**Chemistry & Biology**

**Small-Molecule Modulators of Methyl-Lysine Binding for the CBX7 Chromodomain**

**Graphical Abstract**

**Highlights**
- CBX7 represses gene transcription through histone H3 lysine 27 methylation
- Small molecules compete with methylated H3 lysine 27 binding to CBX7 chromodomain
- Chemical inhibition of CBX7 chromodomain derepresses p16/CDKN2A in PC3 prostate cancer cells

**Authors**
Chunyan Ren, Keita Morohashi, ..., Martin Walsh, Ming-Ming Zhou

**Correspondence**
ming-ming.zhou@mssm.edu

**In Brief**
The Polycomb group protein CBX7 functions to regulate gene transcriptional repression via its chromodomain binding to methylated lysine 27 of histone H3. Ren et al. report a small molecule that inhibits the CBX7 chromodomain and derepresses target gene p16/CDKN2A in PC3 prostate cancer cells.

Ren et al., 2015, Chemistry & Biology 22, 161–168
February 19, 2015 ©2015 Elsevier Ltd All rights reserved
http://dx.doi.org/10.1016/j.chembiol.2014.11.021
Small-Molecule Modulators of Methyl-Lysine Binding for the CBX7 Chromodomain

Chunyan Ren, Keita Morohashi, Alexander N. Plotnikov, Jean Jakoncic, Steven G. Smith, Jiaojie Li, Lei Zeng, Yoel Rodriguez, Vivian Stojanoff, Martin Walsh, and Ming-Ming Zhou

1Department of Structural and Chemical Biology, Icahn School of Medicine at Mount Sinai, 1425 Madison Avenue, New York, NY 10029, USA
2Brookhaven National Laboratory, National Synchrotron Light Source, Upton, NY 11973, USA
3Department of Natural Sciences, Hostos Community College of CUNY, Bronx, NY 10451, USA

*Correspondence: ming-ming.zhou@mssm.edu
http://dx.doi.org/10.1016/j.chembiol.2014.11.021

SUMMARY

Chromobox homolog 7 (CBX7) plays an important role in gene transcription in a wide array of cellular processes, ranging from stem cell self-renewal and differentiation to tumor progression. CBX7 functions through its N-terminal chromodomain (ChD), which recognizes trimethylated lysine 27 of histone 3 (H3K27me3), a conserved epigenetic mark that signifies gene transcriptional repression. In this study, we report the discovery of small molecules that inhibit CBX7ChD binding to H3K27me3. Our crystal structures reveal the binding modes of these molecules that compete against H3K27me3 binding through interactions with key residues in the methyl-lysine binding pocket of CBX7ChD. We further show that a lead compound, MS37452, derepresses transcription of Polycomb repressive complex target gene p16/CDKN2A by displacing CBX7 binding to the INK4A/ARF locus in prostate cancer cells. These small molecules have the potential to be developed into high-potency chemical modulators that target CBX7 functions in gene transcription in different disease pathways.

INTRODUCTION

Polycomb group (PcG) proteins were first identified in Drosophila as silencers of transcription of the bithorax gene complex (BX-C), a group of homeotic (Hox) genes responsible for controlling segmentation (Lewis, 1978). PcG proteins also play a fundamental role in multicellular organism development and cancer progression (Sauvageau and Sauvageau, 2010). PcG proteins function in two main complexes: Polycomb repressive complex 1 (PRC1) recognizes H3K27me3 through the N-terminal ChD responsible for recognition of the repressive mark, H3K27me3. The methyl-lysine binding pocket (also called the “aromatic cavity”) is formed by three aromatic residues (Phe11, Trp32, and Trp35 in CBX7), which are highly conserved among all ChD proteins (Bernstein et al., 2006; Kaustov et al., 2011). Mutations of these residues cause CBX7 dissociation from chromatin and alter the life-span of prostate epithelial cells, underscoring the importance of ChD/H3K27me3 binding for CBX7 activity (Gil et al., 2004; Morey et al., 2012; Yap et al., 2010). Thus, we postulated that small-molecule disruption of CBX7ChD binding to H3K27me3 would inhibit the transcriptional activity of CBX7 and result in derepression of its target genes. Such small molecules could be used to fine-tune a balance between stem cell self-renewal and differentiation, and could also be developed into potential new therapeutics for cancer treatment. Despite the functional importance of CBX7 in gene regulation, only recently have macrocyclic calixarenes (Tabet et al., 2013) and peptidomimetics (Simhadri et al., 2014) been shown as CBX7ChD antagonists, but no small-molecule chemical inhibitors have been reported for CBX7ChD or other CBX ChDs. Here, we report the discovery and characterization of small-molecule chemical modulators of the CBX7ChD, and demonstrate that a lead compound, MS37452, induces transcriptional derepression of p16/CDKN2A by disrupting CBX7ChD binding to H3K27me3 at the INK4A/ARF locus in PC3 prostate cancer cells.

RESULTS AND DISCUSSION

A key PRC1 protein, CBX7 was first identified in functional cDNA screening, designed to extend the life-span of normal human prostate epithelial cells by repressing the Ink4a/Arf locus (Gil et al., 2004). This locus encodes key regulators of both retinoblastoma tumor suppressor (Rb) and p53 (Sherr, 2001, 2006); p14ARF/p53- and p16INK4A/Rb-dependent impairment of cell growth upon CBX7 knockdown highlights CBX7’s role in tumorigenesis (Bernard et al., 2005). CBX7 is also important in stem cell self-renewal and differentiation (Klauke et al., 2013; Morey et al., 2012).

Similar to other CBX proteins, CBX7 is characteristic of an N-terminal ChD responsible for recognition of the repressive mark, H3K27me3. The methyl-lysine binding pocket (also called the “aromatic cavity”) is formed by three aromatic residues (Phe11, Trp32, and Trp35 in CBX7), which are highly conserved among all ChD proteins (Bernstein et al., 2006; Kaustov et al., 2011). Mutations of these residues cause CBX7 dissociation from chromatin and alter the life-span of prostate epithelial cells, underscoring the importance of ChD/H3K27me3 binding for CBX7 activity (Gil et al., 2004; Morey et al., 2012; Yap et al., 2010). Thus, we postulated that small-molecule disruption of CBX7ChD binding to H3K27me3 would inhibit the transcriptional activity of CBX7 and result in derepression of its target genes. Such small molecules could be used to fine-tune a balance between stem cell self-renewal and differentiation, and could also be developed into potential new therapeutics for cancer treatment. Despite the functional importance of CBX7 in gene regulation, only recently have macrocyclic calixarenes (Tabet et al., 2013) and peptidomimetics (Simhadri et al., 2014) been shown as CBX7ChD antagonists, but no small-molecule chemical inhibitors have been reported for CBX7ChD or other CBX ChDs. Here, we report the discovery and characterization of small-molecule chemical modulators of the CBX7ChD, and demonstrate that a lead compound, MS37452, induces transcriptional derepression of p16/CDKN2A by disrupting CBX7ChD binding to H3K27me3 at the INK4A/ARF locus in PC3 prostate cancer cells.

While an H3K27me3 peptide is a well-established biological ligand for CBX7ChD (Bernstein et al., 2006; Yap et al., 2010), its binding affinity is modest ($K_d = 27.7 \mu M$) (Figure 1A; Table S1), making it less ideal as a probe for high-throughput chemical screening. To address this issue, we evaluated another trimethylated lysine peptide of SETDB1 at K1170 (SETDB1K1170me3), making it less ideal as a probe for high-throughput chemical screening. To address this issue, we evaluated another trimethylated lysine peptide of SETDB1 at K1170 (SETDB1K1170me3).
Figure 1. Structural Analysis of H3K27me3 and SETDB1 Recognition by CBX7ChD

(A) Measurement of CBX7ChD binding to lysine-methylated H3 and SETDB1 peptides using a fluorescence anisotropy assay. Results were plotted from at least two independent experiments and error bars denote the SEM.
reported to bind to CBX7ChD with higher affinity (Kaustov et al., 2011). Despite similar core sequences of ARKme3S and ALKme3S for H3K27me3 and SETDB1, respectively, the SETDB1 peptide binds to CBX7ChD with much higher affinity ($K_d = 1.3$ μM) (Figure 1A). The enhanced binding of the SETDB1 peptide is supported by 2D $^{1}H$-$^{15}N$-heteronuclear single quantum coherence (HSQC) nuclear magnetic resonance (NMR) spectra (Figure 1B), and also confirmed by a fluorescence anisotropy binding study (see Figure S2J), thus making SETDB1-1170me3 an attractive probe for high-throughput screening.

To determine the molecular basis of CBX7ChD binding to these lysine-methylated peptides, we solved 1.45 Å and 1.60 Å resolution crystal structures of CBX7ChD bound to an H3K27me3 or SETDB1-K1170me3 peptide, respectively (Figures 1C and 1D; Table S2). In the methyl-lysine binding aromatic cavity, the indole of Trp35 recognizes three methyl groups of Kme3, while Phe11 and Trp32 form the base to interact with the acyl side chain of the methylated lysine (Figure 1E; Figures S1A–S1C). The H3K27me3 peptide binds across the protein with residues Ala24–Ala25-Arg26-K27me3 forming a two-strand antiparallel β sheet with the N-terminal Val10-Ala12 of the protein (Figure 1F). The CBX7ChD has an expanded interface for the SETDB1 peptide with a longer N-terminal β strand composed of Gly7–Ala12 (Figures 1D and 1G). This antiparallel β sheet is enforced by an extended network of intermolecular hydrogen-bond interactions involving protein residues Gly7 and Asp50, as well as hydrophobic interactions with Leu53 that are absent in the H3K27me3-bound structure (Figure 1H). These findings explain the molecular basis for the enhanced affinity of CBX7ChD for SETDB1-K1170me3 over H3K27me3.

We next conducted high-throughput screening of a library of 2,560 Food and Drug Administration (FDA)-approved drug molecules and an L1 library of 100,160 compounds selected from commercial sources to identify inhibitors for the CBX7ChD using a fluorescence anisotropy binding assay with a fluorescein isothiocyanate (FITC)-labeled SETDB1-K1170me3 peptide as an assay probe (see Supplemental Information). A B score $< -6$ was set as the criterion to select initial hits, of which 16 were identified from the FDA library and 40 from the L1 library (Figure 2A). Of these initial hits, five hits from the FDA and one hit from the L1 library were confirmed as binders of the CBX7ChD by using 2D $^{1}H$-$^{15}N$ HSQC spectra (Figures S2A–S2H and S3) and the fluorescence anisotropy binding assay (Figure 2B).

MS37452 was the only chemical hit confirmed from the L1 library with binding affinity ($K_d$) of 28.90 ± 2.71 μM to the CBX7ChD, as determined by NMR titration (Figure S2I). MS37452 disrupts CBX7-H3K27me3 or H3K9me3 binding with $K_i$ of 43.0 and 55.3 μM, respectively, as shown in a fluorescence anisotropy binding assay (Figure S2J).

We then solved a 2.14 Å resolution crystal structure of the CBX7ChD-MS37452 complex (Table S2). Strikingly, MS37425 adapts two rotamer conformations. The dimethoxybenzene and piperazine moieties of MS37452 are bound in the same orientation interacting with Phe11, Trp32, and Trp35 in the aromatic cavity, or sandwiched between Phe11 and Trp32, respectively, whereas the methylbenzene moiety adopts a cis or trans conformation with respect to the dimethoxybenzene, hinged at the carbonyl that connects the methylbenzene to the piperazine ring (Figure 2C). Notably, in the trans conformation, the methylbenzene is positioned in the ChD similarly to H3K27me3 and SETDB1-K1170me3, forming interactions with His47, Val13, and Glu43 (Figures 2D–2G). On the other hand, in the cis conformation, the methylbenzene swings upward into the protein structure, disrupting the N-terminal β strand and compromising the stability of the complex (Figure 2D, left). The trans conformation is secured by water-mediated hydrogen bonding of the ligand’s two carbonyl groups with protein residues His5 (from His-tag) and Tyr39 (Figure 2D, right). In either conformation, Trp35 is pushed outward to accommodate the dimethoxybenzene moiety within the aromatic cavity (Figures 2E and 2G). Point mutations of Phe11, Trp35, or Tyr39 to Ala almost completely abolished CBX7ChD-MS37452 binding (Figure S2I). These results confirmed that the aromatic cage residues and Tyr39 are critical to the CBX7ChD-MS37452 interaction.

We next examined the selectivity of MS37452 for different CBXChDs. MS37452 binds to a subgroup of CBXChDs, including those of CBX2, CBX4, CBX6, and CBX8, but shows almost no binding to ChDs of CBX1, CBX3, and CBX5 that are also known as heterochromatin 1 (HP1) β, γ, and α (Figures S1B and S2). As determined by HSQC titration, MS37452 has approximately a 3-fold and at least 10-fold weaker affinity for CBX4 or CBX2/CBX6/CBX8 than for CBX7, respectively (Figure S2I). The difference in MS37452 affinity between these two subgroups is most likely due to Trp35, which is replaced by a Phe in the HP1 ChDs (see Figure S1C). Phe has a smaller side chain than Trp, and is not sufficient to accommodate two methoxy groups when the aromatic cage residues adopt an extended conformation upon MS37452 binding (Figure S2D). The difference between CBX7 and CBX2/CBX4/CBX6/CBX8 is possibly due to the collective role of Tyr39, Val13, and His47 (see Figure S1C). Specifically, Tyr39 is substituted to Asp in CBX2/CBX4/CBX6/CBX8. We also collected structure-activity relationship data for MS37452 with chemical analogs either available in the L1 library.
The analogs that contain Cl, Br, or no methyl group substitutions on the methylbenzene ring bind to the CBX7ChD with reduced affinity in the range of a B score less than $-3$ but greater than $-6$. Analogs lacking one of the two methoxyls (or switching their positions) almost abolishes binding, indicating that the size and position of the two methoxyls are critical for maintaining MS37452’s interaction with Trp35 of the protein. However, we...
did not find a compound with affinity greater than MS37452. Collectively, these results indicate that both the dimethoxybenzene and methylbenzene moieties contribute to overall binding. The former is secured by its interaction with the aromatic cavity residues and Tyr39, whereas the latter is bound in a conformation similar to the methylated lysine, sandwiched between β1 and α1.

Among the hits identified and confirmed from the FDA library (Figure 2A; Figure S3) are sennoside A, suramin, aurin tricarboxylic acid, trypan blue, and Evans blue. They bind to the CBX7ChD with affinities (IC₅₀) of 33.3, 8.1, 5.5, 7.5, and 3.9 μM, respectively (Figure 2B). Unlike MS37452, these compounds share common features: relatively large in molecular weight, symmetric, and containing polyhydroxyl groups. Their binding to CBX7ChD also induces severe line broadening in NMR HSQC spectra, indicating ligand binding-induced conformational exchange of the protein (Figure S3). We extended our analysis on suramin, which is colorless, water soluble, and shown to induce protein dimerization (Lima et al., 2009).

We solved a 1.63 Å resolution crystal structure of the CBX7ChD-suramin complex, which strikingly reveals two suramin molecules bound to two CBX7ChD molecules in a unit cell (Figure 3A; Table S2). One suramin molecule lies along the cleft/tunnel between the two β sheets of each CBX7ChD and glues the two domains together, and is thus referred to as “suramin glue” (Figures 3A and 3B, yellow). The other suramin re- sides along the plane created by the two β sheets, orthogonal to suramin glue (Figures 3A and 3B, salmon), and locks the two CBX7ChD molecules in the region of Lys19 to Gly24, where the β sheets start to turn, and thus is named “suramin lock.” For the tunnel where suramin glue lies, Trp32 and Trp35 from two CBX7ChD molecules form the hydrophobic binding pocket. In contrast, Phe11 either rotates (chain B) or points outside the binding tunnel (chain A), so its aromatic ring does not contribute to binding (Figures 3A–3D). A large number of direct or water-mediated hydrogen bonds are established at an extensive interface between suramin and CBX7ChD (Figure 3C). The urea and four amide groups form hydrogen bonds on the bottom surface of the binding tunnel. The sulfonate groups of the suramin glue interact with a stretch of charged residues Arg20, Glu46, and His47 to facilitate the binding via electrostatic interactions. Suramin-lock only involves interactions with basic residues from chain B, namely Lys19 and Lys23 (Figure 3C, right). In terms of how these crystals are packed, CBX7ChD molecules lie in a 3D network weaved by suramin, and they form a compact 2:2 complex.

Aside from the methyl-lysine-based peptide inhibitors of the CBXChDs, several methyl-lysine-binding antagonists, particularly against the MBT (malignant brain tumor) domain, have been reported (James et al., 2013). MBTs have a narrow methyl-lysine binding pocket and bind preferentially to H3K27me1 or me2, instead of me3 (James et al., 2013). Notably, the aromatic cage of L3MBTL3 comprises four residues, Phe387, Phe405, Trp408, and Tyr412, which correspond to Phe11, Trp32, Trp35, and Tyr39 of CBX7, respectively (Figure S1D). L3MBTL3 also contains a critical Asp381, which together with the aromatic cage residues creates a full pocket to accommodate the amine moiety of UNC1215, a potent MBT inhibitor (James et al., 2013). In contrast, aromatic cage residues and Tyr39 of CBX7ChD provide a half-closed pocket for MS37452 binding. Even considering the fact that other residues contribute to MS37452 binding, the methyl-lysine binding pocket of CBX7-MS37452 is still open. In other words, the amine group of UNC1215 binds deep into the pocket in L3MBTL3 while MS37452 lies along the surface groove of CBX7ChD, thus explaining a major difference in affinity between these two methyl-lysine reader domain inhibitors. In addition, similar to the mode adopted by suramin in binding to CBX7ChD, UNC1215 bridges different domains of L3MBT3 when bound to the protein, thus resulting in a major increase in binding affinity compared with UNC669 or similar compounds with different linkers (James et al., 2013).

Given that suramin interacts with different proteins, including thrombin and pyruvate kinase (Lima et al., 2009; Morgan et al., 2011), we focused our study of small-molecule modulation of CBX7ChD activity on transcriptional repression with MS37452. Prior research showed that mutation or deletion of CBX7ChD results in a decrease in binding of CBX7 to the INK4A/ARF (CDKN2A) locus at multiple sites (Figure 4A) (Bracken et al., 2007). We indeed observed that human PC3 prostate cancer cells treated with MS37452 (250 μM) for 2 hr showed reduced CBX7 occupancy across the INK4A/ARF locus, as determined by chromatin immunoprecipitation (ChIP) (Figure 4B). We further examined INK4A/ARF transcript levels in PC3 cells using quantitative PCR upon dose- and time-dependent treatment of MS37452 (Figures 4C and 4D). After treatment with MS37452 for 12 hr, transcription increased about 25% and 60% for 250 μM and 500 μM of MS37452, respectively, compared with the DMSO control.

We further observed reduced occupancy of RING1B, and H3K27me3 at the INK4A/ARF locus by MS37452, but much less, if any, change of CBX4 or CBX8 (Figure 4B). MS37452 has much weaker affinity to CBX4ChD or ChX8ChD than to CBX7ChD, consistent with weaker reduction of the former association with the INK4A/ARF locus. As a PRC1 component, RING1B has been shown to interact with CBX7 and play a critical role in CBX7 deposition at chromatin (Moresy et al., 2013; Wang et al., 2010). Therefore, a reduced RING1B association is consistent with reduced CBX7 occupancy after the compound treatment. The observed reduction of H3K27me3 matches the findings of a previous study showing localized reduction in H3K27me3 signal when CBX7 was silenced (Maertens et al., 2009) without a change of global H3K27me3 level (Figure S4B). Based on these ChIP and mRNA transcription results, we concluded that MS37452 can disrupt CBX7 binding to the INK4A/ARF locus, subsequently resulting in transcriptional derepression of its target genes p14/ARF and p16/INK4a in PC3 cells.

**SIGNIFICANCE**

In this study, we discovered small molecules that inhibit CBX7ChD binding to lysine-methylated peptides with affinity better than its biological target, H3K27me3. One hit, suramin, bridges two CBX7ChD molecules from two perpendicular directions, and another hit, MS37452, directly competes against H3K27me3 binding by occupying both the methyl-lysine binding aromatic cage and an internal portion of the protein. We further demonstrated that MS37452 blocks CBX7 binding to its target INK4A/ARF gene locus, and
Figure 3. Structural Analysis of Suramin Binding to CBX7ChD

(A) Crystal structure of a CBX7ChD dimer bound to suramin glue (yellow) and suramin lock (salmon) molecules.

(B) Surface presentation of CBX7ChD bound to suramin.

(C) Detailed analysis of the interactions between CBX7ChD and suramin.

(D) Structural comparison of CBX7ChD bound to suramin or H3K27me3 peptide. Chain A and suramin glue from the CBX7ChD-suramin complex and chain AC from the CBX7ChD-H3K27me3 complex were aligned.

See also Figure S3 and Table S2.
Figure 4. Modulation of CBX7 Binding to INK4a/ARF Locus by MS37452

(A) Scheme illustrating the genomic organization of the INK4a/ARF/INK4b gene loci color-coded in pink, blue, and green for p16/INK4a, p14/ARF, and p15/INK4b, respectively. Roman numerals and arrows (red) denote sites of quantitative ChIP and PCR primers, respectively, as previously reported (Bracken et al., 2007).

(B) MS37452 reduces CBX7 occupancy at different positions along the INK4a/ARF locus in PC3 cells after 2 hr treatment of MS37452 (250 μM). Plot represents data from one independent experiment (as an example) and error bars denote the SD of technical repeat. Independent ChIP data are listed in Figure S4.

(C and D) MS37452 induces gene transcriptional derepression at the INK4a/ARF locus, in a concentration- and time-dependent manner, in PC3 cells. mRNA levels of p14/ARF and p16/INK4a were tested using SYBR quantitative PCR after MS37452 treatment. Data were plotted from at least three independent experiments, and error bars denote the SEM. *p < 0.05, **p < 0.01.

See also Figure S4 and Table S4.
induces derepression of p14/ARF and p16/INK4a in human PC3 cells. Collectively, our structure-guided discovery and characterization of small molecules for the CBX7ChD demonstrate the promise and feasibility of developing selective chemical inhibitors of this important epigenome reader domain protein. Given the fundamental role of CBX7 in PRC-directed transcriptional repression of the Ink4a/Arf locus in biology and disease, such new inhibitors can potentially be used to provide senescence control, cancer prevention, and stem cell lineage specification.

EXPERIMENTAL PROCEDURES

Procedures for protein expression and purification, biochemical binding assay, and high-throughput chemical screening, protein structural analysis using NMR and X-ray crystallography, and cellular analysis of effects of compound treatment in gene transcription are reported in detail in Supplemental Experimental Methods.

SUPPLEMENTAL INFORMATION

Supplemental Information includes detailed experimental methods, four figures and four tables and can be found with this article online at http://dx.doi.org/10.1016/j.chembiol.2014.11.021.

AUTHOR CONTRIBUTIONS

C.R. and M.-M. Z. conceived and designed the experiments for the study. C.R. and K.M. performed the high-throughput chemical screening. J.L. and S.G.S. performed the chemical synthesis. C.R. and A.N.P. determined the crystal structures with assistance from J.J. and V.S. C.R. and I.Z. conducted the biochemical study. C.R. carried out the molecular and cell biology experiments with M.W.’s assistance. C.R. and M.-M.Z. wrote the manuscript.

ACKNOWLEDGMENTS

We thank X6A Workbench 2012 and RapiData 2013 for the X-ray crystallography training both theoretically and experimentally, and beamline members for technical help in data collection at National Synchrotron Light Source (NSLS) of Brookhaven National Laboratory. We thank G. Zhang, Q. Zhang, K.-L. Cheung, G. Lu, L. Peng, S.D. Li, R. Sharma, and Y.F. Sun for technical discussion and advice. The X6A beamline at the Brookhaven National Laboratory is supported by the National Institute of General Medical Sciences (GM08080) and the NSLS is funded by DOE (DE-AC02-98CH10886). This work is supported in part by the research grants from NIH to M.-M.Z.

REFERENCES


