Published in final edited form as:

Nat Chem Biol. 2013 March ; 9(3): 184–191. doi:10.1038/nchembio.1157.

Discovery of a chemical probe for the L3MBTL3 methyl-lysine reader domain

Lindsey I. James[1], **Dalia Barsyte-Lovejoy**[2], **Nan Zhong**[2], **Liubov Krichevsky**[2,3,4], **Victoria K. Korboukh**[1], **Martin J. Herold**[1], **Christopher J. MacNevin**[1,8], **Jacqueline L. Norris**[1], **Cari A. Sagum**[5], **Wolfram Tempel**[2], **Edyta Marcon**[6], **Hongbo Guo**[6], **Cen Gao**[1], **Xi-Ping Huang**[7,8], **Shili Duan**[4], **Andrew Emili**[6], **Jack F. Greenblatt**[6], **Dmitri B. Kireev**[1], **Jian Jin**[1], **William P. Janzen**[1], **Peter J. Brown**[2], **Mark T. Bedford**[5], **Cheryl H. Arrowsmith**[2,3,4,*], and **Stephen V. Frye**[1,*]

[1]Center for Integrative Chemical Biology and Drug Discovery, Division of Chemical Biology and Medicinal Chemistry, UNC Eshelman School of Pharmacy, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina 27599, USA

[2]Structural Genomics Consortium, University of Toronto, Toronto, Ontario, M5G 1L7, Canada

[3]Department of Medical Biophysics, University of Toronto, 101 College Street, Toronto, Ontario, M5G 1L7, Canada

[4]Ontario Cancer Institute and Campbell Family Cancer Research Institute, University of Toronto, 101 College Street, Toronto, Ontario, M5G 1L7, Canada

[5]M. D. Anderson Cancer Center Department of Carcinogenesis, University of Texas, Smithville, TX, USA

[6]Banting and Best Department of Medical Research, Donnelly Centre, 160 College Street, Toronto, ON, M5S 3E1

[7]National Institute of Mental Health Psychoactive Drug Screening Program, University of North Carolina at Chapel Hill Medical School, Chapel Hill, North Carolina 27599, USA

[*]Correspondence and requests for materials should be addressed to S.V.F. or C.H.A., svfrye@email.unc.edu or carrow@uhnres.utoronto.ca.

Author contributions

L.I.J. synthesized all compounds and related analogs and performed ITC studies; D.B.-L. and L.K. performed immunofluorescence FRAP, affinity purification, and co-immunoprecipitation studies; N.Z. and W.T. solved and analyzed the X-ray crystal structure of the UNC1215-L3MBTL3 complex; V.K.K. and W.P.J. performed and analyzed AlphaScreen studies; C.J.M. synthesized the mero76-UNC1215 conjugate; J.L.N. purified proteins and performed mutagenesis; C.C.S. and M.T.B. performed protein array and protein pull-down experiments; E.M., H.G, A.E., and J.F.G performed MS-based studies; S.D. cloned mammalian expression vectors for all cellular studies; X.-P.H. performed and analyzed GPCR selectivity studies; L.I.J., D.B.-L., N.Z., L.K., J.M.H., V.K.K., C.G., D.B.K., J.J., W.P.J., P.J.B., M.T.B, C.H.A., and S.V.F. designed studies and discussed results; L.I.J, C.H.A., and S.V.F. wrote the paper.

Additional information

Supplementary information is available online at http://www.nature.com/naturechemicalbiology/. Reprints and permissions information is available online at http://npg.nature.com/reprintsandpermissions/.

Competing financial interests

The authors declare no competing financial interests.

[8]Department of Pharmacology, University of North Carolina at Chapel Hill Medical School, Chapel Hill, North Carolina 27599, USA

Abstract

We describe the discovery of UNC1215, a potent and selective chemical probe for the methyl-lysine (Kme) reading function of L3MBTL3, a member of the malignant brain tumor (MBT) family of chromatin interacting transcriptional repressors. UNC1215 binds L3MBTL3 with a K_d of 120 nM, competitively displacing mono- or dimethyl-lysine containing peptides, and is greater than 50-fold selective versus other members of the MBT family while also demonstrating selectivity against more than 200 other reader domains examined. X-ray crystallography identified a novel 2:2 polyvalent mode of interaction. In cells, UNC1215 is non-toxic and binds directly to L3MBTL3 via the Kme-binding pocket of the MBT domains. UNC1215 increases the cellular mobility of GFP-L3MBTL3 fusion proteins and point mutants that disrupt the Kme binding function of GFP-L3MBTL3 phenocopy the effects of UNC1215. Finally, UNC1215 demonstrates a novel Kme-dependent interaction of L3MBTL3 with BCLAF1, a protein implicated in DNA damage repair and apoptosis.

The post-translational modification of lysine residues by acetylation and methylation plays a central role in chromatin function, primarily through the creation of binding sites for the readers of these marks.[1,2] While acetylation of lysine eliminates the residue's positive charge, methylation is a more subtle modification, shifting the chemical properties of lysine toward a more diffuse, polarizable positive charge as methylation proceeds from monomethyl (Kme1) to trimethyl (Kme3).[3,4] Recently, small molecule probes of bromodomains, the readers of acetylated lysine, have been reported, and it would be difficult to overstate the impact of these compounds on the understanding of critical gene transcription events such as regulation of the c-myc oncogene.[5–7] While the current repertoire of acetyl lysine readers in the human genome is limited primarily to the 57 structurally homologous bromodomains and possibly a subset of plant homeodomains (PHDs)[8], there are more than 200 methyl-lysine reader domains described within several protein families: PHDs; the so-called 'royal family' made up of Tudor, Agenet, chromo, PWWP and MBT domains; and the WD40 repeat proteins, WDR5 and EED.[9,10] This large Kme reader family seems destined to expand further as research in epigenetics continues.[11] A unifying feature of all of these domains is the existence of an "aromatic cage" that comprises the Kme binding pocket.[12,13] While computational analysis suggests that this target-class has substantial druggability[13] and we have successfully identified several weak and nonspecific small molecule ligands for MBT domains,[14,15] no high-quality small molecule chemical probes[16] have been reported for any Kme reader domain to date.

MBT domains selectively recognize mono and dimethyl-lysine versus unmethylated and trimethylated lysine and have been functionally associated with repression of gene expression, while their misregulation has been shown to contribute to various disease phenotypes.[17,18] For example, SCML2, L3MBTL2, and L3MBTL3 were found to be homozygously deleted in patients with medulloblastomas,[19] while MBTD1, L3MBTL1, and L3MBTL3 play a role in hematopoiesis.[20–23] Some of the human MBT proteins are also known to be part of larger chromatin remodeling complexes. Recently, a family-wide

systematic analysis of MBT-histone interactions was reported, suggesting that some MBT domains recognize methyl-lysine histone peptides in a sequence-selective fashion, while others, such as L3MBTL3, are more promiscuous.[24]

The discovery of the first small molecule probe to a previously untargeted protein domain is often a significant challenge, especially for domains that mediate 'protein-protein' interactions. Herein we report the discovery and characterization of UNC1215 (**1**) as a first in class chemical probe for the Kme1,2 reading function of L3MBTL3, a relatively uncharacterized member of the human MBT family.

RESULTS

Discovery of a potent L3MBTL3 inhibitor, UNC1215

Our work toward chemical probes for Kme binding domains initially focused on MBT domains due to the potential druggability of their 'cavity insertion' mode of Kme recognition that buries most of the lysine side-chain in contrast to the more solvent exposed, 'surface recognition' mode of many other Kme readers.[12,13] We were also intrigued by the potential polyvalency of MBT domains due to the presence of multiple repeats within each of the nine human MBT domain containing proteins, although no definitive evidence for polyvalent binding to histone peptides was available.[25] As our first ligands for L3MBTL1 exhibited only modest affinities (K_d = 5 μM)[15], we designed a series of ligands that incorporated two Kme mimics in the hope that higher affinity might result for Kme binders such as 53BP1, which also recognizes an adjacent unmodified arginine in addition to the H4K20me2 modification.[26] Towards this end, we synthesized a series of dibasic compounds, including UNC1021 (**2**) (Fig. 1a). While the affinity of UNC1021 for 53BP1 remained in the micromolar range, our target-class cross-screening approach revealed high affinity for L3MBTL3 (Supplementary Results, Supplementary Table 1). A series of structure-activity relationship (SAR) studies aimed at improving affinity and cellular activity towards L3MBTL3 led us to append an aniline ring to the core of the molecule, resulting in UNC1215 (Fig. 1a). The IC_{50} of UNC1215 for L3MBTL3 was 40 nM as determined by an AlphaScreen methylated histone peptide competition assay (Supplementary Table 1 and Supplementary Fig. 1),[27] and the direct binding affinity of UNC1215 was confirmed by isothermal titration calorimetry (ITC), yielding a K_d of 120 nM (Fig. 1b). Given the novelty and high-affinity of UNC1215 for L3MBTL3, we pursued its characterization as a potential chemical probe.[16]

Based on previous MBT structure activity relationships (SAR) that established size constraints for the Kme binding pocket,[14] we also synthesized UNC1079 (**3**) (Fig. 1a), the piperidine analog of UNC1021, as a structurally similar but significantly less potent inhibitor for use as a negative control in cellular studies. The low anticipated affinity of UNC1079 was confirmed, as it demonstrated an activity versus L3MBTL3 of > 10 μM by AlphaScreen, which is >1000-fold weaker than UNC1215. UNC1079 also displayed weak binding by ITC (Supplementary Fig. 2).

Mechanism of action studies

MBT domains utilize an aromatic cage that has evolved to recognize Kme1,2 modifications in a relatively non-sequence selective fashion.[28,29] In addition to this π-cation binding mode, a critical aspartic acid residue in the binding pocket hydrogen bonds to the protonated amine. L3MBTL3 contains three tandem MBT domain repeats (Fig. 1c,d), and we sought to determine whether only one or more than one of the potential Kme binding sites was interacting with UNC1215. The aspartic acid that mediates Kme binding of the second domain in the closely related Kme reader, L3MBTL1, is conserved in L3MBTL3 (D381), and it was predicted that this key residue similarly hydrogen bonds to one of the pyrrolidine nitrogens of UNC1215. Using a construct containing the 3 MBT domains of L3MBTL3 (3MBT), a D381A mutant was prepared and its interaction with UNC1215 was examined by ITC (Fig. 1b), revealing that binding to UNC1215 is completely abolished. Interestingly, when the corresponding mutation was made to the first domain of 3MBT (D274A), binding was diminished about 30-fold relative to wildtype (Fig. 1b), suggesting that domain 1 contributes to the potency of UNC1215 for wild-type L3MBTL3. The corresponding aspartic acid in domain 3 (D485A) was also mutated, but the resulting His-tagged protein did not bind to nickel affinity resin, and was therefore not successfully purified. As UNC1215 exhibits about a 75-fold selectivity for L3MBTL3 over L3MBTL1, this further suggested that the two highly homologous proteins differ significantly in their mode of recognition of UNC1215.

Co-crystal structure of the UNC1215-L3MBTL3 complex

To further investigate the interaction of UNC1215 with L3MBTL3, an X-ray crystal structure of the complex was solved at a resolution of 2.6 Å, providing structural insight into the binding mode (Fig. 2). Like L3MBTL1, L3MBTL3 presents a three-leaved propeller overall architecture with all potential ligand binding pockets (domains 1–3) facing in the same direction.[30] Surprisingly, the co-crystal structure revealed a novel 2:2 complex of UNC1215-L3MBTL3, involving both domain 1 and domain 2 binding pockets, which is consistent with the affinities of the aromatic cage mutants discussed above. Analysis of the X-ray structure reveals that UNC1215 binding is primarily mediated by interaction with the aromatic cage of domain 2, consisting of F387, F405, W408, and Y412, and via a key hydrogen bond between the pyrrolidine nitrogen and D381. While the pyrrolidine amine meta to the aniline substituent is deeply buried in domain 2 of one L3MBTL3 protein (similar to the mono-basic, UNC669/L3MBTL1 structure[15]), the pyrrolidine ortho to the aniline bridges the two proteins by interacting with domain 1 of a second L3MBTL3 molecule. Another UNC1215 molecule then binds in a reciprocal fashion to the remaining available domain 2 and domain 1 binding pockets, resulting in the observed 2:2 dimer. While the D381A mutant completely abrogates binding (Fig. 1b), mutagenesis of the aromatic cage residue F387 to alanine weakens binding about 10-fold (K_d = 1.0 μM, Supplementary Fig. 2). The X-ray structure also suggests a possible stacking interaction between UNC1215 and H277, and upon mutagenesis, H277A 3MBT showed a marked loss of affinity relative to wild-type as measured by ITC (Supplementary Fig. 2). Size exclusion chromatography indicated that in the absence of compound, 3MBT migrates as a monomer, but in the presence of compound, 3MBT behaves as a dimer (Supplementary Fig. 3). This

strongly supports the mechanism of inhibition observed in the crystal structure as being operative in solution as well as in the crystal lattice.

UNC1215 is a selective inhibitor of L3MBTL3

To evaluate the selectivity of UNC1215, the compound's affinity for a wide range of epigenetic and non-epigenetic targets was surveyed (Supplementary Table 2). First, the selectivity against a small set of Kme readers was evaluated by AlphaScreen, including both related proteins from the MBT family as well as representative proteins from other Kme reader families. The selectivity of UNC1215 for L3MBTL3 over other MBT domain containing proteins (L3MBTL4, L3MBTL1, MBTD1, and SFMBT) was 50-fold or greater, while 100-fold selectivity was observed for L3MBTL3 over the tandem Tudor domain of 53BP1 which also recognizes H4K20me2 (Supplementary Table 1). UNC1215 showed no activity at concentrations up to 30 μM against the Tandem Tudor domain of UHRF1, the chromodomain of CBX7, or the PHD domain of JARID1A.

In order to better assess the selectivity of UNC1215 against Kme and methylarginine reader proteins, protein microarrays containing over 250 protein domains of chromatin-associated effector proteins spotted in duplicate onto a nitrocellulose-coated membrane were probed with a biotinylated analog of UNC1215 (Fig. 3a, Supplementary Fig. 4).[31] Biotin was covalently attached to UNC1215 via a PEG linker so as to minimize potential hydrophobic linker effects, and high-affinity binding of the biotinylated analog to L3MBTL3 was confirmed by ITC (Supplementary Fig. 2). In order to visualize the binding interactions of the compound after incubation, the membrane was probed with fluorescently tagged streptavidin, revealing positive interactions with only a small number of proteins including the tudor and/or MBT domains of L3MBTL1, PHF20, PHF20L1, 53BP1, and SPF30(S.p), as well as the chromodomain MRG15. Interestingly, with the exception of MRG15, each is a known H4K20me1 or H4K20me2 reader, although other H4K20me2 reader domains such as the BAH domain of ORC1[11] did not bind UNC1215 (Supplementary Fig. 5). This was confirmed by subsequently probing the array with a fluorescent H4K20me2 peptide, and as expected, binding was observed to each of these proteins, excluding MRG15 (Supplementary Fig. 4). It must be noted that in the case of both UNC1215 and H4K20me2, binding to the MBT domains of L3MBTL3 was not observed, likely due to instability or misfolding of the protein on the nitrocellulose membrane or precipitation upon freezing and thawing. Therefore, selectivity estimates from this technique can only be made for the readers where a positive interaction with a histone control peptide is observed.

The affinity of proteins positive for binding to biotinylated-UNC1215 on the membrane array was next quantified by ITC. Both L3MBTL1 and the Tudor domain of PHF20 bind UNC1215 weakly by ITC, with K_d values of 9.4 ± 1.7 μM and 5.6 ± 0.5 μM respectively, while other representative binding proteins had even lower affinities (Supplementary Table 2, Supplementary Fig. 2). The interaction of UNC1215 with purified L3MBTL3, L3MBTL1, 53BP1, and the Tudor domain of PHF20, as well as selectivity versus SPF30 and HP1γ, was also confirmed by pull-down experiments with biotinylated-UNC1215 (Supplementary Fig. 5). Competition studies in an H4K20me2 pull-down experiment also demonstrated nanomolar potency for antagonism of L3MBTL3 pull-down by UNC1215;

however, micromolar antagonism was observed in the case of 53BP1 (Fig. 3b). These studies confirm the potency and high degree of selectivity of UNC1215 for L3MBTL3 over other available methyl-lysine and methyl-arginine binding proteins.

The selectivity of UNC1215 for L3MBTL3 over other epigenetic targets was also evaluated. Notably, UNC1215 was inactive against a panel of histone methyltransferases as evaluated by monitoring the transfer of a radioactive methyl group from *S*-adenosylmethionine (SAM) to lysine or arginine (Supplementary Fig. 6). Furthermore, the stabilization of various bromodomains and lysine demethylases by UNC1215 was analyzed by differential scanning fluorimetry, in each case producing Tm shifts <0.5 °C, consistent with very weak or no interaction. Due to the distinctly 'aminergic' pharmacophore of UNC1215, we also surveyed a panel of G-protein coupled receptors (GPCRs) and ion channels (Supplementary Table 5). Specifically, we tested UNC1215 in the US National Institute of Mental Health's Psychoactive Drug Screening Program (PDSP) Selectivity Panel, revealing >50% inhibition of 8 targets in the panel at 10 μM. Using a radioligand competition binding assay, K_i values were determined for each of the 8 initial hits, revealing affinity for M_1 (K_i = 97 nM) and M_2 (K_i = 72 nM) muscarinic receptors. As radioligand competition binding assays can overestimate functional potencies, functional assays were performed for M_1 and M_2 muscarinic receptors; UNC1215 had no agonist activity against either target and weak antagonist activity against M_1 (IC_{50} = 3.5 μM) and M_2 (40% inhibition at 30 μM) (Supplementary Fig. 7). Lastly, UNC1215 was tested against a panel of 50 diverse kinases and only showed weak potency against FLT3 (64% inhibition at 10 μM). Thus, we conclude that within the molecular targets surveyed, UNC1215 is highly selective for L3MBTL3, and that any effects observed in cells upon treatment with UNC1215 but absent upon treatment with the negative control, UNC1079, could reasonably be attributed to antagonism of Kme binding by L3MBTL3.

UNC1215 is a potent antagonist of L3MBTL3 in cells

Prior to assessing the cellular potency of UNC1215, we evaluated the cytotoxicity of both UNC1215 and the negative control compound UNC1079 to ensure future cellular studies would not be limited by compound toxicity. Both compounds showed no observable level of cytotoxicity up to 100 μM as measured by a CellTiter-Glo luminescent cell viability assay (Supplementary Fig. 9). Furthermore, incubation of UNC1215 with or without HEK293 cells in cell media for 72 hours did not produce degradation products of UNC1215, indicating high stability under cellular assay conditions.

While the cellular function of L3MBTL3 is largely unknown, MBT domains are thought to assist in targeting chromatin regulatory complexes to appropriate genomic loci via binding to specific mono- or dimethylated lysine residues within histone tails.[32–36] We reasoned that UNC1215 would compete with histone Kme marks or other Kme sites in chromatin and thereby "release" L3MBTL3 from the immobile chromatin component of the nucleus. In order to test this hypothesis, we performed fluorescence recovery after photobleaching (FRAP) experiments to assess the relative cellular mobility of the MBT domains of L3MBTL3 in the presence and absence of UNC1215. Human embryonic kidney 293 cells transfected with a GFP fusion protein of the 3 MBT domains of L3MBTL3 (GFP-3MBT)

demonstrated a time-dependent recovery of fluorescence intensity following photobleaching (Fig. 4a). Treatment with UNC1215 decreased the recovery time in a dose responsive manner, resulting in an estimated EC_{50} of 50–100 nM and indicating that UNC1215 promotes diffusibility of GFP-3MBT within the nucleus. These results suggest that UNC1215 competes with cellular factors for binding of the MBT domains at concentrations well below 1 μM. Importantly, the effects on GFP-3MBT mobility are limited to UNC1215, as UNC1079 showed no effect on FRAP at concentrations up to 0.1 mM. Furthermore, GFP-3MBT containing either the domain 2 D381A mutation or the domain 1 D274A mutation also showed higher mobility than wild-type protein (Supplementary Fig. 10 and Supplementary Table 6), providing additional evidence that both domains are engaged in chromatin recognition and that the binding mode observed in the co-crystal structure may be relevant to L3MBTL3 binding to methylated protein targets in cells. The magnitude of the effect of these mutations on mobility also mimicked their effect on UNC1215 binding (Fig. 1b). In addition, in HEK293 cells co-transfected with GFP-3MBT and flag-3MBT, GFP-3MBT co-immunoprecipitated with flag-3MBT in both the absence and presence of UNC1215 (Supplementary Fig. 20). Interestingly, this result supports the notion that the MBT domains may have a tendency to dimerize in cells.

In U2OS and 293 cells, the N-terminal GFP-3MBT fusion protein and the C-terminal 3MBT-GFP fusion protein were both found to localize to the nucleus and form distinct nuclear foci (Fig. 4b and Supplementary Fig. 11). However, the m-Cherry 3MBT fusion protein does form foci under similar conditions (Supplementary Fig. 11). Although the physiological relevance of the GFP-3MBT foci are not known, in order to further probe the ability of UNC1215 to promote a cellular response, we examined the capacity of UNC1215 to inhibit GFP-3MBT foci formation. UNC1215 showed potent effects on the subnuclear localization of GFP-3MBT, demonstrating an IC_{50} of approximately 500 nM for disruption of foci formation (Fig. 4b, Supplementary Fig. 12). UNC1215 did not affect the overall level of GFP-3MBT fusion proteins. The inactive compound UNC1079 was ineffective in reducing the extent of foci formation, while the domain 2 D381A mutant showed no foci formation and the domain 1 D274 mutant showed a reduction in foci formation of about 50 percent (Fig. 4c and Supplementary Fig. 13). Thus, although the foci formation is restricted to specific GFP fusions of 3MBT, inhibition by UNC1215 phenocopies the L3MBTL3 mutants, consistent with a cellular mechanism in which UNC1215 antagonizes the interaction of the MBT domains with one or more cellular Kme targets. The fact that both domain 1 and domain 2 aromatic cage mutants disrupt foci formation again supports the hypothesis that both domains are engaged in recognition of protein targets *in vivo* and that UNC1215 effectively competes with these interactions.

Interestingly, in 293 or U2OS cells transfected with an N-terminal GFP fusion of the full length L3MBTL3 protein (GFP-FLMBT), distinct foci-like subnuclear localization was observed and these foci could no longer be disrupted by treatment with UNC1215 (Fig. 4b). In this case, neither the domain 1 nor the domain 2 mutant showed a decrease in the extent of nuclear foci formation in the context of the full length protein relative to wildtype (Supplementary Fig. 13). The full length C-terminal L3MBTL3 fusion localized to the DNA-rich regions of the nucleus and not the nucleolus (Supplementary Fig. 11). Since

protein tags and other domains or regions of the full length protein appear to contribute to foci formation, we investigated the contribution of other domains to the localization of L3MBTL3. In addition to the MBT domains, L3MBTL3 harbors a Zn finger, a putative RNA-binding domain, and a C-terminal sterile alpha motif (SAM) domain (Fig. 1c). SAM domains are protein-protein interaction domains that can homo- or hetero-dimerize and often play a role in the formation of cellular protein complexes.[37] Recently a polymer forming ability of the SAM domain of L3MBTL3 was proposed.[38] We therefore hypothesized that, in addition to the binding of the MBT domains to Kme, the SAM domain of L3MBTL3 may contribute to the interaction of the full length protein with chromatin, resulting in resistance to UNC1215 inhibition of foci formation. Deletion of the SAM domain in N-terminally tagged FLMBT (GFP-FLMBTΔSAM) resulted in foci elimination (Supplementary Fig. 14), suggesting that the SAM domain also contributes to foci formation in the context of the full length protein. This may explain why the full length protein foci are relatively insensitive to UNC1215 inhibition of the Kme-binding activity of the MBT domains, and suggests that other protein and/or chromatin components of the GFP-FLMBT foci may also mediate localization. Having demonstrated the cellular potency of UNC1215 for GFP-3MBT by FRAP and foci inhibition, we were interested in confirming the binding of UNC1215 to the full length protein given its relative resistance to relocalization and an uncertain role of the SAM domain and other domains within L3MBTL3.

UNC1215 binds and co-localizes with full length L3MBTL3

With its propensity to form foci, we used the N-terminally tagged GFP full length L3MBTL3 fusion protein as a model system to further map the interaction between L3MBTL3 and UNC1215 in living cells. In order to address whether the MBT domains of GFP-FLMBT were available for interaction with UNC1215, the cell membrane permeable and long wavelength emitting merocyanine dye, mero76, was appended via a hexadiamine linker to the aniline ring of UNC1215. Upon treatment of HEK293 cells with the mero76–UNC1215 conjugate, the fluorescent inhibitor co-localized with GFP-FLMBT (Fig. 5a), confirming that the probe effectively binds the full length protein in cells. When the cells were treated with dye alone (not covalently linked to UNC1215), no co-localization was observed indicating that the dye has no independent effect on the observed results (Supplementary Fig. 15). In addition, the fluorescent compound shows no foci formation in the presence of the domain 2 D381A mutant (Supplementary Fig. 15). By FRAP, the mobility of GFP-FLMBT was also found to increase upon treatment with UNC1215, albeit to a lesser extent than GFP-3MBT, and both GFP-FLMBT D274A and D381A mutants from domains 1 and 2, respectively show faster molecular movement (Supplementary Fig. 16), once again consistent with the notion that both domain 1 and domain 2 are utilized for recognition *in vivo*. Overall, these results support a cellular binding mechanism of UNC1215 that mirrors our biophysical data, namely that UNC1215 binds to L3MBTL3 via both the domain 1 and domain 2 aromatic cage pockets thereby displacing native cellular Kme containing targets.

As final confirmation of cellular binding of UNC1215 to full length L3MBTL3, we investigated whether biotinylated-UNC1215 could affinity-purify L3MBTL3 from treated cells. Indeed, overexpressed full length flag-L3MBTL3 from HEK293 cell lysates could be

efficiently affinity-purified with this reagent and competed with non-biotinylated UNC1215 (Fig. 5b). Taken together, these experiments provide convincing evidence that UNC1215 is a high affinity Kme competitive ligand for the MBT domains of L3MBTL3 in the context of the full length protein in cells.

BCLAF1 is a novel target of L3MBTL3

In order to better characterize the functional role of UNC1215, we sought to identify cellular interacting partners of L3MBTL3 whose interaction may be mediated by lysine methylation. We first analyzed several commercial antibodies to histone marks and common chromatin factors (Supplementary Table 7) to look for colocalization with GFP-FLMBT or GFP-3MBT foci. However, we did not observe any clear colocalization with several mono or dimethyl histone marks or with other proteins forming distinct nuclear structures in immunofluorescence experiments, possibly because of epitope unavailability in some cases.

We next investigated potential L3MBTL3 interacting proteins using affinity purification and mass spectrometry. After generating HEK293 cells transiently transfected with GFP-3MBT or GFP-FLMBT, the GFP-fused L3MBTL3 proteins were purified from cells cultured in the presence or absence of UNC1215, and the co-purifying proteins were identified by mass spectrometry. Among a number of proteins identified from this screen as potential L3MBTL3 interacting partners, one whose interaction with L3MBTL3 was inhibited by UNC1215 was BCL2-associated transcription factor 1 (BCLAF1), a transcriptional repressor that interacts with members of the BCL2 family of proteins and has been shown to localize inside the nucleus in focal structures during DNA damage repair.[39,40] Using a commercially available BCLAF1 antibody, we determined that BCLAF1 nuclear foci closely co-localize with GFP-3MBT foci (Fig. 6a) and to a lesser extent with GFP-FLMBT foci (Supplementary Fig. 18). To validate this interaction further, immunoprecipitation experiments were performed with flag-L3MBTL3 proteins, confirming a cellular interaction between BCLAF1 and the MBT domains of L3MBTL3 and a somewhat reduced interaction with the full length protein (Fig. 6b). Upon treatment with 1 μM UNC1215, the interaction between flag-3MBT and BCLAF1 was substantially reduced, while the interaction between flag-FLMBT and BCLAF1 was reduced to a lesser extent. Similarly, the full length domain 2 D381A mutant demonstrated reduced binding to BCLAF1 (Supplementary Fig. 19). This strongly suggests that BCLAF1 is a methyl-lysine mediated interaction partner of L3MBTL3.

Indeed, a search of the PhosphoSitePlus database (www.phosphosite.org/) reveals that publicly available mass spectrometric studies have previously identified K445me as a methylation mark on BCLAF1. To confirm this interaction, we tested the binding of a K445me containing BCLAF1 peptide (residues 438-452) to L3MBTL3 by AlphaScreen, demonstrating that the BCLAF1 Kme peptide effectively displaces H4K20me$_2$ and binds L3MBTL3 with a similar affinity to H4K20me histone peptides (Supplementary Fig. 1). Based on these results, we propose that BCLAF1 is a novel methylation-dependent interaction partner of L3MBTL3 whose recognition can be antagonized by UNC1215.

Discussion

The primary qualification criterion for a chemical probe is provision of sufficient data that associates its *in vitro* activity to its *in vivo* (minimally cellular) activity.[16] UNC1215 demonstrates submicromolar *in vitro* and cellular potency against L3MBTL3 Kme binding, is non-toxic to cells at above 100-fold its $EC_{50,}$ and is selective over a broad panel of other Kme binding domains, other chromatin regulators, and targets that might be modulated by its pharmacophoric properties such as ion channels and GPCRs. Our mechanistic studies of L3MBTL3 antagonism by UNC1215 revealed an unusual and intriguing polyvalent mechanism in which both domain 1 and domain 2 are involved in recognition of the probe. This binding mechanism of UNC1215, combined with the effects of point mutants in domain 1 and domain, hints at a potential polyvalent mechanism for substrate recognition by L3MBTL3. Our data supports a model in which the Kme binding function of L3MBTL3 contributes to, but is not the sole determinant of, its nuclear localization. It is highly likely that additional domains such as the SAM domain contribute along with the MBT domains to orchestrate the cellular localization and function of L3MBTL3. Finally, we showed that UNC1215 antagonized the recognition of BCLAF1 by L3MBTL3, a novel target of this Kme reader. Thus, UNC1215 is a potent, selective, cell-active molecule that can be used with confidence to explore the cellular role of the Kme binding function of L3MBTL3. Moreover, UNC1215 is the first chemical probe for a Kme reader domain, demonstrating the chemical tractability of this target class and opening new opportunities for chemical biology and, potentially, drug discovery.

Online Methods

Protein pull-down experiments using compounds

GST fusion proteins were isolated from bacterial lysates using glutathione sepharose beads (GE Healthcare). *In vitro* compound pull-down assays were performed by pre-binding 5 μg of biotinylated UNC1215 to 2 μg of GST fusion protein in 500 μl of 1x binding buffer (50 mM Tris HCl pH 7.5, 150 mM NaCl, 0.1% NP-40, 5 mM EDTA, 5 mM EGTA, 15 mM $MgCl_2$) giving a final compound concentration of 0.01 μg/μl (6.4 μM). Incubation was done overnight rocking at 4 °C. Next, 25 μL of streptavidin agarose beads (Millipore) were pre-washed with binding buffer and incubated with the compound-protein mix for 1 h rocking at 4 °C. After three washes with 500 μl binding buffer, 30 μl of 2x protein loading buffer was added to the beads and boiled. The samples were resolved on a SDS-PAGE and western blots were performed using anti-GST antibody.

Protein pull-down experiments using the H4K20me2 peptide

In vitro peptide pull-down assays were performed as described above by pre-binding 1 μg of biotinylated H4K20me2 peptide (MW 2.416 kDa) with 1 μg of GST fusion protein in 500 μL of 1x binding buffer.

Crystallization

L3MBTL3 protein was crystallized at a concentration of 10 mg/mL using the sitting drop vapor diffusion method at 18 °C. The reservoir solution contained 25% PEG 3350, 0.2 M

NH_4Ac, 0.1 M Bis-Tris buffer at pH 5.5. Using a nylon loop, crystals were passed through reservoir solution additionally containing 20% glycerol, flash-frozen and stored in liquid nitrogen until data collection.[41]

Data collection/structure determination and refinement

A continuous sweep of 180 single-degree oscillation images was collected at beam line 19ID of the Advanced Photon Source at a wavelength of 0.9793 Å. Data were reduced using XDS, [42] POINTLESS,[43] and SCALA.[44] The structure was solved with the program PHASER[45] and a model related to Protein Data Bank (PDB)[46] entry 3UT1 (to be published). Geometry restraints for the inhibitor were prepared with PRODRG.[47] The atomic model was re-built, refined and validated using COOT,[48] REFMAC[49]/autoBUSTER (Buster version 2.10.0. (Global Phasing Ltd, Cambridge, UK, 2011)) and the MOLPROBITY server,[50] respectively. For data and model statistics, refer to Supplementary Table 4.

Immunofluorescence

Images were acquired using a Quorum Spinning Disk Confocal microscope equipped with 405, 491, 561, and 642 nm lasers. Image analysis and foci quantification were performed using Volocity 5.4.1. Foci were quantified using the % GFP intensity function.

For the localization experiments with GFP-3MBT, GFP-FLMBT, 3MBT-GFP, FLMBT-GFP, mCherry-3MBT, mCherry-FLMBT, GFP-FLMBTÄSAM, and the site-directed mutants, 100,000 HEK293 cells/1mL/well were seeded on coverslips in 12 well plates. 16–24 hours later, cells were transfected with the appropriate tagged construct using GeneJuice reagent, according to the manufacturer's instructions, for an additional 48 hours. Prior to imaging, cells were treated with Hoechst 33342 dye for 30 min, washed with PBS, and incubated with the appropriate media.

For foci elimination experiments, the same transfection procedure was followed. In these experiments, cells were treated with the indicated concentrations of inhibitors for at least 24 hours before imaging.

For merocyanine-UNC1215 and GFP-FLMBT co-localization experiments, cells were seeded as described above and transfected with wildtype GFP-FLMBT, D274A GFP-FLMBT, or D381A GFP-FLMBT for 16–24 hours, followed by a 2 hour incubation with 2.5 μM of the merocyanine dye (mero76) or the merocyanine-UNC1215 conjugate/1 mL/well at 37 °C. After incubation with Hoechst 33342 dye and PBS washes, the media was changed to DMEM/F-12 without phenol red (11039) prior to imaging.

U2OS cells were similarly transfected with GFP fusion constructs. Cells were fixed with 2% formaldehyde in PBS for 10 minutes, permeabilized with 0.1% Triton X-100 PBS for 15 minutes, blocked with 5% goat serum, 3% BSA, 0.1% Tween in PBS for 1 hr and incubated with a BCLAF antibody (Abcam) overnight at 4 °C in the blocking buffer. A secondary Alexa488 antibody (Cell Signaling Technology) was used and coverslips were mounted with Fluoroshield DAPI (Sigma).

FRAP assay

Transfected cells were treated with inhibitors for 3–6 hours and analyzed on a Zeiss laser scanning confocal microscope LSM510 META Confocal using 488 nm laser bleaching and subsequent image acquisition every 2 seconds. The data was processed in Volocity 5.4.1 software, normalized to the signal in an unbleached area, and expressed as a percent of the initial signal. The exponential curve fitting and half-life calculations were done using GraphPad Prism4 software. The dissociation constants were determined as reported by Yguerabide et al.[51]

Affinity purification with biotin-UNC1215 conjugate

HEK293 cells were transiently transfected with Flag-FLMBT using GeneJuice reagent according to manufacturer's instructions. Approximately 10 million cells were used per immunoprecipitation. Frozen cell pellets were lysed in lysis buffer containing 50 mM Tris-HCl, pH 8.0, 500 mM NaCl, 2 mM EDTA, 1% Triton X-100, and complete protease inhibitor cocktail (Roche). After vortexing and centrifugation at 13,200 × g, 4 °C, 5 min to pellet the DNA, the supernatant was diluted to 150 mM NaCl (while keeping other reagents' concentrations constant) and was precleared with Streptavidin Mag Sepharose™ magnetic beads (GE Healthcare) for 1 hour at 4 °C. Streptavidin beads were incubated with 5 nmol biotin-UNC1215 (or incubated in the absence of biotin-UNC1215 as a negative control) in 300 μL IP buffer (50 mM Tris-HCl, pH 8.0, 150 mM NaCl, 2 mM EDTA, 1% Triton X-100, and protease inhibitor cocktail (Roche) for 30 min at room temperature. Precleared lysate was added to the biotin-UNC1215 prebound streptavidin beads, in the absence or presence of 5 nmol or 50 nmol cold, untagged UNC1215 overnight at 4 °C. Three washes were performed with IP buffer, followed by elution with buffer containing 100 mM Tris-HCl, pH 8.0, 150 mM NaCl, 2 mM EDTA, 1% Triton, 3% SDS, 5 mM DTT, and 15 mM ME, analysis by SDS-PAGE, and immunoblotting.

Co-immunoprecipitation experiments

HEK293 cells were transfected with Flag or GFP tagged L3MBTL3, treated with inhibitor (1 μM) for 18 hours, and lysed in 50 mM TrisHCl, 450 mM NaCl, 1 mM EDTA, 1% Triton X-100, and protease inhibitors. The lysates were diluted to adjust the NaCl concentration to 200 mM, and then an anti-Flag antibody (Sigma) and protein A Dynabeads (Life Technologies) were added for incubation at 4 °C overnight. The beads were washed 3 times with the lysis buffer, eluted using 100 mM $Na_2H_2PO_4$, 150 mM NaCl, 2 mM EDTA, 5 mM dithiothreitol, 1% Triton X-100, 3% SDS, 1 mM sodium orthovanadate, and 150 mM β-mercaptoethanol. SDS-PAGE separated proteins were transferred to PVDF membrane (Millipore) and immunoblotted using BCLAF1 (BTF1) antibody (Abcam).

Protein interactor screen and mass spectroscopy analysis

5 × 15 cm plates were used for one biological replica, with two biological replicas performed for each set of conditions. HEK293 cells transiently transfected with GFP-3MBT or GFP-FLMBT were lysed in high salt AFC buffer (10 mM Tris-HCL, pH 7.9, 420 mM NaCl, 0.1% NP-40) by freeze-thawing on a dry ice/ethanol mixture and in a 37 °C water bath. Freeze-thawing was repeated three times, after which cells were sonicated on ice (8

times, 0.3 s on, 0.7 s off), incubated with benzonase nuclease (to give a final concentration of 12.5–25 units/mL) for 30 min at 4°C, and centrifuged for 1 hr at 40,000 rpm. The cell lysates were then incubated with an anti-GFP antibody (Abcam) at 4°C overnight, protein G (Sigma) at 4 °C for 4 hours, and washed with low salt AFC buffer (10 mM Tris-HCl, pH 7.9, 100 mM NaCl, 0.1% NP-40), 5 × 1 mL each time. Bound proteins were eluted with 500 mM ammonium hydroxide, dried, resuspended in 50 mM ammonium bicarbonate, trypsin digested, and subjected to MS analysis. X tandem search algorithm was used to identify the peptides. The total spectral counts were normalized to the molecular weight of proteins and the complex was visualized using cytoscape.[52] For cytoscape analysis, only those interactors were included that were not present in the vector alone purification, and had spectral counts equal to or higher than six.

Additional methods

Synthesis of UNC1021, UNC1079, UNC1215, and other related analogs are described in Supplementary Note 1. Additional methods are presented in Supplementary Information, including biochemical assays, selectivity assays, protein expression and purification, protein mutagenesis, and compound toxicity studies. Additional results are also included in Supplementary Information.

Accession codes

Protein Data Bank: coordinates and structure factors for the co-crystal structure of the L3MBTL3-UNC1215 complex have been deposited with accession code 4FL6.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We thank Masoud Vedadi, Greg Wasney, and Fengling Li (SGC Toronto) for support with the PKMT selectivity screening, Olog Fedorov (SGC Oxford) for support with the bromodomain selectivity screening, Anthony Tumber (SGC Oxford) for support with the lysine demethylase selectivity screening, Emily Hull-Ryde (UNC) for support with the CellTiter-Glo cell viability assay, Klaus Hahn (UNC) for providing mero76, Bryan Roth for helpful discussion regarding the GPCR selectivity studies, and Greg Wang (UNC) for providing PHF23 and JARID1 proteins. Results shown in this report are derived from work performed at Argonne National Laboratory, Structural Biology Center at the Advanced Photon Source. Argonne is operated by UChicago Argonne, LLC, for the U.S. Department of Energy, Office of Biological and Environmental Research under contract DE-AC02-06CH11357. The research described here was supported by the National Institute of General Medical Sciences, US National Institutes of Health (NIH, grant RC1GM090732 and R01GM100919), the Carolina Partnership and the University Cancer Research Fund, University of North Carolina at Chapel Hill, the Center for Environmental and Molecular Carcinogenesis at MDACC, the National Institute of Mental Health Psychoactive Drug Screening Program, the Ontario Research Fund (grant ORF-GL2), the Natural Sciences and Engineering Research Council of Canada, the Ontario Ministry of Health and Long Term Care, the American Cancer Society (CJM 119169-PF-10-183-01-TBE), and the Structural Genomics Consortium which is a registered charity (number 1097737) that receives funds from Canadian Institutes of Health Research, Eli Lilly Canada, Genome Canada, GlaxoSmithKline, the Ontario Ministry of Economic Development and Innovation, the Novartis Research Foundation, Pfizer, Abbott, Takeda, and the Wellcome Trust. M.T.B. is supported by an institutional grant NIEHS ES007784 and CPRIT funding (RP110471). C.H.A. holds a Canada Research Chair in Structural Genomics.

References

1. Wang Z, et al. Combinatorial patterns of histone acetylations and methylations in the human genome. Nat Genet. 2008; 40:897–903. [PubMed: 18552846]
2. 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; 14:1025–40. [PubMed: 17984965]
3. Hughes RM, Wiggins KR, Khorasanizadeh S, Waters ML. Recognition of trimethyllysine by a chromodomain is not driven by the hydrophobic effect. Proc Natl Acad Sci U S A. 2007; 104:11184–8. [PubMed: 17581885]
4. Zacharias N, Dougherty DA. Cation-pi interactions in ligand recognition and catalysis. Trends Pharmacol Sci. 2002; 23:281–7. [PubMed: 12084634]
5. Dawson MA, et al. Inhibition of BET recruitment to chromatin as an effective treatment for MLL-fusion leukaemia. Nature. 2011; 478:529–533. [PubMed: 21964340]
6. Delmore, Jake E., et al. BET Bromodomain Inhibition as a Therapeutic Strategy to Target c-Myc. Cell. 2011; 146:904–917. [PubMed: 21889194]
7. Filippakopoulos P, et al. Selective inhibition of BET bromodomains. Nature. 2010; 468:1067–1073. [PubMed: 20871596]
8. Zeng L, et al. Mechanism and regulation of acetylated histone binding by the tandem PHD finger of DPF3b. Nature. 2010; 466:258–262. [PubMed: 20613843]
9. Margueron R, et al. Role of the polycomb protein EED in the propagation of repressive histone marks. Nature. 2009; 461:762–767. [PubMed: 19767730]
10. Adams-Cioaba MA, Min J. Structure and function of histone methylation binding proteins. Biochem Cell Biol. 2009; 87:93–105. [PubMed: 19234526]
11. Kuo AJ, et al. The BAH domain of ORC1 links H4K20me2 to DNA replication licensing and Meier-Gorlin syndrome. Nature. 2012; 484:115–119. [PubMed: 22398447]
12. Li H, et al. Structural basis for lower lysine methylation state-specific readout by MBT repeats of L3MBTL1 and an engineered PHD finger. Mol Cell. 2007; 28:677–91. [PubMed: 18042461]
13. Santiago C, Nguyen K, Schapira M. Druggability of methyl-lysine binding sites. J Comput Aided Mol Des. 2011; 25:1171–8. [PubMed: 22146969]
14. Herold JM, et al. Structure-activity relationships of methyl-lysine reader antagonists. MedChemComm. 2012; 3:45–51.
15. Herold JM, et al. Small-Molecule Ligands of Methyl-Lysine Binding Proteins. Journal of Medicinal Chemistry. 2011; 54:2504–2511. [PubMed: 21417280]
16. Frye SV. The art of the chemical probe. Nat Chem Biol. 2010; 6:159–161. [PubMed: 20154659]
17. Bonasio R, Lecona E, Reinberg D. MBT domain proteins in development and disease. Semin Cell Dev Biol. 2010; 21:221–30. [PubMed: 19778625]
18. Addou-Klouche L, et al. Loss, mutation and deregulation of L3MBTL4 in breast cancers. Mol Cancer. 2010; 9:213. [PubMed: 20698951]
19. Northcott PA, et al. Multiple recurrent genetic events converge on control of histone lysine methylation in medulloblastoma. Nat Genet. 2009; 41:465–472. [PubMed: 19270706]
20. Gurvich N, et al. L3MBTL1 polycomb protein, a candidate tumor suppressor in del(20q12) myeloid disorders, is essential for genome stability. Proc Natl Acad Sci U S A. 2010; 107:22552–7. [PubMed: 21149733]
21. Perna F, et al. Depletion of L3MBTL1 promotes the erythroid differentiation of human hematopoietic progenitor cells: possible role in 20q- polycythemia vera. Blood. 2010 blood-2010-02-270611.
22. Honda H, et al. Hemp, an mbt domain-containing protein, plays essential roles in hematopoietic stem cell function and skeletal formation. Proceedings of the National Academy of Sciences. 2011; 108:2468–2473.
23. Arai S, Miyazaki T. Impaired maturation of myeloid progenitors in mice lacking novel Polycomb group protein MBT-1. EMBO J. 2005; 24:1863–73. [PubMed: 15889154]

24. Nady N, et al. Histone recognition by human malignant brain tumor domains. J Mol Biol. 2012; 423:702–18. [PubMed: 22954662]
25. Bonasio R, Lecona E, Reinberg D. MBT domain proteins in development and disease. Seminars in Cell & Developmental Biology. 2010; 21:221–230. [PubMed: 19778625]
26. Botuyan MV, et al. Structural Basis for the Methylation State-Specific Recognition of Histone H4-K20 by 53BP1 and Crb2 in DNA Repair. Cell. 2006; 127:1361–1373. [PubMed: 17190600]
27. Wigle TJ, et al. Screening for Inhibitors of Low-Affinity Epigenetic Peptide-Protein Interactions: An AlphaScreenTM-Based Assay for Antagonists of Methyl-Lysine Binding Proteins. J Biomol Screen. 2010; 15:62–71. [PubMed: 20008125]
28. Gao C, et al. Biophysical Probes Reveal a “Compromise” Nature of the Methyl-lysine Binding Pocket in L3MBTL1. Journal of the American Chemical Society. 2011; 133:5357–5362. [PubMed: 21428286]
29. Nady N, et al. Recognition of Multivalent Histone States Associated with Heterochromatin by UHRF1 Protein. Journal of Biological Chemistry. 2011; 286:24300–24311. [PubMed: 21489993]
30. Wang WK, et al. Malignant Brain Tumor Repeats: A Three-Leaved Propeller Architecture with Ligand/Peptide Binding Pockets. Structure. 2003; 11:775–789. [PubMed: 12842041]
31. Kim J, et al. Tudor, MBT and chromo domains gauge the degree of lysine methylation. EMBO Rep. 2006; 7:397–403. [PubMed: 16415788]
32. Takada Y, et al. Mammalian Polycomb Scmh1 mediates exclusion of Polycomb complexes from the XY body in the pachytene spermatocytes. Development. 2007; 134:579–590. [PubMed: 17215307]
33. Trojer P, et al. L3MBTL2 Protein Acts in Concert with PcG Protein-Mediated Monoubiquitination of H2A to Establish a Repressive Chromatin Structure. Molecular Cell. 2011; 42:438–450. [PubMed: 21596310]
34. Trojer P, et al. L3MBTL1, a histone-methylation-dependent chromatin lock. Cell. 2007; 129:915–28. [PubMed: 17540172]
35. Ogawa H, Ishiguro K-i, Gaubatz S, Livingston DM, Nakatani Y. A Complex with Chromatin Modifiers That Occupies E2F- and Myc-Responsive Genes in G0 Cells. Science. 2002; 296:1132–1136. [PubMed: 12004135]
36. Nady N, et al. Histone Recognition by Human Malignant Brain Tumor Domains. J Mol Biol. 2012
37. Kim CA, Gingery M, Pilpa RM, Bowie JU. The SAM domain of polyhomeotic forms a helical polymer. Nat Struct Biol. 2002; 9:453–7. [PubMed: 11992127]
38. Knight MJ, Leettola C, Gingery M, Li H, Bowie JU. A human sterile alpha motif domain polymerizome. Protein Sci. 2011; 20:1697–706. [PubMed: 21805519]
39. Haraguchi T, et al. Emerin binding to Btf, a death-promoting transcriptional repressor, is disrupted by a missense mutation that causes Emery–Dreifuss muscular dystrophy. European Journal of Biochemistry. 2004; 271:1035–1045. [PubMed: 15009215]
40. Lee YY, Yu YB, Gunawardena HP, Xie L, Chen X. BCLAF1 is a radiation-induced H2AX-interacting partner involved in gammaH2AX-mediated regulation of apoptosis and DNA repair. Cell Death Dis. 2012; 3:e359. [PubMed: 22833098]
41. Hope H. Cryocrystallography of biological macromolecules: a generally applicable method. Acta Crystallogr B. 1988; 44 (Pt 1):22–6. [PubMed: 3271102]
42. Kabsch W. Xds. Acta Crystallogr D Biol Crystallogr. 2010; 66:125–32. [PubMed: 20124692]
43. Evans PR. An introduction to data reduction: space-group determination, scaling and intensity statistics. Acta Crystallogr D Biol Crystallogr. 2011; 67:282–92. [PubMed: 21460446]
44. Evans P. Scaling and assessment of data quality. Acta Crystallogr D Biol Crystallogr. 2006; 62:72–82. [PubMed: 16369096]
45. McCoy AJ, et al. Phaser crystallographic software. J Appl Crystallogr. 2007; 40:658–674. [PubMed: 19461840]
46. Berman HM, et al. The Protein Data Bank. Nucleic Acids Res. 2000; 28:235–42. [PubMed: 10592235]

47. Schuttelkopf AW, van Aalten DM. PRODRG: a tool for high-throughput crystallography of protein-ligand complexes. Acta Crystallogr D Biol Crystallogr. 2004; 60:1355–63. [PubMed: 15272157]

48. Emsley P, Lohkamp B, Scott WG, Cowtan K. Features and development of Coot. Acta Crystallogr D Biol Crystallogr. 2010; 66:486–501. [PubMed: 20383002]

49. Murshudov GN, et al. REFMAC5 for the refinement of macromolecular crystal structures. Acta Crystallogr D Biol Crystallogr. 2011; 67:355–67. [PubMed: 21460454]

50. Chen VB, et al. MolProbity: all-atom structure validation for macromolecular crystallography. Acta Crystallogr D Biol Crystallogr. 2010; 66:12–21. [PubMed: 20057044]

51. Yguerabide J, Schmidt JA, Yguerabide EE. Lateral mobility in membranes as detected by fluorescence recovery after photobleaching. Biophys J. 1982; 40:69–75. [PubMed: 7139035]

52. Shannon P, et al. Cytoscape: a software environment for integrated models of biomolecular interaction networks. Genome Res. 2003; 13:2498–504. [PubMed: 14597658]

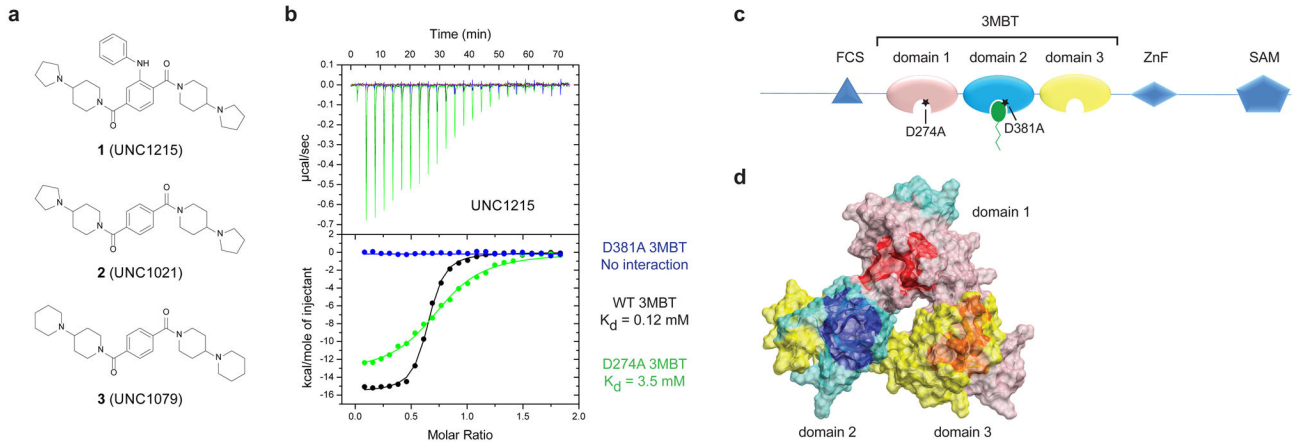

Figure 1. UNC1215 is a potent antagonist of L3MBTL3

(**a**) Structure of UNC1021, a nanomolar antagonist of L3MBTL3 methyl-lysine binding activity. UNC1215 is a more potent L3MBTL3 cellular antagonist, and UNC1079 is a structurally similar but significantly less potent antagonist and negative control. (**b**) ITC analysis of site-directed mutants of L3MBTL3 (3MBT) revealed strong binding of UNC1215 to wild-type protein, but not to domain 2 binding pocket mutant D381A. Domain 1 (D274A) mutant bound to UNC1215 with weaker affinity. (**c**) Domain architecture of the full length L3MBTL3 protein showing a putative FCS nucleic acid binding domain, a Zinc finger (ZnF), and a SAM (sterile alpha motif) domain, in addition to three MBT repeat sequences (pink, blue, and yellow). Mono- or dimethyl-lysine is represented schematically by a green oval. Mutations made in the aromatic cage of the second MBT domain (D381A) and in the aromatic cage of the first MBT domain (D274A), are represented by stars. See Supplementary Figure 21 for a schematic of all protein constructs and tagged fusion proteins used for *in vitro* and cellular studies. (**d**) A surface representation of the of the 3MBT crystal structure (PDB 3UT1), colored as in **c**, but with the residues composing the conserved (domain 2) or semiconserved (domains 1 & 3) aromatic cage of each repeat in red, dark blue, and orange for domain 1, 2 & 3, respectively. Thus, the three presumed binding sites of 3MBT exist in a triangular arrangement, all on the same surface of the molecule.

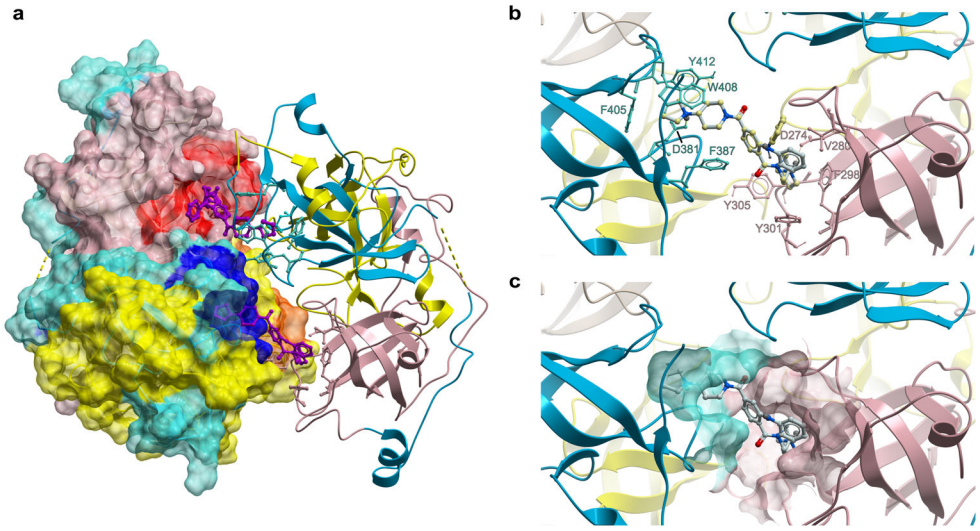

Figure 2. X-ray crystal structure of the UNC1215-3MBT complex

UNC1215 binds in a unique 2:2 binding mode supporting the mechanism of action of UNC1215 (PDB 4FL6). (**a**) Two 3MBT molecules are rotated around a pseudo two-fold axis perpendicular to the plane of the paper. As a consequence, the domain 3 aromatic cages of each protein, domain 3′ and domain 3″, face one another, while domain 1′ faces domain 2″ and domain 2′ faces domain 1″. (**b**) UNC1215 binding is primarily mediated by interaction with the aromatic cage of domain 2, consisting of F362, F380, W383, and Y387, and via a key hydrogen bond between one pyrrolidine nitrogen and D381. The piperidine-pyrrolidine ortho to the aniline ring bridges the two proteins by interacting with domain 1″, forming a salt bridge with D274. A second UNC1215 molecule binds in a reciprocal fashion to domain 1′ and domain 2″ binding pockets. (**c**) Surface representation showing the close association of the two 3MBT subunits and the means by which UNC1215 bridges the aromatic cages of domain 1 (pink) and domain 2 (cyan).

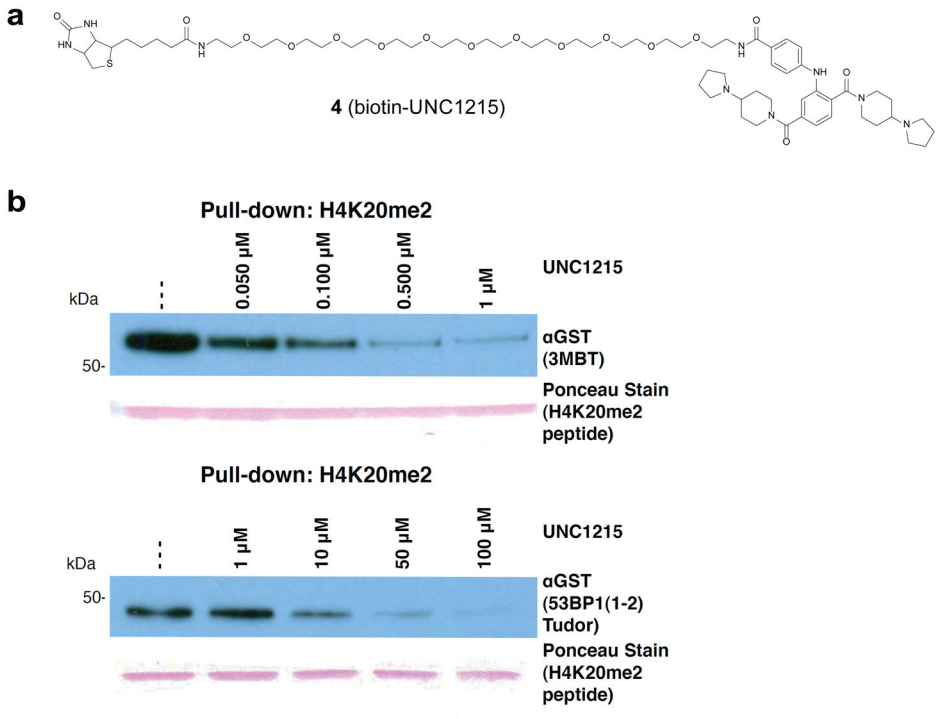

Figure 3. UNC1215 binds a small set of Kme reader proteins with lower affinity than L3MBTL3

(**a**) Structure of a biotinylated analog of UNC1215. (**b**) H4K20me2 pull-down experiments in the presence of increasing concentrations of UNC1215 reveal nanomolar potency for antagonism of 3MBT pull-down, whereas much weaker affinity was observed for 53BP1.

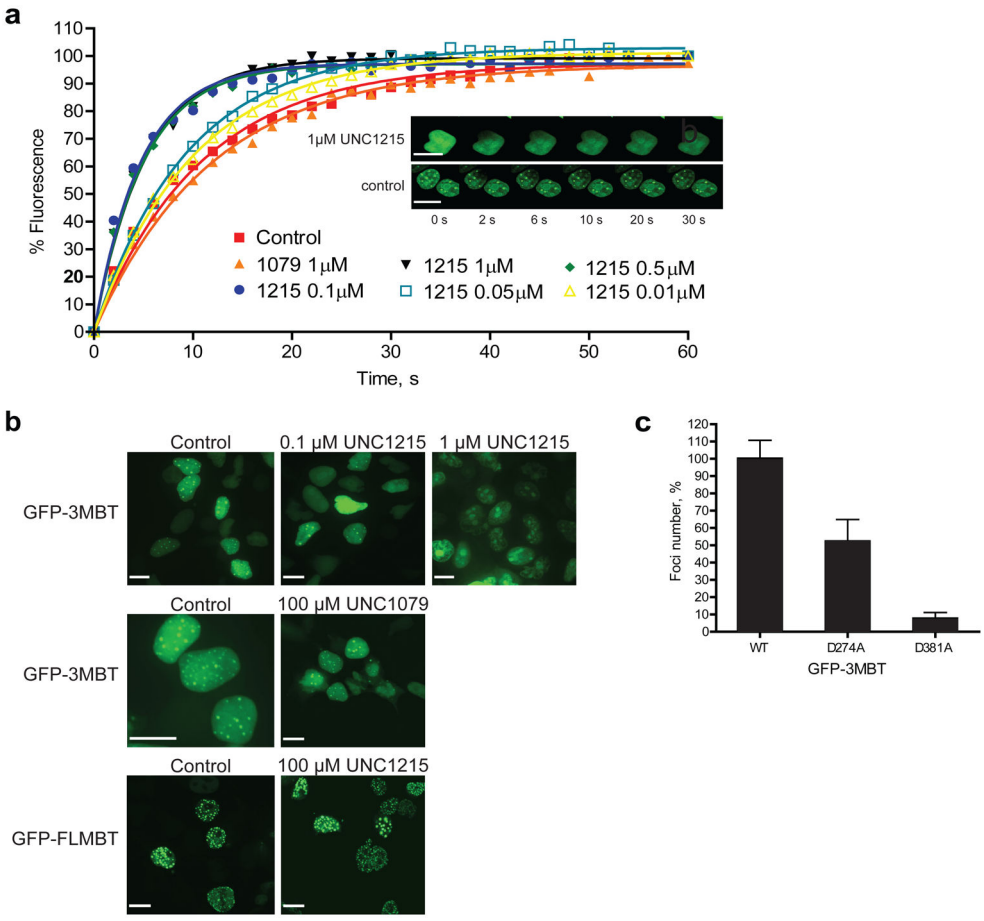

Figure 4. UNC1215 potently antagonizes 3MBT localization in cells

(**a**) Recovery time of a photobleached area in GFP-3MBT expressing cells is reduced upon treatment with UNC1215 in a dose response manner, whereas inactive compound UNC1079 shows no effect. Solid lines represent the exponential fit of the data for 8–10 nuclei with the coefficients of variation ranging from 1–30% for the individual time points. The experiments were performed 3 independent times with similar data resulting. Inset shows time lapse images of photobleached nuclei for UNC1215 (1 μM) and control (water) treatments. (**b**) GFP fusions of 3MBT and FLMBT localize to the nucleus in HEK293 cells. UNC1215 inhibits the foci formation of GFP-3MBT in a dose response fashion, whereas UNC1079 has no effect on the foci. In contrast, UNC1215 is relatively ineffective at inhibiting foci formation of N-terminally tagged GFP-FLMBT (scale bar, 10 μm). (**c**) The GFP-3MBT D274A domain 1 mutant shows a reduction in foci formation, while the GFP-3MBT D381A domain 2 mutant does not form nuclear foci. Error bars in **c** and **d** represent SEM from 15–25 cells.

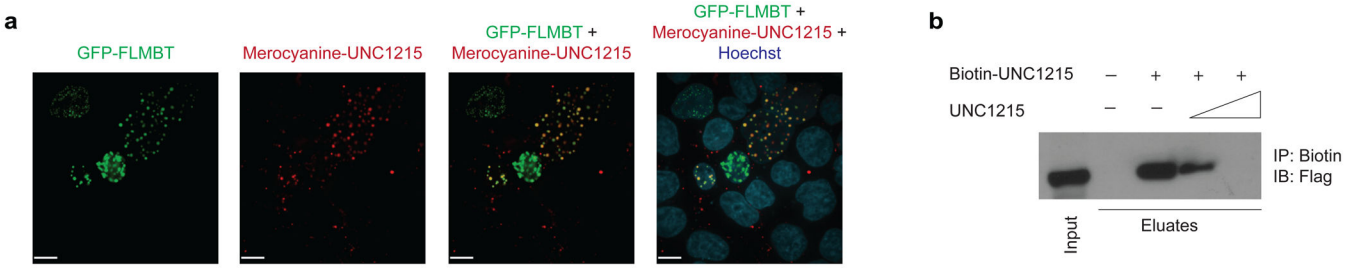

Figure 5. UNC1215 binds and co-localizes with full length L3MBTL3

(**a**) UNC1215 conjugated to the cell-permeable merocyanine dye, mero-76, co-localizes with GFP-FLMBT in HEK293 cells (scale bar, 10 μm; green is GFP-FLMBT, red is merocyanine-UNC1215, and blue is Hoechst dye). (**b**) FLMBT binds to biotin-UNC1215 (5 nmol). The presence of increasing concentrations (1 or 10 equivalents relative to biotin-UNC1215) of untagged UNC1215 results in a decreased amount of bound FLMBT.

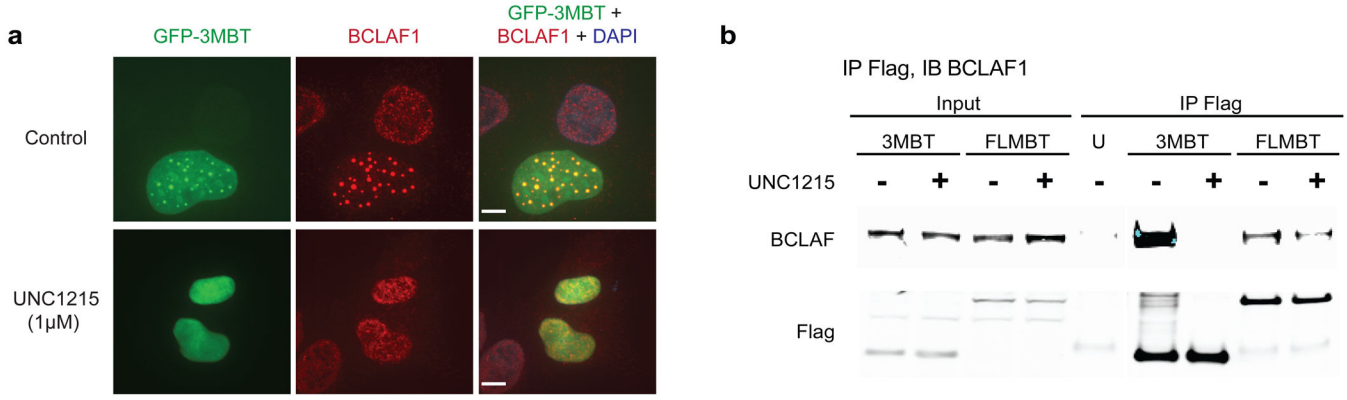

Figure 6. Identification of BCLAF1 as a novel L3MBTL3 protein interactor

(**a**) In U2OS cells, 3MBT colocalizes with BCLAF1 (top panel; green is GFP-3MBT, red is BCLAF, and blue is DAP1). Upon treatment with UNC1215, the 3MBT and BCLAF1 nuclear foci are noticeably disrupted (bottom panel). (**b**) Immunoprecipitation experiments in cells transfected with flag-3MBT or flag-FLMBT show that UNC1215 disrupts the interaction between 3MBT and BCLAF1 and also reduces the interaction between FLMBT and BCLAF1. U denotes untransfected cells.