









ARTICLE

<https://doi.org/10.1038/s41467-019-10324-8>

OPEN

# Selective binding of the PHD6 finger of MLL4 to histone H4K16ac links MLL4 and MOF

Yi Zhang <sup>1,7</sup>, Younghoon Jang <sup>2,7</sup>, Ji-Eun Lee <sup>2</sup>, JaeWoo Ahn<sup>1</sup>, Longxia Xu<sup>3</sup>, Michael R. Holden<sup>1</sup>, Evan M. Cornett<sup>3</sup>, Krzysztof Krajewski<sup>4</sup>, Brianna J. Klein<sup>1</sup>, Shu-Ping Wang<sup>5</sup>, Yali Dou<sup>6</sup>, Robert G. Roeder<sup>5</sup>, Brian D. Strahl <sup>4</sup>, Scott B. Rothbart <sup>3</sup>, Xiaobing Shi <sup>3</sup>, Kai Ge <sup>2</sup> & Tatiana G. Kutateladze <sup>1</sup>

Histone methyltransferase MLL4 is centrally involved in transcriptional regulation and is often mutated in human diseases, including cancer and developmental disorders. MLL4 contains a catalytic SET domain that mono-methylates histone H3K4 and seven PHD fingers of unclear function. Here, we identify the PHD6 finger of MLL4 (MLL4-PHD6) as a selective reader of the epigenetic modification H4K16ac. The solution NMR structure of MLL4-PHD6 in complex with a H4K16ac peptide along with binding and mutational analyses reveal unique mechanistic features underlying recognition of H4K16ac. Genomic studies show that one third of MLL4 chromatin binding sites overlap with H4K16ac-enriched regions *in vivo* and that MLL4 occupancy in a set of genomic targets depends on the acetyltransferase activity of MOF, a H4K16ac-specific acetyltransferase. The recognition of H4K16ac is conserved in the PHD7 finger of paralogous MLL3. Together, our findings reveal a previously uncharacterized acetyllysine reader and suggest that selective targeting of H4K16ac by MLL4 provides a direct functional link between MLL4, MOF and H4K16 acetylation.

<sup>1</sup>Department of Pharmacology, University of Colorado School of Medicine, Aurora, CO 80045, USA. <sup>2</sup>Laboratory of Endocrinology and Receptor Biology, National Institute of Diabetes and Digestive and Kidney Diseases, NIH, Bethesda, MD 20892, USA. <sup>3</sup>Center for Epigenetics, Van Andel Research Institute, Grand Rapids, MI 49503, USA. <sup>4</sup>Department of Biochemistry & Biophysics, The University of North Carolina School of Medicine, Chapel Hill, NC 27599, USA. <sup>5</sup>Laboratory of Biochemistry and Molecular Biology, The Rockefeller University, New York, NY 10065, USA. <sup>6</sup>Department of Pathology, University of Michigan, Ann Arbor, MI 48109, USA. <sup>7</sup>These authors contributed equally: Yi Zhang, Younghoon Jang. Correspondence and requests for materials should be addressed to K.G. (email: [kai.ge@nih.gov](mailto:kai.ge@nih.gov)) or to T.G.K. (email: [tatiana.kutateladze@ucdenver.edu](mailto:tatiana.kutateladze@ucdenver.edu))

**H**uman mixed lineage leukemia 4 (MLL4) or KMT2D is an evolutionarily conserved lysine methyltransferase that plays crucial roles in embryogenesis and development. MLL4 is frequently mutated in diseases, including Kabuki syndrome, congenital heart disorder and various cancers and is implicated in both tumor suppression and cancer cell proliferation<sup>1</sup>. Genomic studies have identified MLL4 as a major histone H3K4 mono-methyltransferase<sup>2</sup>. It catalyzes the transfer of a single methyl group from S-adenosyl-L-methionine (SAM) to the  $\epsilon$ -amino group of Lys4, producing the epigenetic mark H3K4me1 primarily in transcriptional enhancer regions, where MLL4 co-localizes with lineage-determining transcription factors<sup>3,4</sup>. MLL4 is required for enhancer activation and cooperates with the H3K27/K18-specific acetyltransferases p300/CBP in the regulation of cell-type-specific transcriptional programs<sup>4-6</sup>.

MLL4 belongs to a six-member MLL/KMT2 family of enzymes, with each member targeting H3K4 but being capable of generating distinct sets of H3K4 methylation states and at different genomic sites<sup>1</sup>. Much like other MLL/KMT2 family members, MLL4 functions in the context of a large multi-subunit complex consisting of the core subunits WDR5, Ash2L, RBBP5, and DPY30 and several accessory components, including NCOA6, PA1, PTIP, and UTX, with the latter being an H3K27 demethylase (Fig. 1a). MLL4 itself is a massive, ~600 kDa multimodular protein that harbors seven plant homeodomain (PHD) fingers, a pair of the FY-rich (FYRN and FYRC) domains, and the C-terminal catalytic Su(var)3-9, Enhancer of Zeste, Trithorax (SET) domain, also essential for protein stability in cells (Fig. 1b). With the exception of the SET domain, which is conserved in the MLL/KMT family, very little is known regarding functions of the other domains in MLL4<sup>7</sup>. The triple PHD finger cassette (PHD4-6) of MLL4 has recently been shown to associate with histone H4 either unmodified or asymmetrically dimethylated at R3 (H4R3me2a)<sup>8</sup>.

Among the most common epigenetic modifications is acetylation of Lys16 in histone H4 (H4K16ac). This modification is necessary for a wide array of fundamental processes, including chromatin decompaction, DNA damage repair and RNA Polymerase II pause-release<sup>9-11</sup>. In mammals and flies, H4K16ac is generated or “written” by the histone acetyltransferase Males absent on the First (MOF), which is involved in transcriptional regulation, stem cell self-renewal and early embryonic development and is dysregulated in human diseases<sup>12-16</sup>. Despite the crucial role of MOF-dependent H4K16 acetylation in normal cell processes and disease, a specific reader of this mark has not yet been identified.

In this study, we identify the PHD6 finger of MLL4 (MLL4-PHD6) as histone reader that exhibits high selectivity toward H4K16ac. We describe the molecular and structural basis for the recognition of H4K16ac by MLL4-PHD6 and show that MLL4 co-localizes with H4K16ac in an MOF acetyltransferase activity-dependent manner in a set of MLL4 target genomic sites. Our study provides direct evidence of functional correlation between the abilities of MOF and MLL4 to write and read, respectively, the major histone modification H4K16ac.

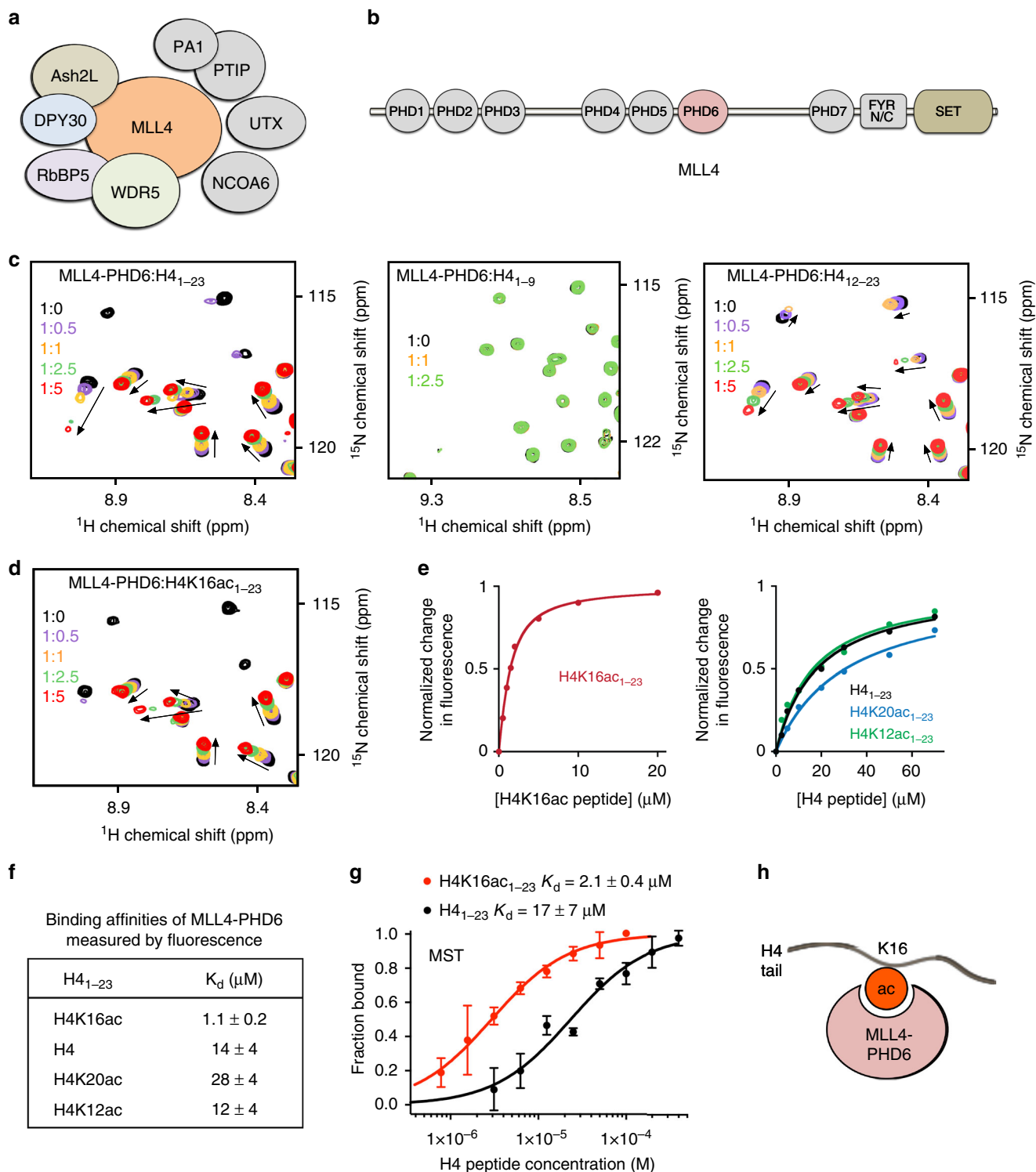
## Results

**The PHD6 finger of MLL4 is a reader of H4K16ac.** The region containing three PHD fingers (PHD4-6) of MLL4 has been shown to associate with histone H4<sup>8</sup>. To define the role of MLL4-PHD6 in this interaction, we expressed it as <sup>15</sup>N-labeled protein and tested by nuclear magnetic resonance (NMR; Fig. 1c). Titration of H4 peptide (residues 1-9 of H4) into the MLL4-PHD6 sample caused no chemical shift perturbations (CSPs) in <sup>1</sup>H,<sup>15</sup>N heteronuclear single quantum coherence (<sup>1</sup>H,<sup>15</sup>N HSQC) spectra,

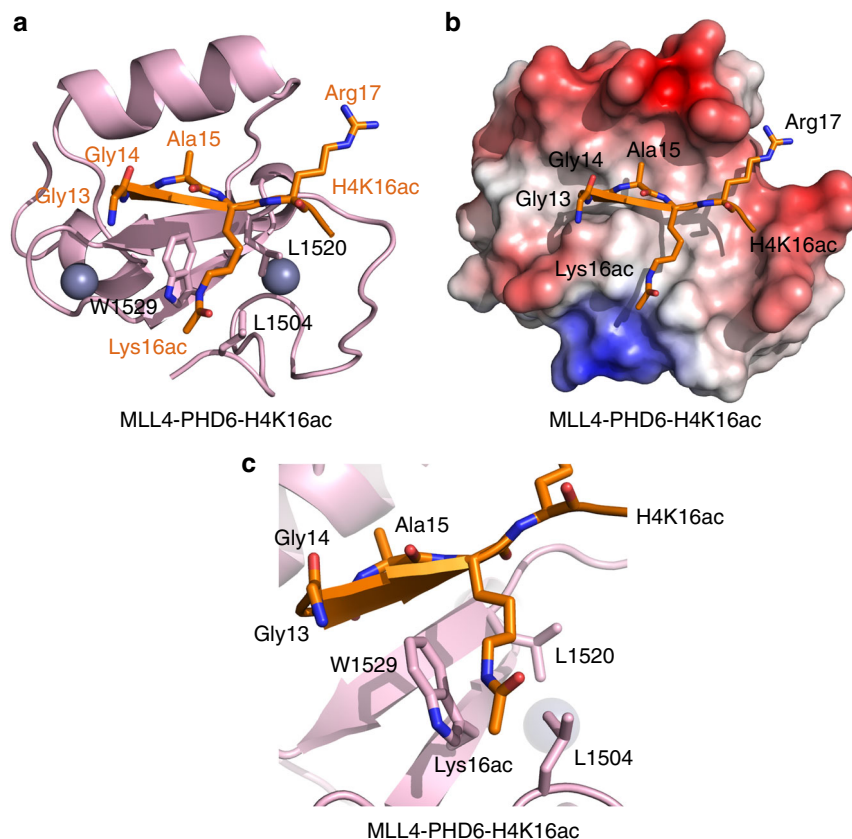
indicating that MLL4-PHD6 does not bind to the nine N-terminal residues of H4 tail. Gradual addition of the longer peptide (residues 1-23 of H4, H4<sub>1-23</sub>), however, resulted in large CSPs, which suggested that the C-terminal part of the histone H4 tail is recognized by MLL4-PHD6. Indeed, peptides encompassing residues 12-23 or 8-25 of H4 induced CSPs similar in directions to CSPs induced by the longer H4<sub>1-23</sub> peptide, supporting the idea that the C-terminal region of H4 is the target sequence of MLL4-PHD6 (Fig. 1c and Supplementary Figs. 1 and 2).

The C-terminal region of the H4 tail contains several lysine residues (K12, K16, and K20) known to be posttranslationally modified, predominantly acetylated. We therefore explored whether MLL4-PHD6 is capable of sensing the acetylation states of these lysine residues. We assessed binding of MLL4-PHD6 to H4K12ac, H4K16ac, and H4K20ac peptides (all, residues 1-23 of H4) using NMR, fluorescence and microscale thermophoresis (MST) assays (Fig. 1d-g and Supplementary Figs. 3-5). An almost identical pattern of CSPs in the intermediate exchange regime on the NMR timescale was observed in <sup>1</sup>H,<sup>15</sup>N HSQC titration experiments with H4K12ac, H4K20ac or H4 peptides, indicating that acetylation of H4K12 or H4K20 does not augment the histone-binding activity of MLL4-PHD6 (Fig. 1c and Supplementary Fig. 4). By contrast, upon titration of the H4K16ac peptide, some resonances of MLL4-PHD6 exhibited more pronounced line broadening, indicative of tight binding (Fig. 1d). In agreement, binding affinity of MLL4-PHD6 to H4K16ac peptide was found to be 1.1  $\mu$ M as measured by tryptophan fluorescence, whereas binding of this domain to H4, H4K12ac, and H4K20ac peptides was 12-28-fold weaker (Fig. 1e, f and Supplementary Fig. 5). The selectivity of MLL4-PHD6 toward H4K16ac was further corroborated by MST experiments. The MST assays yielded a dissociation constant ( $K_d$ ) of 2.1  $\mu$ M for the interaction of MLL4-PHD6 with the H4K16ac peptide, whereas association with the H4 peptide was eight-fold weaker ( $K_d = 17 \mu$ M) (Fig. 1g). Although acetylation of H4K16 substantially enhanced binding, we found that the residues surrounding K16 are also required: a free Kac amino acid failed to induce CSPs in MLL4-PHD6 even at high Kac concentration (Supplementary Fig. 1). We note that the  $K_d$  value for the MLL4-PHD6:H4K16ac interaction was in the range of binding affinities exhibited by the majority of histone-binding modules<sup>17-19</sup>, reinforcing the idea that MLL4-PHD6 is a previously unidentified reader of H4K16ac (Fig. 1h).

**The mechanism of H4K16ac recognition by MLL4-PHD6.** To elucidate the molecular basis for the interaction of MLL4-PHD6 with H4K16ac, we determined the solution structure of the MLL4-PHD6:H4K16ac complex by NMR spectroscopy. The structural ensemble revealed a central double-stranded anti-parallel  $\beta$ -sheet and two  $\alpha$ -helices of MLL4-PHD6 (Supplementary Figs. 6-9 and Supplementary Table 1). Two zinc atoms were coordinated by the C4HC3 motif, characteristic of PHD fingers. Residues Gly14-Arg17 of the H4K16ac peptide adopted an extended conformation and formed a third  $\beta$  strand pairing with the existing  $\beta$ -sheet of MLL4-PHD6 in an antiparallel manner (Fig. 2a). Intermolecular nuclear Overhauser enhancements were identified between residues Gly14, Ala15, Lys16ac and Arg17 of H4 and residues L1519, L1520, I1521, Q1522, W1529, E1540, V1543 and A1547 of MLL4-PHD6. Overall, the PHD6 finger surface where the Gly13-Lys16ac sequence of H4K16ac is bound was largely hydrophobic (Fig. 2b). However, this hydrophobic site was surrounded by the highly negatively charged walls created by the side chains of residues in the loop preceding the  $\beta$ 1 strand and the acidic  $\alpha$ -helix of PHD6, including E1516, E1517, D1518, E1540, E1544, and D1548.



**Fig. 1** MLL4-PHD6 recognizes the histone mark H4K16ac. **a** Schematic of the MLL4 complex. **b** MLL4 domain architecture. **c** Superimposed  $^1\text{H},^{15}\text{N}$  heteronuclear single quantum coherence (HSQC) spectra of MLL4-PHD6<sub>1503-1562</sub> collected upon titration with H4<sub>1-23</sub>, H4<sub>1-9</sub>, and H4<sub>12-23</sub> peptides. Spectra are color coded according to the protein:peptide molar ratio. See also Supplementary Fig. 1, first panel. **d** Superimposed  $^1\text{H},^{15}\text{N}$  HSQC spectra of MLL4-PHD6<sub>1503-1562</sub> collected upon titration with H4K16ac<sub>1-23</sub> peptide. Spectra are color coded according to the protein:peptide molar ratio. **e** Representative binding curves used to determine the  $K_d$  values by fluorescence spectroscopy (also see Supplementary Fig. 5). **f** Binding affinities of wild-type MLL4-PHD6 for the indicated histone peptides measured by tryptophan fluorescence. Error represents s.d. in triplicate measurements. Source data are provided as a Source Data file. **g** Binding curves used to determine the  $K_d$  values by microscale thermophoresis. Error represents s.d. in triplicate measurements. **h** A schematic showing specific reading of the H4K16ac mark by MLL4-PHD6



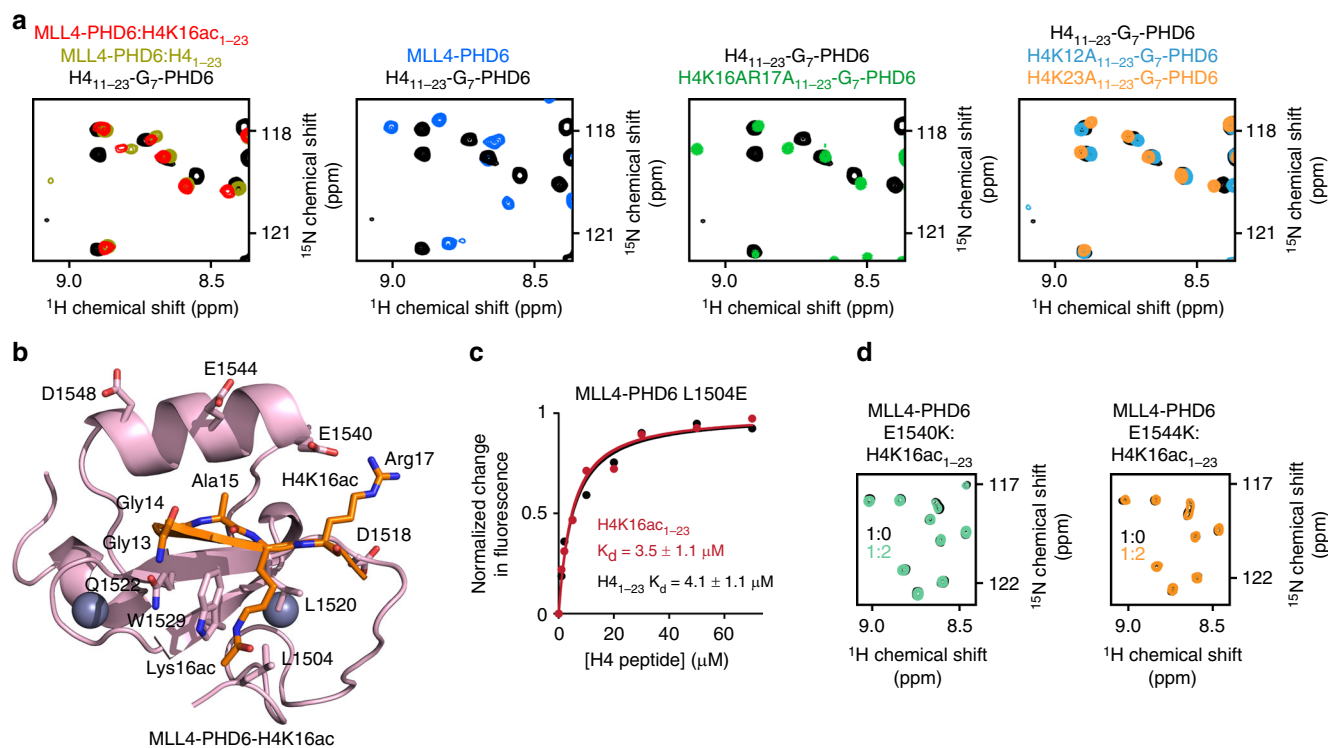
**Fig. 2** Structure of MLL4-PHD6 in complex with H4K16ac peptide. **a** A ribbon diagram of the solution nuclear magnetic resonance structure of MLL4-PHD6 (pink) in complex with the H4K16ac peptide (orange). **b** Electrostatic surface potential of MLL4-PHD6 colored blue and red for positive and negative charges, respectively. The bound H4K16ac peptide is shown in a stick model. **c** A zoom-in view of the H4K16ac-binding pocket

The structure of the complex has illuminated two unique mechanistic features underlying the recognition of H4K16ac. One is the hydrophobic groove formed by the residues L1504, L1520, and W1529 of MLL4-PHD6, which accommodates the non-charged acetylated side chain of Lys16 (Fig. 2b, c). The aromatic side chain of W1529 is positioned orthogonally to the protein surface and is differently perturbed by acetylated and non-acetylated K16 (Supplementary Fig. 10a). The second feature is the presence of a highly flexible Gly13–Gly14 sequence in H4 that ensures proper fitting of the N-terminal part of the H4K16ac peptide in the binding site. Although Gly13 of H4 was not restrained and showed structural flexibility, it formed a sharp turn in the majority of structures in the ensemble, allowing Gly14 to occupy the hydrophobic cavity of MLL4-PHD6 and avoid steric clashes (Supplementary Fig. 9). These structural features point to a previously uncharacterized mechanism for the recognition of histone PTMs by PHD fingers and distinguish MLL4-PHD6 from other histone H4 readers (see discussion below).

**The contribution of the binding site residues.** To determine the role of interfacial residues in the MLL4-PHD6:H4K16ac complex formation, we generated a chimeric construct containing residues 11–23 of H4 coupled to the residues 1503–1560 of MLL4-PHD6 through a hepta-glycine linker (Fig. 3a). The  $^1\text{H}$ ,  $^{15}\text{N}$  HSQC spectrum of the linked H4–MLL4-PHD6 construct showed that the H4 region is tightly bound to the protein. The spectrum (colored black in Fig. 3a) did not overlay with the spectrum of an isolated MLL4-PHD6 finger in the apo state (blue); however, it overlaid well with the spectrum of MLL4-PHD6 bound to H4<sub>1–23</sub> peptide (dark yellow). These data suggest that the linked and

unlinked MLL4-PHD6:H4 complexes have similar structures. Mutation of either H4K12 or H4K23 to alanine in the linked H4–MLL4-PHD6 construct resulted in only small spectral changes, confirming that K12 and K23 of H4 are dispensable in this interaction. In contrast, mutation of K16 and R17 to alanine disrupted formation of the complex, underscoring the critical role of these residues in the binding.

Superimposition of  $^1\text{H}$ ,  $^{15}\text{N}$  HSQC spectra of MLL4-PHD6 bound to H4K16ac<sub>1–23</sub> peptide (red) and H4<sub>1–23</sub> peptide (dark yellow) revealed that the majority of crosspeaks overlap, indicating that both peptides are bound in a similar manner (Fig. 3a and Supplementary Fig. 10b). As expected, most notable differences were observed in resonances of the residues located near the K16-binding-site-forming W1529, L1520, and L1504. Not only the side chain NH $\epsilon$  of W1529 but also backbone amides of W1529, R1528, L1519, and V1505 were differently perturbed by H4K16ac and H4 (Supplementary Fig. 10). Although W1529 and L1520 comprise the hydrophobic core of the domain and are likely necessary for structural stability (W1529A is unfolded, Supplementary Fig. 11), L1504 is more solvent exposed and can be substituted with Glu. We rationalized that mutation of the hydrophobic leucine residue to the negatively charged glutamate would reduce binding of hydrophobic K16ac and increase binding of positively charged unmodified K16. Indeed, compared to wild-type (WT) MLL4-PHD6, the ability of L1504E mutant to interact with H4K16ac was reduced ( $K_d = 3.5 \mu\text{M}$ ), whereas its ability to interact with non-acetylated H4 was enhanced ( $K_d = 4.1 \mu\text{M}$ ) (Fig. 3c and Supplementary Fig. 5). We concluded that hydrophobic nature of the Leu–Leu–Trp binding groove is a driving force in preference of MLL4-PHD6 for the hydrophobic side chain of K16ac.



**Fig. 3** Molecular basis for the specific targeting of H4K16ac by MLL4-PHD6. **a** Superimposed  $^1\text{H},^{15}\text{N}$  heteronuclear single quantum coherence (HSQC) spectra of the linked  $\text{H4}_{11-23}\text{-G}_7\text{-PHD6}$  construct, wild-type (black) and indicated mutants (green, light blue, and orange), the isolated MLL4-PHD6 domain (blue), isolated H4-bound MLL4-PHD6 (dark yellow), and isolated H4K16ac-bound MLL4-PHD6 (red). **b** The structure of the H4K16ac-bound MLL4-PHD6 finger. **c** Representative binding curves used to determine the  $K_d$  values for the L1504E mutant by fluorescence spectroscopy (also see Supplementary Fig. 5). Source data are provided as a Source Data file. **d** Superimposed  $^1\text{H},^{15}\text{N}$  HSQC spectra of MLL4-PHD6 mutants collected upon titration with H4K16ac peptide. Spectra are color coded according to the protein:peptide molar ratio

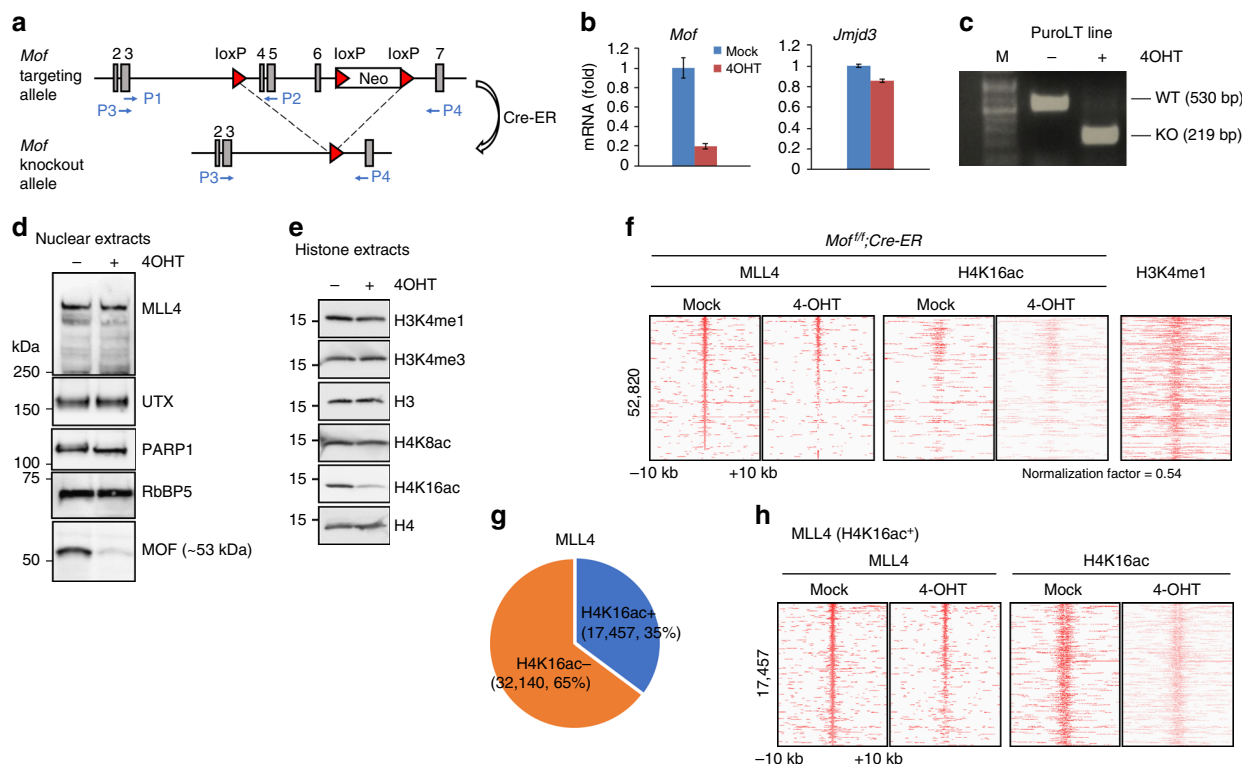
Arg17 of the H4K16ac peptide is bound in a channel formed by two negatively charged clusters of MLL4-PHD6,  $\text{EED}_{1518}$ , and  $\text{EDDVE}_{1544}$  (Fig. 3b). To assess the importance of these clusters in the interaction with H4K16ac, we generated D1518A, E1540K, and E1544K mutants and tested them by NMR. A lack of CSPs in  $^1\text{H},^{15}\text{N}$  HSQC spectra of E1540K and E1544K upon addition of H4K16ac peptide indicated that both glutamates contribute substantially to electrostatic interactions, especially E1540 that is in close contact with Arg17 of the peptide, and both are required for the binding (Fig. 3d). Mutation of D1518, as well as Q1522 and D1548, led to protein unfolding, implying that these residues are necessary for structural stability of MLL4-PHD6 (Supplementary Fig. 11).

**MLL4 colocalizes with H4K16ac in cells.** MLL4 has been shown to be essential for cell differentiation, including adipogenesis<sup>2</sup>, therefore we employed a mouse preadipocytes model system to examine the functional significance of the H4K16ac recognition by MLL4. Because H4K16ac is generated by the MOF acetyltransferase, we isolated brown preadipocytes from  $\text{Mof}^{\text{fl/fl}};\text{Cre-ER}$  mice in order to manipulate H4K16ac levels<sup>10</sup>. In this mouse strain, exons 4–6 of the *Mof* gene are flanked by loxP sites (Fig. 4a). In the presence of 4-hydroxytamoxifen (4OHT), Cre-ER recombinase was able to delete exons 4–6 of the *Mof* locus. The efficient deletion of the loxP-flanked *Mof* exons in the preadipocytes was verified at the mRNA and protein levels. After two rounds of 4OHT treatment, the mRNA level of *Mof* was significantly decreased, whereas mRNA of a control gene *Jmjd3* was unaffected (Fig. 4b). Furthermore, the presence of truncated *Mof* mRNA in the  $\text{Mof}^{\text{fl/fl}};\text{Cre-ER}$  cells confirmed a deletion of *Mof* (Fig. 4c). Western blot analysis also showed reduced MOF protein

levels, whereas the MLL4 complex remained intact (Fig. 4d). Importantly, the MOF knockout (KO) led to a substantial decrease in H4K16ac levels but had no effect on other histone marks, such as H3K4me1, H3K4me3, and H4K8ac (Fig. 4e). Consistent with the previous findings<sup>10</sup>, *Mof* KO also decreased cell growth in preadipocytes (Supplementary Fig. 12).

To determine whether the deletion of *Mof* affects genomic occupancy of MLL4, we performed chromatin immunoprecipitation experiments followed by high-throughput sequencing (ChIP-Seq) in the  $\text{Mof}^{\text{fl/fl}};\text{Cre-ER}$  preadipocytes treated with and without 4OHT. Both MLL4 chromatin association and H4K16ac levels were significantly reduced in the *Mof* KO cells (Fig. 4f). A total of 52,820 MLL4-binding sites were identified from the control (Mock) or *Mof* KO (4OHT) cells using the SICER method<sup>20</sup>. Importantly, these sites also had high level of H3K4me1, which is generated by MLL4<sup>2</sup>. Gene ontology (GO) data showed that MLL4 target genes are involved in a broad range of cellular events, including cell adhesion, metabolism, and mitosis (Supplementary Fig. 13). Further analysis of the 49,597 MLL4 target regions and the H4K16ac mark revealed that 17,457 (~35%) MLL4-binding sites overlapped with H4K16ac-enriched ( $\text{H4K16ac}^+$ ) sites in the control cells (Fig. 4g, h). In this subset of MLL4 target regions, MOF KO led to a concurrent decrease in H4K16ac levels and MLL4 association with chromatin, which is indicative of a direct correlation between MOF, MLL4, and H4K16 acetylation.

**MOF is required for binding of MLL4 to a subset of genes.** The MOF-dependent binding of MLL4 to specific genomic sites was characterized by comparing normalized MLL4 ChIP read counts between WT and *Mof* KO cells. *Mof* KO selectively inhibited



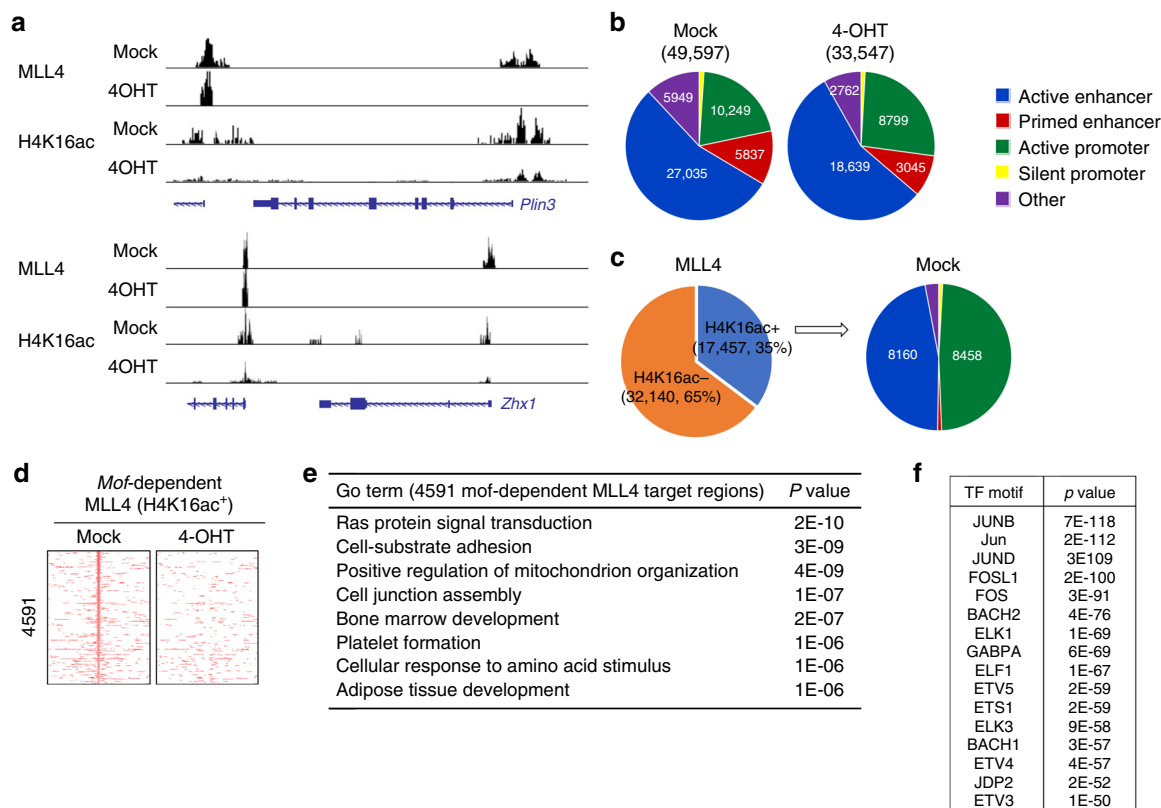
**Fig. 4** A subset of MLL4-bound genomic regions overlaps with H4K16ac. SV40T-immortalized *Mof<sup>fl/fl</sup>;Cre-ER* brown preadipocytes were treated with 4-hydroxytamoxifen to delete *Mof*. Cells were collected for reverse transcriptase PCR (RT-PCR), western blot and chromatin immunoprecipitation–sequencing (ChIP-Seq) analysis. **a** Schematics of targeting allele and knockout allele of *Mof<sup>fl/fl</sup>* mice. In the targeting allele, a single loxP site was inserted in the intron before exon 4 and a neomycin selection cassette flanked by loxP sites was inserted in the intron after exon 6. The locations of RT-PCR primers P1–P2 (used in Fig. 4b) and P3–P4 (used in Fig. 4c) are indicated by arrows. **b** Quantitative RT-PCR confirmation of *Mof* deletion in preadipocytes. The data are presented as mean  $\pm$  s.d. Two technical replicates from a single experiment were performed. Source data are provided as a Source Data file. **c** RT-PCR confirmation of wild-type and truncated *Mof* mRNA. **d** Western blot analysis of MLL4, UTX, and MOF. PARP1 and RbBP5 were used as loading controls. **e** Western blot analysis of histone modifications (also see Supplementary Fig. 15). **f** Heat maps around MLL4-binding sites. H4K16ac ChIP-Seq data were normalized with global H4K16ac levels measured by western blot. Published H3K4me1 data were obtained from GEO database (GSE74189)<sup>49</sup>. **g** Pie chart illustrating that 35% of MLL4-binding sites overlap with H4K16ac. **h** Heat maps of MLL4 and H4K16ac around 17,457 H4K16ac<sup>+</sup> MLL4-binding sites

MLL4 binding to promoters of *Plin3*, the gene encoding a lipid droplet protein, and *Zhx1*, the gene encoding a zinc-finger and homeodomain transcription factor, but not to promoters of nearby genes (Fig. 5a). Overall, the deletion of *Mof* affected genomic occupancy of MLL4 on all sites (Fig. 5b). Interestingly, the H4K16ac<sup>+</sup> MLL4-binding sites were located almost exclusively on active enhancers and active promoters, which is consistent with the previous reports that H4K16ac is enriched around transcription start sites and enhancers of active genes in embryonic stem cells<sup>21</sup> (Fig. 5c). While ~30% of MLL4-binding sites on active enhancers overlapped with H4K16ac, over 82% of MLL4-binding sites on active promoters were H4K16ac positive. Within 17,457 H4K16ac<sup>+</sup> MLL4-binding regions, the deletion of *Mof* resulted in the loss of 4591 (~26%) MLL4-binding sites (Fig. 5d). GO analysis revealed that these 4591 *Mof*-dependent MLL4 target regions are associated with the genes that define mesenchymal lineages (bone marrow development, platelet formation, and adipose tissue development) or specific brown adipose lineages (positive regulation of mitochondrion organization) (Fig. 5e). The 4591 *Mof*-dependent MLL4-binding regions were enriched in motifs of the AP-1 and ETS families of transcription factors (Fig. 5f).

**MOF catalytic activity-dependent binding of MLL4 to chromatin.** To substantiate the biological significance of MOF-mediated H4K16ac recognition by MLL4, we ectopically

expressed WT or an enzymatically dead mutant (K274R) of rat MOF (rMOF) in *Mof<sup>fl/fl</sup>;Cre-ER* preadipocytes and deleted the endogenous *Mof* by treating the cells with 4OHT (Fig. 6a). In total, we identified 5296 MLL4-binding sites that were MOF enzymatic activity dependent (Fig. 6b). MLL4 binding to these 5296 regions was fully rescued by ectopic expression of WT MOF but not by ectopic expression of the catalytically dead MOF K274R mutant. MOF enzymatic activity-dependent MLL4 binding and H4K16ac enrichment was selectively observed in the proximal promoter regions of the *Itf81*, *Ccnd3*, and *Zfp13* genes (Fig. 6c–e). GO analysis indicated that MOF enzymatic activity-dependent MLL4-binding regions are associated with the genes that are necessary to block fat cell differentiation and maintain preadipocyte status, in addition to the genes regulating general biological functions (Fig. 6f). Motif analysis revealed that MOF enzymatic activity-dependent MLL4-binding sites were strongly enriched with AP-1 transcription factor motifs (Fig. 6g).

**H4K16ac recognition is conserved in the PHD7 finger of MLL3.** Much like MLL4 that contains seven PHD fingers, another member of the MLL/KMTD family, MLL3 harbors eight PHD fingers (Fig. 7). Alignment of sequences for all PHD fingers from human MLL4 and MLL3 suggests that MLL4-PHD6 has a high degree sequence similarity to MLL3-PHD7. We tested whether H4K16ac-binding activity is conserved in MLL3-PHD7 by NMR.



**Fig. 5** Targeting of MLL4 to active promoters is dependent on H4K16ac mark. *Mof<sup>f/f</sup>;Cre-ER* brown preadipocytes were treated with 4-hydroxytamoxifen and collected for chromatin immunoprecipitation–sequencing (ChIP–Seq) analysis. **a** ChIP–Seq profiles of MLL4 and H4K16ac on *Plin3* and *Zhx1* gene loci. **b** Deletion of *Mof* decreases MLL4 binding to all sites. **c** H4K16ac-enriched (H4K16ac<sup>+</sup>) MLL4-binding regions are mainly located on active enhancers and active promoters. **d** Heat maps of 4591 *Mof*-dependent H4K16ac<sup>+</sup> MLL4-binding sites. **e** Gene ontology analysis of the genes associated with 4591 *Mof*-dependent H4K16ac<sup>+</sup> MLL4-binding regions identified in **d**. **f** Motif analysis of 4591 *Mof*-dependent H4K16ac<sup>+</sup> MLL4-binding regions identified in **d**

Gradual addition of H4K16ac peptide to the <sup>15</sup>N-labeled MLL3-PHD7 finger led to notable CSPs, thus confirming that, similar to MLL-PHD6, the PHD7 finger of MLL3 is also capable of targeting H4K16ac (Fig. 7a).

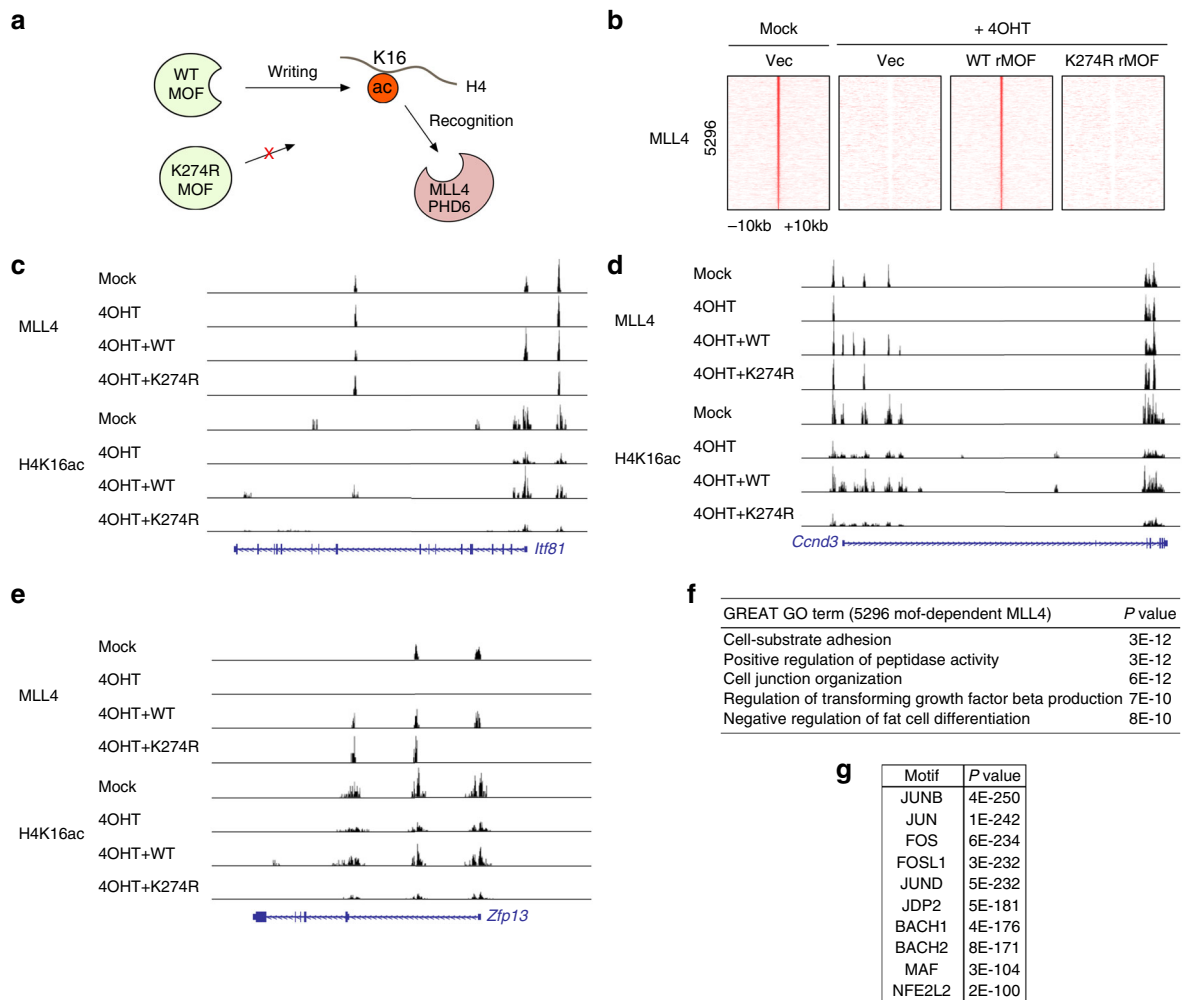
## Discussion

The finding that MLL4-PHD6 and MLL3-PHD7 recognize H4K16ac has expanded the family of acetyllysine readers, which currently includes bromodomain, YEATS, and DPF, and a set of domains in certain proteins, such as a ZZ domain, double bromodomain, and a double PH domain that display selectivity toward acetylated lysine residues<sup>22–28</sup>. Structural comparison of the MLL4-PHD6:H4K16ac complex with acetyllysine-bound BD, YEATS, and DPF reveals distinctly different mechanisms underlying the recognition of acetyllysine. Not only the folds and acetyllysine-binding pockets of these readers are unlike, but the acetylated lysine residues are restrained through different sets of hydrogen bonding,  $\pi$ – $\pi$ – $\pi$ , and hydrophobic interactions (Fig. 7b–d). One similarity perhaps is the overall hydrophobic character of the acetyllysine-binding sites and the presence of aromatic residues that likely facilitate engagement with the hydrophobic side chain of acetyllysine.

The PHD finger has been well established as a reader of the histone H3 tail, either trimethylated at Lys4 (H3K4me3) or unmodified<sup>29–34</sup>, and the finding of a subset of PHD fingers that select for acetylated H4K16 is unexpected. Comparison of the MLL4-PHD6:H4K16ac structure with the ING2-PHD:H3K4me3 structure reveals some similarities in the histone-binding modes (Fig. 7e, f). In both complexes, histone tails are bound in an extended conformation and K16ac of H4 and K4me3

of H3 occupy binding sites made of aromatic and hydrophobic residues (Fig. 7f). The W1529 residue in the MLL4-PHD6:H4K16ac complex that forms one of the walls in the acetyllysine-binding groove is located in a similar position as an invariable Trp residue that forms a wall in the trimethyllysine-binding cage (W238 in ING2) (Fig. 7f, g). However, the aromatic cage for trimethylated lysine is larger in size and contains at least two aromatic residues that are involved in cation– $\pi$  interactions with the trimethylammonium group of K4me3 and are strictly required for the recognition of H3K4me3. In contrast, K16ac lies in a narrow, single aromatic residue-containing groove with the acetyl moiety being placed between the side chains of tryptophan and leucine residues. A critical role of the free amino-terminal Ala1 residue in the recognition of either H3K4me3 or unmodified H3 is well established, but the MLL4 PHD6 finger binds to the middle part of histone H4 tail and does not require the presence of NH<sub>3</sub><sup>+</sup>. Instead, a highly flexible Gly13–Gly14 sequence of H4 allows for the proper fitting of H4K16ac in the binding pocket of MLL4-PHD6. Alignment of amino acid sequences of seven PHD fingers from MLL4 and eight PHD fingers from MLL3 shows that the H4K16ac-binding site residues of MLL4-PHD6 are conserved in MLL3-PHD7 but not in other PHD fingers of MLL4 and MLL3 for which biological roles remain unidentified (Fig. 7g).

Our results demonstrate that MLL4 co-localizes with H4K16ac in vivo and that the association of MLL4 with a set of genomic targets is *Mof* dependent and requires the acetyltransferase activity of MOF. These data imply a hitherto unrecognized pathway for chromatin regulation that is likely central to the control of gene transcription. We propose that MOF-mediated deposition of H4K16ac at certain genomic sites directs the



**Fig. 6** MOF (Males absent On the First) catalytic activity is necessary for *Mof*-dependent chromatin targeting of MLL4. *Mof<sup>f/f</sup>;Cre-ER* brown preadipocytes were infected with lentiviruses expressing wild-type (WT) or enzyme dead mutant (K274R) form of rMOF, followed by 4-hydroxytamoxifen treatment. Cells were collected for chromatin immunoprecipitation–sequencing (ChIP-Seq) analysis of MLL4 and H4K16ac. **a** Schematic representation of WT and enzyme dead MOF. **b** Heat maps of 5296 MOF enzyme activity-dependent MLL4-binding sites. **c–e** ChIP-Seq profiles of MLL4 and H4K16ac on *Zfp13* (**c**), *Ccnd3* (**d**), and *Itf1* (**e**) gene loci. **f** Gene ontology analysis of the genes associated with 5296 MOF catalytic activity-dependent MLL4-binding regions. **g** AP-1 family transcription factor motifs are enriched at MOF catalytic activity-dependent MLL4-binding regions. Top 3000 MLL4-binding regions were used for motif analysis

recruitment or stabilization of MLL4 via its PHD6 finger recognition of H4K16ac. Furthermore, we show that MOF enzymatic activity-dependent MLL4-binding sites are highly enriched in the AP-1 transcription factor motif, therefore AP-1 could cooperate with MOF and H4K16ac in the recruitment of MLL4 to active transcriptional regions. Future studies will be needed to determine the broad significance of this mechanism in gene transcription and whether it is conserved across other eukaryotes.

## Methods

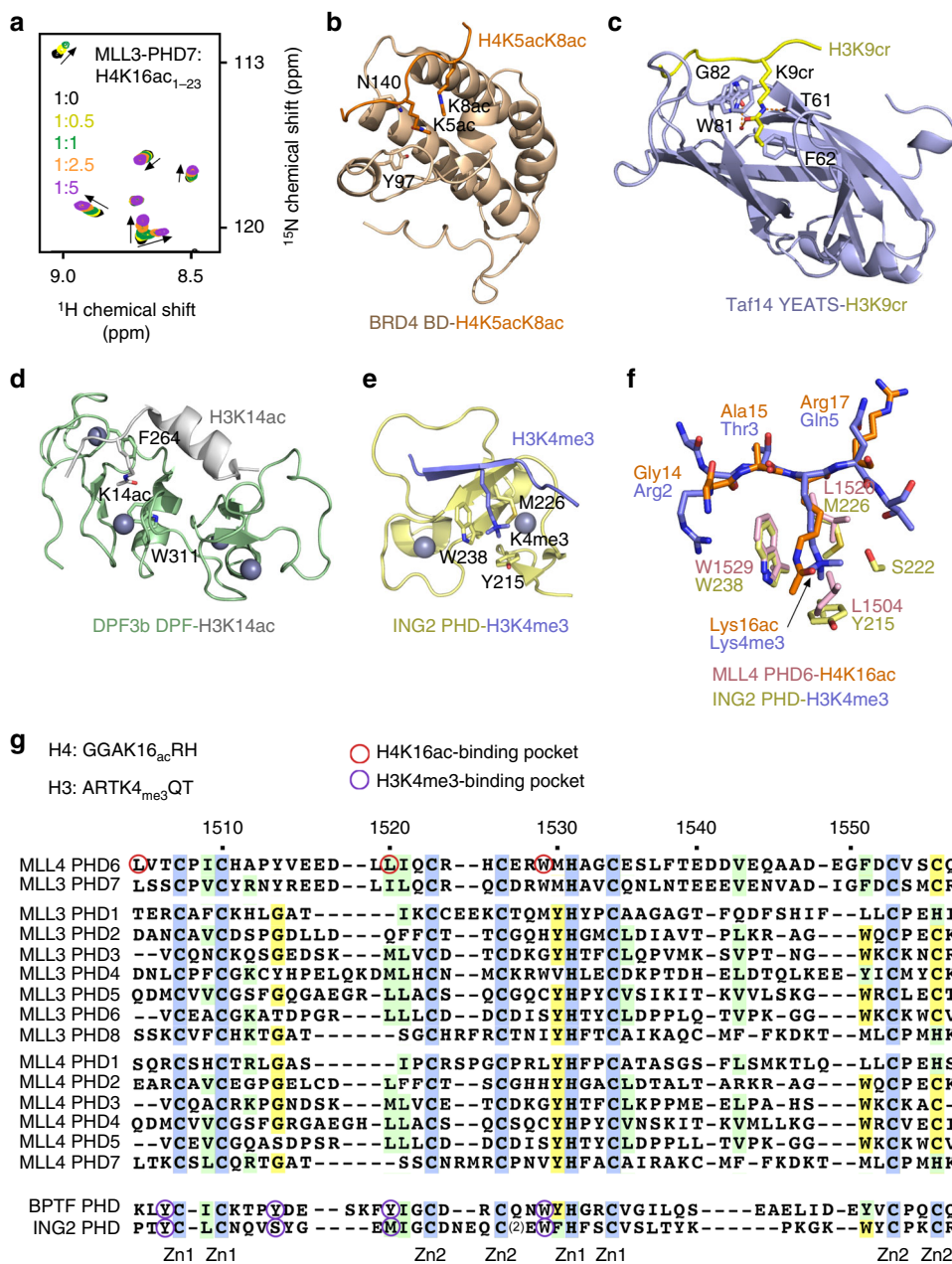
**Plasmids, antibodies, and chemicals.** The lentiviral pLenti-CMV plasmids expressing MYC-tagged WT or K274R mutant of rat MOF (a gift from Jianrong Lu) were described previously<sup>35</sup>. The following homemade antibodies have been described: anti-MLL4#3<sup>36</sup> and anti-UTX<sup>37</sup>. Anti-MOF (A300–992A) and anti-RbBP5 (A300–109A) were from Bethyl Laboratories. Anti-H3 (ab1791), anti-H3K4me1 (ab8895), and anti-H4 (ab7311) were from Abcam. Anti-H3K4me3 (07–473), anti-H4K8ac (07–328), and anti-H4K16ac (07–329) were from Millipore. Anti-PARP1 (556362) was from BD Bioscience. Anti-MYC (sc40) was from Santa Cruz and 4OHT (H7904) was from Sigma. A 1:1000 dilution of antibodies was used for western blot.

**Mouse strains and immortalization of primary brown preadipocytes.** *Mof<sup>f/f</sup>* mice were crossed with *Cre-ER* (Jackson no. 008463) to generate *Mof<sup>f/f</sup>;Cre-ER*

mice<sup>10</sup>. Primary brown preadipocytes were isolated from interscapular brown adipose tissue of newborn (P0) *Mof<sup>f/f</sup>;Cre-ER* pups and immortalized by SV40T<sup>38</sup>. All mouse experiments were performed in accordance with the NIH Guide for the Care and Use of Laboratory Animals and approved by the Animal Care and Use Committee of NIDDK, NIH. For rescue experiment by ectopic expression of rat MOF, immortalized *Mof<sup>f/f</sup>;Cre-ER* preadipocytes stably expressing WT or K274R MOF were established using lentivirus-mediated gene transfer. Briefly, 293FT lentiviral packaging cells were transfected with pLenti-CMV plasmids expressing MYC-tagged WT or K274R rat MOF. Preadipocytes were infected with those lentivirus-containing media, followed by puromycin selection for two passages. Those expressing MYC-tagged rat MOF were confirmed by western blot.

**Protein expression and purification.** The PHD6 domain of human MLL4 (aa 1503–1562) was cloned into a pCool vector. The H4-G7-MLL4-PHD6 construct (aa 11–23 of H4, hepta-glycine (GGGGGG) linker, aa 1503–1562 of MLL4-PHD6) was cloned into a pDEST-15 vector. The human MLL3 PHD7 domain (aa 1083–1143) was cloned into a pGEX-6p vector. All constructs were expressed in Rosetta2 (DE3) pLysS cells. Protein production was induced with 0.5 mM IPTG overnight at 16 °C in Luria broth (LB) or minimal media (M9) supplemented with <sup>15</sup>NH<sub>4</sub>Cl and 0.06 mM ZnCl<sub>2</sub> or <sup>15</sup>NH<sub>4</sub>Cl, <sup>13</sup>C-glucose, and 0.06 mM ZnCl<sub>2</sub>. The GST-tagged PHD proteins were purified on glutathione Sepharose 4B beads (GE Healthcare) in 50 mM Tris-HCl (pH 7.5) buffer, supplemented with 250 mM NaCl and 2 mM dithiothreitol (DTT). The GST tag was cleaved with Thrombin or TEV protease overnight at 4 °C. Unlabeled proteins were further purified by size exclusion chromatography and concentrated in Millipore concentrators. All





**Fig. 7** Comparative analysis of the H3K4me3 and H4K16ac recognition by histone readers. **a** Superimposed  $^1\text{H},^{15}\text{N}$  heteronuclear single quantum coherence spectra of MLL3-PHD7<sub>1083-1143</sub> collected upon titration with H4K16ac peptide. Spectra are color coded according to the protein:peptide molar ratio. **b-d** Structures of the BRD4 BD, Taf14 YEATS, and DPF3b DPF readers in complex with the indicated histone peptides<sup>24,50,51</sup>. PDB IDs: 3UUV (b), 5I0K (c), and 5I3I (d). **e** Structure of the ING2 PHD finger in complex with H3K4me3<sup>30</sup>. PDB ID: 2G6Q. **f** A zoom-in view of the overlaid histone tail-binding sites of the ING2 PHD-H3K4me3 and MLL4 PHD6-H4K16ac complexes. **g** Alignment of amino acid sequences of the PHD fingers from human MLL4, MLL3, BPTF, and ING2. Absolutely, moderately, and weakly conserved residues are colored blue, green, and yellow, respectively. Zinc-coordinating cysteine/histidine residues are labeled

mutants were generated by site-directed mutagenesis using the Stratagene Quik-Change mutagenesis protocol, grown, and purified as WT proteins.

**NMR experiments.** NMR experiments were carried out at 298 K on Varian INOVA 600 and 900 MHz spectrometers. NMR samples for structure determination were prepared in 20 mM Tris-HCl (pH 7.0) buffer, supplemented with 100 mM NaCl, 5 mM DTT, and 8% D<sub>2</sub>O. Backbone and side chain chemical shift assignments for the H4K16ac-bound MLL4-PHD6 finger were obtained by collecting and processing a set of triple resonance experiments (HNCACB, CBCA (CO)NH, CC(CO)NH, HBHA(CO)NH, HNCA, H(CCO)NH)<sup>39</sup> with Non Linear Sampling<sup>40</sup>. Chemical shift assignments were obtained for 96% of backbone amides (minus proline residues). The NMR sample contained 2.5 mM  $^{13}\text{C}/^{15}\text{N}$ -labeled MLL4-PHD6 and 7.5 mM H4K16ac (11–21) peptide (the 3-fold excess of peptide

almost saturated the 2.5 mM  $^{13}\text{C}/^{15}\text{N}$ -labeled MLL4-PHD6 sample based on  $^1\text{H},^{15}\text{N}$  HSQC and  $^1\text{H},^{13}\text{C}$  HSQC titration experiments with H4K16ac (11–21) peptide (Supplementary Fig. 6). Backbone chemical shift assignments for the apo-state were obtained through  $^1\text{H},^{15}\text{N}$  HSQC titration experiments using backbone chemical shift assignments for the H4K16ac-bound MLL4 PHD6 finger. Three-dimensional (3D)  $^{15}\text{N}$ - and  $^{13}\text{C}$ -edited NOESY-HSQC spectra (mixing time of 100 ms) were collected to obtain distance restraints. Two-dimensional (2D)  $^{15}\text{N}/^{13}\text{C}$ -filtered NOESY-HSQC and TOCSY-HSQC spectra were collected to obtain chemical shift assignments for the H4K16ac peptide. NMR experiments, including 2D NOESY, TOCSY, and reverse HSQC titration experiments, were performed on a synthetic  $^{15}\text{N}/^{13}\text{C}$ -labeled H4K16ac peptide (residues 11–21, synthesized by Bio-Synthesis) to validate the assignments. 3D  $^{15}\text{N}/^{13}\text{C}$ -filtered  $^{13}\text{C}$ - and  $^{15}\text{N}$ -edited NOESY-HSQC spectra (mixing time of 150 ms) were collected to obtain inter-molecular distance restraints.

For NMR titration experiments,  $^1\text{H}$ ,  $^{15}\text{N}$  HSQC spectra of 0.15 mM uniformly  $^{15}\text{N}$ -labeled WT or mutated MLL4-PHD6 were recorded while peptides (synthesized by SynPeptide) were added stepwise. All experiments were carried out in phosphate-buffered saline buffer (pH 6.5).

**Structure determination for the MLL4-H4 complex.** Calculation of the structure of MLL4-PHD6 (aa 1503–1562) in complex with H4K16ac (11–21) peptide was carried out using interproton NOE-derived distance restraints and dihedral angle restraints. NMR spectra were processed and analyzed with NMRDraw and CcpNmr Suite<sup>41</sup>. The program DANGLE in CcpNmr Suite was used to predict dihedral angle  $\psi$  and  $\phi$  restraints. Hydrogen bonds were derived from characteristic NOE patterns in combination with dihedral angles. The zinc-coordinating residues were initially identified through sequence alignment with other PHD fingers, and all cysteines but C1526 showed characteristic  $^{13}\text{C}\alpha$  and  $^{13}\text{C}\beta$  chemical shifts of zinc-ligating cysteine residues<sup>42</sup>. To confirm that C1526 but not the preceding H1525 residue coordinates the zinc ion, H1525 was mutated to alanine.  $^1\text{H}$ ,  $^{15}\text{N}$  HSQC spectrum of H1525A MLL4-PHD6 reveals that this mutant is well folded and therefore H1525 is not a zinc-ligating residue (Supplementary Fig. 14). The zinc-coordinating residues were further confirmed by intramolecular distance restraints. During the calculation, sequence region G<sub>13</sub>GAK<sub>ac</sub>RH<sub>18</sub> of the H4 peptide was used with N-terminal acetyl and C-terminal N-Me amide capping groups. The structures were calculated initially with XPLOR-NIH and refined with AMBER<sup>43,44</sup>. Hundred structures were calculated, and the ensemble of 15 conformers with the lowest total energy was selected to represent the complex between MLL4-PHD6 and H4K16ac peptide. The quality of the structures was validated using the program PROCHECK-NMR. The percentage of residues in the most favored, additionally allowed, generously allowed, and disallowed regions is 79.9, 16.8, 2.6, and 0.7, respectively. The structural statistics is listed in Supplementary Table 1.

**Fluorescence spectroscopy.** Spectra were recorded at 25 °C on a Fluoromax-3 spectrofluorometer (HORIBA). The samples containing 1.0  $\mu\text{M}$  MLL4 PHD6 domain, WT or mutants, and progressively increasing concentrations of peptide were excited at 295 nm. Experiments were performed in buffer containing 20 mM Tris (pH 7.2), 150 mM NaCl, and 1 mM DTT. Emission spectra were recorded over a range of wavelengths between 330 and 360 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:

$$\Delta I = \Delta I_{\max} \frac{\left( ([L] + [P] + K_d) - \sqrt{([L] + [P] + K_d)^2 - 4[P][L]} \right)}{2[P]} \quad (1)$$

where  $[L]$  is the concentration of the peptide,  $[P]$  is the concentration of MLL4-PHD6,  $\Delta I$  is the observed change of signal intensity, and  $\Delta I_{\max}$  is the difference in signal intensity of the free and bound states of MLL4-PHD6. The  $K_d$  value was averaged over three separate experiments, with error calculated as the standard deviation between the runs.

**Fluorescent MST-binding assay.** The MST experiments were performed using a Monolith NT.115 instrument (Nanotemper) as described previously<sup>26</sup>. All experiments were performed with the purified MLL4-PHD6 domain (aa 1503–1562) in a buffer containing 20 mM HEPES (pH 7.5), 150 mM KCl, 10 mM MgCl<sub>2</sub>, 10 mM  $\beta$ -ME, 1% sucrose, and 0.08% Tween-20. The final concentration of the N-terminal fluorescein-labeled histone H4 peptide (1–21, KE BIOCHEM) was kept at 200 nM. Dissociation constants for the interaction of MLL4-PHD6 with unlabeled peptides H4 (1–23) and H4K16ac (1–23) were measured using a displacement assay in which increasing amount of unlabeled peptides were added into a preformed PHD6:FAM-H4 complex prepared by supplementing 20  $\mu\text{M}$  PHD6 into each sample. The measurements were performed at 30% LED and 40% MST power with 3 s laser-on time and 22 s off time. For all measurements, samples were loaded into standard capillaries and 900–1300 counts were obtained for the fluorescence intensity. The  $K_d$  and IC<sub>50</sub> values were determined with the MO. Affinity Analysis software (NanoTemper Technologies GmbH), using two independent MST measurements. The  $K_i$  values for unlabeled peptides with PHD6 were determined from the IC<sub>50</sub> values observed in the displacement assay and converted by the following equation:

$$K_i = [I]_{50} / \left( \frac{[L]_{50}}{K_d} + \frac{[P]_0}{K_d} + 1 \right) \quad (2)$$

where  $[I]_{50}$  is the concentration of free unlabeled ligand at 50% binding and  $[L]_{50}$  is the concentration of free labeled H4 peptide at 50% binding. The  $K_d$  value is the dissociation constant of labeled H4 peptide determined in the direct binding experiment described above. Measurements for H4 unmodified and H4K16ac peptides were done in triplicates.

**Western blot and qRT-PCR.** Western blot of histone modifications using acid extracts or of nuclear proteins using nuclear extracts were done as described<sup>45,46</sup>. Total RNA was extracted using TRIzol (Invitrogen) and reverse transcribed using

the ProtoScript II First-Strand cDNA Synthesis Kit (NEB), following the manufacturers' instructions. qRT-PCR was done using the following SYBR green primers: *Mof* (forward, 5'-TGAGATCAACCATGTGCAGAAG-3', and reverse, 5'-AGGTGAGAAATACCAGGCATC-3'). qPCR was done using Luna<sup>®</sup> Universal qPCR Master Mix according to the manufacturers' instructions (NEB). To detect truncated *Mof* transcripts after 4-OHT treatment (2  $\mu\text{M}$ ), RT-PCR was performed as described<sup>10</sup>.

**Chromatin immunoprecipitation-sequencing.** In all,  $2 \times 10^7$  cells were cross-linked by adding formaldehyde to a final concentration of 1.5% for 10 min. To quench crosslinking, 125 mM glycine was added. After centrifugation, pellet was resuspended in Farmham lysis buffer, and the resulting nuclear pellet was sonicated in 2 ml TE buffer supplemented with protease inhibitors to fragment size of 200–500 bp. For ChIP, 8  $\mu\text{g}$  of MLL4 antibody or 4  $\mu\text{g}$  of H4K16ac antibody was pre-incubated with 50  $\mu\text{l}$  Dynabeads protein A (Life Technologies). Approximately 100–200  $\mu\text{g}$  of chromatin was mixed with preincubated protein A-antibody complex and incubated overnight at 4 °C with rotation. ChIP samples were washed and prepared by reverse-crosslinking. DNA was purified using the QIAquick PCR Purification Kit (QIAGEN). Hundred nanograms of ChIPed DNA was used to construct sequencing library using the NEBNext Ultra II DNA Library Prep Kit for Illumina (NEB). All ChIP-Seq samples were sequenced on the Illumina HiSeq 3000.

**ChIP-Seq data analysis.** Raw sequencing data were aligned to mouse mm9 genome using bowtie2. For peak calling of MLL4 and H4K16ac enrichment, SICER algorithm was used<sup>20</sup>. For MLL4 peak calling, a window size of 50 bp and a gap size of 50 bp was used. For H4K16ac, a window size of 200 bp and a gap size of 200 bp was used. Western blot data were used for normalization of H4K16ac. For GO analysis of the MLL4-associated genomic regions, GREAT version 3.0.0 was used (<http://great.stanford.edu/public/html/>)<sup>47</sup>. For enriched transcription factor motif analysis of the MLL4-binding sites, SeqPos motif tool in the Galaxy Cistrome was used (<http://cistrome.org/ap/root>) with default parameters<sup>48</sup>.

**Reporting summary.** Further information on research design is available in the Nature Research Reporting Summary linked to this article.

## Data availability

Coordinates and structure factors for the MLL4-PHD6:H4K16ac complex have been deposited in the Protein Data Bank (PDB ID 6O7G). NMR spectral parameters for the MLL4-PHD6:H4K16ac complex have been deposited to Biological Magnetic Resonance Data Bank under accession number 30585. ChIP-Seq data were submitted to GEO database under accession number GSE130091. All other relevant data supporting the key findings of this study are available within the article and its Supplementary Information files or from the corresponding authors upon reasonable request. The source data underlying Figs. 1f, 3c, and 4b and Supplementary Fig. 12 are provided as a Source Data file. A reporting summary for this article is available as a Supplementary Information file.

Received: 6 December 2018 Accepted: 30 April 2019

Published online: 24 May 2019

## References

- Froimchuk, E., Jang, Y. & Ge, K. Histone H3 lysine 4 methyltransferase KMT2D. *Gene* **627**, 337–342 (2017).
- Lee, J. E. et al. H3K4 mono- and di-methyltransferase MLL4 is required for enhancer activation during cell differentiation. *eLife* **2**, e01503 (2013).
- Wang, C. et al. Enhancer priming by H3K4 methyltransferase MLL4 controls cell fate transition. *Proc. Natl Acad. Sci. USA* **113**, 11871–11876 (2016).
- Lai, B. et al. MLL3/MLL4 are required for CBP/p300 binding on enhancers and super-enhancer formation in brown adipogenesis. *Nucleic acids Res.* **45**, 6388–6403 (2017).
- Jin, Q. et al. Distinct roles of GCN5/PCAF-mediated H3K9ac and CBP/p300-mediated H3K18/27ac in nuclear receptor transactivation. *EMBO J.* **30**, 249–262 (2011).
- Bonn, S. et al. Tissue-specific analysis of chromatin state identifies temporal signatures of enhancer activity during embryonic development. *Nat. Genet.* **44**, 148–156 (2012).
- Ali, M., Hom, R. A., Blakeslee, W., Ikenouye, L. & Kutateladze, T. G. Diverse functions of PHD fingers of the MLL/KMT2 subfamily. *Biochim. Biophys. Acta* **1843**, 366–371 (2014).
- Dhar, S. S. et al. Trans-tail regulation of MLL4-catalyzed H3K4 methylation by H4R3 symmetric dimethylation is mediated by a tandem PHD of MLL4. *Genes Dev.* **26**, 2749–2762 (2012).
- Shogren-Knaak, M. et al. Histone H4-K16 acetylation controls chromatin structure and protein interactions. *Science* **311**, 844–847 (2006).

10. Li, X. et al. MOF and H4 K16 acetylation play important roles in DNA damage repair by modulating recruitment of DNA damage repair protein Mdc1. *Mol. Cell Biol.* **30**, 5335–5347 (2010).
11. Deng, X. et al. Mammalian X upregulation is associated with enhanced transcription initiation, RNA half-life, and MOF-mediated H4K16 acetylation. *Dev. Cell* **25**, 55–68 (2013).
12. Keller, C. I. & Akhtar, A. The MSL complex: juggling RNA-protein interactions for dosage compensation and beyond. *Curr. Opin. Genet. Dev.* **31**, 1–11 (2015).
13. Li, X. et al. The histone acetyltransferase MOF is a key regulator of the embryonic stem cell core transcriptional network. *Cell Stem Cell* **11**, 163–178 (2012).
14. Thomas, T., Dixon, M. P., Kueh, A. J. & Voss, A. K. Mof (MYST1 or KAT8) is essential for progression of embryonic development past the blastocyst stage and required for normal chromatin architecture. *Mol. Cell Biol.* **28**, 5093–5105 (2008).
15. Nativio, R. et al. Dysregulation of the epigenetic landscape of normal aging in Alzheimer's disease. *Nat. Neurosci.* **21**, 497–505 (2018).
16. Fraga, M. F. et al. Loss of acetylation at Lys16 and trimethylation at Lys20 of histone H4 is a common hallmark of human cancer. *Nat. Genet.* **37**, 391–400 (2005).
17. Andrews, F. H., Strahl, B. D. & Kutateladze, T. G. Insights into newly discovered marks and readers of epigenetic information. *Nat. Chem. Biol.* **12**, 662–668 (2016).
18. Musselman, C. A., Lalonde, M. E., Cote, J. & Kutateladze, T. G. Perceiving the epigenetic landscape through histone readers. *Nat. Struct. Mol. Biol.* **19**, 1218–1227 (2012).
19. Taverna, S. D., Li, H., Ruthenburg, A. J., Allis, C. D. & Patel, D. J. How chromatin-binding modules interpret histone modifications: lessons from professional pocket pickers. *Nat. Struct. Mol. Biol.* **14**, 1025–1040 (2007).
20. Zang, C. et al. A clustering approach for identification of enriched domains from histone modification ChIP-Seq data. *Bioinformatics* **25**, 1952–1958 (2009).
21. Taylor, G. C., Eskeland, R., Hekimoglu-Balkan, B., Pradeepa, M. M. & Bickmore, W. A. H4K16 acetylation marks active genes and enhancers of embryonic stem cells, but does not alter chromatin compaction. *Genome Res.* **23**, 2053–2065 (2013).
22. Dhalluin, C. et al. Structure and ligand of a histone acetyltransferase bromodomain. *Nature* **399**, 491–496 (1999).
23. Li, Y. et al. AF9 YEATS domain links histone acetylation to DOT1L-mediated H3K79 methylation. *Cell* **159**, 558–571 (2014).
24. Andrews, F. H. et al. The Taf14 YEATS domain is a reader of histone crotonylation. *Nat. Chem. Biol.* **12**, 396–U333 (2016).
25. Zeng, L. et al. Mechanism and regulation of acetylated histone binding by the tandem PHD finger of DPF3b. *Nature* **466**, 258–262 (2010).
26. Mi, W. et al. The ZZ-type zinc finger of ZZZ3 modulates the ATAC complex-mediated histone acetylation and gene activation. *Nat. Commun.* **9**, 3759 (2018).
27. Jacobson, R. H., Ladurner, A. G., King, D. S. & Tjian, R. Structure and function of a human TAFII250 double bromodomain module.[see comment]. *Science* **288**, 1422–1425 (2000).
28. Su, D. et al. Structural basis for recognition of H3K56-acetylated histone H3-H4 by the chaperone Rtt106. *Nature* **483**, 104–107 (2012).
29. Li, H. et al. Molecular basis for site-specific read-out of histone H3K4me3 by the BPTF PHD finger of NURF. *Nature* **442**, 91–95 (2006).
30. Peña, P. V. et al. Molecular mechanism of histone H3K4me3 recognition by plant homeodomain of ING2. *Nature* **442**, 100–103 (2006).
31. Shi, X. et al. ING2 PHD domain links histone H3 lysine 4 methylation to active gene repression. *Nature* **442**, 96–99 (2006).
32. Wysocka, J. et al. A PHD finger of NURF couples histone H3 lysine 4 trimethylation with chromatin remodelling. *Nature* **442**, 86–90 (2006).
33. Lan, F. et al. Recognition of unmethylated histone H3 lysine 4 links BHC80 to LSD1-mediated gene repression. *Nature* **448**, 718–722 (2007).
34. Ooi, S. K. et al. DNMT3L connects unmethylated lysine 4 of histone H3 to de novo methylation of DNA. *Nature* **448**, 714–717 (2007).
35. Luo, H. et al. MOF acetylates the histone demethylase LSD1 to suppress epithelial-to-mesenchymal transition. *Cell Rep.* **15**, 2665–2678 (2016).
36. Cho, Y. W. et al. Histone methylation regulator PTIP is required for PPARgamma and C/EBPalpha expression and adipogenesis. *Cell Metab.* **10**, 27–39 (2009).
37. Hong, S. et al. Identification of JmjC domain-containing UTX and JMJD3 as histone H3 lysine 27 demethylases. *Proc. Natl Acad. Sci. USA* **104**, 18439–18444 (2007).
38. Park, Y. K. et al. Distinct roles of transcription factors KLF4, Krox20, and peroxisome proliferator-activated receptor gamma in adipogenesis. *Mol. Cell Biol.* **37**, e00554-16 (2017).
39. Sattler, M., Schleucher, J. & Griesinger, C. Heteronuclear multidimensional NMR experiments for the structure determination of proteins in solution employing pulsed field gradients. *Prog. Nucl. Mag. Res. Spectrosc.* **34**, 93–158 (1999).
40. Hyberts, S. G., Milbradt, A. G., Wagner, A. B., Arthanari, H. & Wagner, G. Application of iterative soft thresholding for fast reconstruction of NMR data non-uniformly sampled with multidimensional Poisson Gap scheduling. *J. Biomol. NMR* **52**, 315–327 (2012).
41. Vranken, W. F. et al. The CCPN data model for NMR spectroscopy: development of a software pipeline. *Proteins* **59**, 687–696 (2005).
42. Kornhaber, G. J., Snyder, D., Moseley, H. N. & Montelione, G. T. Identification of zinc-ligated cysteine residues based on 13CAlpha and 13CBeta chemical shift data. *J. Biomol. NMR* **34**, 259–269 (2006).
43. Schwieters, C. D., Kuszewski, J. J., Tjandra, N. & Clore, G. M. The Xplor-NIH NMR molecular structure determination package. *J. Magn. Reson.* **160**, 65–73 (2003).
44. Case, D. A. et al. The Amber biomolecular simulation programs. *J. Comput. Chem.* **26**, 1668–1688 (2005).
45. Jang, Y., Wang, C., Zhuang, L., Liu, C. & Ge, K. H3K4 methyltransferase activity is required for MLL4 protein stability. *J. Mol. Biol.* **429**, 2046–2054 (2017).
46. Jang, Y. et al. H3.3K4M destabilizes enhancer H3K4 methyltransferases MLL3/MLL4 and impairs adipose tissue development. *Nucleic Acids Res.* **47**, 607–620 (2018).
47. McLean, C. Y. et al. GREAT improves functional interpretation of cis-regulatory regions. *Nat. Biotechnol.* **28**, 495–501 (2010).
48. Liu, T. et al. Cistrome: an integrative platform for transcriptional regulation studies. *Genome Biol.* **12**, R83 (2011).
49. Lai, B. et al. MLL3/MLL4 are required for CBP/p300 binding on enhancers and super-enhancer formation in brown adipogenesis. *Nucleic Acids Res.* **43**, 6388–6403 (2017).
50. Filippakopoulos, P. et al. Histone recognition and large-scale structural analysis of the human bromodomain family. *Cell* **149**, 214–231 (2012).
51. Li, W., Zhao, A., Tempel, W., Loppnau, P. & Liu, Y. Crystal structure of DPF3b in complex with an acetylated histone peptide. *J. Struct. Biol.* **195**, 365–372 (2016).

## Acknowledgements

We thank Jianrong Lu for the lentiviral pLenti-CMV plasmids expressing WT or K274R mutant of MOF. This work was supported by grants from NIH GM106416, GM125195, and GM100907 to T.G.K., GM124736 to S.B.R., GM126900 to B.D.S., and DK071900 and CA129325 to R.G.R. and by the Intramural Research Program of NIDDK, NIH to K.G.

## Author contributions

Y.Z., Y.J., J.-E.L., J.W.A., L.X., M.R.H., E.M.C., K.K., B.J.K., and S.-P.W. performed experiments and, together with Y.D., R.G.R., B.D.S., S.B.R., X.S., K.G., and T.G.K., analyzed the data. Y.Z., J.-E.L., K.G., and T.G.K. wrote the manuscript with input from all authors.

## Additional information

**Supplementary Information** accompanies this paper at <https://doi.org/10.1038/s41467-019-10324-8>.

**Competing interests:** The authors declare no competing interests.

**Reprints and permission** information is available online at <http://npg.nature.com/reprintsandpermissions/>

**Journal peer review information:** *Nature Communications* thanks Stefan Knapp and other anonymous reviewer(s) for their contribution to the peer review of this work. Peer reviewer reports are available.

**Publisher's note:** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2019