Cryo-EM structures of PRC2 simultaneously engaged with two functionally distinct nucleosomes

Simon Poepsel1,2, Vignesh Kasinath1,2 and Eva Nogales1,2,3,4*

Epigenetic regulation is mediated by protein complexes that couple recognition of chromatin marks to activity or recruitment of chromatin-modifying enzymes. Polycomb repressive complex 2 (PRC2), a gene silencer that methylates lysine 27 of histone H3, is stimulated upon recognition of its own catalytic product and has been shown to be more active on dinucleosomes than H3 tails or single nucleosomes. These properties probably facilitate local H3K27me2/3 spreading, causing heterochromatin formation and gene repression. Here, cryo-EM reconstructions of human PRC2 bound to bifunctional dinucleosomes show how a single PRC2, via interactions with nucleosomal DNA, positions the H3 tails of the activating and substrate nucleosome to interact with the EED subunit and the SET domain of EZH2, respectively. We show how the geometry of the PRC2–DNA interactions allows PRC2 to accommodate varying lengths of the linker DNA between nucleosomes. Our structures illustrate how an epigenetic regulator engages with a complex chromatin substrate.

Cytoskeletal modification of the N-terminal tails of histone proteins forming the protein core of the nucleosomes that package DNA in eukaryotes is a fundamental mechanism of epigenetic gene regulation. Histone-modifying enzymes catalyze the deposition or removal of histone marks, which can in turn be bound by specific recognition modules within larger protein assemblies that serve gene regulatory functions1. The faithful orchestration of gene regulatory patterns, for example, during development, critically relies on the interplay of sensing and altering the chromatin state. Consequently, both of these activities are typically found in gene regulatory complexes. The dynamic nature of chromatin poses a challenge to studies aiming to elucidate these processes, both in cellular context and in reconstituted systems, and detailed structural studies of epigenetic regulators have so far been limited to histone peptide-bound complexes or single functional modules bound to nucleosomes2,3.

Trimethylation of lysine 27 on histone H3 (H3K27me3), catalyzed by PRC2, leads to gene silencing of developmental and cell fate-determining genes within multicellular organisms4. All four core PRC2 subunits, enhancer of zeste homolog 2 (EZH2), embryonic ectoderm development (EED), suppressor of zeste 12 (SUZ12) and RBAP46 or RBAP48, have been proposed to contribute to histone-tail binding4–8. Engagement of H3K27me3 by the recognition module EED characteristically leads to allosteric activation of the catalytic su(var)3–9, enhancer-of-zeste and trithorax (SET) domain within EZH2, a mechanism that has been characterized in detail only using peptide ligands. Crystal structures of the catalytic lobes of a fungal9 and human10 PRC2, comprised of EZH2, EED, and the C-terminal VEFS domain of SUZ12, bound to stimulatory methylated peptides have offered clues about the structural rearrangements within PRC2 that lead to activation, which is thought to contribute to local spreading of H3K27me3 and the establishment of heterochromatin domains. Regulatory mechanisms controlling PRC2 function also include inhibition by active chromatin marks such as H3K4me3 (ref. 11) or H3K36me3 (ref. 12) or the interaction with noncoding RNAs13–14 and auxiliary subunits15–17.

Biochemical studies have shown that the activity of PRC2 is significantly higher on dinucleosomes and higher-order chromatin structures than on histone tails or mononucleosomes, a property that is not mechanistically understood but may also contribute to the spreading of the PRC2 silencing mark15–17. Indeed, many questions have remained unanswered. How does PRC2 engage with its natural chromatin substrates, nucleosomal arrays? Does PRC2 interact with core histones via the acidic patch used by many nucleosome-binding proteins? Does the interaction instead involve nucleosomal DNA? Can more than one nucleosome be engaged by a single PRC2? If so, can such engagement occur in the context of neighboring nucleosomes, and how does nucleosome spacing and geometry affect PRC2 engagement?

Here, we provide direct visualization of PRC2–chromatin interactions through cryo-EM structures of PRC2 in the specific context of dinucleosomes containing one unmodified substrate nucleosome and one activating, H3K27me3-containing nucleosome. This combination is particularly relevant for our understanding of H3K27me3 spreading, as it functionally corresponds to a boundary condition in which the state of one nucleosome can directly affect the activity of the complex on the neighboring one. We find that PRC2 interacts with the histone H3 tails and with the nucleosomal DNA, but not with other histones or the histone core. Of special functional relevance, our reconstructions show how the specific geometry of the assembly allows the simultaneous engagement of both nucleosomes by the same complex, even in the context of various linker lengths. Binding of the substrate nucleosome by a rigid DNA-binding interface on the CXC domain of EZH2 positions the H3 tail optimally for modification by EZH2. On the other side of the complex, a more flexible binding surface involving the WD40 domain of EED allows for the recognition of an activating H3K27me3 in the context of geometrically diverse chromatin substrates. Our structures support...
an H3K27me3-based PRC2 activation and spreading mechanism, which had been proposed previously on the basis of biochemical data, linking activation of the SET domain with the right engagement of a new PRC2 substrate. Our work also provides a framework to probe the mechanism of how differences in nucleosome spacing might lead to changes in methyltransferase activity.

Results

Structure of a human PRC2 and its interaction with dinucleosomes. For our structural studies of PRC2 interactions with chromatin, we decided to explore a specific functional configuration, with a minimal, structurally tractable substrate: a dinucleosome that includes an unmodified and an H3K27me3-modified nucleosome. We generated recombinant heterodinucleosomes with 35 base pairs (bp) of linker DNA (DiNcl35) by ligating mononucleosomes harboring pseudotrimethylated H3K27 (Nclmod) and unmodified H3 (Nclsub) (Supplementary Fig. 1) and purified recombinant human PRC2 composed of the core subunits EZH2, EED, SUZ12, and RBAP48 and the cofactor AEBP2 (Fig. 1a and Supplementary Fig. 1c). AEBP2 has been shown to have an overall stabilizing effect on the complex through extensive interactions with other subunits and may play a role in chromatin binding based on its proposed DNA-binding properties. We hereafter refer to this five-subunit PRC2 assembly simply as PRC2. Binding of our dinucleosomes to PRC2 was tested using electrophoretic mobility shift assay (EMSA) (Supplementary Fig. 1c). Reference-free 2D classification of an initial negative-stain EM dataset showed typical views of PRC2 with two nucleosomes bound (Supplementary Fig. 1d). In order to be able to place the nucleosomes in the context of the full PRC2 structure, we also obtained a cryo-EM reconstruction of PRC2 (the same five-subunit complex) without nucleosomal substrate that reached an overall resolution of 4.6 Å (Table 1, Fig. 1b and Supplementary Fig. 2).

It was possible to unambiguously assign the densities of the top catalytic lobe, comprising EZH2, EED, and the VEFS domain of SUZ12, and a bottom lobe density for RBAP48 to published crystal structures (Fig. 1b). AEBP2 and the N-terminal part of SUZ12 probably correspond to the remaining unassigned densities localized to the bottom lobe, in agreement with our previous study using genetic labels and with our more recent high-resolution structures of a six-component complex (PRC2–AEBP2–JARID2)²⁰. The cryo-EM structure of human PRC2 was obtained following mild cross-linking of the complex, which is absolutely necessary to maintain the integrity of the complex during the harsh process of sample blotting and vitrification that is used to generate a frozen hydrated sample for cryo-EM visualization (Methods). Cross-linking, however, proved incompatible with nucleosome binding, as assessed by both EMSA and cryo-EM visualization (data not shown), most likely caused by the titration of key functional lysines needed for nucleosome engagement. Disruption of dinucleosome binding to PRC2 was observed even when cross-linker was added after formation of the PRC2–dinucleosome complex, thus suggesting that the off rate of the dinucleosome was fast enough for the cross-linker to compete for the lysines. Therefore, both the negative-stain analysis and the following cryo-EM study of PRC2–dinucleosome interactions had to be carried out in the absence of cross-linker.

Analysis of our negative-stain 2D class averages showed several distinct populations of PRC2–dinucleosome complexes (Fig. 1c and Supplementary Fig. 3a). In all of the class averages, we observed one of the nucleosomes positioned by the top catalytic lobe of PRC2, in the vicinity of the CXC and SET domains of EZH2. The position of the second nucleosome was variable, and we classified the different positions into three main groups. The first group (~30% of PRC2–DiNcl particles) corresponds to the best-defined structure of the PRC2–dinucleosome complex, with consistent orientations of the
nucleosomes and clearly distinguishable features. In this group, both nucleosomes are placed in a unique position with respect to PRC2 and bound to the catalytic lobe of PRC2 in a characteristic orthogonal orientation of nucleosomes to each other (Fig. 1c). In the second group (~55% of particles), the nucleosome distal from the active site of EZH2 is located proximal to the N-terminal part of SUZ12, by the bottom half of the structure. Notably, this group included different classes in which the location of the second nucleosome varied dramatically (>130 Å), with poorly defined density for that nucleosome. These features are consistent with different orientations and marked flexibility (Fig. 1c) and indicate that this nucleosome is not rigidly engaged with PRC2 or may even be completely unattached (Supplementary Fig. 3a and Discussion). The third group of classes was less populated and showed the binding of one nucleosome near the CXC and SET domains of EZH2 and RBAP48 and the second nucleosome in proximity to EED.

Cryo-EM structure of a stable and active PRC2–dinucleosome complex. In order to visualize and understand the basis for nucleosome–PRC2 interactions in 3D, we collected cryo-EM data on frozen hydrated samples of PRC2 bound to dinucleosomes prepared under the same conditions that showed complete binding of dinucleosomes as assessed by EMSA (Table 1). Reference-free 2D class averages of the non-cross-linked frozen hydrated sample did not show density for the bottom segment of PRC2 (Fig. 2a,c), which is not surprising because of the lack of stability of this lobe in the absence of cross-linking. The position of the nucleosomes observed by cryo-EM resembles that of the group 1 visualized by negative-stain EM (‘first group’ described above), corresponding to the best-defined PRC2–dinucleosome positioning. Accordingly, the cryo-EM 3D reconstruction showed a single arrangement of nucleosomes bound to the catalytic lobe of PRC2 (Fig. 2a). To place the nucleosome binding by the catalytic lobe in the structural context of the complete PRC2, we superimposed our two 3D cryo-EM reconstructions, the non-cross-linked PRC2–DiNcl30 and the cross-linked PRC2 alone (Fig. 2b). The bottom lobe of PRC2 does not clash with the nucleosomes, and the observed nucleosome binding again corresponds well with group 1 in the negative-stain study (Figs. 1c and 2b). This correspondence, together with the high degree of similarity between our structure of the top lobe of PRC2 engaged with the dinucleosome and the crystallographic structure of the catalytically active EZH2, EED, and the SUZ12–VEFS subcomplex, are indicative of the preservation of this biologically important catalytic lobe of the complex and its bona fide interaction with nucleosomes.

The cryo-EM 3D reconstruction of the PRC2–dinucleosome complex, at an overall resolution of 6.2 Å (Supplementary Figs. 4 and 5), shows that the PRC2 catalytic lobe corresponds closely to that previously described by X-ray crystallographic studies that is considered the minimal functional core of the complex.10 Its position between the two nucleosomes, which are connected by clear density corresponding to the linker DNA, reveals a remarkably stable arrangement within PRC2–DiNcl30 (Fig. 2a). On both sides, the catalytic lobe of PRC2 makes contact with the DNA near its exit point from the nucleosome, which coincides with the location of the histone H3 tail emerging from the histone core (Fig. 2a,d). We were able to place a DNA model of 35 bp with a bend angle of 52° into the density corresponding to the linker DNA, suggesting that PRC2 binding does not cause significant displacement of the nucleosome cores from the positioning sequences used for reconstitution. The excellent overall agreement of the cryo-EM density for PRC2 with the previously reported crystal structure of a partial human PRC2 complex10 allowed us to unambiguously assign PRC2 components and subdomains within our structure (Fig. 2a,d). EZH2 contains two SANT domains, which are structurally well-conserved domains found in a number of chromatin-associated factors. In our reconstruction, both domains are resolved at lower resolution than the rest of the complex because of their flexibility (Supplementary Fig. 5d). The N-terminal part of a characteristic helix within EZH2 has been termed the SANT-binding domain (SBD), and the stretch of that helix that traverses the WD40 domain of EED is referred to as the EED-binding domain (EBD) (Supplementary Fig. 6a). In the context of the dinucleosome-bound state, the SBD straightens, and the SANT1 domain moves upward relative to its position within the cryo-EM structure (Supplementary Fig. 6b). The SBD helix clearly makes contact with the DNA of Nclsub. Near the active site, the nucleosomal DNA of Nclsub is contacted by the EZH2 CXC domain, which is defined by two characteristic zinc-binding motifs. The back side of EED and the active site of the EZH2 SET domain show density in agreement with ligand occupation of these sites (Fig. 2d). There is well-resolved density corresponding to the SRM helix, a proposed hallmark of activation in the presence of trimethylated peptides. Correspondingly, our reconstruction of PRC2–AEBP2 without nucleosomes is missing an ordered SRM loop, as well as density for ligands bound to EED or the active site of EZH2 (Fig. 2d).

Substrate nucleosome recognition by the PRC2 SET and CXC domains. Local resolution estimation of our cryo-EM reconstruction shows that Nclsub is less resolved than Nclmod, indicating more flexibility of the former (Supplementary Fig. 5d). Indeed, further 3D classification yielded classes with slightly varying orientations of Nclsub, whereas Nclmod was found to be in a similar position relative to PRC2 in all classes (Fig. 2e and Supplementary Fig. 5e,f). In order to obtain higher-resolution information on the Nclsub interface with PRC2, i.e., the CXC and SET domains of EZH2, we carried out focused refinement after signal subtraction of the more flexible Nclsub from the particle images. This procedure yielded an improved map, with an overall resolution of 4.9 Å (Table 2, Fig. 3a and Supplementary Fig. 7a–c). Flexible fitting of the atomic model of the PRC2 catalytic lobe into the density only required a small displacement of the SET and CXC domains and the tilting of the EZH2 SBD helix, further indicating that these are the main structural changes accompanying engagement of PRC2 with nucleosomes.

Table 1 | Cryo-EM data collection and processing

|               | PRC2–AEBP2 (EMD-7713) | PRC2–DiNcl15 (EMD-7306) | PRC2–DiNcl30 (EMD-7307) | PRC2–DiNcl30 (EMD-7308) |
|---------------|------------------------|--------------------------|--------------------------|--------------------------|
| Magnification | 29,000                 | 37,879                   | 37,879                   | 80,000                   |
| Voltage (kV)  | 300                    | 300                      | 300                      | 120                      |
| Electron exposure (e–/Å²) | 60              | 40                       | 40                       | 25                       |
| Defocus range (μm) | 0.5                | -1.5 to -1.5             | -2.5 to -2.5             | -3.5                     |
| Pixel size (Å) | 1.84                 | 1.32                     | 1.32                     | 1.37                     |
| Symmetry imposed | C1                | C1                       | C1                       | C1                       |
| Initial particle images (no.) | 882,317           | 438,601                   | 190,479                   | 46,581                   |
| Final particle images (no.) | 209,322            | 93,384                    | 27,182                    | 16,333                   |
| Map resolution (Å) | 4.6               | 6.2                      | 8.4                      | 13.3                     |
| FSC threshold | 0.143                 | 0.143                     | 0.143                     | 0.143                     |
| Map resolution range (Å) | 4.0–8.0          | 6–10                     | 7–13                     | -                        |
| Map-sharpening factor (Å²) | -78                 | -260                     | -524                     | -1,331                   |
The CXC zinc-coordinating loops and the interaction of the second zinc cluster with the DNA are well resolved, allowing the CXC–DNA contacts to be narrowed down to the region comprising EZH2 amino acids (aa) 561–570, which forms an arch-like density that, at its base, contacts both phosphodiester backbones at the minor groove of the DNA exiting the nucleosome (Fig. 3b). A large, well-conserved electropositive patch on the surface of the CXC and SET domains is ideally positioned to

**Fig. 2 | Cryo-EM structure of the PRC2–dinucleosome complex.**

- **a**, Cryo-EM reconstruction of the catalytic lobe of PRC2 bound to DiNcl_{55} with fitted crystal structures (nucleosome, PDB 3LZ1 (ref. 28); PRC2, PDB 5HYN (ref. 10)) shown in two different views. Ncl_{mod}, H3K27me3-modified nucleosome; Ncl_{sub}, substrate, unmodified nucleosome.
- **b**, Montage of a full PRC2–dinucleosome structure based on the superposition of the PRC2 and the PRC2-DiNcl_{55} cryo-EM maps to show the consistency of the observed nucleosome binding configuration with the full PRC2 structure. The view shown corresponds to one in between those displayed in a, c, PRC2 can stably bind the bifunctional dinucleosome substrate used in our study without involvement of the bottom lobe in nucleosome interaction, as indicated by negative-stain (left) and cryo-EM (right) analyses. The frozen hydrated sample of PRC2-DiNcl_{55}, missing the bottom lobe in 2D class averages (right), engages in dinucleosome interactions indistinguishable from those visualized by negative stain when the full complex is stable and visible. Top row, montage of a full PRC2-dinucleosome structure (left) and cryo-EM structure of PRC2-DiNcl_{55} (right), corresponding to the views shown below. Middle, reprojections of the densities in the top row. Bottom, matching experimental 2D class averages showing good agreement with the 2D reprojections of the densities.
- **d**, Back view of the PRC2 cryo-EM map and model, either bound to DiNcl_{55} (top) or in the absence of substrate nucleosomes (bottom). PRC2-DiNcl_{55} shows densities in the binding sites for substrate histone H3 tail (H3_{sub}, orange) and the H3K27me3-modified H3 tail (H3_{mod}, orange), as well as density for the EZH2 SRM helix. Both H3 tail densities and the ordered SRM are absent in the unbound PRC2 state.
- **e**, Comparison of 3D subclassified PRC2-DiNcl_{55} complexes to visualize structural variability. Two classes are superimposed as examples. Bottom panels show enlarged side views of Ncl_{sub} (left) and Ncl_{mod} (right).
interact with the DNA phosphodiester backbone (Fig. 3c). On the basis of the arrangement of amino acid side chains in the crystal structure\textsuperscript{10}, residues K563, Q565, K569 and Q570 are most likely to contribute to these interactions (Fig. 3d).

Analysis of the residual EM density not accounted for by the fitted crystallographic models of PRC2 and Ncl\textsubscript{sub} shows a continuous density connecting the H3 tail with the location where substrate peptide was observed in the EZH2 active site in the PRC2 X-ray crystal structure (Fig. 3e and Supplementary Fig. 7f). This density suggests an extended but flexible path of the tail from the nucleosome into the active site and directly supports our assignment of the substrate nucleosome as the one contacting the CXC domain.

At the PRC2–Ncl\textsubscript{sub} interface, bridging the H3 tail, the CXC domain and the nucleosomal DNA, we observed an additional unassigned density (Fig. 3e,f and Supplementary Fig. 7f). A candidate region that could possibly account for this density is a flexible segment of EZH2, corresponding to aa 480–515 connecting the CXC with the SANT2 domain, which has not been resolved by crystallography, but might contribute to the unassigned density (difference density in purple). Three crosslinks between this stretch and AEBP2 have been reported previously (K505, K509 and K510)\textsuperscript{19}. Crystal structures shown have been modified (PDB 3LZ1 (ref. 28); PRC2, PDB 5HYN (ref. 10), see methods and Supplementary Fig. 6).
How PRC2 engages with Nclmod, we carried out alignment-free 3D activation of the H3K27me3 epigenetic mark. In order to better resolve all observed conformational states, the EZH2 SBD makes a clear one (class 1) representing a predominant orientation of Nclmod and selected for closer analysis of their interaction with the nucleosome: carrying the modification.

Whereas Nclsub presents the substrate H3 tail to the EZH2 SET domain, Nclmod on the opposite side of EZH2, provides the PRC2-activating H3K27me3 epigenetic mark. In order to better resolve how PRC2 engages with Nclmod, we carried out alignment-free 3D classification after signal subtraction of Nclsub (the green mask marks the part of the complex retained during signal subtraction). Maps were aligned according to their PRC2 density, superimposed, and nucleosome models (PDB 3LZ1, ref. 28) were rigid-body fitted into the respective densities. Black arrows indicate the changed position of Nclmod in class 3 relative to class 1.

The EZH2 SANT1 domain is of yet unknown function and thus serving as a hinge, whereas, depending on nucleosome orientations, the interaction surface on EZH2 is contacted differentially.

The EZH2 SANT1 domain is of yet unknown function and appears to be one of the most flexible regions of PRC2. Comparison of the six classes hints at the existence of various possible orientations for SANT1, potentially involving contacts with the DNA and EED (Supplementary Fig. 8c). A flexible loop (aa 155–167) at the C terminus of the SRM, connecting it with SANT1, becomes extended (ref. 19 and Fig. 3f). Furthermore, because cross-links between AEBP2 and lysines 505, 509 and 510 of EZH2 have also been found19, it is possible that both EZH2 and AEBP2 contribute interactions in this region.

Binding of the H3K27me3 nucleosome by EED and EZH2. Whereas Nclsub presents the substrate H3 tail to the EZH2 SET domain, Nclmod on the opposite side of EZH2, provides the PRC2-activating H3K27me3 epigenetic mark. In order to better resolve how PRC2 engages with Nclmod, we carried out alignment-free 3D classification after signal subtraction of Nclsub. We obtained six classes showing slightly varying orientations of Nclmod relative to PRC2 (Supplementary Fig. 8a). For one such class, it is possible to directly see density that connects the EED-engaged K27me3 with the core of the H3 in the crystal structure docked within the Nclmod density (Table 2 and Supplementary Fig. 8b, asterisks (*)). EED residues 1–76 have not been resolved in the crystal structures9,10, but the lateral surface of the WD40 domain and the N-terminal stretch of EED potential contacts for the engagement of nucleosomes in a range of positions (Fig. 4b,c and Supplementary Fig. 8b, asterisks (*)). EED residues 1–76 have not been resolved in the crystal structures42, but density extending from the last modeled residue is clearly visible in our reconstructions (Fig. 4b, red asterisk). Owing to its high lysine content (aa 70–79, KGKWKSKKCK), this stretch is likely to bind the DNA phosphodiester backbone. Taken together, the SBD provides the most consistent contact point between PRC2 and Nclmod thus serving as a hinge, whereas, depending on nucleosome orientations, the interaction surface on EZH2 is contacted differentially.

For clarity, only two classes out of the six mentioned above were selected for closer analysis of their interaction with the nucleosome: one (class 1) representing a predominant orientation of Nclmod and the other (class 3) showing the largest observed deviation of the Nclmod density (Fig. 4a). Nclmod of class 3 is tilted downward and slightly towards the back of PRC2 compared to that of class 1. There are two main regions of contact between PRC2 and Nclmod. First, in all observed conformational states, the EZH2 SBD makes a clear contact with the DNA minor groove of the upper DNA gyre of Nclmod (Fig. 4b). The SBD bears two patches of positively charged residues (16–20, RKRVK, and 27–30, RQLKR) that are most likely to mediate DNA contacts (Fig. 4c). The second set of interactions is mediated by the lateral surface of the WD40 domain and the N-terminal stretch of EED. The connectivity of these regions with the nucleosome varies with the relative orientations of Nclmod observed in the different classes, and involves contacts with both DNA gyres (Fig. 4b, red asterisk). Again, patches of positive surface potential line the side of EED and parallel the path of these DNA helices, providing a number of potential contacts for the engagement of nucleosomes in a range of positions (Fig. 4b,c and Supplementary Fig. 8b, asterisks (*)).
in the complex upon the upward tilting of SANT1, while the SRM stays in place (Supplementary Fig. 8c, green dot).

**Tolerance and sensing of varying dinucleosome geometries.** The PRC2–DiNcl 35 described by our cryo-EM study shows how the arrangement of the nucleosomes relative to each other and to PRC2 allows the simultaneous functional engagement of H3 tails at both the allosteric EED-binding site and the active site in EZH2. We therefore asked to what extent the geometrical constraints of the linker DNA dictate this arrangement, as well as whether and how PRC2 can accommodate changes in this geometry. Furthermore, variations in linker length between nucleosomes have been linked to different levels of PRC2 activity 8. We therefore reconstituted dinucleosomes with 30 bp (DiNcl 30) or 40 bp (DiNcl 40) of linker DNA, thus removing or adding half a helical turn with respect to our previous arrangement (Table 1 and Supplementary Figs. 9 and 10). Interestingly, the overall architecture of PRC2–DiNcl 30 observed in our cryo-EM studies resembles the PRC2–DiNcl 35 complex in many respects, with the same regions of PRC2 being engaged in interactions with the nucleosomes (Fig. 5a). However, Nclmod in PRC2–DiNcl 30 is flipped by ~180°, so that the tail of the other copy of H3 within the same histone octamer is bound to EED (Fig. 5b). Consequently, the linker DNA follows a different, straighter path. In agreement with what we observed for the PRC2–DiNcl 35 complex, the orientation of Nclsub is less variable than that of Nclmod (Fig. 5c).

Increasing the linker length to 40 bp gave rise to an Nclmod arrangement that resembles the PRC2–DiNcl 30 complex, but with the DNA exit points being further away from each other, owing to the increased linker length (Fig. 5d and Supplementary Figs. 10 and 11). As an effect of the twisting back of Nclmod of DiNcl 30 relative to PRC2, its interface with PRC2 changes considerably compared to what is seen for the DiNcl 35. The major contact involving the N-terminal region of EED occurs with the upper nucleosomal DNA gyre, rather than the lower gyre seen for PRC2–DiNcl 35 (Supplementary Fig. 9e). The SBD–SANT1 region of EZH2 makes a single contact with the DNA rather than traversing the minor groove, and the SANT1–EED contact seen in a subpopulation of PRC2–DiNcl 35 becomes more prominent (Supplementary Fig. 9e, red dot). Based on our structures, we can conclude that slightly different binding sites on the EED and the flipping of the nucleosome

---

**Table 2 | Cryo-EM data collection and processing**

|                | PRC2–Nclsub-DiNcl 35 (EMD-7311) | PRC2–Nclsub-DiNcl 30 (EMD-7312) | PRC2–Nclsubcl1-DiNcl 35 (EMD-7309) | PRC2–Nclsubcl3-DiNcl 35 (EMD-7310) |
|----------------|---------------------------------|---------------------------------|-------------------------------------|-------------------------------------|
| Magnification  | 37,879                          | 37,879                          | 37,879                              | 37,879                              |
| Voltage (kV)   | 300                             | 300                             | 300                                 | 300                                 |
| Electron exposure (e−/Å²) | 40                    | 40                             | 40                                  | 40                                  |
| Defocus range (μm) | -1.5 to -3.5                      | -1.5 to -3.5                      | -1.5 to -3.5                         | -1.5 to -3.5                        |
| Pixel size (Å) | 1.32                            | 1.32                            | 1.32                                | 1.32                                |
| Symmetry imposed | C1                            | C1                             | C1                                  | C1                                  |
| Initial particle images (no.) | 438,601                          | 190,479                          | 438,601                             | 438,601                             |
| Final particle images (no.) | 93,384                          | 27,182                          | 25,512                              | 15,104                              |
| Map resolution (Å) | 4.9                              | 7.7                            | 7.4                                 | 10.1                                |
| FSC threshold  | 0.143                           | 0.143                           | 0.143                               | 0.143                               |
| Map resolution range (Å) | 4.5–8.5                          | 8–12                           | –                                   | –                                   |
| Map-sharpening B factor (Å²) | -200                           | -495                          | -351                                | -481                                |

© 2018 Nature America Inc., part of Springer Nature. All rights reserved.
to bind the alternative H3, together with slightly different torsions of the linker DNA, should be able to accommodate a range of linkers, thus allowing PRC2 binding to different chromatin geometries. These adjustments might have some effect on affinity (for example, higher for minimal DNA torsion) and activity (see below).

To improve the map in the vicinity of the PRC2 Ncl₁ interface, we also carried out masked refinement after signal subtraction, just as we did for the PRC2–DiNcl₁ complex, leading to an overall resolution of 7.7 Å (Table 2 and Supplementary Fig. 1a–c). The atomic model of PRC2 flexibly fitted into the PRC2–DiNcl₁ map (Methods and Supplementary Fig. 5) deviates from the PRC2–DiNcl₁ map in the region of the EZH2 CXC domain, with a good overall fit in the other parts of the complex (Supplementary Fig. 1d). Our analysis indicates a shift of this domain, together with the interacting DNA, toward the SET domain and EED, whereas the first zinc cluster of the CXC seems more flexible, according to local resolution estimation (Supplementary Fig. 12c,e). It is conceivable that these rearrangements of the CXC domain could contribute to changes in EZH2 activity that have been reported for chromatin substrates with different nucleosome densities.

**Discussion**

**Relevance of the PRC2–dinucleosome structures for epigenetic spreading.** Allosteric activation of EZH2 by its own catalytic product, H3K27me3, has long been proposed to form the basis of local H3K27me3 spreading and epigenetic memory of repression, for example, during rounds of replication. The structures of PRC2–dinucleosome complexes presented here show how one PRC2 can bind an H3K27me3-bearing nucleosome that acts as an allosteric activator to prompt trimethylation of H3K27 on a neighboring nucleosome. The stable, dual engagement of PRC2 defines a functional configuration in its interaction with chromatin in which the modified nucleosome not only activates the SET domain, but positions the substrate nucleosome and H3 tail for interaction with the active site. This PRC2–DiNcl₁ arrangement suggests that local H3K27me3 spreading can be facilitated by a single PRC2 complex that simultaneously senses and modifies its chromatin environment. The interactions revealed by our structural analyses are mediated by PRC2–DNA contacts rather than interfaces with the globular histone core. Recent biochemical studies also reported PRC2–DNA contacts to contribute most substantially to the affinity of the complex toward chromatin. Conserved regions of positive charge on the PRC2 surface follow the path of the DNA strands that we have visualized in our study, thus supporting the notion of a strong contribution of the electrostatic interactions with the phosphate backbone to the nanomolar binding affinity seen for nucleosomes, compared to the micromolar affinity seen for peptides alone.

**A rigid interface near the SET domain holds the substrate nucleosome in place.** Our cryo-EM structures show PRC2 interacting with the substrate nucleosome near the catalytic SET domain of EZH2 through contacts between positively charged surface residues of the EZH2 CXC domain and the phosphodiester backbone of the nucleosomal DNA. In addition to this well-resolved region of the interface, a segment of EZH2 and/or AEBP2 that has yet to be identified appears to contribute additional contacts with the substrate nucleosome. These interactions might add to PRC2-binding affinity toward nucleosomes and could also have repercussions on PRC2 activity regulation. The location of this contact site within the PRC2–DiNcl₁ complex indicates a potential interaction interface with the functionally important lysine 36 of histone H3. Methylated H3K36 has been reported to inhibit PRC2 activity and is thought to mark actively transcribed genes. It has been shown that H3K36me3 inhibits PRC2 in cis when it is present on the same H3 tail harboring the substrate H3K27 residue, which underscores the potential importance of this region. Future mechanistic studies should aim to answer whether and how interactions of the unmodified H3K36 residue with EZH2 at this site might be required to stabilize the active conformation of the EZH2 SET domain.

**Versatility of the H3K27me3-modified nucleosome-binding site.** A versatile nucleosome-binding site, comprised of a hinge formed by the EZH2 SBD and a binding surface on EED, allows for the engagement of H3K27me3-modified nucleosomes in orientations that may vary depending on the conformational context of the chromatin substrate, as shown by various orientations of Ncl₃ relative to PRC2 upon 3D subclassification of our cryo-EM maps (Figs. 2c and 4a and Supplementary Figs. 5e,f and 8). A number of positively charged surface residues of EED and the SBD of EZH2 engage in variable contacts with nucleosomal DNA, depending on the local conformational environment of chromatin. The SANT1 domain of EZH2, which shows a high degree of flexibility and is thus not well resolved, seems to change its positions and undergo rearrangements depending on the nucleosome orientation. Its direct connection with the SRM of EZH2, which has been reported to play a central role in transmitting the activating signal from EED to the SET domain of EZH2 (refs. 1,20), suggests that structural changes of SANT1 in response to engagement of nucleosomes in varying orientations might affect EZH2 catalytic activity through this connection to the SRM. Thus, this region of EZH2 should be an interesting candidate for future studies aiming to elucidate the regulatory mechanisms of the response of PRC2 to conformational variations in its chromatin environment.

**Alternative PRC2–DiNcl configurations.** The dual engagement of nucleosomes by EED and CXC and SET of EZH2 described here is incompatible with the simultaneous engagement of an H3 tail from these nucleosomes by the WD40 domain of RBA48 (as seen in group 1 of Fig. 1c). The distance between the H3 tails exiting the nucleosomes or the EZH2 active site and the binding site on RBA48 would not be bridged by a fully extended peptide (Supplementary Fig. 13). On the other hand, the other, more flexible configurations visualized by our negative-stain EM analysis (group 2, Fig. 1c and Supplementary Fig. 3a) could be consistent with binding to the bottom half of the complex, which includes RBA48. However, we suggest that these looser PRC2–dinucleosome arrangements will involve an alternative state in which there is no nucleosome bound to EED, and thus PRC2 is not activated by the H3K27me3 signal. Furthermore, the position of the nucleosome proximal to the SET domain in this arrangement differs from that in group 1 and that described in detail in our cryo-EM structure, so that the nucleosome binding near the active site is unlikely to be optimally positioned for substrate H3 tail engagement. We propose that the configuration described by our cryo-EM analysis is one in which the H3K27me3-activated PRC2 is engaging a new substrate for methyltransferase activity that will facilitate the spread of this silencing mark. Clearly, other nucleosome binding-sites and alternative nucleosome arrangements may be possible, especially for other combinations of histone modifications, such as active transcription marks.

**PRC2 accommodates different DNA linker lengths between nucleosomes.** Our study shows that PRC2 is able to tolerate the geometrical changes imposed by different lengths of linker DNA by alternatively engaging one or the other H3 copy of the nucleosome bearing the H3K27me3 mark. Furthermore, our structures show that when the length of the linker DNA is shortened from 35 to 30 bp, movement of both nucleosomes toward each other causes a conformational change of the nucleosome-binding CXC domain. The apparent increased flexibility of the first zinc cluster of the CXC domain together with the movement of the DNA binding zinc cluster could conceivably affect the neighboring SET domain as part...
of a response of PRC2 to varying linker lengths (Supplementary Fig. 12). How these changes of the CXC domain may contribute to the modulation of methyltransferase activity in response to changes in nucleosome spacing should be the subject of future biochemical and structural studies. Our structures with different linker lengths between nucleosomes provide a possible mechanistic explanation of how PRC2 engagement of activating and substrate nucleosomes can be maintained in a dynamic and diverse chromatin environment. In light of recent discoveries pointing out the heterogeneity of chromatin structure in vivo, the ability of PRC2 to tolerate diverse chromatin substrates via the nucleosomal DNA engagement we describe here, is probably an important functional aspect of PRC2 structure.

**Methods**

Methods, including statements of data availability and any associated accession codes and references, are available at https://doi.org/10.1038/s41594-018-0023-y.

Received: 13 December 2017; Accepted: 21 December 2017; Published online: 29 January 2018

**References**

1. Lalonde, M. E., Cheng, X. & Côté, J. Histone target selection within chromatin: an exemplary case of teamwork. *Genes. Dev.* 28, 1029–1041 (2014).

2. Liu, Y. Y., Li, M. J., Xia, X., Li, X. M. & Chen, Z. C. Mechanism of chromatin remodelling revealed by the SuR2-nucleosome structure. *Nature* 544, 440–445 (2017).

3. Wilson, M. D. et al. The structural basis of modified nucleosome recognition by 53BP1. *Nature* 536, 100–103 (2016).

4. Cao, R. et al. Role of histone H3 lysine 27 methylation in Polycromb-group silencing. *Science* 299, 1039–1043 (2002).

5. Martin, L., Latypova, X. & Terro, F. Post-translational modifications of tau protein: implications for Alzheimer's disease. *Neurochem. Int.* 58, 458–471 (2011).

6. Son, J., Shen, S. S., Margueron, R. & Reinberg, D. Nucleosome-binding activities within JARID2 and EZH1 regulate the function of PRC2 on chromatin. *Genes. Dev.* 27, 2663–2677 (2013).

7. Margueron, R. et al. Role of the polycomb protein EED in the propagation of repressive histone marks. *Nature* 461, 762–767 (2009).

8. Yuan, W. et al. Dense chromatin activates Polycomb repressive complex 2 to regulate H3 lysine 27 methylation. *Science* 337, 971–975 (2012).

9. Jiao, L. & Liu, X. Structural basis of histone H3K27 trimethylation by an active polycomb repressive complex 2. *Science* 350, aac4383 (2015).

10. Justin, N. et al. Structural basis of oncogenic histone H3K27M inhibition of human polycomb repressive complex 2. *Nat. Commun.* 7, 11316 (2016).

11. Schmitges, F. W. et al. Histone methylation by PRC2 is inhibited by active chromatin marks. *Mol. Cell.* 42, 330–341 (2011).

12. Yuan, W. et al. H3K36 methylation antagonizes PRC2-mediated H3K27 methylation. *J. Biol. Chem.* 286, 7983–7989 (2011).

13. Rinn, J. L. et al. Functional demarcation of active and silent chromatin domains in human HOX loci by noncoding RNAs. *Cell* 129, 1311–1323 (2007).

14. Zhao, J., Sun, B. K., Erwin, J. A., Song, J. J. & Lee, J. T. Polycomb proteins targeted by a short repeat RNA to the mouse X chromosome. *Science* 322, 750–756 (2008).

15. Kim, H., Kang, K. & Kim, J. AEBP2 as a potential targeting protein for Polycomb Repression Complex PRC2. *Nucleic Acids Res.* 37, 2940–2950 (2009).

16. Li, G. et al. Jarid2 and PRC2, partners in regulating gene expression. *Genes. Dev.* 24, 368–380 (2010).

17. Ballare, C. et al. Phf19 links methylated Lys36 of histone H3 to regulation of Polycomb activity. *Nat. Struct. Mol. Biol.* 19, 1257–1265 (2012).

18. Martin, C., Cao, R. & Zhang, Y. Substrate preferences of the EZH2 histone methyltransferase complex. *J. Biol. Chem.* 281, 8365–8370 (2006).

19. Ciferri, C. et al. Molecular architecture of human polycomb repressive complex 2. *eLife* 1, e00005 (2012).

20. Kasinath, et al. Structures of human PRC2 with its cofactors AEBP2 and JARID2. *http://dx.doi.org/10.1126/science.aar5700* (2018).

21. Ketel, C. S. et al. Subunit contributions to histone methyltransferase activities of fly and worm polycomb group complexes. *Mol. Cell. Biol.* 25, 6857–6868 (2005).

22. Boyer, L. A., Latek, R. R. & Peterson, C. L. The SANT domain: a unique histone-tail-binding module? *Nat. Rev. Mol. Cell. Biol.* 5, 158–163 (2004).

23. Bai, X. C., Rajendra, E., Yang, G., Shi, Y. & Scheres, S. H. W. Sampling the conformational space of the catalytic subunit of human γ-secretase. *eLife* 4, e11182 (2015).

24. Hansen, K. H. et al. A model for transmission of the H3K27me3 epigenetic mark. *Nat. Cell. Biol.* 10, 1291–1300 (2008).

25. Choi, J. et al. DNA binding by PHF11 prolongs PRC2 residence time on chromatin and thereby promotes H3K27 methylation. *Nat. Struct. Mol. Biol.* 24, 1039–1047 (2017).

26. Wang, X. et al. Molecular analysis of PRC2 recruitment to DNA in chromatin and its inhibition by RNA. *Nat. Struct. Mol. Biol.* 24, 1028–1038 (2017).

27. Wagner, E. J. & Carpenter, P. B. Understanding the language of Lys36 methylation at histone H3. *Nat. Rev. Mol. Cell. Biol.* 13, 115–126 (2012).

28. Voigt, P. et al. Asymmetrically modified nucleosomes. *Cell* 151, 181–193 (2012).

29. Ou, H. D. et al. ChromEMT: visualizing 3D chromatin structure and compaction in interphase and mitotic cells. *Science* 357, eaag0025 (2017).

**Acknowledgements**

We thank P. Grob and J. Fang for technical support, T. Houweling and A. Chintangal for computer support, and D. King for MS confirmation of histone alkylation. We are thankful to R.M. Glaeser for his support and helpful discussions and to S. M. Sterling, A. Patel, B. LalFrance, D. Lipscomb, R. K. Louder, T. H. D. Nguyen and B. J. Greber for helpful discussions and advice during data acquisition and processing. We also thank T. H. D. Nguyen and B. J. Greber for critical comments on the manuscript. Computational resources were provided in part by the National Energy Research Scientific Computing Center (DE-AC02-05CH11231) and LWRENCIUM computing cluster at Lawrence Berkeley National Laboratory. This work was in part funded by Eli Lilly through the Lilly Research Award Program. S.P. was supported by the Alexander von Humboldt Foundation (Germany) Feodor-Lynen postdoctoral fellowship. V.K. was supported by a postdoctoral fellowship from the Helen Hay Whitney Foundation. E.N. is supported by the Howard Hughes Medical Institute.

**Author contributions**

E.N. supervised the study; S.P. designed and performed the experiments, data collection, processing, and interpretation; V.K. collected and processed the PRC2–AEBP2 data and contributed to the experimental design and data interpretation. E.N. and S.P. wrote the manuscript, with feedback from V.K.

**Competing interests**

The authors declare no competing financial interests.

**Additional information**

Supplementary information is available for this paper at https://doi.org/10.1038/s41594-018-0023-y.

Reprints and permissions information is available at www.nature.com/reprints.

Correspondence and requests for materials should be addressed to E.N.

Publisher’s note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.
Methods

Protein expression and purification. For expression of PRC2, the full-length sequences of human EZH2, EED, SUZ12, RBAp48 and AEBP2 were cloned into pFastbac baculoviral expression vectors. A TEV-cleavable GFP tag was engineered at the N-terminus of AEBP2 for affinity purification. PRC2 was expressed in High Five insect cells for 66 h, and cell pellets from 300-ml batches were stored at −80 °C until purification. Lysis was done by sonication in 25 mM HEPES, pH 7.9, 250 mM NaCl, 5% glycerol, 0.1% NP-40, 1 mM TCEP with added benzonex (Sigma–Aldrich) and protease inhibitor cocktail (Roche). After batch binding to Strep-Tactin Superflow Plus resin (Qiagen), the complex was purified by washing with low-salt (25 mM HEPES, pH 7.9, 150 mM NaCl, 1 mM TCEP, 5% glycerol, 0.01% NP-40) and high-salt buffers (25 mM HEPES, pH 7.9, 1 M NaCl, 1 mM TCEP, 5% glycerol, 0.01% NP-40), then eluted with 20 mM desthiobiotin. Fractions containing PRC2 were incubated overnight at 4 °C with TEV protease to cleave the tag before Superose 6 (GE Healthcare) size-exclusion chromatography. The final samples were stored in 10% glycerol at −80 °C as single-use aliquots, which were thawed just before use.

Recombinant histones H2A, H2B, H3, and H4 were expressed from pET3 plasmids in Escherichia coli BL21 RIL, purified from inclusion bodies, and reconstituted into histone octamers, essentially as described previously30. The expression plasmids of Xenopus laevis histone proteins used for the cryo-electron microscopy of nucleosomes18 were a gift from K. Luger, University of Colorado, Boulder). DNA for the reconstitution of nucleosomes was obtained by PCR, using primers to create the desired linker DNA sequence and include DraI restriction sites in addition to the ‘601 strong nucleosome-positioning sequence’ to allow for the ligation into digested nucleosome reconstitution linkers. The linker sequence contains randomly chosen nucleotide sequences, while avoiding nucleosome-positioning di- and trinucleotide sequences18. The Nl linker was sequenced. The Nl linker was sequenced and correspond to 601 agatcgtattgccctgccgta(601) site underlined in italics), whereas the Nc linker was varied depending on the desired linker length (for DNc, ctgctgtattgcctgccgta(601); for DNc, ctgctgtattgcctgccgta(601); for DNc, ctgctgtattgcctgccgta(601); for DNc, ctgctgtattgcctgccgta(601); for DNc, ctgctgtattgcctgccgta(601); for DNc, ctgctgtattgcctgccgta(601); for DNc, ctgctgtattgcctgccgta(601); for DNc, ctgctgtattgcctgccgta(601); for DNc, ctgctgtattgcctgccgta(601); for DNc, ctgctgtattgcctgccgta(601); for DNc, ctgctgtattgcctgccgta(601); for DNc, ctgctgtattgcctgccgta(601); for DNc, ctgctgtattgcctgccgta(601); for DNc, ctgctgtattgcctgccgta(601); for DNc, ctgctgtattgcctgccgta(601); for DNc, ctgctgtattgcctgccgta(601); for DNc, ctgctgtattgcctgccgta(601); for DNc, ctgctgtattgcctgccgta(601); for DNc, ctgctgtattgcctgccgta(601); for DNc, ctgctgtattgcctgccgta(601); for DNc, ctgctgtattgcctgccgta(601); for DNc, ctgctgtattgcctgccgta(601); for DNc, ctgctgtattgcctgccgta(601); for DNc, ctgctgtattgcctgccgta(601); for DNc, ctgctgtattgcctgccgta(601); for DNc, ctgctgtattgcctgccgta(601); for DNc, ctgctgtattgcctgccgta(601); for DNc, ctgctgtattgcctgccgta(601); for DNc, ctgctgtattgcctgccgta(601); for DNc, ctgctgtattgcctgccgta(601); for DNc, ctgctgtattgcctgccgta(601); for DNc, ctgctgtattgcctgccgta(601); for DNc, ctgctgtattgcctgccgta(601); for DNc, ctgctgtattgcctgc

Recombinant PRC2, full-length expression plasmids of PRC2 without nucleosomes was cross-linked at 1.5–2.0 mg/mL with 1.5 M TCEP and stored on ice until further use. Native PAGE of PRC2 was performed as described previously30. The PRC2 complexes were subsequently used for further analysis. To localize potential interaction interfaces of PRC2 with nucleosomes, the projection-matching approach was employed. Representative classes of the groups of classes shown in Fig. 1c were matched to 2D projections of the PRC2 EM maps shown in Fig. 1b, low-pass filtered to 15 Å, using Fabs2 software package.43

EM sample preparation for PRC2–dinucleosome complexes. To initially screen for formation of the PRC2–DiNcl complex, negative-stain EM experiments were performed. PRC2 and dinucleosomes were incubated on ice for 30–45 min in 25 mM HEPES, pH 7.9, 150 mM KCl, 1 mM TCEP at a concentration of 300 nM PRC2 (PRC2/DiNcl). 100 nM PRC2 without nucleosomes served as a negative control. 4 µl of the sample was added to a continuous carbon grid after glow discharge (Solarus, Gatan), immediately blotted with filter paper and stained with three successive short incubation steps in drops of 2% (wt/vol) uranyl formate. The stain was removed by blotting with filter paper and the grids dried before imaging using a Tecnai 10 microscope operated at 120 kV with a 4k × 4k CCD camera (Gatan UltraScan 4000) at a pixel size of 1.37 Å and 25 e–/Å2 dose per exposure. Data collection was done with Helicon45. Negative-stain data were processed independently. Using CifFind3 for CTF estimation46, DoG picker for particle picking and iterative multivariate statistical analysis/multireference alignment (MSA/MRA) for reference-free 2D classification47, characteristic class averages for both PRC2 and PRC2–nucleosome samples, obtained from datasets of ~25,000 and ~71,000 particles, respectively, are shown in Supplementary Fig. 1d and were used for further analysis. To localize potential interaction interfaces of PRC2 with nucleosomes, the projection-matching approach was employed. Representative classes of the groups of classes shown in Fig. 1c were matched to 2D projections of the PRC2 EM maps shown in Fig. 1b, low-pass filtered to 15 Å, using Fabs2 software package.43

Data acquisition and initial image processing of PRC2–DiNcl complex. Cryo-gids of PRC2–DiNcl were transferred to a 626 Cryo-Transfer Holder (Gatan) and images were recorded with Helicon45 on a low-base FEI Titan electron microscope operated at 300 keV with a K2 Summit direct electron detector camera (Gatan). 30-frame movies were recorded using a dose rate of 4.6 e–/Å2/sec and a total dose of 40 e/Å2, using a 1.32 Å pixel size (37,879 pix magnification) and a defocus range from 2–4 µm. For the PRC2–DiNcl, reconstruction, three datasets of 3800, 3500 and 800 micrographs were collected and processed individually. Micrographs were motion corrected with MotionCor2 (ref. 48) and CTF estimation was done with GCTF.49 Poor-quality micrographs were removed on the basis of visual inspection of the raw micrographs and the quality of the CTF fits, as well as their CTF figure of merit provided by GCTF, reducing the size of the dataset to ~2,100, 2,000 and 500 micrographs for the three different datasets. Patches were picked using Gautomaton (F. Zhang, MRC LMB, Cambridge, UK), with templates generated by reference-free 2D classification from a subset of ~13,000 manually picked particles. This and all subsequent classification and refinement runs were performed in RELION 1.4 (ref. 47). The particle images and orientations of the
individual reconstructions for the three datasets were used later to generate a 3D-DART server51 according to the linker DNA length used for the in vitro. The DiNcl35 model was flexibly fitted into the map with iMODFIT50 to low approximately the length of DNA used for the reconstitution of dinucleosomes in vitro. The DiNcl40 complex from the PRC2–DiNcl35 yielded an improved overall resolution of 4.9 Å, according to the gold standard FSC = 0.143 criterion39,40 (Supplementary Fig. 7a). The map of the PRC2–Nclmod part of the complex, owing to the marked flexibility and low local resolution, we were not able to determine these changes. Therefore, we can only conclude that the linker lengths observed in the reconstructions harbor approximately the length of DNA used for the reconstruction of dinucleosomes in vitro. The DiNcl4 model was flexibly fitted into the map with IMODFIT40 to low resolution (10 Å) after modeling in COOT40 in order to correct for a change in the nucleosome orientation in the overall modeling process.

For the nucleosome model that was fitted into the density of the PRC2–DiNcl4 complex after signal subtraction to focus on the PRC2–Nclmod interactions, 1bp of nucleosomal DNA near the interaction interface with PRC2, missing from the atomic model (PDB 5LZI, ref. 38) but present in the nucleosome, was added manually in COOT40. Additionally, the DNA helix was locally fitted into the density according to a local direct density and crystal structure. To identify the local density potentially representing the histone H3 tail reaching into the EZH2 active site from the nucleosome, a residual density was obtained by subtracting all density from the map within 4 Å of the fitted models using the color zone and split map options in Chimera40 (Fig. S6 and Supplementary Fig. 7). A polypeptide connecting the last residue of histone H3 to the linker DNA was rigid-body fitted into the complete density (PDB 5LZI, ref. 38). Ψ(Ψ) and the first residue resolved in the EZH2 active site of human PRC2 (PDB SHYN (ref. 40)) was manually modeled into the resulting density as an extended chain using COOT40. The presence of additional density, which we speculate may correspond to part of AEBP2 or EZH2 (between aa 480–515), was not excluded by using the same approach of density subtraction as above, this time including the modeled H3 tail in the underlying models for the color zone and split map options in Chimera40. The resulting density was persistent at high threshold for visualization, at which only little other density that might be accounted for by DNA and other histone tails appeared in the vicinity of the nucleosome and PRC2 models (Supplementary Fig. 7).

Molecular modeling for PRC2–dinucleosome complexes. For the interpretation of the global structure of dinucleosome-bound PRC2 based on the initial PRC2–DiNcl4 complex model (Supplementary Fig. 6a), the human PRC2–DiNcl4 complex (PDB SHYN)10, containing EZH2, EED and the VEFS domain of SUZ12, was rigid-body fitted into the density using UCSF Chimera49, except for the electrostatic surface potential calculations and figures, which were done using PyMOL (The PyMOL Molecular Graphics System, Version 1.7.6, Schrödinger, LLC, http://www.pymol.org/).

Analysis of conformational heterogeneity of Nclmod and Nclsub. For initial analysis and visualization of nucleosome motions, the global PRC2–DiNcl4 complex model was subclassified into five classes using small angular search range (5 pixels) and step size (1 pixel) and fine angular sampling (1°), in order to visualize the conformational heterogeneity of the PRC2–DiNcl4 complex and possibly obtain more homogenous reconstructions. For a visual comparison of the degree of flexibility of each nucleosome relative to PRC2, all six classes were aligned in Chimera40 on the basis of their PRC2 densities. Individual nucleosome models were rigid-body fitted into each nucleosome density, and the dyad axis was visualized by defining axes in Chimera40 on the basis of a consistent pair of atoms in the nucleosome models (Supplementary Fig. 13).

To better resolve individual PRC2–nucleosome interfaces and reduce the impact of the en bloc mobility of each nucleosome on the overall angular assignment accuracy, signal subtraction and masked refinement (Nclmod and Nclsub) were performed as described26. The masked refinement of the PRC2–Nclmod partial complex from the PRC2–DiNcl4 yielded an improved overall resolution of 4.9 Å, according to the gold standard FSC = 0.143 criterion39,40 (Supplementary Fig. 7). This map was sharpened with a B factor of 200 and filtered to 5 Å. The same procedure was carried out for the PRC2–DiNcl4 sample, leading also to an improved resolution of the PRC2–Nclmod part of the complex (Supplementary Fig. 12a–c).

The masked refinement after the signal-subtraction approach did not improve the map of the PRC2–Nclsub part of the complex, owing to the marked flexibility in this region. Instead, after signal subtraction, a masked classification was done to better resolve the variable contacts in the EED and SBD DNA interfaces. The resulting classes showed varying resolution and particle occupancy, so that two classes were chosen representing two distinct orientations of Nclsub (Supplementary Fig. 8a).

Life Sciences Reporting Summary. Further information on experimental design is available in the Life Sciences Reporting Summary.

Data availability. Cryo-EM density maps have been deposited in the Electron Microscopy Data Bank (EMDB) for the complete PRC2 (EMD-7313), PRC2–Nclmod (EMD-7506), PRC2–DiNcl4 (EMD-7307), and PRC2–DiNcl4 complexes (EMD-7308), improved maps after signal subtraction of PRC2–DiNcl4 (EMD-7309 and Supplementary Fig. 7), as well as two models obtained by signal subtraction and 3D classification of the PRC2–Nclsub part of PRC2–DiNcl4 (classes 1 and 3, Supplementary Fig. 8) (EMD-7309 and EMD-7310). The models of PRC2–DiNcl4 complexes obtained by rigid-body docking and flexible fitting are available in Supplementary Dataset 1; PyMOL sessions were generated with version 1.7.6.
References
30. Dyer, P. N. et al. Reconstitution of nucleosome core particles from recombinant histones and DNA. Methods Enzymol. 375, 23–44 (2004).
31. Luger, K., Rechsteiner, T. J., Flaus, A. J., Waye, M. M. Y. & Richmond, T. J. Characterization of nucleosome core particles containing histone proteins made in bacteria. J. Mol. Biol. 272, 301–311 (1997).
32. Lowary, P. T. & Widom, J. New DNA sequence rules for high affinity binding to histone octamer and sequence-directed nucleosome positioning. J. Mol. Biol. 276, 19–42 (1998).
33. McGinity, R. K., Kim, J., Chatterjee, C., Roeder, R. G. & Muir, T. W. Chemically ubiquitylated histone H2B stimulates hDot1L-mediated intranucleosomal methylation. Nature 453, 812–816 (2008).
34. Simon, M. D. et al. The site-specific installation of methyl-lysine analogs into recombinant histones. Cell 128, 1003–1012 (2007).
35. Zhang, K. Gctf: Real-time CTF determination and correction. J. Struct. Biol. 193, 1–12 (2016).
36. Mindell, J. A. & Grigorieff, N. Accurate determination of local defocus and specimen tilt in electron microscopy. J. Struct. Biol. 142, 334–347 (2003).
37. Vasudevan, D., Chua, E. Y. D. & Davey, C. A. Crystal structures of nucleosome core particles containing the ‘601’ strong positioning sequence. J. Mol. Biol. 403, 1–10 (2010).
38. Goddard, T. D., Huang, C. C. & Ferrin, T. E. Visualizing density maps with UCSF Chimera. J. Struct. Biol. 157, 281–287 (2007).
39. van Dijk, M. & Bonvin, A. M. 3D-DART: a DNA structure modelling server. Nucleic Acids Res. 37, W235–W239 (2009).
40. Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. Features and development of Coot. Acta Crystallogr. D. Biol. Crystallogr. 66, 486–501 (2010).
Life Sciences Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form is intended for publication with all accepted life science papers and provides structure for consistency and transparency in reporting. Every life science submission will use this form; some list items might not apply to an individual manuscript, but all fields must be completed for clarity.

For further information on the points included in this form, see Reporting Life Sciences Research. For further information on Nature Research policies, including our data availability policy, see Authors & Referees and the Editorial Policy Checklist.

- **Experimental design**

1. **Sample size**
   Describe how sample size was determined.
   N/A

2. **Data exclusions**
   Describe any data exclusions.
   N/A

3. **Replication**
   Describe whether the experimental findings were reliably reproduced.
   N/A

4. **Randomization**
   Describe how samples/organisms/participants were allocated into experimental groups.
   N/A

5. **Blinding**
   Describe whether the investigators were blinded to group allocation during data collection and/or analysis.
   N/A

   Note: all studies involving animals and/or human research participants must disclose whether blinding and randomization were used.

6. **Statistical parameters**
   For all figures and tables that use statistical methods, confirm that the following items are present in relevant figure legends (or in the Methods section if additional space is needed).

   n/a  Confirmed
   - [ ] The exact sample size \((n)\) for each experimental group/condition, given as a discrete number and unit of measurement (animals, litters, cultures, etc.)
   - [ ] A description of how samples were collected, noting whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
   - [ ] A statement indicating how many times each experiment was replicated
   - [ ] The statistical test(s) used and whether they are one- or two-sided (note: only common tests should be described solely by name; more complex techniques should be described in the Methods section)
   - [ ] A description of any assumptions or corrections, such as an adjustment for multiple comparisons
   - [ ] The test results (e.g. \(P\) values) given as exact values whenever possible and with confidence intervals noted
   - [ ] A clear description of statistics including central tendency (e.g. median, mean) and variation (e.g. standard deviation, interquartile range)
   - [ ] Clearly defined error bars

   *See the web collection on statistics for biologists for further resources and guidance.*
Software

Policy information about availability of computer code

7. Software

Describe the software used to analyze the data in this study.

No new software was used in this work

For manuscripts utilizing custom algorithms or software that are central to the paper but not yet described in the published literature, software must be made available to editors and reviewers upon request. We strongly encourage code deposition in a community repository (e.g. GitHub). Nature Methods guidance for providing algorithms and software for publication provides further information on this topic.

Materials and reagents

Policy information about availability of materials

8. Materials availability

Indicate whether there are restrictions on availability of unique materials or if these materials are only available for distribution by a for-profit company.

N/A

9. Antibodies

Describe the antibodies used and how they were validated for use in the system under study (i.e. assay and species).

N/A

10. Eukaryotic cell lines

a. State the source of each eukaryotic cell line used.

N/A

b. Describe the method of cell line authentication used.

N/A
c. Report whether the cell lines were tested for mycoplasma contamination.

N/A
d. If any of the cell lines used are listed in the database of commonly misidentified cell lines maintained by ICLAC, provide a scientific rationale for their use.

N/A

Animals and human research participants

Policy information about studies involving animals; when reporting animal research, follow the ARRIVE guidelines

11. Description of research animals

Provide details on animals and/or animal-derived materials used in the study.

N/A

Policy information about studies involving human research participants

12. Description of human research participants

Describe the covariate-relevant population characteristics of the human research participants.

N/A