Published online 13 June 2017

Nucleic Acids Research, 2017, Vol. 45, No. 13 ***7555–7570***
doi: 10.1093/nar/gkx531

SURVEY AND SUMMARY

Molecular structures guide the engineering of chromatin

Stefan J. Tekel and Karmella A. Haynes*

School of Biological and Health Systems Engineering, Arizona State University, Tempe, AZ 85287, USA

Received March 29, 2017; Revised May 18, 2017; Editorial Decision June 06, 2017; Accepted June 07, 2017

ABSTRACT

Chromatin is a system of proteins, RNA, and DNA that interact with each other to organize and regulate genetic information within eukaryotic nuclei. Chromatin proteins carry out essential functions: packing DNA during cell division, partitioning DNA into sub-regions within the nucleus, and controlling levels of gene expression. There is a growing interest in manipulating chromatin dynamics for applications in medicine and agriculture. Progress in this area requires the identification of design rules for the chromatin system. Here, we focus on the relationship between the physical structure and function of chromatin proteins. We discuss key research that has elucidated the intrinsic properties of chromatin proteins and how this information informs design rules for synthetic systems. Recent work demonstrates that chromatin-derived peptide motifs are portable and in some cases can be customized to alter their function. Finally, we present a workflow for fusion protein design and discuss best practices for engineering chromatin to assist scientists in advancing the field of synthetic epigenetics.

CHROMATIN ENGINEERING: AN IMPORTANT AND CHALLENGING UNDERTAKING

Chromatin is a dynamic nuclear structure that has a central role in eukaryotic development. The mechanics of this ancient, highly conserved system (1,2) are primarily driven by the physical structure and interactions of its components, proteins and nucleic acids. Electrostatic bonds and hydrophobic interactions determine the composition of multipart subunits such as nucleosomes, transcription initiation complexes, and repressive complexes. Because of its impact on tissue development, chromatin has great potential for engineering cell populations. Chromatin proteins exert strong and flexible control over cohorts of genes that determine cell fate and tissue organization. Chromatin states, *i.e.* actively transcribed and silenced, can switch from one to the other. At the same time chromatin-mediated regulation can be very stable, persisting over many cycles of DNA replication and mitosis. The latter property is a mode of epigenetic inheritance, where cellular information that is not encoded in the DNA sequence is passed from mother to daughter cells. The stability of chromatin states allows specific epigenetic programs to scale with tissue development in multicellular organisms.

Early biochemical and protein structure studies of the nucleosome (3) have generated a high resolution model that has persisted over time. A single nucleosome includes a complex of eight histone proteins arranged in a spiral-like disc. Each histone contains a C-terminal globular region composed of helix-turn-helix motifs called the histone fold domain, and an unfolded N-terminal tail (4) (Figure 1A). Within each nucleosome, a tetramer of histones H3 and H4 is stacked between two dimers of histones H2A and H2B (Figure 1B). The stacking model can be viewed as a data-guided 3D animation, created by D. Berry (5). Roughly 200 bp of DNA is wrapped twice around the histone complex. Histones H1/H5 interact with the 'linker' DNA at the entry and exit site of the wrapped DNA (6,7). Nucleosome structures appear repeatedly along each linear chromosome in eukaryotic cells and support higher-order packaging of the entire genome (Figure 1B).

In natural systems, the histone octamer is modular and dynamic. There are four natural variants of H2A and H3 with distinct amino acid sequences (reviewed in 8,9), while histones H4 and H2B are largely invariant (10). Substitutions of H2A and H3 with variants in the octamer complex play critical roles in gene regulation, DNA replication, and chromosome structure (4). Kinetic studies of fluorescently labelled histones have shown that H3 and H4 turnover is very slow. In contrast, exchange of the H2A/H2B dimer occurs within a few minutes to two hours (11). Exchange of histone H1 occurs in under two minutes

*To whom correspondence should be addressed. Tel: +1 480 965 4636; Email: karmella.haynes@asu.edu

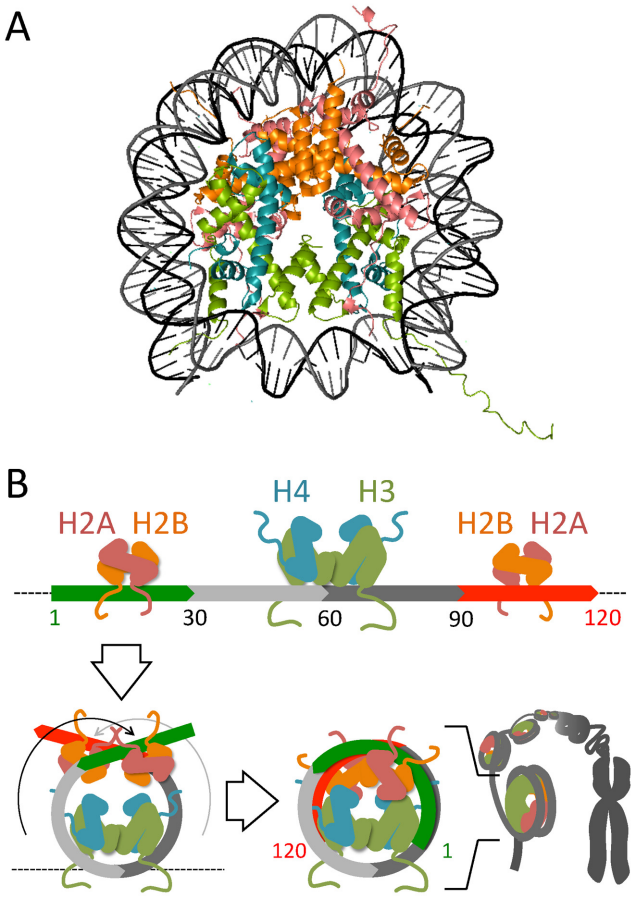

Figure 1. The nucleosome, the core subunit of chromatin, is a modular protein complex. (**A**) 3D model based on K. Luger's 1997 X-ray crystallography data, PDB ID: 1AOI (3). (**B**) The cartoon abstraction shows the general shape of each histone: the Z-shape is the globular region and the thinner line is the unfolded N-terminal tail. A H3/H4 tetramer (green and blue) and two H2A/H2B dimers (orange and pink) bind across ∼120 bp DNA and become stacked to form a solenoid, DNA-wrapped structure (for a data-based animation, see (5)). The lower-right cartoon depicts higher-order packaging of nucleosomes into a metaphase chromosome.

(12). Addition and erasure of post-translational modifications (PTMs) is another highly dynamic feature of nucleosomes. Histone-modifying enzymes covalently link or remove small molecules at the side chains of specific amino acids within each histone. These modifications take place mostly in the unfolded tails, while a few occur in the globular domain. Over 15 known modifications include lysine acetylation (Kac), lysine methylation (Kme), serine phosphorylation (Sp), sumoylation (su), ubiquitination (ub) and crotonylation (cr) (13) (Figure 2). In total, over 50 different amino acid positions are known to be modified. Certain types of PTMs result in transcriptional silencing of a nearby gene, while others enable activation.

Collectively, PTMs make up a rich set of biological information in which single or combinations of molecular tags recruit histone-binding effectors to target genes. By the year 2000, dozens of PTMs had been documented, relationships between histone modifications and gene regulation states began to materialize, and the term 'histone code' was coined by Brian Strahl and David Allis (14–17). Since the first report of the structure of the histone-binding domain from P/CAF in 1999 (18) a plethora of other 3D structures have become available to the scientific community. Investigations of binding pocket specificity support the idea that peptide motifs can distinguish one PTM from another (19–21). These protein structure and interaction data can be used by synthetic biologists to design artificial epigenetic systems and to further confirm or correct aspects of the histone code model. Taverna et al. provide an excellent detailed review of lessons learned from the molecular structures of PTM-binding domains (19).

In spite of its potential usefulness, chromatin is often perceived by biological engineers as an impediment rather than as an enabling tool. Cells are typically engineered by integrating exogenous, recombinant DNA into the chromosomes of the host cell. These transgenes include regulatory components that are carefully designed to operate with predetermined kinetics. However, the transgene often becomes subjected to the surrounding chromatin environment and is mis-regulated (silenced or hyper-activated). The molecular complexity of chromatin may give scientists the impression that chromatin-mediated expression states are impossible to control. Chromatin complexes are often composed of multiple subunits, which have several paralogs in a single organism. For instance, Polycomb Repressive Complex 1 (PRC1) appears as six sub-types that occupy different genomic regions (22). Each of the PRC1 subunits may be one of several distinct paralogs. Furthermore, the core subunit of chromatin known as the nucleosome contains two copies of four types of histones (H2A, H2B, H3, H4) (4), two of which have multiple variants. Histones H3 and H2A have eight and five known variants, respectively. The variants differ in primary sequence, genome distribution, and expression in different tissues and phases of the cell cycle (23,24). Compared to simpler biological principles such as Watson–Crick base-pairing, the complex interactions that govern the behavior of chromatin may seem less amenable to bioengineering.

Is it worthwhile to attempt to engineer multi-layered systems like chromatin within a complex cellular milieu? Synthetic biologists have demonstrated so far that such work produces valuable new knowledge as well as useful innovations (25,26). We believe that the current wealth of information produced by decades of research in chromatin epigenetics provides a sufficient platform to support engineering efforts. In this review, we discuss how proteins and nucleic acids that guide epigenetic regulation in nature have been harnessed for custom-built systems. Specifically, we focus on the molecular structures of chromatin proteins and how our understanding of molecular interactions can be leveraged for chromatin engineering. We discuss best practices for chromatin engineering endeavors and present a flexible, standard workflow for efficient, high-throughput engineering of chromatin-derived proteins.

ENGINEERING NUCLEOSOMES, THE CORE SUBUNITS OF CHROMATIN

Early efforts to engineer nucleosomes used chemical reactions to modify purified histones *in vitro*. Chemical ligation (Figure 2B) has been used to attach synthetic, pre-

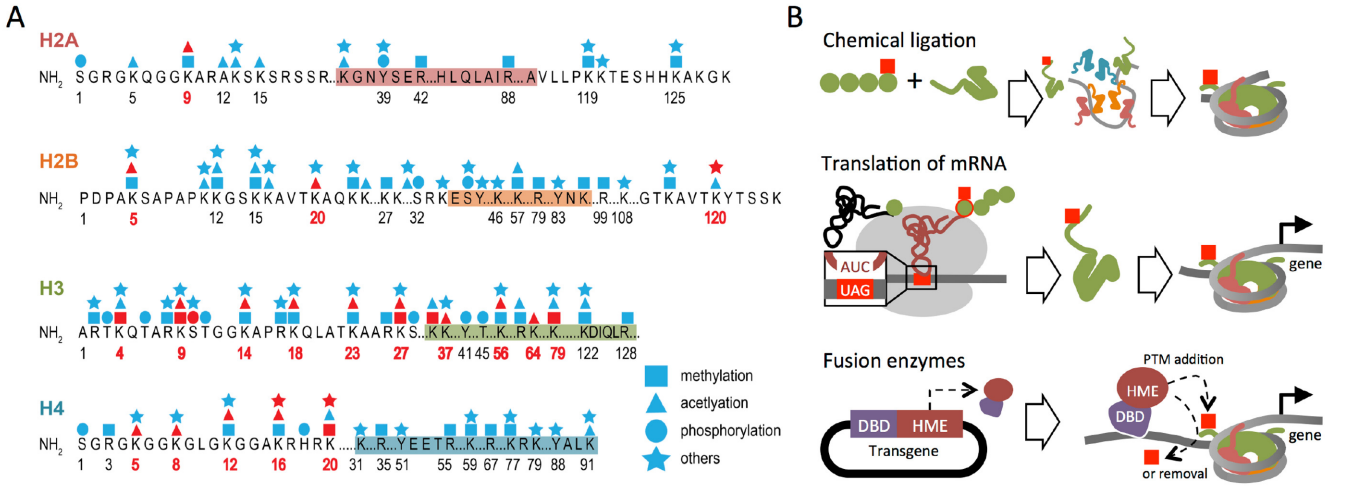

Figure 2. Graphical survey of known histone posttranslational modifications and techniques for generating histone PTMs. (**A**) PTMs that have been artificially generated are shown in red (28–38,40,43–46,49,63,150). Positions of methylation, acetylation, phosphorylation and other PTMs are shown for each canonical histone. Other modifications (stars) include citrullination/deimination, ADP-ribosylation, propionylation, butyrylation, formylation, proline isomerization, hydroxylation, malonylation, glutathionylation, crotonylation and succinylation (166). Globular, folded histone regions are shaded. (**B**) Illustration of three general methods for generating histone PTMs. Chemical ligation adds pre-modified peptides to recombinant histone proteins. The products are assembled with other histones and DNA to generate nucleosomes *in vitro*. Translation of mRNA using engineered tRNA synthetases incorporate modified amino acid residues into genetically encoded amber stop codons (UAG) in the open reading frame of a custom sequence/histone. Fusion enzymes are produced in cells by expressing a DNA-binding domain (DBD) in-frame with a histone-modifying enzyme (HME) or catalytic domain. The fusion protein binds a DNA target and generates or erases PTMs at native nucleosomes.

modified peptides into recombinant histone proteins. A detailed discussion of specific chemical methods is available in (27). Chemical ligation has been used to generate histones bearing acetylation at H3K4, 9, 14, 18, 23, H4K5, 8, 12, 16, tri-methylation at H3K9, and phosphorylation at H3S10 (28–30). Soon after, methylation intermediates (one, two or three methyl groups) were produced on analogs of H3K9, 36, 79 and H4K20 (31). Chemical ligation, and subsequently split inteins,were used to produce ubiquitinated H2B (32,33). Semi-synthetic histones have enabled scientists to build custom nucleosomes and nucleosome arrays to investigate how isolated proteins engage with diverse chromatin templates (27).

Recent advances have allowed scientists to build modified histones via translation of mRNA (Figure 2B). Engineered synthetases have been used to charge amber stop codon-binding tRNAs with pre-modified amino acids. During translation, the modified amino acid is incorporated at any position where a codon has been replaced with UAG in a customized (recoded) mRNA sequence. Acetylation at H2AK9, H2BK5 and K20, H3K56, K23 and K27, and H4K16 were all recombinantly expressed in bacteria with this method (34,35). Kim *et al.* demonstrated successful incorporation of a crotonyl group at position K11 in histone H2B (36). Recently developed synthetases that charge tRNAs with propionyl-lysine, butyryl-lysine (37) and ϵ-*N*-2-hydroxyisobutyryl-lysine (38) further broaden the spectrum of expressed custom histones. In a landmark mammalian cell study, Elsässer *et al.* used a genetically-integrated recoded mRNA translation system to produce pre-modified histones in E14 mouse embryonic stem cells. Six lysine codons in the genes for histones H3.2 and H3.3 were replaced with amber stop codons. Incorporation of acetyllysines at H3.3 positions 9, 23, 27, 37, 56 and 64 resulted in the upregulation of 16 genes, including the locus for the noncoding RNA *Xist*, compared to a wild type H3.3 parental line (39). This technology provides a powerful platform to explore the impact of a variety of histone modifications, natural or novel, in the context of live cells.

Fusion proteins have been used to generate and erase PTMs on endogenous histones, and to enhance or repress transcription at genomic sites in live cells. Histone-modifying enzymes can be fused in-frame with a DNA-binding domain (i.e. dCas9/gRNA, Gal4, LexA, TALE, TetR and Zinc Finger) to generate silencing- or activation-associated PTMs at a gene or non-coding locus (Figure 2B). Fusion enzymes have been used to repress gene expression through methylation of histones: H3K27 via Enhancer of Zeste homologue 2 (EZH2) (40,41), EED (42) and Nuclear Effector (NUE) (43), methylation of H3K9 via Suppressor of Variegation 3–9 homologue 1 (SUV39H1), G9A (44), Kryptonite (KYP) (43), DNA methyltransferase via DMNT3B (42), lysine methyltransferase 1D (EHMT1) (45); and methylation of H4K20 via TgSET8 (43). Gene repression has also been controlled by removal of methyl groups from H3K4 via lysine demethylase 1A (LSD1, KDM1A) (46,47), and removal of acetyl groups from histones via histone deacetylase 4 (HDAC4) (37), HDAC8, RPD3, silent information regulator 2 (Sir2a), and Sin3a (43). Examples of gene-activating PTMs generated by fusion proteins include acetylation of lysines via the catalytic domain of p300 (48,49) and P/CAF (KAT2B) (48). In a creative application of light-regulated peptides, Konermann *et al.* controlled histone PTMs H3K9me1, H4K20me3, H3K27me3, H3K9ac and H4K8ac with a TALE-cryptochrome 2 fusion and its light-induced conju-

gating binding partner CIB1 fused with a variety of catalytic domains (43). Fusion protein-mediated histone modification has been reviewed in detail in other excellent reviews (26,50).

Fusion proteins have demonstrated great utility for testing simple assumptions about the impact of histone PTMs on gene expression. Further progress in the use of histone-modifying fusion proteins to regulate genes requires careful consideration of important contextual aspects: cross-signaling between PTMs and interactions of specific PTM-modifying enzymes with other enzymes. Cross-signaling between PTMs occurs when one modification leads to the generation or erasure of another modification (51). This event is mediated by enzymes that bind pre-existing histone modifications and then catalyze new PTMs (reviewed in 52–54). A classic example is H3K9 methylation by the HP1 SUV39H1 complex. H3K9me3 is recognized by HP1, which recruits SUV39H1, an enzyme that methylates K9 at neighboring H3 histones (55). Other examples of crosstalk occur between distinct types of modifications. Phosphorylation of H3S10 induces Gcn5 to acetylate H3K14 in yeast (56). H2B ubiquitination stimulates methylation of H3K4 and K79 via hDot1L (57). Complexes that contain SGF29, NuA or HBO bind H3K4me3 and acetylate histones H3 and H4 (54). In the case of Rpd3S, cross-signaling promotes the conversion of one gene expression state into another. H3K27me3, which is associated with active transcriptional elongation, is targeted by Rpd3S which removes acetyl groups from histones H3 and H4 to promote silencing (58). Cross-signaling can also be inhibitory, where the presence of one PTM blocks the function of another. The activation-associated mark H3K4me3 blocks interaction of the PRC2 complex with H3, prevents methylation of H3K27, and prevents gene silencing (59). H3S10p disrupts binding of Dido3 with H3K4me3 (60) and enables progression through mitosis. The second contextual aspect we discuss here is the co-recruitment of PTM-modifying enzymes. KRAB is a strong repressor that has been widely used in early and recent work to silence gene targets (e.g. see (41,42,61–64)). The consequences of KRAB fusion-mediated regulation are complicated by interactions with H3K9 methylases (via HP1) as well as histone deacetylases (65). Unintended recruitment of endogenous enzymes can be avoided in some cases by using only the core catalytic domain of the histone modifier (e.g. p300 (48,49)). In summary, the wealth of information from mechanistic studies such as those cited here should be leveraged to design effective chromatin engineering strategies. Pre-existing PTMs at target loci should be determined so that scientists can predict whether fusion enzyme activity will be blocked or enhanced. Recruitment of PTM-modifying binding partners should either be taken into account or avoided in experiments that aim to investigate mechanisms associated with individual PTMs.

ENGINEERING PROTEINS TO RECOGNIZE HISTONE TAGS

PTM-binding proteins can be used to integrate information from histone marks into engineered systems by physically interacting with histone PTMs. In natural systems, PTM-binding proteins act as effectors that regulate gene expression at sites that are enriched for the target histone modification. Here, we highlight four PTM-binding motifs that have been used for chromatin engineering: the chromodomain (CD), bromodomain (BRD), baculovirus inhibitor of apoptosis repeat (BIR) domain, and PHD finger. Several studies have demonstrated that these PTM-binding motifs retain their intrinsic function as isolated peptides, and in some cases after they are incorporated into a fusion protein (Table 1). For a detailed discussion of other well-characterized PTM-binding domains, see Park *et al.* (26).

The chromodomain (CD) motif interacts with methylated histone lysine residues. 3D structure analyses and binding assays have demonstrated that the HP1 chromodomain (66–68) and Polycomb chromodomain paralogs (68–70) show preferential binding to their cognate PTM targets *in vitro*. The CD consists of three β strands packed against a C-terminal α helix, and a hydrophobic pocket that interacts with methyl-lysine (Figure 3A) (66,68,69,71,72). Various CD-containing proteins interact with histone H3 (K4, K9, K27, K36 and K79) and histone H4 (K20) methylated once, twice or three times (me1, me2, me3) (73). Structural studies determined how heterochromatin protein 1 chromodomain (HP1 CD) specifically recognizes trimethylated (me3) histone H3K9. K9me3 is buried in a binding pocket comprised of three aromatic residues, while four amino acids preceding K9 interact with the chromodomain (67). Mutational analysis of a Drosophila HP1 CD showed that the hydrophobic binding cage is necessary for HP1 CD ligand affinity (66). Side chain interactions between the H3 tail and residues from HP1 form a zipper-like β sheet, underscoring the contribution of K9-adjacent histone residues to HP1 CD binding. The Polycomb chromodomain (PCD) is a different type of CD that preferentially binds H3K27me3 and has been shown to cross react with H3K9me3 *in vitro* (71). Although they share high levels of sequence similarity, five mammalian PCDs (CBX2, 4, 6, 7, 8) show significant differences in binding preferences (70). Negatively charged and hydrophobic surfaces distinguish two classes of mammalian PCDs that have high and low affinity, respectively, for their histone ligands (74). Like HP1 CD, the CBX PCDs have a conserved binding motif where a β-strand from the histone tail forms a β sheet with the CD. PCDs share a unique binding site for A25, which fits into a hydrophobic pocket that will not tolerate any other amino acids. In summary, the structural studies of chromodomains suggest two general requirements for CD binding: a hydrophobic pocket and a CD-histone beta sheet (Figure 3). Studies in live cells have demonstrated that CD peptides retain intrinsic PTM-recognition activity within fusion proteins. In Drosophila, fusion proteins containing beta-galactosidase and either the HP1 CD (75) or the Polycomb CD (76) show binding distributions on chromosomes that are similar to the corresponding natural proteins. The Polycomb CD from CBX8 has been used to build a synthetic activator that stimulates gene expression at H3K27me3-enriched genes in human cells (77,78), potentially counteracting oncogenic, repressed chromatin states.

Bromodomains (BRDs) bind acetylated histones tails. These motifs are typically 110 residues in length and appear in histone acetyltransferases and nucleosome remodelling complexes (79–81). Isolated BRD domains from

Table 1. PTM-binding motifs that retain their intrinsic histone specificity in isolation or as a sub-domain within a synthetic or oncogenic fusion protein

Motif	Protein (length, a.a.)	Species	Target PTM(s)	Experiments with isolated peptides *in vitro*	Peptide (a.a.)	Synthetic (S) and oncogenic (O) fusions studied in cells	Peptide (a.a.)
CD	HP1 (206)	*D. mel*	H3K9me2, 3	X-ray crystallography (66,67) FP (68)	17–76	(S) CD-beta-gal (75)	1–95
CD	Pc (390)	*D. mel*	H3K27me3	X-ray crystallography (68,69) FP (68)	15–77, 23–77 1–98	(S) CD-beta-gal (76)	1–117
CD	CBX8 (389)	*H. sap* *M. musc*	H3K27me3 H3K27me3	FP (*M. musc*) (70)	1–62	(S) CD-VP64 (77,78)	1–62
BRD	BRD4 (1362)	*H. sap*	H3Kac, H4Kac	X-ray crystallography, SPOT array (80)	44–168, 333–460	(O) BRD4-NUT (85)	1–719
BRD	BPTF (3046)	*H. sap*	H4K12ac, H4K16ac, H4K20ac	X-ray crystallography, SPOT array, peptide pulldown, FP, ITC (94)	2781–2890	(S) FLAG-PHD-BRD (94)	2722–2890
BIR	BIRC5 (142)	*H. sap*	H3T3p	NMR (89)	1–120	n/r	
PHD	BPTF (3046)	*H. sap*	H3K4me3	X-ray crystallography, SPOT array, peptide pulldown, FP, ITC (94)	2722–2781	(S) FLAG-PHD-BRD (94)	2722–2890
	JARID1A (2221)	*H. sap*	H3K4me3	Peptide pulldown, fluorescence microscopy, Co-IP, ITC (167)	1601–1660	(O) PHD-NUP98 (167)	1489–1690
PHD	Dido3 (2256)	*M. musc*	H3K4me3	NMR, X-Ray crystallography, fluorescence microscopy, peptide pulldown, SPOT array, tryptophan fluorescence (60)	266–325	(S) HA-DIDO (60)	1–528 (Dido1)
	Dido1 (614)						

Representative examples from published reports are included in the table. CD = chromodomain, BRD = bromodomain, FP = fluorescence polarization assay, ITC = isothermal titration calorimetry, n/r = none reported. Lengths of proteins are from Uniprot.

Gcn5p, BRD2 and BRD4 interact with acetylated histone tail peptides in 3D structure studies and in affinity assays (79,80,82,83). Although there is a high degree of sequence variation among the 61 known human BRD proteins, they share a conserved tertiary structure consisting of a left-handed bundle of four α helices linked by variable length loops that surround the histone binding site (Figure 3B). For instance, the BRD region of BRD2 contains a deep hydrophobic pocket contributes to the affinity for acetylated lysines (82,83). Specificity is determined by the primary sequences of the variable loops (80) and by a hydrogen bond between the oxygen of the acetyl carbonyl on the peptide and the amide nitrogen of a conserved asparagine residue (Asn407) (79). In some cases, BRDs recognize PTMs other than acetylation (80). Natural bromodomains from BRD9, CECR2 and TAF1 have been shown to bind to butyryl- and crotonyllysine (84). Flynn *et al.* discovered a gain-of-function mutation in BRD1 that conferred affinity for butyryl, with no loss of the intrinsic acetyl binding (84). Studies of a BRD fusion suggests that BRD can be used to engineer proteins that bind acetyl-histones. BRD domains confer histone-binding activity to the NUT protein in cells where an oncogenic chromosome rearrangement fuses the N-terminal BRD domains of BRD4 to NUT (85). A potential use for BRD domains is a fusion that binds acetyl marks and represses target genes. Aberrant histone acetyltransferase activity has recently become the focus of novel anti-inflammatory therapeutics (reviewed in (86)).

The baculovirus inhibitor of apoptosis repeat (BIR) domain is an approximately 70 amino acid domain containing a zinc coordinated by histidine and cysteine residues, and recognizes phosphorylated histones (87) (reviewed in 88). BIR domain binding has been determined by X-ray crystallography of full-length BIRC5 bound to H3T3p (87,89,90). The BIRC5 (Survivin) protein contains a BIR domain that binds to phosphorylated histone H3 threonine 3 (H3T3p) by surrounding the first four residues of the peptide with 10 negatively charged residues. The residues from BIRC5 that make direct contact with the H3 tail are conserved across vertebrates (89,91). Structural analysis shows that the BIR domain contains an accessible, preformed binding site for H3T3p recognition which suggests a rigid scaffold for binding. This BIR-H3 interaction pattern is highly conserved in other regulatory proteins that contain BIR domains (87,90). Studies revealed a consensus binding motif that includes an N-terminal phosphorylated A-X-S/T-R/K (90) (Figure 3C). Mutational analysis targeting the histone-interacting residues of BIR resulted in reduced binding and reduced centromere localization or poor microtubule depolymerase (MCAK) recruitment in cells (87). Kelley *et al.* showed that BIR interacts with its cognate histone PTM in the absence of other subunits from the chromosomal passenger complex (89). This observation suggests an intrinsic histone-recognition activity that might enable a BIR fusion protein to bind H3T3p. No BIR fusion proteins have been reported, so this is yet to be determined.

Plant homeodomain fingers (PHD) are a class of diverse motifs consisting of two 50–80 amino acid domains that contain a zinc binding site. PHDs generally interact with histone H3 trimethylated at lysine 4 (92,93). Isolated PHD fingers from several proteins, such as Yng1 (94), ING2 (95), TAF3 (96), NURF (97,98), ING3 (99), JARID1A (100) and DIDO (Death Inducer Obliterator) (60) have shown similar interactions with H3K4me3 in different crystal structure studies. The PHD finger of L3MBTL1 recognizes its target via cavity insertion (101). The PHD domain of DIDO (Figure 3D), binds H3K4me3 with a noncanonical aromatic cage that contains a histidine residue. When aligned against other PHD fingers known to bind to H3K4me3, all proteins shared a conserved tryptophan residue (W291 in full length Dido3) in the aromatic cage, as well as other residues known to coordinate zinc ions (60). Using structural information from methyllysine binding, Li *et al.* engineered a plant homeodomain (PHD) finger, which normally binds trimethylated lysines (H3K4me3), to preferentially bind me2 and me1. A single mutation (Y > E) in the aromatic cage responsible for substrate recognition changed the mode of interaction from cavity insertion to surface

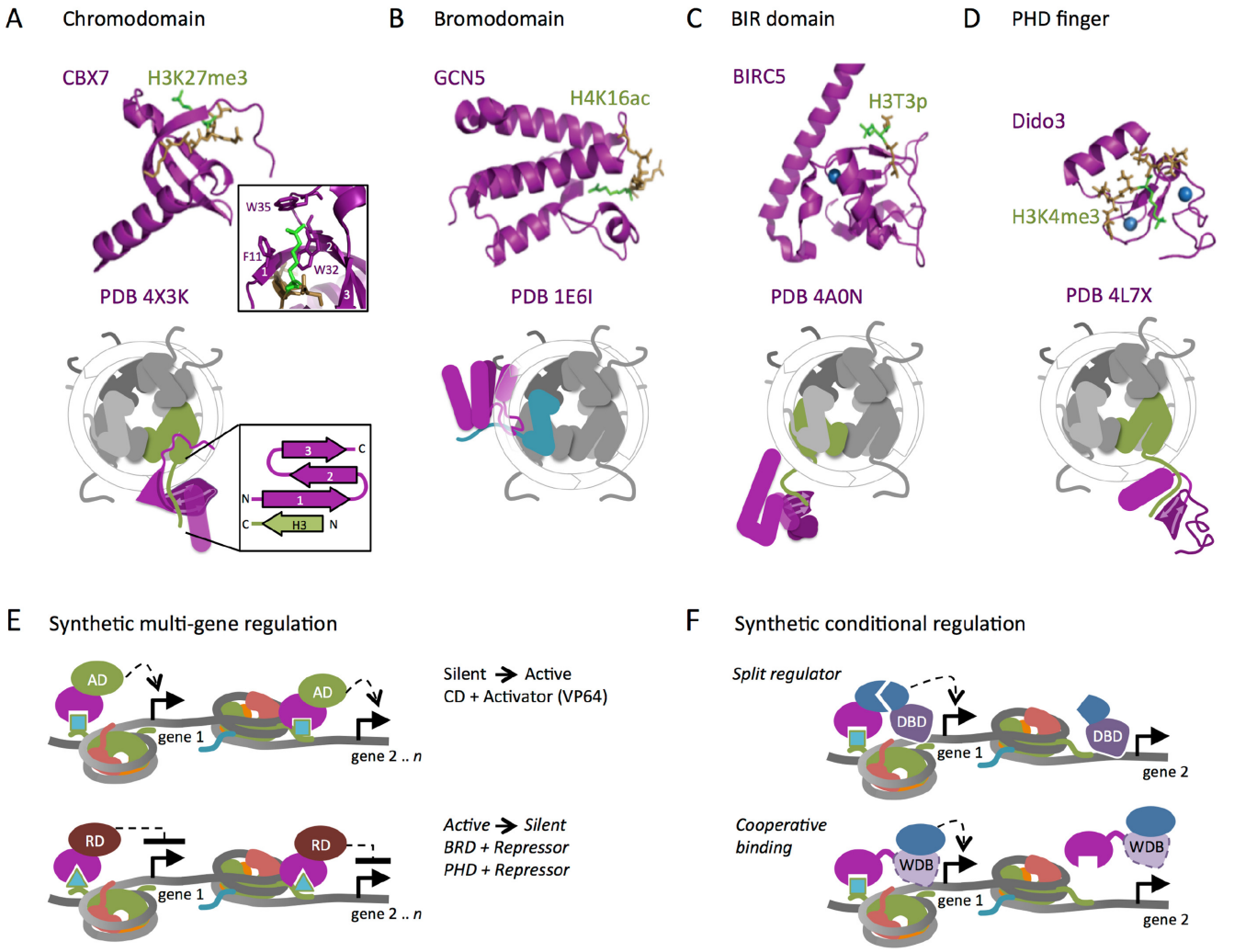

Figure 3. Structural features of representative histone binding domains and their applications in synthetic systems. (**A**) The Chromodomain of CBX7 (116), (**B**) bromodomain of GCN5 (79), (**C**) BIRC5 (Survivin) subunit of the Chromosomal Passenger Complex bound to H3T3p (90) and (**D**) the PHD finger of Dido3 (60). Green = histone sidechain and PTM, blue sphere = coordinating metal ion. Cartoons below each structure depict the binding domain–PTM interaction in the context of a whole nucleosome. (**E**, **F**) Cartoons illustrate uses of histone-binding domains to design synthetic transcriptional regulators. Hypothetical applications have italic captions. (E) Synthetic effectors co-regulate all genes (gene 1, 2, ..., *n*) near a target PTM. This application could be used to convert silencing PTMs into gene activation and activating PTMs into gene repression. The CD from CBX8 has been used to build a synthetic reader that targets the H3K27me3 gene-silencing mark and stimulates gene expression through a C-terminal VP64 domain (77,78). (F) Synthetic regulators could be designed to integrate DNA sequence and PTM information at a specific target gene or allele. (top) Assembly of a split transcriptional activator requires the binding of two fusions. (bottom) Stable binding of weak DNA binding domain (WDB) requires the presence of a specific PTM. AD = transcriptional activator domain, RD = transcriptional repressor domain.

groove recognition. PHD motifs can also recognize acetyllysines, as observed for the protein DPF3b which interacts with H3K14ac (102). PHD motifs retain their intrinsic activity in fusion proteins. The oncogenic fusion of nucleoporin 98 (NUC98) with the PHD domain from JARID1A shows broad colocalization with H3K4me3 and accumulates at the H3K4me3-enriched *HOXA9* gene in human cells (100). In a study using mouse cells, an HA-tagged N-terminal peptide containing the PHD domain from Dido1 (528 aa) was sufficient to rescue the epigenetic activity of the full length Dido3 isoform (2256 aa) (K. van Wely, personal communication) and (60)). A mouse model has been generated to express a DIDO PHD-RFP fusion (60,103), but neither the subcellular localization nor histone interaction of this fusion protein have been reported yet.

Allosteric regulation of PTM-binding proteins by small molecules suggests that chromatin proteins can be designed to switch between binding and nonbinding states at will. Tightly-regulated chromatin systems enable insights into epigenetic dynamics, as seen in other work where a small molecule (doxycycline) was used to control docking of TetR fusions at reporter genes (42,64). Gelato *et al.* described a H3 binding protein, UHRF1, that is allosterically regulated to allow or to block binding to modified histones. UHRF1 has two binding states that target either unmodified H3 tails via a homeodomain or H3K3me3 via a tandem tudor domain (TTD). In the apo state, a polybasic region on the C

terminus sterically occupies a binding groove in the TTD, disrupts interaction with H3K9me3, and allows the homeodomain to bind to unmodified H3 tails. When UHRF1 is bound by phosphatidylinositol phosphate at the polybasic region, a global conformational change is induced, allowing the tudor domain to bind to H3K9me3. This mechanism demonstrates basic allosteric modulation of a histone reader (104). In summary, mechanistic studies of PTM-binding protein folds have provided a useful, relatively untapped resource for fusion protein design. The recognition of specific histone marks could enable broad co-regulation of sets of genes (Figure 3E). Alternatively, DNA and histone recognition could be coupled to conditionally control a single gene target based upon the epigenetic state of the locus or allele (Figure 3F).

MULTIVALENT HISTONE-BINDING DOMAINS INTEGRATE MULTIPLE PTM SIGNALS

Natural multivalent PTM-binding proteins and complexes can interpret multiple epigenetic marks at once and increase binding affinity for a target histone (26). Structures of these multivalent proteins have provided insights into how multivalency can be exploited for synthetic systems. Recognition of combinatorial PTMs could provide greater specificity. For instance, Su *et al.* reported a K_d of 45 nM for the dual spin/Ssty repeat domains of Spindlin1 for histone H3 trimethylated at K4 and dimethylated at R8 (H3K4me3-R8me2) (105). The K_d was lower (139 nM) for an off-target peptide H3K4me3R8me1. Multivalent binding might compensate for the intrinsic low affinity of interactions between single PTM-binding proteins and single targets. In the same investigation, Su *et al.* observed reduced binding affinity (22 μM) of Spindlin for H3R8me2 alone (105). The HP1 protein achieves enhanced affinity through self-dimerization. In a study of HP1 *in vitro*, mutations that disrupted HP1 self-dimerization showed reduced affinity for H3K9me3 (106). The dimerization motif is also required for HP1 activity in cells (107,108).

Research of natural and synthetic multivalent domains suggests principles for designing proteins that recognize more than one PTM. Spacing and orientation of the PTM-binding motifs within the protein appears to be critical for optimal function. The double bromodomain module from $TAF_{II}250$ recognizes H4K5acK12ac (K_d of 1.4 μM). Interestingly, the distance between the tandem bromodomains matches the distance between the acetylated residues (109). In a study of the tandem PHD finger and bromodomain of BPTF, Ruthenburg *et al.* observed that the insertion of two amino acids (QS) into the rigid alpha helical linker region to rotate the domains 200° out of phase disrupted binding to a doubly-modified nucleosome (94). Mutations in the linker that added flexibility also impaired bivalent binding, underscoring the importance of domain orientation in multivalent interactions.

The arrangement of nucleosomes and histone PTMs at genomic target regions influences the function of multivalent PTM-binding proteins. The combined contribution of the PHD finger domain in Rco1 and the chromodomain in Eaf3 direct the histone deacetylase complex Rpd3S to active loci that are enriched H3K36me (58,110). Hu *et al.* proposed that di-nucleosome recognition reinforces Rpd3S binding, thus allowing Rpd3S to tolerate fluctuations in H3K36me levels (111). Optimal binding was achieved *in vitro* when the adjacent nucleosomes were 30–40 bp apart (112). BPTF is another example of a protein that shows a preference for specific PTM placement. In experiments with reconstituted, customized nucleosomes, the divalent PHD-bromodomain from BPTF showed greater interaction with histones when the target PTMs were placed on a single nucleosome, compared to PTMs that were distributed across two nucleosomes (94).

DISRUPTING CHROMATIN: SYNTHETIC ANTAGONISTS OF CHROMATIN COMPLEXES

Chromatin complexes are stabilized by several intramolecular interactions between proteins and histone tails as well as between subunits within the complex. Molecular antagonists have been designed to disrupt chromatin complexes and alter gene expression states by targeting a core subunit that recruits other binding partners to the gene target. Polycomb Repressive Complex 1 types 2 and 4 (PRC1.2, PRC1.4) generally include a histone PTM-binding chromobox (CBX) paralog, RING1A/B, RING1 and YY1-binding protein (RYBP), and polycomb group ring finger 2 (PCGF2, Mel-18) or polycomb group ring finger 4 (PCGF4, BMI1) (22) (Figure 4A). Histone deacetylases (HDACs) support PRC function by removing acetyl groups from H3K27 to allow methylation, which is recognized by CBX. These proteins work in concert to support chromatin compaction and gene silencing, often at tumor suppressors in cancer cells. To perturb CBX7 activity, inhibitors have been designed to bind H3K27me (113) or CBX7 (114–116). Informed by the solved structure of CBX7 in complex with a native peptide or a synthetic inhibitor, Ren *et al.* performed an *in vitro* screen of a library of compounds to identify CBX7 inhibitors. These compounds (see Figure 4B for an example) led to de-repression (transcriptional activation) of CBX7 target genes in human prostate cancer cell lines (72). An *in silico* screen identified a structurally unique compound that demonstrated improved cellular activity. When tested in mouse ES cells, highly specific de-repression of CBX7 target genes was observed, compared to almost no impact on non-target genes (72). Simhadri *et al.* started with trimethyl-lysine peptides and structural data to eventually develop a peptide to directly bind the CBX7 CD and inhibit its function (114) (Figure 4C). In later work, Stuckey *et al.* used a molecular dynamics platform to develop a compound that mimics natural peptide binding to CBX7 (115). Structure-guided work has also led to the discovery of inhibitors of the paralog CBX6 (117). Lastly, molecular structures have enabled the discovery of inhibitors of PTM-modifying enzymes. Whitehead *et al.* identified class 1 HDAC inhibitors that are selective for the HDAC8 isoform (118) (Figure 4D). This work demonstrates that enzyme inhibitors can be highly specific while maintaining biologically-relevant affinities.

BRD2 interacts with acetylated histones, as well as several non-histone proteins including glioma tumor suppressor candidate region gene 1 (GLTSCR1), the histone arginine demethylase JMJD6, and histone methyltransferase

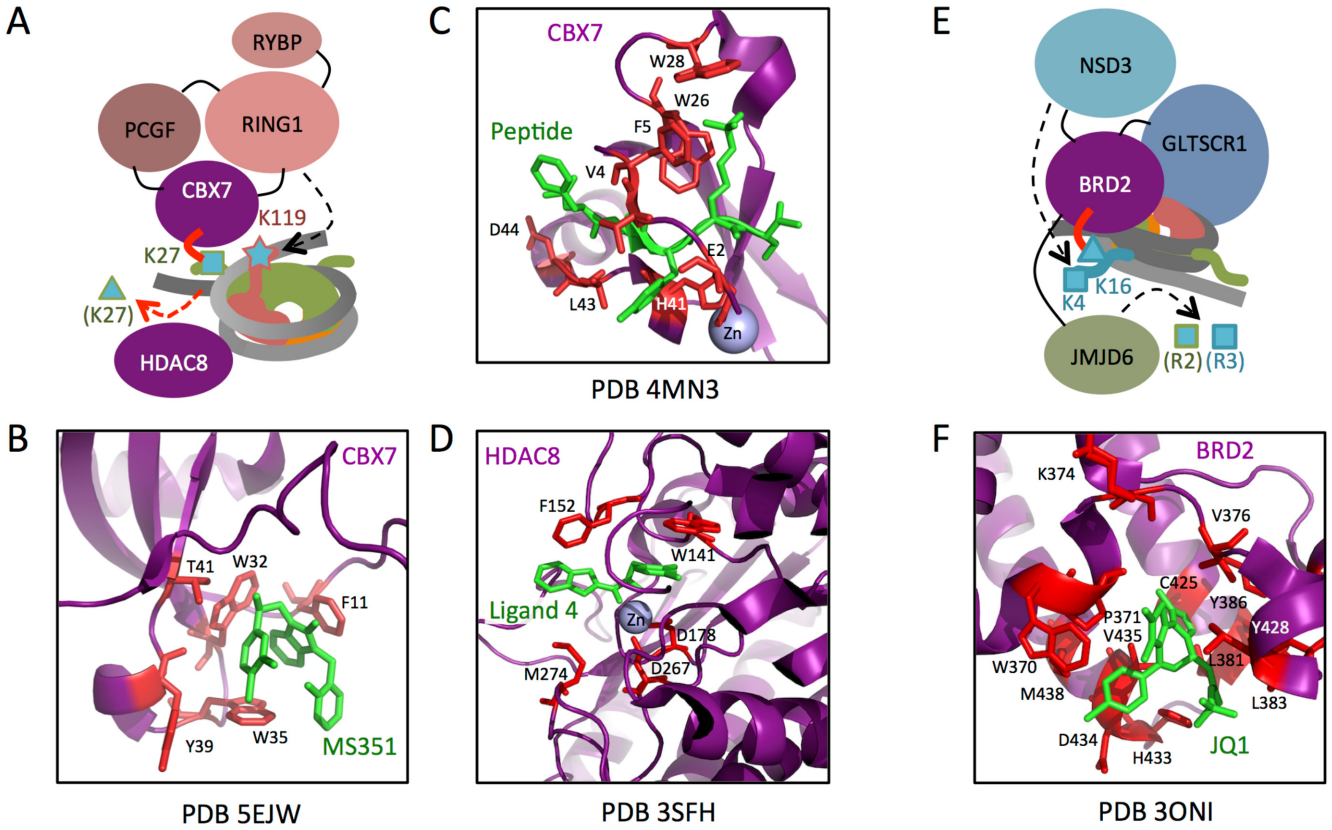

Figure 4. Artificial disruption of multi-protein chromatin complexes. Cartoons depict the context in which inhibitors interact with specific subunits. Shapes are scaled to reflect relative protein sizes (kiloDaltons). Histone PTM symbols are the same as in Figure 2. Arcs = protein interactions (referenced in the text), dashed arrows = histone modifying activity, red lines = inhibitor-mediated disruption. (**A**) The PRC1.2/1.4 type silencing complex can be disrupted by inhibiting the binding of CBX7 CD to H3K27me3 or by inhibiting the removal of acetyl groups from H3K27 by HDAC8. Examples of CBX7 CD inhibitors include MS351 (116) (**B**) and a synthetic peptide (114) (**C**). (**D**) The inhibitor Ligand 4 (118) disrupts HDAC8 activity. (**E**) BRD2 recruits several proteins, including histone modifiers, to stimulate gene activation. Inhibitors such as JQ1 (120) (**F**) have been designed to disrupt the central histone-binding activity. (B–D, F) Boxes show sites within chromatin proteins where low molecular weight inhibitors (green) interact with specific residues (red)

NSD3 to stimulate gene expression (119) (Figure 4E). BRD-containing proteins stimulate MYC oncogene expression and promote the growth of cancer cells (119). Therefore, BRD is an important therapeutic target (81). Bromodomain inhibitors such as JQ1 (120) and I-BET (121) have been identified in high throughput screens to identify small compounds that fit the histone recognition pocket of BRD2 (1-473), BRD3 (1-434) and BRD4 (1-477) (121).

A noteworthy example of structure-guided design come from James *et al.*, who developed strong inhibitors that bind within the methyl-lysine recognition cavity in malignant brain tumor (MBT) repeat domains (122), and others who developed inhibitors of the H3K27 methyltransferase EZH2 (123,124). Horton *et al.* determined the structures of a diverse set of demethylase inhibitors to elucidate the molecular mechanism of binding. Their work provides a general platform for designing epigenetic inhibitors. Structures of the demethylase complexed with various small molecules indicated the potential diversity of inhibitors. The complexes showed a similar binding interfaces but differed in atomic interactions, such as metal coordinating ligands (125).

Since the first reports of the role of Polycomb proteins in oncogenesis and metastasis, a multitude of inhibitors have been developed as epigenetic drugs. Inhibitors of PTM-generating enzymes and their impact on drug discovery is discussed in several excellent reviews (126–131). Structure-focused work has illuminated the molecular interactions between inhibitors and PTM-binding proteins or histone deacetylase enzymes (HDACs). Stuckey *et al.* demonstrated that a PRC1 inhibitor could be designed to inhibit human prostate cancer cell (PC3) proliferation (115). The structure of the inhibitor is similar to the natural ligand H3K27me3 and is highly specific for two chromodomains, CBX4 and CBX7. Inhibitors that bind and interfere with bromodomains (BRDs) have been used to control gene expression in disease-relevant cell culture model systems. Small molecule-mediated inhibition of BRD was used to down-regulate MYC transcription in leukemia and myeloma cells (132,133) and to suppress inflammation in bone marrow-derived macrophages (121). Recently, *in vivo* work has shown that a new BRD-specific inhibitor, MS402, preferentially binds to the first bromodomain motif within BET proteins and is therefore more specific than the broad-acting inhibitor JQ1 (Figure 4F). Cheung *et al.* demonstrated the therapeutic potential of MS402 by preventing and reducing T-cell transfer-induced colitis in mice (134). These studies underscore the clinical potential for synthetic,

epigenetic regulation of genes that have therapeutic properties.

BUILDING CHROMATIN: SCAFFOLDING CHROMATIN COMPLEXES WITH RNA AND DNA

In natural systems, histone-modifying proteins and other chromatin components are spatially arranged along DNA strands. Several studies have demonstrated that specific arrangements of these proteins are non-random, highly regulated, and play a role in epigenetic regulation of gene expression. Chromosome looping, sub-nuclear compartmentalization and gene expression can be altered by manipulating non-coding DNA elements such as the locus control region (LCR), CTCF-bound insulators, and lamina-associated domains (LADs) (reviewed in Park *et al.* (26). Here, we discuss the two classes of epigenetic scaffold elements that have been rigorously analysed within synthetic systems: Polycomb Response Elements and long non-coding RNA.

Polycomb Response Elements (PREs) are *cis*-regulatory DNA sequences that recruit chromatin complexes to clusters of genes. Robust epigenetic control through PRE activity is required to stabilize distinct transcription profiles within subpopulations of cells (135). Drosophila PREs contain clusters of short motifs that interact with several chromatin proteins including Pleiohomeotic (PHO), Zeste protein, GAGA factor (GAF), Dsp1 and others (135). A remarkable characteristic of PREs is 'epigenetic memory', the ability to confer a stable silenced or active state at nearby-genes after removal of the inducer of either state. Early work in Drosophila demonstrated that a PRE called *Fab-7* could be used to switch the expression state at an adjacent reporter gene from active to silenced and vice versa (136–138). These artificially-induced expression states persisted over several rounds of mitosis and meiosis. While Drosophila PREs have been well documented (139), discovery of human PREs has been elusive (140) until recently. In 2010, Woo *et al.* identified a noncoding PRE-like DNA element from the human *HOXD* gene region (141). When this element, D11.12, was placed upstream of a *luciferase* reporter, the transgene became enriched for repressive histone PTMs and silencing-associated proteins and *luciferase* expression was reduced to $<5\%$ of the active state. Since this discovery, other human PREs have been identified (142). Experiments have shown that PRE function can be ported from one metazoan species into another, demonstrating that epigenetic memory is an inherent property of PRE DNA fragments. For instance, mammalian PRE candidates have been validated based on their activity in Drosophila. PREs that were identified in studies of human T cells (143) or the mouse genome (144,145) showed Polycomb protein binding and repression of a reporter gene in transgenic flies.

Investigations of PREs provide some guidance on how these elements can be used to control synthetic genetic systems. In Drosophila, artificially-induced states mediated by PREs can persist over timescales of metazoan tissue development, a characteristic that is critical for the practical use of synthetic gene circuits in multicellular organisms. An important outstanding question for PRE-based genetic engineering is to what extent the behavior of a PRE can be customized to generate distinct gene expression patterns. Mounting evidence from studies in Drosophila show that PREs respond to inducers either early or late in development, and are sensitive or insensitive to their genomic location (reviewed in (140)). Multiple short motifs of varying copy number and order can be found in PREs from Drosophila (146,147) and mammalian genomes (135,141,142) (Figure 5A). Therefore, it is tempting to surmise an underlying protein-scaffolding code. However, current models for PRE function are incomplete. For instance, no human orthologs have been identified for two key Drosophila PRE-recognition proteins, GAF and Zeste (Figure 5A). The Drosophila protein Dsp1 and the human ortholog HMGB2 (hHMG2) have been linked to PRE activity, but there is conflicting evidence for the DNA sequence motif that is recognized by the proteins (135). Finally, although the number of motifs can be identified within a PRE, the stoichiometry of proteins per PRE is not precisely defined. If a deterministic code is eventually identified, artificial PREs could be designed to control the magnitude and dynamics of epigenetic expression in synthetic genetic constructs.

Reminiscent of PREs, synthetic DNA constructs can be designed to include combinations of short DNA motifs that are recognized by transcriptional regulators. Artificial operators have been constructed *de novo* from protein-binding motifs found in well-studied bacterial operons (e.g. Tet, lac, trp) (148) as well as eukaryotic loci (e.g. yeast Gal4 UAS, Zinc finger recognition sites) (149). These motifs can be arrayed within a synthetic DNA fragment and placed upstream of a target gene. Chromatin fusion proteins that recognize the motifs assemble at the scaffold based on the number and arrangement of motifs. Keung *et al.* used two orthogonal Zinc finger (ZF) adapters to co-recruit pairs of gene-regulating proteins to a single promoter (150) (Figure 5A). Although this study was limited to observing synergy or antagonism between single chromatin proteins and the ZF-VP16 activator, the synthetic ZF platform has potential for more general use, i.e. to control spatial arrangements of pairs or larger combinations of chromatin proteins.

Long noncoding RNAs (lncRNAs) also act as modular scaffolds for multi-protein complexes. The underlying design principles of these elements are just beginning to emerge. Engreitz *et al.* provide a superb review of lncRNA structure and function in the context of chromatin and epigenetics (151). Here, we highlight examples of well-defined modular lncRNAs may eventually inform synthetic RNA scaffold design. The lncRNA known as HOTAIR (HOX transcript antisense RNA) has two modules with secondary structures that each interact with distinct histone-modifying complexes to maintain transcriptional repression (152). The 5' domain of HOTAIR binds the PRC2 complex that generates H3K27me3, while the 3' domain binds the LSD1/CoREST/REST complex that stimulates H3K4 demethylation (153). The 5' domain is a distinct module with intrinsic PRC2 binding activity. Structural dissection of HOTAIR revealed that nucleotides 1–300 are sufficient for interaction with PRC2 (153) and that a core motif of 89 nucleotides is required for binding (154). However, recent work has shown that the role of PRC2 is dispensable for HOTAIR-mediated silencing (155). The lncRNA Firre (functional intergenic repeating RNA element) con-

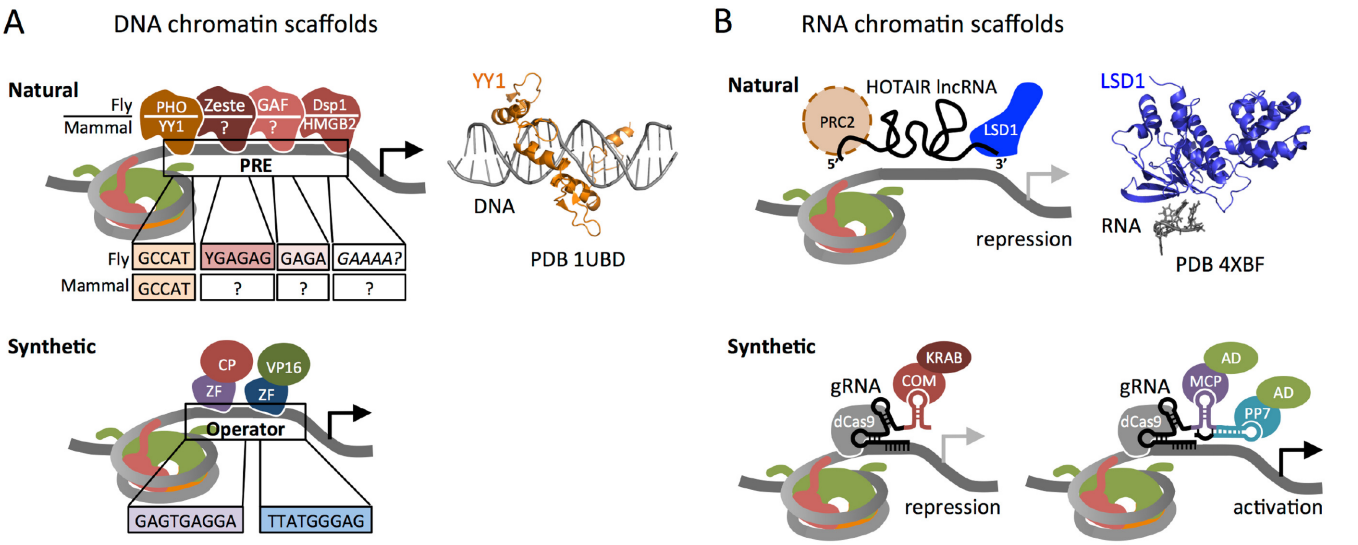

Figure 5. Scaffolding of chromatin proteins on DNA or RNA. (**A**) Natural Polycomb response elements (top) include combinations of short motifs that are recognized by non-histone chromatin proteins. Known PRE-binding proteins and DNA core motifs are shown for Drosophila and mammals. Crystal structure data (168) (top right) shows that YY1 (human PHO homologue) interacts with DNA via a zinc finger motif. A synthetic operator (bottom) designed by Keung *et al.* is shown as an example of a synthetic DNA chromatin scaffold. Nucleotides are named according to the IUPAC code. CP = chromatin protein, ZF = zinc finger. (**B**) HOTAIR (top) is an example of a natural RNA chromatin scaffold. Its 3' and 5' regions act as modules for PRC2 and LSD1/CoREST/REST recruitment. The role of PRC2 (dashed circle) may be dispensable for HOTAIR function (154). LSD1 residues 171–317, 571–654 and 769–836 and RNA (UUAGG) are shown from the crystal structure model (169). Two CRISPR/dCas9-based RNA scaffolds (bottom) are shown as examples of synthetic RNA scaffold systems (158). AD = activation domain.

tains twelve tandem repeats of an RNA motif that binds the nuclear matrix protein HNRNPU, also known as SAFA (scaffold attachment factor A) (156). The structure of Firre demonstrates that tandem replicate RNA modules can operate as independent units, allowing several proteins to co-occupy a single RNA scaffold. Xist is an example of a lncRNA that interacts with a with a DNA-binding protein as well as other chromatin proteins. SAFA acts as a DNA-binding module that tethers the Xist RNA–protein complex to DNA sequences (Figure 5B). The RNA-binding arginine–glycine–glycine (RGG) domain of SAFA is required for interaction with Xist (157). Chromatin proteins may recognize a variety of RNA sequences and secondary structure conformations. Mounting evidence is beginning to shed light on the mechanistic details of protein-lncRNA binding, but there is not sufficient information to engineer lncRNA-derived systems (158).

Scientists must first identify core RNA subunits with intrinsic, portable activity before we can use lncRNAs for synthetic systems. For an immediate solution, researchers have turned to protein-RNA modules from well-understood viral systems to design lncRNA-inspired gene regulators. Zalatan *et al.* used the viral RNA stem-loop hairpins *MS2*, *PP7* and *com* to build a scaffold that was recognized by the proteins MCP, PCP and Com respectively (Figure 5B). They targeted the scaffold to a specific gene by using a CRISPR/dCas system where the DNA-binding guide RNA (gRNA) included a long RNA extension with different combinations of the viral hairpins. The group demonstrated epigenetic repression of a target gene by recruiting KRAB-Com fusion to a *com* RNA hairpin (159). The RNA hairpin adapters are also functional in tandem (Figure 5B), which suggests that this system could be used to recruit combinations of chromatin proteins to a single locus.

CONCLUSIONS AND FUTURE OUTLOOK: BEST PRACTICES FOR CHROMATIN ENGINEERING

We have described recent advances in synthetic, chromatin-derived systems and the fundamental research discoveries that have preceded and enabled the development of these technologies. In this section we will discuss important next steps and opportunities for continued advancement of chromatin engineering.

Deeper understanding of the consequences of artificial post-translational modifications (PTMs) on chromatin is critical to advance synthetic PTM technology. Exploratory studies should determine the impact of customizable parameters such as the relative position of the PTM nucleation site to the target gene, rate of PTM production, and the interactions between PTMs and regulators at single target sites. Recent studies have demonstrated the value of systematically adjusting design parameters. For instance, Hilton *et al.* placed the active core domain of the p300 acetyltransferase at different distances from the target gene and observed significant activation of genes at distal enhancers and regulatory regions up to 46 kb from the transcriptional start site (49). Bintu *et al.* determined the kinetics of four distinct silencing regulators (EED, KRAB, HDAC4, DMNT3B) in single CHO-K1 cells to identify conditions that generally enable silencing of a predetermined degree and duration (42). Amabile *et al.* recruited combinations fusion proteins to a single gene to compare silencing induced by one, two or three repressors (KRAB, DMNT3a, DNMT3L). Two or three repressors generated long-term

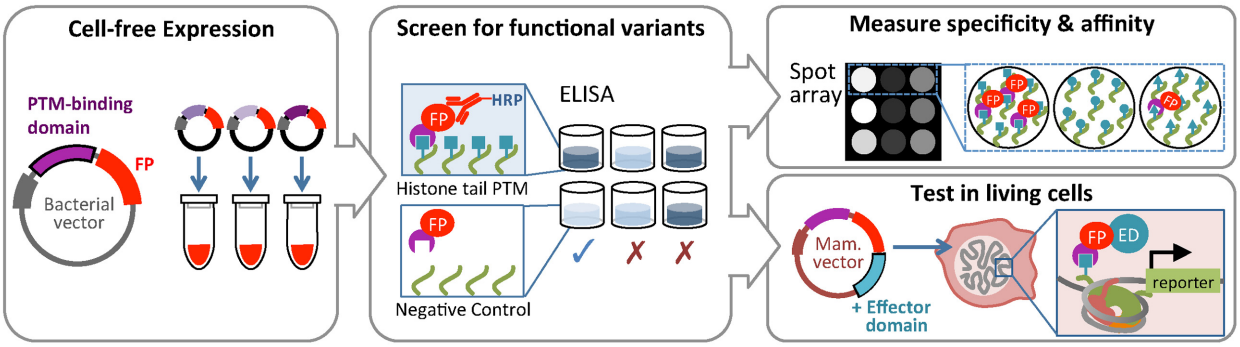

Figure 6. Workflow to discover modular, reusable chromatin-derived peptides that bind histone PTM's. FP = fluorescent protein, HRP = horseradish peroxidase-conjugated antibody, Mam. vector = mammalian expression vector, ED = effector domain.

silencing at endogenous genes in human HEK293T, K562 and B-lymphoblastoid cells, human primary T lymphocytes, and mouse NIH-3T3 cells (64). In a study of KRAB and Sss1 co-recruitment, synergy was dampened by competition of each effector for shared cofactors (41). This work revealed an important caveat in using components from overlapping pathways. Keung *et al.* recently reported an impressive study in which 223 PTM-generating fusion proteins enhanced, antagonized, or did not affect the function of a VP16 activator at a target reporter gene in yeast (150). Pioneering efforts such as the studies highlighted here provide a glimpse of the exciting work yet to be done. In addition to customizable parameters, uncontrollable events such as cross-talk between PTMs, enzyme co-recruitment and the impact of these processes on gene expression states (discussed previously in this review) must also be considered and measured to maximize the value of the results from synthetic studies.

Customization of non-enzymatic chromatin proteins represents an expansive design space that has barely been explored. In order to accelerate the pace of discovery, scientists should develop and share workflows to efficiently parse large libraries of synthetic protein candidates. As an example, we present a workflow that is under development in our lab. We begin with a large library of fusion protein variants and carry out protein–histone peptide interaction tests *in vitro* (ELISA) to identify functional candidates (Figure 6). Candidates that show preference for the target histone PTM in the first-pass test are used to build synthetic gene regulators. We then expose a PTM-bearing reporter gene to the synthetic regulator to validate its function in live cells. This workflow allows us to identify modular, reusable PTM-binding domains to aid in the design, construction, and application of synthetic, chromatin-derived proteins. In addition, we use peptide arrays to determine cross-reactions with various histone PTMs and to calculate affinities for each target. So far, our workflow has demonstrated the interaction of a Polycomb Chromodomain fusion (PCD-mCherry) with H3K27me3 *in vitro* (not published), which corroborates previous studies of this fusion protein in live cells (77,78). We expect ongoing work to identify new variants with interesting properties, such as enhanced affinity for H3K27me3.

A challenge for practical and reliable use of PTM-binding domains in living cells (e.g. as illustrated in Figure 3E and F) is the broad and varying distribution of histone modifications throughout the genome over time and at different stages of cell development. Therefore, it is important to carefully consider the histone PTM target from a systems viewpoint rather than as a single gene, as is the case for DNA-binding regulators. For instance H3K27me3, the target of the PCD motif, appears at thousands of genes and many non-coding regions in human cells. The real utility of regulating a cohort of this size may not be immediately obvious. However, a single PTM might not be sufficient to support artificial regulation of the entire set of genes. We have observed that a regulator fused to the PCD from CBX8 affects only a subset of all H3K27me3-positive genes (77,78) and that many of these genes are near bivalent H3K27me3/H3K4me3-marked promoters (78). The key lesson from this example is that simple assumptions about epigenetic targets should be tested by using integrated transcriptomic and epigenomic analyses. The identification of predictive epigenetic signatures at target genes will enable practical use of PTM-binding proteins for synthetic systems.

Eventually, new chromatin engineering tools may become available for human and animal health applications. The safety and efficacy of these tools will need to be determined in the context of a complex epigenome. Therefore, it is critical to determine the global impact of the synthetic component on genome-wide expression levels in order to identify off-target or broad effects. Evidence from HDAC inhibition experiments in pancreas cells (160–164) suggest that generating activation-associated PTMs at key genes could transdifferentiate alpha cells into beta cells, boost insulin production in pancreatic tissue, and cure diabetes. The RNA-seq data from Bramswig et al. show that dozens of genes that are not involved in the insulin production pathway are also affected (162). Nyer *et al.* used RNA-seq and ChIP-seq to determine that a broad-acting H3K27me3-binding synthetic regulator reactivated silenced tumor suppressors as well as hundreds of other genes (78). The data from the pancreas and cancer cell studies should be further explored to determine the long-term impact of broad changes in gene expression on cell phenotype. Even in cases where the synthetic chromatin protein binds a single unique DNA

target, chromatin modifications nucleated at one site can spread into neighboring loci (discussed in (49)). To verify target specificity of a Cas9-p300 histone acetyltransferase fusion protein, Hilton et al. performed RNA-seq to identify cases where only the target gene was activated and no other genes were affected (49). It is imperative that synthetic biologists and chromatin engineers include genome-wide analysis, bioinformatics, and gene network analysis in their arsenal of research techniques.

Synthetic epigenetic research is still in its infancy. The plethora of structural and biochemical data that describe histone-modifying enzymes, histone PTM-binding proteins, and modular nucleic acid scaffolds present an opportunity to expand the bioengineering toolbox. Histone PTMs represent a rich set of biological information that can be exploited for gene targeting and cellular regulation. Genomic profiling can be coupled with chromatin protein engineering to identify and manipulate epigenetically-regulated target sites. While some epigenetic fusion constructs have been successfully used to activate or repress target genes, the number of utilized domains represents a small fraction of well-characterized PTM binding and modifying proteins (80,149). So far, most of the reported synthetic epigenetic systems are limited to specific regulation of a single genetic locus using PTM modifying enzymes that are fused to DNA-binding domains. Histone PTM-binding proteins fused with effector domains, which have been explored less, could enable co-regulation of many genes at once. The efforts summarized in this review represent important, initial advances into a vast exploration space of potential chromatin protein designs and applications. Chromatin is a central mechanism for precise and reliable control of genes that drive multicellular development. Maturation of the synthetic epigenetics field will produce new technologies and discoveries that will have significant impacts on genetic research, agricultural science and biomedical engineering.

AVAILABILITY

3D models were rendered with The PyMOL Molecular Graphics System, Version 1.3 Schrödinger, LLC (https://www.pymol.org/). Structure data were downloaded from the RCSB Protein Data Bank (PDB) (www.rcsb.org) (165).

ACKNOWLEDGEMENTS

The authors thank students J. Simper and D. Barclay for early contributions to the literature search.

FUNDING

National Institutes of Health (NIH) National Cancer Institute (NCI) [K01 CA188164 to K.A.H.]; Biological Design Graduate Program at ASU (to S.T.). Funding for open access charge: NIH NCI [K01 CA188164 to K.A.H.].
Conflict of interest statement. None declared.

REFERENCES

1. Turner,B.M. (2005) Reading signals on the nucleosome with a new nomenclature for modified histones. *Nat. Struct. Mol. Biol.*, **12**, 110–112.
2. Schwaiger,M., Schonauer,A., Rendeiro,A.F., Pribitzer,C., Schauer,A., Gilles,A.F., Schinko,J.B., Renfer,E., Fredman,D. and Technau,U. (2014) Evolutionary conservation of the eumetazoan gene regulatory landscape. *Genome Res.*, **24**, 639–650.
3. 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 A resolution. *Nature*, **389**, 251–260.
4. Talbert,P.B. and Henikoff,S. (2010) Histone variants — ancient wrap artists of the epigenome. *Nat. Rev. Mol. Cell Biol.*, **11**, 264–275.
5. Berry,D. (2003) Molecular Visualizations of DNA. http://www.wehi.edu.au/wehi-tv/molecular-visualisations-dna.
6. Lyubitelev,A. V, Nikitin,D. V, Shaytan,A.K., Studitsky,V.M. and Kirpichnikov,M.P. (2016) Structure and functions of linker histones. *Biochem.*, **81**, 213–223.
7. Zhu,P. and Li,G. (2016) Structural insights of nucleosome and the 30-nm chromatin fiber. *Curr. Opin. Struct. Biol.*, **36**, 106–115.
8. Sarma,K. and Reinberg,D. (2005) Histone variants meet their match. *Nat. Rev. Mol. Cell Biol.*, **6**, 139–149.
9. Talbert,P.B. and Henikoff,S. (2016) Histone variants on the move: substrates for chromatin dynamics. *Nat. Rev. Mol. Cell Biol.*, **18**, 115–126.
10. Campos,E.I. and Reinberg,D. (2009) Histones: annotating chromatin. *Annu. Rev. Genet.*, **43**, 559–599.
11. Kimura,H. and Cook,P.R. (2001) Kinetics of core histones in living human cells. *J. Cell Biol.*, **153**, 1341–1354.
12. Lever,M.a, Th'ng,J.P., Sun,X. and Hendzel,M.J. (2000) Rapid exchange of histone H1.1 on chromatin in living human cells. *Nature*, **408**, 873–876.
13. Tan,M., Luo,H., Lee,S., Jin,F., Yang,J.S., Montellier,E., Buchou,T., Cheng,Z., Rousseaux,S., Rajagopal,N. *et al.* (2011) Identification of 67 histone marks and histone lysine crotonylation as a new type of histone modification. *Cell*, **146**, 1016–1028.
14. Strahl,B.D. and Allis,C.D. (2000) The language of covalent histone modifications. *Nature*, **403**, 41–45.
15. Kouzarides,T. (2007) Chromatin modifications and their function. *Cell*, **128**, 693–705.
16. Bannister,A.J. and Kouzarides,T. (2011) Regulation of chromatin by histone modifications. *Cell Res.*, **21**, 381–395.
17. Huang,H., Lin,S., Garcia,B.A. and Zhao,Y. (2015) Quantitative proteomic analysis of histone modifications. *Chem. Rev.*, **115**, 2376–2418.
18. Dhalluin,C., Carlson,J.E., Zeng,L., He,C., Aggarwal,a.K. and Zhou,M.M. (1999) Structure and ligand of a histone acetyltransferase bromodomain. *Nature*, **399**, 491–496.
19. 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.
20. Yun,M., Wu,J., Workman,J.L. and Li,B. (2011) Readers of histone modifications. *Cell Res.*, **21**, 564–578.
21. Musselman,C.A., Lalonde,M.-E., Côté,J. and Kutateladze,T.G. (2012) Perceiving the epigenetic landscape through histone readers. *Nat. Struct. Mol. Biol.*, **19**, 1218–1227.
22. Gao,Z., Zhang,J., Bonasio,R., Strino,F., Sawai,A., Parisi,F., Kluger,Y. and Reinberg,D. (2012) PCGF homologs, CBX proteins, and RYBP define functionally distinct PRC1 family complexes. *Mol. Cell*, **45**, 344–356.
23. Kamakaka,R.T. (2005) Histone variants: deviants? *Genes Dev.*, **19**, 295–316.
24. Taguchi,H., Xie,Y., Horikoshi,N., Maehara,K., Harada,A., Nogami,J., Sato,K., Arimura,Y., Osakabe,A., Kujirai,T. *et al.* (2017) Crystal structure and characterization of novel human histone H3 variants, H3.6, H3.7, and H3.8. *Biochemistry*, **56**, 2184–2196.
25. Haynes,K.A. and Silver,P.A. (2009) Eukaryotic systems broaden the scope of synthetic biology. *J. Cell Biol.*, **187**, 589–596.
26. Park,M., Keung,A.J. and Khalil,A.S. (2016) The epigenome: the next substrate for engineering. *Genome Biol.*, **17**, 183.
27. Pick,H., Kilic,S. and Fierz,B. (2014) Engineering chromatin states: chemical and synthetic biology approaches to investigate histone modification function. *Biochim. Biophys. Acta - Gene Regul. Mech.*, **1839**, 644–656.
28. He,S., Bauman,D., Davis,J.S., Loyola,A., Nishioka,K., Gronlund,J.L., Reinberg,D., Meng,F., Kelleher,N. and McCafferty,D.G. (2003) Facile synthesis of site-specifically

acetylated and methylated histone proteins: Reagents for evaluation of the histone code hypothesis. *Proc. Natl. Acad. Sci. U.S.A.*, **100**, 12033–12038.
29. 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.
30. Shogren-Knaak,M. (2006) Histone H4-K16 acetylation controls chromatin structure and protein interactions. *Science*, **311**, 844–847.
31. 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.
32. McGinty,R.K., Kim,J., Chatterjee,C., Roeder,R.G. and Muir,T.W. (2008) Chemically ubiquitylated histone H2B stimulates hDot1L-mediated intranucleosomal methylation. *Nature*, **453**, 812–816.
33. David,Y., Vila-Perelló,M., Verma,S. and Muir,T.W. (2015) Chemical tagging and customizing of cellular chromatin states using ultrafast trans-splicing inteins. *Nat. Chem.*, **7**, 394–402.
34. 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.
35. Kallappagoudar,S., Dammer,E.B., Duong,D.M., Seyfried,N.T. and Lucchesi,J.C. (2013) Expression, purification and proteomic analysis of recombinant histone H4 acetylated at lysine 16. *Proteomics*, **13**, 1687–1691.
36. Kim,C.H., Kang,M., Kim,H.J., Chatterjee,A. and Schultz,P.G. (2012) Site-specific incorporation of ϵ-N-crotonyllysine into histones. *Angew. Chem. Int. Ed.*, **51**, 7246–7249.
37. Wilkins,B.J., Hahn,L.E., Heitmüller,S., Frauendorf,H., Valerius,O., Braus,G.H. and Neumann,H. (2015) Genetically encoding lysine modifications on histone H4. *ACS Chem. Biol.*, **10**, 939–944.
38. Xiao,H., Xuan,W., Shao,S., Liu,T. and Schultz,P.G. (2015) Genetic incorporation of ϵ-*N*-2-hydroxyisobutyryl-lysine into recombinant histones. *ACS Chem. Biol.*, **10**, 1599–1603.
39. Elsässer,S.J., Ernst,R.J., Walker,O.S. and Chin,J.W. (2016) Genetic code expansion in stable cell lines enables encoded chromatin modification. *Nat. Methods*, **13**, 158–164.
40. Hansen,K.H., Bracken,A.P., Pasini,D., Dietrich,N., Gehani,S.S., Monrad,A., Rappsilber,J., Lerdrup,M. and Helin,K. (2008) A model for transmission of the H3K27me3 epigenetic mark. *Nat. Cell Biol.*, **10**, 1291–1300.
41. Ma,A.-N., Wang,H., Guo,R., Wang,Y.-X., Li,W., Cui,J., Wang,G., Hoffman,A.R. and Hu,J.-F. (2014) Targeted gene suppression by inducing de novo DNA methylation in the gene promoter. *Epigenet. Chromatin*, **7**, 20.
42. Bintu,L., Yong,J., Antebi,Y.E., McCue,K., Kazuki,Y., Uno,N., Oshimura,M. and Elowitz,M.B. (2016) Dynamics of epigenetic regulation at the single-cell level. *Science*, **351**, 720–724.
43. Konermann,S., Brigham,M.D., Trevino,A., Hsu,P.D., Heidenreich,M., Le,Cong, Platt,R.J., Scott,D.a, Church,G.M. and Zhang,F. (2013) Optical control of mammalian endogenous transcription and epigenetic states. *Nature*, **500**, 472–476.
44. Snowden,A.W., Gregory,P.D., Case,C.C. and Pabo,C.O. (2002) Gene-specific targeting of H3K9 methylation is sufficient for initiating repression in vivo. *Curr. Biol.*, **12**, 2159–2166.
45. Kungulovski,G., Nunna,S., Thomas,M., Zanger,U.M., Reinhardt,R. and Jeltsch,A. (2015) Targeted epigenome editing of an endogenous locus with chromatin modifiers is not stably maintained. *Epigenet. Chromatin*, **8**, 12.
46. Mendenhall,E.M., Williamson,K.E., Reyon,D., Zou,J.Y., Ram,O., Joung,J.K. and Bernstein,B.E. (2013) Locus-specific editing of histone modifications at endogenous enhancers. *Nat. Biotechnol.*, **31**, 1133–1136.
47. Kearns,N.a, Pham,H., Tabak,B., Genga,R.M., Silverstein,N.J., Garber,M. and Maehr,R. (2015) Functional annotation of native enhancers with a Cas9–histone demethylase fusion. *Nat. Methods*, **12**, 401–403.
48. Kwaks,T.H.J., Sewalt,R.G.A.B., van Blokland,R., Siersma,T.J., Kasiem,M., Kelder,A. and Otte,A.P. (2005) Targeting of a histone acetyltransferase domain to a promoter enhances protein expression levels in mammalian cells. *J. Biotechnol.*, **115**, 35–46.
49. Hilton,I.B., D'Ippolito,A.M., Vockley,C.M., Thakore,P.I., Crawford,G.E., Reddy,T.E. and Gersbach,C.A. (2015) Epigenome editing by a CRISPR-Cas9-based acetyltransferase activates genes from promoters and enhancers. *Nat. Biotechnol.*, **33**, 510–517.
50. Cano-Rodriguez,D. and Rots,M.G. (2016) Epigenetic editing: on the verge of reprogramming gene expression at will. *Curr. Genet. Med. Rep.*, **4**, 170–179.
51. Suganuma,T. and Workman,J.L. (2011) Signals and combinatorial functions of histone modifications. *Annu. Rev. Biochem.*, **80**, 473–499.
52. Suganuma,T. and Workman,J.L. (2008) Crosstalk among histone modifications. *Cell*, **135**, 604–607.
53. Smith,E. and Shilatifard,A. (2010) The chromatin signaling pathway: diverse mechanisms of recruitment of histone-modifying enzymes and varied biological outcomes. *Mol. Cell*, **40**, 689–701.
54. Zhang,T., Cooper,S. and Brockdorff,N. (2015) The interplay of histone modifications—writers that read. *EMBO Rep.*, **16**, 1467–1481.
55. 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.
56. VanDemark,A.P., Kasten,M.M., Ferris,E., Heroux,A., Hill,C.P. and Cairns,B.R. (2007) Autoregulation of the Rsc4 tandem bromodomain by Gcn5 acetylation. *Mol. Cell*, **27**, 817–828.
57. McGinty,R.K., Kim,J., Chatterjee,C., Roeder,R.G. and Muir,T.W. (2008) Chemically ubiquitylated histone H2B stimulates hDot1L-mediated intranucleosomal methylation. *Nature*, **453**, 812–816.
58. Li,B., Gogol,M., Carey,M., Lee,D., Seidel,C. and Workman,J.L. (2007) Combined action of PHD and chromo domains directs the Rpd3S HDAC to transcribed chromatin. *Science*, **316**, 1050–1054.
59. Schmitges,F.W., Prusty,A.B., Faty,M., Stützer,A., Lingaraju,G.M., Aiwazian,J., Sack,R., Hess,D., Li,L., Zhou,S. *et al.* (2011) Histone methylation by PRC2 is inhibited by active chromatin marks. *Mol. Cell*, **42**, 330–341.
60. Gatchalian,J., Fütterer,A., Rothbart,S.B., Tong,Q., Rincon-Arano,H., SánchezdeDiego,A., Groudine,M., Strahl,B.D., Martínez-A,C., Van Wely,K.H.M. *et al.* (2013) Dido3 PHD modulates cell differentiation and division. *Cell Rep.*, **4**, 148–158.
61. Pengue,G., Calabró,V., Bartoli,P.C., Pagliuca,A. and Lania,L. (1994) Repression of transcriptional activity at a distance by the evolutionary conserved KRAB domain present in a subfamily of zinc finger proteins. *Nucleic Acids Res.*, **22**, 2908–2914.
62. Pengue,G. and Lania,L. (1996) Krüppel-associated box-mediated repression of RNA polymerase II promoters is influenced by the arrangement of basal promoter elements. *Proc. Natl. Acad. Sci. U.S.A.*, **93**, 1015–1020.
63. Thakore,P.I., D'Ippolito,A.M., Song,L., Safi,A., Shivakumar,N.K., Kabadi,A.M., Reddy,T.E., Crawford,G.E. and Gersbach,C.A. (2015) Highly specific epigenome editing by CRISPR-Cas9 repressors for silencing of distal regulatory elements. *Nat. Methods*, **12**, 1143–1149.
64. Amabile,A., Migliara,A., Capasso,P., Biffi,M., Cittaro,D., Naldini,L. and Lombardo,A. (2016) Inheritable silencing of endogenous genes by hit-and-run targeted epigenetic editing. *Cell*, **167**, 219–232.
65. Lupo,A., Cesaro,E., Montano,G., Zurlo,D., Izzo,P. and Costanzo,P. (2013) KRAB-zinc finger proteins: a repressor family displaying multiple biological functions. *Curr. Genomics*, **14**, 268–278.
66. Jacobs,S.A. (2002) Structure of HP1 chromodomain bound to a lysine 9-methylated histone H3 tail. *Science*, **295**, 2080–2083.
67. Nielsen,P.R., Nietlispach,D., Mott,H.R., Callaghan,J., Bannister,A., Kouzarides,T., Murzin,A.G., Murzina,N.V. and Laue,E.D. (2002) Structure of the HP1 chromodomain bound to histone H3 methylated at lysine 9. *Nature*, **416**, 103–107.
68. Fischle,W., Wang,Y., Jacobs,S.A., Kim,Y., Allis,C.D. and Khorasanizadeh,S. (2003) Molecular basis for the discrimination of repressive methyl-lysine marks in histone H3 by polycomb and HP1 chromodomains. *Genes Dev.*, **17**, 1870–1881.
69. Min,J. (2003) Structural basis for specific binding of Polycomb chromodomain to histone H3 methylated at Lys 27. *Genes Dev.*, **17**, 1823–1828.

70. Bernstein,E., Duncan,E.M., Masui,O., Gil,J., Heard,E. and Allis,C.D. (2006) Mouse polycomb proteins bind differentially to methylated histone H3 and RNA and are enriched in facultative heterochromatin. *Mol. Cell. Biol.*, **26**, 2560–2569.
71. Eissenberg,J.C. (2012) Structural biology of the chromodomain: Form and function. *Gene*, **496**, 69–78.
72. Ren,C., Morohashi,K., Plotnikov,A.N., Jakoncic,J., Smith,S.G., Li,J., Zeng,L., Rodriguez,Y., Stojanoff,V., Walsh,M. *et al.* (2015) Small-molecule modulators of methyl-lysine binding for the CBX7 chromodomain. *Chem. Biol.*, **22**, 161–168.
73. Black,J.C., Van Rechem,C. and Whetstine,J.R. (2012) Histone lysine methylation dynamics: establishment, regulation, and biological impact. *Mol. Cell*, **48**, 491–507.
74. Kaustov,L., Ouyang,H., Amaya,M., Lemak,A., Nady,N., Duan,S., Wasney,G.A., Li,Z., Vedadi,M., Schapira,M. *et al.* (2011) Recognition and specificity determinants of the human Cbx chromodomains. *J. Biol. Chem.*, **286**, 521–529.
75. Platero,J.S., Hartnett,T. and Eissenberg,J.C. (1995) Functional analysis of the chromo domain of HP1. *EMBO J.*, **14**, 3977–3986.
76. Messmer,S., Franke,A. and Paro,R. (1992) Analysis of the functional role of the polycomb chrom domain in Drosophila melanogaster. *Genes Dev.*, **6**, 1241–1254.
77. Haynes,K.A. and Silver,P.A. (2011) Synthetic reversal of epigenetic silencing. *J. Biol. Chem.*, **286**, 27176–27182.
78. Nyer,D.B., Daer,R.M., Vargas,D., Hom,C. and Haynes,K.A. (2017) Regulation of cancer epigenomes with a histone-binding synthetic transcription factor. *npj Genomic Med.*, **2**, 1.
79. Owen,D.J. (2000) The structural basis for the recognition of acetylated histone H4 by the bromodomain of histone acetyltransferase Gcn5p. *EMBO J.*, **19**, 6141–6149.
80. Filippakopoulos,P., Picaud,S., Mangos,M., Keates,T., Lambert,J.-P., Barsyte-Lovejoy,D., Felletar,I., Volkmer,R., Müller,S., Pawson,T. *et al.* (2012) Histone recognition and large-scale structural analysis of the human bromodomain family. *Cell*, **149**, 214–231.
81. Fujisawa,T. and Filippakopoulos,P. (2017) Functions of bromodomain-containing proteins and their roles in homeostasis and cancer. *Nat. Rev. Mol. Cell Biol.*, **18**, 246–262.
82. Huang,H., Zhang,J., Shen,W., Wang,X., Wu,J., Wu,J. and Shi,Y. (2007) Solution structure of the second bromodomain of Brd2 and its specific interaction with acetylated histone tails. *BMC Struct. Biol.*, **7**, 57.
83. Umehara,T., Nakamura,Y., Jang,M.K., Nakano,K., Tanaka,A., Ozato,K., Padmanabhan,B. and Yokoyama,S. (2010) Structural basis for acetylated histone H4 recognition by the human BRD2 bromodomain. *J. Biol. Chem.*, **285**, 7610–7618.
84. Flynn,E.M., Huang,O.W., Poy,F., Oppikofer,M., Bellon,S.F., Tang,Y. and Cochran,A.G. (2015) A subset of human bromodomains recognizes butyryllysine and crotonyllysine histone peptide modifications. *Structure*, **23**, 1801–1814.
85. Wang,R. and You,J. (2015) Mechanistic analysis of the role of bromodomain-containing protein 4 (BRD4) in BRD4-NUT oncoprotein-induced transcriptional activation. *J. Biol. Chem.*, **290**, 2744–2758.
86. Dekker,F.J., van den Bosch,T. and Martin,N.I. (2014) Small molecule inhibitors of histone acetyltransferases and deacetylases are potential drugs for inflammatory diseases. *Drug Discov. Today*, **19**, 654–660.
87. Niedzialkowska,E., Wang,F., Porebski,P.J., Minor,W., Higgins,J.M.G. and Stukenberg,P.T. (2012) Molecular basis for phosphospecific recognition of histone H3 tails by Survivin paralogues at inner centromeres. *Mol. Biol. Cell*, **23**, 1457–1466.
88. Rossetto,D., Avvakumov,N. and Côté,J. (2012) Histone phosphorylation: a chromatin modification involved in diverse nuclear events. *Epigenetics*, **7**, 1098–1108.
89. Kelly,A.E., Ghenoiu,C., Xue,J.Z., Zierhut,C., Kimura,H. and Funabiki,H. (2010) Survivin reads phosphorylated histone H3 threonine 3 to activate the mitotic kinase aurora B. *Science*, **330**, 235–239.
90. Jeyaprakash,A.A., Basquin,C., Jayachandran,U. and Conti,E. (2011) Structural basis for the recognition of phosphorylated histone H3 by the surviving subunit of the chromosomal passenger complex. *Structure*, **19**, 1625–1634.
91. Wang,F., Dai,J., Daum,J.R., Niedzialkowska,E., Banerjee,B., Stukenberg,P.T., Gorbsky,G.J. and Higgins,J.M.G. (2010) Histone H3 Thr-3 phosphorylation by haspin positions aurora B at centromeres in mitosis. *Science*, **330**, 231–235.
92. Sanchez,R. and Zhou,M.-M. (2011) The PHD finger: a versatile epigenome reader. *Trends Biochem. Sci.*, **36**, 364–372.
93. Musselman,C.A. and Kutateladze,T.G. (2011) Handpicking epigenetic marks with PHD fingers. *Nucleic Acids Res.*, **39**, 9061–9071.
94. Ruthenburg,A.J., Li,H., Milne,T.A., Dewell,S., McGinty,R.K., Yuen,M., Ueberheide,B., Dou,Y., Muir,T.W., Patel,D.J. *et al.* (2011) Recognition of a mononucleosomal histone modification pattern by BPTF via multivalent interactions. *Cell*, **145**, 692–706.
95. Peña,P. V., Davrazou,F., Shi,X., Walter,K.L., Verkhusha,V. V, Gozani,O., Zhao,R. and Kutateladze,T.G. (2006) Molecular mechanism of histone H3K4me3 recognition by plant homeodomain of ING2. *Nature*, **442**, 100–103.
96. van Ingen,H., van Schaik,F.M.A., Wienk,H., Ballering,J., Rehmann,H., Dechesne,A.C., Kruijzer,J.A.W., Liskamp,R.M.J., Timmers,H.T.M. and Boelens,R. (2008) Structural insight into the recognition of the H3K4me3 mark by the TFIID subunit TAF3. *Structure*, **16**, 1245–1256.
97. Li,H., Ilin,S., Wang,W., Duncan,E.M., Wysocka,J., Allis,C.D. and Patel,D.J. (2006) Molecular basis for site-specific read-out of histone H3K4me3 by the BPTF PHD finger of NURF. *Nature*, **442**, 91–95.
98. Wysocka,J., Swigut,T., Xiao,H., Milne,T.A., Kwon,S.Y., Landry,J., Kauer,M., Tackett,A.J., Chait,B.T., Badenhorst,P. *et al.* (2006) A PHD finger of NURF couples histone H3 lysine 4 trimethylation with chromatin remodelling. *Nature*, **442**, 86–90.
99. Kim,S., Natesan,S., Cornilescu,G., Carlson,S., Tonelli,M., McClurg,U.L., Binda,O., Robson,C.N., Markley,J.L., Balaz,S. *et al.* (2016) Mechanism of histone H3K4me3 recognition by the plant homeodomain of inhibitor of growth 3. *J. Biol. Chem.*, **291**, 18326–18341.
100. Wang,G.G., Song,J., Wang,Z., Dormann,H.L., Casadio,F., Li,H., Luo,J., Patel,D.J. and Allis,C.D. (2009) Haematopoietic malignancies caused by dysregulation of a chromatin-binding PHD finger. *Nature*, **459**, 847–851.
101. Li,H., Fischle,W., Wang,W., Duncan,E.M., Liang,L., Murakami-Ishibe,S., Allis,C.D. and Patel,D.J. (2007) Structural basis for lower lysine methylation state-specific readout by MBT repeats of L3MBTL1 and an engineered PHD finger. *Mol. Cell*, **28**, 677–691.
102. Zeng,L., Zhang,Q., Li,S., Plotnikov,A.N., Walsh,M.J. and Zhou,M. (2010) Mechanism and regulation of acetylated histone binding by the tandem PHD finger of DPF3b. *Nature*, **466**, 258–262.
103. Fütterer,a, Raya,Á., Llorente,M., Izpisúa-Belmonte,J.C., de la Pompa,J.L., Klatt,P. and Martínez-A,C. (2012) Ablation of Dido3 compromises lineage commitment of stem cells in vitro and during early embryonic development. *Cell Death Differ.*, **19**, 132–143.
104. Gelato,K.A., Tauber,M., Ong,M.S., Winter,S., Hiragami-Hamada,K., Sindlinger,J., Lemak,A., Bultsma,Y., Houliston,S., Schwarzer,D. *et al.* (2014) Accessibility of different histone H3-binding domains of UHRF1 is allosterically regulated by phosphatidylinositol 5-phosphate. *Mol. Cell*, **54**, 905–919.
105. Su,X., Zhu,G., Ding,X., Lee,S.Y., Dou,Y., Zhu,B., Wu,W. and Li,H. (2014) Molecular basis underlying histone H3 lysine-arginine methylation pattern readout by Spin/Ssty repeats of Spindlin1. *Genes Dev.*, **28**, 622–636.
106. Hiragami-Hamada,K., Soeroes,S., Nikolov,M., Wilkins,B., Kreuz,S., Chen,C., De La Rosa-Velázquez,I.A., Zenn,H.M., Kost,N., Pohl,W. *et al.* (2016) Dynamic and flexible H3K9me3 bridging via HP1β dimerization establishes a plastic state of condensed chromatin. *Nat. Commun.*, **7**, 11310.
107. Norwood,L.E., Moss,T.J., Margaryan,N. V., Cook,S.L., Wright,L., Seftor,E.A., Hendrix,M.J.C., Kirschmann,D.A. and Wallrath,L.L. (2006) A requirement for dimerization of HP1Hs in suppression of breast cancer invasion. *J. Biol. Chem.*, **281**, 18668–18676.
108. Trembecka-Lucas,D.O., Trembecka-Lucas,D.O., Szczurek,A.T., Szczurek,A.T., Dobrucki,J.W. and Dobrucki,J.W. (2013) Dynamics of the HP1β-PCNA-containing complexes in DNA replication and repair. *Nucleus*, **4**, 74–82.
109. Jacobson,R.H., Ladurner,A.G., King,D.S. and Tjian,R. (2000) Structure and function of a human TAF II 250 double bromodomain module. *Science*, **288**, 1422–1425.

110. McDaniel,S.L., Fligor,J.E., Ruan,C., Cui,H., Bridgers,J.B., DiFiore,J.V., Guo,A.H., Li,B. and Strahl,B.D. (2016) Combinatorial histone readout by the dual plant homeodomain (PHD) fingers of Rco1 mediates Rpd3S chromatin recruitment and the maintenance of transcriptional fidelity. *J. Biol. Chem.*, **291**, 14796–14802.
111. Huh,J., Wu,J., Lee,C., Yun,M., Gilada,D., Brautigam,C.A. and Li,B. (2012) Multivalent di-nucleosome recognition enables the Rpd3S histone deacetylase complex to tolerate decreased H3K36 methylation levels. *EMBO J.*, **31**, 3564–3574.
112. Lee,C.-H., Wu,J. and Li,B. (2013) Chromatin remodelers fine-tune H3K36me-directed deacetylation of neighbor nucleosomes by Rpd3S. *Mol. Cell*, **52**, 255–263.
113. Tabet,S., Douglas,S.F., Daze,K.D., Garnett,G.A.E., Allen,K.J.H., Abrioux,E.M.M., Quon,T.T.H., Wulff,J.E. and Hof,F. (2013) Synthetic trimethyllysine receptors that bind histone 3, trimethyllysine 27 (H3K27me3) and disrupt its interaction with the epigenetic reader protein CBX7. *Bioorg. Med. Chem.*, **21**, 7004–7010.
114. Simhadri,C., Daze,K.D., Douglas,S.F., Quon,T.T.H., Dev,A., Gignac,M.C., Peng,F., Heller,M., Boulanger,M.J., Wulff,J.E. *et al.* (2014) Chromodomain antagonists that target the polycomb-group methyllysine reader protein chromobox homolog 7 (CBX7). *J. Med. Chem.*, **57**, 2874–2883.
115. Stuckey,J.I., Dickson,B.M., Cheng,N., Liu,Y., Norris,J.L., Cholensky,S.H., Tempel,W., Qin,S., Huber,K.G., Sagum,C. *et al.* (2016) A cellular chemical probe targeting the chromodomains of Polycomb repressive complex 1. *Nat. Chem. Biol.*, **12**, 180–187.
116. Ren,C., Smith,S.G., Yap,K., Li,S., Li,J., Mezei,M., Rodriguez,Y., Vincek,A., Aguilo,F., Walsh,M.J. *et al.* (2016) Structure-guided discovery of selective antagonists for the chromodomain of polycomb repressive protein CBX7. *ACS Med. Chem. Lett.*, **7**, 601–605.
117. Milosevich,N., Gignac,M.C., McFarlane,J., Simhadri,C., Horvath,S., Daze,K.D., Croft,C.S., Dheri,A., Quon,T.T.H., Douglas,S.F. *et al.* (2016) Selective inhibition of CBX6: a methyllysine reader protein in the polycomb family. *ACS Med. Chem. Lett.*, **7**, 139–144.
118. Whitehead,L., Dobler,M.R., Radetich,B., Zhu,Y., Atadja,P.W., Claiborne,T., Grob,J.E., McRiner,A., Pancost,M.R., Patnaik,A. *et al.* (2011) Human HDAC isoform selectivity achieved via exploitation of the acetate release channel with structurally unique small molecule inhibitors. *Bioorg. Med. Chem.*, **19**, 4626–4634.
119. Rahman,S., Sowa,M.E., Ottinger,M., Smith,J.A., Shi,Y., Harper,J.W. and Howley,P.M. (2011) The Brd4 extraterminal domain confers transcription activation independent of pTEFb by recruiting multiple proteins, including NSD3. *Mol. Cell. Biol.*, **31**, 2641–2652.
120. Filippakopoulos,P., Qi,J., Picaud,S., Shen,Y., Smith,W.B., Fedorov,O., Morse,E.M., Keates,T., Hickman,T.T., Felletar,I. *et al.* (2010) Selective inhibition of BET bromodomains. *Nature*, **468**, 1067–1073.
121. Nicodeme,E., Jeffrey,K.L., Schaefer,U., Beinke,S., Dewell,S., Chung,C., Chandwani,R., Marazzi,I., Wilson,P., Coste,H. *et al.* (2010) Suppression of inflammation by a synthetic histone mimic. *Nature*, **468**, 1119–1123.
122. James,L.I., Barsyte-Lovejoy,D., Zhong,N., Krichevsky,L., Korboukh,V.K., Herold,J.M., MacNevin,C.J., Norris,J.L., Sagum,C.A., Tempel,W. *et al.* (2013) Discovery of a chemical probe for the L3MBTL3 methyllysine reader domain. *Nat. Chem. Biol.*, **9**, 184–191.
123. Verma,S.K., Tian,X., LaFrance,L. V., Duquenne,C., Suarez,D.P., Newlander,K.A., Romeril,S.P., Burgess,J.L., Grant,S.W., Brackley,J.A. *et al.* (2012) Identification of potent, selective, cell-active inhibitors of the histone lysine methyltransferase EZH2. *ACS Med. Chem. Lett.*, **3**, 1091–1096.
124. Konze,K.D., Ma,A., Li,F., Barsyte-Lovejoy,D., Parton,T., MacNevin,C.J., Liu,F., Gao,C., Huang,X.-P., Kuznetsova,E. *et al.* (2013) An orally bioavailable chemical probe of the lysine methyltransferases EZH2 and EZH1. *ACS Chem. Biol.*, **8**, 1324–1334.
125. Horton,J.R., Liu,X., Gale,M., Wu,L., Shanks,J.R., Zhang,X., Webber,P.J., Bell,J.S.K., Kales,S.C., Mott,B.T. *et al.* (2016) Structural basis for KDM5A histone lysine demethylase inhibition by diverse compounds. *Cell Chem. Biol.*, **23**, 769–781.
126. Xu,W.S., Parmigiani,R.B. and Marks,P.A. (2007) Histone deacetylase inhibitors: molecular mechanisms of action. *Oncogene*, **26**, 5541–5552.
127. Kim,H. and Bae,S. (2011) Histone deacetylase inhibitors: molecular mechanisms of action and clinical trials as anti-cancer drugs. *Am. J. Transl. Res.*, **3**, 166–179.
128. Arrowsmith,C.H., Bountra,C., Fish,P. V, Lee,K. and Schapira,M. (2012) Epigenetic protein families: a new frontier for drug discovery. *Nat. Rev. Drug Discov.*, **11**, 384–400.
129. Falkenberg,K.J. and Johnstone,R.W. (2014) Histone deacetylases and their inhibitors in cancer, neurological diseases and immune disorders. *Nat. Rev. Drug Discov.*, **13**, 673–691.
130. Oike,T., Ogiwara,H., Amornwichet,N., Nakano,T. and Kohno,T. (2014) Chromatin-regulating proteins as targets for cancer therapy. *J. Radiat. Res.*, **55**, 613–628.
131. West,A.C. and Johnstone,R.W. (2014) New and emerging HDAC inhibitors for cancer treatment. *J Clin Invest.*, **124**, 30–39.
132. Delmore,J.E., Issa,G.C., Lemieux,M.E., Rahl,P.B., Shi,J., Jacobs,H.M., Kastritis,E., Gilpatrick,T., Paranal,R.M., Qi,J. *et al.* (2011) BET bromodomain inhibition as a therapeutic strategy to target c-Myc. *Cell*, **146**, 904–917.
133. Mertz,J.A., Conery,A.R., Bryant,B.M., Sandy,P., Balasubramanian,S., Mele,D.A., Bergeron,L. and Sims,R.J. (2011) Targeting MYC dependence in cancer by inhibiting BET bromodomains. *Proc. Natl. Acad. Sci. U.S.A.*, **108**, 16669–16674.
134. Cheung,K., Lu,G., Sharma,R., Vincek,A., Zhang,R., Plotnikov,A.N., Zhang,F., Zhang,Q., Ju,Y., Hu,Y. *et al.* (2017) BET N-terminal bromodomain inhibition selectively blocks Th17 cell differentiation and ameliorates colitis in mice. *Proc. Natl. Acad. Sci. U.S.A.*, **114**, 2952–2957.
135. Kassis,J.A. and Brown,J.L. (2013) Polycomb group response elements in Drosophila and vertebrates. In: *Advances in Genetics*. Elsevier Inc., Vol. **81**, pp. 83–118.
136. Cavalli,G. and Paro,R. (1998) The Drosophila Fab-7 chromosomal element conveys epigenetic inheritance during mitosis and meiosis. *Cell*, **93**, 505–518.
137. Cavalli,G. and Paro,R. (1999) Epigenetic inheritance of active chromatin after removal of the main transactivator. *Science*, **286**, 955–958.
138. Déjardin,J. and Cavalli,G. (2004) Chromatin inheritance upon Zeste-mediated Brahma recruitment at a minimal cellular memory module. *EMBO J.*, **23**, 857–868.
139. Müller,J. and Kassis,J.A. (2006) Polycomb response elements and targeting of Polycomb group proteins in Drosophila. *Curr. Opin. Genet. Dev.*, **16**, 476–484.
140. Bauer,M., Trupke,J. and Ringrose,L. (2016) The quest for mammalian Polycomb response elements: are we there yet? *Chromosoma*, **125**, 471–496.
141. Woo,C.J., Kharchenko,P. V, Daheron,L., Park,P.J. and Kingston,R.E. (2010) A region of the human HOXD cluster that confers Polycomb-group responsiveness. *Cell*, **140**, 99–110.
142. Woo,C.J., Kharchenko,P. V, Daheron,L., Park,P.J. and Kingston,R.E. (2013) Variable requirements for DNA-binding proteins at Polycomb-dependent repressive regions in human HOX clusters. *Mol. Cell. Biol.*, **33**, 3274–3285.
143. Cuddapah,S., Roh,T.-Y., Cui,K., Jose,C.C., Fuller,M.T., Zhao,K. and Chen,X. (2012) A novel human Polycomb binding site acts as a functional Polycomb response element in Drosophila. *PLoS One*, **7**, e36365.
144. Sing,A., Pannell,D., Karaiskakis,A., Sturgeon,K., Djabali,M., Ellis,J., Lipshitz,H.D. and Cordes,S.P. (2009) A vertebrate Polycomb response element governs segmentation of the posterior hindbrain. *Cell*, **138**, 885–897.
145. Vasanthi,D., Nagabhushan,A., Matharu,N.K. and Mishra,R.K. (2013) A functionally conserved Polycomb response element from mouse HoxD complex responds to heterochromatin factors. *Sci. Rep.*, **3**, 3011.
146. Ringrose,L., Rehmsmeier,M., Dura,J. and Paro,R. (2003) Genome-wide prediction of polycomb/trithorax response elements in Drosophila melanogaster. *Dev. Cell*, **5**, 759–771.
147. Hauenschild,A., Ringrose,L., Altmutter,C., Paro,R. and Rehmsmeier,M. (2008) Evolutionary plasticity of polycomb/trithorax response elements in Drosophila species. *PLoS Biol.*, **6**, 2130–2143.

148. Gilman,J. and Love,J. (2016) Synthetic promoter design for new microbial chassis. *Biochem. Soc. Trans.*, **44**, 731–737.
149. Qin,J.Y., Zhang,L., Clift,K.L., Hulur,I., Xiang,A.P., Ren,B.-Z. and Lahn,B.T. (2010) Systematic comparison of constitutive promoters and the doxycycline-inducible promoter. *PLoS One*, **5**, e10611.
150. Keung,A.J., Bashor,C.J., Kiriakov,S., Collins,J.J. and Khalil,A.S. (2014) Using targeted chromatin regulators to engineer combinatorial and spatial transcriptional regulation. *Cell*, **158**, 110–120.
151. Engreitz,J.M., Ollikainen,N. and Guttman,M. (2016) Long non-coding RNAs: spatial amplifiers that control nuclear structure and gene expression. *Nat. Rev. Mol. Cell Biol.*, **17**, 756–770.
152. Somarowthu,S., Legiewicz,M., Chillón,I., Marcia,M., Liu,F. and Pyle,A.M. (2015) HOTAIR forms an intricate and modular secondary structure. *Mol. Cell*, **58**, 353–361.
153. Tsai,M.-C., Manor,O., Wan,Y., Mosammaparast,N., Wang,J.K., Lan,F., Shi,Y., Segal,E. and Chang,H.Y. (2010) Long noncoding RNA as modular scaffold of histone modification complexes. *Science*, **329**, 689–693.
154. Wu,L., Murat,P., Matak-Vinkovic,D., Murrell,A. and Balasubramanian,S. (2013) Binding interactions between long noncoding RNA HOTAIR and PRC2 proteins. *Biochemistry*, **52**, 9519–9527.
155. Portoso,M., Ragazzini,R., Brenčič,Ž., Moiani,A., Michaud,A., Vassilev,I., Wassef,M., Servant,N., Sargueil,B. and Margueron,R. (2017) PRC2 is dispensable for HOTAIR -mediated transcriptional repression. *EMBO J.*, **36**, 981–994.
156. Hacisuleyman,E., Shukla,C.J., Weiner,C.L. and Rinn,J.L. (2016) Function and evolution of local repeats in the Firre locus. *Nat. Commun.*, **7**, 11021.
157. Hasegawa,Y., Brockdorff,N., Kawano,S., Tsutui,K., Tsutui,K. and Nakagawa,S. (2010) The matrix protein hnRNP U is required for chromosomal localization of Xist RNA. *Dev. Cell*, **19**, 469–476.
158. Hendrickson,D.G., Kelley,D.R., Tenen,D., Bernstein,B. and Rinn,J.L. (2016) Widespread RNA binding by chromatin-associated proteins. *Genome Biol.*, **17**, 28.
159. Zalatan,J.G., Lee,M.E., Almeida,R., Gilbert,L.A., Whitehead,E.H., La Russa,M., Tsai,J.C., Weissman,J.S., Dueber,J.E., Qi,L.S. *et al.* (2015) Engineering complex synthetic transcriptional programs with CRISPR RNA scaffolds. *Cell*, **160**, 339–350.
160. Wang,X., Wei,X., Pang,Q. and Yi,F. (2012) Histone deacetylases and their inhibitors: molecular mechanisms and therapeutic implications in diabetes mellitus. *Acta Pharm. Sin. B*, **2**, 387–395.
161. Haumaitre,C., Lenoir,O. and Scharfmann,R. (2008) Histone deacetylase inhibitors modify pancreatic cell fate determination and amplify endocrine progenitors. *Mol. Cell. Biol.*, **28**, 6373–6383.
162. Bramswig,N.C., Everett,L.J., Schug,J., Dorrell,C., Liu,C., Luo,Y., Streeter,P.R., Naji,A., Grompe,M. and Kaestner,K.H. (2013) Epigenomic plasticity enables human pancreatic α to β cell reprogramming. *J. Clin. Invest.*, **123**, 1275–1284.
163. Khan,S. and Jena,G.B. (2014) Protective role of sodium butyrate, a HDAC inhibitor on beta-cell proliferation, function and glucose homeostasis through modulation of p38/ERK MAPK and apoptotic pathways: Study in juvenile diabetic rat. *Chem. Biol. Interact.*, **213**, 1–12.
164. Khan,S. and Jena,G. (2014) Sodium butyrate, a HDAC inhibitor ameliorates eNOS, iNOS and TGF-β1-induced fibrogenesis, apoptosis and DNA damage in the kidney of juvenile diabetic rats. *Food Chem. Toxicol.*, **73**, 127–139.
165. Berman,H.M., Westbrook,J., Feng,Z., Gilliland,G., Bhat,T.N., Weissig,H. and Shindyalov,I.N. (2000) The Protein Data Bank. *Nucleic Acids Res.*, **28**, 235–242.
166. Andrews,F.H., Strahl,B.D. and Kutateladze,T.G. (2016) Insights into newly discovered marks and readers of epigenetic information. *Nat. Chem. Biol.*, **12**, 662–668.
167. Wang,G.G., Song,J., Wang,Z., Dormann,H.L., Casadio,F., Li,H., Luo,J., Patel,D.J. and Allis,C.D. (2009) Haematopoietic malignancies caused by dysregulation of a chromatin-binding PHD finger. *Nature*, **459**, 847–851.
168. Houbaviy,H.B., Usheva,A., Shenk,T. and Burley,S.K. (1996) Cocrystal structure of YY1 bound to the adeno-associated virus P5 initiator. *Proc. Natl. Acad. Sci. U.S.A.*, **93**, 13577–13582.
169. PDB ID: 4XBF, Luka,Z., Loukachevitch,L.V., Martin,W.J., Wagner,C. and Reiter,N.J. (2016) Structure of LSD1:CoREST in complex with ssRNA.