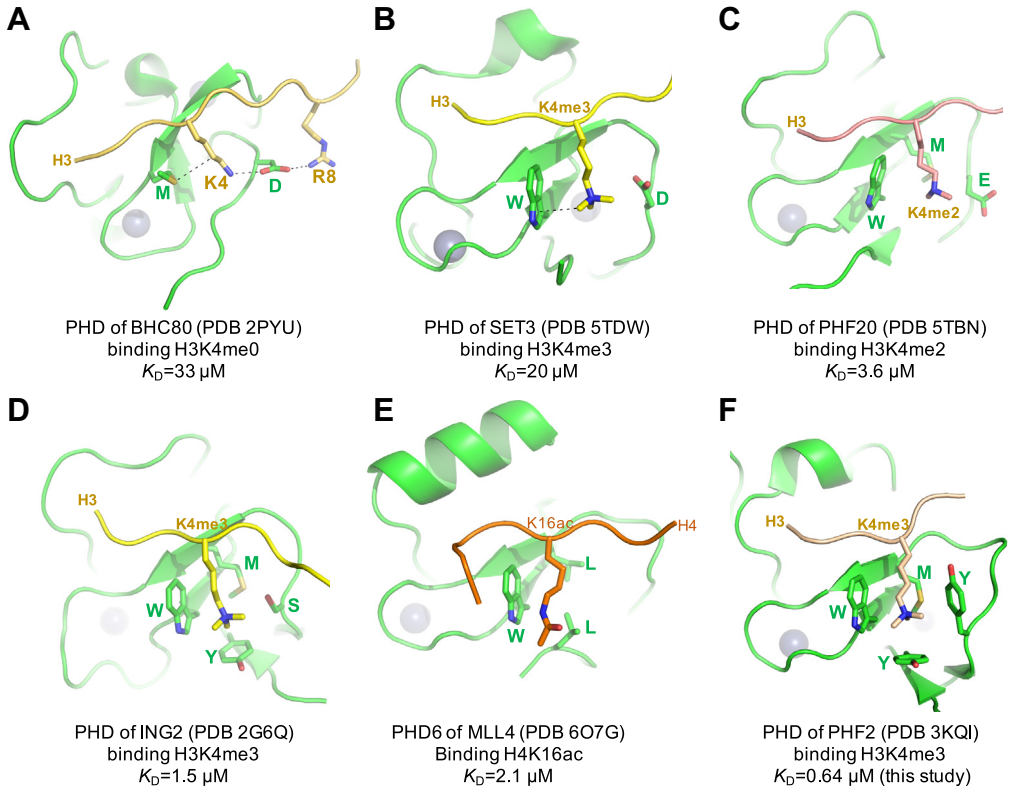


Figure 4. Examples of PHD domain interactions with different states of H3K4 methylation (me0 to me3) or H4K16 acetylation. *A*, human BHC80 (PDB 2PYU), (*B*) *Saccharomyces cerevisiae* Set3 (PDB 5TDW), (*C*) human PHF20 (PDB 5TBN), (*D*) mouse ING2 (PDB 2G6Q), (*E*) human MLL4 (PDB 6O7G), and (*F*) human PHF2 (PDB 3KQI). The structures have been aligned so that the target lysine residues are in a similar orientation, to highlight similarities and differences. M, methionine, W, tryptophan, Y, tyrosine, D, aspartate, E, glutamate, and S, serine. The aliphatic portion of lysine side chain is stacked against hydrophobic methionine (*A*) or aromatic tryptophan (*B*) or combination of both M and W (*C*). The positive charge of lysine or methyl-lysine is balanced by an acidic residue. Addition of one or two tyrosine residues completes the aromatic cage formation (*D* and *F*). The K_D values were taken from respective studies, except for $PHF2_{(1–70)}$ (this study). PHD, plant homeodomain.

auto-induction medium (44). Briefly, an overnight cell culture in MDAG medium was inoculated into ZYM-5052 medium and cultured at 37 °C until the A_{600} equaled 1. Temperature was changed to 22 °C and the cells were cultured for 18 h for autoinduction of target protein. Protein was purified as previously described (28). The protein was concentrated to ∼38 mg/ml in 20 mM Hepes, pH 8.0, 300 mM NaCl, 5% glycerol, and 0.5 mM tris(2-carboxyethyl)phosphine and flash frozen for storage at −80 °C.

For expression of $PHF2_{(1–70)}$, a stop codon was inserted using PCR techniques for the histidine codon at residue 71 in pXC816 to create the new construct pXC2227. The PHD domain was expressed and purified similarly to that described above except that the PHD domain does not bind tandem HiTrap Q/SP columns during purification. Therefore, the flow-through was collected, concentrated, and loaded onto the S200 Sepharose sizing column with a buffer containing 20 mM Tris–HCl (pH 7.5) and 200 mM NaCl. The one major peak from the sizing column was collected and concentrated to approximate concentrations of 10 and 30 mg/ml and flash frozen for storage at −80 °C prior to use.

Isothermal titration calorimetry

Experiments were performed with a MicroCal PEAQ-ITC automated system (Malvern). Experiments were conducted at 25 °C with a reference power of 10 mcal/s. Nineteen injections of peptide of VRK1 or H3 with initial injection of 0.2 ml followed by 18 injections (each of 2 ml) were titrated into PHF2 proteins in 300 mM NaCl, 20 mM Hepes pH 8.0, 0.5 mM tris(2-carboxyethyl) phosphine, and 5% glycerol. ITC data were fitted as 'one site' with the offset subtracted. Binding constants were calculated using the ITC analysis module supplied by the manufacturer.

Crystallography

Concentrated $PHF2_{(1–451)}$ was thawed, and $MnCl_2$ and α-ketoglutarate were added such that the molar ratio of protein:Mn(II):α-ketoglutarate was 1:10:10 and allowed to sit on ice for ∼30 min. At this point, either VRK1(1-12)K4me3, H3(1-12) K4me3 (or H3(1-24)K4me3K9me2) peptide was added to produce a mixture with an approximate 1:2 protein:peptide molar ratio and let to sit on ice for at least an additional hour. Just before crystallization trials, the mixture was diluted with protein storage buffer such that the protein concentration was between 10 to 20 mg/ml. The Art Robbins Gryphon Crystallization Robot was used to set up 0.4-μl sitting drops (0.2 μl of complex plus 0.2 μl of well solution) at ∼19 °C utilizing commercial screens from Hampton Research and Molecular Dimensions. The only observed crystals were conglomerations of inseparable and twinned needles in conditions with >1.9 M ammonium sulfate including buffers with pH 5.5 to 7.0. However, further screening gave some small single crystals that could be separated from the clusters with a well solution of 2.8 M ammonium sulfate and

100 mM Bis–Tris propane buffer (pH 6.3); the addition of 4% of the additive pentaerythritol ethoxylate (3/4 EO/OH) (Hampton Research, Inc) seemed to give a slight increase in the appearance of single needles. These needles were picked up quickly in a nylon loop and momentarily placed into well solution supplemented with 20% (v/v) ethylene glycol to limit ammonium sulfate crystal formation before plunging into liquid nitrogen for cryoprotection.

Crystals were screened and X-ray diffraction data were collected at the SER-CAT beamline 22ID of the Advanced Photon Source at Argonne National Laboratory. Crystallographic datasets were processed with HKL2000 (45). For diffraction from the crystal of PHF2 with H3 peptide, processing with the DIALS package (46) offered a dataset with slightly better statistics which was utilized for structure refinement. Molecular replacement was performed with the PHENIX PHASER module (47) by using the known structure of the PHF2 PHD domain (PDB 7M10) and PHF2 Jumonji domain (PDB 3PUS) as search models. Structure refinement was performed with PHENIX Refine (48) with 5% randomly chosen reflections for the validation by the free R value (49). COOT (50) was used for the initial manual building of peptide using obvious difference electron density and then for corrections of protein and peptide between rounds. Structure quality was analyzed during PHENIX refinements and finally validated by the PDB validation server (51). Molecular graphics were generated by using PyMol (Schrodinger, LLC).

PHF2 activity on H3K9me2 by Western blotting

The *in vitro* demethylase assay used conditions similar to those described (27). A master mix was made, such that upon addition of enzyme, a 40 μl reaction contained 20 mM Tris–HCl at pH 7.5, 150 mM KCl, 50 μM $(NH_4)_2Fe(SO_4)_2$-$6H_2O$, 1 mM α-ketoglutarate, 1 mM ascorbate, and ∼4 μg core histones. Each reaction contained either no enzyme, 5 μM $PHF8_{(80–447)}$, $PHF8_{(1–447)}$, $PHF2_{(1–451)}$, with or without 25 μm K4me3 peptide. The PHF8 enzymes were previously purified in-house (35). Reactions were incubated at 37 °C overnight. The reaction mixture was denatured by heating in SDS sample buffer for 5 min and subjected to Western blotting. Briefly, the denatured samples were separated with a 4 to 20% precast polyacrylamide gel (BioRad, Cat. #4561096), which was then transferred to a precut low fluorescence polyvinylidene difluoride membrane (BioRad, #1620261). After blocking with 5% nonfat dry milk in Tris-buffered saline with Tween 20 at room temperature for 1 h, the membrane was incubated with primary antibody (1:1000 dilution) against H3K9me1 (Invitrogen, MA5-33385), H3K9me2 (Millipore, 05-1249) and H3K9me3 (Abcam, ab176916), and secondary antibody (1:5000 dilution) against Mouse-IgG (Abcam, ab6820) or Rabbit-IgG (Cell Signaling, #7074), respectively.

Data availability

The X-ray structures (coordinates and structure factor files) of the PHF2 (PHD-Jumonji) with bound peptides have been deposited to PDB and are publicly available as of the date of publication. Accession numbers are PDB 8F8Y (VRK1) and PDB 8F8Z (H3).

Supporting information—This article contains supporting information.

Acknowledgments—We thank Dr Briana Dennehey for editing the article and her insightful comments. We thank the beamline scientists of the Southeast Regional Collaborative Access Team (SER-CAT) at the Advanced Photon Source (APS), Argonne National Laboratory, USA. The use of SER-CAT is supported by its member institutions and equipment grants (S10_RR25528, S10_RR028976, and S10_OD027000) from the US National Institutes of Health. Use of the APS was supported by the U.S. Department of Energy, Office of Science, Office of Basic Energy Sciences, under contract W-31-109-Eng-38.

Author contributions—J. R. H., J. Z., Q. C., and M. T. B. investigation; X. Z. supervision; X. Z., M. T. B., and X. C. conceptualization; X. Z. project administration; X. C. funding acquisition; X. C. writing–original draft; X. C. and M. T. B. writing–review and editing; X. C. methodology.

Funding and additional information—The work was supported by the US National Institutes of Health (R35GM134744 to X. C.) and the Cancer Prevention and Research Institute of Texas (RR160029 to X. C.). M. T. B. is supported by NIH GM126421 and CPRIT RP180804. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

Conflict of interest—X. C. is a CPRIT Scholar in Cancer Research. M. T. B. is a co-founder of EpiCypher. The other authors declare that they have no conflicts of interests with the contents of this article.

Abbreviations—The abbreviations used are: H3K4me3, histone H3 trimethylated at lysine 4; ITC, isothermal titration calorimetry; PHD, plant homeodomain; VRK1, vaccinia-related kinase 1.

References

1. Strahl, B. D., and Allis, C. D. (2000) The language of covalent histone modifications. *Nature* **403**, 41–45
2. Jenuwein, T., and Allis, C. D. (2001) Translating the histone code. *Science* **293**, 1074–1080
3. Dhalluin, C., Carlson, J. E., Zeng, L., He, C., Aggarwal, A. K., and Zhou, M. M. (1999) Structure and ligand of a histone acetyltransferase bromodomain. *Nature* **399**, 491–496
4. Bannister, A. J., Zegerman, P., Partridge, J. F., Miska, E. A., Thomas, J. O., Allshire, R. C., *et al.* (2001) Selective recognition of methylated lysine 9 on histone H3 by the HP1 chromo domain. *Nature* **410**, 120–124
5. Lachner, M., O'Carroll, D., Rea, S., Mechtler, K., and Jenuwein, T. (2001) Methylation of histone H3 lysine 9 creates a binding site for HP1 proteins. *Nature* **410**, 116–120
6. Jacobs, S. A., and Khorasanizadeh, S. (2002) Structure of HP1 chromodomain bound to a lysine 9-methylated histone H3 tail. *Science* **295**, 2080–2083
7. Huang, Y., Fang, J., Bedford, M. T., Zhang, Y., and Xu, R. M. (2006) Recognition of histone H3 lysine-4 methylation by the double tudor domain of JMJD2A. *Science* **312**, 748–751
8. Botuyan, M. V., Lee, J., Ward, I. M., Kim, J. E., Thompson, J. R., Chen, J., *et al.* (2006) Structural basis for the methylation state-specific recognition

of histone H4-K20 by 53BP1 and Crb2 in DNA repair. *Cell* **127**, 1361–1373
9. Wysocka, J., Swigut, T., Xiao, H., Milne, T. A., Kwon, S. Y., Landry, J., *et al.* (2006) A PHD finger of NURF couples histone H3 lysine 4 trimethylation with chromatin remodelling. *Nature* **442**, 86–90
10. Pena, P. V., Davrazou, F., Shi, X., Walter, K. L., Verkhusha, V. V., Gozani, O., *et al.* (2006) Molecular mechanism of histone H3K4me3 recognition by plant homeodomain of ING2. *Nature* **442**, 100–103
11. Shi, X., Hong, T., Walter, K. L., Ewalt, M., Michishita, E., Hung, T., *et al.* (2006) ING2 PHD domain links histone H3 lysine 4 methylation to active gene repression. *Nature* **442**, 96–99
12. Li, H., Fischle, W., Wang, W., Duncan, E. M., Liang, L., Murakami-Ishibe, S., *et al.* (2007) Structural basis for lower lysine methylation state-specific readout by MBT repeats of L3MBTL1 and an engineered PHD finger. *Mol. Cell* **28**, 677–691
13. Grimm, C., de Ayala Alonso, A. G., Rybin, V., Steuerwald, U., Ly-Hartig, N., Fischle, W., *et al.* (2007) Structural and functional analyses of methyl-lysine binding by the malignant brain tumour repeat protein sex comb on midleg. *EMBO Rep.* **8**, 1031–1037
14. Collins, R. E., Northrop, J. P., Horton, J. R., Lee, D. Y., Zhang, X., Stallcup, M. R., *et al.* (2008) The ankyrin repeats of G9a and GLP histone methyltransferases are mono- and dimethyllysine binding modules. *Nat. Struct. Mol. Biol.* **15**, 245–250
15. Dhayalan, A., Rajavelu, A., Rathert, P., Tamas, R., Jurkowska, R. Z., Ragozin, S., *et al.* (2010) The Dnmt3a PWWP domain reads histone 3 lysine 36 trimethylation and guides DNA methylation. *J. Biol. Chem.* **285**, 26114–26120
16. Qin, S., and Min, J. (2014) Structure and function of the nucleosome-binding PWWP domain. *Trends Biochem. Sci.* **39**, 536–547
17. Rondelet, G., Dal Maso, T., Willems, L., and Wouters, J. (2016) Structural basis for recognition of histone H3K36me3 nucleosome by human de novo DNA methyltransferases 3A and 3B. *J. Struct. Biol.* **194**, 357–367
18. Kuo, A. J., Song, J., Cheung, P., Ishibe-Murakami, S., Yamazoe, S., Chen, J. K., *et al.* (2012) The BAH domain of ORC1 links H4K20me2 to DNA replication licensing and Meier-Gorlin syndrome. *Nature* **484**, 115–119
19. Qian, S., Lv, X., Scheid, R. N., Lu, L., Yang, Z., Chen, W., *et al.* (2018) Dual recognition of H3K4me3 and H3K27me3 by a plant histone reader SHL. *Nat. Commun.* **9**, 2425
20. Ren, W., Fan, H., Grimm, S. A., Kim, J. J., Li, L., Guo, Y., *et al.* (2021) DNMT1 reads heterochromatic H4K20me3 to reinforce LINE-1 DNA methylation. *Nat. Commun.* **12**, 2490
21. Liu, Y., Tempel, W., Zhang, Q., Liang, X., Loppnau, P., Qin, S., *et al.* (2016) Family-wide characterization of histone binding abilities of human CW domain-containing proteins. *J. Biol. Chem.* **291**, 9000–9013
22. Lan, F., Collins, R. E., De Cegli, R., Alpatov, R., Horton, J. R., Shi, X., *et al.* (2007) Recognition of unmethylated histone H3 lysine 4 links BHC80 to LSD1-mediated gene repression. *Nature* **448**, 718–722
23. Ooi, S. K. T., Qiu, C., Bernstein, E., Li, K., Jia, D., Yang, Z., *et al.* (2007) DNMT3L connects unmethylated lysine 4 of histone H3 to de novo methylation of DNA. *Nature* **448**, 714–717
24. Hausinger, R. P., and Schofield, C. J. (2015) *2-Oxoglutarate-Dependent Oxygenases. RSC Metallobiology Series No. 3*, The Royal Society of Chemistry, London, UK
25. Klose, R. J., Kallin, E. M., and Zhang, Y. (2006) JmjC-domain-containing proteins and histone demethylation. *Nat. Rev. Genet.* **7**, 715–727
26. Wen, H., Li, J., Song, T., Lu, M., Kan, P. Y., Lee, M. G., *et al.* (2010) Recognition of histone H3K4 trimethylation by the plant homeodomain of PHF2 modulates histone demethylation. *J. Biol. Chem.* **285**, 9322–9326
27. Bricambert, J., Alves-Guerra, M. C., Esteves, P., Prip-Buus, C., Bertrand-Michel, J., Guillou, H., *et al.* (2018) The histone demethylase Phf2 acts as a molecular checkpoint to prevent NAFLD progression during obesity. *Nat. Commun.* **9**, 2092
28. Horton, J. R., Upadhyay, A. K., Hashimoto, H., Zhang, X., and Cheng, X. (2011) Structural basis for human PHF2 Jumonji domain interaction with metal ions. *J. Mol. Biol.* **406**, 1–8
29. Baba, A., Ohtake, F., Okuno, Y., Yokota, K., Okada, M., Imai, Y., *et al.* (2011) PKA-dependent regulation of the histone lysine demethylase complex PHF2-ARID5B. *Nat. Cell Biol.* **13**, 668–675
30. Chen, J., Horton, J., Sagum, C., Zhou, J., Cheng, X., and Bedford, M. T. (2021) Histone H3 N-terminal mimicry drives a novel network of methyl-effector interactions. *Biochem. J.* **478**, 1943–1958
31. Valbuena, A., Sanz-Garcia, M., Lopez-Sanchez, I., Vega, F. M., and Lazo, P. A. (2011) Roles of VRK1 as a new player in the control of biological processes required for cell division. *Cell Signal* **23**, 1267–1272
32. Hata, K., Takashima, R., Amano, K., Ono, K., Nakanishi, M., Yoshida, M., *et al.* (2013) Arid5b facilitates chondrogenesis by recruiting the histone demethylase Phf2 to Sox9-regulated genes. *Nat. Commun.* **4**, 2850
33. Feng, W., Yonezawa, M., Ye, J., Jenuwein, T., and Grummt, I. (2010) PHF8 activates transcription of rRNA genes through H3K4me3 binding and H3K9me1/2 demethylation. *Nat. Struct. Mol. Biol.* **17**, 445–450
34. Fortschegger, K., de Graaf, P., Outchkourov, N. S., van Schaik, F. M., Timmers, H. T., and Shiekhattar, R. (2010) PHF8 targets histone methylation and RNA polymerase II to activate transcription. *Mol. Cell Biol.* **30**, 3286–3298
35. Horton, J. R., Upadhyay, A. K., Qi, H. H., Zhang, X., Shi, Y., and Cheng, X. (2010) Enzymatic and structural insights for substrate specificity of a family of Jumonji histone lysine demethylases. *Nat. Struct. Mol. Biol.* **17**, 38–43
36. Kleine-Kohlbrecher, D., Christensen, J., Vandamme, J., Abarrategui, I., Bak, M., Tommerup, N., *et al.* (2010) A functional link between the histone demethylase PHF8 and the transcription factor ZNF711 in X-linked mental retardation. *Mol. Cell* **38**, 165–178
37. Shah, R. N., Grzybowski, A. T., Cornett, E. M., Johnstone, A. L., Dickson, B. M., Boone, B. A., *et al.* (2018) Examining the roles of H3K4 methylation states with systematically characterized antibodies. *Mol. Cell* **72**, 162–177.e7
38. Allis, C. D., Jenuwein, T., and Reinberg, D. (2015) *Epigenetics, Second Edition*, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY
39. Patel, D. J., and Wang, Z. (2013) Readout of epigenetic modifications. *Annu. Rev. Biochem.* **82**, 81–118
40. Vann, K. R., Vishweshwaraiah, Y. L., Dokholyan, N. V., and Kutateladze, T. G. (2021) Searching for methyllysine-binding aromatic cages. *Biochem. J.* **478**, 3613–3619
41. Gatchalian, J., Ali, M., Andrews, F. H., Zhang, Y., Barrett, A. S., and Kutateladze, T. G. (2017) Structural insight into recognition of methylated histone H3K4 by Set3. *J. Mol. Biol.* **429**, 2066–2074
42. Klein, B. J., Wang, X., Cui, G., Yuan, C., Botuyan, M. V., Lin, K., *et al.* (2016) PHF20 readers link methylation of histone H3K4 and p53 with H4K16 acetylation. *Cell Rep.* **17**, 1158–1170
43. Zhang, Y., Jang, Y., Lee, J. E., Ahn, J., Xu, L., Holden, M. R., *et al.* (2019) Selective binding of the PHD6 finger of MLL4 to histone H4K16ac links MLL4 and MOF. *Nat. Commun.* **10**, 2314
44. Studier, F. W. (2005) Protein production by auto-induction in high density shaking cultures. *Protein Expr. Purif.* **41**, 207–234
45. Otwinowski, Z., Borek, D., Majewski, W., and Minor, W. (2003) Multiparametric scaling of diffraction intensities. *Acta Crystallogr. A.* **59**, 228–234
46. Winter, G., Waterman, D. G., Parkhurst, J. M., Brewster, A. S., Gildea, R. J., Gerstel, M., *et al.* (2018) Dials: implementation and evaluation of a new integration package. *Acta Crystallogr. D Struct. Biol.* **74**, 85–97
47. McCoy, A. J., Grosse-Kunstleve, R. W., Adams, P. D., Winn, M. D., Storoni, L. C., and Read, R. J. (2007) Phaser crystallographic software. *J. Appl. Crystallogr.* **40**, 658–674
48. Headd, J. J., Echols, N., Afonine, P. V., Grosse-Kunstleve, R. W., Chen, V. B., Moriarty, N. W., *et al.* (2012) Use of knowledge-based restraints in phenix.refine to improve macromolecular refinement at low resolution. *Acta Crystallogr. D Biol. Crystallogr.* **68**, 381–390
49. Brunger, A. T. (1992) Free R value: a novel statistical quantity for assessing the accuracy of crystal structures. *Nature* **355**, 472–475
50. Emsley, P., and Cowtan, K. (2004) Coot: model-building tools for molecular graphics. *Acta Crystallogr. D Biol. Crystallogr.* **60**, 2126–2132
51. Read, R. J., Adams, P. D., Arendall, W. B., 3rd, Brunger, A. T., Emsley, P., Joosten, R. P., *et al.* (2011) A new generation of crystallographic validation tools for the protein data bank. *Structure* **19**, 1395–1412