Supplementary materials for

Bending and looping of long DNA by Polycomb repressive complex 2 revealed by AFM imaging in liquid

Patrick R. Heenan[1,2], Xueyin Wang[3], Anne R. Gooding[3], Thomas R. Cech[3,4] and Thomas T. Perkins[2,5*]

[1]Department of Physics, University of Colorado, Boulder, Colorado 80309, USA

[2]JILA, National Institute of Standards and Technology and University of Colorado, Boulder, Colorado 80309, USA

[3]Departments of Biochemistry & BioFrontiers Institute, University of Colorado, Boulder, Colorado 80309, USA

[4]Howard Hughes Medical Institute, University of Colorado, Boulder, Colorado, USA

[5]Department of Molecular, Cellular, and Developmental Biology, University of Colorado, Boulder, Colorado 80309, USA

* To whom correspondence should be addressed. Tel: +1 303 492 5291; Fax: +1 303 492 5235; Email: tperkins@jila.colorado.edu.

Figures:

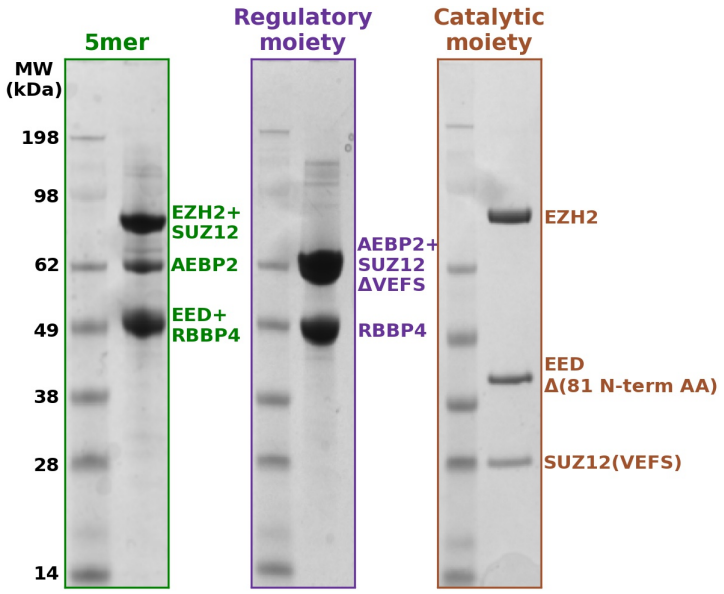

Figure S1. Polyacrylamide gel electrophoresis (PAGE) demonstrates purity of PRC2 moieties. From left to right, PAGE results are shown for the full complex (green text), regulatory moiety (purple text) and catalytic moiety (brown text). For each PRC2 variant, the same ladder is shown on the left lane, and the right lane depicts the PRC2 variant. The ladder is labelled by the molecular weight in kilodaltons, and the major bands for each PRC2 variant are labelled by their constituent components.

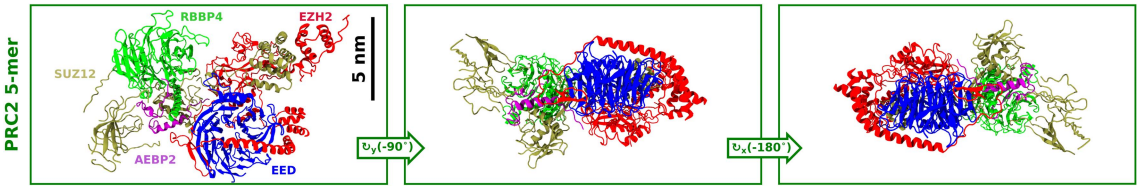

Figure S2. The structure of the PRC2 complex. The structure of the PRC2 "5mer" consisting of enhancer of zeste homolog 2 (EZH2, red), embryonic ectoderm development (EED, blue), suppressor of zeste 12 (yellow, SUZ12), adipocyte enhancer-binding protein 2 (AEBP2, purple), and retinoblastoma-binding protein 4 (RBBP4, green). Images derived from PDB 6C23 (1).

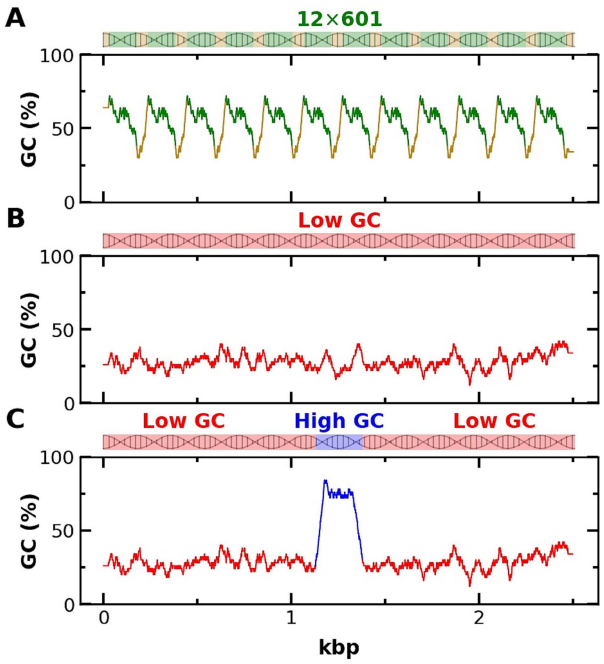

Figure S3. Three 2.5-kbp DNA constructs with spatially varying GC content used for probing the sequence specificity of PRC2 binding to DNA. (**A**) An illustration of the DNA construct (not to scale) and a plot of the averaged GC content versus base pair position. The DNA is color-coded green for the 601 Widom sequence or yellow for a linker sequence. (**B**) A DNA construct with an average GC content of 28% that is color-coded red. (**C**) A DNA construct containing a high GC-rich island region (76%), color-coded blue, inserted at the center of an otherwise low-GC DNA molecule.

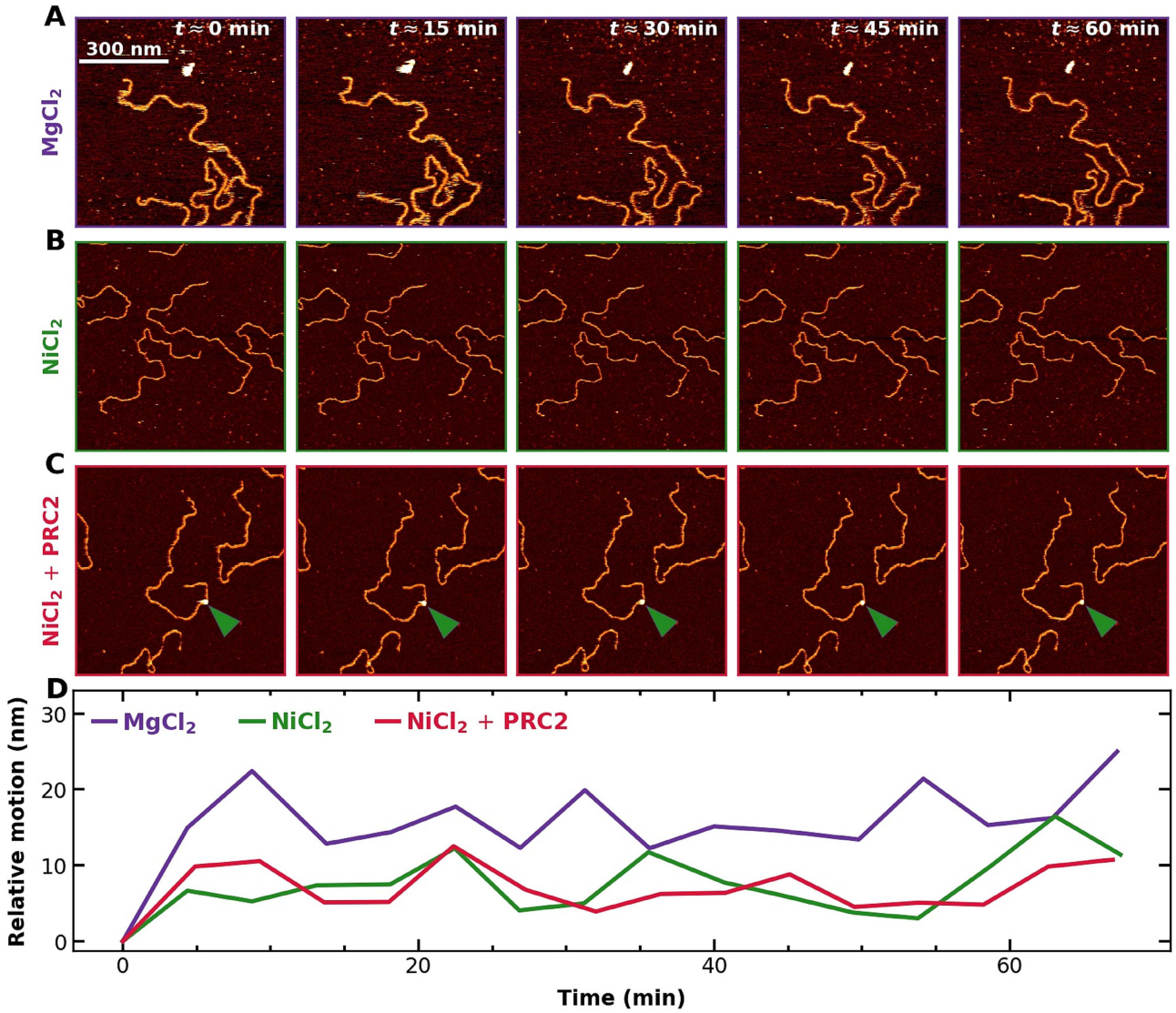

Figure S4. Surface binding of DNA mediated by $MgCl_2$ is similar to surface binding mediated by $NiCl_2$. **(A)** Time-lapsed images of DNA deposited and imaged as described in Materials and Methods, except imaging was done in buffer containing $MgCl_2$ rather than $NiCl_2$. **(B)** Time-lapsed images of DNA deposited and imaged as described in Materials and Methods, with deposition done in $MgCl_2$ and imaging done in buffer containing $NiCl_2$. **(C)** As panel B, except with ~50 nM PRC2 added. Green arrow indicates PRC2 bound to DNA. **(D)** The DNA-protein complex is stably bound to the surface during deposition and imaging, since the root mean squared displacement relative to the center of mass of DNA with PRC2 bound in the standard protocol (red, 7 ± 3 nm, mean ± std. dev., N = 16) is similar to the $MgCl_2$-based deposition and imaging protocol (purple, 16 ± 5 nm, mean ± std. dev., N=16) and similar to the standard protocol with $MgCl_2$-based deposition and $NiCl_2$-based imaging (green, 7 ± 4 nm, mean ± std. dev., N =16). We emphasize that panel *B* and panel *C* were prepared identically, except for the addition of ~50 nM PRC2 in panel *C*. All buffers also contained 25 mM KCl and 10 mM Tris-HCl (pH 7.5).

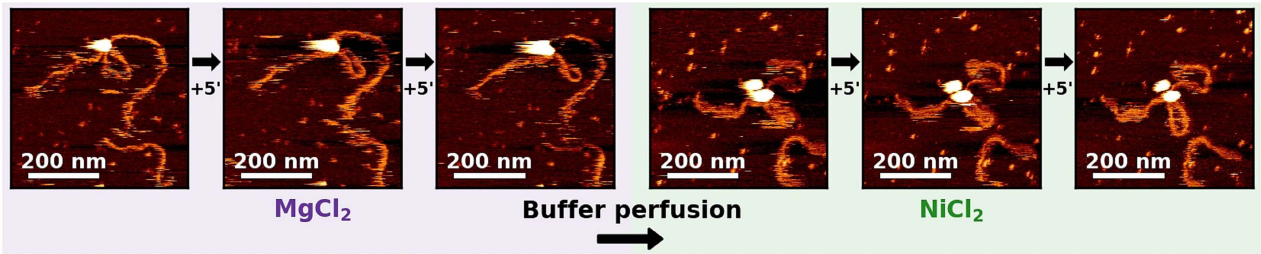

Figure S5. During rinsing, DNA-PRC2 complex remains intact but is capable of nanoscopic rearrangement. Time-lapsed images (600×600 nm^2), with about 5 min between images, of a DNA-PRC2 complex deposited and imaged as described in Materials and Methods, except deposition and imaging were performed in deposition buffer [25 mM KCl, 10 mM $MgCl_2$, 10 mM HEPES (pH 7.5)]. After the first three images, a perfusion system was used to slowly (< 200 µL/minute, 1 mL total) replace the $MgCl_2$ with equimolar $NiCl_2$, and the same complex was imaged another three times after approximately 30 min of settling.

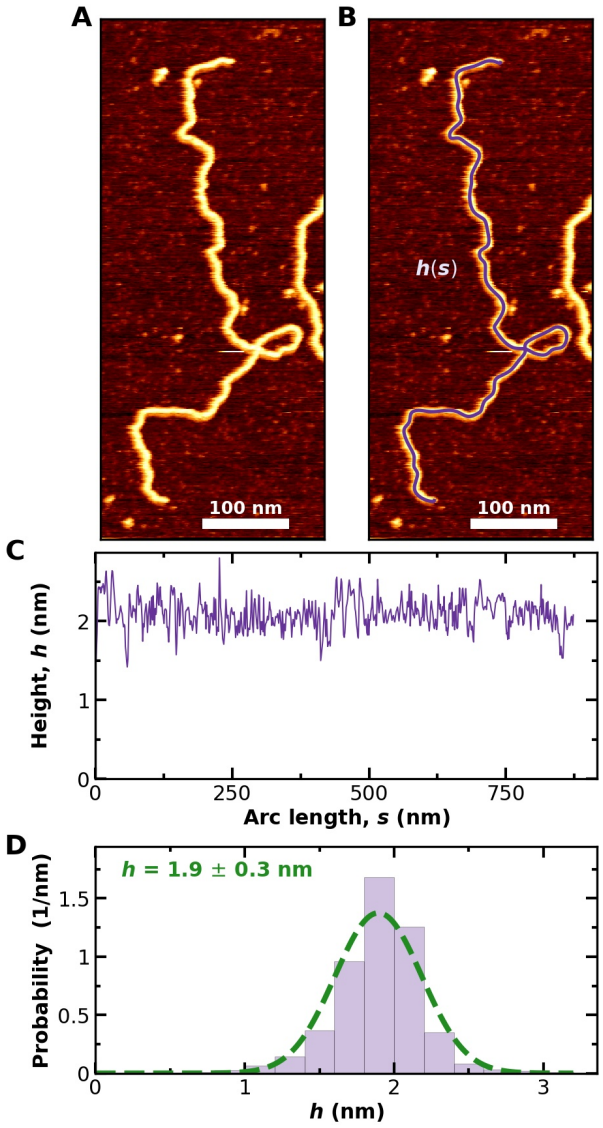

Figure S6. Height of surface-bound DNA is consistent with DNA's crystallographic structure. **(A)** AFM image of naked DNA. **(B)** Image in panel A superimposed with a contour trace. At each point along the arc *s*, the height is measured as *h*(*s*). **(C)** The quantification of the height in panel B yields the expected DNA height (diameter) of 2.0 nm. **(D)** Histogram of all heights along all contours of 52 naked DNA molecules yields an average value of 1.9 ± 0.3 nm (mean ± std. dev.).

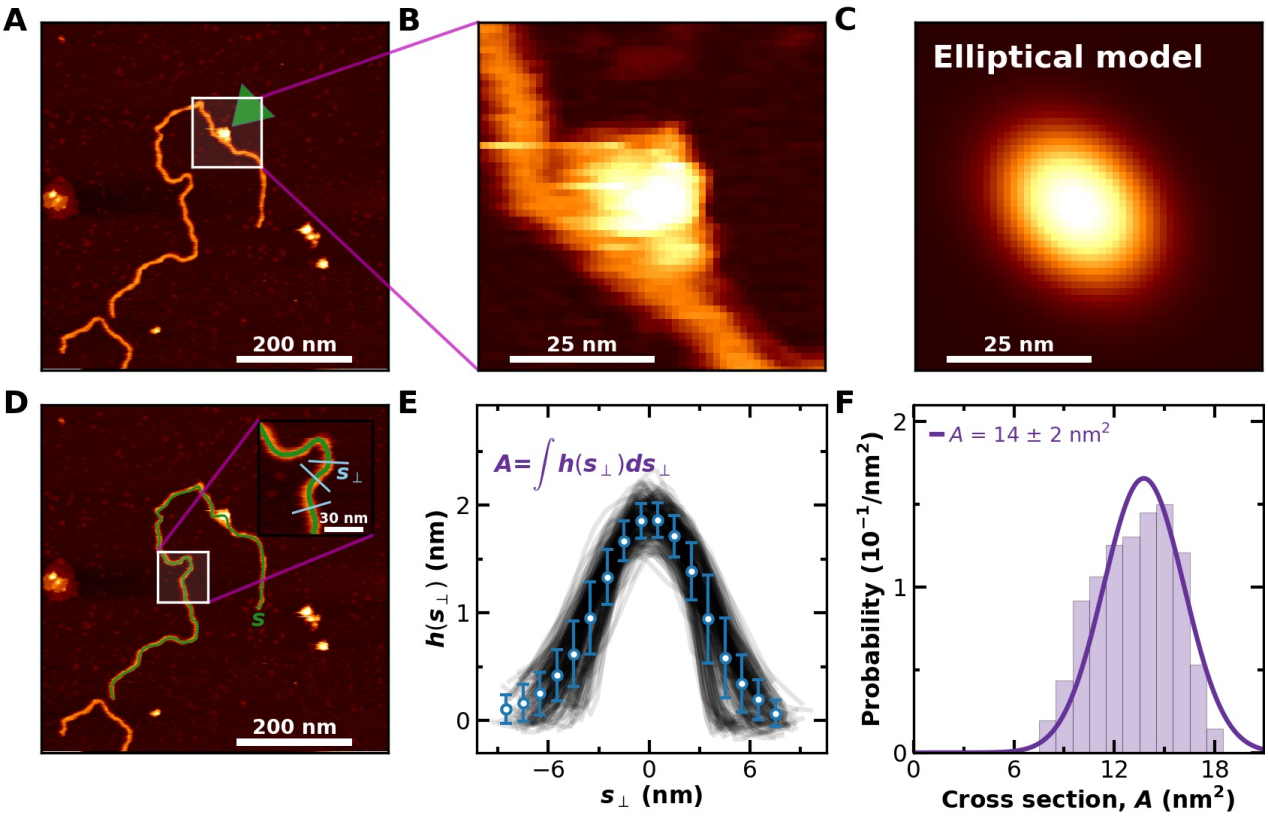

Figure S7. Volume analysis of PRC2-DNA complexes. **(A)** AFM image of DNA with a single PRC2 molecule bound at the position indicated by the green arrow. (**B**) Detailed image centered on the bound PRC2. (**C**) Freely rotating, elliptical Gaussian obtained by fitting to the PRC2-DNA complex of panel *B*. The PRC2-DNA complex volume is obtained by integrating the Gaussian over the region bounded by the complex. (**D**) The images in panel A except with the DNA contour *s* overlayed in green. (**E**) The cross section (*A*) is obtained by integrating the height (*h)* along lines perpendicular to and centered on the DNA arc ($s_\perp$). Only the cross sections further than 50 nm from loops or PRC2 and with maximum heights between 1.5 and 2.5 nm and boundary heights less than 0.5 nm are included (black lines), and the average height at each position is indicated by blue circles with error bars representing the standard deviation (N = 211). (**F**) Histogram of all cross sections obtained by integrating the black lines in panel *E*. The expected cross section for DNA is π nm^2, assuming a rod with radius of 1 nm, with the main contributor to this difference arising from the convolution of the tip radius (r_{nom} = 8 nm for the Olympus BioLever Mini used here).

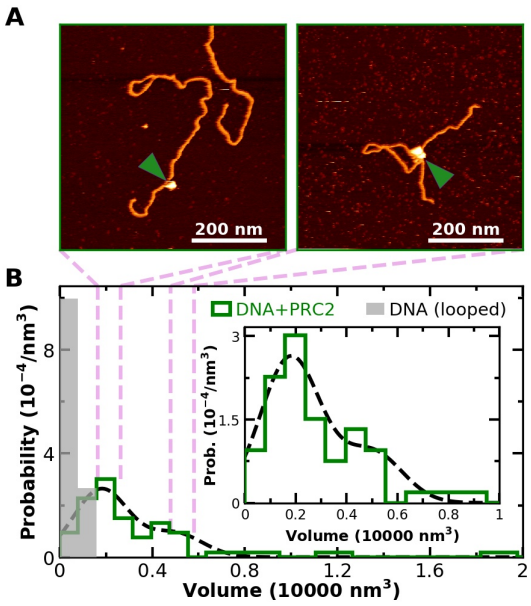

Figure S8. Volume analysis of PRC2-DNA complexes is still consistent with multimeric binding in the absence of deconvolution. This plot is formatted as Figure 7, except the deconvolution procedure detailed in Figure S7 is omitted. In spite of this change, the distribution is still long-tailed, and the distribution at low volumes is still well-described by a sum of two Gaussians.

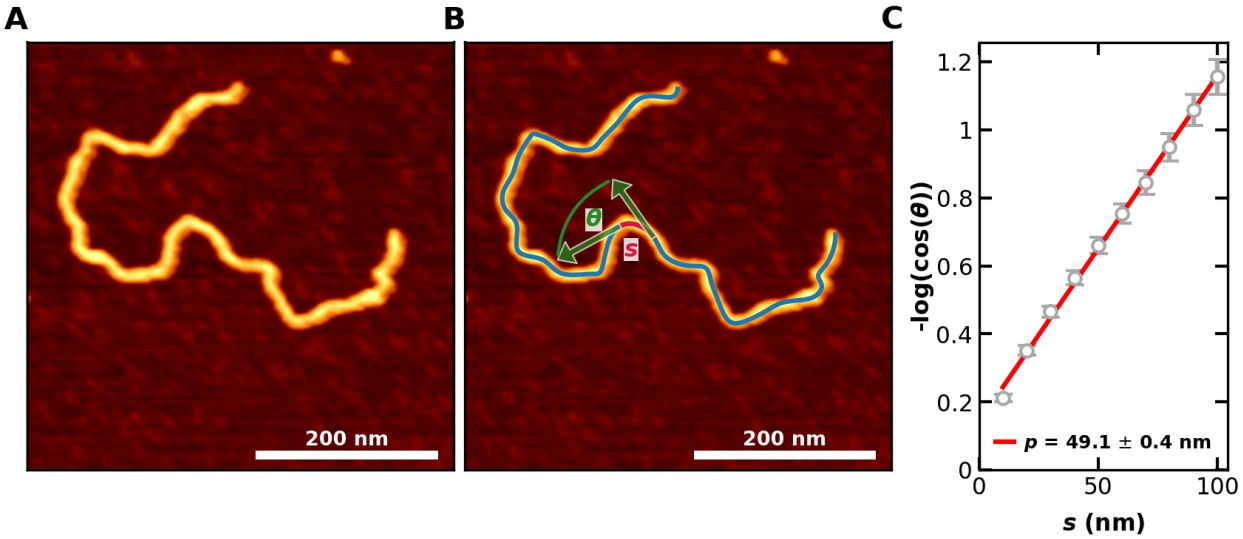

Figure S9. Persistence length (p) analysis agrees with DNA's known value in solution (50 nm) (2). **(A)** AFM image of DNA unbound by PRC2. **(B)** The image in panel A, but with illustrative definitions of the angle between the tangents (θ) at the start and end of an arc length (s). **(C)** The dependence of θ on s is linear when plotted as $-\ln \langle\cos(\theta)\rangle$ vs s. A fit (red line) to data (grey circles) of the equation $\ln(\langle\cos(\theta(s))\rangle) = -s/2p$ yields a persistence length of 49.1 ± 0.4 nm (mean ± std. dev., N = 640) in quantitative agreement with 50 nm. Errors bars are standard deviations. Moreover, this results shows the DNA has equilibrated in 2D with the mica surface, implying a gentle absorption process (3).

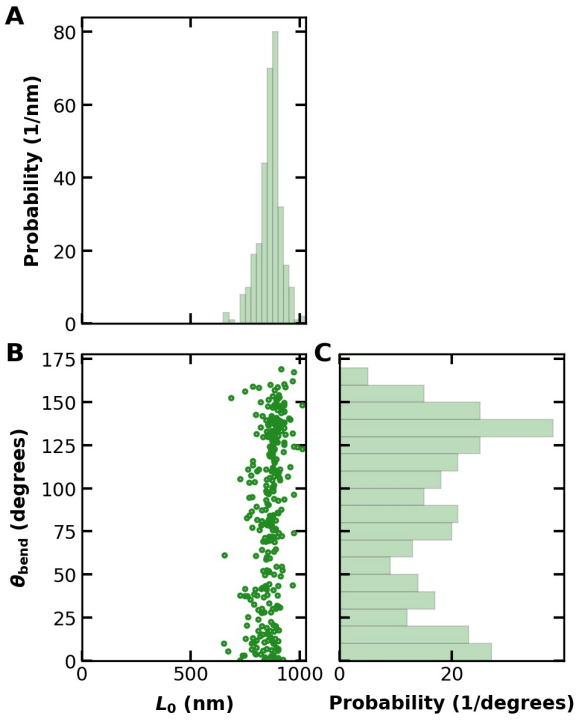

Figure S10. PRC2 bending angles are broadly distributed at a given DNA contour length. **(A)** The probability distribution of contour lengths for DNA-protein complexes where PRC2 is bound to twelve copies of the 601-Widom sequence, where 13 molecules with lengths of less than 650 nm were excluded, out of 331 total. **(B)** Scatter plot of the bending angle versus contour length for the molecules from (A), where the green dots represent individual molecules **(C)** The distribution of bending angles for the data shown in (B). We note that the full width of the contour length distribution is broad and the data is sufficiently noisy that much larger data sets acquired from shorter DNA molecules (*e.g.*, 300 nm) would be suggested for future studies.

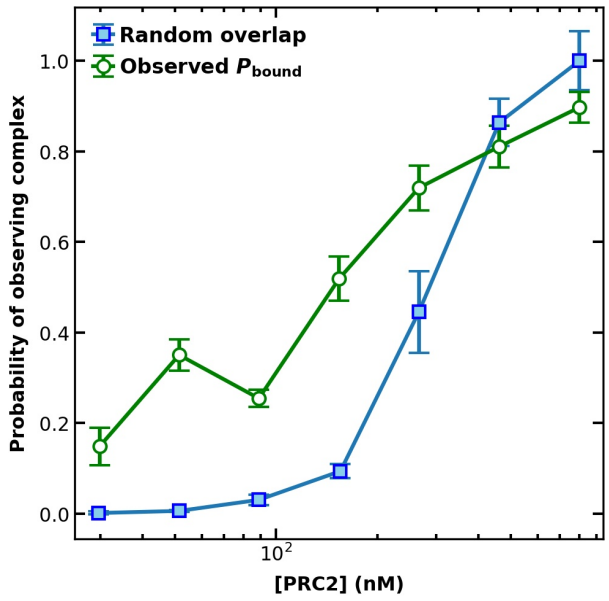

Figure S11. PRC2 concentrations above 300 nM are not useful for measuring PRC2-DNA binding because of PRC2 randomly absorbed onto the surface can accidentally overlap with DNA. The [PRC2]-dependent probability of having DNA bound to PRC2 from Figure 3 (green circles) is plotted alongside the probability of PRC2 randomly overlapping the DNA (blue squares). The bars on the data represent one standard deviation, assuming a binomial distribution. The non-specific adhesion probability was calculated based on at least 3 images at each concentration, using a minimum height of 2 nm for PRC2 and a minimum radius of three pixels (12 nm).

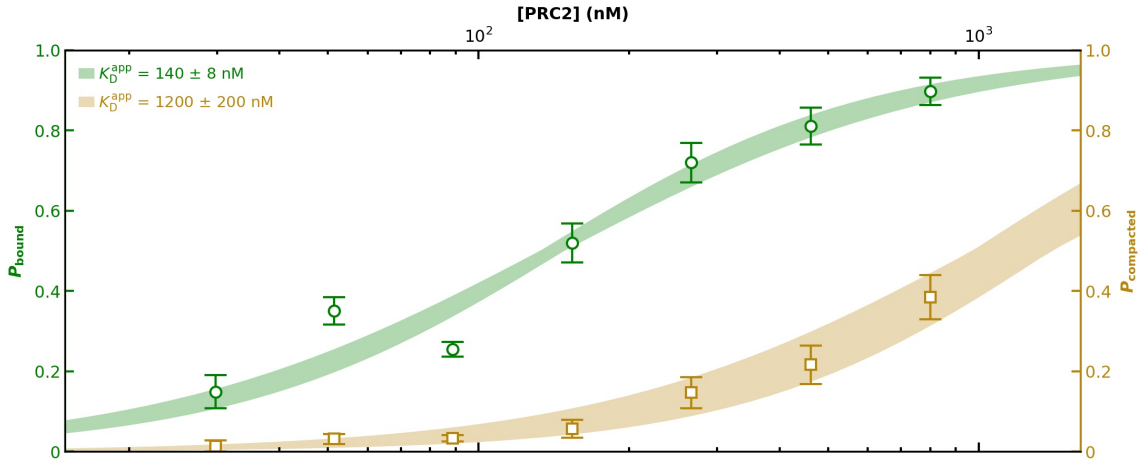

Figure S12. Apparent dissociation constants (K_D^{app}) are not substantially changed when measurements at high protein concentrations are included. This plot and its fits are as Figure 3 A, except the measurements at the highest two concentrations – with the highest protein backgrounds, as shown in Figure S7 – are included in the fitting. The Hill coefficients for the bound (green) and compacted (yellow) curves are 1.2 ± 0.1 and 1.3 ± 0.1, respectively.

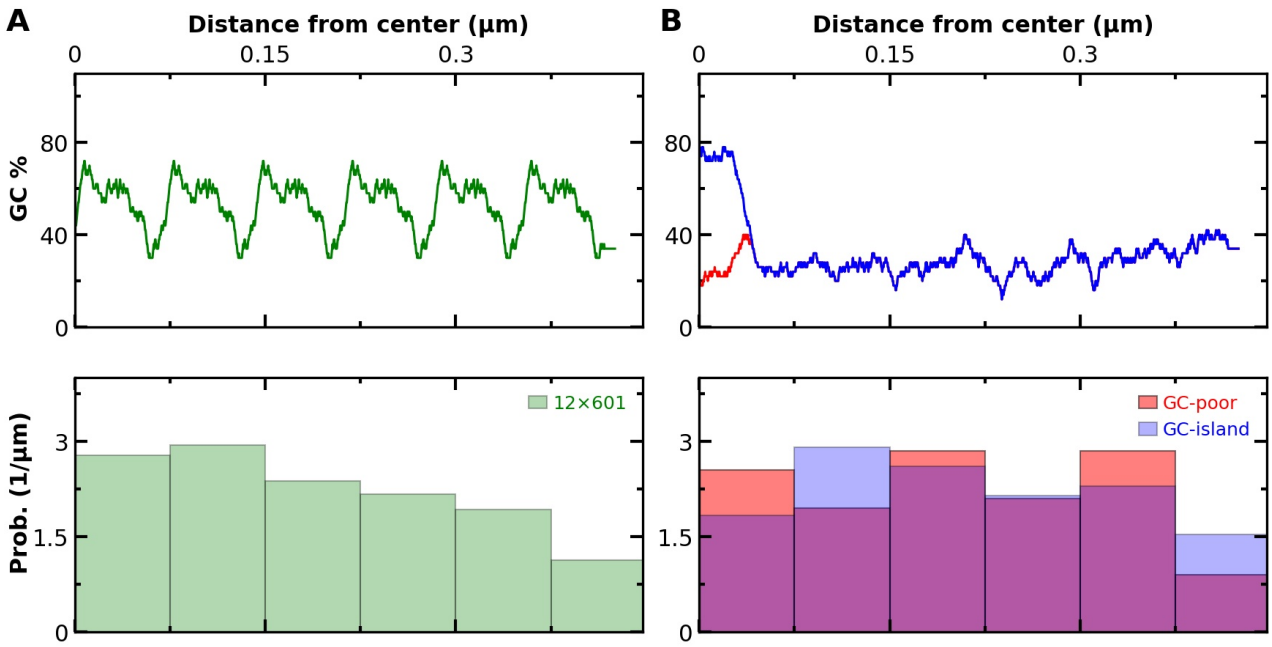

Figure S13. PRC2 binding locations not strongly dependent on GC content when measured relative to the center of the DNA molecule. Plot formatted as Figure 4A–B, except the horizontal axis is the arc length from the center of the PRC2 to the center of the DNA molecules.

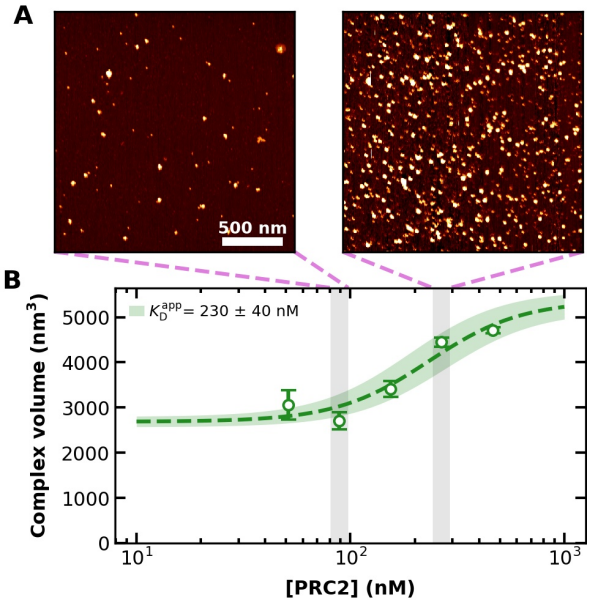

Figure S14. PRC2 apparent volumes are consistent with dimerization at high protein concentrations. **(A)** 2×2 μm^2 images of PRC2 in the absence of DNA at low (left) and high (right) concentrations of PRC2. **(B)** Quantification of average PRC2 volume using the method of Figure S6, except omitting deconvolution (green circles, mean ± std. err.) yields an apparent dissociation constant of dimerization of 230 ± 40 nM (mean ± fitting std. dev). The fit (green shaded area, mean ± std. dev.) is Equation 2, where the Hill coefficient was set to 2, and the equation was multiplied and then offset by the fitted monomer volume (2700 ± 120, mean ± fitting std. dev). We note that the data shown here are derived from the same data used in the blue markers in Figure S11, except the lowest and highest concentration points were omitted due to the small number of protein complexes and the high probability of complex-complex overlap, respectively.

REFERENCES

1. Kasinath, V., Faini, M., Poepsel, S., Reif, D., Feng, X.A., Stjepanovic, G., Aebersold, R. and Nogales, E. (2018) Structures of human PRC2 with its cofactors AEBP2 and JARID2. *Science*, **359**, 940-944.
2. Hagerman, P.J. (1988) Flexibility of DNA. *Annu. Rev. Biophys. Biophys. Chem.*, **17**, 265-286.
3. Heenan, P.R. and Perkins, T.T. (2019) Imaging DNA equilibrated onto mica in liquid using biochemically relevant deposition conditions. *ACS Nano*, **13**, 4220-4229.