Dominant Alleles Identify SET Domain Residues Required for Histone Methyltransferase of Polycomb Repressive Complex 2*

Preeti Joshi†, Elizabeth A. Carrington‡, Liangjun Wang§, Carrie S. Ketel∥, Ellen L. Miller¶, Richard S. Jones†, and Jeffrey A. Simon**

Received for publication, June 10, 2008, and in revised form, August 6, 2008. Published, JBC Papers in Press, August 8, 2008, DOI 10.1074/jbc.M804442200

PcG, Polycomb group; PRC2, Polycomb repressive complex 2; HMTase, histone methyltransferase; PHO, pleiohomeotic.

Polycomb gene silencing requires histone methyltransferase activity of Polycomb repressive complex 2 (PRC2), which methylates lysine 27 of histone H3. Information on how PRC2 works is limited by lack of structural data on the catalytic subunit, Enhancer of zeste (E(Z)), and the paucity of E(z) mutant alleles that alter its SET domain. Here we analyze missense alleles of Drosophila E(z), selected for molecular study because of their dominant genetic effects. Four missense alleles identify key E(Z) SET domain residues, and a fifth is located in the adjacent CXC domain. Analysis of mutant PRC2 complexes in vitro, and H3-K27 methylation in vivo, shows that each SET domain mutation disrupts PRC2 histone methyltransferase. Based on known SET domain structures, the mutations likely affect either the lysine-substrate binding pocket, the binding site for the adenosylmethionine methyl donor, or a critical tyrosine predicted to interact with the substrate lysine e-amino group. In contrast, the CXC mutant retains catalytic activity, Lys-27 specificity, and trimethylation capacity. Deletion analysis also reveals a functional requirement for a conserved E(Z) domain N-terminal to CXC and SET. These results identify critical SET domain residues needed for PRC2 enzyme function, and they also emphasize functional inputs from outside the SET domain.

The Polycomb group (PcG)³ proteins are highly conserved chromatin components that work together to silence target genes during development (for reviews see Refs. 1, 2). The original and best characterized biological function of PcG repressors is to maintain repression of Drosophila Hox genes along the anterior-posterior body axis (3–6). Besides Hox regulation, PcG proteins are implicated in silencing dozens of other developmental control genes in flies and mammals (7–13). A key recent example is the critical role of PcG proteins in keeping differentiation genes silent in human embryonic stem cells, thereby maintaining the self-renewing capacity of these cells (9, 10). Their widespread biological functions and striking evolutionary conservation establish PcG proteins as one of the premier models for deciphering chromatin mechanisms that perpetuate gene silencing during development.

Drosophila PcG proteins sort into at least three discrete chromatin complexes termed PRC1, PRC2, and PHO-RC (14–18). Genetic studies show that individual components of each complex are required for Hox gene silencing in vivo, indicating that these complexes work together to maintain silent chromatin states. Although the integrated mechanisms by which PcG complexes execute gene silencing remain to be determined, subunit compositions and biochemical functions have been defined for individual purified complexes.

Polycomb repressive complex 2 (PRC2) has four core subunits, E(Z), ESC, SU(Z)12, and NURF55, and it possesses an intrinsic enzyme activity that methylates histone H3 on lysine 27 (14, 15, 19, 20). The catalytic subunit is E(Z), which contains a copy of the SET domain commonly found in lysine methyltransferases (21–23). PRC2 is remarkably conserved in multicellular eukaryotes. Fly and human PRC2 have very similar compositions and identical specificity for H3-K27 (14, 15, 19, 20). In mouse and human embryonic stem cells, PRC2 is a key corepressor that methylates H3-K27 and maintains silencing at differentiation gene targets (9, 10, 24, 25). PRC2 function also contributes to gene expression signatures that identify aggressive forms of human cancers (26–30). In addition, PRC2 and Lys-27 methylation are implicated in mammalian X chromosome inactivation (31), and the Caenorhabditis elegans version of PRC2 methylates Lys-27 and functions in germ line silencing (32, 33). Strikingly, plant versions of PRC2, which are required for key events in seed and flower development (34, 35), also share the same subunits and Lys-27 specificity (36). This conservation across kingdoms indicates that PRC2 is part of an ancient strategy for chromatin silencing, which employs H3-K27 methylation as a repressive chromatin mark.

To understand PRC2 function, it is important to dissect and define the subunit contributions of the enzyme complex. The core of the PRC2 active site is provided by the E(Z) SET domain, which is required for methyltransferase activity in vitro and Hox gene silencing in vivo (14, 15, 19, 20). Several additional inputs are also critically required for PRC2 function. Each of the three noncatalytic subunits is needed for full HMTase activity, with
the most crucial contributions provided by the ESC and SU(Z)12 subunits, which both directly contact E(Z) in the enzyme complex (37–40). Single loss of either ESC or SU(Z)12 disrupts PRC2 function in vitro and in vivo (37–42). In addition, several E(Z) domains besides the SET domain are required for PRC2 function (see Fig. 1A). Two of these E(Z) domains function in PRC2 complex assembly, with an N-terminal E(Z) domain (EID) needed for stable binding to ESC (43–45) and homology domain II used to bind SU(Z)12 (38, 46). Although less is known about the molecular role of E(Z) domain I (Fig. 1A), its 77% identity from flies to humans implies that it, too, is functionally important. In addition, the CXC domain immediately preceding the SET domain is implicated in PRC2 catalytic function (20, 38) and in mediating contact with the PcG protein PHO (47), which targets PRC2 to genomic sites in vivo. Thus, PRC2 is a highly collaborative enzyme complex with multiple subunit and domain inputs needed for robust activity.

In light of these multiple inputs, it is important to define E(Z) functional interactions with its partner subunits, and with the nucleosome substrate, in molecular terms. However, relatively little is currently known about workings of the E(Z) SET domain because a three-dimensional structure has not yet been described. Information about functional residues within the E(Z) SET domain is equally limited by the paucity of Drosophila missense mutations that specifically alter this domain. Here we present molecular and biochemical analysis of five E(Z) mutant alleles selected because their genetic properties suggest functionally altered versions of E(Z). Four of these are missense mutations that lie within the E(Z) SET domain. We find that each of these SET mutations disrupt H3-K27 methylation in vitro and in vivo, thereby defining SET domain residues critical for PRC2 activity. Based upon other SET domains with known structures, we suggest that these E(z) mutations affect motifs implicated in either the lysine-substrate binding pocket, the binding site for the adenosylmethionine methyl donor, or a critical tyrosine that may contact the reactive lysine ε-amino group. These missense alleles, together with additional site-directed substitutions, define requirements within the E(Z) SET domain for histone methyltransferase activity. In addition, deletion analysis reveals that E(Z) domain I is also needed for PRC2 enzyme function.

EXPERIMENTAL PROCEDURES

Sequence Determination of Mutant Alleles—Genomic DNA was extracted from larvae or adult flies that were hemizygous for the E(z)1 mutant, E(z)2 mutant, E(z)3 allele, E(z)4, or E(z)62 alleles in trans to a deficiency of the E(z) chromosomal region. The E(z) genomic region was amplified by PCR and sequenced as described previously (48).

Baculovirus Constructs, Site-directed Mutagenesis, and Purification of Recombinant PRC2 Complexes—Full-length cDNAs encoding FLAG-ESC, HA-ESC, E(Z), SU(Z)12, and NURF-55 inserted into pFastBac1 were described previously (14, 38). Site-directed E(Z) mutations were generated in pFastBac1-E(Z) using the QuickChange mutagenesis kit (Stratagene). N-terminal deletions of E(Z) were generated by PCR and were confirmed by DNA sequencing. FLAG tags were inserted at the N termini of E(Z)67–760 and E(Z)161–760, whereas E(Z)Δ86–152 was used in untagged form. Baculovirus expression of recombinant proteins was performed using the Bac-to-Bac system (Invitrogen). Anti-FLAG immunopurification affinity purification of PRC2 complexes was performed as described (38). Complexes bearing E(Z) missense mutations or E(Z)Δ86–152 were purified via FLAG-ESC, whereas E(Z)67–760 and E(Z)161–760 complexes were purified via FLAG tags on the E(Z) subunit after coexpression with HA-ESC. Mutant complexes were prepared in parallel with a wild-type control and purified at least twice independently for each complex. Assembly of complexes was assessed by Coomassie Blue staining after SDS-gel electrophoresis and, in the case of Fig. 5C, by Western blotting. Primary antibodies were anti-FLAG (Sigma) and anti-HA (Covance) mouse monoclonals used at 1:2,000 and 1:5,000, respectively. Anti-NURF55 was a rabbit polyclonal generated using a His6-NURF55 full-length fusion as immunogen, affinity-purified as described (47), and used at 1:2,000.

Histone Methyltransferase Assays and Substrates—Histone methyltransferase assays were performed as described (38) and were repeated at least twice using independently prepared complexes. Polynucleosome substrate, consisting of 8–12-mers purified from HeLa cells, was prepared as described (38). H3/H4 tetramers, which were also used as substrate in HMTase assays, were prepared after coexpression of Drosophila histone H3 and H4 in Escherichia coli (49). The H3/H4 tetramers contained either wild-type or K27A mutant forms of H3.

Generation of E(z) Mutant Embryos and Analysis of in Vivo Histone Methylation—Wild-type embryos were collected from stock y Df(1)w67 -2 at 25 °C. E(z)61 mutant embryos were collected from a homozygous E(z)61 stock at 18 or 29 °C. E(z)mut/E(z)mut adults (where E(z)mut represents E(z)1, E(z)62, E(z)61, E(z)62, or E(z)63) were generated at 18 °C, and their mutant embryo progeny were collected at 29 °C. Embryo extracts were prepared as described (38). Proteins were fractionated on 15% SDS-polyacrylamide gels and transferred to Protran (Schleicher & Schuell). Blots were blocked for at least 30 min in 5% nonfat dry milk and incubated in primary antibody overnight at 4 °C. The dimethyl-H3K27 and trimethyl-H3K27 antibodies were used at 1:500. These antibodies have been described previously (50) and were obtained from Upstate Biotechnology, Inc. A monoclonal antibody against α-tubulin (DM-1A, Sigma) was used at 1:2,000. Horseradish peroxidase-conjugated goat anti-rabbit and goat anti-mouse secondary antibodies (Jackson ImmunoResearch) were used at 1:5,000 and 1:2,000, respectively. Signals were developed using an ECL detection kit (Amersham Biosciences).

Mass Spectrometric Analysis of Lys-27-methylated Histone H3—HMTase assays were performed as described (38) using recombinant H3/H4 tetramers as substrate and 0.6 mM nonradioactive S-adenosylmethionine as methyl donor. The reaction products were electrophoresed on a 15% SDS-polyacrylamide gel and stained with Coomassie Blue, and the histone H3 bands were excised. The H3 gel slices were destained, treated with propionic anhydride, and digested with trypsin as described (50). Peptides were then eluted, purified on a μC18 ZipTip (Millipore), and analyzed using an Applied Biosystems QSTAR pulsar quadrupole time-of-flight mass spectrometer with a
Requirements for PRC2 Histone Methyltransferase

E(z) Mutant Alleles with Dominant Phenotypes Distinct from E(z) Null—The PRC2 catalytic subunit, E(Z), contains multiple functional domains (Fig. 1A). Most of these domains were initially inferred by conservation among E(Z) homologs (51). Definitive in vivo proof of domain contributions can be provided by mutant alleles that alter single residues. However, relatively few E(z) missense alleles have been molecularly characterized to date. Three recessive temperature-sensitive missense mutations, E(z)28A, E(z)32, and E(z)61, provide evidence that domain II and the CXC domain are required for E(Z) function in vivo (3, 48, 52) (see Fig. 1A). The only other E(z) missense allele yet described molecularly is a gain-of-function E(z) mutation, called E(z)Trm, which is located in the C-terminal portion of the SET domain (53). The Trm mutant version of E(Z) retains PRC2 enzyme activity (20).

Many additional fly E(z) mutant alleles have been isolated, but their molecular lesions have not been determined. Because we are most interested in studying mutants that inactivate or alter specific domains rather than cause complete E(Z) loss, we sought E(z) mutants whose phenotypes are distinct from E(z) null. We selected five E(z) mutant alleles, E(z)1, E(z)62, E(z)ω1, E(z)ω2, and E(z)ω3, because genetic studies established that each displayed one or more dominant phenotypes that are more severe than E(z) null or not observed with E(z) null (Table 1). For example, E(z)1 and E(z)62 are dominant modifiers of gene silencing by the zeste chromatin protein (3, 54), and the matrix-assisted laser desorption ionization mass spectrometry. The observed masses (1602 Da for the unmodified H3 peptide, 1616 for the monomethylated H3 peptide, and 1588 for the trimethylated H3 peptide) agree precisely with expected masses for each peptide species bearing the following modifications: unmethylated (H3 peptide 27–40 plus three propionyl groups), monomethylated (peptide 27–40 plus one methyl group), dimethylated (peptide 27–40 plus two propionyl groups plus two methyl groups), and trimethylated (peptide 27–40 plus two propionyl groups plus three methyl groups). The identity of the unmethylated H3 peptide peak at 1601.9 m/z was also confirmed by manual MS/MS sequencing.

RESULTS

E(z) Mutant Alleles with Dominant Phenotypes Distinct from E(z) Null—The PRC2 catalytic subunit, E(Z), contains multiple functional domains (Fig. 1A). Most of these domains were initially inferred by conservation among E(Z) homologs (51).
alleles their phenotypic severities, with all four SET mutants display- ing recessive lethality, whereas son2 is hemizygous viable (Table 1). Intriguingly, each of the SET mutations changes residues associated with conserved motifs (Fig. 1B), whose functions are predicted by analogy to SET domains whose structures are known (22, 23, 56–59). E(z)62 and E(z)61 both affect a hydrophobic motif (FLF in E(z)) that contacts the substrate lysine (22, 58, 60). In contrast, E(z)son1 is associated with a conserved GXG motif implicated in Ado-Met binding (22, 23). Finally, E(z)1 substitutes the tyrosine of a YAG motif, which plays a critical role at the catalytic site. In other SET domains, this tyrosine contacts and orients the lysine ε-amino group for the methyl transfer reaction (22). Thus, these mutations may impair the enzyme in mechanistically different ways by affecting lysine substrate interaction, methyl donor binding, or configuration of the core catalytic site. Indeed, E(z)62 and E(z)61 comprise a phenotypic class distinct from E(z)son1 and E(z)1 (Table 1), which may reflect these underlying differences in altered SET domain subfunctions.

### Methylation Requirements for PRC2 Histone Methyltransferase

**TABLE 1**

| Phenotypes of E(z) mutant alleles | E(z) alleles |
|----------------------------------|-------------|
| Lethal phasea | E(z)61 | 62 | 1 | 63 (null) |
| (in trans to dfl(3)kldad) | L3 | L3 | P | P | V | P |
| Lethal phaseb | P | P | V | V | V | V |
| (in trans to E(z)61) | Meso to pro | + | + | + | + | + |
| Pro to Meso | + | + | + | + | + | + |
| Pro 2nd tarsal segmentc | zeste | eye colord | R | RO | Y | DO | DO |
| Suppression of nos | + | ND | + | + | + | + |

**Figure 2. Histone methyltransferase of E(z) mutants in vitro.** A, assembly. B, methyltransferase activity of the five E(z) missense mutants in context of recombinant PRC2. Substrates were HeLa polycombosomes (top panels) or recombinant H3/H4 tetramers (bottom panels). C, analysis of site-directed E(z) mutations F679Y and F681Y (left panel), which alter the putative lysine-substrate binding pocket. Both mutants assemble into PRC2 (middle panel), but F679Y is highly active, whereas F681Y has little or no HMTase (right panel). WT, wild type.
Functional Tests of N-terminal E(Z) Domains—In addition to the CXC and SET domains, three N-terminal E(Z) domains are recognized based on evolutionary conservation and functional tests. Of these, domain II can mediate E(Z) association with SU(Z)12 in pairwise binding assays (38, 46), and it is required for SU(Z)12 assembly into PRC2 (38). The EID is implicated as an ESC-interacting domain, based on pairwise binding tests (43–45), but the EID requirement for ESC assembly into PRC2 has not been reported. Finally, the role of domain I in either subunit assembly or enzyme activity has not been described. The high conservation of domain I, which is 77% identical from flies to humans, implies a key functional role.

To assess roles of the two most N-terminal E(Z) domains in vitro, we constructed and tested three deletions in this region (Fig. 5A). E(Z)67–760 removes the EID, E(Z)161–760 removes both the EID and domain I, and E(Z)Δ86–152 is an in-frame deletion that removes just domain I. These E(Z) deletion mutants were tested for assembly into recombinant PRC2 and for enzyme activity in histone methyltransferase assays. Purification of wild-type PRC2 and E(Z)Δ86–152 was performed, as above, via FLAG-tagged ESC. However, because the two E(Z) deletions removing the EID seemed likely to disrupt ESC-E(Z) interaction, these complexes were purified using FLAG tags at the N termini of the E(Z) mutant proteins. FLAG tags at the N terminus of either ESC or E(Z) are tolerated without disrupting PRC2 methyltransferase (37, 38).

Fig. 5B shows the subunit compositions of complexes bearing the deleted versions of E(Z). Wild-type and E(Z)Δ86–152 complexes contain all four subunits (Fig. 5B, lanes 1 and 3), indicating that E(Z) domain I is not required for assembly of partner subunits. In contrast, E(Z)67–760 and E(Z)161–760 yield trimeric complexes that specifically lack ESC (Fig. 5B, lanes 2 and 4). Because the NURF55/ESC doublet is not optimally resolved here, we performed Western blot analysis on these same complexes to independently assess the presence or absence of these two subunits. Fig. 5C confirms that all complexes contain NURF55, but only wild-type and E(Z)Δ86–152 retain the ESC subunit (lanes 1 and 3). These results are consistent with ESC assembly into PRC2 requiring a key stable contact with the EID of E(Z). A three-dimensional view of this binding interface has recently been described for the corresponding human PRC2 subunits (43).

Histone methyltransferase assays were performed on these recombinant complexes. As shown in Fig. 5D, despite assembly of a full four-subunit complex, the E(Z)Δ86–152 mutant version of PRC2 shows little or no enzyme activity. As expected because of the loss of ESC, the E(Z)67–760 and E(Z)161–760 trimeric complexes also lacked HMTase (data not shown). These results suggest that domain I of E(Z) is needed for PRC2 enzyme function. Thus, in addition to the key influences of SU(Z)12 and ESC in boosting E(Z) catalysis (37–42), domain I may provide another functional input from outside the SET domain.
requirements for PRC2 histone methyltransferase

FIGURE 4. Trimethylation capacity and lysine specificity of E(z)-F679Y and E(z)son2 mutant PRC2 complexes. A, mass spectrometric detection of unmodified and Lys-27-methylated forms of histone H3 tail tryptic peptide 27–40 (KSAPATGGKPHRE), determined as described previously (50). Drosophila H3/H4 tetramers, purified from E. coli, were either untreated or methylated by wild-type (WT) or indicated PRC2 mutant complexes. Histone H3 was then gel-excised, propionylated, and trypsin-digested, and peptide masses were determined by matrix-assisted laser desorption ionization time-of-flight. Only unmodified and monomethyl lysine get propionylated, which leads to larger peptide masses as compared with their di- and trimethyl counterparts. The labeled peaks correspond to singly charged ions observed at 1601.9 m/z for the unmodified H3 peptide, 1615.9 m/z for the monomethylated peptide, 1573.9 m/z for the dimethylated peptide, and 1587.9 m/z for the trimethylated peptide. B, lysine specificity of mutant PRC2 complexes bearing either the E(z)son2 (left panel) or E(z)-F679Y (right panel) mutations. Both mutant complexes resemble wild-type PRC2 in ability to methylate wild-type recombinant H3/H4 tetramers and failure to methylate tetramers bearing the K27A mutant version of histone H3.

Discussion

Motifs and Subfunctions within the E(z) SET Domain—Three functional sites required in histone lysine methyltransferases are a lysine substrate-binding site, a methyl donor cofactor-binding site, and a catalytic site that promotes methyl transfer from donor to substrate. Different SET domains configure these enzyme sites using a combination of highly conserved motifs and less conserved regions across the 130-amino acid expanse of the domain. Structural studies show that the SET domain active site features an unusual protein fold termed a pseudoknot (for reviews see Refs. 22, 23, 64). The pseudoknot is formed from the two most highly conserved motifs, RFXNH-

SCXPN and EELXFDY, located in the C-terminal portion of the SET domain (motifs IV and V in Fig. 6), with the latter peptide element threading the loop created by the former. Two additional motifs from the N-terminal portion of the domain, featuring conserved GXG and YXG elements (motifs I and II in Fig. 6), are typically juxtaposed to the pseudoknot to jointly create the architecture of the three functional sites. The adenosylmethionine binding pocket is contributed by the GXG motif plus the RFXNHS portion of motif IV. Consequently, we suggest that the E(z)son1 (Fig. 1B) mutation, which closely flanks this GXG motif, alters E(z) binding to S'-adenosylmethionine cofactor. We note that a methylation-defective mutant of MES-2, the worm E(z) homolog, changes two residues that immediately flank the son1 substitution, suggesting a similar biochemical explanation (32, 33). In contrast, the E(z)1 mutation (Y655N; Fig. 1B) alters the YXG motif tyrosine, which interacts with the substrate ε-amino group at the catalytic site. This critical tyrosine is analogous to Tyr-178 in DIM-5 and Tyr-245 in SET7/9, which hydrogen bond to the ε-amino group to align it favorably for methyl transfer (22, 57, 65, 66). Thus, the simplest explanation for E(z)-Y655N loss of enzyme activity is a defective catalytic site.

The lysine substrate binding pocket is lined by large hydrophobic residues, some of which are donated by motif V of the pseudoknot (EELXFDY). However, the remaining composition of the lysine binding channel is more variable, which is thought to reflect differing substrate specificities of SET domain enzymes for different histone lysine residues (67). A more variable middle portion of the SET domain, sometimes called SET-I (for SET-insert), also contributes hydrophobic residues to the lysine-binding pocket (58, 64). This SET-I region encompasses motif III (FLF in E(z)) and is implicated in helping determine lysine choice (64, 67). In DIM-5, SET7/9, SET8, and vSET, the aromatic residues of motif III form one side of the lysine channel wall, which parallels the methylene portion of the lysine side chain (22, 56, 58, 60, 66). Thus, the E(z)62 (S678L) and E(z)son1 (F681I) mutations (Fig. 1B) likely exert their effects by altering this lysine binding pocket. We note that E(z)son1 renders PRC2 inactive despite the
substitution of one large hydrophobic residue for another. Our side-directed F679Y and F681Y mutations (Fig. 2C) further address requirements for particular large hydrophobes in these locations. E(Z)-F679Y displays robust HMTase activity, within ~2-fold of wild type (Fig. 2C), and it retains substrate specificity for lysine 27 (Fig. 4B). The fact that worm MES-2 normally features Tyr at this position, but is defective with His at this position, provides independent evidence of flexibility limits at this residue (32, 33). In contrast, there is little flexibility at Phe-681 because activity is disrupted by substitutions to either Tyr or Ile (Fig. 2B and C). In agreement with this, Phe is conserved at the corresponding position in 17/21 characterized SET domains (60).

Mutational analyses of other SET domains show that Phe versus Tyr residues at certain positions of the lysine substrate binding pocket can influence product specificity (62). Thus, Tyr at these positions is a feature of mono-/dimethyltransferases, whereas Phe promotes activity as a trimethyltransferase. This Phe/Tyr switch has been demonstrated for DIM-5, G9a, SET7/9, and vSET (56, 58, 62). We tested whether Phe-to-Tyr mutations at Phe-679 or Phe-681 altered the methylation multiplicity of E(Z), because mutations at the corresponding positions of the Lys-27-specific HMTase, vSET, had this effect (58, 60). However, F681Y reduced overall HMTase and F679Y retained trimethylation capacity (Fig. 4A). This suggests that E(Z) residues that dictate mono/dimethylation are located elsewhere and that this determinant is not configured similarly in E(Z) and vSET. The E(Z) and vSET enzymes also differ in requirements for noncatalytic subunits (37, 38, 58, 63) and in conservation of the YAG element (motif II; see Fig. 1B). We note that the experimentally determined Phe/Tyr switch residues are located in different positions within the SET domain primary sequence in different HMTases (58, 62). It is also possible that E(Z) methylation multiplicity is influenced by one or more of its noncatalytic partners.

Multiple E(Z) Phenotypes and Biological Functions in Drosophila—In addition to Hox gene silencing, E(Z) is implicated in chromatin regulation at many other fly loci (11–13), and also in mitotic chromosome function (52, 68–70). Certain of these non-Hox regulatory functions provided phenotypic readouts exploited in the isolation and characterization of the five E(z) missense alleles described here. One such phenotype, modification of zeste" yellow eye color, is a hallmark of mutations in several PcG genes (54). Although not understood in molecular terms, dominant suppression of zeste" eye color because of PcG gene loss is conveniently exploited in PcG genetic assays (54, 71, 72). Indeed, the PRC2 subunit SU(z)12 was originally identified as a zeste" modifier (73), hence the name Suppressor of zeste-12, E(z) null mutations, such as E(z)" (63), are weak dominant suppressors of zeste"; they shift eye color from light to slightly darker orange when tested with the + w/+ chromosome used here (Table 1). In contrast, E(z)" suppresses more strongly than null, yielding a reddish eye color, and E(z)" actually has the opposite effect; it enhances zeste" silencing of white to yield a yellow eye color (3) (Table 1).

Another phenotypic readout is the ability of E(z) alleles to suppress nanos mutations (55). Nanos is a maternally provided protein required for development of fly abdominal segments (74). Nanos represses translation of hunchback (HB) in posterior segments, which helps produce the HB gradient that defines head-to-tail body regions in early fly embryos (74, 75). HB, in turn, defines the spatial expression domains of subordi-
nate gap gene products, including Kruppel, knirps, and giant (76). This process occurs hours before PcG silencing of Hox genes. It was therefore unexpected that E(z) alleles (son1, son2, and son3) were isolated in a genetic screen for suppressors of ganoz (55). This suppression occurs through altered levels of knirps and giant expression; these gap genes are abnormally silenced in ganoz mutants, but their expression is restored in nanos;E(z)0o0 double mutants (55). Thus, knirps and giant are, like Hox genes, silenced by E(z) in vivo. Although we now appreciate that E(z) silences many distinct target genes, this Pelegri and Lehmann study (55) provided one of the first examples of PcG regulation of non-Hox target genes.

Phenotypic Differences among E(z) SET Mutants—All four of the E(z) SET domain missense alleles show dramatically reduced histone methyltransferase activity in vitro (Fig. 2B) and in vivo (Fig. 3B). Thus, the simplest molecular explanation for their genetic effects would be that they act as dominant-negatives. In this scenario, if these catalytically inactive versions of E(z) assemble into PRC2 in vivo, as they do in vitro (Fig. 2A), then they should poison these complexes and cause loss-of-function. This type of mechanism could explain why these mutants show stronger dominant effects than E(z) null (Table 1). However, there are phenotypic differences among the four E(z) mutants, which indicates that they do not behave identically in vivo. In particular, the two lysine substrate pocket mutants, E(z)0o2 and E(z)0o1, appear to define a phenotypic sub-class. These two mutants die earlier as hemizygotes and their Hox developmental defects are distinct from the other three mutants and nulls (Table 1). To decipher in vivo mechanisms of these mutants, it will be necessary to determine whether and how mutant PRC2 complexes are directly associated with individual target loci. In particular, the consequences of bound but inactive PRC2 might be different at a target gene, such as giant, with PcG complexes bound outside the transcription unit (13), as opposed to a Hox gene like Ubx, which accumulates PcG complexes at sites both upstream and within the transcription unit (47, 77, 78). It is also difficult to interpret phenotypic differences among E(z) mutants without knowing if PRC2 exerts functions in vivo that are independent of its methyltransferase. Further work will be needed to determine in vivo mechanisms of these E(z) SET mutants.

E(Z) Functional Inputs from Outside the SET Domain—The E(z)0o2 mutant is different in several ways from the other four missense alleles analyzed here. First, it is the only allele of the five to retain adult viability as a hemizygote (Table 1), which indicates that Son2 mutant protein retains substantial partial E(Z) function. Second, son2 is located in the CXC domain (Fig. 1A), a highly conserved cysteine-rich domain that flanks the SET domain. In vitro tests have shown that other E(Z) CXC missense mutants assemble into PRC2 but display reduced enzyme activity (20, 38). Thus, the E(Z) CXC domain could parallel the zinc-binding PRE-SET domains of the SUV39 family of HMTases (23, 64), which are similarly needed for HMTase function (21). In these HMTases, such as DIM-5 and CLR4, the PRE-SET domain contacts the adjacent SET domain, which may provide structural stabilization (64, 65). However, a catalytic defect is not apparent with the E(z)0o2 CXC mutant, which retains HMTase levels, Lys-27 specificity, and trimethylation capacity similar to wild-type (Fig. 2B, Fig. 3B, and Fig. 4). An additional function of the E(Z) CXC domain is to mediate interaction with another PcG protein called pleiohomeotic (PHO) (47). PHO is a sequence-specific DNA-binding protein (79) implicated in targeting PRC2 to chromatin sites in vivo (47). Thus, it is possible that the son2 defect could reflect altered PRC2 targeting in vivo rather than defects in intrinsic enzyme function. However, the viability of son2 hemizygotes suggests that any targeting defects must be much more subtle than the wholesale E(Z) dislodging seen with stronger loss-of-function CXC mutants (48).

In addition to the CXC domain, our analysis also suggests that domain I of E(Z) is important for PRC2 function. Despite its 77% identity from flies to humans, the contribution of this domain has received scant attention. Because there are no known E(z) missense alleles located in domain I, we removed it by in-frame deletion and tested this E(Z) mutant protein for PRC2 assembly and activity in vitro. As shown in Fig. 5, E(Z)Δ86–152 can assemble into a four-subunit complex, but its enzyme activity is dramatically reduced. This suggests that the main function of domain I is not to bind and recruit core subunits into PRC2. Consistent with this idea, domain I is conserved in the divergent C. elegans version of PRC2, which lacks subunits that resemble SU(Z)12 or NURF-55 (32, 38, 80). One possibility is that domain I might directly contact and influence the organization or function of the E(Z) SET domain. Intriguingly, domain I coincides with a binding site for another PcG protein, called PCL in flies and PHF1 in humans (81). PCL is required for PRC2 chromosome binding in vivo (82) and also has been shown to stimulate PRC2 enzyme activity (61, 83, 84). Perhaps PCL association with E(Z) domain I optimizes or stabilizes an interaction with the SET domain.

Clearly, more work is needed to determine the mechanisms of PRC2 function in vitro and in vivo. At the core of the enzyme, our analysis shows that discrete motifs within the E(Z) SET domain, highlighted by conservation and structural studies of other methyltransferases, are critical for PRC2 function. However, multiple inputs from subunits and domains outside the SET domain also need to be deciphered before the inner workings of this highly collaborative and functionally critical enzyme complex are fully revealed.

Acknowledgments—We are grateful to Francisco Pelegri and Ruth Lehmann for providing the E(z)0o mutants. We thank Lori Anderson for guidance and input on mass spectrometry and Nicole Francis for providing constructs to produce recombinant H3/H4 tetramers. We also thank Mike O’Connor for discussions and comments on the manuscript.

REFERENCES

1. Sparmann, A., and van Lohuizen, M. (2006) Nat. Rev. Cancer 6, 846–856
2. Schwartz, Y. B., and Pirrota, V. (2007) Nat. Rev. Genet. 8, 9–22
3. Jones, R. S., and Gelbart, W. M. (1990) Genetics 126, 185–199
4. Struhl, G., and Akam, M. (1985) EMBO J. 4, 3259–3264
5. Lewis, E. B. (1978) Nature 276, 565–570
6. Simon, J., Chiang, A., and Bender, W. (1992) Development (Camb.) 114, 493–505
7. Squazzo, S. L., O’Geen, H., Komashko, V. M., Krig, S. R., Jin, V. X., Jang, S. W., Margueron, R., Reinberg, D., Green, R., and Farnham, P. J. (2006)
Requirements for PRC2 Histone Methyltransferase

and Pimpinelli, S. (2008) Chromosoma (Berl.) 117, 25–39
71. Kalisch, W. E., and Rasmuson, B. (1974) Hereditas 78, 97–104
72. Bornemann, D., Miller, E., and Simon, J. (1996) Development (Camb.) 122, 1621–1630
73. Birve, A., Sengupta, A. K., Beuchle, D., Larsson, J., Kennison, J. A., Rasmuson-LeStander, A., and Muller, J. (2001) Development (Camb.) 128, 3371–3379
74. Wang, C., and Lehmann, R. (1991) Cell 66, 637–647
75. Wharton, R. P., and Struhl, G. (1991) Cell 67, 955–967
76. Hulskamp, M., and Tautz, D. (1991) BioEssays 13, 261–268
77. Papp, B., and Muller, J. (2006) Genes Dev. 20, 2041–2054
78. Kahn, T. G., Schwartz, Y. B., Dellino, G. I., and Pirrotta, V. (2006) J. Biol. Chem. 281, 29064–29075
79. Brown, J. L., Mucci, D., Whiteley, M., Dirksen, M. L., and Kassis, J. A. (1998) Mol. Cell 1, 1057–1064
80. Xu, L., Fong, Y., and Strome, S. (2001) Proc. Natl. Acad. Sci. U. S. A. 98, 5061–5066
81. O’Connell, S., Wang, L., Robert, S., Jones, C. A., Saint, R., and Jones, R. S. (2001) J. Biol. Chem. 276, 43065–43073
82. Savla, U., Benes, J., Zhang, J., and Jones, R. S. (2008) Development (Camb.) 135, 813–817
83. Cao, R., Wang, H., He, J., Erdjument-Bromage, H., Tempst, P., and Zhang, Y. (2008) Mol. Cell. Biol. 28, 1862–1872
84. Sarma, K., Margueron, R., Ivanov, A., Pirrotta, V., and Reinberg, D. (2008) Mol. Cell. Biol. 28, 2718–2731
85. Xiao, B., Jing, C., Kelly, G., Walker, P. A., Muskett, F. W., Frenkiel, T. A., Martin, S. R., Sarma, K., Reinberg, D., Gamblin, S. J., and Wilson, J. R. (2005) Genes Dev. 19, 1444–1454