Published in final edited form as:

Nat Methods. 2014 August ; 11(8): 834–840. doi:10.1038/nmeth.3022.

Accelerated Chromatin Biochemistry Using DNA-Barcoded Nucleosome Libraries

Uyen T. T. Nguyen[1,2], **Lenka Bittova**[1], **Manuel M. Müller**[1], **Beat Fierz**[1,3], **Yael David**[1], **Brian Houck-Loomis**[4], **Vanessa Feng**[1], **Geoffrey P. Dann**[1], and **Tom W. Muir**[1,5]

[1]Department of Chemistry, Princeton University, Princeton, NJ 08544, United States

[4]The Rockefeller University, New York, NY 10065, United States; current address: New York Genome Center, New York, NY 10013, United States

Abstract

Elucidating the molecular details of how chromatin-associated factors deposit, remove and recognize histone posttranslational modification ('PTM') signatures remains a daunting task in the epigenetics field. Here, we introduce a versatile platform that greatly accelerates biochemical investigations into chromatin recognition and signaling. This technology is based on the streamlined semi-synthesis of DNA-barcoded nucleosome libraries with distinct combinations of PTMs. Chromatin immunoprecipitation of these libraries treated with purified chromatin effectors or the combined chromatin recognizing and modifying activities of the nuclear proteome is followed by multiplexed DNA-barcode sequencing. This ultrasensitive workflow allowed us to collect thousands of biochemical data points revealing the binding preferences of various nuclear factors for PTM patterns and how pre-existing PTMs, alone or synergistically, affect further PTM deposition via crosstalk mechanisms. We anticipate that the high-throughput and -sensitivity of the technology will help accelerate the decryption of the diverse molecular controls that operate at the level of chromatin.

Introduction

Chromatin functions as an information hub that integrates diverse biochemical stimuli to orchestrate fine-tuned changes in the downstream transcriptional program [1]. These signals are dynamically inscribed into chromatin through covalent modifications of DNA and histones by 'writer' and 'eraser' enzymes [2,3], while 'readers' further convert this chromatin landscape into defined transcriptional outputs [4]. The complex interplay of these effectors

[5]Corresponding author: muir@princeton.edu.
[2]Current address: Bayer Pharma AG, 13353 Berlin, Germany
[3]Current address: Ecole polytechnique fédérale de Lausanne, CH-1015 Lausanne, Switzerland

Author contributions U.T.T.N., M.M.M. and T.W.M. conceived the research. U.T.T.N. and T.W.M. designed the research with substantial contribution by B.F. U.T.T.N. and L.B. prepared reagents and performed experiments. M.M.M., B.F., Y.D. and B.H.L. provided reagents with the help of L.F. and G.P.D. U.T.T.N., M.M.M. and T.W.M. wrote the manuscript.

Competing financial interests

U.T.T.N., M.M.M. and T.W.M. are co-inventors of a patent application on DNA-barcoding of nucleosomes for profiling chromatin effectors.

with chromatin remains poorly understood despite their potential as drug targets for various human pathologies [5] and the body of correlative information generated by top-down 'omics' efforts [6]. Extracting mechanistic details from large-scale datasets requires biochemical approaches that recapitulate the structural, chemical and functional complexity of chromatin, yet are simple and sensitive enough to allow high-resolution and -throughput data collection [7]. Numerous studies have shown that modified nucleosomes, the minimal repeating units of chromatin, recapitulate well the functional properties of chromatin towards the activity of trans-acting nuclear factors (for example, [8,9]). In particular, they are indispensible substrates in biochemical studies that require the three-dimensional architecture of the nucleosome and where, as a consequence, histone-derived peptides are poor probes of chromatin signaling events. Examples include multivalent interactions involving multiple histone tails [10] or PTM crosstalk mechanisms that depend upon the location of pre-installed marks on different histones [11]. Modified nucleosomes of defined chemical composition can be assembled from the corresponding modified histones prepared by protein semi-synthesis [7]. However, the low throughput of their manufacture and functional implementation, requiring time- and material-consuming processes, has failed to keep nucleosome-based approaches apace with the numerous biochemical questions raised by genome-wide 'omics' initiatives [6].

To address these issues, we apply the concept of DNA barcoding, which has found utility in several areas including small-molecule libraries [12] and tissue-specific antibody sensing [13], to chromatin biochemistry. In our approach, DNA-barcoded libraries of modified nucleosomes, assembled in a streamlined fashion, are treated with purified effectors or the combined chromatin recognizing and modifying activities of the nuclear proteome. Subsequently, the desired products are isolated by chromatin immunoprecipitation, followed by multiplexed DNA-barcode next generation sequencing (*in vitro* ChIP-Seq, Fig. 1a). This workflow allowed us to investigate PTM-based recruitment and modulation of histone mark readers and writers, respectively, and to investigate how PTM signals, alone or synergistically, result in composite systems-level signal outputs through the combined action of the nuclear proteome.

Results

Streamlined preparation of DNA-barcoded nucleosomes

We envisioned a technology based on the use of chemically defined DNA-barcoded nucleosome libraries ('DNLs', Fig. 1a). In this approach, each library member is primed with a distinct combination of histone PTMs, incorporated by protein semi-synthesis, and this PTM pattern, representing a module of a transcriptionally active or repressive chromatin state, is then encoded in a unique barcode appended on the nucleosomal DNA. DNLs allow (i) competitive biochemical assays with the entire pooled collection of nucleosome variants in solution; (ii) an ultrasensitive readout of the modified reaction products by *in vitro* ChIP-Seq; and, as a consequence, (iii) a large number of *in vitro* ChIP-Seq assays to be performed and analyzed in parallel (Fig. 1a). Importantly, because of the strict 1:1 molar correspondence between the barcode and nucleosome it encodes, the sequencing data can be

used to derive relative levels of each member in the library before and after a selection experiment.

We tailored the DNLs in the current study specifically towards the elucidation of chromatin regulation focusing on the bromodomain (BD), a reader module that associates weakly with acetyl-lysine (Kac) modifications [14]. We were particularly interested in how the magnitude and specificity of BD-mediated interactions are fine-tuned by higher order histone PTM patterns and by the presence of reader modules adjacent to the BD, such as the plant homeodomain (PHD) finger, which often associates strongly with trimethylated lysines [4]. Accordingly, a collection of unmodified and modified histones, bearing single or multiple site-specific acetylation, trimethylation and ubiquitylation marks (Fig. 1a and Supplementary Fig. 1), were prepared by solid-phase peptide synthesis, recombinant protein production and traceless protein ligation [15]. Distinct combinations of these histones were combined with 5′-barcoded versions of the strong nucleosome positioning sequence 601 [16] (Supplementary Fig. 2) to yield, in the first version ('DNL-1'), 40 different nucleosomes, containing up to 10 site-specific PTMs and heterotypic modifications on multiple histones (Fig. 1b and Supplementary Table 1). Importantly, we developed a protocol that substantially cuts down the assembly time (3 days) and production scale (0.5 nmole per nucleosome variant) compared to the state-of-the-art approach for nucleosome preparation [17]. This streamlined method is based on a novel affinity-tagged buffer DNA strategy (Supplementary Fig. 3a), which eliminates the need for time- and material-consuming octamer purification and DNA-octamer ratio tests. Indeed, our method permitted all 40 nucleosomes in DNL-1 to be prepared simultaneously in 30 % overall yield (Supplementary Fig. 3b–d). Subsequently, the nucleosome variants were combined in an equimolar ratio to yield the final DNL-1 (Fig. 1c). Although shown for tens of variants in the current study, this method could be extended easily to larger numbers of histone combinations within the same time window and assembly scale.

In vitro ChIP-Seq with DNL-1

To evaluate the fidelity of the barcoding approach and downstream DNA processing, we performed *in vitro* ChIP-Seq experiments using a series of antibodies directed against specific histone PTMs. Simultaneous analysis of many library experiments by Ion Torrent sequencing [18] was achieved by further tagging (multiplexing) the isolated DNA sequences, corresponding to the affinity-captured nucleosomes of each experiment, with a second, manipulation-specific barcode using PCR (Supplementary Fig. 4 and Supplementary Table 2). Importantly, we employed PCR reaction conditions for both the multiplexing and sequencing steps that retain the original DNA ratios in the initial pull-down mixtures (Supplementary Fig. 5a). The sequencing results revealed that, for the antibodies tested, only nucleosomes carrying the expected PTM were isolated (Supplementary Fig. 5b–d, for input-normalized values, see Source Data Table 1), indicating that the barcoding system is robust and that the library members retained their integrity, even when exposed to freeze-thawing or extended storage for > 6 months at 4 °C (Supplementary Fig. 5e). Moreover, these initial control experiments highlighted the sensitivity of the method: each *in vitro* ChIP-Seq experiment consumed 12 fmoles of each library member, meaning that a typical library

production scale (yielding ~150 pmoles of each nucleosome variant) provides enough material for up to 12,500 of such *in vitro* ChIP-Seq library experiments.

Profiling the substrate preference of histone mark readers

Next we probed the three-dimensional contact requirement of single or combinations of reader modules towards patterns of histone PTMs. Specifically, we employed nuclear factors carrying either a BD and PHD finger (the BD PHD finger transcription factor, BPTF, and the transcriptional co-activator, p300) or two consecutive BDs (the transcriptional co-activator, Brd4).

Immobilized BPTF-derived constructs containing the complete PHD-BD module or the respective single reader domains (Supplementary Fig. 6a,b) were incubated with DNL-1 and the affinity-captured nucleosomes analyzed by sequencing. In agreement with peptide and nucleosome binding studies [10], we found that only the BPTF-PHD-BD fusion associates bivalently to substrates carrying a H3K4me3 mark in conjunction with a mono-acetylated lysine residue on H4 (Fig. 2a and Supplementary Fig. 6c). A 2–3-fold increase in nucleosomal binding had been observed previously with nucleosomes modified with H3K4me3-H4K16ac [10], and the contribution of Kac to multivalency was less pronounced at adjacent sites in the H4 tails. In our competitive binding experiments, this positional dependence is flattened out while a strong enhancement of about 7-fold was found with $H4Kac_5$-containing nucleosomes, a variant not examined in previous work. Since the Kac-binding pocket of BPTF's BD can only accommodate one acetylated lysine [10], the enhanced affinity is likely to be a consequence of binding avidity resulting from multiple acetylated lysines in the H4 tails. The increase in association was confirmed using standard pull-down experiments employing individual nucleosomes (Supplementary Fig. 6d).

DNL-1 library recognition experiments employing p300-BD, -PHD or -BD-PHD modules as baits (Supplementary Fig. 7a,b) revealed weak binding of the BD and BD-PHD constructs to nucleosomes containing H4K12ac and robust association to substrates with hyperacetylated H4 tails (Fig. 2b), which is consistent with data reported recently using a modified histone peptide array [19]. The magnitude of this interaction was not influenced by the presence of other types of PTMs, arguing for a BD-mediated interaction. Indeed, the isolated p300-PHD module did not associate with any of the DNL-1 members (Supplementary Fig. 7c), which agrees well with the atypical structure of this domain in p300 [19].

DNL-1 was next incubated with an immobilized construct spanning the continuous BD modules (BD1 and BD2) of the Brd4 (Supplementary Fig. 8a), a promising drug target for treating multiple myeloma and acute myeloid leukemia [20]. Consistent with histone peptide interaction studies [21], Brd4 associated robustly with nucleosomes bearing hyperacetylated H4 tails, and the extent of this binding was not influenced by the co-occurrence of trimethylated or ubiquitylated lysines (Fig. 2c and Supplementary Fig. 8b). In contrast, a 4.5-fold increase in recruitment was observed when the substrate was hyperacetylated on both H3 and H4 tails, compared to a hyperacetylated H4 alone.

To further investigate the relationship between the degree of acetylation and nucleosome binding, we assembled a second more focused nucleosome library, DNL-2. In addition to

the acetylated substrates employed in DNL-1, we included 14 new nucleosome variants containing hyperacetylated H4 tails and iterative acetyl combinations of mono- and di-acetylated H3, as well as substrates carrying 2 or 3 Kac residues distributed on both H3 and H4 tails (Supplementary Fig. 8c,d and Supplementary Table 3). Incubation of DNL-2 with p300-BD-PHD and Brd4-BD1-BD2 underlined, in both cases, the dominant role of hyperacetylation on H4 in driving nucleosome association (Supplementary Fig. 8e and 9a, respectively). Both p300 and Brd4 exhibited individual preferences for mono- and di-acetylated substrates, albeit subtle in nature and, in the case of Brd4, modest in comparison to the binding enhancements associated with hyperacetylated H4 substrates. Consistent with peptide binding studies [21], we observed that the isolated BD1 of Brd4 had much higher affinity for acetylated nucleosomes than BD2 (Supplementary Fig. 9b and 9c, respectively), although the presence of the latter did improve association of Brd4 with nucleosomes in several contexts. In particular, the Brd4-BD1-BD2 construct had a 2.6-fold increase in affinity for nucleosomes hyperacetylated on both H3 and H4 tails compared to Brd4-BD1. This situation is reminiscent of the effect of the BD in BPTF binding and, as there, argues for a bivalent interaction. Indeed, the data is in agreement with a model in which Brd4 recognizes hyperacetylated H4 nucleosomes primarily through its BD1, with this association being enhanced through further interactions mediated by BD2 [21].

The DNL-2 experiment revealed a striking interplay between hyperacetylated H3 and H4 tails in augmenting p300 binding. In particular, the p300 BD module showed strong affinity for hyperacetylated H4 tails, whereas this domain has little affinity for nucleosomes containing only hyperacetylated H3 tails. Remarkably, the presence of both hyperacetylated histones in the same nucleosome - the $H3Kac_5$-$H4Kac_5$ variant- resulted in a dramatic 16-fold binding enhancement as compared to $H4Kac_5$ (Supplementary Fig. 8e). Mono- and diacetylation of H3 sites also augmented recruitment in the presence of H4 hyperacetylation, although to a lesser extent (1.5–4.5 fold). Additional studies will be needed to understand why the effect of H3 acetylation on p300 association should be so context-dependent. Perhaps more relevant to the current study, these observations further illustrate the advantages of employing nucleosomes in binding studies since this inter-tail enhancement would not be detected using peptide-derived ligands.

Profiling the regulation of histone mark writers

We modulated the *in vitro* ChIP-Seq workflow to investigate whether readout of a modified nucleosome by a chromatin effector leads to downstream signal propagation. Specifically, we asked how recognition of hyperacetylated H4 by p300 affects its histone acetyltransferase (HAT) activity. DNL-1 was incubated with purified full-length human p300 (Supplementary Fig. 9d) in the presence of acetyl-coenzyme A (acetyl-CoA). The resulting *de novo* acetylated reaction products were isolated using several antibodies that recognize site-specific acetylated lysines (Fig. 3a). Importantly, only antibodies were employed that showed little or no cross-reactivity with the same PTM at other histone sites or to other types of PTMs, as tested in *in vitro* ChIP-Seq experiments in the absence of enzyme and substrate (Fig. 3b, Supplementary Fig. 10, and Supplementary Table 4). All nucleosomes were modified, albeit to different levels, on multiple lysine residues in the H2A, H3 and H4 tails (Fig. 3b, and Supplementary Fig. 10a–i), which agrees well with

previous mass spectrometry data [22]. Additionally, we observed a substantial increase in global acetylation on substrates carrying pre-installed mono- and hyperacetylation signatures on H3 and H4, compared to the unmodified substrates. This is similar to the cooperative acetylation behavior described for yeast Gcn5 [23]. In line with our binding data, the observed stimulation of HAT activity was not detected with substrates containing only trimethylation or ubiquitination marks. Antibody-independent HAT activity measurements with selected members of our library confirmed these key findings (Fig. 3c and Supplementary Fig. 11a, b).

Additional *in vitro* ChIP-Seq library experiments allowed us to elucidate the molecular basis of the observed signal amplification. Enzymatic assays with a BD-crippled p300 construct [24] diminished acetylation of all nucleosomes pre-acetylated on H4 while deletion of the BD [25] abolished p300 HAT activity in all nucleosomal contexts (Fig. 3b, and Supplementary Fig. 11c). Stimulation of acetylation on unmodified nucleosomes was not observed in the presence of the acetyl-lysine analog, Nω-acetylhistamine [24] (Supplementary Fig. 12a), or a tetra-acetylated H4 tail peptide (Supplementary Fig. 12b). By contrast, both these additives reduced the histone acetylation feedback. Collectively, these data suggest that the p300 positive feedback loop is mediated mainly by nucleosomal recruitment through the BD as opposed to an allosteric process (Fig. 3d).

Profiling histone PTM crosstalk relationships on a systems-level

Next we turned to the question of how acetylation is regulated on a systems-level, i.e. in the context of the manifold signal inputs resulting from the joint chromatin recognition and enzyme activities of the nuclear proteome (Fig. 4a). DNL enzyme assays were performed with nuclear extracts from various cell lines, followed by isolation of the reaction products with a H3K14ac-specific antibody. Subsequent readout by sequencing revealed robust *de novo* acetylation of several members of the library (Fig. 4b, and Supplementary Fig. 13). Interestingly, we observed a positive acetylation feedback loop that was similar to the behavior of p300 *in vitro* (Supplementary Fig. 14a), that is to say increased H3K14ac on nucleosomes carrying pre-installed mono- and hyperacetylation H4 signatures. On a systems-level, the acetylation signal was also regulated by H3 trimethylation in a position-dependent manner: trimethylation at H3K4 robustly stimulated acetylation, a finding that has been reported previously [8], whereas methylation on other H3 sites had a minimal effect (Fig. 4b). These results were confirmed using Western blot and radioactive HAT assays with selected library members (Supplementary Fig. 14b,c). Moreover, our library experiment allowed us to explore synergies between known crosstalk phenomena. For example, H3K4me3 augmented the stimulatory effect of H4 hyperacetylation, collectively resulting in a burst-type signaling behavior, whereas H3K27me3 attenuated the effect.

Since we observed stimulation of histone acetylation by methylation of histone 3 at lysine 4, we repeated the ChIP-Seq experiment with nuclear extract employing an antibody specific towards H3K4me1, a mark which is localized to enhancer elements of genes and which represents the most abundant methylation state at this position *in vivo* [26,27]. In stark contrast to acetylation, this experiment revealed that methylation was modulated by a large number of PTMs, each conferring a subtle difference in activity (Fig. 4c and Supplementary Fig.

14d). Gratifyingly, we detected known crosstalk relationships involving this modification (i.e. stimulation by H3Kac$_5$ and H2BK120ub as well as repression by H3K27me3 [28–30]). Furthermore, a notable repression of H3K4 monomethylation by H2AK119ub was detected, consistent with its role in H3K4 di- and trimethylation [31]. The opposing effects of H2AK119ub and H2BK120ub on the installation of H3K4me1 were confirmed by Western blot analysis with selected library members (Supplementary Fig. 14e). Similar to acetylation, a synergistic modulation of these PTMs was observed; for example, hyperacetylated H3 was found to amplify the effect of H2BK120ub, whereas H3K27me3 antagonized this stimulation. Thus, a key conclusion of our nuclear extract experiments is that the *de novo* deposition of histone marks is highly dependent on pre-existing nucleosome modifications. That these nucleating marks can act in a cooperative fashion to fine-tune the activity of histone 'writers', points to the existence of additional layers of regulation in such PTM crosstalk processes and highlights the critical need for defined chromatin templates to tease them apart.

Discussion

The technology described herein has several advantages over existing methods for the biochemical analysis of chromatin. Compared to the use of histone-derived peptides [32,33], nucleosomes provide a more physiological substrate, enabling the identification of regulatory processes that operate through inter-histone or histone-DNA mechanisms. At the same time, the speed and sensitivity of our method allows an orders of magnitude higher throughput, lower material consumption and the possibility to perform competitive experiments compared to traditional single nucleosome experiments. Indeed, in the course of the current study, we generated 54 chemically defined, modified nucleosomes and performed the equivalent of ~4500 individual nucleosome assays utilizing only 4.5 pmoles (~900 ng) of each nucleosome variant. Acquisition of such a data set would be impractical, both in terms of time and material consumption, using standard biochemistry efforts employing individual nucleosomes. In this sense, our method bridges the gap between top-down '-omics' technologies, which afford large amounts of mostly correlative data, and traditional biochemical approaches, which explore direct causal relationships in a quantitative fashion.

Several of the experiments described in this work employed modification-specific antibodies during the affinity capture step. As with all ChIP-Seq studies [6], the value of the resultant data sets is very much dependent on the quality of such antibodies. An important feature of our technology is therefore the ability to assess the efficiency and specificity of antibodies for ChIP-Seq applications using the input library of nucleosomes with pre-installed PTMs. In our work, we have used the sequencing data to determine the relative enrichment of each library member following affinity capture (i.e. from the sequencing reads of the barcodes with input correction). While it is straightforward to make comparisons within data sets and between those derived from reader binding experiments using appropriate normalization (Supplementary Fig. 4), in some instances, e.g. those from different antibody ChIP-Seq experiments, correlating such data sets may not be possible. Even with this constraint, our method allows rapid access to reliable mechanistic information. For example, we were able to extract the binding preferences of various nuclear factors for modified chromatin,

information that was consistent with previous binding studies using peptides or individual modified nucleosomes. In addition, our approach enabled us to detect a number of known histone PTM crosstalk pathways, while allowing the synergism of these relationships to be explored on a systems-level, often for the first time. While the current experiments were geared towards investigations of BD-mediated chromatin recruitment and systems-level *de novo* acetylation and methylation, the next generations of DNLs (encompassing, e.g., other PTMs, DNA modifications, histone variants and/or mutants) will open up new avenues for a fast and comprehensive characterization of global chromatin signal integration networks. We anticipate that this knowledge will contribute to ongoing efforts towards the development of novel epigenetic drugs [5] and cancer diagnostic tools [34].

Materials and Methods

General laboratory methods

Amino acid derivatives, coupling reagents and resins were purchased from Novabiochem. All commonly used chemical reagents and solvents were purchased from Sigma-Aldrich Chemical Company, Fisher Scientific, or Thermo Scientific. ^{3}H-Acetyl coenzyme A and ^{3}H-S-adenosyl methionine were purchased from Perkin Elmer. Amplify solution, glutathione agarose resins, and Sephacryl S-200 resins were obtained from GE Healthcare. Sf9 cells, MyOne streptavidin magnetic beads, SybrGold, Qubit high sensitivity ds DNA quantification kit, and ethidium bromide DNA stain were purchased from Invitrogen. Chemically competent DH5α and Rosetta™ 2(DE3)pLysS cells were purchased from Novagen. T4 DNA ligase, T4 Polynucleotide Kinase, Phusion, Proteinase K, and restriction enzymes were obtained from New England BioLabs. Slide-A-Lyzer MINI Dialysis Devices (cutoff 3,500 Da) were purchased from Thermo Scientific. Primer synthesis and gene sequencing were performed by Integrated DNA Technologies and Genewiz, respectively. Gene mutagenesis was achieved using a QuikChange Site-Directed Mutagenesis kit. Criterion 15% polyacrylamide Tris-HCl and 5% polyacrylamide TBE gels were purchased from BioRad. Centrifugal filtration units were from Sartorius, and MINI dialysis units, ECL solution, BCA protein quantification kit and GelCode protein stain were from Pierce. Kodak XAR film, anti-Flag M2 affinity gel, and anti-Flag peptide was purchased from Sigma Aldrich Chemical Company. Qiafilter Plasmid Giga, PCR purification, and gel extraction kits were purchased from Qiagen and Denville Scientific. Antibodies for mono-, trimethylated, and acetylated histones (Supplementary Table 4) as well as the anti-H3 antibody were purchased from Active Motif, Abcam, and Millipore. Size exclusion chromatography was performed on an AKTA FPLC system from GE Healthcare equipped with a P-920 pump and UPC-900 monitor. Analytical reversed-phase HPLC (RP-HPLC) was performed on a Agilent 1100 series instrument with a Vydac C18 column (5 micron, 4 × 150 mm), employing 0.1% TFA in water (HPLC buffer A), and 90% acetonitrile, 0.1% TFA in water (HPLC buffer B), as the mobile phases. Typical analytical gradients were 30–70% HPLC buffer B over 30 min at a flow rate of 1 mL/min. Preparative scale purifications were conducted on a Waters DeltaPrep 4000 system equipped with a Waters 486 tunable detector (Waters, Milford, MA). A Vydac C18 preparative coluNUC (15–20 micron, 20 × 250 mm) or a semi-preparative column (12 micron, 10 mm × 250 mm) was employed at a flow rate of 20 mL/min or 4 mL/min, respectively. ESI-MS analysis was conducted on a MicrOTOF-Q

II ESI-Qq-TOF mass spectrometer. UV spectrometry was performed on an Agilent 8453 UV-Vis spectrophotometer. All protein starting materials and ligation products were analyzed by C18 analytical RP-HPLC and ESI-MS.

Cell culture and nuclear extract preparation

U2OS and HFF cells were a kind gift from the Cristea lab (Princeton University). MDA-MB and NMuMG cells were a kind gift from the Kang lab (Princeton University). All mammalian cells (U2OS, 293T, HFF, MDA-MB, NMuMG) were cultured as monolayers in Dulbecco's Minimal Essential Medium, supplemented with 10% fetal calf serum and penicillin, streptomycin, and L-glutamine to near confluency in a 37°C incubator in an atmosphere of 5% CO_2 in air. Nuclear extract preparation was performed as described[35], resulting in about ~ 2 mg/mL total protein (quantified using a BCA protein quantification kit) in 20 mM HEPES, pH 7.9, 20 % (v/v) glycerol, 0.1 M KCl, 0.2 mM EDTA, 1 mM DTT, 1x Roche Complete Protease Inhibitor Cocktail.

Cloning, expression and purification of human GST-BPTF and GST-p300 constructs

Human BPTF cDNA (NP_872579.2) encoding residues 2722–2890 (GST-BPTF-PHD-BD), 2722–2781 (GST-BPTF-PHD), or the 2781–2890 (GST-BPTF-BD) [10] carrying an *N*-terminal Glutathione-*S*-transferase tag was cloned into a pGEX-2TPL vector by standard cloning techniques and verified by DNA sequencing. Human p300 cDNA (Q09472.2) encoding residues 1039–1285 (GST-p300-BD-PHD), 1039–1196 (GST-p300-BD), or 1185–1285 (GST-p300-PHD) [24] carrying an *N*-terminal Glutathione-*S*-transferase tag was cloned into a pGEX-2TPL vector by standard cloning techniques and verified by DNA sequencing. BPTF and p300 GST fusion proteins were expressed following the published protocol for BPTF expression [10].

Cloning, expression and purification of the human Flag-Brd4 construct

pFlag-CMV2-Brd4 (1–1362) encoding the cDNA of full-length human Brd4 protein was obtained from Addgene (plasmid #22304). Nucleotides corresponding to residues 1–722 (encoding the BDs [36]) and an additional N-terminal Flag-tag were cloned into the pFastBac1 vector by standard cloning techniques and verified by DNA sequencing. Virus generation and amplification were performed following the manufacturer's instructions (Invitrogen). Sf9 cells were transfected with the virus in mid-log phase from a 3 day suspension, and expression was allowed to proceed for 60 h at 27 °C. Cells were harvested and lysed using a Dounce homogenizer. All buffers used in the following steps were identical to those used for GST-BPTF and GST-p300 purifications. The clarified lysates were incubated with anti-Flag M2 affinity gel (Sigma) for 4 h at 4 °C on a tube rotator. The resin was washed (3 × 15 min), and the protein was eluted with 0.4 mg/mL Flag peptide (Sigma) and analyzed on a 15 % Tris-HCl Criterion gel (BioRad) followed by Coomassie Brilliant Blue (CBB) staining.

Cloning, expression and purification the Brd4 single bromodomains, His_6-Brd4-BD1-Flag and His_6-Brd4-BD2-Flag

Human Brd4 cDNA encoding residues 44–168 (for BD1) or 287–530 (for BD2) carrying an *N*-terminal His_6 tag and a *C*-terminal Flag tag was cloned into a pET15b vector by standard

cloning techniques and verified by DNA sequencing. Protein expression and purification was performed as described [21], except that we performed a second purification step by size exclusion chromatography. The final products were analyzed on a 15 % Tris-HCl Criterion gel (BioRad) followed by Coomassie Brilliant Blue (CBB) staining.

Expression of p300 in Sf9 cells

FastBac1 plasmids encoding an N-terminally Flag-tagged full-length human p300 (p300) and the corresponding bromodomain-deletion mutant (p300-ΔBD) (Q09472.2) were kind gifts from Dr. Robert G. Roeder (Rockefeller University) and Dr. W. Lee Kraus (UT Southwestern) [28,37]. The point mutant (p300-Y1089F1090) [24] was generated by PCR using the QuikChange Site-Directed Mutagenesis kit. The proteins were expressed recombinantly in Sf9 cells and purified on anti-Flag M2 affinity gel resin as described [38].

Expression of recombinant wild-type histones

Unmodified, recombinant human histones H2A, H2B, H3.2(C110A) and H4 were expressed and purified as described for the corresponding *Xenopus laevis* variants [11,39] with some minor modifications. In brief, *E. coli* BL21(DE3)pLysS cells were transfected with histone expression plasmids, grown in 6 L LB medium at 37 °C until OD_{600} of 0.6 and protein expression was induced by addition of 0.5 mM IPTG for 2–3 h. Cells were harvested by centrifugation at 4500 × g, resuspended in 40 mL of lysis buffer (50 mM Tris-HCl, 100 mM NaCl, 1 mM EDTA, 1 mM 2-mercaptoethanol, pH 7.5) homogenized by sonication, and lysed by 5 passages through a French Press. Inclusion bodies were washed twice with 40 mL lysis buffer (20 mM Tris-HCl, 200 mM NaCl, 1 mM EDTA, 1 mM 2-mercaptoethanol, pH 7.5) containing 1% Triton X-100, and once without detergent. The pellets were soaked in 0.5 mL DMSO and extracted with extraction buffer (6 M guanidine hydrochloride, 20 mM Tris-HCl, 1 mM EDTA, 100 mM NaCl, 1 mM DTT, pH 7.5) under agitation for 2 h. The suspension was clarified by centrifugation and dialyzed against unfolding buffer A (10 mM Tris-HCl, 7 M urea, 1 mM EDTA, 0.1 M NaCl, 1 mM DTT, pH 7.5). The solution was then applied to a HiTrap SP HP 5 mL column, equilibrated in unfolding buffer A. The histones were eluted using a gradient over 15 column volumes (CV) from unfolding buffer A to unfolding buffer B (10 mM Tris-HCl, 7 M urea, 1 mM EDTA, 1 M NaCl, 1 mM DTT, 0.2 mM PMSF, pH 8.0). The histones were further purified using preparative C-18 RP-HPLC and a gradient of 40–70% HPLC solvent B. The unmodified recombinant histones were analyzed by analytical HPLC and ESI-MS (H2A $(M+H)^+$ observed: 13964 Da, expected 13964 Da; H2B $(M+H)^+$ observed: 13758 Da, expected 13759 Da; H3.2(C110A) $(M+H)^+$ observed: 15256 Da, expected 15257 Da; H4 $(M+H)^+$ observed: 11235 Da, expected 11236 Da).

Synthesis of H2AK119ub(G76A) and H2BK120ub(G76A)

Human H2AK119ub and H2BK120ub were produced following the general sequential expressed protein ligation route developed for the corresponding *Xenopus laevis* variants [40,41] with some minor modifications as indicated below.

H2AK119ub—The protein was assembled from three pieces, a recombinant fragment of H2A (residues 1–114, **1**) containing an α-thioester derived from thiolysis of the

corresponding intein fusion, a synthetic peptide fragment of H2A (residues 115–130 and containing an A115C mutation, **2**) containing a cysteine residue attached to the ε-amino group of lysine-119 through an isopeptide bond, and a recombinant fragment of ubiquitin (residues 1–75, **3**) containing an α-thioester derived from thiolysis of the corresponding intein fusion. Proteins **1** and **3** were expressed in *E. coli* and processed according to published protocols [40]. Synthetic peptide **2** was assembled using automated solid-phase peptide synthesis with a 9-fluorenylmethyl-oxycarbonyl (Fmoc) Nα protection strategy and using 2-(1H-benzotriazole-1-yl)-1,1,3,3-tetramethyluronium hexafluorophosphate (HBTU) for amino acid activation on a CEM liberty microwave-assisted peptide synthesizer (CEM; Matthews, NC, USA). Orthogonal 4-methyltrityl (Mtt) protection of the ε-amino group of lysine-119 allowed regioselective attachment of cysteine to this side-chain, whilst the *N*-terminal cysteine residue was protected as a thiazolidine derivative to ensure the desired regioselectivity in the first ligation. Assembly of H2AK119ub proceeded by expressed protein ligation of protein **3** and peptide **2** to give the ubiquitylated conjugate of peptide **2**. Deprotection of the thiazolidine group within this intermediate using methoxylamine under acidic conditions allowed the attachment of protein **1** through a second expressed protein ligation step. Finally, the synthesis was completed by radical-initiated desulfurization of the two cysteine residues at the ligation junctions. The HPLC purified ubiquitylated histone was characterized by analytical RP-HPLC and ESI-MS (H2AK119ub $(M+H)^+$ observed: 22526 Da; expected 22525 Da; see Supplementary Fig. 1).

H2BK120ub—The protein was assembled from three pieces, a recombinant fragment of H2B (residues 1–116, **4**) containing an α-thioester derived from thiolysis of the corresponding intein fusion, a synthetic peptide fragment of H2B (residues 117–125 and containing an A117C mutation, **5**) containing a cysteine residue attached to the ε-amino group of lysine-120 through an isopeptide bond, and a recombinant fragment of ubiquitin (residues 1–75, **3**) containing an α-thioester derived from thiolysis of the corresponding intein fusion. Protein **4** was expressed in *E. coli* and processed according to standard protocols [40]. Synthetic peptide **5** was assembled using automated solid-phase peptide synthesis with an Fmoc Nα protection strategy and using HBTU for amino acid activation. Orthogonal protection of the ε-amino group of lysine-120 with the Mtt group allowed regioselective attachment of cysteine to this side-chain, whilst the N-terminal cysteine residue was protected as a thiazolidine derivative to set the desired regioselectivity in the first ligation. Assembly of H2BK120ub proceeded by expressed protein ligation of protein **3** and peptide **5** to give the ubiquitylated conjugate of peptide **5**. Deprotection of the thiazolidine group within this intermediate using methoxylamine under acidic conditions allowed the attachment of protein **4** through a second expressed protein ligation step. Finally, the synthesis was completed by radical-initiated desulfurization of the two cysteine residues at the ligation junctions. The HPLC purified ubiquitylated histone was characterized by analytical RP-HPLC and ESI-MS (H2BK120ub $(M+H)^+$ observed: 22320 Da; expected 22320 Da; see Supplementary Fig. 1).

Synthesis of modified H3 α-thioester peptides

H3 peptides carrying a trimethylated lysine (H3K4me3, **6**, H3K9me3, **7**, and H3K27me3, **8**) or 5 acetylated lysines (H3K9/14/18/23/27ac, $H3Kac_5$, **9**) corresponding to residues 1–28 of

human H3.1 were synthesized on a 0.1 mmole scale by a combination of manual and automated synthesis with a Fmoc Nα protection strategy and using HBTU for amino acid activation. Standard *tert*-butyl side-chain protection was used throughout, and Fmoc-Lys(me3)-OH or Fmoc-Lys(ac)-OH were used to introduce the modifications at the desired positions. The C-terminal residue, Fmoc-Ser(tBu)-OH, was coupled to diaminobenzoic acid in solution as described [42]. The resulting amino-acyl-linker was then loaded onto a rink amide ChemMatrix resin by using HBTU and the remainder of the sequence was assembled by using Fmoc-protocols. Following chain assembly, the free amino-group of the diaminobenzoyl linker was acylated by treating the resin with molar excess of p-nitrophenylchloroformate (PNCF) for 2 × 40 min at room temperature (RT). Each resin was then incubated with 0.5 M *N*,*N*-diisopropylethylamine (DIEA) in DMF to yield a cyclic *N*-acylurea intermediate. The peptides were then cleaved from the resin with TFA/triisopropylsilane (TIS)/H_2O (95:2.5:2.5) for 3 h. Crude peptides were incubated in reducing buffer (100 mM phosphate buffer, 2-mercaptoethanesulfonate (MESNa), pH 7.5) for 1 h at RT to generate the desired α-thioester prior to purification. The crude peptides were purified by RP-HPLC on a preparative scale using a 5–20 % HPLC buffer B gradient over 60 min, yielding 8–22 mg pure peptide. The α-thioester peptides were characterized by ESI-MS (H3K4me3-MESNa $(M+H)^+$ observed: 3165 Da; expected: 3165 Da; H3K9me3-MESNa $(M+H)^+$ observed: 3165 Da; expected: 3165 Da; H3K27me3-MESNa $(M+H)^+$ observed: 3165 Da; expected: 3165 Da; H3Kac_5-MESNa $(M+H)^+$ observed: 3332 Da; expected: 3332 Da).

H3 peptides carrying a mono- (H3K9ac, **10**, H3K14ac, **11**, and H3K18ac, **12**, H3K23ac, **13**, H3K27ac, **14**,) or diacetylated lysines (H3K9acK14ac, **15**, H3K14acK18ac, **16**, H3K18acK23ac, **17**) were synthesized on a 0.24 mmole scale on a derivatized 2-chloro-trityl chloride resin to obtain a C-terminal hydrazide using manual addition of the reagents or on a Liberty Peptide Synthesizer equipped with a Discovery microwave module (CEM, Matthews, NC). For H3K9ac, H3K14ac and H3K9acK14ac, residues 1–14 were synthesized, while for H3K18ac, H3K23ac, H3K27ac, H3K14acK18ac and H3K18acK23ac residues 1–28 were assembled. A reported protocol for hydrazide incorporation into peptides [43] was followed with minor modifications. A four-fold excess (relative to resin loading) of hydrazine monohydrate was added to 2 mL of DMF along with an 8-fold excess of DIPEA, and the resulting reaction mixture was then placed in an ice water bath. After addition of the hydrazine solution to the resin, the reaction was stirred for one hour. The resin was subsequently washed three times with DMF, followed by repetition of the reaction three times with a fresh aliquot of hydrazine for quantitative loading. Peptide chain assembly was achieved through conventional Fmoc solid phase synthesis protocols. Note that the first residue was coupled manually irrespective of whether manual or machine peptide synthesis was used for the remaining residues. The peptides were cleaved from the resin with 95% TFA, 2.5% TIS and 2.5% H_2O, precipitated with diethyl ether, dissolved in water with 0.1% TFA and analyzed via RP-HPLC. Preparative HPLC purification (5–25% HPLC buffer B) was then used to isolate the peptide of interest (yields: 5–10 mg), and the final products were analyzed by ESI-MS (monoacetylated H3 hydrazide peptides (residue 1–14) $(M+H)^+$ observed: 1546 Da; expected: 1546 Da; diacetylated H3 hydrazide peptides (residue 1–14) $(M+H)^+$ observed: 1588 Da; expected: 1588 Da; monoacetylated H3 hydrazide peptides (residue 1–28) $(M+H)^+$ observed: 3052 Da; expected: 3053 Da;

diacetylated H3 hydrazide peptides (residue 1–28) $(M+H)^+$ observed: 3095 Da; expected: 3095 Da).

Synthesis of acetylated H4 α-thioester peptides

Synthesis of acetylated α-thioester H4-derived peptides (H4K5ac, **18**, H4K8ac, **19**, H4K12ac, **20**, H4K16ac, **21**, H4K20ac, **22**, and H4K5ac/K8ac/K12ac/K16ac/K20ac [$H4Kac_5$] **23**; all peptides carried an N-terminal acetyl group) encompassing residues 1–37 of human H4 was performed using a Nα-^{t}butoxycarbonyl-leucyl-3-mercaptopropionamide-MBHA resin on a 0.125 mmole scale with a ^{t}butoxycarbonyl (Boc) Nα protection strategy using an *in situ* neutralization/HBTU amino-acid activation protocol according to published protocols [44]. The acetyl group at the N-terminus was introduced through treatment with acetic anhydride, whereas Boc-Lys(Ac)-OH was used to introduce the modification into the desired side-chain(s). For the mono-acetylated H4 peptides, the trityl-associated mercaptopropionic acid-leucine (TAM-PAL) method of resin attachment was employed [45]. Following cleavage with liquid hydrofluoric acid containing 5% (v/v) cresol as a scavenger, the crude peptides were purified by RP-HPLC on a preparative scale using a 15–40 % ($H4Kac_5$) or 17–42 % (mono-acetylated peptides) HPLC buffer B gradient over 60 min, yielding 10–15 mg pure peptide. The acH4(1–37)-COSR thioester peptides were characterized by ESI-MS ($H4Kac_5$ $(M+H)^+$ observed: 4279 Da; expected: 4278 Da; H4K5ac $(M+H)^+$ observed: 4179 Da; H4K8ac $(M+H)^+$ observed: 4179 Da; H4K12ac $(M+H)^+$ observed: 4179 Da; H4K16ac $(M+H)^+$ observed: 4179 Da; H4K20ac $(M+H)^+$ observed: 4179 Da; mono-acetylated peptides $(M+H)^+$ expected: 4178 Da).

Synthesis of acetylated human Bio-H4Kac$_4$ peptide

An H4 peptide carrying 4 acetylated lysines (H4K5/8/12/16ac) as well as an N-terminal acetyl group (Bio-$H4Kac_4$, **24**) corresponding to residues 1–20 of human H4 with a C-terminal Gly-Lys(biotin)-Gly extension was synthesized on a 0.1 mmole scale on a rink amide ChemMatrix resin by a combination of manual and automated Fmoc-synthesis using HBTU for amino acid activation. Standard *tert*-butyl side-chain protection was used throughout. The acetyl group at the N-terminus was introduced through treatment with acetic anhydride. Fmoc-Lys(ac)-OH was employed to install the modified acetylated residues, whereas Fmoc-Lys(Alloc)-OH was used at residue 22 to incorporate the biotin moiety after Alloc deprotection [46]. Following chain assembly, the peptide was cleaved from the resin with TFA/TIS/H_2O (95:2.5:2.5) for 3 h. The crude peptide was purified by RP-HPLC on a preparative scale using a 5–25 % HPLC buffer B gradient over 60 min, yielding 8 mg pure peptide. The peptide was characterized by ESI-MS (Bio-$H4Kac_4$ $(M+H)^+$ observed: 2670 Da; expected: 2670 Da).

Preparation of recombinant H3 C-terminal fragment

Using the human H3.1 expression plasmid as a template, an H3 gene fragment encompassing residues 15–135 with an A29C, C96A, and C110A mutation (**25**) or 29–135 with an A29C, C96A and C110A mutation (**26**) was PCR amplified and inserted in-frame into a pET30 6x-His-Sumo plasmid to afford the corresponding H3 fragment fused to an N-terminal 6x-His tag followed by a SUMO tag. The correct sequence was confirmed by gene

sequencing. For protein expression, *E. coli* BL21(DE3) cells, transformed with the H3 plasmid, were grown in auto-inducing medium overnight at 37 °C. Cells were harvested by centrifugation and lysed in lysis buffer (50 mM Tris-HCl, 150 mM NaCl, 1% (v/v) Triton X-100, 0.1 mM EDTA, 5 mM 2-mercaptoethanol, pH 7.9) by passage through a French Press (4x). Pelleted inclusion bodies were washed once with lysis buffer containing 1% (v/v) Triton X-100 and once with lysis buffer without detergent. Inclusion bodies were extracted with extraction buffer (50 mM Tris-HCl, 150 mM NaCl, 6 M GdmHCl, 5 mM imidazole, pH 7.9) and applied to Ni:NTA resin. The resin was washed with 2 × 3 CV extraction buffer containing 5 mM imidazole and 1 × 3 CV of urea buffer (50 mM Tris-HCl, 150 mM NaCl, 6 M urea, pH 8.2). The protein was eluted with 5 × 1 CV elution buffer (37.5 mM Tris-HCl, 112.5 mM NaCl, 4.5 M urea, 500 mM imidazole, pH 8.2) and analyzed by SDS-PAGE. SUMO refolding and protease cleavage were performed simultaneously by dilution into buffer containing 50 mM Tris-HCl, 150 mM NaCl, 1.5 M urea, 100 mM L-arginine, 10 mM L-cysteine, 2 mM DTT, pH 7.5 to give a final concentration of 1 mg/mL total protein in the presence of 0.03 mg/mL Ulp1 Sumo protease at RT. The cleavage reaction was monitored by RP-HPLC (45–60 % B gradient over 30 min) and proceeded to near completion after 2 h. The proteins were precipitated with 15% (w/v) trichloroacetic acid (TCA) on ice for 30 min. The precipitate was isolated and washed with cold Milli-Q water. The pelleted protein was re-dissolved in 45 % HPLC buffer B, and the insoluble material removed by centrifugation (15,000 g, 15 min). 10 mM tris(2-carboxyethyl)phosphine (TCEP) was added, and the reaction mixtures were purified by preparative RP-HPLC (45–60 % HPLC buffer B gradient, 20 mL/min, over 60 min). The final products were characterized by ESI-MS (H3.1 (15–135, A29C/C96A/C110A) $(M+H)^+$ observed: 13769 Da, calculated: 13769 Da; H3.1 (29–135, A29C/C96A/C110A) $(M+H)^+$ observed: 12261 Da, calculated: 12261 Da). Typical yields were 20 or 6 mg (15–135 or 29–135, respectively) pure protein per liter cell culture.

Preparation of recombinant H4 C-terminal fragment

Using the human H4 expression plasmid as a template [11,39,44], an H4 gene fragment encompassing residues 37–102 with an A37C mutation (**27**) was PCR amplified and subcloned into the pET30 6x-His-Sumo plasmid described above. The correct sequence was verified by gene sequencing. For protein expression, *E. coli* Rosetta(DE3) cells, transformed with the H4 plasmid, were grown in LB medium to an OD_{600} of 0.65 and induced by the addition of 1 mM IPTG. Protein expression was allowed to proceed for 1.5 h at 37°C after which cells were harvested by centrifugation and lysed in lysis buffer (50 mM Tris-HCl, 150 mM NaCl, 0.5 % (v/v) Triton X-100, 0.1 mM EDTA, Roche protease inhibitors, 1 mM 2-mercaptoethanol, pH 7.5) by passage 4x through a French Press. Pelleted inclusion bodies were washed once with lysis buffer containing 1% (v/v) Triton X-100 and once with lysis buffer without detergent. Inclusion bodies were extracted with extraction buffer (50 mM Tris-HCl, 6 M GdmHCl, pH 7.9), applied to Ni:NTA resin, and nutated overnight at 4 °C. The resin was washed with 5 CV of extraction buffer containing 5 mM imidazole, followed by 5 CV of urea buffer (50 mM Tris-HCl, 6 M urea, pH 7.5). The protein was eluted with 5 × 1 CV urea buffer with 200 mM imidazole and analyzed by SDS-PAGE. The combined elution fractions were diluted 3-fold with protease cleavage buffer (50 mM Tris-HCl, 2 M urea, 150 mM NaCl, 90 mM L-Arginine, 2 mM DTT, pH 7.5) containing 0.03 mg/mL Ulp1 Sumo protease. The cleavage reaction was allowed to proceed at RT and was deemed

complete after 1 hour as indicated by RP-HPLC (30–70 % HPLC buffer B gradient over 30 min). The reaction was quenched by adding GdmHCl to a final concentration of 2 M, which redissolved precipitates that had formed during the reaction. Any other insoluble material was removed by centrifugation. The protein was purified by preparative RP-HPLC (30–60 % HPLC B gradient over 60 min, 20 mL/min). The final product was characterized by ESI-MS ($(M+H)^+$ observed: 7,329 Da, calculated: 7,329 Da). Typical yields were 2 mg pure protein per liter cell culture.

Generation of modified human H3.1

In a typical ligation reaction, 3.7 mg of H3(29–135, **26**) protein (0.3 μmole) and 2.0 mg of the respective (1–28) H3 thioester peptide (H3K4me3, **6**, H3K9me3, **7**, H3K27me3, **8**, or $H3Kac_5$, **9**, 0.6 μmole, 2 eq.) were dissolved in ligation buffer (100 mM phosphate, 6 M guanidine hydrochloride, pH 7.8) in the presence of 100 mM 4-mercaptophenylacetic acid (MPAA) and 50 mM TCEP to a final concentration of 0.5 mM and 1.0 mM protein and peptide, respectively. For all H3Kme3-containing ligations, the concentration was 2.0 mM and 4.0 mM protein and peptide, respectively. The ligations proceeded rapidly and yielded full conversion within 3–6 h. An equal volume of HPLC buffer B was added to quench the reaction, and the product was purified by semi-preparative C-18 RP-HPLC using a 47–62 % HPLC buffer B gradient over 45 min, yielding typically 2.5–3.0 mg of pure modified H3 proteins. The products were characterized by ESI-MS and desulfurized to the native alanine residue using a free-radical based approach [47]. In a typical reaction, the modified H3 proteins were dissolved in solubilization buffer (100 mM phosphate, 6 M GdmHCl, pH 6.8) at a concentration of 10 mg/mL protein. 40 mM glutathione and 200 mM TCEP were added followed by addition of 16 mM VA-061 in methanol, and the mixture was incubated under argon atmosphere at 37 °C for 18 h. The resulting modified desulfurized proteins were purified by semi preparative C-18 RP-HPLC over a 47–62 % HPLC buffer B, yielding about 1.3–1.6 mg of pure protein. The product was characterized by analytical RP-HPLC and ESI-MS - see Supplementary Fig. 1).

For mono- and diacetylated H3.1 proteins, the general ligation method reported in ref [43] was followed with slight modifications described herein. The respective peptide acyl-hydrazide (H3K9ac, **10**, H3K14ac, **11**, H3K18ac, **12**, H3K23ac, **13**, H3K27ac, **14**, H3K9acK14ac, **15**, H3K14acK18ac, **16**, H3K18acK23ac, **17**, 1.3 μmoles, final concentration ~ 2.5 mM) was dissolved in oxidation buffer (20 mM $NaNO_2$, 6 M GdmHCl, 0.2 M phosphate buffer, pH 3.0) and allowed to react for 45 minutes in an ice water bath. To form the *in situ* thioester, an equal volume (relative to the oxidation reaction) of 300 mM MPAA, 6 M GdmHCl, 0.2 M phosphate buffer, pH 7.0, was added to the oxidation reaction. The pH was then adjusted with NaOH to give a final pH of 7.5 and the reaction incubated for 15 minutes at room temperature for full conversion to the peptide thioester. Finally, the thioester reaction was added to either 0.65 μmoles of (**25**) or (**26**). The pH was adjusted to 7.5, and the reaction was allowed to proceed for 18 hours at RT. After one hour of reaction time, TCEP was added to a final concentration of 20 mM and additional aliquots of TCEP were added as required to maintain complete reduction of thiols. The crude proteins were purified by semi-preparative RP-HPLC using a 30–70 % HPLC buffer B gradient over 45 min. The products were characterized by ESI-MS and desulfurized to the native alanine residue as described

above, yielding typically 1–2 mg of pure modified H3 proteins. The product was characterized by analytical RP-HPLC and ESI-MS - see Supplementary Fig. 1.

Generation of acetylated human H4

Semi-synthesis of acetylated H4 proteins was performed essentially identically as described above for H3Kac$_5$. A typical scale was 0.4 μmole of H4 fragment (**27**) and 0.8 μmole of the respective peptides (H4K5ac, **18**, H4K8ac, **19**, H4K12ac, **20**, H4K16ac, **21**, H4K20ac, **22**, or H4Kac$_5$, **23**). Following ligation and desulfurization, semisynthetic H4 proteins were purified by semi-preparative C-18 RP-HPLC using either a 39–57 % (mono-acetylated versions) or 40–60 % (H4Kac$_5$) buffer B gradient over 45 min. Typical yields were about 1.0 mg protein. The final products were then characterized by analytical RP-HPLC and ESI-MS - see Supplementary Fig. 1.

Octamer formation

Octamers containing the desired histone variants (Supplementary Table 1) were assembled as previously described [11,48] with the following modifications. Recombinant and semi-synthetic histones H2A, H2B, H3 and H4 were dissolved in histone unfolding buffer (20 mM Tris-HCl, 6 M GdmHCl, 5 mM DTT, pH 7.5) and combined at equimolar ratios. The total histone concentration was adjusted to 1 mg/mL. Note that for downstream single nucleosome experiments we used 5 nmoles of each histone in the assembly, whereas for library experiments only 1 nmole of each histone was employed. The mixtures were placed in Slide-A-Lyzer MINI Dialysis devices (3,500 Da cutoff) and dialyzed at 4 °C against 3× 600 mL of octamer refolding buffer (10 mM Tris-HCl, 2 M NaCl, 1 mM EDTA, 1 mM DTT, pH 7.5) for at least 4 h for each step, with one dialysis step overnight. The mixtures were then transferred to clean microcentrifuge tubes and spun down at 17,000 g for 5 min at 4 °C to remove precipitates. Supernatants were transferred into fresh microcentrifuge tubes. 50 % (v/v) glycerol was added, and the octamer concentrations were measured by UV spectroscopy. A fraction of the octamers was directly processed for nucleosome assemblies, and the rest was stored at −20 °C. Typical yields were 60 %, resulting in about 300 pmoles of histone octamers.

DNA preparation

The 147 bp 601 sequence containing a 5′-BsaI and a 3′-DraIII overhang (601) was prepared by digestion of either; (i) a DNA fragment produced by PCR using Phusion Polymerase and a 601 DNA template [16] and primers 1 (FW) and 2 (RV, Supplementary Table 5) or, (ii) a plasmid containing 12 copies of the desired sequence (flanked by EcoRV sites on either site; full sequence of 1 repetitive unit:

5′–
ATATCACCCTAGGTCTCTGATGCTGGAGAATCCCGGTGCCGAGGCCGCTCAATT
GGTCGTAGACAGCTCTAGCACCGCTTAAACGCACGTACGCGCTGTCCCCCGCGT
TTTAACCGCCAAGGGGATTACTCCCTAGTCTCCAGGCACGTGTCAGATATATAC
ATCCTGTCACGCGGTGAACAGCGATATC-3′

The PCR product was produced on large-scale, yielding ~500 μg of crude material, which was purified using a Qiagen PCR purification kit. The plasmid was produced in DH5α competent cells on a 1 L scale and purified using a Qiagen Qiafilter Plasmid Giga kit yielding typically 5 mg DNA. Both DNA fragments were digested with BsaI and DraIII (total DNA concentration of 1 mg/mL in the presence of 100 U/mL of each restriction enzyme) at 20 °C for 20 h to release the 147 bp 601 sequence with the desired overhangs (601). The digested PCR product was purified using a Qiagen PCR purification kit, ethanol (70 % (v/v)) precipitated, re-dissolved in DNA buffer (10 mM Tris-HCl, pH 8.5), quantified by UV spectroscopy at 260 nm, and stored in aliquots at −80 °C. The digested plasmid was phenol/chloroform-extracted and purified by acrylamide electrophoresis using a preparative cell (Biorad, Hercules, CA), followed by precipitation with 70 % (v/v) ethanol. The pellet was re-dissolved in DNA buffer, quantified by UV spectroscopy at 260 nm, and stored in aliquots at −80 °C. Barcoded nucleosomal DNA (BC-601) was generated by DNA ligation of a barcoding module (BC) to the 601 DNA. The double-stranded BC DNA module containing both the 6 bp barcode (upper strand: HNNNNH, where H = A/T/C; N = any nucleotide) and nucleotides 10–30 of the forward Ion Torrent Adaptor A (FW-Adaptor A) was produced by hybridizing stoichiometric amounts of

upper strand: 5′-CCTGCGTGTCTCCGACTCAGHNNNNH-3′

and

bottom strand: 5′-CATCDNNNNDCTGAGTCGGAGACACGCAGGAA-3′

(for a list of barcodes, Supplementary Table 1) at 95 °C for 5 min and slowly cooling down to RT for 1 h. In a typical ligation reaction, 0.41 μM of 601 DNA was combined with 1.1 eq. ds ‘BC’ DNA and incubated with 0.1 U/μL Polynucleotide kinase (PNK) in a total volume of 100 μL in 1x T4 ligase buffer (NEB) for 1 h at 37 °C. Subsequently, 10 U/μL T4 DNA ligase was added, and the mixture was incubated for 1 h at RT. The ligation reaction was monitored by native gel polyacrylamide electrophoresis (5 % polyacrylamide, 0.5 × Tris/Borate/EDTA (TBE), 200 V, 40 min) followed by ethidium bromide DNA staining (Supplementary Fig. 2b). The final product was purified using a Qiagen PCR purification kit and quantified by UV spectroscopy at 260 nm. The biotinylated buffer DNA was prepared by PCR using a MMTV DNA template [49] with the primers 3 and 4 (Supplementary Table 5). The PCR product was purified using a Qiagen PCR purification kit. The final product was quantified by UV spectroscopy at 260 nm, and stored in aliquots at −80 °C.

Nucleosome reconstitution

In a typical reconstitution, a crude histone octamer preparation (50 pmoles) was combined with the appropriate BC-601 DNA (30 pmoles, for a list of nucleosome barcodes, see Supplementary Table 1) and biotinylated buffer DNA (20 pmoles) in 70 μL octamer refolding buffer (10 mM Tris-HCl, 2 M NaCl, 0.1 mM EDTA, 1 mM DTT, pH 7.8; Supplementary Fig. 3a). The mixture was transferred to a Slide-A-Lyzer MINI Dialysis devices (3,500 Da cutoff) and dialyzed against 200 mL nucleosome reconstitution start buffer (10 mM Tris-HCl, 1.4 M KCl, 0.1 mM EDTA, 1 mM DTT, pH 7.8) for 1 h at 4 °C. Subsequently, 330 mL of nucleosome reconstitution end buffer (10 mM Tris-HCl, pH 7.8, 10 mM KCl, 0.1 mM EDTA, 1 mM DTT) was added to the extended solution at a rate of 1

mL/min using a peristaltic pump, followed by two final dialysis steps against nucleosome reconstitution end buffer (1 h and overnight). The dialysis mixture was transferred to a microcentrifuge tube, and removal of buffer DNA-associated side-products was achieved by affinity depletion using 35 μL MyOne streptavidin-coated Dynabeads for 1 h at RT. The unbound material in the supernatant was transferred to a clean microcentrifuge tube, and any precipitate was removed by centrifugation. The final nucleosomes were quantified by UV spectroscopy at 260 nm. The quality of the individual nucleosomes was assessed by polyacrylamide electrophoresis (5 % acrylamide gel, 0.5 × TBE, 200 V, 40 min), followed by ethidium bromide staining. This resulted in a main band of > 90 % purity migrating between the 500 and 800 bp DNA markers, depending on the nucleosome variant (Supplementary Fig. 3b,c and Supplementary Fig. 8c,d). Additionally, the nucleosomes were analyzed on a denaturing 15 % Tris-HCl polyacrylamide gel (200 V, 40 min), followed by Coomassie blue staining using the GelCode stain (Supplementary Fig. 3d). To form the library (DNL-1 and DNL-2), the barcoded nucleosomes (prepared in parallel) were combined at equimolar ratios and concentrated to a concentration of approximately 1 μM of total nucleosomes using Vivaspin 500 centrifugal filter units (10,000 Da molecular weight cut-off). The nucleosome library was analyzed by native gel (Fig. 1c). For storage, 20 % (v/v) glycerol and 0.5 mM phenylmethylsulfonyl fluoride (PMSF) were added, and the library was shock-frozen in aliquots and stored at − 80 °C. Typical yields for the nucleosome assembly step were ~ 50 %, mainly determined by the limiting amount of BC-601, which resulted in an overall yield of nucleosome assembly of about 30 %, starting from individual histones.

Analysis of DNL-1 stability towards DNA scrambling

In order to test whether DNA scrambling occurred among the nucleosome library members, DNL-1 (12 fmoles of total nucleosomes) were combined with the appropriate modification specific antibody (α-H3K4me3, α-H3K9ac, α-H4K8ac, Supplementary Table 4) in a total volume of 100 μL antibody (AB) buffer (20 mM Tris-HCl, pH 7.5, 50 mM NaCl, 5 mM EDTA) to give a final antibody concentration of 15 μg/mL. This mixture was incubated at RT for 1 h on a nutator, after which 100 μL AB buffer as well as 5 μL of protein G resin was added, and the mixture was incubated for a further 1 h at RT on a tube rotator. The beads were washed 4 times with 200 μL AB buffer. The DNA was eluted using 100 μL of DNA elution buffer (100 mM Tris-HCl, 10 mM EDTA, 1 % SDS, 10 mM 2-mercaptoethanol, 200 μg/mL proteinase K, pH 7.8) for 1.5 h at 50 °C and purified using the Qiagen PCR purification kit. The resulting DNA was quantified using the Qubit high sensitivity ds DNA quantification kit. For the subsequent multiplexing PCR step, the DNA was diluted with H_2O to a final concentration of 2 pg/μL in H_2O (the dilution factors of the respective experiment were considered later during sequencing data analysis, see below). Input samples (5–50 % of the initial library input, depending on the experiment) were processed identically. Each experiment was performed in independent triplicates, unless otherwise noted, and normalized against input and, within each experiment, the indicated variant (see Supplementary Fig. 4). Values are shown as ‘mean ± SD’. Data from experiments performed in duplicate were plotted individually.

BPTF, p300 and Brd4 binding pull-down experiments with DNL-1 and DNL-2

GST pull-down experiments were performed as described [10] with the following modifications. Glutathione Sepharose 4B beads (GE Healthcare, 3 μL) were incubated for 1.5 h at RT with an excess of purified GST-tagged BPTF or p300 construct (BPTF-PHD-BD, BPTF-PHD, BPTF-BD, p300-BD-PHD, p300-BD, or p300-PHD) in 200 μL GDP300 buffer (20 mM HEPES, pH 7.9, 300 mM KCl, 0.2 mM EDTA, 1 mM DTT, 0.01 % NP-40) on a tube rotator. For the Brd4 constructs (Brd4-BD1-BD2, Brd4-BD1, Brd4-BD2), 10 μL of anti-Flag M2 resin (Sigma) were used, and the GDP300 buffer contained 10 % (v/v) glycerol. After washing with 4 × 200 μL GDP300 buffer to remove unbound protein, the slurry was incubated with the nucleosome library (containing 0.125 pmoles of each of the nucleosome library member) in 200 μL GDP300 buffer (25 nM or 14 nM of total nucleosome concentration for DNL-1 or DNL-2, respectively) on an tube rotator for 4 h at 4 °C. After washing with 4 × 200 μL GDP300 buffer, the DNA was eluted as described above. To prepare DNA samples for library input normalization, 20 μL, i.e. 10 %, of the input nucleosome mixture were treated with the DNA elution buffer. Each pull-down was performed in triplicate (unless otherwise noted), normalized against input and, within each experiment, the indicated variant (see Supplementary Fig. 4). Values are shown as 'mean ± SD'. Data from experiments performed in duplicate were plotted individually.

Individual BPTF pull-down experiments

GST-BPTF pull-down experiments with individual nucleosome variants were performed similarly to those described above for the library pull-downs, but on a larger scale using 15 μL of glutathione sepharose 4B resin and 1 pmole of the respective nucleosome (unmodified, H3K4me3, H3K4me3-H4K16ac, or H3K4me3-H4Kac$_5$). Pull-down efficiency was assessed on a 5% acrylamide gel (0.5 × TBE buffer, 200 V, 40 min), followed by DNA staining with SybrGold, imaging on an ImageQuant (GE Healthcare), and quantification of the bands using the software ImageJ. Each pull-down was performed in triplicate, and the values of each replicate were normalized against the unmodified variant, set as 1. Values are shown as (mean ± SD).

p300 acetylation of the DNL followed by antibody immunoprecipitation

The nucleosome library (30 nM total, 12 fmoles of each member) was incubated with Flag-tagged p300 (3 nM) or variants thereof in the presence of 10 μM acetyl-CoA (Sigma) in 15 μL HAT buffer (50 mM Tris-HCl, 0.1 mM EDTA, 100 mM Na-butyrate, 10 % glycerol, 1 mM DTT, pH 8.0) for 1 h at 30 °C on a tube rotator. To isolate the reaction products, the reaction mixture was then added to 85 μL AB buffer containing 15 μg/mL of modification specific antibody (α-H2AK5ac, α-H3K9ac, α-H3K14ac, α-H3K18ac, α-H3K23ac, α-H3K27ac, α-H4K5ac, α-H4K8ac, α-H4K12ac, α-H4K16ac, Supplementary Table 4) and incubated for 1 h at RT. All subsequent steps were performed as described for the nucleosome stability experiments. Input samples (5–50 % of the nucleosome library, depending on the experiment) were processed identically. When Nω-Acetylhistamine (NAH) or the biotinylated tetra-acetylated H4 peptide (Bio-H4Kac$_4$) was included in the reaction, the acetyl-CoA concentration was increased to 200 μM. Additionally, Bio-H4Kac$_4$ was removed by streptavidin affinity purification with 100 μL MyOne streptavidin magnetic

beads for 30 min at 4 °C prior to antibody incubation. Each experiment was performed in triplicate and normalized against input and, within each experiment, the indicated variant (see Supplementary Fig. 4), unless otherwise noted. Values are shown as 'mean ± SD'. Data from experiments performed in duplicate were plotted individually.

p300 acetylation of individual nucleosomes

Individual nucleosome variants (4.8 pmoles) were incubated with wild-type p300 (0.12 pmoles) in the presence of 5 μM ^{3}H-acetyl-CoA (Perkin-Elmer) in 20 μL HAT buffer at 30 °C. The reactions were quenched after 0, 5, 15, and 45 min with 4x SDS sample buffer (final concentration: 50 mM Tris-HCl pH 6.8, 2 (w/v) % SDS, 10 (v/v) % glycerol, 1 (v/v) % 2-mercaptoethanol, and 12.5 mM EDTA, 0.02 (w/v) % bromophenol blue), and analyzed on a 15 % polyacrylamide Tris-HCl gel (Biorad) followed by protein staining using CBB and imaging. Additionally, a sample at the 45 min time-point was analyzed by native polyacrylamide gel electrophoresis (5 % acrylamide gel, 0.5 × TBE, 200 V, 40 min), followed by ethidium bromide staining and imaging. The gels were fixed (40 % EtOH, 10 % AcOH, 45 min, RT), incubated in Amplify solution (GE healthcare, 30 min, RT), and dried (2 h, 70 °C). ^{3}H-acetyl incorporation was assessed by fluorography using a light-sensitive film. The experiment was performed in triplicate.

p300 acetylation of individual nucleosomes followed by characterization using mass spectrometry

HAT assays were performed essentially the same as outlined above, but using non-radioactive acetyl-CoA for 60 min at 30 °C and a protocol described in ref. [50] with the following modifications. Propionylated histone tryptic peptides were desalted using StageTip micro-scale reversed-phase chromatography [51], and then subjected to reversed-phase nano-UPLC-MS and MS/MS performed on an Easy-nLC Ultra 1000 nanoflow capillary UPLC system (Proxeon, Odense, Denmark) coupled to a VelosPro-Orbitrap Elite hybrid mass spectrometer (ThermoFisher, Bremen, Germany) using a Flex ion source (Proxeon). Sample concentration and washing was performed online using a trapping capillary column (100 μm × ca. 40 mm, packed with 3 μm, 100 Å Magic AQ C18 resin (Bruker-Michrom, Auburn, CA)) at a flow rate of 4 μL/min for 9 min. Separation was achieved using an analytical capillary column (75 μm × ca. 20 cm, packed with 1.7 μm 100 Å BEH C18 resin (Waters, Billerica, MA)), and a linear gradient from 5% to 35% solutions A and B (solution A: 100 % water/0.1 % FA; solution B: 100 % ACN/0.1 % FA) applied over 90 min at a flow rate of approximately 250 nL/min. Nano-electrospray ionization was accomplished at 2.1 kV using a capillary temperature of 275 °C. Full-scan mass spectra were acquired in the Orbitrap in positive-ion mode over the m/z range of 335–1800 at a resolving power of 120,000. MS/MS spectra were acquired using collision-induced dissociation in the VelosPro linear trap for the 15 most abundant multiply charged species in the full-scan spectrum having signal intensities of >1000 NL. All spectra were acquired in profile mode. Dynamic exclusion was set such that MS/MS was acquired only once for each species over a period of 120 seconds. The resulting LC-MS/MS data were processed into peak-list files (mgf) using ProteomeDiscoverer (v. 1.4, ThermoFisher), and then searched against a database consisting of human histone H2A, H2B, H4 sequences, and that of the H3.2 variants, hH3.2 C110A and hH3.2 C96A; C110A using the Mascot search engine (v.

2.2.7, Matrix Science), allowing for a mass error of 6 ppm for precursor and 1.2 Da for fragment ion species, ≤ 8 missed trypsin cleavages, and variable modifications of methionine oxidation, lysine and peptide N-terminal propionylation, and lysine acetylation. Aggregate search results were collated and consolidated using Scaffold software (Proteome Software), according to the PeptideProphet [52, 53] parsimony algorithms and filtered to the 90 % peptide and 95 % protein probability levels. Ascore derived PTM site localization probabilities [54] were calculated for each PTM using the implementation of this algorithm in the ScaffoldPTM software (v. 2.0, Proteome Software). Fragmentation spectral assignments were subject to manual inspection and validation using the original tandem mass spectra acquired in profile mode using Xcalibur software (ThermoFisher).

Incubation of DNL-1 with nuclear extracts

DNL-1 (30 nM total, 12 fmoles of each member) was incubated with 1.9–7.5 μL nuclear extract, depending on the enzyme activity of the respective preparation (extracts from U2OS, 293T, HFF, MDA-MB, or NMuMG cells), in the presence or absence of 20 μM acetyl-CoA, 30 μM *S*-adenosyl methionine (SAM), and 10 μM adenosine triphosphate (ATP) in HAT buffer supplemented with 10 mM $MgCl_2$ (final total volume 15 μL) for 1 h at 30 °C. To isolate the reaction products, the reaction mixture was then added to 85 μL AB buffer containing 15 μg/mL of modification specific antibody (α-H4K14ac (MP07-353) or α-H3K4me1 (MP-07436)) and incubated for 1 h at RT. All subsequent steps were performed as described above. The experiment was performed in duplicate or triplicate, and the obtained sequencing reads were normalized against the input and within each experiment against the $H3Kac_5$ variant (see Supplementary Fig. 4). The final data is shown as 'mean ± SD'. Data from experiments performed in duplicate were plotted individually.

Acetylation and methylation of individual nucleosomes in nuclear extract

Western blot—Individual nucleosome variants (1.1 pmoles) were incubated with 6 μL of 293T nuclear extract in the presence of 20 μM acetyl-CoA or 30 μM SAM (Sigma) in HAT buffer (12 μL total volume) at 30 °C. The reactions were quenched after 60 mins with 4x SDS sample buffer and analyzed on a 4–20 % polyacrylamide Tris-HCl gel (Biorad) followed by Western blotting using a Trans-Blot® SD Semi-Dry Transfer Cell (BioRad) and Towbin buffer (25 mM Tris, 192 mM glycine, 20 % methanol (v/v), 0.025 % SDS, pH 8.3; PDVF membrane, 25 V, 45 min). The membrane was blocked with 3 % (w/v) milk/TBST (50 mM Tris, 150 mM NaCl, 0.05 % Tween-20) for 30 min, and incubated with the appropriate antibody (α-H3K14ac (MP07-353) or α-H3K4me1 (MP-07436)) in 3 % (w/v) milk/TBST (dilution: 1:1000) at 4 °C overnight. After incubation with an anti-rabbit IgG antibody (BioRad, 1:1000 dilution) for 45 min at RT followed by TBST washes (3 × 5 min) and treatment with ECL solution (Pierce), the chemiluminescent signal was recorded using an ImageQuant (GE Healthcare). To check for equal loading, the antibodies were stripped with stripping buffer (200 mM glycine, 3.5 % (w/v) SDS, 1 % (v/v) Tween-20, pH 2.2, 2 × 10 min), and the membrane was washed with PBS (2 × 10 min) and TBST (2 × 5 min). Antibody incubation with an anti-H3 antibody (ab1791, 1:2500 dilution) followed by anti-rabbit IgG was performed identically as described for the α-H4K14ac or α-H3K4me1 treatments.

Fluorography—Individual nucleosome variants (1.6 pmoles) were incubated with 9 μL of 293T nuclear extract in the presence of 3 μM (^{3}H)-acetyl-CoA (Perkin Elmer) in HAT buffer (18 μL total volume) at 30 °C. The reactions were quenched after 60 min with 4x SDS sample buffer and analyzed on a 15 % polyacrylamide Tris-HCl gel (Biorad) followed by protein CBB staining and fluorescence imaging. The gels were fixed (40 % EtOH, 10 % AcOH, 45 min, RT), incubated in Amplify solution (GE healthcare, 30 min, RT), and dried (2 h, 70 °C). ^{3}H-acetyl incorporation was assessed by fluorography using a light-sensitive film.

Preparation of standards A-D by PCR for DNA standard mixture

An internal barcoded DNA standard mixture was prepared in order to assess the fidelity of all steps subsequent to the pull-down experiments (i.e. the experiment multiplexing PCR, Ion Torrent emulsion-PCR, and Ion Torrent sequencing). This standard set was included in each multiplexing PCR experiment in order to check for any unwanted changes in the DNA ratios during DNA manipulations and sequencing (see below). The standard mixture contained four DNAs, A-D, and was produced by Phusion PCR using a 601 template [16] and the primers 5–8 (FW) and 9 (RV, Supplementary Table 5). The PCR products were purified using a Qiagen purification kit, eluted with 50 μL DNA buffer and quantified using the Qubit high sensitivity DNA quantification kit. Standards A-D were mixed in DNA buffer to a total DNA concentration of 0.2 pg/μL with the following distribution: 1 eq. standard A, 10 eq. standard B, 100 eq. standard C, and 1,000 eq. standard D.

Multiplex PCR and addition of Ion Torrent A and P1 adaptors

In a typical multiplex PCR reaction, DNA (9 pg) isolated from a ChIP experiment was combined with 1 pg of the internal standard mixture (containing standards A-D, see above) in the presence of 0.01 U/μL Phusion DNA polymerase, 0.2 mM dNTPs using 0.5 μM of primers 10 (FW) and 11 (RV, Supplementary Table 5; for a list of multiplex barcodes, see Supplementary Table 2). PCR cycling conditions were chosen to ensure the reaction remained in the exponential phase (PCR cycle program: initial denaturation, 30 s/98 °C; denaturation, 10 s/98 °C; annealing, 15 s/62 °C; extension, 5 s/72 °C; 15 cycles total; final extension, 7 min/72°C). The PCR products were purified using a Qiagen PCR purification kit and eluted with 50 μL DNA buffer. The multiplexed DNA sequences were pooled at equal volumes and subjected to Ion Torrent sequencing. In the experiment shown in Supplementary Fig. 5a, the multiplexing PCR reactions were performed with 10 pg of standard in the absence of any pulldown DNA. In this particular case, the raw DNA sequencing reads were averaged from three independent experiments and plotted against the equivalents (eq.) of the respective barcoded DNA standard sequences.

Ion Torrent sequencing

Ion Torrent sequencing was performed on the Personal Genome Machine (PGM) at the Sequencing Core Facility at Princeton University according to the Ion Torrent protocol (Life Technologies). The multiplexed DNA library from the multiplex PCR was diluted to 30 pM and subjected to emulsion PCR using the OneTouch 200 Template Kit v2 DL (Invitrogen). Sequencing primers and DNA polymerases were added to the final enriched Ion spheres and

loaded onto a 314 or 316 chip, depending on the experiment. Base calls were generated using Ion Torrent Suite version 3.2.1 and the resulting FASTQ files were used for further analysis.

Ion Torrent data analysis

The Ion Torrent sequencing data was de-multiplexed using the barcode splitter tool Fastx Toolkit (hannonlab.cshl.edu/fastx toolkit/ind3ex.html). The DNA reads were first sorted according to the list of experimental multiplexing barcodes located at the 3′-end of the sequences (Supplementary Table 2). Subsequently, they were sorted according to the list of nucleosome barcodes located at the 5′-end of the sequences (Supplementary Table 1 and 3 for DNL-1 and DNL-2, respectively), followed by multiplying the experiment-specific and nucleosome variant-specific reads by the dilution factor (see above). Finally, the sequencing reads for each variant were normalized against those of the respective nucleosome variants from the input sequencing data (typically averaged from two input samples). This procedure resulted in nucleosome variant-specific sequencing reads expressed as 'Pulldown efficiency (% IN)' for each protein or antibody pull-down experiment. If no further normalization was performed, the reads were averaged from 3 independent experiments, unless otherwise noted, and the values were expressed as (mean ± SD) 'Pull-down efficiency (% IN)'. In all other cases, the reads within each experiment were further normalized towards one specific nucleosome variant (which was set as 1, as indicated in the respective figure legends). These normalized values were plotted individually (n=3) or averaged from 3 independent experiments and expressed as (mean ± SD) 'Normalized reads (a.u.)'. For more information, see Supplementary Fig. 4. Note that the internal PCR standard (see above) was used to assess the linearity of the PCR and sequencing experiment. In cases where the expected ratio (1:10:100:1000) of the barcoded standards was not observed, the data was discarded (observed in less than 5 % of the samples).

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

The authors thank the current and former members of the Muir laboratory for many valuable discussions. We further thank S. A. Blythe and S. Josefowicz for critically reading the manuscript. We would like to thank W. Wang and D. Storton (High Throughput Sequencing and MicroArray Facility, Princeton University) for Ion Torrent Sequencing, D. H. Perlman (Princeton Proteomics & Mass Spectrometry Core, Princeton University) for the mass spectrometry data, the Roeder lab (Rockefeller University) for providing the Flag-p300 plasmid, the Allis lab (Rockefeller University) for the BPTF plasmid, the Kraus lab (UT Southwestern) for the Flag-ΔBD-p300 plasmid, the Kang lab (Princeton University) for providing the MDA-MB and NMuMG cells, and the Cristea lab (Princeton University) for providing the U2OS and HFF cells. This research was supported by the U.S. National Institutes of Health (grants R37-GM086868 and R01 GM107047). U.T.T.N. was funded by a postdoctoral fellowship of the Deutsche Forschungsgemeinschaft. Funding by the Swiss National Science foundation (postdoctoral fellowships to M.M.M. and B.F.) is gratefully acknowledged.

References for Main Text

1. Badeaux AI, Shi Y. Emerging roles for chromatin as a signal integration and storage platform. Nat Rev Mol Cell Biol. 2013; 14:211–224.

2. Suzuki MM, Bird A. DNA methylation landscapes: provocative insights from epigenomics. Nat Rev Genet. 2008; 9:465–476. [PubMed: 18463664]
3. Kouzarides T. Chromatin modifications and their function. Cell. 2007; 128:693–705. [PubMed: 17320507]
4. Taverna SD, Li H, Ruthenburg AJ, Allis CD, Patel DJ. How chromatin-binding modules interpret histone modifications: lessons from professional pocket pickers. Nat Struct Mol Biol. 2007; 14:1025–1040. [PubMed: 17984965]
5. Helin K, Dhanak D. Chromatin proteins and modifications as drug targets. Nature. 2013; 502:480–488. [PubMed: 24153301]
6. ENCODE Project Consortium et al. An integrated encyclopedia of DNA elements in the human genome. Nature. 2012; 489:57–74. [PubMed: 22955616]
7. Fierz B, Muir TW. Chromatin as an expansive canvas for chemical biology. Nat Chem Biol. 2012; 8:417–427. [PubMed: 22510649]
8. Hung T, et al. ING4 mediates crosstalk between histone H3 K4 trimethylation and H3 acetylation to attenuate cellular transformation. Mol Cell. 2009; 33:248–256. [PubMed: 19187765]
9. Bartke T, et al. Nucleosome-interacting proteins regulated by DNA and histone methylation. Cell. 2010; 143:470–484. [PubMed: 21029866]
10. Ruthenburg AJ, et al. Recognition of a mononucleosomal histone modification pattern by BPTF via multivalent interactions. Cell. 2011; 145:692–706. [PubMed: 21596426]
11. McGinty RK, Kim J, Chatterjee C, Roeder RG, Muir TW. Chemically ubiquitylated histone H2B stimulates hDot1L-mediated intranucleosomal methylation. Nature. 2008; 453:812–816. [PubMed: 18449190]
12. Mannocci L, Leimbacher M, Wichert M, Scheuermann J, Neri D. 20 years of DNA-encoded chemical libraries. Chem Commun. 2011; 47:12747–12753.
13. Ullal AV, et al. Cancer cell profiling by barcoding allows multiplexed protein analysis in fine-needle aspirates. Sci Transl Med. 2014; 6:219ra9.
14. Filippakopoulos P, Knapp S. The bromodomain interaction module. FEBS Lett. 2012; 586:2692–2704. [PubMed: 22710155]
15. Muir TW. Semisynthesis of proteins by expressed protein ligation. Annu Rev Biochem. 2003; 72:249–289. [PubMed: 12626339]
16. Lowary PT, Widom J. New DNA sequence rules for high affinity binding to histone octamer and sequence-directed nucleosome positioning. J Mol Biol. 1998; 276:19–42. [PubMed: 9514715]
17. Dyer PN, et al. Reconstitution of nucleosome core particles from recombinant histones and DNA. Methods Enzymol. 2004; 375:23–44. [PubMed: 14870657]
18. Rothberg JM, et al. An integrated semiconductor device enabling non-optical genome sequencing. Nature. 2011; 475:348–352. [PubMed: 21776081]
19. Delvecchio M, Gaucher J, Aguilar-Gurrieri C, Ortega E, Panne D. Structure of the p300 catalytic core and implications for chromatin targeting and HAT regulation. Nat Struct Mol Biol. 2013; 20:1040–1046. [PubMed: 23934153]
20. Delmore JE, et al. BET bromodomain inhibition as a therapeutic strategy to target c-Myc. Cell. 2011; 146:904–917. [PubMed: 21889194]
21. Filippakopoulos P, et al. Histone recognition and large-scale structural analysis of the human bromodomain family. Cell. 2012; 149:214–231. [PubMed: 22464331]
22. Schiltz RL, et al. Overlapping but distinct patterns of histone acetylation by the human coactivators p300 and PCAF within nucleosomal substrates. J Biol Chem. 1999; 274:1189–1192. [PubMed: 9880483]
23. Li S, Shogren-Knaak MA. The Gcn5 bromodomain of the SAGA complex facilitates cooperative and cross-tail acetylation of nucleosomes. J Biol Chem. 2009; 284:9411–9417. [PubMed: 19218239]
24. Ragvin A, et al. Nucleosome binding by the bromodomain and PHD Finger of the transcriptional cofactor p300. J Mol Biol. 2004; 337:773–788. [PubMed: 15033350]

25. Kraus WL, Manning ET, Kadonaga JT. Biochemical analysis of distinct activation functions in p300 that enhance transcription initiation with chromatin templates. Mol Cell Biol. 1999; 19:8123–8135. [PubMed: 10567538]
26. Calo E, Wysocka J. Modification of enhancer chromatin: what, how, and why? Mol Cell. 2013; 49:825–837. [PubMed: 23473601]
27. LeRoy G, et al. A quantitative atlas of histone modification signatures from human cancer cells. Epigenetics Chromatin. 2013; 6:20. [PubMed: 23826629]
28. Tang Z, et al. SET1 and p300 act synergistically, through coupled histone modifications, in transcriptional activation by p53. Cell. 2013; 154:297–310. [PubMed: 23870121]
29. Kim J, et al. The n-SET domain of Set1 regulates H2B ubiquitylation-dependent H3K4 methylation. Mol Cell. 2013; 49:1121–1133. [PubMed: 23453808]
30. Kim DH, et al. Histone H3 K27 trimethylation inhibits H3 binding and function of SET1-like H3K4 methyltransferase complexes. Mol Cell Biol. 2013; 33:4936–4946. [PubMed: 24126056]
31. Nakagawa T, et al. Deubiquitylation of histone H2A activates transcriptional initiation via trans-histone cross-talk with H3K4 di- and trimethylation. Genes & Dev. 2008; 22:37–49. [PubMed: 18172164]
32. Garske AL, et al. Combinatorial profiling of chromatin binding modules reveals multisite discrimination. Nat Chem Biol. 2010; 6:283–290. [PubMed: 20190764]
33. Rothbart SB, Krajewski K, Strahl BD, Fuchs SM. Peptide microarrays to interrogate the “histone code”. Methods Enzymol. 2012; 512:107–135. [PubMed: 22910205]
34. Sandoval J, Peiró-Chova L, Pallardó FV, García-Giménez JL. Epigenetic biomarkers in laboratory diagnostics: emerging approaches and opportunities. Expert Rev Mol Diagn. 2013; 5:457–471. [PubMed: 23782253]
35. Dignam JD, Lebovitz RM, Roeder RG. Accurate transcription initiation by RNA polymerase II in a soluble extract from isolated mammalian nuclei. Nucleic Acids Res. 1983:11, 1475–1489. [PubMed: 6306560]
36. Wu SY, Lee AY, Lai HT, Zhang H, Chiang CM. Phospho Switch Triggers Brd4 Chromatin Binding and Activator Recruitment for Gene-Specific Targeting. Mol Cell. 2013; 49:843–857. [PubMed: 23317504]
37. Manning ET, Ikehara T, Ito T, Kadonaga JT, Kraus WL. p300 forms a stable, template-committed complex with chromatin: role for the bromodomain. Mol Cell Biol. 2001; 21:3876–3887. [PubMed: 11359896]
38. An W, Kim J, Roeder RG. Ordered cooperative functions of PRMT1, p300, and CARM1 in transcriptional activation by p53. Cell. 2004; 117:735–748. [PubMed: 15186775]
39. Chatterjee C, McGinty RK, Fierz B, Muir TW. Disulfide-directed histone ubiquitylation reveals plasticity in hDot1L activation. Nat Chem Biol. 2010; 6:267–269. [PubMed: 20208522]
40. Fierz B, Kilic S, Hieb AR, Luger K, Muir TW. Stability of nucleosomes containing homogenously ubiquitylated H2A and H2B prepared using semisynthesis. J Am Chem Soc. 2012; 134:19548–19551. [PubMed: 23163596]
41. McGinty RK, et al. Structure–activity analysis of semisynthetic nucleosomes: mechanistic insights into the stimulation of Dot1L by ubiquitylated histone H2B. ACS Chem Biol. 2009; 4:958–968. [PubMed: 19799466]
42. Canosa JB, Dawson PE. An efficient Fmoc-SPPS approach for the generation of thioester peptide precursors for use in native chemical ligation. Angew Chem Intl Edit. 2008; 47:6851–6855.
43. Fang GM, Li YM, Shen F, Huang YC. Protein chemical synthesis by ligation of peptide hydrazides. Angew Chem Intl Edit. 2011; 50:7645–7659.
44. Fierz B, et al. Histone H2B ubiquitylation disrupts local and higher-order chromatin compaction. Nat Chem Biol. 2011; 7:113–119. [PubMed: 21196936]
45. Hackeng TM, Griffin JH, Dawson PE. Protein synthesis by native chemical ligation: expanded scope by using straightforward methodology. Proc Natl Acad Sci USA. 1999; 96:10068–10073. [PubMed: 10468563]
46. Grieco P, Gitu PM, Hruby VJ. Preparation of ‘side-chain-to-side-chainco, Preparation of ‘sidehesis by native chemical. J Pep Res. 2001; 57:250–256.

47. Geiermann AS, Micura R. Selective desulfurization significantly expands sequence variety of 3′-peptidyl-tRNA mimics obtained by native chemical ligation. Chembiochem. 2012; 13:1742–1745. [PubMed: 22786696]

48. Dyer PN, Edayathumangalam RS, White CL, Bao Y, Chakravarthy S, Muthurajan UM, Luger K. Reconstitution of nucleosome core particles from recombinant histones and DNA. Methods Enzymol. 2003; 375:23–44. [PubMed: 14870657]

49. Flaus A, Richmond TJ. Positioning and stability of nucleosomes on MMTV 3′ LTR sequences. J Mol Biol. 1998; 275:427–441. [PubMed: 9466921]

50. Garcia BA, Mollah S, Ueberheide BM, Busby SA. Chemical derivatization of histones for facilitated analysis by mass spectrometry. Nat Protoc. 2007; 2:933–938. [PubMed: 17446892]

51. Rappsilber J, Ishihama Y, Mann M. Stop and go extraction tips for matrix-assisted laser desorption/ionization, nanoelectrospray, and LC/MS sample pretreatment in proteomics. Anal Chem. 2003; 75:663–670. [PubMed: 12585499]

52. Nesvizhskii AI, Kolker E, Aebersold R. Empirical statistical model to estimate the accuracy of peptide identifications made by MS/MS and database search. Anal Chem. 2002; 74:5383–5392. [PubMed: 12403597]

53. Nesvizhskii AI, Keller A, Kolker E. A statistical model for identifying proteins by tandem mass spectrometry. Anal Chem. 2003; 75:4646–4658. [PubMed: 14632076]

54. Beausoleil SA, Villén J, Gerber SA, Rush J, Gygi SP. A probability-based approach for high-throughput protein phosphorylation analysis and site localization. Nat Biotechnol. 2006; 24:1285–1292. [PubMed: 16964243]

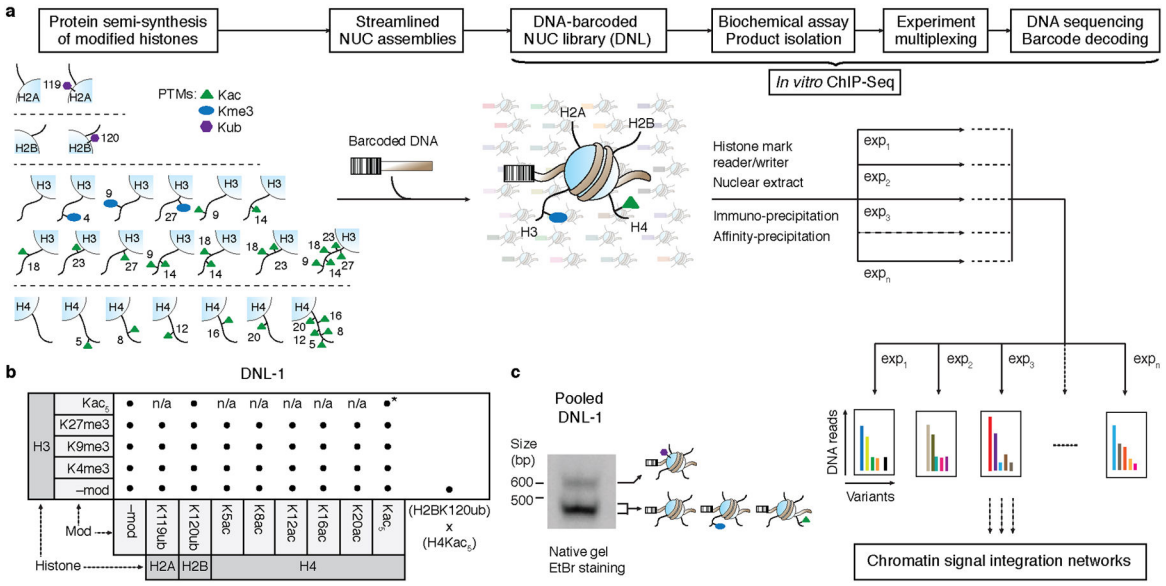

Figure 1.

Preparation of DNL and its use in *in vitro* ChIP-Seq experiments. (**a**) Modified histone variants prepared by protein semi-synthesis are assembled with the respective barcoded DNA into a barcoded nucleosome (‘NUC’) library (‘DNL’). After biochemical assays with a writer, reader or nuclear extract, the binders and reaction products are isolated by affinity- or immunoprecipitation, followed by DNA experiment multiplexing. NUC identity and abundance is analyzed by next generation sequencing (NGS). (**b**) Combinations of histone modifications (‘mod’) selected for the first version of the library (‘DNL-1’). Unmodified (‘–mod’) or modified H3 proteins (vertical axis) were combined with otherwise unmodified histones (‘–mod’), H2AK119ub, H2BK120ub, or mono-/hyperacetylated H4 (horizontal axis). Additionally, a NUC bearing H2BK120ub and H4Kac_5 was prepared. Asterisk: this variant was employed in the Brd4 experiment (Fig. 2c). (**c**) Analysis of the combined DNL-1 by native gel electrophoresis and ethidium bromide (EtBr) DNA staining. The bands from NUCs containing combinations of unmodified, Kac or Kme3 histones overlap, whereas the shifted fainter upper band represents NUCs containing ubiquitylated H2A or H2B. For details on modified histones, DNA preparation, NUC assembly and NGS, see Supplementary Figs. 1–4.

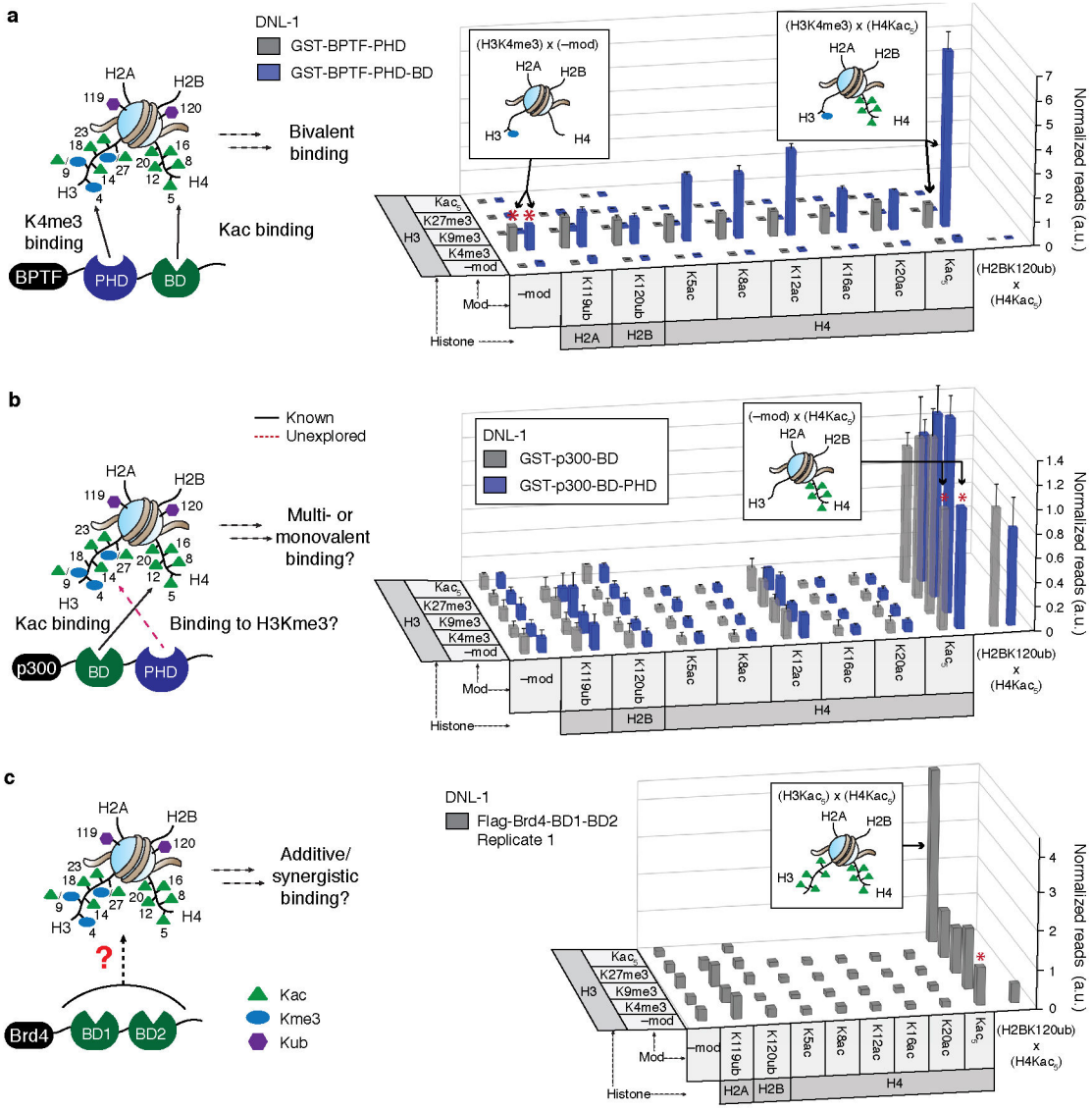

Figure 2.

Profiling the substrate specificity of histone mark readers. (**a**) Left: Model showing bivalent nucleosomal recognition behavior of BPTF. Right: Immobilized GST-tagged BPTF-PHD (gray) and BPTF-PHD-BD (blue) were incubated with DNL-1, followed by DNA isolation and experiment multiplexing. The resulting DNA was analyzed by next generation sequencing, and the raw sequencing reads were normalized against input and the H3K4me3-containing variant, set as 1 (red asterisks). Values (mean ± SD, n=3) for the respective NUC variants are plotted as a 3D bar graph using the same grid as shown in Fig. 1b. (**b**) Left: Model depicting potential binding modes of p300. Right: Library incubation of resin-bound GST-tagged p300-BD (gray) and p300-BD-PHD (blue). All subsequent steps were performed as in (**a**). Internal normalization: $H4Kac_5$ variant, set as 1 (red asterisks). Values are shown as (mean ± SD, n=3). (**c**) Left: Model depicting potential interaction between Brd4 and acetylated NUC. Right: Library incubation of resin-bound Flag-Brd4-BD1-BD2.

All subsequent steps were performed as in (**a**) except for pulldown with an anti-Flag antibody. Internal normalization: H4Kac$_5$ variant, set as 1 (red asterisk). Experiment was performed in duplicate (replicate shown in Supplementary Fig. 8b). See Source Data Table 1 for input-normalized sequencing reads.

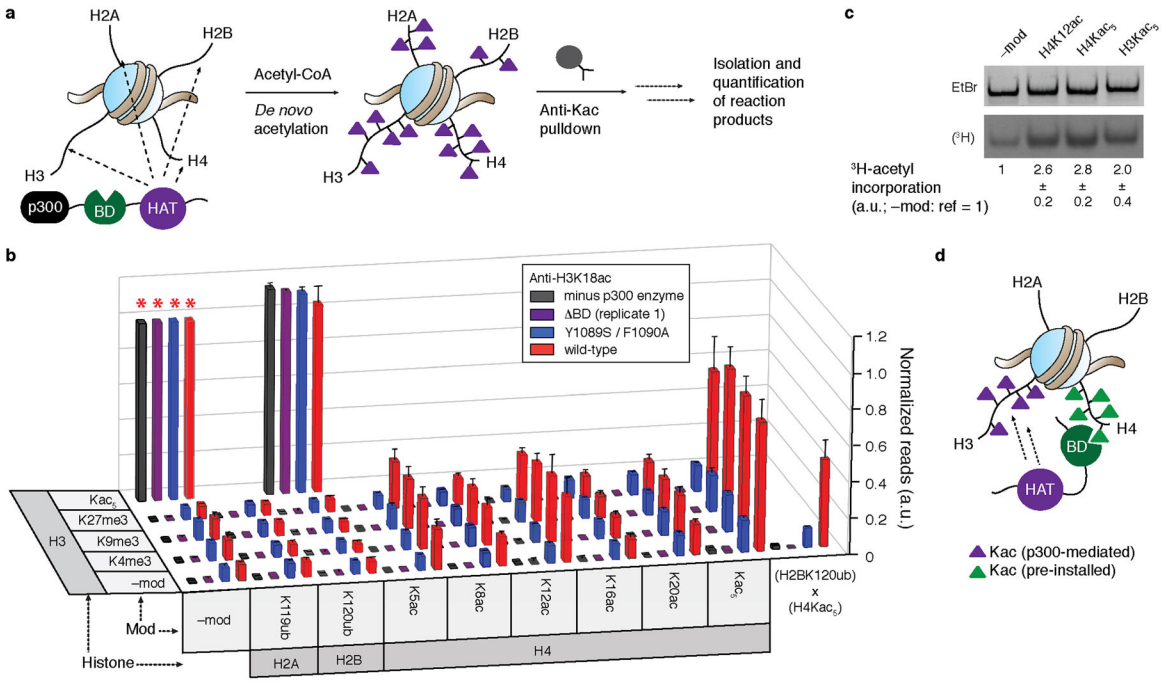

Figure 3.

Histone acetylation signaling is propagated through a positive feedback loop. (**a**) Schematic overview of the *in vitro* ChIP-Seq p300 enzyme assay. The library (exemplified by one unmodified variant) is incubated with recombinant p300 and acetyl-CoA, followed by ChIP-Seq with a site-specific antibody. (**b**) 3D bar plot of sequencing results for HAT assays, employing α-H3K18ac for the ChIP, in the absence (gray) or presence of p300 enzyme (purple: ΔBD; blue: Y1089AF1090S mutant; red: wt). Internal normalization: $H3Kac_5$ variant, set as 1 (red asterisks). Values are shown as (mean ± SD, n=3), except for ΔBD, which was performed in duplicate (replicate shown in Supplementary Fig. 11c). (**c**) p300 experiments employing individual NUCs and ^{3}H-acetyl-CoA. The extent of ^{3}H-acetyl incorporation was determined by native gel electrophoresis and EtBr staining (top) followed by fluorography (bottom). Band intensity of the unmodified NUC was set as 1, and values are shown as (mean ± SD, n=3). (**d**) Model for p300's positive feedback loop resulting from BD-mediated recruitment to hyperacetylated NUCs.

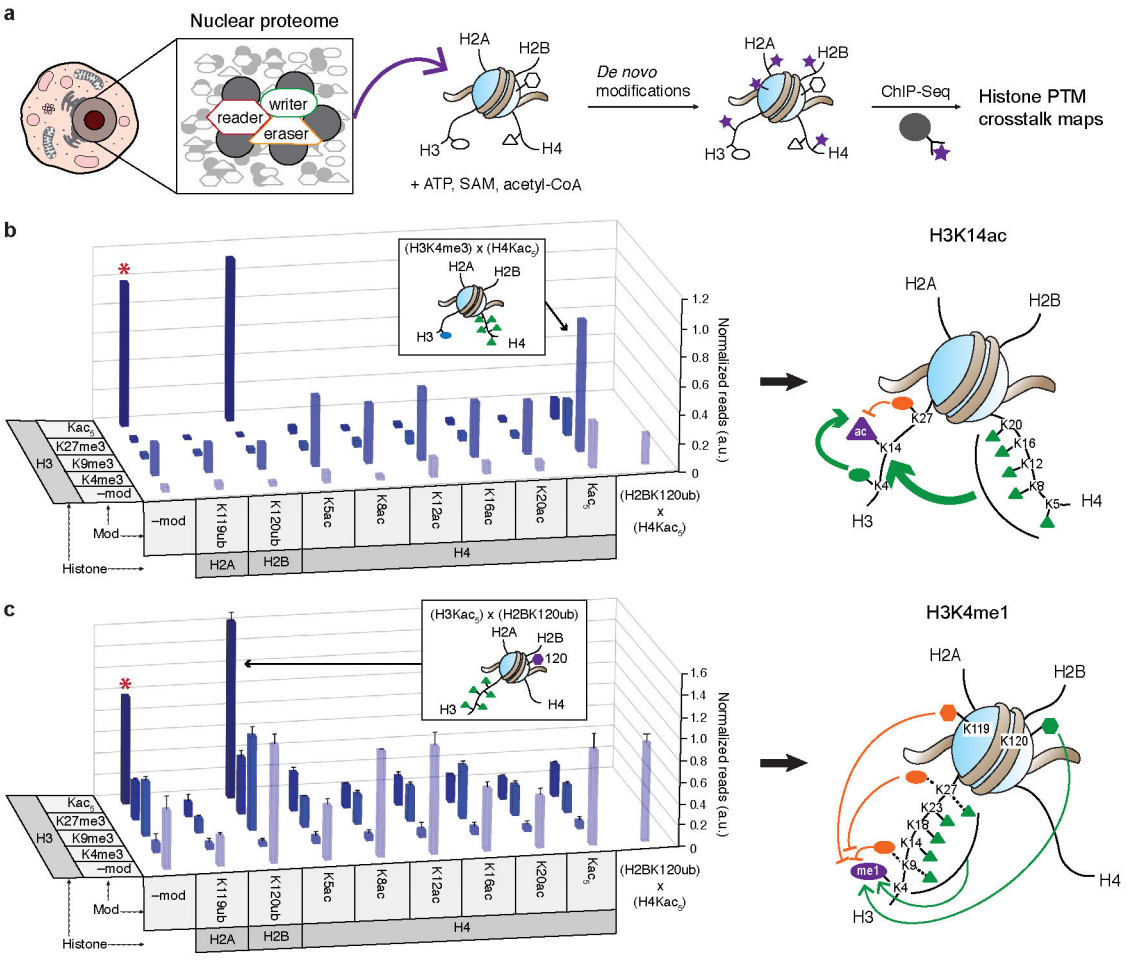

Figure 4. Systems-level synergistic modulation of histone PTMs

(**a**) DNL-1 (exemplified by one modified variant) is exposed to the entire nuclear proteome, containing complexes of readers, writers, erasers and other accessory proteins (gray circles), in the presence of ATP, SAM and acetyl-CoA. The extent of *de novo* chromatin modification (purple stars) is analyzed by *in vitro* ChIP-Seq as described in Fig. 3a. (**b,c**) ChIP-Seq results employing U2OS cell nuclear extracts and either α-H3K14ac (**b**) or α-H3K4me1 (**c**) antibodies. Internal normalization: $H3Kac_5$ variant (red asterisks), set as 1. Data in panel (**b**) were performed in duplicate (replicate shown in Supplementary Fig. 13a). Values shown in panel c represent (mean ± SD, n=3). Different shades of blue were only used for illustration purposes. For AB specificity, see Supplementary Figs. 10c and 14d. Histone crosstalk maps based on the respective results are shown on the right. The pre-installed modified lysine residues on H3 and H4 that stimulate or repress the deposition of H3K14ac (**b**, purple triangle) or H3K4me1 (**c**, purple oval) are colored in green or orange, respectively. Dotted black lines indicate pre-modification with either Kac or Kme3. The thickness of the arrows represents qualitatively the strength of the observed crosstalk.