Polycomb paralogs chromodomain inhibitors active against both CBX6 and CBX8

Natalia Milosevich,^{[a],[†]} Chelsea R. Wilson,^{[a],[†]} Tyler M. Brown,^[a] Aktan Alpsoy,^[b] Sijie Wang,^[b] Katelyn E. Connelly,^[b] Kirsten A. D. Sinclair,^[a] Felino R. Ponio,^[a] Rebecca Hof,^[a] Emily C. Dykhuizen,^[b] and Fraser Hof^{*[a]}

In memory of François Diederich: enthusiastic supporter, motivator, and mentor

[a] N. Milosevich, C.R. Wilson, T.M. Brown, K.A.D Sinclair, F.R. Ponio, R. Hof, Prof. F. Hof
Department of Chemistry
University of Victoria
Victoria, BC, V8W 3V6, Canada
fhof@uvic.ca

[b] A. Alpsoy, S. Wang, K.E. Connelly, Prof. E.C. Dykhuizen
Department of Medicinal Chemistry and Molecular Pharmacology
College of Pharmacy, Purdue University and Purdue University Center for Cancer Research
575 Stadium Mall Drive, West Lafayette, Indiana 47906, United States

[†] Equal Contribution

Supporting information for this article is given via a link at the end of the document

Abstract: Methyllysine reader proteins bind to methylated lysine residues and alter gene transcription by changing the compaction state of chromatin or by the recruitment of other multiprotein complexes. The polycomb paralogs family of methyllysine readers bind to trimethylated lysine on the tail of histone 3 via a highly conserved aromatic cage located in their chromodomains. Each of the polycomb paralogs are implicated in several disease states. CBX6 and CBX8 are members of the polycomb paralogs family with two structurally similar chromodomains. By exploring the structure-activity relationships of a previously reported CBX6 inhibitor we have discovered more potent and cell permeable analogs. Our current report includes potent, dual-selective inhibitors of CBX6 and CBX8. We have shown that the –2 position in our scaffold is an important residue for selectivity amongst the polycomb paralogs. Preliminary cell-based studies show that the new inhibitors impact cell proliferation in a rhabdoid tumor cell line.

Introduction

Post-translational modifications (PTMs) can be used as epigenetic signatures that control and regulate gene expression pathways. Lysine methylation is one such PTM that is important in several malignancies.^[1] Lysine methylation on histones, which package and order DNA, can activate or repress transcription of the associated genes. Methyllysine ‘reader’ proteins that recognize this mark form multi-protein complexes that regulate fundamental pathways critical for health and development.^[2]

Chromodomains (chromatin organization modifier domains) are a family of methyllysine readers that include proteins belonging to the Heterochromatin protein-1 (HP1) and polycomb sub-families. Both families recognize higher methylation states on H3K9 and H3K27.^[3] The polycomb family of methyllysine readers consists of five human paralogs: CBX2, CBX4, CBX6, CBX7, and CBX8. The human CBX polycombs are important regulatory proteins that are responsible for cellular differentiation during development, stem cell maintenance via transcriptional repression, and cancer progression.^[4] Each member can participate in different versions of the multi-protein complex called polycomb repressive complex 1 (PRC1). Each is functionally unique^[4b, 5], and while the biological data are growing rapidly the

phenotypes of chemical inhibition are only recently beginning to be discovered.

The challenge in targeting the polycomb paralogs with selective chemical inhibitors is due to the high degree of sequence similarity within the chromodomains. The polycomb paralogs chromodomains all possess a conserved aromatic cage made up of 3–4 aromatic residues that recognize the side chain of Kme3. They achieve selectivity by recognizing residues adjacent to the Kme3 in a “β-groove” that engages the Kme3-containing partner in a beta-sheet like structure. The main differences between polycomb chromodomains are relatively small variations in the side chains that line the β-groove.

CBX7-selective inhibitors are by now well known,^[6] but other CBX’s are also arising as promising new drug targets. CBX6 and CBX8 are both involved in regulating stem cell differentiation. CBX6 is an essential regulator of embryonic stem cells (ESCs).^[7] Depletion of CBX6 in ESCs leads to rapid cell differentiation. Cell-based studies with CBX6 mutants lacking the two methyllysine-binding tryptophan residues in the aromatic cage showed that the function of CBX6 in ESCs is strongly linked to the methyllysine reading function of the chromodomain.^[7] Unlike CBX7 (and apparently all other human polycomb paralogs), CBX6 does not act through binding H3K27me3.^[3, 8] Ablation of the H3K27 methyl-depositing PRC2 complex in ESCs did not change CBX6 recruitment to chromatin. CBX8 plays a critical role in dictating cell fate decisions and activation of differentiation genes.^[9] CBX8 acts as part of the PRC1 complex and is upregulated during ESC differentiation. CBX7 in ESCs acts to maintain pluripotency and is downregulated upon increased expression of CBX8 during differentiation.^[4b, 5b, 9]

Both CBX6 and CBX8 have arisen as potential targets in oncology. Overexpression of CBX6 promotes tumor growth and is linked to poor patient outcomes in hepatocellular carcinoma (HCC).^[10] *In vitro* studies in multiple HCC cell lines showed increased CBX6 expression to promote cell growth whereas knockdown of CBX6 showed reductions in proliferation. *In vivo* studies also showed CBX6 overexpression to promote HCC tumour growth.^[10] CBX6 is also upregulated in rhabdoid tumours, an aggressive pediatric cancer.^[11] The oncogenic role of CBX8 has been demonstrated in several malignancies including hepatocellular carcinoma (HCC),^[12] acute myeloid leukemia (AML),^[13] breast cancer^[14] and esophageal carcinomas.^[15] CBX8 is also overexpressed in bladder cancer,^[16] choriocarcinomas,^[17] and glioblastoma multiforme.^[18]

Isolated examples of selective inhibition of CBX6 and CBX8 have recently been reported. A peptidic CBX6 inhibitor (**3**) developed by structure-based design achieved selectivity over the other polycomb paralogs due to installation of a bulky group directed into protein's (−2) pocket, which is larger in CBX6 than other polycombs.^[19] In separate work, a DNA-encoded library showed that CBX8-selective inhibition was achievable by combining large groups in the (−2) pocket with other changes in the ligand structure along the β-groove.^[20] Similarly a CBX8-selective positive allosteric modulator was found by incorporating larger groups in the (−2) pocket along with other structural changes.^[21]

In this work, we thoroughly define the structure-activity relationships (SAR) of CBX6 inhibition, and also report a second generation of CBX6 inhibitors with improved cell permeability. As we modified ligands in the β-groove region, we arrived at a novel class of dual-active CBX6- and CBX8-inhibiting ligands and identified the structural features that give rise to this dual selectivity. We provide preliminary biological data that demonstrate that the inhibitors have functional effects in a cancer cell line.

Results

We started our efforts with modifications to previously reported CBX6-targeting peptidic inhibitors,^[19, 22] and chose to screen each compound against CBX1/6/7/8. This panel was chosen in order to efficiently obtain data on CBX6- and CBX8-selectivity over CBX7, which is the CBX that intrinsically binds ligands most strongly. It also includes CBX1 (HP1β) as a representative of the HP1 class of chromodomains. Our previous report on CBX6 inhibition included the initial scaffold **1**, the promiscuous inhibitor **2**, CBX6-selective inhibitor **3** that is 6–90-fold selective for CBX6 over CBX1, **7** and **8**, **4** the dye-free analog of **2**; and compound **5** the dye-free analog of **3** (Figure 1).^[19, 22] As a first step towards making lower molecular weight compounds, we investigated **1**, which is an analog of **2** that lacks the C-terminal Glu residue and dye label (Figure 1, Table 1). Compound **1** is 10-fold selective for CBX7 over CBX1 and 6- and 29-fold selective over CBX6 and CBX8, respectively.

To improve the cell-permeability and drug-like properties of our inhibitors we first focused on understanding the roles of the dye and the (−2) substituent in CBX6/8 selectivity (Table 1). Substituting the alanine in compound **4** to a valine in compound **5** produced a >100-fold decrease in binding to CBX7. Removal of the FITC-labeled lysine on the scaffold containing the (−2) valine residue (**2** compared to **4**) resulted in a 20-, 40- and 3-fold decrease in binding to CBX6, CBX7 and CBX8, respectively.^[19, 22] We knew from previous reports that removal of the FITC dye label makes these compounds significantly less potent. From the current experiments we can also see that the FITC removal affects CBX6 and CBX8 binding differently. Although **5** is still selective for CBX6 over CBX7, it is no longer selective for CBX6 over CBX8. Compound **5** is equipotent against CBX6/8 (IC₅₀ values of ~17 μM) and is 12-fold selective over CBX7. Compound **6**, in which an isoleucine residue is present at the (−2) position, displayed no measurable binding to CBX1 or CBX7 (>500 μM). Compound **6** displayed similar potency for CBX6 and CBX8 (IC₅₀ values of 9 and 14 μM).

Figure 1. Chemical structures of compounds **1-6**.

Table 1. IC₅₀ values for compounds **1**, **4**, **5** and **6** (μM).^[a]

Compound	CBX1	CBX6	CBX7	CBX8
1	14 ^[b]	8 ^[b]	1.4 ^[b]	41 ^[b]
4	N.T. ^[c]	5	1.8	28
5	N.T. ^[c]	17	210 ^[b]	18 ^[b]
6	>500 ^[b]	9	>500	14

[a] Competitive FP assays performed in triplicate unless stated otherwise. Raw data and asymmetric 95% confidence intervals are reported in SI.

[b] Competitive FP assays performed in duplicate

[c] Not tested

We focused next on optimization of compound **5** by removing the quaternary ammonium ion (Kme3) and the Glu residue, which we hypothesized would each cause poor cell permeability. We removed the C-terminal glutamic acid residue to reduce overall charge, leading to compound **7**, and made an analogous compound **8** that swapped trimethyllysine (Kme3) for diethyllysine (Ket2) (Figure 2).^[6f] Removal of the C-terminal Glu gave decreased binding to CBX7 and no significant change in potency was observed for CBX6/8 (Table 2). Replacement of Kme3 with Ket2 gave a 7-fold increase in binding to CBX6, and no change in binding to CBX8. Compound **8** is 6-fold selective for CBX6 over CBX8 and displays no binding to CBX1/7. We next sought to substitute the (−1) residue in **8** with a more hydrophobic leucine residue (**9**, Figure 3). Compound **9** showed a slight decrease in potency for CBX6 and a 4-fold increase in binding to CBX8 (Table 3). Neither compound showed activity against CBX1 or CBX7.

Figure 2. Structure of compounds **7** and Ket2 analog **8**

Table 2. Binding affinity for compounds **7** and **8** (IC₅₀ values in μ M).^[a]

Compound	CBX1	CBX6	CBX7	CBX8
7	> 500	24	> 250	18
8	> 500	3.2	> 500	18

[a] Competitive FP assays performed in triplicate unless stated otherwise. Raw data and asymmetric 95% confidence intervals are reported in SI.

We next explored the roles of N-terminal capping (N-cap) residues on CBX6/8 selectivity, following leads from our collaborative studies of CBX8-selective inhibitors generated by the Krusemark lab.^[20] Using the scaffold with leucine at the (−1) position (**9**), we substituted the N-cap residue with two distinct heterocycles. Substitution of the *p*-bromobenzamide group with 1*H*-pyrazole-3-carboxamide (**10**) resulted in a 10-fold decrease in binding to CBX6 and a 15-fold decrease in binding to CBX8, reminiscent of results previously seen with a DEL (Table 3).^[20a] Substitution with the 5-methylisoxazole-3-carboxamide (**11**) gave a similar binding profile as **9** with CBX6/8 and showed activity against CBX7 with an IC₅₀ value of 105 μ M (Table 3). A similar profile was observed for a 5-methylisoxazole N-cap ligand with a tyrosine residue at the (−1) position (**12**).

Figure 3. Chemical structures of compounds **8–12** with (−1) and N-cap substitutions.

Table 3. Binding affinity for compounds **8–12** (IC₅₀ values in μ M).^[a]

	R (N)	R (−3)	CBX1	CBX6	CBX7	CBX8
8			> 500	3.2	> 500	18
9			> 300	6.6 ^[b]	> 250 ^[b]	4.5

10			> 300	71	162	68
11			> 300	12	105	7
12			> 500	16	>300	6

[a] Competitive FP assays performed in triplicate unless stated otherwise. Raw data and asymmetric 95% confidence intervals are reported in SI.

[b] Competitive FP assays performed in duplicate

To explore the SAR in the β -groove we combined the N-cap heterocyclic substitutions with aromatic and hydrophobic rings in the (−3) position (Figure 4, Table 4). Leucine was held constant as the (−1) residue. Replacement of the phenylalanine residue in **10** with a cyclohexyl group (**13**) showed a slight 2–3-fold increase in binding to CBX6, **7** and **8**. The same substitution in the 5-methylisoxazole-3-carboxamide containing **11** to give compound **14** did not significantly change binding to CBX6 and CBX8 but did show a decrease in binding to CBX7 (Table 4).

Modifying the aromatic residues at the (−3) position while holding the N-cap methylisoxazole constant showed a decrease in binding to CBX7. Replacement of the phenyl moiety (**11**) with a 4-methoxyphenyl group (**15**) resulted in no significant change in binding to CBX6. CBX7 binding was decrease 2-fold while CBX8 was increase 12-fold. Substituting the methoxy for fluorine (**16**) and iodine (**17**) abolished binding to CBX7. The *p*-fluorophenyl (**16**) analog showed no significant change in CBX6 and 3-fold decrease in CBX8 binding, while the *p*-iodophenyl (**17**) showed a 6- and 33- fold decrease in binding to CBX6 and CBX8. Exchanging the phenyl moiety (**11**) with a pyridine-3-yl group at the (−3) position decreased binding to CBX6/7/8 by a factor of 3, 4 and 2 respectively (**18**, Table 4). Replacing the pyridine-3-yl with a thiophene (**19**) showed a 7-, 2- and 10- fold increase to CBX6/7/8 binding.

Figure 4. Overall chemical structures of compounds **10**, **11** and **13–19** with (−3) and N-cap substitutions. See Table 4 for compound numbers

Table 4. Binding affinity for compounds **10**, **11** and **13–19** (IC₅₀ values in μ M).^[a]

	R (N)	R (−3)	CBX1	CBX6	CBX7	CBX8
10			> 300	71	162	68
13			> 300	32	104	25
11			> 300	12	105	7
14			> 300	8	171	6
15			> 500	12	175	0.6 ^[b]
16			N.T. ^[c]	9.3	>500	1.8

17		N.T. [c]	68	>500	20
18		> 500	33	400	17
19		N.T. [c]	5.0	209	1.6

[a] Competitive FP assays performed in triplicate unless stated otherwise. Raw data and asymmetric 95% confidence intervals are reported in SI.

[b] Estimated value because IC_{50} is outside of linear range of assay.^[22-23]

[c] Not tested

Substitution of the methylisoxazole N-cap gave large improvements in potency. We synthesized compounds **20** and **21**, containing a propyl substitution in the 5-position of the isoxazole ring (Figure 5, Table 5). The 5-propylisoxazole N-cap combined with the (-3) cyclohexyl moiety (**20**) displayed potent binding to CBX6 with an IC_{50} value of 2.4 μ M and is 17- and 3-fold selectivity over CBX7 and CBX8. 5-propylisoxazole (R_N) with a phenylalanine (-3) and threonine (-1) lead to a potent CBX8 inhibitor (**21**) with an IC_{50} of 2.2 μ M. This inhibitor was 3- and 120-fold selective for CBX8 over CBX6 and CBX7. The 5-phenylisoxazole N-cap substitution in compounds **22** and **23** lead to the most potent inhibitors. The addition of a 5-phenylisoxazole N-cap substitution with a (-3) phenyl group provided **22**, which is the most potent compound of the series (Table 5). Compound **22** is equipotent for CBX6 and CBX8 with IC_{50} values of 200 nM and is 18-fold selective over CBX7. We then swapped the (-1) leucine for a tyrosine to obtain compound **23**. Compound **23** was one of the more potent compounds however it had no significant selectivity amongst CBX6/7/8. Substitution of the methyl group in the 5-methylisoxazole N-cap with larger alkyl chains and aromatic groups improves potency due to increased engagement with the extended β -groove of CBX6/7/8.

Figure 5. Chemical structure of peptidic ligands **20-23** with isoxazoles, (-1) and (-3) substitutions.

Table 5. Binding affinity for compounds **20-23**. (IC_{50} values in μ M).^[a]

Compound	CBX1	CBX6	CBX7	CBX8
20	> 300	2.4	40	7.4
21	N.T. [b]	7.5	268	2.2
22	> 500	0.2	3.6	0.2 [c]
23	N.T. [b]	2.2	5.4	1.0

[a] Competitive FP assays performed in triplicate unless stated otherwise. Raw data and asymmetric 95% confidence intervals are reported in SI.

[b] Not tested

[c] Estimated value because IC_{50} is outside of linear range of assay.^[22-23]

Larger hydrophobic substitutions in the (-2) position of the ligand allow for CBX8 selective inhibitors. We synthesized analogs of compound **11** containing a cyclopropyl, cyclopentyl and sec-butyl group at the (-2) position (Figure 6). Compound **24** containing a cyclopropyl group was nearly equipotent for CBX6/8 and displayed between 6- and 7-fold selectivity over CBX7 (Table 6). Compared to the isopropyl-containing compound (**11**), the cyclopropyl substitution increased binding to CBX6/7/8 by factors of 2–5. Replacing the cyclopropyl with a cyclopentyl group (**25**) dramatically reduced binding to CBX6/7 and decreased binding to CBX8 by a factor of 6 (IC_{50} of 18 μ M, Table 6). Compound **25** is 10x and 20x selective for CBX8 over CBX6 and CBX7 respectively. Addition of a sec-butyl group (**26**) further decreased binding to CBX6 and abolished binding to CBX7 at the concentrations tested. All CBX polycomb paralogs tested showed decreased binding with increasing size of the substituent at the (-2) position.

Figure 6. Chemical structure of peptidic ligands **11, 24-26** with substitutions at the (-2) position.

Table 6. Binding affinity for compounds **11, 24-26**. (IC_{50} values in μ M).^[a]

Compound	R (-2)	CBX1	CBX6	CBX7	CBX8
11		> 300	12	105	7
24		> 500	3.6	22	3.0
25		> 500	178	372	18
26		> 500	250	> 500	53

[a] Competitive FP assays performed in triplicate unless stated otherwise. Raw data and asymmetric 95% confidence intervals are reported in SI.

CBX8 was best able to tolerate larger substitutions directed into the (-2) pocket of the protein. We and others have reported CBX8-selective ligands that engage the large size (-2) pocket within CBX8 compared to other members of the polycomb paralog CBX family.^[20-21] The SAR reported demonstrates the (-2) pocket of CBX8 to be capable of engaging and binding to ligands with bulkier (-2) substitutions. The residues lining the (-2) pocket are nearly identical in CBX6/8 with the exception of an isoleucine residue in CBX6 that is replaced with a leucine in CBX8. Compound **25**, containing a cyclopentyl group at the (-2) position displays promising selectivity for CBX8. We predict that

combining the (–2) cyclopentyl group with other substitutions could improve the potency and selectivity of **25** for CBX8.

Figure 7. Inhibitors are cell permeable and able to engage CBX proteins. A) G401 rhabdoid cells are affected by treatment of **8**, **11**, **22**, **23** and show decreased cell proliferation. Cell counts were analyzed using CellTiter-Glo following 72-hour treatment. B) Treatment with compound **23** with wild type G401 decreased cell proliferation while no effect was observed in CBX6 KO. P value calculated using students t-test, n = 6. *** p<.001

CBX6 is upregulated in rhabdoid tumours,^[11] a highly aggressive pediatric malignancy driven by loss of *SMARCB1* subunit of the SWI/SNF chromatin remodeling complex.^[11, 24] We used immunoblot analysis to profile CBX6 expression in a panel of kidney cell lines, including non-transformed epithelial cells (HK-2), clear cell renal cancers (Caki1, Caki2, 786O), kidney adrenal cancer (SW-13) and kidney rhabdoid tumor cell line G401, as well as muscle rhabdoid cell line A204. There was no detectable CBX6 protein in the majority of the cell lines, while the kidney rhabdoid line G401 displayed a high expression level of CBX6, consistent with patient transcriptional data (Figure S47a). We used CRISPR-Cas9 to knock out CBX6 in the G401 cell line (Figure S47b) and determined that loss of CBX6 significantly reduces the growth rate (Figure S47c). To investigate whether CBX chromodomain inhibition can recapitulate the effects of CBX6 knockout in rhabdoid cells, we measured cell proliferation of G401 rhabdoid cells following treatment with **8**, **11**, **22** and **23**. Wild-type G401 cells were treated with the inhibitors for 72 hours and cell counts were measured using CellTiter-Glo (Figure 7a). While all four inhibitors displayed significant toxicity at 100 μM, only compound **23** reduced growth to a similar degree as CBX6 knockout and was active at 50 μM. To define whether compound **23** is targeting CBX6 function, we treated wild-type and CBX6 KO G401 cells (Figure 7b) with 50 μM compound **23**. Compound **23** decreased cell proliferation in wild-type but does not affect proliferation in KO cells, and this provide indirect evidence that it is targeting CBX6. However, inhibition of other CBX chromodomains, particularly with 100 μM treatment, cannot be excluded. Further experiments are needed to determine whether CBX7 and CBX8 play a role in rhabdoid tumor viability, and whether these ligands might also be inhibiting them.

Discussion

These studies show the critical role the (–2) position of CBX ligands for selectivity within the polycomb paralogs family. The ability of ligands with larger groups in the (–2) position to bind the CBX proteins depends on the peptide scaffold and neighboring residues. Previous work demonstrated dye-labeled ligands with a valine residue in the (–2) position to be selective for CBX6. Removal of the dye from these scaffolds showed reduced selectivity for CBX6 over CBX8, however CBX6 was still better

able to tolerate ligands containing bulkier groups within the (–2) pocket (compound **3**, **6**, Table 1). Larger groups in the (–2) position within the scaffolds reported demonstrate selectivity for CBX8. The aromatic cage, hydrophobic clasp and (–2) pocket are tightly connected and residues in the ligand to the right and left of the Ket2 can change the shape and size of the (–2) pocket. Despite the highly similar sequence homology of the residues in the (–2) pocket in CBX6/8, our results show a clear difference in the binding effects within this pocket depending on the chemical structure of the ligand and resulting protein-ligand interactions.

Aromatic and hydrophobic residues are well tolerated at the (–3) position of the ligand. The greatest selectivity for CBX8 within the (–3) series was achieved through the addition of a 4-methoxy phenyl moiety. Further exploration of substituted aromatic rings in the (–3) position of the ligand will aid in a better understanding of selectivity determinants.

The 5-phenylisoxazole N-cap moiety in compound **22** was the most potent dye-free CBX6/8 inhibitor in our series of compounds. Replacement of the 5-methyl to the 5-phenyl group on the terminal isoxazole ring improved binding to CBX6 and **8** by a factor of 60 and 35. These results clearly demonstrate the potential for potent inhibitors to be developed through extension of the ligand to engage the proteins' peptide binding β-groove.

Developing cell-active chromodomain inhibitors is challenging and there are limited examples that have achieved this goal. In spite of the significant interest in CBX's as drug targets, very few phenotypes for CBX inhibition have been reported. Of the 15 reports of chromodomain inhibitors,^[6, 19-22, 25] there are 5 examples of cell-active CBX inhibitors.^[6a, 6f, 6h, 20b, 21] The small molecule inhibitor MS37452 (K_d 30 μM) was shown to be active in cells but required compound treatment of >250 μM to observe decreased amounts of CBX7 on a known target locus.^[6a] The aforementioned UNC3866, with approximately 5% cell permeability, was the first example of a cell-active peptidic chromodomain inhibitor.^[6f] The use of DNA-encoded libraries to target the CBX proteins resulted in the discovery of a CBX8-selective and cell-permeable peptidic inhibitor.^[20b] In addition, the first positive allosteric modulator of CBX7 was reported and shown to regulate PRC1 activity in cellular assays.^[6h] A second allosteric modulator was recently published for CBX8 that showed activity in two different cancer cell lines.^[21] The cell-based studies reported demonstrate that our second generation inhibitors are able to enter cells and engage CBX proteins. Further studies will be required to determine the observed effects are due only to engagement of CBX6 or also through CBX7/CBX8, and to what extent they are inhibited in a cellular context.

Conclusion

CBX6 and CBX8 are emerging therapeutic targets for multiple aggressive malignancies. We have reported the first dual-selective inhibitors for CBX6 and CBX8, and demonstrated their entry into cells. These are among the first examples of inhibitors that bind specific CBX polycomb paralogs with no measurable activity against CBX7. The inhibitors and SAR described in this work will pave the way for future efforts in chromodomain inhibition. Cell permeable chemical probes will aid in unraveling the mysteries of CBX6 and elucidate the paralog-specific roles of both CBX6 and CBX8.

Experimental Section

All experimental procedures along with compound characterizations are reported in the Supporting Information.

Funding

Canadian Cancer Society Research Institute (Innovation Grant 703789); Prostate Cancer Foundation of British Columbia (Pioneer Award to N.M.), NIH (U01CA207532).

Acknowledgements

F.H. thanks the Canada Research Chairs program for ongoing support.; N.M thanks the West Coast Motorcycle Ride to Live and CIHR for scholarship support. We thank Casey Krusemark and James McFarlane for helpful discussions. We thank Michael Gignac and Alok Shaurya for technical assistance.

Abbreviations

CBX, chromobox; DEL, DNA-encoded library; DNA, deoxyribonucleic acid; ESC, embryonic stem cell; FITC, fluorescein isothiocyanate; FP, fluorescence polarization; Glu, glutamic acid; H3, histone 3; H3K27, lysine 27 on histone 3; H3K9, lysine 9 on histone 3; H3K9me3, trimethylated lysine 9 on histone 3; H3K27me3, trimethylated lysine 27 on histone 3; HCC, hepatocellular carcinoma; HP1, heterochromatin protein; IC₅₀, inhibitory concentration that reduces effect by 50 percent; Ket2, di-ethyllysine; Kme3, tri-methyllysine; KO, knockout; N.T., not tested; PRC1, polycomb repressive complex 1; PRC2, polycomb repressive complex 2; PTM, post-translational modification; SAR, structure activity relationship.

Keywords: Methyllysine reader proteins • polycomb paralogs • CBX inhibitors • CBX6 • CBX8

- [1] a) Z. Chen, L. Wang, Q. Wang, W. Li, *Epigenomics* **2010**, 2, 551-560; b) E. L. Greer, Y. Shi, *Nat. Rev. Genet.* **2012**, 13, 343-357; c) R. Hamamoto, V. Saloura, Y. Nakamura, *Nat Rev Cancer* **2015**, 15, 110-124.
- [2] a) N. Milosevich, F. Hof, *Biochemistry* **2016**, 55, 1570-1583; b) T. Wagner, D. Robaa, W. Sippl, M. Jung, *ChemMedChem* **2014**, 9, 466-483.
- [3] L. Kaustov, H. Ouyang, M. Amaya, A. Lemak, N. Nady, S. Duan, G. A. Wasney, Z. Li, M. Vedadi, M. Schapira, J. Min, C. H. Arrowsmith, *The Journal of biological chemistry* **2011**, 286, 521-529.
- [4] a) A. P. Bracken, K. Helin, *Nat. Rev. Cancer* **2009**, 9, 773-784; b) L. Morey, G. Pascual, L. Cozzuto, G. Roma, A. Wutz, S. A. Benitah, L. Di Croce, *Cell Stem Cell* **2012**, 10, 47-62; c) H. Richly, L. Aloia, L. Di Croce, *Cell. Death Dis.* **2011**, 2, e204.
- [5] a) K. Klauke, V. Radulovic, M. Broekhuis, E. Weersing, E. Zwart, S. Olthof, M. Ritsema, S. Bruggeman, X. Wu, K. Helin, L. Bystrykh, G. de Haan, *Nat Cell Biol* **2013**, 15, 353-362; b) A. O'Loughlin, A. M. Munoz-Cabello, A. Gaspar-Maia, H. A. Wu, A. Banito, N. Kunowska, T. Racek, H. N. Pemberton, P. Beolchi, F. Lavial, O. Masui, M. Vermeulen, T. Carroll, J. Graumann, E. Heard, N. Dillon, V. Azuara, A. P. Snijders, G. Peters, E. Bernstein, J. Gil, *Cell Stem Cell* **2012**, 10, 33-46; c) H. L. Chan, L. Morey, *Trends Biochem Sci* **2019**, 44, 688-700.
- [6] a) C. Ren, K. Morohashi, A. N. Plotnikov, J. Jakoncic, S. G. Smith, J. Li, L. Zeng, Y. Rodriguez, V. Stojanoff, M. Walsh, M. M. Zhou, *Chem. Biol.* **2015**, 22, 161-168; b) C. Ren, S. G. Smith, K. L. Yap, S. Li, J. Li, M. Mezei, A. S. Vincek, F. Aguilo, M. J. Walsh, M.-M. Zhou, *ACS Med. Chem. Lett.* **2016**, 7, 601-605; c) C. Simhadri, K. D. Daze, S. F. Douglas, N. Milosevich, L. Monjas, A. Dev, T. M. Brown, A.

- K. H. Hirsch, J. E. Wulff, F. Hof, *ChemMedChem* **2019**, 14, 1444-1456; d) C. Simhadri, K. D. Daze, S. F. Douglas, T. T. Quon, A. Dev, M. C. Gignac, F. Peng, M. Heller, M. J. Boulanger, J. E. Wulff, F. Hof, *J Med Chem* **2014**, 57, 2874-2883; e) C. Simhadri, M. C. Gignac, C. J. Anderson, N. Milosevich, A. Dheri, N. Prashar, R. T. Flemmer, A. Dev, T. G. Henderson, S. F. Douglas, J. E. Wulff, F. Hof, *ACS Omega* **2016**, 1, 541-551; f) J. I. Stuckey, B. M. Dickson, N. Cheng, Y. Liu, J. L. Norris, S. H. Cholensky, W. Tempel, S. Qin, K. G. Huber, C. Sagum, K. Black, F. Li, X.-P. Huang, B. L. Roth, B. M. Baughman, G. Senisterra, S. G. Pattenden, M. Vedadi, P. J. Brown, M. T. Bedford, J. Min, C. H. Arrowsmith, L. I. James, S. V. Frye, *Nat. Chem. Biol.* **2016**, 12, 180-187; g) J. I. Stuckey, C. Simpson, J. L. Norris-Drouin, S. H. Cholensky, J. Lee, R. Pasca, N. Cheng, B. M. Dickson, K. H. Pearce, S. V. Frye, L. I. James, *Journal of Medicinal Chemistry* **2016**, 59, 8913-8923; h) K. N. Lamb, D. Bsteh, S. N. Dishman, H. F. Moussa, H. Fan, J. I. Stuckey, J. L. Norris, S. H. Cholensky, D. Li, J. Wang, C. Sagum, B. Z. Stanton, M. T. Bedford, K. H. Pearce, T. P. Kenakin, D. B. Kireev, G. G. Wang, L. I. James, O. Bell, S. V. Frye, *Cell chemical biology* **2019**, 26, 1365-1379.e1322.
- [7] A. Santanach, E. Blanco, H. Jiang, K. R. Molloy, M. Sansó, J. LaCava, L. Morey, L. Di Croce, *Nature Communications* **2017**, 8, 1235.
- [8] E. Bernstein, E. M. Duncan, O. Masui, J. Gil, E. Heard, C. D. Allis, *Mol Cell Biol* **2006**, 26, 2560-2569.
- [9] C. Creppe, A. Palau, R. Malinverni, V. Valero, M. Buschbeck, *PLOS Genetics* **2014**, 10, e1004851.
- [10] H. Zheng, W.-h. Jiang, T. Tian, H.-s. Tan, Y. Chen, G.-l. Qiao, J. Han, S.-y. Huang, Y. Yang, S. Li, Z.-g. Wang, R. Gao, H. Ren, H. Xing, J.-s. Ni, L.-H. Wang, L.-j. Ma, W.-p. Zhou, *Oncotarget* **2017**, 8, 18872-18884.
- [11] S. Gadd, S. T. Sredni, C. C. Huang, E. J. Perlman, G. Renal Tumor Committee of the Children's Oncology, *Lab Invest* **2010**, 90, 724-738.
- [12] a) C. Z. Zhang, S.-L. Chen, C.-H. Wang, Y.-F. He, X. Yang, D. Xie, J.-P. Yun, *Cancer Research* **2017**, canres.0700.2017; b) B. Tang, Y. Tian, Y. Liao, Z. Li, S. Yu, H. Su, F. Zhong, G. Yuan, Y. Wang, H. Yu, S. Tomlinson, X. Qiu, S. He, *Cell Death & Disease* **2019**, 10, 52.
- [13] J. Tan, M. Jones, H. Koseki, M. Nakayama, A. G. Muntean, I. Maillard, J. L. Hess, *Cancer Cell* **2011**, 20, 563-575.
- [14] C.-Y. Chung, Z. Sun, G. Mullokandov, A. Bosch, Z. A. Qadeer, E. Cihan, Z. Rapp, R. Parsons, J. A. Aguirre-Ghiso, E. F. Farias, B. D. Brown, A. Gaspar-Maia, E. Bernstein, *Cell Reports* **2016**, 16, 472-486.
- [15] W. Xiao, C. Ou, J. Qin, F. Xing, Y. Sun, Z. Li, J. Qiu, *International journal of clinical and experimental pathology* **2014**, 7, 4817-4826.
- [16] G. J. Yuan, X. Chen, J. Lu, Z. H. Feng, S. L. Chen, R. X. Chen, W. S. Wei, F. J. Zhou, D. Xie, *Cancer Sci* **2017**, 108, 2166-2175.
- [17] J. B. Mello, P. D. Ramos Cirilo, O. C. Michelin, M. A. Custodio Domingues, M. V. Cunha Rudge, S. R. Rogatto, I. Maesta, *Placenta* **2017**, 50, 8-15.
- [18] G. Li, C. Warden, Z. Zou, J. Neman, J. S. Krueger, A. Jain, R. Jandial, M. Chen, *PLoS One* **2013**, 8, e80970.
- [19] N. Milosevich, M. C. Gignac, J. McFarlane, C. Simhadri, S. Horvath, K. D. Daze, C. S. Croft, A. Dheri, T. T. H. Quon, S. F. Douglas, J. E. Wulff, I. Paci, F. Hof, *ACS Medicinal Chemistry Letters* **2016**, 7, 139-144.
- [20] a) K. E. Denton, S. Wang, M. C. Gignac, N. Milosevich, F. Hof, E. C. Dykhuizen, C. J. Krusemark, *SLAS DISCOVERY: Advancing Life Sciences R&D* **2018**, 23, 417-428; b) S. Wang, K. E. Denton, K. F. Hobbs, T. Weaver, J. M. B. McFarlane, K. E. Connelly, M. C. Gignac, N. Milosevich, F. Hof, I. Paci, C. A. Musselman, E. C. Dykhuizen, C. J. Krusemark, *ACS Chemical Biology* **2019**.
- [21] J. L. Suh, D. Bsteh, Y. Si, B. Hart, T. M. Weaver, C. Pribitzer, R. Lau, S. Soni, H. Ogana, J. M. Rectenwald, J. L. Norris, S. H. Cholensky, C. Sagum, J. D. Umana, D. Li, B. Hardy, M. T. Bedford, S. M. Mumenthaler, H.-J. Lenz, Y.-m. Kim, G. G. Wang, K. H. Pearce, L. I. James, D. B. Kireev,

-
- C. A. Musselman, S. V. Frye, O. Bell, *bioRxiv* **2021**, 2021.2002.2023.432388.
- [22] N. Milosevich, J. McFarlane, M. C. Gignac, J. Li, T. M. Brown, C. R. Wilson, L. Devorkin, C. S. Croft, R. Hof, I. Paci, J. J. Lum, F. Hof, *Bioorg Med Chem* **2020**, *28*, 115176.
- [23] X. Huang, *J Biomol Screen* **2003**, *8*, 34-38.
- [24] J. I. Geller, J. J. Roth, J. A. Biegel, *Crit Rev Oncog* **2015**, *20*, 199-216.
- [25] a) K. D. Barnash, K. N. Lamb, J. I. Stuckey, J. L. Norris, S. H. Cholensky, D. B. Kireev, S. V. Frye, L. I. James, *ACS Chemical Biology* **2016**, *11*, 2475-2483; b) L. Yang, Y. Liu, M. Fan, G. Zhu, H. Jin, J. Liang, Z. Liu, Z. Huang, L. Zhang, *Eur J Med Chem* **2019**, *182*, 111656.

Entry for the Table of Contents

Dual active CBX inhibitor: A library of peptidomimetic based inhibitors for chromobox homolog 6 and 8 (CBX6/8) was developed. Through SAR key residues important for selectivity and potency were identified. We arrived at a dual active CBX6/8 inhibitor with sub micromolar affinity and 18-fold selectivity over CBX7. Preliminary cell studies show our compounds are cell permeable and have an effect on cell proliferation.

Institute and/or researcher Twitter usernames: @Hof_group_news

Supplementary information for:

Polycomb paralog chromodomain inhibitors active against both CBX6 and CBX8

Natalia Milosevich,^{[a],[†]} Chelsea R. Wilson,^{[a],[†]} Tyler M. Brown,^[a] Aktan Alpsoy,^[b] Sijie Wang,^[b]
Katelyn E. Connelly,^[b] Kirsten A. D. Sinclair,^[a] Felino R. Ponio,^[a] Rebecca Hof,^[a] Emily C.
Dykhuisen,^[b] and Fraser Hof^{*[a]}

Table of Contents

Experimental methods and data	2
General synthesis	2
General peptide synthesis methods	2
Synthesis of FITC labeled peptide probe for FP	3
Synthesis of diethyl-lysine containing peptides	4
Characterization data	5
Supplementary Figure S1	5
Supplementary Figure S2	6
Supplementary Figure S3	7
Supplementary Figure S4	8
Supplementary Figure S5	9
Supplementary Figure S6	10
Supplementary Figure S7	11
Supplementary Figure S8	12
Supplementary Figure S9	13
Supplementary Figure S10	14
Supplementary Figure S11	15
Supplementary Figure S12	16
Supplementary Figure S13	17
Supplementary Figure S14	18
Supplementary Figure S15	19
Supplementary Figure S16	20
Supplementary Figure S17	21
Supplementary Figure S18	22
Supplementary Figure S19	23
Supplementary Figure S20	24
Supplementary Figure S21	25
Supplementary Figure S22	26
Supplementary Figure S23	27
Protein expression and purification	28
Fluorescence polarization methods	28
Fluorescence polarization data	31
Supplementary Figure S24	31
Supplementary Figure S25	32

Supplementary Figure S26.....	33
Supplementary Figure S27.....	34
Supplementary Figure S28.....	35
Supplementary Figure S29.....	36
Supplementary Figure S30.....	37
Supplementary Figure S31.....	38
Supplementary Figure S32.....	39
Supplementary Figure S33.....	40
Supplementary Figure S34.....	41
Supplementary Figure S35.....	42
Supplementary Figure S36.....	43
Supplementary Figure S37.....	44
Supplementary Figure S38.....	45
Supplementary Figure S39.....	46
Supplementary Figure S40.....	47
Supplementary Figure S41.....	48
Supplementary Figure S42.....	49
Supplementary Figure S44.....	51
Supplementary Figure S45.....	52
Supplementary Figure S46.....	53
Methods for cellular Assays.....	54
CellTiter-Glo luminescent cell viability assay.....	54
Supplementary Figure S47.....	55

Experimental methods and data

General synthesis

All natural and unnatural amino acids and coupling agents were purchased from ChemImpex, Enamine and Sigma Aldrich. Acetylaldehyde and sodium cyanoborohydride were purchased from Sigma Aldrich and carboxylic acids used for N-terminal modifications were purchased from Enamine, ChemImpex and Sigma Aldrich. Fmoc-L-Lys(Me)₃-OH was purchased from GL Biochem (Shanghai).

General peptide synthesis methods

Peptides were either synthesized using an automated peptide synthesizer, or by manual peptide synthesis using Fmoc solid phase peptide synthesis protocols. Automated synthesis was achieved on Rink amide resin (ChemImpex International) using CEM Liberty-1 microwave synthesizer with standard Fmoc protocols using HBTU as the coupling agent. For manual peptide synthesis, Fmoc-protected Rink amide resin (ChemImpex International) was swelled in CH₂Cl₂ for 20 minutes. Deprotection of Fmoc was done using 20% piperidine in DMF (10 mL, 5 minutes, 3

times) and washed with DMF (10 mL, 3 times following each incubation). Fmoc- and side chain protected amino acids as well as N-capping carboxylic acids (5 eq.) were dissolved in DMF (10 mL) with HBTU (5 eq.), and DIPEA (10 eq.), added to the Fmoc deprotected resin and/or Fmoc deprotected peptide on resin, bubbled with N₂ for one hour at room temperature. The resin was washed with DMF (10 mL, 3 times), and washed with CH₂Cl₂ (10 mL, 3 times) after the final coupling.

Peptide products were cleaved from resin with 10 mL of 95:2.5:2.5 TFA/H₂O/triisopropylsilane for 2.5 h. The solution was then concentrated in vacuo and added to cold diethyl ether to yield a crude yellow precipitate that was collected by centrifugation. Purifications by preparative HPLC were carried out using a Phenomenex Luna, 5 µm, C-18 column, 250 x 21.20 mm column.

Peptides were characterized using UPLC, HPLC, and ESI-MS and purity was determined to be >95% by analytical HPLC or analytical UPLC (exceptions are noted in the characterization data). Retention times reported arise from analytical traces done on either a Thermo Scientific C-18 column (4.6 mm x 250 mm, 5 µm), Waters UPLC-MS Acquity UPLC BEH C18 column (2.1 x 50 mm, 1.7 µm), or Shimadzu HPLC C-18 column (4.6 mm x 250 mm, 5 µm). Traces done on the Thermo or Shimadzu HPLC were at a flow rate of 1.5 mL/min, gradient running from 90:10 water (0.1 % TFA) and MeCN (0.1 % TFA) to 10:90 water (0.1 % TFA) and MeCN (0.1 % TFA) over 30-38 min. Traces done on the Waters UPLC-MS were at a flow rate of 0.5 mL/min with a gradient running from 90:10 water (0.1 % TFA) and MeCN (0.1 % TFA) to 10:90 water (0.1 % TFA) and MeCN (0.1 % TFA) over five minutes.

Synthesis of FITC labeled peptide probe for FP

For side-chain functionalization with fluorescein isothiocyanate (FITC), a side-chain (Mtt)-protected lysine was used. For functionalization of side-chain Mtt-protected lysine, the Mtt protecting group was cleaved after complete synthesis on resin. Selective cleavage was achieved with treatment of 2:2:96 TFA/triisopropylsilane/CH₂Cl₂ for 20 minutes (×4). The resin was filtered and washed with CH₂Cl₂ (3 × 5 mL) and DMF (3 × 5 mL). FITC (5 equiv) was added to a 10 mL solution of 1:1 DMF and pyridine. This solution was added to the resin and bubbled with N₂ at ambient temperature overnight. Following the functionalization with FITC, all

peptide/resin solutions were filtered and the resin washed with DMF (3×5 mL) and CH_2Cl_2 (3×5 mL) and air-dried.

Synthesis of diethyl-lysine containing peptides

Side-chain (Mtt)-protected lysine was used in the synthesis protocols, and cleaved as described previously in Chapter 2, section 2.8.2 (treatment with 2:2:96 TFA/triisopropylsilane/ CH_2Cl_2 for 20 minutes, 4 times). The resin was filtered and washed with CH_2Cl_2 (3×5 mL) and DMF (3×5 mL). A solution of acetaldehyde (50-100 eq.), sodium cyanoborohydride (10 eq.), and 1 drop of glacial acetic acid in DMF (10 mL) was added to the peptide on resin. The solution was filtered and the resin washed with DMF (3×5 mL) and CH_2Cl_2 (3×5 mL) and air-dried. Crude peptides were dissolved in a mixture of H_2O and MeCN or DMSO, depending on solubility. Peptide cleavage from resin and purification are the same as reported above in section 4.7.3. Diethyl peptides with poor solubility were purified from DMSO stocks with a gradient running from 60:40 water (0.1 % TFA) and MeCN (0.1 % TFA) to 10:90 water (0.1 % TFA) and MeCN (0.1 % TFA) over 30-38 min.

Characterization data

Supplementary Figure S1: Characterization data for compound 1. A) Analytical HPLC trace, B) Low resolution mass spectrum. LR-ESI-MS: $[m/z]$ calcd. for $C_{40}H_{53}BrN_7O^+$: 838.31; found: 838.53.

No.	Ret.Time min	Peak Name	Height mAU	Area mAU*min	Rel.Area %	Amount μl	Type
1	24.02	n.a.	0.664	0.058	0.16	n.a.	BMB*
2	24.25	n.a.	292.917	37.070	99.78	n.a.	bBMB*
3	25.24	n.a.	0.384	0.024	0.06	n.a.	BMB*
Total:			293.965	37.152	100.00	0.000

Supplementary Figure S2: Characterization data for compound **5**. A) Analytical HPLC trace, B) Low resolution mass spectrum. LR-ESI-MS: $[m/z]$ calcd. for $C_{47}H_{64}BrN_8O_{11}^+$: 995.39; found: 995.67.

Supplementary Figure S3: Characterization data for compound **6**. A) Analytical HPLC trace, B) Low resolution mass spectrum. LR-ESI-MS: $[m/z]$ calcd. For $C_{48}H_{66}BrN_8O_{11}^+$: 1011.41; found: 1011.4.

Supplementary Figure S5: Characterization data for compound **8**. A) Analytical HPLC trace, B) Low resolution mass spectrum. LR-ESI-MS: [m/z] calcd. for: $C_{43}H_{58}BrN_7O_8^+$: 880.36; found: 880.47.

Supplementary Figure S6: Characterization data for compound **9**. A) Analytical UPLC trace. Product peak is at 2.82 minutes with area of 98 %. B) Low resolution mass spectrum. LR-ESI-MS: $[m/z]$ calcd. for: $C_{40}H_{61}BrN_7O_7^+$: 830.38; found: 830.33.

Supplementary Figure S7: Characterization data for compound **10**. A) Analytical HPLC trace, B) Low resolution mass spectrum. LR-ESI-MS: $[m/z]$ calcd. for: $C_{37}H_{60}N_9O_7^+$: 742.46; found: 742.60.

Supplementary Figure S9: Characterization data for compound **12**. A) Analytical HPLC trace, B) Low resolution mass spectrum. LR-ESI-MS: $[m/z]$ calcd. for: $C_{41}H_{59}N_8O_9^+$: 807.44; found: 807.5.

Supplementary Figure S10: Characterization data for compound **13**. A) Analytical HPLC trace, B) Low resolution mass spectrum. LR-ESI-MS: $[m/z]$ calcd. for: $C_{37}H_{66}N_9O_7^+$: 748.51; found: 748.60.

Supplementary Figure S11: Characterization data for compound **14**. A) Analytical HPLC trace, B) Low resolution mass spectrum. LR-ESI-MS: $[m/z]$ calcd. for: $C_{38}H_{67}N_8O_8^+$: 763.51; found: 763.47.

Supplementary Figure S12: Characterization data for compound **15**. A) Analytical HPLC trace, B) Low resolution mass spectrum. LR-ESI-MS: $[m/z]$ calcd. for: $C_{39}H_{63}N_8O_9^+$: 787.47; found: 787.5.

A)

16:21:24

20190331_Tbrown19

31-May-2019

2: Diode Array

Range: 1.308e+2

Time	Height	Area	Area%
2.1517	6736433	143337.13	4.93
2.1892	75017800	2762469.00	95.07

B)

20190331_Tbrown19 1052 (2.191)

1: Scan ES+
5.31e7

Supplementary Figure S13: Characterization data for compound **16**. A) Analytical HPLC trace, B) Low resolution mass spectrum. LR-ESI-MS: $[m/z]$ calcd. for: $C_{38}H_{60}FN_8O_8^+$: 775.45; found: 775.6

Supplementary Figure S14: Characterization data for compound **17**. A) Analytical HPLC trace, B) Low resolution mass spectrum. LR-ESI-MS: $[m/z]$ calcd. for: $C_{38}H_{60}IN_8O_8^+$: 883.36; found: 883.5

Supplementary Figure S15: Characterization data for compound **18**. A) Analytical HPLC trace, B) Low resolution mass spectrum. LR-ESI-MS: $[m/z]$ calcd. for: $C_{37}H_{60}N_9O_8^+$: 758.46; found: 758.4.

Supplementary Figure S16: Characterization data for compound **19**. A) Analytical HPLC trace, B) Low resolution mass spectrum. LR-ESI-MS: $[m/z]$ calcd. for: $C_{36}H_{59}N_8O_8S^+$: 763.42; found: 763.5

Supplementary Figure S17: Characterization data for compound **20**. A) Analytical UPLC trace. Product peak is at 3.65 minutes with an area of 92%. B) Low resolution mass spectrum. LR-ESI-MS: $[m/z]$ calcd. for: $C_{40}H_{71}N_8O_8^+$: 791.54; found: 791.53.

Supplementary Figure S18: Characterization data for compound **21**. A) Analytical HPLC trace, B) Low resolution mass spectrum. LR-ESI-MS: $[m/z]$ calcd. for: $C_{41}H_{67}N_8O_8^+$: 773.46; found: 773.6

Supplementary Figure S19: Characterization data for compound **22**. A) Analytical UPLC trace. Product peak is at 2.87 minutes with an area of 97%. B) Low resolution mass spectrum. LR-ESI-MS: $[m/z]$ calcd. for: $C_{43}H_{63}N_8O_8^+$: 819.48; found: 819.4.

Supplementary Figure S20: Characterization data for compound **23**. A) Analytical HPLC trace, B) Low resolution mass spectrum. LR-ESI-MS: $[m/z]$ calcd. for: $C_{46}H_{61}N_8O_9^+$: 869.46; found: 869.6

Supplementary Figure S21: Characterization data for compound **24**. A) Analytical UPLC trace, B) Low resolution mass spectrum. LR-ESI-MS: $[m/z]$ calcd. for: $C_{38}H_{59}N_8O_8^+$: 755.45; found: 755.5.

Supplementary Figure S22: Characterization data for compound **25**. A) Analytical HPLC trace, B) Low resolution mass spectrum. LR-ESI-MS: $[m/z]$ calcd. for: $C_{40}H_{63}N_8O_8^+$: 783.48; found: 783.6.

Supplementary Figure S23: Characterization data for compound **26**. A) Analytical HPLC trace, B) Low resolution mass spectrum. LR-ESI-MS: $[m/z]$ calcd. for: $C_{39}H_{63}N_8O_8^+$: 771.48; found:

Protein expression and purification

E. coli pellets were thawed overnight and sequentially lysed via a 30 minute incubation with a CHAPS/dH₂O solution (Biobasic) and sonication. After centrifugation to remove cell debris, the purification procedure involved two chromatographic steps. First, an affinity chromatographic step on a nickel-nitrilotriacetic acid chelating column (Qiagen) followed by a gel filtration step using an FPLC outfitted with a Hiload 16/600 Superdex 75 pg size exclusion column (GE). The FPLC step performed a buffer exchange into our ‘minimal’ fluorescence polarization (FP) buffer consisting of 20 mM Tris HCl, 250 mM NaCl, and 0.01% Tween-20. Purified protein was then concentrated to the desired volume using Amicon centrifugal filter units. Proteins were flash frozen and stored at –80°C while purity was assessed by SDS page. His tags were cleaved off overnight after the nickel column step using our in-house TEV protease.

Fluorescence polarization methods

Direct FP method

Direct FP was done by titration of CBX proteins into FITC-labelled probes. Binding assays were performed in black 384 well plates with optical bottoms at a total volume of 50 µL per well. All solutions were dissolved in ‘full’ FP buffer made from the addition of DTT, PMSF and benzamidine (1 mM final concentration) into ‘minimal’ FP buffer (20 mM Tris HCl, 250 mM NaCl, 0.01% Tween-20). A solution containing 100 nM of the labelled probe and a varying concentration of the CBX protein in question, was titrated into wells containing a solution with a constant concentration (100 nM) of FITC-labelled probe. 100 µL of the CBX-containing solution was added to well 19. Wells 2-18 contained 50 µL of the FITC-probe only solution. 50 µL of the well 19 solution was added to well 18 and mixed 4 times. These dilutions continued until well 3, whereupon 50 µL was discarded after mixing, leaving the total volume of all wells at 50 µL. Well 1 contained only full FP buffer and well 2 contained only the FITC-probe in full FP buffer. The maximum concentrations used for each CBX protein varied (0.001 µM – 600 µM) based on availability and solubility of the protein at high concentration, as well as the expected strength of the FITC-probe being tested. Assays were performed in triplicate or duplicate depending on CBX protein availability (noted in raw data). After removing any bubbles with a needle, plates were incubated in the dark at RT for 15 minutes and then read with a SpectraMax M5 plate reader (Molecular Devices). Raw data for perpendicular (P) and parallel (S) light was collected

with SoftMax Pro software using a fluorescence polarization endpoint read. 100 reads were made and averaged per well with 5 second automix before reading and medium PMT sensitivity. Excitation was 450 nm for FITC and emission was 530 nm (with a 515 nm cut-off).

Analysis of direct FP data

The background of the blank buffer was subtracted from the parallel and perpendicular intensities of emission. Average values for parallel and perpendicular emission intensities (minus the average values for buffer control wells) were determined for sets of replicates using Microsoft Excel and millipolarization units (mP) were calculated using the formula $((\text{Parallel} - \text{Perpendicular})/(\text{Parallel} + \text{Perpendicular})) \times 1000$. Data were then plotted using Graphpad Prism 8 and fitted using a 'one-site' total binding model. The mathematical model for the fit was described by the equation $y = (B_{\text{max}} \times x)/(K_d + x) + NS \times x + \text{background}$ (B_{max} = maximum specific binding, NS = slope of nonspecific binding and background = amount of nonspecific binding with no added labeled ligand). For certain assays, some of the data points (0 to 5 points) representing the highest CBX concentrations had to be excluded due to abnormally high mP values, likely arising from protein aggregation.

Competitive FP method

Competitive FP analysis of **1** and **4** binding to CBX6/7/8 was used to determine IC_{50} values as with compound **2** as the competitive binding probe. Compound **2** and peptide inhibitors were dissolved in DMSO and then diluted into FP buffer. DMSO was kept below 2.5% for the majority of the assays, with the exception of **6**, **8**, **22** and **23** (5%DMSO). The assay was carried out in NUNC black 96-well plates and 'full' FP buffer was used as described above (20 mM Tris-HCl, pH = 8.0, 250 mM NaCl, 1 mM DTT, 1 mM benzamidine, 1 mM PMSF, 0.01% Tween). Two solutions were used, one containing protein, FITC probe **2**, and peptide inhibitor and another with just the protein and FITC probe **2**. Probe **2** was at a constant concentration of 100 nM in both solutions. Each CBX protein was at a constant concentration dependent on the K_d of the protein to probe **2**. CBX1 was used at 10 μM , CBX6 at 1 μM , CBX7 at 0.4 μM and CBX8 at 4 μM . Inhibitor concentrations varied between 0-500 μM . All wells in the plate were made to final volume of 100 μL . 200 μL of the inhibitor-containing solution was added to well 12. Wells 2-11 contained 100 μL of the CBX and **2** only solution. 100 μL of the well 12 solution

was added to well 11 and mixed. These dilutions continued to well 3, whereupon 100 μ L was discarded after mixing, leaving the total volume of all wells at 100 μ L. Well 1 contained only full FP buffer and well 2 contained only CBX protein and **2** in full FP buffer. Following serial dilution of the inhibitor across the plate, plates were incubated for 15 minutes in the dark and then were read with Cytation 5 cell imaging and multi-mode reader plate reader. A polarized filter cube was used with filters for excitation at 485 nm and emission at 520 nm (with a cut-off of 510 nm). All plates were read at 28°C.

Analysis of competitive FP data

The background of the blank buffer was subtracted from the parallel and perpendicular intensities of emission. Values were graphed using Graphpad Prism 8 and fitted using a sigmoidal curve function from which IC₅₀ values were extrapolated. The equation used is described as follows: $y = \text{bottom} + (\text{top} - \text{bottom}) / (1 + 10^{(X - \text{LogIC}_{50})})$. Errors are reported as 95% confidence intervals. Experiments were done in duplicate or triplicate and this is indicated in the raw data shown.

Fluorescence polarization data

Supplementary Figure S24: Competitive fluorescence polarization data of compound **1** with CBX proteins preformed in duplicate. A) CBX1 ($\log IC_{50} = -4.845$ M, 95% CI -5.137 M to -4.510 M, $R^2 = 0.928$), B) CBX6 ($\log IC_{50} = -5.113$ M, 95% CI -5.222 M to -5.00 M, $R^2 = 0.990$), C) CBX7 ($\log IC_{50} = -5.862$ M, 95% CI -5.983 M to -5.746 M, $R^2 = 0.983$), D) CBX8 ($\log IC_{50} = -4.385$ M 95% CI -4.593 M to -4.162 M, $R^2 = 0.967$). Error bars reported as asymmetrical 95 % confidence intervals.

Supplementary Figure S25: Competitive fluorescence polarization data of compound **5** with CBX proteins. CBX6 preformed in triplicate, CBX7 and CBX8 preformed in duplicate. A) CBX6 ($\log IC_{50} = -4.756$ M, 95% CI -4.834 M to -4.678 M, $R^2 = 0.990$), B) CBX7 ($\log IC_{50} = -3.672$ M, 95% CI -3.750 M to -3.590 M, $R^2 = 0.968$), C) CBX8 ($\log IC_{50} = -4.778$, 95% CI -4.949 M to -4.602 M, $R^2 = 0.969$). Error bars reported as asymmetrical 95 % confidence intervals. CBX7 was fitted constraining the mP value of the endpoint to 100 millipolarization units.

Supplementary Figure S26: Competitive fluorescence polarization data of compound 6 with CBX proteins. CBX1 preformed in duplicate; CBX6, CBX7 and CBX8 performed in triplicate. A) CBX1 ($\log IC_{50} = > -3.301$ M), B) CBX6 ($\log IC_{50} = -5.072$ M, 95% CI -5.276 M to -4.861 M, $R^2 = 0.931$), C) CBX7 ($\log IC_{50} = > -3.301$ M), D) CBX8 ($\log IC_{50} = -4.871$ M, 95% CI -4.997 M to -4.745 M, $R^2 = 0.973$). Error bars reported as asymmetrical 95% confidence intervals.

Supplementary Figure S27: Competitive fluorescence polarization data of compound **7** with CBX proteins performed in triplicate. A) CBX1 ($\log\text{IC}_{50} = > -3.301$ M), B) CBX6 ($\log\text{IC}_{50} = -4.625$ M, 95% CI -4.754 M to -4.490 M, $R^2 = 0.972$), C) CBX7 ($\log\text{IC}_{50} = > -3.301$ M), D) CBX8 ($\log\text{IC}_{50} = -4.734$ M, 95% CI -4.856 M to -4.4605 M, $R^2 = 0.984$). Error bars reported as asymmetrical 95 % confidence intervals.

Supplementary Figure S28: Competitive fluorescence polarization data of compound **8** with CBX proteins performed in triplicate. A) CBX1 ($\log IC_{50} = > -3.69$ M) B) CBX6 ($\log IC_{50} = -5.495$ M, 95% CI -5.591 M to -5.398 M, $R^2 = 0.985$), C) CBX7 ($\log IC_{50} = > -3.69$ M), D) CBX8 ($\log IC_{50} = -4.746$ M, 95% CI -4.896 M to -4.586 M, $R^2 = 0.969$). Error bars reported as asymmetrical 95 % confidence intervals. Each assay was performed with 5% DMSO due to solubility issues of the inhibitor.

Supplementary Figure S29: Competitive fluorescence polarization data of compound **9** with CBX proteins. CBX1 and CBX8 preformed in triplicate, CBX6 and CBX7 preformed in triplicate. A) CBX1 ($\log\text{IC}_{50} = > -3.522$ M) B) CBX6 ($\log\text{IC}_{50} = -5.178$ M, 95% CI -5.235 M to -5.120 M, $R^2 = 0.995$), C) CBX7 ($\log\text{IC}_{50} = > -3.602$ M), D) CBX8 ($\log\text{IC}_{50} = -5.342$ M, 95% CI -5.525 M to -5.147 M, $R^2 = 0.947$). Error bars reported as asymmetrical 95 % confidence intervals.

Supplementary Figure S30: Competitive fluorescence polarization data of compound **10** with CBX proteins performed in triplicate. A) CBX1 ($\log \text{IC}_{50} = > -3.522 \text{ M}$) B) CBX6 ($\log \text{IC}_{50} = -4.143 \text{ M}$, 95% CI -4.205 M to -4.080 M , $R^2 = 0.994$), C) CBX7 ($\log \text{IC}_{50} = -3.791 \text{ M}$, 95% CI -3.892 M to -3.683 M , $R^2 = 0.989$), D) CBX8 ($\log \text{IC}_{50} = -4.169 \text{ M}$, 95% CI -4.238 M to -4.099 M , $R^2 = 0.993$). Error bars reported as asymmetrical 95 % confidence intervals.

Supplementary Figure S31: Competitive fluorescence polarization data of compound **11** with CBX proteins performed in triplicate. A) CBX1 ($\log IC_{50} = > -3.522$ M) B) CBX6 ($\log IC_{50} = -4.982$ M, 95% CI -4.983 M to -4.872 M, $R^2 = 0.995$), C) CBX7 ($\log IC_{50} = -3.978$ M, 95% CI -4.013 M to -3.941 M, $R^2 = 0.998$), D) CBX8 ($\log IC_{50} = -5.186$ M, 95% CI -5.265 M to -5.106 M, $R^2 = 0.987$). Error bars reported as asymmetrical 95 % confidence intervals.

Supplementary Figure S32: Competitive fluorescence polarization data of compound **12** with CBX proteins performed in triplicate. A) CBX1 ($\log\text{IC}_{50} = > -3.301$ M) B) CBX6 ($\log\text{IC}_{50} = -4.781$ M, 95% CI -4.804 M to -4.758 M, $R^2 = 0.999$), C) CBX7 ($\log\text{IC}_{50} = > -3.522$ M), D) CBX8 ($\log\text{IC}_{50} = -5.207$ M, 95% CI -5.284 M to -5.130 M, $R^2 = 0.991$). Error bars reported as asymmetrical 95 % confidence intervals.

Supplementary Figure S33: Competitive fluorescence polarization data of compound **13** with CBX proteins performed in triplicate. A) CBX1 ($\log\text{IC}_{50} = > -3.522$ M) B) CBX6 ($\log\text{IC}_{50} = -4.502$ M, 95% CI -4.528 M to -4.475 M, $R^2 = 0.999$), C) CBX7 ($\log\text{IC}_{50} = -3.981$ M, 95% CI -4.030 M to -3.932 M, $R^2 = 0.997$), D) CBX8 ($\log\text{IC}_{50} = -4.594$ M, 95% CI -4.673 M to -4.515 M, $R^2 = 0.992$). Error bars reported as asymmetrical 95 % confidence intervals.

Supplementary Figure S34: Competitive fluorescence polarization data of compound **14** with CBX proteins performed in triplicate. A) CBX1 ($\log IC_{50} = > -3.522$ M) B) CBX6 ($\log IC_{50} = -5.080$ M, 95% CI -5.120 M to -5.039 M, $R^2 = 0.997$), C) CBX7 ($\log IC_{50} = -3.676$ M, 95% CI -3.833 M to -3.698 M, $R^2 = 0.996$), D) CBX8 ($\log IC_{50} = -5.234$ M, 95% CI -5.317 M to -5.150 M, $R^2 = 0.985$). Error bars reported as asymmetrical 95 % confidence intervals.

Supplementary Figure S35: Competitive fluorescence polarization data of compound **15** with CBX proteins performed in triplicate. A) CBX1 ($\log IC_{50} = > -3.301$ M) B) CBX6 ($\log IC_{50} = -4.934$ M, 95% CI -4.969 M to -4.899 M, $R^2 = 0.998$), C) CBX7 ($\log IC_{50} = > -3.602$ M), D) CBX8 ($\log IC_{50} = -6.189$ M, 95% CI -6.239 M to -6.138 M, $R^2 = 0.992$). Error bars reported as asymmetrical 95 % confidence intervals.

Supplementary Figure S36: Competitive fluorescence polarization data of compound **16** with CBX proteins performed in triplicate. A) CBX6 ($\log\text{IC}_{50} = -5.030$ M, 95% CI -5.073 M to -4.986 M, $R^2 = 0.996$), B) CBX7 ($\log\text{IC}_{50} = -3.216$ M, 95% CI -3.414 M to -2.955 M, $R^2 = 0.977$), C) CBX8 ($\log\text{IC}_{50} = -5.734$, 95% CI -5.824 M to -5.641 M, $R^2 = 0.974$). Error bars reported as asymmetrical 95 % confidence intervals.

Supplementary Figure S37: Competitive fluorescence polarization data of compound **17** with CBX proteins performed in triplicate. A) CBX6 ($\log\text{IC}_{50} = -4.165$ M, 95% CI -4.270 M to -4.058 M, $R^2 = 0.984$), B) CBX7 ($\log\text{IC}_{50} = > -3.301$ M), C) CBX8 ($\log\text{IC}_{50} = -4.699$, 95% CI -4.773 M to -4.624 M, $R^2 = 0.994$). Error bars reported as asymmetrical 95 % confidence intervals.

Supplementary Figure S38: Competitive fluorescence polarization data of compound **18** with CBX proteins performed in triplicate. A) CBX1 ($\log IC_{50} = > -3.301$ M) B) CBX6 ($\log IC_{50} = -4.480$ M, 95% CI -4.535 M to -4.425 M, $R^2 = 0.995$), C) CBX7 ($\log IC_{50} = > -3.602$ M), D) CBX8 ($\log IC_{50} = -4.758$ M, 95% CI -4.800 M to -4.715 M, $R^2 = 0.996$). Error bars reported as asymmetrical 95 % confidence intervals.

Supplementary Figure S39: Competitive fluorescence polarization data of compound **19** with CBX proteins performed in triplicate. A) CBX6 ($\log IC_{50} = -5.303$ M, 95% CI -5.379 M to -5.224 M, $R^2 = 0.988$), B) CBX7 ($\log IC_{50} = -3.679$ M, 95% CI -3.823 M to -3.523 M, $R^2 = 0.979$), C) CBX8 ($\log IC_{50} = -5.792$ M, 95% CI -5.846 M to -5.737 M, $R^2 = 0.990$). Error bars reported as asymmetrical 95 % confidence intervals.

Supplementary Figure S40: Competitive fluorescence polarization data of compound **20** with CBX proteins performed in triplicate. A) CBX1 ($\log IC_{50} = > -3.301$ M) B) CBX6 ($\log IC_{50} = -5.546$ M, 95% CI -5.592 M to -5.499 M, $R^2 = 0.996$), C) CBX7 ($\log IC_{50} = -4.401$ M, 95% CI -4.438 M to -4.364 M, $R^2 = 0.998$), D) CBX8 ($\log IC_{50} = -5.125$ M, 95% CI -4.989 M to -5.301 M, $R^2 = 0.964$). Error bars reported as asymmetrical 95 % confidence intervals.

Supplementary Figure S41: Competitive fluorescence polarization data of compound **21** with CBX proteins performed in triplicate. A) CBX6 ($\log\text{IC}_{50} = -5.126$ M, 95% CI -5.244 M to -5.006 M, $R^2 = 0.973$), B) CBX7 ($\log\text{IC}_{50} = -3.572$ M, 95% CI -3.721 M to -3.406 M, $R^2 = 0.979$), C) CBX8 ($\log\text{IC}_{50} = -5.656$, 95% CI -5.732 M to -5.578 M, $R^2 = 0.984$). Error bars reported as asymmetrical 95 % confidence intervals.

Supplementary Figure S42: Competitive fluorescence polarization data of compound **22** with CBX proteins performed in triplicate. A) CBX1 ($\log\text{IC}_{50} = > -3.301$ M) B) CBX6 ($\log\text{IC}_{50} = -6.793$ M, 95% CI -6.866 M to -6.719 M, $R^2 = 0.990$), C) CBX7 ($\log\text{IC}_{50} = -5.440$ M, 95% CI -5.600 M to -5.278 M, $R^2 = 0.960$), D) CBX8 ($\log\text{IC}_{50} = -6.809$ M, 95% CI -6.945 M to -6.672 M, $R^2 = 0.957$). Error bars reported as asymmetrical 95 % confidence intervals.

Supplementary Figure S43: Competitive fluorescence polarization data of compound **23** with CBX proteins performed in triplicate. A) CBX6 ($\log\text{IC}_{50} = -5.667$ M, 95% CI -5.895 M to -5.412 M, $R^2 = 0.961$), B) CBX7 ($\log\text{IC}_{50} = -5.270$ M, 95% CI -5.360 M to -5.177 M, $R^2 = 0.985$), C) CBX8 ($\log\text{IC}_{50} = -5.999$, 95% CI -6.147 M to -5.849 M, $R^2 = 0.975$). Error bars reported as asymmetrical 95 % confidence intervals.

Supplementary Figure S44: Competitive fluorescence polarization data of compound **24** with CBX proteins performed in triplicate. A) CBX1 ($\log\text{IC}_{50} = > -3.301$ M) B) CBX6 ($\log\text{IC}_{50} = -5.300$ M, 95% CI -5.362 M to -5.237 M, $R^2 = 0.993$), C) CBX7 ($\log\text{IC}_{50} = -4.654$ M, 95% CI -4.768 M to -4.538 M, $R^2 = 0.977$), D) CBX8 ($\log\text{IC}_{50} = -5.518$ M, 95% CI -5.620 M to -5.414 M, $R^2 = 0.976$). Error bars reported as asymmetrical 95 % confidence intervals.

Supplementary Figure S45: Competitive fluorescence polarization data of compound **25** with CBX proteins performed in triplicate. A) CBX1 ($\log IC_{50} = > -3.699$ M) B) CBX6 ($\log IC_{50} = -3.750$ M, 95% CI -3.837 M to -3.659 M, $R^2 = 0.989$), C) CBX7 ($\log IC_{50} = -3.43$ M, 95% CI -3.458 M to -3.400 M, $R^2 = 0.983$), D) CBX8 ($\log IC_{50} = -4.751$ M, 95% CI -4.824 M to -4.678 M, $R^2 = 0.989$). Error bars reported as asymmetrical 95 % confidence intervals.

Supplementary Figure S46: Competitive fluorescence polarization data of compound **26** with CBX proteins performed in triplicate. A) CBX1 (logIC₅₀ = > -3.301 M), B) CBX6 (logIC₅₀ = -3.582 M, 95% CI -3.631 M to -3.522 M, $R^2 = 0.973$), C) CBX7 (logIC₅₀ = > -3.301 M), D) CBX8 (logIC₅₀ = -4.315 M, 95% CI -4.356 M to -4.274 M, $R^2 = 0.997$). Error bars reported as asymmetrical 95 % confidence intervals.

Methods for cellular Assays

Generation of CBX6 KO cell lines:

Oligonucleotides for control guide RNA (sgControl: 5'-GTAGCGAACGTGTCCGGCGT-3') and CBX6 targeting guide RNAs (sgCBX6_1: 5'-CATCAAACGGCGGATCCGAA-3' ; sgCBX6_2: 5'-TCGGATCCGCCGTTTGATGA-3') were cloned into lenticrispr v2.0 vector (a gift from Feng Zhang Addgene plasmid no. 52961). HEK293T cells were used to make lentiviral particles for each construct. G401 cells were transduced with concentrated virus particles and selected with puromycin to generate stable cell lines.

CellTiter-Glo luminescent cell viability assay

The effect of **8**, **11**, **22** and **23** on cell viability was determined using a CellTiter-Glo ATP detection system (#G7573, Promega). G401 cells were seeded in 0.1×10^6 cells/mL density in 96-well clear bottom white microplate (#655098, Greiner Bio-One). Peptides (50, 100 μ M) were added to the cells with 0.5% DMSO. After 3d, CellTiter-Glo reagent was added to cells, and incubated with gentle shake for 15 minutes in dim light at R.T. Luminescence was read on an GloMax® microplate reader.

Supplementary Figure S47: Generation of CBX6 knockout in G401 Rhabdoid tumor cell line. A) Immunoblot analysis of whole cell lysates from kidney epithelial cells (HK2), clear cell renal cancer cell lines (Caki1, Caki2, 786O), kidney adrenal cancer (SW13), kidney rhabdoid tumor (G401, and muscle rhabdoid tumor A204. Top band is CBX6 and bottom band is non-specific. Beta actin was included as a loading control. B) CRISPR-Cas9 was used to knock out CBX6 in G401 cells. Immunoblot analysis confirms a reduction in CBX6, but not other CBX subunit protein levels. C) G401 cells were grown for 72h and luminescence was measured after addition of CellTiter Glo®.