Supporting Information

for

Chromodomain antagonists that target the polycomb-group methyllysine reader protein Chromobox homolog 7 (CBX7)

Chakravarthi Simhadri,^{1,†} Kevin D. Daze,^{1,†} Sarah F. Douglas,¹ Taylor T.H. Quon,¹ Amarjot Dev,¹ Michael Gignac,¹ Fangni Peng,² Markus Heller,³ Martin J. Boulanger,² Jeremy E. Wulff,¹ and Fraser Hof^{1,*}

¹ Department of Chemistry, University of Victoria, PO Box 3065, Victoria, BC, V8W 3V6 Canada

² Department of Biochemistry and Microbiology, University of Victoria, PO Box 3055, Victoria, BC, V8W 3V6 Canada

³ Centre for Drug Research and Development, 2405 Wesbrook Mall, Vancouver, BC, V6T 1Z3 Canada

* To whom correspondence should be addressed: fhof@uvic.ca

Table of Contents

Section	Pages
Experimental procedures for peptide synthesis	S2-S3
NMR spectra for Figure 3 compounds	S4-S7
Sequence information and alignments, Phylogenetic tree	S8
X-ray crystallographic data	S9
2D-NMR studies	S10
Structural comparison of human chromodomain structures	S11
Isothermal titration calorimetry data	S12-S14

Experimental procedures for peptide synthesis

Peptide Synthesis — C-terminal Amides

Fmoc-Lys(Me)₃-OH was purchased from PL Labs. All other natural and non-natural amino acids were purchased in their enantiopure, Fmoc-protected forms from Chemimpex. Carboxylic acids or sulfonyl chloride for N-terminal modifications were supplied from TCI America, Chemimpex, Sigma-Aldrich or Matrix Scientific.

Peptide synthesis was performed using standard Fmoc solid phase synthesis protocols. Briefly, Fmoc-protected Rink Amide resin (Chemimpex) was swollen in CH₂Cl₂ for 30 minutes, deprotected by treatment with 20% piperidine in DMF (x3), and washed with DMF. N-Fmoc and side-chain protected amino acids (5 eq.) were mixed with HBTU (4.9 eq.) and DIEA (10 eq.) in DMF (3 mL), loaded onto the resin, and shaken at room temperature for 1 hour. Fmoc deprotection was performed using 20% piperidine in DMF (3 x 10 minutes). N-terminal acetylation was achieved by treatment with a 30:20:50 pyridine/Ac₂O/CH₂Cl₂ mixture for 2 hours (other N-terminal modifications are reported below). The peptide was cleaved from the resin by stirring with a 95:2.5:2.5 solution of TFA/H₂O/triisopropylsilane over 2 hours, and the resin was washed with additional TFA (3 x 5 mL). This TFA solution was concentrated *in vacuo* and the crude peptide precipitated with the addition of cold diethyl ether. After centrifugation and decanting, crude peptides were dried then purified by preparative RP-HPLC (Alltech Apollo or Phenomenex Luna columns; C18, 10 μm particle size, 250 mm x 22 mm) using a gradient of 0.1% TFA in H₂O and 0.1% TFA in MeCN (90:10 to 10:90, H₂O:MeCN). Product purities and identities were determined by LC-MS, with purities >95% for all samples. Masses were subsequently re-confirmed by separate LR-ESI-MS measurements (mass agreement ≤0.5 amu). Select compounds **61–64** were studied by ¹H NMR (see main text for data, and see below for spectra).

Peptide synthesis — C-terminal Acids

As above, except that peptides were prepared as C-terminal acids by initial synthesis on a 2-Cl-Trt resin (Chemimpex). Coupling of the first amino acid (Fmoc-Lys(Me)₃-OH, 0.6 eq.) was done in 95:5 CH₂Cl₂:DMF, DIEA (2.4 eq.) as a base and stirred at ambient temperature for 2.5 hours. Further amino acids attachments, deprotection, and acetylation on N-terminus were performed as above. After N-termini acetylation, the resin cleavage from the synthesized peptide was carried out by shaking with 1%TFA in dichloromethane for 10 min (3x10ml) and resin was with DCM (3x10ml). This DCM solution was concentration under vacuo and the crude peptide was purified by reverse phase HPLC as described above and product identities were confirmed by ESI-MS.

Synthesis of C-terminal modifications

Isolated peptide carboxylic acid (1eq), HCTU (3eq), Amine (3eq), and DIEA (10eq) were taken in DMF and reaction mixture was stirred at 70°C overnight under nitrogen atmosphere. Reaction mixture was diluted with ethyl acetate and water, extracted with ethyl acetate (3x 4ml) to remove unreacted amine and other impurities. Aqueous layer was distilled off and obtained crude product was purified by preparative reverse phase-HPLC using linear 0.1% TFA in water/Acetonitrile gradient. The product identities were confirmed by ESI-MS.

Synthesis of N-terminal modifications

Peptides were synthesized as outlined above. Treatment with 20% piperidine in DMF, followed by washing with DMF (3 x 10 mL) and CH₂Cl₂ (2 x 10 mL) provided the N-terminal deprotected peptide on resin.

Amide-linked N-terminal groups. R-CO₂H (5 eq.), HBTU (5 eq.) and DIEA (10 eq.) were dissolved in 5 mL DMF and added to the resin bearing N-terminal deprotected peptide. The resin and coupling solution were stirred overnight, then drained and washed with DMF (3 x 10 mL) and DCM (2 x 10 mL). Cleavage from resin was done as outlined above.

Sulfonamide linked N-terminal groups. R-SO₂Cl (5 eq.) and pyridine (10 eq.) were dissolved in 5 mL DMF and added to the resin bearing N-terminal deprotected peptide. The resin and coupling solution were stirred overnight, then drained and washed with DMF (3 x 10 mL) and DCM (2 x 10 mL). Cleavage from resin was done as outlined above.

Carbamate linked N-terminal group. Benzyl chloroformate (5 eq.) and Et₃N (10 eq.) were dissolved in 5 mL DMF and added to the resin bearing N-terminal deprotected peptide. The resin and coupling solution were stirred overnight, then drained and washed with DMF (3 x 10 mL) and DCM (2 x 10 mL). Cleavage from resin was done as outlined above.

Alkyl linked N-terminal groups. (Compound **52**) R-CHO (5 eq.) and sodium triacetoxyborohydride (5 eq.) in 10 mL DMF (with 1 drop of glacial AcOH) was added to the resin bearing N-terminal deprotected peptide, and stirred overnight. Resin was drained and washed with DMF (3 x 10 mL) and DCM (2 x 10 mL). Cleavage from resin was done as outlined above. (Compound **53**) R-CH₂Br (5 eq.) and pyridine (5 eq.) in 5 mL DMF was added to resin bearing N-terminal deprotected peptide. The resin and coupling solution were stirred overnight, then drained and washed with DMF (3 x 10 mL) and DCM (2 x 10 mL). Cleavage from resin was done as outlined above.

NMR spectra for Figure 3 compounds (500 MHz)

¹H NMR spectrum of compound **61**.

COMPOUND-62

^1H NMR spectrum of compound **62**.

¹H NMR spectrum of compound **63**.

¹H NMR spectrum of compound **64**.

Sequence information and alignments

10 20 30 40 50
 GEQVF AVESI RKKRV RKGKV EYLVK WKGWP PKYST WEPEE HILDV RLVMA YEEKE E

Figure S1. Sequence and numbering of the CBX7 chromodomain construct used for all experiments

A)

CBX1	sp	P83916	21-79	YVVEKVLDRR	VVKGKVEYLL	KWKGFSDEDN	TWEPEENLD	CPDLIAEFLO	SQKTAHETDK	71%
CBX2	sp	Q14781	12-70	FAAECILSKR	LRKGKLEYLV	KWRGWSSKHN	SWEPEENIL	DPRLLLAFQK	KEHEKEVQNR	88%
CBX3	sp	Q13185	30-88	FVVEKVLDRR	VVNGKVEYFL	KWKGFTDADN	TWEPEENLD	CPELIEAFLN	SQKAGKEKDG	62%
CBX4	sp	O00257	11-69	FAVESIEKKR	IRKGRVEYLV	KWRGWSPKYN	TWEPEENIL	DPRLLIQFQN	RERQEQLMGY	90%
CBX5	sp	P45973	20-78	YVVEKVLDRR	VVKGQVEYLL	KWKGFSSEHN	TWEPEKNLD	CPELISEFMK	KYKKMKEGEN	74%
CBX6	sp	O95503	11-69	FAAESIIKRR	IRKGRIEYLV	KWKGWAIKYS	TWEPEENIL	DSRLIAAFEQ	KEREREYLYP	80%
CBX7	sp	O95931	11-69	FAVESIRKKR	VRKGKVEYLV	KWKGWPPKYS	TWEPEEHIL	DPRLVMAYEE	KEERDRASGY	—
CBX8	sp	Q9HC52	11-69	FAAEALLKRR	IRKGRMEYLV	KWKGWSQKYS	TWEPEENIL	DARLLAAFEQ	REREMELYGP	88%
				:. . *	: . : *	: : : : : *	: : : : : *	: : : : : *	: : : : : *	

B)

CBX4	sp	O00257	11-69	FAVESIEKKR	IRKGRVEYLV	KWRGWSPKYN	TWEPEENIL	DPRLLIQFQN	RERQEQLMGY	90%
CBX7	sp	O95931	11-69	FAVESIRKKR	VRKGKVEYLV	KWKGWPPKYS	TWEPEEHIL	DPRLVMAYEE	KEERDRASGY	—
				*****	***	:***:*****	**:*:*:*:	*****:*	*****:***	:*::**

Figure S2. A) Aligned sequences of all human CBX chromodomains (ClustalW2 alignment conducted using sequences as reported in Uniprot database). The red bar highlights the unique presence of the metal-binding His residue in CBX7 that is Asn in all other CBX proteins. B) ClustalW2 sequence alignment of only CBX4 and CBX7 chromodomains highlight their extreme similarity. An * (asterisk) indicates positions which have a single, fully conserved residue. A : (colon) and . (period) indicate conservation between groups of strongly and weakly similar properties, respectively. Percent values at far right of each sequence are EMBL EMBOSS Matcher pairwise similarity scores for each chromodomain sequence relative to CBX7.

Figure S3. Phylogenetic tree for human chromodomains, as obtained from CHROMOHUB, the Structural Genomics Consortium: Liu, L.; Zhen, X. T.; Denton, E.; Marsden, B. D.; Schapira, M. *Bioinformatics* **2012**, 28, 2205

X-ray crystallographic data

Table S1. Data collection and refinement statistics.

	CBX7
Data collection	
Spacegroup	C 1 2 1
a, b, c (Å)	63.29, 31.86, 38.39
α , β , γ (deg.)	90, 119.02, 90
Wavelength (Å)	0.9786
Measured reflections	21792
Unique reflections	9357
Redundancy	2.3 (2.2)
Completeness (%)	94.4 (96.8)
$I/\sigma(I)$	6.0 (2.8)
R_{merge}^a	0.082 (0.308)
Refinement Statistics	
Resolution range (Å)	29.50 – 1.54 (1.63-1.54)
R_{work}^b	0.172 (0.231)
R_{free}^c	0.212 (0.280)
No. of atoms	
Protein (CBX7)	481
Peptide	50
Solvent	40
EDO	12
Magnesium	1
PEG	13
B-values (Å ²)	
Protein (CBX7)	22.23
Peptide	16.75
Solvent	30.84
EDO	37.36
Magnesium	8.27
PEG	36.51
r.m.s. deviation from ideality	
Bond lengths (Å)	0.02
Bond angles (deg.)	2.13
Ramachandran statistics (%)	
Most favoured	98.1
Allowed	1.9
Disallowed	0.0
Values in parentheses are for the highest resolution shell	
^a $R_{\text{merge}} = \sum_{hkl} \sum_i I_{hkl,i} - [I_{hkl}] / \sum_{hkl} \sum_i I_{hkl,i}$, where $[I_{hkl}]$ is the	
is the average of symmetry related observations of a unique reflection	
^b $R_{\text{work}} = \sum F_{\text{obs}} - F_{\text{calc}} / \sum F_{\text{obs}}$, where F_{obs} and F_{calc} are the observed and the	
calculated structure factors, respectively	
^c R_{free} is R using 5% of reflections randomly chosen and omitted from	
refinement	

2D-NMR studies

Experimental Procedure

Globally ^{15}N -labeled CBX7 chromodomain was expressed and purified exactly as described for unlabeled protein, except that *E. coli* were grown in minimal media containing $^{15}\text{NH}_4\text{Cl}$. NMR Samples were prepared in the same buffer used for FP experiments (see above), with 10% D_2O added as a lock reference. ^1H , ^{15}N HSQC spectra were collected for samples of ^{15}N -labeled CBX7 chromodomain (1 mM) containing 0, 0.5, 1.0, 2.0, and 5.0 equivalents of Ac-FAYKme3S- NH_2 . Peaks were assigned using the published assignments for free CBX7 (BMRB entry 15674) and CBX7 saturated with H3K27me2 peptide (BMRB entry 16778). Chemical shift perturbations (CSPs) were observed upon addition of ligand for both backbone and side-chain ^1H , ^{15}N HSQC resonances and were grouped into large, medium, and small chemical shifts as described in the manuscript. Shifts were color-mapped onto the X-ray structure for the complex of the same ligand as shown in the Figure, below.

Figure S4. 2D-NMR Chemical Shift Perturbation (CSP) data support the binding mode observed in the solid state. The X-ray structure of the complex between CBX7 and Ac-FAYKme3S- NH_2 is colored according to the degree of CSP upon binding of ligand in solution, as determined by ^1H , ^{15}N HSQC. Red = large CSP; orange = medium CSP; yellow = small CSP; blue = no CSP or peaks unassigned.

Comparison of human chromodomain structures

Polycomb homologs — hydrophobic clasp (HC) motif

HP1 homologs — polar clasp (PC) motif

Figure S5. Comparison of Polycomb and HP1 family chromodomain structures with a focus on hydrophobic clasp (HC) and polar clasp (PC) motifs. Top: All known ligand-bound structures of Polycomb homologs show an intact hydrophobic clasp. Only the co-structure determined here (top left) shows the additional Mg^{2+} -mediated Glu-His bridge that significantly influences the structure of the ligand binding pocket. Bottom: the ligand-bound structures of HP1 homologs (incl. close relative MPP8) show a variable degree of engagement of polar clasp residues, and also show generally open binding sites. Surfaces are colored by atom type: green = carbon, red = oxygen, blue = nitrogen. The regions corresponding to the Glu-His- Mg^{2+} bridge are open and highlighted with a pink arrow in all other structures.

Exemplary Isothermal Titration Calorimetry Data

Figure S6. ITC data for compound **61** binding to CBX7 (left) and CBX8 (right).

Figure S7. ITC data for compound **62** binding to CBX7 (left) and CBX8 (right).

Figure S8. ITC data for compound **63** binding to CBX7 (left) and CBX8 (right).

Figure S9. ITC data for compound **64** binding to CBX7 (left) and CBX8 (right).

Table S2. Fitted data arising from ITC titration of compounds into CBX7 and CBX8.^a

Compound	Protein	N	K _d (μM)	ΔH (kJ mol ⁻¹)	-TΔS (kJ mol ⁻¹)	CBX7 Selectivity
1	CBX7	1.37	2.0 ± 0.20	-32.9 ± 0.44	0.39	7.1
	CBX8	1.06	14.2 ± 1.97	-10.5 ± 0.39	-17.1	
11	CBX7	1.05	1.77 ± 0.12	-37.2 ± 0.33	-4.42	6.9
	CBX8	1.03	12.3 ± 1.44	-12.3 ± 0.36	-15.7	
23	CBX7	1.01	5.36 ± 0.90	-21.2 ± 0.90	-8.88	2.8
	CBX8	1.03	15.0 ± 2.81	-5.72 ± 0.49	-21.7	
54	CBX7	0.947	0.281 ± 0.05	-35.1 ± 0.41	-2.28	5.4
	CBX8	0.957	1.52 ± 0.26	-9.94 ± 0.21	-23.3	
61	CBX7	0.996	4.10 ± 0.18	-26.1 ± 0.19	-4.60	3.7
	CBX8	1.055	15.0 ± 2.97	-6.6 ± 0.48	-20.7	
62	CBX7	0.867	0.216 ± 0.02	-47.3 ± 2.4	9.5	2.2
	CBX8	1.33	0.481 ± 0.04	-10.4 ± 0.06	-25.7	
63	CBX7	1.29	0.198 ± 0.03	-20.1 ± 0.18	-18.0	9.0
	CBX8	0.949	1.78 ± 0.25	-9.85 ± 0.17	-22.8	
64	CBX7	1.35	0.197 ± 0.04	-27.2 ± 0.22	-13.4	9.6
	CBX8	0.995	1.89 ± 0.16	-11.3 ± 0.12	-21.2	

a) Averaged values determined by duplicate ITC titrations at 298 K in buffered H₂O (20 mM TRIS, 250 mM NaCl, 1 mM DTT, pH 8).

Table S3. Fitted data arising from ITC titration of compounds into CBX1 and CBX4.^a

Compound	Protein	N	K _d (μM)	ΔH (kJ mol ⁻¹)	-TΔS (kJ mol ⁻¹)
1	CBX1	1.09	43 ± 2.8	-39.8 ± 1.9	14.8
	CBX4	1.13	3.7 ± 0.46	-8.9 ± 0.18	-22.1
64	CBX1	1	>500 ^b	n.a. ^b	n.a. ^b
	CBX4	1.01	0.287 ± 0.02	-23.2 ± 0.25	-14.1

a) Averaged values determined by duplicate ITC titrations at 298 K in buffered H₂O (20 mM TRIS, 250 mM NaCl, 1 mM DTT, pH 8). b) No binding is evident at the concentrations tested, and we set 500 μM as a lower limit for K_d. It is impossible to derive thermodynamic parameters from such a weak binding event.