# Molecular Cell

The BET Protein BRD2 Cooperates with CTCF to Enforce Transcriptional and Architectural Boundaries

Graphical Abstract

Authors

Sarah C. Hsu, Thomas G. Gilgenast, Caroline R. Bartman, ..., Arjun Raj, Jennifer E. Phillips-Cremins, Gerd A. Blobel

Correspondence

jcremins@seas.upenn.edu (J.E.P.-C.), blobel@email.chop.edu (G.A.B.)

In Brief

BET family proteins impact gene expression in multiple ways, and they are pharmacologic targets for various diseases. Hsu et al. show that BRD2 is recruited to CTCF sites where it contributes to CTCF boundary formation and enhancer insulation.

Highlights

- BRD2 and to a lesser extent BRD3, but not BRD4, co-localize with CTCF genome-wide

- CTCF recruits BRD2 to chromatin, but not vice versa, as shown by genome editing

- Disruption of a CTCF/BRD2-occupied element enables inappropriate enhancer activity

- BRD2 depletion broadly impairs the integrity of CTCF/BRD2 co-occupied boundaries

CrossMark

CellPress

The BET Protein BRD2 Cooperates with CTCF to Enforce Transcriptional and Architectural Boundaries

Sarah C. Hsu,[1,2] Thomas G. Gilgenast,[3] Caroline R. Bartman,[1,2] Christopher R. Edwards,[1] Aaron J. Stonestrom,[1,2] Peng Huang,[1] Daniel J. Emerson,[3] Perry Evans,[1] Michael T. Werner,[1,2] Cheryl A. Keller,[4] Belinda Giardine,[4] Ross C. Hardison,[4] Arjun Raj,[3] Jennifer E. Phillips-Cremins,[3,*] and Gerd A. Blobel[1,2,5,*]

[1]Division of Hematology, Children's Hospital of Philadelphia, Philadelphia, PA 19104, USA
[2]Perelman School of Medicine
[3]Department of Bioengineering
University of Pennsylvania, Philadelphia, PA 19104, USA
[4]Department of Biochemistry and Molecular Biology, Pennsylvania State University, University Park, PA 16802, USA
[5]Lead Contact
*Correspondence: jcremins@seas.upenn.edu (J.E.P.-C.), blobel@email.chop.edu (G.A.B.)
http://dx.doi.org/10.1016/j.molcel.2017.02.027

SUMMARY

Bromodomain and extraterminal motif (BET) proteins are pharmacologic targets for the treatment of diverse diseases, yet the roles of individual BET family members remain unclear. We find that BRD2, but not BRD4, co-localizes with the architectural/insulator protein CCCTC-binding factor (CTCF) genome-wide. CTCF recruits BRD2 to co-bound sites whereas BRD2 is dispensable for CTCF occupancy. Disruption of a CTCF/BRD2-occupied element positioned between two unrelated genes enables regulatory influence to spread from one gene to another, suggesting that CTCF and BRD2 form a transcriptional boundary. Accordingly, single-molecule mRNA fluorescence in situ hybridization (FISH) reveals that, upon site-specific CTCF disruption or BRD2 depletion, expression of the two genes becomes increasingly correlated. HiC shows that BRD2 depletion weakens boundaries co-occupied by CTCF and BRD2, but not those that lack BRD2. These findings indicate that BRD2 supports boundary activity, and they raise the possibility that pharmacologic BET inhibitors can influence gene expression in part by perturbing domain boundary function.
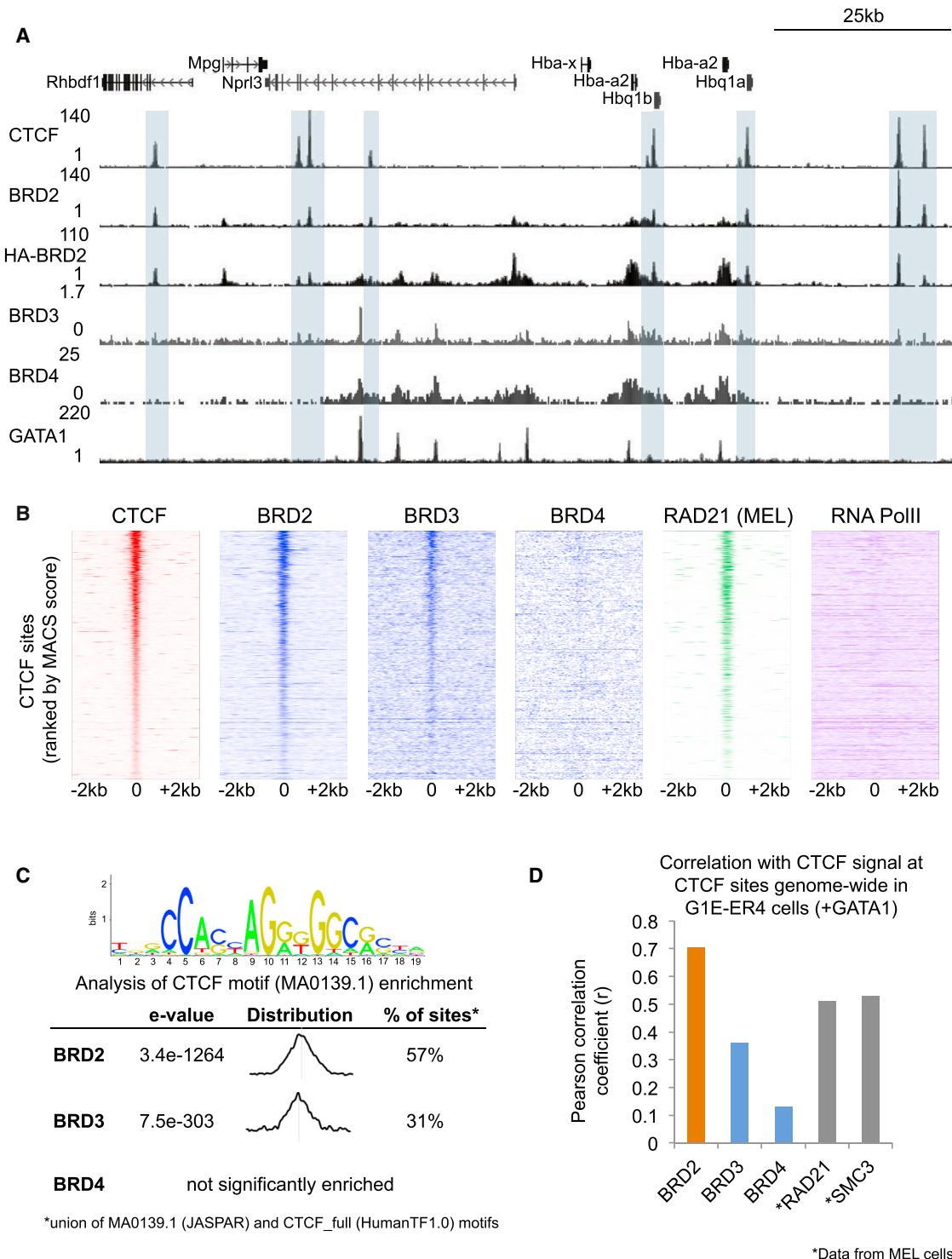
INTRODUCTION

BRD2, BRD3, and BRD4 are members of the bromodomain and extraterminal motif (BET) family of proteins that are critical regulators of transcription. BET inhibitors that competitively target the bromodomain-acetyl lysine interaction have shown great promise in treating cancer and other pathologies and are currently in clinical development. However, much of how BET proteins function remains to be explored. BET proteins are ubiquitously expressed and bind to chromatin via bromodomain association with acetylated histones and/or transcription factors. A leading model is that BET proteins serve as general adaptors that recruit regulatory proteins to chromatin; however, BET inhibition does not result in global, but rather targeted and context-dependent effects on gene expression. Such selective sensitivity to BET inhibition has been attributed to BET protein occupancy at enhancers (Lovén et al., 2013) and recruitment by lineage-specific transcription factors (Asangani et al., 2014; Huang et al., 2009; Lamonica et al., 2011; Roe et al., 2015; Stonestrom et al., 2015). However, since most BET inhibitors indiscriminately target all family members (Filippakopoulos et al., 2010), the degree to which BRD2, BRD3, and BRD4 function with distinct regulatory complexes or play unique roles in gene regulation has not been well characterized. Most studies have focused on BRD4 as the primary transcriptional effector, yet several reports have also identified roles for BRD2 in transcriptional activation (Belkina et al., 2013; Stonestrom et al., 2015; Surface et al., 2016; Vardabasso et al., 2015). How BRD2 functions distinctly from BRD3 and BRD4 remains largely unclear.

Several studies point to a role for BET proteins in both large-scale nuclear structure and chromatin organization. The testes-specific BET, BRDT, is required to maintain centromeric heterochromatin foci during spermatogenesis (Berkovits and Wolgemuth, 2011; Shang et al., 2007), while BRD4 has been linked to regulating both global (Devaiah et al., 2016; Wang et al., 2012) and gene-specific (Zhao et al., 2011) chromatin compaction. Notably, the yeast BET protein Bdf1 maintains heterochromatin-euchromatin boundaries at telomeres and mating loci, suggesting that BETs can help assemble physical barriers (Ladurner et al., 2003). In addition, Fs(1)h, a *Drosophila* BET protein, forms complexes with multiple insulator proteins (Kellner et al., 2013). Fs(1)h binding increases upon heat shock, a stress that is accompanied by changes in 3D genomic architecture (Li et al., 2015). Yet, whether mammalian BETs also participate directly in insulator or architectural functions remains unclear.

In vertebrates, CCCTC-binding factor (CTCF) contributes to the partitioning of the genome into discrete globular structures termed topologically associating domains (TADs) (Dixon et al., 2012; Nora et al., 2012). TADs represent megabase-sized

Figure 1. High Correlation of BRD2 and CTCF Occupancy Genome-wide

(A) Genome browser tracks showing ChIP-seq signal for CTCF, BRD2, HA-BRD2, BRD3, BRD4, and GATA1 at the α-globin (*Hba*) locus. Gray boxes highlight regions of BRD2/CTCF co-localization. All tracks shown are in the estradiol-induced (+GATA1) condition.

(B) ChIP-seq signal for the indicated proteins in a 4 kb window centered on CTCF-binding sites in induced G1E-ER4 cells. Each row represents a single peak, ranked from highest to lowest CTCF signal (MACS score).

(legend continued on next page)

regions within which loci are more likely to interact with each other than with genomic segments that lie outside. A number of observations suggest that TADs constrain the activity of transcriptional regulatory elements. For example, in some cases, genes within TADs exhibit more correlated expression patterns with each other than with genes in other domains (Flavahan et al., 2016; Nora et al., 2012), and enhancers typically exert their activity within TAD boundaries (Symmons et al., 2014). CTCF was initially implicated in the formation of TAD structures due to its strong enrichment at TAD boundaries (Dixon et al., 2012). Upon deletion of a CTCF site at a TAD boundary, Narendra et al. (2015) observed the spreading of positive regulatory influence across the boundary and upregulation of previously inactive genes. Global depletion of CTCF increases genomic contacts across domain boundaries (Zuin et al., 2014). Mutations that abrogate CTCF binding at specific boundaries enable ectopic enhancer-promoter contacts between adjacent architectural domains, leading to aberrant gene activation in models of cancer and limb malformation (Flavahan et al., 2016; Lupiáñez et al., 2015). Taken together, these observations suggest that CTCF both functionally and physically divides the genome into structural domains. However, because the majority of CTCF-bound sites are not associated with TAD boundaries (Dixon et al., 2012; Phillips-Cremins et al., 2013), the context and mechanisms by which CTCF performs its insulating or boundary function at select sites are unknown.

Here we show that the BET protein BRD2 co-localizes with CTCF genome-wide. CTCF is required to recruit BRD2 to co-occupied sites, while CTCF binding is largely BRD2 independent. We present evidence that CTCF forms a functional boundary at the Mitoferrin 1 (*Slc25a37*) locus by restricting the activity of the *Slc25a37* enhancer to prevent aberrant upregulation of a nearby gene. Using single-molecule mRNA fluorescence in situ hybridization (FISH), we find that CTCF, in concert with BRD2, limits the correlation in expression of two genes flanking the boundary. HiC experiments show that BRD2 contributes to chromatin domain boundary function by limiting chromatin contacts across boundaries specifically occupied by BRD2, but not those lacking BRD2. These findings reveal BRD2 as a CTCF cofactor, and they suggest that BET proteins may impact transcription in part through regulating higher-order chromatin architecture.

RESULTS

BRD2 Co-localizes with CTCF Genome-wide
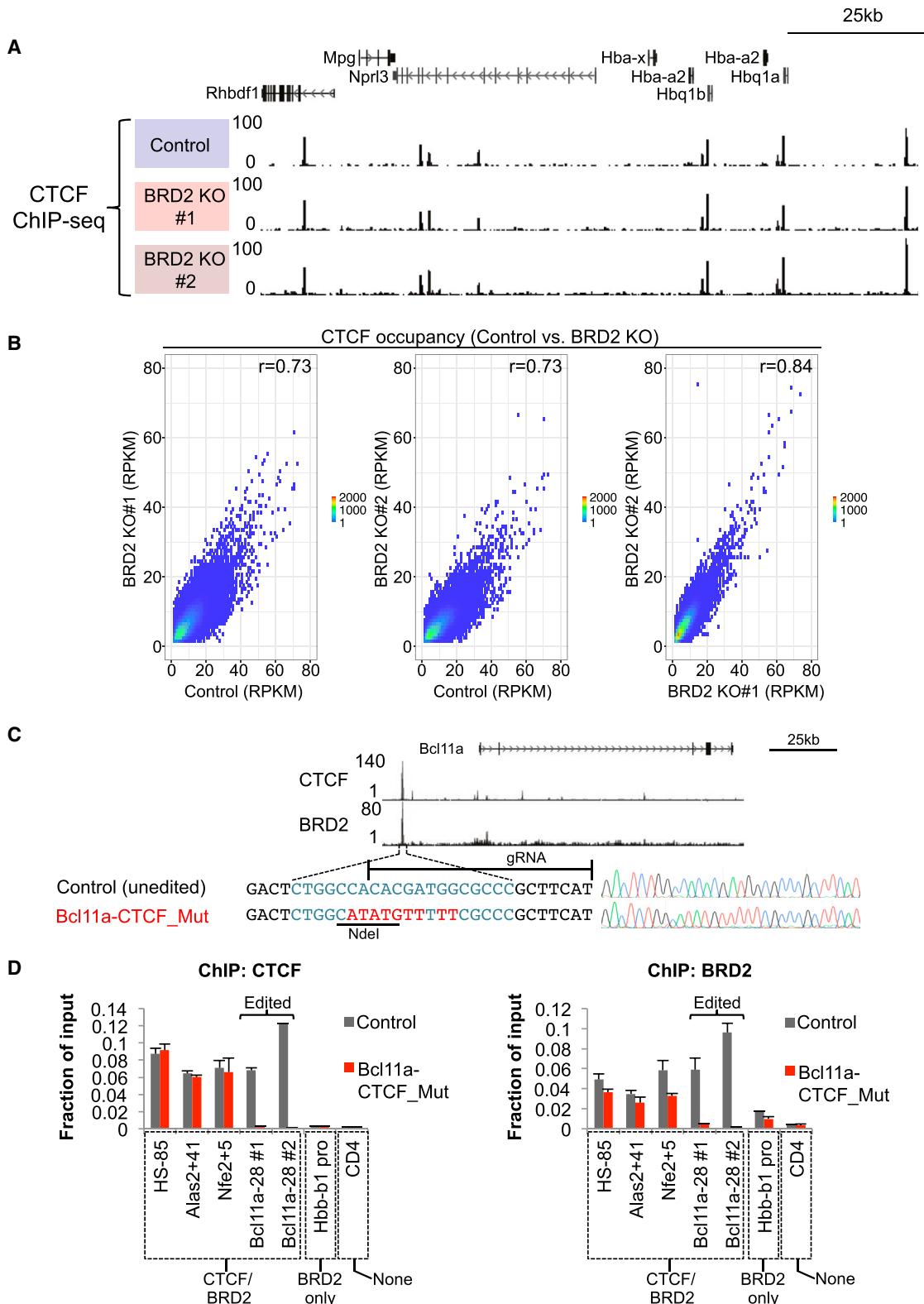
BRD2 and BRD4 are each required for gene activation during erythroid maturation driven by the hematopoietic transcription factor GATA1 (Stonestrom et al., 2015). However, genome-wide analysis of BET localization patterns in the erythroblast cell line G1E-ER4 indicated that they may promote erythroid maturation through distinct mechanisms. Specifically, in contrast to BRD3 and BRD4, BRD2 genome-wide chromatin occupancy does not significantly overlap with that of GATA1, leaving its mechanism of action unresolved (Stonestrom et al., 2015). We analyzed BRD2 occupancy patterns in greater depth in relation to chromatin immunoprecipitation sequencing (ChIP-seq) datasets in G1E-ER4 cells from our laboratory (Stonestrom et al., 2015) and from the Mouse Encyclopedia of DNA Elements (ENCODE) Consortium. G1E-ER4 cells lack endogenous GATA1 and instead express a fusion protein of GATA1 and the ligand-binding domain of the estrogen receptor. The addition of estradiol induces erythroid maturation and activates a red cell-specific gene expression program (Welch et al., 2004) (here we refer to the differentiated state as induced G1E-ER4 or +GATA1 and the undifferentiated state as uninduced G1E-ER4 or –GATA1). In induced cells, we observed a striking overlap between BRD2 and the architectural protein CTCF (Figure 1A). Overexpression of HA-tagged BRD2 tended to produce broader signals, but displayed a similar pattern of CTCF co-localization. Ranking of CTCF peaks by signal intensity (model-based analysis of ChIP-seq [MACS] score) revealed a strong correlation with BRD2 occupancy (Figures 1B and 1D). To a lesser extent BRD3 was also enriched at CTCF sites, possibly reflecting functional overlap between BRD2 and BRD3 (Stonestrom et al., 2015) (Figure 1B). We observed a similar pattern of BRD2 co-localization with CTCF in undifferentiated cells (Figures S1A and S1B). Analysis of DNA sequences under BRD2 peaks using the MEME suite (Machanick and Bailey, 2011) identified a CTCF consensus at 57% of BRD2 sites, supporting extensive BRD2-CTCF co-localization (Figure 1C).

Among the known CTCF-binding partners, the cohesin complex in particular occupies a large fraction of CTCF sites in diverse cell types, and it participates in CTCF-mediated insulator function (Parelho et al., 2008; Wendt et al., 2008) and loop formation (Seitan et al., 2013; Sofueva et al., 2013). Using ChIP-seq datasets from the Mouse ENCODE Consortium in murine erythroleukemia cells (MEL), a related murine erythroid cell line, we found that the occupancy of cohesin subunits RAD21 and SMC3 was highly correlated with CTCF at CTCF-occupied sites in induced G1E-ER4 cells, consistent with their established functional relationship (Figure 1D). The correlation between BRD2 and CTCF binding was similar to that observed between cohesin and CTCF, despite possible cell line-specific differences between MEL and G1E-ER4, suggesting that BRD2 may play a similarly important role in mediating CTCF function. At CTCF sites, BRD2 and cohesin also exhibited a strong correlation (Figures 1B and S1C), suggesting that many CTCF sites are occupied by both BRD2 and cohesin. Even though CTCF and BRD2 can associate with RNA polymerase II (Pol II) (Chernukhin et al., 2007; Denis et al., 2006), this shared property does not account for their co-localization, as there is little overlap between

(C) CTCF motif enrichment analysis using the MEME suite, performed on BRD2-, BRD3-, or BRD4-occupied sites in induced G1E-ER4 cells. The percentage of sites containing a CTCF consensus motif (union of both the MA0139.1 [JASPAR] and CTCF_full [HumanTF1.0] motifs), as well as the distribution of the motif relative to the BET protein peak, are indicated.

(D) Pearson correlation coefficients between CTCF (reads per kilobase per million mapped reads [RPKM]) and indicated proteins (RPKM) at CTCF-binding sites. RAD21 and SMC3 ChIP-seq data are from ENCODE datasets in MEL cells.

See also Figure S1.

A

B CTCF occupancy (Control vs. BRD2)

C

D ChIP: CTCF   ChIP: BRD2

(legend on next page)

Pol II (datasets from Hsiung et al., 2016) and BRD2 occupancy patterns at CTCF-occupied sites (Figures 1B and S1B).

Consistent with previous reports (Anders et al., 2014), we also noted some places, such as the α-globin genes (*Hba*), at which BRD2, BRD3, and BRD4 were all bound (Figure 1A). However, BRD4 was not highly enriched at CTCF sites genome-wide (Figure 1B), and it did not correlate well with CTCF occupancy (Figure 1D), pointing to possible unique functional roles for BRD2 and BRD4.

CTCF Is Required for BRD2 Occupancy at Co-bound Sites

The significant co-occupancy of BRD2 and CTCF suggests they may facilitate each other's association with chromatin, similar to what has been described for BET proteins in the context of GATA1 or the androgen receptor (Asangani et al., 2014; Lamonica et al., 2011; Stonestrom et al., 2015). To test if BRD2 promotes CTCF binding, we performed ChIP-seq of CTCF in two independent G1E-ER4 clonal sub-lines, in which BRD2 was depleted by CRISPR/Cas9 (BRD2 knockout [KO] cell lines 1 and 2; Figure S2A), and unedited control G1E-ER4 cells. CTCF occupancy was unchanged at the majority of sites examined, including the α-globin locus (Figure 2A). When we compared CTCF signal intensities in control and BRD2-depleted cells genome-wide (Figure 2B), we found they were very similar among the two BRD2 KO cell lines and control cells (Pearson correlation coefficients: BRD2 KO 1 versus BRD2 KO 2 = 0.84; control versus BRD2 KO 1 = 0.73; control versus BRD2 KO 2 = 0.73) (Figure 2B), indicating that CTCF chromatin occupancy is largely BRD2 independent. These results were confirmed using ChIP-qPCR at select CTCF sites (Figure S2B). Given that cohesin is present at many CTCF/BRD2 co-occupied sites, we tested whether BRD2 is required for cohesin recruitment at CTCF sites. ChIP-qPCR using antibodies against the cohesin subunit SMC1 showed that its occupancy was unaffected by BRD2 depletion at the majority of CTCF/BRD2 sites examined (Figure S2C), suggesting that BRD2 is dispensable for cohesin binding at CTCF sites.

BET proteins are recruited by transcription factors in diverse contexts (Asangani et al., 2014; Huang et al., 2009; Lamonica et al., 2011; Roe et al., 2015; Shi et al., 2014; Stonestrom et al., 2015; Wu et al., 2013). To determine if BRD2 requires CTCF to occupy chromatin, we used CRISPR/Cas9-mediated homology-directed repair to mutate a specific CTCF site upstream of the *Bcl11a* locus (Figure 2C). We introduced nine point mutations in the CTCF consensus sequence (Bcl11a-CTCF_Mut) that created an NdeI restriction site previously

shown to eliminate CTCF binding in vitro (Szabó et al., 2004). This approach avoids the confounding effects of larger deletions or insertions that may further alter the structure of the locus. As predicted, these mutations abrogated CTCF binding in vivo (Figure 2D). ChIP-qPCR revealed that BRD2 occupancy was lost in the absence of CTCF at this site, but not at control regions where CTCF binding remained intact (Figure 2D). The expression of the surrounding genes was unaffected (Figure S2D), indicating that loss of BRD2 at this site was not a result of transcriptional changes at the locus. In sum, these data indicate that CTCF mediates BRD2 chromatin association, but not vice versa.
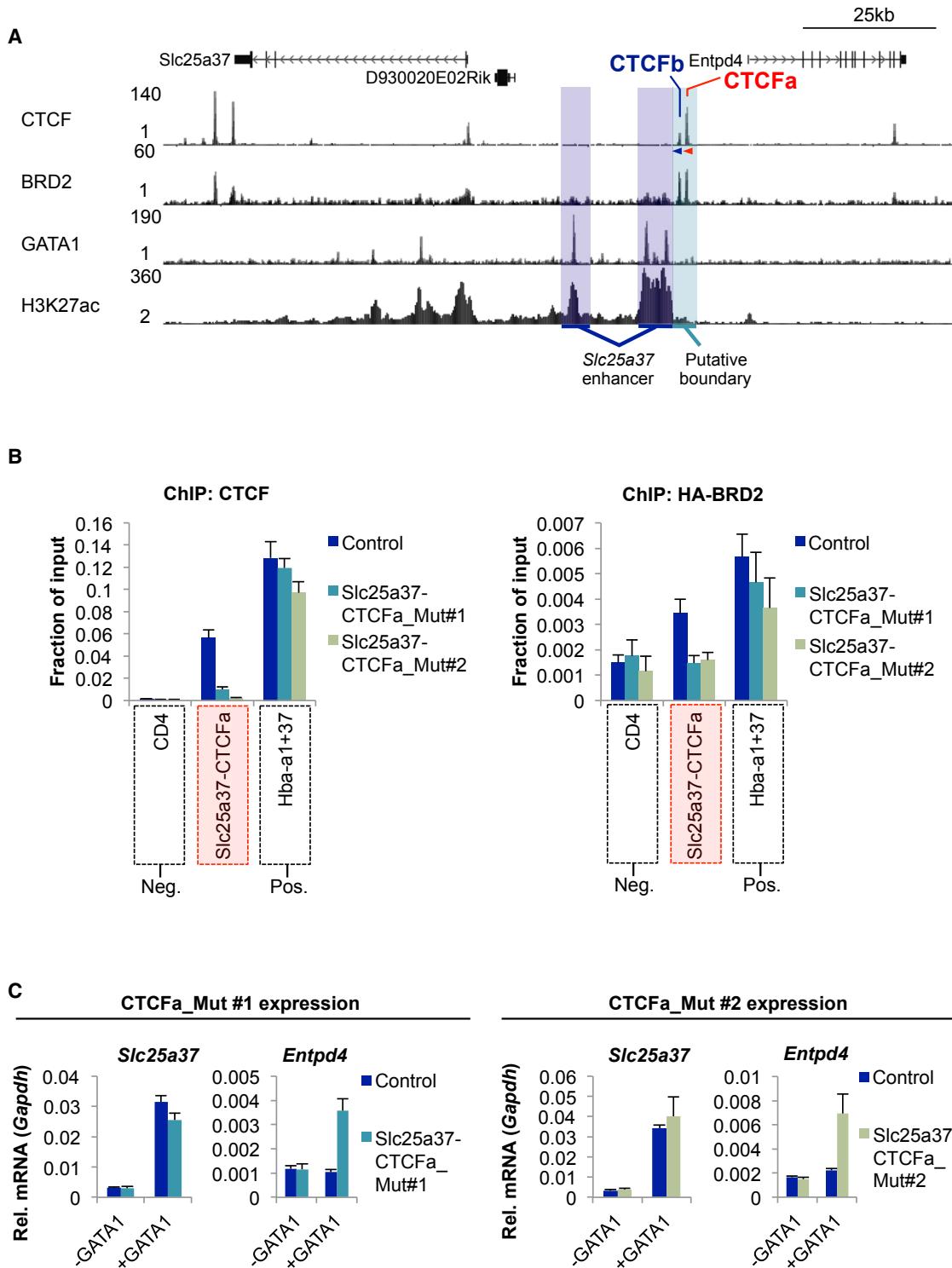
CTCF Defines a Functional Boundary at the *Slc25a37* Locus

CTCF has diverse roles in transcriptional regulation, including gene activation and repression, enhancer blocking, splicing, and forging long-range chromatin interactions (Phillips and Corces, 2009). To determine if BRD2 participates in any of these CTCF functions, we first perturbed CTCF in a locus- and site-specific manner. We focused on the *Slc25a37* locus, or Mitoferrin 1, which contains a gene encoding a mitochondrial iron transporter required for erythroid maturation (Shaw et al., 2006). Three GATA1-occupied regulatory regions downstream of the *Slc25a37* gene function as enhancer elements and are in physical proximity with the *Slc25a37* promoter (Huang et al., 2016; Hughes et al., 2014). Two CTCF sites (CTCFa and CTCFb; Figure 3A) separate the enhancer from a neighboring gene, *Entpd4*, suggesting that they may function as boundary or insulator elements to prevent the enhancer from inappropriately regulating *Entpd4*. Consistent with this possibility, chromatin interactions with the *Slc25a37* promoter drop off significantly at these CTCF sites in Capture-C experiments (Hughes et al., 2014) (Figure S3A), and disruption of the GATA1-dependent enhancer elements reduces *Slc25a37* expression, but has no effect on the expression of *Entpd4* (Huang et al., 2016). BRD2 co-localizes with CTCF at both CTCFa and CTCFb (Figure 3A).

Using CRISPR/Cas9, we generated two independent clones with either biallelic deletions or combined deletions/insertions disrupting all or part of the CTCFa consensus sequence, Slc25a37-CTCFa_Mut1 and Mut2 (Figure S3B). ChIP-qPCR confirmed loss of occupancy by both CTCF and an HA-tagged form of BRD2 (Figure 3B). CTCF and HA-BRD2 binding at the second site, CTCFb, were unaffected by these mutations (Figure S3C). Attempts to mutate both CTCFa and CTCFb failed for unknown reasons. To examine the transcriptional consequences of disrupting CTCFa, we measured the expression of the surrounding genes. *Slc25a37* levels changed little, if at all,

Figure 2. CTCF Recruits BRD2
(A) Genome browser tracks showing CTCF ChIP-seq at the α-globin (*Hba*) locus in induced control G1E-ER4 and BRD2 KO 1 and 2 cells.
(B) Scatterplot (binned 2D density plot) comparing CTCF ChIP-seq read density (RPKM) at CTCF peaks in control and BRD2 KO cells. Pearson correlation coefficients are indicated in each panel. The color scale represents the density of data points within each bin.
(C) CRISPR-Cas9-mediated mutation of the CTCF-binding site upstream of the *Bcl11a* gene. The mutant sequence creates an NdeI site. Blue and red nucleotides annotate the intact and mutant CTCF motifs, respectively. The guide RNA (gRNA) target sequence is also shown, along with sequence traces of both control cells and a mutant clone (Bcl11a-CTCF_Mut).
(D) ChIP-qPCR of CTCF (left) and BRD2 (right) in control or Bcl11a-CTCF_Mut cells (all + GATA1). Bcl11a-28 1 and 2 indicate two independent primer pairs to the region surrounding the mutated CTCF site (error bars represent SD, n = 2).
See also Figure S2.

Figure 3. CTCF Supports a Boundary at the *Slc25a37* Locus

(A) Genome browser tracks of the indicated proteins at the Mitoferrin 1 (*Slc25a37*) locus in induced G1E-ER4 cells. Arrowheads below show orientation of the CTCF motifs. Blue box, GATA1-dependent *Slc25a37* enhancer; green box, CTCFa and CTCFb sites co-occupied by BRD2.

(legend continued on next page)

in both clones. However, the expression of *Entpd4* was upregulated ~3-fold (Figure 3C). Notably the increase in *Entpd4* expression was GATA1 dependent, suggesting it results from inappropriate activation by the GATA1-regulated *Slc25a37* enhancer. Other genes normally activated during erythroid maturation were unaffected (Figure S3D). These results are consistent with a role for CTCF in enhancer blocking or boundary formation.

To directly test whether CTCF restricts the *Slc25a37* enhancer from contacting the *Entpd4* gene, we performed chromosome conformation capture (3C). Both Slc25a37-CTCFa mutant cell lines exhibited increased interaction frequency between the *Slc25a37* enhancer and the *Entpd4* promoter (Figure S4B), indicating that CTCF serves to block enhancer contacts at this locus. One mechanism by which CTCF may perform its insulating function is through the pairing with other convergently oriented CTCF sites that demarcate a given contact domain (de Wit et al., 2015; Guo et al., 2015; Rao et al., 2014). The CTCFa and CTCFb motifs are both oriented in the reverse (left) direction (Figure 3A), suggesting their loop orientation is similar. We speculate that the effect sizes of inappropriate upregulation of *Entpd4* expression and contacts with the *Slc25a37* enhancer would have been greater had we been able to disrupt both CTCF sites simultaneously, similar to what has been observed at the *Hox* gene clusters (Narendra et al., 2015).

A direct test of BRD2's role in the function of this boundary is hampered by the lack of a means to perturb BRD2 binding in a site-specific fashion. Global BRD2 depletion reduced expression of *Slc25a37* (Figure S4A), likely due to BRD2 enrichment at the *Slc25a37* promoter and enhancer (Figure 3A), while *Entpd4* levels were relatively unaffected. Moreover, the *Slc25a37* enhancer-promoter contacts were diminished in the absence of BRD2 as measured by 3C (Figure S4C), indicating that the *Slc25a37* enhancer requires BRD2 for its full activity. The weakening of the *Slc25a37* enhancer in the absence of BRD2 complicates the assessment of whether BRD2, similar to CTCF, limits the range of *Slc25a37* enhancer activity. Using 3C we observed a subtle trend toward increased contacts between the *Slc25a37* enhancer and the *Entpd4* promoter upon BRD2 loss, but it did not reach statistical significance (Figure S4C). It is possible that direct or indirect effects from BRD2 depletion obscure potential increases in both *Entpd4* levels and contacts between the *Slc25a37* enhancer and *Entpd4* promoter. Hence, we pursued an independent strategy to detect perturbed boundary strength by examining the degree of correlation in expression of genes flanking the boundary, as described in the following section.

BRD2 Potentiates CTCF Boundary Function in Single Cells

Loss of the Slc25a37-CTCFa site renders *Entpd4* GATA1 responsive and increases contacts between the *Slc25a37* enhancer and *Entpd4*, suggesting that the two genes now share the regulatory influence of this enhancer. To test whether *Slc25a37* and *Entpd4* are co-regulated in the absence of CTCFa, we examined their expression in individual cells by quantitative single-molecule mRNA FISH (Femino et al., 1998; Raj et al., 2008). This method is based on the use of fluorescent probes targeting exons and allows the counting of individual *Slc25a37* and *Entpd4* mRNAs. The absolute mRNA counts of *Slc25a37* and *Entpd4* reflected gene expression changes observed in bulk populations (Figure S5A), with cells from both Slc25a37-CTCFa mutated lines exhibiting elevated *Entpd4* levels compared to control cells, and the levels of *Slc25a37* remaining unchanged. We next assessed the correlation in mRNA levels in single cells, as depicted in the model shown in Figure 4A. A low correlation suggests that *Slc25a37* and *Entpd4* have little regulatory relationship, while a high correlation implies that the two genes share regulatory control. We note that such a relationship cannot be observed in bulk mRNA measurements in which, for example, two cell populations could have equal total mRNA levels of two genes but exhibit distinct levels of each transcript in individual cells (as illustrated in Figure 4A). To test this we plotted the number of mRNA molecules per cell of each gene from either control or Slc25a37-CTCFa mutant populations, and we measured the correlation coefficient across three biological replicates (Figures 4B–4D). In control cells, expression of *Slc25a37* and *Entpd4* exhibited a low level of correlation. Deletion of the intervening CTCFa site significantly increased their correlation (Figures 4B–4D) in both Slc25a37-CTCFa mutant lines.

We considered that the observed changes in correlated gene expression could result from other confounding factors. For example, assume that two genes have equal transcription rates in every cell, but that this rate varies from cell to cell. At high transcript levels, a strong correlation between transcript counts would be observed. However, at low transcript abundances, the effects of random counting noise would play a disproportionate role, resulting in a decreased observed correlation despite the exact equivalence in transcription rate in every cell. In the present case, *Entpd4* expression was lower in control cells, raising the possibility that a potential correlation with *Slc25a37* expression was masked due to low molecule number noise. To test this possibility, we scaled down the levels of *Entpd4* in the Slc25a37-CTCFa-mutated lines such that the means of control and mutant populations were equivalent. We then added random sampling error to the number of *Entpd4* mRNA molecules, thus mimicking counting errors at low expression levels. We found that, while some of the increased correlation in the mutant could be explained this way, the difference in correlation between the Slc25a37-CTCFa mutant and control populations remained intact (Figure S5B). To increase the analytical stringency, we scaled the levels of *Entpd4* in the Slc25a37-CTCFa mutated lines down to half that of control levels. While the correlation in the Slc25a37-CTCFa mutated cells decreased further, it still remained higher than in the control setting (unpublished data).

(B) ChIP-qPCR of CTCF and HA-BRD2 in control induced G1E-ER4 cells and Slc25a37-CTCFa_Mut1 and Mut2 cells (error bars represent SEM, n = 3). The CD4 gene serves as a negative control, while Hba-a1+37 serves as a positive control (unedited) CTCF site.
(C) qRT-PCR of indicated transcripts in either uninduced (−GATA1) or GATA1-induced (+GATA1) cells. Transcripts were normalized to *Gapdh* (error bars represent SEM, n = 3).
See also Figures S3 and S4.
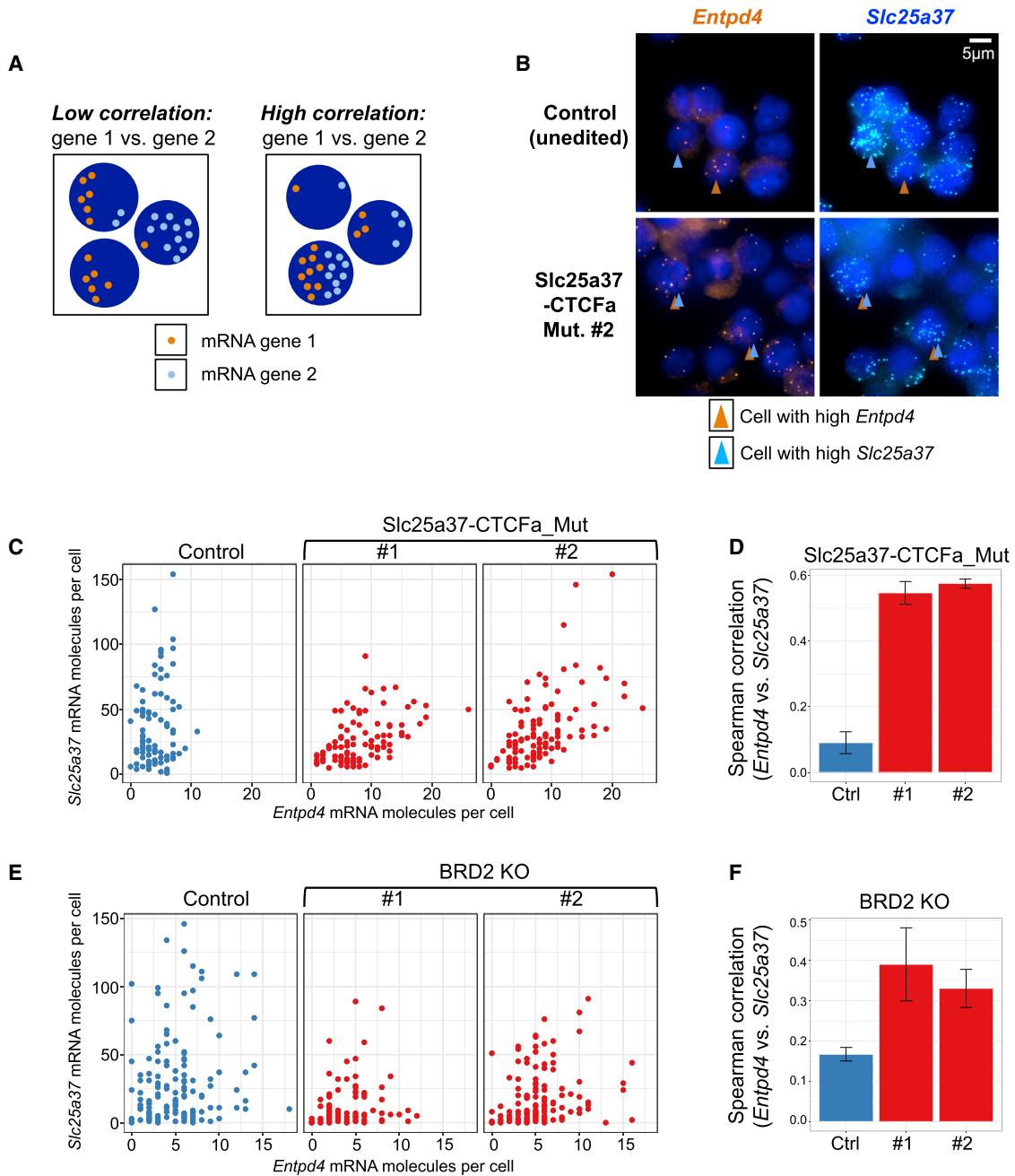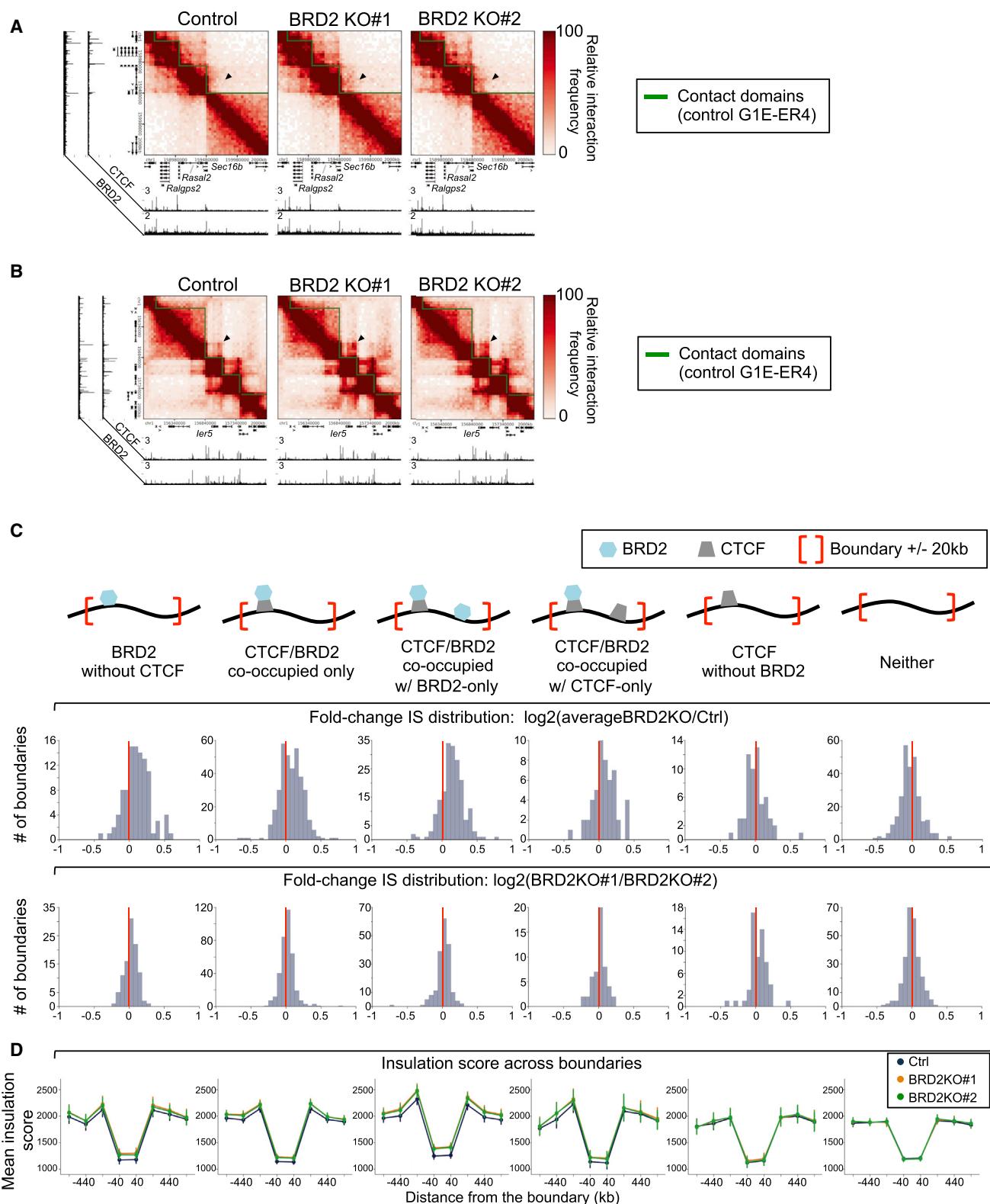
Figure 4. CTCF and BRD2 Maintain the Integrity of a Transcriptional Boundary in Single Cells

(A) Model illustrating scenarios in which two genes exhibit either low (left) or high (right) correlation on a per-cell basis.

(B) Single-molecule RNA FISH using probes to exons in either *Slc25a37* or *Entpd4* in induced control or boundary-mutated Slc25a37-CTCFa_Mut cells. Each spot represents a single mRNA molecule. Arrows indicate cells with high levels of either *Slc25a37* or *Entpd4* mRNA.

(C) Scatterplots of a representative mRNA FISH experiment showing *Slc25a37* and *Entpd4* mRNA counts per cell in control and both Slc25a37-CTCFa_Mut cell lines (+GATA1).

(D) Average Spearman correlations of *Slc25a37* and *Entpd4* transcripts from three biological replicates of mRNA FISH in control or Slc25a37-CTCFa_Mut cell lines 1 and 2 (error bars represent SEM, n = 3).

(E) Scatterplots of a representative mRNA FISH experiment as in (C) for induced control or BRD2 KO 1 and 2 cells.

(F) Average Spearman correlations of *Slc25a37* and *Entpd4* quantified over four biological mRNA FISH replicates in control or BRD2 KO 1 and 2 cells (error bars represent SEM, n = 4).

See also Figure S5.

A

Control   BRD2 KO#1   BRD2 KO#2

Contact domains
(control G1E-ER4)

B

Control   BRD2 KO#1   BRD2 KO#2

Contact domains
(control G1E-ER4)

C

BRD2   CTCF   [ ] Boundary +/- 20kb

BRD2   CTCF/BRD2   CTCF/BRD2   CTCF/BRD2   CTCF   Neither
without CTCF co-occupied only   co-occupied   co-occupied   without BRD2
                                 w/ BRD2-only   w/ CTCF-only

Fold-change IS distribution: log2(averageBRD2KO/Ctrl)

Fold-change IS distribution: log2(BRD2KO#1/BRD2KO#2)

D

Insulation score across boundaries

Distance from the boundary (kb)

Ctrl
BRD2KO#1
BRD2KO#2

(legend on next page)

Transcript abundance can correlate with cell size (Padovan-Merhar et al., 2015). To control for possible effects of cell size differences between populations, we normalized *Entpd4* mRNA levels to cell area, and we measured the correlation between *Slc25a37* and *Entpd4*. The correlation between these two genes in the boundary-mutated cells remained higher, demonstrating that possible cell size differences were not a main contributing factor (Figure S5C). In concert, these results further support that CTCF prevents the *Slc25a37* enhancer from acting on the *Entpd4* gene.

To test whether BRD2 participates in CTCF's boundary function at this locus, we performed mRNA FISH on BRD2-depleted cells (Figure 4E). Absolute mRNA measurements of *Slc25a37* and *Entpd4* were consistent with qRT-PCR analysis (Figures S4A and S5D). Notably, in both BRD2-deficient clones the two genes displayed an increased correlation in their expression (Figures 4E and 4F) relative to control cells. In comparing the difference between correlation coefficients, we found that one BRD2-depleted clone reached statistical significance relative to control cells (BRD2 KO 2, p = 0.018, t test), while the other showed the same trend but fell short of statistical significance (BRD2 KO 1, p = 0.069, t test). Within each of the four biological replicates of this experiment, we noted that, with one exception, the correlation between the two genes was increased in the BRD2-depleted clones relative to control cells (Figure S5E). While we included all replicates in our pooled comparisons, thus explaining why there is some degree of variability in the data, the consistency of the trend within the majority of experiments supports that the two genes are indeed more correlated in the absence of BRD2. The increased correlation in BRD2-depleted cells remained intact after normalizing for cell area (Figure S5C). We noted that performing a similar simulation as described above, in this case scaling down the levels of *Slc25a37* in control cells such that the means of control and BRD2 KO populations were equivalent, did not significantly alter the results (unpublished data), suggesting that changes in overall transcript abundances were not driving the differences in correlation that we observed.

While we cannot rule out potentially confounding secondary effects from sustained BRD2 depletion, these results suggest that BRD2 supports CTCF's ability to partition transcriptional regulatory elements at this locus.

Loss of BRD2 Weakens BRD2-Occupied Boundaries

CTCF's function as an insulating or boundary element may be linked to its role in organizing genome architecture. To assess whether BRD2 regulates chromatin structure and whether this function overlaps with that of CTCF, we generated 40 kb resolution in situ HiC maps in control G1E-ER4 cells and two BRD2-depleted cell lines. Two biological replicates for each cellular condition generated a total of ~934 million unique interaction pairs (control, ~323 million; BRD2 KO 1, ~301 million; BRD2 KO 2, ~310 million) (Figure S6A; Table S1). Since BRD2 is required for erythroid differentiation, experiments were performed in uninduced cells (−GATA1) to avoid potentially confounding effects of differences in maturation. The raw reads from biological replicates were highly correlated with one another, demonstrating the reproducibility of the experiment and allowing us to merge the two replicates for all subsequent analysis. We corrected for inherent biases in HiC maps via matrix balancing, and we visualized contact frequency heatmaps binned at 40 kb resolution. We used the *Kit* locus as a benchmark to determine whether our HiC data captured known structural features (Figure S6B). In the 2 Mb region surrounding *Kit*, we observed domain structure highly similar to published HiC data in mouse embryonic stem cells (mESCs) (Dixon et al., 2012) and CH12-LX cells (Rao et al., 2014). While the larger domain structure was similar between cell lines, cell type-specific differences surrounding *Kit* emerged at the sub-domain level (Figure S6B).

CTCF has been implicated in the formation and maintenance of architectural domain boundaries. To define such boundaries, we computed the directionality index for each 40-kb bin in the genome by comparing the bin's interactions with the 50 bins (2 Mb) immediately upstream and downstream. We then applied the Hidden Markov Model approach proposed by Ren and colleagues (Dixon et al., 2012) to call 1,814 domains (median size 920 kb) in our control cells. Supporting the validity of this method, domain calls showed strong correlation with the visual representation of large-scale domain structure in the heatmaps (Figures 5A, 5B, and S6B). We note that these regions are likely comparable to megabase-scale TADs identified in published work (Dixon et al., 2012; Nora et al., 2012); however, we refer to them as domains and domain boundaries for simplicity.

We next determined whether domain boundary structure was affected by BRD2 depletion. Contact maps uncovered a spectrum of structural perturbations at individual boundaries, with the most common being an increase in ectopic, short-range interactions across boundaries (Figures 5A and 5B) and a blurring of the transitions between domains, including rare but dramatic boundary disruptions (Figure S6C). To determine whether domain boundary disruption was associated with

---

Figure 5. Effects of BRD2 Depletion on Boundary Strength

(A) Contact frequency heatmaps at 40 kb resolution of the *Sec16b* locus in control and BRD2 KO cells. Color bars range from low (white) to high (red) contact frequency. Black arrows point to areas of increased cross-boundary interactions in BRD2 KO cells. All heatmaps are framed by CTCF and BRD2 ChIP-seq tracks from control undifferentiated cells (−GATA1).

(B) Contact frequency heatmaps for the locus surrounding *Ier5* with areas of increased interaction frequency indicated by arrows.

(C) Top: cartoon illustrating boundary classifications. Bottom: histograms show the distribution of insulation score changes at boundaries called at 40 kb resolution, either between control cells and an average of the two BRD2 KO cell lines (top row, log2[averageBRD2 KO/control]) or between BRD2 KO 1 and BRD2 KO 2 (bottom row, log2[BRD2KO 1/BRD2KO 2]). The red line indicates no change.

(D) Average insulation scores centered on boundaries called at 40 kb resolution in control G1E-ER4 cells or either of the two BRD2 KO cell lines (1 and 2) at the indicated boundary category (error bars represent 2.5th and 97.5th percentiles of the distribution of the sampling means generated via a bootstrapping procedure in which the insulation score values of boundaries in a particular class were sampled with replacement 1,000 times).

See also Figure S6 and Tables S1–S3, S4, and S5.

altered transcriptional regulation, we measured expression of the genes surrounding the boundaries in Figures 5A and S6C (Figure S6G). Some genes (*Sec16b* from the locus in Figure 5A and *Xrcc4* from the locus in Figure S6C) were downregulated, while others were either unaffected (*Rasal2* and *Tmem167*) or increased (*Ralgps2* from the Figure 5A locus) upon BRD2 depletion (Figure S6G). We note that we cannot distinguish whether the changes in chromatin structure causally underlie the changes in gene expression at these regions, nor can we rule out a role for BRD2 in regulating transcription independent of any architectural function. However, the observation that transcription can be upregulated suggests that domain perturbations might lead to enhancer miswiring across domains.

To measure the contributions of CTCF and BRD2 to boundary formation genome-wide, we classified boundaries by their BRD2 and CTCF occupancy within ±20 kb of the boundary position (Table S4). We called a set of high-confidence BRD2 and CTCF peaks in control G1E-ER4 cells, and we intersected these to find regions of the genome with high BRD2/CTCF occupancy. To distinguish sites uniquely occupied by either BRD2 or CTCF, we additionally called a set of low-confidence peaks, and we used these to rule out occupancy by the other factor. Boundaries were parsed into the following categories: (1) those with BRD2 without CTCF, (2) those in which CTCF and BRD2 were directly co-localized with no additional CTCF-only or BRD2-only sites (CTCF/BRD2 co-occupied only), (3) those in which CTCF and BRD2 were directly co-localized with additional BRD2-only sites (CTCF/BRD2 co-occupied with BRD2-only), (4) those in which CTCF and BRD2 were directly co-localized with additional CTCF-only sites (CTCF/BRD2 co-occupied with CTCF-only), (5) those with CTCF without BRD2, and (6) those that lacked either protein (neither) (see Figure 5C; Tables S2 and S4). All boundaries containing multiple classes of co-localized sites and/or peaks that did not meet our stringent thresholds were classified as "other" and were not included in the analysis. We validated this strategy by examining ChIP-seq peak enrichment in each category (Figure S6H). Our classification scheme suggests domain boundaries exhibit diverse signatures of BRD2 and CTCF occupancy.

To quantify changes in boundary strength upon BRD2 depletion, we computed the insulation score for every boundary within a given classification in control and BRD2 KO cells. The insulation score aggregates the number of interactions in a sliding window along the linear chromosome (Crane et al., 2015). A less well-insulated region, characterized by a high frequency of interactions passing over, would have a concomitantly high insulation score. A 5-by-5 bin (200-by-200 kb) sliding window was used to quantify insulation scores across boundaries. The relative change (log2 scale) in insulation score for each boundary between control and BRD2 KO cells was plotted as a histogram for each boundary category (Figure 5C, top panels). We compared these to empiric null distributions created by computing the log2 fold changes in boundary insulation score between BRD2 KO 1 and BRD2 KO 2 cells (Figure 5C, bottom panels). When each category was examined individually, we observed an increase in insulation score (corresponding to a gain of interactions across the boundary) at boundary classes containing BRD2 (Figure 5C; Table S2). The boundaries most affected by

BRD2 loss had CTCF/BRD2 co-occupied sites with additional BRD2-only sites. Effects were also present at boundaries containing BRD2 without CTCF, boundaries with CTCF/BRD2 co-occupied only sites, or CTCF/BRD2 co-occupied with CTCF-only sites (Figure 5C; Table S2). Importantly, CTCF-only boundaries and boundaries devoid of CTCF and BRD2 were essentially unperturbed in the BRD2 KO cells, suggesting the effect is specific to BRD2-occupied regions. The shift in insulation score was mostly driven by relatively modest changes at a substantial proportion of BRD2/CTCF boundaries rather than large alterations at a small fraction (Figure 5C). These data indicate that BRD2 contributes to the demarcation of domain boundaries.

To better understand the link between BRD2 and higher-order chromatin architecture, we visualized the average insulation scores in each boundary category as line graphs in which boundaries appear as local minima. Again, the most perturbed boundaries in BRD2-depleted cells were those containing BRD2, as evidenced by an increase in insulation score (or interaction frequency) across boundaries. Boundaries lacking BRD2 remained intact (Figure 5D), suggesting that BRD2 potentiates domain boundary integrity. To ensure that our results were robust to the resolution of our HiC data, we also conducted our insulation score analyses at 10-kb resolution (Tables S3 and S5), and we observed very similar effects of BRD2 depletion on the shifts in insulation score (Figures S6D–S6F). Even at higher resolution, we could still observe a substantial number of significantly weakened BRD2-occupied boundaries, whereas the impact on boundaries without BRD2 was negligible (Figures S6D–S6F; Table S3). We also observed that, at both resolutions, the strongest effects on insulation score tended to occur at domain boundaries.

We note that the datasets generated in this study will allow for much broader analyses of domain architecture, as well as intra-domain looping interactions at various length scales. Space constraints precluded us from presenting more permutations of the analyses, but the datasets are publicly available for further study. Taken together, our results indicate that BRD2 is required to maintain chromatin domain integrity and may operate through both CTCF-dependent and CTCF-independent mechanisms.

DISCUSSION

Our work uncovers a functional link between the architectural protein CTCF and the BET protein BRD2 in large-scale chromatin organization. BRD2 and to a lesser extent BRD3 are co-localized with CTCF genome-wide. Loss of BRD2 is accompanied by widespread alterations in domain boundaries, characterized by an increase in ectopic cross-boundary chromatin contacts, particularly involving regions at which both BRD2 and CTCF are present. While BET proteins are well known for their ability to recruit transcriptional co-regulators to acetylated chromatin, a role in maintaining domain-level chromatin architecture has not previously been appreciated.

BET proteins can facilitate chromatin occupancy of DNA-binding proteins (Lamonica et al., 2011; Stonestrom et al., 2015), but CTCF binding occurs normally in the absence of BRD2, suggesting that BRD2 functions downstream of CTCF. However, given
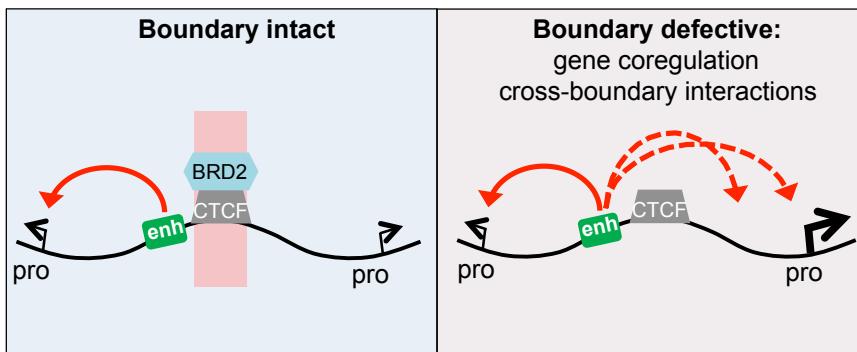
Figure 6. BRD2 Maintains the Integrity of Transcriptional and Architectural Boundaries

Model depicting potential mechanisms by which BRD2 regulates boundary function with CTCF. CTCF recruits BRD2 to co-bound sites. Loss of BRD2 leads to both inappropriate co-regulation of genes normally insulated from one another and an increase in ectopic interactions across domain boundaries.

that BRD3 can also occupy CTCF sites, a compensatory role for BRD3 in aiding CTCF DNA binding remains possible. Indeed, the yeast BET homolog Bdf2 responds to Bdf1 loss by shifting its occupancy to Bdf1 sites (Durant and Pugh, 2007).

BET proteins associate with chromatin not only via acetylated histones but also acetylated transcription factors (Asangani et al., 2014; Huang et al., 2009; Lamonica et al., 2011; Roe et al., 2015; Shi et al., 2014). Whether the CTCF-BRD2 interaction is direct and whether it requires acetylation of CTCF or another CTCF cofactor remain open questions. Conventional co-immunoprecipitation experiments failed to detect an association between BRD2 and CTCF (unpublished data), but it is possible that the interaction occurs in the context of chromatin or is sensitive to extraction conditions.

Disruption of one of two CTCF sites that normally separate the erythroid-specific gene *Slc25a37* from its independently regulated neighbor *Entpd4* enabled the activity of the GATA1-controlled *Slc25a37* enhancer to inappropriately contact the *Entpd4* promoter and activate *Entpd4* expression in a GATA1-dependent fashion. Thus, at this site, CTCF functions as part of a boundary mechanism that constrains enhancer action. Since BRD2 occupancy cannot be easily disrupted locally, we examined the activity of this boundary in BRD2-depleted cells. BRD2 deficiency does not augment *Entpd4* expression and, thus, does not phenocopy the bulk transcriptional effects of the specific CTCF site mutation. However, the interpretation of this experiment is confounded by the reduced expression of *Slc25a37* in BRD2-depleted cells and the decreased contacts between the *Slc25a37* enhancer and *Slc25a37* promoter, indicating that the enhancer is less active in this context. Population-based gene expression studies are limited in that they might obscure transcriptional relationships that occur at the single-cell level. We therefore used single-molecule mRNA FISH to quantify the number of *Slc25a37* and *Entpd4* mRNA molecules per cell, and we measured their correlation. CTCF site mutation at the boundary as well as global BRD2 loss elevated the correlation in expression of both genes. Although global BRD2 loss could have indirect effects, such as the depletion of other factors involved in maintaining this boundary element, these observations suggest that, in the absence of CTCF or BRD2, the genes share transcriptional regulatory mechanisms. This supports the model that BRD2 contributes to CTCF's ability to limit enhancer activity, and it highlights the power of RNA FISH to probe correlative relationships of gene expression. The effect of BRD depletion on the *Slc25a37/Entpd4* boundary was not as strong as that following the disruption of CTCF. This might be due to partial compensation by BRD3 or to CTCF exerting some of its activity independently of BET proteins. Conversely, a fraction of BRD2-occupied sites lacks CTCF, indicating that BRD2 can also function in a CTCF-independent manner.

CTCF's role as a boundary or insulator element may be related to its ability to organize higher-order chromatin architecture (Phillips and Corces, 2009). CTCF anchors long-range chromatin interactions (Phillips-Cremins et al., 2013), and its disruption increases inter-domain chromatin contacts (Narendra et al., 2015; Zuin et al., 2014). Our HiC analysis revealed that loss of BRD2 similarly increases contact frequencies across BRD2-occupied domain boundaries genome-wide, demonstrating that BRD2 globally contributes to domain insulation. The mechanism by which BRD2 functions in this context remains an open question. CTCF is thought to form chromatin loops, which might structurally partition the genome. BRD2 has a putative dimerization domain (Garcia-Gutierrez et al., 2012) that could aid in forging chromatin contacts. Alterations in chromatin interactions in BRD2-deficient cells were accompanied by changes in gene expression, suggesting at a correlative level that the relaxing of boundaries enables enhancer rewiring or enhancer promiscuity. However, given that these experiments involved sustained BRD2 depletion, it is also possible that the architectural changes are a corollary of disrupted transcription or indirectly result from loss of factors involved in looping or insulation.

Previous reports suggested that BRD2, BRD3, and BRD4 bind to overlapping sets of transcriptional regulatory complexes (Dawson et al., 2011; Rahman et al., 2011) and exhibit significantly overlapping occupancy at active genes (Anders et al., 2014). Our study underscores key differences between the functions of individual BET family members by showing that, in contrast to BRD2 and BRD3, BRD4 is not enriched at CTCF sites. Also, a prior study in erythroid cells found that BRD2, BRD3, and BRD4 exhibit distinct localization patterns in relation to the erythroid transcription factor GATA1 and, thus, are likely to exert disparate functions in promoting erythroid maturation (Stonestrom et al., 2015). That study, similar to the present one, also suggested functional overlap between BRD2 and BRD3. Several important questions remain for future studies. What is the molecular basis underlying the selective recruitment of BET proteins by distinct sets of transcription factors? Do CTCF sites co-bound by BRD2 have functions disparate from

CTCF sites that do not recruit BRD2? To what degree can BRD2 form boundaries or architectural domains at places where CTCF is not also present? And, more broadly, what distinguishes CTCF and BRD2 sites at boundaries from those that lie in non-boundary chromatin? Finally, to what extent does this function of BRD2 relate to the organization of smaller scale structures, such as enhancer-promoter loops?

In concert, our findings indicate that BRD2 acts to augment the boundary function of CTCF both by limiting the spread or range of enhancer activity and by physically preventing the formation of cross-boundary contacts genome-wide (Figure 6). This implies that pharmacologic BET inhibitors may alter transcriptional regulation on a broader scale as a result of domain boundary perturbations.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- CONTACT FOR REAGENT AND RESOURCE SHARING
  - Experimental Model and Subject Details
- METHOD DETAILS
  - Retroviral infection of murine cells
  - CRISPR/Cas9 genome editing
  - ChIP
  - RT-qPCR
  - 3C
  - Immunoblotting
  - Single-molecule RNA FISH imaging
  - ChIP-seq
  - In situ HiC
- QUANTIFICATION AND STATISTICAL ANALYSIS
  - ChIP-seq analysis: binding correlation and motifs
  - RNA FISH image analysis
  - HiC read alignment and processing
  - Identifying BRD2, CTCF sites to classify boundaries
  - Contact domain calling and categorization
  - Insulation score analysis
- DATA AND SOFTWARE AVAILABILITY

SUPPLEMENTAL INFORMATION

Supplemental Information includes six figures and six tables and can be found with this article online at http://dx.doi.org/10.1016/j.molcel.2017.02.027.

AUTHOR CONTRIBUTIONS

S.C.H., J.E.P.-C., and G.A.B. conceived the study and designed experiments. S.C.H., C.R.B., C.R.E., P.H., M.T.W., and C.A.K. performed experiments. S.C.H., A.J.S., C.R.B., T.G.G., D.J.E., P.E., B.G., R.C.H., A.R., J.E.P.-C., and G.A.B. analyzed data. S.C.H., J.E.P.-C., and G.A.B. wrote the manuscript with input from all authors.

ACKNOWLEDGMENTS

REFERENCES

Anders, L., Guenther, M.G., Qi, J., Fan, Z.P., Marineau, J.J., Rahl, P.B., Lovén, J., Sigova, A.A., Smith, W.B., Lee, T.I., et al. (2014). Genome-wide localization of small molecules. Nat. Biotechnol. *32*, 92–96.

Asangani, I.A., Dommeti, V.L., Wang, X., Malik, R., Cieslik, M., Yang, R., Escara-Wilke, J., Wilder-Romans, K., Dhanireddy, S., Engelke, C., et al. (2014). Therapeutic targeting of BET bromodomain proteins in castration-resistant prostate cancer. Nature *510*, 278–282.

Belkina, A.C., Nikolajczyk, B.S., and Denis, G.V. (2013). BET protein function is required for inflammation: Brd2 genetic disruption and BET inhibitor JQ1 impair mouse macrophage inflammatory responses. J. Immunol. *190*, 3670–3678.

Berkovits, B.D., and Wolgemuth, D.J. (2011). The first bromodomain of the testis-specific double bromodomain protein Brdt is required for chromocenter organization that is modulated by genetic background. Dev. Biol. *360*, 358–368.

Bolstad, B.M., Irizarry, R.A., Astrand, M., and Speed, T.P. (2003). A comparison of normalization methods for high density oligonucleotide array data based on variance and bias. Bioinformatics *19*, 185–193.

Bullard, J.H., Purdom, E., Hansen, K.D., and Dudoit, S. (2010). Evaluation of statistical methods for normalization and differential expression in mRNA-seq experiments. BMC Bioinformatics *11*, 94.

Chernukhin, I., Shamsuddin, S., Kang, S.Y., Bergström, R., Kwon, Y.W., Yu, W., Whitehead, J., Mukhopadhyay, R., Docquier, F., Farrar, D., et al. (2007). CTCF interacts with and recruits the largest subunit of RNA polymerase II to CTCF target sites genome-wide. Mol. Cell. Biol. *27*, 1631–1648.

Cong, L., and Zhang, F. (2014). Genome engineering using CRISPR-Cas9 system. In Methods in Molecular Biology (Springer), pp. 197–217.

Crane, E., Bian, Q., McCord, R.P., Lajoie, B.R., Wheeler, B.S., Ralston, E.J., Uzawa, S., Dekker, J., and Meyer, B.J. (2015). Condensin-driven remodelling of X chromosome topology during dosage compensation. Nature *523*, 240–244.

Dawson, M.A., Prinjha, R.K., Dittmann, A., Giotopoulos, G., Bantscheff, M., Chan, W.-I., Robson, S.C., Chung, C.-W., Hopf, C., Savitski, M.M., et al. (2011). Inhibition of BET recruitment to chromatin as an effective treatment for MLL-fusion leukaemia. Nature *478*, 529–533.

de Wit, E., Vos, E.S.M., Holwerda, S.J.B., Valdes-Quezada, C., Verstegen, M.J.A.M., Teunissen, H., Splinter, E., Wijchers, P.J., Krijger, P.H.L., and de Laat, W. (2015). CTCF binding polarity determines chromatin looping. Mol. Cell *60*, 676–684.

Deng, W., Lee, J., Wang, H., Miller, J., Reik, A., Gregory, P.D., Dean, A., and Blobel, G.A. (2012). Controlling long-range genomic interactions at a native locus by targeted tethering of a looping factor. Cell *149*, 1233–1244.

Denis, G.V., McComb, M.E., Faller, D.V., Sinha, A., Romesser, P.B., and Costello, C.E. (2006). Identification of transcription complexes that contain the double bromodomain protein Brd2 and chromatin remodeling machines. J. Proteome Res. *5*, 502–511.

Devaiah, B.N., Case-Borden, C., Gegonne, A., Hsu, C.H., Chen, Q., Meerzaman, D., Dey, A., Ozato, K., and Singer, D.S. (2016). BRD4 is a histone

acetyltransferase that evicts nucleosomes from chromatin. Nat. Struct. Mol. Biol. 23, 540–548.

Dixon, J.R., Selvaraj, S., Yue, F., Kim, A., Li, Y., Shen, Y., Hu, M., Liu, J.S., and Ren, B. (2012). Topological domains in mammalian genomes identified by analysis of chromatin interactions. Nature 485, 376–380.

Durant, M., and Pugh, B.F. (2007). NuA4-directed chromatin transactions throughout the Saccharomyces cerevisiae genome. Mol. Cell. Biol. 27, 5327–5335.

Femino, A.M., Fay, F.S., Fogarty, K., and Singer, R.H. (1998). Visualization of single RNA transcripts in situ. Science 280, 585–590.

Feng, J., Liu, T., Qin, B., Zhang, Y., and Liu, X.S. (2012). Identifying ChIP-seq enrichment using MACS. Nat. Protoc. 7, 1728–1740.

Filippakopoulos, P., Qi, J., Picaud, S., Shen, Y., Smith, W.B., Fedorov, O., Morse, E.M., Keates, T., Hickman, T.T., Felletar, I., et al. (2010). Selective inhibition of BET bromodomains. Nature 468, 1067–1073.

Flavahan, W.A., Drier, Y., Liau, B.B., Gillespie, S.M., Venteicher, A.S., Stemmer-Rachamimov, A.O., Suvà, M.L., and Bernstein, B.E. (2016). Insulator dysfunction and oncogene activation in IDH mutant gliomas. Nature 529, 110–114.

Garcia-Gutierrez, P., Mundi, M., and Garcia-Dominguez, M. (2012). Association of bromodomain BET proteins with chromatin requires dimerization through the conserved motif B. J. Cell Sci. 125, 3671–3680.

Giardine, B., Riemer, C., Hardison, R.C., Burhans, R., Elnitski, L., Shah, P., Zhang, Y., Blankenberg, D., Albert, I., Taylor, J., et al. (2005). Galaxy: a platform for interactive large-scale genome analysis. Genome Res. 15, 1451–1455.

Guo, Y., Xu, Q., Canzio, D., Shou, J., Li, J., Gorkin, D.U., Jung, I., Wu, H., Zhai, Y., Tang, Y., et al. (2015). CRISPR inversion of CTCF sites alters genome topology and enhancer/promoter function. Cell 162, 900–910.

Hsiung, C.C.S., Bartman, C.R., Huang, P., Ginart, P., Stonestrom, A.J., Keller, C.A., Face, C., Jahn, K.S., Evans, P., Sankaranarayanan, L., et al. (2016). A hyperactive transcriptional state marks genome reactivation at the mitosis-G1 transition. Genes Dev. 30, 1423–1439.

Huang, B., Yang, X.D., Zhou, M.M., Ozato, K., and Chen, L.F. (2009). Brd4 coactivates transcriptional activation of NF-kappaB via specific binding to acetylated RelA. Mol. Cell. Biol. 29, 1375–1387.

Huang, J., Liu, X., Li, D., Shao, Z., Cao, H., Zhang, Y., Trompouki, E., Bowman, T.V., Zon, L.I., Yuan, G.-C., et al. (2016). Dynamic control of enhancer repertoires drives lineage and stage-specific transcription during hematopoiesis. Dev. Cell 36, 9–23.

Hughes, J.R., Roberts, N., McGowan, S., Hay, D., Giannoulatou, E., Lynch, M., De Gobbi, M., Taylor, S., Gibbons, R., and Higgs, D.R. (2014). Analysis of hundreds of cis-regulatory landscapes at high resolution in a single, high-throughput experiment. Nat. Genet. 46, 205–212.

Imakaev, M., Fudenberg, G., McCord, R.P., Naumova, N., Goloborodko, A., Lajoie, B.R., Dekker, J., and Mirny, L.A. (2012). Iterative correction of Hi-C data reveals hallmarks of chromosome organization. Nat. Methods 9, 999–1003.

Kellner, W.A., Van Bortle, K., Li, L., Ramos, E., Takenaka, N., and Corces, V.G. (2013). Distinct isoforms of the Drosophila Brd4 homologue are present at enhancers, promoters and insulator sites. Nucleic Acids Res. 41, 9274–9283.

Ladurner, A.G., Inouye, C., Jain, R., and Tjian, R. (2003). Bromodomains mediate an acetyl-histone encoded antisilencing function at heterochromatin boundaries. Mol. Cell 11, 365–376.

Lamonica, J.M., Deng, W., Kadauke, S., Campbell, A.E., Gamsjaeger, R., Wang, H., Cheng, Y., Billin, A.N., Hardison, R.C., Mackay, J.P., and Blobel, G.A. (2011). Bromodomain protein Brd3 associates with acetylated GATA1 to promote its chromatin occupancy at erythroid target genes. Proc. Natl. Acad. Sci. USA 108, E159–E168.

Langmead, B., and Salzberg, S.L. (2012). Fast gapped-read alignment with Bowtie 2. Nat. Methods 9, 357–359.

Li, L., Lyu, X., Hou, C., Takenaka, N., Nguyen, H.Q., Ong, C.-T., Cubeñas-Potts, C., Hu, M., Lei, E.P., Bosco, G., et al. (2015). Widespread rearrangement

of 3D chromatin organization underlies polycomb-mediated stress-induced silencing. Mol. Cell 58, 216–231.

Liu, T., Ortiz, J.A., Taing, L., Meyer, C.A., Lee, B., Zhang, Y., Shin, H., Wong, S.S., Ma, J., Lei, Y., et al. (2011). Cistrome: an integrative platform for transcriptional regulation studies. Genome Biol. 12, R83.

Lovén, J., Hoke, H.A., Lin, C.Y., Lau, A., Orlando, D.A., Vakoc, C.R., Bradner, J.E., Lee, T.I., and Young, R.A. (2013). Selective inhibition of tumor oncogenes by disruption of super-enhancers. Cell 153, 320–334.

Lupiáñez, D.G., Kraft, K., Heinrich, V., Krawitz, P., Brancati, F., Klopocki, E., Horn, D., Kayserili, H., Opitz, J.M., Laxova, R., et al. (2015). Disruptions of topological chromatin domains cause pathogenic rewiring of gene-enhancer interactions. Cell 161, 1012–1025.

Machanick, P., and Bailey, T.L. (2011). MEME-ChIP: motif analysis of large DNA datasets. Bioinformatics 27, 1696–1697.

Meyer, L.R., Zweig, A.S., Hinrichs, A.S., Karolchik, D., Kuhn, R.M., Wong, M., Sloan, C.A., Rosenbloom, K.R., Roe, G., Rhead, B., et al. (2013). The UCSC Genome Browser database: extensions and updates 2013. Nucleic Acids Res. 41, D64–D69.

Narendra, V., Rocha, P.P., An, D., Raviram, R., Skok, J.A., Mazzoni, E.O., and Reinberg, D. (2015). CTCF establishes discrete functional chromatin domains at the Hox clusters during differentiation. Science 347, 1017–1021.

Nora, E.P., Lajoie, B.R., Schulz, E.G., Giorgetti, L., Okamoto, I., Servant, N., Piolot, T., van Berkum, N.L., Meisig, J., Sedat, J., et al. (2012). Spatial partitioning of the regulatory landscape of the X-inactivation centre. Nature 485, 381–385.

Padovan-Merhar, O., Nair, G.P., Biaesch, A.G., Mayer, A., Scarfone, S., Foley, S.W., Wu, A.R., Churchman, L.S., Singh, A., and Raj, A. (2015). Single mammalian cells compensate for differences in cellular volume and DNA copy number through independent global transcriptional mechanisms. Mol. Cell 58, 339–352.

Paralkar, V.R., Taborda, C.C., Huang, P., Yao, Y., Kossenkov, A.V., Prasad, R., Luan, J., Davies, J.O.J., Hughes, J.R., Hardison, R.C., et al. (2016). Unlinking an lncRNA from its associated cis element. Mol. Cell 62, 104–110.

Parelho, V., Hadjur, S., Spivakov, M., Leleu, M., Sauer, S., Gregson, H.C., Jarmuz, A., Canzonetta, C., Webster, Z., Nesterova, T., et al. (2008). Cohesins functionally associate with CTCF on mammalian chromosome arms. Cell 132, 422–433.

Phillips, J.E., and Corces, V.G. (2009). CTCF: master weaver of the genome. Cell 137, 1194–1211.

Phillips-Cremins, J.E., Sauria, M.E.G., Sanyal, A., Gerasimova, T.I., Lajoie, B.R., Bell, J.S.K., Ong, C.-T., Hookway, T.A., Guo, C., Sun, Y., et al. (2013). Architectural protein subclasses shape 3D organization of genomes during lineage commitment. Cell 153, 1281–1295.

Quinlan, A.R., and Hall, I.M. (2010). BEDTools: a flexible suite of utilities for comparing genomic features. Bioinformatics 26, 841–842.

Rahman, S., Sowa, M.E., Ottinger, M., Smith, J.A., Shi, Y., Harper, J.W., and Howley, P.M. (2011). The Brd4 extraterminal domain confers transcription activation independent of pTEFb by recruiting multiple proteins, including NSD3. Mol. Cell. Biol. 31, 2641–2652.

Raj, A., Peskin, C.S., Tranchina, D., Vargas, D.Y., and Tyagi, S. (2006). Stochastic mRNA synthesis in mammalian cells. PLoS Biol. 4, e309.

Raj, A., van den Bogaard, P., Rifkin, S.A., van Oudenaarden, A., and Tyagi, S. (2008). Imaging individual mRNA molecules using multiple singly labeled probes. Nat. Methods 5, 877–879.

Raj, A., Rifkin, S.A., Andersen, E., and van Oudenaarden, A. (2010). Variability in gene expression underlies incomplete penetrance. Nature 463, 913–918.

Rao, S.S.P., Huntley, M.H., Durand, N.C., Stamenova, E.K., Bochkov, I.D., Robinson, J.T., Sanborn, A.L., Machol, I., Omer, A.D., Lander, E.S., and Aiden, E.L. (2014). A 3D map of the human genome at kilobase resolution reveals principles of chromatin looping. Cell 159, 1665–1680.

Roe, J.-S., Mercan, F., Rivera, K., Pappin, D.J., and Vakoc, C.R. (2015). BET bromodomain inhibition suppresses the function of hematopoietic transcription factors in acute myeloid leukemia. Mol. Cell 58, 1028–1039.

Seitan, V.C., Faure, A.J., Zhan, Y., McCord, R.P., Lajoie, B.R., Ing-Simmons, E., Lenhard, B., Giorgetti, L., Heard, E., Fisher, A.G., et al. (2013). Cohesin-based chromatin interactions enable regulated gene expression within preexisting architectural compartments. Genome Res. 23, 2066–2077.

Servant, N., Varoquaux, N., Lajoie, B.R., Viara, E., Chen, C.-J., Vert, J.-P., Heard, E., Dekker, J., and Barillot, E. (2015). HiC-Pro: an optimized and flexible pipeline for Hi-C data processing. Genome Biol. 16, 259.

Shang, E., Nickerson, H.D., Wen, D., Wang, X., and Wolgemuth, D.J. (2007). The first bromodomain of Brdt, a testis-specific member of the BET sub-family of double-bromodomain-containing proteins, is essential for male germ cell differentiation. Development 134, 3507–3515.

Shaw, G.C., Cope, J.J., Li, L., Corson, K., Hersey, C., Ackermann, G.E., Gwynn, B., Lambert, A.J., Wingert, R.A., Traver, D., et al. (2006). Mitoferrin is essential for erythroid iron assimilation. Nature 440, 96–100.

Shi, J., Wang, Y., Zeng, L., Wu, Y., Deng, J., Zhang, Q., Lin, Y., Li, J., Kang, T., Tao, M., et al. (2014). Disrupting the interaction of BRD4 with diacetylated Twist suppresses tumorigenesis in basal-like breast cancer. Cancer Cell 25, 210–225.

Sofueva, S., Yaffe, E., Chan, W.-C., Georgopoulou, D., Vietri Rudan, M., Mira-Bontenbal, H., Pollard, S.M., Schroth, G.P., Tanay, A., and Hadjur, S. (2013). Cohesin-mediated interactions organize chromosomal domain architecture. EMBO J. 32, 3119–3129.

Stonestrom, A.J., Hsu, S.C., Jahn, K.S., Huang, P., Keller, C.A., Giardine, B.M., Kadauke, S., Campbell, A.E., Evans, P., Hardison, R.C., and Blobel, G.A. (2015). Functions of BET proteins in erythroid gene expression. Blood 125, 2825–2834.

Storey, J.D. (2002). A direct approach to false discovery rates. J. R. Stat. Soc. Series B Stat. Methodol 64, 479–498.

Surface, L.E., Fields, P.A., Subramanian, V., Behmer, R., Udeshi, N., Peach, S.E., Carr, S.A., Jaffe, J.D., and Boyer, L.A. (2016). H2A.Z.1 monoubiquitylation antagonizes BRD2 to maintain poised chromatin in ESCs. Cell Rep. 14, 1142–1155.

Symmons, O., Uslu, V.V., Tsujimura, T., Ruf, S., Nassari, S., Schwarzer, W., Ettwiller, L., and Spitz, F. (2014). Functional and topological characteristics of mammalian regulatory domains. Genome Res. 24, 390–400.

Szabó, P.E., Tang, S.H.E., Silva, F.J., Tsark, W.M.K., and Mann, J.R. (2004). Role of CTCF binding sites in the Igf2/H19 imprinting control region. Mol. Cell. Biol. 24, 4791–4800.

Tripic, T., Deng, W., Cheng, Y., Zhang, Y., Vakoc, C.R., Gregory, G.D., Hardison, R.C., and Blobel, G.A. (2009). SCL and associated proteins distinguish active from repressive GATA transcription factor complexes. Blood 113, 2191–2201.

Vardabasso, C., Gaspar-Maia, A., Hasson, D., Pünzeler, S., Valle-Garcia, D., Straub, T., Keilhauer, E.C., Strub, T., Dong, J., Panda, T., et al. (2015). Histone variant H2A.Z.2 mediates proliferation and drug sensitivity of malignant melanoma. Mol. Cell 59, 75–88.

Wang, R., Li, Q., Helfer, C.M., Jiao, J., and You, J. (2012). Bromodomain protein Brd4 associated with acetylated chromatin is important for maintenance of higher-order chromatin structure. J. Biol. Chem. 287, 10738–10752.

Weiss, M.J., Yu, C., and Orkin, S.H. (1997). Erythroid-cell-specific properties of transcription factor GATA-1 revealed by phenotypic rescue of a gene-targeted cell line. Mol. Cell. Biol. 17, 1642–1651.

Welch, J.J., Watts, J.A., Vakoc, C.R., Yao, Y., Wang, H., Hardison, R.C., Blobel, G.A., Chodosh, L.A., and Weiss, M.J. (2004). Global regulation of erythroid gene expression by transcription factor GATA-1. Blood 104, 3136–3147.

Wendt, K.S., Yoshida, K., Itoh, T., Bando, M., Koch, B., Schirghuber, E., Tsutsumi, S., Nagae, G., Ishihara, K., Mishiro, T., et al. (2008). Cohesin mediates transcriptional insulation by CCCTC-binding factor. Nature 451, 796–801.

Wickham, H. (2009). ggplot2: Elegant Graphics for Data Analysis (Springer).

Wu, S.-Y., Lee, A.-Y., Lai, H.-T., Zhang, H., and Chiang, C.-M. (2013). Phospho switch triggers Brd4 chromatin binding and activator recruitment for gene-specific targeting. Mol. Cell 49, 843–857.

Zhang, Y., Liu, T., Meyer, C.A., Eeckhoute, J., Johnson, D.S., Bernstein, B.E., Nusbaum, C., Myers, R.M., Brown, M., Li, W., and Liu, X.S. (2008). Model-based analysis of ChIP-seq (MACS). Genome Biol. 9, R137.

Zhao, R., Nakamura, T., Fu, Y., Lazar, Z., and Spector, D.L. (2011). Gene bookmarking accelerates the kinetics of post-mitotic transcriptional re-activation. Nat. Cell Biol. 13, 1295–1304.

Zuin, J., Dixon, J.R., van der Reijden, M.I.J.A., Ye, Z., Kolovos, P., Brouwer, R.W.W., van de Corput, M.P.C., van de Werken, H.J.G., Knoch, T.A., van IJcken, W.F.J., et al. (2014). Cohesin and CTCF differentially affect chromatin architecture and gene expression in human cells. Proc. Natl. Acad. Sci. USA 111, 996–1001.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Antibodies		
Rabbit polyclonal anti-CTCF	Millipore	Cat#07-729; Lot#2517762; RRID: AB_441965
Rabbit polyclonal anti-BRD2	Bethyl	Cat#A302-583A; Lot#4; RRID: AB_2034829
Rabbit polyclonal anti-SMC1	Bethyl	Cat#A300-055A; Lot#5; RRID: AB_2192467
Mouse monoclonal anti-HA (clone 12CA5)	Laboratory of Gerd Blobel	N/A
Rabbit monoclonal anti-BRD2 (clone D89B4)	Cell Signaling Technology	Cat#5848
Mouse monoclonal anti-B-actin-Peroxidase (clone AC-15)	Sigma	Cat#A3854; RRID: AB_262011
Chemicals, Peptides, and Recombinant Proteins		
Protein A agarose beads	Thermo Fisher	Cat#15918014
Protein G agarose beads	Thermo Fisher	Cat#15920010
iScript Reverse Transcription Supermix	Bio-Rad	Cat#1708841
Trizol	Thermo Fisher	Cat#15596026
Power SYBR Green PCR Master Mix	Thermo Fisher	Cat#4367660
BglII	New England BioLabs	Cat#R0144S
T4 DNA Ligase	New England BioLabs	Cat#M0202S
DpnII	New England BioLabs	Cat#R0543M
Biotin-14-dATP	Thermo Fisher	Cat#19524-016
Dynabeads MyOne Streptavidin C1 beads	Thermo Fisher	Cat#65002
AMPureXP beads	Beckman Coulter	Cat#A63881
Critical Commercial Assays		
QIAGEN PCR Purification Kit	QIAGEN	Cat#28106
QIAGEN RNeasy Kit	QIAGEN	Cat#74106
TruSeq ChIP Sample Preparation Kit	Illumina	Cat# IP 202-1012
NEBNextQ5 Hot Start HiFi PCR Master Mix	New England BioLabs	Cat#M0543L
NEBNext DNA Library Prep Master Mix for Illumina	New England BioLabs	Cat#E6040S
NEBNext Multiplex Oligos for Illumina Set1	New England BioLabs	Cat#E7335S
Cell Line Nucleofector Kit R	Lonza	Cat#VVCA-1001
Deposited Data		
ChIP-seq sequencing files	This paper	GEO: GSE80527
HiC sequencing files	This paper	GEO: GSE95476
Experimental Models: Cell Lines		
G1E-ER4	Laboratory of Mitchell J. Weiss	Weiss et al., 1997
Oligonucleotides		
ChIP-qPCR primers	This paper	Table S6
qRT-PCR pimers	This paper	Table S6
3C primers	This paper	Table S6
BRD2_gRNA: ATTAGGACAATATCATCGGT	This paper	N/A
Bcl11a_CTCF_gRNA: ATGAAGCGGGCGCCATCGTG	This paper	N/A
Bcl11a_CTCF_Mut_RepairTemplate: CATTCTCGTGG TCCCGCCGTGCCACCTGCACTGCAGTACCAGACT CTGGCatatgttTttCGCCCGCTTCATGCAGACCCGGA TCACGGCGGCGGGGCAGACAAAGGCGG	This paper	N/A
Slc25a37_CTCFa_Mut#1_gRNA: ACTCTTCTTTGAGC CACTAG	This paper	N/A

(Continued on next page)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Slc25a37_CTCFa_Mut#2_gRNA: CAAACATCAGCGCC CTCTAG	This paper	N/A
Single-molecule RNA FISH probes	This paper	Table S6
Recombinant DNA		
MigR1 HA-BRD2	Laboratory of Gerd A. Blobel	Stonestrom et al., 2015
MigR1 Cas9 mCherry	Laboratory of Gerd A. Blobel	Stonestrom et al., 2015
MigR1 guide RNA GFP	Laboratory of Gerd A. Blobel	Stonestrom et al., 2015
pX330 GFP	Paralkar et al., 2016	Addgene 42230, Vikram Paralkar
BAC RP23-14E1 (Clone ID: 14E1)	Thermo Fisher	Cat#RPCI23.C
Software and Algorithms		
CRISPR Design Tool	Cong and Zhang., 2014	http://crispr.mit.edu
UCSC Genome Browser	Meyer et al., 2013	http://genome.ucsc.edu
Cistrome	Liu et al., 2011	http://cistrome.org/
Galaxy	Giardine et al., 2005	https://galaxyproject.org/
Bowtie2	Langmead and Salzberg., 2012	http://bowtie-bio.sourceforge.net/bowtie2/index.shtml
MACS2	Zhang et al., 2008	https://github.com/taoliu/MACS
BEDTools	Quinlan and Hall., 2010	https://github.com/arq5x/bedtools2
MEME Suite	Machanick and Bailey, 2011	meme-suite.org
Single-molecule RNA FISH image analysis MATLAB software	Laboratory of Arjun Raj	https://bitbucket.org/arjunrajlaboratory/rajlabimagetools/wiki/Home
ICED matrix balancing	Imakaev et al., 2012	https://github.com/hiclib/iced
Directionality Index and Hidden Markov Modeling	Dixon et al., 2012	http://chromosome.sdsc.edu/mouse/hi-c/download.html
Insulation Score	Crane et al., 2015	http://www.nature.com/nature/journal/v523/n7559/full/nature14450.html
Quantile normalization	Bolstad et al., 2003	https://www.ncbi.nlm.nih.gov/pubmed/12538238

CONTACT FOR REAGENT AND RESOURCE SHARING

Further information and requests for resources and reagents should be directed to and will be fulfilled by the Lead Contact, Gerd A. Blobel (blobel@email.chop.edu).

Experimental Model and Subject Details

G1E-ER4 cells are a sub-line of G1E cells, (derived from $GATA1^{-/-}$ murine embryonic stem cells; Weiss et al., 1997), which express GATA1 fused to the ligand binding domain of the estrogen receptor (GATA1-ER) introduced by a retrovirus (Weiss et al., 1997). GATA1 activation and erythroid maturation are induced by the addition of 100nM estradiol to the media for 24 hr ("+GATA1"). Where indicated, cells were used in the undifferentiated ("-GATA1") state, and are referred to as "uninduced." G1E-ER4 cells were grown in IMDM+15% FBS, penicillin/streptomycin, Kit ligand, monothioglycerol and epoetin alpha in a standard tissue culture incubator at 37C with 5% CO_2.

For G1E-ER4 lines expressing HA-BRD2, cells were infected with MigR1-HA-BRD2 retrovirus, which contains IRES-GFP. HA-BRD2 was cloned into the MigR1 vector from mouse cDNA (Stonestrom et al., 2015). Following infection, cells were expanded and GFP+ cells were sorted on a FACS Aria II (BD Biosciences) to obtain pure populations.

BRD2-depleted ("BRD2 KO") cell lines were derived by cotransfection of G1E-ER4 cells with MigR1-Cas9 (mCherry) and MigR1-guide RNA (GFP) targeting the *Brd2* locus (BRD2_gRNA) using the Amaxa electroporator (Lonza). Single mCherry+/GFP+ cells were sorted 24 hr later into 96-well plates. Cells were expanded and clones were screened with a combination of PCR, Sanger sequencing and immunoblot.

Bcl11a-CTCF_Mut cell line was generated by co-transfection of G1E-ER4 cells with pX330-U6-Chimeric_BB-CBh-hSpCas9 Cas9/guide RNA plasmid (with GFP) (Paralkar et al., 2016) (a gift from Feng Zhang and Vikram Paralkar, Addgene plasmid 42230) with the Bcl11a_CTCF_gRNA cloned in, along with a mutated repair template (Bcl11a_CTCF_Mut_RepairTemplate, ssODN, 4nM Ultramer,

IDT). Single cells were sorted, expanded as above, and screened using a combination of PCR, restriction digest and Sanger sequencing.

Slc25a37_CTCFa_Mut cell lines were derived by transfection with the pX330 Cas9/guide RNA plasmid as above with Slc25a37_CTCFa_Mut#1 or Mut#2 guide RNAs cloned in, and sorted and expanded as above. Clones were screened using PCR and Sanger sequencing, followed by ChIP-qPCR to confirm CTCF binding disruption. Allele sequences were determined by TOPO-TA cloning a PCR product encompassing the edited region, followed by sequencing.

See Key Resources Table for guide RNA and repair template sequences.

METHOD DETAILS

Retroviral infection of murine cells

To generate G1E-ER4 cells stably expressing HA-BRD2, cells were infected with a MigR1-HA-BRD2 retrovirus. Retrovirus was produced as previously described in HEK293T (Tripic et al., 2009). Briefly, HEK293T were plated, grown to ~90% confluence in a 10cm dish, and transfected using the calcium phosphate method to co-introduce MigR1 and pCL-Eco packaging plasmids. 500ul of 2X transfection buffer (2X transfection buffer: 0.5mL 0.5M HEPES pH7.1, 4.05mL dH$_2$O, 0.45mL 2M NaCl, 100ul of 100mM Na$_2$HPO$_4$) was mixed with 500ul DNA mixture (62.5ul 2M CaCl$_2$, 50ul 10X NTE, 15ug MigR1 plasmid, 15ug pCL-Eco, up to 500ul with dH2O) drop-wise. This mixture was then added to cells and incubated at 37C. Media was changed 6 hr post-transfection. Viral supernatant was harvested 48 hr post-transfection. For retroviral infection, ~1-2 million G1E-ER4 cells were plated in 6-well plates with 1mL of media and 1mL of viral supernatant. The cationic polymer polybrene (8ug/ul) and 10mM HEPES were added and cells were spun for 90 min at 3200RPM at room temperature. Following infection, cells were replated in fresh media and expanded for sorting.

CRISPR/Cas9 genome editing

G1E-ER4 cells were transfected with Cas9 and guide RNA plasmids or a combined vector harboring both (see Experimental Models and Subject Details above), as well as a repair template if indicated, using the Amaxa II electroporator (Lonza) with program G-016 and reagent Kit R. Single cells were sorted into 96 well plates using a FACS Aria II (BD Biosciences), expanded, and screened by PCR, DNA sequencing, and ChIP or immunoblot as appropriate. Individual allele sequences were determined by PCR of the region surrounding the mutated site, followed by TOPO-TA cloning and Sanger sequencing. Guide RNA and repair template sequences (4nM Ultramer, IDT) are provided in the Key Resources Table. All guide RNA sequences were designed using the CRISPR design tool (crispr.mit.edu) (Cong and Zhang, 2014).

ChIP

Chromatin immunoprecipitation (ChIP) was performed using the following antibodies: CTCF (Millipore 07-729), BRD2 (Bethyl A302-583A), SMC1 (Bethyl A300-055A), and HA (12CA5).

G1E-ER4 cells ($\geq$ 10 million per sample) were fixed in 1% formaldehyde in PBS at room temperature with agitation for 10 min, then quenched with 1M glycine for 5 min. Fixed cells were resuspended in 1mL Cell Lysis Buffer (10mM Tris pH 8, 10mM NaCl, 0.2% NP-40/Igepal) prepared fresh with protease inhibitors (Sigma P8340) and 1mM phenylmethylsulfonyl fluoride (PMSF) and incubated on ice for 10 min. Nuclei were pelleted and resuspended in 1mL Nuclear Lysis Buffer (50mM Tris pH 8, 10mM EDTA, 1% SDS, prepared fresh with protease inhibitors and PMSF), and incubated on ice for 20 min. Samples were then diluted with 0.6mL IP Dilution Buffer (20mM Tris pH 8, 2mM EDTA, 150mM NaCl, 1% Triton X-100, 0.01% SDS, prepared fresh with protease inhibitors and PMSF), and sonicated at 4C for 45 min (Epishear, Active Motif). After sonication, samples were spun at 21130xg for 5 min at 4C to remove debris and added to preclearing reactions containing 3.4mL IP Dilution Buffer, protein A/G agarose beads (agarose beads slurry was prepared by mixing Protein A (ThermoFisher 15918014) and Protein G (ThermoFisher 15920010) agarose beads at 1:1 ratio) and 50ug of isotope-matched IgG. Samples were precleared for $\geq$ 2 hr.

Prior to setting up immunoprecipitation ("IP") reactions, 200ul of precleared chromatin was removed as "Input." To set up IP reactions, precleared chromatin was added to protein A/G beads pre-bound with antibody and rotated overnight at 4C.

Beads were washed once with IP Wash 1 (20mM Tris pH 8, 2mM EDTA, 50 mM NaCl, 1% Triton X-100, 0.1% SDS), twice with High Salt Buffer (20 mM Tris pH 8, 2mM EDTA, 500 mM NaCl, 1% Triton X-100, 0.01% SDS), once with IP Wash Buffer 2 (10 mM Tris pH 8, 1mM EDTA, 0.25 M LiCl, 1% NP-40/Igepal, 1% sodium deoxycholate), and twice with TE (10mM Tris pH 8, 1mM EDTA pH 8). All washes were performed on ice. Following the final wash, beads were moved to room temperature and eluted twice with 100ul of Elution Buffer (100mM NaHCO3, 1%SDS, prepared fresh) for a final eluate volume of 200ul. The following were added to each IP and input sample: 12ul of 5M NaCl, 2ul RNaseA (10mg/ml, 10109169001 BMB) and samples were incubated at 65C for $\geq$ 1 hr. 3ul of proteinase K (20mg/ml, 3115879 BMB) was added and samples were incubated at 65C overnight.

Following overnight incubation, 10ul of 3M sodium acetate pH 5.2 was added to each sample and DNA was purified using the QIAquick PCR Purification kit (QIAGEN 28106) per the manufacturer's instructions.

ChIP-qPCR was performed with Power SYBR Green (ThermoFisher). Standard curves were constructed for each input sample and used to calculate the IP quantities for each primer set. Table S6 contains ChIP-qPCR primer sequences.

RT-qPCR

Cells were resuspended in 1mL Trizol (Life Technologies). RNA was isolated using the RNeasy Mini Kit (QIAGEN). Reverse transcription was performed using 5x iScript Reverse Transcription Supermix (Bio-Rad). Quantitative PCR (qPCR) was performed using Power SYBR Green (ThermoFisher). Transcripts were normalized to *Gapdh* or *β-actin* as indicated. Table S6 contains RT-qPCR primer sequences.

3C

3C was performed as in (Deng et al., 2012). Briefly, 10 million cells were treated with estradiol for 24 hr, fixed with 1.5% formaldehyde at room temperature for 10 min, quenched with 250mM glycine, permeabilized in ice-cold hypotonic cell lysis buffer (10 mM Tris pH 8.0, 10 mM NaCl, 0.2% NP-40, protease inhibitors, 1 mM PMSF), and lysed in a dounce with 10 strokes of a type A pestle. Nuclei were spun down at 800xg, washed in ice-cold 1.2x digestion buffer 3.1 (New England BioLabs), and then resuspended in this same buffer. SDS was added to 0.3% and nuclei were shaken in a thermomixer (Eppendorf) at 950 rpm at 37C for 1 hr. Triton X-100 was added to 1.8% and shaken at 950 rpm at 37C for 1 hr. To digest genomic DNA, 800 units of BglII (New England BioLabs) were added and samples were shaken overnight at 950 rpm at 37C. Next, SDS was added to 1.6% and samples were shaken at 950 rpm at 65C for 25 min. Samples were diluted in ligation buffer (50 mM Tris pH 7.5, 10 mM $MgCl_2$, 1 mM ATP, 10 mM DTT, 1% Triton X-100, and 100 ug/mL BSA) and incubated at 37C with occasional mixing. Genomic DNA fragments were ligated by adding 4,000 units of T4 DNA ligase (New England BioLabs) and incubating at 16C for 4 hr. EDTA was added to 10 mM to stop the reaction. Crosslinks and proteins were removed by incubating overnight at 65C with proteinase K. Nucleic acids were purified with phenol-chloroform, precipitated with ethanol, and resuspended in 10 mM Tris pH 8.0. RNA was removed by treating with RNaseA at 37C for 30 min. Finally, DNA was purified with phenol-chloroform, precipitated with ethanol, and resuspended in 10 mM Tris pH 8.0 by incubating at 55C for 10 min and 4C overnight.

Crosslinking frequencies between the anchor and baits were analyzed via Taqman qPCR. BglII digestion was assessed via qPCR (Power SYBR Green, ThermoFisher) of amplicons spanning the BglII sites and control amplicons adjacent to these sites. Primers and probes were designed using the online tool Primer3Plus (http://primer3plus.com). Specificity was confirmed using UCSC Genome Browser in silico PCR and NCBI Primer-BLAST. The presence of SNPs in G1E-ER4 cells relative to the mm9 reference genome was accounted for when selecting oligos and BglII sites. The Taqman qPCR signals were quantitated using a library of BglII digestion-ligation products derived from a BAC (BAC RP23-14E1 (Life Technologies, Cat # RPCI23.C, Clone ID: 14E1) spanning the relevant genomic region. Table S6 contains oligo sequences for 3C.

Immunoblotting

Nuclear lysates were prepared from indicated cell lines. Antibodies used were: BRD2 (Cell Signaling D89B4), and β-Actin-HRP (Sigma A3854).

Single-molecule RNA FISH imaging

We performed single-molecule RNA FISH as described previously (Femino et al., 1998; Raj et al., 2006, 2008). Briefly, we fixed cells in 1.85% formaldehyde in PBS for 10 min at room temperature, and stored them in 70% ethanol at 4C until imaging. FISH probes consisted of oligonucleotides complimentary to target mRNA, conjugated to fluorescent dyes: anti-Entpd4 oligos in Cy5, anti-Slc25a37 oligos in Cy3. We hybridized pools of FISH probes to samples, followed by DAPI staining and wash steps performed in suspension. Samples were cytospun onto slides for imaging on a Nikon Ti-E inverted fluorescence microscope. Table S6 contains sequences for RNA FISH probes.

ChIP-seq

ChIP-seq samples were prepared as described above for ChIP. All ChIP-seq samples were processed for library construction for Illumina sequencing using Illumina's ChIP-seq Sample Preparation Kit. In brief, ChIP-enriched, fragmented DNA was subjected to end repair to generate blunt-end double stranded DNA, adenylation of 3′ ends, and adaptor ligation. Following ligation, SPRIselect (Beckman Coulter) beads were used at 0.9X and 0.6X for left and right side selection, respectively, to obtain an average library target size of ∼300 bp. After size selection, fragments were amplified for 16 cycles, and PCR products were purified using Agencourt AMPureXP beads (Beckman Coulter). Sequencing was performed on an Illumina NextSeq 500 using Illumina's reagents and kits as appropriate. Reads were mapped to mouse genome assembly mm9 using Bowtie. Reads were extended in the 3′ direction to 200bp using MACS to generate bigwig format files for browser display.

In situ HiC

10 million uninduced G1E-ER4 cells (-GATA1), either control or BRD2 KO clonal sublines, were fixed in 2% formaldehyde at room temperature for 10 min then quenched with 0.25M glycine for 5 min. Cells were washed with PBS, pelleted, resuspended in 1mL cold Cell Lysis Buffer (10mM Tris pH 8, 10mM NaCl, 0.2% NP-40/Igepal) and incubated on ice for 10 min. Nuclei were pelleted and washed once with 800ul of cold 1.2X NEB Buffer DpnII, pelleted and resuspended in 500ul 1.2X NEB Buffer DpnII. SDS was added to a final concentration of 0.3% and samples were incubated with shaking at 950rpm at 37C. 40ul of water was added to each sample, followed by 300U of DpnII (NEB R0543M), and samples were incubated overnight at 37C with shaking. An additional

300U of DpnII was added, and samples were incubated for 2 more hours at 37C, followed by incubation at 65C for 20 min. Nuclei were cooled on ice, pelleted and resuspended in 1X NEB Buffer2 with biotin-14-dATP, dCTP, dGTP, dTTP and DNA Polymerase I, Large (Klenow) Fragment (NEB M0210), and incubated at 37C with shaking for 1.5 hr. Fragments were ligated in a total volume of 1.2mL with T4 DNA Ligase (NEB) for 4 hr at 16C, then at room temperature for 30 min. Proteinase K (20ul of 20mg/mL) and SDS (120ul of 10% SDS) were added and crosslinks were reversed overnight at 65C. An additional 10ul of proteinase K was added and samples were incubated at 55C for 2 hr. DNase-free RNaseA was added and samples were incubated at 37C for 30 min. DNA was purified by phenol chloroform extraction.

To prepare the sequencing library, DNA was sonicated to 200-300bp fragments and purified using AMPure XP beads (Beckman Coulter). Ligation junctions were pulled down using streptavidin beads (50ul of Dynabeads MyOne Streptavidin C1 beads/10 million cells). Libraries were prepared using the NEBNext DNA Library Prep Master Mix Set for Illumina (NEB E6040, M0543L, E7335S) as per manufacturer's instructions. Briefly, fragmented DNA was subjected to blunt end repair and dA tailing, followed by adaptor ligation. Following ligation, DNA was purified using AMPure XP beads (Beckman Coulter) and eluted in 15ul of water. Libraries were amplified for 6 cycles and PCR products were purified using AMPure XP beads. Sequencing was performed on an Illumina NextSeq 500.

QUANTIFICATION AND STATISTICAL ANALYSIS

ChIP-seq analysis: binding correlation and motifs

Mouse CTCF (in G1E or G1E-ER4 cells), RAD21 (MEL cells), and SMC3 (MEL cells) peak and aligned read (bam) files were downloaded from the Mouse ENCODE Project. BRD2, BRD3, and BRD4 ChIP-seq data were previously published by our laboratory (Stonestrom et al., 2015), as were RNA Pol II ChIP-seq data (Hsiung et al., 2016). Bigwig format files were visualized with the UCSC genome browser for track display (Meyer et al., 2013). Heatmap visualization of multiple tracks at CTCF sites was performed using the Cistrome (Liu et al., 2011) Galaxy (Giardine et al., 2005) analysis tool "Heatmap."

Correlation of binding signal between factors at CTCF sites was determined by computing the RPKM for the indicated factors over CTCF peaks (+GATA1) using BEDTools (Quinlan and Hall, 2010) and calculating a Pearson correlation coefficient. Binned 2D scatterplots were generated using the function geom_bin2d in the R package ggplot2 (Wickham, 2009).

Motif analysis was performed using MEME-ChIP (Machanick and Bailey, 2011) for combined de novo motif analysis, comparison to a known motif database, and analysis of central enrichment in 500bp regions (centered at the middle of each peak using the UCSC genome browser mouse reference genome mm9). Motif analysis was performed on peak sets called using MACS (Feng et al., 2012; Zhang et al., 2008) from induced G1E-ER4 cells (+GATA1).

RNA FISH image analysis

We manually segmented boundaries of cells from bright field images and localized RNA spots using custom software written in MATLAB (Raj et al., 2010) with subsequent analyses performed in R. All correlations were determined by the Spearman correlation test.

To examine whether unedited control G1E-ER4 cells could have a high biological correlation between the *Entpd4* and *Slc25a37* genes, which we might have failed to detect due to the relatively high impact of noise on the low numbers of *Entpd4* RNA molecules per cell, we first computationally normalized the *Entpd4* RNA counts from the control and CTCF mutant clones to the same low level. Since we observed that the change in mRNA level between control and Slc25a37-CTCFa-mutated cells was driven mainly by a change in burst frequency (unpublished data), the increase in noise due to decreased burst frequency can be approximately modeled as a Poisson distribution (Raj et al., 2006). We then added increased noise to the computationally normalized data by drawing values from a Poisson distribution centered on the observed values. We then computed the correlation of *Slc25a37* with these new, noisier values, and repeated this 10,000 times such that we could generate confidence intervals.

HiC read alignment and processing

Paired-end reads were aligned independently to mm9 mouse genome using bowtie2 (Langmead and Salzberg, 2012) (global parameters:–very-sensitive –L 30 –score-min L,-0.6,-0.2 –end-to-end–reorder; local parameters:–very-sensitive –L 20 –score-min L,-0.6,-0.2 –end-to-end–reorder) through the HiC-Pro software (Servant et al., 2015). Unmapped reads, non-uniquely mapped reads and PCR duplicates were filtered and uniquely aligned reads were paired (Table S1). We initially performed two runs of sequencing on these libraries and found that the raw reads from biological replicates for each condition were highly correlated (Control_rep1 versus Control_rep2 Spearman correlation coefficient: 0.83, BRD2KO#1_rep1 versus BRD2KO#1_rep2 Spearman correlation coefficient: 0.77, BRD2KO#2_rep1 versus BRD2KO#2_rep2 Spearman correlation coefficient: 0.77), demonstrating the reproducibility of the experiment and allowing us to merge the two replicates of each sample. To increase our sequencing depth we ultimately performed five total sequencing runs on these libraries. All subsequent analysis presented in the paper was performed on the full datasets pooling five sequencing runs with merged replicates for each sample. HiC maps were generated at 40kb and 10 kb matrix resolution and balanced using the iterative correction and eigenvector decomposition (ICED) technique (Imakaev et al., 2012). The lowest-interacting 10% of bins were removed prior to balancing to facilitate convergence. Matrices for all three conditions converged to a tolerance of 0.1 within 200 iterations.

Identifying BRD2, CTCF sites to classify boundaries

ChIP-seq peak calling was performed with MACS2 (Zhang et al., 2008) using parameters for punctate (CTCF; -p 1E-8) and diffuse (BRD2; -p 1E-8–broad–broad-cutoff 1e-4) chromatin marks. To address the challenge of peak calling BRD2 signal, we performed intersections to keep the enriched signal called with–broad-cutoff 1e-2 that also intersected peaks called with–broad-cutoff 1e-4. The result of this intersection is the detection of both punctate and diffuse BRD2 sites via the capture of the breadth of moderate-confidence BRD2 signal surrounding narrower high-confidence BRD2 peaks. To find genomic loci co-occupied by high-confidence CTCF and BRD2, we intersected CTCF and BRD2 peaks and vice-versa, and then the results of the intersections were concatenated and merged. To identify exclusively occupied sites, we subtracted low-confidence occupancy regions for CTCF from the high-confidence peak calls for BRD2 and vice-versa. We parsed low-confidence occupancy regions in the genome as zones of possible but low-signal binding using -p 1e-2 (CTCF) or–broad-cutoff 1e-2 (BRD2). The subtraction operations ensured that regions of moderate-confidence binding of either protein are excluded from the stringent exclusive site list.

Contact domain calling and categorization

Contact domains were identified in control uninduced G1E-ER4 cells by applying the directionality index (DI) and a Hidden Markov Model (Dixon et al., 2012). We called two sets of domains: one set on the 40kb interaction matrices, using a 50 bin (2Mb) window in both the upstream and downstream direction to compute the directionality index; and a second set on the 10kb interaction matrices, using a 100 bin (1Mb) directionality index window (boundary locations are provided in Tables S4 and S5). We defined domain boundaries as the endpoints of each domain called, except in cases where the next domain was over 400kb away, suggesting a stretch of generally disorganized chromatin rather than a specific boundary. Disorganized chromatin regions were not considered in downstream analyses. We detected 1,814 domains (median size 920kb) in our 40 kb contact matrices from control uninduced G1E-ER4 cells. Moreover, by using a smaller directionality index computation window (100 bins) on our higher resolution 10 kb contact matrices, we identified a larger set of 3,214 domains of smaller size (median size 470kb) in control cells, These results are consistent with the notion that large domains such as TADs are further arranged into a nested set of overlapping sub-domains.

Control cell contact domain boundaries were then classified based on the presence or absence of nearby high-confidence CTCF and BRD2 ChIP-seq peaks. Boundaries were parsed into categories based on which combinations of these factors were found near the boundary, including: (1) those with BRD2 without CTCF, (2) those in which CTCF and BRD2 are directly colocalized, with no additional CTCF-only or BRD2-only sites ("CTCF/BRD2 co-occupied only"), (3) those in which CTCF and BRD2 are directly colocalized with additional BRD2-only sites ("CTCF/BRD2 co-occupied with BRD2-only"), (4) those in which CTCF and BRD2 are directly colocalized with additional CTCF-only sites ("CTCF/BRD2 co-occupied with CTCF-only"), (5) those with CTCF without BRD2, (6) those that lack either protein ("Neither"), and (7) all "Other" boundaries containing multiple classes of colocalized sites and/or peaks that did not meet our stringent thresholds. This "Other" category ensures the stringency of categories (1)-(6) by including cases where high-confidence BRD2 was present together with moderate-confidence CTCF (or vice-versa), as well as cases where exclusive CTCF and exclusive BRD2 sites were both found near the boundary call. When categorizing the boundaries of domains called on 40kb contact matrices, we considered ChIP-seq peaks within 20kb of the boundary coordinate, in accordance with the ± 20kb uncertainty inherent in identifying a coordinate on a 40kb binned matrix. Similarly, when categorizing the boundaries of domains called on the 10kb contacted matrices, we considered ChIP-seq peaks within ± 20kb of the boundary coordinate. The number of boundaries classified into each category is shown in Tables S2 and S3.

Insulation score analysis

To enable comparison between samples, ICED corrected HiC libraries were parsed into contact matrices for each chromosome and then quantile normalized (Bolstad et al., 2003; Bullard et al., 2010). Bin-bin pairs with ICED values lower than 0.01 were excluded from quantile normalization and downstream analyses. Our normalization approach resulted in the equalization of global counts distributions across different samples while preserving underlying biological differences unique to each condition.

To assess the relative strength of contact domain boundaries using our 40kb contact matrices, we computed an insulation score as the sum of a 5 bin by 5 bin (200kb by 200kb) window positioned adjacent to the diagonal at the bin of interest as in (Crane et al., 2015). We positioned two windows over the bins immediately upstream and downstream of the boundary ("-1" and "+1," respectively), and then slid the windows away from the boundary until they no longer overlapped a previous window. This allowed the boundary position to emerge as a relative minimum in insulation score relative to the surrounding chromatin context. To understand global trends, the insulation scores for each window position across boundaries were averaged within each boundary category and condition. Error bars were computed by taking the 2.5th and 97.5th percentile of the distribution of the sampling means generated via a bootstrapping procedure in which the insulation score values of boundaries in a particular class were sampled with replacement 1,000 times. A high insulation score over a bin implies that many contacts between the 200kb of chromatin directly downstream of the bin and the 200kb directly upstream of the bin pass over the bin, or equivalently, that the bin does not prevent contacts at this range from passing over the bin. Low values of insulation score correspond to well-insulated/strong boundaries. We repeated this insulation score analysis on our 10kb contact matrices, using a variety of window sizes. The results presented in Figures S6E and S6F were obtained by precisely matching the insulation score window size (200kb by 200kb) and non-overlapping sliding window genomic positions to those used in the 40kb analysis.

To better understand the contributions of individual domain boundaries to the genome-wide trends shown in the insulation score line plots, the log2 fold change in insulation score between the BRD2 knockout (averaging over the two biological replicates) and control conditions was computed for each boundary, and these differences were collected for each boundary category. The collected quantities were then plotted as histograms (top row of Figure 5C, analysis of 40kb matrices, and Figure S6E, analysis of 10kb matrices). For comparison, we generated an empirical null distribution representing log2 fold changes in insulation score between the two BRD2 depleted conditions (BRD2 KO#1 and BRD2 kO#2), allowing the evaluation of the former histograms with respect to the level of noise in the assay (bottom row of Figure 5C, analysis of 40kb matrices, and Figure S6E, analysis of 10kb matrices). Within each category, we used this empirical null distribution to assign one-sided p values for the log2 fold change in insulation score of each individual boundary in that category between BRD2 depleted and control conditions. To correct for multiple hypothesis testing, we applied the direct FDR procedure (Storey, 2002) and called boundaries as significantly strengthening or significantly weakening upon BRD2 depletion at 15% pFDR. The numbers of significantly changing boundaries within each category are provided in Table S2 (based on the 40kb contact matrices) and Table S3 (based on the 10kb contact matrices).

DATA AND SOFTWARE AVAILABILITY

The accession number for the ChIP-seq sequencing data reported in this paper is GEO: GSE80527. The accession number for the HiC sequencing data reported in this paper is GEO: GSE95476.
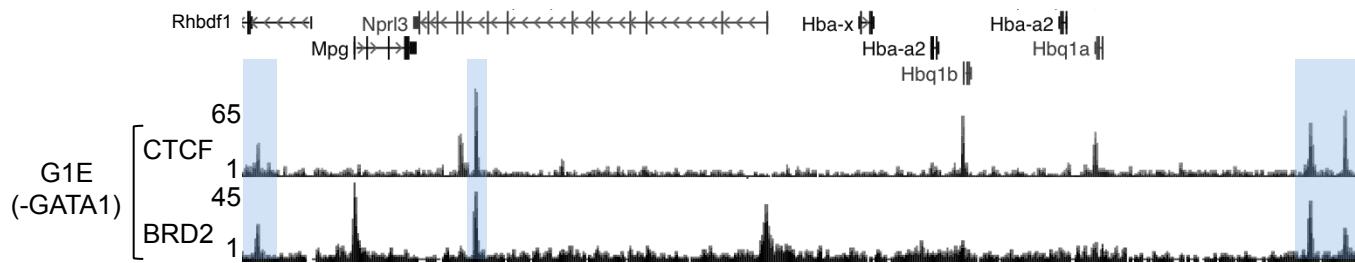
Supplemental Information

The BET Protein BRD2 Cooperates

# with CTCF to Enforce Transcriptional

and Architectural Boundaries

Sarah C. Hsu, Thomas G. Gilgenast, Caroline R. Bartman, Christopher R. Edwards, Aaron J. Stonestrom, Peng Huang, Daniel J. Emerson, Perry Evans, Michael T. Werner, Cheryl A. Keller, Belinda Giardine, Ross C. Hardison, Arjun Raj, Jennifer E. Phillips-Cremins, and Gerd A. Blobel

Figure S1

A

B
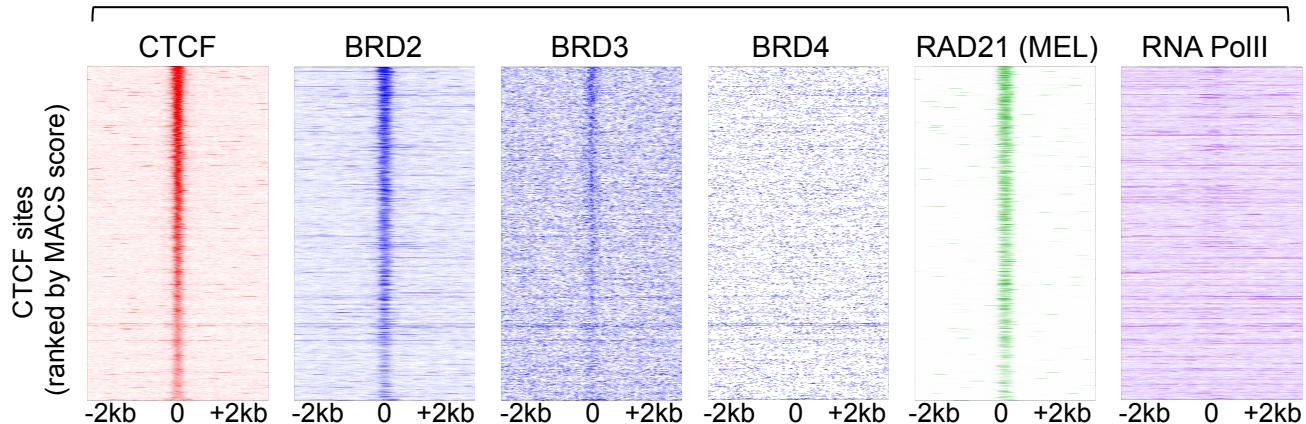
C
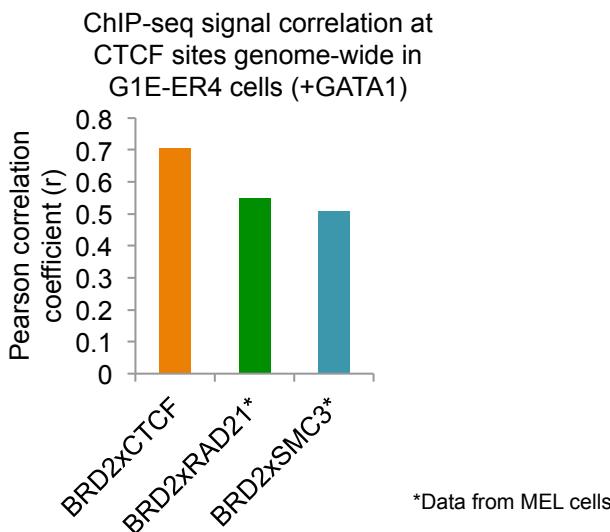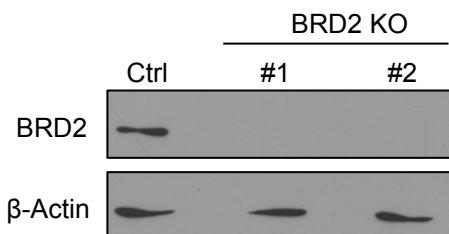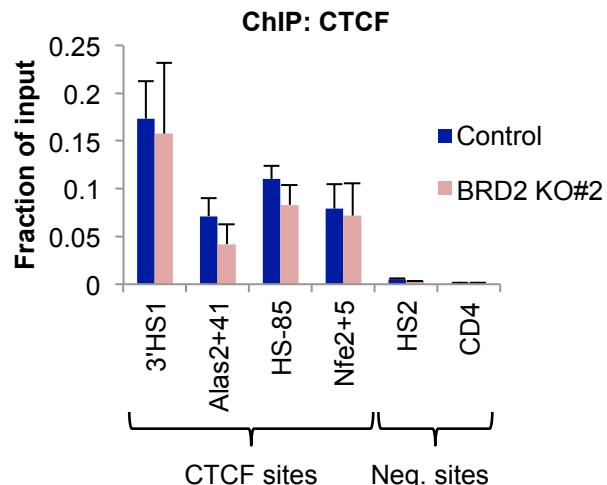
Figure S1 (related to Figure 1). (A) Genome browser tracks showing ChIP-seq for CTCF and BRD2 in G1E cells (undifferentiated, -GATA1) at the α-globin (*Hba*) locus. Gray boxes highlight regions of BRD2/CTCF colocalization. (B) ChIP-seq signal for the indicated proteins in a 4kb window centered on CTCF binding sites in G1E cells (undifferentiated, -GATA1) for CTCF, BRD2, BRD3, BRD4 and RNA PolII (-GATA1). RAD21 data are from MEL cells. Each row represents a single peak, ranked from highest to lowest CTCF signal (MACS score). (C) Correlation coefficient (Pearson) between the ChIP-seq signal (RPKM) of the indicated factors at CTCF binding sites defined in G1E-ER4 cells (+GATA1). RAD21 and SMC3 datasets are from MEL cells (ENCODE). The CTCF vs. BRD2 correlation is replotted here from Figure 1D as a reference.

Figure S2
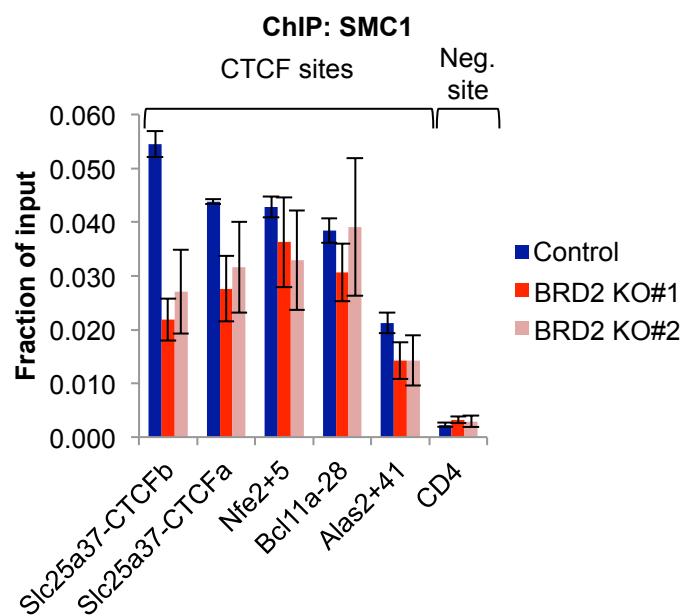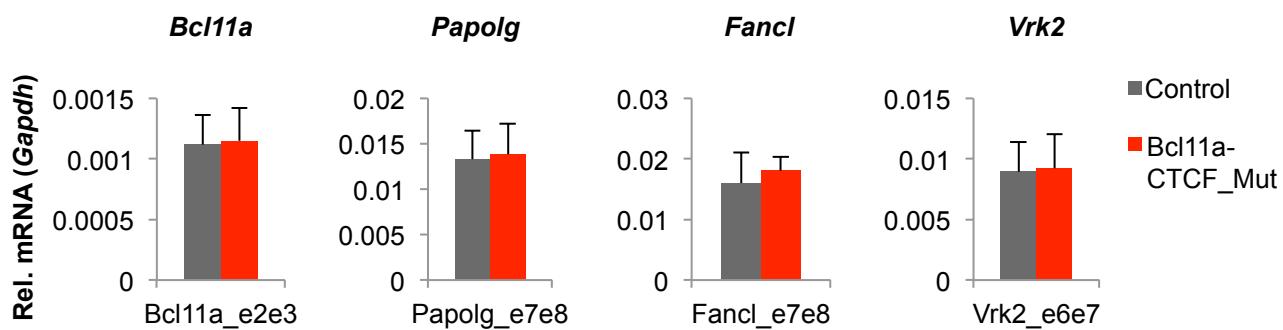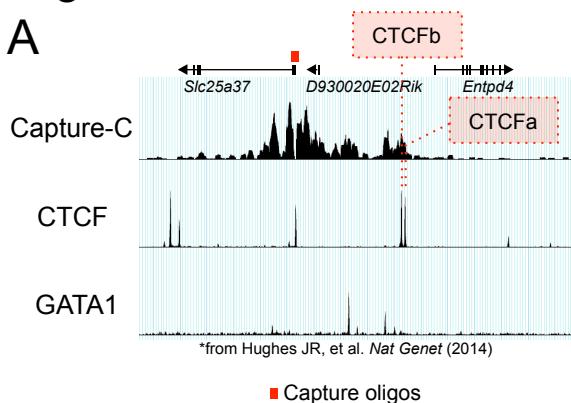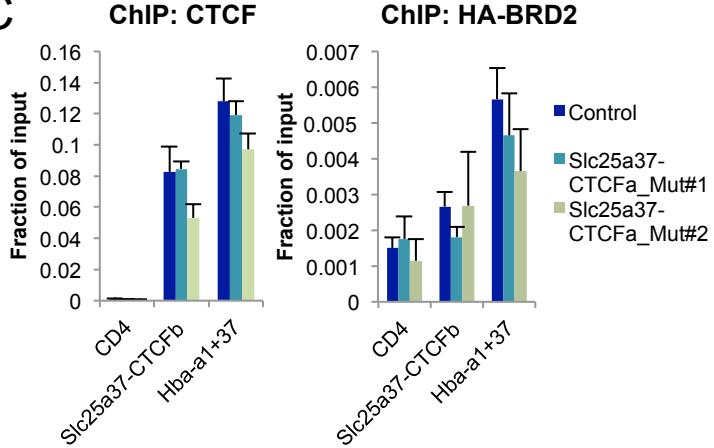
Figure S2 (related to Figure 2) . (A) Western blot of BRD2 in either unedited control G1E-ER4 cells (Ctrl) or two BRD2 knockout clones generated with CRISPR/Cas9 (BRD2 KO #1 and #2). (B) CTCF ChIP-qPCR in control or BRD2 KO G1E-ER4 cells (+GATA1) (n≥3, error bars represent SD). (C) SMC1 ChIP-qPCR in control or BRD2 KO G1E-ER4 clonal sublines (+GATA1) (n=3, error bars represent SEM) (D) RT-qPCR of indicated genes near the edited *Bcl11a* CTCF site in control or Bcl11a-CTCF_Mut cells (+GATA1). Transcripts were normalized to *Gapdh* (n=3, error bars represent SD).

Figure S3
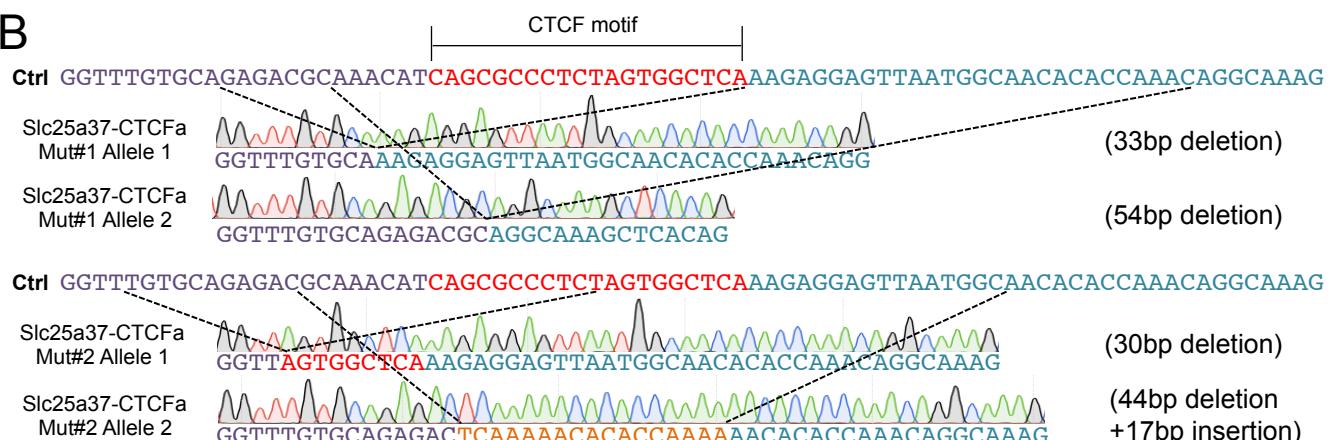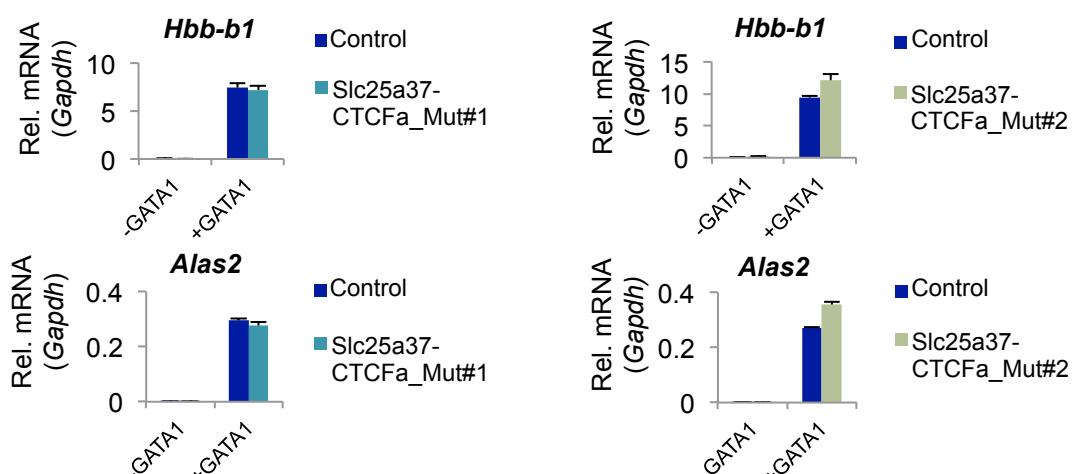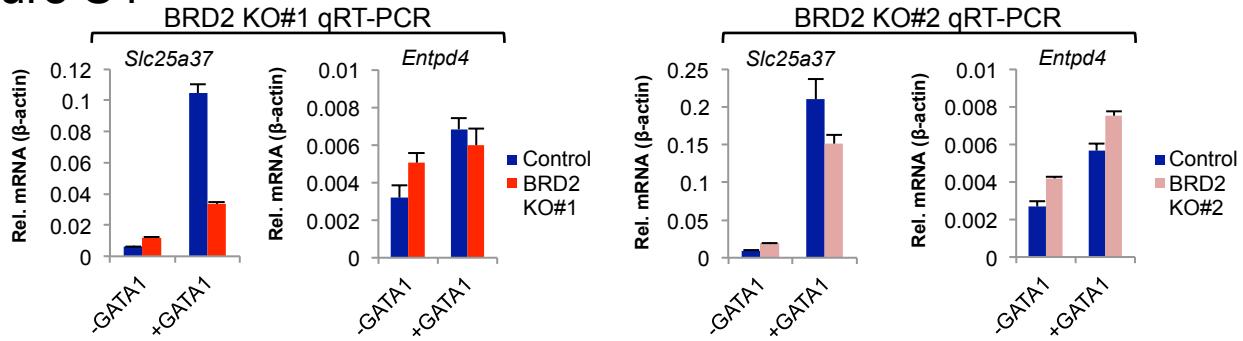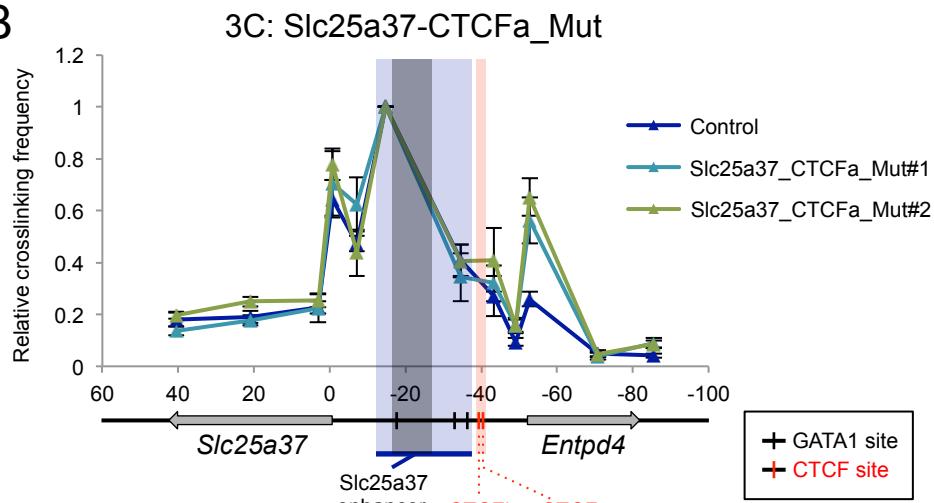
**Figure S3 (related to Figure 3) .** (A) Capture-C of the *Slc25a37* promoter and associated CTCF and GATA1 ChIP-seq tracks in mouse erythroid Ter119+ cells (from Hughes, et al. *Nat Genet* (2014), available from the Capture-C mouse erythroid database (https://gbrowse2.molbiol.ox.ac.uk/gbrowse2/MM9_CapC_erythroid/capc_full.html). CTCF sites Slc25a37-CTCFa and -CTCFb are labeled as indicated. (B) CRISPR/Cas9-mediated editing of the CTCF site, Slc25a37-CTCFa with unedited control (Ctrl) sequences and resulting insertions/deletions indicated for each allele. (C) ChIP-qPCR of either CTCF or HA-BRD2 at Slc25a37-CTCFb in control or Slc25a37-CTCFa mutated cells. CD4 (negative control region) and Hba-a1+37 (positive control region) are replotted from Figure 3B and shown as references (all were measured in the same experiment). (D) RT-qPCR of genes activated by GATA1 in Slc25a37-CTCFa mutated cell lines. Transcripts were normalized to *Gapdh* (n=3, error bars represent SEM).

Figure S4
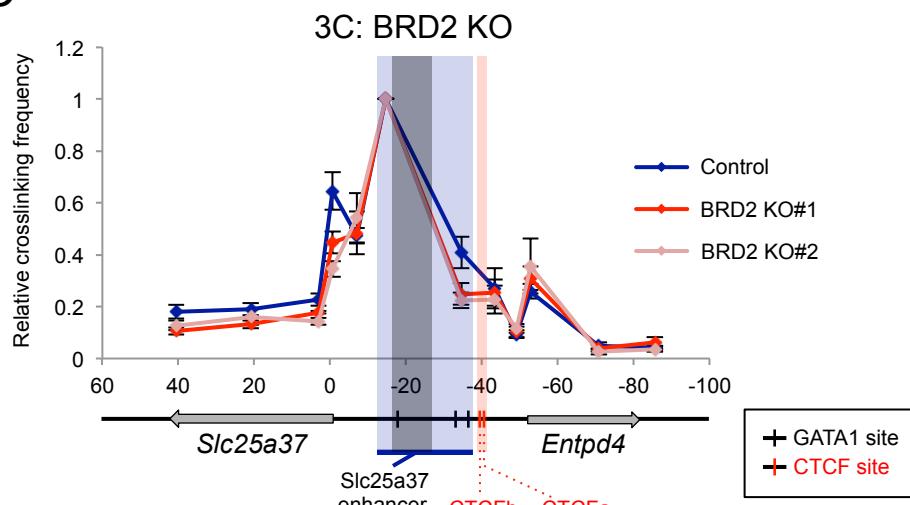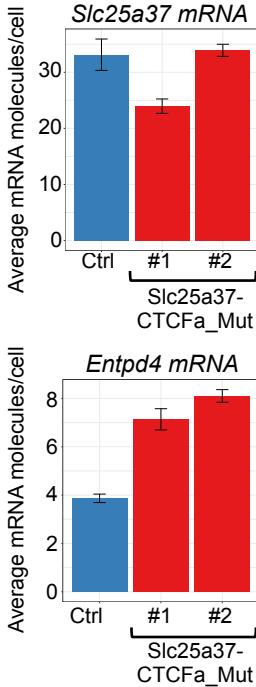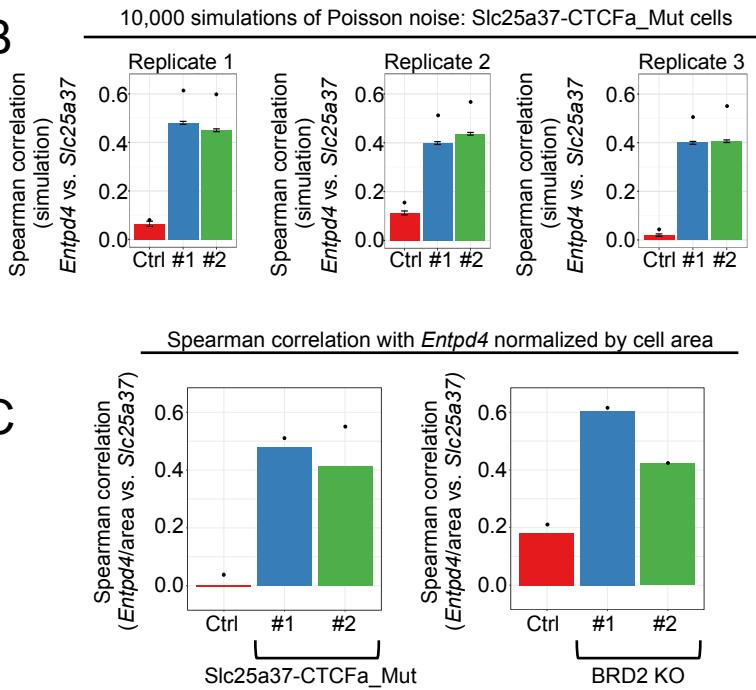
Figure S4 (related to Figure 3). (A) RT-qPCR for *Slc25a37* and *Entpd4* in unedited control G1E-ER4 and BRD2 KO cell lines in –GATA1 and +GATA1 conditions. Transcripts were normalized to *β-actin* (n=3, error bars represent SEM). (B) 3C of the *Slc25a37/Entpd4* locus measuring cross-linking frequencies between BglII-digested fragments in either unedited control or Slc25a37_CTCFa_Mut#1 or #2 cells (+GATA1, n>3 for each data point). Gray bar denotes the anchor region, near the first GATA1 site in the *Slc25a37* enhancer. Cross-linking frequencies are normalized to the bait nearest the anchor. The structure of the locus is shown below the graph, with GATA1 sites in black and CTCFa/b sites in red. The x-axis indicates the distance in kilobases from the *Slc25a37* TSS which represents zero. (C) 3C as in (B) for control and BRD2 KO#1 and #2 cells. 3C experiments for (B) and (C) were performed at the same time but are plotted separately for purposes of clarity; the same control data is plotted in both figures.

Figure S5
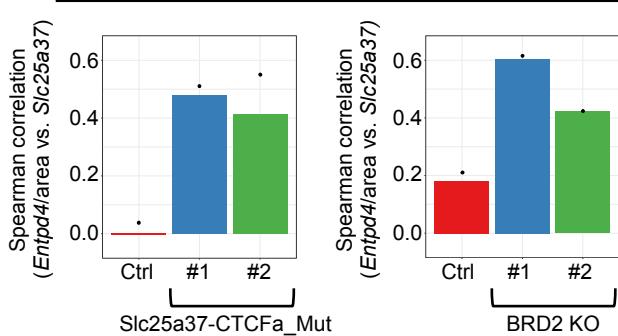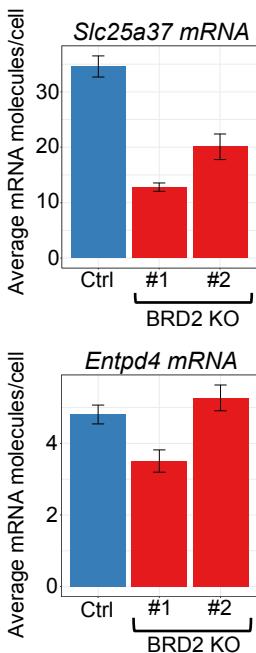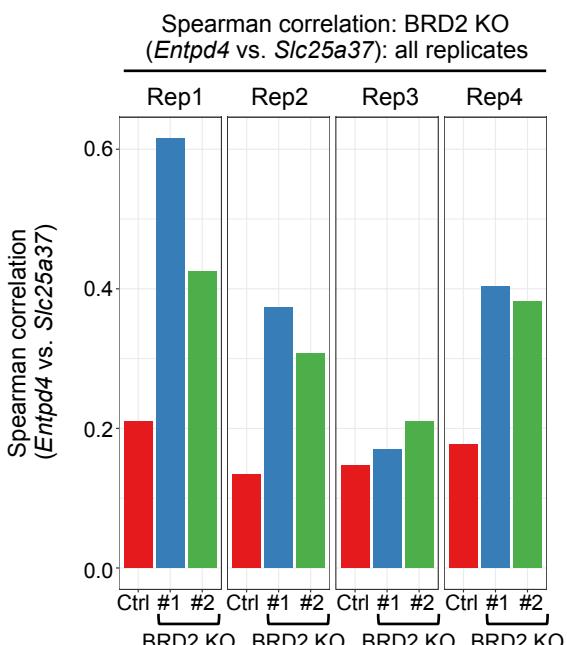
Figure S5 (related to Figure 4). (A) Average mRNA counts per cell for *Slc25a37* and *Enptd4* from single-molecule RNA FISH in unedited control G1E-ER4 (Ctrl) and Slc25a37-CTCFa-Mut. #1 and #2 cell lines (+GATA1) (n=3, error bars represent SEM). (B) Spearman correlations between *Slc25a37* and *Entpd4* calculated by (1) computationally rendering the mean *Enptd4* expression in Ctrl and Slc25a37-CTCFa-Mut cells equivalent, and (2) adding random sampling error from a Poisson distribution 10,000 times to the number of *Entpd4* molecules observed. All three biological replicates are shown (error bars represent SEM, the dot is the observed correlation for each sample). (C) Spearman correlations between *Entpd4* and *Slc25a37* with *Entpd4* levels normalized for cell area (representative replicate is shown for both Slc25a37-CTCFa-Mut and BRD2 KO cells, the dot is the observed correlation for each sample). (D) Average mRNA counts per cell as in (A) for Ctrl and BRD2 KO cell lines #1 and #2 (+GATA1) (n=4, error bars represent SEM) (E) Observed Spearman correlation between *Entpd4* and *Slc25a37* in Ctrl and BRD2 KO cells for each replicate, plotted individually.

Figure S6

A Types of HiC interactions

- Cis long-range contacts (>20kb)
- Cis short-range contacts (<20kb)
- Trans Contacts

B

G1E-ER4 (Control) 40kb — *Data from this study*
G1E-ER4 (Control) 10kb — *Data from this study*
mESC 40kb — *Data from Dixon et al. Nature (2012)*
CH12-LX 10kb — *Data from Rao et al. Cell (2014)*

C Other perturbations observed upon BRD2 depletion

Control   BRD2 KO#1   BRD2 KO#2

— Contact domains at 40kb resolution (control G1E-ER4)

D Heat maps binned at 10kb resolution

Control   BRD2 KO#1   BRD2 KO#2

— Contact domains at 10kb resolution (control G1E-ER4)

E

- BRD2
- CTCF
- [ ] Boundary +/- 20kb

BRD2 without CTCF | CTCF/BRD2 co-occupied only | CTCF/BRD2 co-occupied w/ BRD2-only | CTCF/BRD2 co-occupied w/ CTCF-only | CTCF without BRD2 | Neither

Fold-change IS distribution: $\log_2(\text{averageBRD2KO/Ctrl})$

Fold-change IS distribution: $\log_2(\text{BRD2KO\#1/BRD2KO\#2})$

F Insulation score across 10kb boundaries

- Ctrl
- BRD2KO#1
- BRD2KO#2

Distance from the boundary (kb)

Figure S6, continued

G

Gene expression at *Sec16b* locus in Fig. 5A

Gene expression at *Xrcc4* locus in Fig. S6C

H

Validation of boundary classification strategy: ChIP-seq peak enrichment at boundary classes (boundaries called at 40kb)

Distance from the boundary (kb)
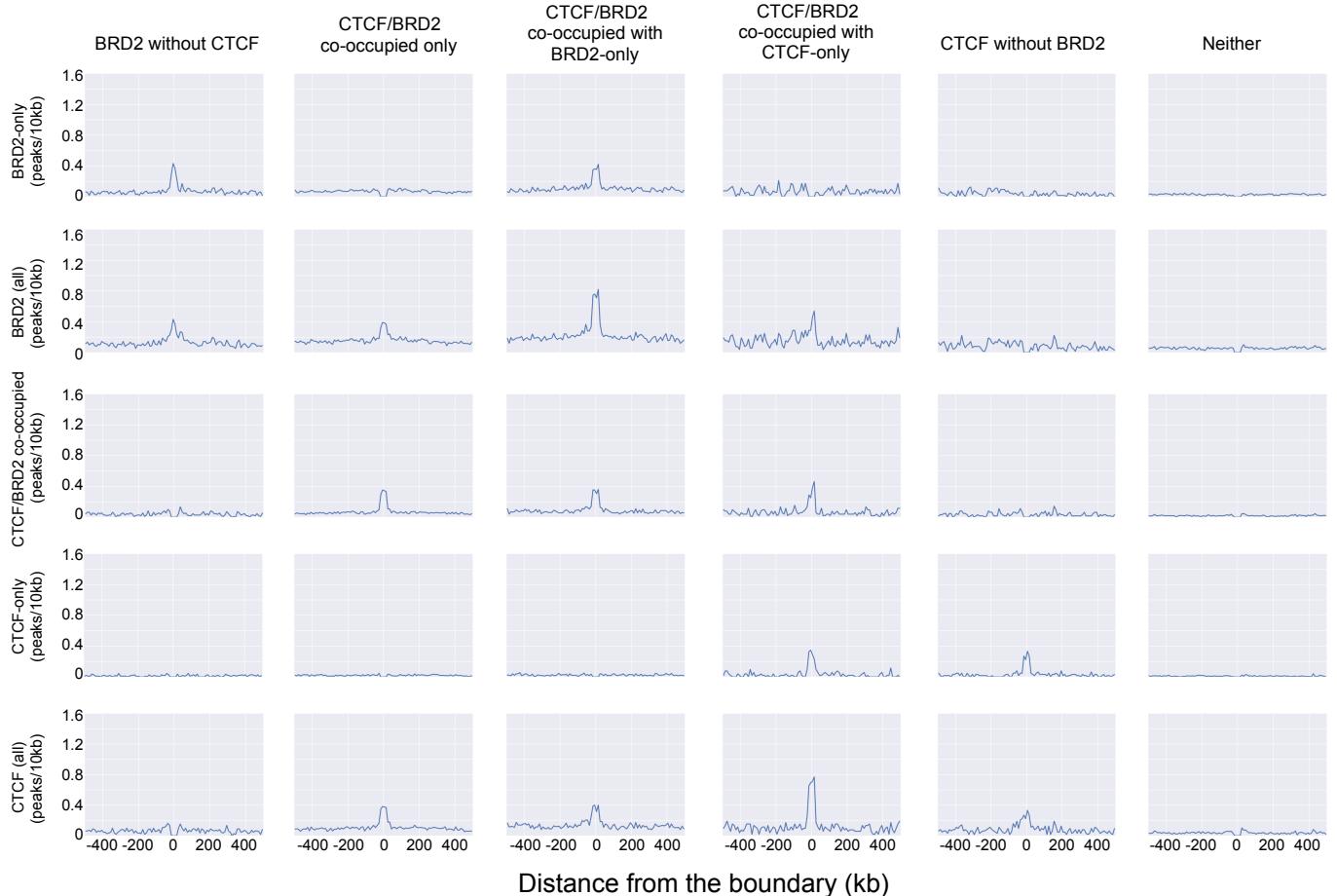
**Figure S6 (related to Figure 5) .** (A) Breakdown of HiC interactions obtained for control G1E-ER4, BRD2 KO#1 and BRD2 KO#2 samples (all are uninduced, -GATA1). (B) Comparison of published HiC data from mESC (Dixon, et al *Nature* (2012)) and CH12-LX (Rao, et al *Cell* (2014)) with our G1E-ER4 (control) dataset at the *Kit l*ocus. Both 40kb and 10kb resolutions are shown. Lines drawn on the heat maps indicate the boundaries called for each dataset and at each resolution. (C) Contact frequency heat maps of the *Xrcc4* locus at 40kb resolution showing other types of domain perturbations (D) Contact frequency heat maps binned at 10kb resolution of the *Ier5* locus from Figure 5B, with areas of increased interaction frequency indicated by black arrows. Color bars range from low (white) to high (red) interaction frequency. (E) Histograms showing the distribution of the change in insulation score at boundaries called at 10kb resolution either between control cells and an average of the two BRD2 KO cell lines (top row, log2(averageBRD2 KO/control)), or between BRD2 KO #1 and BRD2 KO #2 (bottom row, log2(BRD2KO #1/BRD2KO #2)). The red line indicates no change. (F) Average insulation score centered on the boundaries called at 10kb resolution in control G1E-ER4 cells or either of the two BRD2 KO cell lines (#1 and #2) at the indicated boundary category (error bars represent the 2.5th and 97.5th percentile of the distribution of the sampling means generated via a bootstrapping procedure in which the insulation score values of boundaries in a particular class were sampled with replacement 1,000 times). (G) RT-qPCR of the indicated transcripts in control or BRD2 KO cells (-GATA1) for either the *Sec16b* locus from Figure 5A (left), or the *Xrcc4* locus from Figure S6C (right). Transcripts were normalized to β-actin (n=3, error bars represent SEM). (H) CTCF and BRD2 ChIP-seq enrichment at the indicated boundary classes called at 40kb resolution (signal plotted as peaks per 10kb centered over the boundary).

Table S1

	Valid interaction pairs	Self-circle pairs	Dumped pairs	Dangling end pairs
Control_merged	323102926	3628768	640174	38994512
BRD2KO1_merged	301803454	4616571	720014	54858433
BRD2KO2_merged	310088615	5459898	751417	69005157

Table S1 (related to Figure 5). Valid HiC interaction pairs for merged replicates of control G1E-ER4 and BRD2 KO #1 and #2 cells

Table S2

	Total	Significantly stronger boundary in BRD2 KO vs. control*	Significantly weaker boundary in BRD2 KO vs. control*
CTCF Brd2 co-occupied only	360	0	75
CTCF Brd2 co-occupied with Brd2 only	187	0	113
CTCF Brd2 co-occupied with CTCF only	51	0	13
CTCF without Brd2	59	0	0
Brd2 without CTCF	101	0	33
Neither CTCF nor Brd2	267	0	0

*15% False Discovery Rate threshold as described in STAR Methods

Table S2 (related to Figure 5). Number of boundaries in each domain boundary class in control 40 kb contact matrices and their changes in BRD2 KO cells

Table S3

	Total	Significantly stronger boundary in BRD2 KO vs. control*	Significantly weaker boundary in BRD2 KO vs. control*
CTCF Brd2 co-occupied only	617	0	170
CTCF Brd2 co-occupied with Brd2 only	463	0	299
CTCF Brd2 co-occupied with CTCF only	120	0	42
CTCF without Brd2	51	0	0
Brd2 without CTCF	114	0	40
Neither CTCF nor Brd2	113	0	0

*15% False Discovery Rate threshold as described in STAR Methods

Table S3 (related to Figure 5). Number of boundaries in each domain boundary class in control 10 kb contact matrices and their changes in BRD2 KO cells

Table S4 (related to Figure 5). Domain boundaries called at 40kb resolution in control G1E-ER4 cells

Table S5 (related to Figure 5). Domain boundaries called at 10kb resolution in control G1E-ER4 cells

Table S6 (related to STAR Methods). Oligo sequences for ChIP-qPCR, RT-qPCR, 3C, and single-molecule mRNA FISH.