

# Nucleosome-Interacting Proteins Regulated by DNA and Histone Methylation

Till Bartke,<sup>1</sup> Michiel Vermeulen,<sup>2,3</sup> Blerta Xhemalce,<sup>1</sup> Samuel C. Robson,<sup>1</sup> Matthias Mann,<sup>2</sup> and Tony Kouzarides<sup>1,\*</sup>

<sup>1</sup>The Gurdon Institute and Department of Pathology, Tennis Court Road, Cambridge CB2 1QN, UK

<sup>2</sup>Department of Proteomics and Signal Transduction, Max-Planck-Institute for Biochemistry, D-82152 Martinsried, Germany

<sup>3</sup>Present address: Department of Physiological Chemistry and Cancer Genomics Centre, University Medical Center Utrecht, 3584 CX Utrecht, The Netherlands

\*Correspondence: [t.kouzarides@gurdon.cam.ac.uk](mailto:t.kouzarides@gurdon.cam.ac.uk)

DOI 10.1016/j.cell.2010.10.012

## SUMMARY

Modifications on histones or on DNA recruit proteins that regulate chromatin function. Here, we use nucleosomes methylated on DNA and on histone H3 in an affinity assay, in conjunction with a SILAC-based proteomic analysis, to identify “crosstalk” between these two distinct classes of modification. Our analysis reveals proteins whose binding to nucleosomes is regulated by methylation of CpGs, H3K4, H3K9, and H3K27 or a combination thereof. We identify the origin recognition complex (ORC), including LRWD1 as a subunit, to be a methylation-sensitive nucleosome interactor that is recruited cooperatively by DNA and histone methylation. Other interactors, such as the lysine demethylase Fbx11/KDM2A, recognize nucleosomes methylated on histones, but their recruitment is disrupted by DNA methylation. These data establish SILAC nucleosome affinity purifications (SNAP) as a tool for studying the dynamics between different chromatin modifications and provide a modification binding “profile” for proteins regulated by DNA and histone methylation.

## INTRODUCTION

Most of the genetic information of eukaryotic cells is stored in the nucleus in the form of a nucleoprotein complex termed chromatin. The basic unit of chromatin is the nucleosome, which consists of 147 bp of DNA wrapped around an octamer made up of two copies each of the core histones H2A, H2B, H3, and H4 (Luger et al., 1997). Nucleosomes are arranged into higher-order structures by additional proteins, including the linker histone H1, to form chromatin. Because chromatin serves as the primary substrate for all DNA-related processes in the nucleus, its structure and activity must be tightly controlled.

Two key mechanisms known to regulate the functional state of chromatin in higher eukaryotes are the C<sub>5</sub> methylation of

DNA at cytosines within CpG dinucleotides and the posttranslational modification of amino acids of histone proteins. Whereas DNA methylation is usually linked to silent chromatin and is present in most regions of the genome (Bernstein et al., 2007), the repertoire and the location of histone modifications are much more diverse, with different modifications associated with different biological functions (Kouzarides, 2007). Most modifications can also be removed from chromatin, thus conferring flexibility in the regulation of its activity. Due to the large number of possible modifications and the enormous diversity that can be generated through combinatorial modifications, epigenetic information can be stored in chromatin modification patterns. Several chromatin-regulating factors have recently been identified that recognize methylated DNA or modified histone proteins. Such effector molecules use a range of different recognition domains such as methyl-CpG-binding domains (MBD), zinc fingers (ZnF), chromo-domains, or plant homeodomains (PHD) in order to establish and orchestrate biological events (Sasai and Defossez, 2009; Taverna et al., 2007). However, most of these studies were conducted using isolated DNA or histone peptides and cannot recapitulate the situation found in chromatin. Considering the three-dimensional organization of chromatin in the nucleus, DNA methylation and histone modifications most likely act in a concerted manner by creating a “modification landscape” that must be interpreted by proteins that are able to recognize large molecular assemblies (Ruthenburg et al., 2007).

In an effort to increase our understanding of how combinatorial modifications on chromatin might modulate its activity, we set out to identify factors that recognize methylated DNA and histones in the context of nucleosomes. We reasoned that using whole nucleosomes would enable us to find factors that integrate the folded nucleosomal structure with modifications on the DNA and on histones. Here, we describe a SILAC nucleosome affinity purification (SNAP) approach for the identification of proteins that are influenced by CpG methylation and histone H3 K4, K9, or K27 methylation (or a combination thereof) in the context of a nucleosome. Our results reveal many proteins and complexes that can read the chromatin modification status. These results establish SNAP as a valuable approach in defining the chromatin “interactome.”

## RESULTS

### The SILAC Nucleosome Affinity Purification

Proteins recognize modifications of chromatin in the context of a nucleosome. However, to date, modification-interacting proteins have been identified using modified DNA or modified histone peptides as affinity columns. We set out to identify proteins that can sense the presence of DNA and histone methylation within the physiological background of a nucleosome. To this end, we reconstituted recombinant nucleosomes containing combinations of CpG-methylated DNA and histone H3 trimethylated at lysine residues 4, 9, and 27 (H3K4me<sub>3</sub>, H3K9me<sub>3</sub>, or H3K27me<sub>3</sub>). These modified nucleosomes were immobilized on beads and used to affinity purify interacting proteins from SILAC-labeled HeLa nuclear extracts (Figure 1A). Bound proteins regulated by the different modification patterns were identified by mass spectrometry (MS).

The methylation of lysines in H3 was accomplished by native chemical ligation (Muir, 2003). An existing protocol (Shogren-Knaak et al., 2003) was adapted to develop an improved method that allows the purification of large quantities of recombinant tail-less human H3.1 (Figure 1B). This method employs the coexpression of tobacco etch virus (TEV) protease and a modified TEV cleavage site (Tolbert and Wong, 2002) to expose a cysteine in front of the histone core sequence in *E. coli* (Figure S1A available online). The tail-less H3.1 starting with a cysteine at position 32 was ligated to thioester peptides spanning the N terminus of histone H3.1 (residues 1–31) and containing the above-mentioned methylated lysines (Figure S1B). The resulting full-length modified H3.1 proteins (Figure S1C) were subsequently refolded into histone octamers together with recombinant human histones H2A, H2B, and H4 (Figure 1C).

As nucleosomal DNAs, we used two biotinylated 185 bp DNA fragments containing either the 601 or the 603 nucleosome positioning sequences (Lowary and Widom, 1998). Both DNAs have similar nucleosome-forming properties, albeit with different sequences (Figure S1D), which allows us to test for sequence specificities of methyl-CpG interactors. The nucleosomal DNAs were treated with recombinant prokaryotic M.SssI DNA methyltransferase, which mimics the methylation pattern found at CpG dinucleotides in eukaryotic genomic DNA (Figures S1E and S1F). Finally, nucleosomal core particles were reconstituted from the nucleosomal DNAs and octamers and were immobilized on streptavidin beads via the biotinylated DNAs. All assembly reactions were quality controlled on native PAGE gels (Figure S1G).

The immobilized modified nucleosomes were incubated in HeLaS3 nuclear extracts and probed for the binding of known modification-interacting factors to make sure that the nucleosomal templates were functional. Figure 1D shows that, as expected, PHF8, HP1 $\alpha$ , and the polycomb repressive complex 2 (PRC2) subunit SUZ12 (Bannister et al., 2001; Hansen et al., 2008; Kleine-Kohlbrecher et al., 2010) specifically bind to H3K4me<sub>3</sub>-, H3K9me<sub>3</sub>-, and H3K27me<sub>3</sub>-modified nucleosomes, respectively. In addition, we did not detect any modification of the immobilized nucleosomal histones by modifying activities present in the nuclear extract (Figure S1H).

In order to identify proteins that bind to chromatin in a modification-dependent manner, we utilized a SILAC pull-down

approach that we have developed to identify interactors of histone modifications (Vermeulen et al., 2010). We simply replaced immobilized peptides with complete reconstituted modified nucleosomes (Figure 2A). All pull-downs were repeated in two experiments. In a “forward” experiment, the unmodified nucleosomes were incubated with light (R<sub>0</sub>K<sub>0</sub>) extracts, and the modified nucleosomes were incubated with heavy-labeled (R<sub>10</sub>K<sub>8</sub>) extracts, as depicted in Figure 2A. In an independent “reverse” experiment, the extracts were exchanged. Bound proteins were identified and quantified by high-resolution MS for both pull-down experiments. A logarithmic (Log<sub>2</sub>) plot of the SILAC ratios heavy/light (ratio H/L) of the forward (x axis) and reverse (y axis) experiments for each identified protein allows the unbiased identification of proteins that specifically bind to the modified or the unmodified nucleosomes. Proteins that preferentially bind to the modified nucleosomes show a high ratio H/L in the forward and a low ratio H/L in the reverse experiment and can, therefore, be identified as outliers in the bottom-right quadrant. Proteins that are excluded by the modification have a low ratio H/L in the forward experiment and a high ratio H/L in the reverse experiment and appear in the top-left quadrant. Background binders have a ratio H/L of around 1:1 and cluster around the intersection of the x and y axes. Outliers in the bottom-left quadrant are contaminating proteins. Outliers in the top-right quadrant are false positives. An enrichment/exclusion ratio of 1.5 in both directions generally identifies outliers outside of the background cluster. We consider a protein to be significantly regulated when it is enriched/excluded at least 2-fold. Higher ratios H/L in the forward and lower ratios H/L in the reverse experiments indicate stronger binding, whereas stronger exclusion is indicated by lower ratios H/L in the forward and higher ratios H/L in the reverse experiments.

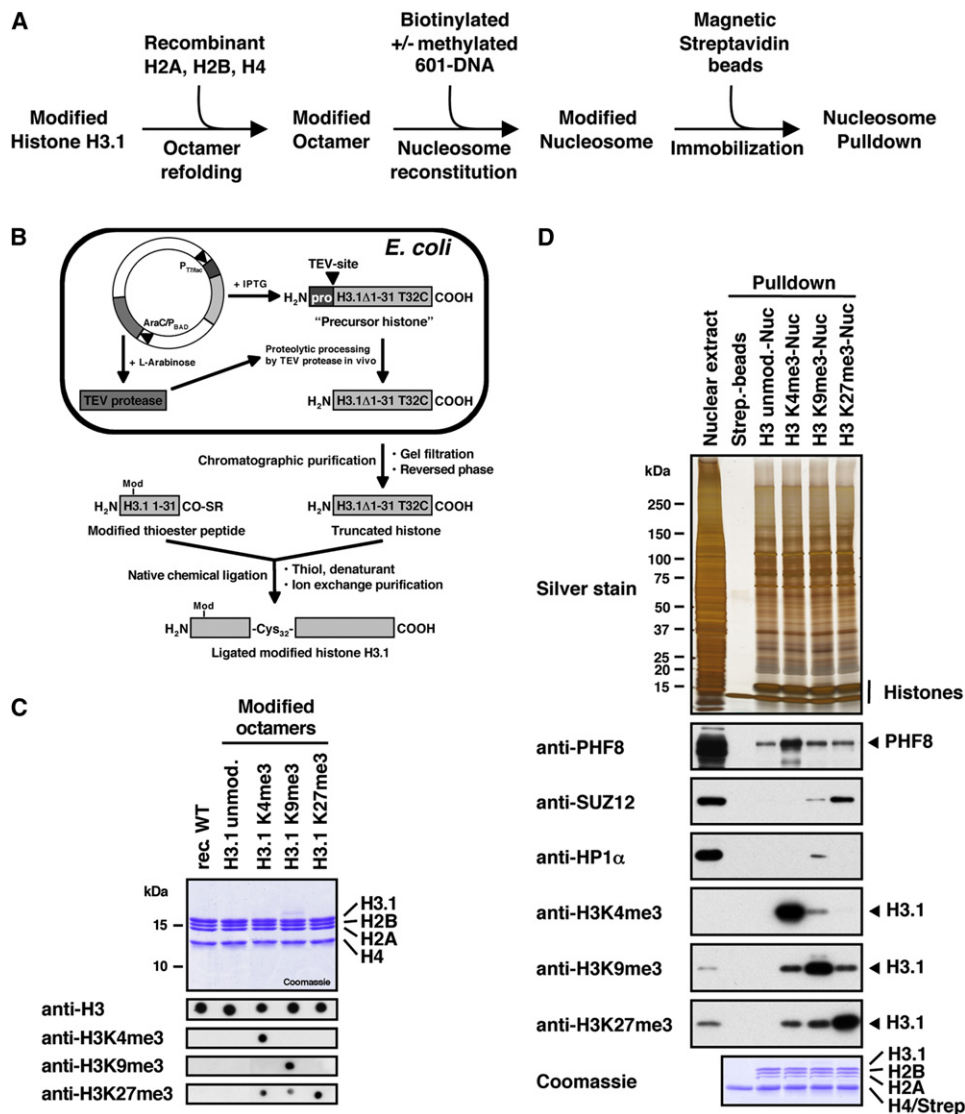
### Proteins Identified by SNAP

The SNAP approach was used to identify proteins that are recruited or excluded by DNA methylation, histone H3 methylation, or a combination of both (Figures 2B and 2C and Figure S2). In Table 1, Table 2, and Table S2, we summarize the proteins that display a regulation of at least 1.5 in both the forward and reverse experiments, thus defining the proteins that are enriched or excluded by the modified nucleosomes. The complete MS analysis defining all interacting proteins in all pull-down reactions is summarized in Table S1.

The data set includes a number of proteins (about 20%) that are already known to bind methyl-DNA and methyl-H3, as well as many proteins whose regulation by modifications had not been previously defined. The presence of many known methyl-binding proteins validates our approach. The database provides a complex “profile” for the modulation of proteins by DNA and histone methylation that have the potential to recognize specific “chromatin landscapes.” Below, we highlight several interactions with modified nucleosomes, which exemplify the different modes of regulation that we observe (summarized in Figures 2D and 2E).

### Regulation by CpG Methylation

Table 1 shows DNA- and nucleosome-binding proteins regulated by CpG methylation. The two different methylated DNAs



**Figure 1. Preparation of Reconstituted Modified Nucleosomes**

(A) Experimental strategy for the preparation of immobilized and modified nucleosomes for pull-down studies.

(B) The native chemical ligation strategy for generating posttranslationally modified histone H3.1. We bacterially express an IPTG-inducible truncated histone precursor containing a modified TEV-cleavage site (ENLYFQ↓C) followed by the core sequence of histone H3.1 starting from glycine 33. The plasmid also contains TEV-protease under the control of the AraC/P<sub>BAD</sub> promoter. TEV-protease accepts a cysteine instead of glycine or serine as the P1' residue of its recognition site, and upon arabinose induction, it processes the precursor histone into the truncated form (H3.1Δ1-31 T32C), which is purified and ligated to modified thioester peptides spanning the N-terminal residues 1 to 31 of histone H3.1. All ligated histones contain the desired modification and a T32C mutation.

(C) Summary of the modified histone octamers. The top panel shows 1 μg of each octamer separated by SDS-PAGE and stained with Coomassie. For the bottom panel, octamers were dot blotted on PVDF membranes and probed with modification-specific antibodies as indicated. The anti-H3K27me3 antibody shows slight cross-reactivity with H3K4me3 and H3K9me3.

(D) Functional test of the nucleosome affinity matrix. R<sub>10</sub>K<sub>8</sub>-labeled nuclear extract was incubated with immobilized modified nucleosomes as indicated. Binding of PHF8, HP1α, and SUZ12 was detected by immunoblot. Equal loading was confirmed by silver and Coomassie staining. Modification of histone H3 was verified by immunoblot against H3 trimethyl lysine marks. All three antibodies show slight cross-reactivity with the other histone marks.

See also Figure S1.

were subjected to SNAP analysis either on their own (601<sup>me</sup> DNA and 603<sup>me</sup> DNA) or assembled into nucleosomes (601<sup>me</sup> Nuc and 603<sup>me</sup> Nuc). We identify several well-characterized methyl-binding proteins such as MBD2 (Sasai and Defossez, 2009) to be enriched on the 601<sup>me</sup> and 603<sup>me</sup> DNAs. MBD2 is enriched

on both DNAs and exemplifies a form of methyl-CpG binding that is not sequence selective. In contrast, other proteins (e.g., ZNF295) display sequence specificity toward only one of the methylated DNAs, suggesting that they may recognize CpG methylation in a sequence-specific manner.

We also identify many proteins that preferentially recognize nonmethylated DNA and are excluded by CpG methylation. The most prominent example is the general RNA polymerase III transcription factor TFIIIC. All subunits of the TFIIIC complex show specific exclusion from the 603<sup>me</sup> DNA (e.g., GTF3C5 shown in Figure 2D), most likely because this DNA (unlike the 601<sup>me</sup> DNA) contains two putative B box elements (Figure S1D), sequences that are known TFIIIC-binding sites. This defines a form of methyl-CpG-dependent exclusion that is sequence specific.

CpG methylation can have a distinct influence on protein binding when it is present within a nucleosomal background. Factors such as MeCP2 are specifically enriched on CpG-methylated DNA only in the context of a nucleosome, but not on free DNA (Figure 2D). Other factors, such as L3MBTL3, show nucleosome-dependent exclusion by CpG methylation. These two factors are influenced by DNA methylation regardless of DNA sequence. Several proteins, such as the DNA-binding factor USF2, are specifically excluded only from 601<sup>me</sup> nucleosomes. This is most likely due to an E box motif in the 601 DNA (Figure S1D), which is recognized by USF2.

One final example of the effect of nucleosomes on DNA-binding proteins is demonstrated by the observation that many proteins such as TFIIIC bind free DNA but cannot recognize the DNA when it is assembled into nucleosomes. This is probably due to binding motifs (such as the B box motif) being occluded by the histone octamer (Figure 2D and Table S2). This type of interaction may identify proteins that need nucleosome-remodeling activities to bind their DNA element. Together, these examples highlight the additional constraints forced on protein-DNA interactions by the histone octamer.

### Regulation by H3 Lysine Methylation

Table 2 shows a summary of the proteins enriched or excluded by nucleosomes trimethylated at H3K4, H3K9, or H3K27 in the presence or absence of DNA methylation. Trimethylation of H3K4 is primarily associated with active promoters, whereas trimethyl H3K9 and H3K27, as well as methyl-CpG, are hallmarks of silenced regions of the genome (Kouzarides, 2007).

We identify several known histone methyl-binding proteins in our screen, such as the H3K4me3-interactor CHD1, the H3K9me3-binder UHRF1, and the H3K27me3-interacting polycomb group protein CBX8 (Hansen et al., 2008; Karagianni et al., 2008; Pray-Grant et al., 2005). In addition, a number of uncharacterized factors were identified. For example, Spindlin1 binds strongly to H3K4me3. Spindlin1 is a highly conserved protein consisting of three Spin/Ssty domains that have recently been shown to fold into Tudor-like domains (Zhao et al., 2007), motifs known to bind methyl lysines on histone proteins. Most notably, we identify the origin recognition complex (Orc2, Orc3, Orc4, Orc5, and to a lesser extent Orc1) to be enriched on both H3K9me3- and H3K27me3-modified nucleosomes. Because no binding was detected on H3K4me3 nucleosomes, the origin recognition complex (ORC) seems to specifically recognize heterochromatic modifications (Figure 2E). One protein, PHF14, and, to a lesser extent, HMG20A and HMG20B are excluded by the H3K4me3 modification. Of interest, these factors represent the only significant examples

of proteins excluded from nucleosomes by methylation of histones, including methylation at H3K9 and H3K27.

### Crosstalk between DNA and Histone Methylation

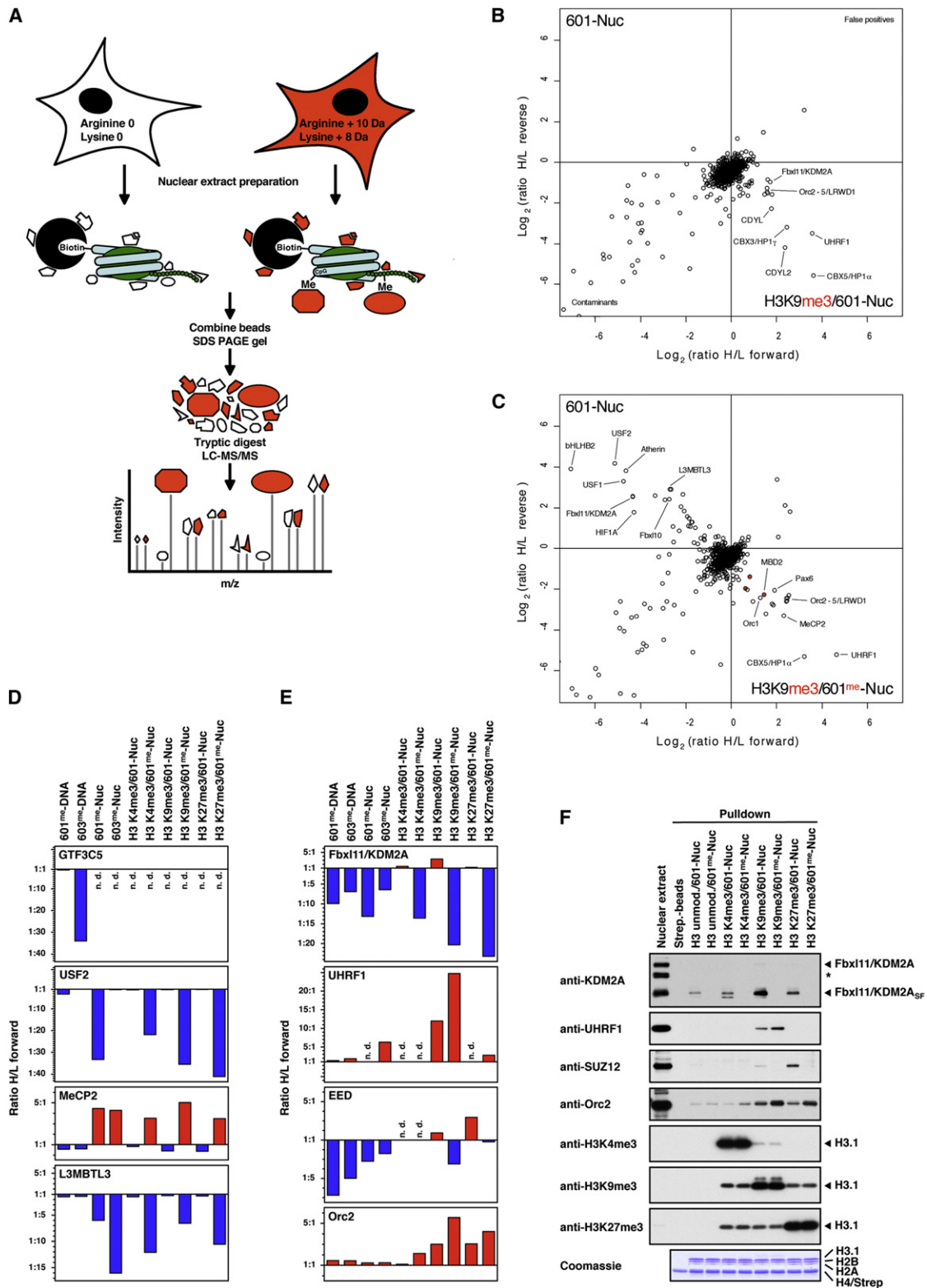
The SNAP approach allows us to investigate cooperative effects between DNA methylation and histone modifications on the recruitment of proteins to chromatin. Analysis of our data reveals several examples of such a regulation (Figures 2E and 2F). We observe a cooperative stronger binding of UHRF1 to H3K9me3-modified nucleosomes in the presence of CpG methylation. Similarly, the ORC (as shown for the Orc2 subunit) can recognize nucleosomes more effectively if CpG methylation coincides with the repressive histone marks H3K9me3 or H3K27me3. This might explain its preferential localization to heterochromatic regions in the nucleus (Pak et al., 1997; Prasanth et al., 2004). In contrast, the H3K36 demethylase Fbx11/KDM2A is enriched by H3K9 methylation but excluded by DNA methylation. Finally, the PRC2 complex is enriched on H3K27me3 nucleosomes (and to a lesser extent on H3K9me3 nucleosomes), but incorporation of methyl-CpG DNA counteracts this recruitment, as shown for the EED (Figure 2E) and the SUZ12 (Figure 2F) subunits. These findings demonstrate the ability of these factors to simultaneously monitor the methylation status of both histones and DNA on a single nucleosome.

### Identification of Complexes Regulated by Chromatin Modifications

The proteins regulated by nucleosome modifications in the SNAP experiments were subjected to a cluster analysis in order to define common features of regulation. In this analysis, the SILAC enrichment values are represented as a heat map in which proteins with similar interaction profiles group into clusters that may be indicative of protein complexes. Figure 3 shows that members of several known complexes cluster together in this analysis, including the BCOR and the NuRD corepressor complexes (Gearhart et al., 2006; Le Guezennec et al., 2006).

### Identification of LRWD1 as an ORC-Interacting Protein

The cluster analysis also identifies the ORC based on the similar interaction profiles of the ORC subunits. Of interest, an uncharacterized protein termed LRWD1 closely associates with the ORC cluster (see also Figures 2B and 2C and Figures S2G and S2H), suggesting that this protein may be a component of ORC. To test this hypothesis, we raised an antibody against LRWD1 (Figure S3A) and used it to probe for colocalization with the ORC by immunofluorescence (IF) staining of MCF7 cells. Figure 4A indicates that LRWD1 colocalizes with the ORC at a subset of nuclear foci marked by strong staining with an antibody against the Orc2 subunit. As previously shown for Orc2 (Prasanth et al., 2004), these foci often colocalize with HP1 $\alpha$ , a marker for H3K9me3-containing heterochromatin (Figure S3B). In addition, endogenous LRWD1 and Orc2 can be coimmunoprecipitated from extracts prepared from MCF7 and HeLaS3 cells (Figure 4B and Figure S3C). We further expressed various truncated variants of FLAG-tagged LRWD1 in 293T cells and immunoprecipitated them using an anti-FLAG antibody. The



**Figure 2. Identification of Nucleosome-Interacting Proteins Regulated by DNA and Histone Methylation Using SNAP**

(A) Experimental design of the SILAC nucleosome affinity purifications. Nuclear extracts are prepared from HeLaS3 cells grown in conventional “light” medium or medium containing stable isotope-labeled “heavy” amino acids. The resulting “light” and “heavy” labeled proteins can be distinguished and quantified by MS.

coimmunoprecipitation of Orc1 and Orc2 indicates that LRWD1 interacts with ORC via its WD40 domain (Figures 4C and 4D and Figure S3D). Similar to Orc3 (Prasanth et al., 2004), expression of LRWD1 depends on Orc2 because reducing Orc2 expression in MCF7 cells by siRNA treatment also reduces LRWD1 protein levels (Figure 4E) without perturbing its transcription (data not shown). These experiments establish LRWD1 as an ORC component and demonstrate the potential of the modification interaction profiling for the identification of protein complex subunits.

### Recognition of Nucleosome Modification Status by Fbx11/KDM2A

To provide independent validation of the SNAP approach, we investigated in greater detail the modulation of binding of Fbx11/KDM2A by DNA and histone methylation. This enzyme is a JmjC domain protein that demethylates lysine 36 on histone H3 (Tsukada et al., 2006). Our data show that KDM2A is enriched on H3K9me3-modified nucleosomes, but its recruitment is disrupted by CpG-methylation on either free or nucleosomal DNA (Figure 2E).

KDM2A has several described isoforms, and in our initial SNAP experiments, some identified KDM2A peptides showed a markedly lower enrichment than others. The H3K9me3-nucleosome SILAC pull-down was repeated to assign the identified peptides to gel bands covering different molecular weights. Most peptides were detected in a band corresponding to a molecular weight of 60–75 kDa and mapped to the C-terminal half of KDM2A (Figures S4A and S4B). Probing for the binding of KDM2A to modified nucleosomes by immunoblot also showed enrichment of a lower molecular weight isoform (Figure 2F and Figure S4C). Immunoprecipitating KDM2A from nuclear extracts confirmed the presence of this isoform (Figure S4D). This variant corresponds to the recently described 70 kDa isoform KDM2A<sub>SF</sub> that is transcribed from an alternative promoter and spans the C-terminal half of KDM2A from position 543 (Tanaka et al., 2010).

We next sought to verify the recruitment of KDM2A to the H3K9me3 modification seen by SNAP in a different biochemical assay. To this end, various methylated and unmethylated nucleosomes or histone H3 peptides were used to isolate FLAG-tagged full-length KDM2A from transfected 293T cell extracts. The SILAC experiments indicated a moderate enrichment of

KDM2A on H3K9me3-nucleosomes (Figure 2E). However, we could not detect substantial binding to either H3K9me3-modified nucleosomes (Figure 5A, lane 5) or peptides (Figure 5A, lane 8) with the overexpressed protein. This result suggested the possibility that KDM2A may need a second factor in order to recognize H3K9me3. A recent study reporting the interaction of KDM2A with all HP1 isoforms (Frescas et al., 2008) prompted us to test whether the binding was mediated by HP1. Indeed, addition of purified HP1 $\alpha$  to the pull-down reactions strongly stimulated the association of KDM2A to H3K9me3 nucleosomes (Figure 5A, lane 13). Using HP1 $\alpha$ , - $\beta$ , and - $\gamma$  showed that the interaction could be mediated by all HP1 isoforms (Figure 5B).

We next verified the disruptive effect of DNA methylation seen in the SNAP experiments. KDM2A harbors a DNA-binding module consisting of a CXXC-type zinc finger domain that was recently demonstrated to bind unmethylated CpG residues and to be sensitive to DNA methylation (Blackledge et al., 2010). When FLAG-tagged KDM2A was isolated from extracts with immobilized 601 DNA (Figure S4E), binding was abolished by CpG methylation as expected. We also sought to establish whether the recruitment of KDM2A to H3K9me3 nucleosomes in the presence of HP1 could be disrupted by DNA methylation. Lane 14 in Figure 5A clearly shows that KDM2A cannot recognize H3K9me3 nucleosomes when the DNA is methylated. The simultaneous recognition of DNA and HP1 leads to a stronger association with nucleosomes. This is indicated by a more effective recruitment of KDM2A to H3K9me3 nucleosomes compared to H3K9me3-modified peptides in the presence of HP1 (compare lanes 13 and 16 in Figure 5A).

To confirm that the recruitment of KDM2A to nucleosomes through HP1 also occurs in a physiological context, we investigated whether the recently reported localization of KDM2A to ribosomal RNA genes (rDNA) in MCF7 cells (Tanaka et al., 2010) is dependent on HP1. Indeed, downregulation of HP1 $\alpha$  by siRNA results in a specific decrease of HP1 $\alpha$  and KDM2A binding, as assessed by chromatin immunoprecipitation (ChIP) analysis (Figures 5C and 5D).

Together, these experiments confirm the observations made using SNAP and show that KDM2A recognizes H3K9me3 via HP1 and that an additional interaction component is conferred by its recognition of DNA, which is sensitive to the state of methylation.

Immobilized unmodified or modified nucleosomes are separately incubated with light or heavy extracts, respectively. Both pull-down reactions are pooled, and eluted proteins are separated by SDS-PAGE. After in-gel trypsin digestion, peptides are analyzed by high-resolution MS.

(B) Results of SNAP performed with H3K9me3-modified nucleosomes containing unmethylated 601 DNA. Shown are the Log<sub>2</sub> values of the SILAC ratios (ratio H/L) of each identified protein for the forward (x axis) and the reverse (y axis) experiments. The identities of several interacting proteins are indicated. Subunits of the MBD2/NuRD complex are labeled in orange.

(C) Results of SNAP performed with H3K9me3-modified nucleosomes containing CpG-methylated 601 DNA. For additional SNAP results, see Figure S2 and Table S1.

(D) Differential recognition of nucleosomes. The graphs show the forward SILAC enrichment values (ratio H/L forward) of MeCP2, L3MBTL3, USF2, and the TFIIC subunit GTF3C5 on CpG-methylated DNAs and modified nucleosomes. Binding to the modified nucleosomes or DNAs is indicated in red; exclusion is indicated in blue. If proteins were not detected (n.d.), no value is assigned.

(E) Crosstalk between DNA and histone methylation. The graphs show the SILAC enrichment values of the proteins KDM2A, UHRF1, the PRC2 subunit EED, and the ORC subunit Orc2 as described in (D).

(F) Immobilized modified nucleosomes were incubated with an independently prepared R<sub>0</sub>K<sub>0</sub> nuclear extract as indicated. Binding of KDM2A, UHRF1, Orc2, and the PRC2 subunit SUZ12 was detected by immunoblot. Equal loading and modification of histone H3 were verified as in Figure 1D. The asterisk marks a cross-reactive band recognized by the KDM2A antibody.

**Table 1. Proteins Enriched or Excluded by CpG-Methylated DNA and Nucleosomes as Identified by SNAP**

Enrichment/Exclusion (Ratio H/L Forward)	601 <sup>me</sup> DNA	603 <sup>me</sup> DNA	601 <sup>me</sup> Nuc	603 <sup>me</sup> Nuc
Enriched Proteins				
very strong enrichment (>10)	ZBTB33	ZBTB33		ZHX2
strong enrichment (5–10)	ZHX1	ZHX1 MBD2 <sup>b</sup> HOMEZ		UHRF1
moderate enrichment (2–5)	ZBTB9 ZHX2 ZHX3 MBD2 <sup>b</sup> MTA2 <sup>b</sup> CDK2AP1 <sup>b</sup> GATAD2A <sup>b</sup> FOXA1 CHD4 <sup>b</sup> ZNF295 MTA3 <sup>b</sup> HOMEZ MTA1 <sup>b</sup> GATAD2B <sup>b</sup> MBD4	ZHX2 MTA2 <sup>b</sup> GATAD2A <sup>b</sup> MTA3 <sup>b</sup> ZHX3 CDK2AP1 <sup>b</sup> FOXA1 CHD4 <sup>b</sup> RFXANK <sup>d</sup> RFXAP <sup>d</sup> MTA1 <sup>b</sup> PBX1 RFX5 <sup>d</sup> PKNOX1 FIZ1 TRIM28 ZBTB40	MeCP2 PAX6 MTERF MBD2 <sup>b</sup> MBD4 ZBTB12 CHD4 <sup>b</sup> MeCP2 GATAD2B <sup>b</sup> ZHX3 ZHX1 C14orf93 RBBP4 <sup>b</sup> RBBP7 <sup>b</sup> MTERF PAX6 LCOR	GATAD2A <sup>b</sup> MTA2 <sup>b</sup> MBD2 <sup>b</sup> MBD4 ZBTB12 CHD4 <sup>b</sup> MeCP2 GATAD2B <sup>b</sup> ZHX3 ZHX1 C14orf93 RBBP4 <sup>b</sup> RBBP7 <sup>b</sup> MTERF PAX6 LCOR
weak enrichment (1.5–2)	PAX9 CHD3 <sup>b</sup> CUX1 ZNF740*	RBBP7 <sup>b</sup> POGZ KIAA1958 UHRF1 ZNF787 MBD4 CHD3 <sup>b</sup> ZFH3 ZBTB9* NR2C1 MAD2B	MTA2 <sup>b</sup> MBD4 CHD4 <sup>b</sup> GATAD2A <sup>b</sup> PIIB JUNB	ACTR5 ZBED5 AURKA HOXC10 JUNB
Excluded Proteins				
weak exclusion (0.5–0.67)		ANKRD32	Atherin* SKP1* <sup>a</sup> RBBP5 NUFIP1 CBFB	MSH3 RBBP5
moderate exclusion (0.2–0.5)	RB1 TFEB SIX4 HES7 ZFP161 YAF2 TIGD5 ARID4B CXXC5 SKP1 <sup>a</sup> JRK USF2 USF1 FBXW11 RAD1 ZBTB2 MLX	SP3 HES7 TCOF1* TFDP1 ATF1 MLL SKP1 <sup>a</sup> RECQL ONECUT2 ZFP161 TIGD1 RB1 E2F3 CUX1 EED <sup>c</sup>	RUNX RNF2 <sup>a</sup> RING1 <sup>a</sup> BANP PRDM11 SUZ12 <sup>c</sup> NAIF1 MYC SUB1 NAIF1 RING1 <sup>a</sup> BANP*	RMI1 TOP3A RPA2 <sup>e</sup> NAIF1 RPA1 <sup>e</sup> RPA3 <sup>e</sup> KIAA1553 TCF7L2 RNF2 <sup>a</sup> BCOR <sup>a</sup> RING1 <sup>a</sup> BANP*

**Table 1. Continued**

Enrichment/Exclusion (Ratio H/L Forward)	601 <sup>me</sup> DNA	603 <sup>me</sup> DNA	601 <sup>me</sup> Nuc	603 <sup>me</sup> Nuc
		BCORL1 ZNF639		
strong exclusion (0.1–0.2)	ZBTB25 PURB RPA1 <sup>e</sup> RPA3* <sup>e</sup> RPA2 <sup>e</sup> MNT UBF1 UBF2 EED <sup>c</sup> SUZ12 <sup>c</sup> VHL E2F4 BCOR <sup>a</sup> FBXL10 <sup>a</sup> FBXL11	SUZ12 <sup>c</sup> RPA3 <sup>e</sup> SSBP1 RPA2 <sup>e</sup> RPA1 <sup>e</sup> CGGBP1 UBF2 FBXL11 PURA UBF1 ZBTB2 ZNF639 RAD1 HUS1 PURB BCORL1 OLA1	MAX L3MBTL3 BCOR <sup>a</sup> FBXL10 <sup>a</sup> PCGF1 <sup>a</sup>	FBXL11 SUB1 FBXL10 <sup>a</sup>
very strong exclusion (<0.1)	E2F1 PCGF1 <sup>a</sup> ZNF395 TIMM8A KIAA1553 bHLHB2 CGGBP1 GMEB2	GTF3C2 <sup>f</sup> BCOR <sup>a</sup> GTF3C4 <sup>f</sup> FBXL10 <sup>a</sup> PCGF1 <sup>a</sup> GTF3C1 <sup>f</sup> E2F1 DEAF1 GTF3C3 <sup>f</sup> GTF3C6 <sup>f</sup> GTF3C5 <sup>f</sup>	HIF1A CXXC5 BCORL1* FBXL11 Syntenin1 ARNT HES7 USF2 bHLHB2 USF1	PCGF1 <sup>a</sup> Atherin L3MBTL3 FLYWCH1 Syntenin1 ZFP161

Table 1 shows the proteins that were enriched or excluded by CpG-methylated DNA or nucleosomes compared to the respective unmodified species at least 1.5-fold in both the forward and reverse pull-down experiments. Proteins are grouped according to their ratio H/L in the forward experiments. Proteins marked by an asterisk are just below the threshold. For the values of the SILAC ratios, see Table S1 and Table S2.

<sup>a</sup>BCOR complex.

<sup>b</sup>NuRD complex.

<sup>c</sup>PRC2 complex.

<sup>d</sup>Regulatory factor X.

<sup>e</sup>Replication factor A complex.

<sup>f</sup>TFIIIC complex.

**DISCUSSION**

Proteins are localized on chromatin depending on a complex set of cues derived from the recognition of histones and DNA in a modified or unmodified form. Here, we present an approach (SNAP) that allows the identification of proteins that recognize distinct chromatin modification patterns. The SNAP method employs modified recombinant nucleosomes to isolate proteins from SILAC-labeled nuclear extracts and to identify them by mass spectrometry. In this study, we have used nucleosomes containing a combination of methylation events on DNA (CpG) and histone H3 (K4, K9, and K27). It is apparent from our results that proteins recognizing methylated nucleosomes can be

influenced by (1) the DNA sequence (in a modified and unmodified form), (2) the configuration of the histone octamer, and (3) the precise combination of histone and DNA modifications. Below, we discuss these modes of engagement.

### (1) Recognition of DNA

The use of two distinct DNA sequences (601 or 603) in our SNAP experiments has identified proteins that recognize methyl-CpGs in a sequence-specific way (e.g., ZNF295) as well as proteins that are not sequence selective (e.g., MBD2). This suggests that some proteins may have a promiscuous methyl-DNA recognition domain (i.e., recognizing methylated CpG dinucleotides regardless of the surrounding DNA sequence), whereas others require a specific motif surrounding the methylated CpG site. Analysis of factors recognizing CpG methylation for the presence of known domains identifies a striking number of zinc finger-containing proteins (Table S2). Our data indicate that around 50% of proteins binding to methyl-CpG and 20% of proteins excluded from methylated DNA and nucleosomes harbor a zinc finger domain, a motif already known to have methyl-CpG binding potential (Sasai and Defossez, 2009). Of interest, the second most prevalent domain in methyl-CpG-binding proteins (20%) is a homeobox (e.g., in HOMEZ, PKNOX1, and ZHX proteins). Homeoboxes are known DNA-binding domains but have not previously been demonstrated to bind methyl-CpG. These data raise the possibility that homeoboxes may possess a methyl-CpG recognition function.

### (2) Influence of Nucleosomes

When methylated 601 or 603 DNA is incorporated into nucleosomes, the histone octamer appears to have an effect on the binding of certain proteins. The TFIIC complex cannot bind a B box effectively in the presence of an octamer, suggesting the need for remodeling activities for full access. The methyl-CpG-binding protein MeCP2 is seen to bind DNA-methylated nucleosomes but showed no binding to methyl-DNA in the absence of a histone octamer. The USF2 transcription factor is excluded from its binding site in the 601 DNA more strongly in the presence of histone octamers. These examples indicate that the histone octamer may have a steric effect on the DNA binding of such factors or that these factors contain additional contact points with histones, which results in an increased affinity to nucleosomes compared to free DNA.

### (3) Regulation by a Combination of DNA and Histone Methylation

Proteins are able to associate with nucleosomes depending on the precise status of DNA and histone methylation. UHRF1, which binds cooperatively to methyl-DNA and H3K9me3, may represent a class of proteins that have an intrinsic capacity to recognize both modifications directly because it contains an SRA domain that binds methylated DNA and a tandem Tudor and a PHD domain that can bind methylated H3K9 (Hashimoto et al., 2009). In the case of protein complexes, the recognition of each modification may reside on separate subunits. We identified two protein complexes, ORC and PRC2, that are influenced by both types of modification in opposite ways. The ORC, including the LRWD1 protein, recognizes H3K9 and

H3K27 methylation in a cooperative manner with DNA methylation. This may allow for a stronger interaction of ORC with heterochromatic regions (Pak et al., 1997; Prasanth et al., 2004). The PRC2 complex, which recognizes H3K27 methylation, is negatively regulated by DNA methylation. This may enable this transcriptional repressor to associate preferentially with a specific chromatin state that is not silenced completely and can respond to external stimuli, such as poised genes. Finally, the KDM2A histone H3K36 demethylase can recognize H3K9me3 indirectly via its association with HP1, and recruitment is blocked when DNA is methylated. This disruptive effect would allow the demethylase to distinguish between distinct chromatin landscapes: it will recognize silenced genes that are marked by H3K9 methylation and HP1, but it will not dock on heterochromatic regions that carry both H3K9me3 and DNA methylation. Together, these examples provide evidence that proteins can monitor the methylation state of both histones and DNA in order to discriminate between distinct states of repressed chromatin.

### SNAP as a Tool for Studying Chromatin Modification Crosstalk

SNAP has several advantages over the current approaches using peptides and oligonucleotides to identify chromatin-binding factors. One advantage is that nucleosomes provide a more physiological substrate. Proteins may have a number of contact points to chromatin (histone tails, histone core, DNA) and may recognize more than one histone at a time. As a result of this multiplicity of possible interactions, SNAP will allow the identification of proteins whose affinity may be too weak to be selected for by the current methods. Our results clearly identify proteins, such as KDM2A, whose binding depends on such a physiological nucleosomal context. A second powerful advantage of SNAP is that it allows the identification of proteins that recognize multiple independent modifications on chromatin. In this study, we have analyzed histone modifications in combination with DNA methylation. But it is equally possible to monitor the binding of proteins to combinations of histone modifications either on the same histone or on different histones or to use multiple nucleosomes. The SNAP approach is also suitable for modified histones generated using methyl-lysine analogs (Simon et al., 2007). But because binding affinities might be crucial for the identification of interacting proteins, natural modified amino acids might be more desirable. In this regard, recent successful attempts to genetically install modified amino acids in recombinant histones are very promising (Neumann et al., 2009; Nguyen et al., 2009). In summary, our findings demonstrate that chromatin modification-binding proteins can recognize distinct modification patterns in a chromatin landscape. The SNAP approach is therefore a valuable tool for studying the mechanisms by which epigenetic information encoded in chromatin modifications can be interpreted by proteins.

### EXPERIMENTAL PROCEDURES

#### Extract Preparation and Immunoprecipitation

HeLa S3 cells were grown in suspension in RPMI 1640 medium containing 5% FBS and normal arginine and lysine or 5% dialyzed FBS and heavy



**Table 2. Nucleosome-Binding Proteins Regulated by CpG and Lysine Methylation as Identified by SNAP**

Enrichment/Exclusion (Ratio H/L Forward)	H3K4me3/601 Nuc	H3K4me3/601 <sup>me</sup> Nuc	H3K9me3/601 Nuc	H3K9me3/601 <sup>me</sup> Nuc	H3K27me3/601 Nuc	H3K27me3/601 <sup>me</sup> Nuc	
Enriched Proteins	very strong enrichment (>10)	Spindlin1	IWS1 <sup>h</sup> Spindlin1	CBX5/HP1 $\alpha$ UHRF1	UHRF1		
	strong enrichment (5–10)	PHF8	PHF8	CBX3/HP1 $\gamma$	CBX5/HP1 $\alpha$		
		CHD1		CDYL2	Orc4 <sup>c</sup> Orc2 <sup>c</sup> Orc3 <sup>c</sup> Orc5 <sup>c</sup> LRWD1 MeCP2		
	moderate enrichment (2–5)	DIDO1	PAX6	LRWD1	PAX6	C17orf96	LRWD1
		UBF1	CHD1	CDYL	CBX3/HP1 $\gamma$	LRWD1	Orc2 <sup>c</sup>
		Sin3A <sup>f</sup>	MeCP2	<i>FBXL11</i>	CDYL	EED <sup>d</sup>	Orc3 <sup>c</sup>
			MTERF	UBF1	MTERF	Orc4 <sup>c</sup>	Orc4 <sup>c</sup>
			MBD2 <sup>b</sup>	Orc2 <sup>c</sup>	MBD2 <sup>b</sup>	Orc5 <sup>c</sup>	Orc5 <sup>c</sup>
			DIDO1	Orc4 <sup>c</sup>	Orc1 <sup>c</sup>	SUZ12 <sup>d</sup>	MeCP2
			Orc2 <sup>c</sup>	Orc5 <sup>c</sup>		Orc2 <sup>c</sup>	CBX8
		Orc4 <sup>c</sup>	Orc3 <sup>c</sup>		Orc3 <sup>c</sup>	UHRF1	
		MBD4			EZH2 <sup>d</sup>	PAX6	
					MTF2 CBX8	MTERF Orc1 <sup>c</sup>	
weak enrichment (1.5–2)	SAP30 <sup>f</sup>	Orc5 <sup>c</sup>	CHD1	MTA2 <sup>b</sup>	PIIB	CDCA7L	
	WDR82	LRWD1	SUZ12 <sup>d</sup>	MBD4		BMI1	
	EMG1	PPIB	EED <sup>d</sup>	ZSCAN21		PPIB	
	TAF9B	ING4	PPIB	CHD4 <sup>b</sup>		MTA2 <sup>b</sup>	
	PPIB	TOX4	NONO	NSD3		MBD4*	
	VRK2	MTA2 <sup>b</sup>	MTF2				
	HNRNPA1*	CHD4 <sup>b</sup>	SUB1				
	HNRNPA2B1*	ZSCAN21					
	ING4	Orc3 <sup>c</sup>					
	WDR61	NONO					
	HNRNPA0*	CDCA7L*					
	FLYWCH1	WDR82*					
	BUB3						
	FUBP3						
Excluded Proteins	weak exclusion (0.5–0.67)		SKP1 <sup>a</sup>	SKP1 <sup>a</sup>		HCFC1	
			RCOR1	CREB1		PHF14 SKP1 <sup>a</sup>	
moderate exclusion (0.2–0.5)	HMG20A HMG20B MTF2*	RING1 <sup>a</sup>	IMP4	RCOR1		SPTH16 <sup>g</sup>	
		SUB1		BANP		SSRP1 <sup>g</sup>	
		HMG20B		RING1 <sup>a</sup>		TCF7L2	
		NAIF		SUB1		BANP*	
		MYC		EED <sup>d</sup>		PRDM11	
				TIGD5		NAIF1	
				RNF2 <sup>a</sup>		RPA1 <sup>e</sup>	
				MYC		BANP*	
				NAIF1		SUB1	
				ARNT			
		TCF7L2 HES7					
strong exclusion (0.1–0.2)	PHF14	FBXL10 <sup>a</sup>		MAX		RPA2 <sup>e</sup>	
		PHF14		CXXC5		BCOR <sup>a</sup>	
		BCOR <sup>a</sup>		L3MBTL3		MYC	
		PCGF1 <sup>a</sup>		FBXL10 <sup>a</sup> BCOR <sup>a</sup>		FBXL10 <sup>a</sup> PCGF1 <sup>a</sup> MAX	
very strong exclusion (<0.1)		L3MBTL3		PCGF1 <sup>a</sup>		L3MBTL3	
		ARNT		HIF1A		HES7	
		<i>FBXL11</i>		Syntenin1		Syntenin1	

**Table 2. Continued**

Enrichment/Exclusion (Ratio H/L Forward)	H3K4me3/601 Nuc	H3K4me3/601 <sup>me</sup> Nuc	H3K9me3/601 Nuc	H3K9me3/601 <sup>me</sup> Nuc	H3K27me3/601 Nuc	H3K27me3/601 <sup>me</sup> Nuc
		Syntenin1		<i>FBXL11</i>		HIF1A
		Atherin		Atherin		Atherin
		USF2		USF1		ARNT
		USF1		USF2		<i>FBXL11</i>
		HIF1A*		bHLHB2		USF1
		bHLHB2				USF2
						bHLHB2

Table 2 shows the proteins that were enriched or excluded by modified nucleosomes compared to unmodified nucleosomes at least 1.5-fold in both the forward and reverse pull-down experiments. Proteins are grouped according to their ratio H/L in the forward experiments. Proteins marked by an asterisk are just below the threshold. For the values of the SILAC ratios, see Table S1 and Table S2. Fbxl11/KDM2A is italicized.

<sup>a</sup>BCOR complex.

<sup>b</sup>NuRD complex.

<sup>c</sup>ORC complex.

<sup>d</sup>PRC2 complex.

<sup>e</sup>Replication factor A complex.

<sup>f</sup>Sin3A complex.

<sup>g</sup>FACT.

<sup>h</sup>IWS should be treated with caution because it was found as a false positive outlier in the 601me-Nuc pull-down.

arginine-<sup>13</sup>C<sub>6</sub>, <sup>15</sup>N<sub>4</sub> and lysine-<sup>13</sup>C<sub>6</sub>, <sup>15</sup>N<sub>2</sub> (Isotec). Cells were harvested at a density of 0.5–0.8 × 10<sup>6</sup> cells/ml, and nuclear extracts were essentially prepared as described (Dignam et al., 1983). For both SILAC extracts, three independent nuclear extracts were prepared and pooled to yield an “average” extract that compensates for differences in each individual preparation. 293T and MCF7 cells were grown in DMEM medium supplemented with 10% FBS. 293T cells were transfected using a calcium phosphate protocol. Whole-cell extracts were prepared ~36 hr after transfection by rotating the cells in extraction buffer (20 mM HEPES [pH 7.5], 300 mM NaCl, 1 mM EDTA, 20% Glycerol, 0.5% NP40, 1 mM DTT, and complete protease inhibitors [Roche]) for 1 hr at 4°C. HeLa S3 nuclear extracts and 293T or MCF7 whole-cell extracts were snap frozen and stored in aliquots at –80°C. For coimmunoprecipitations, extracts were prepared without DTT and diluted 1:1 with 20 mM HEPES (pH 7.5), 1 mM EDTA, and 20% Glycerol containing complete protease inhibitors. Extracts were precleared and proteins immunoprecipitated with typically 5 μg of antibody and Protein-G Sepharose (GE Healthcare) or 20 μl anti-FLAG M2 agarose (Sigma).

#### Chromatin Immunoprecipitation and Immunofluorescence

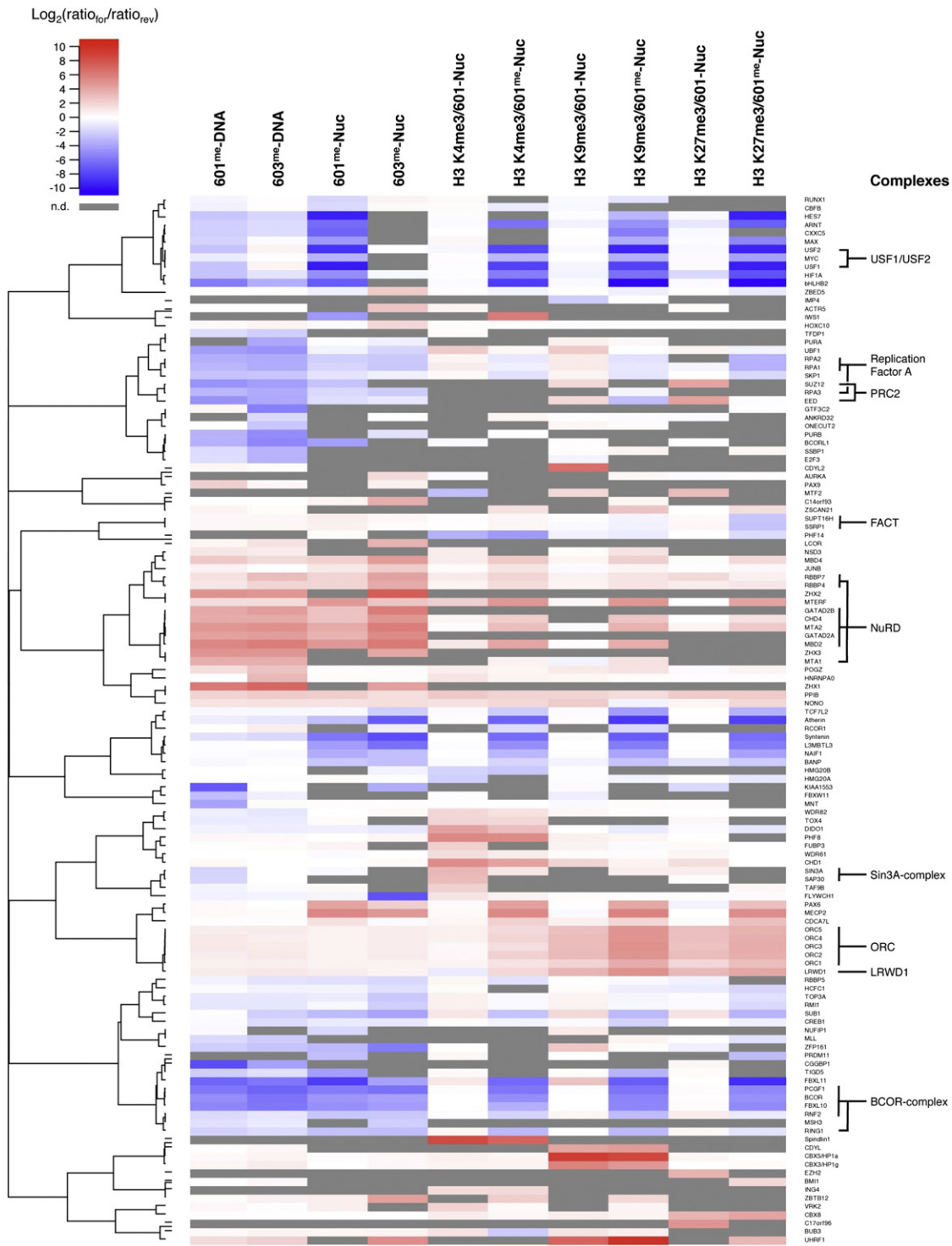
For ChIPs, MCF7 cells were reverse transfected with siRNAs against HP1α or negative control siRNA using Lipofectamine RNAiMAX (Invitrogen) according to the manufacturer’s protocol. At 48 hr after transfection, cells were washed twice with PBS, fixed with 1% formaldehyde (Sigma) in PBS at room temperature for 10 min, and quenched with 125 mM Glycine for 5 min. After three washes with 10 ml of cold PBS, cells were harvested in cold PBS supplemented with complete protease inhibitor cocktail by scraping. Pellets from two 10 cm dishes were suspended in 1.6 ml of RIPA buffer (50 mM Tris-HCl (pH 8), 150 mM NaCl, 2 mM EDTA, 1% NP-40, 0.5% sodium deoxycholate, and 0.1% SDS supplemented with EDTA-free complete protease inhibitors), sonicated in 15 ml conical tubes three times for 10 min at high 30 s on/off cycles in a cooled Bioruptor (Diagenode), and cleared by centrifugation for 15 min at 13,000 rpm. ChIPs were then performed as described (Xhemalce and Kouzarides, 2010). The PCR analysis was performed on a StepOnePlus Real-Time PCR System using Fast SYBR Green (Applied Biosystems). For IFs, MCF7 cells were grown in slide flasks, washed with PBS, treated for 5 min on ice with CSK buffer (10 mM PIPES [pH 6.8], 100 mM NaCl, 300 mM sucrose, 3 mM MgCl<sub>2</sub>, 1 mM EGTA, and 0.5% Triton), washed again with PBS, and fixed with 5% Formalin solution (Sigma) in PBS/2% sucrose. The fixed cells were incubated O/N at 4°C with 0.5 μg/ml of each primary antibody and for 1 hr at RT with DAPI and the secondary antibodies. Images were acquired with an Olympus FV1000 Upright confocal microscope and processed using Adobe Photoshop CS software.

#### Protein Expression and Purification

Recombinant histone proteins were expressed in *E. coli* BL21(DE3)/RIL cells from pET21b(+) (Novagen) vectors and purified by denaturing gel filtration and ion exchange chromatography essentially as described (Dyer et al., 2004). Truncated H3.1Δ1-31T32C protein was generated in vivo by expressing a H3.1Δ1-31T32C precursor in the presence of TEV-protease. For this purpose, *E. coli* cells harboring the pET28a(+)-AraC-P<sub>BAD</sub>-His<sub>6</sub>TEV/pro-H3.1Δ1-31T32C plasmid were grown in LB medium containing 0.25% L-arabinose to keep TEV-protease induced. At an OD<sub>600</sub> of 0.6 the expression of pro-hH3.1Δ1-31T32C was induced for 3 hr at 37°C with 50 μM IPTG. TEV-protease processes the precursor histone H3.1 into tail-less H3.1Δ1-31T32C. The insoluble protein was extracted from inclusion bodies with solubilization buffer (20 mM Tris [pH 7.5], 7 M Guanidine HCl, and 100 mM DTT) for 1 hr at RT and passed over a Sephacryl S200 gel filtration column (GE Healthcare) in SAU-200 (20 mM NaAcetate [pH 5.2], 7 M Urea, 200 mM NaCl, and 1 mM EDTA) without any reducing agents. Positive fractions were directly loaded onto a reversed-phase ResourceRPC column (GE Healthcare) and eluted with a gradient of 0%–65% B (A: 0.1% TFA in water, B: 90% Acetonitrile; 0.1% TFA) over 20 column volumes. Fractions containing pure H3.1Δ1-31T32C were pooled and lyophilized. All histone proteins were stored lyophilized at –80°C. Recombinant HP1 GST-fusion proteins were expressed in *E. coli* BL21(DE3)/RIL cells and purified by glutathione Sepharose (GE Healthcare) chromatography. HP1 proteins were cleaved off the beads with biotinylated thrombin (Novagen). After removal of thrombin with streptavidin Sepharose, HP1 proteins were dialyzed into TBS/10% glycerol, snap frozen, and stored at –80°C.

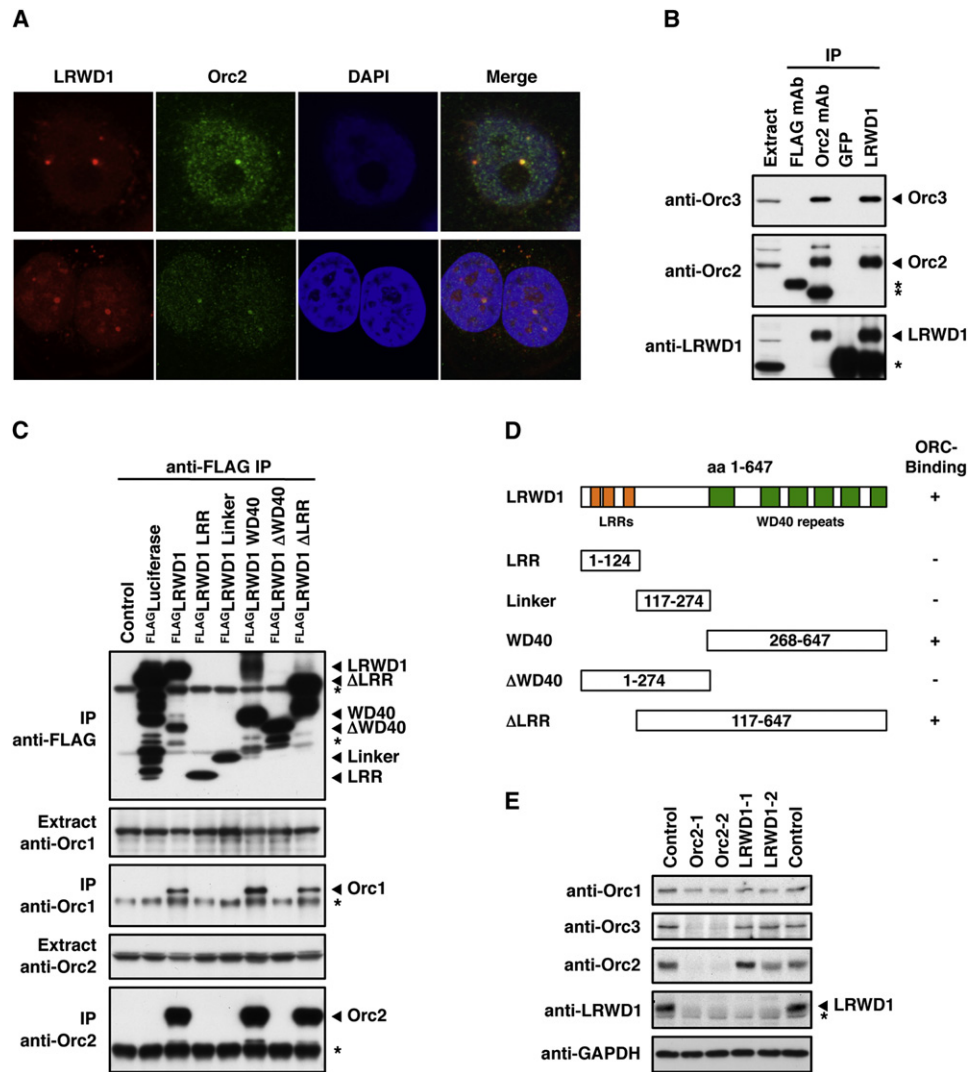
#### Preparation of Modified Histones and Nucleosomal DNAs

For native chemical ligations, lyophilized modified H3.1 1-31 thioester peptide (Almac) was incubated at a concentration of 0.56 mg/ml (~0.167 mM) and truncated H3.1Δ1-31T32C protein at 4 mg/ml (~0.333 mM) and thiophenol at 2% (v/v) in ligation buffer (6 M Guanidine HCl and 200 mM KPO<sub>4</sub> [pH 7.9]). The cloudy mixture was left shaking vigorously at RT for 24 hr. The reaction was stopped by adding DTT to a final concentration of 100 mM, dialyzed three times against SAU-200 buffer containing 5 mM 2-Mercaptoethanol, and then loaded onto a Hi-Trap SP HP column (GE-Healthcare). The ligated Histone H3 was eluted with a linear gradient from SAU-200 to SAU-600 buffer (20 mM NaAcetate [pH 5.2], 7 M Urea, 600 mM NaCl, 1 mM EDTA, and 5 mM 2-Mercaptoethanol). Positive fractions were pooled, diluted 3-fold in SAU-0 buffer (20 mM NaAcetate [pH 5.2], 7 M Urea, 1 mM EDTA, and 5 mM 2-Mercaptoethanol) to reduce the NaCl concentration, and reloaded onto the column. Three rounds of purification were needed to yield sufficiently



**Figure 3. Interaction Profiles of Chromatin Modification-Binding Proteins**

Agglomerative hierarchical clustering was performed on the SILAC enrichment values of proteins regulated by DNA and histone methylation to identify proteins with related binding profiles. This analysis includes proteins based on an enrichment/exclusion of at least 1.5-fold in both directions in one of the nucleosome pull-down experiments and excludes factors that were found solely in the DNA pull-downs.  $\text{Log}_2(\text{ratio}_{\text{for}}/\text{ratio}_{\text{rev}})$  is the  $\text{log}_2$  ratio between the SILAC values (ratio H/L) of the forward and reverse experiments. Enrichment by modifications is indicated in red; exclusion is indicated in blue. Gray bars indicate whether proteins were not detected (n.d.) in particular experiments. These incidences were not included in the cluster analysis. Clusters of several known protein complexes and their respective subunits are indicated on the right. For values, see Table S2.



**Figure 4. LRWD1 Interacts with the Origin Recognition Complex**

(A) LRWD1 colocalizes with Orc2. IF staining of MCF7 cells with LRWD1 (2527) and Orc2 antibodies following pre-extraction shows colocalization at distinct nuclear foci.

(B) LRWD1 and ORC coimmunoprecipitate. LRWD1 and Orc2 were immunoprecipitated from MCF7 whole-cell extracts, and interacting proteins were detected by immunoblot as indicated. LRWD1 was immunoprecipitated using anti-LRWD1 (A301-867A) and detected using anti-LRWD1 (2527) antibodies. Anti-FLAG and anti-GFP antibodies were used as IgG negative controls. Asterisks mark bands derived from antibody heavy chains.

(C) FLAG-tagged full-length and truncated versions of LRWD1 were overexpressed in 293T cells and immunoprecipitated using an anti-FLAG antibody. 1% of the input and 10% of the IP were separated by SDS-PAGE, and Orc1, Orc2, and the FLAG fusions were detected by immunoblot. The asterisks mark bands derived from the anti-FLAG IP antibody.

(D) Identities of the LRWD1 truncation constructs. Only deletions containing the WD40 repeats interact with ORC.

(E) LRWD1 expression is Orc2 dependent. Expression levels of LRWD1 and ORC proteins in MCF7 cells were detected by immunoblot after transfection with siRNAs against LRWD1 and Orc2 as indicated. Cells were reverse transfected twice, 56 hr and 28 hr before harvesting. GAPDH serves as a loading control. The asterisk marks a cross-reactive band detected by the anti-LRWD1 (2527) antibody.

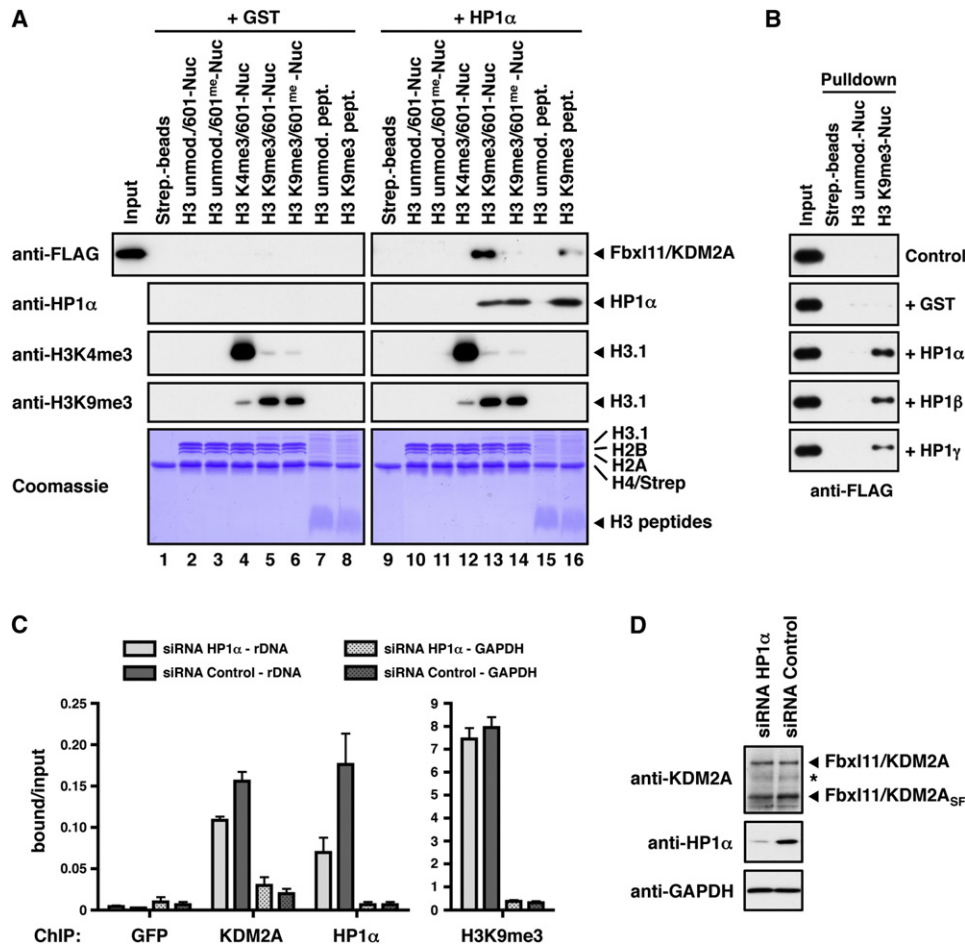
See also Figure S3.

pure ligated histone. Following ion exchange purification, the ligated histone was dialyzed against water containing 1 mM DTT, lyophilized, and stored at  $-80^{\circ}\text{C}$ . Nucleosomal 601 or 603 DNAs were excised from purified plasmid DNAs (Plasmid Giga Kit, QIAGEN) by digestion with EcoRV and separated from the vector by PEG precipitation as described (Dyer et al., 2004). For end biotinylation, the DNA was further digested with EcoRI and the overhangs filled in with biotin-11-dUTP (Yorkshire Bioscience) using Klenow ( $3' \rightarrow 5'$  exo<sup>-</sup>) polymerase (NEB). Nucleosomal biotinylated DNAs were then sepa-

rated by PEG precipitation or further methylated with M.SssI CpG Methyltransferase (NEB) and then PEG precipitated to remove small cleavage products.

#### Reconstitution of Nucleosomes and Nucleosome Pull-Downs

Octamers were refolded from purified histones and assembled into nucleosomes with biotinylated nucleosomal DNAs by salt deposition as described (Dyer et al., 2004). Optimal reconstitution conditions were determined by titration and then kept constant for all nucleosome assembly reactions.



**Figure 5. Fbx11/KDM2A Integrates DNA Methylation and H3K9me3 Modification Signals on Nucleosomes**

(A) In vitro binding of KDM2A to modified nucleosomes. Whole-cell extracts prepared from transiently transfected 293T cells overexpressing FLAG-tagged KDM2A were incubated with immobilized modified nucleosomes or modified H3 peptides as indicated. Binding reactions were supplemented with recombinant purified HP1 $\alpha$  or GST as a control. Binding was detected by immunoblot against the FLAG tag or HP1 $\alpha$ . Equal loading of the nucleosomes and peptides and modification of histone H3 were verified as in Figure 1D.

(B) KDM2A binding to H3K9me3 nucleosomes is mediated by HP1 $\alpha$ , - $\beta$ , and - $\gamma$ . Unmodified or H3K9me3-modified nucleosomes were immobilized on streptavidin beads and incubated with 293T whole-cell extracts overexpressing FLAG-tagged KDM2A. Pull-down reactions were supplemented with recombinant purified HP1 $\alpha$ , - $\beta$ , or - $\gamma$  or GST as indicated. Binding of KDM2A was detected by immunoblot against the FLAG tag.

(C) Recruitment of KDM2A to the rDNA locus is augmented by HP1 $\alpha$ . MCF7 cells were transfected with HP1 $\alpha$ -specific siRNAs and analyzed for the enrichment of the H13 region of the rDNA locus by ChIP using antibodies against KDM2A, HP1 $\alpha$ , and histone H3K9me3. Shown are the mean  $\pm$  SD of the signals normalized to input of three independent experiments. KDM2A shows only little enrichment at the GAPDH locus.

(D) Analysis of KDM2A and HP1 $\alpha$  expression in siRNA-treated MCF7 cells by immunoblot. GAPDH serves as loading control.

See also Figure S4.

Nucleosomes were checked on 5% native PAGE gels. For SILAC pull-downs, nucleosomes corresponding to 12.5  $\mu$ g of octamer were immobilized on 75  $\mu$ l Dynabeads Streptavidin MyOne T1 (Invitrogen) in the final reconstitution buffer (10 mM Tris [pH 7.5], 250 mM KCl, 1 mM EDTA, and 1 mM DTT; supplemented with 0.1% NP40) and then rotated with 0.5 mg HeLa S3 nuclear extract in 1 ml of binding buffer (20 mM HEPES [pH 7.9], 150 mM NaCl, 0.2 mM EDTA, 20% Glycerol, 0.1% NP40, 1 mM DTT, and complete protease inhibitors) for 4 hr at 4 $^{\circ}$ C. After five washes with 1 ml of binding buffer, the beads from both SILAC pull-downs were pooled, and bound proteins were eluted in sample buffer and analyzed on 4%–12% gradient gels by colloidal blue staining (NuPAGE/NOVEX, Invitrogen). For DNA and peptide pull-downs, streptavidin-coated magnetic beads were saturated with either biotinylated 601 DNA or H3 peptides (residues 1–21) and then used as described for the nucleosome beads.

#### Mass Spectrometry of Proteins and Computational Analyses

Nucleosome-bound proteins resolved on SDS-PAGE gels were subjected to in-gel trypsin digestion as described (Vermeulen et al., 2010). Peptide identification experiments were performed using an EASY nLC system (Proxeon) connected online to an LTQ-FT Ultra mass spectrometer (Thermo Fisher, Germany). Tryptic peptide mixtures were loaded onto a 15 cm long 75  $\mu$ m ID column packed in house with 3  $\mu$ m C18-AQUA-Pur Reasil reversed-phase beads (Dr. Maisch GmbH) and eluted using a 2-h linear gradient from 8% to 40% acetonitrile. The separated peptides were electrosprayed directly into the mass spectrometer, which was operated in the data-dependent mode to automatically switch between MS and MS<sup>2</sup>. Intact peptide spectra were acquired with 100,000 resolution in the FT cell while acquiring up to five tandem mass spectra in the LTQ part of the instrument. Proteins were identified and quantified by analyzing the raw data files using the MaxQuant

software, version 1.0.12.5, in combination with the Mascot search engine (Matrix Science), essentially as described (Vicent et al., 2009). The raw data from all forward and reverse pull-downs were processed together and filtered such that a protein was only accepted when it was quantified with at least two peptides, both in the forward and the reverse pull-down. Results from the pull-downs were visualized using the open-source software package R. For the cluster analysis, the log<sub>2</sub> ratio between the forward and reverse SILAC values (ratio H/L) of each protein was calculated. These data were clustered to identify related clades of proteins. Clustering was performed in R using the hopach package (van der Laan and Pollard, 2003). The distance between pairwise log<sub>2</sub> ratio values was calculated using the absolute uncentered correlation distance, and agglomerative hierarchical clustering using complete linkage was performed.

#### Deposition of MS-Related Data

The MS raw data files for nucleosome pull-downs can be accessed via TRANCHE (<https://proteomecommons.org/>) under the name "SILAC Nucleosome Affinity Purification."

#### SUPPLEMENTAL INFORMATION

Supplemental Information includes Extended Experimental Procedures, four figures, and two tables and can be found with this article online at doi:10.1016/j.cell.2010.10.012.

#### ACKNOWLEDGMENTS

We would like to thank Kevin Ford, Timothy Richmond, Bruce Stillman, Jonathan Widom, and Yi Zhang for providing materials; Helder Ferreira and Tom Owen-Hughes for advice on native chemical ligations; and Peter Tessarz and Emmanuelle Viré for experimental help. This work was supported by post-doctoral fellowships to T.B. from EMBO and HFSP and by a fellowship to M.V. from the Dutch Cancer Society. The M.M. laboratory is supported by the Max-Planck Society for the Advancement of Science and HEROIC, a grant from the European Union under the 6<sup>th</sup> Research Framework Programme. The T.K. lab is funded by grants from Cancer Research UK and the European Union (Epitron, HEROIC, and SMARTER). T.K. is a director of Abcam Ltd.

Received: February 10, 2010

Revised: September 28, 2010

Accepted: October 8, 2010

Published: October 28, 2010

#### REFERENCES

- Bannister, A.J., Zegerman, P., Partridge, J.F., Miska, E.A., Thomas, J.O., Allshire, R.C., and Kouzarides, T. (2001). Selective recognition of methylated lysine 9 on histone H3 by the HP1 chromo domain. *Nature* 410, 120–124.
- Bernstein, B.E., Meissner, A., and Lander, E.S. (2007). The mammalian epigenome. *Cell* 128, 669–681.
- Blackledge, N.P., Zhou, J.C., Tolstorukov, M.Y., Farcas, A.M., Park, P.J., and Klose, R.J. (2010). CpG islands recruit a histone H3 lysine 36 demethylase. *Mol. Cell* 38, 179–190.
- Dignam, J.D., Lebovitz, R.M., and Roeder, R.G. (1983). Accurate transcription initiation by RNA polymerase II in a soluble extract from isolated mammalian nuclei. *Nucleic Acids Res.* 11, 1475–1489.
- Dyer, P.N., Edayathumangalam, R.S., White, C.L., Bao, Y., Chakravarthy, S., Muthurajan, U.M., and Luger, K. (2004). Reconstitution of nucleosome core particles from recombinant histones and DNA. *Methods Enzymol.* 375, 23–44.
- Frescas, D., Guardavaccaro, D., Kuchay, S.M., Kato, H., Poleshko, A., Basrur, V., Elenitoba-Johnson, K.S., Katz, R.A., and Pagano, M. (2008). KDM2A represses transcription of centromeric satellite repeats and maintains the heterochromatic state. *Cell Cycle* 7, 3539–3547.
- Gearhart, M.D., Corcoran, C.M., Wamstad, J.A., and Bardwell, V.J. (2006). Polycomb group and SCF ubiquitin ligases are found in a novel BCOR complex that is recruited to BCL6 targets. *Mol. Cell Biol.* 26, 6880–6889.
- Hansen, K.H., Bracken, A.P., Pasini, D., Dietrich, N., Gehani, S.S., Monrad, A., Rappilber, J., Lerdrup, M., and Helin, K. (2008). A model for transmission of the H3K27me3 epigenetic mark. *Nat. Cell Biol.* 10, 1291–1300.
- Hashimoto, H., Horton, J.R., Zhang, X., and Cheng, X. (2009). UHRF1, a modular multi-domain protein, regulates replication-coupled crosstalk between DNA methylation and histone modifications. *Epigenetics* 4, 8–14.
- Karagianni, P., Amazit, L., Qin, J., and Wong, J. (2008). ICBP90, a novel methyl K9 H3 binding protein linking protein ubiquitination with heterochromatin formation. *Mol. Cell Biol.* 28, 705–717.
- Kleine-Kohlbrecher, D., Christensen, J., Vandamme, J., Abarategui, I., Bak, M., Tommerup, N., Shi, X., Gozani, O., Rappilber, J., Salcini, A.E., and Helin, K. (2010). A functional link between the histone demethylase PHF8 and the transcription factor ZNF711 in X-linked mental retardation. *Mol. Cell* 38, 165–178.
- Kouzarides, T. (2007). Chromatin modifications and their function. *Cell* 128, 693–705.
- Le Guezennec, X., Vermeulen, M., Brinkman, A.B., Hoeijmakers, W.A., Cohen, A., Lasonder, E., and Stunnenberg, H.G. (2006). MBD2/NuRD and MBD3/NuRD, two distinct complexes with different biochemical and functional properties. *Mol. Cell Biol.* 26, 843–851.
- Lowary, P.T., and Widom, J. (1998). New DNA sequence rules for high affinity binding to histone octamer and sequence-directed nucleosome positioning. *J. Mol. Biol.* 276, 19–42.
- Luger, K., Mäder, A.W., Richmond, R.K., Sargent, D.F., and Richmond, T.J. (1997). Crystal structure of the nucleosome core particle at 2.8 Å resolution. *Nature* 389, 251–260.
- Muir, T.W. (2003). Semisynthesis of proteins by expressed protein ligation. *Annu. Rev. Biochem.* 72, 249–289.
- Neumann, H., Hancock, S.M., Buning, R., Routh, A., Chapman, L., Somers, J., Owen-Hughes, T., van Noort, J., Rhodes, D., and Chin, J.W. (2009). A method for genetically installing site-specific acetylation in recombinant histones defines the effects of H3 K56 acetylation. *Mol. Cell* 36, 153–163.
- Nguyen, D.P., Garcia Alai, M.M., Kapadnis, P.B., Neumann, H., and Chin, J.W. (2009). Genetically encoding N(epsilon)-methyl-L-lysine in recombinant histones. *J. Am. Chem. Soc.* 131, 14194–14195.
- Pak, D.T., Pflumm, M., Chesnokov, I., Huang, D.W., Kellum, R., Marr, J., Romanowski, P., and Botchan, M.R. (1997). Association of the origin recognition complex with heterochromatin and HP1 in higher eukaryotes. *Cell* 91, 311–323.
- Prasanth, S.G., Prasanth, K.V., Siddiqui, K., Spector, D.L., and Stillman, B. (2004). Human Orc2 localizes to centrosomes, centromeres and heterochromatin during chromosome inheritance. *EMBO J.* 23, 2651–2663.
- Pray-Grant, M.G., Daniel, J.A., Schieltz, D., Yates, J.R., III, and Grant, P.A. (2005). Chd1 chromodomain links histone H3 methylation with SAGA- and SLIK-dependent acetylation. *Nature* 433, 434–438.
- Ruthenburg, A.J., Li, H., Patel, D.J., and Allis, C.D. (2007). Multivalent engagement of chromatin modifications by linked binding modules. *Nat. Rev. Mol. Cell Biol.* 8, 983–994.
- Sasai, N., and Defossez, P.A. (2009). Many paths to one goal? The proteins that recognize methylated DNA in eukaryotes. *Int. J. Dev. Biol.* 53, 323–334.
- Shogren-Knaak, M.A., Fry, C.J., and Peterson, C.L. (2003). A native peptide ligation strategy for deciphering nucleosomal histone modifications. *J. Biol. Chem.* 278, 15744–15748.
- Simon, M.D., Chu, F., Racki, L.R., de la Cruz, C.C., Burlingame, A.L., Panning, B., Narlikar, G.J., and Shokat, K.M. (2007). The site-specific installation of methyl-lysine analogs into recombinant histones. *Cell* 128, 1003–1012.
- Tanaka, Y., Okamoto, K., Teye, K., Umata, T., Yamagiwa, N., Suto, Y., Zhang, Y., and Tsuneoka, M. (2010). JmjC enzyme KDM2A is a regulator of rRNA transcription in response to starvation. *EMBO J.* 29, 1510–1522.

- Taverna, S.D., Li, H., Ruthenburg, A.J., Allis, C.D., and Patel, D.J. (2007). How chromatin-binding modules interpret histone modifications: lessons from professional pocket pickers. *Nat. Struct. Mol. Biol.* *14*, 1025–1040.
- Tolbert, T.J., and Wong, C.H. (2002). New methods for proteomic research: preparation of proteins with N-terminal cysteines for labeling and conjugation. *Angew. Chem. Int. Ed. Engl.* *41*, 2171–2174.
- Tsukada, Y., Fang, J., Erdjument-Bromage, H., Warren, M.E., Borchers, C.H., Tempst, P., and Zhang, Y. (2006). Histone demethylation by a family of JmjC domain-containing proteins. *Nature* *439*, 811–816.
- van der Laan, M.J., and Pollard, K.S. (2003). A new algorithm for hybrid hierarchical clustering with visualization and the bootstrap. *J. Statist. Plann. Inference* *117*, 275–303.
- Vermeulen, M., Eberl, H.C., Matarese, F., Marks, H., Denisov, S., Butter, F., Lee, K.K., Olsen, J.V., Hyman, A.A., Stunnenberg, H.G., and Mann, M. (2010). Quantitative interaction proteomics and genome-wide profiling of epigenetic histone marks and their readers. *Cell* *142*, 967–980.
- Vicent, G.P., Zaurin, R., Nacht, A.S., Li, A., Font-Mateu, J., Le Dily, F., Vermeulen, M., Mann, M., and Beato, M. (2009). Two chromatin remodeling activities cooperate during activation of hormone responsive promoters. *PLoS Genet.* *5*, e1000567.
- Xhemalce, B., and Kouzarides, T. (2010). A chromodomain switch mediated by histone H3 Lys 4 acetylation regulates heterochromatin assembly. *Genes Dev.* *24*, 647–652.
- Zhao, Q., Qin, L., Jiang, F., Wu, B., Yue, W., Xu, F., Rong, Z., Yuan, H., Xie, X., Gao, Y., et al. (2007). Structure of human spindlin1. Tandem tudor-like domains for cell cycle regulation. *J. Biol. Chem.* *282*, 647–656.