

# Structure of the Polycomb Group Protein PCGF1 in Complex with BCOR Reveals Basis for Binding Selectivity of PCGF Homologs

Sarah E. Junco,<sup>1</sup> Renjing Wang,<sup>1</sup> John C. Gaipa,<sup>1</sup> Alexander B. Taylor,<sup>1</sup> Virgil Schirf,<sup>1</sup> Micah D. Gearhart,<sup>2</sup> Vivian J. Bardwell,<sup>2</sup> Borries Demeler,<sup>1</sup> P. John Hart,<sup>1,3</sup> and Chongwoo A. Kim<sup>1,\*</sup>

<sup>1</sup>Department of Biochemistry and CTCRC, University of Texas Health Science Center at San Antonio, MSC 7760, 7703 Floyd Curl Drive, San Antonio, TX 78229-3990, USA

<sup>2</sup>Department of Genetics, Cell Biology and Development, Masonic Cancer Center and Developmental Biology Center, University of Minnesota, Minneapolis, MN 55455, USA

<sup>3</sup>Geriatric Research, Education, and Clinical Center, Department of Veterans Affairs, South Texas Veterans Health Care System, San Antonio, TX 78229, USA

\*Correspondence: [chong@biochem.uthscsa.edu](mailto:chong@biochem.uthscsa.edu)  
<http://dx.doi.org/10.1016/j.str.2013.02.013>

## SUMMARY

Polycomb-group RING finger homologs (PCGF1, PCGF2, PCGF3, PCGF4, PCGF5, and PCGF6) are critical components in the assembly of distinct Polycomb repression complex 1 (PRC1)-related complexes. Here, we identify a protein interaction domain in BCL6 corepressor, BCOR, which binds the RING finger- and WD40-associated ubiquitin-like (RAWUL) domain of PCGF1 (NSPC1) and PCGF3 but not of PCGF2 (MEL18) or PCGF4 (BMI1). Because of the selective binding, we have named this domain PCGF Ub-like fold discriminator (PUFD). The structure of BCOR PUFD bound to PCGF1 reveals that (1) PUFD binds to the same surfaces as observed for a different Polycomb group RAWUL domain and (2) the ability of PUFD to discriminate among RAWULs stems from the identity of specific residues within these interaction surfaces. These data show the molecular basis for determining the binding preference for a PCGF homolog, which ultimately helps determine the identity of the larger PRC1-like assembly.

## INTRODUCTION

Polycomb group (PcG) proteins form repressive complexes that mediate epigenetic modifications of histones (Simon and Kingston, 2009). The *Drosophila* PcG protein PSC is a key player in multiple PcG complexes with different functions. As a member of *Drosophila* Polycomb repression complex 1 (PRC1) (Figure 1A), PSC helps to create a compacted chromatin state that is inaccessible to chromatin remodeling enzymes (Shao et al., 1999), a function that PSC can carry out on its own (Francis et al., 2001). In a related complex, the N-terminal RING finger heterodimer of PSC and dRING1 act as a ubiquitin ligase (E3) to modify histone H2A (Wang et al., 2004). The dRAF complex is yet another PSC-containing complex that includes dKDM2,

whose demethylation of dimethylated histone 3 Lys 36 (H3K36me2) is coupled to H2A E3 activity (Lagarou et al., 2008). PSC is also part of a cell cycle regulatory complex (Mohd-Sarip et al., 2012). In humans, the expansion to six different PSC orthologs (PCGF1/NSPC1, PCGF2/MEL18, PCGF3, PCGF4/BMI1, PCGF5, and PCGF6/MBLR) has resulted in the functional diversification of the complexes they assemble into (Figure 1B). While PCGF2 and PCGF4 are part of complexes most similar to *Drosophila* PRC1 (Gao et al., 2012; Levine et al., 2002), PCGF1 assembles into a dRAF-like complex that includes KDM2B, a H3K36me2 demethylase, and BCOR (Gao et al., 2012; Gearhart et al., 2006; Sánchez et al., 2007).

A closer look at the individual protein-protein interactions within these PcG complexes provides clues indicating how distinct larger assemblies could emerge (Figure 1B). In particular, a ubiquitin-like fold domain involved in protein interactions called ring finger- and WD40-associated ubiquitin-like (RAWUL) (Sanchez-Pulido et al., 2008) appears to be key in determining the identity of the larger assembly. A recent study has shown that the binding partners to the RING1B RAWUL and the identity of the PCGF homolog define the larger assemblies (Gao et al., 2012). The contribution made by the RING1B RAWUL (also called C-RING1B) toward forming a distinct larger assembly likely stems from its ability to utilize a mutually exclusive binding surface to associate with short peptides within different proteins that can be quite diverse in sequence (Wang et al., 2010). The RAWUL domains of the PCGF proteins appear to play a similar role, given that the RING finger domain of the PCGF homologs plays the singular role of heterodimerizing with RING1B to facilitate the H2A E3 activity (Buchwald et al., 2006; Li et al., 2006; Wang et al., 2004), thus leaving the RAWUL domain to play the part of helping to define the larger assembly. The ability of the PCGF RAWUL to bind different proteins (Alkema et al., 1997; Gunster et al., 1997; Kyba and Brock, 1998; Lee et al., 2007) supports this notion.

The direct binding partner of PCGF1 is BCOR (Gearhart et al., 2006), which has emerged as an important player in development and health. Mutation of *BCOR* results in oculofaciocardio-dental syndrome (Ng et al., 2004). *Bcor* is required for appropriate regulation of gene expression during differentiation mouse embryonic stem cells (Wamstad et al., 2008). BCOR is

involved in mediating transcriptional repression by the oncoprotein BCL6 and thus likely plays an active role in the development of BCL6-driven lymphomas (Gearhart et al., 2006; Huynh et al., 2000). In two other cancers, *BCOR* is involved in translocations; a fusion of *BCOR* with retinoic acid receptor alpha has been observed in an acute promyelocytic leukemia (Yamamoto et al., 2010), while a *BCOR-CCNB3* gene fusion has been found in Ewing sarcomas that are distinct from those caused by the more familiar *EWS-ETS* fusions (Pierion et al., 2012). In contrast, emerging data suggest that, in other cancers, *BCOR* plays a tumor suppressive role. Mutation of *BCOR* has been found in brain tumors (Pugh et al., 2012), retinoblastomas (Zhang et al., 2012), and acute myeloid leukemia (AML) (Grossmann et al., 2011). The closely related homolog *BCOR-like 1* (*BCORL1*) has also been implicated in AML (Li et al., 2011).

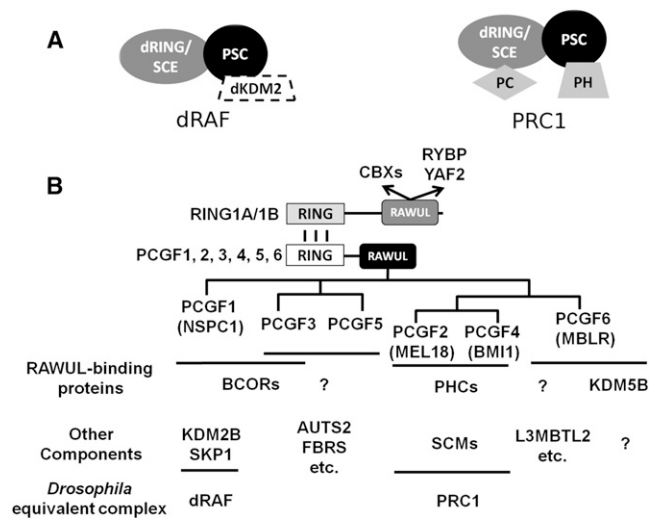
The assembly of the different PcG complexes is dependent on the binding preferences of the PCGF homologs. More specifically, the binding of PCGF1 to BCOR will lead to assembly of the dRAF-like complex, while PCGF2 and PCGF4 binding to the PcG protein called polyhomeotic (PHC1, PHC2, and PHC3) results in formation of the mammalian PRC1 equivalents (Figure 1). To better understand the molecular determinants underlying the formation of different human PcG complexes, we investigated the binding selectivity and the structural basis for how PCGF1 RAWUL binds directly to a previously unidentified protein interaction domain at the C terminus of BCOR.

## RESULTS

### PUFD Is a Protein Interaction Domain Selective for PCGF1 and PCGF3

We first mapped the minimum essential binding regions within PCGF1 and BCOR (Figures 2A and 2B). Using the bacterial two-hybrid assay, in which the interaction between two target proteins results in expression of the  $\beta$ -lactamase reporter gene (Dove and Hochschild, 2004), the PCGF1 RAWUL (Figure 2C) was found to be sufficient to obtain a positive signal of bacterial growth in ampicillin media when coexpressed with BCOR 1451–1755 (Figure 2B, top left panel). BCOR 1451–1755 is the previously identified region required for binding PCGF1 (Gearhart et al., 2006). Further dissection of this interaction revealed that BCOR 1634–1748 is both necessary and sufficient for interaction with the PCGF1 RAWUL. The 115 BCOR residues were a surprisingly large region to be bound by a RAWUL domain as the one in RING1B (also called C-RING1B; Wang et al., 2010), PCGF2 and PCGF4 (see below) form stable complexes with approximately 30 residue peptides of their respective binding partners. A BLAST search of BCOR 1634–1748 finds matches only in the related human protein called BCORL1 and in BCOR orthologs in other species (Figure 2D). We have named this protein interaction domain PCGF Ub-like fold discriminator (PUFD).

We were unable to isolate the PCGF1 RAWUL alone, precluding a more thorough investigation of its interaction with the PUFD or any other potential binding partners. We were, however, able to isolate the PCGF1/PUFD complexes through coexpression of both components in bacteria and perform sedimentation velocity experiments. At 8.5  $\mu$ M, the lowest possible concentration allowable for 280 nm absorption detection in our



**Figure 1. PRC1-like Assemblies**

(A) Two PSC-containing *Drosophila* PcG complexes are shown. The dashed line indicates that the protein-protein interactions required to assemble dKDM2 into the complex are unknown.

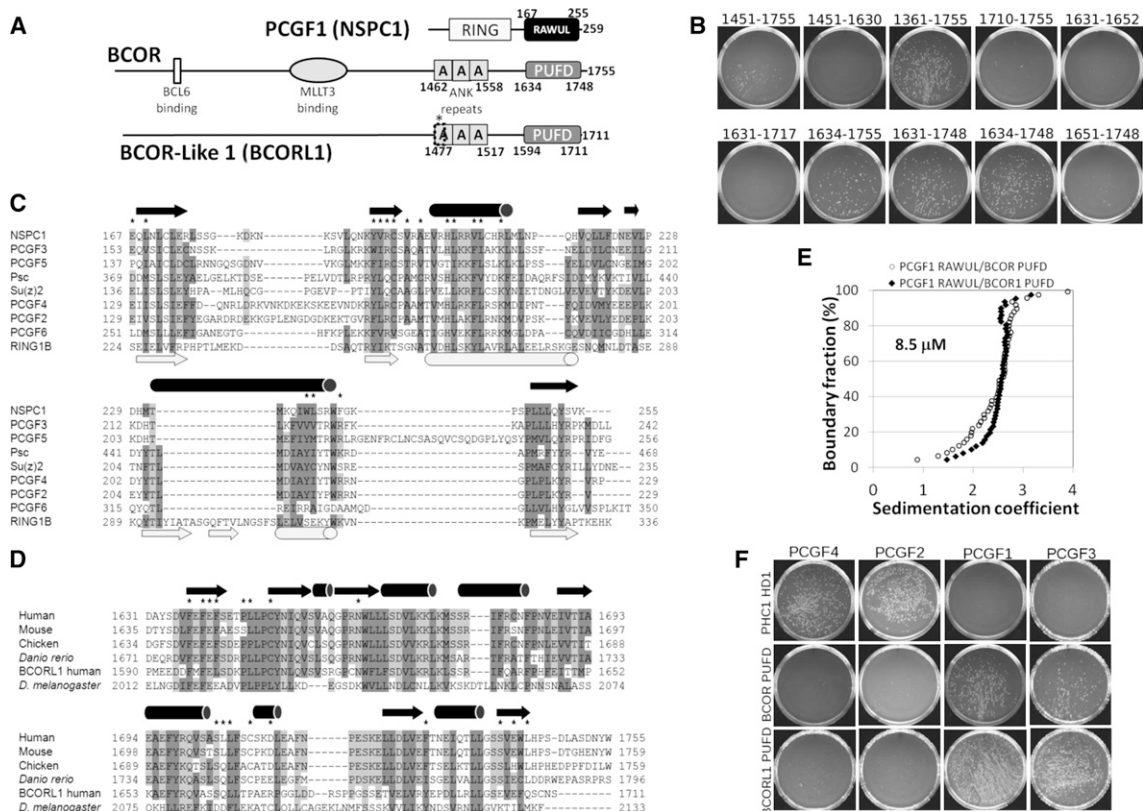
(B) Distinct mammalian PcG complexes are in part defined by the RAWUL binding reactions. The figure is similar to the one proposed by Gao et al. (Gao et al., 2012), which identified the RING1B RAWUL-binding partners along with the PCGF homologs as playing a key role in defining distinct PRC1-like assemblies. The phylogenetic tree of the PCGF proteins is based on the sequence alignment of their RAWUL domains.

chosen buffer, an *s* value of  $\sim 2.5$  for the complexes was observed (Figure 2E). Using a genetic algorithm-Monte Carlo analysis (Brookes and Demeler, 2006; Demeler and Brookes, 2008), the molecular weights were calculated to be 22 and 26 kDa for the PCGF1 RAWUL in complex with the PUFD from BCOR and BCORL1, respectively (Figure S1 available online). This closely matches the expected 25 kDa for the complex, indicating that PCGF1 RAWUL forms a stable complex with the PUFDS with an equilibrium dissociation constant under 8.5  $\mu$ M.

We next investigated whether the PCGF RAWULs could discriminate between their binding partners (Figure 2F). In bacterial two-hybrid assays, the PCGF2 and PCGF4 RAWULs exhibit a positive binding signal with PHC1 HD1 but not with BCOR or BCORL1 PUFDS. In contrast, the PCGF1 and PCGF3 RAWULs bind the PUFDS of BCOR and BCORL1 but do not interact with PHC1 HD1. Consistent with this result, PCGF3 has been found in the same complex with BCORL1 (Gao et al., 2012). These results highlight the different binding selectivities of the PCGF RAWULs.

### PCGF1 RAWUL/PUFD Structure Reveals Conserved Binding Surfaces on the RAWUL

To probe the molecular basis for how the RAWULs discriminate between binding partners, we determined the crystal structures of the PCGF1 RAWUL in complex with the PUFD of BCOR and also with BCORL1 (Figures 3A and S2; Table S1). The structure of the PUFD core shows similarity to the dachshund homology domain (Kim et al., 2002; Wilson et al., 2004; Figure S3A; comparison of the RING1B and PCGF1 RAWULs is shown in Figure S4). However, the PUFD displays clear deviations from



**Figure 2. BCOR PUFD Binds PCGF1 Ub Fold**

(A) Domains of PCGF1, BCOR, and BCORL1 are shown. \*, Although the RefSeq for BCORL1 is lacking an exon encoding a portion of the first ankyrin repeat, this exon is almost always included in human and other species' complementary DNAs.

(B) The minimum BCOR region required to bind PCGF1 was determined using bacterial two-hybrid assay. The PCGF1 RAWUL (167–255) was coexpressed with the indicated residues of BCOR.

(C and D) The sequence alignments of (C) RAWULs and (D) PUFDs are shown. The NCBI gene ID for the *Drosophila melanogaster* protein is CG14073. Sheets and helices are indicated by arrows and cylinders. The secondary structure of RING1B is indicated below its sequence. Conserved hydrophobic and polar residues are dark- and light-shaded, respectively. \*, Residues that make intermolecular contacts are highlighted.

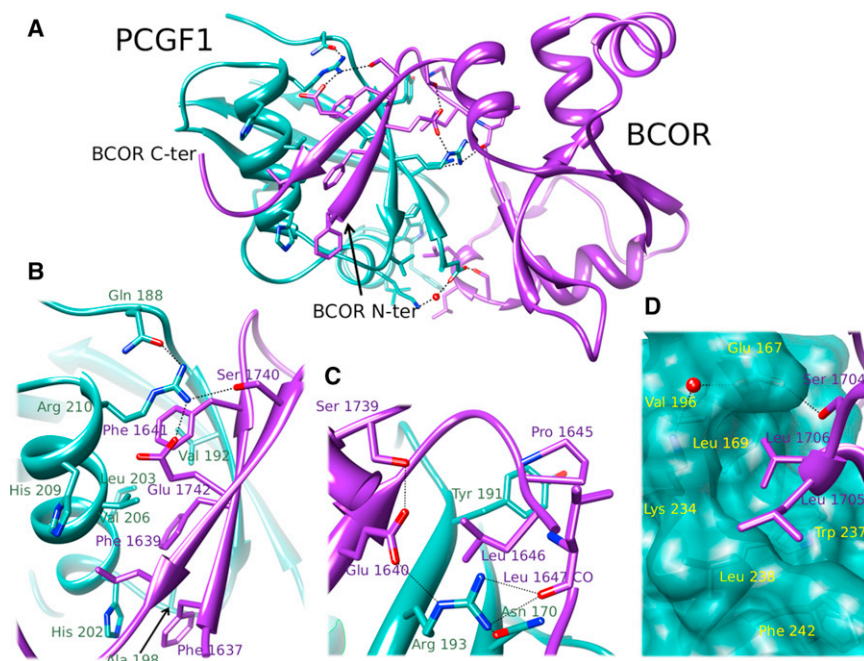
(E) van-Holder-Weischet integral distribution plot (Demeler and van Holde, 2004) of the PCGF1 RAWUL/PUFD complexes.

(F) Bacterial two-hybrid assay of the RAWULs of PCGF2, PCGF4, PCGF1, and PCGF3 show selective binding abilities. PCGF5 RAWUL produced all negative results. Due likely to its toxicity, PCGF6 RAWUL could not be subcloned and was thus not tested.

See also Figure S1.

the core of this fold that affords PUFD the unique ability to bind PCGF1 (Figure S3B). PCGF1 RAWUL and BCOR PUFD form an intermolecular  $\beta$  sheet similar to that observed between RING1B and either RYBP or CBX7 (Wang et al., 2010; Figure 3B). However, while both RYBP and CBX7 form their  $\beta$  sheets from ~20 consecutive residues, the BCOR PUFD is unique in that its N and C termini come together to form the  $\beta$  sheet. A key residue at the PCGF1  $\beta$  sheet surface is Val 206, which packs against BCOR Phe 1639. In the RING1B RAWUL, a Tyr residue is at a position in the RAWUL equivalent to Val 206 (RING1B Tyr 262). This position may play a key discriminatory role in selecting the binding partner of the RAWUL (see below). As in the RING1B RAWUL, a second surface, the loop-binding surface, is present on PCGF1 (Figure 3C). Like RING1B Tyr 247, which is central to creating its loop-binding surface, PCGF1 Tyr 191 is at the equivalent position and plays a similar role. A third binding surface is present in PCGF1, which we refer to as the Leu cage because it completely encases BCOR Leu 1706 (and the equivalent BCORL1 Leu 1665) (Figure 3D).

As noted above, we were unable to isolate the PCGF1 RAWUL alone or the RAWULs from any of the other PCGF homologs in order to quantitatively assess their binding interactions. As far as we are aware, no such studies have yet to be reported for any PCGF RAWUL, likely reflecting their instability, which we have observed. However, we could make qualitative assessments regarding the function of the RAWULs using the bacterial two-hybrid assay. This assay presents an advantage in sensitivity in that even proteins that express insolubly in bacteria can still yield a binding signal. Using this assay, we assessed the contribution made by each of the binding surfaces (Figures 4A and 4B). Consistent with the important role for Val 206 observed in the crystal structure, the PCGF1 Val206Asp mutant, a  $\beta$  sheet surface mutation, shows no positive interaction signal. Similarly, loop-binding surface mutations Tyr191Ala and Arg193Ala show hindered binding to BCOR. We cannot rule out the possibility that the negative results are the consequence of the mutations hindering the ability of the PCGF1 RAWUL to form the proper three-dimensional structure rather than



**Figure 3. Structure of PCGF1 RAWUL/BCOR PUF D Complex**

(A) The asymmetric unit of the PCGF1 RAWUL/BCOR PUF D crystal structure is shown. (B–D) Views are close-ups of (B)  $\beta$  sheet-, (C) loop-, and (D) Leu cage-binding surfaces. Residues are labeled in different colors for PCGF1 and BCOR. See also Figures S2–S4 and Table S1.

a consequence of disrupting specific contacting residues required for stable complex formation. However, Try191, Arg193, and Val206 are all mostly solvent exposed in the absence of PUF D (83%, 81%, and 62%, respectively), and thus these mutations would be expected to have less of an impact on folding than mutations within the core of the RAWUL. It is also worth noting that PCGF1 Tyr 191 and Val 206 are equivalent to RING1B Tyr 247 and Tyr 262. When these RING1B residues are mutated to Ala, it did not result in disrupting the tertiary structure of the RING1B RAWUL (Wang et al., 2010), thus providing an additional assurance that the PCGF RAWUL structures would likely tolerate the mutations. The Leu cage appears to contribute less to binding compared to the other surfaces, as BCOR Leu1706Asp or Arg (also BCORL1 Leu1665Asp or Arg) showed little difference from wild type.

#### RAWUL Binding Selectivity Requires Matched Surface Complementarity

Because equivalent binding surfaces on the RAWULs of both PCGF1 and RING1B are utilized for binding, it would seem likely that PCGF2 and PCGF4 also bind similarly to their binding partners. Indeed, PCGF2 and PCGF4 mutations predicted to be at both the  $\beta$  sheet and loop-binding surfaces resulted in far fewer colonies than wild type in their interaction with PHC1 HD1 (Figure 4C). The unavailability of the structure of PCGF2 or PCGF4 bound to HD1 precludes knowing the precise structural features that provide the binding selectivity of the PCGF proteins. However, comparative analysis with RING1B indicates that binding surfaces of the interacting partners must be complementary without steric clashes or gaps (Figure 5). Unlike the PCGF1 RAWUL, the RING1B RAWUL-binding partners are much shorter in length, binding to a stretch of 30 residues within the CBX homologs, called the cbox domain, and also to 35 residues within a protein called RYBP (Wang et al., 2010; Figure 5A,

compare the CBX7 cbox in yellow to the purple PUF D). Despite this difference, the binding modes of the RAWULs show many similarities. RING1B Tyr 262 is in the equivalent position as PCGF1 Val 206. When RING1B RAWUL is overlaid onto PCGF1 RAWUL, the large aromatic side chain of RING1B Try 262 would sterically clash (Figure 5B, red arrow) with the BCOR Phe 1641, thereby precluding binding between RING1B and BCOR. In contrast, the smaller side chain of PCGF1 Val 206 can accommodate BCOR Phe 1641. The lack of a large hydrophobic residue in CBX7 (or RYBP)

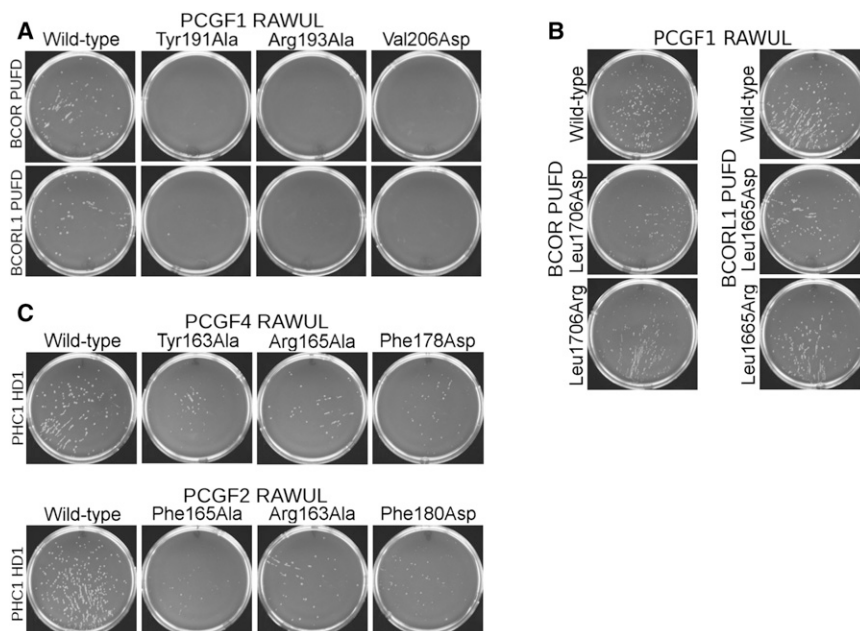
equivalent to BCOR Phe 1641 would also result in a large gap at the interface (Figure 5C, red shading), disfavoring interactions between PCGF1 and CBX7. The binding selectivity predicted by this overlay analysis was confirmed in the bacterial two-hybrid assay (Figure 5D). The RAWULs of PCGF1 and RING1B show no binding signal with the other RAWUL-binding partner. These results suggest that surface complementarity likely plays an important role in determining binding selectivity.

#### DISCUSSION

In sum, we present structural and functional data that reveal how a RAWUL of a PCGF homolog recognizes its binding partner. The PCGF1 RAWUL uses the same binding surfaces as the RING1B RAWUL, suggesting other PCGF RAWULs interact with their respective binding partners in a similar fashion. Despite this similarity in binding modes, very small differences in the contacting surfaces of the PCGF RAWULs are likely to result in completely changing their binding preference and, consequently, the identities of the larger multiprotein complex.

Our study suggests that specific residues on both the  $\beta$ -sheet and loop interactions surfaces are critical in selectively binding their partners. Whereas modeling PCGF4 and PCGF2 residues over PCGF1 can be informative in showing the differences with PCGF1, determining the precise basis of the binding selectivity remains a difficult task. The structures of these other RAWUL complexes will have to be determined to understand the precise structural determinants required for PCGF2 and PCGF4 to bind the PHC proteins.

The identity of the PCGF homologs helps distinguish the larger multiprotein complex that assembles (Gao et al., 2012; Figure 1). RING1B in all these complexes can easily be accounted for through its RING finger domain interaction with the RING finger domain of the PCGF homologs. We have shown that the PCGF



**Figure 4. Bacterial Two-Hybrid Assay Testing Binding of Structure-Guided Mutants**

(A) PCGF1 RAWUL housing the indicated mutations were coexpressed with the PUFIDs of BCOR and BCORL1.

(B) Mutations to the PUFID Leu residue encased by the PCGF1 RAWUL “Leu cage” did not disrupt binding.

(C) PCGF2 and PCGF4 RAWUL residues equivalent to those in PCGF1 that are important for binding the BCOR PUFID are required to bind PHC1 HD1. The panels are aligned with the equivalent positions shown in (A).

RAWUL performs a selective binding function, which, given the PCGF homolog role in defining the larger complex, could further play a key role in the assembly of the distinct higher order PRC1-related complex. It is possible that the RAWUL interactions may be more than a simple 1:1 protein interaction module but also assist in bringing together the other components, leading to assembly of the distinct PRC1-related complex. Future investigation of these proteins will be required to determine the role that the PCGF1/BCOR interaction plays, if any, in the overall assembly of the particular higher order complex that includes KDM2B and SKP1.

## EXPERIMENTAL PROCEDURES

### Bacterial Two-Hybrid Assays

All genes of PCGF RAWULs (encompassing the following residues: PCGF1 167–255, PCGF2 121–229, PCGF3 152–242, and PCGF4 127–228) were cloned into pTRG as a fusion with the  $\alpha$ -subunit of RNA polymerase. All other genes were cloned into pBT, creating a fusion with bacteriophage  $\lambda$  repressor protein ( $\lambda$ cl). PHC1 HD1 encompasses PHC1 residues 720–760. Results using genes cloned in the opposite vector often produced negative results and are, thus, not reported. Mutations were introduced using QuikChange Site-Directed Mutagenesis Kit (Agilent). Chemical competent BacterioMatch (Stratagene) reporter cells were transformed with miniprep plasmid DNA and plated out onto LB agar plates containing 100  $\mu$ g/ml ampicillin, 50  $\mu$ g/ml tetracycline, and 25  $\mu$ g/ml chloramphenicol. Images of the plates were taken after 16 hr incubation at 37°C.

### Protein Expression/Purification

The genes encoding BCOR 1634–1748 and BCORL1 1594–1711 were cloned into a modified pET-30a (Novagen) vector, which contained a METR leader sequence. The gene-encoding PCGF1 RAWUL, residues 167–255, was cloned into a modified pET-3c (Novagen) vector with a leader sequence containing a hexahistidine tag and a tobacco etch virus (TEV) recognition site (MHHHHHHAMKGVDSPEAELDKKAENLYFQGTR). Either BCOR or BCORL1 were coexpressed with PCGF1 in BL21 Gold (DE3) cells (Stratagene) pretransformed with the pRARE plasmid (Novagen), maintaining both expression plasmids in the presence of kanamycin and ampicillin. Selenomethionine (SeMet) was incorporated following a previously published protocol (Van Duyne et al.,

1993). Briefly, proteins were purified by resuspending cells from 1 l of culture in 10 ml of buffer containing 50 mM Tris pH 8.0, 200 mM NaCl, 10 mM imidazole pH 7.5, 10 mM  $\beta$ -mercaptoethanol, and 5% glycerol. Cells were lysed by sonication and proteins partially purified using Ni affinity chromatography. The N-terminal leader sequence on PCGF1 was cleaved using TEV and subtilisin for the native PCGF1/BCOR and PCGF1/BCORL1 complexes, respectively. The SeMet PCGF1/

BCOR complex was not subjected to protease treatment. Additional purification steps included ion exchange and gel filtration chromatography.

### Analytical Ultracentrifugation

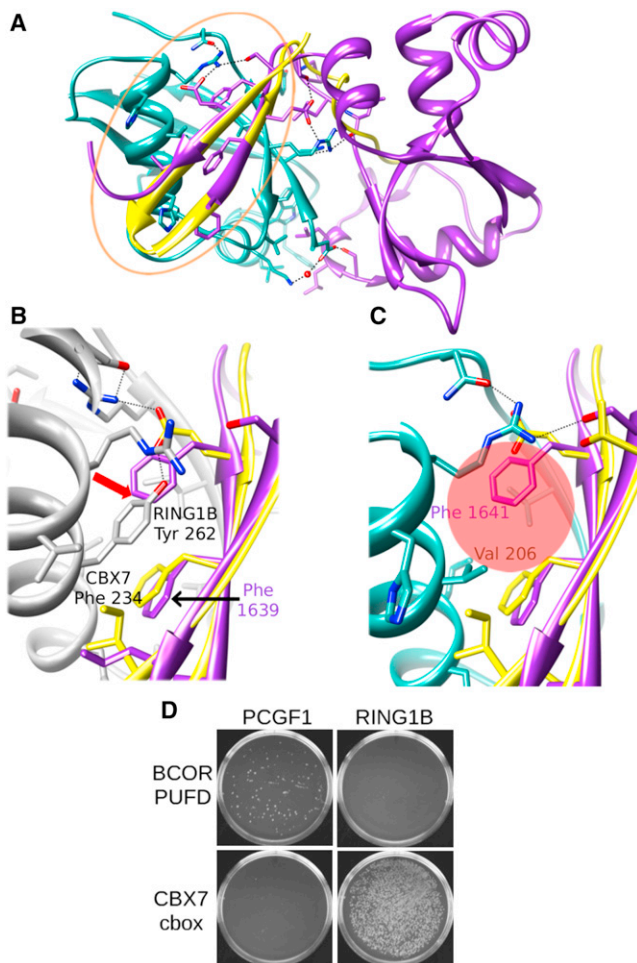
Experiments were conducted using a Beckman Optima XL-I analytical ultracentrifuge. The PCGF1 RAWUL complexes with the PUFIDs were prepared to a concentration of 8.5  $\mu$ M in 10 mM Tris pH 8.0, 50 mM NaCl, and 1 mM TCEP. Experiments were scanned in intensity mode at 280 nm, 20°C, 35 krpm in standard two-channel epon centerpieces (Beckman-Coulter). Ultrascan III (Demeler, 2005) (version 2.0 release 1498, January 2013) was used to analyze the data.

### X-Ray Crystallography

Native PCGF1/BCOR complex crystals were grown in 50 mM 2-(N-morpholino) ethanesulfonic acid pH 6.6, 8.5% polyethylene glycol (PEG) 8000, 300 mM sodium bromide, and 10% ethylene glycol using hanging drop vapor diffusion. SeMet-substituted PCGF1/BCOR crystals were grown in 100 mM Tris pH 8.5, 10% PEG 8000, and 200 mM sodium acetate. The PCGF1/BCORL1 complex crystallized in 100 mM sodium citrate pH 4.5, 19% PEG 1500, and 100 mM dibasic sodium phosphate. A single-wavelength anomalous diffraction data set for the SeMet-substituted PCGF1/BCOR crystal was collected at the Advanced Photon Source (APS) beamline 21-ID-F, processed using XDS (Kabsch, 2010), and input into autoSHARP (Vornrhein et al., 2007) to generate an interpretable electron density map. The overall figure of merit to 2.45 Å resolution was 0.38 before density modification. An initial model was built using Coot (Emsley and Cowtan, 2004) and then refined against the native 2.1 Å data set collected at the Advanced Light Source beamline 4.2.2. The refined PCGF1/BCOR structure was used as the search model for the 1.9 Å native PCGF1/BCORL1 data set collected at APS beamline 21-D-F. Molecular replacement using PHASER (Storoni et al., 2004) identified a solution containing two PCGF1/BCORL1 complexes in the asymmetric unit. All structures were refined using PHENIX (Adams et al., 2002). Ramachandran statistics (Chen et al., 2010) are as follows: PCGF1/BCOR: favored 98.42%, allowed 1.05%, outliers 0.53%; PCGF1/BCORL1: favored 98.94%, allowed 1.06%, 0% outliers. All figures were prepared using Chimera (Pettersen et al., 2004).

### ACCESSION NUMBERS

Protein Data Bank (PDB) IDs are 4HPL for PCGF1/BCOR and 4HPM for PCGF1/BCORL1.



**Figure 5. Binding Specificity between RAWULs and Binding Partners Require Matched Surface Complementarity**

(A) The asymmetric unit of the PCGF1 (cyan)/BCOR (purple) structure is shown. The entire CBX7 cbox (yellow) from the RING1B/CBX7 structure (PDB ID: 3GS2) is overlaid, showing the contiguous nature of its  $\beta$  sheet in contrast to that of BCOR, which is formed by bringing together the BCOR PUF termini. The region highlighted by the oval is equivalent to the areas that are magnified in (B) and (C).

(B) The BCOR  $\beta$  sheet (purple) is overlaid on the CBX7  $\beta$  sheet of the RING1B RAWUL (gray)/CBX7 cbox structure (Wang et al., 2010). The red arrow indicates potential steric clash.

(C) CBX7 cbox is overlaid on BCOR of the PCGF1/BCOR structure. The red shading shows potential gap that could disfavor the interaction between PCGF1 and CBX7.

(D) Cross-binding between the RAWULs (PCGF1 and RING1B) and their respective binding partners was tested using the bacterial two-hybrid assay.

#### SUPPLEMENTAL INFORMATION

Supplemental Information includes four figures and one table and can be found with this article online at <http://dx.doi.org/10.1016/j.str.2013.02.013>.

#### ACKNOWLEDGEMENTS

This work was supported by the American Heart Association (0830111N to C.A.K.), the American Cancer Society (RSG-08-285-01-GMC to C.A.K.), the Department of Defense Breast Cancer Research Program (BC075278 to C.A.K.), the National Institutes of Health (5R01CA071540 to V.J.B.), NIH-NCI

(2P30 CA054174-17 to B.D. and P.J.H.), XSEDE (TG-MCB070039 to B.D.), and the Welch Foundation (AQ-1399 to P.J.H.). UTHSCSA core facilities are supported by the UTHSCSA Executive Research Committee and the Cancer Therapy Research Center. This work is based upon research at Beamline 4.2.2 of the Molecular Biology Consortium at the Advance Light Source (ALS). ALS is supported by the Director, Office of Science, Office of Basic Energy Sciences, of the U.S. Department of Energy under Contract No. DE-AC02-05CH11231. Advanced Photon Source is supported by the U.S. Department of Energy, Office of Science, Office of Basic Energy Sciences, under Contract No. DE-AC02-06CH11357. LS-CAT Sector 21 was supported by the Michigan Economic Development Corporation and the Michigan Technology Tri-Corridor (Grant 085P1000817). We also thank Dr. Jay Nix (ALS) and Dr. Dmitri Ivanov (APS) for assistance with the data collection.

Received: November 12, 2012

Revised: February 5, 2013

Accepted: February 11, 2013

Published: March 21, 2013

#### REFERENCES

- Adams, P.D., Grosse-Kunstleve, R.W., Hung, L.W., Ioerger, T.R., McCoy, A.J., Moriarty, N.W., Read, R.J., Sacchettini, J.C., Sauter, N.K., and Terwilliger, T.C. (2002). PHENIX: building new software for automated crystallographic structure determination. *Acta Crystallogr. D Biol. Crystallogr.* **58**, 1948–1954.
- Alkema, M.J., Bronk, M., Verhoeven, E., Otte, A., van 't Veer, L.J., Berns, A., and van Lohuizen, M. (1997). Identification of Bmi1-interacting proteins as constituents of a multimeric mammalian polycomb complex. *Genes Dev.* **11**, 226–240.
- Brookes, E., and Demeler, B. (2006). Genetic algorithm optimization for obtaining accurate molecular weight distributions from sedimentation velocity experiments. In *Analytical Ultracentrifugation VIII*, Progr. Colloid Polym. Sci., C. Wandrey and H. Colfen, eds. (Berlin-Heidelberg: Springer), p. 78.
- Buchwald, G., van der Stoop, P., Weichenrieder, O., Perrakis, A., van Lohuizen, M., and Sixma, T.K. (2006). Structure and E3-ligase activity of the Ring-Ring complex of polycomb proteins Bmi1 and Ring1b. *EMBO J.* **25**, 2465–2474.
- Chen, V.B., Arendall, W.B., 3rd, Headd, J.J., Keedy, D.A., Immormino, R.M., Kapral, G.J., Murray, L.W., Richardson, J.S., and Richardson, D.C. (2010). MolProbity: all-atom structure validation for macromolecular crystallography. *Acta Crystallogr. D Biol. Crystallogr.* **66**, 12–21.
- Demeler, B. (2005). UltraScan A comprehensive data analysis software package for analytical ultracentrifugation experiments. In *Modern Analytical Ultracentrifugation: Techniques and Methods*, D.J. Scott, S.E. Harding, and A.J. Rowe, eds. (UK: Royal Society of Chemistry), pp. 210–229.
- Demeler, B., and Brookes, E. (2008). Monte Carlo analysis of sedimentation experiments. *Colloid Polym. Sci.* **286**, 129–137.
- Demeler, B., and van Holde, K.E. (2004). Sedimentation velocity analysis of highly heterogeneous systems. *Anal Biochem.* **335**, 279–288.
- Dove, S.L., and Hochschild, A. (2004). A bacterial two-hybrid system based on transcription activation. *Methods Mol. Biol.* **261**, 231–246.
- Emsley, P., and Cowtan, K. (2004). Coot: model-building tools for molecular graphics. *Acta Crystallogr. D Biol. Crystallogr.* **60**, 2126–2132.
- Francis, N.J., Saurin, A.J., Shao, Z., and Kingston, R.E. (2001). Reconstitution of a functional core polycomb repressive complex. *Mol. Cell* **8**, 545–556.
- Gao, Z., Zhang, J., Bonasio, R., Strino, F., Sawai, A., Parisi, F., Kluger, Y., and Reinberg, D. (2012). PCGF homologs, CBX proteins, and RYBP define functionally distinct PRC1 family complexes. *Mol. Cell* **45**, 344–356.
- Gearhart, M.D., Corcoran, C.M., Wamstad, J.A., and Bardwell, V.J. (2006). Polycomb group and SCF ubiquitin ligases are found in a novel BCOR complex that is recruited to BCL6 targets. *Mol. Cell. Biol.* **26**, 6880–6889.
- Grossmann, V., Tiacci, E., Holmes, A.B., Kohlmann, A., Martelli, M.P., Kern, W., Spanhol-Rosseto, A., Klein, H.U., Dugas, M., Schindela, S., et al. (2011). Whole-exome sequencing identifies somatic mutations of BCOR in acute myeloid leukemia with normal karyotype. *Blood* **118**, 6153–6163.

- Gunster, M.J., Satijn, D.P., Hamer, K.M., den Blaauwen, J.L., de Bruijn, D., Alkema, M.J., van Lohuizen, M., van Driel, R., and Otte, A.P. (1997). Identification and characterization of interactions between the vertebrate polycomb-group protein BMI1 and human homologs of polyhomeotic. *Mol. Cell Biol.* **17**, 2326–2335.
- Huynh, K.D., Fischle, W., Verdin, E., and Bardwell, V.J. (2000). BCoR, a novel corepressor involved in BCL-6 repression. *Genes Dev.* **14**, 1810–1823.
- Kabsch, W. (2010). Xds. *Acta Crystallogr. D Biol. Crystallogr.* **66**, 125–132.
- Kim, S.S., Zhang, R.G., Braunstein, S.E., Joachimiak, A., Cvekl, A., and Hegde, R.S. (2002). Structure of the retinal determination protein Dachshund reveals a DNA binding motif. *Structure* **10**, 787–795.
- Kyba, M., and Brock, H.W. (1998). The Drosophila polycomb group protein Psc contacts ph and Pc through specific conserved domains. *Mol. Cell Biol.* **18**, 2712–2720.
- Lagarou, A., Mohd-Sarip, A., Moshkin, Y.M., Chalkley, G.E., Bezstarosti, K., Demmers, J.A., and Verrijzer, C.P. (2008). dKDM2 couples histone H2A ubiquitylation to histone H3 demethylation during Polycomb group silencing. *Genes Dev.* **22**, 2799–2810.
- Lee, M.G., Norman, J., Shilatifard, A., and Shiekhhattar, R. (2007). Physical and functional association of a trimethyl H3K4 demethylase and Ring6a/MBLR, a polycomb-like protein. *Cell* **128**, 877–887.
- Levine, S.S., Weiss, A., Erdjument-Bromage, H., Shao, Z., Tempst, P., and Kingston, R.E. (2002). The core of the polycomb repressive complex is compositionally and functionally conserved in flies and humans. *Mol. Cell Biol.* **22**, 6070–6078.
- Li, Z., Cao, R., Wang, M., Myers, M.P., Zhang, Y., and Xu, R.M. (2006). Structure of a Bmi-1-Ring1B polycomb group ubiquitin ligase complex. *J. Biol. Chem.* **281**, 20643–20649.
- Li, M., Collins, R., Jiao, Y., Ouillette, P., Bixby, D., Erba, H., Vogelstein, B., Kinzler, K.W., Papadopoulos, N., and Malek, S.N. (2011). Somatic mutations in the transcriptional corepressor gene BCORL1 in adult acute myelogenous leukemia. *Blood* **118**, 5914–5917.
- Mohd-Sarip, A., Lagarou, A., Doyen, C.M., van der Knaap, J.A., Aslan, U., Bezstarosti, K., Yassin, Y., Brock, H.W., Demmers, J.A., and Verrijzer, C.P. (2012). Transcription-independent function of Polycomb group protein PSC in cell cycle control. *Science* **336**, 744–747.
- Ng, D., Thakker, N., Corcoran, C.M., Donnai, D., Perveen, R., Schneider, A., Hadley, D.W., Tiff, C., Zhang, L., Wilkie, A.O., et al. (2004). Oculofaciocardiodental and Lenz microphthalmia syndromes result from distinct classes of mutations in BCOR. *Nat. Genet.* **36**, 411–416.
- Pettersen, E.F., Goddard, T.D., Huang, C.C., Couch, G.S., Greenblatt, D.M., Meng, E.C., and Ferrin, T.E. (2004). UCSF Chimera—a visualization system for exploratory research and analysis. *J. Comput. Chem.* **25**, 1605–1612.
- Pierron, G., Tirode, F., Lucchesi, C., Reynaud, S., Ballet, S., Cohen-Gogo, S., Perrin, V., Coindre, J.M., and Delattre, O. (2012). A new subtype of bone sarcoma defined by BCOR-CCNB3 gene fusion. *Nat. Genet.* **44**, 461–466.
- Pugh, T.J., Weeraratne, S.D., Archer, T.C., Pomeranz Krummel, D.A., Auclair, D., Bochicchio, J., Carneiro, M.O., Carter, S.L., Cibulskis, K., Erlich, R.L., et al. (2012). Medulloblastoma exome sequencing uncovers subtype-specific somatic mutations. *Nature* **488**, 106–110.
- Sánchez, C., Sánchez, I., Demmers, J.A., Rodríguez, P., Strouboulis, J., and Vidal, M. (2007). Proteomics analysis of Ring1B/Rnf2 interactors identifies a novel complex with the Fbx10/Jhdml1B histone demethylase and the Bcl6 interacting corepressor. *Mol. Cell. Proteomics* **6**, 820–834.
- Sanchez-Pulido, L., Devos, D., Sung, Z.R., and Calonje, M. (2008). RAWUL: a new ubiquitin-like domain in PRC1 ring finger proteins that unveils putative plant and worm PRC1 orthologs. *BMC Genomics* **9**, 308.
- Shao, Z., Raible, F., Mollaaghababa, R., Guyon, J.R., Wu, C.T., Bender, W., and Kingston, R.E. (1999). Stabilization of chromatin structure by PRC1, a Polycomb complex. *Cell* **98**, 37–46.
- Simon, J.A., and Kingston, R.E. (2009). Mechanisms of polycomb gene silencing: knowns and unknowns. *Nat. Rev. Mol. Cell Biol.* **10**, 697–708.
- Storoni, L.C., McCoy, A.J., and Read, R.J. (2004). Likelihood-enhanced fast rotation functions. *Acta Crystallogr. D Biol. Crystallogr.* **60**, 432–438.
- Van Duyn, G.D., Standaert, R.F., Karplus, P.A., Schreiber, S.L., and Clardy, J. (1993). Atomic structures of the human immunophilin FKBP-12 complexes with FK506 and rapamycin. *J. Mol. Biol.* **229**, 105–124.
- Vonrhein, C., Blanc, E., Roversi, P., and Bricogne, G. (2007). Automated structure solution with autoSHARP. *Methods Mol. Biol.* **364**, 215–230.
- Wamstad, J.A., Corcoran, C.M., Keating, A.M., and Bardwell, V.J. (2008). Role of the transcriptional corepressor Bcor in embryonic stem cell differentiation and early embryonic development. *PLoS ONE* **3**, e2814.
- Wang, H., Wang, L., Erdjument-Bromage, H., Vidal, M., Tempst, P., Jones, R.S., and Zhang, Y. (2004). Role of histone H2A ubiquitination in Polycomb silencing. *Nature* **431**, 873–878.
- Wang, R., Taylor, A.B., Leal, B.Z., Chadwell, L.V., Ilangovan, U., Robinson, A.K., Schirf, V., Hart, P.J., Lafer, E.M., Demeler, B., et al. (2010). Polycomb group targeting through different binding partners of RING1B C-terminal domain. *Structure* **18**, 966–975.
- Wilson, J.J., Malakhova, M., Zhang, R., Joachimiak, A., and Hegde, R.S. (2004). Crystal structure of the dachshund homology domain of human SKI. *Structure* **12**, 785–792.
- Yamamoto, Y., Tsuzuki, S., Tsuzuki, M., Handa, K., Inaguma, Y., and Emi, N. (2010). BCOR as a novel fusion partner of retinoic acid receptor alpha in a t(X;17)(p11;q12) variant of acute promyelocytic leukemia. *Blood* **116**, 4274–4283.
- Zhang, J., Benavente, C.A., McEvoy, J., Flores-Otero, J., Ding, L., Chen, X., Ulyanov, A., Wu, G., Wilson, M., Wang, J., et al. (2012). A novel retinoblastoma therapy from genomic and epigenetic analyses. *Nature* **481**, 329–334.