Supporting Information for:

Nucleation and propagation of heterochromatin by the histone methyltransferase PRC2: geometric constraints and impact of the regulatory subunit JARID2

Eva J. Ge[1], Krupa S. Jani[1], Katharine L. Diehl[1], Manuel M. Müller[1,2], and Tom W. Muir[1*]

Affiliations:

[1]Department of Chemistry, Princeton University, Princeton, NJ 08544, United States.

[2]Present Address: Department of Chemistry, King's College London, London SE1 1DB, UK

Correspondence to: muir@princeton.edu.

Supporting Information Table of Contents

Methods

General laboratory materials and methods

Commonly used chemical reagents and solvents including Dimethylformamide (DMF), dichloromethane (DCM), and triisopropylsilane (TIS) were purchased from Sigma-Aldrich or Fisher Scientific and used without further purification. Trifluoroacetic acid (TFA) was purchased from Halocarbon (North Augusta, SC). Analytical reversed-phase HPLC (RP-HPLC) was performed on an Agilent 1100 or 1260 series instrument employing a Vydac C18 column (5 μm, 4 × 150 mm) at a flow rate of 1 mL/min. Semipreparative scale purifications were performed on an Agilent 1260 series instrument employing a Vydac C18 semipreparative column (12 μm, 10 mm × 250 mm) at a flow rate of 4 mL/min. Preparative RP-HPLC was performed on a Waters prep LC system comprised of a Waters 2545 Binary Gradient Module and a Waters 2489 UV detector. Purifications were carried out on a C18 Vydac 218TP1022 column (10 μM; 22 x 250 mm) at a flow rate of 18 mL/min. For all RP-HPLC purifications and analyses, 0.1% TFA in water (HPLC solvent A) and 90% acetonitrile, 0.1% TFA in water (HPLC solvent B) were used as the mobile phases. Electrospray ionization mass spectrometry (ESI-MS) analysis was conducted on a MicroTOF-Q II ESI-Qq-TOF mass spectrometer (Bruker Daltonics).

Cloning

General cloning:

Restriction enzymes were obtained from New England BioLabs. PCR amplification reactions were performed using the PfuUltra II Fusion HS DNA Polymerase (Agilent Technologies) according to manufacturer's protocols. Primers were ordered from Integrated DNA Technologies (IDT, Coralville, IA) or Sigma Aldrich. PCR purification kits were obtained from Qiagen. All plasmids were sequence verified by GENEWIZ (South Plainfield, NJ).

Histone cloning:

The H3 K27R point mutant was generated by Quik Change Site-Directed Mutagenesis (Agilent) using the following primers: CGCCGCTCTGCGCCTGCGACCGGTGG and AGAGCGGCGTGCCGCTTTGGTCGCCAGC.

Modified tJARID2 cloning:

Truncated tJARID2(108-450) and tJARID2(137C-450)L137C were PCR amplified out of an existing construct containing full-length JARID2 and inserted into a vector containing a His-SUMO sequence using Gibson Assembly with the following primers:

Primers for tJARID2(108-450) amplification:

CGCCCGCGCCTGCAGGCTCAGCGTAAGTTC
TTATTGCAACCCTTCGCGTGCTTGATGCGC

Primers for tJARID2(137C-450)L137C amplification:

TGTCCCCCTCCCGCAACTCAGATTAGCGACTTG
TTATTGCAACCCTTCGCGTGCTTGATGCGC

Primers for Gibson Assembly of tJARID2(108-450):

ACCACCAATCTGTTCTCTG
acagagaacagattggtggtCGCCCGCGCCTGCAGGCT
tgtcgacggagctcgaattcTTATTGCAACCCTTCGCGTGCTTGATGCGC
GAATTCGAGCTCCGTCGAC

Primers for Gibson Assembly of tJARID2(137C-450)L137C:

ACCACCAATCTGTTCTCTG
acagagaacagattggtggtTGTCCCCCTCCCGCAACT
tgtcgacggagctcgaattcTTATTGCAACCCTTCGCGTGC
GAATTCGAGCTCCGTCGAC

A Quik Change Site-Directed Mutagenesis (Agilent) cloning strategy was used to generate tJARID2 point mutants using the following primers:

tJARID2(108-450)L137C:

GATTGTCGAACCGCTTTGTCCCCCTCCCGC
AGTTGCGGGAGGGGGACAAAGCGGTTCGAC

tJARID2(108-450)K116R:

CAGCGTCGATTCGCTCAGTCTCAACCTAAC
CGAAGCTACGCTGAGCCTGCAGGCGCGGG

Solid phase peptide synthesis

Fmoc amino acids were purchased from Novabiochem (Darmstadt, Germany) or Bachem (Torrance, CA). Peptides were synthesized on ChemMatrix Trityl-OH PEG resin (0.41 mmol/g) purchased from Biotage (Charlotte, NC), either using manual addition of reagents (with a stream of dry N_2 to agitate the reaction) or on a Liberty Peptide Synthesizer equipped with a Discovery microwave module (CEM). For manually synthesized peptides, each cycle was: i) FMOC group deprotection with 20% piperidine in DMF (1x5 min, 1x10 min) and ii) coupling of 5 eq. of each amino acid with 4.9 eq. of HATU or HOBt/HBTU and 10 eq. DIPEA for 30 min. Double couplings were used for the first two amino acids and any special modified amino acids. Resin was thoroughly washed with DMF between steps. Finished peptides were cleaved from resin and deprotected with 95% TFA, 2.5% TIPS, 2.5% water for 3 hours at RT. Cleavage products were precipitated with cold diethyl ether, dissolved in HPLC Solvent A (0.1% TFA in water), and analyzed by RP-HPLC. Crude peptides were purified by RP-HPLC on a preparative scale. Fractions were analyzed by ESI-MS, and pure fractions were pooled, lyophilized, and stored at -20ºC.

Peptide sequences

<u>Peptides for histone semisynthesis:</u>

H3(1-28)K27me1: ARTKQTARKSTGGKAPRKQLATKAARK(me1)S
H3(1-28)K27me2: ARTKQTARKSTGGKAPRKQLATKAARK(me2)S
H3(1-28)K27me3: ARTKQTARKSTGGKAPRKQLATKAARK(me2)S
H3(1-14)K4me3: ARTK(me3)QTARKSTGGK

<u>Peptides for JARID2 semisynthesis:</u>

*JARID2(108-117)K116me2: Ac-RPRLQAQRK(me2)F
*JARID2(108-117)K116me3: Ac-RPRLQAQRK(me3)F
JARID2(118-136)pSer120: CQS(phos)QPNSPSTTPVKIVEPL
JARID2(118-136)pSer124: CQSQPNS(phos)PSTTPVKIVEPL
JARID2(118-136)pSer126: CQSQPNSPS(phos)TTPVKIVEPL
*JARID2(108-136)K116me3: Ac-RPRLQAQRK(me3)FAQSQPNSPSTTPVKIVEPL
*JARID2(108-136)pSer120: Ac-RPRLQAQRKFAQS(phos)QPNSPSTTPVKIVEPL
*JARID2(108-136)pSer124: Ac-RPRLQAQRKFAQSQPNS(phos)PSTTPVKIVEPL

*JARID2(108-136)pSer126: Ac-RPRLQAQRKFAQSQPNSPS(phos)TTPVKIVEPL
(*)Denotes JARID2 peptides that are N-terminally acetylated

Purification of histone proteins

Wild type or mutant histone expression plasmids were transformed into Rosetta(DE3) cells and grown at 37ºC to OD 600 = 0.6. Protein expression was induced with 1 mM IPTG for 3 hours. Cells were lysed by French press in lysis buffer containing 40 mM Tris, pH 7.5, 300 mM NaCl, 1 mM EDTA, 1 mM DTT, and 0.5 mM PMSF. The insoluble fraction was isolated by centrifugation at 17,000 RCF for 30 minutes. The insoluble fraction was washed twice with lysis buffer containing 1% Triton X-100, and then once more with lysis buffer without Triton X-100. Histone proteins were extracted from the insoluble fraction by an extraction buffer containing 40 mM Tris pH 7.5, 300 mM NaCl, 6 M guanidine hydrochloride, and 1 mM DTT for 3 hours at room temperature, and then centrifuged at 17,000 RCF for 30 minutes. The soluble fraction was passed through a 0.45 μM filter (Millipore), injected onto a preparative RP-HPLC column, and eluted over a 20-80% solvent B gradient in 45 minutes. Fractions were analyzed by ESI-MS, and pure fractions were pooled, lyophilized, and stored at -80°C.

Truncated histone constructs 6xHis-Sumo-H3 (29-135, A29C) or (15-135, A15C) were expressed and extracted as described above. After extraction, the histone was immobilized on Ni-affinity resin, washed with wash buffer containing 50 mM Tris, 7.5, 300 mM NaCl, 6 M urea, 20 mM imidazole, 1 mM DTT, and eluted with wash buffer supplemented with 300 mM imidazole. Eluted protein was dialyzed for 4 hours at 4ºC into dialysis buffer (50 mM Tris, pH 7.5, 300 mM NaCl, 1.5 M urea, 1 mM DTT), and then once more for 16 hours in the presence of Ulp1 to cleave the Sumo tag. Following dialysis, solid urea was added to dissolve any precipitated solids, and the sample was purified by RP-HPLC as described above.

Purification and refolding of truncated JARID2 proteins

Truncated JARID2 (108-450) sequence:
RPRLQAQRKFAQSQPNSPSTTPVKIVEPLLPPPATQISDLSKRKPKTEDFLTFLCLRGSPAL
PNSMVYFGSSQDEEEVEEEDDETEDVKTATNNASSSCQSTPRKGKTHKHVHNGHVFNG
SSRSTREKEPVQKHKSKEATPAKEKHSDHRADSRREQASANHPAAAPSTGSSAKGLAAT
HHHPPLHRSAQDLRKQVSKVNGVTRMSSLGAGVTSAKKMREVRPSPSKTVKYTATVT
KGAVTYTKAKRELVKDTKPNHHKPSSAVNHTISGKTESSNAKTRKQVLSLGGASKSTGP
AVNGLKVSGRLNPKSCTKEVGGRQLREGLQLREGLRNSKRRLEEAHQA

Wild type or mutant His-SUMO-tJARID2 expression plasmids were transformed into Rosetta(DE3) cells and grown at 37ºC to OD600 = 0.6. Protein expression was induced with 1 mM IPTG for 3 hours. Cells were lysed by French press in lysis buffer containing 40 mM Tris pH 7.5, 750 mM NaCl, 1 mM DTT, and 0.5 mM PMSF. The soluble fraction was isolated by centrifugation at 17,000 RCF for 30 minutes. The protein was immobilized on Ni-affinity resin, washed with wash buffer containing 40 mM Tris pH 7.5, 750 mM NaCl, 20 mM imidazole, and 1 mM DTT, and eluted with wash buffer supplemented with 300 mM imidazole. Eluted protein was dialyzed for 2 hours at 4ºC into dialysis buffer (50 mM Tris pH 7.5, 350 mM NaCl, 1 mM DTT), and then once more for 16 hours in the presence of Ulp1 to cleave the Sumo tag. Following dialysis, the sample was purified by RP-HPLC using a 10-70% solvent B gradient over 45 minutes. Fractions were analyzed by ESI-MS, and pure fractions were pooled, lyophilized, and stored at -80 °C.

For refolding, 1 mg of protein was dissolved in 500 uL of unfolding buffer containing 40 mM HEPES pH 7.4, 7 M urea, and 1 mM DTT. The sample was loaded into a 0.1-0.5mL 10,000 MWCO Slide-A-Lyzer dialysis cassette (Thermo Fisher Scientific) and dialyzed into 1 L of HEGN350 (40 mM HEPES pH 7.4, 350 mM NaCl, 0.1% NP-40, 10% glycerol, 3 mM $MgCl_2$) for 4 hours at 4ºC, and then twice more for 4 hours and then 16 hours. The protein sample was then recovered from the cassette, spun down at 17,000 RCF for 10 minutes, and the final concentration was quantified by Bradford Assay using the Pierce Coomassie Plus (Bradford) Assay Kit (Thermo Scientific). The protein was then aliquoted, flash frozen, and stored at -80 °C.

Protein semisynthesis

Histone semisynthesis:

Human H3.1K27me3 was prepared by protein semisynthesis. A peptide containing residues 1-28 of the protein, including K27me3, was synthesized on solid phase with a C-terminal hydrazine. This handle enabled conversion of the peptide into a thioester for native chemical ligation[1]. K27me3 was incorporated with FMOC-Lysine(me3,Cl)-OH. Peptide synthesis procedures were as described above.

Truncated H3.1 protein (residues 29-135) containing A29C, C96A, and C110A mutations as well as a SUMO fusion was produced in *E. coli* as described above. H3(1-28)K27me3 peptide (10mg, 16 eq.) was dissolved in degassed acyl hydrazide oxidation buffer (100 mM K_2HPO_4/KH_2PO_4 pH 3.0, 6 M guanidine HCl, 75 mM $NaNO_2$) and reacted in a -15°C ice slurry for 20 min. Truncated H3 (5.5 mg, 1 eq.) was dissolved in degassed ligation buffer (200 mM K_2HPO_4/KH_2PO_4 pH 7.5, 6 M guanidine HCl) with 2-3% v/v 2,2,2-trifluoroethanethiol (TFET), and added to the peptide[2]. The reaction was carefully pH-adjusted to pH 7.5, and allowed to proceed at RT overnight. Reaction progress was checked by analytical RP-HPLC. Following, the ligation, $NaNO_2$ was removed by dialysis into ligation buffer for 6h.

Stock solutions of TCEP, glutathione, and the radical initiator VA-061 were added to the ligation mixture for final concentrations of 5 M guanidine HCl, 250 mM TCEP, 100 mM K_2HPO_4/KH_2PO_4, 20 mM VA-061, 40 mM glutathione. The desulfurization reaction was allowed to proceed at 37ºC for 16 hours. Reaction progress was monitored by analytical RP-HPLC. Desulfurized histone product was purified by RP-HPLC (40-70% HPLC solvent B gradient over 45 minutes). Fractions were analyzed by ESI-MS. Pure fractions were pooled, lyophilized, and stored at -20ºC.

Truncated JARID2 semisynthesis:

A two-piece ligation strategy was employed to generate tJARID2-K116me2, tJARID2-K116me3, tJARID2-pS120, tJARID2-pS124, and tJARID2-pS126. For the ligation, modified JARID2(108-136) thioester (4 mg, 8.4 eq) was dissolved in degassed acyl hydrazide oxidation buffer (100 mM K_2HPO_4/KH_2PO_4 pH 3.0, 6 M guanidine HCl, 75 mM $NaNO_2$) and reacted in a -15°C ice slurry for 20 min. Truncated JARID2(137C-450) protein (4.1 mg, 1 eq.) was dissolved in degassed ligation buffer (200 mM K_2HPO_4/KH_2PO_4 pH 7.5, 6 M guanidine HCl), and mixed with the oxidized peptide. 100 equivalents of solid 4-mercaptophenylacetic acid (MPAA) were added, and the pH was adjusted to 7.0. The reaction was allowed to proceed at room temperature for 60 min, and then TCEP was added to a final concentration of 20 mM. The reaction was then allowed to proceed at RT overnight. Reaction progress was checked by analytical RP-HPLC. The final reaction product was purified by semipreparative RP-HPLC (20-55% HPLC solvent B gradient

over 40 minutes). Fractions were analyzed by ESI-MS. Pure fractions were pooled, lyophilized, and stored at -20ºC.

A three-piece ligation strategy was employed to generate tJARID2-K116me3-pS120. Following SPPS and cleavage from resin, crude JARID2(108-117)K116me3 hydrazide peptide was oxidized by dissolving in acyl hydrazide oxidation buffer (100 mM K_2HPO_4/KH_2PO_4 pH 3.0, 6 M guanidine HCl, 15 mM $NaNO_2$) and reacting in a -15°C ice slurry for 20 min. After 20 min, solid sodium 2-mercaptoethanesulfonate (MESNA) was added to a final concentration, the reaction was pH adjusted to 7.0, and allowed to proceed at room temperature for 30 minutes. Crude peptide was purified by RP-HPLC on a preparative scale. Fractions were analyzed by ESI-MS, and pure fractions were pooled, lyophilized, and stored at -20ºC until the first ligation step. For the first ligation step, the N-terminal JARID2(108-117)K116me3 thioester (5.5 mg, 1.5 eq.) and the C-terminal JARID2(118-136)pSer120 were dissolved together in degassed ligation buffer (200 mM K_2HPO_4/KH_2PO_4 pH 7.5, 6 M guanidine HCl) with 5% v/v TFET and 20 mM TCEP, and the reaction was adjusted to pH 7.0. Reaction progress was checked by analytical RP-HPLC. Following the first ligation, the product was desulfurized. Stock solutions of TCEP, glutathione, and the radical initiator VA-061 were added to the ligation mixture for final concentrations of 5 M guanidine HCl, 250 mM TCEP, 100 mM K_2HPO_4/KH_2PO_4, 20 mM VA-061, 40 mM glutathione. The desulfurization reaction was allowed to proceed at 37ºC for 16 hours. Reaction progress was monitored by analytical RP-HPLC. The desulfurized reaction product was purified by semipreparative RP-HPLC (0-50% HPLC solvent B gradient over 40 minutes). Fractions were analyzed by ESI-MS. Pure fractions were pooled, lyophilized, and stored at -20ºC until the second ligation step. The peptide product obtained from the first ligation was then used for a second ligation following a procedure analogous to the two-piece ligation strategy described above.

Histone octamer preparation

Histone octamers were prepared as previously published[3]. In brief, purified unmodified and semi-synthetic histones were dissolved in histone unfolding buffer (6 M guanidinium HCl, 20 mM Tris, pH 7.5, 10 mM DTT, and combined (1.1 eq. H2A, 1.1 eq. H2B, 1.0 eq H3 variant, 1.0 eq H4), and total histone concentration was adjusted to 1 mg/mL. Octamers were assembled by dialysis (3 x 1 L, 4°C) into octamer refolding buffer (10 mM Tris, pH 7.5, 2 M NaCl, 0.5 mM EDTA, 1 mM DTT). The mixture was then concentrated on a spin concentrated, and purified by gel filtration on a Superdex 200 Increase 10/300 column. Fractions were analyzed by SDS-PAGE and using the FPLC elution chromatogram, and fractions containing octamers were combined, concentrated to 30-60 µM as quantified by A 280 (extinction coefficient = 44,700), diluted with glycerol to a final 50% v/v. and stored at -20ºC.

Expression and purification of PRC2 complexes.

All PRC2 proteins used in this study were prepared in Sf9 cells using a baculovirus system. Flag-tagged EZH2, and His-tagged EED, RBBP4, and SUZ12 were cloned into a pACEBac1 vector using the Multibac system[4]. His-AEBP2 and Flag-JARID2 were prepared in their own individual pACEBac1 vectors.

Plasmids were used to generate bacmids according to manufacturer protocol (Multibac, Geneva Biotech). 2.5 ug bacmid was transfected into $1x10^6$ attached Sf9 cells in a 6-well plate. Following transfection, cells were overlaid with 2 mL fresh medium (Sf-900III SFM, Thermo Fisher Scientific) and incubated at 27°C for 96 hours in the dark. The supernatant was collected, filtered through 0.2 µM, and supplemented with 2% FBS v/v to produce the P1 virus. P2 virus was

generated by infection of 10 mL of Sf9 cells (1.5 x 10^6 cells/mL) with 1 mL of P1 virus. Cells were grown at 27°C in suspension culture until they reached <50% viability as monitored by trypan blue staining. Culture supernatant was collected, filtered through 0.22 μm, and supplemented with 2% v/v FBS. To generate P3 virus, 300 uL of P2 virus was added to 50mL of Sf9 cells (1.5 x 10^6 cells/mL). Cells were grown at 27°C in suspension culture until they reached <50% viability as monitored by trypan blue staining. Culture supernatant was collected, filtered through 0.22 μm, and supplemented with 2% v/v FBS. During all steps of viral amplification, Sf9 cell density was maintained at 1.5-2 x 10^6 cells/mL, diluting if needed until growth arrested.

For protein expression, a 1:100 dilution of P3 virus was added to an Sf9 cell culture at 2.0x10^6 cells/mL density. After 48 h of incubation at 28°C in the dark, cells were harvested by centrifugation. Cells were lysed using a Dounce homogenizer (Wheaton) in HEGN600 buffer (25 mM HEPES pH7.0, 600 mM NaCl, 1 mM EDTA, 10% glycerol, 0.02% NP-40), and lysate was clarified by centrifugation. Soluble extracts were incubated with anti-Flag M2 affinity gel (100 uL resin per 100mL cell culture) in HEGN 350 (25 mM HEPES pH7.0, 350 mM NaCl, 1 mM EDTA, 10% glycerol, 0.02% NP-40) for 2 hours at 4°C. Resin was collected by centrifugation and washed three times with HEGN350. Bound proteins were eluted with HEGN350 containing 0.25 mg/mL Flag peptide (3 x 20 min at 4°C). Eluted proteins were pooled, spin concentrated in a VivaSpin centrifugal concentrator (MWCO 30,000, Viva Products), and purified by size exclusion chromatography on a Superose 6 column (GE Healthcare) on an ACKTA FPLC system from GE Healthcare equipped with a P-920 pump and a UPC-900 monitor. Fractions containing monomeric PRC2 (as analyzed by SDS-PAGE) were pooled, flash frozen with liquid N_2, and stored at -80°C. Prior to use, protein concentration was quantified by A280 (PRC2 core complex extinction coefficient = 296910, PRC2-JARID2 extinction coefficient = 394200, PRC2-JARID2-AEBP2 extinction coefficient = 429150).

Full-length JARID2 PTM analysis by mass spectrometry

Following Coomassie staining the band corresponding to the molecular weight of the product was excised and subjected to in-gel thiol reduction/alkylation and trypsin Gold (Promega) digestion overnight as previously published.[5] Samples were dried completely in a Speedvac and resuspended with 21 μL of 0.1% formic acid pH 3. 2 μL was injected per run using an Easy-nLC 1200 UPLC system. Samples were loaded directly onto a 45cm long 75um inner diameter nano capillary column packed with 1.9 μm C18-AQ (Dr. Maisch, Germany) mated to metal emitter in-line with an Orbitrap Fusion Lumos (Thermo Scientific, USA). The mass spectrometer was operated in data dependent mode with the 120,000 resolution MS1 scan (AGC 4e5, Max IT 50ms, 400-1500 m/z) in the Orbitrap followed by up to 20 MS/MS scans with CID fragmentation in the ion trap. Dynamic exclusion list was invoked to exclude previously sequenced peptides for 60s if sequenced within the last 30s and maximum cycle time of 3s was used. Peptide were isolated for fragmentation using the quadrupole (1.2 Da window). Ion-trap was operated in Rapid mode with AGC 2e3, maximum IT of 300ms and minimum of 5000 ions.

Raw files were searched using Byonic [6] and Sequest HT algorithms [7] within the Proteome Discoverer 2.1 suite (Thermo Scientific, USA). 10 ppm MS1 and 0.4 Da MS2 mass tolerances were specified. Caramidomethylation of cysteine was used as fixed modification, oxidation of methionine, acetylation of protein N-termni, conversion of glutamine to pyro-glutamate and deamidation of asparagine were specified as dynamic modifications. Trypsin digestion with maximum of 2 missed cleavages were allowed. Files searched against custom database of common contaminants and human histone proteins (124 entries total).

Scaffold (version Scaffold_4.8.4, Proteome Software Inc., Portland, OR) was used to validate MS/MS based peptide and protein identifications. Peptide identifications were accepted if they could be established at greater than 90.0% probability by the Scaffold Local FDR algorithm. Protein identifications were accepted if they could be established at greater than 99% probability and contained at least 2 identified peptides. Protein probabilities were assigned by the Protein Prophet algorithm [8]. Proteins that contained similar peptides and could not be differentiated based on MS/MS analysis alone were grouped to satisfy the principles of parsimony.

Skyline software (MacCoss Lab., University of Washington) was used to performed area based quantitation of mono-, di- and tri- methylation of Lysine.

DNA preparation

601 DNA sequence:

ACAGGATGTATATATCTGACACGTGCCTGGAGACTAGGGAGTAATCCCCTTGGCGGT
TAAAACGCGGGGGACAGCGCGTACGTGCGTTTAAGCGGTGCTAGAGCTGTCTACGA
CCAATTGAGCGGCCTCGGCACCGGGATTCTCCAG

MMTV DNA sequence:

ACTTGCAACAGTCCTAACATTCACCTCTTGTGTGTTTGTGTCTGTTCGCCATCCCGT
CTCCGCTCGTCACTTATCCTTCACTTTCCAGAGGGTCCCCCCGCAGACCCCGGCGA
CCCTGGTCGGCCGACTGCGGC ACAGTTTTTTG

Plasmid DNA containing 12 repeats of the 601 sequence (147-bp repeat with 30-bp linker) was purified from DH5alpha cell culture according to previously published procedures[3]. The 12x601 was excised from the plasmid using treatment with the EcoRV restriction enzyme (500 U/mL of EcoRV, plasmid concentration 1 mg/mL) at 37°C for 24 hours. Excised 12x601 was purified from plasmid backbone fragments using 6% PEG-6000 precipitation. Following purification, proteins and PEG were removed by phenol-chloroform extraction and ethanol precipitation. The remaining 12x601 DNA pellet was dissolved in 10 mM Tris pH 8.0, 0.1mM EDTA buffer, aliquoted, and stored at -20°C.

DNA containing tetrameric repeats of the 601 sequence (4x601) was prepared analogously to the 12x601 with an additional prep cell electrophoresis purification step using a Model 491 Prep Cell (Biorad). Yields of 3-6 mg of DNA were obtained per prep cell run (120V, 8 h elution).

A similar protocol with minor modifications was used to prepare the weak nucleosome binding mouse mammary tumor virus (MMTV) DNA sequence referred to as "buffer DNA." In the modified protocol, 7% PEG-6000 was used for purification, and the 155-bp MMTV DNA fragment remained in the supernatant while the plasmid backbone was separated into the pellet fraction.

Formation of homogeneous and scrambled nucleosome arrays

In a typical array assembly, FPLC-purified octamers were combined with 12x601 DNA in 70 µL buffer (2 M NaCl, 20 mM Tris, 0.1 mM EDTA, 1 mM DTT, pH 7.4 at 4°C). Mixtures were placed in Slide-A-Lyzer MINI dialysis buttons (3.5 kDa MW cutoff), and dialyzed at 4°C into array assembly start buffer (10 mM Tris, 1.4 M KCl, 0.1 mM EDTA, 1 mM DTT, pH 7.4 at 4°C) for 1 hour. Next, 330 mL of array assembly end buffer (10 mM Tris, 10 mM KCl, 0.1 mM EDTA, 1 mM DTT, pH 7.4 at 4°C) was slowly added using a peristaltic pump (1 mL/min), followed by two final 2-hour dialyses steps into array assembly end buffer. Assembly mixtures were transferred to

microcentrifuge tubes and centrifuged at 17,000 xg for 10 min at 4°C to remove precipitates. Arrays were purified by magnesium precipitation. $MgCl_2$ was added to a final concentration of 3.5mM, and samples were incubated on ice for 15 min and then centrifuged at 17,000 xg for 10 min at 4°C. The supernatant containing MMTV DNA and MMTV mononucleosomes was removed, and the pellet was re-dissolved in array assembly end buffer. Final array preparations were quantified by UV spectroscopy at 260 nM. Quality of array assemblies were assessed by native agarose polyacrylamide gel electrophoresis followed by SYBR Gold or ethidium bromide staining.

Formation of heterotypic nucleosome arrays by DNA ligation

Heterotypic arrays were assembled as previously described[9] by DNA ligation of mononucleosomes and/or homotypic 4-mer fragments with unique non-palindromic overhangs generated by AlwNI, BsaI, and DraIII. A one-pot ligation contains all three fragments (1.25 equivalents of outer fragments and 1.0 equivalent of middle fragment) and T4 DNA ligase (5 U/uL, NEB) in ligation buffer (70 mM Tris-HCl, pH 7.5, 6 mM $MgCl_2$, 1 mM ATP, and 10 mM DTT). Ligation reactions proceeded for 2 h at 16ºC, and then additional $MgCl_2$ was added to a final concentration of 10mM and the reaction was centrifuged at high speed to produce a pellet of predominantly 9-mers or 12-mers. The pellet was redissolved in TEK buffer (10 mM Tris pH 8.0, 0.1 mM EDTA, 10 mM KCl) and purified by sucrose gradient centrifugation (10-40% sucrose in TEK buffer, centrifugation at 35,000 xg for 4 hours). Sucrose gradient fractions were analyzed by agarose gel electrophoresis. Pure fragments were pooled, buffer exchanged, and concentrated in a VivaSpin centrifugal concentrator (Viva Products).

12mer array Mg^{2+} precipitation assay

Chromatin self-assembly was assayed as previously published[9]. Precipitation of 12mer arrays in HMT assay buffer with different concentrations of Mg^{2+} was assessed by incubating arrays in HMT assay buffer with varying concentrations of $MgCl_2$ (0-4 mM) for 10 min at 4 ºC. Precipitated arrays were pelleted by centrifugation (10 min at 17,000 g at 4 ºC), and the remaining chromatin in the supernatant fraction was quantified using SYBR Gold fluorescence (SpectraMax m3 plate reader, Molecular Devices, Sunnyvale, CA, at 495 nm excitation, 540 nm emission).

***In vitro* spreading assays for scintillation counting**

PRC2 methyltransferase activity was measured using a scintillation based assay. In a typical assay, arrays (480 nM 601 sites) were incubated with 15 nM PRC2 in 10 µL of HMT assay buffer (50 mM HEPES pH 8.0, 35 mM NaCl, 0.5 mM Mg^{2+}, 0.1% Tween-20, 5 mM DTT, 1 mM PMSF and 0.66 µM [^{3}H]SAM (Perkin Elmer). Where indicated, tJARID2 variants were pre-mixed with PRC2 core complex and added in a final concentration of 15 nM. Reactions were allowed to proceed for 40 min at 30°C, and were then quenched by spotting on Whatman P81 phosphocellulose filter paper (Sigma). Filters were dried for 45 min at RT, washed 3 x 15 minutes with 0.2 M $NaHCO_3$ pH 9.0, and dried for 45 min using a gel dryer at 40 ºC. Dried filters were placed in 1 mL Ultima Gold scintillation cocktail and scintillation counting was performed on a MicroBeta scintillation counter (Perkin Elmer). All counts were corrected for background using a reaction in the absence of enzyme.

Kinetic measurements

Kinetic measurements were obtained using the same scintillation based assays as described above. A substrate titration was performed with a fixed concentration of PRC2 (15 nM) in 10 µL of HMT assay buffer. Kinetic values were obtained by plotting data in GraphPad Prism 6.01, and applying non-linear regression analysis to determine individual apparent k_{cat} ($k_{cat, app}$) and $K_{m, substrate}$ values. In cases where the enzyme was not saturated, low substrate data points representing conditions of [S] << $K_{m, substrate}$ were used to plot initial rate vs. initial substrate condition, and the slope was used to determine $k_{cat, app}/K_{m,susbtrate}$.

PRC2 inhibition assays

PRC2 inhibition was determined by performing the scintillation assay described above in the presence of EED226 and EPZ6438 inhibitors. Counts were normalized to values in the absence of inhibitors, and corrected for background counts in the absence of chromatin substrate. All data was collected in triplicate, and relative IC_{50} values were determined from dose-response curves using the equation:

$$Y = \frac{1}{1 + {}^{[I]}/_{IC_{50}}}$$

Where Y is the relative PRC2 activity, and [I] is the inhibitor concentration.[10]

Western blot protocol

SDS-PAGE gels (12% Bis-Tris; Biorad) were transferred onto PVDF membranes by semi-dry transfer (20 V, 20 min), and blocked in 3% bovine serum albumin (BSA) in TBST (50 mM Tris, pH 7.5, 150 mM NaCl, 0.1% Tween-20) for 1 hour at room temperature. Membranes were then incubated in primary antibody in 3% BSA in TBST overnight at 4 ºC. Membranes were then washed 3 x for 5 minutes in TBST, and incubated in secondary antibody in 3% BSA in TBST for 45 minutes. Membranes were washed 3 x for 5 minutes in TBST and then imaged with HRP Pierce ECL Western Blotting Substrate (Thermo Scientific).

<u>Antibody list:</u>
Anti-histone H4, mouse, Cell Signaling Technology L64C1, 1:1000
Anti-H3K27me1, rabbit, Millipore 07-448, 1:1000
Anti-H3K27me2, rabbit, Active Motif 61435, 1:1000
Anti-H3K27me3, rabbit, Active Motif 39156, 1:2000
Goat Anti-Mouse IgG (H+L)-HRP Conjugate, Bio-Rad, 1:3000

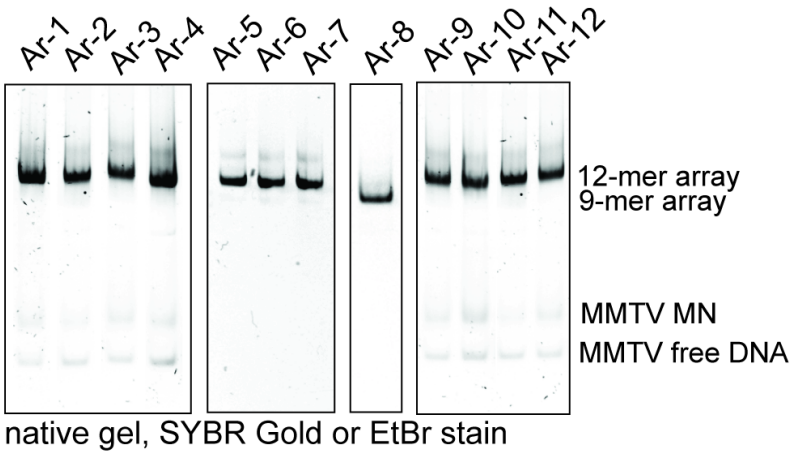
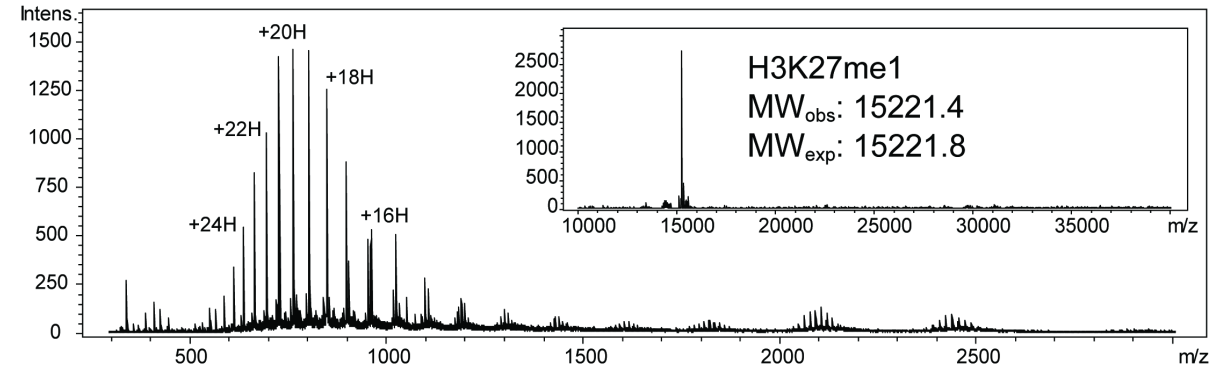
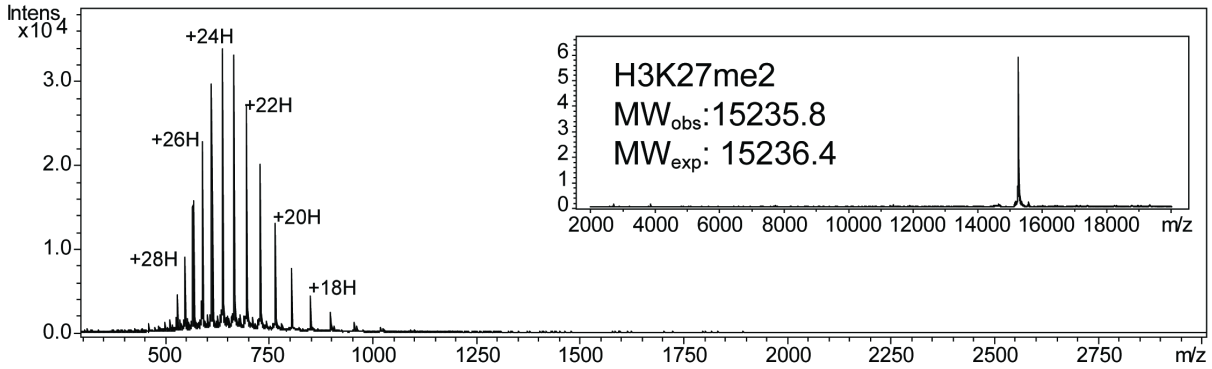
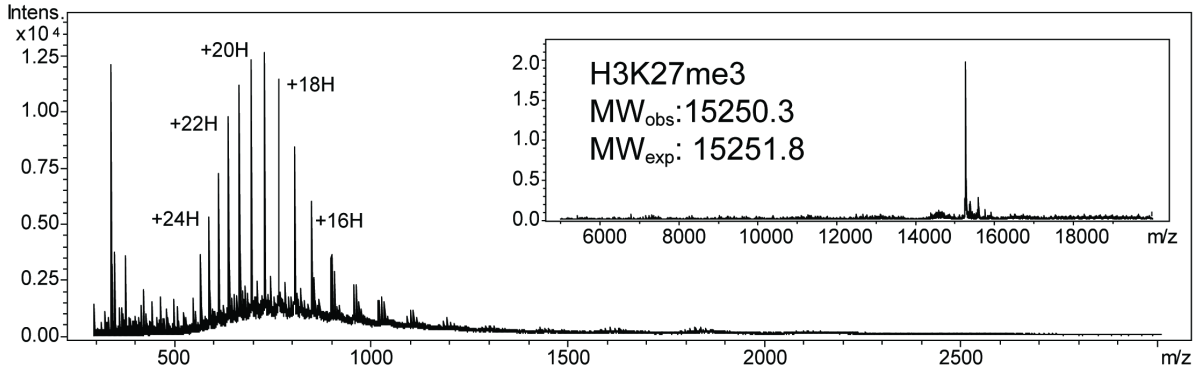


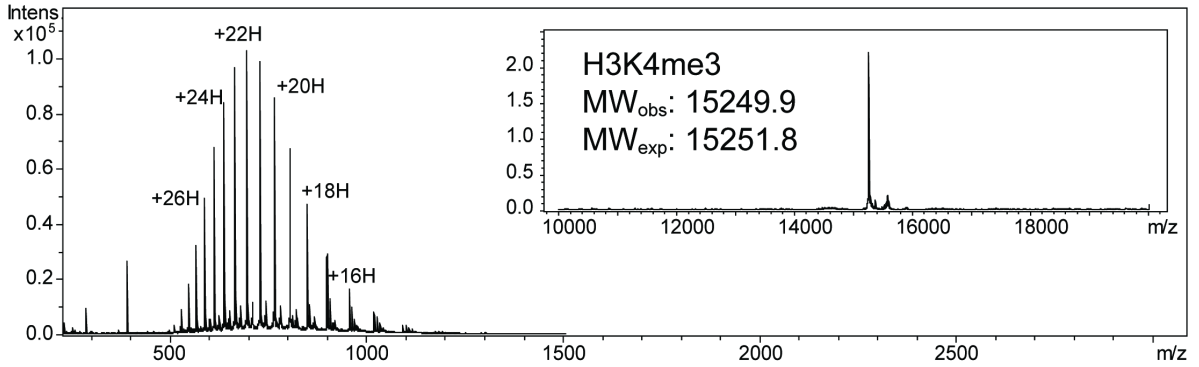
Figure S1: SYBR Gold or ethidium bromide staining of native agarose polyacrylamide gel electrophoresis (APAGE) of purified 9mer and 12mer nucleosome arrays used in this study. MMTV = mouse mammary tumor virus DNA (used to scavenge excess histones). MN = mononucleosome.

a)


Intens.
+20H
+18H
+22H
+24H
+16H
H3K27me1
MWobs: 15221.4
MWexp: 15221.8
m/z


Intens.
x10 4
+24H
+22H
+26H
+20H
+28H
+18H
H3K27me2
MWobs:15235.8
MWexp: 15236.4
m/z


Intens.
x10 4
+20H
+18H
+22H
+24H
+16H
H3K27me3
MWobs:15250.3
MWexp: 15251.8
m/z


Intens.
x10 5
+22H
+24H
+20H
+26H
+18H
+16H
H3K4me3
MWobs: 15249.9
MWexp: 15251.8
m/z

b)

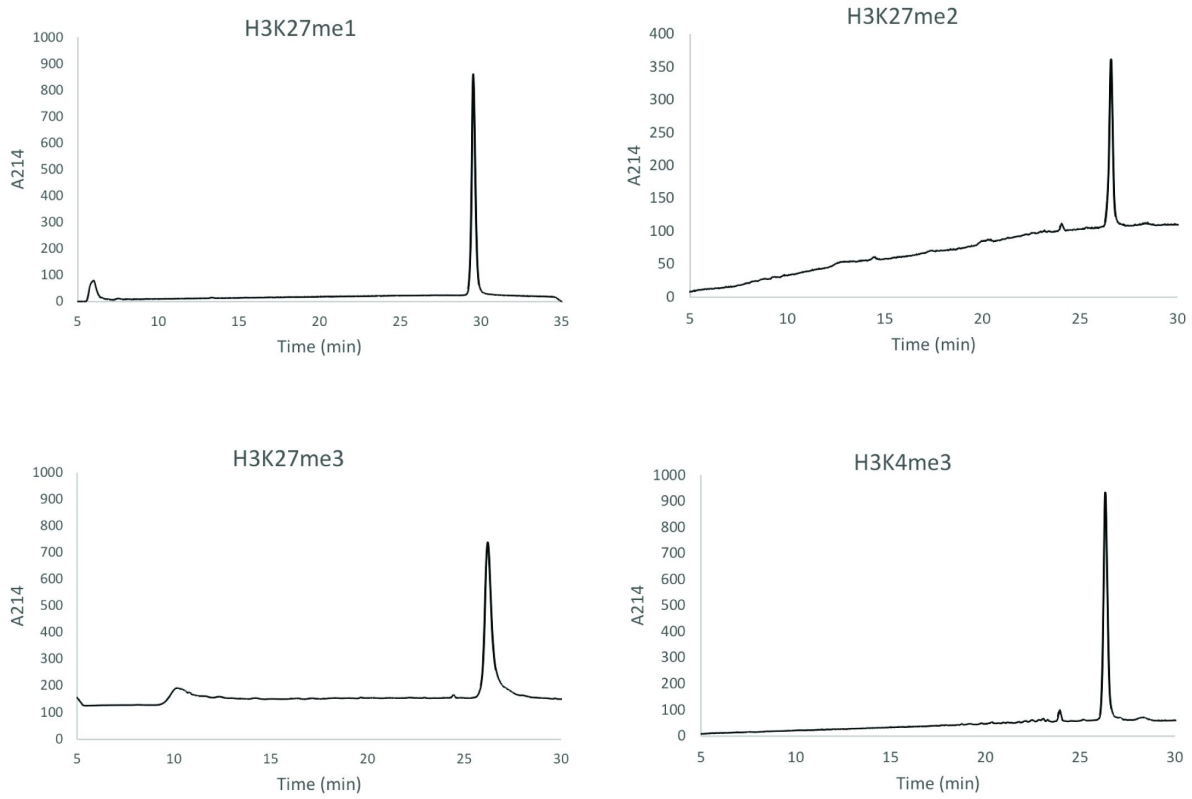


Figure S2. Analytics of semi-synthetic modified histones. (a) ESI mass spectra of purified H3K27me1, H3K27me2, H3K27me3, and H3K4me3. Deconvoluted mass spectra are depicted in the insets. (b) RP-HPLC traces using a 0-73% gradient on a C18 column for purified H3K27me1, H3K27me2, H3K27me3, and H3K4me3.

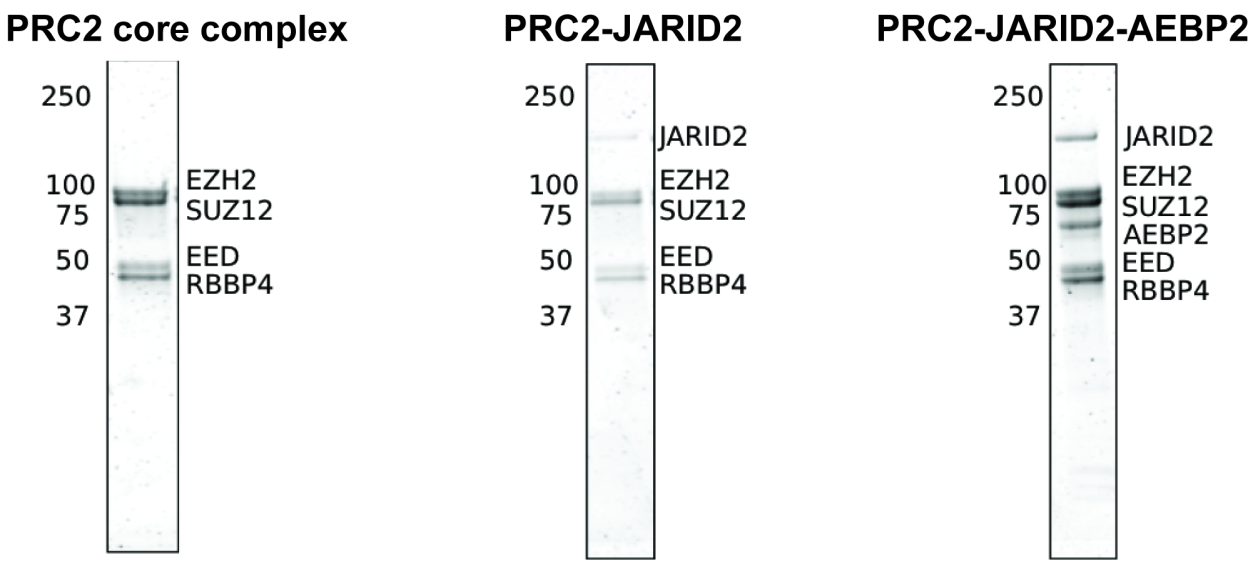


Figure S3: Coomassie staining of SDS-PAGE showing purified PRC2 with various subunit compositions.

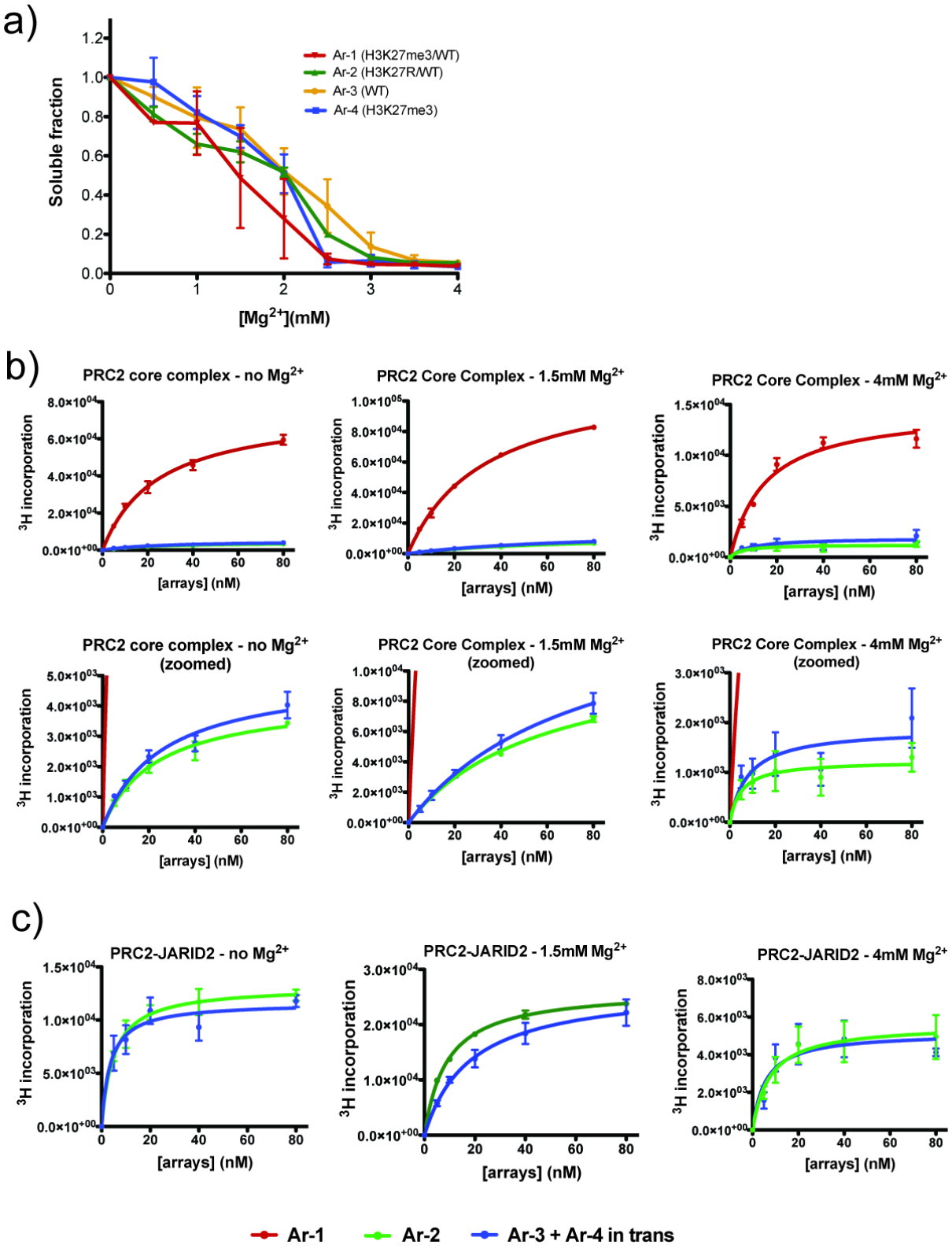

a)
Soluble fraction
[Mg2+](mM)
Ar-1 (H3K27me3/WT)
Ar-2 (H3K27R/WT)
Ar-3 (WT)
Ar-4 (H3K27me3)
b)
PRC2 core complex - no Mg2+
PRC2 Core Complex - 1.5mM Mg2+
PRC2 Core Complex - 4mM Mg2+
PRC2 core complex - no Mg2+ (zoomed)
PRC2 Core Complex - 1.5mM Mg2+ (zoomed)
PRC2 Core Complex - 4mM Mg2+ (zoomed)
3H incorporation
[arrays] (nM)
c)
PRC2-JARID2 - no Mg2+
PRC2-JARID2 - 1.5mM Mg2+
PRC2-JARID2 - 4mM Mg2+
Ar-1
Ar-2
Ar-3 + Ar-4 in trans

Figure S4: (a) Mg^{2+} precipitation curves showing the fraction of various 12mer nucleosome arrays remaining in the soluble fraction after incubation with increasing Mg^{2+} buffer concentrations. Error = s.e.m (n = 4-5). (b) PRC2 core complex HMT activity as measured by incorporation of ^3H (from [^3H]SAM cofactor) on array substrates at indicated concentration of Mg^{2+} (from main text Figure 2) with zoomed in views beneath each original graph. Errors = s.e.m. (n = 3). (c) PRC2-JARID2 HMT activity as measured by incorporation of ^3H (from [^3H]SAM cofactor) on array substrates at indicated concentration of Mg^{2+}. Errors = s.e.m. (n = 3).

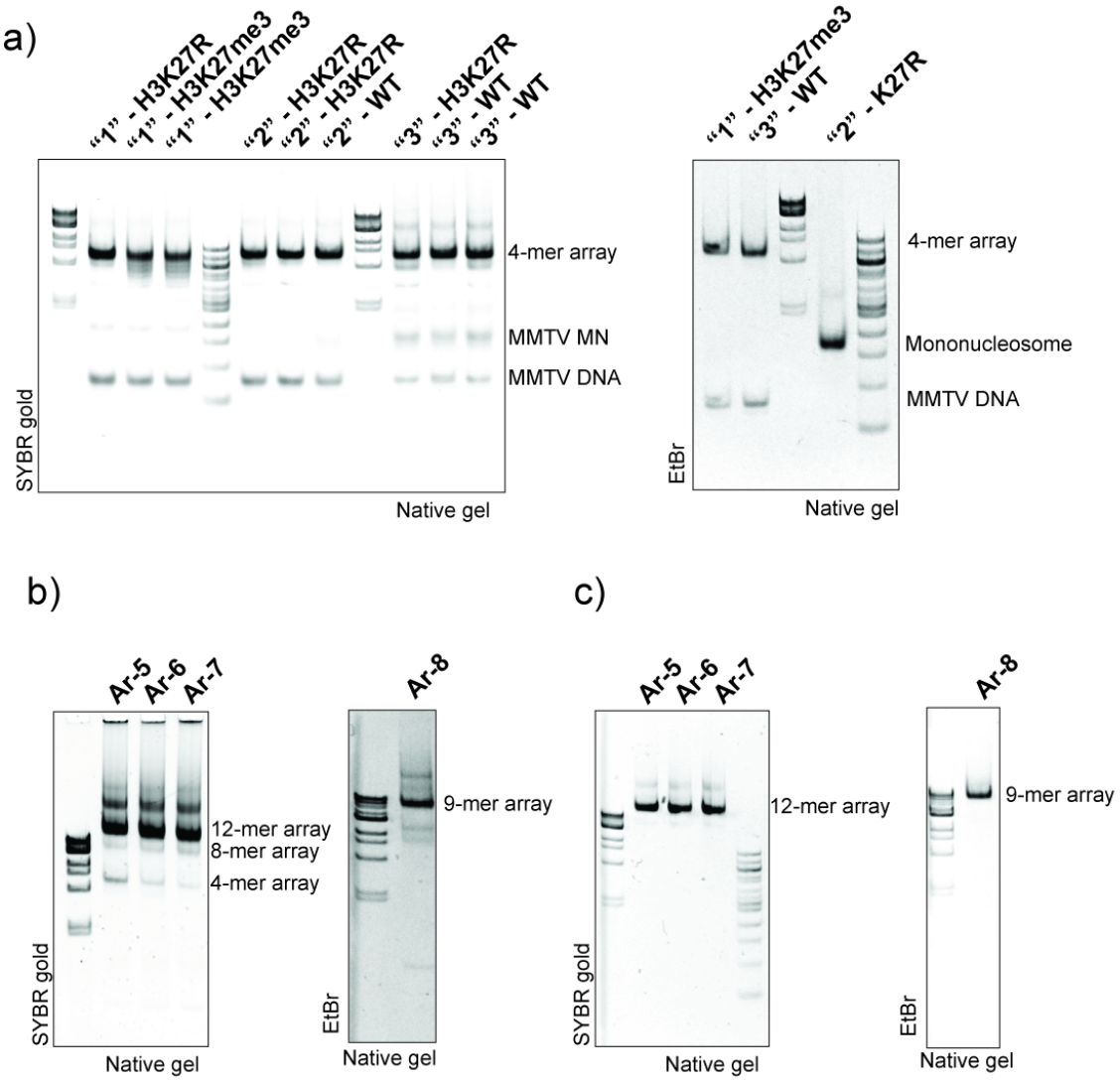


Figure S5: (a) SYBR gold or EtBr stain of native gel electrophoresis of crude tetranucleosome array building blocks used to make heterotypic designer 12mer arrays. "1" denotes the first tetranucleosome building block, "2" denotes the second tetranucleosome building block or the second mononucleosome building block, and "3" denotes the third tetranucleosome building block. MMTV = mouse mammary tumor virus DNA (used to scavenge excess histones). MN = mononucleosome. (b) SYBR Gold or EtBr stain of native gel electrophoresis of crude, unpurified heterotypic designer 9mer and 1-mer array ligation products. (c) SYBR gold or EtBr stain of native gel electrophoresis heterotypic designer 9-mer and 12-mer arrays following sucrose gradient purification.

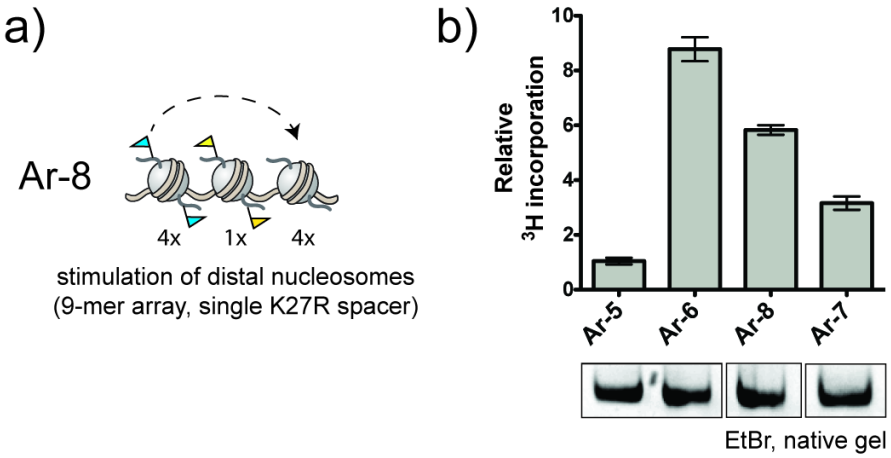


Figure S6: (a) Cartoon representation of heterotypic designer chromatin array. **Ar-8**. (b) HMT activity of PRC2 core complex on various designer chromatin substrates. Data from radio-HMT assays are plotted relative to PRC2 core on **Ar-5**. Errors = s.e.m. (*n*=3). Array substrate loading is shown below.

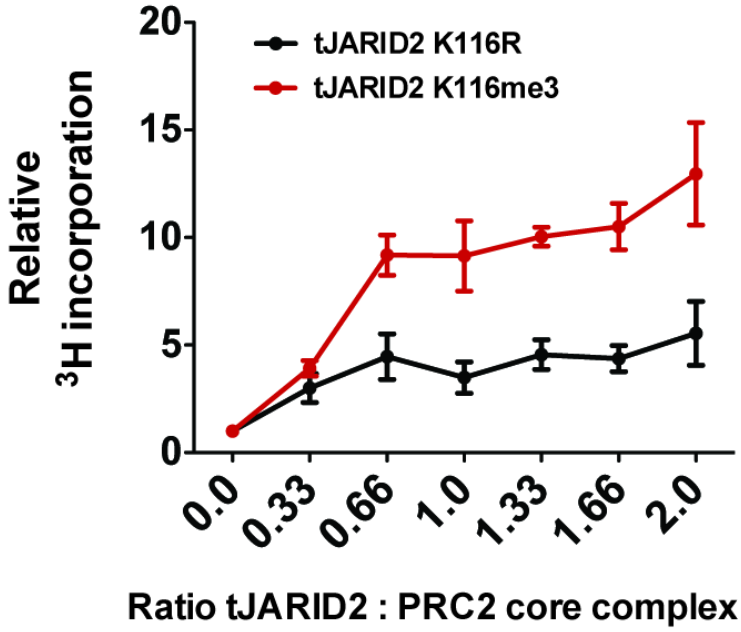


Figure S7: Titration of tJARID2 into the PRC2 core complex. HMT activity of PRC2 containing ratios of either tJARID2 K116R or tJARID2-K116me3 on unmodified (WT) 12-mer array substrates. Errors = s.e.m. (n = 3).

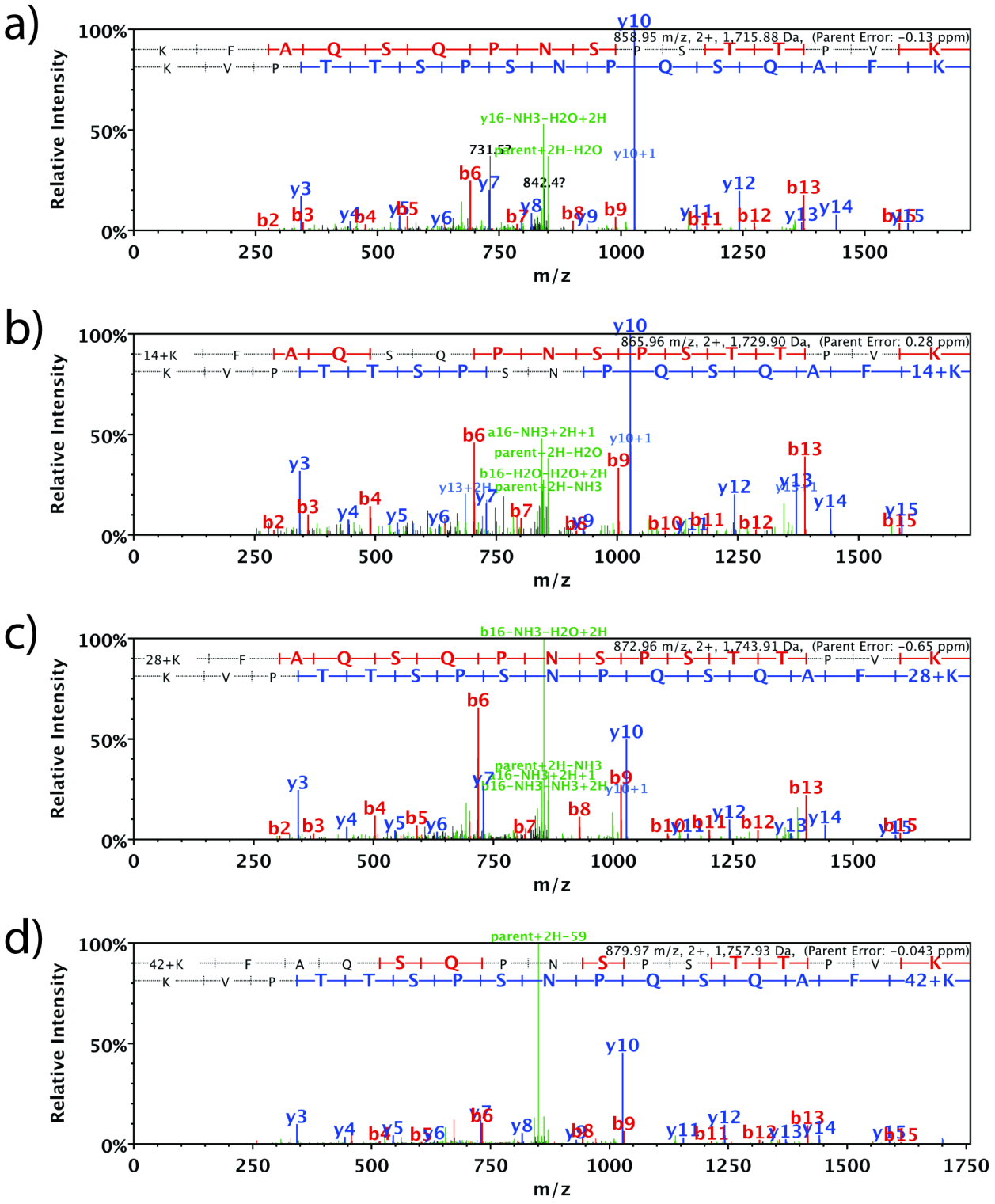


Figure S8: Scaffold software (version Scaffold 4.8.4) spectra showing MS/MS fragmentation patterns with b and y ions of the JARID2 peptide sequence KTAQSQPNSPSTTPVK corresponding to (a) unmodified, (b) monomethylated +14, (c) dimethylated +28, and (d) trimethylated +42 states of Lys-116. Some parent ions and characteristic ammonia loss ions are displayed as well. For quantification, see **Table S1**.

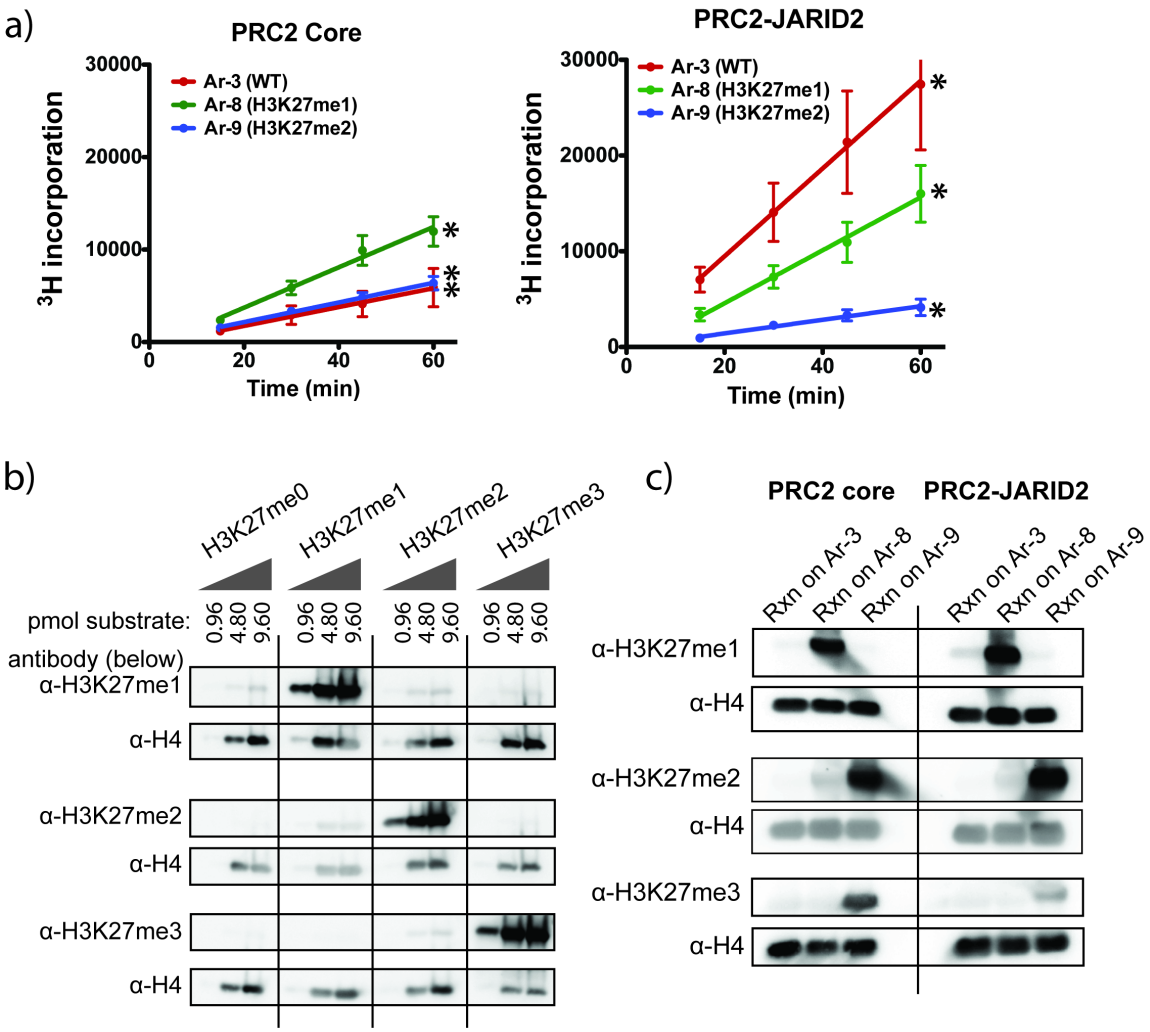


Figure S9: (a) Scintillation-based time-courses of PRC2 methyltransferase activity on various methylation state 12mer nucleosome array substrates. Errors ± s.e.m (n = 3). Data is fitted to a linear regression. Data points marked by asterisks were repeated under non-radioactive conditions for Western blot analysis shown in panel c. (b) Profiling of antibodies specific to H3K27me1, H3K27me2, and H3K27me3 epitopes. Antibodies were used to blot against 0.960 pmol, 4.80 pmol, or 9.60 pmol standards of purified H3K27me1, H3K27me2, or H3K27me3 octamers (representing quantities 10%, 50%, and 100% substrate turnover relative to the amount of substrate used in a typical PRC2 HMT assay). Anti-H4 signal serves as the loading control. Antibodies show limited cross-reactivity between different methylation state epitopes. (c) Methylation state specific antibody analysis of PRC2 core or PRC2-JARID2 activity on H3K27me0, H3K27me1, or H3K27me2 12mer nucleosome arrays (corresponds to the asterisk-labeled data points in panel a). An anti-H4 signal serves as the loading control in each case.

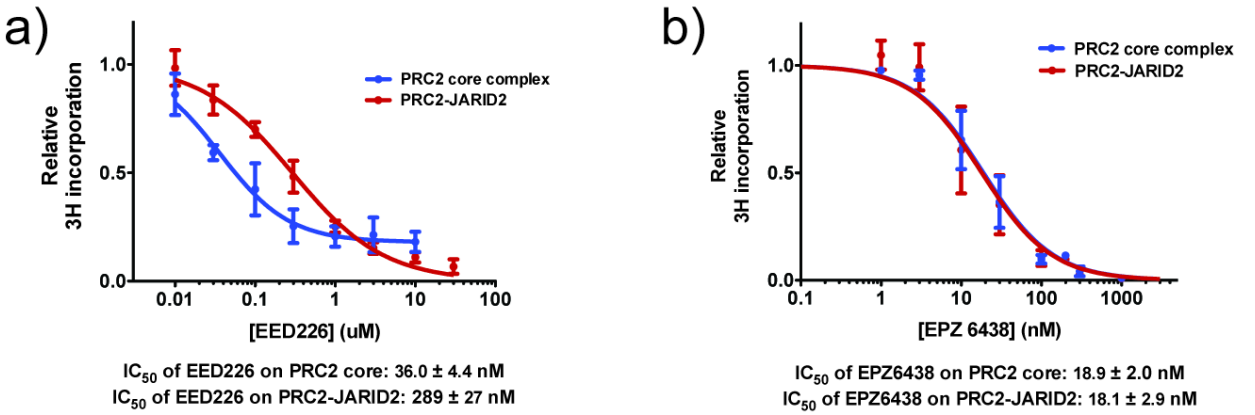
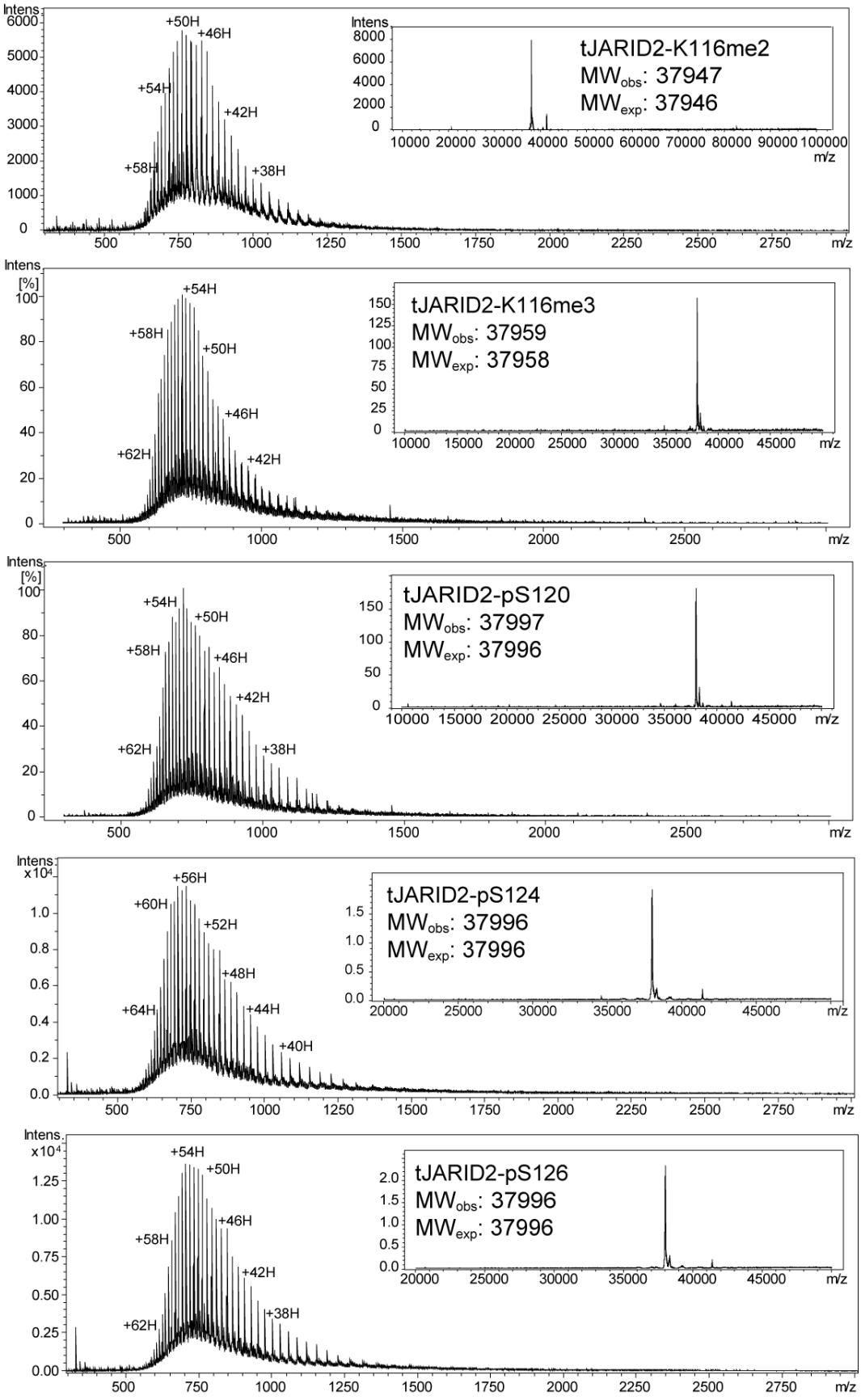


Figure S10: Dose-response curves for EED226 and EPZ 6438. (a) EED226 inhibition of PRC2 core complex and PRC2-JARID2 at fixed concentration of SAM and substrate WT (unmodified) 12mer arrays. IC_{50} values are reported. As previously reported, EED226 behaves as a partial inhibitor in the absence of H3K27me3 (or JARID2).[11] Errors = s.e.m. (n = 3). (b) EPZ6438 inhibition of PRC2 core complex and PRC2-JARID2 JARID2 at fixed concentration of SAM and substrate WT (unmodified) 12-mer arrays. IC_{50} values are reported. Errors = s.e.m. (n = 2).

a)


Intens
+50H
+46H
+54H
+42H
+58H
+38H
tJARID2-K116me2
MWobs: 37947
MWexp: 37946
m/z
Intens
[%]
+54H
+58H
+50H
+46H
+62H
+42H
tJARID2-K116me3
MWobs: 37959
MWexp: 37958
m/z
Intens
[%]
+54H
+50H
+58H
+46H
+42H
+62H
+38H
tJARID2-pS120
MWobs: 37997
MWexp: 37996
m/z
Intens
x10⁴
+56H
+60H
+52H
+48H
+64H
+44H
+40H
tJARID2-pS124
MWobs: 37996
MWexp: 37996
m/z
Intens
x10⁴
+54H
+50H
+46H
+58H
+42H
+62H
+38H
tJARID2-pS126
MWobs: 37996
MWexp: 37996
m/z

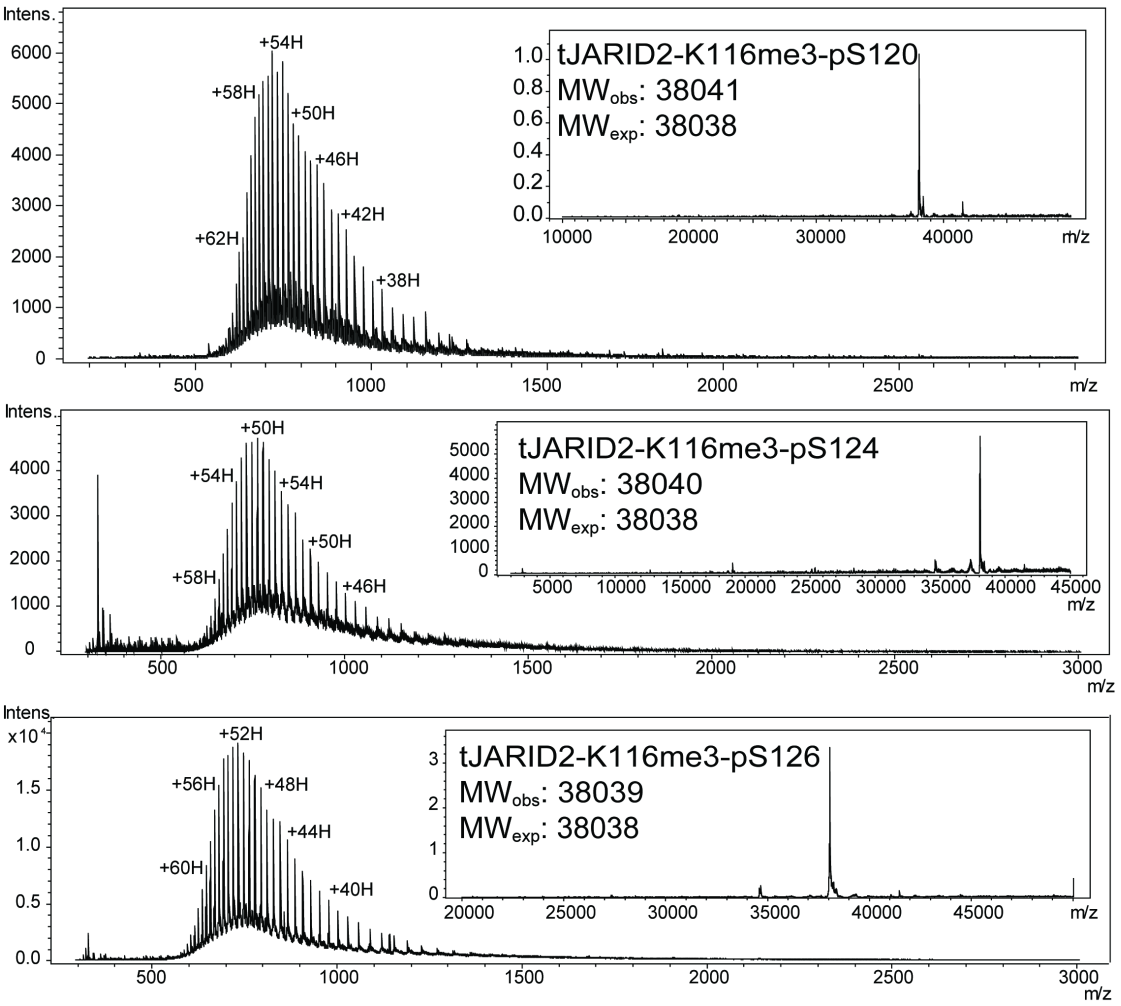


Figure S11: Analytics of semisynthetic purified truncated JARID2 constructs. (a) ESI mass spectra of purified tJARID2-K116me2, tJARID2-K116me3, tJARID2-pS120, tJARID2-pS124, tJARID2-pS126, tJARID2-K116me3-pS120, tJARID2-K116me3-pS124, and tJARID2-K116me3-pS126. The deconvoluted mass spectra are depicted in the insets.

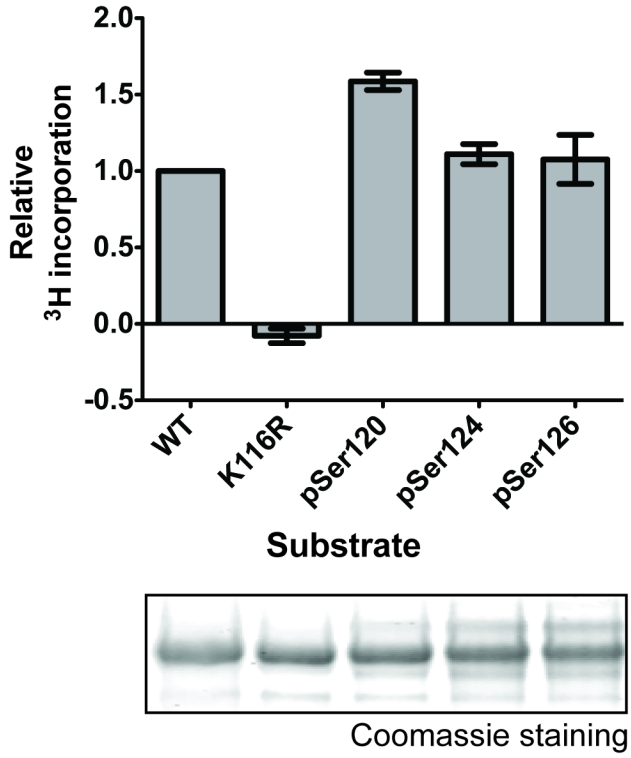


Figure S12: Relative methyltransferase activity of the PRC2 core complex on a substrate of either WT (unmodified) tJARID2 or tJARID2 containing different phosphoserine modifications (tJARID2 pS120, tJARID2 pS124, tJARID2 pS126). Data from radio-HMT assays are plotted relative to tJARID2(WT). Errors s.e.m. (n = 3).

JARID2 (115-130): R K F A Q S Q P N S P S T T P V
ATM: S Q
GSK3: S X X X S P
CDK family: P X S P

Figure S13: Alignment of JARID2 residues 115-130 with known consensus sequences of three candidate kinases (Ref 10).[12] X represents any amino acid. Underlined residues represent phosphoserine sites.

Tryptic peptide	Peak Area	Percent methylated	Modification
KFAQSQPNSPSTTPVK	2.70E+08	0.38%	unmodified
K[+14]FAQSQPNSPSTTPVK	9.96E+09	13.87%	methyl
K[+28]FAQSQPNSPSTTPVK	5.14E+10	71.61%	dimethyl
K[+42]FAQSQPNSPSTTPVK	1.01E+10	14.14%	trimethyl

Table S1: Peak area-based quantification of mono-, di-, and tri-methylation of Lys-116 within the JARID2 tryptic peptide KFAQSQPNSPSTTPVK using Skyline software (MacCoss Lab, University of Washington).

References:

1. Zheng, J. S.; Tang, S.; Qi, Y. K.; Wang, Z. P.; Liu, L., Chemical synthesis of proteins using peptide hydrazides as thioester surrogates. *Nat. Protoc.* **2013,** *8* (12), 2483-95.
2. Thompson, R. E.; Liu, X.; Alonso-Garcia, N.; Pereira, P. J.; Jolliffe, K. A.; Payne, R. J., Trifluoroethanethiol: an additive for efficient one-pot peptide ligation-desulfurization chemistry. *J. Am. Chem. Soc.* **2014,** *136* (23), 8161-4.
3. Dyer, P. N.; Edayathumangalam, R. S.; White, C. L.; Bao, Y.; Chakravarthy, S.; Muthurajan, U. M.; Luger, K., Reconstitution of nucleosome core particles from recombinant histones and DNA. *Methods Enzymol.* **2004,** *375*, 23-44.
4. Berger, I.; Garzoni, F.; Chaillet, M.; Haffke, M.; Gupta, K.; Aubert, A., The multiBac protein complex production platform at the EMBL. *J. Vis. Exp.* **2013,** (77), e50159.
5. Shevchenko, A.; Tomas, H.; Havlis, J.; Olsen, J. V.; Mann, M., In-gel digestion for mass spectrometric characterization of proteins and proteomes. *Nat. Protoc.* **2006,** *1* (6), 2856-2860.
6. Bern, M.; Kil, Y. J.; Becker, C., Byonic: advanced peptide and protein identification software. *Curr. Protoc. Bioinformatics* **2012,** *Chapter 13*, Unit13 20.
7. Eng, J. K.; McCormack, A. L.; Yates, J. R., An approach to correlate tandem mass spectral data of peptides with amino acid sequences in a protein database. *J. Am. Soc. Mass. Spectrom.* **1994,** *5* (11), 976-89.
8. Nesvizhskii, A. I.; Keller, A.; Kolker, E.; Aebersold, R., A statistical model for identifying proteins by tandem mass spectrometry. *Anal. Chem* **2003,** *75* (17), 4646-58.
9. Muller, M. M.; Fierz, B.; Bittova, L.; Liszczak, G.; Muir, T. W., A two-state activation mechanism controls the histone methyltransferase Suv39h1. *Nat. Chem. Biol.* **2016,** *12* (3), 188-93.
10. Neubig, R. R.; Spedding, M.; Kenakin, T.; Christopoulos, A.; International Union of Pharmacology Committee on Receptor, N.; Drug, C., International Union of Pharmacology Committee on Receptor Nomenclature and Drug Classification. XXXVIII. Update on terms and symbols in quantitative pharmacology. *Pharmacol. Rev.* **2003,** *55* (4), 597-606.
11. Qi, W.; Zhao, K.; Gu, J.; Huang, Y.; Wang, Y.; Zhang, H.; Zhang, M.; Zhang, J.; Yu, Z.; Li, L.; Teng, L.; Chuai, S.; Zhang, C.; Zhao, M.; Chan, H.; Chen, Z.; Fang, D.; Fei, Q.; Feng, L.; Feng, L.; Gao, Y.; Ge, H.; Ge, X.; Li, G.; Lingel, A.; Lin, Y.; Liu, Y.; Luo, F.; Shi, M.; Wang, L.; Wang, Z.; Yu, Y.; Zeng, J.; Zeng, C.; Zhang, L.; Zhang, Q.; Zhou, S.; Oyang, C.; Atadja, P.; Li, E., An allosteric PRC2 inhibitor targeting the H3K27me3 binding pocket of EED. *Nat. Chem. Biol.* **2017,** *13* (4), 381-388.
12. Kim, S. T.; Lim, D. S.; Canman, C. E.; Kastan, M. B., Substrate specificities and identification of putative substrates of ATM kinase family members. *J. Biol. Chem.* **1999,** *274* (53), 37538-43.