

**PHS PUBLIC ACCESS**

Author manuscript

Science. Author manuscript; available in PMC 2016 May 20.

Published in final edited form as:

Science. 2015 November 20; 350(6263): 978–981. doi:10.1126/science.aad3346.

Transcription factor trapping by RNA in gene regulatory elements

Alla A. Sigova¹, Brian J. Abraham¹, Xiong Ji¹, Benoit Molinie⁴, Nancy M. Hannett¹, Yang Eric Guo¹, Mohini Jangi^{2,3,†}, Cosmas C. Giallourakis⁴, Phillip A. Sharp^{2,3}, and Richard A. Young^{1,2,*}

¹Whitehead Institute for Biomedical Research, Cambridge, MA, 02142

²Department of Biology, Massachusetts Institute of Technology, Cambridge, MA, 02142

³David H. Koch Institute for Integrative Cancer Research, Cambridge, MA, 02140

⁴Massachusetts General Hospital, Harvard Medical School, Boston, MA, 02114

Abstract

Transcription factors (TFs) bind specific sequences in promoter-proximal and distal DNA elements in order to regulate gene transcription. RNA is transcribed from both of these DNA elements, and some DNA-binding TFs bind RNA. Hence, RNA transcribed from regulatory elements may contribute to stable TF occupancy at these sites. We show that the ubiquitously expressed TF YY1 binds to both gene regulatory elements and also to their associated RNA species genome-wide. Reduced transcription of regulatory elements diminishes YY1 occupancy whereas artificial tethering of RNA enhances YY1 occupancy at these elements. We propose that RNA makes a modest but important contribution to the maintenance of certain TFs at gene regulatory elements and suggest that transcription of regulatory elements produces a positive feedback loop that contributes to the stability of gene expression programs.

Active promoters and enhancer elements are transcribed bi-directionally (Fig. 1A) (1–3). Although various models have been proposed for the roles of RNA species produced from these regulatory elements, their functions are not fully understood (4–13). Evidence that some DNA-binding transcription factors (TFs) also bind RNA (14, 15) led us to consider the possibility that there might be a direct and general role for promoter-proximal and distal enhancer RNA in the binding and maintenance of TFs at regulatory elements.

We sequenced nascent transcripts (GRO-seq) in murine embryonic stem cells (ESCs) at great depth, which confirmed that active promoters and enhancer elements are generally transcribed bi-directionally (Fig. 1B, fig. S1A, table S1). We then focused our studies on the

*Correspondence to: young@wi.mit.edu.

†Current address: Biogen, Cambridge, MA, 02142

Supplementary materials:

Materials and Methods

Figs. S1 to S16

Tables S1 to S3

References (25–52)

TF Yin-Yang 1 (YY1) because it is ubiquitously expressed in mammalian cells, plays key roles in normal development, and can bind RNA species *in vitro* (15, 16). ChIP-seq analysis in ESCs revealed that YY1 binds to both active enhancers and promoters, with some preference for promoters (Fig. 1C, and D, fig. S1, table S2). In contrast, the pluripotency TF OCT4 preferentially occupies enhancers (fig. S1B). Consistent with this, YY1 sequence motifs were enriched at promoters, whereas OCT4 motifs were enriched at enhancers (fig. S1B). Neither YY1 nor OCT4 occupied the promoter-proximal sequences of inactive genes (fig. S2). These results establish that YY1 generally occupies active enhancer and promoter-proximal elements in ESCs.

We next investigated YY1 binding to RNA *in vivo* by using CLIP-seq in ESCs (fig. S3, S4, table S3). The results showed that YY1 binds RNA species at the active enhancer and promoter regions where it is bound to DNA (Fig. 1, C and D, fig. S1C). At promoters, YY1 preferentially occupied RNA downstream rather than upstream of transcription start sites (fig. S1B), consistent with YY1 motif distribution and evidence that upstream ncRNA is unstable (3, 17, 18). In similar experiments with OCT4, significant levels of RNA binding were not observed (fig. S5). These results suggest that YY1 generally binds to RNA species transcribed from enhancers and promoters *in vivo*.

The DNA and RNA binding properties of YY1 were further investigated *in vitro* (Fig. 2, fig. S6 to S8). Recombinant murine YY1 protein bound both DNA and RNA probes in electrophoretic mobility shift assays (EMSA), showing higher affinity for DNA than RNA. There was variation in the affinity of YY1 for different RNA sequences (fig. S8). The four YY1 zinc-fingers can bind DNA (19), but the portion of YY1 that interacts with RNA is unknown. The zinc-finger –containing C-terminal region and the N-terminal region of YY1 were purified and their DNA and RNA binding properties were further investigated (fig. S9). The zinc-finger region of YY1 bound to DNA, but not to RNA, whereas the N-terminal region of YY1 bound to RNA (fig. S9). Furthermore, the DNA probe did not compete efficiently with the RNA probe for YY1 binding (fig. S7C, S8C). These results suggest that different regions of YY1 are responsible for binding to DNA and RNA.

The observation that YY1 binds to enhancer and promoter-proximal elements and to RNA transcribed from those regions led us to postulate that nascent RNA contributes to stable TF occupancy at these regulatory elements (Fig. 3A). If this model is correct, then reduced levels of nascent RNA at promoters and enhancers might lead to reduced YY1 occupancy at these sites. We briefly inhibited transcription elongation with the reversible inhibitor D-rybofuranosylbenzimidazole (DRB) to reduce RNA levels at promoters and enhancers without causing changes in the steady-state levels of YY1 (fig. S10, S11). DRB treatment reduced transcription at promoters and enhancers and this caused small but significant decrease in the levels of YY1 at these regions (fig. S10). Super-enhancers are clusters of enhancers that are highly transcribed (20), and DRB treatment had a profound effect on transcription at these sites (fig. S10). Similar results were observed with additional inhibitors (fig. S10). When transcription was allowed to resume after DRB removal, the levels of YY1 increased at promoters and enhancers (Fig. 3B, fig. S10A). These results suggest that nascent RNA produced at promoters and enhancers contributes to YY1 binding to these elements.

The exosome reduces the levels of enhancer RNAs once they are released from Pol II (degradation is 3' to 5') (21), so knockdown of an exosome component will cause an increase in untethered enhancer RNA, which might titrate some YY1 away from enhancers. Indeed, exosome knockdown led to increased steady state levels of enhancer RNAs and a decrease in the levels of YY1 bound to enhancers (Fig. 3C, fig. S12). These results are consistent with the model that YY1 binding to DNA is stabilized by binding to nascent RNA.

If YY1 binding to DNA is stabilized by its binding to RNA, then RNase treatment of chromatin should reduce YY1 occupancy. Chromatin was extracted from ESC nuclei and the levels of YY1 in the chromatin preparation were compared with and without RNase A treatment (Fig. 3D). The results show that the levels of YY1 bound to chromatin were significantly decreased when the chromatin preparation was treated with RNase, consistent with the idea that RNA contributes to the stability of YY1 in chromatin.

To test the idea that RNA near regulatory elements can contribute to stable TF occupancy *in vivo*, we tethered RNA in the vicinity of YY1 binding sites at six different enhancers in ESCs using the CRISPR/Cas9 system and determined whether the tethered RNA increases the occupancy of YY1 at these enhancers (Fig. 4). We generated stable murine ESC lines expressing both the catalytically inactive form of bacterial endonuclease Cas9 (dCas9) and a fusion RNA composed of guide RNA (sgRNA), tracrRNA, and a 60-nt RNA derived from the promoter sequence of *Arid1a* compatible with YY1 binding *in vitro* (fig. S8). For controls, stable cell lines were created that express dCas9 and sgRNA fused to tracrRNA for the six enhancers. Tethering the *Arid1a* RNA at each enhancer led to increased binding of YY1 to the targeted enhancer as measured by ChIP-qPCR (Fig. 4B). This elevation in YY1 binding was specific to the targeted locus and the sequence of tethered RNA as there was no observable increase in YY1 binding at the enhancers not targeted in the same cells (Fig. 4B) or targeted with tethered RNA not compatible with YY1 binding *in vitro* (fig. S13). These results show that RNA tethered near regulatory elements *in vivo* can enhance the level of YY1 occupancy at these elements.

To corroborate the *in vivo* RNA tethering results, a competition EMSA was used to test whether tethered RNA increases the apparent binding affinity of YY1 to its motif in DNA (fig. S14, S15). A short 30-bp labeled DNA probe containing a consensus YY1 binding motif was incubated with recombinant murine YY1 protein in the presence of increasing concentrations of cold competitor DNA with tethered or untethered RNA, and the amount of radiolabeled DNA that remained bound was quantified (fig. S15). This analysis revealed that DNA containing tethered RNA outcompetes the DNA without tethered RNA for YY1 binding. These results indicate that tethering RNA near the YY1 binding motif in DNA leads to increased binding of YY1 to DNA *in vitro*.

In summary, our results are consistent with the proposal that RNA enhances the level of YY1 occupancy at active enhancer and promoter-proximal regulatory elements (Fig. 3A). We suggest that nascent RNA produced in the vicinity of enhancer and promoter elements captures dissociating YY1 via relatively weak interactions, which allows this TF to rebind to nearby DNA sequences, thus creating a kinetic sink that increases YY1 occupancy on the

regulatory element. The observation that YY1 occupies active enhancers and promoters throughout the ESC genome where RNA is produced, coupled with evidence that YY1 is expressed in all mammalian cells, suggests that this model is general. There are additional DNA-binding TFs that can bind RNA (fig. S16) (14), so transcriptional control may generally involve a positive feedback loop, where YY1 and other TFs stimulate local transcription, and newly transcribed nascent RNA reinforces local TF occupancy. This model helps explain why TFs occupy only the small fraction of their consensus motifs in the mammalian genome where transcription is detected and suggests that bidirectional transcription of active enhancers and promoters evolved, in part, to facilitate trapping of TFs at specific regulatory elements. The model also suggests that transcription of regulatory elements produces a positive feedback loop that may contribute to the stability of gene expression programs in cells. The significance of this TF trapping mechanism in cellular regulation has yet to be established, but will be important for future study because much of disease-associated sequence variation occurs in enhancers (20, 22) and may thus affect both DNA and RNA sequences that interact with gene regulators.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

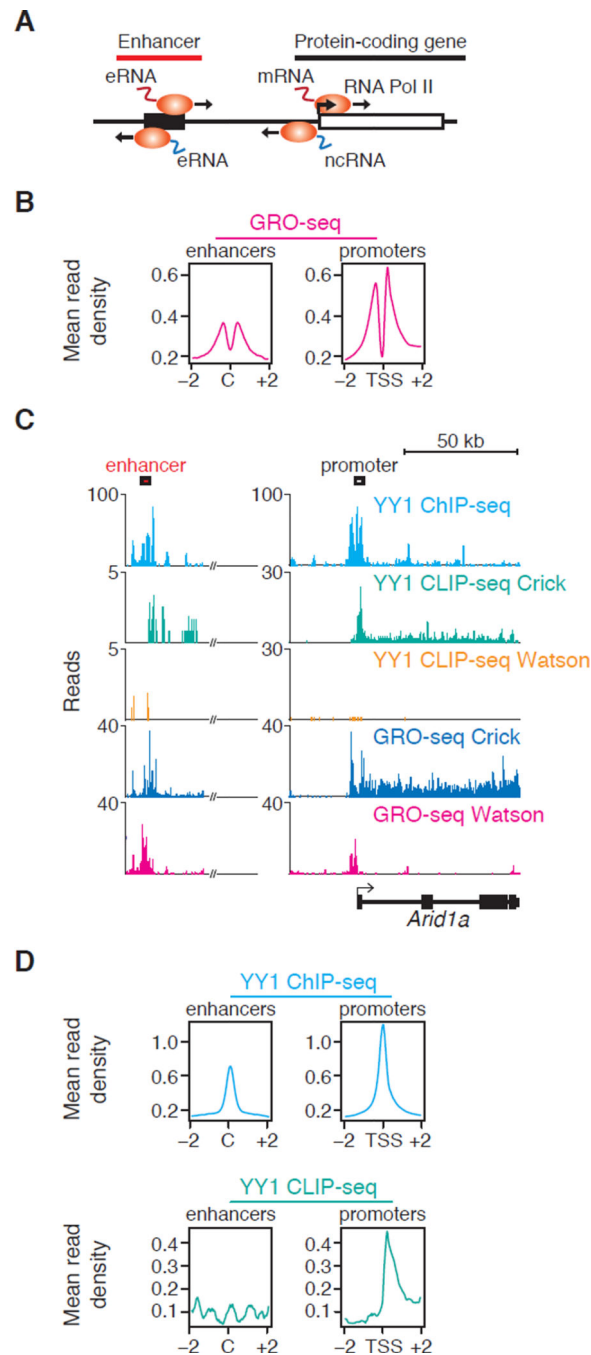
Acknowledgments

We thank D. Orlando, C. Lin, V. Saint-André, Z. P. Fan, L. Zhang, and A. Chiu for help with computational analysis and advice. The lentiviral dCas9 and sgRNA expression plasmids are available from Addgene under the Uniform Biological Material Transfer Agreement. Supported by NIH HG002668 (R.A.Y.), the Hope Funds for Cancer Research (B.J.A.), the Cancer Research Institute (Y.E.G), and Biogen (R.A.Y.). The Whitehead Institute intends to file a patent application that relates to transcription factor trapping by RNA. R.A.Y. is a founder of Syros Pharmaceuticals.

References

1. Core LJ, Waterfall JJ, Lis JT. *Science*. 2008; 322:1845–1848. [PubMed: 19056941]
2. Seila AC, et al. *Science*. 2008; 322:1849–1851. [PubMed: 19056940]
3. Sigova AA, et al. *Proc Natl Acad Sci U S A*. 2013; 110:2876–2881. [PubMed: 23382218]
4. Kim TK, et al. *Nature*. 2010; 465:182–187. [PubMed: 20393465]
5. Wang D, et al. *Nature*. 2011; 474:390–394. [PubMed: 21572438]
6. Melo CA, et al. *Mol Cell*. 2013; 49:524–535. [PubMed: 23273978]
7. Lai F, et al. *Nature*. 2013; 494:497–501. [PubMed: 23417068]
8. Lam MT, et al. *Nature*. 2013; 498:511–515. [PubMed: 23728303]
9. Li W, et al. *Nature*. 2013; 498:516–520. [PubMed: 23728302]
10. Kaikkonen MU, et al. *Mol Cell*. 2013; 51:310–325. [PubMed: 23932714]
11. Mousavi K, et al. *Mol Cell*. 2013; 51:606–617. [PubMed: 23993744]
12. Di Ruscio A, et al. *Nature*. 2013; 503:371–376. [PubMed: 24107992]
13. Schaukowitch K, et al. *Mol Cell*. 2014; 56:29–42. [PubMed: 25263592]
14. Cassidy LA, Maher LJ 3rd. *Nucleic Acids Res*. 2002; 30:4118–4126. [PubMed: 12364590]
15. Jeon Y, Lee JT. *Cell*. 2011; 146:119–133. [PubMed: 21729784]
16. Gordon S, Akopyan G, Garban H, Bonavida B. *Oncogene*. 2006; 25:1125–1142. [PubMed: 16314846]
17. Flynn RA, Almada AE, Zamudio JR, Sharp PA. *Proc Natl Acad Sci U S A*. 2011; 108:10460–10465. [PubMed: 21670248]

18. Preker P, et al. *Science*. 2008; 322:1851–1854. [PubMed: 19056938]
19. Houbaviy HB, Usheva A, Shenk T, Burley SK. *Proc Natl Acad Sci U S A*. 1996; 93:13577–13582. [PubMed: 8942976]
20. Hnisz D, et al. *Cell*. 2013; 155:934–947. [PubMed: 24119843]
21. Lubas M, et al. *Cell Rep*. 2015; 10:178–192. [PubMed: 25578728]
22. Maurano MT, et al. *Science*. 2012; 337:1190–1195. [PubMed: 22955828]
23. Whyte WA, et al. *Cell*. 2013; 153:307–319. [PubMed: 23582322]
24. Wang L, et al. *Nature*. 2015; 523:621–625. [PubMed: 26123024]
25. Vella P, Barozzi I, Cuomo A, Bonaldi T, Pasini D. *Nucleic Acids Res*. 2012; 40:3403–3418. [PubMed: 22210892]
26. Kim J, Chu J, Shen X, Wang J, Orkin SH. *Cell*. 2008; 132:1049–1061. [PubMed: 18358816]
27. Boyer LA, et al. *Nature*. 2006; 441:349–353. [PubMed: 16625203]
28. Schmierer R, Edwards R. *Bioinformatics*. 2011; 27:863–864. [PubMed: 21278185]
29. Kim J, Cantor AB, Orkin SH, Wang J. *Nature Protocols*. 2009; 4:506–517. [PubMed: 19325547]
30. Langmead B, Trapnell C, Pop M, Salzberg SL. *Genome Biol*. 2009; 10:R25. [PubMed: 19261174]
31. Zhang Y, et al. *Genome biology*. 2008; 9:R137. [PubMed: 18798982]
32. Kent WJ, et al. *Genome research*. 2002; 12:996–1006. [PubMed: 12045153]
33. Grant CE, Bailey TL, Noble WS. *Bioinformatics*. 2011; 27:1017–1018. [PubMed: 21330290]
34. Bryne JC, et al. *Nucleic acids research*. 2008; 36:D102–D106. [PubMed: 18006571]
35. Jolma A, et al. *Cell*. 2013; 152:327–339. [PubMed: 23332764]
36. Trapnell C, Pachter L, Salzberg SL. *Bioinformatics*. 2009; 25:1105–1111. [PubMed: 19289445]
37. Wang L, Wang S, Li W. *Bioinformatics*. 2012; 28:2184–2185. [PubMed: 22743226]
38. Loven J, et al. *Cell*. 2012; 151:476–482. [PubMed: 23101621]
39. Lin CY, et al. *Cell*. 2012; 151:56–67. [PubMed: 23021215]
40. Quinlan AR, Hall IM. *Bioinformatics*. 2010; 26:841–842. [PubMed: 20110278]
41. Sigova AA, et al. *Proc Natl Acad Sci U S A*. 2013; 110:2876–2881. [PubMed: 23382218]
42. Wang D, et al. *Nature*. 2011; 474:390–394. [PubMed: 21572438]
43. Kopylova E, Noe L, Touzet H. *Bioinformatics*. 2012; 28:3211–3217. [PubMed: 23071270]
44. Kent WJ, Zweig AS, Barber G, Hinrichs AS, Karolchik D. *Bioinformatics*. 2010; 26:2204–2207. [PubMed: 20639541]
45. Jangi M, Boutz PL, Paul P, Sharp PA. *Genes & development*. 2014; 28:637–651. [PubMed: 24637117]
46. Martin M. *EMBnet journal*. 2011; 17:10–12.
47. Langmead B, Salzberg SL. *Nature Methods*. 2012; 9:357–359. [PubMed: 22388286]
48. Li H, et al. *Bioinformatics*. 2009; 25:2078–2079. [PubMed: 19505943]
49. Cernilogar FM, et al. *Nature*. 2011; 480:391–U151. [PubMed: 22056986]
50. Mullen AC, et al. *Cell*. 2011; 147:565–576. [PubMed: 22036565]
51. Wang H, et al. *Cell*. 2013; 153:910–918. [PubMed: 23643243]
52. Kearns NA, et al. *Development*. 2014; 141:219–223. [PubMed: 24346702]

**Fig. 1.**

YY1 binds to DNA and RNA at transcriptional regulatory elements. **(A)** Cartoon depicting divergent transcription at enhancers and promoters in mammalian cells. **(B)** Alignment of GRO-seq reads at all enhancers and promoters in ESCs. Enhancers were defined as in (23). The x-axis indicates distance from either the enhancer center (C) or the transcription start site (TSS) in kilobases. The y-axis indicates average density of uniquely mapped GRO-seq reads per genomic bin. **(C)** Gene tracks for the *Arid1a* gene and enhancer showing ChIP-seq and CLIP-seq data for bio-YY1 cells, as well as GRO-seq reads for mESC. **(D)** Mean read

density of YY1 ChIP-seq and CLIP-seq reads at enhancers and promoters of all RefSeq genes in ESCs.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

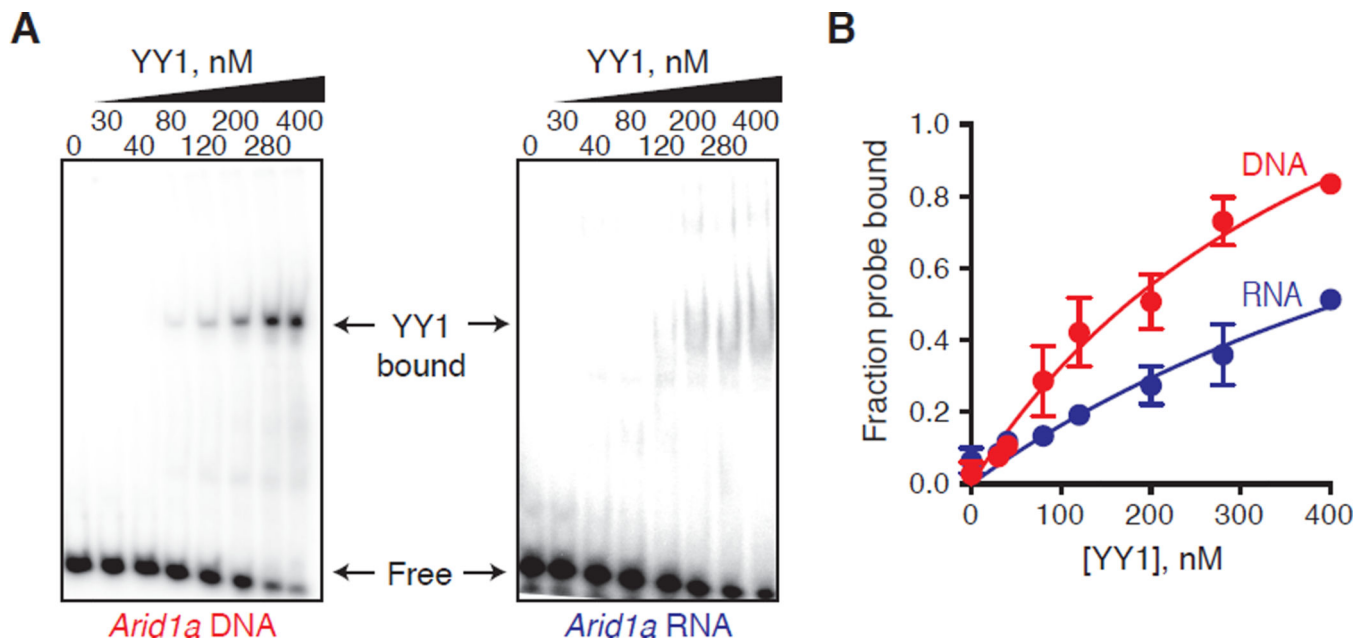


Fig. 2. YY1 binds to DNA and RNA *in vitro*. (A) Left panel: EMSA of YY1-DNA complexes at different concentrations of recombinant YY1. 5 nM of radioactively labelled 30-bp DNA probe derived from the promoter region of *Arid1a* gene containing a consensus YY1 binding motif (CTCTTCTCTTAAAATGGCTGCCTGTCTG) was incubated with increasing concentrations of recombinant murine YY1 protein. Right panel: EMSA of YY1-RNA complexes at different concentrations of recombinant YY1. 5 nM of radioactively labeled 30-nt RNA probe derived from the same region of the *Arid1a* gene was incubated with increasing concentrations of recombinant YY1 protein. (B) Graph depicting relationship between the fraction of radioactively labeled DNA or RNA probe bound and the concentration of recombinant YY1 in the binding reaction.

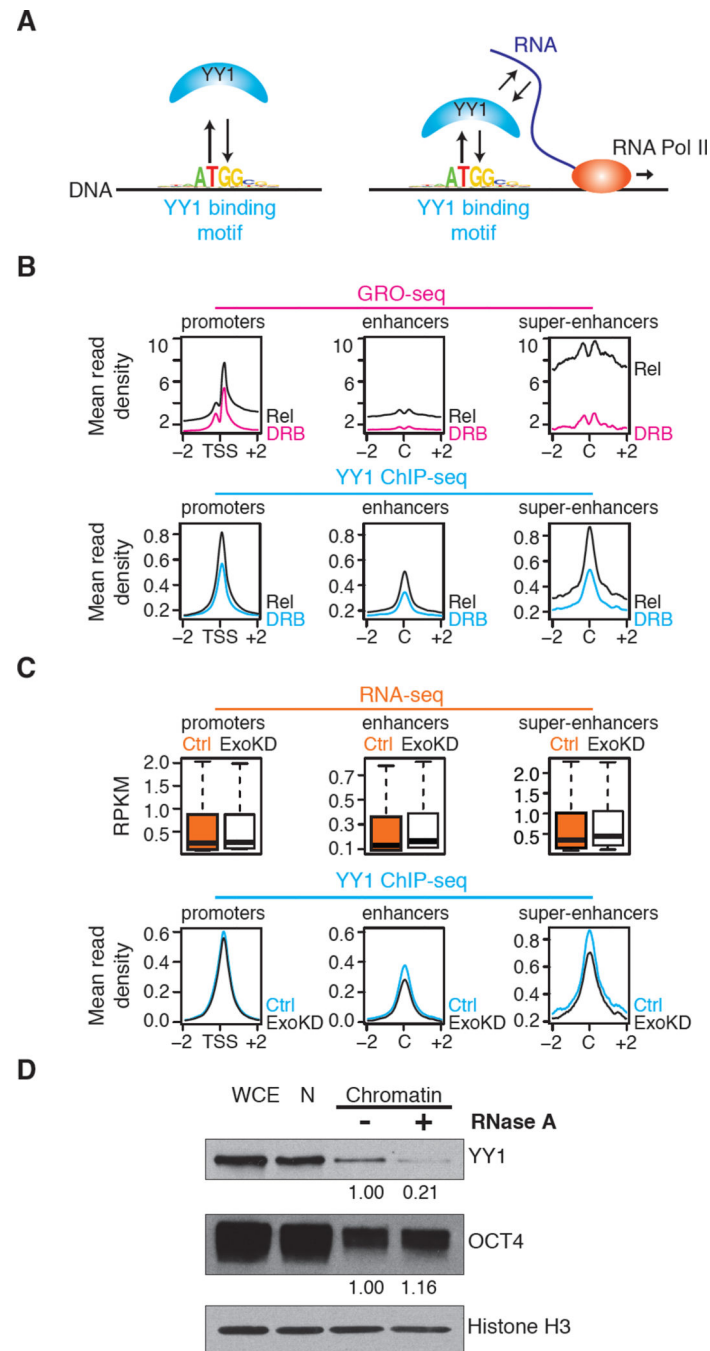


Fig. 3. Perturbation of RNA levels affects YY1 binding to DNA. **(A)** Cartoon depicting hypothesis that RNA transcribed from regulatory elements enhances occupancy of these elements by TFs capable of binding both DNA and RNA. **(B)** (Top) GRO-seq reads (24) at promoters, enhancers, and super-enhancer constituents in cells before (DRB) and after release (Rel) from transcriptional inhibition by DRB. (Bottom) YY1 ChIP-seq reads at promoters, enhancers and super-enhancer constituents in cells before (DRB) and after release (Rel) from transcriptional inhibition by DRB. Increase in YY1 binding after release from DRB

inhibition is significant: p -value $< 3.6 \times 10^{-207}$ for promoters, p -value $< 1.6 \times 10^{-214}$ for enhancers, p -value $< 9.8 \times 10^{-37}$ for super-enhancers. **(C)** (Top) Box plots depicting RNA-seq data for ribo-depleted total RNA at promoters, enhancers, and super-enhancers in ESCs after targeting with control (Ctrl) or *Exosc3* (ExoKD) shRNA. (Bottom) Alignment of YY1 ChIP-seq reads at promoters, enhancers and super-enhancers in ESCs after targeting with control (Ctrl) or *Exosc3* (ExoKD) shRNA. The decrease in YY1 binding in ExoKD ESCs is significant: p -value $< 8.1 \times 10^{-9}$ for promoters p -value $< 1.8 \times 10^{-27}$ for enhancers, p -value $< 3.3 \times 10^{-5}$ for super-enhancers. **(D)** Western blot analysis of YY1, OCT4, and histone H3 levels in whole-cell extracts (WCE), nuclei (N), and a nuclear chromatin preparation before and after RNase A treatment. Histone H3 serves as a loading control and OCT4 serves as a negative control. Quantitation of the relative levels of YY1 and OCT4 are noted.

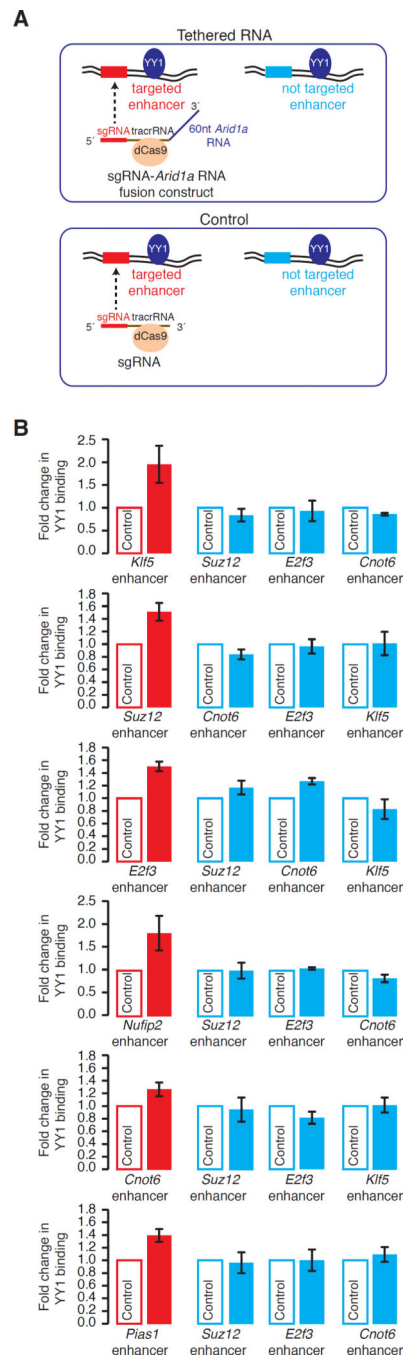


Fig. 4. Tethering of RNA adjacent an YY1 DNA binding site enhances binding of YY1 to the genome *in vivo*. (A) Strategy for tethering of RNA in the vicinity of an YY1 binding site at enhancers *in vivo*. (B) ChIP-qPCR analysis of YY1 binding at six targeted (red) and three not targeted (blue) enhancers in three independent experiments. The y-axis indicates fold change in YY1 binding in ESCs expressing the sgRNA-*Arid1a* RNA fusion construct relative to cells expressing the control sgRNA targeted to the same locus. The difference in YY1 binding was significant for the targeted enhancers: *Klf5* (p -value=0.03), *Suz12* (p -

value=0.01), *E2f3* (p -value=0.01), *Nufip2* (p -value=0.03), *Cnot6* (p -value=0.03), and *Pias1* (p -value=0.01), but not for the not targeted enhancers.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript