Published in final edited form as:

Science. 2015 October 16; 350(6258): aac4383. doi:10.1126/science.aac4383.

Structural basis of histone H3K27 trimethylation by an active polycomb repressive complex 2

Lianying Jiao and **Xin Liu***

Cecil H. and Ida Green Center for Reproductive Biology Sciences and Division of Basic Research, Department of Obstetrics and Gynecology and Department of Biophysics, University of Texas Southwestern Medical Center, Dallas, TX 75390, USA

Abstract

Polycomb repressive complex 2 (PRC2) catalyzes histone H3K27 trimethylation (H3K27me3), a hallmark of gene silencing. Here we report the crystal structures of an active PRC2 complex of 170 kilodaltons from the yeast *Chaetomium thermophilum* in both basal and stimulated states, which contain Ezh2, Eed, and the VEFS domain of Suz12 and are bound to a cancer-associated inhibiting H3K27M peptide and a S-adenosyl-L-homocysteine cofactor. The stimulated complex also contains an additional stimulating H3K27me3 peptide. Eed is engulfed by a belt-like structure of Ezh2, and Suz12(VEFS) contacts both of these two subunits to confer an unusual split active SET domain for catalysis. Comparison of PRC2 in the basal and stimulated states reveals a mobile Ezh2 motif that responds to stimulation to allosterically regulate the active site.

A tripartite gene silencing complex

The formation of specialized cell types during development involves the silencing of genes not required in those cell types. An important player in this silencing process is the polycomb repressive complex 2 (PRC2), which methylates histone H3 on lysine residue 27 (H3K27me). Jiao and Liu determined the x-ray crystal structure of a functional PRC2 complex from a thermophilic yeast species (see the Perspective by Schapira). The intimate association of the three subunits confers stability to PRC2. The structure also reveals how the reaction product, H3K27me, stimulates PRC2 allosterically and how a cancer-associated histone mutation blocks the PRC2 active site.

Polycomb-group proteins mediate gene silencing as multisubunit protein complexes by modifying histone tails and altering high-order chromatin structure. Polycomb repressive complex 2 (PRC2) catalyzes trimethylation of histone H3 at lysine 27 (H3K27me3), an epigenetic hallmark of repressed chromatin (1–6). PRC2 consists of four core subunits—Ezh2, Eed, Suz12, and Rbbp4. In addition, auxiliary subunits—such as Aebp2, Jarid2, and mammalian orthologs of *Drosophila melanogaster* polycomb-like (Pcl) protein (Phf1, Phf19,

*Corresponding author. xin.liu@utsouthwestern.edu.

SUPPLEMENTARY MATERIALS

www.sciencemag.org/content/350/6258/aac4383/suppl/DC1

Figs. S1 to S16

Table S1

References (45, 46)

and Mtf2)—associate with the core PRC2, modulate its enzyme activity, and facilitate its recruitment to target genomic loci (1–3, 5, 6).

A catalytic SET [su(var)3-9, enhancer-of-zeste and trithorax] domain is located at the C terminus of Ezh2, which minimally requires the Eed subunit and the VEFS [Vrn2-Emf2-Fis2-Su(z)12] domain from the C terminus of Suz12 to confer catalytic activity toward H3K27me3 (7). PRC2 is likely also responsible for the deposition of mono- and dimethyl marks on H3K27 (8). Whereas H3K27me1 is accumulated within actively transcribed genes, H3K27me2 is pervasive throughout large chromatin domains (8).

The end product of PRC2 catalysis, H3K27me3, interacts with Eed and stimulates the successive methyltransferase activity of PRC2, a mechanism believed to account for the propagation of the repressive H3K27me3 histone mark and hence the spreading of the facultative heterochromatin (9–11). Moreover, chromatin context–dependent regulation of the H3K27me3 deposition by methylated Jarid2 is achieved, at least in part, with a similar Eed-bridged mechanism (12).

PRC2 and, in particular, some of its core components have been previously subjected to structural analyses. An earlier negative-stain electron microscopy study defined the overall structural architecture of a human holo-PRC2 (i.e., Ezh2-Eed-Suz12-Rbbp4-Aebp2) at low resolution (13). In addition, a series of crystal structures of Eed in complex with an Ezh2 peptide and a variety of trimethylated histone and nonhistone peptides highlighted the critical roles of Eed in mediating Ezh2 binding and allosteric regulation of PRC2 (10–12, 14). Moreover, the recent crystal structures of an isolated inactive catalytic domain of Ezh2 revealed an autoinhibited conformation, implying that structural rearrangement of this domain is likely required for an active PRC2 (15, 16). Finally, the crystal structures of Nurf55, a *Drosophila* homolog of Rbbp4, bound to a Suz12 peptide and an N-terminal histone H3 peptide provided insights into PRC2 association with nucleosomal substrates (17).

Aberrant PRC2 activity, in particular that caused by Ezh2 mutations, is broadly linked to human diseases, including hematological malignancies and Weaver syndrome (18–21). In addition, a histone H3K27M missense mutation is present in some pediatric brain cancers and leads to a global decrease in the amount of H3K27me3, possibly by targeting the catalytic domain of Ezh2 to mediate PRC2 enzyme inhibition (22). Mechanistic understanding of the complex PRC2 function and regulation has so far been limited. We report here the high-resolution crystal structures of an active PRC2 from a thermophilic yeast, *C. thermophilum* (hereafter referred to as "*ct*"), containing the Ezh2-Eed-Suz12(VEFS) ternary complex of 170 kD in both basal and stimulated states.

Characterization and crystallization of the active *ct*PRC2 complex

On the basis of sequence alignment with the corresponding human copies, we identified sequences of Ezh2, Eed, and Suz12(VEFS) for *C. thermophilum* (fig. S1) (23). We overexpressed the Ezh2-Eed-Suz12(VEFS) ternary complex in *Saccharomyces cerevisiae*. The regions of Ezh2 (residues 191 to 950) and Suz12(VEFS) (residues 530 to 691) were

chosen for complex formation on the basis of their sequence homology to the full-length human Ezh2 and the reported minimal *Drosophila* Suz12 construct required for PRC2 catalysis, respectively (7). To maintain the stoichiometry of the complex, these Ezh2 and Suz12(VEFS) regions were expressed as a single fusion protein, which was coexpressed with Eed (residues 1 to 565) to allow reconstitution and purification of the ternary complex. We established biochemical assays to show that the ternary complex, but not the binary Ezh2-Eed or Ezh2-Suz12(VEFS) complexes, displayed robust enzyme activity toward H3K27me3 (fig. S2, A and B). In addition, we demonstrated that the basal enzyme activity of the complex is stimulated by an H3K27me3 product peptide and inhibited by a substrate H3 peptide harboring a K27M cancer mutation (fig. S2C), indicating functional conservation of *ct*PRC2 with its human and *Drosophila* orthologs. Indeed, a stimulating H3K27me3 product peptide, a cancer-associated inhibiting H3K27M peptide, and a S-adenosyl-L-homocysteine (SAH) cofactor were required for generating initial crystallization hits and for crystal refinement. Whereas the PRC2 complex lacking the H3K27me3 peptide was less prone to crystallize, omitting the H3K27M peptide from the complex made it impossible for us to generate crystals of sufficient quality for structural determination.

The overall structure

The crystal structure of the Ezh2-Eed-Suz12(VEFS) ternary complex bound to a stimulating H3K27me3 product peptide was determined at 2.3 Å resolution, and that of the same complex lacking the peptide and representing the basal state was determined at 2.7 Å resolution (Fig. 1, A and B, and table S1). An SAH cofactor and an inhibiting H3K27M peptide bound to the catalytic SET domain were also captured in both crystal structures (Fig. 1, A and B). Both complexes adopt a similar overall structure except for an exposed Ezh2 motif, which directly binds to the stimulating peptide and becomes disordered in its absence. Most of the structural analyses are for the stimulated complex at 2.3 Å resolution, except for when the structures of the two states are compared.

Nearly 17,000 $Å^2$ of solvent-accessible surface area is buried upon formation of the ternary complex. The complex adopts an overall compact structural arrangement, which is roughly divided into the regulatory moiety and the catalytic moiety (Fig. 1B). Sites of stimulation and catalysis are separated by more than 30 Å, suggesting an allosteric mechanism for product stimulation through long-range communication. The atomic structure of the complex fits well into corresponding regions of the reported negative-stain electron microscopy map of a human holo-PRC2 (13), suggesting conservation across species of the overall structural architecture (fig. S3).

Ten structurally and functionally discrete Ezh2 domains are dispersed across the entire active complex and compose an extended structural scaffold to accommodate juxtaposed Eed and Suz12(VEFS) (Fig. 1, B and C). The N terminus of Ezh2 forms an intramolecular interaction with the SANT1-like (SANT1L) domain of Ezh2 and is referred to as the SANT1L-binding domain (SBD). The Eed-binding domain (EBD) was previously characterized for mouse Ezh2 (14) and is found next to the SBD in the current structure, where the EBD similarly occupies a surface groove across the bottom face of the Eed WD40 domain (24) (fig. S4). Continuation of the EBD on the side face of Eed leads to the β-

addition motif (BAM), consisting of three β strands that are added to the β-propeller fold of the WD40 repeats of Eed. Following the BAM, an Ezh2 loop region migrates away from the Eed surface, extends to the back of the SET domain of the catalytic moiety, and is referred to as SET activation loop (SAL). The buried SAL emerges from between the SET domain and Suz12(VEFS) to directly connect to the stimulation-responsive motif (SRM), which sits on the stimulating peptide at the center of the top face of Eed and forms a sandwich-like assembly with Eed and the H3K27me3 peptide. Following the SRM, the SANT1L domain adopts a helix bundle structure, analogous to the predicted SANT1 domain of human Ezh2 (fig. S1), and binds to the SBD. Together, the SBD, EBD, BAM, SAL, SRM, and SANT1L regions complete the belt-like structure of Ezh2 that surrounds Eed (Fig. 1C and fig. S5).

Entering the catalytic moiety, the SANT2-like (SANT2L) domain is linked to the SANT1L domain through the motif connecting SANT1L and SANT2L (MCSS), which harbors two helices, a zinc-binding motif, and a β hairpin, and interacts with multiple surfaces. The SANT2L domain coordinates a zinc ion as well and is analogous to the SANT2 domain of human Ezh2 based on secondary-structure prediction (fig. S1). The CXC and SET domains resemble human counterparts (15, 16). The CXC domain contains a Zn_3Cys_8His and a Zn_3Cys_9 cluster. The SET domain is positioned above Eed and Suz12(VEFS) and adjacent to the SRM. The overall structure of the active ternary complex, and in particular the location of the catalytic domain, underlies allosteric modulation of the enzyme activity through various functional surfaces within the complex.

Association of the ternary complex

Comparison of the current structure to the structure of the mouse Eed-EBD binary complex reveals that the helix and loop regions of the EBD of Ehz2, which vary in lengths, govern Eed binding on its bottom face in both structures (Fig. 2A; the close-up views hereafter are mapped to the overall structure as in fig. S6) (14). The seven-bladed β-propeller WD40 domains of Eed from both species are well aligned with a root mean square deviation (RMSD) of 0.9 Å, except for a yeast-specific insertion domain, which protrudes away from the WD40-containing region (Fig. 2A). Following the EBD, the three β strands of the BAM are added to the first and seventh blade of the β propeller, further securing the Ezh2-Eed association (Fig. 2A).

Both SBD and SANT1L of Ezh2 abut Eed, and the SBD-SANT1L binary complex appears to tighten the Ezh2 belt-like structure around Eed like a buckle (Fig. 2B and fig. S5). Compared to an SBD-like region of human Ezh2 that also immediately precedes the EBD, some hydrophobic residues of the SBD are aligned and involved in binding to a surface groove on the SANT1L domain (Fig. 2B and fig. S1). Although poorly conserved in sequence, the SANT1L shares a key feature of the canonical SANT domain, which is the enrichment of bulky aromatic and hydrophobic residues in the core of the α-helix bundle (Fig. 2B) (25). The SBD-SANT1L intramolecular interaction of Ezh2 may provide stability for Eed in the complex in addition to that offered by the EBD and BAM of Ezh2. Moreover, because the SANT1L is directly connected to the SRM of Ezh2 at the site of stimulation, the SBD-SANT1L assembly may play a role in enzyme regulation as well. Ezh2-S21, a previously mapped phosphorylation site of human Ezh2, is located in the middle of the SBD

based on sequence and structural homology, raising a possibility that the reported diminution of enzyme activity for the Ezh2-S21–phosphorylated human PRC2 is caused allosterically by a disturbed SBD-SANT1 interaction (26).

Suz12(VEFS) is located in the space behind the MCSS, SANT2L, CXC, and SET of Ezh2 (Figs. 1C and 2, C and D). Similar to the canonical SANT domain, the SANT2L structurally aligns well with a Myb DNA binding domain, except that the SANT2L also contains a zinc-binding motif (fig. S7). While the helical contents of Suz12(VEFS) form the 10-helix bundle with the MCSS and SANT2L, the N-terminal loop region of Suz12(VEFS) cements together Eed and the SAL and SET regions of Ezh2 (Figs. 1C and 2D). Disease-causing mutations of several conserved residues of Suz12(VEFS)—including L593, F596, D598, E603, W610, andD611 (numbered by *ct*Suz12) (fig. S1)—result in impaired complex integrity and weakened methyltransferase activity (27–29). These residues are scattered around the vicinity of the MCSS of Ezh2, in particular its conserved zinc-binding motif, and make various interactions in the catalytic moiety (Fig. 2C). The previously proposed noncoding RNA binding region of Ezh2 (30), though not conserved in the current structure, is projected to be situated in a conserved zinc-coordinating loop of the SANT2L domain, neighboring to Suz12(VEFS) and the MCSS (figs. S1, S6A, and S7).

Although the Eed-Suz12(VEFS) interface is relatively small compared with that for the Ezh2-Eed and Ezh2-Suz12(VEFS) pairs, it nonetheless appears to be important for optimal enzyme activity of PRC2. Conserved W584 residue at the junction of the helical and loop contents of Suz12(VEFS) is involved in hydrophobic interaction with Y231 and Y232 of Eed (Fig. 2D), and mutation of this tryptophan residue leads to a reduced enzyme activity of *Drosophila* PRC2 (29).

Split catalytic domain

The catalytic SET domain of Ezh2 preserves most of the essential sequence and structural features observed for other SET-containing methyltransferases (fig. S8). An isolated CXC-SET domain of human Ezh2 displays an autoinhibited state of the SET domain (15, 16). The SET-I and post-SET regions adopt atypical conformations, with an inaccessible substrate-binding groove and an incomplete cofactor-binding pocket, which is in line with the requirement of Eed and Suz12(VEFS) for catalysis. Structural comparison of the inactive SET domain with the current structure captured in the active state provides a mechanism for SET activation. The overall structures of the SET domains in both states are aligned with an RMSD of 1.6 Å (Fig. 3, A and B). During activation, the SET-I region is rotated over 20° counterclockwise, such that it swings away from the otherwise blocked peptide substrate-binding groove on one side and close to the cofactor-binding pocket on the other side (Fig. 3B). The conformational change of SET-I exposes the peptide substrate binding site and results in the completion of the cofactor-binding pocket, with placement of the post-SET region in a typical conformation for cofactor SAH binding (Fig. 3B). Specifically, in the active state of PRC2, the cofactor SAH is stabilized by extensive hydrophobic and hydrogen-bonding interactions with residues from all subregions of the Ezh2-SET domain, including the SET-N, SET-I, SET-C, and post-SET (Fig. 3C and fig. S8). Many of these

residues may also be involved in orienting the methyl donor S-adenosylmethionine (SAM) to allow catalysis.

The SET activation loop SAL from the N-terminal portion of Ezh2 sequence extends from the Eed surface toward the SET domain and exits the catalytic moiety along the side of the last β strand of SET-N (Fig. 3A). The SAL of Ezh2 is engaged in extensive interactions with the SET, the MCSS, Eed, and Suz12(VEFS), whereby it maintains the active conformation of the SET domain (fig. S9). Consistently, deletion or mutation of a number of six semi-conserved amino acids (residues 310 to 315) in the SAL abolishes enzyme activity completely, with little effect on complex assembly (Fig. 3D). Furthermore, alteration of only three amino acids in this region (residues 313 to 315) is sufficient to abrogate catalysis (fig. S10). Therefore, at least two split regions of the Ezh2 sequence, the SAL segment and the SET domain, are required to compose the active catalytic domain (Figs. 1A and 3, A and B). Further structural alignment with SET domains from other methyltransferase families reveals that the SAL mimics a similar loop structure found in the N-flanking regions of some SET domains, including G9a, Ash1l, and Atxr5 (fig. S11). These N-flanking regions diverge structurally and are thought to stabilize SET (31); however, whether the SAL-aligned regions in those SET domains are similarly required for catalysis remains to be tested. Lastly, whereas Eed seems to be critical for the overall positioning of the SAL onto the back of the SET domain, the MCSS and Suz12(VEFS) may guide formation of the unique local SAL conformation for SET interaction (Fig. 3, A and B, and fig. S9).

Inhibition of PRC2 active site by the H3K27M cancer mutant

A histone H3 substrate peptide harboring a K27M cancer mutation is captured bound to the canonical substrate-binding cleft of the SET domain. This mutant peptide was shown to inhibit enzyme activity of human PRC2 by targeting Ezh2 directly (22). Residue H3R26 of the mutant peptide occupies the lysine-binding channel of the active site and occludes substrate binding (Fig. 3E), reminiscent of autoinhibition of human Setd2 mediated by an arginine residue from the post-SET loop (32). Whereas the aliphatic portion of H3R26 side chain is surrounded by a group of aromatic residues, the guanidinium part is trapped by the invariant Y645 and the carbonyl cage at the active site (Fig. 3E). The same set of residues may also direct H3K27 methylation when wild-type substrate is bound (15, 16, 31). In addition, residue H3A25 sits in a shallow pocket of SET-I, which would exclude a bulkier side chain at this position of H3, a feature also contributing to the binding specificity of the inhibitory H3K27M mutant peptide. Consistently, the H3K27M mutant peptide harboring an additional H3R26A mutation is incapable of mediating PRC2 enzyme inhibition, confirming the critical role of residue H3R26 in such a process in solution (Fig. 3F).

The side chain of the mutated methionine is not resolved in the current structure, leaving open the question of why a H3K27M mutation is more potent than other amino acid substitutions for PRC2 inhibition (fig. S12) (22).

Mobile stimulation-responsive motif

The PRC2 enzyme activity is stimulated by its own reaction product, generating a positive feedback loop. The H3K27me3 product peptide is directly recognized by theWD40 region of Eed and leads to allosteric stimulation of PRC2 activity, a mechanism believed to account for propagation of the H3K27me3 histone marks in the repressive chromatin domains (9–11). Moreover, K116 of Jarid2 is also trimethylated by PRC2 and is thought to facilitate H3K27me3 deposition at loci devoid of such a histone mark in a similar manner (12). However, the mechanism whereby Ezh2, the catalytic subunit of PRC2, receives and translates the stimulating signal remains unknown. The crystal structure of the H3K27me3-bound ternary complex, together with the basal state of the complex in the absence of the stimulating peptide determined at 2.7 Å resolution, provides a molecular mechanism for this allosteric regulation. In the stimulated state, the H3K27me3 peptide is sandwiched between Eed and an exposed Ezh2 motif, which immediately follows the SAL of Ezh2 and is referred to as the stimulation-responsive motif (SRM) (Fig. 1, A and B). The SRM is not visible in the electron density map of the ternary complex in the basal state, suggesting that it is highly flexible in the absence of the stimulating H3K27me3 peptide (fig. S13). Because both basal and stimulated PRC2 complexes were crystallized in the same crystal lattice under the same crystallization conditions, it is unlikely that the observed dramatic conformational change of the SRM of Ezh2 is due to differences in crystal packing. A structural alignment indicates that the entire catalytic moiety—including the MCSS, SANT2L, CXC, SET, and Suz12(VEFS)—is rotated counterclockwise toward the SRM along the interface of the two moieties from the basal to stimulated state, which implies that a large domain-wise motion is likely associated with stimulation in solution (Fig. 4A).

Analysis of the sandwich structure at the site of stimulation reveals that the trimethylated H3K27 side chain is bound within an aromatic cage on the Eed surface, akin to that in the previous crystal structure of the Eed-H3K27me3 binary complex (fig. S14). Main-chain atoms of the A25, K27me3, and S28 residues of the H3K27me3 peptide provide a unique docking surface for the SRM of Ezh2, by making hydrogen-bonding interactions with the side chains of H326 andD329 of the SRM (Fig. 4B). Conserved residues from this region of the SRM, including P325 and L327, are hotspots for disease mutations in human Ezh2 (21, 33).

Most of the sequence contents, including the signature A-R-Kme3-S sequence of the H3K27me3 peptide, are discriminated by Eed through hydrogen bonding or steric exclusion, except that residue H3A24 is bound by a part of the SRM surface that excludes bulky side chains (Fig. 4B). This structural mechanism may account for the observed specificity of enzyme stimulation. A linker histone H1K26me3 peptide contains a bulkier lysine residue at the equivalent position of residue H3A24 relative to the trimethylated lysine (fig. S15). Although the H1K26me3 peptide binds to Eed with an affinity comparable to that of the stimulating H3K27me3 peptide, it is incompetent for PRC2 stimulation (11), probably owing to the steric exclusion that prevents binding of the SRM of Ezh2 to the H1K26me3 peptide.

On the back of the Eed-H3K27me3-SRMsandwich structure and at the border of the two moieties, the SRM makes extensive interactions, mostly hydrophobic ones, with the helix region of SET-I (Fig. 4C). A short loop region following the SET-I helix is also involved in binding of the SRM and becomes disordered together with the latter in the basal state (Fig. 4C).

The SRM of Ezh2 establishes a possible allosteric pathway to communicate with the catalytic site for stimulation, which involves the H3K27me3-SRM and SRM-SET-I interactions. At the H3K27me3 binding site, mutations of P325, H326, and D329 of the SRM largely abolish enzyme stimulation while resulting in moderately reduced basal activities, likely due to an impaired binding of the SRM to the docking surface on the stimulating peptide (Fig. 4, C and D, and fig. S16). Furthermore, we show that deletion of L350 in the SRM on the SRM-SET-I binding interface slightly decreases the H3K27me3-mediated stimulation as well (Fig. 4, C and D, and fig. S16). Along this pathway, R839 and D868 of SET-I form a strictly conserved salt bridge, connecting the endpoint of the pathway directly to the catalytic site (Fig. 4C and fig. S8). Mutating R839 to an aspartate blocks stimulation completely (Fig. 4D and fig. S16). A subtle conformational change of the active site may have a large effect on catalysis. Indeed, an A677G cancer-related mutation of human Ezh2 results in an enhanced catalytic efficiency in converting H3K27me2 to H3K27me3 without compromising the mono- and dimethylation steps, probably owing to a slightly expanded substrate lysine channel that is otherwise spatially restrictive for lysine trimethylation (34). A677 of human Ezh2 is the equivalent of A869 in yeast Ezh2 in the current structure, and in this regard, allosteric effects exerted on D868 through the salt bridge may spread to the next residue A869 to allow enzyme stimulation with an analogous mechanism employed by the A677G human cancer-causing mutant.

Allosteric regulation may not be limited to the SRM-mediated pathways. A small acidic patch of Suz12(VEFS) is known to be important for the basal enzyme activity and for the enzymatic stimulation of *Drosophila* PRC2 by a fragment of histone H3 (residues 31 to 42) in a dense chromatin environment (7). Though not conserved in the active ternary complex here, these acidic amino acids next to the W584 residue of Suz12(VEFS) are predicted to be located adjacent to the SAL of Ehz2 on the back of the SET domain based on the current structure (Fig. 2D), through which allosteric stimulation also becomes possible.

Conclusions

Ezh2, Eed, and Suz12(VEFS) associate intimately, and the intramolecular and intermolecular binding interface of these subunits seems to underlie many aspects of enzyme regulation. The ternary complex undergoes structural interconversion between the autoinhibited, basal, and stimulated states of catalysis, during which local and global conformational changes are observed. In addition, the SAL and SET regions together compose a split catalytic domain of Ezh2, and this domain is efficiently inhibited by the H3K27M cancer mutant, for which the H3R26 residue plays an important role. The SRM, a structural element of Ezh2, mediates the H3K27me3 product stimulation through an allosteric mechanism. Both SRM and SAL bind to Eed and Suz12(VEFS), transmit regulatory signals from different distant surfaces, and are therefore at the center of PRC2

enzyme regulation pathways. Cooperative and antagonizing allosteric effects are integrated at the active site and result in distinct cellular output.

PRC2 displays a high degree of regulatory plasticity through numerous auxiliary factors. Many of these factors act on Ezh2 directly to modulate PRC2 enzyme activity. The SRM is well positioned on the surface of PRC2 to allow access by other nuclear factors. In addition, at least one of the SRM-mediated allosteric pathways involving a conserved salt bridge of the SET-I region may tip the balance of the three methylation states of H3K27 toward H3K27me3, the functional histone mark for gene repression, in a manner similar to that of the A677G cancer mutation of human Ezh2. Lastly, the SRM and Eed together dictate the specificity and possibly the strength of stimulation, which adds a second layer of control for H3K27me3 deposition besides that offered by the catalytic SET domain. Since allosteric stimulation of PRC2 is probably necessary for formation of the repressive H3K27me3 domains (12), H3K27 trimethylation initiated at the SET domain of Ehz2 may be promoted or suppressed by distinct cellular mechanisms through the SRM surface of Ehz2 during the stimulation step to generate desired biological outcomes.

Dinucleosomes with four copies of histone H3 are known to be much better substrates for PRC2 than mononucleosomes (7, 35, 36). Coincidentally, this and previous studies have suggested that at least four distinct PRC2 surfaces are available to interact with histone H3, including the SET domain of Ezh2 for an H3K27 substrate, the SRM of Ezh2 and Eed for a stimulating H3K27me3 histone tail (10), Suz12(VEFS) for residues 31 to 42 of a neighboring histone H3 in a dense chromatin environment (7), and Rbbp4 for an unmodified N terminus of histone H3 (17). Therefore, a dinucleosome may represent a minimal unit for optimal chromatin context-dependent PRC2 binding and catalysis (13). The current structure reveals the spatial arrangement of the first three H3-binding surfaces of PRC2 mentioned above, which may correlate with a defined conformation of a dinucleosome substrate with a preferred linker DNA length for PRC2 catalysis (7). Mechanistic analysis of PRC2-mediated methylation of chromatin templates will entail high-resolution structural information of PRC2 fully engaged on a dinucleosome substrate.

Materials and methods

Cloning

DNA encoding Ezh2, Eed, and Suz12 was amplified from the extracted genomic DNA of *Chaetomium thermophilum* var. *thermophilumDSM1495* or synthesized by Integrated DNA Technologies (IDT). A DNA fragment corresponding to the sequence, 2×Protein A-SSGENLYFQSNHHHHHHA, was added N-terminal to the fused Ezh2 (residues 191 to 950)-LVPRGS-Suz12 (residues 530 to 691) (hereafter referred to as Ezh2-VEFS) fragment, and the Ezh2-VEFS construct was sub-cloned into the p416GAL1 vector. Full-length Eed (residues 1 to 565) was tagged with an N-terminal twin-strep tag (IBA Lifesciences) and was cloned into a modified p416GAL1 plasmid with the selection marker replaced by TRP1. All the constructs harboring mutations were generated with the site-directed mutagenesis strategy.

Protein expression and purification

Constructs of Ezh2-VEFS and EED were cotransformed into *S. cerevisiae* CB010 strain (*MATa pep4::HIS3 prb1::LEU2 prc::HISG can1 ade2 trp1 ura3 his3 leu2-3,112*). Cell starters were grown in synthetic complete drop out medium without uracil and tryptophan (Sc-Ura-Trp) supplemented with 2% raffinose at 30°C for 24 hours. Protein expression was induced by inoculating the cell starters into the Sc-Ura-Trp medium supplemented with 2% galactose at 30°C for an additional 24 hours. Cells were harvested by centrifugation and resuspended in lysis buffer containing 100 mM Tris-HCl (pH 8.5), 300 mM NaCl, 2.5 mM dithiothreitol (DTT), 10% glycerol, 0.5 mM phenylmethylsulfonyl fluoride, and protease inhibitors. After lysis by bead beating and clarification by centrifugation, the supernatant was mixed with equilibrated immunoglobulin G (IgG) resin and allowed to bind for 2 hours at 4°C. The IgG resin was made by coupling purified rabbit IgG to the CNBr-activated Sepharose 4 Fast Flow resin according to the manufacturer's protocol (GE Healthcare). The binding mixture was then applied to an empty column to capture the IgG resin with the bound ternary complex. The resin was washed stepwise by buffer A [50 mM Tris-HCl (pH 8.0), 500 mM NaCl, 2.5 mM DTT, 10% glycerol, and 0.1% NP40], buffer B [50 mM Tris-HCl (pH 8.0), 1 M NaCl, 2.5 mM DTT, and 10% glycerol), and buffer C [50 mM Tris-HCl (pH 8.0), 100 mM NaCl, 2.5 mM DTT, and 10% glycerol]. Tobacco etch virus (TEV) protease was added to the column and mixed thoroughly with the resin. After cleavage at 4°C for 12 hours, the flow-through was collected and the protein was eluted with an additional five bed volumes of buffer C. The eluted protein was then applied to a gravity column with Strep-Tactin resin equilibrated with buffer C. A quick buffer C wash was performed, and the ternary complex was eluted with buffer C supplemented with 5 mM d-desthiobiotin. Protein-containing fractions were pooled, concentrated, and loaded onto a Superdex S200 preparative size exclusion chromatography column (GE Healthcare) equilibrated with 20 mM Tris-HCl (pH 8.0), 100 mM NaCl, and 2 mM DTT. The peak fractions were concentrated to about 20 mg/ml and stored at −80°C. Thirty liters of yeast cells will typically yield 10 mg of purified ternary complex.

Crystallization and structure determination

Prior to crystallization, two peptides, H3K27M (KQLATKAAR**M**SAPATGGVKK) and H3K27me3 (TKAAR**Kme3**SAPAT) (the mutated or modified residue is highlighted in bold), and SAH were added to the protein in a molar ratio of 1:5:5:10 (protein:H3K27M:H3K27me3:SAH). Initial crystallization conditions were screened by sitting drop vapor diffusion method at 22°C. Crystals were refined by mixing equal volumes of the protein solution at 10 to 25 mg/ml with the reservoir solution in 2-μl drops over 500 μl of the latter, containing 15% polyethylene glycol (PEG) 4000 and 175 mM ammonium citrate (pH 7.0). Crystals were harvested after 2 to 3 weeks, cryoprotected by mother liquor with 35% xylitol, and flash frozen in liquid nitrogen.

Diffraction data were collected at Advanced Light Source (ALS) beamline BL5.0.2 and Advanced Photo Source (APS) beamline 19ID. Data sets were indexed, integrated, and scaled with HKL2000 (37) and further processed with the CCP4 suite of programs (38). The structure was determined in $P2_1$ space group with one complex in an asymmetric unit by the single-wavelength anomalous dispersion method with data collected from platinum

derivatized crystals. Briefly, Pt sites were identified with Phenix_autosol (39) and a partial model was generated by Phenix_autobuild (39). Further model building and iterative refinement were carried out with Coot (40), Phenix_refine (39), Refmac (41) and autoBUSTER (42). Phases were extended to native data sets by molecular replacement with Phaser (43). The final models were obtained by using TLS (translation, libration and screw-motion) refinement with autoBUSTER (42). Molecular graphics were generated with Pymol (44). Statistics for data collection, phase calculation, and refinement are summarized in table S1.

Histone methyltransferase assay

Mutant and wild-type ternary complexes for the methyltransferase assay were purified from100-ml cell culture with a protocol similar to that described above. Concentration of the purified proteins was normalized visually by SDS–polyacrylamide gel electrophoresis (SDS-PAGE). Proteins were aliquoted, flash frozen, and stored at −80°C. For assays with Western blotting, about 0.4 μg of purified ternary complex was incubated with 600 nM histone H3 and 32 μM SAM in 20 μl of histone methyltransferase assay buffer [50 mM Tris-HCl (pH 8.0), 100 mM NaCl, 2.5 mM $MgCl_2$, 1 mM EDTA, and 2.5 mM DTT] at 30°C for 2 hours. The reaction was stopped by adding 6 μl of 4× sample loading dye and boiling at 95°C for 5 min. Trimethylated H3 product was separated by SDS-PAGE, transferred to polyvinylidene difluoride membrane, and detected by Western blotting with anti-H3K27me3 primary antibody (Cell Signaling, catalog no. 9733S). For assays with the enzyme-linked immunosorbent assay (ELISA) method, 0.2 or 0.4 μg of enzyme was incubated with 1 μM biotinylated histone H3 (residues 21 to 44) peptide (AnaSpec, catalog no. AS-64641) and 32 μM SAM in 20 μl of the histone methyltransferase assay buffer with or without 50 or 100 μM synthesized stimulating or inhibiting histone peptides at 30°C for 2 hours. The reaction was stopped by adding 80 μl of 100 mM sodium acetate (pH 4.5). Methylated product was captured by a streptavidin-coated plate (Pierce, catalog no. 15502) and detected by standard ELISA procedure with the same anti-H3K27me3 primary antibody as used for the Western blot method.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

This research was supported by Welch Foundation research grant I-1790, CPRIT research grant R1119, Rita Allen Foundation research grant, University of Texas Southwestern Medical Center Endowed Scholar fund, and NIH grant GM114576 to X.L. X.L. is a W. W. Caruth, Jr. Scholar in Biomedical Research. This research also received support from the Cecil H. and Ida Green Center Training Program in Reproductive Biology Sciences Research. This research used resources of the Advanced Photon Source, a U.S. Department of Energy (DOE) Office of Science User Facility operated for the DOE Office of Science by Argonne National Laboratory under contract no. DE-AC02-06CH11357. The Advanced Light Source is supported by the Director, Office of Science, Office of Basic Energy Sciences, of the U.S. DOE under contract no. DE-AC02-05CH11231. We thank R. Kornberg, R. Marmorstein, and L. Kraus for critical reading of the manuscript. The crystal structures described in this study have been deposited in the Protein Data Bank under accession number 5CH1 (PRC2 in the stimulated state) and 5CH2 (PRC2 in the basal state).

REFERENCES AND NOTES

1. Cao R, Zhang Y. The functions of E(Z)/EZH2-mediated methylation of lysine 27 in histone H3. Curr. Opin. Genet. Dev. 2004; 14:155–164. pmid: 15196462. [PubMed: 15196462]
2. Simon JA, Kingston RE. Mechanisms of polycomb gene silencing: Knowns and unknowns. Nat. Rev. Mol. Cell Biol. 2009; 10:697–708. pmid: 19738629. [PubMed: 19738629]
3. Margueron R, Reinberg D. The Polycomb complex PRC2 and its mark in life. Nature. 2011; 469:343–349. pmid: 21248841. [PubMed: 21248841]
4. Delest A, Sexton T, Cavalli G. Polycomb: A paradigm for genome organization from one to three dimensions. Curr. Opin. Cell Biol. 2012; 24:405–414. pmid: 22336329. [PubMed: 22336329]
5. Di Croce L, Helin K. Transcriptional regulation by Polycomb group proteins. Nat. Struct. Mol. Biol. 2013; 20:1147–1155. pmid: 24096405. [PubMed: 24096405]
6. Schwartz YB, Pirrotta V. A new world of Polycombs: Unexpected partnerships and emerging functions. Nat. Rev. Genet. 2013; 14:853–864. pmid: 24217316. [PubMed: 24217316]
7. Yuan W, et al. Dense chromatin activates Polycomb repressive complex 2 to regulate H3 lysine 27 methylation. Science. 2012; 337:971–975. pmid: 22923582. [PubMed: 22923582]
8. Ferrari KJ, et al. Polycomb-dependent H3K27me1 and H3K27me2 regulate active transcription and enhancer fidelity. Mol. Cell. 2014; 53:49–62. pmid: 24289921. [PubMed: 24289921]
9. Hansen KH, et al. A model for transmission of the H3K27me3 epigenetic mark. Nat. Cell Biol. 2008; 10:1291–1300. pmid: 18931660. [PubMed: 18931660]
10. Margueron R, et al. Role of the polycomb protein EED in the propagation of repressive histone marks. Nature. 2009; 461:762–767. pmid: 19767730. [PubMed: 19767730]
11. Xu C, et al. Binding of different histone marks differentially regulates the activity and specificity of polycomb repressive complex 2 (PRC2). Proc. Natl. Acad. Sci. U.S.A. 2010; 107:19266–19271. pmid: 20974918. [PubMed: 20974918]
12. Sanulli S, et al. Jarid2 methylation via the PRC2 complex regulates H3K27me3 deposition during cell differentiation. Mol. Cell. 2015; 57:769–783. pmid: 25620564. [PubMed: 25620564]
13. Ciferri C, et al. Molecular architecture of human polycomb repressive complex 2. eLife. 2012; 1:e00005. pmid: 23110252. [PubMed: 23110252]
14. Han Z, et al. Structural basis of EZH2 recognition by EED. Structure. 2007; 15:1306–1315. pmid: 17937919. [PubMed: 17937919]
15. Wu H, et al. Structure of the catalytic domain of EZH2 reveals conformational plasticity in cofactor and substrate binding sites and explains oncogenic mutations. PLOS ONE. 2013; 8:e83737. pmid: 24367611. [PubMed: 24367611]
16. Antonysamy S, et al. Structural context of disease-associated mutations and putative mechanism of autoinhibition revealed by X-ray crystallographic analysis of the EZH2-SET domain. PLOS ONE. 2013; 8:e84147. pmid: 24367637. [PubMed: 24367637]
17. Schmitges FW, et al. Histone methylation by PRC2 is inhibited by active chromatin marks. Mol. Cell. 2011; 42:330–341. pmid: 21549310. [PubMed: 21549310]
18. Simon JA, Lange CA. Roles of the EZH2 histone methyltransferase in cancer epigenetics. Mutat. Res. 2008; 647:21–29. pmid: 18723033. [PubMed: 18723033]
19. Ernst T, et al. Inactivating mutations of the histone methyltransferase gene EZH2 in myeloid disorders. Nat. Genet. 2010; 42:722–726. pmid: 20601953. [PubMed: 20601953]
20. Nikoloski G, et al. Somatic mutations of the histone methyltransferase gene EZH2 in myelodysplastic syndromes. Nat. Genet. 2010; 42:665–667. pmid: 20601954. [PubMed: 20601954]
21. Gibson WT, et al. FORGE Canada Consortium, Mutations in EZH2 cause Weaver syndrome. Am. J. Hum. Genet. 2012; 90:110–118. pmid: 22177091. [PubMed: 22177091]
22. Lewis PW, et al. Inhibition of PRC2 activity by a gain-of-function H3 mutation found in pediatric glioblastoma. Science. 2013; 340:857–861. pmid: 23539183. [PubMed: 23539183]
23. Amlacher S, et al. Insight into structure and assembly of the nuclear pore complex by utilizing the genome of a eukaryotic thermophile. Cell. 2011; 146:277–289. pmid: 21784248. [PubMed: 21784248]

24. Stirnimann CU, Petsalaki E, Russell RB, Müller CW. WD40 proteins propel cellular networks. Trends Biochem. Sci. 2010; 35:565–574. pmid: 20451393. [PubMed: 20451393]
25. Boyer LA, Latek RR, Peterson CL. The SANT domain: A unique histone-tail-binding module? Nat. Rev. Mol. Cell Biol. 2004; 5:158–163. pmid: 15040448. [PubMed: 15040448]
26. Cha TL, et al. Akt-mediated phosphorylation of EZH2 suppresses methylation of lysine 27 in histone H3. Science. 2005; 310:306–310. pmid: 16224021. [PubMed: 16224021]
27. Ntziachristos P, et al. Genetic inactivation of the polycomb repressive complex 2 in T cell acute lymphoblastic leukemia. Nat. Med. 2012; 18:298–301. pmid: 22237151. [PubMed: 22237151]
28. Score J, et al. Inactivation of polycomb repressive complex 2 components in myeloproliferative and myelodysplastic/myeloproliferative neoplasms. Blood. 2012; 119:1208–1213. pmid: 22053108. [PubMed: 22053108]
29. Rai AN, et al. Elements of the polycomb repressor SU(Z)12 needed for histone H3-K27 methylation, the interface with E(Z), and in vivo function. Mol. Cell. Biol. 2013; 33:4844–4856. pmid: 24100017. [PubMed: 24100017]
30. Kaneko S, et al. Phosphorylation of the PRC2 component Ezh2 is cell cycle-regulated and up-regulates its binding to ncRNA. Genes Dev. 2010; 24:2615–2620. pmid: 21123648. [PubMed: 21123648]
31. Xiao B, Wilson JR, Gamblin SJ. SET domains and histone methylation. Curr. Opin. Struct. Biol. 2003; 13:699–705. pmid: 14675547. [PubMed: 14675547]
32. Zheng W, et al. Sinefungin derivatives as inhibitors and structure probes of protein lysine methyltransferase SETD2. J. Am. Chem. Soc. 2012; 134:18004–18014. pmid: 23043551. [PubMed: 23043551]
33. Bejar R, et al. Clinical effect of point mutations in myelodysplastic syndromes. N. Engl. J. Med. 2011; 364:2496–2506. pmid: 21714648. [PubMed: 21714648]
34. McCabe MT, et al. Mutation of A677 in histone methyltransferase EZH2 in human B-cell lymphoma promotes hypertrimethylation of histone H3 on lysine 27 (H3K27). Proc. Natl. Acad. Sci. U.S.A. 2012; 109:2989–2994. pmid: 22323599. [PubMed: 22323599]
35. Kuzmichev A, Jenuwein T, Tempst P, Reinberg D. Different EZH2-containing complexes target methylation of histone H1 or nucleosomal histone H3. Mol. Cell. 2004; 14:183–193. pmid: 15099518. [PubMed: 15099518]
36. Martin C, Cao R, Zhang Y. Substrate preferences of the EZH2 histone methyltransferase complex. J. Biol. Chem. 2006; 281:8365–8370. pmid: 16431907. [PubMed: 16431907]
37. Otwinowski Z, Minor W. Processing of X-ray diffraction data collected in oscillation mode. Methods Enzymol. 1997; 276:307–326.
38. Winn MD, et al. Overview of the CCP4 suite and current developments. Acta Crystallogr. D Biol. Crystallogr. 2011; 67:235–242. pmid: 21460441. [PubMed: 21460441]
39. Adams PD, et al. PHENIX: A comprehensive Python-based system for macromolecular structure solution. Acta Crystallogr. D Biol. Crystallogr. 2010; 66:213–221. pmid: 20124702. [PubMed: 20124702]
40. Emsley P, Cowtan K. Coot: Model-building tools for molecular graphics. Acta Crystallogr. D Biol. Crystallogr. 2004; 60:2126–2132. pmid: 15572765. [PubMed: 15572765]
41. Murshudov GN, Vagin AA, Dodson EJ. Refinement of macromolecular structures by the maximum-likelihood method. Acta Crystallogr. D Biol. Crystallogr. 1997; 53:240–255. pmid: 15299926. [PubMed: 15299926]
42. Bricogne, G., et al. BUSTER version 2.10.1 Ed. Cambridge, UK: Global Phasing.; 2011.
43. McCoy AJ, et al. Phaser crystallographic software. J. Appl. Crystallogr. 2007; 40:658–674. pmid: 19461840. [PubMed: 19461840]
44. The PyMOL Molecular Graphics System, version no. 1.7.4.3. Schrodinger, LLC; 2010.
45. Thompson JD, Higgins DG, Gibson TJ. CLUSTAL W: Improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Res. 1994; 22:4673–4680. pmid: 7984417. [PubMed: 7984417]

46. Pettersen EF, et al. UCSF Chimera—a visualization system for exploratory research and analysis. J. Comput. Chem. 2004; 25:1605–1612. pmid: 15264254. [PubMed: 15264254]

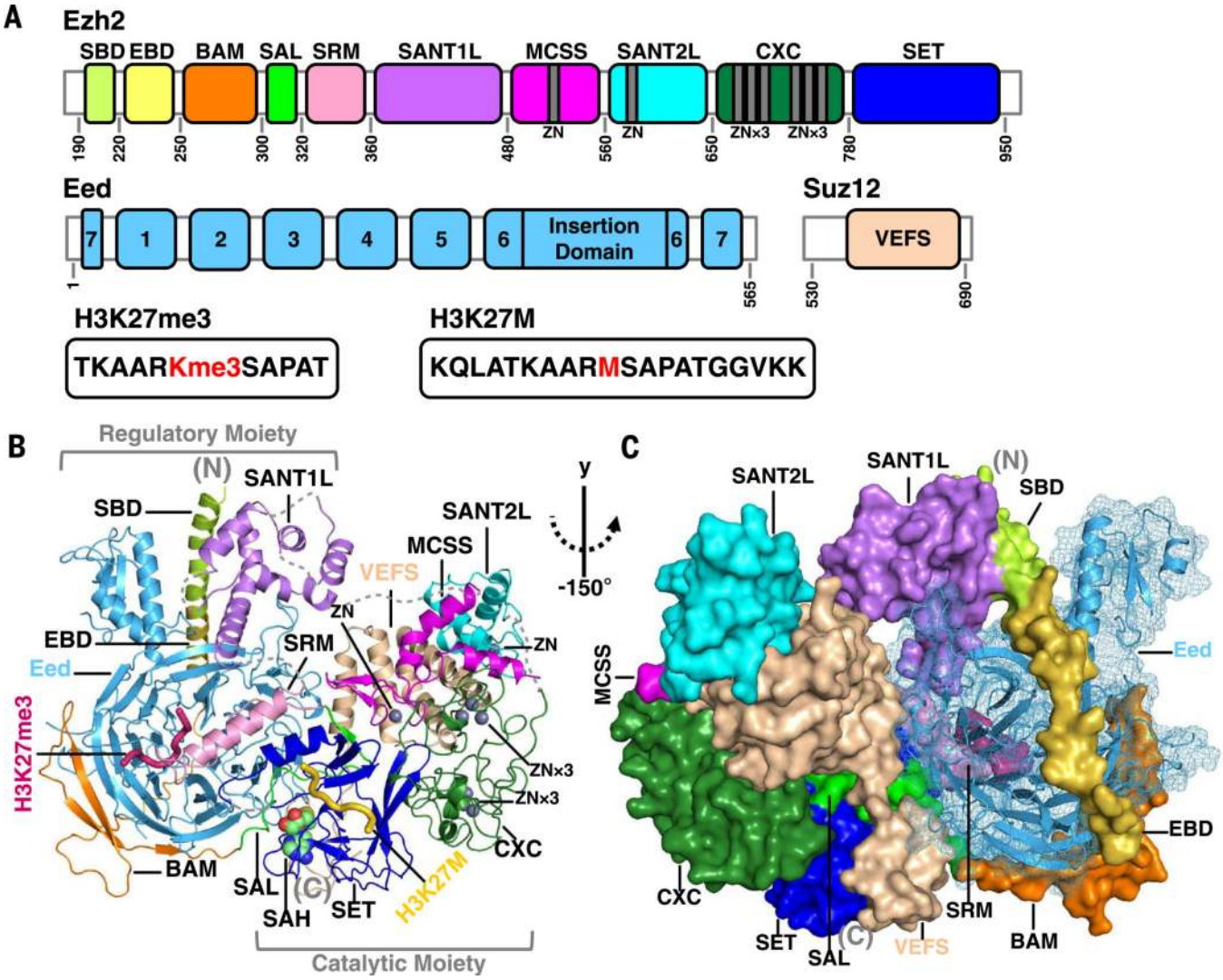

Fig. 1. Overall structure of the active Ezh2-Eed-Suz12 (VEFS) ternary complex

(**A**) Schematic domain structures of Ezh2, Eed, and Suz12 (VEFS) are shown. Individual functional domains of Ezh2 are labeled, and locations of zinc-coordinating structures are indicated. Seven blades of the β-propeller structure of the Eed WD40 domain are numbered, and an insertion domain specific to *C. thermophilum* Eed is indicated. Peptides included for crystallization are also shown, with modified or mutated residues colored in red. (**B**) The overall structure of the active ternary complex bound to the stimulating H3K27me3 peptide, inhibiting H3K27M cancer mutant peptide, and cofactor SAH is shown in cartoon representation. Functional domains are colored as in (A). Zn^{2+} ions and SAH are shown as spheres, and the two peptide ligands are shown as thick ribbons. The regulatory and catalytic moieties of the complex are also indicated. (**C**) Ezh2 and Suz12(VEFS) subunits of the ternary complex are shown in surface representation, and Eed is in mesh. Eed is engulfed by the belt-like structure of Ezh2 within the regulatory moiety.

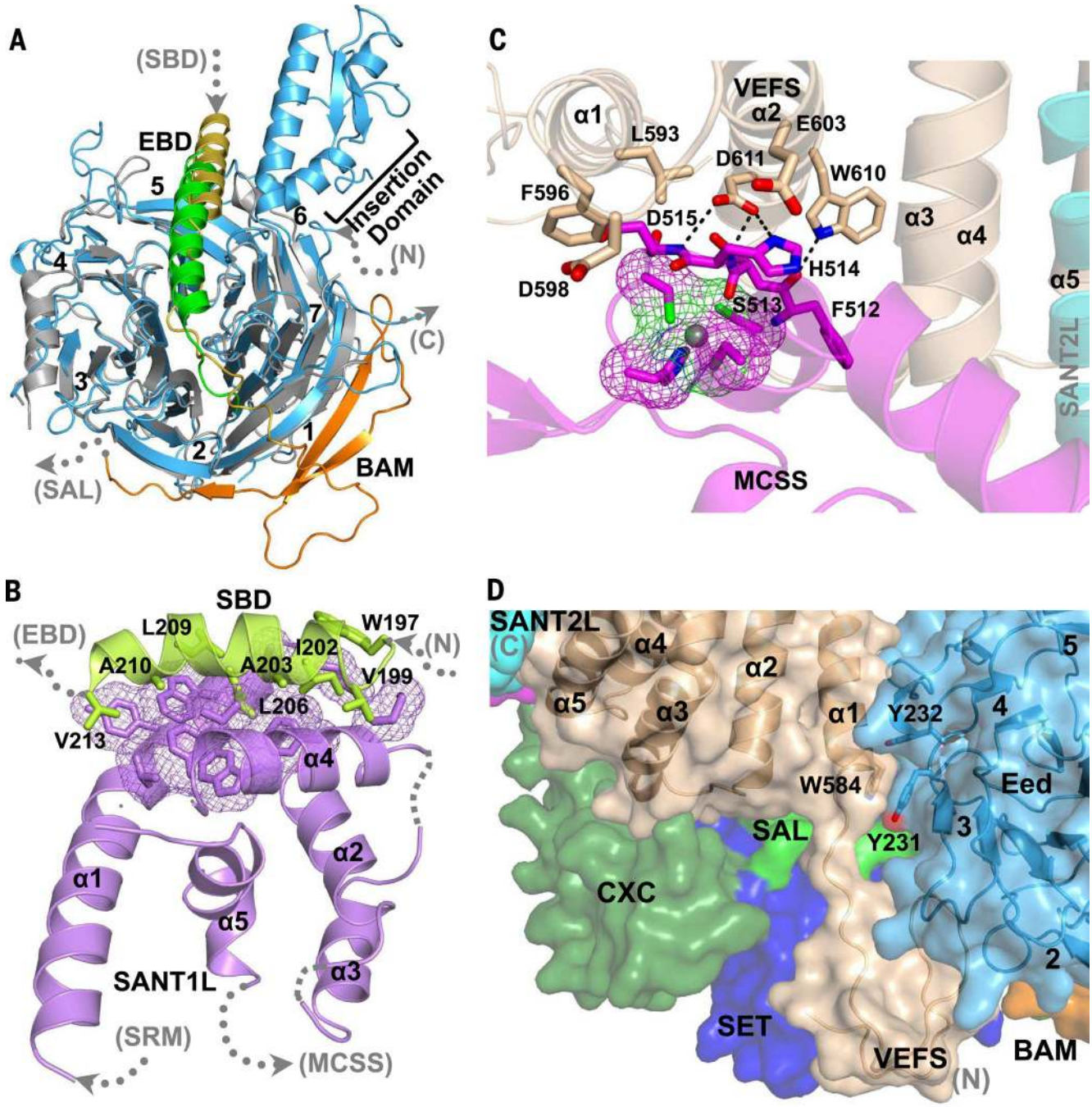

Fig. 2. Close-up views of the intra- and intermolecular interactions in the active ternary complex

(**A**) The crystal structure of the stimulated ternary complex is superimposed with the mouse Eed-EBD binary complex (PDB: 2QXV). Mouse Eed and EBD are shown in gray and green, respectively. The BAM of Ezh2 in orange is also shown interacting with the WD40 domain of Eed. To indicate the close-up views in the context of the overall structure, Ezh2 domains preceding and following the shown ones are hereafter denoted by the dotted gray arrows. A map of the close-up views is also provided in fig. S6 for the same purpose. (**B**) Intramolecular hydrophobic interactions between the SBD and SANT1L regions of Ezh2 are

shown, with involved residues highlighted as sticks. (**C**) Interactions between the MCSS region of Ezh2 and Suz12(VEFS) are shown, with the disease mutation residues of Suz12(VEFS) and their interacting residues in the MCSS highlighted as sticks. The zinc-coordinating motif of the MCSS is shown as mesh. (**D**) Suz12(VEFS) sits in the interface of the regulatory and catalytic moieties on the back of the SET domain. Conserved interacting residues of Suz12(VEFS) and Eed are highlighted as sticks.

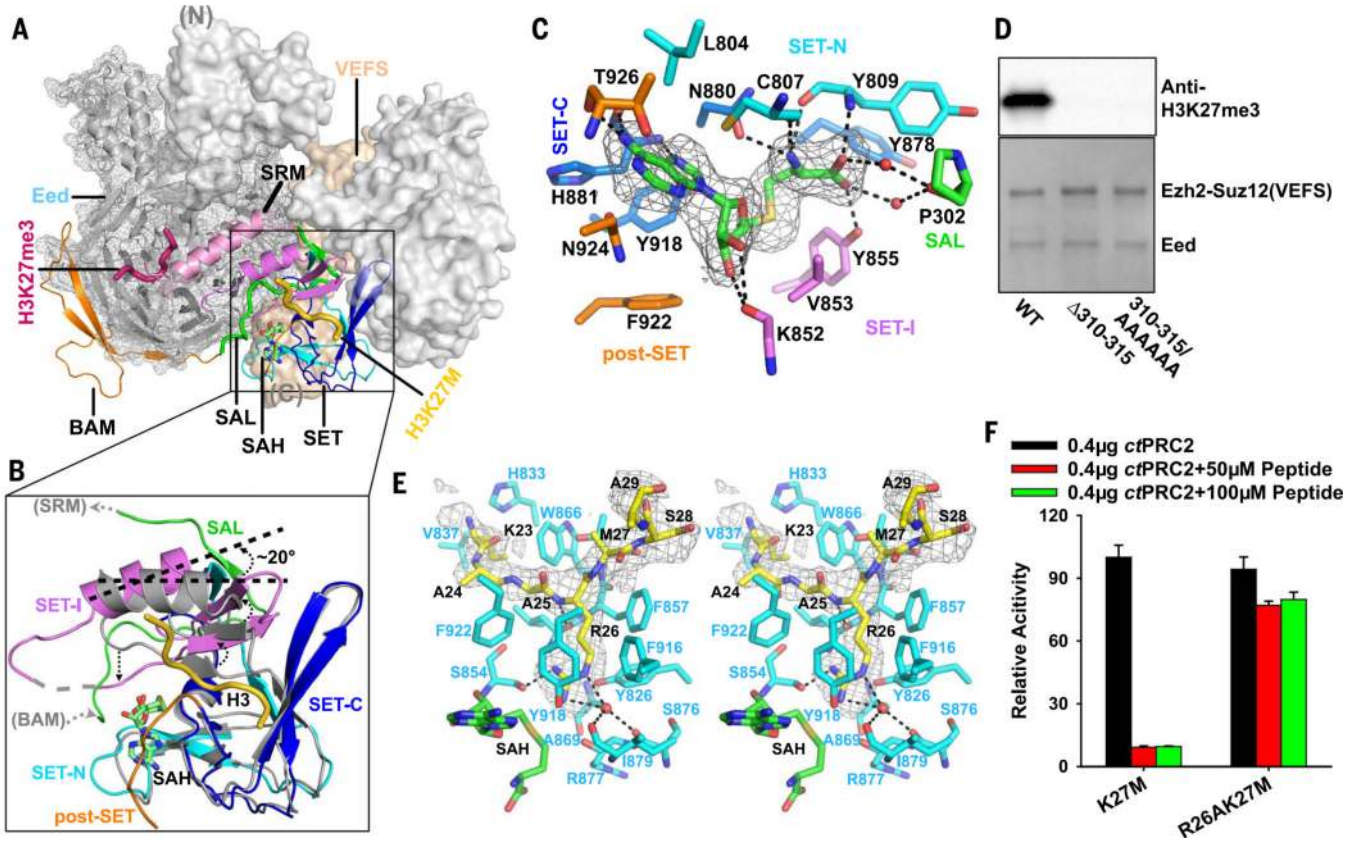

Fig. 3. Conformation of the active catalytic SET domain of Ezh2

(**A**) The trajectory of the SAL of Ezh2 is highlighted as a green ribbon in the context of the overall structure. The catalytic SET domain of Ezh2 is zoomed in and shown in Fig. 3B. (**B**) The catalytic SET domain from the current structure is superimposed with the isolated inactive SET domain of human Ezh2 (PDB:4MI0, shown in gray). SET-I swings about 20° counterclockwise relative to the inactive conformation. The SAL region of Ezh2 required to maintain the active conformation is also shown as a green ribbon. (**C**) Close-up view of Ezh2 residues involved in SAH interaction. The Fo − Fc omit electron density map for SAH contoured to 2.2σ is also shown as a gray mesh. (**D**) Deletion or mutation of residues 310 to 315 in the SAL of Ezh2 abolishes enzyme activity. Purified wild-type and mutant ternary complexes are shown below the result of the Western blot–based activity assay. (**E**) The lysine access channel of the active SET domain is shown as a stereo pair. The Fo − Fc omit electron density map contoured at 2.5σ that corresponds to the inhibiting H3K27M peptide is indicated by a gray mesh. Residue H3R26 occupying the lysine channel is unambiguously defined by the electron density map. (**F**) Inhibition of PRC2 enzyme activity by the H3K27M and H3R26AK27M mutant histone peptides. Relative enzyme activities from the ELISA-based assays are shown. Data are mean ± SD based on three trials.

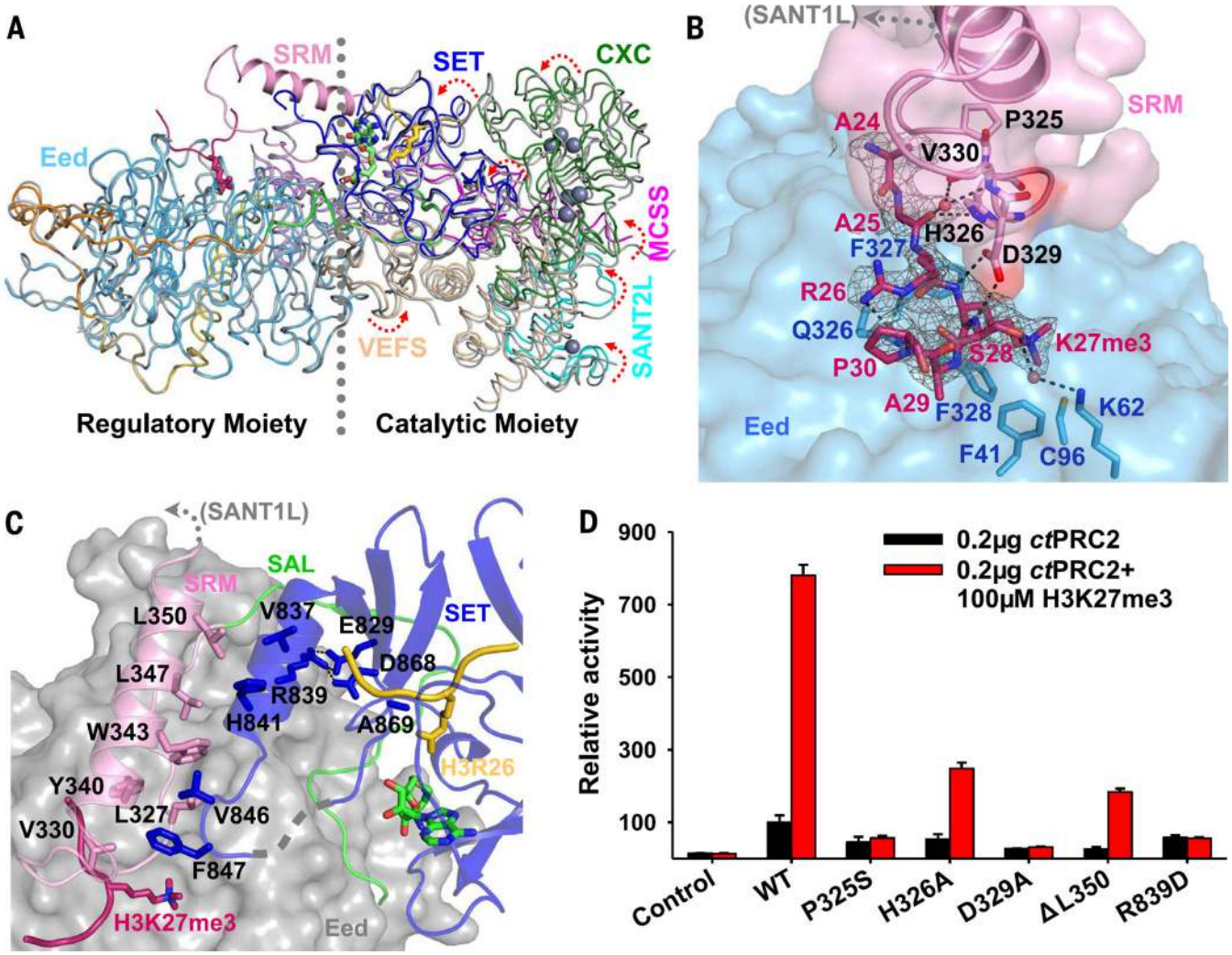

Fig. 4. Allosteric regulation of the PRC2 enzyme activity

(**A**) The crystal structures of the ternary complexes in the basal (gray ribbons) and stimulated (colored ribbons) states are aligned on the basis of Eed. The catalytic moiety is rotated counterclockwise toward the SRM from the basal to the stimulated state. The SRM of Ezh2, which experiences disorder-to-order conformational transition, is highlighted in cartoon representation. (**B**) Close-up view of interactions between the H3K27me3 peptide (red), the SRM (pink), and Eed (light blue). Interacting residues are shown as sticks. The Fo – Fc omit electron density map corresponding to the H3K27me3 peptide is contoured at 2.5σ and is shown as a gray mesh. (**C**) Close-up view of interactions between the SRM (pink) and the SET domain (deep blue) of Ezh2. (**D**) Enzyme stimulation assay of the wild-type and mutant ternary complexes. Relative enzyme activities from the triplicated ELISA-based assays are shown. Data are mean ± SD based on three trials.