HHS Public Access

Author manuscript

Nat Chem Biol. Author manuscript; available in PMC 2018 March 01.

Published in final edited form as:

Nat Chem Biol. 2017 July ; 13(7): 750–756. doi:10.1038/nchembio.2377.

Developing Spindlin1 Small Molecule Inhibitors Using Protein Microarrays

Narkhyun Bae[1,§], **Monica Viviano**[2,§], **Xiaonan Su**[3,4,§], **Jie Lv**[5], **Donghang Cheng**[1], **Cari Sagum**[1], **Sabrina Castellano**[2,6], **Xue Bai**[3,4], **Claire Johnson**[1], **Mahmoud Ibrahim Khalil**[1,7], **Jianjun Shen**[1], **Kaifu Chen**[5], **Haitao Li**[3,4,*], **Gianluca Sbardella**[2,*], and **Mark T. Bedford**[1,*]

[1]Department of Epigenetics and Molecular Carcinogenesis, The University of Texas MD Anderson Cancer Center, Smithville, TX 78957, USA

[2]Dipartimento di Farmacia, Epigenetic Med Chem Lab, Università degli Studi di Salerno, Via Giovanni Paolo II 132, I-84084 Fisciano (SA), Italy

[3]Beijing Advanced Innovation Center for Structural Biology, MOE Key Laboratory of Protein Sciences, Department of Basic Medical Sciences, School of Medicine, Tsinghua University, Beijing 100084, China

[4]Tsinghua-Peking Joint Center for Life Sciences, Beijing, 100084, P.R. China

[5]Center for Regenerative Medicine, Department of Cardiovascular Sciences, Houston Methodist Research Institute, Houston, TX 77030, USA & Department of Cardiothoracic Surgery, Weill Cornell Medical College, Cornell University

[6]Dipartimento di Medicina e Chirurgia, Università degli Studi di Salerno, Via Salvador Allende, I-84081 Baronissi (SA), Italy

[7]Molecular Biology Unit, Department of Zoology, Faculty of Science, Alexandria University, Egypt

Abstract

The discovery of inhibitors of methyl- and acetyl-binding domains has provided evidence for the “druggability” of epigenetic effector molecules. The small molecule probe UNC1215 prevents methyl-dependent protein-protein interactions by engaging the aromatic cage of MBT domains, and with lower affinity, Tudor domains. Using a library of tagged UNC1215 analogs we screened

*Correspondence should be addressed to - M.T.B. (mtbedford@mdanderson.org), G.S. (gsbardella@unisa.it), H.L. (lht@tsinghua.edu.cn).

§These authors contributed equally to this work.

ACCESSION CODES

The atomic coordinates and structure factors of Spindlin1 bound to EML405 and EML631 have been deposited in the Protein Data Bank under accession codes of **5JSG** and **5JSJ**, respectively. The GEO accession number for our RNA-Seq data is **GSE92547**.

CONTRIBUTIONS

M.T.B. and G.S. conceived the project. N.B. and D.C. carried out the cell-based and competition experiments. C.S., C.J. and M.I.K. constructed, maintained and probed the protein domain microarrays. M.V., S.C. and G.S. designed and synthesized the compounds used in this study. X.S, X.B. and H.L. performed the structural studies and ITC experiments. J.S. performed the deep-sequencing. K.C. and J.L. performed the bioinformatics analysis of the RNA-Seq data sets. M.T.B., G.S. and H.L. wrote the manuscript and supervised the work in their respective fields.

COMPETING FINANCIAL INTERESTS

M.T.B. is a cofounder of EpiCypher.

Other authors have no competing financial interests.

a protein domain microarray of human methyl-lysine effector molecules to rapidly detect compounds with novel binding profiles - either improved or loosened specificity. Using this approach, we identified a compound (EML405) that acquired a novel interaction with the Tudor domain-containing protein Spindlin1 (SPIN1). Structural studies facilitated the rational synthesis of more selective SPIN1 inhibitors (EML631–633), which engage SPIN1 in cells, block its ability to "read" H3K4me3 marks, and inhibit its transcriptional coactivator activity. Protein microarrays can thus be used as a platform to "target hop" and identify small molecules that bind and compete with domain–motif interactions.

Introduction

Post-translational modifications (PTMs) are dynamic processes that frequently occur on both the globular domain of histone proteins and on their protruding tails regions[1], and give rise to a complex pattern referred to as the "histone code"[2]. These chemical marks (the most prominent being acetyl, methyl, ubiquitinyl and phosphate groups) are added or removed by various enzyme families, and different combinations of PTMs are recognized by specific binding modules or domains[3–5]. This recognition of different PTMs on the histone tail, by effector molecules and their corresponding protein complexes, is critical for both the activation and repression of gene expression. Importantly, histone tail modifications and their effector molecules are often misregulated in diseases including cancer[6–8]. As a consequence, epigenetic regulators have become major targets for drug development[9]. Lysine and arginine methylation play central roles in this "histone code" theory, and these two residues can accept more than one methyl group, with different protein methyltransferases able to execute varying degrees of methylation[10]. Methyl reader domains are clustered into eight major families, including plant homeodomains (PHDs), WD-40 domains, chromatin organization modifier domains (chromodomains), Tudor domains, Agenet domains, proline-tryptophan-tryptophan-proline (PWWP) domains, Bromo adjacent homology (BAH) domains, and malignant brain tumor (MBT) domains[11,12].

The recent discoveries of compounds that prevent the binding of acetyl-lysine motifs with bromodomains have clearly demonstrated the feasibility of targeting histone code reader domains[13,14]. Domains that bind methylated motifs are also likely good targets for competitive small molecule ligands. Indeed, all eight domain-types that bind methyllysine marks do so through an "aromatic cage"[15]. Thus, if lead compounds can be found that dock into these cages, then chemical space can likely be explored to identify specific inhibitors for the different domain types[16]. Moreover, many of these domain types are predicted to be very druggable[17]. Thus, there has been a focused attempt by a number of groups to identify compounds that can inhibit methyl-dependent protein-protein interactions, including small molecules that competitively inhibit PHD finger binding[18,19], the development of potent H3K27me3 peptide mimetics which selectively inhibit protein interactions that are Chromo domain mediated[20,21], and the employment of virtual screening strategy to identify small-molecule ligands for MBT domains[22] and Tudor domains[23]. The MBT domain ligands are a series of nicotinamides, which do not bind PHD or Chromo domains[24]. Optimization studies resulted in dibasic ligands with improved affinity, comprising their lead compound UNC1215, endowed with a high binding affinity for L3MBTL3[25]. Yet, when this compound

is tagged with biotin and used to challenge a protein domain microarray, it binds not only MBT domains, but also Tudor domains in a K_d range of 30 μM–100 nM[25].

Starting from UNC1215, we herein describe the development of a library of biotin-tagged analogues that we used to screen an array of protein domains. This “library-on-library” screening approach not only distinguished compounds that selectively bound the Tudor domains of PHF20, but also identified a compound that gained binding affinity towards additional “aromatic cage”-containing domains, including the Tudor-domains of Spindlin1 (SPIN1). Structural optimization studies led to the identification of selective inhibitors of SPIN1 that are active in cell-based assays.

Results

Microarray probing with tagged small molecules

To screen for novel protein-chemical interactions, we generated a protein domain microarray that harbored 98 GST fusion proteins, including 41 Tudor domains and 31 Chromo domains. Representative PHD, BHA, MBT, PWWP, ANK, AGENET and HEAT domains were also added to the array (Supplemental Results, Supplementary Fig. 1). Thus, the majority of the protein domains harbored aromatic cages with potential methyl “reading” ability. We used a robotic pin arrayer to spot the GST fusions, in duplicate, onto nitrocellulose-coated glass slides and used fluorescently tagged streptavidin, which was pre-conjugated to the biotinylated small molecules, to visualize binding interactions.

Many methyl reader proteins are characterized by the presence of repeated effector domains (e.g. PHD, MBT, Tudor, Chromo)[17,26,27], thus they are potentially polyvalent binders. As a consequence, ligands incorporating two Kme and/or Rme mimics could provide higher affinity. UNC1215 was originally designed based on this hypothesis, and structural studies revealed that it engages L3MBTL3 in a unique 2:2 binding mode[25]. It contains a central moiety symmetrically decorated with two nitrogen-containing groups that are a mimetic of substituted lysine residues (Supplementary Fig. 2a). Therefore, with the aim to discover molecular probes for various aromatic cage-containing effector domains, we designed a series of ligands (Supplementary Fig. 2b) featuring a central core decorated, both symmetrically and asymmetrically, with basic nitrogen-containing “privileged structures”[25]. To facilitate the screening of the protein domain microarray, we tagged all of the compounds (47) with a biotinylated PEG linker to the central aromatic core (Supplementary Table 1). These tagged compounds displayed varied binding profiles on the array (Supplementary Fig. 3). Among the newly synthesized compounds (Supplementary Note), derivatives EML405 and EML 417 (Fig. 1a) generated novel interaction patterns. EML405 bound strongly to the Tudor domains of PHF20 and its homolog PHF20L1, the Tudor domains of 53BP1, a subset of MBT domains, a yeast Tudor domain ($SPF30^{sp}$) and to Spindlin1, which harbors three Tudor domains. Thus, EML405 displayed broader binding specificity than UNC1215. On the other hand, EML417 selectively interacted with the Tudor domain of PHF20 (Fig. 1b).

SPIN1/EML co-crystallographic studies

To independently confirm the interaction of EML405 with SPIN1, we used isothermal titration calorimetry (ITC) and an *in vitro* thermal shift assay (TSA) (Fig. 2a). We measured a dissociation constant (K_d) of 15 μM between SPIN1 and EML405 compound (Fig. 2a, left). Consistent with the ITC assay, EML405 displayed a stabilization effect on recombinant SPIN1. The melting temperature of SPIN1 was increased by 5°C in the presence of EML405 (Fig. 2a, right). To understand the underlying molecular mechanism, we solved the cocrystal structure of SPIN1 bound to EML405 at 2.5 Å resolution (Supplementary Table 2). In the complex structure, EML405 was positioned within the surface grooves of SPIN1 and bridged the first and the second Tudor domains. The two arms ("left arm" and "right arm") were inserted into the aromatic cages of Tudor 1 and Tudor 2, which recognize histone H3R8me2a and H3K4me3[28], respectively (Fig. 2b). Buried surface area analysis revealed that the "left-arm" of EML405 covered 306 $Å^2$ within the first Tudor. However, the "right-arm" of EML405 had far more wiggle room with a 364 $Å^2$ surface burial (Fig. 2b,c). Indeed, the benzene group of EML405 was located in close proximity (~ 10 Å) to the acidic loop region (E142, E144, D145, T143, and D149) of the second Tudor domain, which could be exploited to gain affinity and selectivity (Fig. 2c). In particular, we hypothesized that introducing an additional nitrogen containing moiety, linked through a spacer of the proper length, to the lead structure could establish extra binding interactions with negatively charged E142, E144, and D145 residues, just above the "right-arm" of EML405 (Fig. 2c). We reasoned that an additional protonable substituent in this region of the molecule could occupy this pocket and result in a gain of the binding affinity. Therefore, we introduced a pyrrolidine group at the 3′ and 4′ positions of the aniline phenyl ring, through a flexible ethylene or oxymethylene linker to avoid potential steric hindrance, and synthesized compounds EML631, EML632, and EML633 (Fig. 2d).

To determine whether the EML405 analogs bound SPIN1 with higher affinity, we performed ITC experiments. We measured a binding K_d of ~ 3 μM for EML631, ~ 7 μM for EML632, and ~ 2 μM for EML633, which was a 2- to 7- fold enhancement compared to EML405 (Fig. 3a). We next attempted to obtain the cocrystal structures of SPIN1 with each of EML405 analogs, but were only successful with SPIN1-EML631 (Supplementary Table 3). Unexpectedly, EML631 did not form interactions with the acidic loop region of the second Tudor domain of SPIN1. Instead, the added pyrrolidine group folded into a negatively charged groove between the first and second Tudor domains, forming two hydrogen bonds with D173 and H252 (Fig. 3b). Importantly, the addition of this bulky pyrrolidine group, not only enhanced interaction with SPIN1, but also dramatically improved the specificity for SPIN1. ITC binding studies showed that derivatives EML631–633 bound to the Tudor domains of 53BP1 and PHF20, and the MBT domains of L3MBTL1 much less well (Supplementary Table 4). However, EML632 and EML633 bound to L3MBTL3 nearly as well as SPIN1. Most importantly, EML631 was a good SPIN1 binder (K_d = 3 μM) while binding all proteins in the counter screen much more weakly (<100 μM) (Supplementary Table 4 and Supplementary Fig. 4a).

EML compounds block the SPIN1-H3K4me3 interaction

SPIN1 is a transcriptional coactivator that binds the H3K4me3 mark through one of its three Tudor domains[28,29]. Structural studies indicate that the EML compounds engage the aromatic cage of the first and second Tudor domains of SPIN1 and likely block the ability of the H3K4me3 to bind the second Tudor domain (Fig. 2b, 3b). To test this, we performed a competition pulldown assay using bead-bound H3K4me3 peptide and GST-SPIN1$^{(50\text{-}262)}$, in the presence of increasing concentrations of untagged small molecule inhibitors (Fig. 4a). As expected, in these competition pulldown assays, UNC1215 was not active, and higher concentrations of EML405 than EML631–633 were needed to inhibit the interaction. It has been shown that SPIN1 also interacts directly with the transcription factor TCF4[39]. This interaction involves the region of SPIN1 that harbors the first two Tudor domains[28], and both EML405 and EML631 blocked this interaction as well (Supplementary Fig. 4b). We further demonstrated that tagged EML compounds not only interacted with SPIN1 fused to the GST protein, but also with a GFP-fusion that was expressed in mammalian cells. GFP-SPIN1$^{(FL)}$ was transfected into HEK293 cells and whole cell lysates were passed over beads with immobilized compounds (Supplementary Fig. 5). Upon chemiprecipitation, we detected the pulldown of GFP-SPIN1 by EML405, EML631–633, but not by UNC1215 (Fig. 4b). Tagged UNC1215 was functional, because it was able to interact with GFP-53BP1, which also weakly interacted with the EML series. It should be noted that the tagged form of EML632 interacted much more weakly with GFP-SPIN1$^{(FL)}$ than did EML631 and EML633, even though the untagged version of these three compounds behaved similarly in both ITC assays (Fig. 3a) and methyl-peptide competition assays (Fig. 4a). To reaffirm that EML632 can interact with SPIN1, we performed a pulldown assay with tagged and bead-immobilized EML405, and efficiently competed it away from GST-SPIN1 with increasing concentrations of untagged EML632 (Supplementary Fig. 6a). It is thus likely that the addition of the tag onto EML632 interferes with its ability to interact with SPIN1. There are five members of the SPIN protein family (SPIN1, 2a, 2b, 3 and 4)[29], and upon chemiprecipitation with EML631 we detected the efficient pulldown of GFP-SPIN1, reduced pulldown of SPIN2a/b and no interaction with either SPIN3 or SPIN4 (Supplementary Fig. 6b). The EML631 probe is thus relatively specific for SPIN1.

EML compounds engage SPIN1 in cells

To investigate whether the EML compounds are cell permeable, we used the recently developed Cellular Thermal Shift Assay (CETSA)[30,31]. In this assay, cultured cells are incubated in the presence of a small molecule and then heated, which denatures and precipitates proteins. Proteins that interact with the small molecule in the cell are more stable and precipitate at higher temperatures than the unbound proteins. The amount of protein in the soluble fraction, at different temperatures, can then be gauged by Western analysis. When the CETSA assay was performed on GFP-SPIN1 transfected cells, we detected a stabilization of this fusion protein at high temperatures in the presence of EML405 and to a great degree in the presence of EML631–633, but not with UNC1215, which does not interact with SPIN1 *in vitro* (Fig. 4c).

EML compounds inhibit SPIN1 coactivator activity

As a transcriptional coactivator, SPIN1 has been reported to regulate the expression of rRNA genes[29]. In an effort to identify additional genes that are regulated by SPIN1, we performed RNA-seq experiments on T778 liposarcoma cells that were either knocked down for SPIN1 or treated with EML631. We performed RNA-seq analysis on biological triplicates only after verifying the efficiency of the SPIN1 knockdown. We identified 655 transcripts that were down regulated in both the SPIN1 knockdown and EML631-treated cells (Fig. 5a). An example of the graph tracks of the RNA sequencing data shows reduced SPIN1 levels in the knockdown sample, and reduced IL1B and BST2 levels in both knockdown and EML631-treated samples (Supplementary Fig. 7a). We selected five candidate loci that displayed at least a 0.5-fold expression change for further analysis – IL1B, BST2, C1QTNF1, ALDH1A3 and IFI44L, as well as the rDNA locus. We validated that SPIN1 regulates these six transcripts by knocking down SPIN1 in T778 liposarcoma cells and performing quantitative RT-PCR analysis (Fig. 5b). Next, we treated T778 cells with EML405 and EML631 and observed a decrease in the expression levels of all six candidate genes after compound treatment, without a dramatic effect on endogenous SPIN1 levels (Fig. 5c and Supplementary Fig. 7b). Further, using a ChIP-qPCR approach, we found that GFP-SPIN1 was recruited to the promoter regions of rDNA and IL1B, and that this recruitment was blocked by treating cells will EML631 (Fig. 5d). We also found, as previously reported[29], that ectopic overexpression of SPIN1 elevated the expression of rRNA, as well as the expression of IL1B, BST2, C1QTNF1, ALDH1A3 and IFI44L (Fig. 5e), and this elevated expression could be dramatically attenuated by treating T778 cells with EML405 and EML631 shortly after GFP-SPIN1 transfection (Fig. 5e and Supplementary Fig. 7c). Finally, the inhibitory effect of EML631 on the transcriptional response of all six SPIN1-regulated genes was time- and concentration-dependent (Supplementary Fig. 8a & b). Thus, EML405 and EML631 are cell permeable and they engage SPIN1 and block its recruitment to chromatin, thereby inactivating it.

Discussion

We initially developed protein domain microarrays to identify novel signal transduction pathways that involve proline-rich interacting domains and phospho-binding domains[32,33]. We then expanded this approach to investigate methyl-reading domains, which could interpret the histone code[34]. More recently, we have used this domain array approach to establish the selectivity of small molecules that bind protein domains[21,25]. In these latter screens, we developed lead compounds that inhibited MBT domains (UNC1215) and chromodomain (UNC3866) methyl-dependent interactions. These compounds were biotinylated and used to probe protein domain microarrays to establish the specificity of the developed compound. In the present study, we converted the protein domain microarray into a screening platform to “target hop” with tagged small molecules. After screening a small library of about fifty UNC1215 analogs, we identified compound EML417 that selectively bound PHF20, and EML405 that acquired an expanded binding spectrum, including the capacity to interact with SPIN1 (Fig. 1), thus validating this screening approach. However, it should be noted that some interactions may be missed because we cannot be sure that all the arrayed recombinant proteins are correctly folded into functional domains. In the future, as

we expand our biotin-tagged small molecule library, we may be able to reverse the screening approach by immobilizing the compounds on streptavidin-coated slides that could be probed with recombinant protein domains, similar to small-molecule arrays that have been previously developed[35,36]. Moreover, since many of the arrayed domains have identified methyl-peptide ligands that correspond to different marks on the histone tail, it may be possible to adapt this protein domain microarray screening approach for use with small molecule libraries that are not biotinylated by developing a tagged-peptide/untagged-compound competition approach.

We focused our attention on optimizing specific small molecule antagonists of the SPIN1/H3K4me3 interaction because of the role this transcriptional coactivator plays in cell transformation and cancer. SPIN1 was originally identified in a screen for genes involved in ovarian cancer[37]. Overexpression of SPIN1 induces transformation of NIH3T3 cells that acquire the ability to grow in soft agar and in nude mice[38]. SPIN1 protein levels are elevated in a number of different cancers[39,40] and the protein displays a diffuse nuclear localization and is enriched in nucleoli[29]. Mechanistically, SPIN1 functions as a “reader” of the histone H3K4me3 mark and coactivates transcription of rRNA genes, genes regulated by the MAZ transcription factor and Wnt target genes[29,40]. The interaction of SPIN1 with H3K4me3 is stabilized when a second Tudor domain engages the H3R8me2a mark[28]. SPIN1 null mice exhibit early post-natal lethality and display a defect in meiosis[41]. Finally, SPIN1 promotes RET signaling in liposarcoma[40], where it was found that: 1) knockdown of SPIN1 in a liposarcoma cell line (T778) reduces cell proliferation, 2) proliferation rates are rescued with wild-type SPIN1, but not with a mutant in the Tudor domain that binds H3K4me3 (the interaction we are targeting), and 3) SPIN1 knockdown dramatically reduces tumor weight in a xenograft mouse model. This last finding could be adapted in the future, as an *in vivo* screening approach for the efficacy of the EML compounds.

Because of the emerging role of SPIN1 in transformation and cancer, there is a growing interest in developing inhibitors of the protein-protein interactions that it mediates. Recently, a suite of assays were developed to facilitate the high-throughput screening of small molecule libraries for SPIN1 inhibitors[42]. The pilot screen of a small library of bioactive compounds, using an AlphaLISA assay, identified a lead compound, A366 (which was previously reported as selective inhibitor of lysine methyltransferase G9a[43]), as endowed with good *in vitro* inhibition of the Spindlin1-H3K4me3 interaction (IC_{50} about 200 nM), but poor cell-based activity (100 μM). Moreover, similarly to what was previously reported against G9a[43], even minor structural modifications of A366 led to significant reduction of SPIN1 binding affinity[42]. Therefore, there is still a need to develop novel scaffolds capable of inhibiting the recognition of H3K4me3 by SPIN1 in cells, which could then be further optimized. More recently, the same group used a virtual screening approach to identify a few new potential lead scaffolds for the development of SPIN1 inhibitors, but these compounds displayed only moderate activity *in vitro*. Interestingly, they found that an appropriate lysine-mimetic group is crucial for inhibitory activity[44].

We identified EML405 as a lead molecule that bound SPIN1 (Fig. 1). Structural analysis revealed that the symmetric character of EML405 allows it to engage both the first and second Tudor domains of SPIN1 (Fig. 2b), and also revealed a large unoccupied pocket that

we exploited in a subsequent round of compound synthesis to generate EML631–633. The introduction of the pyrrolidine-containing additional arm in these three analogs (Fig. 2d) was designed to engage this negatively charged pocket, and binding studies revealed an improved binding constant of the three derivatives over EML405 (3 μM for EML631 *vs* 14 μM for EML405) (Fig. 3a). However, SPIN1/EML631 structural studies revealed an unpredicted interaction of the pyrrolidine group with a negatively charged groove that lies between the first two Tudor domains (Fig. 3b). Importantly, this addition also dramatically improved the specificity of EML631 SPIN1, which now displays weaker than 100 μM binding for the Tudor domain-containing proteins 53BP1 and PHF20, and for the MBT domain-containing proteins L3MBTL1 and L3MBTL3 (Supplementary Table. 4 & Supplementary Fig. 4). Furthermore, EML631 was active in competition assays, chemiprecipitation experiments, and CETSA experiments (Fig. 4), as well as in cell-based coactivator assays (Fig. 5). When tagged, EML632 did not function well in chemiprecipitation experiments, and both EML632 and EML633 cross-reacted with L3MBTL3. Thus, we consider EML631 the most specific and versatile SPIN1 inhibitor identified in this study.

Here we demonstrated that protein microarrays can be used as a platform to “target hop” and identify small molecules that bind and compete with domain–motif interactions. This approach could easily be adapted to the screening and profiling of other domain inhibitors, like those that target bromodomains.

ONLINE METHODS

Antibodies, Plasmids and Peptides

Rabbit polyclonal SPIN1 antibody was purchased from Abcam (ab118784, 1:2000) and Mouse polyclonal GFP antibody was obtained from Santa Cruz Biotech (sc-9996, 1:2000). Rabbit GST-specific antibody was generated in the Bedford laboratory (1:3,000). Secondary mouse HRP-IgG and rabbit HRP-IgG were purchased from GE Healthcare Life Sciences. Complementary DNA encoding full length SPIN1, SPIN2A, SPIN2B, SPIN3, SPIN4 and 53BP1 fragment (aa 1144~1709) were cloned into pEGFP-C1. Biotinylated H3K4me3 peptide (MW 2.416kDa) was synthesized by the W.M. Keck center (New Haven, CT).

Cell culture and RNAi interference

HEK 293T, HeLa, and T778 cells were purchased from ATCC. All cell lines used in this study were tested for mycoplasma by using MycoAlert™ mycoplasma detection Kit (Lonza) and found to be uncontaminated. HEK 293T and HeLa cells were cultured in DMEM supplemented with 10% fetal bovine serum (FBS), penicillin/streptomycin, nonessential amino acids and glutamine. T778 cells were cultured in RPMI medium supplemented 10% FBS, penicillin/streptomycin and glutamine. All cells were maintained in a humidified 37° C incubator with 5% CO_2. T778 cells were transfected with 15 nM SPIN1 siRNA (Silencer® Selcet Pre-designed siRNA, Ambion/Life technologies, Cat#: 4392420) or control siRNA (Silencer® Select negative control #1 siRNA, Ambion/Life technologies, Cat#: 4390843) using Lipofectamine RNAiMAX for 24 h as per manufacturer’s instructions.

Quantitative Real-time PCR (RT-qPCR)

Total RNA was purified from cells using TRIzol reagent (Invitrogen) and reverse transcribed using Superscript III First Strand Synthesis kit (Invitrogen). Quantitative RT-PCR was performed with the Applied Biosystems 7900HT RT-PCR instrument using the iTaq Universal SYBR Green Supermix (Bio-Rad) with primers for the indicated genes (Listed below). Gene expression was calculated following normalization to GAPDH levels using the comparative Ct (Cycle threshold) method and is shown as fold relative to the expression of each gene in the control cells. T778 cells were grown to 30% confluency, at which point cells were treated or untreated with EML631 (10 μM) and EML405 (20 μM) for 4 days respectively. T778 cells were transfected with SPIN1-GFP for 6 h, and then immediately incubated with small molecules at 10 μM for 24 h after washing cells twice with PBS. The primer sequences used for qPCR analyses are listed in below.

GAPDH Forward: 5’-GAGTCCACTGGCGTCTTCAC-3’

GAPDH Reverse: 5’-CATGCCCACCATCACGCCCTGG-3’

β-Actin Forward: 5’-GCGAGCACAGAGCCTCGCCTT-3’

β-Actin Reverse: 5’-CATGCCCACCATCACGCCCTGG-3’

SPIN1 Forward: 5’-CTGTGGACTGCAGCCTCGGCG-3’

SPIN1 Reverse: 5’-TGGCCTGCATCAGCTCTGGACC-3’

rRNA Forward: 5’-TGTCAGGCGTTCTCGTCTC-3’

rRNA Reverse: 5’- AGCACGACGTCACCACATC-3’

IL1B Forward: 5’-ATGCACCTGTACGATCACTG-3’

IL1B Reverse: 5’-ACAAAGGACATGGAGAACACC-3’

BST2 Forward: 5’-TGATGGAGTGTCGCAATGTC-3’

BST2 Reverse: 5’-GTCCTTGGGCCTTCTCTG-3’

C1QTNF1 Forward: 5’-GCCCCAGATCAACATCACTAT-3’

C1QTNF1 Reverse: 5’-GCCTGTTTTGCCATATTTCCC-3’

ALDH1A3 Forward: 5’-CTTCTGCCTTAGAGTCTGGAAC-3’

ALDH1A3 Reverse: 5’-CGTATTCACCTAGTTCTCTGCC-3’

IFI44L Forward: 5’-CCGTCAGTATTTGGAATGTGAAG-3’

IFI44L Reverse: 5’-TGAAACCAAGTCTGCATAGGG-3’

Chromatin Immunoprecipitation and Quantitative Real Time PCR (Chip-qPCR)

SPIN1-GFP transfected T778 cells were treated with EML631 (10 μM) overnight. And empty GFP vector transfected cells were used as control. The Chip assay was performed following EZ-Chip™ (Millipore, Catalog# 17-371) assay kit protocol. Briefly, Cells were crosslinked with 1% formaldehyde for 10 min at room temperature, and the reaction was stopped with 125mM glycine. Chromatin was sheared by using a Bioruptor sonication

device (Diagenode) and subjected to immunoprecipitation overnight at 4 °C by using 2 μg of GFP-antibody (Lifetechnologies, catalog# A6455). Immune complexes were incubated with 30 μL of a mix of Protein A/G agarose for overnight at 4 °C. After reverse crosslinking was performed, the DNA was eluted and purified using a PCR purification kit (Qiagene). The primer sequences used for qPCR analyses have previously been reported (PRM3[29], IL1B[45] & rDNA[46]) and are listed in below.

PRM3 Forward: 5'-GAAGTTATCCTGACTCACAC-3'

PRM3 Reverse: 5'-CCAGAGCCCAGGCCACAGCC-3'

IL1B Promoter (-199 to -109) Forward: 5'-AACGATTGTCAGGAAAACAATG-3'

IL1B Promoter (-199 to -109) Reverse: 5'-CTGGTTCATGGAAGGGC-3'

rDNA loci (12,855-12,970) Forward: 5'-ACCTGGCGCTAAACCATTCGT-3'

rDNA loci (12,855-12,970) Reverse: 5'-GGACAAACCCTTGTGTCGAGG-3'

RNA sequencing

Total RNA from control, SPIN1 siRNA-mediated knockdown (24 hour post-transfection) and EML631 treated (10 μM, 3 days) T778 cells were isolated by RNeasy mini kit (Qiagen). The experiment was performed with three independent biological replicates of each condition. The nine libraries were constructed by using the Illumina TruSeq stranded total RNA preparation kit (Illumina, cat #: RS-122-2301), which contained Ribo-Zero Gold that facilitates the depletion of rRNA. Importantly, we only amplified our libraries with 8 PCR cycles to minimize amplification induced noise. Purified libraries were quantified using a KAPA library quantification kit (KAPA Biosystems, Wilmington, MA), and then loaded on cBot (Illumina, San Diego, CA) at final concentration of 10 pM to perform cluster generation, followed by 2×76 bp sequencing on HiSeq 2500 (Illumina).

RNA-Seq bioinformatic analysis

For RNA-seq data, we used TopHat v2.0.12 to map the raw reads in FASTQ format to the hg19 human reference genome with the following parameters settings: --mate-std-dev 200 -p 8 -r 203. The mapped reads for each sample were saved in a BAM format file. We downloaded the reference genes UCSC KnownGenes from http://genome.ucsc.edu/cgi-bin/hgTables. The BAM file and reference genes were subjected to the Cuffdiff function in Cufflink suite v2.2.1 to calculate read counts and gene expression (fragments per kilobases per million, or FPKM). To identify differentially expressed genes based on read counts between different RNA-Seq samples, we used the normalizeQuantiles, estimateCommonDisp and estimateTagwiseDisp functions in the R package edgeR v3.14.0 to normalize the read counts, estimate common dispersion and estimate moderated tag-wise dispersion, respectively. The edgeR then defined differential genes based on a negative binomial test. In the final list of differential genes, we required each differential gene to have a differential FDR value smaller than 0.05 and FPKM value larger than 1 in at least one sample. To calculate RNA-Seq read density at each base pair in the genome, we used the genomecov function in Bedtools v2.17.0 to convert the mapped reads in a BAM file to read density values saved in a BedGraph format file. After normalizing the total density values in

each sample to 10 billions using a custom python script, we used the tool bedGraphToBigWig v4 to convert the normalized BedGraph file to a bigWig format file. We then subject the bigWig file to the Integrative Genomics Viewer (IGV) v2.3.67 to plot the read density figures for the genes SPIN1, IL1B and BST2.

In vitro pull down assay with GFP fusion proteins

HEK 293T cells were transiently tranfected with 53BP1-GFP or SPIN1-GFP using Lipofectamine 2000 according to manufacturer's instructions. Cells were lysed in RIPA buffer containing protease inhibitor cocktail (Roche). 30 μL of streptavidin agarose beads (Millipore) were prewashed with binding buffer (50 mM Tris HCl pH 7.5, 150 mM NaCl, 0.1% NP-40, 5 mM EDTA, 5 mM EGTA, 15 mM $MgCl_2$) and incubated with 10 μg of biotinylated compounds for 2 h with rocking at 4 °C. After three washing with 500 μL binding buffer to remove unbound compound, the compound-streptavidin agarose mix was incubated with 2 μg SPIN1-GST overnight with rocking at 4 °C. After three washes with 500 μL binding buffer, 30 μL of 2× protein loading buffer was added to the beads and boiled. The samples were loaded on a SDS-PAGE and detected by Western blotting.

In vitro GST pull down assay with GST fusion proteins

SPIN1-GST proteins were purified from *Escherichia coli* BL21 cell lysates using glutathione sepharose beads (GE Healthcare). In vitro GST pull-down assays were performed as described above by prebinding biotinylated H3K4me3 peptide with 1 μg GST fusion protein in 500 μL binding buffer. The samples were loaded on a SDS-PAGE and detected by Western blotting.

Cellular thermal shift assay

SPIN1-GFP transfected HeLa cells were treated with UNC1215, EML405, EML631, EML632 and EML633 (20 μM) overnight. Before treating with these small molecules, cells were washed twice in PBS to remove transfection reagents completely. They were then aliquoted for heating to different temperatures for 3 min, cooled at room temperature for 2 min and placed on ice. Cells were lysed by three freeze/thaw cycles in liquid nitrogen. Insoluble proteins were separated by centrifugation, and the soluble fractions were used for SDS PAGE and Western blotting.

Source of Chemicals used for Synthesis

Details on the synthesis for the tagged EML small molecule library are presented as supplemental data (Supplementary Note). All chemicals were purchased from Aldrich Chimica (Milan, Italy) and were of the highest purity. All solvents were reagent grade. Reactions were routinely monitored by TLC performed on aluminum-backed silica gel plates (Merck DC, Alufolien Kieselgel 60 F254) with spots visualized by UV light (λ = 254, 365 nm). Solvents were removed using a rotary evaporator operating at a reduced pressure of ~10 Torr.

High performance liquid chromatography (HPLC) was performed on a Shimadzu SPD 20A UV/VIS detector (λ= 220 nm) using Phenomenex Synergi Fusion – RP 80A (75× 4.60 mm; 4 μm) C18 column at room temperature, using a mobile phase A (water + 0.1% trifluoracetic

acid (TFA)) and B (ACN + 0.1% TFA) at a flow rate of 1 mL/min. The following gradient was applied: isocratic elution for 1 min at 10% of solvent B, linear increase from 10% to 95% of solvent B over 10 min, hold at 95% solvent B for 3 min. Preparative HPLC was performed using an Shimadzu Prominence LC-20AP with the UV detector set to 220 nm and 254 nm. Samples were injected onto a Phenomenex Synergi Fusion – RP 80A (150× 21 mm; 4 μm) C18 column at room temperature. Mobile phases of A (water + 0.1% trifluoracetic acid (TFA)) and B (ACN + 0.1% TFA) were used with a flow rate of 20 mL/min. A general gradient of 0–3 minutes at 5% B, 3–14 minutes increasing from 5 to 35% B, and 14–18 minutes increasing from 35 to 90% B was used, followed by a 90% B flush for another 2 minutes. Small variations in this purification method were made as needed to achieve ideal separation for each compound.

^{1}H and ^{13}C NMR spectra were recorded at 400 MHz and 100 MHz respectively with a Bruker Avance 400 spectrometer. Chemical shifts are reported in δ (ppm) relative to the internal reference tetramethylsilane (TMS). Mass spectra were recorded on a Finnigan LCQ DECA TermoQuest (San Jose, CA) mass spectrometer in electrospray positive and negative ionization modes (ESI-MS). All compounds that were evaluated in biochemical and biophysical assays had >95% purity as determined by ^{1}H NMR and HPLC.

Protein expression and purification

Human Spindlin1 ($SPIN1_{50–262}$) was cloned on the pRSFDuet vector with an N-terminal 6×His tag. Tudor domains of PHF20, 53BP1, and MBT domains of L3MBTL1 and L3MBTL3 were cloned on a pGEX4T vector containing an N-terminal GST tag. All recombinant proteins were overexpressed in *E. coli* strain BL21(DE3) at 16 °C for 18 h and induced by 0.2 mM isopropyl-β-D-thiogalactopyranoside (IPTG). Cells were harvested and lysed in lysis buffer: 0.2 M NaCl, 20 mM Tris, pH 8.0. After centrifugation, the 6×His-tagged SPIN1 was purified to homogeneity over successive HisTrap, anion exchange Q, and Superdex G75 columns (GE Healthcare). The GST-tagged PHF20, 53BP1, L3MBTL1, and L3MBTL3 were incubated with GST resin for 4 hours at 4 °C. Unbound protein was washed away with lysis buffer, and free proteins were collected by overnight PreScission protease cleavage at 4 °C. The free proteins were further polished through Superdex G75 column (GE Healthcare). All the proteins were concentrated to 20 mg/ml in 0.2 M NaCl, 20 mM Tris, pH 8.0 for future use.

Crystallization, data collection and structure determination

SPIN1 was incubated with EML405, EML631, EML632, or EML633 in a molar ratio of 1:2 for 30 min at 4 °C. Complex crystals were grown at 18 °C by mixing 1 μL of protein complex with 1 μL of reservoir solution using the sitting drop vapor diffusion method. The SPIN1-EML405 crystal was grown in buffer consisting 0.1 M HepesNa, pH 7.5, 10% PEG10000, 0.2 M $MgCl_2$. The crystal of SPIN1-EML631 was obtained from a reservoir solution containing 0.1 M HepesNa, pH 7.5, 25% PEG3350, 0.2 M $MgCl_2$.

Crystals were flash-frozen in liquid nitrogen under cryoprotectant conditions (reservoir solution supplemented with 10% glycerol). Diffraction data were collected at beamlines BL17U and BL18U of the Shanghai Synchrotron Radiation Facility (SSRF). All data sets

were indexed, integrated and scaled with the HKL2000 suite[47]. Crystal structures were determined by molecular replacement using MOLREP in CCP4[48] with previous published Spindlin1 structure (PDB code: 4MZH) as the search model. Model building and refinement were performed with COOT[49] and PHENIX[50], respectively. Data processing and refinement statistics are summarized in Table S2 & S3. The refined structures were validated by Procheck[51]. Ramachandran plot analysis showed that all residues of SPIN1-EML405 and SPIN1-EML631 complex structures are within the most favored or allowed regions. Structural analysis and figure preparation were mostly performed using PyMol (http://www.pymol.org). Electrostatic potential surfaces were calculated by the APBS tool[52] within PyMol.

Isothermal titration calorimetry

Calorimetric experiments were conducted at 25 °C with a MicroCal iTC200 instrument (GE Healthcare). All protein samples including SPIN1, PHF20, 53BP1, L3MBTL1 and L3MBTL3, were exchanged to the ITC buffer containing 150 mM NaCl and 20 mM HepesNa (pH 7.5) by gel filtration. Protein concentration was determined by its 280 nm absorbance. EML analogs and UNC1215 were lyophilized, quantified by weighing on a large scale and dissolved in the same ITC buffer for titration. Acquired calorimetric titration data were analyzed with Origin 7.0 (GE Healthcare) using the One Set of Binding Sites fitting model.

In vitro Thermal shift assay

The thermal shift assay (TSA) was performed with a CFX96™ real-time PCR instrument (Bio-Rad). A typical TSA solution is composed of 1 mg/mL SPIN1, 100 μM EML405 compound, 1× Sypro Orange (Invitrogen). All solutions were prepared in 25 μL under the buffer (150 mM NaCl, 20 mM Tris-Na at pH 8.0). During TSA assays, all samples were heated from 25 °C to 90 °C at a rate of 0.5 °C per minute. Protein denaturation was monitored by increased fluorescence signal of Sypro Orange, which captures exposed hydrophobic residues during thermal unfolding. The recorded curves were analyzed by the software CFX-Manager (Bio-Rad). The temperature corresponding to the inflection point was defined as T_m.

Statistical analysis

Statistical analyses were performed with GraphPad Prism 6.0 (GraphPad Software, Inc., San Diego, CA). Data are presented as means ±S.D., and *P*-values were calculated using Student's *t*-test for comparisons involving two groups and one-way ANOVA with Tukey's *post hoc* test for comparisons involving more than two groups. Furthermore, *P*-, values <0.05 are considered statistically significant.

Protein Domain Microarray probing

Protein-chemical interactions were screened using a protein domain microarray containing 98 GST fusion proteins. GST fusion proteins were codon optimized for bacterial expression and cloned into the pGEX4T vector. Recombinant proteins were expressed in *E. coli* BL21 and induced by 0.1 mM isopropyl-β-D-thiogalactopyranoside (IPTG) at 30 °C for 4 h. GST

fusion proteins were purified using glutathione sepharose beads (GE Healthcare). Equal amounts of GST fusion proteins were printed onto nitrocellulose-coated glass slides (Grace Bio-labs) using an array printer (Aushon). For compound screening, protein domain microarrays were first blocked for 1 h at RT with blocking buffer (1x PBS, 0.1% Tween 20, 3% milk, 3% bovine serum albumin). Streptavidin conjugated to Cy3 was incubated with 1 μM of biotinylated compound for 30 min on ice. Excess Cy3-streptavidin was removed with biotin agarose beads (Sigma). Cy3-labeled compound was incubated with microarray overnight at 4 °C. Microarray was washed three times with washing buffer (1x PBS, 0.1% Tween 20) for 10 min each. Microarrays were dried and scanned using an array scanner (Axon Instruments).

CODE AVAILABILITY STATEMENT

The atomic coordinates and structure factors of Spindlin1 bound to EML405 and EML631 have been deposited in the Protein Data Bank under accession codes of **5JSG** and **5JSJ**, respectively. The GEO accession number for our RNA-Seq data is **GSE92547**.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

M.T.B. is supported by an NIH grant (DK062248) and a CPRIT grant (RP130432) for the protein array analysis. The deep sequencing is supported by a CPRIT grant (RP120348) to J.S. The Center for Cancer Epigenetics at MDACC also supported this study. G.S. is supported by grants from the Italian Ministero dell'Istruzione, dell'Università e della Ricerca (MIUR), Progetti di Ricerca di Interesse Nazionale (PRIN 2012ZHN9YH), and from the Università di Salerno (Italy), and by European Cooperation in Science and Technology (COST Action CM1406). H.L. is supported by grants from the Major State Basic Research Development Program in China (2015CB910503 and 2016YFA0500700) and the Tsinghua University Initiative Scientific Research Program. N.B. is supported by the Odyssey Fellowship Program at The University of Texas MD Anderson Cancer Center. We thank the staff members at beamline BL17U and BL18U of the Shanghai Synchrotron Radiation Facility for assistance in data collection.

References

1. Huang H, Lin S, Garcia BA, Zhao Y. Quantitative Proteomic Analysis of Histone Modifications. Chemical Reviews. 2015; 115:2376–2418. [PubMed: 25688442]
2. Strahl BD, Allis CD. The language of covalent histone modifications. Nature. 2000; 403:41–45. [PubMed: 10638745]
3. Jenuwein T, Allis CD. Translating the Histone Code. Science. 2001; 293:1074–1080. [PubMed: 11498575]
4. Wozniak GG, Strahl BD. Hitting the 'mark': Interpreting lysine methylation in the context of active transcription. Biochimica et Biophysica Acta (BBA) - Gene Regulatory Mechanisms. 2014; 1839:1353–1361. [PubMed: 24631869]
5. Su Z, Denu JM. Reading the Combinatorial Histone Language. ACS Chemical Biology. 2016; 11:564–574. [PubMed: 26675328]
6. McGrath J, Trojer P. Targeting histone lysine methylation in cancer. Pharmacology & therapeutics. 2015; 150:1–22. [PubMed: 25578037]
7. Watson IR, Takahashi K, Futreal PA, Chin L. Emerging patterns of somatic mutations in cancer. Nature reviews. Genetics. 2013; 14:703–718.
8. Yang Y, Bedford MT. Protein arginine methyltransferases and cancer. Nat Rev Cancer. 2013; 13:37–50. [PubMed: 23235912]

9. Helin K, Dhanak D. Chromatin proteins and modifications as drug targets. Nature. 2013; 502:480–488. [PubMed: 24153301]
10. Beaver JE, Waters ML. Molecular Recognition of Lys and Arg Methylation. ACS Chem Biol. 2016; 11:643–653. [PubMed: 26759915]
11. Musselman CA, Lalonde M-E, Cote J, Kutateladze TG. Perceiving the epigenetic landscape through histone readers. Nat Struct Mol Biol. 2012; 19:1218–1227. [PubMed: 23211769]
12. Taverna SD, Li H, Ruthenburg AJ, Allis CD, Patel DJ. How chromatin-binding modules interpret histone modifications: lessons from professional pocket pickers. Nat Struct Mol Biol. 2007; 14:1025–1040. [PubMed: 17984965]
13. Filippakopoulos P, et al. Selective inhibition of BET bromodomains. Nature. 2010; 468:1067–1073. [PubMed: 20871596]
14. Palmer WS, et al. Structure-Guided Design of IACS-9571, a Selective High-Affinity Dual TRIM24-BRPF1 Bromodomain Inhibitor. J Med Chem. 2015; 59:1440–1454. [PubMed: 26061247]
15. Musselman CA, Khorasanizadeh S, Kutateladze TG. Towards understanding methyllysine readout. Biochim Biophys Acta. 2014; 1839:686–693. [PubMed: 24727128]
16. Herold JM, Ingerman LA, Gao C, Frye SV. Drug discovery toward antagonists of methyl-lysine binding proteins. Curr Chem Genomics. 2011; 5:51–61. [PubMed: 22145013]
17. Santiago C, Nguyen K, Schapira M. Druggability of methyl-lysine binding sites. Journal of computer-aided molecular design. 2011; 25:1171–1178. [PubMed: 22146969]
18. Miller TC, et al. Competitive binding of a benzimidazole to the histone-binding pocket of the Pygo PHD finger. ACS chemical biology. 2014; 9:2864–2874. [PubMed: 25323450]
19. Wagner EK, Nath N, Flemming R, Feltenberger JB, Denu JM. Identification and characterization of small molecule inhibitors of a plant homeodomain finger. Biochemistry. 2012; 51:8293–8306. DOI: 10.1021/bi3009278 [PubMed: 22994852]
20. Simhadri C, et al. Chromodomain antagonists that target the polycomb-group methyllysine reader protein chromobox homolog 7 (CBX7). Journal of medicinal chemistry. 2014; 57:2874–2883. [PubMed: 24625057]
21. Stuckey JI, et al. A cellular chemical probe targeting the chromodomains of Polycomb repressive complex 1. Nat Chem Biol. 2016; 12:180–187. [PubMed: 26807715]
22. Kireev D, et al. Identification of non-peptide malignant brain tumor (MBT) repeat antagonists by virtual screening of commercially available compounds. J Med Chem. 2010; 53:7625–7631. [PubMed: 20931980]
23. Perfetti MT, et al. Identification of a Fragment-like Small Molecule Ligand for the Methyl-lysine Binding Protein, 53BP1. ACS chemical biology. 2015; 10:1072–1081. [PubMed: 25590533]
24. Herold JM, et al. Small-molecule ligands of methyl-lysine binding proteins. J Med Chem. 2011; 54:2504–2511. [PubMed: 21417280]
25. James LI, et al. Discovery of a chemical probe for the L3MBTL3 methyllysine reader domain. Nat Chem Biol. 2013; 9:184–191. [PubMed: 23292653]
26. Milosevich N, Hof F. Chemical Inhibitors of Epigenetic Methyllysine Reader Proteins. Biochemistry. 2015; 55:1570–1583. [PubMed: 26650180]
27. Yun M, Wu J, Workman JL, Li B. Readers of histone modifications. Cell Res. 2011; 21:564–578. [PubMed: 21423274]
28. Su X, et al. Molecular basis underlying histone H3 lysine-arginine methylation pattern readout by Spin/Ssty repeats of Spindlin1. Genes & development. 2014; 28:622–636. [PubMed: 24589551]
29. Wang W, et al. Nucleolar protein Spindlin1 recognizes H3K4 methylation and stimulates the expression of rRNA genes. EMBO reports. 2011; 12:1160–1166. [PubMed: 21960006]
30. Jafari R, et al. The cellular thermal shift assay for evaluating drug target interactions in cells. Nature protocols. 2014; 9:2100–2122. [PubMed: 25101824]
31. Martinez Molina D, et al. Monitoring drug target engagement in cells and tissues using the cellular thermal shift assay. Science. 2013; 341:84–87. [PubMed: 23828940]
32. Espejo A, Cote J, Bednarek A, Richard S, Bedford MT. A protein-domain microarray identifies novel protein-protein interactions. Biochem J. 2002; 367:697–702. [PubMed: 12137563]

33. Espejo A, Bedford MT. Protein-domain microarrays. Methods Mol Biol. 2004; 264:173–181. [PubMed: 15020789]
34. Kim J, et al. Tudor, MBT and chromo domains gauge the degree of lysine methylation. EMBO Rep. 2006; 7:397–403. [PubMed: 16415788]
35. Barnes-Seeman D, Park SB, Koehler AN, Schreiber SL. Expanding the functional group compatibility of small-molecule microarrays: discovery of novel calmodulin ligands. Angew Chem Int Ed Engl. 2003; 42:2376–2379. [PubMed: 12783500]
36. Koehler AN, Shamji AF, Schreiber SL. Discovery of an inhibitor of a transcription factor using small molecule microarrays and diversity-oriented synthesis. J Am Chem Soc. 2003; 125:8420–8421. [PubMed: 12848532]
37. Yue W, Sun LY, Li CH, Zhang LX, Pei XT. [Screening and identification of ovarian carcinomas related genes]. Ai zheng = Aizheng = Chinese journal of cancer. 2004; 23:141–145. [PubMed: 14960231]
38. Gao Y, et al. Spindlin1, a novel nuclear protein with a role in the transformation of NIH3T3 cells. Biochemical and biophysical research communications. 2005; 335:343–350. [PubMed: 16098913]
39. Wang JX, et al. SPINDLIN1 promotes cancer cell proliferation through activation of WNT/TCF-4 signaling. Molecular cancer research : MCR. 2012; 10:326–335. [PubMed: 22258766]
40. Franz H, et al. The histone code reader SPIN1 controls RET signaling in liposarcoma. Oncotarget. 2015; 6:4773–4789. [PubMed: 25749382]
41. Chew TG, et al. A tudor domain protein SPINDLIN1 interacts with the mRNA-binding protein SERBP1 and is involved in mouse oocyte meiotic resumption. PLoS One. 2013; 8:e69764. [PubMed: 23894536]
42. Wagner T, et al. Identification of a small-molecule ligand of the epigenetic reader protein Spindlin1 via a versatile screening platform. Nucleic Acids Res. 2016; 44:e88. [PubMed: 26893353]
43. Sweis RF, et al. Discovery and Development of Potent and Selective Inhibitors of Histone Methyltransferase G9a. ACS Medicinal Chemistry Letters. 2014; 5:205–209. [PubMed: 24900801]
44. Robaa D, et al. Identification and Structure-Activity Relationship Studies of Small-Molecule Inhibitors of the Methyllysine Reader Protein Spindlin1. ChemMedChem. 2016; 11:2327–2338. [PubMed: 27634332]
45. Liang MD, Zhang Y, McDevit D, Marecki S, Nikolajczyk BS. The interleukin-1beta gene is transcribed from a poised promoter architecture in monocytes. J Biol Chem. 2006; 281:9227–9237. [PubMed: 16439360]
46. O'Sullivan AC, Sullivan GJ, McStay B. UBF binding in vivo is not restricted to regulatory sequences within the vertebrate ribosomal DNA repeat. Mol Cell Biol. 2002; 22:657–668. [PubMed: 11756560]
47. Otwinowski, Z., Minor, W. Methods in Enzymology. Vol. 276. Academic Press; 1997. p. 307-326.
48. Vagin A, Teplyakov A. Molecular replacement with MOLREP. Acta crystallographica. Section D, Biological crystallography. 2010; 66:22–25. [PubMed: 20057045]
49. Emsley P, Cowtan K. Coot: model-building tools for molecular graphics. Acta crystallographica. Section D, Biological crystallography. 2004; 60:2126–2132. [PubMed: 15572765]
50. Adams PD, et al. PHENIX: a comprehensive Python-based system for macromolecular structure solution. Acta crystallographica. Section D, Biological crystallography. 2010; 66:213–221. [PubMed: 20124702]
51. Laskowski RA, Moss DS, Thornton JM. Main-chain bond lengths and bond angles in protein structures. Journal of molecular biology. 1993; 231:1049–1067. [PubMed: 8515464]
52. Baker NA, Sept D, Joseph S, Holst MJ, McCammon JA. Electrostatics of nanosystems: application to microtubules and the ribosome. Proceedings of the National Academy of Sciences of the United States of America. 2001; 98:10037–10041. [PubMed: 11517324]

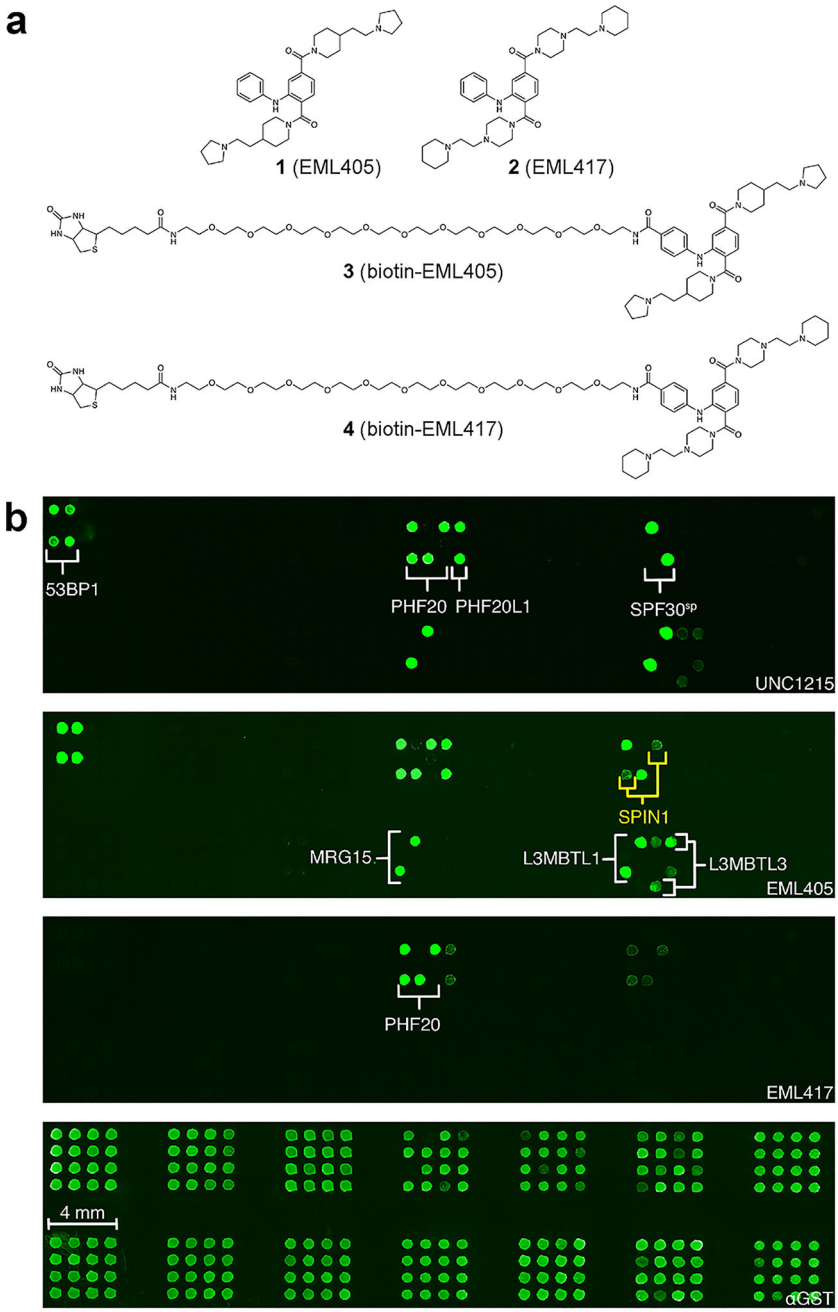

Figure 1.

Protein microarrays identify domain-binding compounds. (**a**) The structure of the EML405 and EML417 molecules and their biotinylated forms that were used to screen the protein domain microarrays. Similar tagged derivatives were generated for the full EML series. (**b**) EML405 selectively interacts with a number of aromatic cage-containing proteins on a microarray containing 96 epigenetic reader proteins across multiple domain types. The upper panel displays interactions of UNC1215 with domains on the array. The second panel shows EML405 interactions. The third panel demonstrates specific interactions of EML417 with the Tudor domains of PHF20. The lower panel demonstrates roughly equal deposition

of the GST fusion proteins on the array, as detected with an anti-GST antibody and Cy3-labeled secondary antibody.

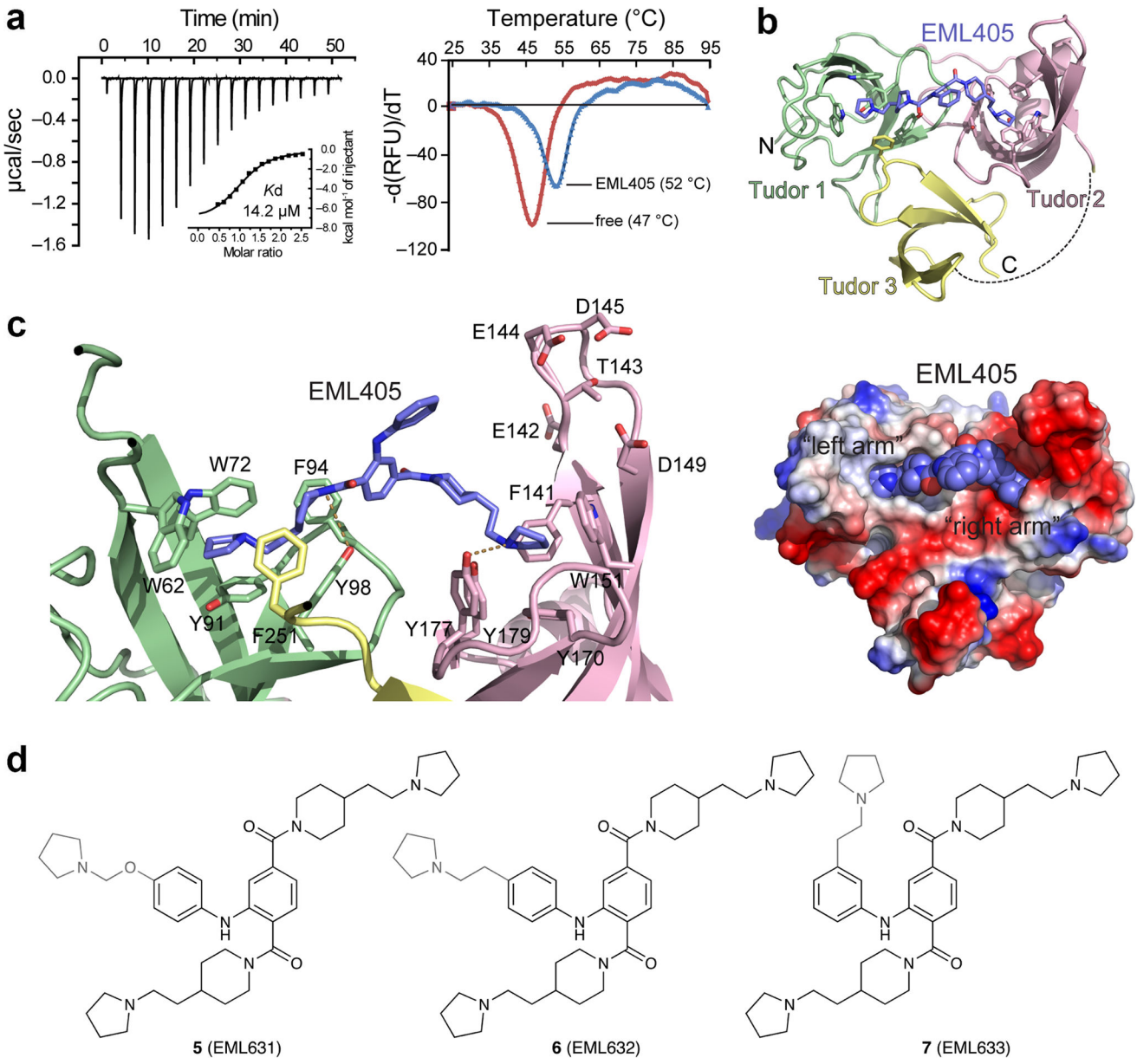

Figure 2.

Calorimetric and crystallographic studies of SPIN1-EML405. (**a**) ITC titration curve of SPIN1 with EML405 (left) and Thermo-fluor shift melting curve of apo-SPIN1 and SPIN1-EML405 (right). ITC and TSA curves were confirmed by three independent experiments. (**b**) Overall structure of SPIN1 bound to EML405 in ribbon view (top) and electrostatic potential surface view (bottom). EML405 is depicted as blue sticks. Electrostatic potential is expressed as a spectrum ranging from −6 kT/e (red) to +6 kT/e (blue). EML405 was excluded for electrostatic potential calculation. (**c**) Details of EML405 compound recognition by SPIN1. Dashes, hydrogen bonds. (**d**) Structures of three EML405 analogs EML631, EML632 and EML633.

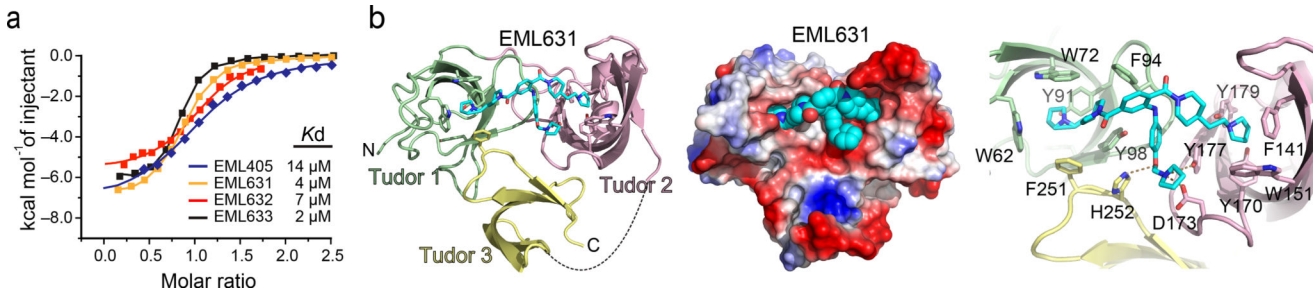

Figure 3.

Calorimetric and crystallographic studies of SPIN1-EML631. (**a**) ITC fitting curves of SPIN1 with EML405, EML631, EML632 and EML633. ITC titrations were confirmed by three independent experiments. (b) Overall view (left and middle) and close-up view (right) of SPIN1-EML631 complex structure. EML631 is depicted as cyan sticks or space-filling spheres. Electrostatic potential is expressed as a spectrum ranging from −6 kT/e (red) to +6 kT/e (blue). Dashes, hydrogen bonds.

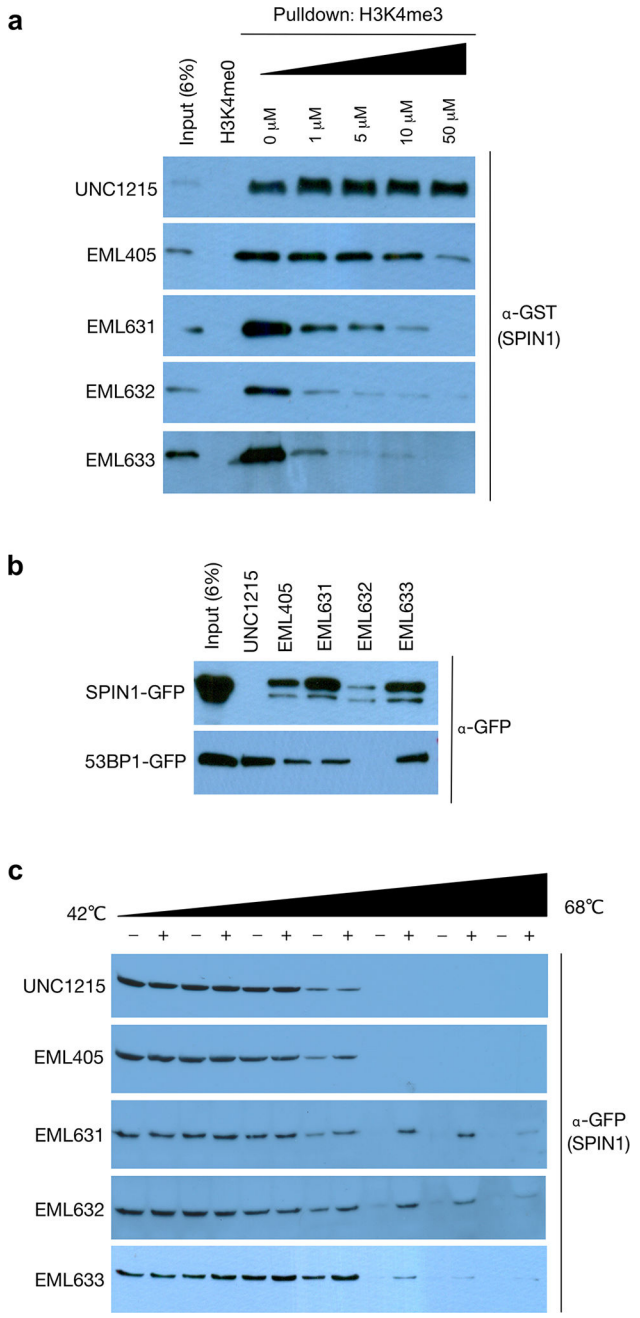

Figure 4.

EML631-633 block the "reader" ability of SPIN1 and are cell permeable. (a) *In vitro* pull-down assays were performed by using purified SPIN1-GST tudor and biotinylated H3K4me3 peptides in the presence of increasing concentration of UNC1215, EML405, EML 631, EML632 and EML633. Biotinylated H3K4me0 peptides were used as negative control. (b) Whole cell lysates of transiently transfected HEK293T cells with SPIN1-GFP and 53BP1-GFP were used to pull down with biotinylated UNC1215, EML405, EML 631, EML632 and EML633. (c) CETSA was carried out with SPIN1-GFP transfected HeLa cells grown in the presence of UNC1215, EML 405, EML631, EML632 and EML633 (20 μM).

CETSA data was confirmed by two independent experiments. Original uncropped Western blots are provided in Supplementary Figure 9–11.

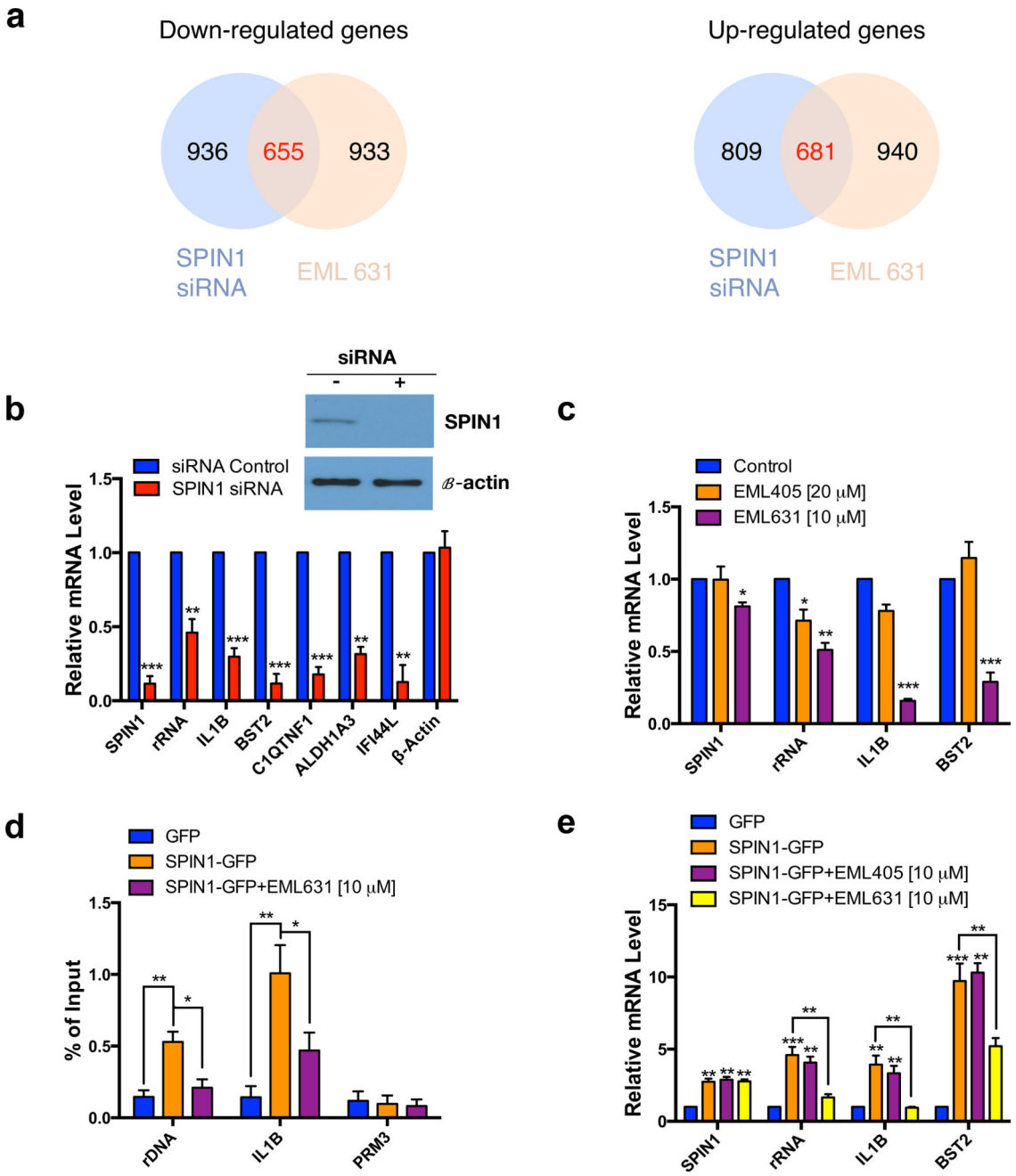

Figure 5.

RNA-Seq identification of SPIN1 regulated transcripts and inhibition of this co-activator activity by EML631. (**a**) Venn diagrams showing the numbers of down-regulated and up-regulated genes, as determined by RNA-Seq, in response to SPIN1 knockdown and EML631 treatment. Randomly expected overlap sizes are 88 (down) and 85 (up), which in both cases represent over 7-fold enrichment (P ‹ 2.2e-16, Fisher's exact test). (**b**) The mRNA levels of SPIN1 target genes, in control and SPIN1 knockdown T778 cells, were analyzed by RT-qPCR. The knockdown level of SPIN1 protein was detected by Western blotting. (**c**) RT-qPCR analysis of T778 cells after treatment with or without EML 405 (20 μM) and EML631

(10 μM) for 4 days. (**d**) The effect of EML631 (10 μM) on blocking SPIN1 chromatin association was assessed by α-GFP ChIP-qPCR analysis of two active loci (rDNA and IL1B) and a control locus (PRM3). ChIP-qPCR data is shown as a ratio relative to input. (**e**) The effect of EML405 (10 μM) and EML631 (10 μM) was monitored, analyzing the gene expression changes in up-regulated SPIN1 target genes. Gene expression was normalized to GAPDH. The mean value for the control groups was arbitrarily set as 1. All data represent the average of three independent experiments (biological replicates), each subjected to three independent RT-qPCR (total of n=9) and Chip-qPCR reactions (total of n=9). S.D. is denoted by error bar. *P<0.05, **P<0.01, ***P<0.001 (Student's t-test). Original uncropped Western blots are provided in Supplementary Figure 12.