PORTLAND PRESS

Research Article

Histone H3 N-terminal mimicry drives a novel network of methyl-effector interactions

Jianji Chen[1,2]**, John Horton**[3]**, Cari Sagum**[1]**, Jujun Zhou**[3]**, Xiaodong Cheng**[3]** and ⓘ Mark T. Bedford**[1]

[1]Department of Epigenetics and Molecular Carcinogenesis, The University of Texas MD Anderson Cancer Center, Smithville, TX 78957, U.S.A.; [2]Graduate Program in Genetics & Epigenetics, The University of Texas MD Anderson Cancer Center UTHealth Graduate School of Biomedical Sciences, Houston, TX 77030, U.S.A.; [3]Department of Epigenetics and Molecular Carcinogenesis, The University of Texas MD Anderson Cancer Center, Houston, Texas 77030, U.S.A.

Correspondence: Mark T. Bedford (mtbedford@mdanderson.org)

OPEN ACCESS

The reader ability of PHD fingers is largely limited to the recognition of the histone H3 N-terminal tail. Distinct subsets of PHDs bind either H3K4me3 (a transcriptional activator mark) or H3K4me0 (a transcriptional repressor state). Structural studies have identified common features among the different H3K4me3 effector PHDs, including (1) removal of the initiator methionine residue of H3 to prevent steric interference, (2) a groove where arginine-2 binds, and (3) an aromatic cage that engages methylated lysine-4. We hypothesize that some PHDs might have the ability to engage with non-histone ligands, as long as they adhere to these three rules. A search of the human proteome revealed an enrichment of chromatin-binding proteins that met these criteria, which we termed H3 N-terminal mimicry proteins (H3TMs). Seven H3TMs were selected, and used to screen a protein domain microarray for potential effector domains, and they all had the ability to bind H3K4me3-interacting effector domains. Furthermore, the binding affinity between the VRK1 peptide and the PHD domain of PHF2 is ∼3-fold stronger than that of PHF2 and H3K4me3 interaction. The crystal structure of PHF2 PHD finger bound with VRK1 K4me3 peptide provides a molecular basis for stronger binding of VRK1 peptide. In addition, a number of the H3TMs peptides, in their unmethylated form, interact with NuRD transcriptional repressor complex. Our findings provide *in vitro* evidence that methylation of H3TMs can promote interactions with PHD and Tudor domain-containing proteins and potentially block interactions with the NuRD complex. We propose that these interactions can occur *in vivo* as well.

Introduction

Serving as a central hub and carrying a variety of histone post-translational modifications (PTMs), histone N-terminal tails orchestrate the recruitment of protein complexes that change the chromatin landscape at the target gene loci [1–5]. A combination of different PTMs on histone tails, including methylation, acetylation, and phosphorylation, has been termed the 'histone code' [1,5]. Single or combinatorial histone codes dictate the formation of specific chromatin-associated protein complexes that are involved in chromatin remodeling and transcriptional regulation. Histone H3K4me3 is found at the transcription start site (TSS) of almost every active gene, and is therefore recognized as a transcriptional activation mark [6]. Transcriptional initiation encoded by H3K4me3 is accomplished by the direct recruitment of reader modules from the transcription machinery and chromatin remodeling complexes [7–9]. There are two major H3K4me3 recognition module families: the royal superfamily of Chromo and Tudor-domains, and the PHD-finger family [10–14]. This study is focused on the PHD-finger superfamily because of its unique ability to recognize the tip of histone tails in a sequence-specific and modification-specific manner.

PHD fingers are characterized by a conserved zinc-coordinating Cys4-His-Cys3 motif [15]. On one hand, PHD fingers exhibit high sequence variability, providing various binding specificity towards H3K4me0 and K3K4me2/3 as the two major PHD ligand classes [7,16–19]. On the other hand, PHD

fingers share two adjacent ligand-binding sites that help stabilize the positively charged side chains of R2 and K4 within the histone H3 sequence (ARTKS) [16–19]. A clear example of this can be observed in BPTF (PHD), whose R2 is anchored in place by hydrogen bonds and electrostatic interactions with negatively charged D27 while its K4me3 is stabilized by van der Waals and cation-π interactions within an aromatic cage [18]. Notably, the free N-terminal amine forms a pair of hydrogen bonds with adjacent backbone carbonyls, enhancing the binding specificity of BPTF–H3K4me3 [18]. The free N-terminus of histone H3 is generated by enzymes that facilitate the cleavage of the initial methionine (iMet) [20].

Initiator methionine cleavage is a co-translational process that is mediated by the methionine aminopeptidases (MAP). There are two MAPs in yeast and humans, MAP1 and MAP2 [21,22]. Around two-thirds of proteins in the human proteome undergo initial methionine cleavage. The specificity of this process is determined by the size of the amino acid that is immediately adjacent to the iMet [23]. The initiator methionine is usually cleaved when the next residue has a small side chain: glycine, alanine, serine, threonine, cysteine, proline, and valine. Histone H3 has alanine at the N-terminus; therefore, its iMet is cleaved, generating a free N-terminal amino group that stabilizes the H3–PHD finger interactions.

Non-histone proteins with a histone-like sequence motif are called histone mimics. These proteins have the ability to mimic the H3 amino tail's interaction with modification-specific recognition modules [24–26]. One example of these mimics is the histone H3 methyltransferase G9a, which carries a ^{163}ARKT^{166} sequence that is similar to the ^{7}ARKS^{10} of histone H3. Similar to H3K9, the K165 residue can be self-methylated by G9a, and K165me2/3 interacts with the H3K9me2/3 effector HP1 [25,27,28]

Another example is the nonstructural protein 1 (NS1) of the influenza A virus. NS1 has a C-terminal ^{226}ARSK^{229} sequence that resembles the N-terminal ^{1}ARTK^4 of histone H3. This C-terminal sequence of NS1 interacts with the human PAF1 transcription elongation complex (hPAF1c), suppressing antiviral gene expression by blocking hPAF1c-mediated transcription elongation [24]. The dimethylated K229 of NS1, like dimethylated H3K4, interacts with the double chromodomains of chromatin remodeling protein CHD1 [29], and the CW domain of the MORC3 ATPase [26]. In addition, we have identified an internal H3-like ^{173}ARTK^{176} motif in PRMT6 that can be automethylated at the R residue. Notably, all the H3-mimicry proteins discussed above do not have the histone mimetic sequence at their N-termini, which could potentially undermine their ability to hijack PHDs that prefer binding the H3 N-terminal tail, and require iMet cleavage of their ligands.

This study seeks to address whether there are proteins in the human proteome that have an H3-like N-terminal motif and whether they can interact with known H3K4me3-reader proteins in a methylation-dependent manner, which could expand the PHD interaction network. H3 N-terminal mimicry proteins (H3TMs) are defined as proteins that start with an iM-Z1-R2-X3-K4 sequence motif. In this motif, Z1 represents one of the seven smaller amino acids that allows for iMet cleavage, and X3 represents any amino acid. First, we searched the database for proteins that harbor this relaxed motif when it is positioned just after the iMet residue, and we identified 48 H3TMs. We chose seven of these H3TMs as candidates for further *in vitro* validation experiments. Next, we used an in-house protein domain microarray, which harbors roughly 200 potential methyl-effector domains, to identify trimethylation-specific interaction profiles between the seven H3TMs and eight arrayed PHD fingers (and one Tudor domain-containing protein). We validated a subset of these interactions by peptide pull-down of recombinant GST-fusion proteins and their endogenous counterparts from total cell lysates. Using the same pull-down approach, we further show that a number of the unmethylated H3TMs interact with the NuRD complex. We also explored the possibility that the K4 site of H3TMs could be post-translationally modified by three known H3K4 methyltransferases: Set7/9, PRDM9 and the MLL1 complex. Furthermore, structural studies reveal how the PHD of PHF2 interacts with one of the methylated H3TMs (VRK1). This body of work demonstrates that K4 methylation of H3TMs promotes a new interaction network with PHD fingers, and potentially blocks their interactions with the repressive NuRD complex.

Materials and methods
Generation of methyl reader domain array and peptide synthesis

The current methyl reader domain array contains 108 PHD, 40 Tudor, and 31 Chromo domains. Protein domain microarray generation and probing have been published [30]. Briefly, recombinant proteins were synthesized and cloned into pGEX-4T-1 vector by Biomatik Corporation. These GST-fused reader domains were subsequently expressed, purified, and spotted in duplicate onto nitrocellulose-coated glass slides (Oncyte

PORTLAND
PRESS

Avid slides, Grace Bio-Labs) using a pin arrayer (Aushon 2470, Quanterix). For probing, microarray slides were blocked with 3% milk, 3% bovine serum albumin, 0.1% Tween 20 in PBS.

Methylated and unmodified forms of the following peptides were supplied by CPC Scientific Inc (China): VRK1 (2–13): PRVK(me3)AAQAGRQS-Biotin, BCL11B (2–13): SRRK(me3)QGNPQHLS-Biotin,

TSHZ1 (2–13): PRRK(me3)QQAPRRSA-Biotin, HIVEP1 (2–13): PRTK(me3)QIHPRNLR-Biotin,

PCLAF (2–13):VRTK(me3)ADSVPGTY-Biotin, MAPK8(2–13): SRSK(me3)RDNNFYSV-Biotin,

MEF2C(2–13):GRKK(me3)IQITRIMD-Biotin, HIRIP3(2–13): AREKme3EMQEFTRS-Biotin,

Histone H3 (1-18):ARTK(me3)QTARKSTGGKAPRK-biotin.

Biotinylated peptides were pre-labeled with streptavidin-Cy3 fluorophore (GE Healthcare) and incubated with the blocked array slides. Slides were then washed with PBST and allowed to air dry. Fluorescent interactions were visualized using a GenePix 4200A Microarray Scanner (Molecular Devices) and quantified using the Mapix software platform (Innopsys Inc).

Peptide pull-down assays

Biotinylated peptides (6 µg) were immobilized on 25 µl of streptavidin beads (Sigma, St Lous, MO, U.S.A.) in 500 µl of binding buffer (50 mM Tris-HCL pH 7.5, 150 mM NaCl, 1 mM EDTA, 1 mM EGTA, 0.5% NP-40) at 4°C for 1 h. The beads were washed with binding buffer three times and incubated with 10 µg of GST fusion protein or 30 µg 293T cell lysates for 1 h with rotation at 4°C. After three washes with binding buffer, the beads were boiled in protein loading buffer.

Western blot

Western blot analysis was performed as described previously [31]. Primary antibodies against PHF2 and MTA1 were purchased from Cell Signaling (Danvers, MA). The anti-ING2 antibody was purchased from Proteintech Group (Chicago, IL). The anti-GST antibodies was purchased from Covance (Berkeley, CA). The anti-SPIN1 antibody and RP-conjugated streptavidin antibody were purchased from ThermoFisher Scientific (Rockford, IL).

In vitro methylation assay

The GST-Set7/9 was expressed and purified, as described previously [32]. The MLL1 complex and PRDM9 protein were purchased from Reaction Biology Corp (Malvern, PA) and Active Motif (Carlsbad, CA), respectively. In vitro methylation reactions were performed in a final volume of 30 µl of 50 mM Tris–HCl (pH 8.5), 5 mM MgCl2, 4 mM DTT, and 0.42 µM ^{3}H-labeled S-adenosyl-L-[methyl ^{3}H] methionine (PerkinElmer Life Sciences) [33,34].

Isothermal titration calorimetry

GST tagged human PHF2 (residues 1–451; pXC816 or residues 1–71; pXC2227) was expressed in BL21(DE3) CodonPlus cells using auto-induction medium [35]. Briefly, an overnight cell culture in MDAG medium was inoculated to ZYM-5052 medium and cultured at 37°C until the OD_{600} reached to 1. Temperature was changed to 22°C and cells were cultured for 18 h for autoinduction of target protein. Protein was purified as previously described [36]. Two peptides were purchased from BioMatik representing the N-termini of VRK1 (PRV[Kme3] AAQAGRQS) and BCL11B (SRR[Kme3]QGNPQHLS). Isothermal titration calorimetry (ITC) experiments were performed with MicroCal PEAQ-ITC automated system (Malvern). Experiments were conducted at 25°C with a reference power of 10 µcal/s. Nineteen injections of 250 µM of a peptide of VRK1 or BCL11B with initial injection of 0.2 µl followed by eighteen injections (each of 2 µl) were titrated into 20 µM PHF2 (residues 1–451) in 300 mM NaCl, 20 mM HEPES pH 8.0, 0.5 mM TCEP and 5% glycerol. ITC data were fitted as 'one site' with the offset subtracted and binding constants were calculated using ITC analysis module supplied by the manufacturer.

Crystallography

Concentrated PHD domain of PHF2 (10–20 mg/ml in 200 mM NaCl and 20 mM Tris–HCl, pH 7.5) was mixed with the VRK1 peptide at an approximate 1 : 1.5 molar ratio (PHD/peptide) for ~1 h on ice. The Art Robbins Gryphon Crystallization Robot set up 0.4 ml sitting drops (0.2 ml of complex plus 0.2 ml of well solution) at ~19°C utilizing screens from Hampton Research and Molecular Dimensions. The initial observed crystals were conglomerations of needles and plates. Subsequent crystallization using hanging drop technique at

3.5 M sodium formate and 100 mM Tris–HCl (pH 8.5) generated larger, but were still not single, crystals. The crystals were manually broken into smaller single pieces, picked up in nylon loops and directly flash frozen in liquid nitrogen as high concentrations of organic salts are themselves cryoprotectants (PMID: 20606259).

Crystals were screened and X-ray diffraction data were collected at the SER-CAT beamline 22ID of the Advanced Photon Source at Argonne National Laboratory. Crystallographic datasets were processed with HKL2000 [37]. Molecular replacement was performed with the PHENIX PHASER module [38] by using the known structure of the PHD domain (PDB ID 3KQI) as a search model. Structure refinement was performed with PHENIX Refine [39] with 5% randomly chosen reflections for the validation by the R_{free} value. The crystallographic temperature B-factors for individual atoms (except for water) were refined anisotropically in final refinement rounds. COOT [40] was used for the initial manual building and corrections between refinement rounds. Structure quality was analyzed during PHENIX refinements and finally validated by the PDB validation server. Molecular graphics were generated by using PyMol (Schrodinger, LLC).

Results

Bioinformatics search for histone H3 N-terminal mimicry proteins

To investigate whether a certain class of protein might be enriched for H3TMs, we focused on the over 20 thousand proteins that are functionally annotated in the Panther database (http://pantherdb.org/) (Figure 1A). To identify H3 N-terminal mimicry proteins that undergo initial methionine cleavage and contain RxK motifs, we searched for M-Z-R-X-K motif-containing proteins using the Motif Search database (https://www.genome. jp/tools/motif/MOTIF2.html). Z represents alanine, glycine, serine, cysteine, threonine, proline, or valine, which is a critical component of this motif because these seven amino acids are permissive for initiator methionine cleavage [23 41,42]. X represents all 20 possible amino acid residues. We excluded histone H3 variants from the results and identified 257 M-Z-R-X-K motif-containing proteins, which harbored this motif in either the body of the protein (Figure 1B) or at the protein's N-terminus (Figure 1C). About 1/5 of these proteins have the motif restricted to their N-terminus; specifically, 48 of these proteins possessed the motif at their N-termini (Figure 1C and Supplementary Table S1), while the remaining 209 proteins possessed the motif at the internal regions (Figure 1B). Functional annotation shows that the slice of the pie chart represented by 'chromatin-binding protein' does not increase if the M-Z-R-X-K motif is internal (Figure 1A,B). However, if the iM-Z-R-X-K motif was restricted to the N-terminus of proteins, there was a dramatic enrichment in the fraction of proteins that are functionally annotated as H3TMs (Figure 1C, blue wedge). These results reveal selective enrichment of H3TMs proteins in the 'chromatin-binding protein' class, which indicates that these types of protein (transcription factors and transcriptional co-regulators) may engage the PHD interaction network.

Phylogenetic analysis of the 48 proteins that harbor iM-Z-R-X-K motifs, identified several families of transcriptional factors with conserved N-terminal sequences: MEF2A/B/C/D, SALL1/2/3/4, BCL11A/B, FOG1/2, and TSHZ1/2/3 (Figure 1D). Notably, SALL1[43], FOG1 [44,45], and BCL11A/B [46,47] are known to interact with the NuRD corepressor complex via their N-termini. We chose to perform follow-up studies on one protein from each of the families, MEF2C, BCL11B, and TSHZ1. VRK1 is a nuclear serine/threonine kinase that phosphorylates several essential proteins including p53/TP53, histone H3, BANF1, and ATF2 [30,48–50]. Importantly, VRK1 K4me1 has been identified using mass spectrometry (MS) implying K4 residue is likely a target for methylation *in vivo* [51]. PCLAF interacts with PCNA to regulate DNA repair during DNA replication [52,53]. MAPK8 is another serine/threonine kinase that involves cell proliferation, differentiation, migration, and apoptosis [54–59]. HIVEP1 is a zinc-finger protein that binds to the HIV-1 enhancer to activate HIV-1 gene expression [60]. Thus, we selected seven proteins from distinct families of transcription factors/co-regulators as H3TMs candidates for further evaluation, because of their distribution across the phylogenetic tree and their roles in transcriptional regulation: VRK1, BCL11B, TSHZ1, MEF2C, PCLAF, MAPK8, and HIVEP1 (Figure 1E). Then, we screened for potential novel interactions between these candidates and PHD finger proteins.

Identification of H3TMs interactions with methyllysine-effector domains

To identify potential interactions between the seven candidate proteins and methyl-binding domains, we synthesized N-terminal peptides (1–12 amino acids excluding the initial methionine) with and without tri-methylation at the K4 position. We used these biotinylated peptides as probes to interrogate an in-house
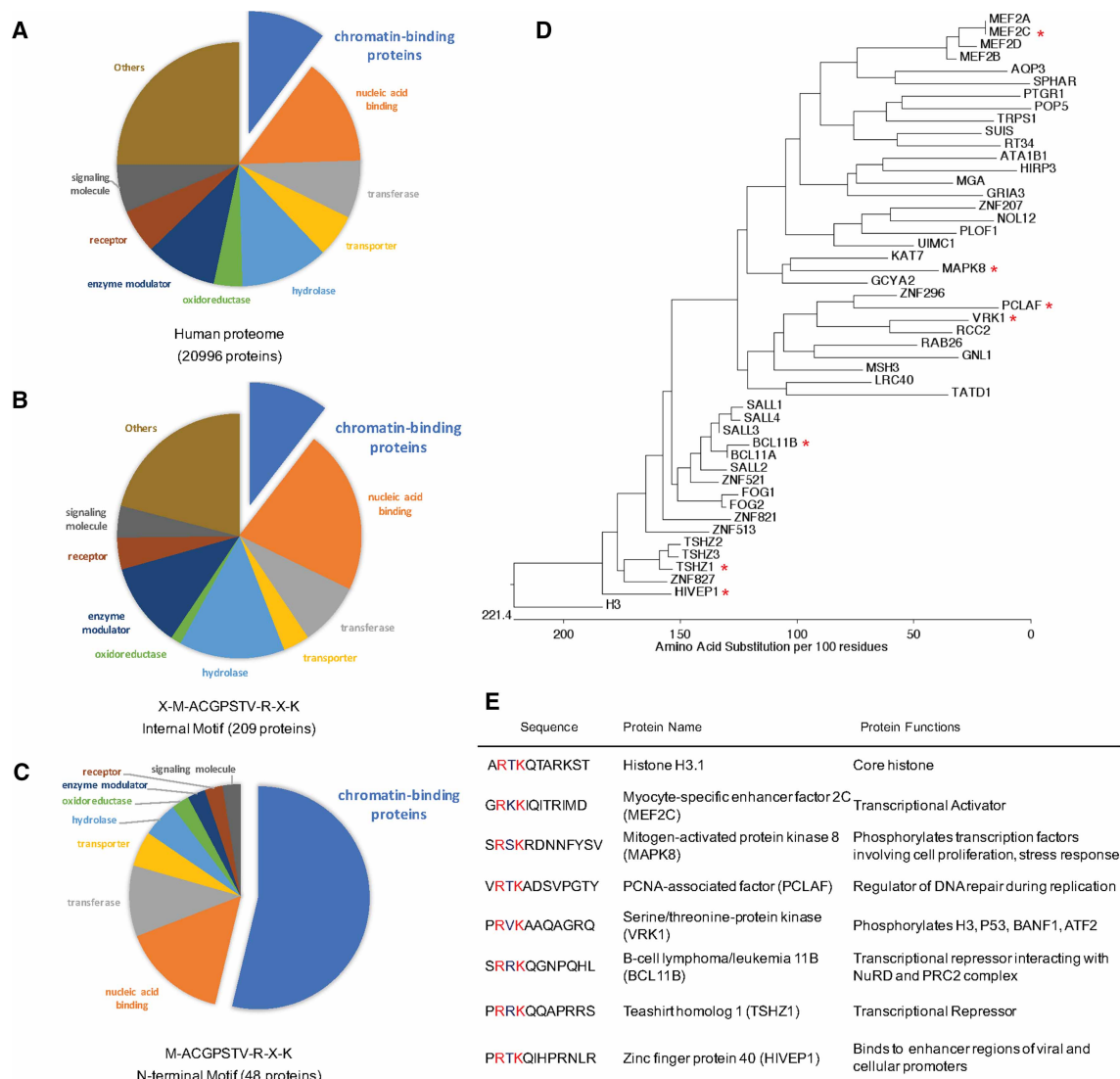
Figure 1. Identification of H3 N-terminal mimicry proteins.

PANTHER Protein Class annotation of 20996 proteins in the human proteome (**A**), 209 internal motif-containing proteins (**B**), and 48 N-terminal motif-containing proteins (excluding histone H3 and variants) (**C**). (**D**) Phylogenetic analysis of the N-terminal sequences of 48 H3TMs and histone H3. (**E**) Red asterisks indicate seven H3TMs candidates for follow-up study and their N-terminal sequences and functions are summarized in Figure S1.

methyl-reader microarray that harbors GST fusions of known and predicted binders of methyl-lysine and methyl-arginine marked motifs [61]. The current array contains 108 PHD, 40 Tudor, and 31 Chromo domains. As positive and negative controls, the histone H3K4me3 peptide interacted with known H3K4me3-interacting PHD finger domains and Tudor domains, while the unmodified peptide did not (Figure 2, top left panel). Interestingly, each of the seven peptide probes that harbored the K4me3 mark (Figure 2), displayed a unique interaction profile with known H3K4me3-interacting PHD finger domains [62], and many of these peptides also bound the Tudor domain-containing protein SPIN1, which is a well-characterized H3K4me3 reader [63]. The relative intensities of these interactions on the arrays were quantified (Supplementary Figure S1). Notably, compared with the other probes, VRK1 and BCL11B had a greater number of interacting reader domains. When the K4 position was tri-methylated, both VRK1 and BCL11B interacted with PHF2, TAF3, and KDM7A PHD fingers; however, VRK1 selectively interacted with DIDO1, ING1/2, and MLL5 PHD fingers, while BCL11B selectively interacted with ING3 and PHF23 PHD fingers. The TSHZ1 K4me3 peptide bound the
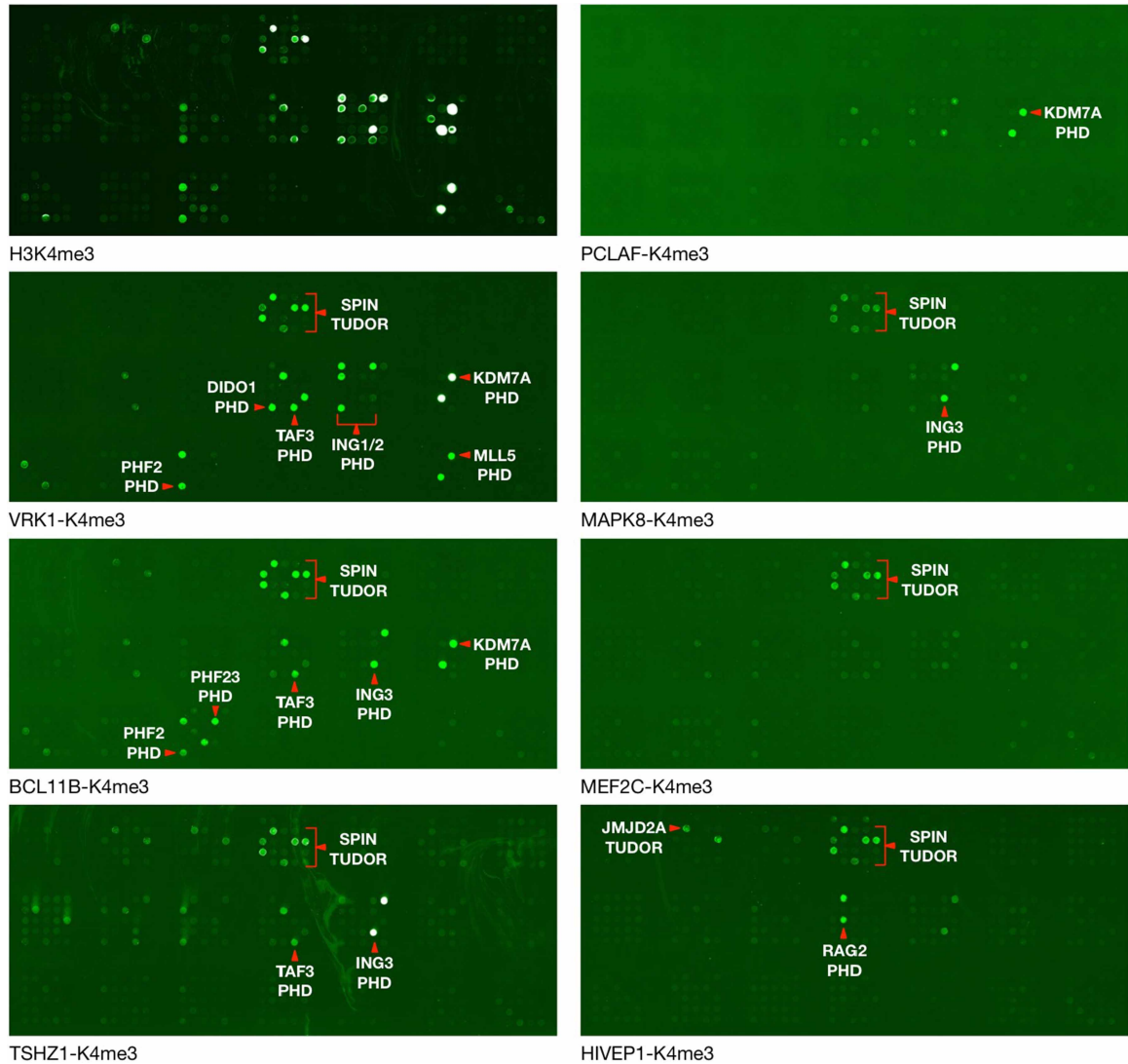
Figure 2. Identification of novel methylation-dependent interactions between H3TMs and PHD fingers.

N-terminal peptide (1–12 amino acids) were generated with K4me3 modified, for seven H3TMs and histone H3, and used to probe our methyl-reader microarray that contains 108 PHD fingers, 40 TUDOR and 31 Chromo domains. Duplicates of each protein domain were used to facilitate identification and reproducibility. The histone H3K4me3 peptide was used as positive control to visualize H3K4me3-interacting methyl reader domains.

PHD fingers of TAF3 and ING3, and PCLAF only recognized KDM7A. All the PHD fingers listed above have been shown to bind the histone H3K4me3 mark [7,8,19,64–67]. Also, as six of the seven methylated peptide probes interacted well with the arrayed SPIN1 Tudor domains, we included this domain as a positive control in the downstream validation experiments.

Validation of novel interactions using *in vitro* interaction assays

We used a peptide pulldown assay to confirm a small subset of the interactions we observed on the protein domain microarrays, and to validate the methyl-dependent nature of these bindings. GST fusions of PHD domains from PHF2, ING2 and ING3, and the Tudor domains of SPIN1 were used to perform pulldowns with biotinylated VRK1, BCL11B, TSHZ1 and histone H3 peptide pairs, which were both methylated and unmethylated. The PHF2 PHD finger strongly interacted with methylated peptides derived from VRK1, BCL11B, and

H3; whereas, it weakly interacted with the methylated form of TSHZ1. The ING2 PHD finger only recognized methylated VRK1 and H3 peptides. ING3 showed methylation-dependent interactions with TSHZ1 and H3, but could also interact with VRK1 in a methyl-independent fashion. SPIN1 interacted with VRK1, BCL11B, TSHZ1, and H3 peptides in their K4me3 form (Figure 3A).

Next, we asked if the same peptide sets could be used to pulldown the methyl-effector proteins from whole-cell lysates. We reasoned that the isolated recombinant protein domain may behave differently to the same domain when it is in the context of the full-length endogenous protein. Using antibodies that could recognize endogenous PHF2, ING2 and SPIN1, and whole-cell lysates pulldowns using the VRK1 peptides, we found that endogenous PHF2, ING2 and SPIN1 were indeed enriched with the methylated peptide. The BCL11B methyl-peptide enriched for endogenous SPIN1, and very weakly endogenous PHF2. The TSHZ1 methyl-peptide only enriched for endogenous SPIN1 (Figure 3B).

A number of the H3TMs listed in the phylogenetic tree (Figure 1D), harbor the ability to bind to the NuRD repressor complex, through their N-termini. For example, SALL1 [43], FOG1 [44,45], and BCL11A/B [46,47] all possess this property, and it is thought methylation at their K4 position will block this interaction just as it does for the interaction between NuRD and Histone H3 [68]. It has been shown that the MTA1 is a core subunit of the NuRD complex, and the C-terminal of MTA1 (454–715) recognizes the N-terminus of BCL11B [69]. In addition, the tandem PHD fingers of CHD4, which is also found in the NuRD complex, binds unmethylated histone H3 tails [70]. We next wanted to determine whether K4me3 has the ability to disrupt the interactions between the NuRD complex and VRK1, BCL11B, and TSHZ1. To do this, we again use the peptides (with and without methylation) to perform pulldowns from whole-cell lysates, and we probed these pull-downs with antibodies against MTA1 and RBAP46/48 (components of the NuRD complex). Indeed, as expected, we find that BCL11B and histone H3 peptides engage NuRD when unmethylated, and so does TSHZ1, which has not been reported before. The unmethylated VRK1 peptide does not have the ability to interact with the NuRD complex (Figure 3B). Thus, just like histone H3, a number of H3TMs can switch from binding activator complexes (in their methylated state) to binding the NuRD repressor complex (in their unmethylated state).
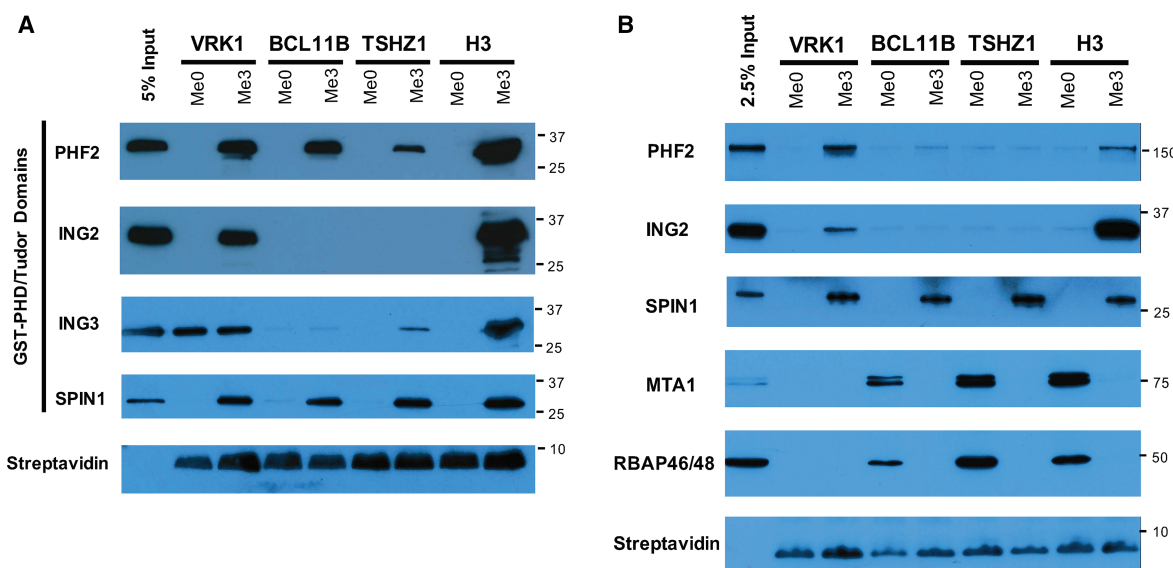

Figure 3. Validation of novel interactions using *in vitro* interaction assays.
(**A**) Peptides pull down GST-tagged PHD fingers. Unmodified and K4me3 peptide probes for VRK1, BCL11B, TSHZ1, and H3 were used to pull down GST-tagged PHF2 PHD, ING2 PHD, ING3 PHD, and SPIN1 TUDOR. (**B**) Peptides pull down endogenous PHD fingers and NuRD complex from 293T whole cell lysates. Unmodified and K4me3 peptide probes for VRK1, BCL11B, TSHZ1, and H3 were used to pull down PHF2, ING2, ING3, SPIN1 and NuRD complex components including MTA1 and RBAP46/48. Immunoblot using streptavidin-HRP was used to ensure equal loading of all biotin-tagged peptide probes. Molecular Weight markers (kDa) are shown to the right of each panel.
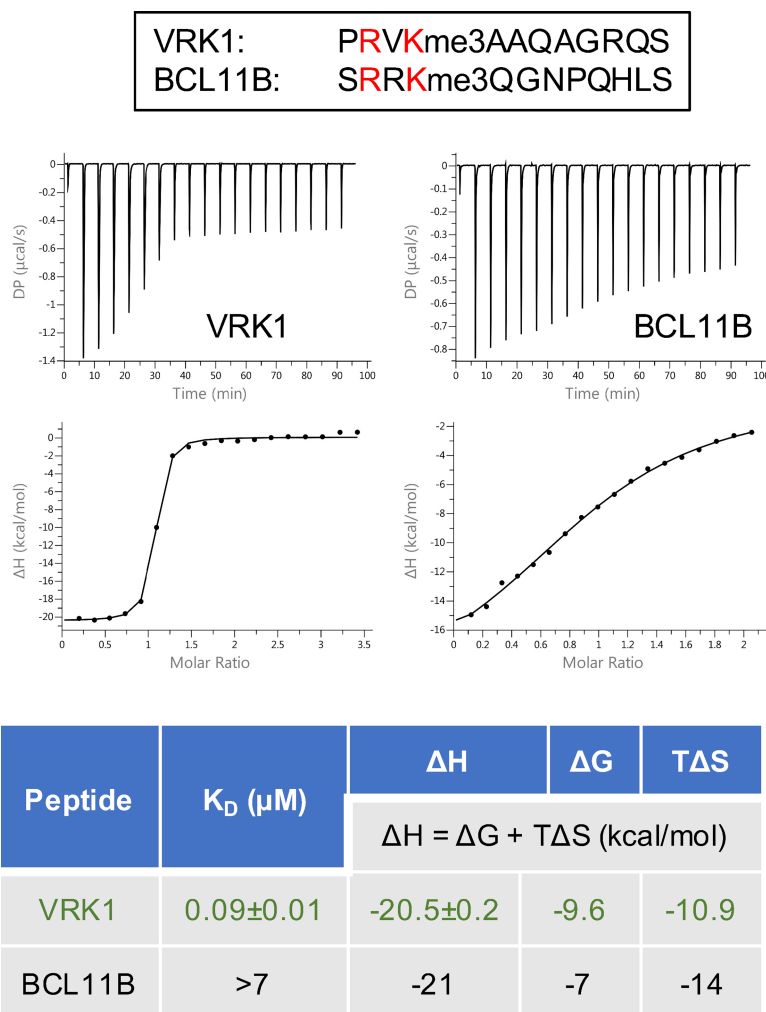
VRK1:   PRVKme3AAQAGRQS
BCL11B: SRRKme3QGNPQHLS

Peptide	K_D (µM)	ΔH	ΔG	TΔS
		ΔH = ΔG + TΔS (kcal/mol)		
VRK1	0.09±0.01	-20.5±0.2	-9.6	-10.9
BCL11B	>7	-21	-7	-14

Figure 4. Investigating the strength of the interaction between PHF2 and the N-termini of VRK1 and BCL11B.
Isothermal titration calorimetry measurements were performed with recombinant PHF2 (residues 1–451) and indicated peptides of VRK1 and BCL11B. The top panel is the raw data showing the heat released and measured by the sensitive calorimeter during gradual titration of the peptide into the sample cell containing PHF2 until the binding reaction has reached equilibrium. The bottom panel shows each peak in the raw data is integrated and plotted versus the molar ratio of peptide to protein. The resulting isotherm can be fitted to a binding model from which the affinity (K_D) is derived. The Y-axis measures enthalpy change (ΔH) using the relationship ΔH = ΔG + TΔS where ΔG is the Gibbs free energy, ΔS is the entropy change and T is the absolute temperature.

In our pulldown assays, performed from whole-cell lysates, we noticed that substantially more PHF2 was enriched using the VRK1-K4me3 peptide as compared with the H3K4me3 peptide (Figure 3B). PHF2 fragment (residues 1–451) containing both PHD domain and Jumonji domain has been reported to have a binding affinity (K_D) of 230 nM for H3K4me3 peptide [36], and we performed isothermal titration calorimetry (ITC) to determine the binding affinity of the same PHF2 fragment for the methylated N-terminus of VRK1. As suggested by the pulldown, we observe an almost three-fold stronger binding of PHF2 for VRK1, than for H3K4me3 (90 nM vs 230 nM). The binding affinity of PHF2 for BCL11B was much weaker, at greater than 7 µM (Figure 4). Thus, there are a myriad of interactions that can potentially occur with H3TMs, either in their methylated or unmethylated states, and the high binding affinity between the VRK1 K4me3 peptide and the PHF2 PHD domain, suggests a competition model where PHF2 might be recruited to chromatin by H3K4me3, and then handed off to VRK1-K4me3.

Structural comparison of PHF2 PHD binding to VRK1-K4me3 and histone H3K4me3

We next sought to understand how PHF2 recognizes the VRK1 tri-methylated peptide. Accordingly, we purified the PHD domain of PHF2 (residues 1–70) and grew co-crystals with VRK1 peptide (residues 1–12). The complex crystallized in space group *I*222, resulting in a structure determined to a resolution of 1.2 Å (Table 1). Although 12 residues of VRK1 peptide was used for crystallization, we observed the N-terminal seven residues

Table 1 Summary of X-ray diffraction and refinement

PHF2 with VRK1 peptide	PDB 7M10		
Wavelength (Å)	1.0000		
Space Group	*I*222		
Cell dimensions (Å)	34.24, 63.72, 103.15		
α, β, γ (°)	90, 90, 90		
Resolution (Å)	32.15–1.15	(1.24–1.20)	(1.19–1.15)
R_{merge}	0.106	(0.579)	(0.656)
R_{pim}	0.034	(0.310)	(0.370)
CC1/2	-	(0.740)	(0.810)
$\langle I/\sigma I \rangle$	20.1	(1.9)	(1.5)
Completeness (%)	81.5	(51.7)	(24.7)
Redundancy	9.1	(3.7)	(2.3)
Unique reflections	33 230	(1839)	(991)
Observed reflections	316 530		
Refinement			
Resolution (Å)	1.15		
No. reflections	33 196		
R_{work}/R_{free}	0.157/0.165		
No. Atoms			
Protein	585		
Peptide	81		
Zinc	2		
Solvent	142		
B Factors (Å^2)			
Protein	17.6		
Peptide	20.5		
Zinc	13.1		
Solvent	33.1		
R.m.s. deviations			
Bond lengths (Å)	0.004		
Bond angles (°)	0.8		

*Values in parenthesis correspond to the two highest resolution shells. The shell of 1.19–1.15 Å has ~25% completeness and its 991 reflections were included in the refinement.
$R_{merge} = \Sigma|I - \langle I \rangle|/\Sigma I$, where I is the observed intensity and $\langle I \rangle$ is the averaged intensity from multiple observations.
$\langle I/\sigma I \rangle$ = averaged ratio of the intensity (I) to the error of the intensity (σI).
$R_{work} = \Sigma|F_o - F_c|/\Sigma|F_o|$, where F_o and F_c are the observed and calculated structure factors. R_{free} was calculated using a randomly chosen subset (5%) of the reflections not used in refinement.
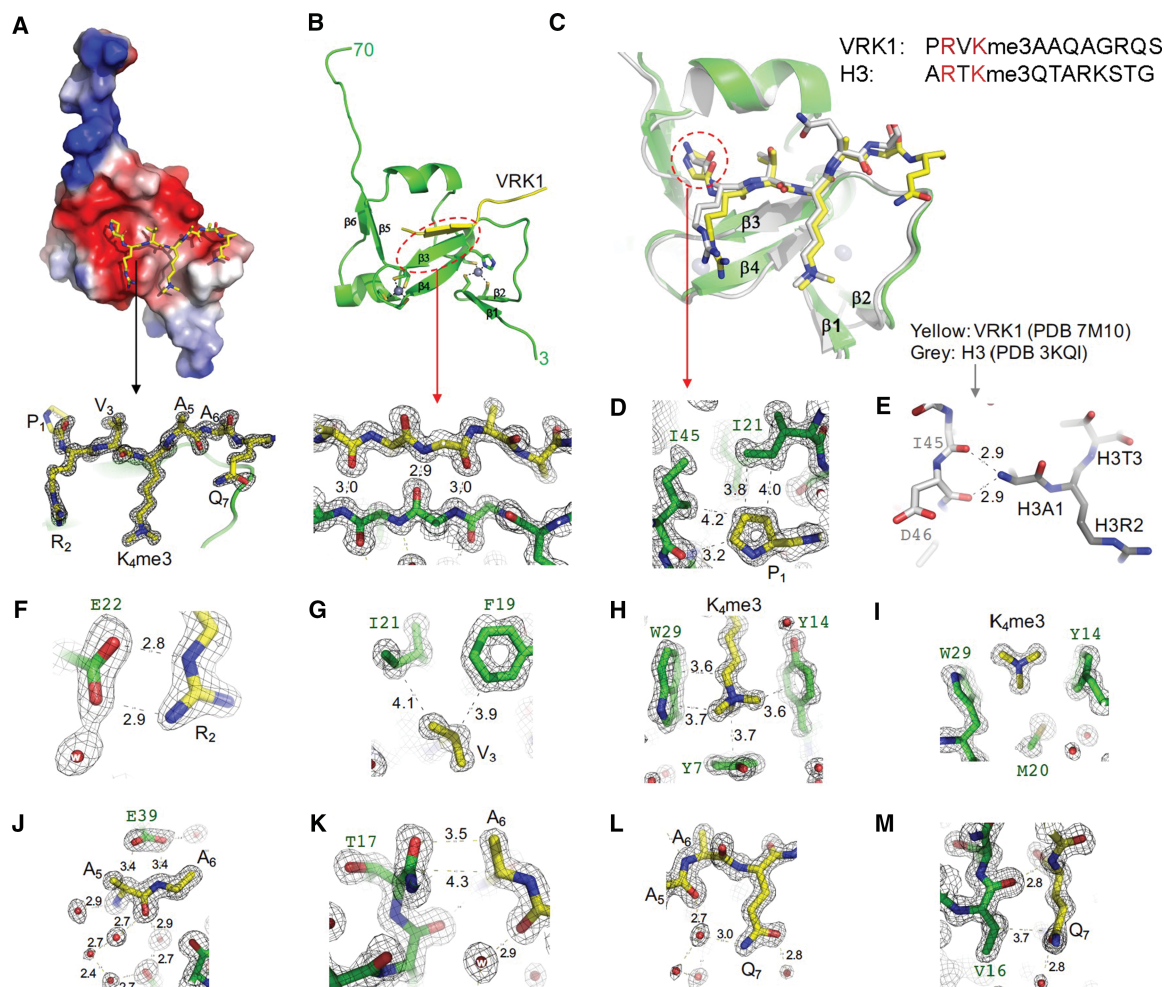
Figure 5. Structure of PHD domain of PHF2 (residues 1–70) in complex with VRK1 peptide.
(**A**) A surface representation of the PHD domain displayed in red for negative, blue for positive and white for neutral. The bound VRK1 peptide is shown as a stick model in yellow. The bottom panel shows Fo-Fc omit electron density for the peptide, contoured at 5s above the mean. (**B**) The PHD domain in green is displayed in a ribbon model with the bound VRK1 peptide in yellow. The two zinc ions are coordinated with Cys4 and His-Cys3, respectively. The bottom panel shows antiparallel arrangement between VRK1 peptide (in yellow) and the PHF2 strand (in green). The 2Fo-Fc electron density is contoured at 2s above the mean. (**C**) Superimposition of PHF2 with bound VRK1 peptide (yellow) and histone H3 peptide (grey). (**D**) Interaction with VRK1 P1 residue. (**E**) Interaction with histone H3 A1 residue (PDB 3KQI). (**F**) Interaction with VRK1 R2 residue. (**G**) Interaction with VRK1 V3 residue. (**H**) Aromatic cage for binding K4me3. (**I**) The bound K4me3 from the viewpoint of Tyr7 of PHF2. (**J** and **K**) Interactions with VRK1 A5 and A6 residues. (**L** and **M**) Interactions with VRK1 Q7.

(Figure 5A). Like other structurally characterized PHD domains [67], VRK1 peptide in an extended conformation lies in an acidic surface of the PHD (Figure 5A), and is inserted antiparallel to the strand β3 of PHD via main-chain hydrogen bonds (Figure 5B), effectively expanded the two-stranded sheet (β3 and β4) to three-stranded β sheet. The structure is highly similar to that of previously determined PHD-H3K4me3 peptide [67], with a root-mean-square-deviation of <0.5 Å when pairwise superimposed (Figure 5C). The largest difference between the two bound peptides is the interactions with the first residue. In VRK1, the bulkier proline forms van der Waals contacts with two hydrophobic residues Ile21 and Ile45 (Figure 5D). In addition, the main chain carbonyl oxygen of Ile45 interacts with the carbon Cδ of pyrrolidine ring via an O⋅⋅H-C type hydrogen bond [71]. We note that the secondary amine of the VRK1 proline is not involved in direct interaction with the protein. In histone H3 peptide, the N-terminal amino group NH_3 forms two hydrogen bonds with the backbone carbonyl oxygen atoms of Ile45 and Asp46 (Figure 5E).

As noted above, the positively charged guanidino group of VRK1 R2 interacts with the negatively charged Glu22 of PHF2 (Figure 5F). VRK1 V3 forms van der Waals contacts with Ile21 and Phe19 (Figure 5G). The tri-methylated K4 of VRK1 is bound in a three-sided cage formed by Tyr14, Met20 and Trp29, and capped with Tyr7 of PHF2 (Figure 5H,I). Glu39 of PHF2 forms two weak interactions with A5 (via the main-chain Cα atom) and A6 (via the main-chain amide nitrogen) of VRK1 (Figure 5J). The associated main-chain amide nitrogen and carbonyl oxygen of A5 of VRK1 is also involved in a network of water-mediated interactions (Figure 5J). The hydrophobic C_α-C_β carbon atoms of VRK1 A6 are in van der Waals contacts with Thr17 of PHF2 (Figure 5K). The last ordered side chain of VRK1 Q7 has a water-mediated intramolecular interaction with main-chain carbonyl oxygen of A5 (Figure 5L) and a van der Waals contact with Val16 of PHF2 (Figure 5M). In sum, PHF2 provides extensive interactions with the first seven residues of VRK1, involving electrostatic, hydrogen bonds, aromatic cage, van der Waals contacts in both side-chain and main-chain atoms.

The K4 residue of three H3TMs can be methylated by H3K4 methyltransferases *in vitro*

Importantly, VRK1-K4me1 has been identified using MS implying K4 residue is likely a target for methylation *in vivo* [51]. In addition, three independent MS studies performed by Cell Signaling Technology (CST), and publicly available through the PhosphoSitePlus website (www.phosphosite.org), have also identified mono-methylation at the VRK1-K4 position. All of these reported studies focused on enriching for Kme1 peptides, either using methyl-specific Kme1 antibodies (CST) or using a unique chemical derivation approach followed by antibody enrichment [51]. These studies suggest that VRK1 is indeed targeted for methylation at its K4 position, just like H3K4, and that higher degrees of methylation (me2/3) may occur.

To experimentally test if H3TMs have the potential of being methylated by enzymes that modify the H3K4 site, we set up a number of *in vitro* methylation reactions. Three different lysine methyltransferases that have been reported to deposit methyl marks at the histone H3K4 position were used: Set7/9 [72], and MLL1 complex [73,74] and PRDM9 [75]. We performed *in vitro* methylation using the same set of peptide pairs used in previous experiments. For each *in vitro* methylation reaction, enzyme and ^{3}H-labeled SAM in the absence of peptide served as a negative control, while H3K4 peptide served as a positive control. We showed that Set7/9 could methylate unmodified peptides from all three candidate proteins VRK1, BCL11B and TSHZ1, but could not methylate the K4me3 peptides, indicating the methylation is focused at the K4 sites (Figure 6). MLL1 and
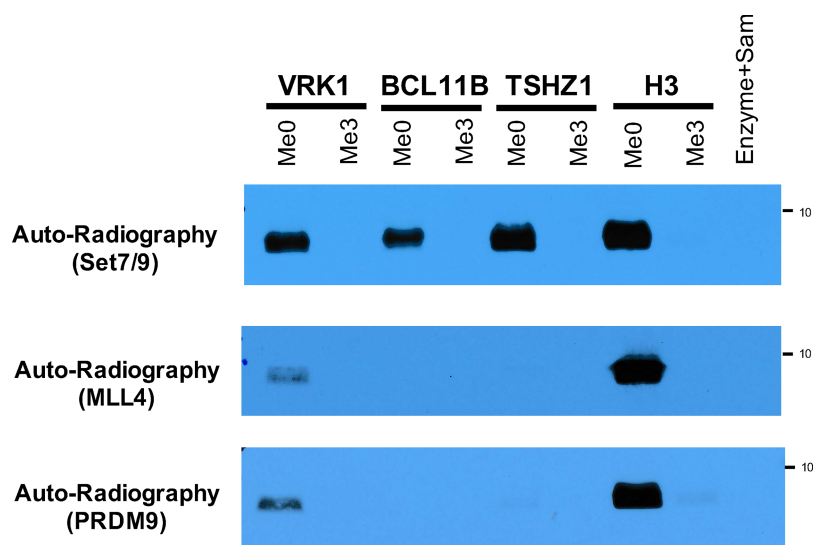

Figure 6. Three H3TMs can be methylated at Lys4 by H3K4 Methyltransferases.
Unmodified and K4me3 peptide probes for VRK1, BCL11B, TSHZ1, and H3 were incubated with three H3K4-specific methyltransferases, Set7/9 (upper), MLL1 (middle) and PRDM9 (lower). Enzymes and ^{3}H-labeled Sam were used as a negative control (most right lane). Immunoblot using streptavidin-HRP was used to ensure equal loading of all biotin-tagged peptide probes. Molecular Weight markers (kDa) are shown to the right of each panel.

PRDM9 can only methylate VRK1 with moderate activity, compared with H3 peptide, suggesting VRK1 is a weaker substrate than histone H3 for MLL1 and PRDM9 *in vitro*. These data support the hypothesis the at least a subset of H3TMs can be targeted by both the readers and writers of the histone H3K4me3.

Discussions

In this study, we proposed the concept of histone H3 N-terminal mimicry and provided *in vitro* biochemical evidence that H3TM N-terminal K4me3 serves as a molecular switch that engages different protein interaction networks. These findings may dramatically expand the scope of functions that have been bestowed on methyl-effector proteins and the NuRD complex beyond the original histone code hypothesis, by providing evidence that PHD fingers (and the Tudor domain of SPIN1) recognize not only histone proteins, but also non-histone proteins.

For these findings to be relevant, they of course need to also occur in cells and *in vivo*. However, we did not find direct evidence for the existence of the H3TM-K4me3 methylation state *in vivo*. This is not through lack of trying. First, using MS (data not shown), we were unable to map the N-terminal PTMs for the six tested H3TM candidates (VRK1, BLC11B, MEF2C, TSHZ1, PCLAF and MAPK8), because, like histone H3, the H3TMs are enriched for lysine or arginine residues in the vicinity of the K4 residue, generating tryptic peptides
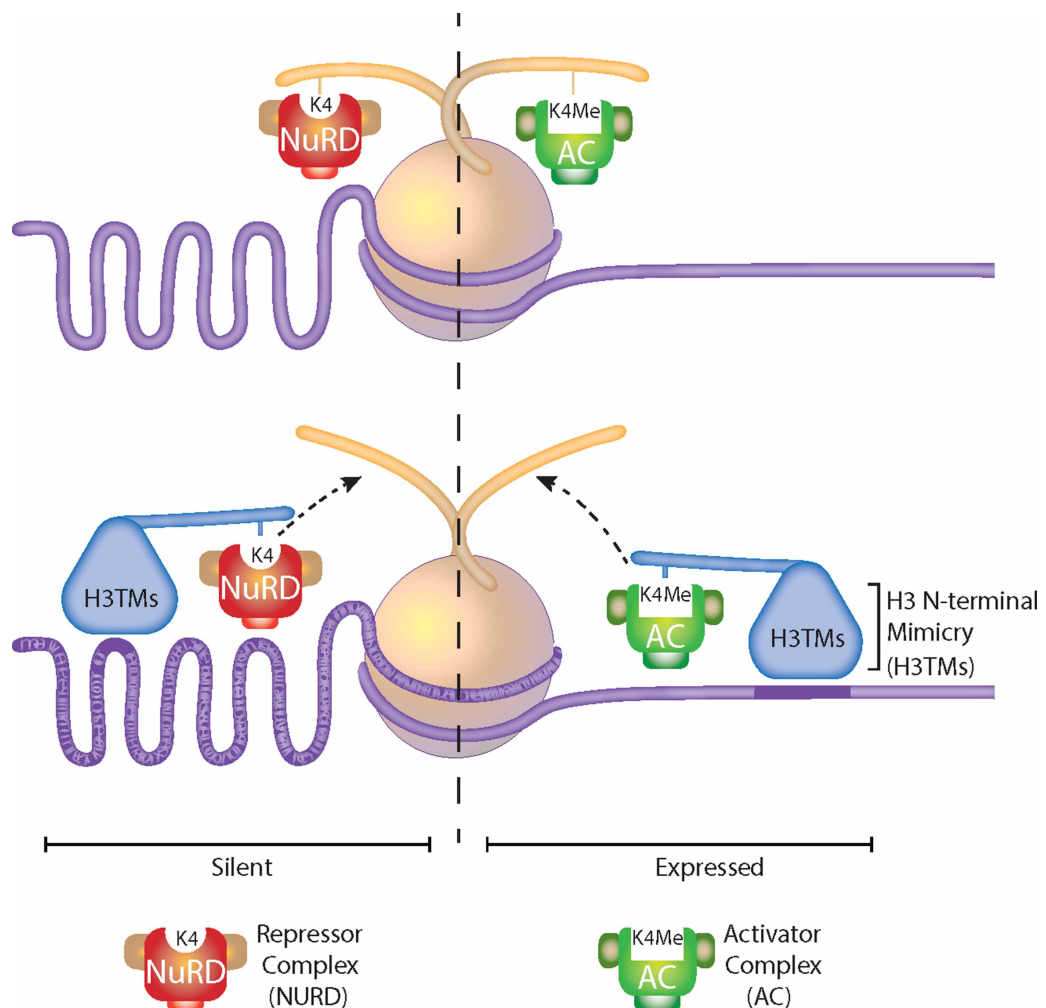

Figure 7. The Model of H3TM interactions.
In their methylated state, H3TMs can bind PHD and Tudor domains that are implicated in maintaining an open chromatin state, and are generally involved in promoting transcription. If the H3TMs are unmethylated they have the potential of interacting with the NuRD transcriptional repressor complex.

that are often very short and incompatible with MS analysis. Direct evidence for the occurrence of H3TM K4me3 will likely come from middle-/top-down mass spectrometry studies [76]. Second, we developed methyl-specific antibodies to K4me3 sites on VRK1, BCL11B and TSHZ1 (data not shown). While these antibodies selectively recognize the K4me3 peptides, and not the K4me0 peptides, in dot-blot assays, they were unable to detect these three endogenous H3TMs. This could be due to a number of reasons: (1) the antibodies are just bad; (2) the K4me3 mark occurs at very low levels, which make it difficult for the antibodies to detect; (3) the H3TMs may carry K4me1/2 marks that are not detected by our K4me3 antibodies, and (4) the K4me3 mark on these H3TMs may be cell type specific, tissue specific, cell cycle specific, or development specific, and we have missed them.

Although the cellular context of K4me3 remains unknown, the N-termini of H3TMs are disordered (using https://iupred2a.elte.hu/) as the H3 N-terminus is, implying that the N-terminal regions need to be stabilized by PTMs or by interacting with binding partners [77]. Indeed, H3TMs with unmodified N-terminal tails mimic unmodified histone H3. Several of the 48 H3TM candidates interact with the repressive NuRD complex [43–47,78,79]. Interestingly, the NuRD complex also recognizes the unmodified histone H3 N-terminus, but this interaction is disrupted by H3K4 trimethylation [68,80]. K4me3 prohibits the chromatin-association of the repressive NuRD complex, regulating local chromatin status to facilitate transcriptional activation. Among the 48 candidate proteins, SALL1/2/3/4, FOG1/2, and BCL11A/B are known to interact with the NuRD complex via their N-termini, repressing target gene expressions [43–47,78,79]. However, the mechanism by which these interactions are turned off remains unknown. Among the seven candidates that were used to probe the protein domain microarray (Figure 2), VRK1 K4me1, MEF2C K4ac, TSHZ1 K4ac, and BCL11B S2ph have been identified in various mass spectrometry studies [46,51,81,82]. Notably, BCL11B serine 2 phosphorylation has been shown to negatively regulate BCL11B's interaction with the NuRD complex [46]. Given that the N-terminal tail of histone H3 is heavily modified by various PTMs, such as phosphorylation, methylation, and acetylation, it is likely that the N-termini of these H3TM proteins are also subject to PTMs that positively or negatively regulate their ability to interact with recognition modules.

We have proposed a model that H3TMs bind the repressive NuRD complex in their unmethylated (K4me0) state, which is already established, and in their K4me3 state these same H3TMs directly interact with effectors (protein that harbor PHD fingers and Tudor domains) that have transcriptional activator activity (Figure 7). In cases where the binding affinity of the effector is greater for the H3TM than the for H3K4me3, then these effectors may already reside on active chromatin and then be transferred to the H3TM. If the binding affinities are reversed (higher for H3K4me3 than H3TM-Kme3) then the H3TM may 'deliver' the effector to chromatin.

Conclusions

As search for proteins that harbor N-termini with sequence similarity to histone H3, identified a striking enrichment (5-fold) of fraction of proteins with chromatin binding functions (Figure 1). Many of these H3TM proteins have already be shown to interact with the NuRD transcriptional repressor complex in their unmethylated states. We tested 7 of these methylated H3TMs for their ability to interact with known histone H3K4me3 binders, and they all do, to varying degrees. We thus hypothesize that the 48 H3TMs that we identified will be a major class of interacting proteins for what are traditionally thought of as H3K4me effector proteins.

Data Availability
All the data needed to evelute the conclusions that are drawn in this manuscript are available in the main text. There is no supplemental section for this manuscript. The X-ray structure (coordinates) and the source data (structure factor file) of PHF2 PHD domain with bound VRK1 peptide have been submitted to the PDB under accession numbers 7M10.

Competing Interests
M.T.B. is a co-founder of EpiCypher.

Open Access

Open access for this article was enabled by the participation of the University of Texas Southwestern Medical Centre in an all-inclusive *Read & Publish* pilot with Portland Press and the Biochemical Society under a transformative agreement with EBSCO.

CRediT Author Contribution

Mark T. Bedford: Conceptualization, Funding acquisition, Project administration, Writing — review and editing. **Jianji Chen:** Conceptualization, Formal analysis, Validation, Methodology, Writing — original draft. **John Horton:** Data curation, Investigation, Methodology. **Cari Sagum:** Data curation, Methodology. **Jujun Zhou:** Data curation, Validation, Methodology. **Xiaodong Cheng:** Resources, Funding acquisition, Validation, Methodology, Writing — review and editing.

Abbreviations

H3TM, H3 N-terminal mimicry proteins; iMet, initial methionine; ITC, isothermal titration calorimetry; MAP, methionine aminopeptidase; NS1, nonstructural protein 1; NuRD, the nucleosome remodeling and deacetylase complex; PHD, plant homeodomain finger; PTM, post-translational modification.

References

1 Jenuwein, T. and Allis, C.D. (2001) Translating the histone code. *Science* **293**, 1074–1080 https://doi.org/10.1126/science.1063127
2 Kouzarides, T. (2007) Chromatin modifications and their function. *Cell* **128**, 693–705 https://doi.org/10.1016/j.cell.2007.02.005
3 Ruthenburg, A.J., Li, H., Patel, D.J. and Allis, C.D. (2007) Multivalent engagement of chromatin modifications by linked binding modules. *Nat. Rev. Mol. Cell Biol.* **8**, 983–994 https://doi.org/10.1038/nrm2298
4 Taverna, S.D., Li, H., Ruthenburg, A.J., Allis, C.D. and Patel, D.J. (2007) How chromatin-binding modules interpret histone modifications: lessons from professional pocket pickers. *Nat. Struct. Mol. Biol.* **14**, 1025–1040 https://doi.org/10.1038/nsmb1338
5 Strahl, B.D. and Allis, C.D. (2000) The language of covalent histone modifications. *Nature* **403**, 41–45 https://doi.org/10.1038/47412
6 Santos-Rosa, H., Schneider, R., Bannister, A.J., Sherriff, J., Bernstein, B.E., Emre, N.T. et al. (2002) Active genes are tri-methylated at K4 of histone H3. *Nature* **419**, 407–411 https://doi.org/10.1038/nature01080
7 Shi, X., Hong, T., Walter, K.L., Ewalt, M., Michishita, E., Hung, T. et al. (2006) ING2 PHD domain links histone H3 lysine 4 methylation to active gene repression. *Nature* **442**, 96–99 https://doi.org/10.1038/nature04835
8 Vermeulen, M., Mulder, K.W., Denissov, S., Pijnappel, W.P., van Schaik, F.M., Varier, R.A. et al. (2007) Selective anchoring of TFIID to nucleosomes by trimethylation of histone H3 lysine 4. *Cell* **131**, 58–69 https://doi.org/10.1016/j.cell.2007.08.016
9 Wysocka, J., Swigut, T., Xiao, H., Milne, T.A., Kwon, S.Y., Landry, J. et al. (2006) A PHD finger of NURF couples histone H3 lysine 4 trimethylation with chromatin remodelling. *Nature* **442**, 86–90 https://doi.org/10.1038/nature04815
10 Flanagan, J.F., Mi, L.-Z., Chruszcz, M., Cymborowski, M., Clines, K.L., Kim, Y. et al. (2005) Double chromodomains cooperate to recognize the methylated histone H3 tail. *Nature* **438**, 1181–1185 https://doi.org/10.1038/nature04290
11 Huang, Y., Fang, J., Bedford, M.T., Zhang, Y. and Xu, R.-M. (2006) Recognition of histone H3 lysine-4 methylation by the double tudor domain of JMJD2A. *Science* **312**, 748–751 https://doi.org/10.1126/science.1125162
12 Maurer-Stroh, S., Dickens, N.J., Hughes-Davies, L., Kouzarides, T., Eisenhaber, F. and Ponting, C.P. (2003) The Tudor domain 'Royal Family': tudor, plant agenet, chromo, PWWP and MBT domains. *Trends Biochem. Sci.* **28**, 69–74 https://doi.org/10.1016/S0968-0004(03)00004-5
13 Sims, R.J., Chen, C.-F., Santos-Rosa, H., Kouzarides, T., Patel, S.S. and Reinberg, D. (2005) Human but not yeast CHD1 binds directly and selectively to histone H3 methylated at lysine 4 via its tandem chromodomains. *J. Biol. Chem.* **280**, 41789–41792 https://doi.org/10.1074/jbc.C500395200
14 Musselman, C.A. and Kutateladze, T.G. (2011) Handpicking epigenetic marks with PHD fingers. *Nucleic Acids Res.* **39**, 9061–9071 https://doi.org/10.1093/nar/gkr613
15 Aasland, R., Gibson, T.J. and Stewart, A.F. (1995) The PHD finger: Implications for chromatin-mediated transcriptional regulation. *Trends Biochem. Sci.* **20**, 56–59 https://doi.org/10.1016/S0968-0004(00)88957-4
16 Lan, F., Collins, R.E., De Cegli, R., Alpatov, R., Horton, J.R., Shi, X. et al. (2007) Recognition of unmethylated histone H3 lysine 4 links BHC80 to LSD1-mediated gene repression. *Nature* **448**, 718–722 https://doi.org/10.1038/nature06034
17 Ooi, S.K., Qiu, C., Bernstein, E., Li, K., Jia, D., Yang, Z. et al. (2007) DNMT3L connects unmethylated lysine 4 of histone H3 to de novo methylation of DNA. *Nature* **448**, 714–717 https://doi.org/10.1038/nature05987
18 Li, H., Ilin, S., Wang, W., Duncan, E.M., Wysocka, J., Allis, C.D. et al. (2006) Molecular basis for site-specific read-out of histone H3K4me3 by the BPTF PHD finger of NURF. *Nature* **442**, 91–95 https://doi.org/10.1038/nature04802
19 Pena, P.V., Davrazou, F., Shi, X., Walter, K.L., Verkhusha, V.V., Gozani, O. et al. (2006) Molecular mechanism of histone H3K4me3 recognition by plant homeodomain of ING2. *Nature* **442**, 100–103 https://doi.org/10.1038/nature04814
20 Song, O.K., Wang, X., Waterborg, J.H. and Sternglanz, R. (2003) An Nα-acetyltransferase responsible for acetylation of the N-terminal residues of histones H4 and H2A. *J. Biol. Chem.* **278**, 38109–38112 https://doi.org/10.1074/jbc.C300355200
21 Giglione, C., Boularot, A. and Meinnel, T. (2004) Protein N-terminal methionine excision. *Cell. Mol. Life Sci.* **61**, 1455–1474 https://doi.org/10.1007/s00018-004-3466-8
22 Bradshaw, R.A., Brickey, W.W. and Walker, K.W. (1998) N-terminal processing: the methionine aminopeptidase and Nα-acetyl transferase families. *Trends Biochem. Sci.* **23**, 263–267 https://doi.org/10.1016/S0968-0004(98)01227-4
23 Wingfield, P.T. (2017) N-terminal methionine processing. *Curr. Protoc. Protein Sci.* **88**, 6.14.1–6.14.3 https://doi.org/10.1002/cpps.29

24 Marazzi, I., Ho, J.S.Y., Kim, J., Manicassamy, B., Dewell, S., Albrecht, R.A. et al. (2012) Suppression of the antiviral response by an influenza histone mimic. *Nature* **483**, 428–433 https://doi.org/10.1038/nature10892

25 Sampath, S.C., Marazzi, I., Yap, K.L., Sampath, S.C., Krutchinsky, A.N., Mecklenbräuker, I. et al. (2007) Methylation of a histone mimic within the histone methyltransferase G9a regulates protein complex assembly. *Mol. Cell* **27**, 596–608 https://doi.org/10.1016/j.molcel.2007.06.026

26 Zhang, Y., Ahn, J., Green, K.J., Vann, K.R., Black, J., Brooke, C.B. et al. (2019) MORC3 is a target of the influenza A viral protein NS1. *Structure* **27**, 1029–1033.e3 https://doi.org/10.1016/j.str.2019.03.015

27 Bannister, A.J., Zegerman, P., Partridge, J.F., Miska, E.A., Thomas, J.O., Allshire, R.C. et al. (2001) Selective recognition of methylated lysine 9 on histone H3 by the HP1 chromo domain. *Nature* **410**, 120–124 https://doi.org/10.1038/35065138

28 Chang, Y., Sun, L., Kokura, K., Horton, J.R., Fukuda, M., Espejo, A. et al. (2011) MPP8 mediates the interactions between DNA methyltransferase Dnmt3a and H3K9 methyltransferase GLP/G9a. *Nat. Commun.* **2**, 533 https://doi.org/10.1038/ncomms1549

29 Qin, S., Liu, Y., Tempel, W., Eram, M.S., Bian, C., Liu, K. et al. (2014) Structural basis for histone mimicry and hijacking of host proteins by influenza virus protein NS1. *Nat. Commun.* **5**, 3952 https://doi.org/10.1038/ncomms4952

30 Nichols, R.J., Wiebe, M.S. and Traktman, P. (2006) The vaccinia-related kinases phosphorylate the N′ terminus of BAF, regulating its interaction with DNA and its retention in the nucleus. *Mol. Biol Cell* **17**, 2451–2464 https://doi.org/10.1091/mbc.e05-12-1179

31 Kim, J., Daniel, J., Espejo, A., Lake, A., Krishna, M., Xia, L. et al. (2006) Tudor, MBT and chromo domains gauge the degree of lysine methylation. *EMBO Rep.* **7**, 397–403 https://doi.org/10.1038/sj.embor.7400625

32 Cheng, D., Côté, J., Shaaban, S. and Bedford, M.T. (2007) The arginine methyltransferase CARM1 regulates the coupling of transcription and mRNA processing. *Mol. Cell* **25**, 71–83 https://doi.org/10.1016/j.molcel.2006.11.019

33 Nishioka, K., Rice, J.C., Sarma, K., Erdjument-Bromage, H., Werner, J., Wang, Y. et al. (2002) PR-Set7 is a nucleosome-specific methyltransferase that modifies lysine 20 of histone H4 and is associated with silent chromatin. *Mol. Cell* **9**, 1201–1213 https://doi.org/10.1016/S1097-2765(02)00548-8

34 Hamidi, T., Singh, A.K., Veland, N., Vemulapalli, V., Chen, J., Hardikar, S. et al. (2018) Identification of Rpl29 as a major substrate of the lysine methyltransferase Set7/9. *J. Biol. Chem.* **293**, 12770–12780 https://doi.org/10.1074/jbc.RA118.002890

35 Studier, F.W. (2005) Protein production by auto-induction in high density shaking cultures. *Protein Expr. Purif.* **41**, 207–234 https://doi.org/10.1016/j.pep.2005.01.016

36 Horton, J.R., Upadhyay, A.K., Hashimoto, H., Zhang, X. and Cheng, X. (2011) Structural basis for human PHF2 Jumonji domain interaction with metal ions. *J. Mol. Biol.* **406**, 1–8 https://doi.org/10.1016/j.jmb.2010.12.013

37 Otwinowski, Z., Borek, D., Majewski, W. and Minor, W. (2003) Multiparametric scaling of diffraction intensities. *Acta Crystallogr. A* **59**, 228–234 https://doi.org/10.1107/S0108767303005488

38 McCoy, A.J., Grosse-Kunstleve, R.W., Adams, P.D., Winn, M.D., Storoni, L.C. and Read, R.J. (2007) Phaser crystallographic software. *J. Appl. Crystallogr.* **40**, 658–674 https://doi.org/10.1107/S0021889807021206

39 Afonine, P.V., Grosse-Kunstleve, R.W., Echols, N., Headd, J.J., Moriarty, N.W., Mustyakimov, M. et al. (2012) Towards automated crystallographic structure refinement with phenix.refine. *Acta Crystallogr. D Biol. Crystallogr.* **68**, 352–367 https://doi.org/10.1107/S0907444912001308

40 Emsley, P. and Cowtan, K. (2004) Coot: model-building tools for molecular graphics. *Acta Crystallogr. D Biol. Crystallogr.* **60**, 2126–2132 https://doi.org/10.1107/S0907444904019158

41 Sherman, F., Stewart, J.W. and Tsunasawa, S. (1985) Methionine or not methionine at the beginning of a protein. *Bioessays* **3**, 27–31 https://doi.org/10.1002/bies.950030108

42 Frottin, F., Martinez, A., Peynot, P., Mitra, S., Holz, R.C., Giglione, C. et al. (2006) The proteomics of N-terminal methionine cleavage. *Mol. Cell. Proteomics* **5**, 2336–2349 https://doi.org/10.1074/mcp.M600225-MCP200

43 Lauberth, S.M. and Rauchman, M. (2006) A conserved 12-amino acid motif in Sall1 recruits the nucleosome remodeling and deacetylase corepressor complex. *J. Biol. Chem.* **281**, 23922–23931 https://doi.org/10.1074/jbc.M513461200

44 Hong, W., Nakazawa, M., Chen, Y.-Y., Kori, R., Vakoc, C.R., Rakowski, C. et al. (2005) FOG-1 recruits the NuRD repressor complex to mediate transcriptional repression by GATA-1. *EMBO J.* **24**, 2367–2378 https://doi.org/10.1038/sj.emboj.7600703

45 Svensson, E.C., Huggins, G.S., Dardik, F.B., Polk, C.E. and Leiden, J.M. (2000) A functionally conserved N-terminal domain of the friend of GATA-2 (FOG-2) protein represses GATA4-dependent transcription. *J. Biol. Chem.* **275**, 20762–9 https://doi.org/10.1074/jbc.M001522200

46 Dubuissez, M., Loison, I., Paget, S., Vorng, H., Ait-Yahia, S., Rohr, O. et al. (2016) Protein kinase C-mediated phosphorylation of BCL11B at serine 2 negatively regulates its interaction with NuRD complexes during CD4+ T-cell activation. *Mol. Cell. Biol.* **36**, 1881–1898 https://doi.org/10.1128/MCB.00062-16

47 Moody, R.R., Lo, M.-C., Meagher, J.L., Lin, C.-C., Stevers, N.O., Tinsley, S.L. et al. (2018) Probing the interaction between the histone methyltransferase/deacetylase subunit RBBP4/7 and the transcription factor BCL11A in epigenetic complexes. *J. Biol. Chem.* **293**, 2125–2136 https://doi.org/10.1074/jbc.M117.811463

48 Lopez-Borges, S. and Lazo, P.A. (2000) The human vaccinia-related kinase 1 (VRK1) phosphorylates threonine-18 within the mdm-2 binding site of the p53 tumour suppressor protein. *Oncogene* **19**, 3656–3664 https://doi.org/10.1038/sj.onc.1203709

49 Sevilla, A., Santos, C.R., Vega, F.M. and Lazo, P.A. (2004) Human vaccinia-related kinase 1 (VRK1) activates the ATF2 transcriptional activity by novel phosphorylation on Thr-73 and Ser-62 and cooperates with JNK. *J. Biol. Chem.* **279**, 27458–27465 https://doi.org/10.1074/jbc.M401009200

50 Kang, T.-H., Park, D.-Y., Choi, Y.H., Kim, K.-J., Yoon, H.S. and Kim, K.-T. (2007) Mitotic histone H3 phosphorylation by vaccinia-related kinase 1 in mammalian cells. *Mol. Cell. Biol.* **27**, 8533–8546 https://doi.org/10.1128/MCB.00018-07

51 Wu, Z., Cheng, Z., Sun, M., Wan, X., Liu, P., He, T. et al. (2015) A chemical proteomics approach for global analysis of lysine monomethylome profiling. *Mol. Cell. Proteomics* **14**, 329–339 https://doi.org/10.1074/mcp.M114.044255

52 Povlsen, L.K., Beli, P., Wagner, S.A., Poulsen, S.L., Sylvestersen, K.B., Poulsen, J.W. et al. (2012) Systems-wide analysis of ubiquitylation dynamics reveals a key role for PAF15 ubiquitylation in DNA-damage bypass. *Nat. Cell Biol.* **14**, 1089–1098 https://doi.org/10.1038/ncb2579

53 Kais, Z., Barsky, S.H., Mathsyaraja, H., Zha, A., Ransburgh, D.J., He, G. et al. (2011) KIAA0101 interacts with BRCA1 and regulates centrosome number. *Mol. Cancer Res.* **9**, 1091–1099 https://doi.org/10.1158/1541-7786.MCR-10-0503

54 Kurokawa, M., Mitani, K., Yamagata, T., Takahashi, T., Izutsu, K., Ogawa, S. et al. (2000) The evi-1 oncoprotein inhibits c-Jun N-terminal kinase and prevents stress-induced cell death. *EMBO J.* **19**, 2958–2968 https://doi.org/10.1093/emboj/19.12.2958

55 Fleming, Y., Armstrong, C.G., Morrice, N., Paterson, A., Goedert, M. and Cohen, P. (2000) Synergistic activation of stress-activated protein kinase 1/ c-Jun N-terminal kinase (SAPK1/JNK) isoforms by mitogen-activated protein kinase kinase 4 (MKK4) and MKK7. *Biochem. J.* **352**, 145–154 https://doi. org/10.1042/bj3520145

56 Gupta, S., Barrett, T., Whitmarsh, A.J., Cavanagh, J., Sluss, H.K., Derijard, B. et al. (1996) Selective interaction of JNK protein kinase isoforms with transcription factors. *EMBO J.* **15**, 2760–2770 https://doi.org/10.1002/j.1460-2075.1996.tb00636.x

57 Dai, R., Frejtag, W., He, B. and Zhang, Y. (2000) Mivechi NF. c-Jun NH2-terminal kinase targeting and phosphorylation of heat shock factor-1 suppress its transcriptional activity. *J. Biol. Chem.* **275**, 18210–8 https://doi.org/10.1074/jbc.M000958200

58 Miotto, B. and Struhl, K. (2011) JNK1 phosphorylation of Cdt1 inhibits recruitment of HBO1 histone acetylase and blocks replication licensing in response to stress. *Mol. Cell* **44**, 62–71 https://doi.org/10.1016/j.molcel.2011.06.021

59 Tomlinson, V., Gudmundsdottir, K., Luong, P., Leung, K., Knebel, A. and Basu, S. (2010) JNK phosphorylates Yes-associated protein (YAP) to regulate apoptosis. *Cell Death Dis.* **1**, e29 https://doi.org/10.1038/cddis.2010.7

60 Maekawa, T., Sakura, H., Sudo, T. and Ishii, S. (1989) Putative metal finger structure of the human immunodeficiency virus type 1 enhancer binding protein HIV-EP1. *J. Biol. Chem.* **264**, 14591–3 https://doi.org/10.1016/S0021-9258(18)63733-0

61 Chen, J., Sagum, C. and Bedford, M.T. (2020) Protein domain microarrays as a platform to decipher signaling pathways and the histone code. *Methods* **184**, 4–12 https://doi.org/10.1016/j.ymeth.2019.08.007

62 Jain, K., Fraser, C.S., Marunde, M.R., Parker, M.M., Sagum, C., Burg, J.M. et al. (2020) Characterization of the plant homeodomain (PHD) reader family for their histone tail interactions. *Epigenetics Chromatin* **13**, 3 https://doi.org/10.1186/s13072-020-0328-z

63 Yang, N., Wang, W., Wang, Y., Wang, M., Zhao, Q., Rao, Z. et al. (2012) Distinct mode of methylated lysine-4 of histone H3 recognition by tandem tudor-like domains of Spindlin1. *Proc. Natl Acad. Sci. U.S.A.* **109**, 17954–9 https://doi.org/10.1073/pnas.1208517109

64 Ali, M., Rincon-Arano, H., Zhao, W., Rothbart, S.B., Tong, Q., Parkhurst, S.M. et al. (2013) Molecular basis for chromatin binding and regulation of MLL5. *Proc. Natl Acad. Sci. U.S.A.* **110**, 11296–11301 https://doi.org/10.1073/pnas.1310156110

65 Gatchalian, J., Futterer, A., Rothbart, S.B., Tong, Q., Rincon-Arano, H., Sanchez de Diego, A. et al. (2013) Dido3 PHD modulates cell differentiation and division. *Cell Rep.* **4**, 148–158 https://doi.org/10.1016/j.celrep.2013.06.014

66 Pena, P.V., Hom, R.A., Hung, T., Lin, H., Kuo, A.J., Wong, R.P. et al. (2008) Histone H3K4me3 binding is required for the DNA repair and apoptotic activities of ING1 tumor suppressor. *J. Mol. Biol.* **380**, 303–312 https://doi.org/10.1016/j.jmb.2008.04.061

67 Wen, H., Li, J., Song, T., Lu, M., Kan, P.-Y., Lee, M.G. et al. (2010) Recognition of histone H3K4 trimethylation by the plant homeodomain of PHF2 modulates histone demethylation. *J. Biol. Chem.* **285**, 9322–9326 https://doi.org/10.1074/jbc.C109.097667

68 Zegerman, P., Canas, B., Pappin, D. and Kouzarides, T. (2002) Histone H3 lysine 4 methylation disrupts binding of nucleosome remodeling and deacetylase (NuRD) repressor complex. *J. Biol. Chem.* **277**, 11621–4 https://doi.org/10.1074/jbc.C200045200

69 Wu, M., Wang, L., Li, Q., Li, J., Qin, J. and Wong, J. (2013) The MTA family proteins as novel histone H3 binding proteins. *Cell Biosci.* **3**, 1 https://doi. org/10.1186/2045-3701-3-1

70 Musselman, C.A., Ramirez, J., Sims, J.K., Mansfield, R.E., Oliver, S.S., Denu, J.M. et al. (2012) Bivalent recognition of nucleosomes by the tandem PHD fingers of the CHD4 ATPase is required for CHD4-mediated repression. *Proc. Natl Acad. Sci. U.S.A.* **109**, 787–792 https://doi.org/10.1073/pnas.1113655109

71 Horowitz, S. and Trievel, R.C. (2012) Carbon-oxygen hydrogen bonding in biological structure and function. *J. Biol. Chem.* **287**, 41576–41582 https://doi.org/10.1074/jbc.R112.418574

72 Wang, H., Cao, R., Xia, L., Erdjument-Bromage, H., Borchers, C., Tempst, P. et al. (2001) Purification and functional characterization of a histone H3-lysine 4-specific methyltransferase. *Mol. Cell* **8**, 1207–1217 https://doi.org/10.1016/S1097-2765(01)00405-1

73 Dou, Y., Milne, T.A., Ruthenburg, A.J., Lee, S., Lee, J.W., Verdine, G.L. et al. (2006) Regulation of MLL1 H3K4 methyltransferase activity by its core components. *Nat. Struct. Mol. Biol.* **13**, 713–719 https://doi.org/10.1038/nsmb1128

74 Song, J.-J. and Kingston, R.E. (2008) WDR5 interacts with mixed lineage leukemia (MLL) protein via the histone H3-binding pocket. *J. Biol. Chem.* **283**, 35258–35264 https://doi.org/10.1074/jbc.M806900200

75 Wu, H., Mathioudakis, N., Diagouraga, B., Dong, A., Dombrovski, L., Baudat, F. et al. (2013) Molecular basis for the regulation of the H3K4 methyltransferase activity of PRDM9. *Cell Rep.* **5**, 13–20 https://doi.org/10.1016/j.celrep.2013.08.035

76 Sidoli, S. and Garcia, B.A. (2017) Middle-down proteomics: a still unexploited resource for chromatin biology. *Expert Rev. Proteomics* **14**, 617–626 https://doi.org/10.1080/14789450.2017.1345632

77 Bah, A. and Forman-Kay, J.D. (2016) Modulation of intrinsically disordered protein function by post-translational modifications. *J. Biol. Chem.* **291**, 6696–6705 https://doi.org/10.1074/jbc.R115.695056

78 Cismasiu, V.B., Adamo, K., Gecewicz, J., Duque, J., Lin, Q. and Avram, D. (2005) BCL11B functionally associates with the NuRD complex in T lymphocytes to repress targeted promoter. *Oncogene* **24**, 6753–6764 https://doi.org/10.1038/sj.onc.1208904

79 Xu, J., Bauer, D.E., Kerenyi, M.A., Vo, T.D., Hou, S., Hsu, Y.-J. et al. (2013) Corepressor-dependent silencing of fetal hemoglobin expression by BCL11A. *Proc. Natl Acad. Sci. U.SA.* **110**, 6518–6523 https://doi.org/10.1073/pnas.1303976110

80 Allen, H.F., Wade, P.A. and Kutateladze, T.G. (2013) The NuRD architecture. *Cell. Mol. Life Sci.* **70**, 3513–3524 https://doi.org/10.1007/s00018-012-1256-2

81 Angelelli, C., Magli, A., Ferrari, D., Ganassi, M., Matafora, V., Parise, F. et al. (2008) Differentiation-dependent lysine 4 acetylation enhances MEF2C binding to DNA in skeletal muscle cells. *Nucleic Acids Res.* **36**, 915–928 https://doi.org/10.1093/nar/gkm1114

82 Schwer, B., Eckersdorff, M., Li, Y., Silva, J.C., Fermin, D., Kurtev, M.V. et al. (2009) Calorie restriction alters mitochondrial protein acetylation. *Aging Cell* **8**, 604–606 https://doi.org/10.1111/j.1474-9726.2009.00503.x