Polycomb purification by in vivo biotinylation tagging reveals cohesin and Trithorax group proteins as interaction partners

Gero Strübbe[a], Christian Popp[b], Alexander Schmidt[c,1], Andrea Pauli[d,2], Leonie Ringrose[e], Christian Beisel[a], and Renato Paro[a,f,3]

[a]Department of Biosystems Science and Engineering, Swiss Federal Institute of Technology Zürich, Mattenstrasse 26, 4058 Basel, Switzerland; [b]The Babraham Institute, Babraham Research Campus, Cambridge CB22 3AT, United Kingdom; [c]Institute of Molecular Systems Biology, Swiss Federal Institute of Technology Zürich, Wolfgang-Pauli-Strasse 16, 8093 Zürich, Switzerland; [d]Department of Biochemistry, University of Oxford, South Parks Road, Oxford, OX1 3QU, United Kingdom; [e]Institute of Molecular Biotechnology of the Austrian Academy of Sciences, 1030 Vienna, Austria; and [f]Faculty of Sciences, University of Basel, 4056 Basel, Switzerland

Edited by Igor B. Dawid, National Institute of Child Health and Human, Bethesda, MD, and approved February 8, 2011 (received for review June 4, 2010)

The maintenance of specific gene expression patterns during cellular proliferation is crucial for the identity of every cell type and the development of tissues in multicellular organisms. Such a cellular memory function is conveyed by the complex interplay of the Polycomb and Trithorax groups of proteins (PcG/TrxG). These proteins exert their function at the level of chromatin by establishing and maintaining repressed (PcG) and active (TrxG) chromatin domains. Past studies indicated that a core PcG protein complex is potentially associated with cell type or even cell stage-specific sets of accessory proteins. In order to better understand the dynamic aspects underlying PcG composition and function we have established an inducible version of the biotinylation tagging approach to purify Polycomb and associated factors from *Drosophila* embryos. This system enabled fast and efficient isolation of Polycomb containing complexes under near physiological conditions, thereby preserving substoichiometric interactions. Novel interacting proteins were identified by highly sensitive mass spectrometric analysis. We found many TrxG related proteins, suggesting a previously unrecognized extent of molecular interaction of the two counteracting chromatin regulatory protein groups. Furthermore, our analysis revealed an association of PcG protein complexes with the cohesin complex and showed that Polycomb-dependent silencing of a transgenic reporter depends on cohesin function.

biotagging | BirA | pairing-sensitive silencing | proteomics

Polycomb group (PcG) proteins maintain repressed states of gene expression, while Trithorax group (TrxG) proteins counteract silencing and maintain active states in a heritable manner during development. These regulators play important roles in epigenetic processes such as stem cell maintenance, genomic imprinting, and X chromosome inactivation (reviewed in refs. 1 and 2). PcG/TrxG proteins confer their transcriptional function through cis-regulatory elements (PREs), which are composite sequences containing complex sets of conserved short motifs recognized by known DNA-binding proteins (3). PcG silencing depends primarily on the activities of Polycomb Repressive Complex 1 and 2 (PRC1 and PRC2). Additionally, PhoRC, containing as key factor the Pho protein, is involved in the recruitment of PRC1 and PRC2 to PREs (4). PcG silencing of target genes is likely to be cell lineage-specific; little is known about the tissue-specific assembly and distribution of PcG complexes and how they are recruited to PREs in particular cell lineages, however.

In this work, we established the in vivo biotinylation tagging approach in transgenic *Drosophila*. The system is based on the bacterial BirA biotin ligase, which recognizes and biotinylates a short 23aa peptide termed BioTag, allowing subsequent high stringency streptavidin affinity chromatography for purification of bait proteins (Fig. 1*A*). Biotin tagging has been used for purification of proteins and protein complexes in yeast, insect cells, and mice (5–9). Furthermore, the approach has been extended to BioChIP–Chip applications to characterize the genome-wide distribution of histones, transcription factors, and chromatin-binding proteins (10–13). As an extension of the methodology, we report here biotinylation in the fruit fly by combining biotinylation tagging with the versatile Gal4/UAS system for BirA expression (14). We use the system to identify partners of the PRC1 component Polycomb (Pc) by MS, revealing many previously unrecognized interactors. This strategy opens the perspective for cell type-specific protein purification experiments by using tissue-specific BirA driver lines.

Results

Establishment of in Vivo Inducible Biotinylation Tagging in Transgenic *Drosophila*. To produce a Pc-BioTag fusion protein we used a genomic Pc clone containing the Pc promoter and further regulatory sequences resulting in a transgene that mimics expression of the endogenous gene (15) (Fig. 1*A*). To complete the in vivo tagging system we cloned BirA cDNA into pUAS vectors for Gal4 inducible expression and generated transgenic *Drosophila* (Table S1).

The Pc-Bio fusion gene produces a functional Pc protein as judged by three independent tests: First, Pc-Bio had the same elution profile as Pc when nuclear extracts were assayed by gel filtration chromatography, indicating that the fusion protein was incorporated into corresponding Pc complexes (Fig. 4*E*). Second, we consistently observed down-regulation of the endogenous protein upon expression of the fusion construct, indicating that Pc-Bio feeds back on the expression of endogenous Pc (Fig. 1*B*). Third, we directly tested for protein functionality by introducing Pc-Bio into the genetic background of animals with transheterozygous combinations of the lethal alleles Pc^1, Pc^3, and Pc^{XL5}. Embryonic lethality was successfully rescued, indicating that the Pc-Bio fusion protein can substitute for its wild-type counterpart (Fig. S1).

Author contributions: G.S., L.R., C.B., and R.P. designed research; G.S., C.P., A.S., and A.P. performed research; G.S., A.S., A.P., C.B., and R.P. analyzed data; and G.S., C.B., and R.P. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

Freely available online through the PNAS open access option.

[1]Present address: Biozentrum, University of Basel, Klingelbergstrasse 50/70, 4057 Basel, Switzerland.

[2]Present address: Department of Molecular and Cellular Biology, Harvard University, Cambridge, MA 02138.

[3]To whom correspondence should be addressed. E-mail: renato.paro@bsse.ethz.ch.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1007916108/-/DCSupplemental.

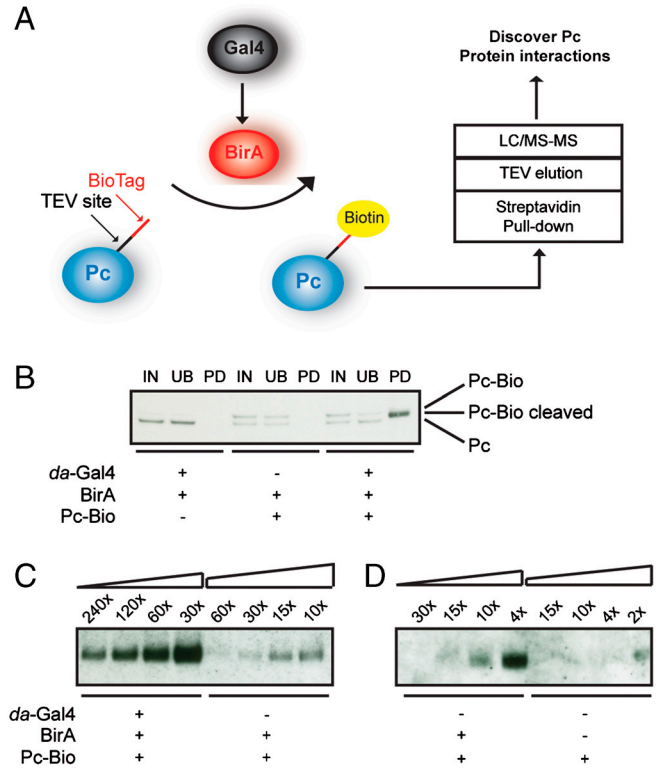

Fig. 1. Establishment of in vivo biotinylation tagging in *Drosophila*. (*A*) Pc is fused at its C terminus to a TEV-cleavable BioTag allowing proteolytic elution from streptavidin-coated beads. Gal4 induction of BirA biotin ligase expression leads to biotinylation of the BioTag. Biotinylated Pc-Bio is incorporated into the PRC1 complex and binds to chromatin. Streptavidin pull-down (PD) allows purification of Pc-Bio and detection of interaction partners by mass spectrometry. Experiments *B*, *C*, and *D* were performed with 7.5 mg of input nuclear extract. (*B*) Analysis of Pc-Bio PD efficiency comparing 30 μg of IN (input) and UB (unbound) with 1/5 of a PD eluate. The lower band corresponds to endogenous Pc, the upper band to the Pc-Bio fusion protein, and the middle band to Pc-Bio after TEV cleavage. (*C* and *D*) Quantification of background biotinylation after streptavidin PDs. Numbers indicate the fold-dilution of the eluate obtained from an IP of 7.5 mg nuclear extract that were loaded to the gel. (*C*) Comparing *da*-Gal4 induced versus uninduced strains reveals leaky biotinylation in the absence of the driver. (*D*) Comparing uninduced strains with or without UAS-BirA transgenes identifies the latter as the source of leaky biotinylation.

In order to evaluate biotinylation of Pc-Bio, we combined the ubiquitous *daughterless*-Gal4 driver (*da-Gal4*) with BirA and Pc-Bio. After pull-down of biotinylated proteins from embryonic extracts with streptavidin beads, we analyzed depletion of Pc-Bio from input (IN) nuclear extracts (NE) by comparing them to equal loads of unbound (UB) fractions on Western blots and estimated that approximately 65% of Pc-Bio becomes biotinylated after *da*-Gal4 induction of BirA and can be pulled down by streptavidin (Fig. 1*B*). However, prolonged exposure of Western blots revealed also weak signals after streptavidin pull-down from controls lacking Gal4 induction. In comparison to *da*-Gal4 induced lines, this signal was 44-fold lower (Fig. 1*C*). Because lines lacking UAS-BirA do not show any detectable biotinylation, this background biotinylation is caused by leaky expression of BirA from the UAS construct (Fig. 1*D*). In summary, Gal4 induction of BirA ligase leads to efficient biotinylation of Pc-Bio significantly above uninduced controls.

Purification of Pc-Bio Reveals Previously Undescribed Interaction Partners. We used the in vivo biotinylation system to identify previously undescribed interaction partners of Pc. The PRC1 complex has previously been purified employing a combination of ion exchange-, affinity-, and size-exclusion chromatography using Flag-tagged Ph as bait (16, 17). Here, we optimized our strategy toward identification of weakly associated interaction partners by avoiding multiple chromatography steps. A major obstacle for purification of biotinylated bait proteins is their separation from endogenously biotinylated proteins. We solved this problem by introducing a TEV protease cleavage site between the bait protein and the C-terminally located BioTag, allowing to specifically cleave off immobilized Pc-Bio from streptavidin beads leaving endogenously biotinylated proteins bound to the beads (Fig. S2). Parallel purification of control preparations lacking the BioTag (untagged) or missing *da*-Gal4 induction (uninduced) enabled us to identify BioTag-specific entities. Eluates were separated by electrophoresis and Western blots were performed with antibodies against all PRC1 core components confirming their specific copurification with the bait protein (Fig. 2*A*). Silver staining indicated the complexity of the final protein sample including the corresponding controls (Fig. 2*B*). Proteins were in-gel digested with trypsin, and peptides were analyzed by nanoelectrospray tandem mass spectrometry. We obtained a list of 454 copurifying proteins for our BioTag test sample, excluding keratin (Fig. 2*C*). To distinguish specific interacting proteins from unspecific contaminants, we first removed all proteins that copurified in the untagged control, except for Pc and dRing, due to strong enrichments over this control. Leaky BirA expression led to low-level enrichments of specific Pc-interaction partners in the uninduced control. Thus only proteins were kept in the hit list that were at least twofold enriched over the uninduced control. Of the remaining 92 proteins, exclusion of all proteins identified with less than four peptides total (see Fig. S3 for estimation of cutoff) led to a final hit list of 20 proteins, specifically enriched in the sample over controls (Fig. 2*C* and Fig. 3). As expected, PRC1 core components Psc, Ph-p, Ph-d, dRing, Su(z)2 and Pc rank highest in the list of identified proteins, demonstrating that our approach is capable of identifying true Pc-interacting proteins.

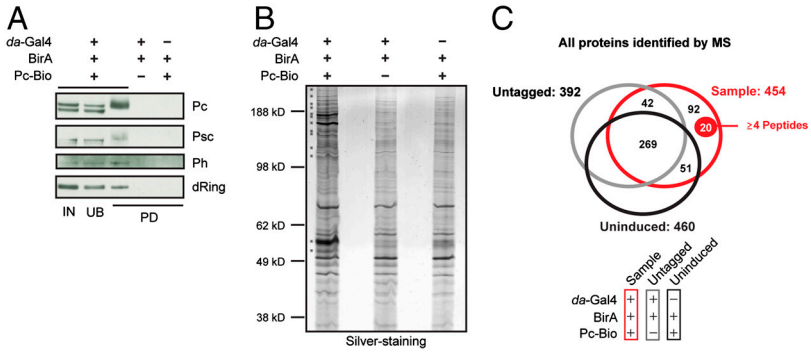

Fig. 2. Identification of Pc-interaction partners. (*A*) After streptavidin PD of 7.5 mg NE, TEV eluates of a Pc-Bio purification were separated by SDS/PAGE and Western blotting of PRC1 core proteins shows their specific enrichment over uninduced and untagged controls. Again, equivalent amounts of IN and UB were loaded (30 μg) together with 1/5 of the PD eluates. (*B*) Silver staining of a representative Pc-Bio purification. Controls missing Pc-Bio or Gal4 induction are shown in the right two lanes as controls. Asterisks indicate specific bands in the Pc-Bio test sample. A gel run in parallel was used to excise whole lanes for tryptic in-gel digestion and subsequent LC-MS/MS analysis. (*C*) Venn diagram depicting the overlap of proteins identified in the BioTag test sample (red), in the control without BioTag (gray), and in the control without the Gal4 driver (black). The final group of 20 specific interaction partners of Pc-Bio identified with ≥4 peptides total is highlighted.

	Name	Peptides (total)			Vertebrate homologue	Biochemical interactions	Classification by genetic interaction	Functions
		Sample	NoTag	NoDriver				
1	Psc	145		21	Bmi1, Mel-18	SPOP, Ph, Pc	PcG (Jurgens et al., 1985)	Contributes to ubiquitination by dRing
2	Ph-p	138		8	Hph1, Hph2	Pc, Psc	PcG (Dura et al., 1987)	Tumor suppressor
3	Sce/dRing	102	1	17	RNF2, RING1A/1B	Grh, BMI1	PcG (Fritsch et al., 2003)	Ubiquitinates H2A
4	Ph-d	100		8	Phc2	Pc, Psc	PcG (Dura et al., 1987)	Tumor suppressor
5	Su(z)2	81		2	Bmi1, Mel-18	SPOP, Ph, Pc	PcG (Soto et al., 2005)	Psc homologue
6	Pc	79	2	17	Hph1, Hph2	Psc, Ph	PcG (Lewis, 1978)	Binds H3K27me3
7	Smc1	20		6	Smc1a, Smc1b	Smc3, Rad21	-	Cohesion, transcriptional regulation
8	Smc3/Cap	18		5	Smc3	Smc1, Rad21	-	Cohesion & segregation timing
9	dSfmbt	10			Sfmbt1	Pho, Grh,	PcG (Klymenko et al., 2006)	Binds H3K9me2/3 and H4K20me1/2
10	Pp1-87B	7			PP1	Trx	PcG (Rudenko et al., 2003)	Colocalizes with Trx on polytenes
11	Ebi	6			Tbl1x	Ttk88	TrxG (This study)	Ubiquitination targeting factor
12	Enok	6			Moz/Morf	BRPF1-3	-	HAT domain
13	Grh	5		1	CP2, LBP-1a, LBP-9	Sce/dRing, Pho	PcG (Blastyak et al., 2006)	DNA-binding
14	Eaf6	5			Eaf6	BRPF1-3	-	-
15	Pho	5			YY1	dSfmbt, Grh	PcG (Girton et al., 1994)	DNA-binding
16	Fs(1)h	5			Brd2-4, Brdt, (Ring3)	-	TrxG (Shearn et al., 1989)	DNA-binding
17	CG1845	4			Brpf1-3	MOZ/MORF; Eaf6	TrxG (Laue et al., 2008)	Histone-binding, HAT cofactor
18	Rdx	4			Spop, Mel-26	BMI1 (Psc)	TrxG (This study)	Ubiquitination targeting factor
19	Rad21	4			Rad21	Smc1, Smc3	TrxG (Hallson et al., 2008)	Cohesion, transcriptional regulation
20	26-29kd prot.	4			-	-	-	-

PRC1 Core Complex

Cohesin Complex

PhoRc Complex

Moz/Morf like complex

Fig. 3. MS results for the identification of Pc-interaction partners. List of identified Pc-Bio interacting proteins ranked by the total number of peptides found by mass spectrometry. The six top scoring proteins are known members of the PRC1 core complex. Note that all core and many auxiliary proteins show peptides in the uninduced control sample due to leaky biotinylation by low-level UAS-BirA expression (Fig. 1*C*). Other proteins specifically enriched in the Pc-Bio pull-down are members of the cohesin complex (blue), PhoRC complex (green) and a Moz/Morf-like complex (yellow). Homologous proteins from vertebrates were retrieved using HomoloGene (http://www.ncbi.nlm.nih.gov/pubmed/). Known biochemical interactions among the recognized proteins identified by literature mining are indicated. Genetic interactions with PcG/TrxG mutations are listed including references. A protein/gene is classified as a PcG member if it enhances PcG and suppresses TrxG phenotypes. Conversely, a protein behaves as a TrxG protein if it genetically interacts in the opposite manner. The last column lists known molecular functions of the respective proteins.

One class of previously found interactors consists of DNA-binding proteins that have been linked to PcG silencing, like Pho and dSfmbt, which form the PhoRC complex (18). We verified the interaction of Pc with Pho by streptavidin pull-down, IP, and Gel filtration experiments (Fig. 4 *A*–*C* and *E*). Grainyhead (Grh), another DNA-binding protein, has previously been reported to associate with the dRing subunit of PRC1, enhancing Pho's recruitment to DNA (19). Additionally, we find other candidate proteins interacting that show DNA-binding activity such as the TrxG protein Female sterile homeotic (Fs(1)h) (Fig. 4*A*). Fs(1)h was previously reported to bind to Zeste motifs upstream of the *Ubx* gene supporting *Ubx* expression (20). Another protein with a putative DNA-binding domain is the histone acetyltransferase (HAT) Enoki mushroom (Enok) of the MYST family. Its unique HAT domain is conserved in the vertebrate proteins Moz and Morf, where it forms complexes with Eaf6 and Brpf1-3 (21, 22). The fact that we also detect the *Drosophila* homologues Eaf6 and CG1845 (23) suggests that formation of the Moz/Morf HAT complex identified in mammalian cells may also be conserved in *Drosophila* and potentially linked to PcG silencing.

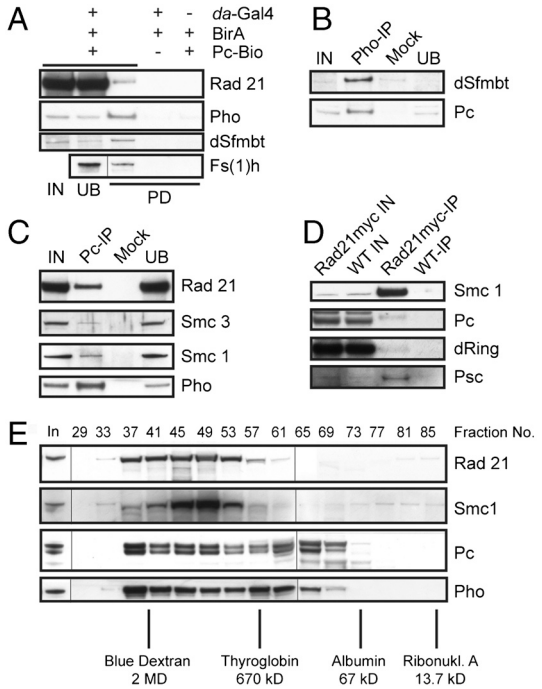

Fig. 4. Verification of Pc-interaction partners by Co-IP. Equal amounts of IN and UB (30 μg) were loaded to gels together with 1/6 of the PD eluate. (*A*) Western blot analysis after streptavidin PD of 7.5 mg of NE using antibodies against Rad21, Pho, dSfmbt, and Fs(1)h. In the following Co-IPs, 1 mg of NE was used as IN and 50% of the PD eluates were loaded to the gels. (*B*) Wild-type nuclear extracts were immunoprecipitated with Pho antibody. Immunoblotting revealed copurification of dSfmbt and Pc. (*C*) Wild-type nuclear extracts were immunoprecipitated with Pc antibody. Immunoblotting revealed copurification of Pho as well as the three core cohesin complex members Rad21, Smc3, and Smc1. (*D*) Coimmunoprecipitation experiments using nuclear extract from flies expressing myc-tagged Rad21, which was captured on myc-sepharose. Immunoblotting revealed copurification of Smc1, Pc, dRING, and Psc. (*E*) Gel filtration chromatography on a Superose 6 column for size estimation of the fly cohesin complex and comparison to the elution profile of PcG proteins. The Pc profile shows multiple peaks of which the middle peak coelutes with cohesin at about 1–1.5 MD.

Pc Interacts with the Cohesin Complex. The strongest Pc-interactors besides PRC1 core proteins were obtained for the structural maintenance of chromatin proteins Smc1 and Smc3 (Fig. 3). Smc1 and Smc3 are best known for their role in sister chromatid cohesion, where they associate with a third protein called Rad21/SCC1 to form a tripartite, ring-shaped protein complex called cohesin (reviewed in ref. 26). The fact that we also identified Rad21/SCC1 as an interaction partner of Pc further substantiated the previously unrecognized physical interaction between PRC1 and cohesin. Besides its role in sister chromatid cohesion there is accumulating evidence from multiple organisms that cohesin and associated factors have diverse roles in gene regulation, organization of chromatin structure and development (27, 28). One example is the genetic interaction of Rad21 with TrxG mutations in *Drosophila* (29). In addition, we show that the *Rad21ex3* and

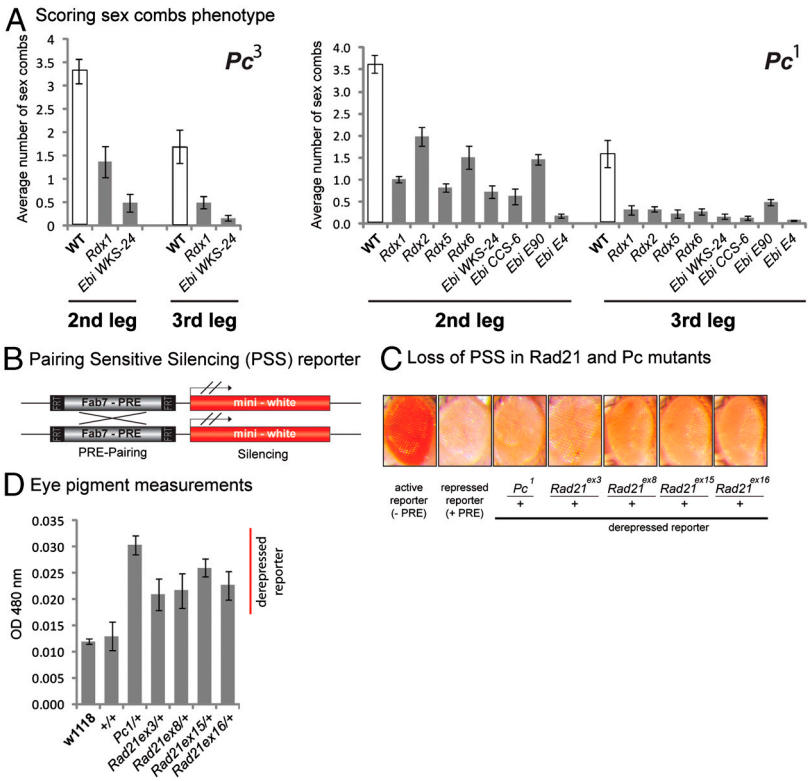

Fig. 5. Genetic interaction and suppression of pairing-sensitive silencing (PSS). (*A*) Genetic interactions of several *roadkill* and *ebi* mutants with the Pc^1 and Pc^3 alleles revealed by suppression of the extra sex combs Pc phenotype identifies these genes as TrxG members. (*B*) Illustration of the transgenic reporter used to analyze PSS. (*C* and *D*) Heterozygous Rad21 and Pc mutations suppress PSS of the transgenic miniwhite reporter containing a Fab7-PRE. In the absence of the PRE, the homozygous transgene is fully expressed (dark red eye). PSS is observed for the homozygous reporter in a PRE-dependent manner (white eye). (*E*) Quantitative spectrometric measurements of eye pigment levels confirm derepression of the reporter.

vtd4 alleles also genetically interact with Polycomb by suppression of the Pc4 extra sex combs phenotype. In case of the Rad21ex3 allele, this effect is clearly specific to the truncated Rad21 gene as expression of full length myc-tagged Rad21 rescues the extra sex combs phenotype (Fig. S4). In order to further validate the biochemical interaction of Pc with cohesin subunits, we performed a series of immunoprecipitation experiments using embryonic extracts, confirming interaction of all cohesin components with the endogenous Pc protein in wild-type extracts (Fig. 4 *A*, *C*, and *D*). We also generated samples from animals expressing myc-tagged Rad21 in a Rad21 mutant background (27) and confirmed interaction of Rad21 with Pc. To further explore this biochemical association, we separated embryonic nuclear extract on a Superose 6 gel filtration column and followed the elution of Rad21, Smc1, and Pc (Fig. 4*E*). As expected, Rad21 and Smc1 elute together as one large complex corresponding to a mass of about 1–1.5 MD. Consistent with a previous study (30), Pc fractionates in at least three peaks. Most interestingly for our work, the middle peak was found to coelute with the cohesin complex. Taken together with the IP results, these findings support a physical interaction between a specific subgroup of PcG complexes and the cohesin complex.

The genome-wide binding profiles of cohesin in *Drosophila*, humans, and mouse cells reveal a bias for binding to 5′-untranslated regions of active genes as well as exclusion from H3K27me3 domains. Due to the fact that Pc-binding largely coincides with H3K27me3, coimmunostainings of polytene chromosomes for Pc and Rad21 are in agreement with these previous studies showing that these two proteins do not colocalize (Fig. S5). This suggests that the reported biochemical interaction of cohesins and Pc either does not take place on chromatin of salivary gland cells or might be restricted to early embryonic stages. Furthermore, removal of Rad21 from polytene chromosomes does not alter Pc distribution, rendering a role in recruiting Pc to chromatin unlikely in this tissue (Fig. S5).

To further test a potential functional link we checked the role of cohesin in pairing sensitive silencing, a phenomenon characteristic of PcG mediated repression in *Drosophila*. Insertion of a PRE next to a miniwhite reporter gene induces partial silencing of the transgene in heterozygous flies and results in reduced red eye color in adult flies (31). Homozygous flies show a much stronger silencing of the reporter gene. This phenomenon depends on pairing of homologous PREs on two somatic chromosomes and is referred to as pairing-sensitive silencing (PSS). Mutations in PcG genes lead to loss of PSS. The degree of PSS and its dependency on PcG proteins varies considerably with the location of the transgene insertion site in the genome. Pairing of homologous chromosomes is a fundamental function of nuclear organization especially in dipteran insects. In other species, it has often been observed during meiosis and sometimes affecting mitotic chromatin (32). Interestingly, cohesin is known to be essential for pairing of meiotic sister chromosomes, which is furthermore required for pairing of homologous chromosomes thus connecting these two processes. Mutations in cohesin proteins lead to loss of pairing and aberrant meiotic chromosome structure (33).

To address whether cohesin might contribute to pairing of homologous interphase chromosomes and might therefore play a role in PSS, we employed a previously established transgenic combination of the Fab-7 PRE with a miniwhite reporter as a readout system for PcG-dependent PSS (Fig. 5*B* and Fig. S6*D*) (34). These reporter lines are uniquely suited for this test, because the PRE is flanked by FRT sequences allowing removal of the PRE sequence by Flp recombination, thereby excluding position-specific effects. As expected, silencing is not observed in heterozygous and homozygous animals if the Fab-7 PRE is absent (Fig. S6 *A*–*C*). Moreover, Pc and Rad21 mutations do

not change reporter gene expression in the absence of the Fab-7 PRE, excluding direct effects of these alleles on the transgene (Fig. S6*B*). In contrast, homozygous animals with PREs show strong PSS, leading to almost complete silencing of the reporter. Importantly, silencing was partially relieved by crossing the reporter to different *Rad21* alleles (Fig. 5 *C* and *D*), demonstrating that Rad21, an essential subunit of the cohesin complex, contributes together with PRC1 to PSS.

Discussion

We report the identification of previously undescribed Pc-interacting proteins using a newly established protein tagging system, which is inducible in animals. Combinations of one-step capture with streptavidin, low stringency washes, specific elution, and detection of peptides using a highly sensitive LTQ-FT-ICR mass spectrometer enabled the identification of even labile and transient interactions. It has been well recognized that PcG and TrxG proteins exert their counteracting activities at the level of chromatin by employing various biochemical activities directed against histones, like methylation, acetylation, and chromatin remodeling (4). Indeed, our study reveals a substantial number of Pc-interacting proteins implicated in TrxG action. The genes encoding for *Rdx, Ebi, CG1845, Rad21*, and *Fs(1)h* have been shown genetically to belong to the TrxG suppressing PcG mutant phenotypes and activating HOX gene expression, for example. Additionally, Pp1-87B has been found to interact with Trx or its homologue MLL (35, 36). These data indicate that Pc and specific members of the TrxG may physically cooperate to maintain the on/off state of genes.

So far, the DNA-binding proteins Zeste, Gaf, Pho, Dsp1, Sp1/Klf family members, Psq, and Grh have been connected to PcG function on the basis of genetic interactions, biochemical copurification, functional assays, and/or colocalization on PREs. Here we find direct biochemical interactions of Grh and Pho with Pc. Moreover, we identified a Pc-interacting protein called Fs(1)h that might, as well, contribute to recruitment of PRCs to chromatin. *Fs(1)h* interacts strongly with *Ubx, trx*, and *ash1* mutations and leads to homeotic phenotypes when overexpressed (20). Fs(1)h is essential for development and conserved in mammals. Whether Pc is recruited by Fs(1)h or opposes its function in gene activation needs to be established. Beside the aforementioned DNA-binding proteins, Enok is a Pc interactor with a putative DNA-binding domain. Enok forms part of the MYST domain family of histone acetyl transferases (HATs), and mutants with defects in the HAT domain show retarded development and pupal lethality (37). Enok's HAT domain is conserved in the vertebrate Moz/Morf proteins. They typically form complexes comprising one protein per BRPF-, ING-, and EAF family member. In *Drosophila*, a Moz/Morf like complex may consist of Enok, CG1845 (homologue of Brpf1-3), and Eaf6 as all these proteins copurified with Pc and were detected with high confidence (Fig. 3). Moz and Brpf1 are TrxG proteins required for HOX gene expression in vertebrates. Although MYST-domain-containing HATs have generally been associated with transcriptional activation, there are also examples with a link to HOX gene repression in *Drosophila*.

Our work uncovered a connection between Pc and the cohesin complex. Cohesin has been described in detail for its roles in mitosis and meiosis (26), embracing sister chromatids in mitotic cells. Interestingly, mutations in Ph-p, Psc, and Pc have been reported to result in chromosome missegregation phenotypes in embryos (38). Besides its traditional role in sister chromatid cohesion, cohesin has also been implicated in both activation and repression of transcription (28). Furthermore, mutations in the Rad21 subunit of the cohesin complex strongly enhance TrxG (29) and suppress PcG loss of function phenotypes. Pc and cohesins are not colocalized on salivary gland chromatin, and removal of cohesin does not affect Pc binding. We cannot rule out that Pc is needed for recruitment of cohesin, however. For example, chromatin binding of cohesin in *S. pombe* depends on formation of heterochromatin, requiring another chromo domain protein, HP1.

A hallmark of PcG repression in flies is PSS, depending on pairing of homologous chromosomes in interphase chromatin (31). It is known that multiple copies of a transgenic PRE interact with each other if inserted on the same or even on different chromosomes. Because cohesin plays a role in pairing of homologous chromosomes in meiosis and has been suggested to facilitate long-range DNA interactions, it may also facilitate PRE pairing. The transgenic reporter for PSS used in this study only showed PRE-dependent silencing upon PRE pairing. The observation that cohesin mutant alleles reduce PSS supports a model in which cohesins contribute to PRE pairing in interphase chromatin.

The identification of Pc-interacting proteins was made possible by employing the in vivo biotinylation system combined with highly sensitive mass spectrometric analysis, thereby preserving near physiological conditions for protein purification. The identification of substoichiometric levels of interacting proteins shows that in vivo biotinylation was effective in capturing even weakly or underrepresented associated proteins. Inducible biotinylation tagging is currently limited to the use of Gal4 drivers that trigger biotinylation well above the background levels. Generation of libraries of different UAS-BirA transgenic lines with less leaky expression and flies carrying BirA under direct control of tissue-specific promoters will further improve and expand this tool, making it a versatile system for proteomic and genomic studies in specialized cell types. As a major advantage over tissue-specific expression of tagged bait proteins, biotin tagging allows to express the bait protein under control of endogenous promoter sequences, whereas the induction of the BirA ligase can be independently induced via the Gal4/UAS system avoiding bait protein misexpression artifacts. This work opens the perspective for tissue-specific applications, potentially enabling a systems analysis on how protein networks can control subsets of genes in specialized cells.

Experimental Procedures.

Cloning of Pc-Bio and UAS-BirA. A linker encoding the TEV protease cleavage site, the BioTag, and Flag-tag was inserted in front of the Pc open reading frame in the genomic Pc-Pst1 fragment (15). The genomic clone (3712 nucleotides) was PCR amplified and subcloned in two parts into pCRII-TOPO (Invitrogen) to generate an artificial SmaI-site in front of the stop codon for subsequent insertion of the Tag-linker. PCR oligo fragment PstI1-2225SmaI:Pc-BioI + Pc-BioCtermII; PCR oligo fragment SmaI2226-3712PstI:Pc-BioCtermIII + Pc-BioIV. Both fragments were cloned together in pCRII-TOPO via BamHI/SmaI to generate the full-length clone (pCRII-TOPO-PcCtermSmaI). Two Oligos encoding the epitope tag TevBioFlag were phosphorylated, annealed, and inserted into the SmaI cut pCRII-TOPO-PcCtermSmaI. The linker containing Pc fragment was cloned via PstI into the *Drosophila* transformation vector pCaSpeR4. All constructs were verified by sequencing. BirA ligase cDNA pBS-BirA was obtained from J. Strouboulis (Erasmus Medical Center). BirA was cloned via BglII into the *Drosophila* transformation vector pUASp. In order to express Flag-tagged BirA in *Drosophila*, we PCR amplified pBS-BirA with oligos NotI-Flag-BirA and BirA-Stop-Xba1 into the transformation vectors pUASp and pUASty. BirA lines are listed in Table S1, and fly strains used in this study can be found in Table S2.

PRC1 Purification. For PRC1 Purification, a large scale collection of 0–12 h embryos and preparation of nuclear extracts was performed essentially as previously described (39), with the following modifications: Isolated nuclei were resuspended in 1 mL 15 mM HEPES pH 7.6, 110 mM NaCl, 5 mM $MgCl_2$, 0.1 mM EDTA, 1 mM DTT and protease inhibitors per g embryos. One tenth

of 4M $(NH_2)_4SO_4$ pH 7.9 was added and nuclear proteins were extracted for 20 min on a roller. Clear supernatants without the top lipid layer were precipitated with 0.3 g/mL $(NH_2)_4SO_4$ and centrifuged at 15 k in a SS34 rotor. Pellets were dissolved in Dialysis buffer (DB, 30 mM Hepes pH7.6, 150 mM NaCl, 10 mM $MgCl_2$, 0.2 mM EDTA, 1 mM DTT, 10% Glycerol, protease inhibitors (Complete, EDTA free Roche) and dialyzed twice for 2 h against the same buffer. 7.5 mg of nuclear extracts (NE) were incubated overnight with 120 μl of Streptavidin Sepharose 50% slurry (GE) supplemented with 0.1% NP-40. Bound material was washed 2x with DB-1 (0.1% NP-40), 2x with DB-2 (0.05% NP-40, 350 mM NaCl) and 2x with DB-3 (0.05% NP-40). Bound material was eluted from beads in 100 μl DB-3 supplemented with 1.5 μl TEV Protease (Invitrogen) at 16 °C for 3 h on a roller. Eluates were acetone precipitated and pellets were dissolved in 1x LDS Sample Buffer (Invitrogen) prior to electrophoresis. For MS analysis, 7 IPs were combined and eluates were separated on a 4–12% Bis-Tris Gradient Gel (Invitrogen). Proteins were in-gel digested, peptides were extracted and cleaned on reverse phase C18 chromatography columns (Waters) prior to MS Analysis. Eluates were analyzed using a μRPLC-MS system (*SI Text* section S2). A detailed list of peptides identified can be found in Table S3. Analytical gel filtration experiments were performed on a Superose 6 HR column (23.56 mL, 1 cm × 30 cm, GE Healthcare) as specified in *SI Text* section S3.

ACKNOWLEDGMENTS. We thank M. Seimiya for valuable technical support and J. Fonseca for performing imaging experiments of Pc bodies in Rad21 depleted cells. We are grateful to R. Jones, J. Müller, C. Lehner, C. Sunkel, S. Bickel, T. Kornberg, J. Hooper, L. Tsuda, and I. Dawid for generous gifts of antibodies and J. Stroubulis for BirA containing plasmids. We greatly appreciate comments on the manuscript from A. Hochheimer, A. Monqaut, and P. Talbert. We thank the members of the SystemsX.ch RTD WingX for helpful discussions and valuable advice. G.S. was supported by a Böhringer Ingelheim Fonds fellowship. Research by A.P. was conducted in the lab of Kim Nasmyth, who is acknowledged for generous support and funding. A.P. acknowledges the EMBO Long-Term Fellowship program and Human Frontier Science Program for current funding. R.P. was supported by funds from the SFB/Transregio 5 of the Deutsche Forschungsgemeinschaft and the Swiss Federal Institute of Technology Zürich.

1. Simon JA, Kingston RE (2009) Mechanisms of polycomb gene silencing: Knowns and unknowns. *Nat Rev Mol Cell Biol* 10:697–708.
2. Surface LE, Thornton SR, Boyer LA (2010) Polycomb group proteins set the stage for early lineage commitment. *Cell Stem Cell* 7:288–298.
3. Ringrose L, Paro R (2007) Polycomb/Trithorax response elements and epigenetic memory of cell identity. *Development* 134:223–232.
4. Muller J, Verrijzer P (2009) Biochemical mechanisms of gene regulation by polycomb group protein complexes. *Curr Opin Genet Dev* 19:150–158.
5. de Boer E, et al. (2003) Efficient biotinylation and single-step purification of tagged transcription factors in mammalian cells and transgenic mice. *Proc Natl Acad Sci USA* 100:7480–7485.
6. Rodriguez P, et al. (2006) Isolation of transcription factor complexes by in vivo biotinylation tagging and direct binding to streptavidin beads (Translated from eng). *Methods Mol Biol* 338:305–323 (in eng).
7. Tagwerker C, et al. (2006) A tandem affinity tag for two-step purification under fully denaturing conditions: Application in ubiquitin profiling and protein complex identification combined with in vivocross-linking. *Mol Cell Proteomics* 5:737–748.
8. Furuyama T, Henikoff S (2006) Biotin-tag affinity purification of a centromeric nucleosome assembly complex. *Cell Cycle* 5:1269–1274.
9. Wang J, et al. (2006) A protein interaction network for pluripotency of embryonic stem cells. *Nature* 444:364–368.
10. Lee TI, et al. (2006) Control of developmental regulators by Polycomb in human embryonic stem cells. *Cell* 125:301–313.
11. Boyer LA, et al. (2006) Polycomb complexes repress developmental regulators in murine embryonic stem cells. *Nature* 441:349–353.
12. Kim J, Chu J, Shen X, Wang J, Orkin SH (2008) An extended transcriptional network for pluripotency of embryonic stem cells. *Cell* 132:1049–1061.
13. Mito Y, Henikoff JG, Henikoff S (2005) Genome-scale profiling of histone H3.3 replacement patterns. *Nat Genet* 37:1090–1097.
14. Duffy JB (2002) GAL4 system in *Drosophila*: A fly geneticist's Swiss army knife. *Genesis* 34:1–15.
15. Paro R, Hogness DS (1991) The Polycomb protein shares a homologous domain with a heterochromatin-associated protein of *Drosophila*. *Proc Natl Acad Sci USA* 88:263–267.
16. Shao Z, et al. (1999) Stabilization of chromatin structure by PRC1, a Polycomb complex. *Cell* 98:37–46.
17. Saurin AJ, Shao Z, Erdjument Bromage H, Tempst P, Kingston RE (2001) A *Drosophila* Polycomb group complex includes Zeste and dTAFII proteins. *Nature* 412:655–660.
18. Klymenko T, et al. (2006) A Polycomb group protein complex with sequence-specific DNA-binding and selective methyl-lysine-binding activities. *Genes Dev* 20:1110–1122.
19. Tuckfield A, et al. (2002) Binding of the RING polycomb proteins to specific target genes in complex with the grainyhead-like family of developmental transcription factors. *Mol Cell Biol* 22:1936–1946.
20. Chang YL, King B, Lin SC, Kennison JA, Huang DH (2007) A double-bromodomain protein, FSH-S, activates the homeotic gene ultrabithorax through a critical promoter-proximal region. *Mol Cell Biol* 27:5486–5498.
21. Yang XJ, Ullah M (2007) MOZ and MORF, two large MYSTic HATs in normal and cancer stem cells. *Oncogene* 26:5408–5419.
22. Laue K, et al. (2008) The multidomain protein Brpf1 binds histones and is required for Hox gene expression and segmental identity. *Development* 135:1935–1946.
23. Ullah M, et al. (2008) Molecular architecture of quartet MOZ/MORF histone acetyltransferase complexes. *Mol Cell Biol* 28:6828–6843.
24. Pintard L, Willems A, Peter M (2004) Cullin-based ubiquitin ligases: Cul3-BTB complexes join the family. *EMBO J* 23:1681–1687.
25. Boulton SJ, Brook A, Staehling Hampton K, Heitzler P, Dyson N (2000) A role for Ebi in neuronal cell cycle control. *EMBO J* 19:5376–5386.
26. Nasmyth K, Haering CH (2009) Cohesin: Its roles and mechanisms. *Annu Rev Genet* 43:525–58.
27. Pauli A, et al. (2008) Cell-type-specific TEV protease cleavage reveals cohesin functions in *Drosophila* neurons. *Dev Cell* 14:239–251.
28. Dorsett D (2009) Cohesin, gene expression and development: Lessons from *Drosophila*. *Chromosome Res* 17:185–200.
29. Hallson G, et al. (2008) The *Drosophila* cohesin subunit Rad21 is a trithorax group (trxG) protein. *Proc Natl Acad Sci USA* 105:12405–12410.
30. Lagarou A, et al. (2008) dKDM2 couples histone H2A ubiquitylation to histone H3 demethylation during Polycomb group silencing. *Genes Dev* 22:2799–2810.
31. Kassis JA (2002) Pairing-sensitive silencing, polycomb group response elements, and transposon homing in *Drosophila*. *Adv Genet* 46:421–438.
32. Lewis JP, Tanke HJ, Raap AK, Beverstock GC, Kluin-Nelemans HC (1993) Somatic pairing of centromeres and short arms of chromosome 15 in the hematopoietic and lymphoid system. *Hum Genet* 92:577–582.
33. Revenkova E, Jessberger R (2005) Keeping sister chromatids together: cohesins in meiosis. *Reproduction* 130:783–790.
34. Schmitt S, Prestel M, Paro R (2005) Intergenic transcription through a polycomb group response element counteracts silencing. *Genes Dev* 19:697–708.
35. Rudenko A, Bennett D, Alphey L (2003) Trithorax interacts with type 1 serine/threonine protein phosphatase in *Drosophila*. *EMBO Rep* 4:59–63.
36. Yokoyama A, et al. (2004) Leukemia proto-oncoprotein MLL forms a SET1-like histone methyltransferase complex with menin to regulate Hox gene expression. *Mol Cell Biol* 24:5639–5649.
37. Scott EK, Lee T, Luo L (2001) enok encodes a *Drosophila* putative histone acetyltransferase required for mushroom body neuroblast proliferation. *Curr Biol* 11:99–104.
38. O'Dor E, Beck SA, Brock HW (2006) Polycomb group mutants exhibit mitotic defects in syncytial cell cycles of *Drosophila* embryos. *Dev Biol* 290:312–322.
39. Kamakaka RT, Tyree CM, Kadonaga JT (1991) Accurate and efficient RNA polymerase II transcription with a soluble nuclear fraction derived from *Drosophila* embryos. *Proc Natl Acad Sci USA* 88:1024–1028.