Published in final edited form as:
Nat Struct Mol Biol. 2012 December ; 19(12): 1257–1265. doi:10.1038/nsmb.2434.

Phf19 links methylated Lys36 of histone H3 to regulation of Polycomb activity

Cecilia Ballaré[1,2,8], **Martin Lange**[1,2,8], **Audrone Lapinaite**[3,8], **Gloria Mas Martin**[4], **Lluis Morey**[1,2], **Gloria Pascual**[1,2], **Robert Liefke**[5,6], **Bernd Simon**[3], **Yang Shi**[5,6], **Or Gozani**[4], **Teresa Carlomagno**[3], **Salvador Aznar Benitah**[1,2,7], and **Luciano Di Croce**[1,2,7]

[1]Department of Gene Regulation and Stem Cells, Centre for Genomic Regulation (CRG), Barcelona, Spain.

[2]Universitat Pompeu Fabra, Barcelona, Spain.

[3]Structural and Computational Biology Unit, European Molecular Biology Laboratory, Heidelberg, Germany.

[4]Department of Biological Sciences, Stanford University, Stanford, California, USA.

[5]Division of Newborn Medicine and Program in Epigenetics, Department of Medicine, Children's Hospital, Harvard Medical School, Boston, Massachusetts, USA.

[6]Department of Cell Biology, Harvard Medical School, Boston, Massachusetts, USA.

[7]Institució Catalana de Recerca i Estudis Avançats (ICREA), Barcelona, Spain.

Abstract

Polycomb-group proteins are transcriptional repressors with essential roles in embryonic development. Polycomb repressive complex 2 (PRC2) contains the methyltransferase activity for Lys27. However, the role of other histone modifications in regulating PRC2 activity is just beginning to be understood. Here we show that direct recognition of methylated histone H3 Lys36 (H3K36me), a mark associated with activation, by the PRC2 subunit Phf19 is required for the full enzymatic activity of the PRC2 complex. Using NMR spectroscopy, we provide structural evidence for this interaction. Furthermore, we show that Phf19 binds to a subset of PRC2 targets in mouse embryonic stem cells and that this is required for their repression and for H3K27me3 deposition. These findings show that the interaction of Phf19 with H3K36me2 and H3K36me3 is essential for PRC2 complex activity and for proper regulation of gene repression in embryonic stem cells.

Correspondence should be addressed to L.D.C. (luciano.dicroce@crg.es)..

[8]These authors contributed equally to this work.

Note: Supplementary information is available in the online version of the paper.

Accession codes. Protein Data Bank: Structural data and coordinates for Phf19-Tudor in complex with H3K36me3 have been deposited with accession number 4BD3. Gene Expression Omnibus: Genome-wide data on PRC2 binding have been deposited with accession numbers GSE41277 and GSE41589.

AUTHOR CONTRIBUTIONS

C.B., M.L., A.L., G.M.M., L.M., G.P., R.L., B.S., Y.S., O.G., T.C., S.A.B. and L.D.C. designed, executed and analyzed the experiments. C.B., M.L., O.G., T.C., S.A.B. and L.D.C. wrote the manuscript.

COMPETING FINANCIAL INTERESTS

The authors declare no competing financial interests.

Polycomb group proteins have important roles in regulating embryonic development[1,2] and have been implicated in embryonic stem cell pluripotency[3-7]. The two Polycomb complexes, PRC1 and PRC2, have been characterized in depth. The PRC2 complex trimethylates histone H3 Lys27 (H3K27me3)[8], providing a docking site for proteins with a chromobox (Cbx) domain[9]. Proteins in the Cbx family are subunits of the PRC1 complex, and they facilitate PRC1 recruitment to target genes. The two complexes can silence genes either synergistically or independently of each other. Thus, regulation of H3K27 methylation represents an essential step in gene regulation by Polycomb proteins. In flies, the Polycomb complexes are recruited to chromatin at the Polycomb responsive elements (PRE). Although a few mammalian PRE sequences have been identified[10,11], it is still unclear how Polycomb group proteins are recruited to genome loci. Recent data suggest that mammalian PRC2 preferentially binds to CpG islands[12]. Moreover, long noncoding RNAs and transcription factors have also been implicated in modulating Polycomb group protein occupancy[13]. Among those, Jarid2 has been implicated in regulating the binding of the PRC2 complex to genomic targets in mouse embryonic stem (mES) cells; however, there are likely to be other mechanisms for fine-tuning[14-17]. For example, in *Drosophila melanogaster*, the Polycomb-like protein (Pcl) associates with PRC2 and is required for global H3K27 trimethylation[18,19]. A role for Polycomb-like proteins in regulating H3K27 trimethylation seems to be conserved in higher vertebrates[20-23]. Notably, all members of the human PCL family of proteins (Phf1, MTF2 and Phf19) contain several domains that have been implicated in recognizing methylated histone residues.

In this study, we have investigated the role of Phf19 in regulating PRC2 activity in mES cells. We report that Phf19 is associated with PRC2 proteins in a complex lacking Jarid2. Biophysical and structural data show that the Tudor domain of Phf19 binds with high affinity to methylated H3K36, a characteristic shared by all members of the PCL family. Depletion of Phf19 caused a global reduction of H3K27me3 and impaired the occupancy of PRC2 complex genome-wide. We observed similar effects in cells ectopically expressing Phf19 carrying a point mutation in its Tudor domain. Phf19 knockdown ES cells showed loss of pluripotency and defects in differentiation. Mechanistically, we further show that Phf19 has a pivotal role in silencing active promoters by directly targeting the PRC2 complex and H3K36 histone demethylases. Our data indicate that Phf19 is an important regulator of PRC2 function in mES cells.

RESULTS

Phf19 is a component of the PRC2 complex

We initially purified proteins that associate with Phf19, one of the Pcl orthologs in vertebrates, from 293T cells using tandem-affinity immunopurification (TAP) followed by SDS-PAGE and silver staining (Fig. 1a). MS analysis and western blotting of the TAP purification showed that PHF19 associates with core subunits EZH2, SUZ12 and EED of PRC2 (Fig. 1b,c). Moreover, we identified the proteins RBBP4 and RBBP7, which are also part of PRC2 (Fig. 1b,c). We confirmed the interaction of PHF19 with endogenous components of the PRC2 complex using co-immunoprecipitation (co-IP) experiments in mES cells (Fig. 1d). Size-exclusion chromatography[24] followed by western blotting showed that the majority of Phf19 co-migrates with the PRC2 (Fig. 1e), indicating that Phf19 is stably associated with this Polycomb complex. Notably, Jarid2, a protein recently identified as a PRC2 subunit, did not co-immunoprecipitate with Phf19, and the corresponding elution profile of Jarid2 peaked at a different fraction by size-exclusion chromatography. This suggests that Jarid2 and Phf19 reside in different PRC2 subcomplexes. Of note, none of the previous Jarid2 purifications has identified Phf19 as a subunit of PRC2 (refs. 14-17).

Polycomb-like proteins are evolutionarily conserved from flies to humans, and they contain a Tudor domain, two adjacent plant-homeodomain (PHD) fingers and a chromo-like domain (Fig. 1f). To begin to map the interaction of PHF19 with PRC2, we performed semi-endogenous co-immunoprecipitation using different parts of PHF19 as bait. We identified the C-terminal end of PHF19 as necessary and sufficient to co-immunoprecipitate endogenous SUZ12 (Fig. 1f,g). This suggests that this region facilitates binding to one of the PRC2 core components. Furthermore, the short isoform of human PHF19 (PHF19s) did not pull down SUZ12, indicating that it does not interact with PRC2. Further analysis of the interacting region showed that it is partially homologous to chromodomains first identified in the HP1 proteins (Fig. 1h). Secondary-structure prediction using PSIPRED[25] suggests that, in contrast to other described chromodomains[26], the α-helix of PHF19 is N-terminal to the β-sheets. We propose, therefore, to call this domain the 'reversed chromodomain'. Notably, the α-helix seems to be conserved among Mtf2, Phf19 and the *D. melanogaster* homolog Pcl, but it was not predicted for Phf1.

Phf19's Tudor domain binds to di- and trimethylated H3K36

To determine the domains necessary for regulating PRC2 activity at chromatin, we tested the purified, recombinantly expressed Phf19 domains for binding to a histone peptide array[27] (Fig. 2a). The Tudor domain of Phf19 bound to dimethylated (me2) and trimethylated (me3) H3K36, but not to numerous other methylated peptides. Histone peptide pull-down assays confirmed the specificity of the interaction (Fig. 2b).

We next asked whether the H3K36 interaction is conserved among proteins of this family, including the mammalian homologs Phf1 (Pcl1) and Mtf2 (Pcl2) and *D. melanogaster* Pcl. We found that the Tudor-H3K36 interaction is conserved among all proteins of the family, but *D. melanogaster* Pcl did not bind H3K36me2 or H3K36me3. Sequence comparison of the Tudor domain among the Pcl family members showed that several amino acids are conserved, including a set of aromatic residues (Supplementary Fig. 1a). These amino acids commonly give rise to the conserved 'aromatic cage' that accommodates methylated histone residues. Notably, the Tudor domain of *D. melanogaster* Pcl lacks one of these aromatic residues and fails to bind methylated H3K36 as an isolated polypeptide (Fig. 2b and Supplementary Fig. 1a). In line with this, a recently published NMR structure of *D. melanogaster* Pcl reveals a lack of a well-defined cage in this protein[28]. Nevertheless, using a Pcl Tudor–PHD1 construct, we were able to rescue the binding to H3K36me2 and H3K36me3 (Fig. 2b). Of note, none of the other Phf19 domains showed detectable binding to methylated H3K36.

Structural analysis of the Phf19 Tudor domain

We measured the binding of human PHF19-Tudor to an 11-mer H3K36me3-derived peptide (31-ATGGVKme3KPHRY-41) by NMR spectroscopy in 2D ^{15}N-^{1}H and 2D ^{13}C-^{1}H correlation spectra (Fig. 3a and Supplementary Fig. 2). Analysis of the ligand concentration–dependent chemical shift changes gave a K_d for the complex of 29 μM (Fig. 3b), which is in good agreement with the value derived by calorimetric analysis (data not shown). Similar to other trimethyllysine-binding domains, including single Tudor domains, double and tandem Tudor domains, chromodomains and PHD fingers[29-36], the H3K36me3 peptide maps to the region of the putative aromatic cage. We observed substantial chemical shift changes after binding for the H_N resonances of residue stretches 65–67 and 72–81, as well as for the aromatic resonances of Trp50, Tyr56, Phe74 and Tyr80 and for the methyl groups of Leu47 (Supplementary Fig. 2). Titration of Phf19-Tudor with the peptide (31-ATGGVKKPHRY-41), as well as surface plasmon resonance and pull-down assays with peptides with H3K36 or a modified H3K36me1, H3K36me2 or H3K36me3 residue, showed no substantial binding for unmodified or monomethylated peptides, but peptides with either

di- or trimethylated H3K36 supported the interaction (Fig.2b and Supplementary Fig. 3a,b). Indeed, NMR analysis with H3K36me2-containing peptide indicated a K_d for the complex of 295 μM (Supplementary Fig. 3c,d), one-tenth as strong as binding of the H3K36me3 peptide (K_d = 29 μM).

To gain insight into the specificity of the interaction between Phf19-Tudor and the H3K36me3 histone tail, we solved the solution structure of the complex by NMR (Table 1). A detailed description of the structure determination and data statistics is found in the supporting online material. Phf19-Tudor contains five antiparallel β-sheets, which form the characteristic β-barrel. Trimethyllysine is inserted into a deep hydrophobic pocket, formed by the aromatic cage residues Trp50 (β1-β2 loop), Tyr56 (β1), Phe74 (β3) and Tyr80 (β4), with Val82 (β4) delineating the bottom of the pocket (Fig. 3c). In agreement with this, the W50A, Y56A, F74A and Y80A mutants were unable to bind the H3K36me3 peptide (Supplementary Fig. 3e). The trimethyllysine residue was recognized by a classical π-cation interaction, aided by additional hydrophobic contacts with the floor of the pocket (Val82) and coulombic interactions with the negatively charged residues at the edge of the pocket (Asp76 and Ser78 from the β3-β4 loop). Additional negatively charged side chains (Glu75 from the same loop, and Thr51 and Asp52 from the β1-β2 loop) surround the pocket entry site (Fig. 3c,d) and compensate for the positive charge of the peptide. The mutant D52R did not bind the H3K36me3 peptide, either because of strong coulombic repulsion with the positively charged peptide or because of partial occlusion of the methyllysine-binding pocket by the long arginine side chain.

The peptide exists in an extended conformation, as is often observed for other complexes involving histone tails (see also Supplementary Note)[29,31,33,34,37]. The C-terminal half of the peptide interacts firmly with the Tudor domain, whereas the histone H3 residues 31–34 do not show any NOE connectivities to the protein. Accordingly, a peptide lacking the C-terminal part but containing the trimethyllysine ($H3_{31–36}$) was unable to bind Phf19-Tudor, whereas the 7-mer peptide histone $H3_{35–41}$ had a similar affinity to the 11-mer histone $H3_{31–41}$ peptide (K_d = 44 μM, Fig. 3b). Thus, the specificity of Phf19-Tudor for the H3K36me3 histone tail resides in the amino acid sequence following the methylated lysine. The side chain of H3K37, which lies next to H3K36me3, is involved in electrostatic interactions with the side chain of Phf19-Tudor residue Glu75. The rest of the peptide (H3 Pro38–Tyr41) fits snugly to a hydrophobic surface consisting of β1 residue Leu47, β2 residues Leu54, Tyr55, Tyr56 and Leu57, and the C-terminal residue Gln88. H3 Pro38 packs neatly against the side chain of Tyr56 and Leu54, whereas H3 His39 contacts both Leu47 and Leu57. Accordingly, the methyl resonances of Leu47 show a prominent chemical shift change upon binding the peptide, and several NOE connectivities to both H3 His39 and H3 Tyr41 (Supplementary Fig. 2b). H3 Arg340 points away from the protein, whereas H3 Tyr41 interacts with Leu47, Tyr55 and possibly Gln88. Although we observed no NOE connectivities between H3 Tyr41 and Gln88, the chemical shift change of the Gln88-Hβs upon peptide binding supports this interaction. The H_N of H3 His39 is close to the carbonyl of Tyr55 in most structures of the ensemble and could further stabilize the complex by an intermolecular hydrogen bond.

We confirmed this network of intermolecular interactions by inserting point mutations into histone H3 at amino acids 37–41. Using NMR, we measured the K_d of the complex with Phf19-Tudor for each of these alanine mutants (Fig. 3b). The H3 R10A mutation did not increase the complex K_d, in agreement with the idea that its side chain is directed toward the solvent. In contrast, we observed the most prominent effects for H3 K37A and H3 Y41A, with K_d values increased by a factor of 2.1 and 1.6, respectively.

The network of interactions seen in the complex structure provides a rationale for the selectivity of the Phf19 Tudor domain for the H3K36me histone tail, as observed in the peptide array (Fig. 2a). H3K36me3 uniquely contains a long positively charged side chain (Lys37) and two bulky hydrophobic side chains at the +1, +3 and +5 position from the trimethyllysine, respectively. A positively charged amino acid occupies position +1 from the methyllysine in only three histone-derived peptides, H3K36me3, H2AK75me3 and H2BK23me3 (Supplementary Fig. 1b). However, neither H2AK75me3 nor H2BK23me3 have bulky hydrophobic side chains at either the +3 or +5 position.

Phf19 colocalizes with PRC2 and regulates its occupancy

To investigate the physiological link between H3K36 methylation, Phf19 and PRC2, we first analyzed the expression profile of Phf19 during mES cell differentiation. Phf19 expression is downregulated after induction of differentiation by treatment with retinoic acid (Supplementary Fig. 4a). This pointed toward a potential role for Phf19 in both mES cell self-renewal and differentiation mechanisms. To identify direct target genes of Phf19 in mES cells, we performed chromatin immunoprecipitation followed by high-throughput sequencing (ChIP-seq). We identified 2,334 Phf19 target genes with peaks mostly found in close vicinity to the transcription start site (Fig. 4a, Supplementary Fig. 4b and Supplementary Table 1). Consistent with the finding that Phf19 associates with PRC2, 85% of Phf19 target genes were also bound by the PRC2 subunit Suz12, the vast majority of which were also decorated with H3K27me3 (Fig. 4b,c). In line with being Polycomb targets, gene ontology analysis revealed that many genes encoded signaling molecules and transcription factors required for embryonic development that are normally repressed in mES cells (Fig. 4d). We validated the ChIP-seq results and the specificity of the Phf19 antibody by performing ChIP-qPCR in control cells (shRandom) and cells knocked down for Phf19 (shPhf19). Target genes obtained from the ChIP-seq analysis were also bound in control cells, but knockdown of Phf19 led to a marked reduction in their occupancy (Fig. 4e and Supplementary Fig. 4a). Notably, the occupancy of Phf19 was substantially reduced in differentiated mES cells (Fig. 4f), consistent with its downregulation upon retinoic acid–induced differentiation of ES cells (retinoic acid treatment 72 h). We then asked whether an intact PRC2 is required for Phf19 to bind to chromatin. To this end, we performed ChIP-qPCR in cells in which Eed, a core subunit of the PRC2 (ref. 38), is deleted ($Eed^{-/-}$). Binding of Phf19 to target genes was almost entirely absent in these cells, suggesting that Phf19 binds to target genes as an integral component of PRC2 (Fig. 4g). To test whether Phf19 is required for PRC2 binding to target genes, we looked at genome-wide H3K27me3 and PRC2 subunits in cells following Phf19 knockdown. Compared to control cells, Phf19 knockdown cells showed a substantial reduction of Suz12 levels in 82% of the Phf19, Suz12 and H3K27me3 co-targets. Moreover, 71% of the target genes completely lost their H3K27me3 mark upon Phf19 depletion (Fig. 5a,b). However, in contrast to the effect on the core subunits of PRC2 (ref. 39), knockdown of Phf19 did not affect the protein levels of PRC2 components or the stability of the PRC2 complex (Supplementary Fig. 4c).

Overall, these results confirm that Phf19 binds to target genes only in the context of an intact PRC2 and that it is required for a stable association of PRC2 at target genes. In line with this, we observed a global reduction of H3K27me3 in three independent Phf19 knockdown cell lines, with a corresponding increase of H3K27me1 and acetylated H3K27 (H3K27ac, Fig. 5c).

These data are in good agreement with the phenotype observed upon loss of Pcl in *D. melanogaster*, as well as that observed with deletion of Phf1 (Pcl1) in mammals[18,22,23]. Of note, it has been recently shown that the loss of Mtf2 (Pcl2) in mES cells leads to enhanced

pluripotency and to sustained levels of H3K27me3, suggesting functional diversification among members of the Pcl family[21].

Role of Phf19 in mES cell self-renewal and differentiation

Global mRNA expression analysis using microarrays showed that an approximately equal number of genes were upregulated and downregulated upon depletion of Phf19 (Fig. 6a,b and Supplementary Table 2). Of the upregulated genes, 20% were direct Phf19 targets and 44% were PRC2 targets, as determined by Suz12 occupancy (Supplementary Fig. 4d,e). As expected, gene ontology annotation showed that the genes differentially expressed in the Phf19 knockdown are highly enriched in developmental regulators (Supplementary Fig. 4f).

Notably, the cohort of downregulated genes in Phf19-knockdown ES cells included those encoding several pluripotency markers, such as Nanog and Klf4 (Fig. 6b). In agreement with this, we observed a loss of pluripotency in Phf19-knockdown ES cells, as judged by the number of alkaline phosphatase–positive colonies (Supplementary Fig. 4g). Although many cells remained alkaline phosphatase positive (Supplementary Fig. 4g, above), a small but substantial number of them became alkaline phosphatase negative upon knockdown of Phf19 (Supplementary Fig. 4g, below).

We next investigated whether Phf19 has a role during differentiation of mES cells into embryoid bodies. In Phf19-depleted embryoid bodies, we observed that the pluripotency markers (such as *Oct4* and *Nanog*) were not completely silenced (Supplementary Fig. 4h), in agreement with the role of Polycomb complexes in directly regulating repression of these genes during differentiation[9,40]. Moreover, Phf19-depleted embryoid bodies showed a differentiation delay, as indicated by the sustained expression of early-differentiation markers for ectoderm (*Fgf5*) and mesoderm (Brachyury/T) lineages (Supplementary Fig. 4h). When subcutaneously injected into immunodeficient mice, Phf19-depleted cells produced teratomas that contained tissues representative of the three germ layers, with an overrepresentation of ectoderm tissues (Supplementary Fig. 5a–c). Indeed, analysis of the gene expression profile comparing wild-type and Phf19-knockdown mES cells (Fig. 6a) indicated that several ectoderm and trophoectoderm markers were upregulated, whereas endoderm and mesoderm markers were either unaffected or downregulated (Supplementary Fig. 5d). Notably, the sizes of the PHF19-depleted teratomas were, on average, substantially smaller than the controls (Fig. 6c and Supplementary Fig. 5e).

Phf19 facilitates re-silencing of transcribed genes

It has been recently shown that PRC2 enzymatic activity is inhibited by the presence of the H3K36me mark on mononucleosome substrates[8,41]. This might prevent the spreading of repressive chromatin states to active promoters. The observation that loss of Phf19 leads to reduced H3K27me3 levels (Fig. 5c) prompted us to test whether the Tudor-H3K36me interaction might relieve the H3K36me2- and H3K36me3-dependent inhibition of PRC2 activity at Phf19 target promoters to facilitate re-silencing of transcribed genes. Unbiased analysis of Phf19-bound peaks showed an enrichment of H3K36me2 and H3K36me3 marks when compared to an equal number of random sequences. Validation of Phf19 targets confirmed an enrichment of H3K36me3 at the transcription start site (TSS) when compared to the corresponding 2-kb upstream regions (Supplementary Fig. 5f).

The moderate genome-wide overlapping observed between Phf19 and the H3K36me2 and H3K36me3 marks suggested that H3K36me2 and H3K36me3 histone demethylases (KDMs) might be involved in Phf19-mediated gene silencing. We therefore analyzed the occupancy of several H3K36 KDMs at Phf19 target genes. We focused on the KDM4 family, which demethylate H3K9me3 and (with a much lower affinity) H3K36me3, and on

the KDM2 family, which has been shown to demethylate H3K36me2 but could theoretically also act on H3K36me3 (ref. 42). We found that KDM2b occupied several Phf19 target genes (Fig. 7a). To get further insight into the mechanist links between Phf19, KDM2b and Polycomb complexes, we triggered ES cell differentiation with retinoic acid and investigated several promoters whose activity depends on Phf19. In all cases, retinoic acid administration led to promoter activation, a reduction in Polycomb occupancy concomitant with Phf19 and KDM2b displacement, and an accumulation of H3K36me2 and H3K36me3 marks. Notably, when Phf19 was knocked down in mES cells, the occupancy of KDM2b was also substantially reduced (data not shown). Together, these data suggest that KDM2b participates directly in regulating Phf19 target genes, although we cannot exclude the possibility that other KDMs might also have a role.

The Phf19 Tudor domain is required for PRC2 function

Finally, to show that recognition of H3K36me by the Tudor domain of Phf19 is required for Polycomb function and for H3K27me3 deposition, we knocked down Phf19 in mES cells and replaced it with the human form of Phf19 (hPhf19), which is insensitive to the knockdown shRNA used for the endogenous Phf19. The ectopically expressed human form of Phf19 was either wild type (WT.hPhf19) or mutated in its Tudor domain (mut.hPhf19), which rendered it incapable of binding to H3K36me (Supplementary Fig. 3e) but still able to interact with the PRC2 complex (Supplementary Fig. 5g–i). ChIP analysis indicated that binding of endogenous Phf19 was diminished in knockdown cells; however, binding of tagged wild-type human Phf19 was detected in rescued cells and, to a lesser extent, in cells in which Tudor-mutated Phf19 was re-introduced (Fig. 7b). Notably, the frequencies of PRC2 binding and H3K27me3 marks were substantially reduced in Phf19 knockdown cells. Re-expression of the human wild-type Phf19 in knockdown cells rescued PRC2 binding, the frequency of H3K27me3 marks and promoter silencing; in contrast, this rescue was not observed in cells expressing Tudor-mutated Phf19 (Fig. 7b,c). Together, these data suggest that Phf19 facilitates and/or stabilizes binding of PRC2 complex at H3K36-methylated promoters. Binding of the Phf19 Tudor domain to H3K36me2 and H3K36me3 might serve to overcome the allosteric inhibition of the PRC2 observed *in vitro*[8], thus facilitating the recruitment of Polycomb proteins and of KDM2b, which in turn causes gene silencing.

DISCUSSION

In yeast, the H3K36me2 and H3K36me3 marks are implicated in recruiting a histone deacetylase–containing complex (Rpd3s) during RNA polymerase II–mediated transcription, which prevents aberrant transcription initiation within coding sequences. In higher eukaryotes, methylated H3K36 has been associated with active chromatin and with alternative splicing, but its potential roles in gene regulation and promoter silencing have so far been elusive[43]. Here we show that Phf19 binds to methylated H3K36 and recruits the PRC2 complex to promoters for silencing transcribed genes. Previously, it was reported that the PWWP domain of BRPF1 and the chromodomain of EAF3 (MRG15) bind specifically to methylated H3K36, but in both cases their affinity for H3K36me2 and H3K36me3 is much lower than that of Phf19-Tudor[37,44]. Hence, Phf19-Tudor is by far the strongest and most specific recognition module for methylated H3K36. Our structural data confirm the amazing diversity in the binding modes of histone tails containing methyllysines. This diversity revolves around the trimethyllysine recognition motif (the aromatic cage), which provides a conserved anchor (Supplementary Figs. 6 and 7). Furthermore, an assorted ensemble of intermolecular interactions around this anchor provides a means for specific recognition and regulation.

Analysis of Phf19 targets indicated that the H3K36me2 and H3K36me3 marks were enriched at the TSS, although with a lower intensity when compared to the corresponding coding regions. The high affinity of Phf19-Tudor for H3K36me2 and H3K36me3 might therefore compensate for a reduced level of H3K36me2 and H3K36me3, whereas the specificity of binding at the TSS is likely to be provided by other domains of Phf19 for noncoding RNA molecules or for combinations of histone marks. Notably, all members of the mammalian Pcl family posses a conserved aromatic cage within the Tudor domain. Our *in vitro* experiments further support their ability to recognize methylated H3K36. Thus, it seems that all mammalian PCL proteins have a pivotal role in silencing active promoters by directly targeting the PRC2 complex, although with different promoter specificity. In *D. melanogaster*, this function is likely to be performed by a dedicated transcription factor, because the fly Pcl-Tudor does not bind to methylated H3K36, nor to any of the histone modifications we tested.

Together with the PRC2 complex, Phf19 also recruits a specific demethylase for methylated H3K36, the KDM2b enzyme. In agreement with our data, it has been reported that KDM2b associates with Polycomb proteins[45] and is implicated in regulating senescence[46], leukemia progression[47] and reprogramming[48]. The presence of PRC2 complex and KDM2b would ensure that promoter silencing is achieved by the deposition of repressive marks (such as H3K27me3), and removal of the H3K36 methylation marks, which are associated with activation (Fig. 7d). We cannot exclude the possibility that other KDMs might also have a role.

There is a strong genome-wide overlap between Phf19 and PRC2, with 85% of Phf19 target genes being co-regulated by PRC2. Deletion of Eed in mES cells, which causes disassembly and degradation of the PRC2 proteins, led to a strong reduction of Phf19 occupancy to promoters. This suggests that a stable genome-wide occupancy of Phf19 is achieved only in the context of an intact PRC2 complex. Conversely, depletion of Phf19 in mES cells has a strong impact both on global levels of H3K27me3 and on PRC2 occupancy. Taken together, these data provide evidence for an intricate contribution of Phf19 and PRC2 core components to proper gene regulation, mES cell pluripotency and lineage choices.

This study thus unveils a previously unanticipated role of methylated H3K36 in modulating PRC2 function and promoter activity.

METHODS

Methods and any associated references are available in the online version of the paper.

ONLINE METHODS

Antibodies

An antibody against mouse Phf19 was generated by immunizing rabbits with a peptide corresponding to the N terminus of mouse Phf19 (ETQALEPGTLEAFGATSPNK). Sera were obtained and the antibody was purified using the same peptide and the AminoLink Kit (Thermo Scientific) according to the manufacturer's instructions. For details about other antibodies, see Supplementary Table 3.

Protein purification

We transfected cells from the human 293T cell line with the pcDNA3-CTAP vector containing full-length human PHF19. TAP and MS were done as described[49].

Glutathione S-transferase (GST) fusion proteins were purified from BL21 *Escherichia coli* according to the manufacturer's instructions. Briefly, cells were transformed with the respective pGEX3 expression vector, and expression was induced with 0.5 mM IPTG for 4 h at 37 °C.

Gel filtration

Gel filtration was done using 1 mg of nuclear extract from 293T cells fractionated on the AEKTA-Explorer system (Amersham), using a Superose 6 column, followed by western blotting.

Peptide pull-down assays

For histone peptide–binding assays, 1 μg of biotinylated histone peptide (Yale Peptide Facility) was incubated with 1 μg of purified GST fusion protein in binding buffer (50 mM Tris-HCl, pH 7.5, 150 mM NaCl, 0.1% (v/v) NP-40) for 2 h at 4 °C with rotation. Streptavidin beads (Dynabeads) were added and incubated for 1 h at 4 °C with rotation, followed by four washes in binding buffer. Bound proteins were analyzed by western blotting.

Immunoprecipitation

Tap-tagged PHF19 fragments were retrovirally introduced into HeLa-S cells to obtain stably expressing cells. Whole-cell extracts were made using CHAPS buffer (50 mM Tris-HCl, pH 7.8, 350 mM NaCl, 0.01 M CHAPS, 1 mM DTT). For immunoprecipitation, samples were incubated with anti-Flag beads (Sigma) for 3 h at 4 °C. After washing the beads with CHAPS buffer, the precipitated proteins were visualized by western blotting using antibodies against hemagglutinin (HA) and Suz12.

Biacore analysis

Biacore analyses were performed by immobilizing biotinylated histone peptides to streptavidin chips, to a response unit (RU) of approximately 300. Purified GST-labeled Tudor domain was used as an analyte in serial dilution, and K_d values were obtained using the Biacore analysis software.

Protein expression, purification and sample preparation for NMR experiments

The Tudor domain DNA sequence (residues 37–95) of the hs-Phf19 protein (UniProtKB Q5T6S3) was amplified by PCR and cloned into the expression vector pET-MCN (Novagen) using the NdeI-MfeI restriction sites. The resulting construct, encoding an N-terminal His_6-tagged Tudor domain with a tobacco etch virus (TEV) protease-cleavage site between them, was transformed into *E. coli* BL21 Rosetta2 cells (Novagen). Cells were grown at 37 °C in M9 minimal medium containing ^{15}N-labeled NH_4Cl, ^{13}C-labeled glucose, 33 mg ml^{-1} chloramphenicol and 50 mg ml^{-1} ampicillin. Protein expression was induced with 0.5 mM IPTG for 20 h at 22 °C. Harvested cells were resuspended in buffer A (50 mM Tris-HCl, pH 7.5, 30 mM imidazole, 500 mM NaCl) containing 0.2 mg ml^{-1} lysozyme, 2 U ml^{-1} RNase-free DNase (Promega), 10 mM β-mercaptoethanol and EDTA-free protease inhibitors (Roche) and then lysed by sonication. The lysate was spun down and the supernatant was loaded onto a His-TRAP FF column (GE Healthcare), washed with buffer A and buffer B (50 mM Tris-HCl, pH 7.5, 30 mM imidazole, 500 mM NaCl, 2 M LiCl, 5 mM β-mercaptoethanol) and eluted with a gradient of buffer C (50 mM Tris-HCl, pH 7.5, 500 mM imidazole, 500 mM NaCl, 5 mM β-mercaptoethanol). After exchanging to buffer A containing 1 mM DTT and TEV protease, the N-terminal His_6 tag was cleaved overnight at room temperature. The uncleaved protein and the TEV protease were removed by reverse purification with the His-TRAP FF column. The cleaved Tudor domain was further purified

by size-exclusion chromatography in buffer D (20 mM Tris-HCl, pH 7.0, 200 mM NaCl, 1 mM DTT) using the HiLoad 16/60 Superdex75 gel filtration column (GE Healthcare). Phf19-Tudor was eluted in 100 ml. Samples for NMR experiments were concentrated to 0.1–0.58 mM.

Peptides for NMR studies were dissolved in buffer D or buffer E (20 mM D_2-Tris-DCl, pH 7.0, 200 mM NaCl, 1 mM DTT, in H_2O or D_2O).

Peptides used were as follows: 31-ATGGV-Kme3-KPHRY-41 (H3K36me3); 31-ATGGVKKPHRY-41 (H3K36me0); 1-ART-Kme3-QTARK-9 (H3K4me3); 31-ATGGV-Kme3-K-37, 34-GV-Kme3-KPHRY-41; 31-ATGGV-Kme3-KPHDY-41 (H3K36me3-R40D); DANIK-Kme3-AMACI (random sequence); 34-GV-Kme3-APHRY-41 (H3K36me3-K37A); 34-GV-Kme3-KAHRY-41 (H3K36me3-P38A); 34-GV-Kme3-KPARY-41 (H3K36me3-H39A); 34GV-Kme3-KPHAY-41 (H3K36me3-R40A); 34-GV-Kme3-KPHRA-41 (H3K36me3-Y41A).

Samples for protein and peptide (H3K36me3) resonance assignment and protein-detected NOESY experiments were prepared with different molar ratios of protein to peptide (1:0.5, 1:1, 1:1.3, 1:2, 1:5).

Samples for hydrogen–deuterium (H–D) exchange experiments were prepared by lyophilizing free Tudor and Tudor-peptide complex followed by dissolving in D_2O.

NMR experiments

NMR experiments were performed on Bruker Avance III 600 MHz and 800 MHz spectrometers equipped with HCN triple-resonance cryo-probes. Spectra were acquired at a temperature of 298 K. Protein backbone assignments for free Phf19-Tudor and the Phf19-Tudor-H3K36me3 peptide complex were obtained using a combination of triple resonance HNCA[50-52], HNCACB[53,54] and CBCA(CO)NH[55] experiments. Protein side chain resonances in both free Phf19-Tudor and the complex were assigned from (H)CCH-TOCSY and H(C)CH-TOCSY[51] spectra, and amide-detected (H)CC(CO)NH-TOCSY and H(CCCO)NH-TOCSY[56,57] experiments. For the Phf19-Tudor–H3K36me3 peptide complex, (HB)CB(CD)HD and (HB)CB(CDCE)HE experiments[58] were performed to aid assignment of aromatic protein resonances.

Peptide resonances in the Phf19-Tudor–H3K36me3 peptide complex were assigned from double-^{12}C,^{14}N-filtered 2D NOESY (τ_m = 120 ms) and double-^{12}C,^{14}N-filtered 2D TOCSY spectra[59].

The binding of the peptides was monitored by simultaneous 2D $^{13}C/^{15}N$ HSQC correlations. The dissociation constants of different peptides were obtained following chemical shift perturbations of the H_N resonances upon addition of the ligand.

Analysis of the chemical shift index[60] and NOE cross-peak pattern allowed identification of secondary-structure elements. Distance constraints were collected from 3D ^{15}N-NOESY (τ_m = 150 ms), 3D ^{13}C-edited NOESY (τ_m = 130 ms), 3D ^{13}C-edited/^{12}C-filtered NOESY (τ_m = 150 ms), aromatic 3D ^{13}C-HMQC-NOESY and double-^{12}C,^{14}N-filtered 2D ^{1}H-^{1}H NOESY spectra in H_2O and in D_2O (τ_m = 150 ms)[59].

H/D experiments allowed us to detect amides involved in the formation of hydrogen bonds. The 2D ^{1}H-^{15}N HSQC spectra were recorded immediately after dissolving lyophilized samples in D_2O. Data were processed with NMRPipe[61] and analyzed using NMRView[62].

Structure calculation and refinement

The experimentally determined distance and dihedral angle restraints were applied in a simulated annealing protocol with CNS/ARIA1.2 (refs. 63,64) with modified annealing protocols[65]. All NOE connectivities were manually assigned and automatically calibrated using 8 iterations, calculating 20 structures in iterations 1–7 and 100 structures in iteration 8. In the spectra, we did not observe the NMR resonances of the Hε and the methyl hydrogens of the H3K36me3 residue. This is due to exchange broadening, as a consequence of the conformational exchange process between the bound and free form of the peptide in conjunction with the large ring-current shift imposed to the Hε and methyl hydrogens of bound H3K36me3 by the presence of the aromatic cage. In the absence of NOE connectivities between the H3K36me3 methyl hydrogens and the rest of the protein, H3K36me3 occupies the binding pocket of the aromatic cage, but its position is not well defined. To overcome this problem, we added four additional ambiguous interaction restraints between all aromatic carbons of the exchange-broadened residues Trp50, Tyr56, Phe74 and Tyr80 and the Nζ of H3K36me3 of 4.0 ± 1 Å, respectively. These nonexperimental restraints are justified by the large chemical shifts experienced by the aromatic side chains of Trp50, Tyr56, Phe74 and Tyr80, which clearly indicate their involvement in binding H3K36me3. The CNS E_{repel} function was used to simulate van der Waals interactions with an energy constant of 25.0 kcal mol^{-1} $Å^{-4}$, using 'PROLSQ' van der Waals radii[66]. Distance restraints were used with a soft square-well potential, using an energy constant of 50 kcal mol^{-1} $Å^{-2}$. For hydrogen bonds, distance restraints with bounds of 1.8–2.3 Å (H-O) and 2.8–3.3 Å (N-O) were imposed for slowly exchanging amide protons. Dihedral angle restraints derived from TALOS[67] were applied to ϕ and ψ backbone angles using energy constants of 200 kcal mol^{-1} rad^{-2}. The final ensemble of NMR structures was refined in a shell of water molecules[68,69]. Of 100 structures calculated in the last iteration, 30 were refined in explicit water. The Ramachandran statistics for the ten lowest-energy structures including residues 40–90 of PHF19-Tudor and residues 35–41 of H3K36me3 peptide are 94% in core, 5.3% in allowed, 0.2% in generous and 0.6% in disallowed regions (Table 1). Structural quality was analyzed using PROCHECK[70] and WHAT IF[71], and accessed through the icing web server (http://nmr.cmbi.ru.nl/icing).

The protein structure can be determined with high precision (backbone r.m.s. deviation, r.m.s. deviation = 0.33 Å), whereas the peptide is less well defined (backbone r.m.s. deviation = 0.85 Å) (Table 1 and Supplementary Fig. 5). This is due to the low affinity of the complex (K_d in the micromolar range), which reduces the number of intermolecular distances that can be collected. Nevertheless, the position of the H3K36m3 peptide on Phf19-Tudor is very well-defined and unique.

Embryonic stem cells

mES cells (E14 clone) were grown and differentiated as described[9]. mES cell differentiation was performed by addition of all-*trans* retinoic acid (1 μM) for 72 h. shRNA-mediated knockdown was performed using plKO.1 vectors obtained from Sigma.

Gene expression analyses

RNA was extracted using the RNeasy mini kit (Qiagen) and transcribed to cDNA by reverse transcription using the first-strand cDNA synthesis kit (Fermentas).

Real-time PCR analysis was performed using SYBR Green I PCR Master Mix (Roche) and the Roche LightCycler 480. Expression values were normalized to the housekeeping gene *Rpo*. Primer sequences are given in Supplementary Table 3. Microarray analysis was performed using Agilent Mouse Gene Expression Microarrays, and genes with fold change > 1.5 and adjusted $P < 0.01$ were considered to be significantly expressed. We used biomaRt

(Bioconductor package) to convert both unique Affymetrix IDs into MGI (Mouse Genome Informatics) IDs. We used a Bioconductor annotation packages (MmAgilentDesign026655.db) to convert unique Agilent probe IDs into MGI IDs.

Chromatin immunoprecipitation and chromatin immunoprecipitation–sequencing

ChIP was carried out essentially as described[9]. ChIP-seq was performed using 10 ng of immunoprecipitated DNA from each sample, followed by library amplification and sequencing on the Solexa system.

Mice injection

Cells were collected and resuspended in cold 0.25% (w/v) Matrigel (BD Biosciences)/PBS. Swiss Nude mice were injected at two flanks with 1×10^6 cells per flank. Two weeks after the injections, the mice were killed and their tumors were collected.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We thank V.A. Raker for help in manuscript preparation, D. Patel, G. Castellano, A. Ladurner and members of the Di Croce laboratory for discussions, and the Centre for Genomic Regulation (CRG) Genomic, Bioinformatic and Proteomic Units. Peptides for NMR studies were supplied by R. Pipkorn, DKFZ (German Cancer Research Center). This work was supported by grants from the Spanish 'Ministerio de Educación y Ciencia' (BFU2010-18692), from AGAUR (Agency for Administration of University and Research Grants), from the European Commission's 7th Framework Program 4DCellFate grant number 277899 to L.D.C. and from the US National Institutes of Health (NCI118487 and GM071004) to Y.S. M.L. was supported by a Juan de la Cierva fellowship, R.L. was supported by a research fellowship from the German research foundation (DFG, LI 2057/1-1), and L.M. was supported by a postdoctoral CRG-Novartis fellowship. T.C., B.S. and A.L. were supported by the European Molecular Biology Laboratory.

References

1. Sparmann A, van Lohuizen M. Polycomb silencers control cell fate, development and cancer. Nat. Rev. Cancer. 2006; 6:846–856. [PubMed: 17060944]
2. Schwartz YB, Pirrotta V. Polycomb silencing mechanisms and the management of genomic programmes. Nat. Rev. Genet. 2007; 8:9–22. [PubMed: 17173055]
3. Bracken AP, Dietrich N, Pasini D, Hansen KH, Helin K. Genome-wide mapping of Polycomb target genes unravels their roles in cell fate transitions. Genes Dev. 2006; 20:1123–1136. [PubMed: 16618801]
4. Boyer LA, et al. Polycomb complexes repress developmental regulators in murine embryonic stem cells. Nature. 2006; 441:349–353. [PubMed: 16625203]
5. O'Carroll D, et al. The polycomb-group gene Ezh2 is required for early mouse development. Mol. Cell. Biol. 2001; 21:4330–4336. [PubMed: 11390661]
6. Surface LE, Thornton SR, Boyer LA. Polycomb group proteins set the stage for early lineage commitment. Cell Stem Cell. 2010; 7:288–298. [PubMed: 20804966]
7. Sauvageau M, Sauvageau G. Polycomb group proteins: multi-faceted regulators of somatic stem cells and cancer. Cell Stem Cell. 2010; 7:299–313. [PubMed: 20804967]
8. Schmitges FW, et al. Histone methylation by PRC2 is inhibited by active chromatin marks. Mol. Cell. 2011; 42:330–341. [PubMed: 21549310]
9. Morey L, et al. Nonoverlapping functions of the Polycomb group Cbx family of proteins in embryonic stem cells. Cell Stem Cell. 2012; 10:47–62. [PubMed: 22226355]
10. Sing A, et al. A vertebrate Polycomb response element governs segmentation of the posterior hindbrain. Cell. 2009; 138:885–897. [PubMed: 19737517]

11. Woo CJ, Kharchenko PV, Daheron L, Park PJ, Kingston RE. A region of the human HOXD cluster that confers polycomb-group responsiveness. Cell. 2010; 140:99–110. [PubMed: 20085705]
12. Mendenhall EM, et al. GC-rich sequence elements recruit PRC2 in mammalian ES cells. PLoS Genet. 2010; 6:e1001244. [PubMed: 21170310]
13. Richly H, et al. Transcriptional activation of polycomb-repressed genes by ZRF1. Nature. 2010; 468:1124–1128. [PubMed: 21179169]
14. Landeira D, et al. Jarid2 is a PRC2 component in embryonic stem cells required for multi-lineage differentiation and recruitment of PRC1 and RNA Polymerase II to developmental regulators. Nat. Cell Biol. 2010; 12:618–624. [PubMed: 20473294]
15. Pasini D, et al. JARID2 regulates binding of the Polycomb repressive complex 2 to target genes in ES cells. Nature. 2010; 464:306–310. [PubMed: 20075857]
16. Peng JC, et al. Jarid2/Jumonji coordinates control of PRC2 enzymatic activity and target gene occupancy in pluripotent cells. Cell. 2009; 139:1290–1302. [PubMed: 20064375]
17. Shen X, et al. Jumonji modulates polycomb activity and self-renewal versus differentiation of stem cells. Cell. 2009; 139:1303–1314. [PubMed: 20064376]
18. Nekrasov M, et al. Pcl-PRC2 is needed to generate high levels of H3–K27 trimethylation at Polycomb target genes. EMBO J. 2007; 26:4078–4088. [PubMed: 17762866]
19. Savla U, Benes J, Zhang J, Jones RS. Recruitment of *Drosophila* Polycomb-group proteins by Polycomblike, a component of a novel protein complex in larvae. Development. 2008; 135:813–817. [PubMed: 18216170]
20. Casanova M, et al. Polycomblike 2 facilitates the recruitment of PRC2 Polycomb group complexes to the inactive X chromosome and to target loci in embryonic stem cells. Development. 2011; 138:1471–1482. [PubMed: 21367819]
21. Walker E, et al. Polycomb-like 2 associates with PRC2 and regulates transcriptional networks during mouse embryonic stem cell self-renewal and differentiation. Cell Stem Cell. 2010; 6:153–166. [PubMed: 20144788]
22. Cao R, et al. Role of hPHF1 in H3K27 methylation and Hox gene silencing. Mol. Cell. Biol. 2008; 28:1862–1872. [PubMed: 18086877]
23. Sarma K, Margueron R, Ivanov A, Pirrotta V, Reinberg D. Ezh2 requires PHF1 to efficiently catalyze H3 lysine 27 trimethylation *in vivo*. Mol. Cell. Biol. 2008; 28:2718–2731. [PubMed: 18285464]
24. Koh AS, et al. Aire employs a histone-binding module to mediate immunological tolerance, linking chromatin regulation with organ-specific autoimmunity. Proc. Natl. Acad. Sci. USA. 2008; 105:15878–15883. [PubMed: 18840680]
25. McGuffin LJ, Bryson K, Jones DT. The PSIPRED protein structure prediction server. Bioinformatics. 2000; 16:404–405. [PubMed: 10869041]
26. Eissenberg JC. Structural biology of the chromodomain: form and function. Gene. 2012; 496:69–78. [PubMed: 22285924]
27. Bua DJ, et al. Epigenome microarray platform for proteome-wide dissection of chromatin-signaling networks. PLoS ONE. 2009; 4:e6789. [PubMed: 19956676]
28. Friberg A, Oddone A, Klymenko T, Müller J, Sattler M. Structure of an atypical Tudor domain in the *Drosophila* Polycomblike protein. Protein Sci. 2010; 19:1906–1916. [PubMed: 20669242]
29. Lee J, Thompson JR, Botuyan MV, Mer G. Distinct binding modes specify the recognition of methylated histones H3K4 and H4K20 by JMJD2A-Tudor. Nat. Struct. Mol. Biol. 2008; 15:109–111. [PubMed: 18084306]
30. Botuyan MV, et al. Structural basis for the methylation state-specific recognition of histone H4–K20 by 53BP1 and Crb2 in DNA repair. Cell. 2006; 127:1361–1373. [PubMed: 17190600]
31. Huang Y, Fang J, Bedford MT, Zhang Y, Xu RM. Recognition of histone H3 lysine-4 methylation by the double Tudor domain of JMJD2A. Science. 2006; 312:748–751. [PubMed: 16601153]
32. Flanagan JF, et al. Double chromodomains cooperate to recognize the methylated histone H3 tail. Nature. 2005; 438:1181–1185. [PubMed: 16372014]
33. Jacobs SA, Khorasanizadeh S. Structure of HP1 chromodomain bound to a lysine 9-methylated histone H3 tail. Science. 2002; 295:2080–2083. [PubMed: 11859155]

34. Nielsen PR, et al. Structure of the HP1 chromodomain bound to histone H3 methylated at lysine 9. Nature. 2002; 416:103–107. [PubMed: 11882902]
35. Peña PV, et al. Molecular mechanism of histone H3K4me3 recognition by plant homeodomain of ING2. Nature. 2006; 442:100–103. [PubMed: 16728977]
36. Li H, et al. Molecular basis for site-specific read-out of histone H3K4me3 by the BPTF PHD finger of NURF. Nature. 2006; 442:91–95. [PubMed: 16728978]
37. Vezzoli A, et al. Molecular basis of histone H3K36me3 recognition by the PWWP domain of Brpf1. Nat. Struct. Mol. Biol. 2010; 17:617–619. [PubMed: 20400950]
38. Margueron R, et al. Role of the polycomb protein EED in the propagation of repressive histone marks. Nature. 2009; 461:762–767. [PubMed: 19767730]
39. Pasini D, Bracken AP, Jensen MR, Lazzerini Denchi E, Helin K. Suz12 is essential for mouse development and for EZH2 histone methyltransferase activity. EMBO J. 2004; 23:4061–4071. [PubMed: 15385962]
40. Li G, et al. Jarid2 and PRC2, partners in regulating gene expression. Genes Dev. 2010; 24:368–380. [PubMed: 20123894]
41. Yuan W, et al. H3K36 methylation antagonizes PRC2-mediated H3K27 methylation. J. Biol. Chem. 2011; 286:7983–7989. [PubMed: 21239496]
42. Cloos PA, Christensen J, Agger K, Helin K. Erasing the methyl mark: histone demethylases at the center of cellular differentiation and disease. Genes Dev. 2008; 22:1115–1140. [PubMed: 18451103]
43. Wagner EJ, Carpenter PB. Understanding the language of Lys36 methylation at histone H3. Nat. Rev. Mol. Cell Biol. 2012; 13:115–126. [PubMed: 22266761]
44. Sun B, et al. Molecular basis of the interaction of *Saccharomyces cerevisiae* Eaf3 chromo domain with methylated H3K36. J. Biol. Chem. 2008; 283:36504–36512. [PubMed: 18984594]
45. Sánchez C, et al. Proteomics analysis of Ring1B/Rnf2 interactors identifies a novel complex with the Fbxl10/Jhdm1B histone demethylase and the Bcl6 interacting corepressor. Mol. Cell. Proteomics. 2007; 6:820–834. [PubMed: 17296600]
46. He J, Kallin EM, Tsukada Y, Zhang Y. The H3K36 demethylase Jhdm1b/ Kdm2b regulates cell proliferation and senescence through $p15^{nk4b}$. Nat. Struct. Mol. Biol. 2008; 15:1169–1175. [PubMed: 18836456]
47. He J, Nguyen AT, Zhang Y. KDM2b/JHDM1b, an H3K36me2-specific demethylase, is required for initiation and maintenance of acute myeloid leukemia. Blood. 2011; 117:3869–3880. [PubMed: 21310926]
48. Liang G, He J, Zhang Y. Kdm2b promotes induced pluripotent stem cell generation by facilitating gene activation early in reprogramming. Nat. Cell Biol. 2012; 14:457–466. [PubMed: 22522173]
49. Lange M, et al. Regulation of muscle development by DPF3, a novel histone acetylation and methylation reader of the BAF chromatin remodeling complex. Genes Dev. 2008; 22:2370–2384. [PubMed: 18765789]
50. Grzesiek S, Bax A. Improved 3D triple-resonance NMR techniques applied to a 31-kDa protein. J. Magn. Reson. 1992; 96:432–440.
51. Kay LE, Xu GY, Singer AU, Muhandiram DR, Formankay JD. A gradient-enhanced HCCH-TOCSY experiment for recording side-chain H-1 and C-13 correlations in H2O samples of proteins. J. Magn. Reson. B. 1993; 101:333–337.
52. Schleucher J, Sattler M, Griesinger C. Coherence selection by gradients without signal attenuation – application to the 3-dimensional HNCO experiment. Angew. Chem. 1993; 32:1489–1491.
53. Wittekind M, Mueller L. HNCACB, a high-sensitivity 3D NMR experiment to correlate amide proton and nitrogen resonances with the alpha-carbon and beta-carbon resonances in proteins. J. Magn. Reson. B. 1993; 101:201–205.
54. Muhandiram DR, Kay LE. Gradient-enhanced triple-resonance 3-dimensional NMR experiments with improved sensitivity. J. Magn. Reson. B. 1994; 103:203–216.
55. Yamazaki T, Lee W, Arrowsmith CH, Muhandiram DR, Kay LE. A suite of triple-resonance NMR experiments for the backbone assignment of N-15, C-13, H-2 labeled proteins with high sensitivity. J. Am. Chem. Soc. 1994; 116:11655–11666.

56. Grzesiek S, Anglister J, Bax A. Correlation of backbone amides and aliphatic side-chain resonances in C-13/N-15-enriched proteins by isotropic mixing of C-13 magnetization. J. Magn. Reson. B. 1993; 101:114–119.
57. Montelione GT, Lyons BA, Emerson SD, Tashiro M. An efficient triple resonance experiment using C-13 isotropic mixing for determining sequence-specific resonance assignments of isotopically enriched proteins. J. Am. Chem. Soc. 1992; 114:10974–10975.
58. Yamazaki T, Forman-Kay JD, Kay LE. 2-dimensional NMR experiments for correlating C-13-beta and H-1-delta/epsilon chemical shifts of aromatic residues in C-13-labeled proteins via scalar couplings. J. Am. Chem. Soc. 1993; 115:11054–11055.
59. Zwahlen C, et al. Methods for measurement of intermolecular NOEs by multinuclear NMR spectroscopy: Application to a bacteriophage lambda N-peptide/boxB RNA complex. J. Am. Chem. Soc. 1997; 119:6711–6721.
60. Wishart DS, Sykes BD. The C-13 chemical shift index - a simple method for the identification of protein secondary structure using C-13 chemical shift data. J. Biomol. NMR. 1994; 4:171–180. [PubMed: 8019132]
61. Delaglio F, et al. NMRPIPE—A multidimensional spectral processing system based on UNIX pipes. J. Biomol. NMR. 1995; 6:277–293. [PubMed: 8520220]
62. Johnson BA, Blevins RA. NMR VIEW – A computer program for the visualization and analysis of NMR data. J. Biomol. NMR. 1994; 4:603–614. [PubMed: 22911360]
63. Linge JP, Habeck M, Rieping W, Nilges M. ARIA: automated NOE assignment and NMR structure calculation. Bioinformatics. 2003; 19:315–316. [PubMed: 12538267]
64. Nederveen AJ, et al. RECOORD: a recalculated coordinate database of 500+ proteins from the PDB using restraints from the BioMagResBank. Proteins. 2005; 59:662–672. [PubMed: 15822098]
65. Bernard A, Vranken WF, Bardiaux B, Nilges M, Malliavin TE. Bayesian estimation of NMR restraint potential and weight: a validation on a representative set of protein structures. Proteins. 2011; 79:1525–1537. [PubMed: 21365680]
66. Linge JP, Nilges M. Influence of non-bonded parameters on the quality of NMR structures: A new force field for NMR structure calculation. J. Biomol. NMR. 1999; 13:51–59. [PubMed: 10905826]
67. Cornilescu G, Delaglio F, Bax A. Protein backbone angle restraints from searching a database for chemical shift and sequence homology. J. Biomol. NMR. 1999; 13:289–302. [PubMed: 10212987]
68. Linge JP, Williams MA, Spronk C, Bonvin A, Nilges M. Refinement of protein structures in explicit solvent. Proteins. 2003; 50:496–506. [PubMed: 12557191]
69. Nozinovic S, Furtig B, Jonker HRA, Richter C, Schwalbe H. High-resolution NMR structure of an RNA model system: the 14-mer cUUCGg tetraloop hairpin RNA. Nucleic Acids Res. 2010; 38:683–694. [PubMed: 19906714]
70. Laskowski RA, Macarthur MW, Moss DS, Thornton JM. PROCHECK—A program to check the stereochemical quality of protein structures. J. Appl. Crystallogr. 1993; 26:283–291.
71. Vriend G. WHAT IF: a molecular modeling and drug design program. J. Mol. Graph. 1990; 8:52–56. 29. [PubMed: 2268628]

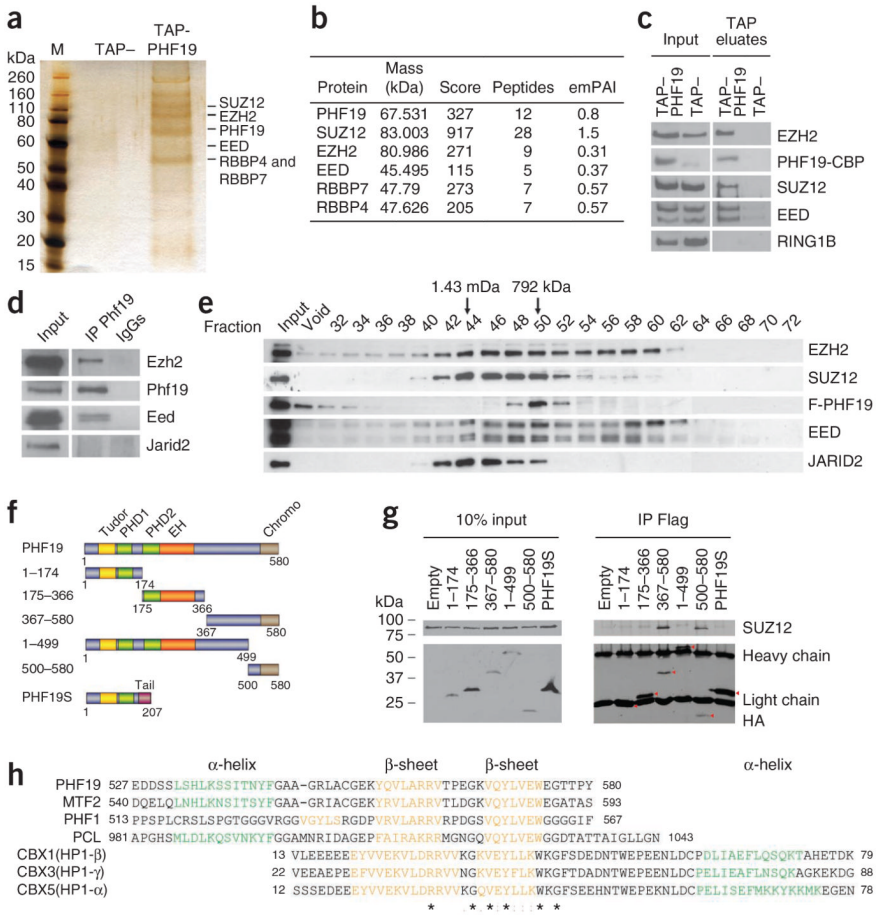

Figure 1.

Phf19 is a component of the PRC2 complex. (**a**) SDS-PAGE followed by silver staining of TAP-purified PHF19 (TAP-PHF19) in 293T cells. TAP−, TAP-empty. (**b**) Summary of peptides identified by MS in eluates from TAP-PHF19 purification. emPAI, experimentally modified protein abundance. (**c**) Western blot analysis of eluates from TAP-PHF19 purification confirming the specific association with the PRC2. Note that the PRC1 subunit RING1b does not interacts with PHF19. (**d**) Endogenous co-immunoprecipitation (IP) in mES cells of Phf19 with PRC2 components. Note that Jarid2 is not present in Phf19–PRC2 complexes. (**e**) Size-exclusion chromatography of 293T nuclear extracts followed by western blotting. Note that PHF19 and JARID2 co-migrate with components of the PRC2 complex but elute at different size fractions—around 790 kDa and 1.4 mDa, respectively. (**f**) Schematic representation of the domain architecture of human PHF19, depicting the Tudor domain, two plant homeodomain (PHD) fingers, the extended homology (EH) domain, and a chromo-like domain. (**g**) The chromo-like domain is necessary and sufficient to co-immunoprecipitate the endogenous PRC2 core component SUZ12. (**h**) Predicted secondary structures of the chromo-like domains compared to the first chromodomain of the HP1 proteins.

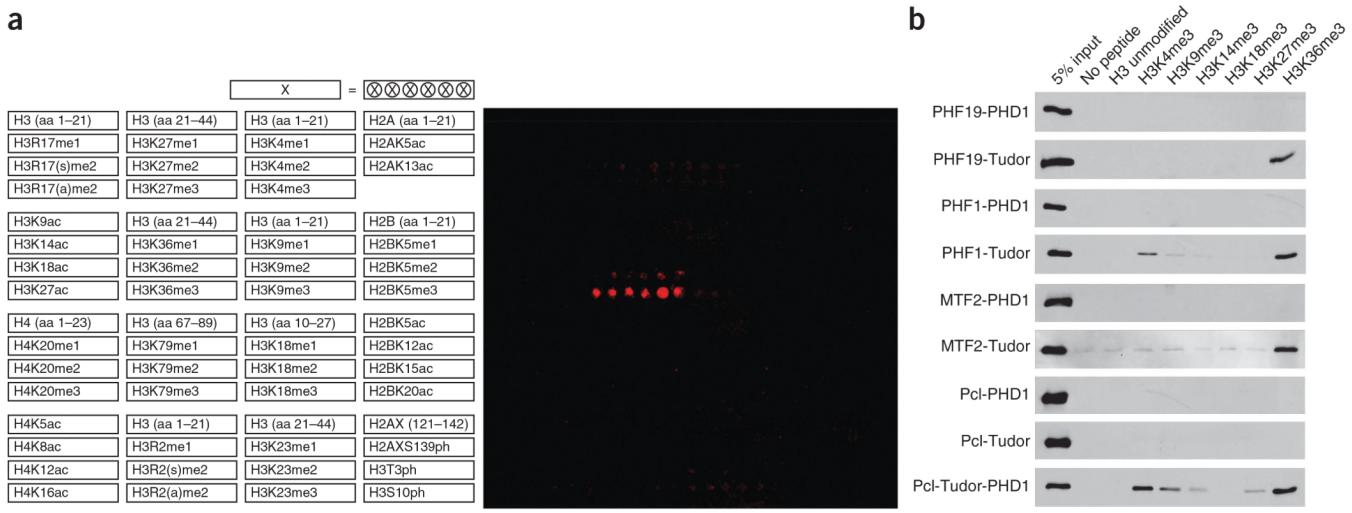

Figure 2.

Binding of Phf19-Tudor to methylated H3K36. (**a**) Histone peptide array showing specific binding of GST-labeled Phf19-Tudor to H3K36me2 and H3K36me3 peptides. aa, amino acids. (**b**) Histone peptide pull-down assay using recombinant Phf19, Phf1, MTF2 and Pcl domains with methylated histone peptides. Phf19-Tudor, Phf1-Tudor and MTF2-Tudor all bind to H3K36me2 and H3K36me3 peptides. *D. melanogaster* Pcl-Tudor, which lacks an otherwise conserved aromatic amino acid, requires the addition of the PHD1 for binding.

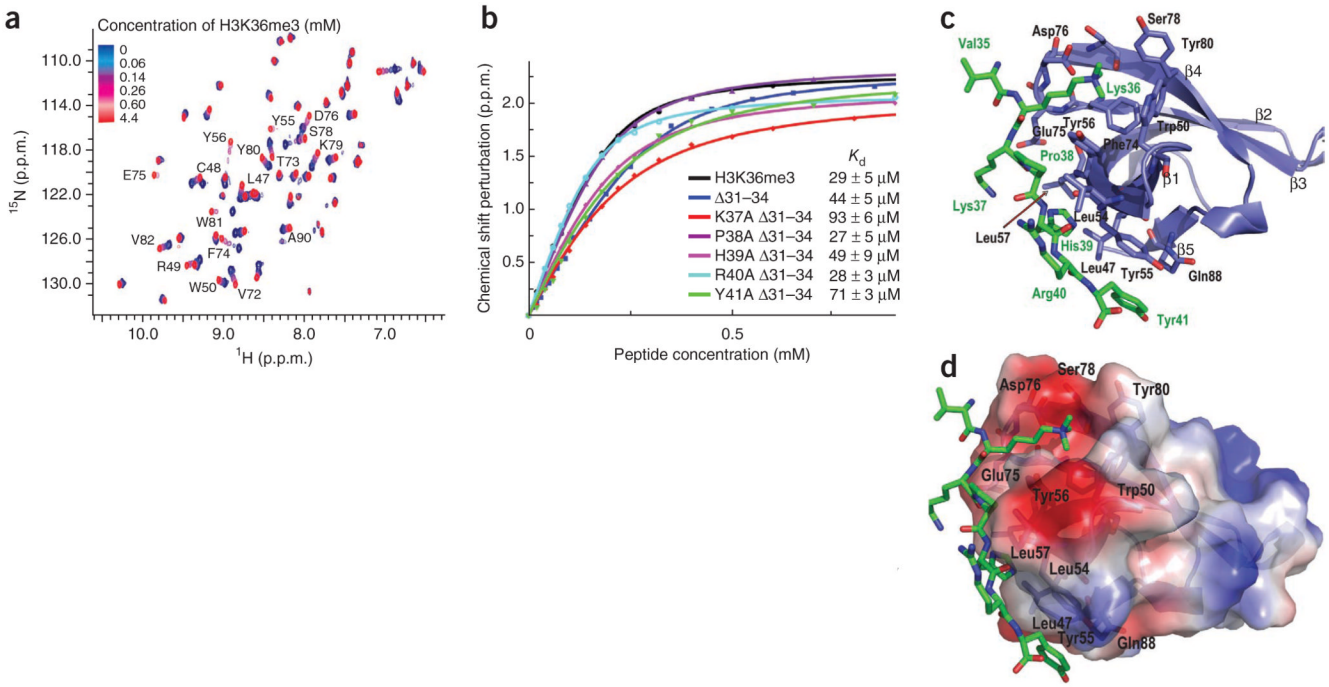

Figure 3.

NMR-based structural analysis of the complex between Phf19-Tudor and an H3 peptide methylated at Lys36. (**a**) NMR 2D ^{15}N-^{1}H correlation spectrum showing the chemical shift changes of the H_N resonances of Phf19-Tudor upon addition of the 11-mer H3K36me3 peptide (31-ATGGVKme3KPHRY-41; from blue to red). The protein concentration was 0.2 mM. Amino acid single-letter code is used for brevity. (**b**) Plots of the changes in chemical shifts induced upon complex formation as a function of the concentration of the H3K36me3 wild-type (amino acids 31–41) and mutant peptides for the amide resonance of Asp76. The values of K_d are calculated as average values for 8 to 14 well-resolved H_N resonances of the complexes. (**c**) Ribbon representation of the lowest-energy structure of the Phf19-Tudor complex with the H3K36me3 peptide. Phf19-Tudor side chains in contact with the H3K36me3 peptide are shown as sticks. Phf19-Tudor: C, violet; O, red; N, blue. H3K36me3 peptide: C, green; O, red; N, blue. Amino acids of Phf19-Tudor are labeled in black and those of the H3K36me3 peptide in green (PDB 4BD3). (**d**) Surface representation of the Phf19-Tudor domain colored by electrostatic charge. White, neutral; blue, positive; red, negative. Side chains of the protein in contact with the peptide are shown in violet. Only residues 35–41 are shown for the histone H3K36me3 peptide. Phf19-Tudor: C, violet; O, red; N, blue; H3K36me3 peptide: C, green; O, red; N, blue.

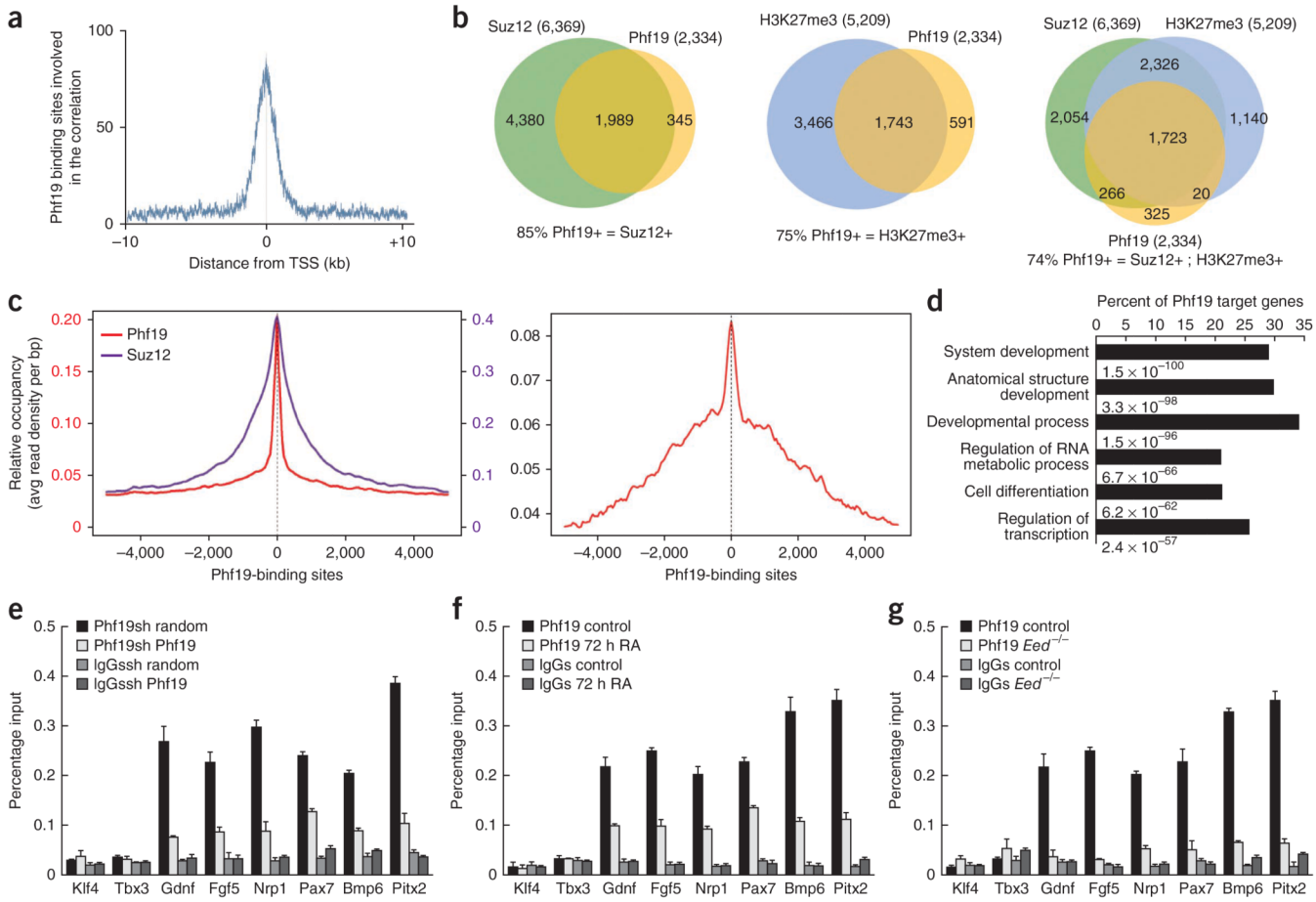

Figure 4.

Phf19 is in integral part of the PRC2 complex in mES cells. (**a**) Histogram showing the distribution of Phf19 ChIP-seq peaks relative to the TSS. Almost all (93%) of Phf19 peaks were found within ±10 kb of a known TSS. (**b**) Venn diagrams showing the overlap of Phf19, Suz12 and H3K27me3 target genes in mES cells. (**c**) Suz12, Phf19 and H3K27me3 occupancy in Phf19 binding sites are plotted as the average profile of ChIP-seq reads (read density per base pair) around the summit of Phf19 peaks. (**d**) Gene ontology analysis of Phf19 target genes in mES cells. (**e–g**) ChIP-qPCR assays using an antibody anti-Phf19 and IgG as negative control. Results are presented as a percentage of the input material. Values represent the average and s.d. of three independent experiments. (**e**) Validation of Phf19 target genes. Phf19 knockdown cells showed reduced binding, demonstrating the specificity of the antibody. Klf4 and Tbx3 served as negative controls. (**f**) ChIP-qPCR analysis shows a reduced Phf19 binding to target genes after differentiation induction by all-*trans* retinoic acid (RA) supplemented at 1×10^{-6} M for 72 h. (**g**) ChIP-qPCR analysis in cells deleted for Eed, a core subunit of the PRC2, indicates that Phf19 binding is dependent on a functional PRC2. $n = 3$ in **e–g**.

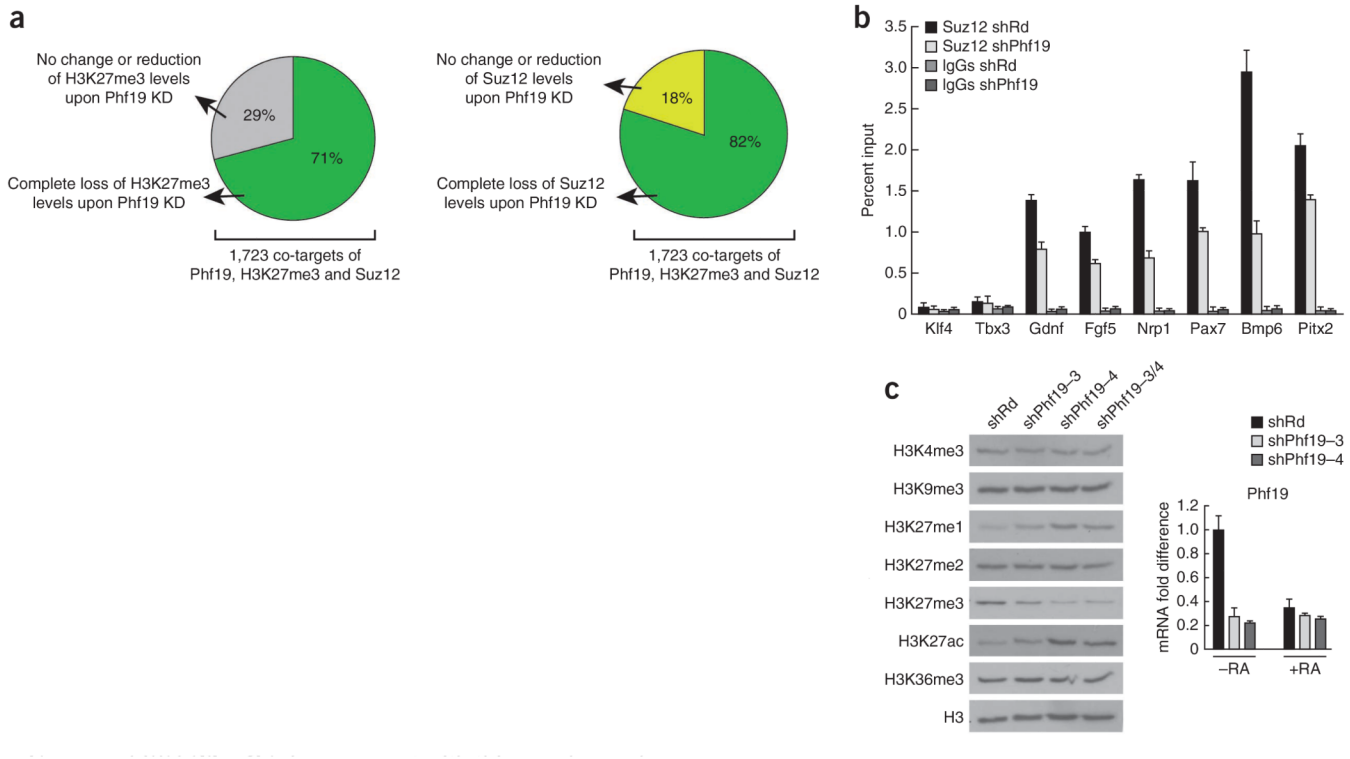

Figure 5.

Phf19 is required for PRC2 binding to target genes. (**a**) Changes in H3K27me3 and Suz12 levels evaluated by ChIP-seq analysis in Phf19, H3K27me3 and Suz12 co-targets following Phf19 knockdown (KD). (**b**) ChIP-qPCR analysis shows that binding of the PRC2 component Suz12 is reduced in Phf19 knockdown cells. Data represent average and s.d. of three independent experiments. (**c**) Western blot analysis of histone modifications in Phf19 knockdown cells shows a global reduction of H3K27me3 and a concomitant increase in H3K27me1 and H3K27ac. Two different Phf19-depleted cell lines were analyzed (shPhf19–3 and shPhf19–4) (left). Expression levels of Phf19 were determined by RT-qPCR in the wild type and in the cell lines depleted of shPhf19–3 and shPhf19–4. Data represent the average and s.d. of three independent experiments (right).

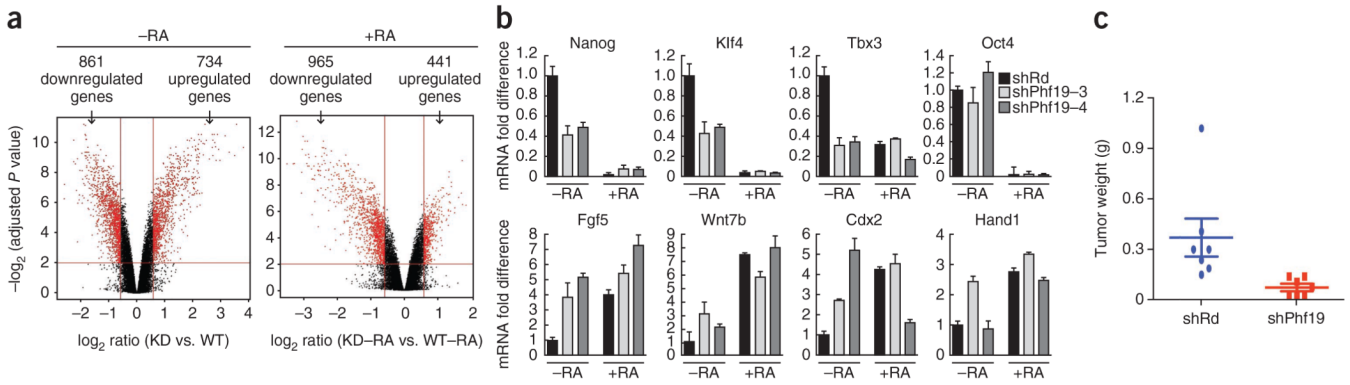

Figure 6.

Role of Phf19 in pluripotency. (**a**) Microarray expression analysis of Phf19 knockdown cells. In cells maintained under pluripotency conditions (−RA), 734 genes were upregulated and 861 were downregulated (fold change >1.5 and adjusted, $P < 0.01$). After treatment with all-*trans* RA (10^{-6} M) for 72 h, 441 genes were upregulated and 965 were downregulated (fold change >1.5 and adjusted, $P < 0.01$). (**b**) Expression levels of pluripotency genes and several Phf19 target genes were determined by RT-qPCR under the experimental conditions described in **a**. Two different Phf19-depleted cell lines were analyzed (shPhf19–3 and shPhf19–4). Expression levels were normalized to the those of *Rpo* housekeeping gene. Data represent the average of three independent experiments ($n = 3$). (**c**) Control ES cells (shRd) or cells depleted of Phf19 (shPhf19) were injected into Swiss Nude mice at two flanks (1×10^6 cells per flank). Two weeks after injection, mice were killed and tumors were collected. The tumor weights for both sets of teratomas are shown. The differences in size were statistically significant ($n = 8$, Student's *t*-test, $P < 0.05$).

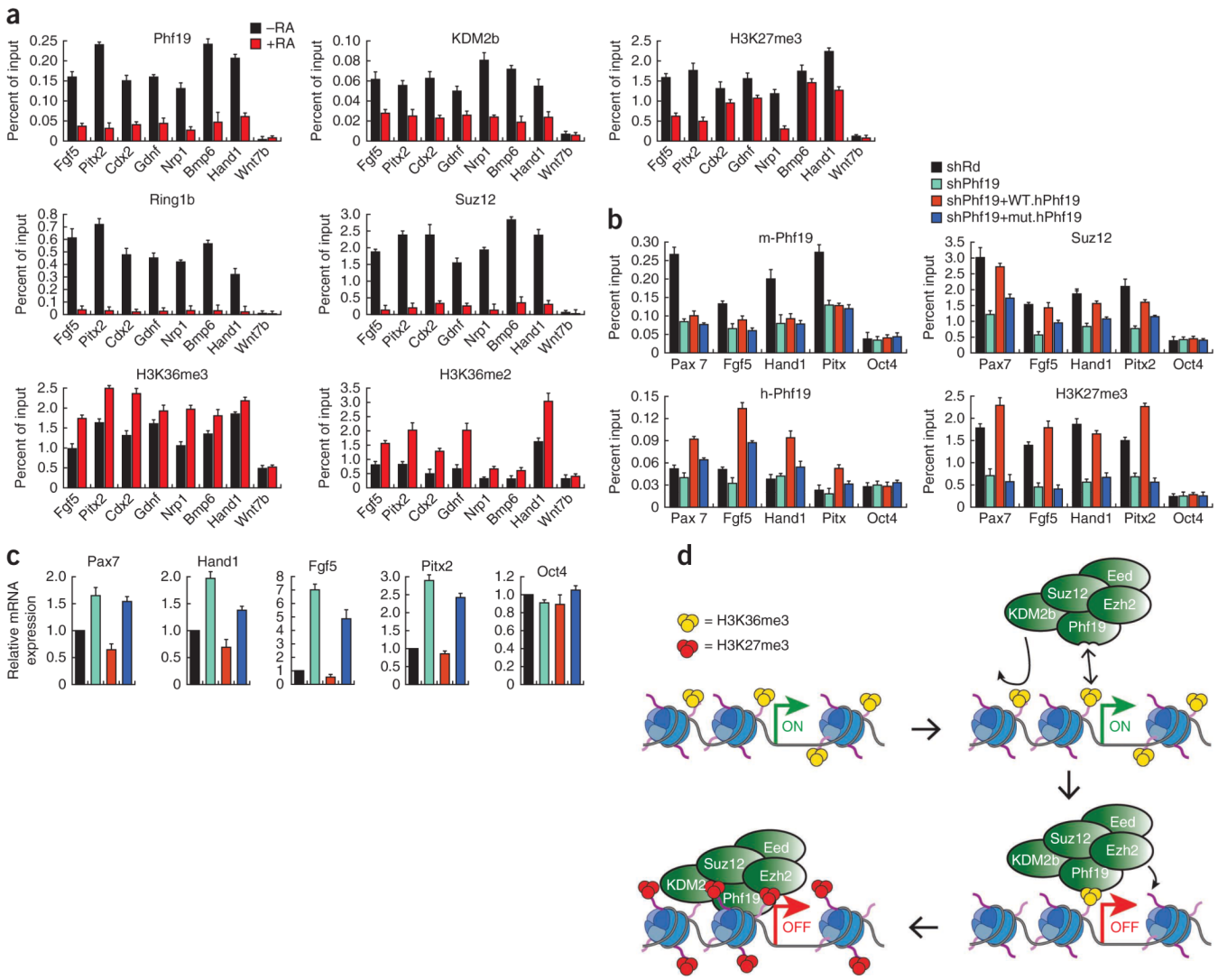

Figure 7.

Molecular mechanism of Phf19-mediated gene repression. (**a**) ChIP-qPCR analyses were performed in mES cells maintained under pluripotency conditions (−RA) or after treatment with all-*trans* RA (+RA), supplemented at 1×10^{-6} M for 5 d. Data represent the average and s.d. of three independent experiments ($n = 3$). (**b**) ChIP-qPCR experiments in control and Phf19 knockdown cells, and in knockdown cells stably transfected with the Flag-tagged human wild-type Phf19 (WT.hPhf19) or with Phf19 mutated (mut.hPhf19) in the Tudor domain (mutation W50A, which renders it unable to bind H3K36me3). Amplification of the *Oct4* promoter was used as a negative control. Data represent the average and s.d. from triplicate experiments. (**c**) Expression levels of Phf19 target genes were determined by RT-qPCR under the experimental conditions described in **b**. Data represent the average and s.d. from triplicates. (**d**) Phf19 facilitates and/or stabilizes binding of the PRC2 complex at H3K36-methylated promoters. Binding of Phf19-Tudor to H3K36me2 and H3K36me3 mediates the recruitment of Polycomb proteins and of the H3K36 demethylase KDM2b, which in turn causes gene silencing.

Table 1

NMR and refinement statistics for H3 K36me3–PHF19-Tudor

	H3 K36me3 bound to PHF19-Tudor
NMR distance and dihedral constraints	
Distance constraints	
Total NOE	1,762
Intra-residue	720
Inter-residue	
Sequential ($\|i - j\| = 1$)	267
Medium-range ($\|i - j\| < 4$)	157
Long-range ($\|i - j\| > 5$)	598
Intermolecular	34
Hydrogen bonds	20
Total dihedral angle restraints	
φ	49
ψ	49
Structure statistics	
Violations (mean ± s.d.)	
Distance constraints (Å)	0.008 ± 0.001
Dihedral angle constraints (°)	0.30 ± 0.06
Max. dihedral angle violation (°)	2.2
Max. distance constraint violation (Å)	0.24
Deviations from idealized geometry	
Bond lengths (Å)	0.0035 ± 0.0001
Bond angles (°)	0.410 ± 0.008
Impropers (°)	0.98 ± 0.08
Average pairwise r.m.s. deviation** (Å)	
Heavy	0.73 ± 0.16
Backbone	1.39 ± 0.24

** Pairwise r.m.s. deviation was calculated for residues 40–90 of PHF19-Tudor and 35–41 of H3K36me3 among 10 refined structures.