Published in final edited form as:

Nature. 2023 April ; 616(7957): 574–580. doi:10.1038/s41586-023-05914-y.

Tracking chromatin state changes using nanoscale photo-proximity labeling

Ciaran P. Seath[1,2,3,*], **Antony J. Burton**[2,4,*], **Xuemeng Sun**[2], **Gihoon Lee**[2], **Ralph E. Kleiner**[2], **David W. C. MacMillan**[1,2,†], **Tom W. Muir**[2,†]

[1]Merck Center for Catalysis at Princeton University, Princeton, New Jersey 08544, USA.

[2]Department of Chemistry, Princeton University, Princeton, New Jersey 08544, USA.

[3]Present address: Department of Chemistry, Scripps-UF, Jupiter, Florida, FL, 33518.

[4]Present address: Discovery Biology, Discovery Sciences, Biopharmaceuticals R&D, AstraZeneca, Waltham, MA 02451.

Abstract

Interactions between biomolecules underlie all cellular processes, and ultimately control cell fate. Perturbation of native interactions through mutation, changes in expression levels, or external stimuli leads to altered cellular physiology and can result in either disease or therapeutic effects.[1,2] Mapping these interactions and determining how they respond to stimulus is the genesis of many drug development efforts, leading to new therapeutic targets and improvements in human health.[1] However, in the complex environment of the nucleus it is challenging to determine protein-protein interactions due to low abundance, transient or multi-valent binding, and a lack of technologies that are able to interrogate these interactions without disrupting the protein binding surface under study.[3] Here, we describe a method for the traceless incorporation of Ir-photosensitizers into the nuclear microenvironment using engineered split inteins. These Ir-catalysts can activate diazirine warheads via Dexter energy transfer to form reactive carbenes within a ~10 nm radius, cross-linking with proteins within the immediate microenvironment (a process termed μMap) for analysis via quantitative chemoproteomics.[4] Additionally, we show that this nanoscale proximity labeling method can reveal the critical changes in interactomes in the presence of cancer-associated mutations, as well as treatment with small-molecule inhibitors. μMap improves our fundamental understanding of nuclear protein-protein interactions, and in doing so is expected

[†]Correspondence and requests for materials should be addressed to D.W.C.M (dmacmill@princeton.edu), and T.W.M (muir@princeton.edu).

[*]Authors contributed equally.

Author Contributions. CPS, AJB, TWM, DWCM conceived the work. CPS, AJB, XS, GL, REK, TWM, DWCM designed and executed the experiments. CPS, AJB, TWM, DWCM prepared this manuscript.

Supplementary Information. Supplementary information is linked to the online version of the paper at www.nature.com/nature

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

Competing interests

A provisional U.S. patent has been filed by DWCM and CPS based on photocatalysts used in this work, 62/982,366; 63/076,658. International Application No. PCT/US2021/019959. DWCM declares an ownership interest, and CPS declares an affiliation interest, in the company Dexterity Pharma LLC, which has commercialized materials used in this work.

The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper.

to have a significant impact on the field of epigenetic drug discovery in both academia and industry.

Mapping protein-protein interactions (PPIs) is central to our understanding of cellular biology.[5] The enormous challenges associated with this undertaking are magnified in the nucleus where transient and multi-valent interactions, fine-tuned by post-translational modifications (PTMs), combine to choreograph DNA-templated processes such as transcription.[6] Perturbations to these regulatory mechanisms often leads to disease,[7] for example, somatic mutations that alter the composition and activity of chromatin-associated protein complexes are implicated in many human cancers and developmental disorders.[8 9,10] Moreover, recent studies have shown that histone proteins are themselves frequently mutated in cancers.[11-13] Understanding how these mutations lead to, or perpetuate disease, is the focus of intense investigation,[11-13] work that necessitates accurate comparative mapping of chromatin-associated PPI networks as a function of altered cell states.[1]

The elucidation of nuclear PPIs has typically been performed by immunoprecipitation/mass spectrometry (IP/MS) workflows, where antibodies recognizing selected proteins are used to enrich their target along with direct interactors.[14] However, IP/MS approaches rely on nuclear lysates as input, which may not be ideal for every system, [15,16] especially when the interactions are transient in nature (e.g. PTM-driven) or require multi-protein complexes that bridge native chromatin.[17,18,19] This has fuelled the development of chemo-proteomics approaches such as those based on photocrosslinking[20,21] or proximity labeling[22-24] technologies that seek to capture PPIs in a native-like environment. Despite these ongoing advances, no single method exists to map chromatin interactomes in a general and unbiased manner, and particularly to discern how such interactions are affected by perturbations such as mutation or drug treatment (Figure 1a). To address this need we envisioned the union of technologies recently disclosed by our two laboratories, µMap[4] and *in nucleo* protein *trans*-splicing,[20] to enable a traceless, short-range proximity labeling method that can be readily deployed to any nuclear protein target. Protein *trans*-splicing using ultra-fast split inteins facilitates the installation of Ir-photocatalysts onto the N- or C-termini of target proteins. Upon irradiation with blue LEDs in the presence of a biotin-diazirine probe, localized carbene generation (through Dexter energy transfer from photoexcited Ir to aryl diazirines) allows interactomes to be determined, specifically those within ~10 nm of the iridium-centered photocatalyst (Figure 1b).

Our proposed workflow offers distinct advantages for the elucidation of subtly perturbed chromatin interactomes through mutation or ligand binding. Principally, the short radius afforded by this technology limits labelling to the close vicinity of a designated chromatin factor or nucleosome, only identifying proteins that are affected by a particular mutation or pharmacological intervention. This is critically important given the structural and functional heterogeneity of chromatin.[25] Furthermore, the incorporation of the µMap catalyst is designed to be almost traceless and its small size is expected to minimize disruption to the native environment (e.g. in comparison to fusion proteins), allowing the study of modifications that only minimally change the native interactome.

We began our studies by preparing an Ir-conjugated intein fragment (Cfa^{C}-Ir), using a combination of solid phase peptide synthesis (SPPS) and click chemistry (Figure 2a-b). We next fused the N-terminal fragment of the engineered Cfa split intein (Cfa^{N})[26,27] to the C-terminus of histone H3.1. For analytical convenience, we also included HA and FLAG epitope tags flanking the intein (Figure 2c). We then treated nuclei isolated from these transfected cells with the complementary split intein fragment (Cfa^{C}) linked to the Ir photocatalyst. This resulted in the site-selective incorporation of the photocatalyst onto the C-terminus of H3.1 through *in nucleo* protein trans-splicing (Figure 2e). Irradiation of these nuclei in the presence of the diazirine-biotin probe (Dz-Bt, Figure 2d) led to dramatically enhanced protein labeling compared to control reactions that excluded Cfa^{C}-Ir or Bt-Dz (Figure 2f). Importantly, elution from streptavidin beads showed strong enrichment of histone H3 by western blot (Figure 2f).

Encouraged by this, we established a tandem mass tag (TMT)-based quantitative chemo-proteomics workflow to determine the interactome for H3.1 versus the centromere-specific H3 variant CENP-A (Figure 2g-h, Extended Data Figure 1) to assess whether our workflow could delineate between nucleosomal interactomes at specific regions of chromatin. We used cut-offs of > 0.5 Log_2FC and FDR-corrected p-value < 0.05, and hits were found in all replicates. Note, additional analyses of the proteomic datasets in this study are provided in Extended Data Figure 10. Comparison of the H3.1 vs CENP-A interactomes returned established H3.1 modifying enzymes (e.g. EHMT2, SUV39H1) and reader proteins (e.g. HP1 isoforms) (Figure 2b, Extended Data Figure 10, see SI for full data table).[28] CENP-A interacting proteins included both members of the FACT complex (SSRP1 & SPT16), involved in the deposition of CENP-A into chromatin (Figure 2b).[30] Consistent with the centromeric localization of CENP-A, we see enrichment of transcriptional regulators (CBX6, DNMT3A, CHD4, RBBP4, KAT7, KMT2A) only in the H3.1 samples. Furthermore, GO analysis of the H3.1 hits showed strong enrichment of chromatin organization ($-Log_{10}P = 56$), histone lysine methylation ($-Log_{10}P = 11$), and epigenetic regulation of gene expression ($-Log_{10}P = 25$) (Extended Data Figure 1), consistent with known role of H3.1 as a platform for transcriptional regulation.

With the basic chemo-proteomics workflow established we questioned whether our method would be able to sense subtle changes in nucleosomal structure, such as somatic mutations. Recent sequencing of patient tumour samples identified >4,000 histone mutations associated with a wide range of cancers.[11-13,31] Determining whether such mutations drive oncogenesis and cancer progression, or are simply passenger mutations, is critical to identifying novel opportunities for therapeutic intervention. We sought to apply our method to the cancer-associated histone mutation H2A E92K, which is correlated with a range of cancers.[13] This mutation introduces a charge swap within the critical acidic patch interaction motif on the nucleosome, into which arginine residues of interacting proteins are known to anchor.[32-35] We hoped that our methodology could shed light on to what extent chromatin PPIs are perturbed by this oncohistone mutation.

We found the local chromatin microenvironment is indeed sensitive to the H2A E92K mutation (Figure 3a, Extended Data Figure 1, see SI for full data table). GO analysis revealed diminished enrichment of proteins related to chromatin modification and

reorganization as a function of the perturbed acidic patch (Figure 3b). Comparison of the dataset to the interactome of a construct with a triply mutated acidic patch (E61A, E90A, E92A) suggests that E92K impacts a specific subset of acidic patch binding proteins (Extended Data Figure 2).

Intriguingly, SIRT6, a deacetylase, was highly enriched in the wild-type sample, whereas the transcriptional activators, BRD2/3/4 were all enriched by E92K. Taken together, this suggests the mutation acts to block histone lysine deacetylase activity, resulting in increased local acetylation and binding of associated reader proteins. To support this hypothesis, we transiently expressed both wild-type and H2A E92K histones in HEK293T cells and isolated intact mononucleosomes by anti-FLAG immunoprecipitation (Figure 3c). We found global lysine acetylation was increased 2-fold in the mutant nucleosomes over wild-type, with a particular increase observed on histone H4. This effect is striking given that we are likely only enriching one copy of the mutant histone per nucleosome due to presumed stochastic incorporation (Extended Data Figure 3).[36,37]

Following this, we questioned whether this local increase in acetylation would lead to a global effect by weakening DNA/nucleosome interactions, leading to a more accessible chromatin environment and consequently increasing the concentration of chromatin open reading frames. To probe this, we performed ATAC-seq analysis[38] on HEK293T cell lines stably expressing either wild-type H2A or the E92K mutant (Extended Data Figure 4). Consistent with our hypothesis, we found 5699 ORFs that showed a significant (FDR <= 0.05) change in concentration across four replicates. Of these, 90% indicate regions of higher accessibility in E92K in comparison to the wild-type H2A (Figure 3d&e, Extended Data Figure 5), consistent with our proposed model.

We also noted that installation of E92K had a negative effect on nucleosomal DNMT3A binding (Figure 3a). Recent structural and biochemical work has demonstrated that DNMT3 binds to the acidic patch through two arginine residues (R740 & R743 in DNMT3B).[39] Charge reversal mutants (e.g. R740E) of these residues were shown to diminish nucleosomal binding and de novo methylation. Our data indicates that E92K may disrupt binding in the same manner. To verify this, we performed a mononucleosome FLAG-IP from HEK293T cells stably expressing H2A or H2A E92K and analysed the precipitated DNA for [MeC] (Figure 3f). Consistent with our hypothesis, the data show a ~5% decrease in local DNA methylation in the E92K samples, supporting a model where DNMT3 binding is impaired by E92K, leading to fewer de novo methylation events.

We were able to validate loss of function binding interactions with SIRT6, DNMT3A, INCENP (part of the CEN complex[40]) and the chromatin remodeller SMARCA4[8,41] (BRG1) using the recombinant mononucleosome assay developed by McGinty.[35] Here, biotinylated mononucleosomes were immobilized on streptavidin beads before incubation with HEK293T nuclear lysate. The eluent was subsequently blotted against hits from our photoproximity labelling experiment to assess the effect of the E92K mutation on binding. In all cases, this assay supported the data presented by our µMap experiment (Figure 3g), with comparable enrichment values between western blot and proteomics datasets, further demonstrating that our method can accurately capture interacting proteins in an unbiased

manner. Interestingly, performing the same comparative proteomics experiment with H2A/H2A(E92K)-APEX2 constructs did not reveal any differences in the interactome of mutant and wild-type samples when applying the same fold change and significance cut-offs used previously (Extended Data Figures 6 & 10), presumably due to the increased labelling radius of the peroxidase-based platform (Extended Data Figure 6).

Together, these data demonstrate that the resolution provided by our proximity labelling method can be used to uncover molecular level details for gain-of-function or loss-of-function interactions in the nucleus in a single experiment.

Excited by this, we posited that our method could be used to determine the roles of small molecule ligands in the chromatin microenvironment (Figure 4a). Epigenetic drug discovery has become a critical focus for therapeutic intervention, encompassing dozens of targets across a range of therapeutic areas.[3] We began by examining the effects that the bromodomain inhibitor JQ-1 has on chromatin (Figure 4b). JQ-1 is known to target the BET family of bromodomain containing readers (BRD2/3/4), with analogues progressing through the clinic.[42,43] We theorized that μMap would be sensitive enough to measure the effects that BRD inhibition would have on the chromosomal microenvironment. To probe this, we compared the interactomes of Ir-conjugated H2A in the presence or absence of 1 and 5 μM JQ-1 (Figure 4b, Extended Data Figures 4, 7 & 10, see SI for full data table). BRD2/3/4 were all enriched in the untreated sample ($Log_2FC > 0.5$, FDR < 0.05), consistent with blocking bromodomain-nucleosome interactions. The known JQ-1 off-target SOAT1[44] was also enriched in the untreated sample. Importantly, RNA-seq analysis of the ligand treated cells showed that the identified interactors are not enriched based on transcriptional changes associated with drug treatment (Extended Data Figure 8).

Performing the same experiment using existing methods (APEX2, as above) showed enrichment of BRD2 and no enrichment of BRD3/BRD4 (Extended Data Figure 7).

We then performed a similar experiment with the DOT1L methyltransferase inhibitor, pinometostat, to assess both the selectivity of this ligand and the effect that depletion of H3K79 methylation has on the chromatin microenvironment.[45,46] DOT1L was enriched ($Log_2FC > 0.5$, FDR < 0.05) in the untreated sample, as were several proteins related to transcriptional activation, including BRD2/3/4 and POLR2E (Figure 4c,d. Extended Data Figure 7, see SI for full data table). This observation is consistent with previous reports demonstrating that H3K79 methylation leads to recruitment of the acetyltransferase P300, subsequent BRD recruitment, and transcriptional activation.[46] It was gratifying to see that in a single experiment we can extract both target-ID data for a small molecule ligand along with downstream transcriptional effects.

Finally, we applied our method to RNA polymerase II (POL II), which is responsible for transcribing protein-encoding genes in a highly regulated process involving the sequential association of multiple protein complexes.[47-49] Release of promoter-proximal paused POL II is primarily achieved through phosphorylation of the disordered C-terminal domain (CTD) of the RPB1 subunit, comprised of a repeating unit of Y-S-P-T-S-P-S, by cyclin dependent kinases (CDKs). Inhibition of CDKs has therefore become an attractive therapeutic strategy

to stall POL II and downregulate transcription and therefore cancer proliferation. The small molecule ligand AT7519, a pan-CDK inhibitor, has been developed to probe this strategy in the context of multiple myeloma and is currently in clinical trials.[50] We questioned if we could determine how this ligand affects the POL II interactome, and whether our method could unveil at what stage in its transcriptional cycle POL II is arrested by AT7519 treatment.

To assess this, we expressed RPB1-Cfa^N in HEK 293T cells, where the intein is fused to the C-terminus of the CTD (Extended Data Figure 9). The μMap workflow was then performed in the presence and absence of AT7519 (Figure 4e). Comparing our treated and untreated datasets revealed CDK11, 9, and 1 to be enriched ($Log_2FC > 0.5$, FDR < 0.05) in the untreated sample, consistent with wide-ranging CDK inhibition by AT7519 (Figure 4d, see SI for full data table).[50] Mediator complex subunits in addition to several transcription factors were enriched in the treated sample, suggesting CDK inhibition halts the progression of POL II at the pre-initiation complex, before CTD phosphorylation and association of the NELF and DSIF complexes.[51]

Conclusion

In summary, we have developed a photocatalytic proximity labelling technology that can be deployed across the nuclear proteome. The short range diazirine activation mechanism allows the collection of highly precise interactomics data that is sensitive to single amino acid mutations and can be used to detect changes caused by external stimuli such as ligand incubation. We believe this method will be broadly applicable across nuclear biology for the study of disease-associated mutations. Additionally, this method is an effective tool for ligand target-ID in chromatin, identifying on- and off-target proteins, and revealing how treatment with these molecules impacts local chromatin interactomes.

Methods

Solid Phase Peptide Synthesis

Boc-N^{α}-Cfa^C-CFGSGK(alloc)G-NH_2 was synthesized on a 0.1 mmol scale by standard Fmoc solid-phase peptide synthesis using DIC-Oxyma activation on a CEM Liberty Blue microwave-assisted peptide synthesizer on ChemMatrix Rink amide resin. Each residue was double coupled during the synthesis. Fmoc deprotection was performed at room temperature by the addition of 20% piperidine in DMF with 0.1 M HOBt. Alloc deprotection was performed by the addition of 0.1 eq. $Pd(PPh_3)_4$ and 2.5 eq. N,N'-dimethylbarbituric acid in DCM with nitrogen agitation. Treatment was performed twice at rt for 30 min. The resin was sequentially washed with 3x DCM, 3x DCM:DMF (1:1 v/v), 3x DMF, and 1x 5% w/v sodium diethyldithiocarbamate in DMF.

The resin was then split into 0.02 mmol aliquots and treated as follows: **Cfa^C-biotin**: Biotin (5 eq.) was coupled to the deprotected lysine side chain with PyAOP (4.95 eq.) and DIEA (10 eq.) activation in NMP for 4 h with nitrogen agitation. The resin was washed with 3x NMP, 3x DMF, and 3x DCM. Side-chain deprotection and cleavage from the resin was affected by addition of a 92.5:2.5:2.5:2.5 v/v/v solution of TFA:TIPS:EDT:H_2O for

130 minutes at rt. The cleavage solution was reduced to <5 mL volume under a positive pressure of N_2, and the crude peptide was precipitated using cold diethyl ether. The crude peptide was isolated by refrigerated centrifugation, resuspended in 50/50 v/v H_2O:MeCN with 0.1% TFA, and lyophilized to yield a white solid. **Cfa^C-Ir**: The iridium photocatalyst[52] was conjugated to the deprotected lysine side chain by treatment with NHS-Ir (1.2 eq.) and DIEA (2 eq.) in DMF for 2 h with nitrogen agitation in the absence of light. The resin was washed with 3x DMF, and 3x DCM. Side-chain deprotection and cleavage from the resin was affected by addition of a 95:2.5:2.5 v/v/v solution of TFA:TIPS:H_2O for 130 minutes at rt in the absence of light. The cleavage solution was reduced to <5 mL volume under a positive pressure of N_2, and the crude peptide was precipitated using cold diethyl ether. The crude peptide was isolated by refrigerated centrifugation, resuspended in 50/50 v/v H_2O:MeCN with 0.1% TFA, and lyophilized to yield a pale yellow solid.

Crude lyophilized peptides were purified by preparative scale reversed-phase HPLC and characterized by mass spectrometry.

Cell culture

HEK 293T cells were cultured as a monolayer in DMEM (Thermo Fisher), supplemented with 10% v/v FBS (Thermo Fisher), 100 U ml^{-1} penicillin (Thermo Fisher), and 100 μg ml^{-1} streptomycin (Thermo Fisher). Cells were maintained in an incubator at 37 °C with 5% CO_2.

Each 10 cm plate of HEK293T cells at 70% confluency were transfected with a plasmid encoding POI-HA-Cfa^N-FLAG (5 μg per plate) with Lipofectamine 2000 (12 μl per plate) following the manufacturer's instructions. After 6 h, the media was aspirated and replaced with fresh media. Transfection was performed for 24 or 48 h in an incubator at 37 °C with 5% CO_2.

Protein trans-splicing in isolated nuclei for TMT-10 plex experiments.

3×10^7 HEK 293T cells transfected with POI-HA-Cfa-N-FLAG were lysed by hypotonic lysis in 3 mL RSB buffer (10 mM tris, 15 mM NaCl, 1.5 mM $MgCl_2$, Roche cOmplete EDTA-free protease inhibitors, pH 7.6) for 10 min on ice. The crude nuclei were isolated by centrifugation at 400 g for 5 min at 4 °C. The nuclei were resuspended in 3 ml RSB buffer, and homogenized with ten strokes of a loose pestle Dounce homogenizer, and pelleted at 400 g for 5 min at 4 °C. The nuclei were resuspended in crosslinking buffer (20 mM HEPES, 1.5 mM $MgCl_2$, 150 mM KCl, Roche cOmplete EDTA-free protease inhibitors, pH 7.6) and centrifuged at 400 g for 5 min at 4 °C. Finally, the nuclei were resuspended in 300 μL of crosslinking buffer per 1 $x10^7$ cells. To the isolated nuclei was added Cfa^C-Ir in crosslinking buffer (0.5 μM final concentration). The nuclei were incubated at 37 °C for 1 h.

The nuclei were isolated by centrifugation at 400 g for 5 min at 4 °C and washed twice with crosslinking buffer (500 μL) to remove excess peptide. The pellets were then resuspended in 3 mL crosslinking buffer containing diazirine-biotin conjugate (200 μM) and irradiated with blue light for between 45-90s in the Penn PhD Photoreactor M2 at 100% light intensity at 4 °C. The nuclei were re-isolated by centrifugation at 400 g for 5 min at 4 °C and washed

once with crosslinking buffer to remove excess biotin-diazirine. The washed pellets were then resuspended in 2 mL LB3 buffer (10 mM tris, 100 mM NaCl, 1 mM EDTA, 0.5 mM EGTA, 0.1% sodium deoxycholate, 0.5% sodium lauroyl sarcosinate, pH 7.5) and sonicated using a Branson probe tip sonicator (12 cycles at 25% amplitude, 15 seconds on 15 seconds off on ice).

The lysed nuclei were then clarified through centrifugation at 15,000 g for 20 mins at 4 °C and the protein concentration of the supernatant was determined by BCA assay. Protein concentration was normalized across all experimental replicates and diluted to 1 mg/mL with binding buffer (25 mM tris, 150 mM NaCl, 0.25% v/v NP-40, pH 7.5). 1 mL of each sample was then incubated with 125 μL of pre-washed magnetic Sepharose streptavidin beads (https://www.cytivalifesciences.com, No: 28985738) for 2 h at rt with end-over-end rotation. The beads were subsequently washed twice with 1% w/v SDS in PBS, twice with 1 M NaCl in PBS, and 10% EtOH in PBS x 3.

Experiments with ligand treatment

Following transfection, plated cells expressing POI-HA-Cfa-N-FLAG were treated with the small molecule ligand for the specified amount of time (Supplementary Table 2). Following this, cells were scraped and pelleted as described previously. All buffers used in cell processing prior to irradiation (RSB, crosslinking) contained the ligand. Following irradiation, samples were treated as normal.

For proteomic analysis

Following streptavidin-enrichment, the beads were resuspended in PBS (300 μL) and transferred to a new 1.5 mL Lo-bind tube. The supernatant was removed and the beads were washed with 3 x PBS (0.5 mL) and 3 x ammonium bicarbonate (100 mM).

The beads were re-suspended in 500 μL 3 M urea in PBS and 25 μL of 200 mM DTT in 25 mM NH_4HCO_3 was added. The beads were incubated at 55 °C for 30 min. Subsequently, 30 μL 500 mM iodoacetamide in 25 mM NH_4HCO_3 was added and incubated for 30 min at room temperature in the dark. The supernatant was removed and the beads washed with 3 x 0.5 mL DPBS and 6 x 0.5 mL triethyl ammonium bicarbonate (TEAB, 50 mM). The beads were resuspended in 0.5 mL TEAB (50 mM) and transferred to a new protein LoBind tube. The beads were resuspended in 40 μL TEAB (50 mM), 1.2 μL trypsin (1 mg/mL in 50 mM acetic acid) was added and the beads incubated overnight with end-over-end rotation at 37 °C. After 16 hours, an additional 0.8 μL trypsin was added and the beads were incubated for an additional 1 hour at 37 °C. Meanwhile, TMT 10-plex label reagents (0.8 mg) (Thermo) were equilibrated to room temperature and diluted with 40 μL of anhydrous acetonitrile (Optima grade; 5 min with vortexing) and centrifuged. 40 μL of each TMT reagent was added to the appropriate sample. The reaction was incubated for 2 hours at room temperature. The samples were quenched with 8 μL of 5% hydroxylamine and incubated for 15 minutes. The samples were pooled in a new Protein LoBind tube and acidified with TFA (16 μL, Optima).

Mass spectra were obtained using an Orbitrap Fusion Lumos mass spectrometer at Princeton Proteomics Facility and analyzed using MaxQuant.[53] TMT labeled peptides were dried

down in a SpeedVac, re-dissolved in 300 µl of 0.1% TFA in water and fractionated into 8 fractions using the Pierce™ High pH Reversed-Phase Peptide Fractionation Kit (#84868). Resulting MS/MS/MS data was searched in Maxquant against the Uniprot human protein database containing common contaminants. The proteinGroups.txt file was subsequently imported into Perseus (https://maxquant.net/perseus/).[54] The data were subsequently filtered based upon the following criteria, 'only identified by site', 'reverse', and 'potential contaminant'. The resulting data was Log_2 transformed and median normalization was performed. FDR-corrected p values were determined by a 2-sample T-test following the Benjamini-Hochberg procedure. The data were visualized by plotting as a volcano plot.

Procedure for mononucleosome IP

Two 10 cm plates of HEK293T cells transfected with H2A-E92K-HA-Cfa^N-FLAG (approx. 30 million cells) and two 10 cm plates of HEK293T cells transfected with H2A-HA-Cfa^N-FLAG (approx. 30 million cells) were lysed in 1 mL hypotonic lysis buffer (10 mM tris, 15 mM NaCl, 1.5 mM $MgCl_2$, Roche cOmplete EDTA-free protease inhibitors, 1 mM DTT, 5 mM sodium butyrate, pH 7.6) for 10 min on ice, and the nuclei were pelleted at 400 g for 5 min at 4 °C.The nuclei were then resuspended in RSB + 1 mM DTT + 0.2% v/v triton + 5 mM Na butyrate (500 µL per condition) and incubated on ice for 5 min. Nuclei were then pelleted at 400 g for 5 min at 4 °C. The nuclei were washed once more with RSB + 1 mM DTT + 5 mM Na butyrate and centrifuged at 600 g for 5 min at 4 °C.

The nuclei were then resuspended in 500 µl MNase digestion buffer (10 mM tris, 60 mM KCl, 15 mM NaCl, 2 mM $CaCl_2$, pH 7.5) and were incubated at 37 °C for 10 min. MNase (5µL, NEB) was then added to each condition for 10 min at 37 °C (*N.B.* digestion time varies by enzyme batch and must be determined for each experiment. 2 µL of the digest was removed every two minutes and quenched by addition of 20 mM EGTA. These aliquots were run on a 1.2% agarose gel and the digestion efficiency visualized with ethidium bromide staining). Once digestion to mononucleosomes was complete, the reaction was quenched with the addition of 20 mM EGTA on ice for 5 min. The sample was spun at 1.3 kg for 5 min at 4 °C and the supernatant was collected (fraction S1). 500 µL of buffer TE + 5 mM Na butyrate (10 mM tris, 1 mM EDTA, pH 8.0) was added to the pellet and the sample was rotated end-over-end at 4 °C for 30 min. The sample was spun at 13,000 g for 5 min at 4 °C and the supernatant was collected (fraction S2). To 475 µL of S1 was added 475 µL of 2× buffer E (30 mM HEPES, 225 mM NaCl, 3 mM $MgCl_2$, 0.4% triton X-100, 20% v/v glycerol, pH 7.5) with constant vortexing dropwise over one minute. To 475 µl of S2 was added 237 µL of 3× buffer D (60 mM HEPES, 450 mM NaCl, 4.5 mM $MgCl_2$, 0.6 mM EGTA, 0.6% triton X-100, 30% v/v glycerol, pH 7.5) with constant vortexing dropwise over one minute. The samples were spun at 13,000 g for 5 min at 4 °C. The two fractions were combined in a 5 ml Lo-bind Eppendorf tube and 30 µΛ of magnetic FLAG beads (Sigma M2 anti-FLAG magnetic beads; pre-washed with 1× buffer D) were added per condition. The FLAG-IP was performed overnight with end-over-end rotation at 4 °C.

The beads were washed sequentially with 1× buffer D and 1× buffer D + 0.5% v/v triton X-100 for 2 min each with rotation. 40 µL 1× SDS loading buffer was added and the beads

were boiled for 10 min. Samples were run on a 15% tris gel for western blotting with appropriate antibodies.

Procedure for assaying local DNA methylation

Mononucleosome isolation was performed as above from HEK293T cells stably expressing H2A-HA-Cfa^{N}-FLAG or H2A-E92K-HA-Cfa^{N}-FLAG. Following this, mononucleosomes were enriched using 30 μL of magnetic FLAG beads (Sigma M2 anti-FLAG magnetic beads; pre-washed with 1× buffer D) per condition. The FLAG-IP was performed overnight with end-over-end rotation at 4 °C.

The beads were washed sequentially with 1× buffer D and 1× buffer D + 0.5% v/v triton X-100 for 2 min each with rotation. The beads were then washed twice with TE and resuspended in 200 μL elution buffer (50 mM Tris-Cl (pH 8.0), 10 mM EDTA, 1% SDS) supplanted with 1 μL Proteinase K (20 μg/μL). The samples were rotated at 55 °C for 2 h. Following this, the samples were centrifuged at 10 x 1000g for 1 min before being pelleted on a magnetic rack. The supernatants were transferred to clean low-bind tubes. The beads were then resuspended in 100 μL elution buffer supplanted with 1 μL Proteinase K (20 μg/μL) and rotated at 55 °C for a further 2 h. Again, the samples were centrifuged at 10,000 g for 1 min before being pelleted on a magnetic rack and the supernatants transferred to clean low-bind tubes. The supernatants were combined (within each replicate) and the eluted DNA was purified using a Monarch® PCR & DNA Cleanup Kit according to the manufacturer's instructions.

The eluted DNA was then digested using the Nucleoside Digestion Mix (NEB: M0649S) following the manufacturer's instructions. Digested nucleotide mixtures were then analyzed by MS and the ratio of dC to methylated dC was calculated and normalized to global methylation levels. Briefly, LC-QQQ-MS quantitation of digested deoxynucleosides was performed following literature precedent (Traube, F.R., et al.)[55] using a dynamic multiple reaction monitoring method (DMRM) on an Agilent 1260 LC Infinity II system coupled to an Agilent 6470 triple quadrupole mass spectrometer in positive ion mode. An InfinityLab Poroshell 120 SB C18 Column (Agilent, 683775-906(T), 2.7 μm particle size, 2.1 x 150 mm) was used for all analyses with a gradient composed of 0.1% formic acid in water (A) and acetonitrile (B) at 0.4 mL/min flow rate. The following mass spectrometer operating parameters were used: gas temperature 325 °C, gas flow rate 12 L/min, nebulizer pressure 20 psi, capillary voltage 2,500 V, and fragmentor voltage 70V; collision energy was set to 14 for dC and 7 for m5dC. The MS1 (parent ion) to MS2 (deglycosylated base ion) transition for dC was set to m/z 228.1 112 and m/z 242.1126 for m5dC. Commercially available deoxyribonucleosides were used to generate standard curves and the concentration of m5dC was normalized to dC concentration. Data quantification was performed with Agilent MassHunter Workstation Data Acquisition v10.0.6

Subcellular fractionation

One 10 cm plate of HEK 293T cells were transfected with the denoted plasmids and harvested as previously described. Cells were lysed in hypotonic lysis buffer (10 mM tris, 15 mM NaCl, 1.5 mM $MgCl_2$, PI, pH 7.6) on ice for 10 min, followed by sonication at 400

g for 5 min at 4 °C. The supernatant was removed (cytosolic fraction) and the pellet was resuspended in hypotonic lysis buffer + 1% v/v Triton X-100. Nuclei were lysed with 10 strokes of a tight pastel homogenizer followed by centrifugation at 10,000 g for 10 min at 4 °C. The supernatant was removed (nucleoplasmic fraction) and RIPA buffer was added to the pellet. The sample was sheared by probe sonication (2 x 10 sec total, 25% amplitude, 1 sec on, 1 sec off) to yield the chromatin fraction. SDS loading buffer was added and the samples were analyzed by western blotting using the indicated antibodies.

Recombinant nucleosome immunoprecipitation

Nuclear lysate was prepared as before. Following this, biotinylated nucleosomes (https://www.epicypher.com/products/nucleosomes) were immobilized on streptavidin T1 magnetic dynabeads (MyOne, Thermofisher) in BB150 (20 mM HEPES pH 7.5, 150 mM NaCl, 10% glycerol, 1 mM DTT, 0.1% NP-40) by rotation for 2 h at 4°C. 60 μg of nucleosomes were incubated with 150 μL resuspended resin in 1 ml of BB150. Streptavidin-bound nucleosomes were rinsed twice and washed for 30 min by rotating with 1 mL of BB150 at 4°C and divided equally into three tubes. Nuclear lysate (700 μg) was mixed with 300 μL of BB150 and centrifuged at 10 000 × g for 10 min at 4°C before adding to nucleosome-bound streptavidin beads.

The mixture was rotated for 2 h at 4°C after which the beads were rinsed twice with 0.5 ml of BB150 and then rotated for 1 h at 4°C with 0.5 mL of BB150. Washed dynabeads were moved to new tubes with ~500 μL BB150, the beads were pelleted on a magnetic rack and the buffer was aspirated, washed beads were then centrifuged at 800 × g for 2 min at 4°C, and aspirated again to remove residual buffer. Beads were then resuspended in 15 μL of 2× gel loading buffer (Bio-Rad) and boiled for 5 minutes before being pelleted on a magnetic rack. Eluted proteins were then analyzed by western blotting using the indicated antibodies.

ATAC-Seq analysis

ATAC-seq analysis was performed by Genewiz. Briefly, cells stably expressing H2A-HA-Cfa^{N}-FLAG or H2A(E92K)-HA-Cfa^{N}-FLAG were harvested and cryopreserved before delivery to Genewiz for analysis. Both conditions were analyzed over 4 biological replicates.

RNA-Seq analysis

RNA-seq analysis was performed by Genewiz. Briefly, cells stably expressing H2A-HA-Cfa^{N}-FLAG were treated with either JQ-1 (1 μM for 3 h) or Pinometostat (2.5 μM for 24 h) harvested and cryopreserved before delivery to Genewiz for analysis. Four biological replicates were performed for each condition.

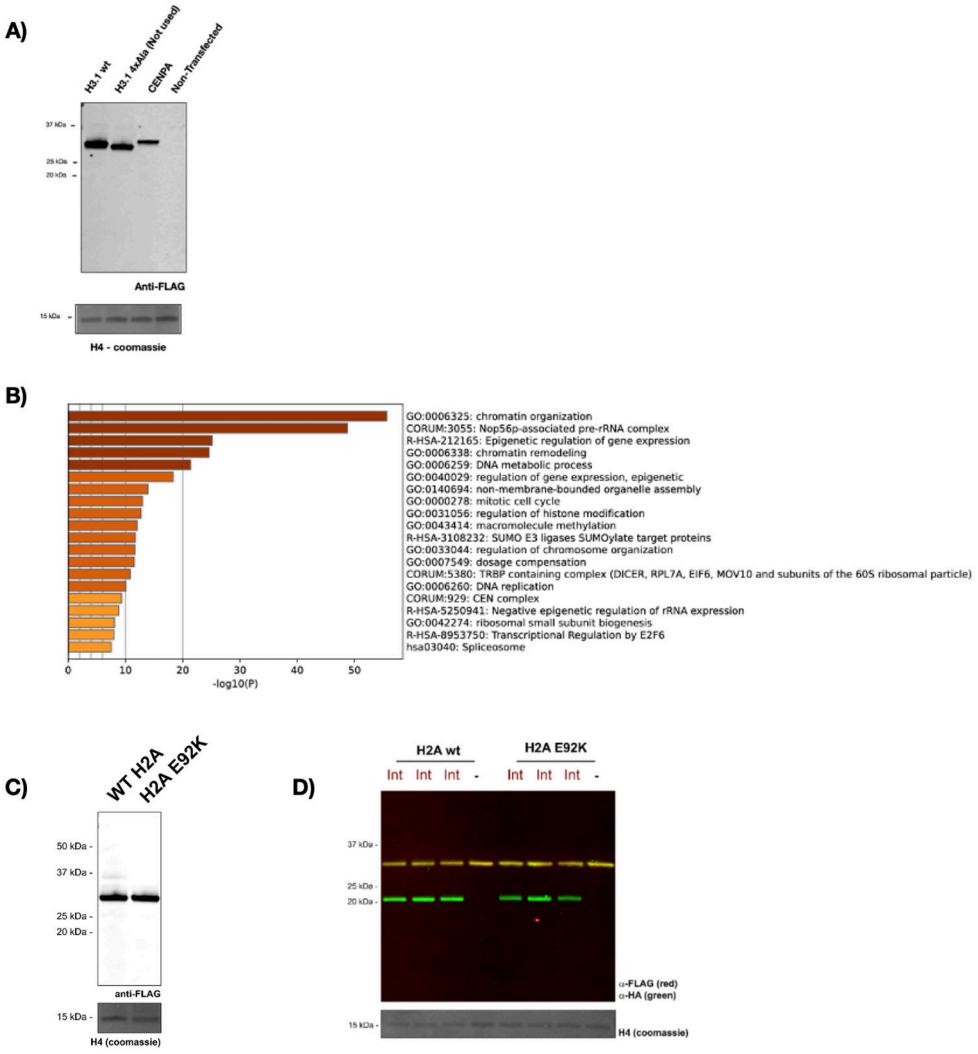
Generation of stable cell lines.

Constructs were cloned into a pCDH-CMV-MCS-EF1-Puro expression vector using standard restriction cloning procedures. Lentiviral cell lines were generated using established methods.

Statistics and Reproducibility

Statistical analyses were performed using Prism (GraphPad). P values were determined by paired T-tests as appropriate and as listed in the figure legends. The statistical significances of differences (*P < 0.05, **P < 0.01, ***P < 0.001, ****P < 0.0001) are specified throughout the figures and legends. GO analysis was performed using Metascape.[56] Heatmap was generated using Heatmapper (Heatmapper.ca).[57] Expression and splicing blots were repeated successfully three times. Proteomics experiments in figures 2H, 3A, 4B, 4C, and 4F have been performed once.

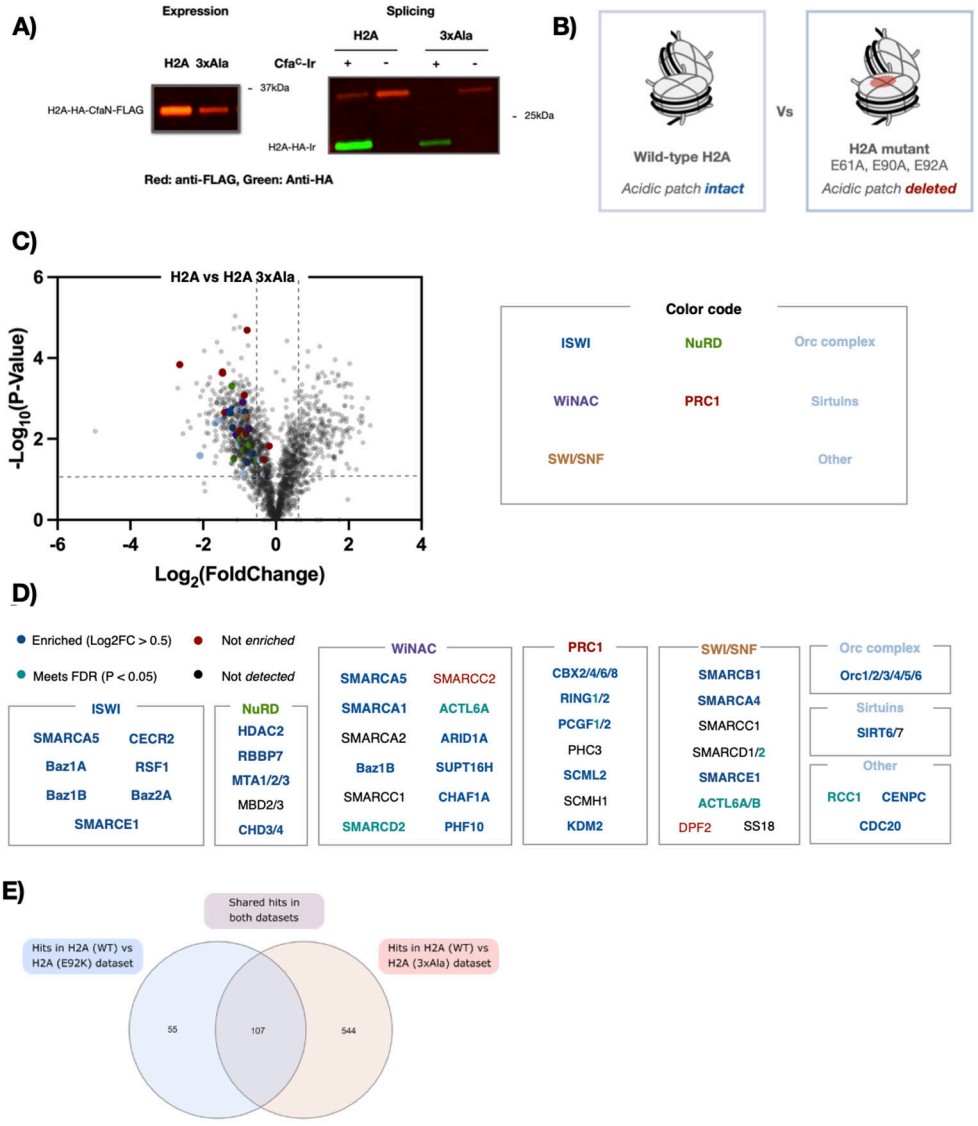
Extended Data



Extended Data Figure 1: Supporting data for H3.1 vs CENP-A and H2A WT vs H2A E92K experiments.

Anti-FLAG western blot for the expression of H3.1-HA-Cfa^{N}-FLAG, H3.1 (4xAla)-HA-Cfa^{N}-FLAG (not used in MS experiment) and CENP-A-HA-Cfa^{N}-FLAG. H4 visualized by coomassie blue stain is provided as a loading control. MW of H3.1-HA-Cfa-N-FLAG = 29,365 Da. B) GO analysis of proteins enriched in H3.1 samples. Data shows significant

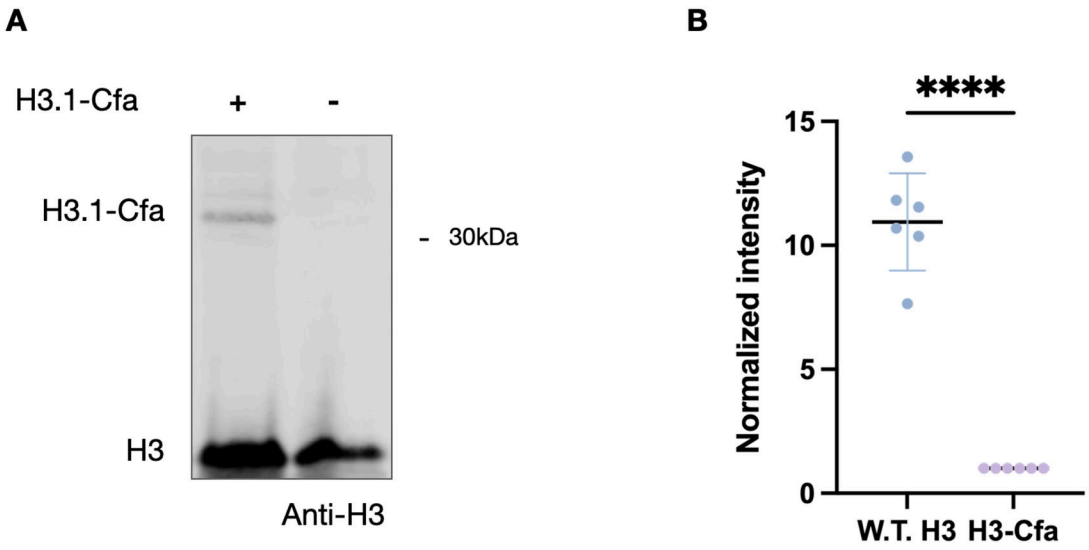
GO terms related to Chromatin organization, remodeling, and epigenetic control of gene expression. Analysis performed using metascape. C) Anti-FLAG western blot for the expression of H2A-HA-Cfa^{N}-FLAG and H2A-E92K-HA-Cfa^{N}-FLAG. H4 visualized by coomassie blue stain is provided as a loading control. D) In-nuclei splicing reactions for H2A (WT) and H2A (E92K) constructs in the presence of Cfa^{C}-Ir (referred to as Int) and control as visualized by anti-HA and ant-FLAG western blot. H4 visualized by coomassie blue stain is provided as a loading control. MW of H2A-HA-Cfa^{N}-FLAG = 28,121 Da. For blot source data, see Supplementary Figure 1.



Extended Data Figure 2: Interactome of acidic patch triple mutant and comparison to H2A E92K.

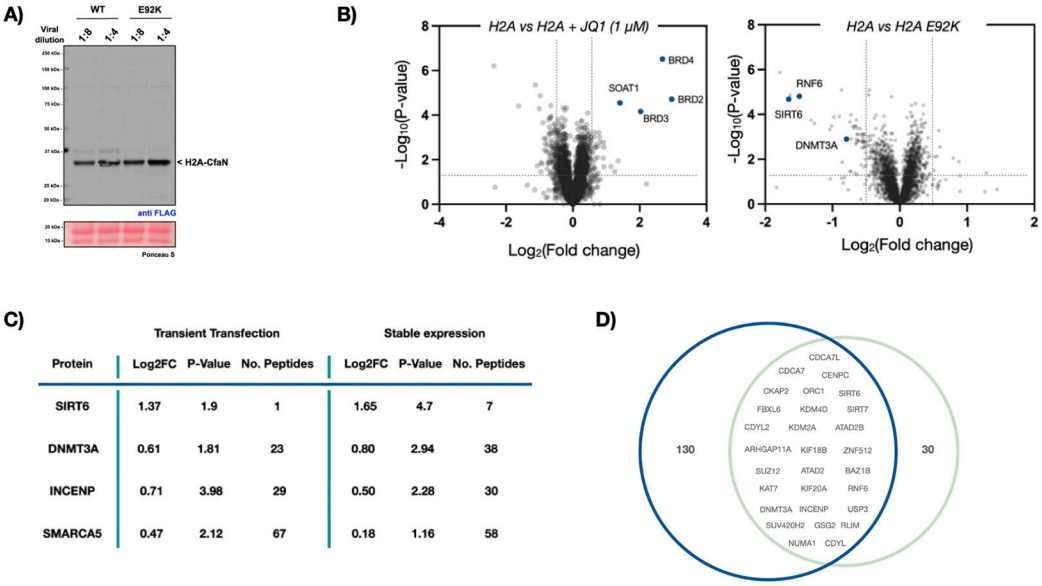
A) Left: Anti-FLAG western blot for the expression of H2A-HA-Cfa^{N}-FLAG and H2A-3xAla-HA-Cfa^{N}-FLAG. Note, 3xAla refers to the acidic patch mutations E61A, E90A and E92A in H2A. Right: In-nuclei splicing reactions for H2A (WT) and H2A (3xAla) constructs in the presence of Cfa^{C}-Ir as visualized by anti-HA and ant-FLAG western

blot. B) Graphic representing a comparative experiment between H2A (WT) and H2A (3xAla). C) Volcano plot derived from a two-sided t-test displaying proteins enriched in a comparative proteomics study between H2A (WT) and H2A (3xAla) using the µMap workflow. Proteins are highlighted that represent acidic patch binders and chromatin remodelers, and are colour coded according to the remodelling complex they belong to. FDR values were calculated using the Benjamini–Hochberg procedure, as described in Methods. D) Table showing known chromatin remodeling complexes that are enriched in the µMap dataset between H2A (WT) and H2A (3xAla). Proteins are color coded by enrichment. Blue = Log_2FC >0.5, FDR <0.05. Cyan = Meets FDR requirement, FDR <0.05. Red = Not enriched. Black = Not in data set. Data shows that nucleosome binding for the majority (80%) of all complex members is affected by the deletion of the H2A acidic patch. E) Venn diagram displaying overlap in the loss of function proteins in the H2A (WT) vs H2A (E92K) and H2A (WT) vs H2A (3xAla) datasets. Approximately two thirds of the hits (Log_2FC >0.5, FDR <0.05) observed in the H2A (E92K) experiment are also hits in the H2A (3xAla) experiment. For blot source data, see Supplementary Figure 1.



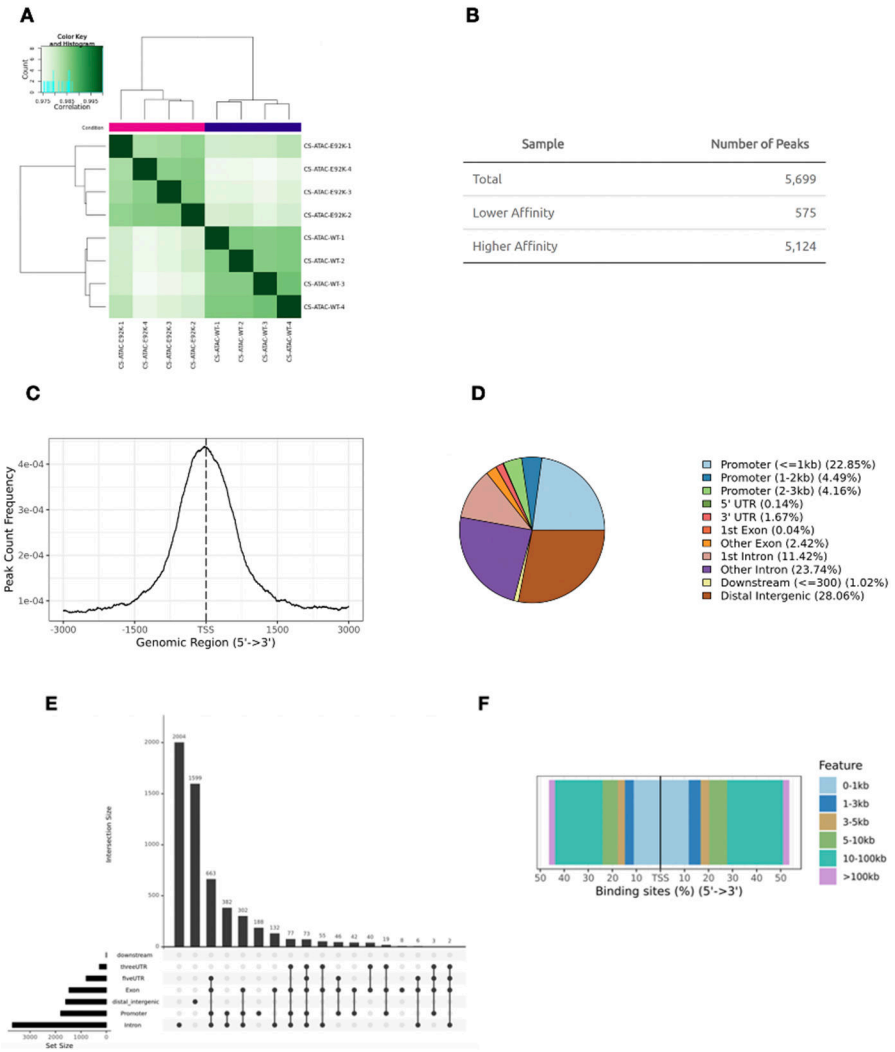
Extended Data Figure 3: Measurement of relative quantities of transfected histone H3.1.

A) Representative western blot showing relative concentrations of native H3 and transfected H3.1-HA-Cfa^N-FLAG. B) Bar graph showing relative intensities of transfected and native H3. N = 6 independent biological replicates, (**** = P<0.0001). Over 6 replicates transfected histones make up ≈1/12 of all histone H3. For blot source data, see Supplementary Figure 1.



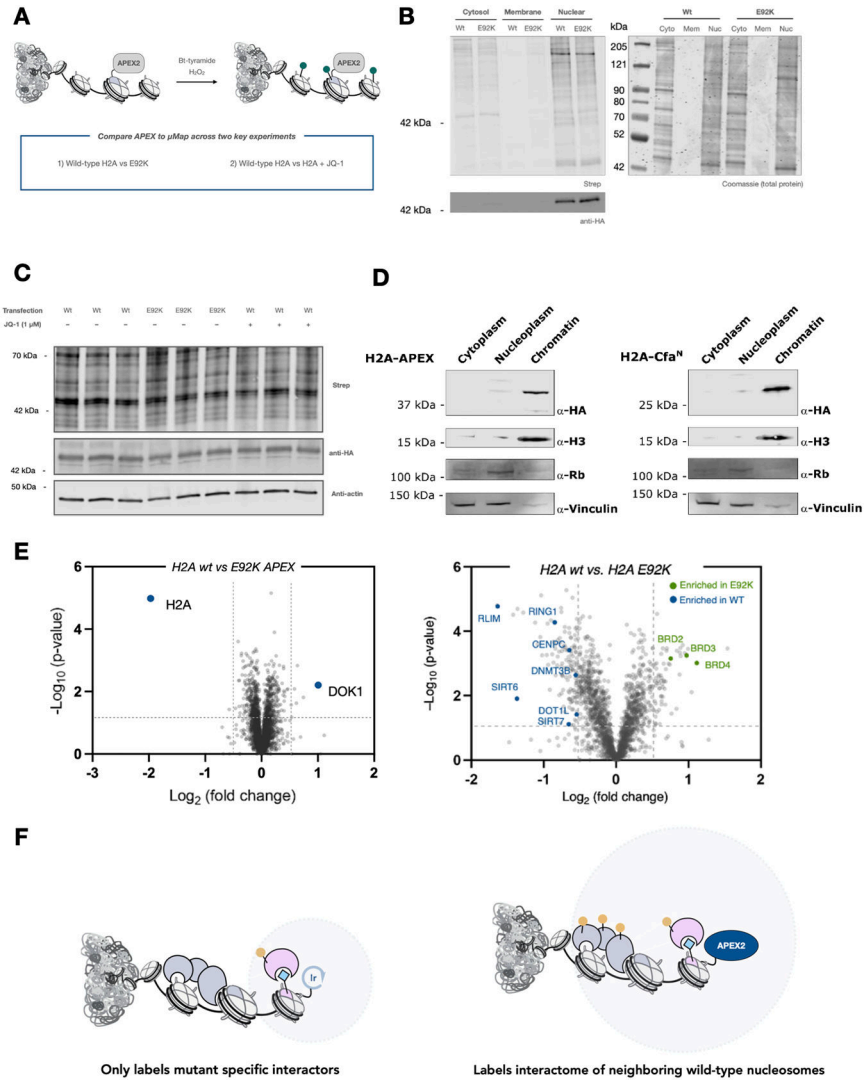
Extended Data Figure 4: Comparison of proteomics data obtained from stable cell lines.

To compare the effects of stable expression on µMap we conducted our two core workflows in HEK23T cells expressing H2A-HA-Cfa^{N}-FLAG and/or H2A(E92K)-HA-Cfa^{N}-FLAG. A) Anti-FLAG western blot showing stable expression of H2A-HA-Cfa^{N}-FLAG and H2A(E92K)-HA-Cfa^{N}-FLAG at the indicated viral titer. Ponceau staining is provided as a loading control. B) Left: Volcano plot derived from a two-sided t-test showing µMap target ID experiment with JQ-1 in stably expressing cells with 1 µM JQ-1. Data is consistent with transiently transfected cell lines. Right: Volcano plot derived from a two-sided t-test showing µMap experiment comparing wild-type and E92K interactomes. C) Comparison of H2A (WT) vs H2A (E92K) proteomics hits from transiently transfected cells compared to stably expressing cells. Table shows the four proteins investigated in this study. D) Venn diagram showing all proteins in common between datasets (Log_2FC >0.5, FDR <0.05). Overlap is observed for key chromatin modifying proteins. For blot source data, see Supplementary Figure 1.



Extended Data Figure 5: ATAC-Seq extended data.

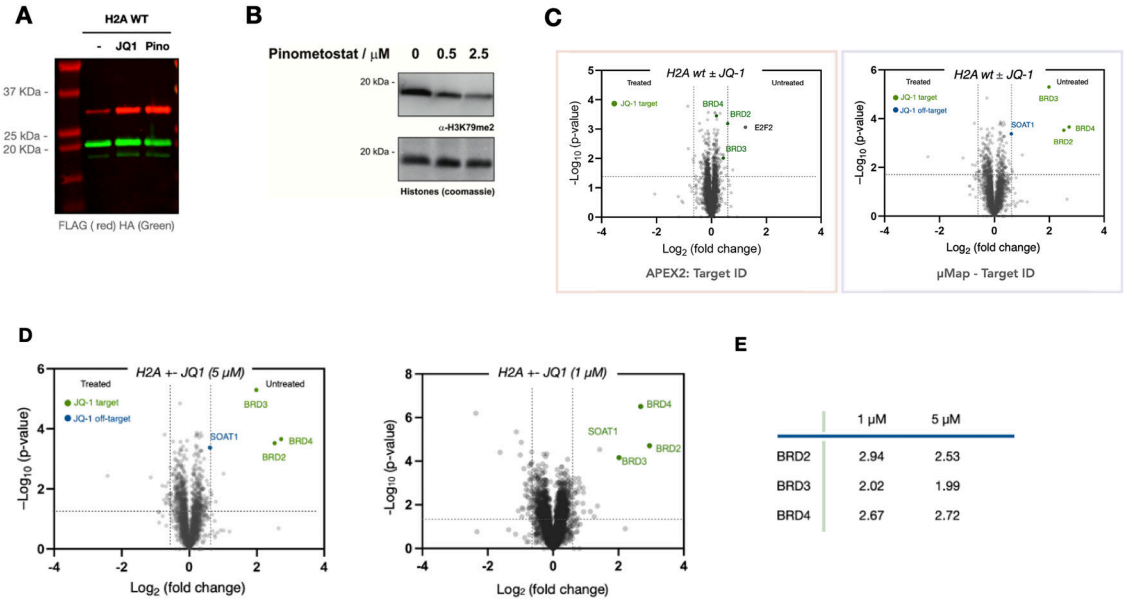
A) Hierarchical clustering analysis: Read counts in merged peaks are used to calculate distances between samples and generate a heatmap of sample-to-sample relationships. Data shows strong association between replicates. B) Differential binding analysis: The R package DiffBind is used for differential region of interest (ROI) detection using FDR <= 0.05 as a cutoff. Peaks that have a fold-change below zero and FDR <= 0.05 indicate regions with lower accessibility. Peaks with a fold change greater than zero and FDR <= 0.05 indicate regions with higher accessibility. C) Average ATAC-seq profile across all samples at annotated transcription start sites (TSS's): Data shows the average profile of ATAC-seq peaks falling within +/− 3Kb of annotated TSS sites genome-wide. D) Differential peak annotation: Peaks are annotated based on genomic features they overlap. When annotations overlap, assignment priorities are Promoter > 5'UTR > 3'UTR > Exon > Intron > Downstream > Intergenic. E) Histogram of overlapping features under merged peaks, ranked by number of peaks. F) Distribution of genomic distance from peak summits to the nearest TSS.



Extended Data Figure 6: APEX2 comparison – Validation.

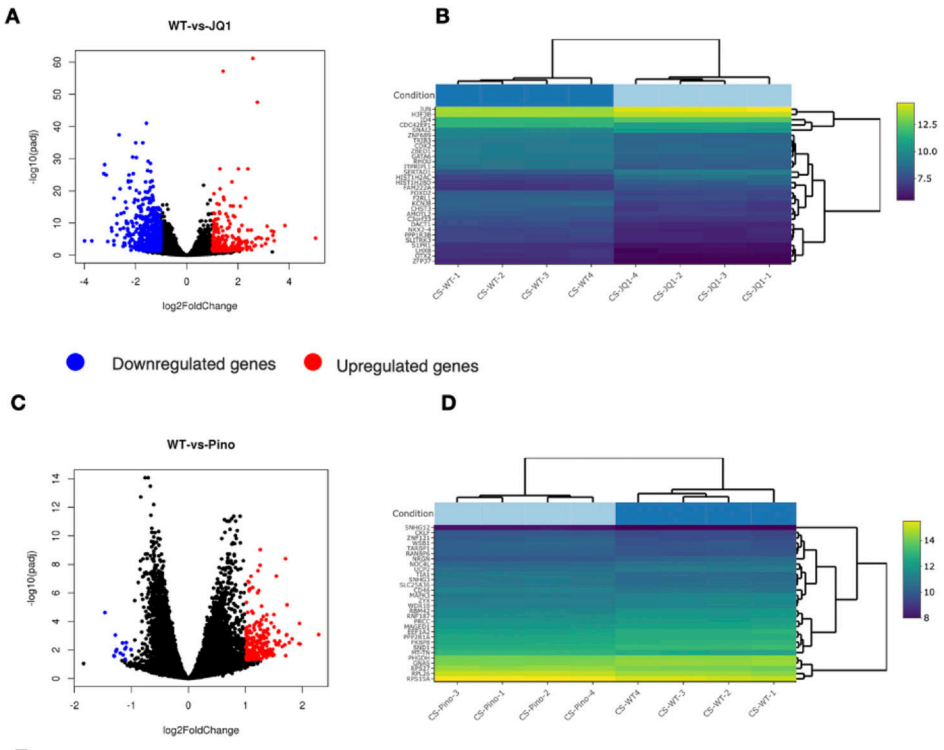
Comparison of µMap to APEX approaches: A) Key experiments in this study were repeated using previously reported APEX2 methodology according to the published procedure: Hung, V., Udeshi, N., Lam, S. et al. Spatially resolved proteomic mapping in living cells with the engineered peroxidase APEX2. *Nat Protoc 11*, 456–475 (**2016**). https://doi.org/10.1038/nprot.2016.018. B) Western blot showing relative levels of biotinylation in the fractionated nuclear, cytosolic and membrane proteome. Data shows that the APEX construct is localized to the nucleus. C) Western blot showing data from a 9-plex experiment representing two different experiments (H2A (WT) vs. H2A (E92K) and H2A ± JQ1; each N=3). H2A-APEX2 expression and levels of biotinylation are even across all replicates. MW: H2A-HA-APEX2-FLAG = 41,981 Da. D) Fractionation of HEK 293T cells transfected with either H2A-APEX or H2A-Cfa^{N} shows that expressed HA-tagged histone fusion proteins are localized to the chromatin fraction. Left) Anti-HA western blot showing H2A-HA-APEX2-FLAG localization in different cellular fractions. Appropriate loading controls are provided. Right) Anti-HA western blot showing H2A-HA-Cfa^{N}-FLAG localization to

different cellular fractions. Appropriate loading controls are provided. E) Volcano plots derived from a two-sided t-test showing comparison of results obtained from APEX2 proximity labeling and µMap experiments for H2A (WT) vs H2A (E92K). F) Model for interactomes detected using APEX2 and Iridium conjugated nucleosomes. The short radius of µMap only allows labeling of proteins directly interacting with the expressed nucleosome. The longer radius of APEX2 leads to additional labeling of the wild-type nucleosomal interactomes that border the expressed histone, leading to significant reduction in signal to noise. For blot source data, see Supplementary Figure 1.



Extended Data Figure 7: Supporting data for H2A ± JQ-1 and Pinometostat.

A) *In nucleo* splicing reactions for H2A-Cfa^{N} constructs in the presence of Cfa^{C}-Ir as visualized by anti-HA/anti-FLAG western blot. Left (untreated cells), middle (cells treated with JQ1), right (cells treated with Pinometostat). MW of H2A-HA-Cfa^{N}-FLAG = 28,121 Da. B) Treatment of HEK 293T cells with pinometostat for 24 h (0, 0.5, 2.5 µM) results in ≈60% decrease in H3K79 dimethylation, as visualized by western blotting with an anti-H3K79me2 antibody. Histones H3, H2A, and H2B, stained by coomassie blue, are provided as a loading control. C) Volcano plots derived from a two-sided t-tests comparing APEX2 target ID proximity labeling to µMap target ID for H2A ± JQ-1. D) Volcano plot derived from a two-sided t-test for comparison of H2A ± JQ1 at two different treatment concentrations (5 µM (left) and 1 µM (right)). E) Chart comparing the observed $\log_2$ (fold change) values returned for BRD2/3/4 at the indicated treatment concentrations. For blot source data, see Supplementary Figure 1.



E

Wt vs JQ-1					
		baseMean	Log2(Fold Change)	pvalue	padj
BRD2	ENSG00000204256	12953.71477	1.060818099	5.30E-16	8.72E-14
BRD3	ENSG00000169925	1685.627352	0.283577424	0.041784882	0.15874942
BRD4	ENSG00000141867	2787.228506	0.434805647	0.090038333	0.26292896

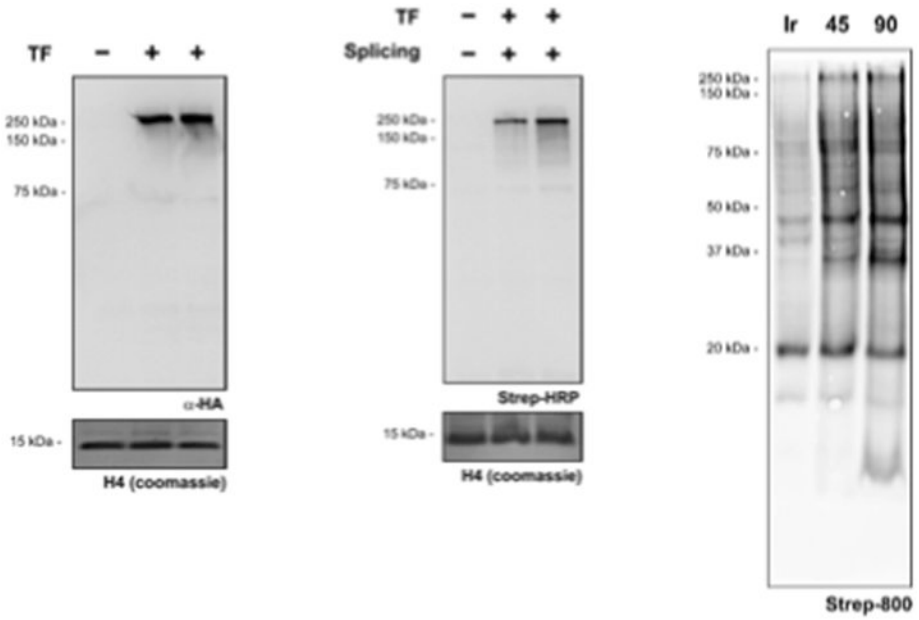
F

WT vs Pino					
		baseMean	Log2(FoldChange)	pvalue	padj
BRD2	ENSG00000204256	8400.58167	-0.26353089	0.00225403	0.011874
BRD3	ENSG00000169925	1467.72258	-0.380993586	0.00120351	0.00734239
BRD4	ENSG00000141867	2234.75844	-0.459723852	0.0001776	0.00164966
DOT1L	ENSG00000104885	1907.79669	-0.283934925	0.05559328	0.13845057
POLR2E	ENSG00000099817	5239.83375	-0.470085386	4.88E-09	4.00E-07

Extended Data Figure 8: RNA-Seq data.

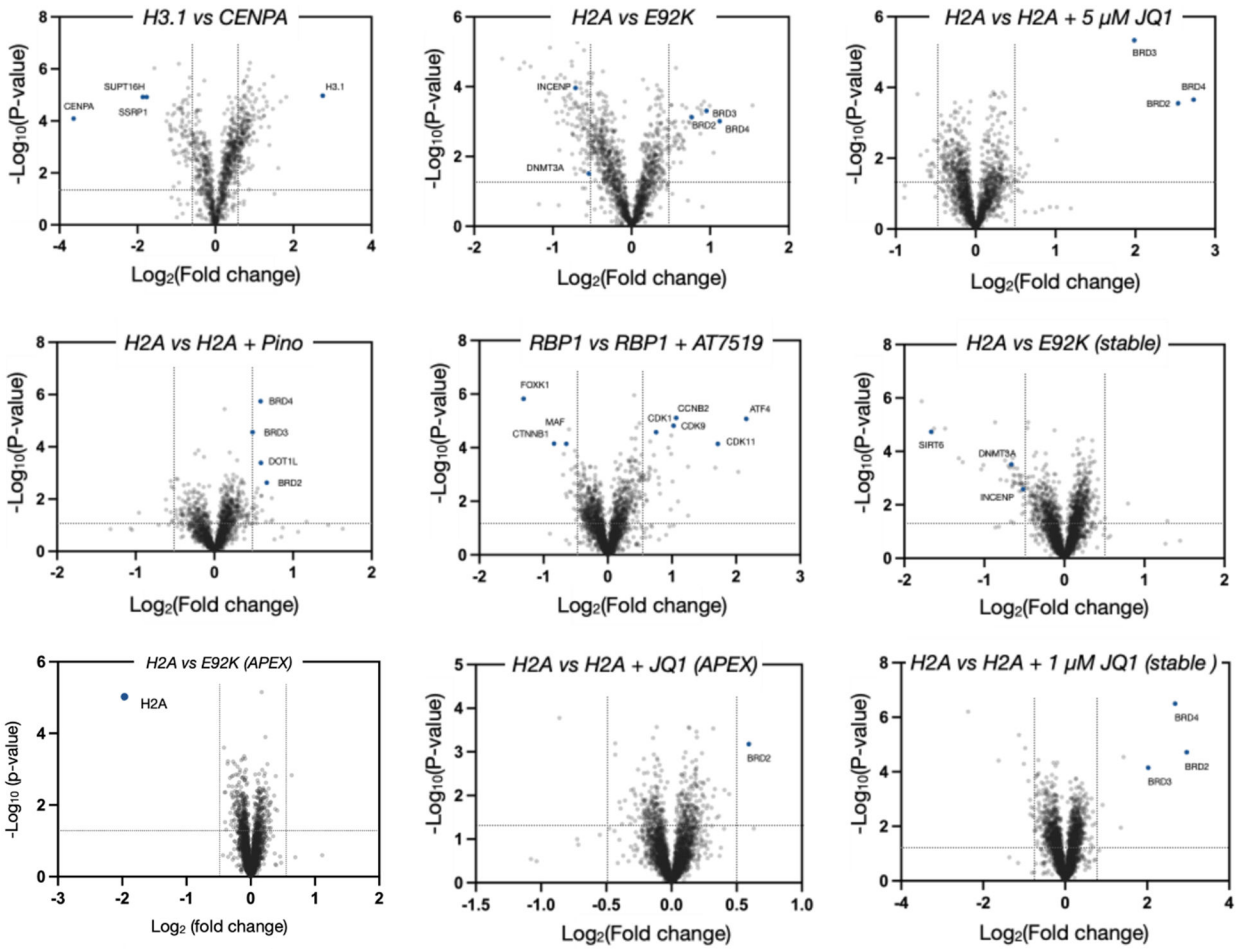
A) The global transcriptional change across the groups (WT vs JQ-1) compared was visualized by a volcano plot derived from a two-sided t-test. Each data point in the scatter plot represents a gene. The $\log_2$ fold change of each gene is represented on the x-axis and the log10 of its adjusted p-value is on the y-axis. Genes with an adjusted p-value less than 0.05 and a $\log_2$ fold change greater than 1 are indicated by red dots. These represent up-regulated genes. Genes with an adjusted p-value less than 0.05 and a log2 fold change less than −1 are indicated by blue dots. These represent down-regulated genes. B) A bi-clustering heatmap was used to visualize the expression profile of the top 30 differentially expressed genes (WT vs JQ-1) sorted by their adjusted p-value by plotting their $\log_2$ transformed expression values in samples. This analysis is useful to identify co-regulated genes across the treatment conditions. C&D) The same analysis was performed on (WT vs Pinometostat)

samples. E) Table showing RNA-seq fold changes and P-values for proteomic hits in from H2A vs H2A +JQ1 experiment. Data strongly suggests that enrichment is due to proximity and not global changes in gene expression. F) Table showing RNA-seq fold changes and P-values for proteomic hits in from H2A vs H2A + Pinometostat experiment. Data strongly suggests that enrichment is due to proximity and not global changes in gene expression Importantly, genes enriched in proteomic analysis were NOT enriched in RNA-seq, showing enrichment is based on proximity, rather than changes in gene expression caused by ligand treatment.



Extended Data Figure 9: Validation of RBP1 expression and splicing.

Left: Anti-HA western blot for the expression of RBP1-HA-Cfa^{N}-FLAG. H4 visualized by coomassie blue stain is provided as a loading control. Middle: *In nucleo* splicing reactions for RBP1-HA-Cfa^{N}-FLAG in the presence of Cfa^{C}-biotin (+) as visualized by streptavidin-800. H4 visualized by coomassie blue stain is provided as a loading control. Right: Labeling of nuclear proteins after installation of Ir photocatalyst (45 and 90 second irradiation with blue LEDs) versus free Ir control as visualized by streptavidin-800. MW of RBP1-HA-Cfa^{N}-FLAG = 231,290 Da. For blot source data, see Supplementary Figure 1.



Extended Data Figure 10: Volcano plots of all proteomics experiments analyzed with two unique peptide filter.

Volcano plots derived from a two-sided t-tests for MS experiments. All proteins presented require two unique peptides found across all replicates. Cut-offs are Log_2FC >0.5 and FDR <0.05. For blot source data, see Supplementary Figure 1.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements.

Research reported in this publication was supported by the NIH National Institute of General Medical Sciences (R35GM134897-01, R01- GM103558-03 and R37-GM086868) and the NIH National Cancer Institute (P01 CA196539). AJB was a Damon Runyon Fellow of the Damon Runyon Cancer Research Foundation (DRG-2283-17). The authors thank Saw Kyin and Henry H. Shwe at the Princeton Proteomics Facility. Elements in Figure 1 were created with Biorender.com.

Data Availability.

All relevant data are included in the manuscript and supplementary information. Mass spectrometry data files have been uploaded to the MassIVE proteomics database

PXD038956/ ftp://massive.ucsd.edu/MSV000090929/. Sequencing data is available on the GEO database (GSE221674).

References

1. Scott DE, Bayly AR, Abell C & Skidmore J Small molecules, big targets: drug discovery faces the protein–protein interaction challenge. Nat. Rev. Drug Discov 15, 533–550 (2016). [PubMed: 27050677]
2. Veltman JA & Brunner HG De novo mutations in human genetic disease. Nat. Rev. Genet 13, 565–575 (2012). [PubMed: 22805709]
3. Campbell RM & Tummino PJ Cancer epigenetics drug discovery and development: The challenge of hitting the mark. J. Clin. Invest 124, 64–69 (2014). [PubMed: 24382391]
4. Geri JB et al. Microenvironment mapping via Dexter energy transfer on immune cells. Science (80-.) 367, 1091–1097 (2020).
5. Ruffner H, Bauer A & Bouwmeester T Human protein-protein interaction networks and the value for drug discovery. Drug Discov. Today 12, 709–716 (2007). [PubMed: 17826683]
6. Kouzarides T. Chromatin Modifications and Their Function. Cell 128, 693–705 (2007). [PubMed: 17320507]
7. Ganesan A, Arimondo PB, Rots MG, Jeronimo C & Berdasco M The timeline of epigenetic drug discovery: from reality to dreams. Clin. Epigenetics 11, 174 (2019). [PubMed: 31791394]
8. Schick S et al. Systematic characterization of BAF mutations provides insights into intracomplex synthetic lethalities in human cancers. Nat. Genet 51, 1399–1410 (2019). [PubMed: 31427792]
9. Cheng F et al. Comprehensive characterization of protein–protein interactions perturbed by disease mutations. Nat. Genet 53, 342–353 (2021). [PubMed: 33558758]
10. Cheng Y et al. Targeting epigenetic regulators for cancer therapy: Mechanisms and advances in clinical trials. Signal Transduct. Target. Ther 4, (2019).
11. Weinberg DN, Allis CD & Lu C Oncogenic mechanisms of histone H3 mutations. Cold Spring Harb. Perspect. Med 7, 1–14 (2017).
12. Bagert JD et al. Oncohistone mutations enhance chromatin remodeling and alter cell fates. Nat. Chem. Biol 17, 403–411 (2021). [PubMed: 33649601]
13. Nacev BA et al. The expanding landscape of ‘oncohistone’ mutations in human cancers. Nature 567, 473–478 (2019). [PubMed: 30894748]
14. Müller MM & Muir TW Histones: At the crossroads of peptide and protein chemistry. Chem. Rev 115, 2296–2349 (2015). [PubMed: 25330018]
15. Vermeulen M & Déjardin J Locus-specific chromatin isolation. Nat. Rev. Mol. Cell Biol 21, 249–250 (2020). [PubMed: 31996790]
16. Van Mierlo G & Vermeulen M Chromatin proteomics to study epigenetics - Challenges and opportunities. Mol. Cell. Proteomics 20, 100056 (2021). [PubMed: 33556626]
17. Ciferri C et al. Molecular architecture of human polycomb repressive complex 2. Elife 2012, 1–22 (2012).
18. Ruthenburg AJ et al. Recognition of a mononucleosomal histone modification pattern by BPTF via multivalent interactions. Cell 145, 692–706 (2011). [PubMed: 21596426]
19. Zhao S, Yue Y, Li Y & Li H Identification and characterization of ‘readers’ for novel histone modifications. Curr. Opin. Chem. Biol 51, 57–65 (2019). [PubMed: 31082667]
20. Burton AJ et al. In situ chromatin interactomics using a chemical bait and trap approach. Nat. Chem 12, 520–527 (2020). [PubMed: 32472103]
21. Kleiner RE, Hang LE, Molloy KR, Chait BT & Kapoor TM A Chemical Proteomics Approach to Reveal Direct Protein-Protein Interactions in Living Cells. Cell Chem. Biol 25, 110–120.e3 (2018). [PubMed: 29104064]
22. Seath CP, Trowbridge AD, Muir TW & MacMillan DWC Reactive intermediates for interactome mapping. Chem. Soc. Rev 50, 2911–2926 (2021). [PubMed: 33458734]

23. Villaseñor R et al. ChromID identifies the protein interactome at chromatin marks. Nat. Biotechnol 38, 728–736 (2020). [PubMed: 32123383]
24. Ummethum H & Hamperl S Proximity Labeling Techniques to Study Chromatin. Front. Genet 11, 1–13 (2020). [PubMed: 32117431]
25. Baldi S, Korber P & Becker PB Beads on a string—nucleosome array arrangements and folding of the chromatin fiber. Nat. Struct. Mol. Biol 27, 109–118 (2020). [PubMed: 32042149]
26. Stevens AJ et al. A promiscuous split intein with expanded protein engineering applications. Proc. Natl. Acad. Sci. U. S. A 114, 8538–8543 (2017). [PubMed: 28739907]
27. Stevens AJ et al. Design of a Split Intein with Exceptional Protein Splicing Activity. J. Am. Chem. Soc 138, 2162–2165 (2016). [PubMed: 26854538]
28. Scott WA & Campos EI Interactions With Histone H3 & Tools to Study Them. Front. Cell Dev. Biol 8, 1–21 (2020). [PubMed: 32117956]
29. Pan D et al. Mechanism of centromere recruitment of the CENP-A chaperone HJURP and its implications for centromere licensing. Nat. Commun 10, 1–18 (2019). [PubMed: 30602773]
30. Chen CC et al. Establishment of Centromeric Chromatin by the CENP-A Assembly Factor CAL1 Requires FACT-Mediated Transcription. Dev. Cell 34, 73–84 (2015). [PubMed: 26151904]
31. Bennett RL et al. A mutation in histone H2B represents a new class of oncogenic driver. Cancer Discov. 9, 1438–1451 (2019). [PubMed: 31337617]
32. McGinty RK & Tan S Nucleosome structure and function. Chem. Rev 115, 2255–2273 (2015). [PubMed: 25495456]
33. McBride MJ et al. The nucleosome acidic patch and H2A ubiquitination underlie mSWI/SNF recruitment in synovial sarcoma. Nat. Struct. Mol. Biol 27, 836–845 (2020). [PubMed: 32747783]
34. Dao HT, Dul BE, Dann GP, Liszczak GP & Muir TW A basic motif anchoring ISWI to nucleosome acidic patch regulates nucleosome spacing. Nat. Chem. Biol 16, 134–142 (2020). [PubMed: 31819269]
35. Skrajna A et al. Comprehensive nucleosome interactome screen establishes fundamental principles of nucleosome binding. Nucleic Acids Res. 48, 9415–9432 (2020). [PubMed: 32658293]
36. Anink-Groenen LCM, Maarleveld TR, Verschure PJ & Bruggeman FJ Mechanistic stochastic model of histone modification pattern formation. Epigenetics and Chromatin 7, 1–16 (2014). [PubMed: 24393457]
37. Tachiwana H et al. Chromatin structure-dependent histone incorporation revealed by a genome-wide deposition assay. Elife 10, 1–30 (2021).
38. Grandi FC, Modi H, Kampman L & Corces MR Chromatin accessibility profiling by ATAC-seq. Nat. Protoc 17, 1518–1552 (2022). [PubMed: 35478247]
39. Xu T et al. Structure of nucleosome-bound DNA methyltransferases DNMT3A and DNMT3B. Nature 586, 151–155 (2020). [PubMed: 32968275]
40. Bolton MA et al. Aurora B Kinase Exists in a Complex with Survivin and INCENP and Its Kinase Activity Is Stimulated by Survivin Binding and Phosphorylation. Mol. Biol. Cell 13, 3064–3077 (2002). [PubMed: 12221116]
41. Hodges C, Kirkland JG & Crabtree GR The Many Roles of BAF (mSWI/SNF) and PBAF Complexes in Cancer. Cold Spring Harb. Perspect. Med 6, a026930 (2016). [PubMed: 27413115]
42. Filippakopoulos P et al. Selective inhibition of BET bromodomains. Nature 468, 1067–1073 (2010). [PubMed: 20871596]
43. Shi J & Vakoc CR The Mechanisms behind the Therapeutic Activity of BET Bromodomain Inhibition. Mol. Cell 54, 728–736 (2014). [PubMed: 24905006]
44. Savitski MM et al. Multiplexed Proteome Dynamics Profiling Reveals Mechanisms Controlling Protein Homeostasis. Cell 173, 260–274.e25 (2018). [PubMed: 29551266]
45. Stein EM et al. The DOT1L inhibitor pinometostat reduces H3K79 methylation and has modest clinical activity in adult acute leukemia. Blood 131, 2662–2669 (2018).
46. Gilan O et al. Functional interdependence of BRD4 and DOT1L in MLL leukemia. Nat. Struct. Mol. Biol 23, 673–681 (2016). [PubMed: 27294782]
47. Osman S & Cramer P Structural Biology of RNA Polymerase II Transcription: 20 Years On. Annu. Rev. Cell Dev. Biol 36, 1–34 (2020). [PubMed: 32822539]

48. Cramer P. Organization and regulation of gene transcription. Nature 573, 45–54 (2019). [PubMed: 31462772]

49. Young RA RNA POLYMERASE II. Annu. Rev. Biochem 60, 689–715 (1991). [PubMed: 1883205]

50. Santo L et al. AT7519, A novel small molecule multi-cyclin-dependent kinase inhibitor, induces apoptosis in multiple myeloma via GSK-3B activation and RNA polymerase II inhibition. Oncogene 29, 2325–2336 (2010). [PubMed: 20101221]

51. Harlen KM & Churchman LS The code and beyond: Transcription regulation by the RNA polymerase II carboxy-terminal domain. Nat. Rev. Mol. Cell Biol 18, 263–273 (2017). [PubMed: 28248323]

Additional references associated with methods

52. Trowbridge AD et al. , Small molecule photocatalysis enables drug target identification via energy transfer. Proc. Natl. Acad. Sci 119, e2208077119 (2022). [PubMed: 35969791]

53. Cox J & Mann M MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. Nat. Biotechnol 26, 1367–1372 (2008). [PubMed: 19029910]

54. Tyanova S et al. The Perseus computational platform for comprehensive analysis of (prote)omics data. Nat. Methods 13, 731–740 (2016). [PubMed: 27348712]

55. Traube FR et al. Isotope-dilution mass spectrometry for exact quantification of noncanonical DNA nucleosides. Nat. Protoc 14, 283–312 (2019). [PubMed: 30559375]

56. Zhou Y, Zhou B, Pache L et al. Metascape provides a biologist-oriented resource for the analysis of systems-level datasets. Nat Commun 10, 1523 (2019). [PubMed: 30944313]

57. Babicki S et al. Heatmapper: web-enabled heat mapping for all. Nucleic acids research 44, 147–153, (2016).

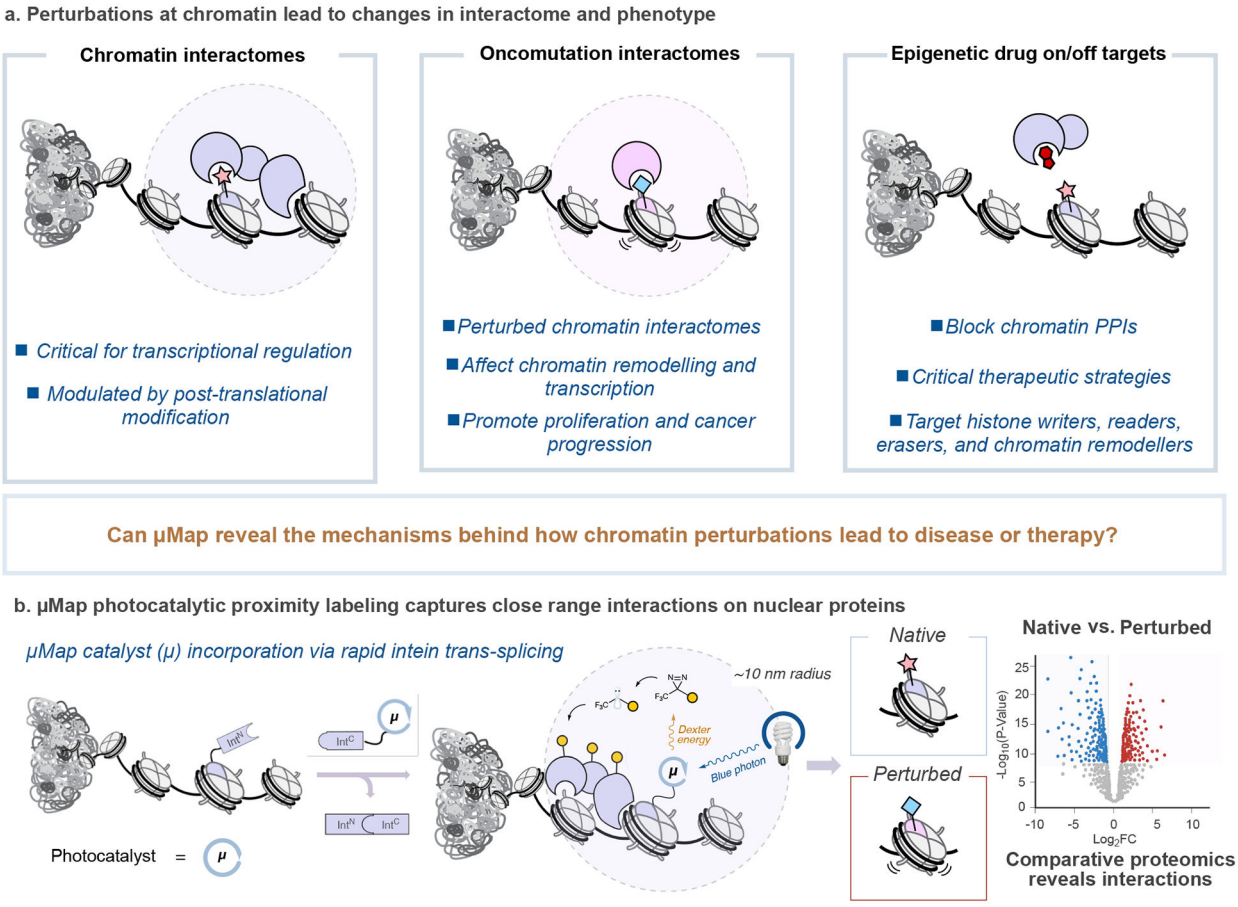


Figure 1 |. Development of a catalytic labeling platform using inteins.

a, Chromatin interactomes can be perturbed through mutation leading to oncogenic phenotypes. Epigenetic drugs alter the chromatin interactome for therapeutic benefit. **b,** Cartoon showing strategy for nuclear photoproximity labeling. Ir-photocatalysts can be incorporated into nuclear proteins via protein *trans*-splicing. The N-terminal fragment of the engineered Cfa split intein (Cfa^N) is fused to a target nuclear protein, while the complementary C-terminal fragment (Cfa^C) is linked to the photocatalyst. *In nucleo* splicing provides Ir-conjugated nuclear proteins. Following excitation under blue light irradiation, the photocatalyst is able to activate diazirines (to form carbenes via loss of molecular nitrogen) through a process called Dexter energy transfer. These carbenes are then poised to insert into neighbouring proteins which can be enriched using streptavidin beads. Comparative chemoproteomic analysis reveals how a given perturbation (e.g. a mutation or drug treatment) affects local interactomes, providing insights into the mechanistic basis for disease or therapy.

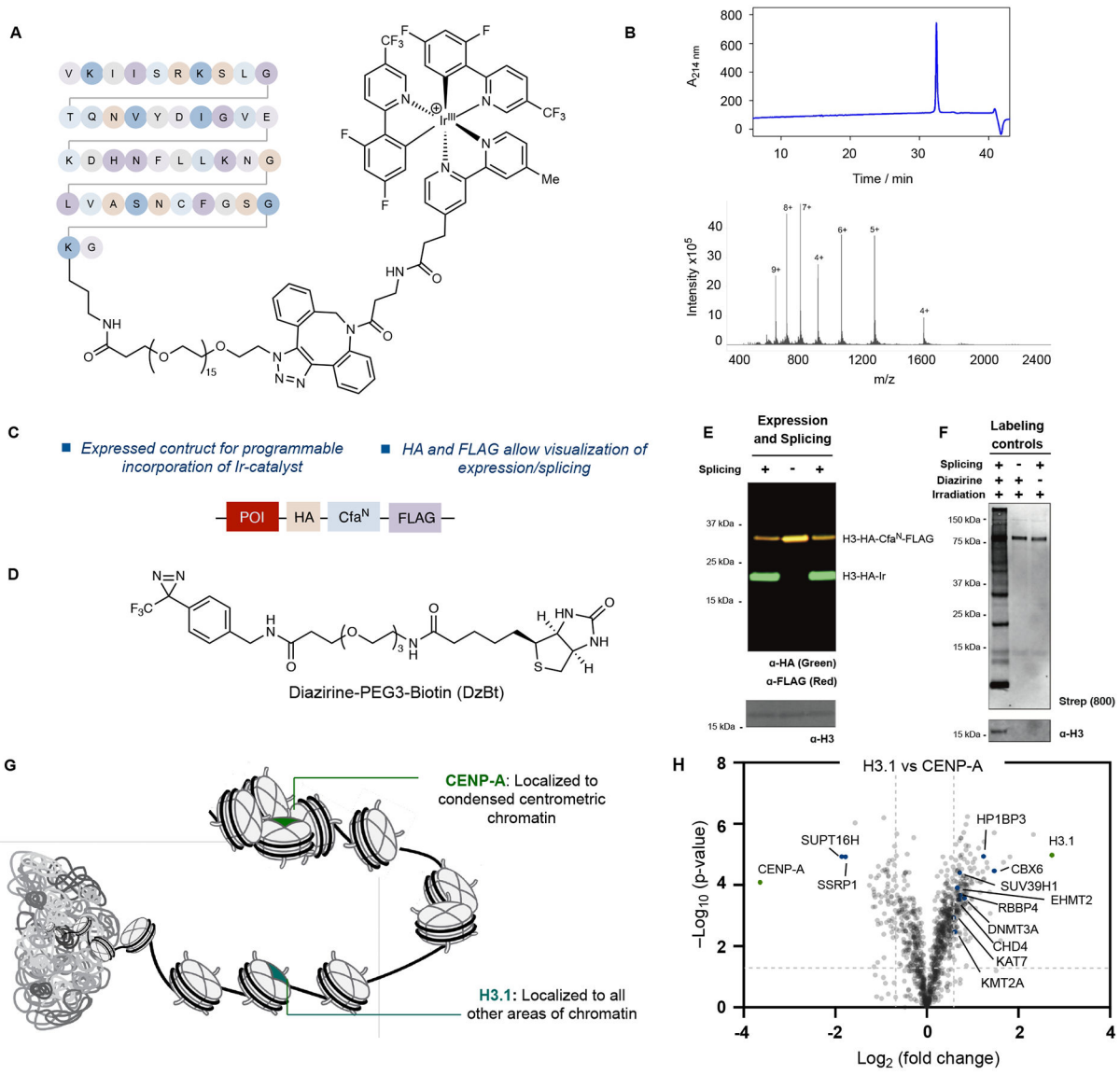


Figure 2 | Development of a chromatin localized µMap proximity labelling platform.

a, Structure of the Cfa^{C}-Ir construct used in this study. **b,** HPLC trace (top) and ESI-MS spectrum (bottom) for purified Cfa^{C}-Ir. **c,** General design of constructs used for µMap. **d,** Structure of diazirine-peg3-biotin probed used in this study. **e,** Validation of the approach using histone H3. In nucleo protein trans-splicing with 0.5 µM Cfa^{C}-Ir for 1 h results in the generation of the H3 spliced product (green) as shown by western blotting. **f,** Biotinylation of nuclear proteins, as detected by streptavidin blotting, is dependent on irradiation, Cfa^{C}-Ir and the diazirine probe. H3 is enriched following labeling and streptavidin enrichment. **g,** Graphic representing differential localization between histone H3.1 and centromeric variant CENP-A. **h,** Volcano plot derived from a two-sided t-test displaying proteins enriched in a comparative proteomics study between H3.1 and CENP-A using the µMap workflow. Selected proteins are highlighted that are discussed in the text. Dotted lines indicate cut-offs employed. FDR values were calculated using the Benjamini–Hochberg procedure, as described in the Methods. For blot source data, see Supplementary Figure 1.

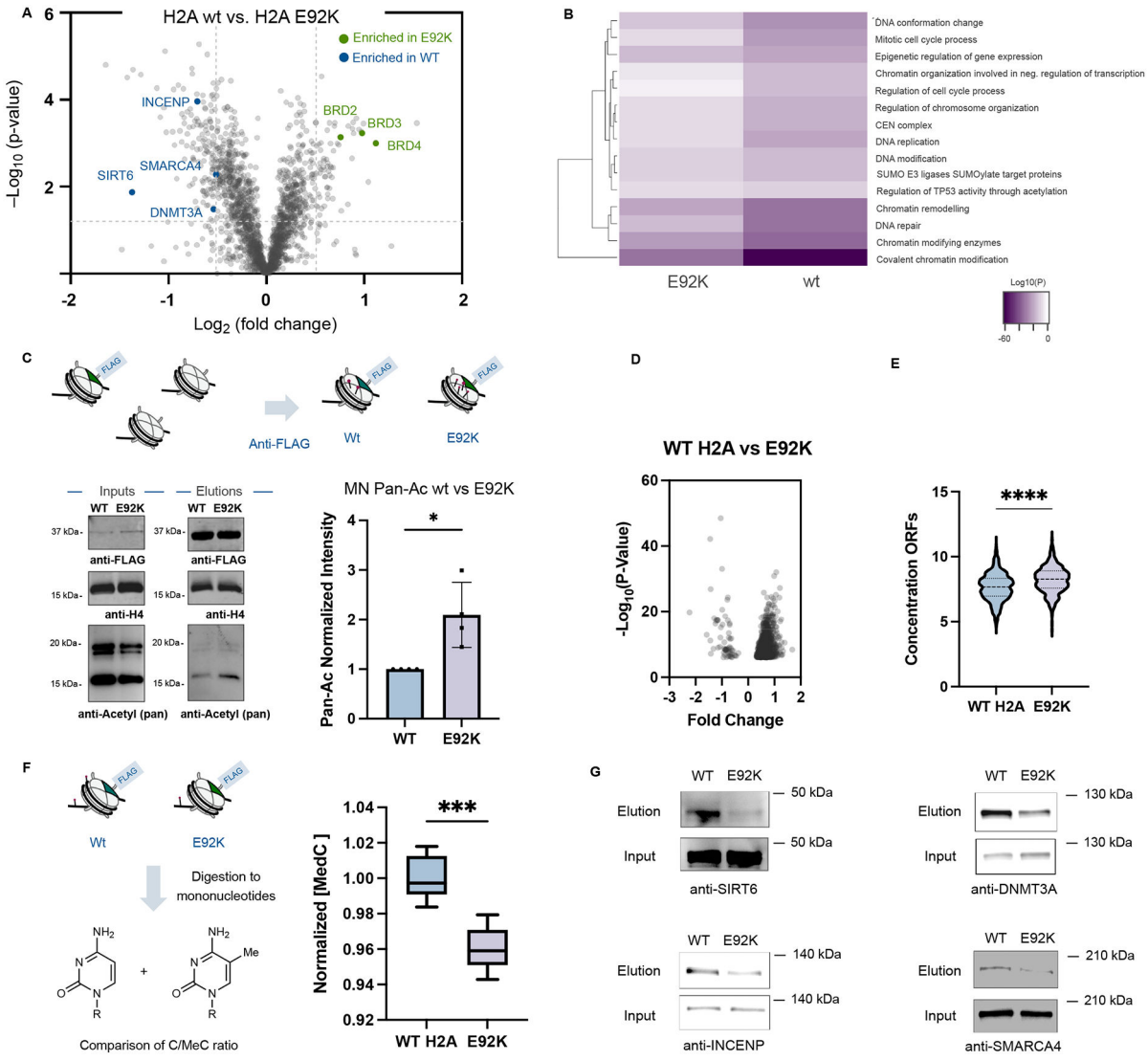


Figure 3 | µMap as a method to uncover oncogenic function of the somatic mutation H2A E92K.

a, A volcano plot derived from a two-sided t-test displaying protein interactors from the µMap method comparing H2A vs H2A E92K. Select proteins are highlighted. Dotted lines indicate cut-offs employed. FDR values were calculated using the Benjamini–Hochberg procedure, as described in the Methods. **b,** Comparative GO analysis for wt vs. E92K hits. GO terms consistent with the role of the acidic patch are de-enriched in the E92K mutant. **c,** Increased local acetylation of histone H4 is observed in the presence of H2A E92K, as determined by mononucleosome immunoprecipitation experiments. Western blot analysis showed an increase (2.1 +/− 0.33) in acetylation in the E92K mutant compared to wild type. Bar plot shows change in acetylation levels on H4 normalized to wild-type as determined by densitometry analysis of western blots (Mean with std. dev. n=4 independent biological replicates, P-Value = 0.0156). **d,** Volcano plot showing the concentration of significant (Two-sided t-test, P > 0.05) open reading frames (ORFs) in a comparative ATAC-seq experiment between HEK293T cells stably expressing wt H2A vs H2A E92K. **e,** Violin plot comparing the overall concentration of ORFs between HEK293T cells

stably expressing wt H2A vs H2A E92K from ATAC-seq analysis (P-Value = <0.0001, Two-sided T-test). The data shows that the E92K population has an increased concentration of ORFs when compared to the wild type sample. **f,** Left: Graphic showing experimental design for measuring local DNA methylation. Right: Box and whisker plots showing local concentration of Me-dC normalized to global Me-dC levels. E92K mutant shows a 5% decrease in methylation (n=6 independent biological replicates, whiskers represent min to max, P = 0.0003, Two-sided t-test). **g,** Western blot analysis comparing binding of selected proteins with either biotinylated wild-type nucleosomes or with biotinylated nucleosomes containing the H2A mutation E92K. *P < 0.05, *** P < 0.001, **** P < 0.0001. For blot source data, see Supplementary Figure 1.

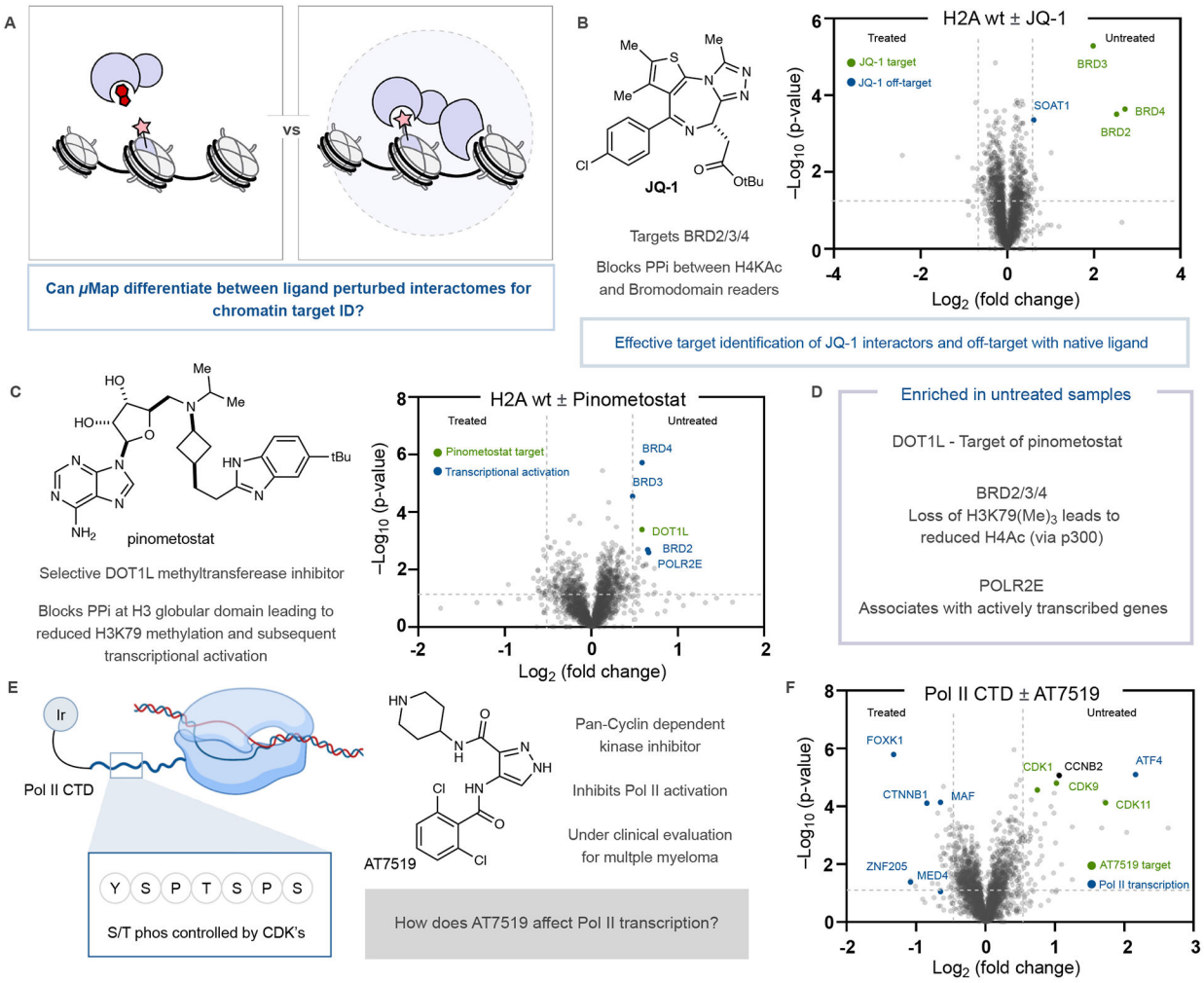


Figure 4 |. Interactome mapping of nuclear proteins to assess ligand interactions.

a, Performing the µMap workflow in the presence of a bioactive ligand unveils protein-protein interactions that are disrupted or promoted by ligand treatment. This method can provide nuclear target identification data. **b,** Volcano plot derived from a two-sided t-test displaying the H2A interactome vs H2A + 5 µM JQ-1 (structure shown at left). This method identifies BRD2/3/4 as JQ-1 target proteins in addition to known off-target SOAT1. **c,** Volcano plot derived from a two-sided t-test displaying the H2A interactome vs H2A + 2.5 µM pinometostat. Pinometostat target DOT1L is enriched in untreated cells, in addition to several proteins associated with loss of H3K79me. **d,** Summary of enriched proteins and their biological relevance is shown. **e,** RNA Pol II transcription is controlled by phosphorylation of the CTD. AT7519 inhibits CTD phosphorylation, stalling transcription. **f,** Volcano plot derived from a two-sided t-test displaying RPB1-Ir vs RPB1-Ir + 2 µM AT7519. Dotted lines indicate cut-offs employed. FDR values were calculated using the Benjamini–Hochberg procedure, as described in the Methods.