Electronic Supplementary Material (ESI) for ChemComm.

Supplementary Information

Polycomb Repressive Complex 2 Binds RNA Irrespective of Stereochemistry

Charles E. Deckard III and Jonathan T. Sczepanski*

Department of Chemistry, Texas A&M University, College Station, Texas 77843, United States

S1. Supplementary Text

Materials and Methods.

General. The DNA and RNA oligonucleotides were either purchased from Integrated DNA Technologies (IDT, Coralville, IA) or prepared using an Expedite 8909 DNA/RNA synthesizer. Oligonucleotide synthesis reagents, D-nucleoside phosphoramidites, and Cyanine 3 (Cy3) phosphoramidite were purchased from Glen Research (Sterling, VA), and L-nucleoside phosphoramidites were purchased from ChemGenes (Wilmington, Ma). All oligonucleotides were purified by polyacrylamide gel electrophoresis (PAGE) and desalted by ethanol precipitation; the purity and integrity of each oligonucleotide was confirmed with mass spectrometry (Novatia, Newton, PA). Polycomb Repressive Complex 2 was purchased from Active Motif (Carlsbad, CA). N-Hydroxysuccinimide (NHS) ester of Cyanine 5 (Cy5) used in the labeling of HOTAIR was acquired from Lumiprobe Life Science Solutions (Hallendale Beach, FL).

Electrophoretic Mobility Shift Assays (EMSA). Prior to use, Cy3-labeled oligonucleotides (Figure S1) were diluted to 100 nM in TE Buffer (10 mM TRIS pH 7.5, 1 mM EDTA) and denatured at 95 °C for 10 minutes before being snap cooled on ice for 5 minutes. The oligonucleotides were then diluted to 50 nM in PRC2 binding buffer (50 mM Tris pH 7.4, 100 mM KCl, 2.5 mM $MgCl_2$, 0.1 mM $ZnCl_2$, 2 mM BME, 0.1 mg/mL BSA, and 5% glycerol) and allowed to fold at 37 °C for 30 minutes. In some instances, the KCl was replaced with LiCl (see Figure S2). The oligonucleotides were then diluted to 2 nM into individual binding reactions (10 µL) containing PRC2 binding buffer and increasing concentrations of PRC2 (0.1–1000 nM). Binding reactions were carried out at 30 °C for 30 minutes and bound and unbound fractions were subsequently separated by 1% agarose gel electrophoresis (0.2 X TBE supplemented with 10 mM KOAc or LiOAc as indicated). Agarose gels were run at 4 °C for 75 minutes at 44V. The gels were visualized using GE Typhoon gel imager using the Cy3-emmision filter (excitation: 532 nm; PMT: 950 V) and quantified using ImageQuant TL software.

We found that the proximity of the Cy3 dye to the terminal guanosines within our G4 RNAs resulted in fluorescent quenching (~2.5-fold). However, upon PRC2-binding, an increased Cy3 emission was observed that we attributed to exclusion of the dye from

proximal guanosine residues. This phenomenon has been observed previously for G-rich sequences.[3,4] To account for this phenomenon in our calculations, we corrected all unbound fluorescent intensities by a factor equal to the maximum Cy3-signal as measured in the presence of saturating PRC2 divided by the fluorescence of unbound Cy3-RNA.

Circular Dichroism (CD) Spectroscopy. For CD experiments, oligonucleotides (9.8 μM) were folded as described above in a buffer containing 2 mM sodium phosphate (pH 7.0), 0.1 mM EDTA, and 100 mM of either KCl or LiCl as indicated. Data were obtained from a 450 μL sample in a quartz cuvette using an Applied Photophysics Chirascan spectrophotometer (Leatherhead, England) at 1 nm intervals from 220 to 370 nm. All data were collected at a constant temperature of 23 °C.

$(GGAA)_{10}$ Competition Assay. Complexes of PRC2 (100 nM) and Cy3-labeled $(GGAA)_{10}$ (10 nM) were pre-formed in PRC2 binding buffer as described for EMSAs (30 minutes at 30 °C). Competitive binding experiments were carried out by adding variable concentrations (10–1300 nM) of unlabeled D-$(GGAA)_{10}$ competitor to the pre-formed PRC2-Cy3-L-$(GGAA)_{10}$ complexes (or vise versa), and the reaction was allowed to proceed for 30 minutes at 30 °C. Bound versus unbound fractions were subsequently separated by 1% agarose gel electrophoresis (0.2 × TBE supplemented with 10 mM KOA) and quantified as described above.

HOTAIR-binding and Competition Assay. A DNA fragment representing the first 300 nt of HOTAIR (HOTAIR-300) was prepared via PCR assembly using gBlocks Gene Fragments (IDT; Coralville, IA). The resulting DNA was added directly into a 100 μL transcription reaction containing 10 U/μL T7 RNA polymerase, 0.001 U/μL Inorganic pyrophosphatase (IPP), 25 mM $MgCl_2$, 2 mM spermidine, 10 mM DTT, 40 mM Tris (pH 7.9), and 5 mM of each of the four NTPs, where 5-aminoallyl-UTP (Thermo Fisher Scientific, Waltham, MA) was supplemented in the transcription reaction at 0.5 mM. The

reaction mixture was incubated at 37 °C for 2 hours, then enzymes, DNA, and unincorporated NTPs were removed using a Quick-RNA Mini Prep Plus Kit (Zymo Research, Irvine, California) and pure HOTAIR RNA was obtained in 1x TE buffer. We then used the internally positioned amine functional groups (on the 5-aminoallyl-UTP) to couple a Cy5 NHS-ester (Lumiprobe Life Science Solutions, Hallendale Beach, FL) using the provided procedure. For the competition experiments, HOTAIR-PRC2 complexes (25 and 250 nM, respectively) were pre-formed in PRC2 binding buffer as described for EMSAs, and unlabeled $(GGAA)_{10}$ or $(A)_{40}$ competitor RNA was added in 3-fold increments from 1 nM to 3 uM. (Figures 3a and S7). Bound versus unbound fractions were subsequently separated by 1% agarose gel electrophoresis (0.2 × TBE supplemented with 10 mM KOA) and quantified as described above.

Assembly and Reconstitution of Cy5-labeled Oligonucleosome Arrays. Human histone proteins were expressed and purified as described previously (Banerjee *et. al*)[2] and the Cy5-labelled nucleosome array was assembled using our recently published "plug and play" approach.[2] Briefly, we utilized two internally positioned nicking endonuclease sites (Nt. BstNBI) within the fifth 601 unit (N5) of the 12×601 array (Figure S8) to generate two single-stranded breaks flanking a region of 28 nucleotides (nt). The dual-nicked 12×601 array DNA was then mixed with 20-fold excess of a Cy5-labelled (internally) oligonucleotide insert consisting of a sequence identical to the 28 nt fragment generated by the Nt. BstNBI nicking endonuclease. The mixture was then heated at 80°C for 20 minutes before being cooled to room temperature at -1°C/min. Following the annealing step (~1 hour), T4 DNA ligase and ATP (2 mM final concentration) were added to the mixture to reseal the nicks and generate an intact DNA strand. The efficiency of the exchange process (nicking, insertion, and ligation) was carefully monitored in order to ensure complete insertion of the modified oligonucleotide (see Figure S8b). Oligonucleosome reconstitutions were carried out via salt dialysis and the arrays were purified by selective Mg^{2+}-induced precipitation. Nucleosome saturation was confirmed by selective restriction enzyme digestion (Figure S8d).

Chromatin Binding and Competition Assay. In order to confirm that PRC2 was capable of binding the Cy5-labled oligonucleosome array, we performed an EMSA using the same conditions described for the $(GGAA)_{10}$ binding experiments (Figure S8e). Using 8 nM arrays, we found that 1:1 PRC2-chromatin complexes were initiated at PRC2 concentrations <100 nM (Figure S8e). At higher concentrations of PRC2 (>500 nM), we observed non-stoichiometric binding by PRC2, resulting higher molecular weight complexes that migrated significantly slower than unbound chromatin when analyzed by agarose gel electrophoresis (0.7%). For the competition assay, we chose a concentration of PRC2 that resulted in a clearly visible interaction by gel electrophoresis (1000 nM PRC2) and generated PRC2-chromatin complexes by incubating PRC2 with the Cy5-labeled array (8 nM) at 30 °C for 30 minutes in PRC2 binding buffer. We then added unlabeled competitor $(GGAA)_{10}$ RNA in 3-fold increments from 1 nM up to 3 uM and analyzed the results by 0.7% agarose gel electrophoresis (0.2 × TBE, 10 mM KOAc, 44 V, 5 hours) (Figure 3b & S9).

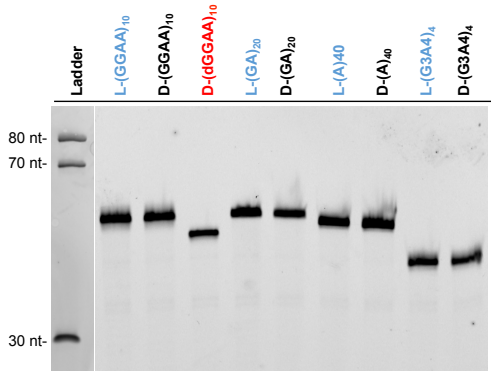
S2. Supplemental Figures

Figure S1

a

	Name	Sequence
D-RNA	D-$(GGAA)_{10}$	5′-Cy3-AA-GGAA-GGAA-GGAA-GGAA-GGAA-GGAA-GGAA-GGAA-GGAA-GGAA
D-RNA	D-$(GA)_{20}$	5′-Cy3-AA-GAGA-GAGA-GAGA-GAGA-GAGA-GAGA-GAGA-GAGA-GAGA-GAGA
D-RNA	D-$(A)_{40}$	5′-Cy3-AA-AAAA-AAAA-AAAA-AAAA-AAAA-AAAA-AAAA-AAAA-AAAA-AAAA
D-RNA	D-$(G3A4)_4$	5′-Cy3-AA-AAA-GGGAAAA-GGGAAAA-GGGAAAA-GGAAAA
DNA	D-$(dGGAA)_{10}$	5′-Cy3-dAdA-dGdGdAdA-dGdGdAdA-dGdGdAdA-dGdGdAdA-dGdGdAdA-dGdGdAdA-dGdGdAdA-dGdGdAdA-dGdGdAdA-dGdGdAdA
L-RNA	L-$(GGAA)_{10}$	5′-Cy3-AA-GGAA-GGAA-GGAA-GGAA-GGAA-GGAA-GGAA-GGAA-GGAA
L-RNA	L-$(GA)_{20}$	5′-Cy3-AA-GAGA-GAGA-GAGA-GAGA-GAGA-GAGA-GAGA-GAGA-GAGA
L-RNA	L-$(A)_{40}$	5′-Cy3-AA-AAAA-AAAA-AAAA-AAAA-AAAA-AAAA-AAAA-AAAA-AAAA
L-RNA	L-$(G3A4)_4$	5′-Cy3-AA-AAA-GGGAAAA-GGGAAAA-GGGAAAA-GGAAAA

b



c

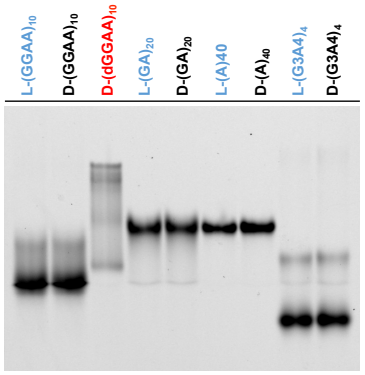


Figure S1. Oligonucleotides used in this work. (a) Sequences of oligonucleotides used for binding EMSAs and competition experiments. Terminal D-deoxyribose residues (D-dA) on the RNA strands are underlined. (b) Denaturing PAGE analysis of the Cy3-labeled oligonucleotides presented in S1a (10%, 29:1 acrylamide:bisacrylamide). (c) Native PAGE analysis (10%, 29:1 acrylamide:bisacrylamide) of the same oligonucleotides in panel (b). The running buffer (1 × TBE) was supplemented with 10 mM KOAc. The increased electrophoretic mobility of $(GGAA)_{10}$ and $(G3A4)_4$ relative to $(A)_{40}$ is indicative of the G4 structure formation by these oligonucleotides in the presence of K^+.

Figure S2

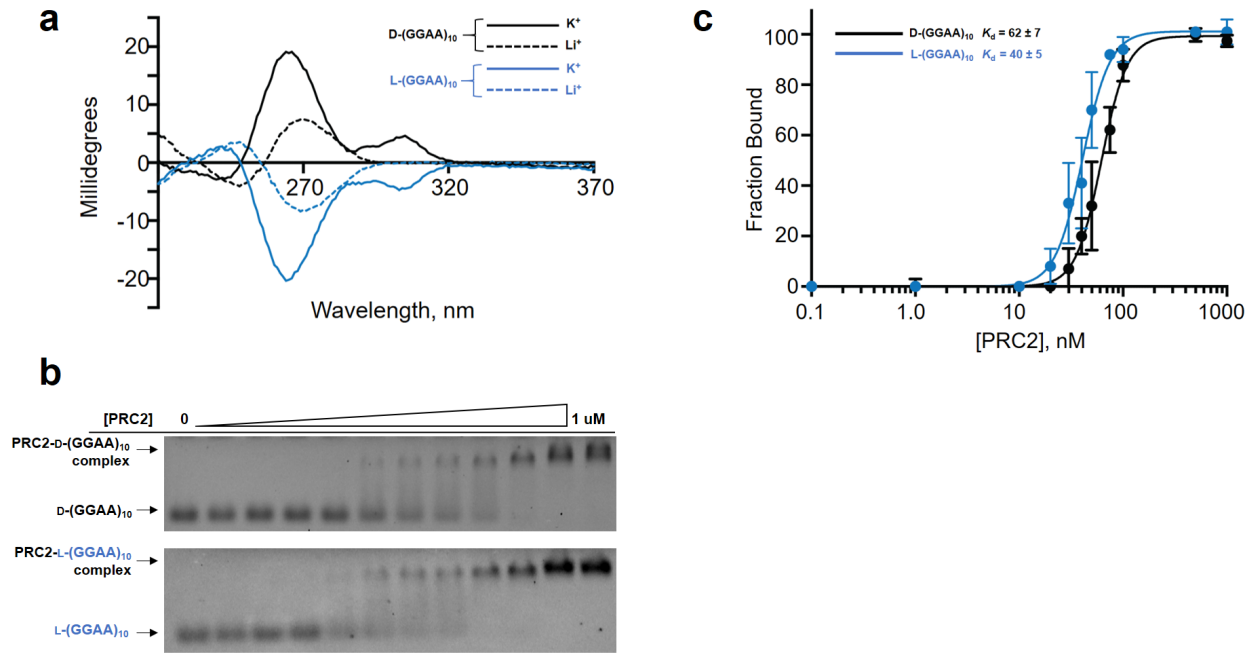


Figure S2. PRC2 binds both D- and L-$(GGAA)_{10}$ RNAs in the presence Li^+. (a) CD spectra of D- and L-$(GGAA)_{10}$ in the presence of either 100 mM KCl or LiCl. (b) Representative EMSA gels (1% agarose, 0.2 × TBE supplemented with 10 mM LiOAc) of $(GGAA)_{10}$ binding to PRC2 (0–1 μM) in the presence of Li^+. Binding conditions were the same as described in Figure 1b (main text), except that the KCl was replaced with LiCl. (c) Saturation plot for binding of either D- or L-$(GGAA)_{10}$ to PRC2 in the presence of Li^+. Error bars show SD (n= 3).

Figure S3

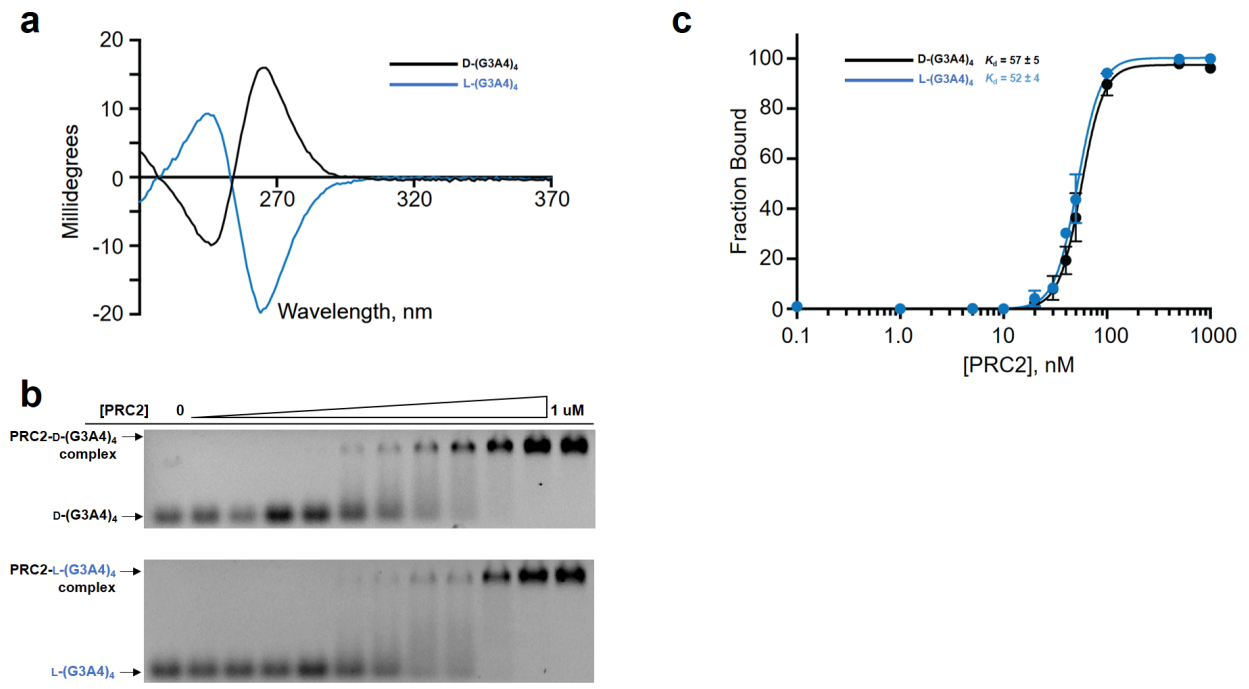


Figure S3. PRC2 binds similarly to both D- and L-$(G3A4)_4$ G4 RNAs. (a) CD spectra of D- and L-$(G3A4)_4$. (b) Representative EMSA gels (1% agarose, 0.2 × TBE supplemented with 10 mM KOAc) of $(GGAA)_{10}$ binding to PRC2 (0–1 μM). Binding conditions were the same as described in Figure 1b (main text). (c) Saturation plot for binding of either D- or L-$(G3A4)_4$ to PRC2. Error bars show SD (n= 3).

Figure S4.

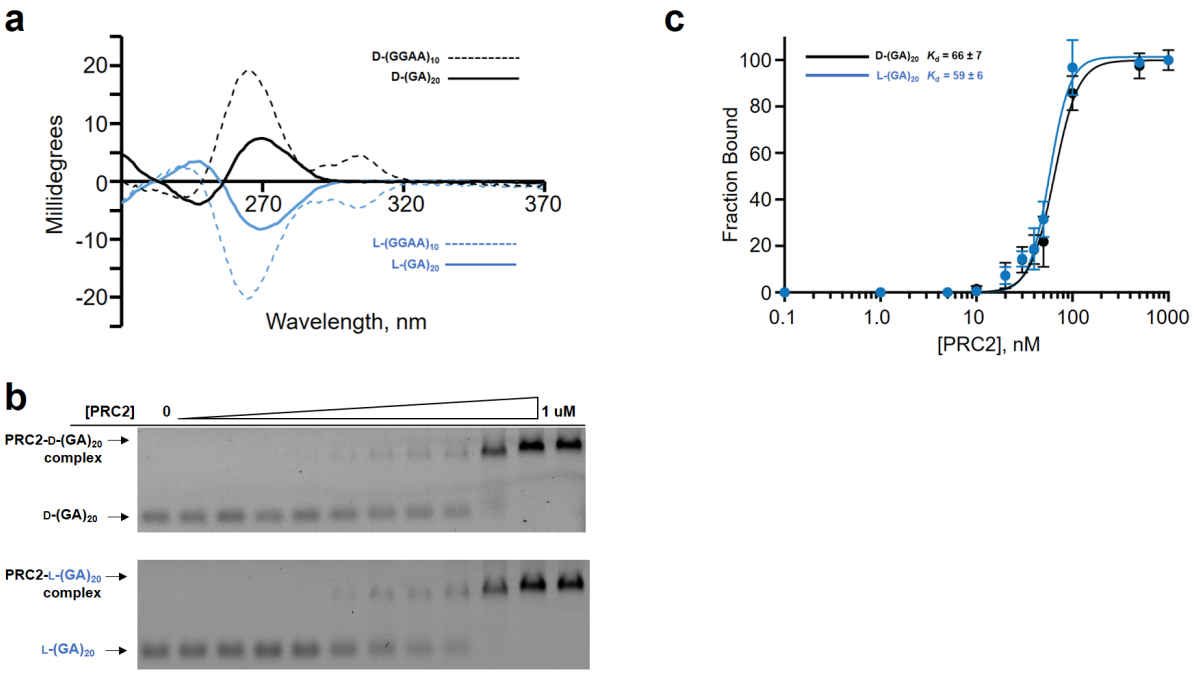


Figure S4. PRC2 binds similarly to both D- and L-$(GA)_{20}$ RNAs. (a) CD spectra of D- and L-$(GA)_{20}$. The spectral data for both D- and L-$(GGAA)_{10}$ RNAs is overlaid for comparison. (b) Representative EMSA gels (1% agarose, 0.2 × TBE supplemented with 10 mM KOAc) of D-$(GA)_{20}$ and L-$(GA)_{20}$ binding to PRC2 (0–1 μM). Binding conditions were the same as described in Figure 1b (main text). (c) Saturation plot for binding of either D- or L- $(GA)_{20}$. Error bars show SD (n= 3).

Figure S5

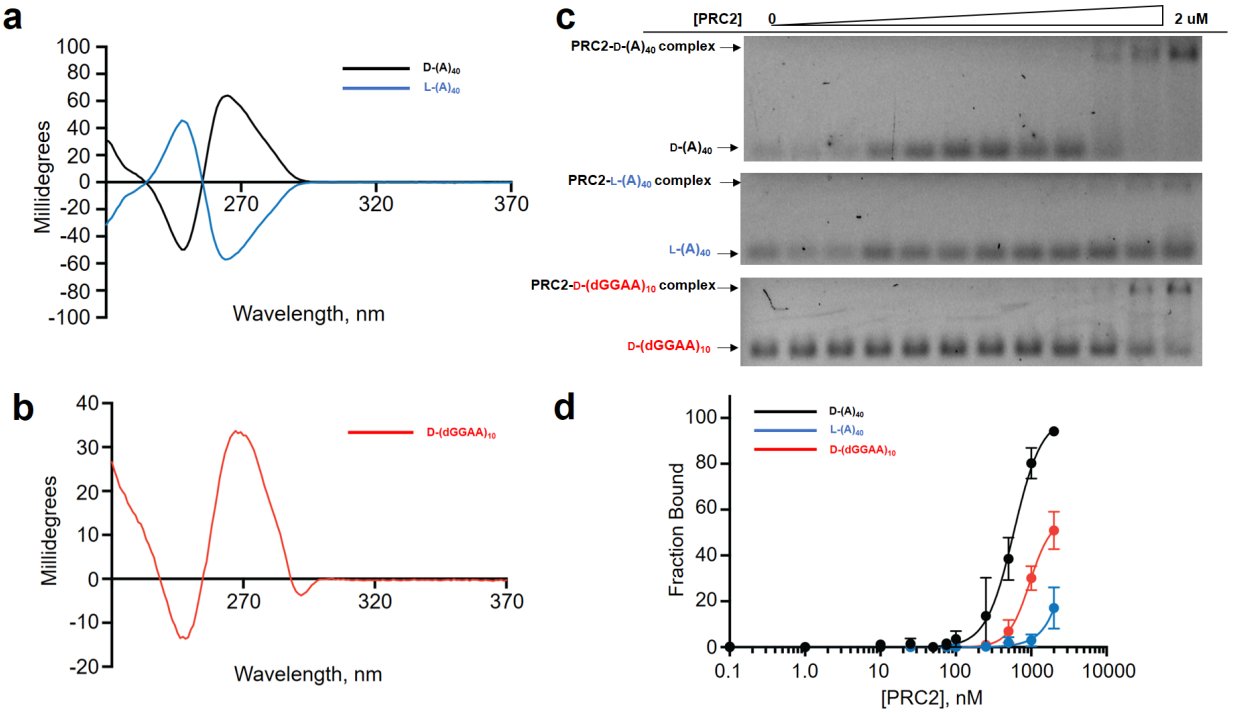


Figure S5. PRC2 binds weakly to $(A)_{40}$ and D-$(dGGAA)_{10}$. (a,b) CD spectra of both enantiomers of $(A)_{40}$ and the D-$(dGGAA)_{10}$, respectively. c) Representative EMSA gels (1% agarose, 0.2 × TBE supplemented with 10 mM KOAc) of D-$(A)_{40}$, L-$(A)_{40}$, and D-$(dGGAA)_{10}$ binding to PRC2 (0–2 μM). Binding conditions were the same as described in Figure 1b (main text). (d) Saturation plot for binding of D-$(A)_{40}$, L-$(A)_{40}$, and D-$(dGGAA)_{10}$ to PRC2. Error bars show SD (n= 3).

Figure S6

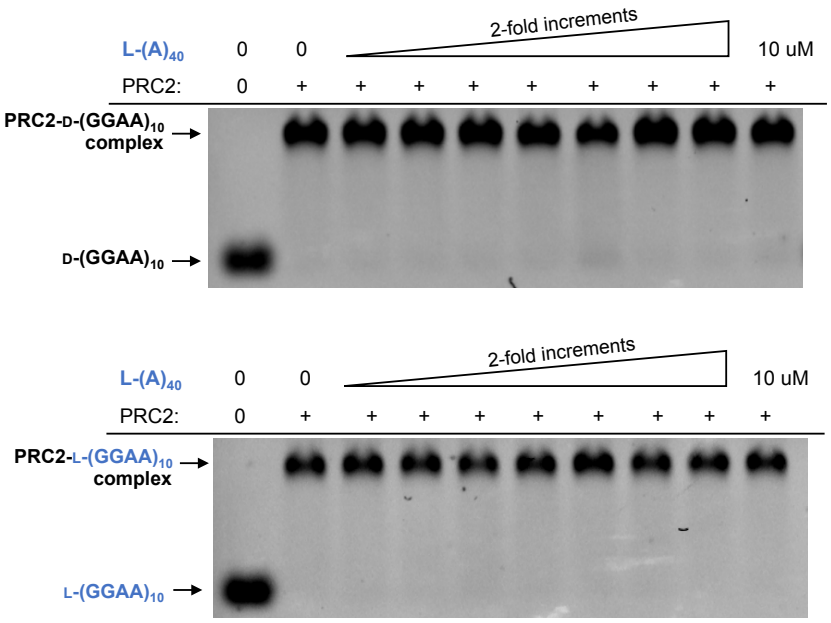


Figure S6. Competitive binding experiments for L-$(A)_{40}$ versus pre-formed PRC2-$(GGAA)_{10}$ complexes. Initial PRC2-RNA complexes were prepared using the binding conditions described in Figure 1b (main text), and competitor $(A)_{40}$ RNA was added in 2-fold increments from 80 nM to 10 uM.

Figure S7

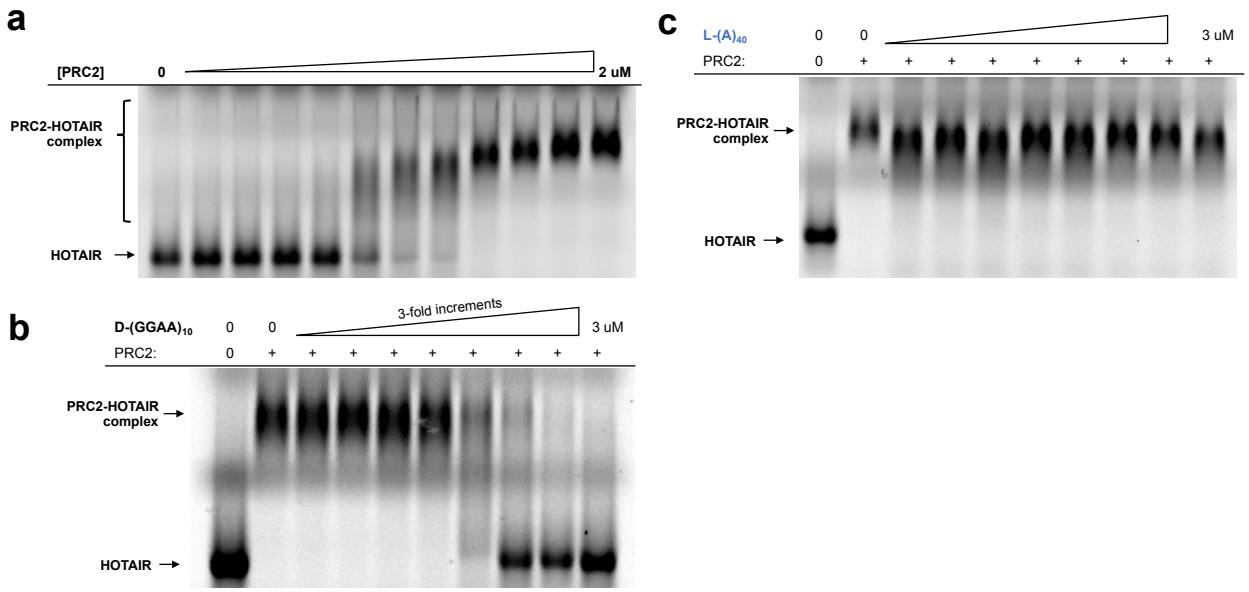


Figure S7. Disruption of PRC2-HOTAIR complexes. (a) EMSA gel (1% agarose, 0.2 × TBE supplemented with 10 mM KOAc) of PRC2 (0.1–2000 nM) binding Cy5-labeled HOTAIR-300 (25 nM). Binding conditions were the same as described in Figure 1b (main text). (b) D-$(GGAA)_{10}$ is able to outcompete HOTAIR-300 for binding to PRC2. Initial PRC2-HOTAIR complexes were prepared using the binding conditions described in Figure 1b (main text) except where indicated (above). Data points represent a 3-fold titration of unlabeled D-$(GGAA)_{10}$ ranging from 0.001–3 μM. (c) L-$(A)_{40}$ is unable to compete with HOTAIR for binding to PRC2.

Figure S8

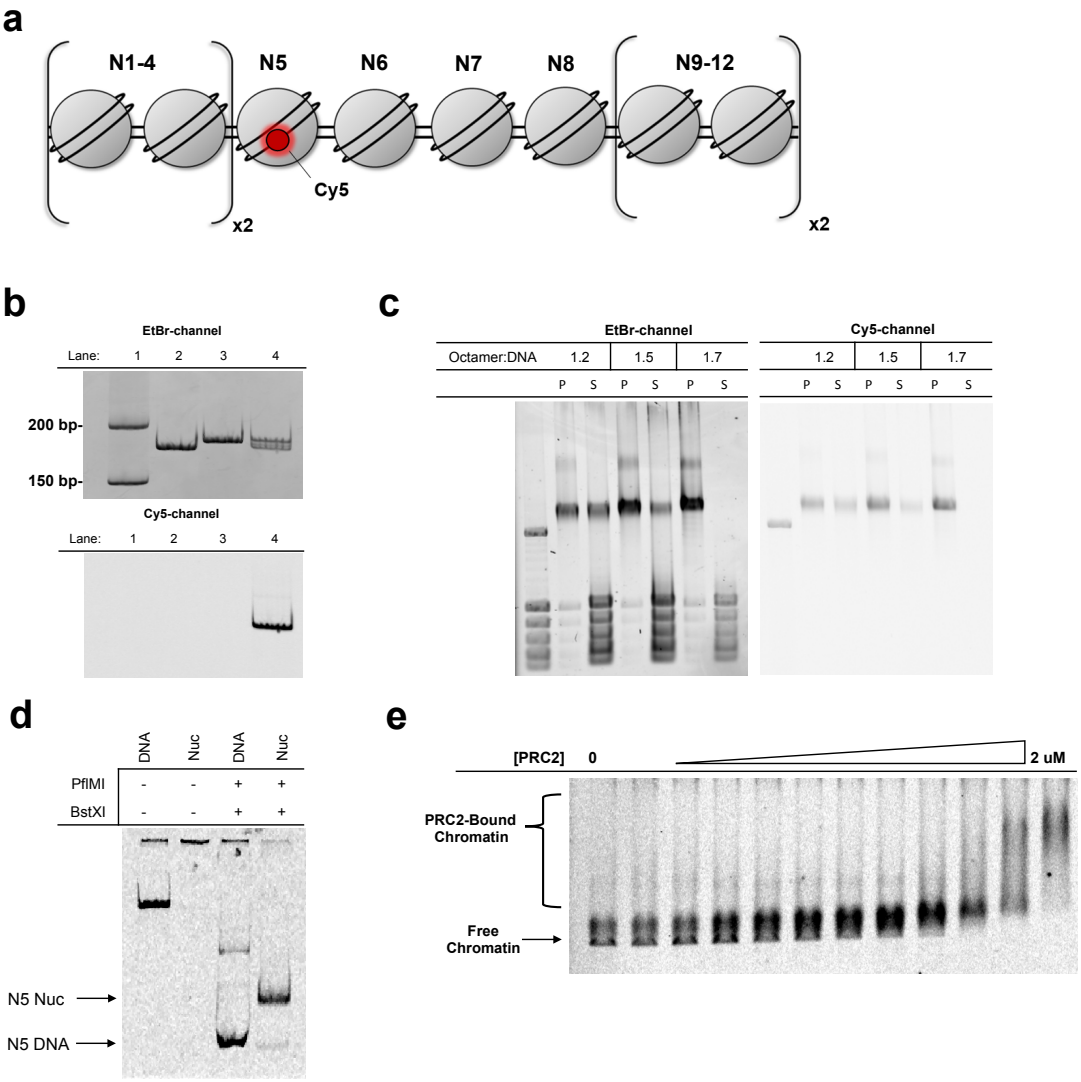


Figure S8. 12-mer oligonucleosome array assembly. (a) Schematic of Cy5-labeled 12-mer oligonucleosome array employed in the PRC2 binding and competition assays. A single Cy5 dye is positioned within the fifth nucleosome unit (N5). (b) Insertion of the Cy5 dye containing oligonucleotide was confirmed by 10% native PAGE (29:1, acrylamide:bisacrylamide). Lane 1, ladder; lane 2, unmodified N5 DNA fragment; lane 3, nicked N5 DNA fragment; lane 4, Cy5-labeled N5 DNA fragment following the strand exchange process. (c) Reconstitution of Cy5-labeled oligonucleosome arrays. Agarose gel (0.6%, 0.2 × TBE) analysis of Mg^{2+}-induced precipitation of reconstitutions for several histone octamer:DNA ratios visualized with different fluorescent channels (see

figure heading). Aliquots from re-suspended nucleosome pellets (P) and the supernatant (S) following Mg^{2+} precipitation are indicated for each octamer:DNA ratio employed. (d) Restriction enzyme digest analysis of the Cy5-containing N5 fragment (5%, 59:1 acrylamide:bisacrylamide). Both naked (DNA) and reconstituted (Nuc) 12-mer arrays were digested similarly and their corresponding N5 fragments analyzed side-by-side. (e) EMSA gel (0.5% agarose, 0.2 × TBE supplemented with 10 mM KOAc) of PRC2 (0.1–2000 nM) binding to the Cy5-labled array (8 nM). Binding conditions were the same as described in Figure 1b (main text) except where indicated.

Figure S9

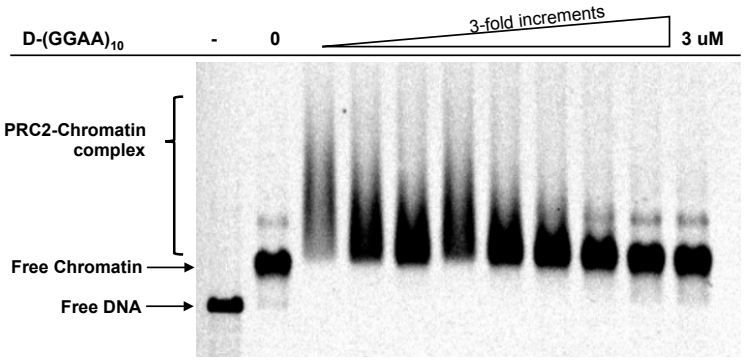
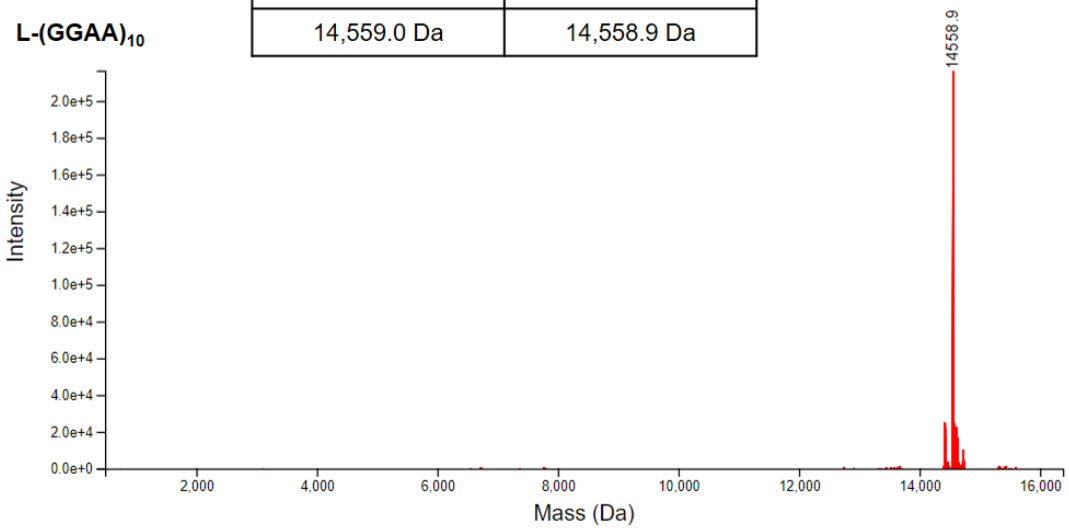


Figure S9. Pre-formed complexes between PRC2 (1000 nM) and a 12-mer oligonucleosome array (8 nM) are disrupted by D-$(GGAA)_{10}$. Initial PRC2-substrate complexes were prepared using the binding conditions described in Figure 1b (main text) except where indicated (above). Data points represent a 3-fold titration of unlabeled L-$(GGAA)_{10}$ ranging from 0.001–3 μM.

S3. Mass Analysis of Synthetic Oligonucleotides

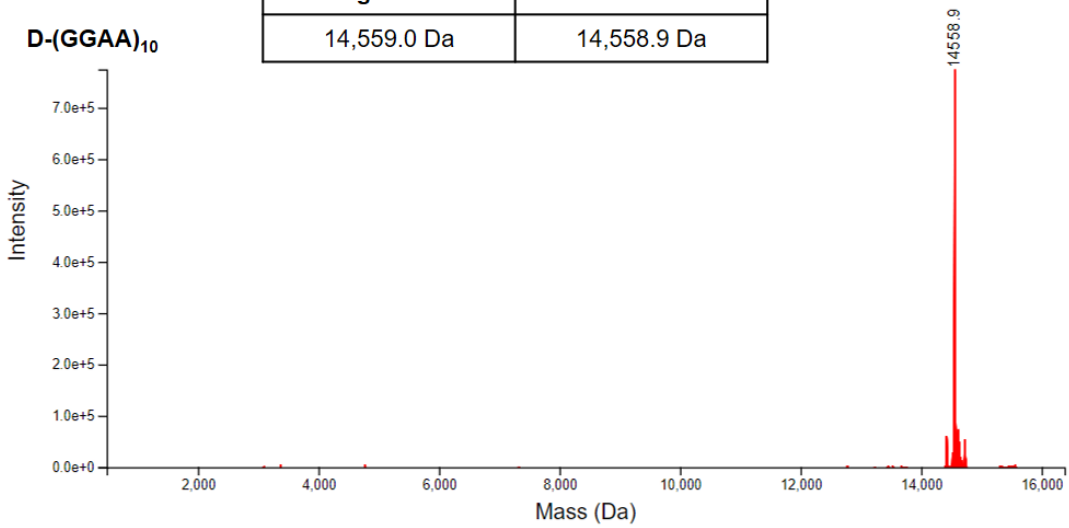
L-(GGAA)$_{10}$

Target Mass	Observed Mass
14,559.0 Da	14,558.9 Da



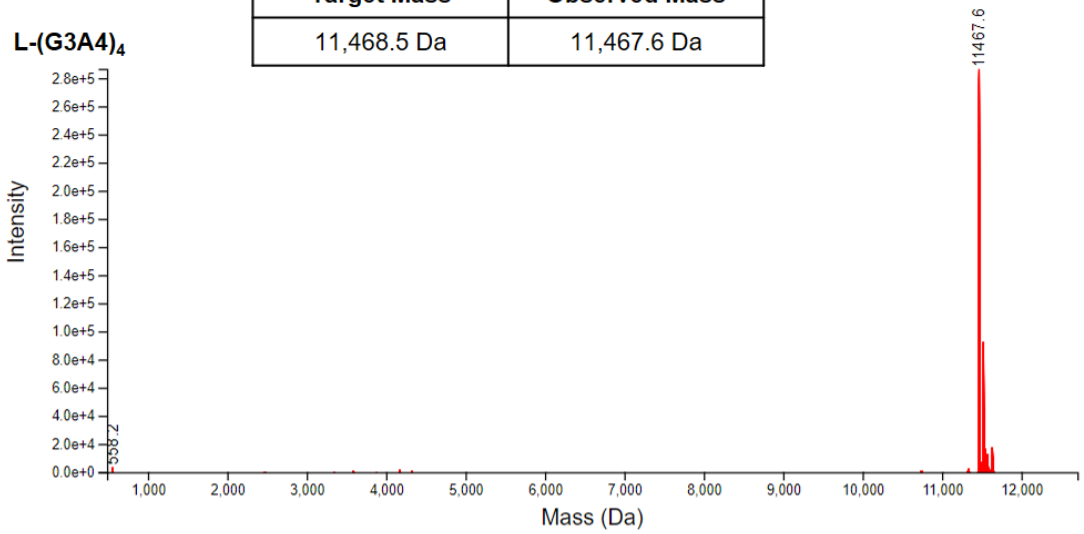
D-(GGAA)$_{10}$

Target Mass	Observed Mass
14,559.0 Da	14,558.9 Da



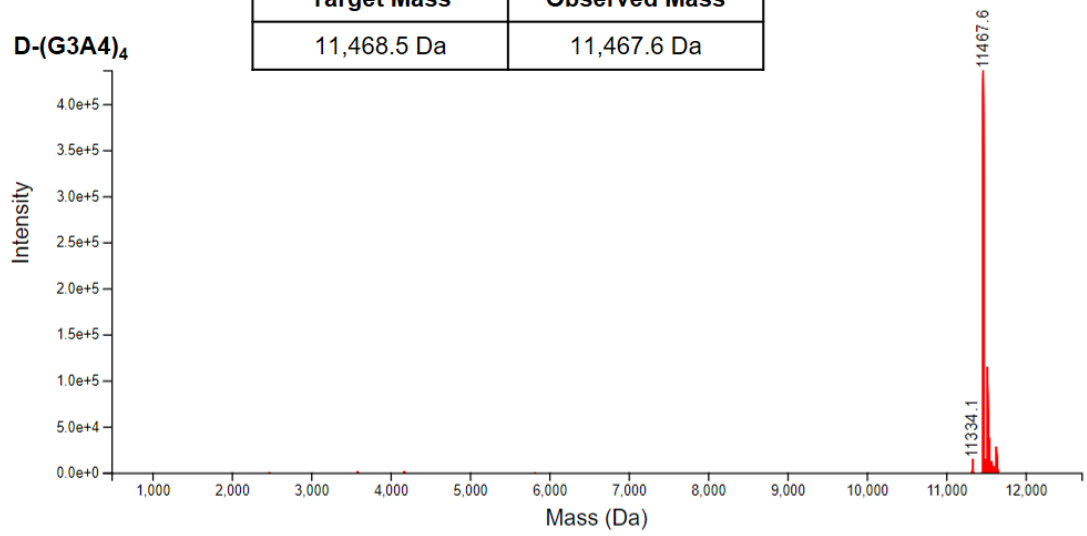
L-(G3A4)$_4$

Target Mass	Observed Mass
11,468.5 Da	11,467.6 Da



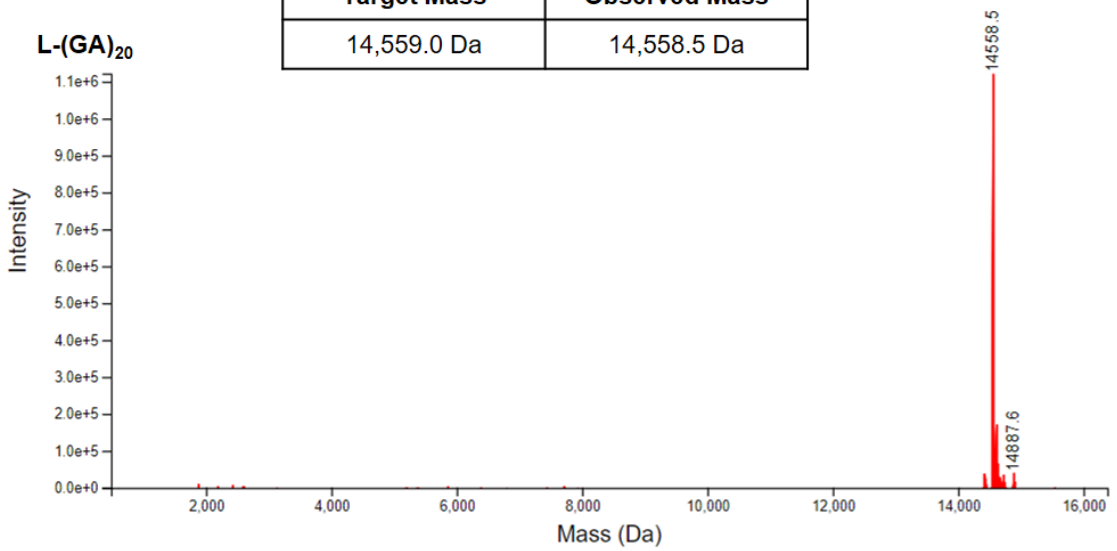
D-(G3A4)$_4$

Target Mass	Observed Mass
11,468.5 Da	11,467.6 Da



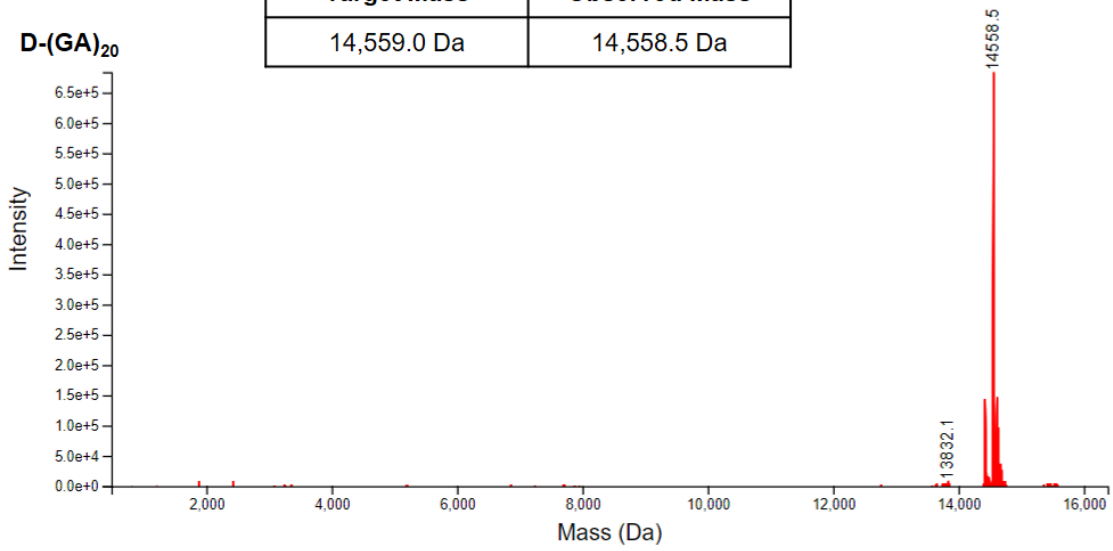
Target Mass	Observed Mass
14,559.0 Da	14,558.5 Da

L-(GA)$_{20}$



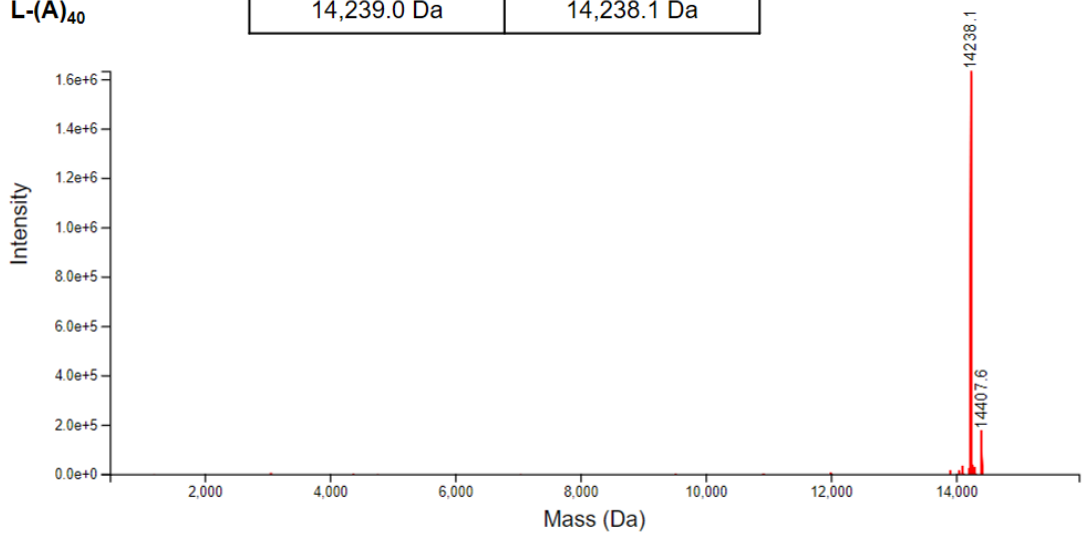
Target Mass	Observed Mass
14,559.0 Da	14,558.5 Da

D-(GA)$_{20}$



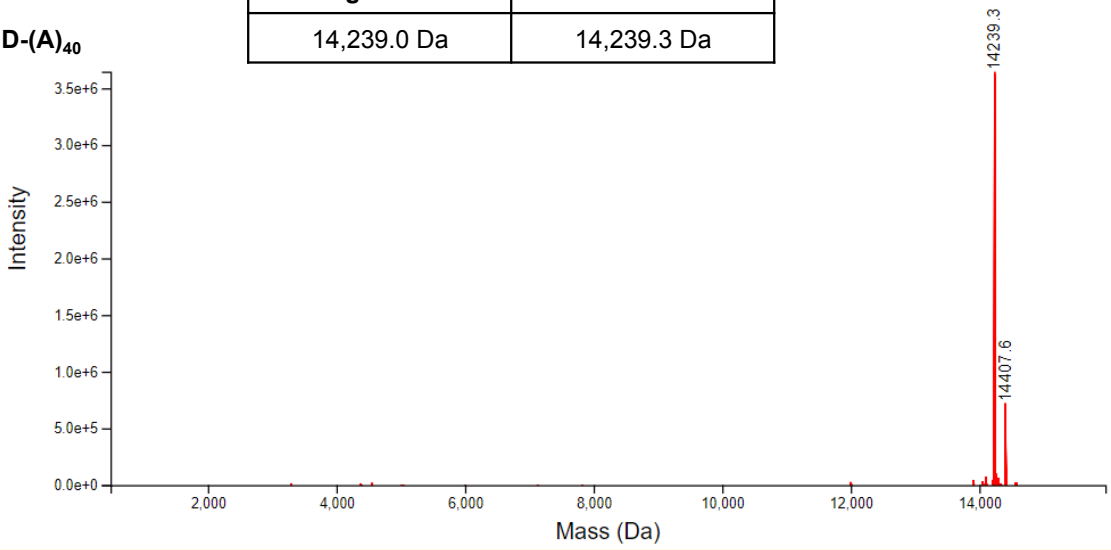
L-$(A)_{40}$

Target Mass	Observed Mass
14,239.0 Da	14,238.1 Da



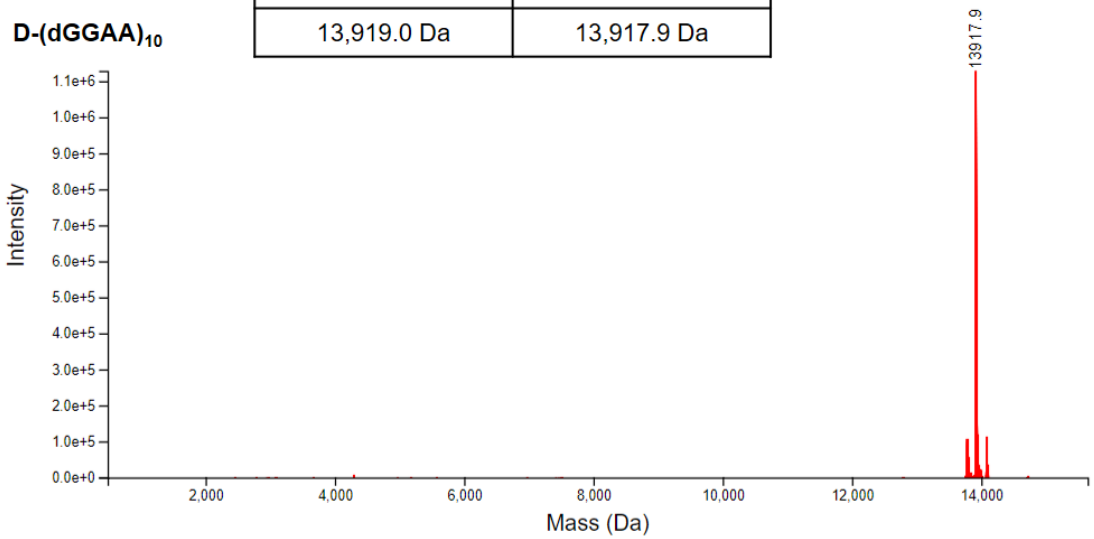
D-$(A)_{40}$

Target Mass	Observed Mass
14,239.0 Da	14,239.3 Da



D-(dGGAA)$_{10}$

Target Mass	Observed Mass
13,919.0 Da	13,917.9 Da



S4. Supplemental References

(1) Wang, X.; Goodrich, K. J.; Gooding, A. R.; Naeem, H.; Archer, S.; Paucek, R. D.; Youmans, D. T.; Cech, T. R.; Davidovich, C. *Mol. Cell* **2017**, *65*, 1056-1067.

(2) Banerjee, D. R.; Deckard III, C. E.; Elinski, M. B.; Buzbee, M. L.; Wang, W. W.; Batteas, J. D.; Sczepanski, J. T. *J. Am. Chem. Soc.* **2018**, *146*, 8260–8267.

(3) Nazarenko, I., Pires, R., Lowe, B., Obaidy, M., Rashtchian, A. *Nucleic Acids Res.* **2002**, *30*, 2089-2195.

(4) Mergny, J.L.; Maurizot, J.C. *Chembiochem* **2002**, *2*, 124-132.