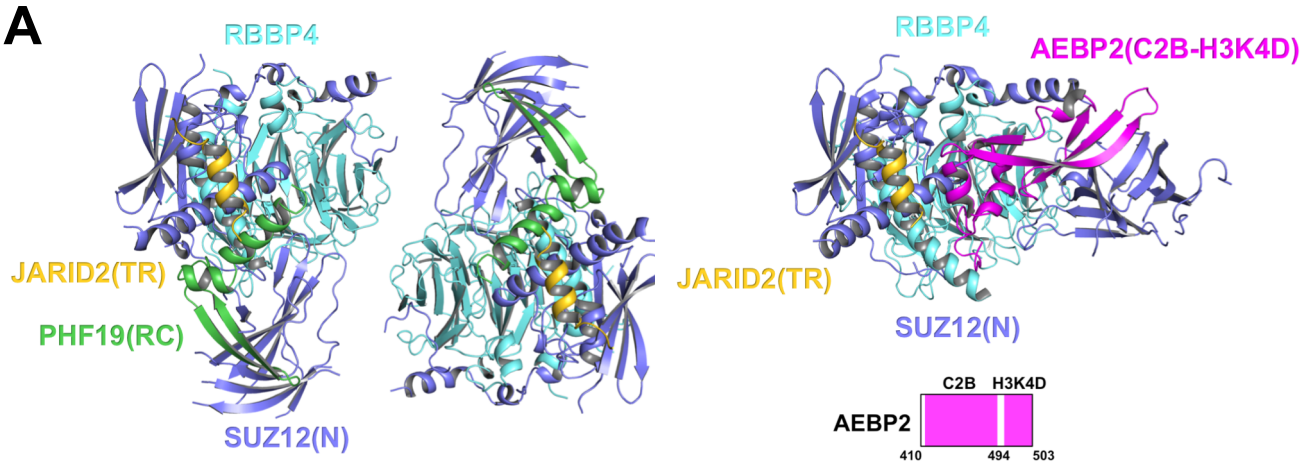
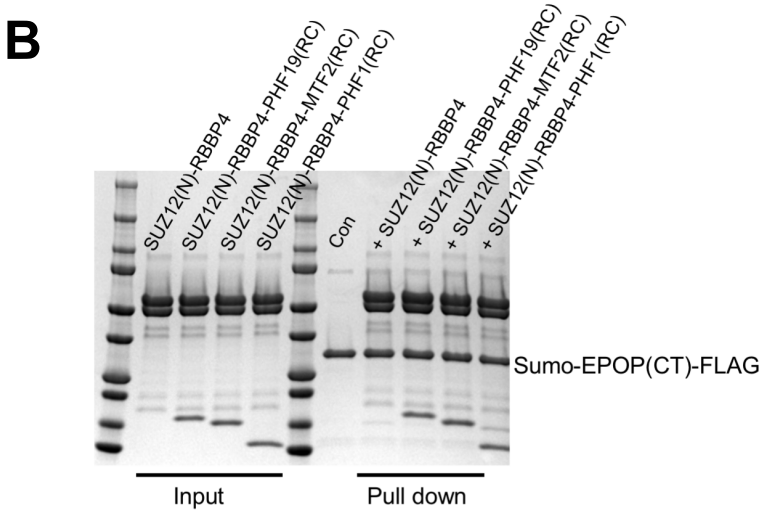
Supplemental Information

A Dimeric Structural Scaffold for PRC2-PCL Targeting to CpG Island Chromatin

Siming Chen, Lianying Jiao, Xiuli Liu, Xin Yang, Xin Liu

Fig. S1 (related to Fig. 1)





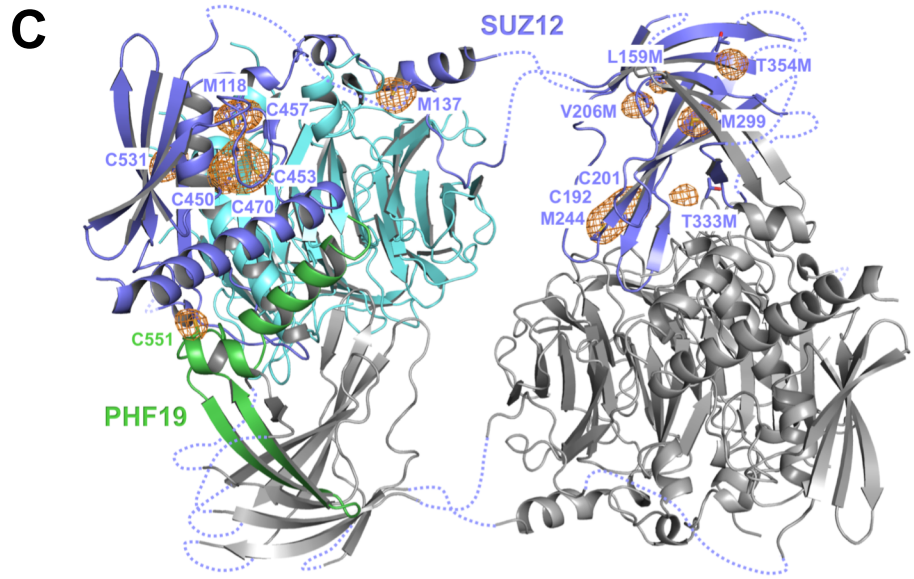


Fig. S1 Crystal structure analysis of SUZ12(N)-RBBP4-PHF19(RC)-JARID2(TR)

(A) Structural comparison of the SUZ12(N)-RBBP4-AEBP2(C2B-H3K4D)-JARID2(TR) and SUZ12(N)-RBBP4-PHF19(RC)-JARID2(TR) complexes

Different proteins are color-coded and labeled. Two complexes are oriented similarly for comparison. The JARID2(TR) in yellow occupies the same surface on the SUZ12(N)-RBBP4 protein body in the two structures, and has little crosstalk with the AEBP2(C2B-H3K4D) in magenta (lower panel) or the PHF19(RC) in green (upper panel). Domain structure of AEBP2 is also provided. "C2B" stands for C2-Binding and "H3K4D" stands for H3K4 Displacement.

(B) Reconstitution of the SUZ12(N)-RBBP4-PHF19(RC)-EPOP(CT) complex

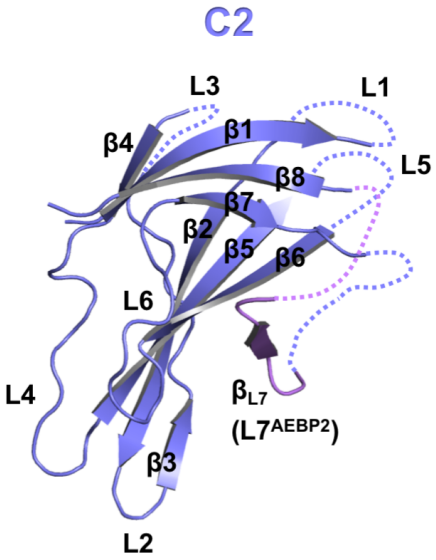
The complex was reconstituted by mixing a FLAG-tagged EPOP(CT) corresponding to residues 285-379 with the purified ternary complexes, including SUZ12(N)-RBBP4-PHF1(RC), SUZ12(N)-RBBP4-MTF2(RC) and SUZ12(N)-RBBP4-PHF19(RC). Anti-FLAG resin IP indicated stoichiometric binding in all cases. Coomassie blue staining of a SDS-PAGE gel is shown.

(C) SAD signals used for model building and validation

A total of sixteen sulfur atoms from both endogenous residues or designed methionine mutants of SUZ12 and PHF19 were located.

Fig. S2 (related to Fig. 2)

A



B

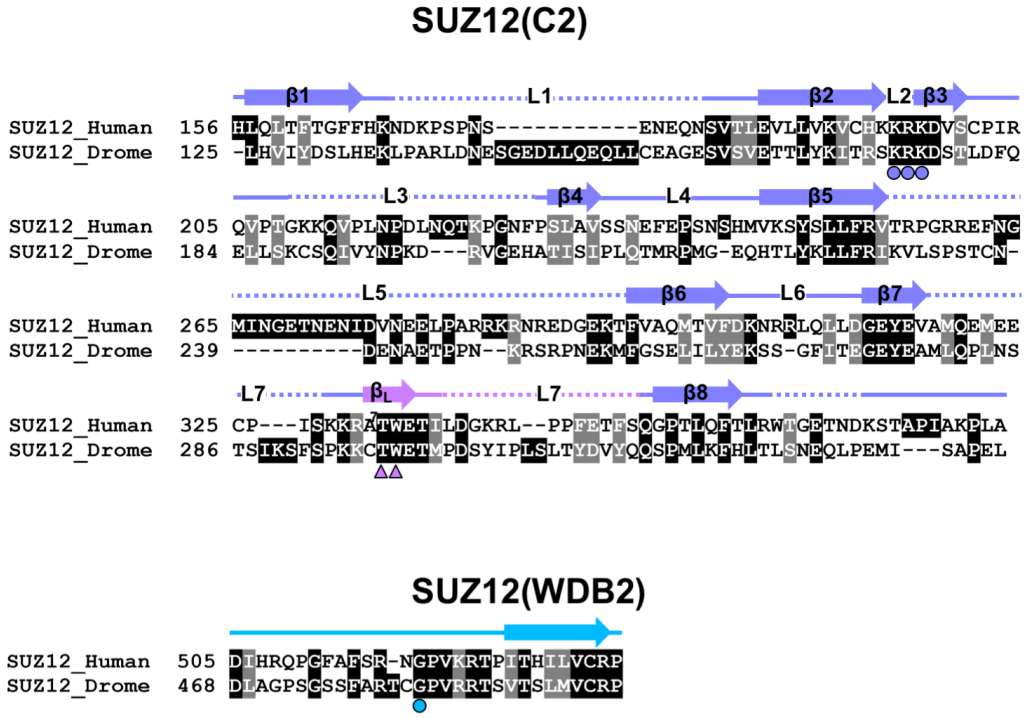


Fig. S2 Structure of the SUZ12(C2) and sequence of the SUZ12(C2) and the SUZ12(WDB2)

(A) Structure of the C2 domain

Zoom-in view of the C2 domain of SUZ12. The L2 loop forms a major dimer interface. The L7 loop is partially folded into the β_{L7} strand that locks the PHF19(RC) (omitted for clarity) in the current structure. The L7 loop is disordered in the AEBP2-bound structure.

(B) Sequence alignment of SUZ12

The C2 and WDB2 domains of SUZ12 from humans and *Drosophila* were aligned. Residues of the C2 domain involved in RBBP4 binding are indicated by filled circles. Residues of the β_{L7} strand involved in PHF19 binding are indicated by filled triangles. Human RBBP4 and its homolog in *Drosophila*, CAF1 (a.k.a. NURF55), share about 90% sequence identifies, with the SUZ12(C2)-interacting residues strictly conserved. The sequence alignment of RBBP4 and CAF1 is not shown.

Fig. S3 (related to Fig. 3)

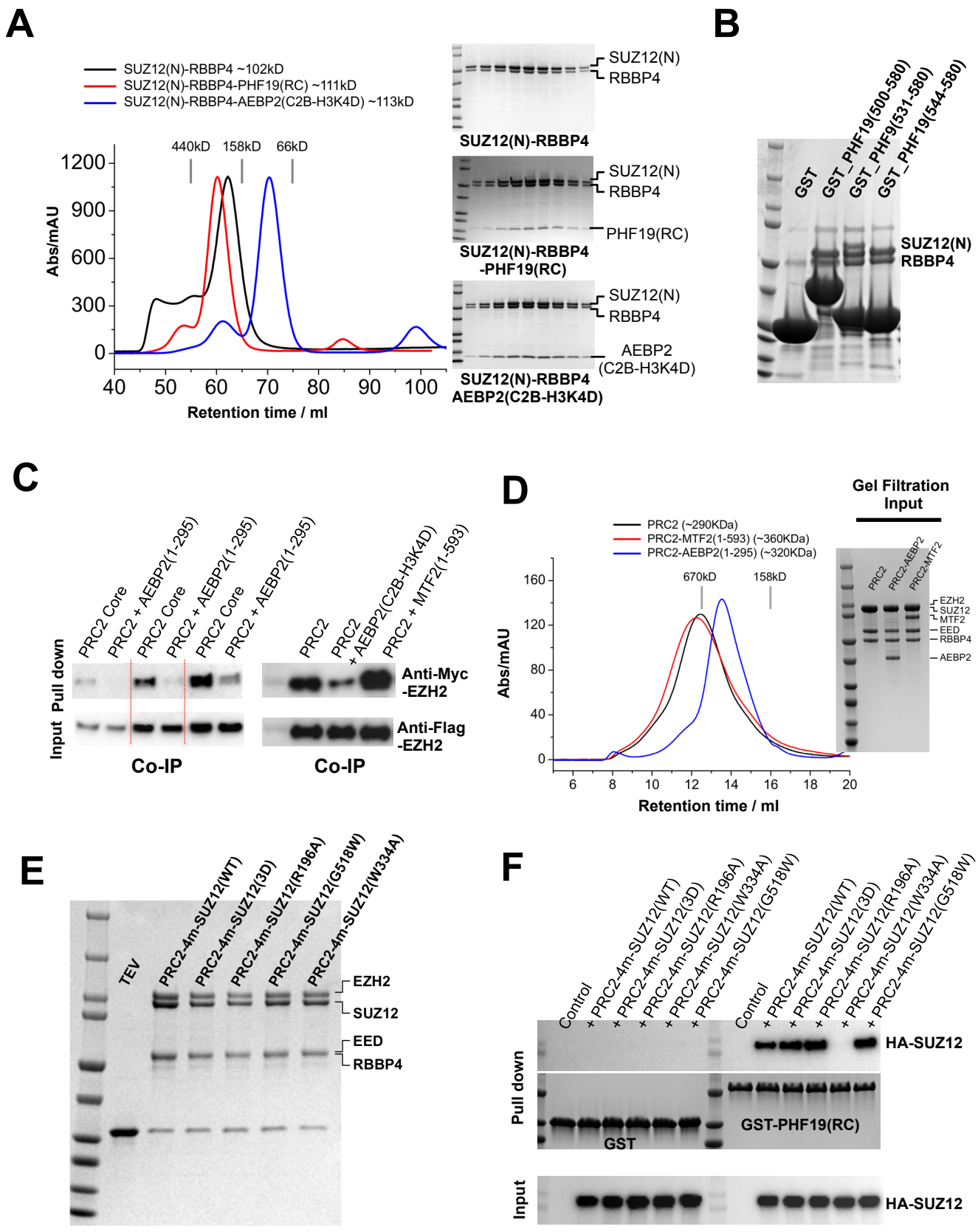


Fig. S3 Biochemical analysis of the PRC2 dimer

(A) SEC elution profiles of SUZ12(N)-RBBP4 bound to the AEBP2(C2B-H3K4D) fragment or the PHF19(RC) fragment

Calculated molecular weights of SUZ12(N)-RBBP4, SUZ12(N)-RBBP4- AEBP2(C2B-H3K4D) and SUZ12(N)-RBBP4-PHF19(RC) are indicated. SDS-PAGE gels corresponding to SEC fractions are shown. SUZ12(N)-RBBP4 and SUZ12(N)-RBBP4-PHF19(RC) behaved as dimers and SUZ12(N)-RBBP4-AEBP2(C2B-H3K4D) behaved as a monomer.

(B) SUZ12(N)-RBBP4 binding by PHF19(RC)

GST-tagged PHF19(RC) (residues 500-580) and truncation variants were used to pull down SUZ12(N)-RBBP4. Residues 531-580 were revealed in the current crystal structure. The DS helix was deleted in the construct containing residues 544-580. The DS helix of PHF19 was not essential for SUZ12(N)-RBBP4 binding.

(C) Disruption and stabilization of PRC2-4m dimer by full-length AEBP2 and MTF2

Co-IP was used to assess the extent of dimer formation. PRC2-4m containing both FLAG-tagged EZH2 and Myc-tagged EZH2 was purified from HEK293T cells through transient expression. Equal amounts of PRC2-4m were subjected to dimer disruption and stabilization by full-length AEBP2 (residues 1-295) (left) and full-length MTF2 (residues 1-593) (right), respectively. Anti-FLAG-EZH2 IP captured Myc-EZH2 through PRC2 dimerization, which was eluted by FLAG peptide. Anti-Myc Western signals represented the extent of PRC2 dimerization. Left: increasing amounts of Co-IP samples (1×, 2× and 4×) were loaded. Right: the AEBP2(C2B-H3K4D) fragment was used as a control that disrupted the intrinsic PRC2 dimer.

(D) SEC elution profiles of PRC2, PRC2-MTF2 and PRC2-AEBP2

Purity of the three complexes prior to SEC was assessed by SDS-PAGE gel. SEC elution profiles were compared. Calculated molecular weights of the complexes are indicated.

(E) Purification of the WT and mutant PRC2 core complexes

Purity of the WT and mutant PRC2 core complexes used for dimer formation Co-IP assay was assessed by SDS-PAGE.

(F) PHF19(RC) binding by the WT and mutant PRC2 core complexes

GST-tagged PHF19(RC) was used to pull down the purified PRC2 core complexes. Bound PRC2 was assessed by anti-HA Western blot signals from HA-SUZ12. Different from other mutations of SUZ12, the W334A mutation disrupted PHF19(RC) binding to the PRC2 core complex.

Fig. S4 (related to Fig. 4)

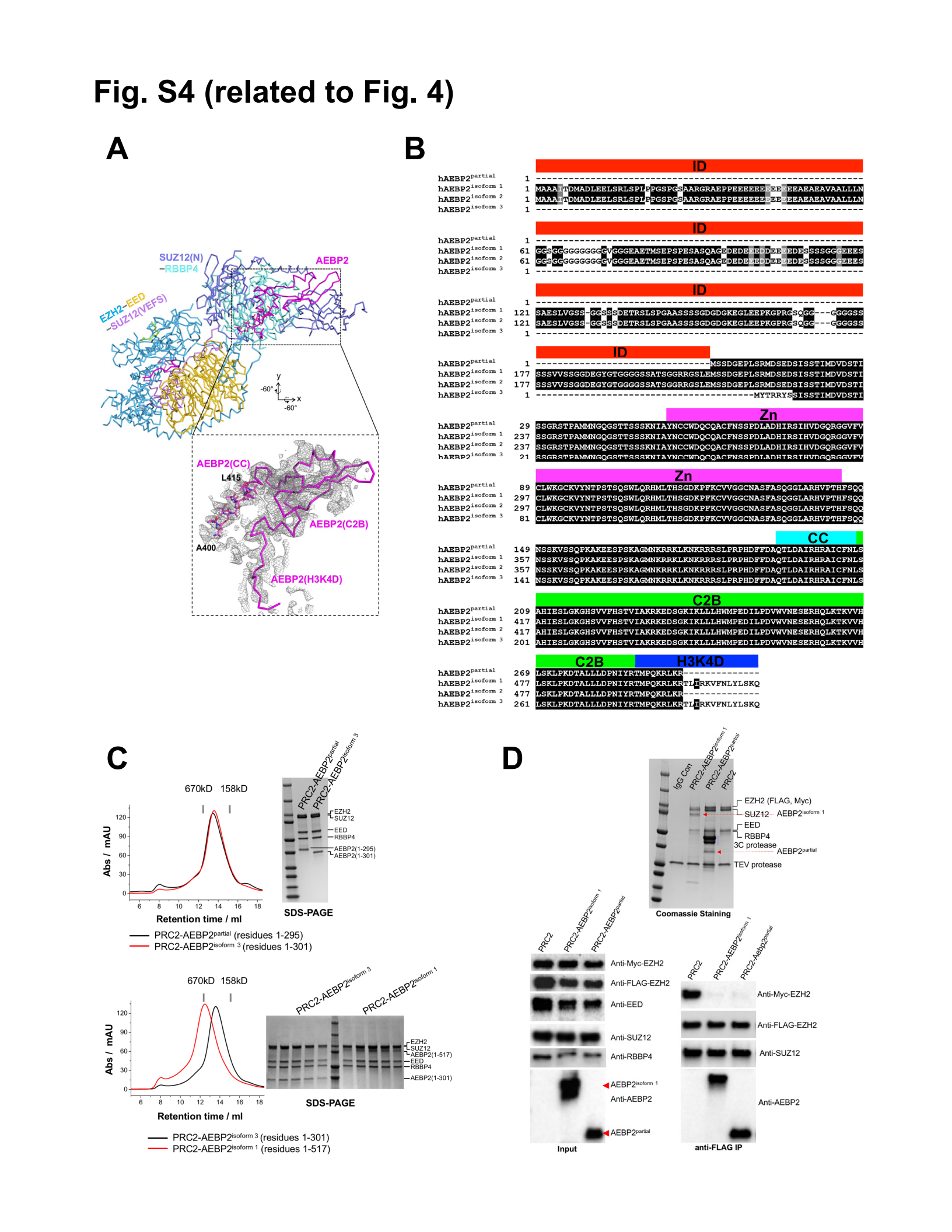


Fig. S4 Structural and biochemical analysis of PRC2-AEBP2

(A) Fitting of the crystal structure into the cryo-EM density map

The crystal structure of SUZ12(N)-RBBP4-AEBP2(C2B-H3K4D)-JARID2(TR) from PDB 5WAI was fitted into the cryo-EM map of a PRC2-AEBP2-JARID2 holo complex. The AEBP2(CC) helix that was missing in the crystal structure was built as a polyalanine model based on the cryo-EM density map EMD-7334 and is shown as sticks. The estimated registers of the AEBP2(CC) are indicated.

(B) Sequence alignment of different versions of AEBP2 used in literatures and from UniProt database

Four versions of human AEBP2 were aligned, including $AEBP2^{partial}$ (residues 1-295), $AEBP2^{isoform\ 1}$ (residues 1-517), $AEBP2^{isoform\ 2}$ (residues 1-503) and $AEBP2^{isoform\ 3}$ (residues 1-301). Various functional domains of AEBP2 are indicated as below: ID = Intrinsically Disordered; Zn = Zinc finger, CC = Central Connecting; C2B = C2-Binding; H3K4D = H3K4 Displacement.

(C) SEC elution profiles of $PRC2\text{-}AEBP2^{isform\ 1}$, $PRC2\text{-}AEBP2^{isoform\ 3}$ and $PRC2\text{-}AEBP2^{partial}$

SEC elution profiles were compared in pairs. Protein quality was assessed by SDS-PAGE. $AEBP2^{isform\ 1}$ = 2×StrepII-AEBP2 (1-517)-3C-FLAG; $PRC2\text{-}AEBP2^{isoform\ 3}$ = 2×StrepII-AEBP2(1-301); $AEBP2^{partial}$ = 2×StrepII-3C-Aebp2(1-295)-His_6.

(D) PRC2 dimerization Co-IP assay in the presence of $AEBP2^{partial}$ (residues 1-295) and $AEBP2^{isoform\ 1}$ (residues 1-517)

The PRC2 core complex and the AEBP2-containing PRC2 core complexes were purified from HEK293T cells that transiently coexpressed the corresponding protein subunits. Purified complexes were assessed by SDS-PAGE (upper panel). Input complexes were checked by Western blot using anti-FLAG (FLAG-EZH2), anti-Myc (Myc-EZH2), anti-EED, anti-SUZ12, anti-RBBP4 and anti-AEBP2 antibodies. The anti-AEBP2 antibody recognized both the longer and shorter versions of AEBP2 (lower left panel). Anti-Myc signals represented Myc-EZH2 captured through PRC2 dimerization during anti-FLAG IP. Both versions of AEBP2 disrupted the intrinsic PRC2 dimer under the assay buffer condition containing 50mM Tris-HCl, pH 8.0, 150mM NaCl, 10% glycerol, 0.1% NP40 and 2mM DTT.

Fig. S5 (related to Fig. 5)

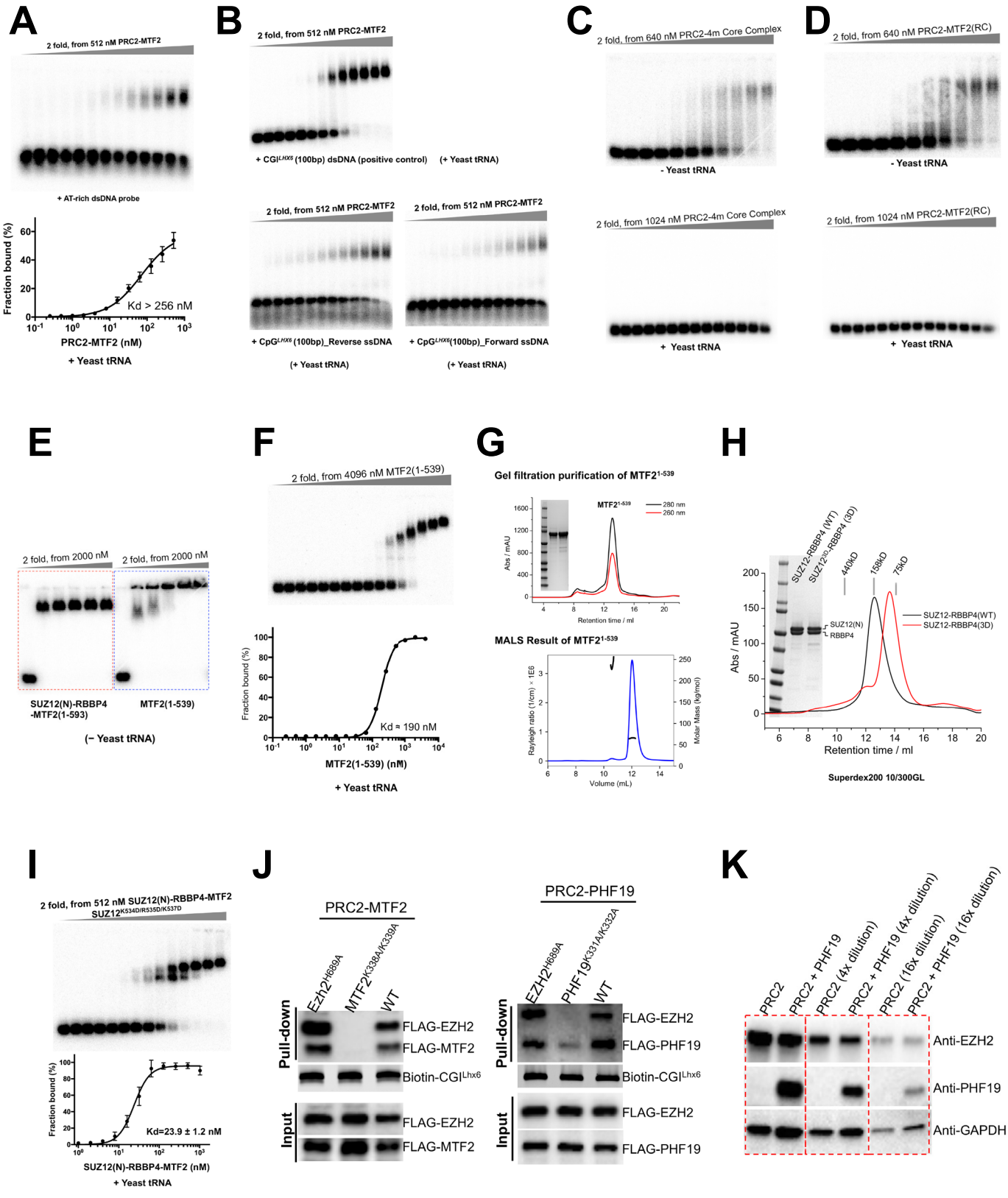


Fig. S5 Role of PRC2 dimerization in CGI DNA binding *in vitro*

(A) EMSA of PRC2-MTF2 binding to an AT-rich DNA probe

EMSA using ^{32}P-labeled AT-rich DNA probe was performed for three times in the presence of yeast tRNA. Binding was not saturated at 512nM of PRC2-MTF2. Sequence of the 100bp AT-rich probe is provided in the Methods section.

(B) EMSA of PRC2-MTF2 binding to single stranded CGI^{LHX6} DNA probes

EMSA using ^{32}P-labeled single-stranded CGI^{LHX6} DNA probes (forward and reverse strands) was performed in the presence of yeast tRNA (lower two panels). PRC2-MTF2 binding to the double-stranded CGI^{LHX6} probe was also performed in parallel as a positive control in the presence of yeast tRNA (upper panel).

(C) DNA^{LHX6} binding by the PRC2 core complex

EMSA were performed both in the absence (upper panel) and presence (lower panel) of yeast tRNA.

(D) DNA^{LHX6} binding by PRC2-MTF2(RC)

PRC2-MTF2(RC) that lacks the DNA binding domain of MTF2 was used in EMSA both in the absence (upper panel) and presence (lower panel) of yeast tRNA.

(E) DNA^{LHX6} binding by MTF2 lacking the RC domain in the absence of yeast tRNA

MTF2$^{\Delta RC}$ corresponding to residues 1-539 did not bind PRC2 and formed aggregates with the DNA probe in the absence of yeast tRNA (right panel). As a positive control, SUZ12(N)-RBBP4-MTF2 displayed specific binding to the DNA probe (left panel).

(F) DNA^{LHX6} binding by MTF2 lacking the RC domain in the presence of yeast tRNA

MTF2$^{\Delta RC}$ bound to the DNA probe specifically, but with a greatly reduced affinity in the presence of yeast tRNA. No aggregates were observed under this assay condition. In both **(E)** and **(F)**, MTF2$^{\Delta RC}$ was diluted in the binding buffer before mixing with the DNA probe.

(G) Quality control of purified MTF2$^{\Delta RC}$ (residues 1-539)

The SEC profile (upper panel) and SEC-MALS result (lower panel) were shown, both supporting that MTF2$^{\Delta RC}$ (residues 1-539) used in EMSA was monodisperse.

(H) SEC elution profiles of the WT and mutant SUZ12(N)-RBBP4

The SEC elution profiles of WT and the K195D/R196D/K197D (3D) mutant SUZ12(N)-RBBP4 are shown. Protein purity was assessed by SDS-PAGE.

(I) EMSA of CGI^{LHX6} DNA probe binding by SUZ12(N)-RBBP4-MTF2 containing a control charge reversal mutant of SUZ12

The K534D/R535D/K537D mutant of SUZ12 was used in DNA binding in the context of SUZ12(N)-RBBP4-MTF2. EMSA was performed for three times in the presence of yeast tRNA. The binding affinity was measured to be 23.9nM, comparable to the WT complex (refer to **Fig. 5E**).

(J) Biotinylated CGI^{LHX6} DNA pull-down of mutant PRC2-MTF2 and PRC2-PHF19 complexes

Two types of mutations were assayed, including a DNA binding defective double mutation of MTF2 and PHF19 (MTF2$^{K338A/K339A}$ and PHF19$^{K331A/K332A}$) and an enzyme inactivating single mutation of EZH2 (EZH2^{H689A}). DNA binding was sensitive to the former type but not the latter type of mutation, highlighting specificity of the pull-down assay. Protein input was quantitated by anti-FLAG Western signals. DNA bait was detected by SYBR Gold. The bound protein complexes were quantitated by anti-FLAG Western signals.

(K) Expression levels of transiently expressed proteins

PRC2 and PRC2-PHF19 were transiently expressed in HEK293T cells. Only overexpressed PHF19 in the latter case but not endogenous PHF19 in the former case was detected by an anti-PHF19 antibody. Different amounts of samples were loaded for an accurate comparison.

Fig. S6 (related to Fig. 6)

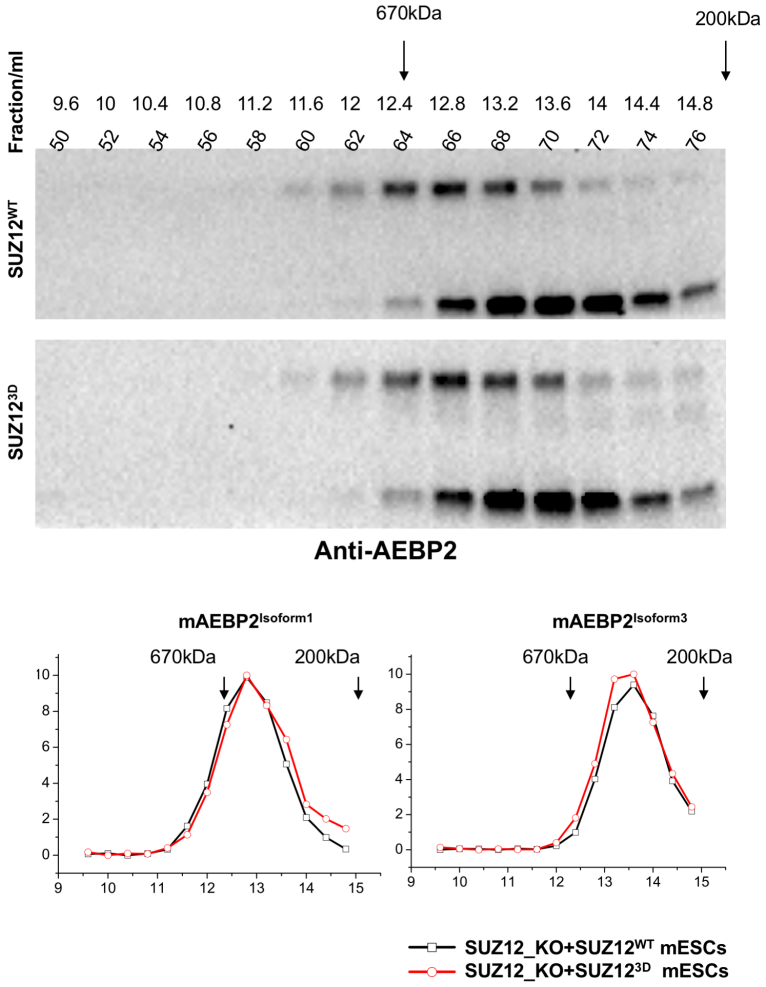


Fig. S6 SEC profiles of endogenous mouse AEBP2 from SUZ12^{3D} and SUZ12^{WT} mESCs

Both the longer (e.g. mAEBP2$^{isoform\ 1}$) and shorter (e.g. mAEBP2$^{isoform\ 3}$) isoforms were detected by anti-AEBP2 antibody. No difference in SEC profiles was noted between SUZ12^{3D} and SUZ12^{WT} mESCs. Quantification of Western blot signals for SEC fractions is provided at the bottom.

Table S1 List of primers used for ChIP-qPCR (related to Fig. 6)

ChIP-qPCR primer pairs		
Gene name	**Forward primer sequence (5' to 3')**	**Reverse primer sequence (5' to 3')**
TBX3	GTTCTAGCCAACCACCCTGC	GGACAGACACACTGACCCTTTTG
SATB2	CTCCTCCCTCACTGCCCC	TTCCTCTTCCCCGGTCCTAG
GATA4	CTCTCCCGAGCTCACTTCAAGG	GGAGAAGGTGACCTCGCACAC
PAX3	AGAAAGGCGGAAAGAGATTAGGAC	GAGCCAGTGAGGATGGAAAAAG
FGF5	AGGGACGGTCAAGATTCCTT	AGAACCAGCAGAGTCCCAGA
HOXA7	GAGAGGTGGGCAAAGAGTGG	CCGACAACCTCATACCTATTCCTG
NANOG	GGCATGGTGGTAGACAAGCC	TTAGTAAGTTGGTCCATGCTTTGG
OCT4	CAGAGCATGGTGTAGGAGCA	GCTGGCGGAAAGACACTAAG