

Chromatin Immunoprecipitation Indirect Peaks Highlight Long-Range Interactions of Insulator Proteins and Pol II Pausing

Jun Liang,^{1,7} Laurent Lacroix,^{1,7} Adrien Gamot,¹ Suresh Cuddapah,² Sophie Queille,¹ Priscillia Lhoumaud,¹ Pierre Lepetit,¹ Pascal G.P. Martin,^{1,5} Jutta Vogelmann,³ Franck Court,¹ Magali Hennion,¹ Gaël Micas,¹ Serge Urbach,⁴ Olivier Bouchez,⁵ Marcelo Nöhlmann,³ Keji Zhao,² Eldon Emberly,⁶ and Olivier Cuvier^{1,*}

¹Laboratoire de Biologie Moléculaire Eucaryote (LBME), CNRS, Université de Toulouse (UPS), 31000 Toulouse, France

²Systems Biology Center, National Heart, Lung and Blood Institute, National Institutes of Health (NIH), Bethesda, MD 20824, USA

³Centre de Biochimie Structurale, CNRS, 34090 Montpellier, France

⁴Platform Proteomic Fonctionnelle, IGF, CNRS, 34060 Montpellier, France

⁵INRA, UMR444- GeT-PlaGe, Genotoul, 31326 Toulouse, France

⁶Physics Department, Simon Fraser University (SFU), Burnaby, BC V5A 1S6, Canada

⁷These authors contributed equally to this work

*Correspondence: cuvier@ibcg.biotoul.fr

<http://dx.doi.org/10.1016/j.molcel.2013.12.029>

SUMMARY

Eukaryotic chromosomes are partitioned into topologically associating domains (TADs) that are demarcated by distinct insulator-binding proteins (IBPs) in *Drosophila*. Whether IBPs regulate specific long-range contacts and how this may impact gene expression remains unclear. Here we identify “indirect peaks” of multiple IBPs that represent their distant sites of interactions through long-range contacts. Indirect peaks depend on protein-protein interactions among multiple IBPs and their common cofactors, including CP190, as confirmed by high-resolution analyses of long-range contacts. Mutant IBPs unable to interact with CP190 impair long-range contacts as well as the expression of hundreds of distant genes that are specifically flanked by indirect peaks. Regulation of distant genes strongly correlates with RNAPII pausing, highlighting how this key transcriptional stage may trap insulator-based long-range interactions. Our data illustrate how indirect peaks may decipher gene regulatory networks through specific long-range interactions.

INTRODUCTION

Eukaryotic chromosomes are physically partitioned into topologically associating domains (TADs) in the space of the nucleus, as shown in human, mouse, and *Drosophila* (Lieberman-Aiden et al., 2009; Dixon et al., 2012; Sexton et al., 2012; Nora et al., 2012; Hou et al., 2012; Phillips-Cremins et al., 2013). These findings were based on genome-wide mapping of long-range interactions by chromosome conformation capture (3C/Hi-C; Dekker, 2006). In the case of *Drosophila*, this provided a global map of contacts for every 10 kbp genomic region (Sexton

et al., 2012; Hou et al., 2012). Such organization may play critical roles in the regulation of multiple processes such as DNA replication, repair, or transcription (Misteli, 2007; Sanyal et al., 2011; Dostie and Bickmore, 2012; Tanay and Cavalli, 2013; Phillips-Cremins and Corces, 2013). How the global landscape of long-range interactions translates into specific regulation is still, however, not well understood.

Of interest, *Drosophila* insulator-binding proteins (IBPs) including CCCTC-binding insulator (dCTCF), Boundary Element-Associated Factor (Beaf32), and cofactors such as centrosomal protein 190 (CP190) were enriched in the borders of TADs (Sexton et al., 2012; Hou et al., 2012). Insulators have the ability upon binding to restrict long-range contacts between enhancers and promoters when interposed (Vogelmann et al., 2011; Ghirlando et al., 2012; Phillips-Cremins and Corces, 2013). The juxtaposition of enhancers to promoters, through long-range contacts, defines a key feature in activating gene expression (Deng et al., 2012) and is required for the regulation of a vast number of genes involving cohesin (Kagey et al., 2010; Xiao et al., 2011). In vertebrates, however, where the only characterized IBP is CTCF, its thousands of binding sites were shown to have little effect on the overall interaction levels between enhancers and promoters (Sanyal et al., 2012). Rather, vertebrate CTCF may participate in functional long-range interactions between distant regulatory elements (Handoko et al., 2011; Li et al., 2011) and act together with cohesin in defining TADs (Phillips-Cremins et al., 2013). Similarly, *Drosophila* IBPs could favor preferential long-range interactions between distant TAD borders (Hou et al., 2012), possibly participating in the clustering of active, gene-dense regions near borders away from silenced regions. Such interactions implicated multiple *Drosophila* IBPs including dCTCF, Beaf32, and cofactors such as cohesin, chromator, or CP190 (Wood et al., 2011; Sexton et al., 2012; Hou et al., 2012). Functional contacts may further depend on cellular or genomic contexts including the presence of nearby regulatory elements and/or of additional *Drosophila* IBPs, GAGA Factor (GAF), Zest-white5 (Zw5), or suppressor of Hairy-wing (Su(Hw)) (Gerasimova et al., 2007; Nègre et al., 2010; Wood et al., 2011;

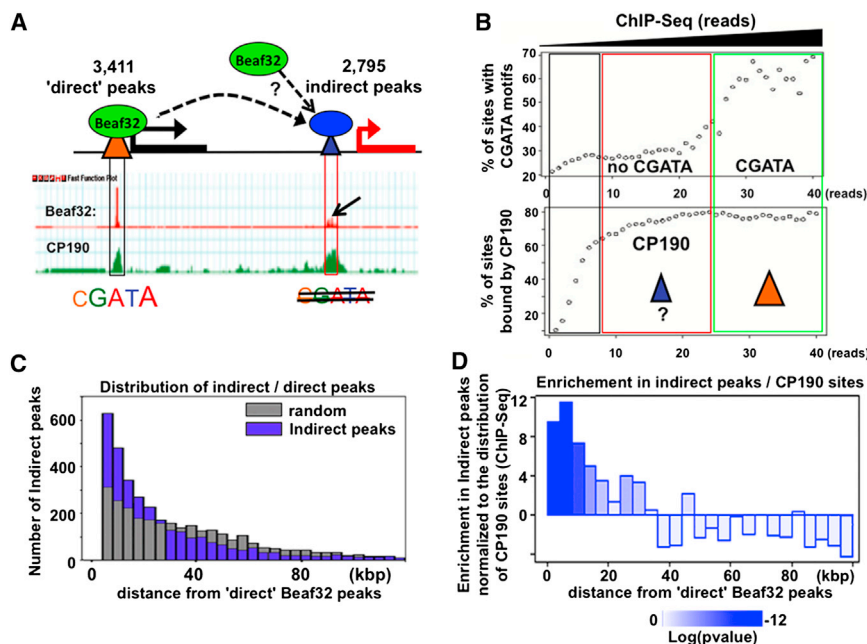


Figure 1. Identification of Beaf32 ChIP-Indirect Peaks in Promoters of Distant Genes

(A) Genomic view accessible through our Gbrowser (http://insulators_chromosome-dynamics.biotoul.fr/) showing ChIP-seq analysis of Beaf32 sites highlighting two possible modes of Beaf32 binding: on the left is one of the 3,411 direct Beaf32 peaks (ChIP-seq peaks > 25 reads; see Supplemental Experimental Procedures) sharing its DNA consensus "CGATA" as indicated, accounting for direct DNA binding to these sites; on the right are four of the 2,795 "indirect" peaks of lower intensity (ChIP-seq peaks 10–25 reads) not enriched in the CGATA consensus. Also shown are the ChIP profile of CP190, whose peaks are enriched in ChIP-indirect peaks of Beaf32 (see C; see text). Note that DREF consensus is not enriched in indirect peaks (tATCGATa; Figure S1C). The dotted arrows represent two hypotheses by which Beaf32 could interact with indirect peaks: either from a soluble pool or from Beaf32 already bound to direct peaks. (B) Percentage of direct Beaf32 peaks harboring the CGATA binding motif (y axis) as a function of peak intensities (x axis, in ChIP-seq reads). Lower panel: percentage of peaks bound by CP190 (y axis; see Experimental Procedures) as a function of peak intensities (x axis). See also Figure S1.

(C) Graph showing the specific enrichment of CGATA-less "indirect" Beaf32 peaks near direct Beaf32 peaks as compared to "random" control (see Experimental Procedures).

(D) Normalized enrichment of CGATA-less Beaf32 ChIP-indirect near direct peaks to the distribution of CP190 sites. CP190 sites harboring an indirect peak were normalized to the total number of CP190 sites within the same interval (Z score; see Supplemental Experimental Procedures; y axis). The colored bar shows the standard deviation in log p value.

Gohl et al., 2011). Insulators were also implicated in mediating specific long-range contacts with "paused" RNA Polymerase II (RNAPII) (Chopra et al., 2009), a key transcriptional stage controlling developmentally regulated genes (Hendrix et al., 2008; Core and Lis, 2008; Gilchrist et al., 2010). High-resolution mapping of insulator-prone long-range contacts may help clarify how multiple *Drosophila* IBPs influence gene expression.

Here, we detect by ChIP-Seq the long-range interaction sites of Beaf32, dCTCF, and GAF as ChIP-indirect peaks. Indirect peaks highlight a network of functional long-range contacts among distinct IBP sites through their common cofactors, CP190, as confirmed by aggregating genome-wide Hi-C data over indirect peaks of IBPs at high resolution. The functional relevance of indirect peaks was further addressed using synthetic IBP mutants that prevented interactions with CP190, which functionally impaired the expression of distant genes associated with indirect peaks. These features are largely dependent on RNAPII pausing, highlighting a functional interplay between IBPs and this key transcriptional stage.

RESULTS

ChIP-Seq Highlights Two Possible Binding Modes of Beaf32 to Chromatin

Clusters (three or more) of CGATA motifs are the hallmark of Beaf32 genomic binding sites (Emberly et al., 2008; Bushey et al., 2009; Nègre et al., 2010). ChIP-Seq confirmed the enrichment of these motifs for Beaf32 binding, and we refer to these CGATA-containing peaks as "direct" Beaf32 peaks (left peak,

Figure 1A; see also Figures S1A and S1B). Further inspection of the ChIP-Seq signal highlighted an additional subset of 2,795 peaks of lower intensity (Figure 1A, right peak), which were previously ignored as being below thresholds for peak detection (Jiang et al., 2009). Such peaks were, however, enriched close to promoters (76.9% < 250 bp from TSS), similarly to the 3,411 direct peaks (91.1%), supporting their significance as compared to background signal (Figure S1C, compare middle and lower panels). Unlike direct peaks, peaks of lower intensities did not share the Beaf CGATA consensus (Figure 1B), and they were called thereafter "indirect peaks," as Beaf32 might not bind directly to DNA at these sites. Less than 0.5% of the indirect peaks overlapped with the binding sites of 32A, an isoform of Beaf that has little influence on its binding to chromatin (Jiang et al., 2009), and less than 5% contained the related DREF consensus ("tATCGATa"; Figures S1B and S1C), showing that these factors may not account for the indirect peaks of Beaf32.

Interestingly, Beaf32 depletion affected the expression of genes whose promoters harbored indirect peaks (p value < 1×10^{-200} ; data not shown), highlighting their potential relevance in regulating gene expression. Furthermore, the averaged Beaf32 binding profiles of all promoters harboring either zero or one CGATA were highly similar (Figure S1D), arguing against the possibility that indirect peaks are solely due to the weaker interaction of soluble Beaf32 with such DNA sites harboring one (or a degenerated) consensus CGATA motif. Indirect peaks might be due to protein interactions of soluble Beaf32 with factors already bound to promoter regions (Figure 1A, see model). The averaged binding levels of Beaf32

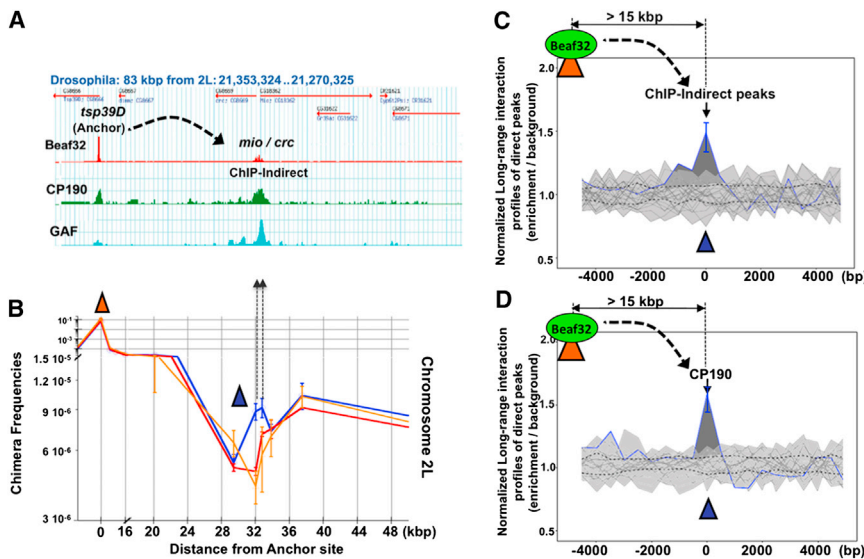


Figure 2. ChIP-Indirect Peaks Reflect Long-Range Interaction Sites

(A) Genomic view from our Gbrower showing the *tsp39D* and *mio/crc* loci from chromosome 2L that harbor direct and indirect peaks, respectively. Also indicated are ChIP data for Beaf32 (red), CP190 (green), and GAF (blue; from Nègre et al., 2010; see also Figures S2 and S3). The arrow represents the interactions as detected by 3C (see B). All genomic features are shown in the flipped orientation.

(B) 3C analysis of long-range interactions between direct Beaf32 peak at *tsp39D* and indirect peaks at *mio/crc*. The graph represents the relative frequency chimera products as measured by qPCR from Beaf32- (red), CP190- (orange), or control mock- (blue) depleted cells (see Experimental Procedures). Chimeras (ligation product) were measured using the HindIII restriction (2L21342022) site prior to the direct Beaf32 peak as anchor site (reverse primer and TaqMan-MGB probe) and primers spanning the whole locus. Variations were tested by Student's t test (see Supplemental Experimental Procedures). Error bars are standard errors to the mean from three independent preparations of 3C DNA (see Figure S2F for 3C using ChIP-indirect peaks as anchor sites).

(C) Normalized Hi-C profiles obtained from Hi-C data (Sexton et al., 2012) probing the long-range interactions of distant/direct Beaf32 peaks (>15 kbp away; orange triangle/green oval) over ChIP-indirect peaks (position 0; x axis) as compared to control windows. Long-range interaction profiles were aggregated over ChIP-indirect peaks or control neighbor windows (see Supplemental Experimental Procedures; as adapted from Jee et al., 2011), resulting in the amplification of the signal allowing us to accurately probe interactions over 500 bp windows. The relative enrichment in interactions (y axis) was obtained by normalizing interaction levels between direct and indirect peaks over background interaction levels between the same direct peaks and the 26 neighboring bins (± 3.5 ; up to 9 kbp). The error bar represents the variation of the measures (see Supplemental Experimental Procedures for details).

(D) Normalized Hi-C profiles through aggregation plots as in (C), except that enrichment in long-range interactions was measured over CP190 binding sites (position 0) as compared to neighbor control sites.

were slightly higher for promoters close to (<20 kbp) direct peaks compared to those localized more distantly (>100 kbp; Figure S1D), suggesting a possible physical link between ChIP-indirect and -direct peaks. Supporting this idea, indirect peaks were largely enriched nearby (<40 kbp) "direct" Beaf32 peaks (Figure 1C). Given the role of IBPs in long-range interactions, these results raised the hypothesis that ChIP-indirect peaks could actually reflect long-range interaction sites of direct Beaf32 binding sites with distant genomic loci.

ChIP-Indirect Peaks Highlight Long-Range Interactions

We performed 3C analysis to examine how long-range interactions arise between direct and indirect peaks. Beaf32 depletion reduced long-range interactions between the direct *tsp39D* peak and ChIP-indirect peaks within *crc/mio* promoters compared to controls (Figures 2A and 2B; see Experimental Procedures), indicating that ChIP-indirect peaks could reflect long-range interaction sites of Beaf32. Further genome-wide identification of ChIP-indirect peaks in embryos (Figures S2A and S2B; see Supplemental Experimental Procedures) prompted us to test their relevance in long-range interactions by inspection of Hi-C data available in embryos (Sexton et al., 2012). Long-range interactions were measured genome-wide between direct peaks and distant (>15 kbp) indirect peaks through aggregation plots (Figure 2C; see Supplemental Experimental Procedures). Such genome-wide analysis showed a clear enrichment for long-range interactions of the direct peaks over the indirect peaks, as compared to background interaction

levels (p value $\sim 1 \times 10^{-7}$; see Supplemental Experimental Procedures) at a high resolution (500 bp; Figure 2C). The higher frequencies of contacts at ChIP-indirect peaks demonstrate that they reflect long-range interactions with direct Beaf32 peaks, in complete agreement with our 3C data. Further inspection of long-range interactions depending on the presence or not of overlapping dCTCF (or GAF) binding sites showed no variations in the proportion of interactions (Figure S2C), showing that Beaf32 itself may be important. Furthermore, the preferential interactions were detected for ChIP-indirect sites more distantly localized from the direct sites of Beaf32 (>100 kbp; Figure S2D), although long-range interactions of direct peaks with the indirect peaks may be favored within TADs (see below).

CP190 Interacts with Beaf32 and Is Required for Long-Range Interactions

We next sought to identify factors that might recruit Beaf32 to ChIP-indirect peaks independently of its DNA-binding activity to CGATA motifs (Emberly et al., 2008). The Beaf32 complex was purified by affinity chromatography (Figures 3A and 3B) using our specific peptide-raised antibodies (Figure S3A). Mass spectrometric analysis showed that the major copurified protein was centrosomal protein 190 (CP190; 154 peptides; Figure 3B), as confirmed by coimmunoprecipitation experiments using anti-CP190 antibodies (Figure S3B). CP190 interacts with other IBPs including dCTCF (Gerasimova et al., 2007), and its recruitment may play a pivotal role in insulator function through chromatin looping (Wood et al., 2011). CP190 binding sites

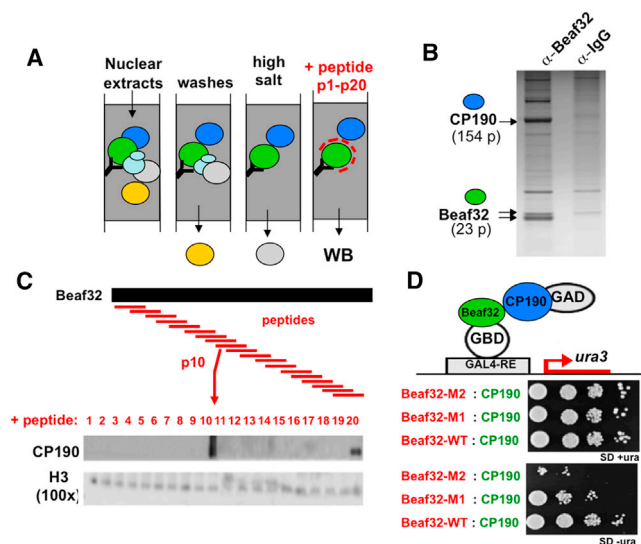


Figure 3. Biochemical Purification of Beaf32 Complex Identifies CP190 as a Major Cofactor

(A) Affinity chromatography procedure to purify the Beaf32 complex followed by competition with either the C-terminal peptide of Beaf32 (B) or peptides covering the entire Beaf32 sequence (C).

(B) Coomassie-stained SDS-PAGE gel after elution of affinity-purified complex with the C-terminal peptide as a competitor added on beads coupled to purified anti-Beaf32 antibodies or IgG control (see Figure S2A; Supplemental Experimental Procedures). The arrow shows the CP190 protein as confirmed by mass spectrometric analysis (144 peptides versus 2 peptides in IgG control) (see also Figures S2B and S2C).

(C) Following affinity chromatography, Beaf32-CP190 interactions were competed by adding one of the (18 aa long) peptides spanning the entire Beaf32 (see Experimental Procedures). The elution by addition of each peptide was assessed by western blotting using anti-CP190 (see Figure S2B) or anti-histone H3 antibodies as loading control. The sequence of peptide number 10 ("p10") is SEDPLCYSPHVMDDDEGL. "100x" indicates the exposure time (~30 min) to detect H3. See also Figure S3.

(D) Two-hybrid assays probing the interactions between wild-type or mutant Beaf32 and CP190. For Beaf32 mutants, 2 or 4 mutations ("Beaf32-M1/2," respectively; see Experimental Procedures) were introduced into the region corresponding to "p10" (C). SD ± ura corresponds to the same batch of cells spot onto media containing uracil (growth control) or not, respectively.

largely overlap with the other insulator proteins, Beaf32 and GAF, suggesting that it is an essential component of all *Drosophila* insulator types (Bartkuhn et al., 2009; Bushey et al., 2009; Nègre et al., 2010). In agreement, a majority (~80%) of the Beaf32 indirect peaks corresponded to regions bound by CP190 (Figure 1C; p value $< 1 \times 10^{-300}$; Table S1), thus raising the possibility that indirect peaks might arise from long-range interactions through CP190. CP190 sites were statistically enriched in ChIP-indirect peaks only if they localized nearby (< 40 kbp) a direct Beaf32 site, not if more distantly localized (> 100 kbp; Figure 1D), arguing against the possibility that indirect peaks could simply result from the recruitment/interaction of soluble Beaf32 to CP190 sites.

Supporting this view, 3C analysis showed that CP190 depletion significantly reduced the levels of long-range interactions between the direct *tsp39D* peak and ChIP-indirect peaks compared to control cells (Figure 2B), as also supported by 3C

analyses of the mirror long-range interactions using the ChIP-indirect peaks as anchor sites (Figure S2F). Strongly supporting these results, genome-wide analyses of Hi-C data between direct Beaf32 peaks and distant CP190 peaks showed a clear enrichment in long-range interactions as compared to neighboring control regions (Figure 2D; p value $\sim 1 \times 10^{-9}$).

Given the involvement of CP190 in long-range contacts, we next sought to test whether its interaction with Beaf was necessary for ChIP-indirect peaks by producing Beaf mutants that impaired such interaction. A peptide array that covers the whole Beaf32 protein sequence (Figure 3C) was tested for its ability to elute CP190 from affinity-purified Beaf32 complexes. Since peptide 10 ("p10") was most efficient in eluting CP190, we introduced four point mutations within the corresponding region of *beaf32* (see Supplemental Experimental Procedures), thereby producing mutants that clearly impaired direct interaction of Beaf32 with CP190, as confirmed using 2-Hybrid (Figure 3D).

Introducing silent mutations within the RNAi targeted region of *beaf32* (NNN \rightarrow NNn over a 800 bp region; see Experimental Procedures) allowed us to express a synthetic *beaf* gene from stably transfected cell lines in the absence of endogenous *beaf32* (Figures S4A–S4D). 3C analysis in stably transfected cells showed that the produced Beaf mutants could specifically impair the long-range interaction observed between the direct and ChIP-indirect peaks of *mio/crc* as compared to WT controls (Figures 4A and 4B). Coimmunoprecipitation experiments performed in the same cellular contexts confirmed that such mutants specifically impaired interactions with CP190 as compared to WT controls (Figure S4C), strongly supporting a role of CP190 as a key cofactor for long-range interactions.

Synthetic Beaf Mutants Impair Binding to ChIP-Indirect Peaks

ChIP analysis in the context of ChIP-direct or -indirect peaks showed that mutant Beaf proteins (thereafter called "mut-Beaf") showed similar levels of binding to all (10/10) direct peaks tested as compared to WT controls (Figures 4C and S3D, left panels) for similar levels of expression (Figure S4B). Supporting this result, in vitro analyses using fluorescence anisotropy confirmed that the DNA-binding activity of mut-Beaf was not impaired (Figures S4E and S4F). In stark contrast, mut-Beaf systematically reduced its binding to all indirect peaks tested (10/10) as compared to WT-Beaf (Figures 4C and S3D, right panels; p value 1.6×10^{-3}), strongly supporting that the interaction with CP190 was required for its binding to the distant ChIP-indirect peaks.

ChIP-indirect peaks localized nearby direct peaks (< 40 kbp) independently of the distribution of CP190 sites (Figure 1D). We thus tested whether direct Beaf32 peaks were needed for the formation of nearby indirect peaks using reporter constructs that harbor a direct binding site flanking the upstream reporter gene (Figure 4D, \pm orange triangle). ChIP analysis showed that Beaf32 binding was detected near the TSS of the distant CGATA-less promoter of *hsp27* (Figure 4E; compare first two bars). By contrast, Beaf32 binding to the distant promoter was abolished in the absence of the upstream binding site (Figure 4E, "no Beaf32 site"), showing that binding to the indirect peak was

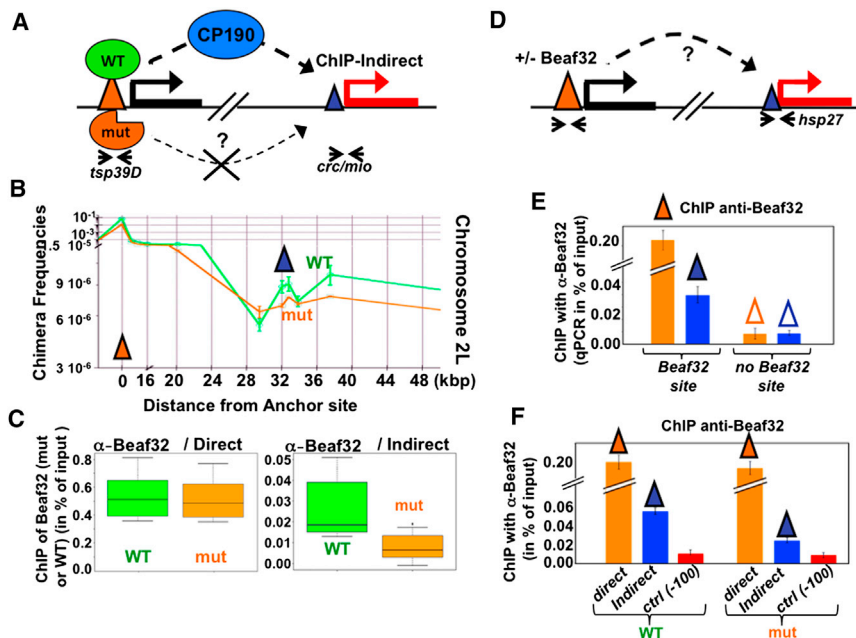


Figure 4. Synthetic Beaf32 Mutants Impair Its Interaction to ChIP-Indirect Sites

(A) Scheme representing WT- and mutant-Beaf proteins (green and orange ovals, respectively) with respect to their predicted binding to direct or indirect peaks (orange and blue triangles, respectively) as tested by 3C and ChIP (see B and C).

(B) Relative frequency chimera products (obtained from 3C data) measured by qPCR analyses in multiple independent 3C assays in Beaf32-depleted cells expressing mutant-Beaf as compared to WT control (orange and green curves, respectively). The x axis represents the distance from anchor site (direct Beaf32 peak at *tsp39D*). Chimeras (ligation product) were measured with a set of primers spanning the whole locus relative to three control sites using TaqMan-MGB probe (see [Supplemental Experimental Procedures](#)). Error bars are standard errors from two independent 3C DNA preparations.

(C) Box plots showing the binding levels of WT/mutant-Beaf proteins (see A) as measured by qPCR of ChIP samples precipitated with anti-Beaf32 or IgG control antibodies, from Beaf32-depleted cells stably transfected with WT/mutant

Beaf. The binding levels were measured corresponding to ten direct Beaf32 sites and ten nearby ChIP-indirect peaks (see [Supplemental Experimental Procedures](#), Figure S3D).

(D) Scheme representing the constructs used to measure the interactions of WT- or mutant-Beaf proteins to the CGATA-less promoter region (blue triangle) of *hsp27* in the presence or absence of an upstream Beaf32 site (orange triangle), as tested by ChIP (E and F).

(E) Histogram showing the binding of Beaf proteins to ChIP-indirect peaks (blue) in the presence or absence of an upstream direct Beaf32 peak (orange triangle \pm) as measured by qPCR of ChIP with anti-Beaf32 or IgG control antibodies.

(F) Histogram showing the binding of WT/mutant-Beaf proteins to ChIP direct (orange) or indirect (blue) peaks as measured by ChIP with anti-Beaf32 or IgG control antibodies, from Beaf32-depleted cells stably transfected with WT/mutant Beaf. See also [Figure S4](#). Error bars are standard errors from three independent measures.

dependent on the upstream direct peak. Moreover, mut-Beaf also impaired binding to the indirect peak, for similar binding levels to the direct peak (Figure 4F). Taken altogether, our data thus strengthened the hypothesis that ChIP-direct and -indirect peaks highlight two modes of binding of Beaf32. The former depends on DNA-binding activity of Beaf32 and the presence of multiple CGATAs, whereas the latter reflects long-range contacts with distant loci depending on interactions with cofactors such as CP190.

ChIP-Indirect Peaks Highlight Functional Long-Range Interactions

To further investigate how the loss of Beaf-CP190 interactions might impact the expression of distant genes, we compared genome-wide expression levels in our stably transfected cell lines expressing synthetic mut-Beaf compared to WT-Beaf. A significant proportion of the differentially expressed (DE) genes corresponded to “primary targets” (Figure 5A, orange pie), i.e., genes whose promoters harbored a direct Beaf32 binding site (>45%; Table S2). By contrast, genes showing no changes in expression were not enriched in direct peaks (Figure 5A, middle panel). As such, impairing Beaf32-CP190 interactions may alter CP190 recruitment to primary targets (see below), which would impair their expression as suggested for other insulators (Gerasimova et al., 2007; Wood et al., 2011).

Curiously, the percentage of DE genes whose promoters were bound by CP190 was higher (>70%; Table S2) compared

to those bound by Beaf32 (<48%). As such, hundreds of genes were differentially regulated in mut-Beaf compared to WT-Beaf, yet they did not harbor a direct Beaf32 binding site. This result suggested that Beaf-CP190 interactions regulate a subset of genes whose promoters were bound by a distinct IBP and CP190, yet not (directly) by Beaf32, hereafter called “long-range targets.” Strikingly, the chance that Beaf32 regulated such genes depended on the presence of indirect peaks in their promoters (Figure 5B, p value $\sim 1 \times 10^{-29}$) that were further enriched in dCTCF or GAF sites, as evidenced from three independent RNA-seq. Long-range targets were also statistically enriched within < 40 kbp distances from a direct peak (Figure 5C). Therefore, dCTCF/GAF may recruit CP190 to promoters as shown (Bushey et al., 2009; Bartkuhn et al., 2009; Nègre et al., 2010; Wood et al., 2011) (see below), thereby promoting long-range contacts with distant Beaf32 promoters through Beaf32-CP190 interactions (Figure 5D). Supporting this view, mutant Beaf did not affect the expression of *ctcf/gaf* (Figures S4G and S4H), and dCTCF/GAF sites were largely enriched in the ChIP-indirect peaks of Beaf32 (~43%; p values < 1×10^{-37}). Moreover, long-range targets whose promoters were not bound by dCTCF or GAF were specifically enriched in the binding sites of Su(Hw) (47 promoters; p value 1×10^{-7}). This represents a fourth family of *Drosophila* insulators that was originally shown to recruit CP190 to chromatin (Gerasimova et al., 2007), confirming the relevance of ChIP-indirect peaks with respect to the regulation of gene expression

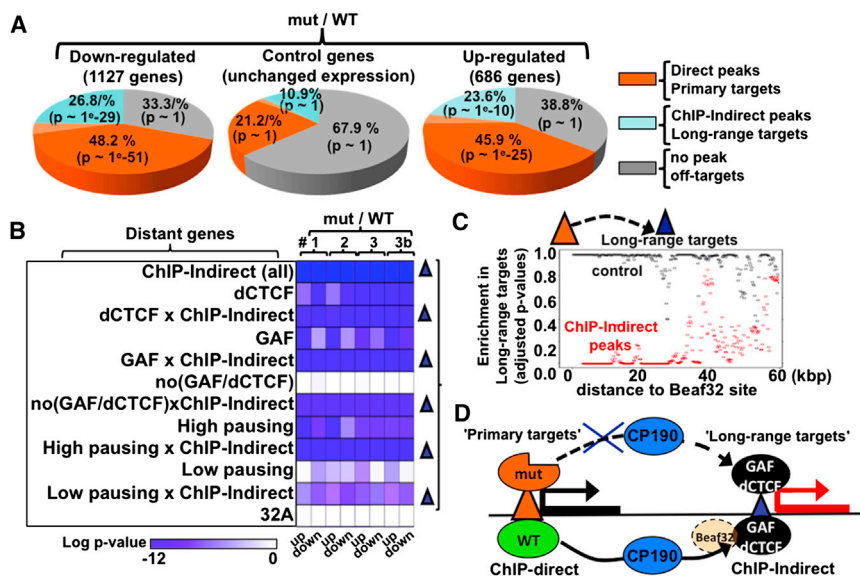


Figure 5. Genome-wide Analyses Highlight a Genome-wide Impact of Beaf32 on Distant Genes through CP190.

(A) Differentially expressed (DE) genes (down-regulated/upregulated) or control genes for cell lines expressing mutant-Beaf compared to WT-Beaf (see Supplemental Experimental Procedures) depending on the presence or not of direct or indirect peaks versus no peak. "Primary targets" (543/315 down-/upregulated, respectively) and "long-range targets" (302/162 down-/upregulated; see Table S3 for a list) correspond to genes whose promoters harbor a direct (CGATA-containing) or a ChIP-indirect (CGATA-less) peak, respectively.

(B) Heat map showing the enrichment (in log p value) of long-range targets as obtained by RNA-seq depending on the presence in their promoters of dCTCF or GAF binding sites and/or of a ChIP-indirect Beaf32 peak (blue triangles). Numbers 1, 2, and 3 represent three independent replicates. Column 3b corresponds to DE genes with a minimal fold change in expression of 1.6x (see Figure S4). Also indicated is the enrichment in

long-range targets as a function of high or low RNAPII pausing (see text) as previously defined (Gilchrist et al., 2010) (see Supplemental Experimental Procedures). Note the absence of enrichment for binding sites unique to the 32A isoform of Beaf ("32A") in agreement with its little influence on long-range interactions (see Figure S2C).

(C) Graphs showing the distribution of DE genes as a function of their distance to a direct Beaf32 peaks (0; x axis). The y axis indicates the enrichment for DE genes (adjusted p values with Benjamini & Hochberg correction) as a function of the presence (red curve) or not (black curve) of a ChIP-indirect peak in promoters.

(D) Model for Beaf32-mediated regulation of primary or long-range targets (see also Figure S5 for a similar analysis of dCTCF-mediated regulation of gene expression through its own ChIP-indirect peaks).

through long-range interactions implicating distinct IBPs (Figure 5D).

A similar approach led us to identify hundreds of ChIP-indirect peaks from previous ChIP-Seq of dCTCF and ChIP-chip of GAF (Figures S5A and S5B; see Supplemental Experimental Procedures). Approximately 40% and 48% of the ChIP-indirect peaks of dCTCF and GAF could be detected as Beaf32 direct peaks, respectively (p values $1 \times 10^{-59}/1 \times 10^{-31}$; Table S4), in total agreement with our data showing that such direct/indirect peaks reflected long-range interactions among various IBPs. Although dCTCF depletion did not affect the expression of *beaf32* or *gaf* (Figure S5C), it may specifically impair the expression of long-range target genes harboring dCTCF ChIP-indirect peaks (Figure S5D), within ~40 kbp distances from its direct binding sites (Figures S5E and S5F; Table S5). These corresponded to genes highly enriched in their promoters in Beaf32 or GAF (direct) binding sites (Figure S5G), thus validating the reciprocity in detection of ChIP-indirect peaks. Furthermore, ChIP-indirect peaks of dCTCF showed preferential long-range interactions with its direct (Figure S5I), confirming the relevance of ChIP-indirect peaks for the detection of specific long-range contacts involving multiple *Drosophila* IBPs together with shared cofactors, such as CP190 (Figure 5D).

Beaf Mutants Regulate CP190 Levels Near TSSs of Long-Range Targets

Given that CP190 recruitment is a major regulatable step in insulator function involving long-range interactions (Wood et al., 2011), we next inspected whether Beaf mutant might

interfere with its binding. CP190 is recruited by various IBPs including dCTCF, Su(Hw) (Gerasimova et al., 2007; Bushey et al., 2009; Wood et al., 2011), and GAF (Bartkuhn et al., 2009) (Figure S6A) and possibly through interactions with Beaf32 (Figure 2). Supporting this view, ChIP-seq analysis showed that CP190 binding levels were higher for gene promoters harboring a direct Beaf32 peak as compared to control promoters (Figure S6A), in complete agreement with their enrichment in CP190 sites (Figure S6B). ChIP-seq analysis of CP190 in stably transfected cell lines showed that mut-Beaf significantly decreased CP190 binding levels as compared to WT-Beaf, provided Beaf32 was the only insulator protein bound to promoters (Figure 6A; p value 1×10^{-5}), as confirmed by qPCR analysis (Figure S6C). By contrast, mut-Beaf did not influence CP190 levels for promoters harboring an overlapping binding site for dCTCF or GAF (Figure 6A; p value ~1), in total agreement with their ability to recruit CP190 independently of Beaf32.

dCTCF/GAF sites were enriched among the indirect peaks of Beaf32 including those flanking long-range targets (Figures 5B and S6D), where these IBPs may recruit CP190 to such promoters independently of Beaf32. Surprisingly, inspection of our genome browser showed that mut-Beaf, however, affected CP190 binding near *mio/crc* as compared to WT-Beaf (Figure 6B), as also evidenced by the overall increase in CP190 levels near the TSSs of all long-range targets (Figure 6C). Such variations of CP190 levels strongly depended on the presence of ChIP-indirect peaks of Beaf32 and/or a dCTCF or a GAF site in promoters (Figure 6D) for genes most often localized < 40 kbp from a direct Beaf32 site (Figure S6E). Given the pivotal role of

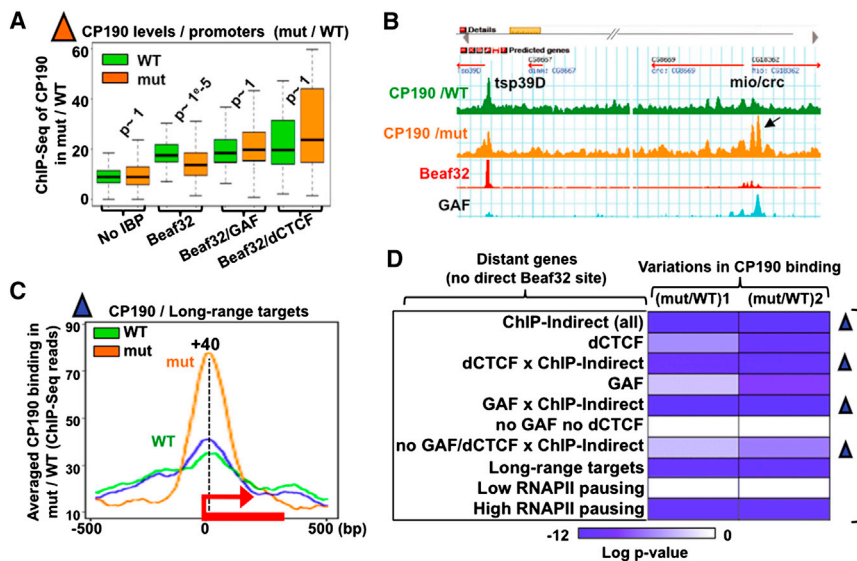


Figure 6. Beaf32 Influences CP190 Binding over Promoters of Long-Range Targets

(A) Boxplot showing our ChIP analysis with anti-CP190 antibodies in (Beaf32-depleted) cells stably transfected with WT/mutant-Beaf. The levels of CP190 binding are indicated as averaged ChIP-seq reads/promoter (y axis) overlapping or not with a dCTCF or GAF peak as indicated. The p values (pairwise Wilcoxon test) indicate the difference in CP190 binding to promoters between cells expressing WT- or mut-Beaf (see text; see also Figure S6).

(B) ChIP-seq analysis of CP190 for the *tsp39D* and *mio/crc* direct and indirect peaks, respectively, in cells expressing WT- or mut-Beaf. Also indicated are ChIP data for Beaf32 (red) and of GAF (blue).

(C) Averaged CP190 binding for promoters of long-range targets in cells expressing mut-Beaf (orange curve) compared to cells expressing WT-Beaf or to wild-type cells (green and blue curves, respectively).

(D) Heat map showing the variations (in log ratio) in CP190 binding levels between cells expressing

WT- versus mut-Beaf for long-range targets (genes without a direct Beaf binding site) in two independent ChIP-seq analyses ("mut/WT1/2") according to the presence or not of a dCTCF or GAF site and/or of a ChIP-indirect Beaf32 peak (blue triangles) in their promoters and/or their low/high degree of RNAPII pausing, as previously defined (Gilchrist et al., 2010) (see Supplemental Experimental Procedures for details). See also Figure S6.

CP190 in insulating function (Wood et al., 2011), these results strengthened our expression data showing the impact of mutants on long-range targets, involving long-range contacts mediated through Beaf32-CP190 interactions to dynamically modulate CP190 binding over their TSSs. Therefore, our results strengthened the functional interplay among multiple *Drosophila* IBPs through long-range interactions.

ChIP-Indirect Peaks Highlight the Link between CP190 and RNAPII Pausing

Our ChIP-seq data highlighted a CP190 peak near +40 after TSSs (Figures 6C and 7A), thus raising the possibility that such binding is linked to RNAPII pausing (Gilchrist et al., 2010). This transcriptional stage was initially observed for GAF/heat shock promoters, and it defines a key mechanism of transcriptional control regulated genome-wide by NELF (Hendrix et al., 2008; Core and Lis, 2008; Gilchrist et al., 2010). Interestingly, this transcriptional stage has been implicated in insulator function involving specific long-range contacts between insulator sites and a paused polymerase (Chopra et al., 2009).

The averaged binding levels of CP190 tightly correlated with the levels of RNAPII pausing (Figure 7A) independently of transcriptional activity (see Supplemental Experimental Procedures). Given the specific long-range contacts of direct Beaf32 sites with CP190 (Figure 2D), our data raised the possibility that long-range interactions implicating IBPs might involve RNAPII pausing. The regulation of long-range targets tightly correlated with RNAPII pausing (Figure 5B), similarly to variations of CP190 levels (Figure 6D). Furthermore, Beaf mutants significantly increased the levels of CP190 over paused genes as compared to WT controls (Figure S6F), which further depended on their proximity to a direct Beaf32 binding site (<40 kbp away). Mut-Beaf may not directly affect the expression of *nelf* as compared to WT-Beaf (Figure S7A; not shown),

yet its influence on long-range targets was largely dependent on their degree of RNAPII pausing, unlike what was found for primary targets or control genes (Figure 7B and Table S6). Therefore, RNAPII pausing may define a key stage for regulations involving CP190-mediated long-range contacts, which are impaired by Beaf mutants. Hi-C data confirmed that highly paused genes define preferential interaction sites with Beaf32 (Figure 7C). In complete agreement, reinspection of the positioning of Beaf32 indirect peaks over long-range targets showed that the average Beaf32 binding profile was shifted downstream of the TSSs of long-range targets (Figure S7C; peaks centered near +30 bp after TSSs). Direct Beaf32 sites could define preferential long-range interaction sites with GAF (Figure S7D), and a high proportion of the long-range targets of Beaf32 harbored GAF binding sites (Table S6 and Figure S7B) as for the *hsp27* reporter gene (Figures 4B and 4C). Mutant-Beaf led to reduce its expression by ~5-fold compared to WT-Beaf (Figure S7E), an effect that was impaired upon removal of the GAF site (Figure S7E; compare "GAF" and "no GAF"). Taken altogether, our data strongly support a role of the distinct insulator proteins in functional long-range interactions through cofactors including CP190 (Figure 7D), illustrating how ChIP-indirect peaks may represent a major tool toward our understanding of the functional higher-order chromatin organization.

DISCUSSION

Our genome-wide analysis of ChIP data highlights that the binding of the insulator proteins dCTCF, Beaf32, and GAF is detected on one side of loops as ChIP-direct peaks, where they bind directly to their DNA binding motifs, as well as on the other side of loops as ChIP-indirect peaks. We show that ChIP-indirect peaks are heterotypic long-range interactions that provide key

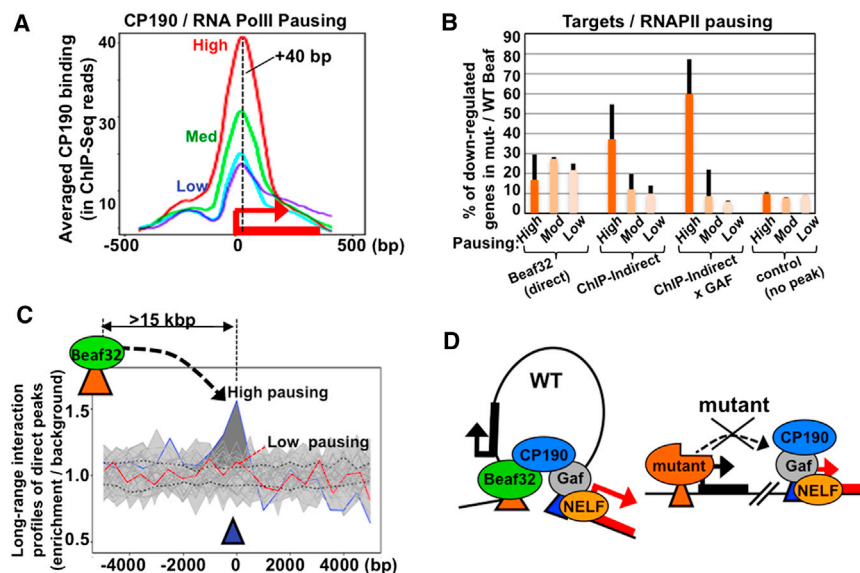


Figure 7. CP190 Binding Marks Highly RNAPII Paused Genes, Thereby Favoring Long-Range Interactions

(A) Averaged binding of CP190 in WT cells as quantified by ChIP-seq (y axis; in ChIP-seq reads normalized to IgG control) as a function of high, medium, or low (red, green, and blue curves) RNAPII pausing as previously defined (Gilchrist et al., 2010) (see Supplemental Experimental Procedures).

(B) Graph showing the percentages of differentially regulated genes in cells expressing mutant compared to WT-Beaf depending on the presence in their promoters of a direct Beaf32 binding site ("primary targets"), a ChIP-indirect site ("long-range targets") with or without a GAF site, as a function of the high, moderate, or low RNAPII pausing as previously defined (Gilchrist et al., 2010). Error bars are standard errors from three independent RNA-seq.

(C) Normalized Hi-C profiles by aggregation plots (see Figure 2C for details) probing the enrichment in long-range interactions between direct peaks of Beaf32 (orange triangle/green oval)

over promoters harboring high or low RNAPII pausing levels (p value of 10^{-7} and 1, respectively). See also Figure S7.

(D) Model: CP190-mediated long-range interactions model for the interplay between insulator-regulated gene expression and long-range interactions targeting genes depending on their degree of RNAPII pausing (see text).

information regarding the factors mediating specific long-range contacts as detected by 3C/Hi-C. In particular, the presence of ChIP-indirect peaks revealing long-range contacts between insulators and RNAPII pausing sites is highly relevant to their function, as pausing has been implicated in insulator activities (Chopra et al., 2009). Moreover, depleting one IBP affects the expression of genes localized on the other side of specific loops that possess its indirect peaks, through long-range contacts implicating a distinct IBP. ChIP-indirect peaks therefore bring valuable information given the massive usage of genome-wide ChIP available in different species, cell types, and/or developmental stages.

Functional long-range interactions among nearby regulatory elements may largely depend on the physical organization of chromosomes into TADs (Sanyal et al., 2012; Sexton et al., 2012; Hou et al., 2012). Very recently, vertebrate CTCF together with cohesin were implicated as architectural proteins contributing to define TADs (Phillips-Cremins et al., 2013). Similarly, combination of the *Drosophila* IBP dCTCF/Beaf32/CP190 sites would favor long-range contacts between distant topological borders, which may contribute to the clustering of active regions (near borders) from silenced chromatin (in the interior) (Hou et al., 2012). Such organization may in turn favor shorter-range, cell type-specific interactions between distant regulatory elements (Phillips-Cremins et al., 2013), possibly through interactions with additional *Drosophila* IBPs. Supporting this model, long-range contacts between direct and indirect peaks can be detected through aggregation of Hi-C data over the later peaks, or over CP190 sites, which appears to reflect functional contacts regulating distant genes. Also, the recruitment (and/or variations) of CP190 to insulator sites may define a key regulatable step in insulating function (Wood et al., 2011). Beaf mutants impact CP190 levels over the paused sites of long-range targets,

providing strong support for the regulation of distant genes through specific long-range contacts. Interactions between direct and indirect peaks are restricted in the presence of intervening topological borders (Figure S7F), supporting the idea that functional interactions may take place within predefined TADs (Sexton et al., 2012; Hou et al., 2012; Phillips-Cremins et al., 2013). TAD borders are further enriched in specific IBP/CP190 sites that may trap long-range contacts, thereby limiting functional interactions across borders. Such insulator-based physical organization may account for why chromatin is more permissive to gene expression near borders, independently of histone modifications (Hou et al., 2012). IBP-prone long-range interactions may thus condition their limited influence on nearby epigenetically defined domains (Cuddapah et al., 2009; Schwartz et al., 2012; Ghirlando et al., 2012; Tanay and Cavalli, 2013).

CP190 distributes into a distinct pattern called insulator "bodies" (Gerasimova et al., 2007) providing an interaction platform among IBPs through specific chromosomal interactions (Wood et al., 2011). The link between insulator bodies and long-range interactions remains not totally clear, however (Vogelmann et al., 2011), maybe in part because long-range interactions among IBPs are highly dynamic (Wood et al., 2011). The dynamic (dis-) assembly of chromatin hubs has been linked to the RNAPII pausing-elongation switch (Vernimmen et al., 2007; Eskiw and Fraser, 2011). Our data indicate that RNAPII pausing regulated by GAF may serve as a bait for long-range contacts with IBP (dCTCF or Beaf32) sites. RNAPII pausing regulates the expression of developmentally regulated genes (Hendrix et al., 2008; Lee et al., 2008; Gilchrist et al., 2010), whereas Beaf32/DREF and/or dCTCF mostly regulate active genes controlling cell proliferation (Emberly et al., 2008; Bushey et al., 2009; Gurudatta et al., 2013). In turn, pre-existing long-range contacts may be

dynamically regulated through recruitment of transcriptional activators, as recently shown (Jin et al., 2013). Long-range contacts may therefore allow crosstalk between these distinct gene regulatory pathways, depending on insulator-prone functional long-range interactions.

EXPERIMENTAL PROCEDURES

RNAi-Mediated Depletions, RTqPCR, Deep-Sequencing, and RNA-Sequencing

Drosophila Schneider S2 cells were treated with specific interfering RNAs to knock down Beaf32, dCTCF, or control RNA as a mock control essentially as described (see Supplemental Experimental Procedures). Stably transfected cell lines expressing synthetic WT/mutant Beaf in the absence of endogenous Beaf were obtained by neomycin selection (see Supplemental Experimental Procedures). RNA-seq analyses were performed using HT-seq and DEG/DE-seq packages to identify DE genes (p value < 0.001) from three independent replicates (see Supplemental Experimental Procedures).

Antibodies, ChIP-Seq and Statistical Analyses

Affinity purified anti-Beaf32, anti-dCTCF, and anti-CP190 antibodies were prepared in the lab (see Supplemental Experimental Procedures). For ChIP, chromatin extracts from S2 cells were prepared by sonication of cells cross-linked with formaldehyde as described (see Supplemental Experimental Procedures). The distribution of peaks was analyzed with respect to TSS after ranking peaks according to density of reads (associated or not with promoter/TSSs). CP190 binding sites and ChIP-indirect and -direct peaks of Beaf32, dCTCF, and GAF were identified as described using ChIP data from S2 cells (see Supplemental Experimental Procedures). ChIP data in transfected cells were analyzed by qPCR and the pyQPCR software (see Supplemental Experimental Procedures).

3C and Hi-C Analyses

3C DNA samples were prepared from Beaf32-depleted, CP190-depleted, or mock-depleted control S2 cells or from stably transfected cell lines expressing synthetic *beaf* genes using TaqMan technology as previously described (Hagège et al., 2007). Chimeras (ligation product) were measured using TaqMan technology after HindIII restriction and proximity-enhanced religation. Anchors were chosen at HindIII sites close to Beaf32 peak in *tsp39D* (2L21342022), between *crc* and *mio* (2L21309216 and 2L21309961). Chimera frequencies were calculated relative to the total DNA amount assessed using three control regions within the same loci (see Supplemental Experimental Procedures). Genome-wide statistical analyses of Hi-C data (Sexton et al., 2012) were performed as described (see Supplemental Experimental Procedures).

ACCESSION NUMBERS

Raw sequencing data are available at the NCBI Sequence Read Archive (accession number: GSE52887).

SUPPLEMENTAL INFORMATION

Supplemental Information includes seven figures, six tables, and Supplemental Experimental Procedures and can be found with this article online at <http://dx.doi.org/10.1016/j.molcel.2013.12.029>.

ACKNOWLEDGMENTS

We thank T. Forné for help and advice regarding 3C; G. Cavalli, T. Sexton, and A. Tanay for Hi-C data and/or for suggestions; D. Gilmour for the generous gift of anti-GAF antibodies; C. Carles for help with web servers; and O.C. lab members including E. Guillou for RNAi and K.Z. lab members for performing Hi-Seq (A. Barski and K. Cui). K.Z.'s lab was supported by the Division of Intramural Research Program of the National Heart, Lung and Blood Institute, NIH.

M.N.'s lab was supported by the European Research Council (Starting Grant 260787 to M.N.). E.E.'s lab was supported by NSERC and the Canadian Institute for Advanced Research (CIFAR). O.C.'s lab was supported by a grant of the ANR, the ARC cancer research funding, and the CNRS-Inserm ATIP-AVENIR program.

Received: February 19, 2013

Revised: August 21, 2013

Accepted: December 27, 2013

Published: January 30, 2014

REFERENCES

- Bartkuhn, M., Straub, T., Herold, M., Herrmann, M., Rathke, C., Saumweber, H., Giffillan, G.D., Becker, P.B., and Renkawitz, R. (2009). Active promoters and insulators are marked by the centrosomal protein 190. *EMBO J.* 28, 877–888.
- Bushey, A.M., Ramos, E., and Corces, V.G. (2009). Three subclasses of a *Drosophila* insulator show distinct and cell type-specific genomic distributions. *Genes Dev.* 23, 1338–1350.
- Chopra, V.S., Cande, J., Hong, J.W., and Levine, M. (2009). Stalled Hox promoters as chromosomal boundaries. *Genes Dev.* 23, 1505–1509.
- Core, L.J., and Lis, J.T. (2008). Transcription regulation through promoter-proximal pausing of RNA polymerase II. *Science* 319, 1791–1792.
- Cuddapah, S., Jothi, R., Schones, D.E., Roh, T.Y., Cui, K., and Zhao, K. (2009). Global analysis of the insulator binding protein CTCF in chromatin barrier regions reveals demarcation of active and repressive domains. *Genome Res.* 19, 24–32.
- Dekker, J. (2006). The three 'C' s of chromosome conformation capture: controls, controls, controls. *Nat. Methods* 3, 17–21.
- 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.
- 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.
- Dostie, J., and Bickmore, W.A. (2012). Chromosome organization in the nucleus - charting new territory across the Hi-Cs. *Curr. Opin. Genet. Dev.* 22, 125–131.
- Emberly, E., Blattes, R., Schuettengruber, B., Hennion, M., Jiang, N., Hart, C.M., Käse, E., and Cuvier, O. (2008). BEAF regulates cell-cycle genes through the controlled deposition of H3K9 methylation marks into its conserved dual-core binding sites. *PLoS Biol.* 6, 2896–2910.
- Eski, C.H., and Fraser, P. (2011). Ultrastructural study of transcription factories in mouse erythroblasts. *J. Cell Sci.* 124, 3676–3683.
- Gerasimova, T.I., Lei, E.P., Bushey, A.M., and Corces, V.G. (2007). Coordinated control of dCTCF and gypsy chromatin insulators in *Drosophila*. *Mol. Cell* 28, 761–772.
- Ghirlando, R., Giles, K., Gowher, H., Xiao, T., Xu, Z., Yao, H., and Felsenfeld, G. (2012). Chromatin domains, insulators, and the regulation of gene expression. *Biochim. Biophys. Acta* 1819, 644–651.
- Gilchrist, D.A., Dos Santos, G., Fargo, D.C., Xie, B., Gao, Y., Li, L., and Adelman, K. (2010). Pausing of RNA polymerase II disrupts DNA-specified nucleosome organization to enable precise gene regulation. *Cell* 143, 540–551.
- Gohl, D., Aoki, T., Blanton, J., Shanower, G., Kappes, G., and Schedl, P. (2011). Mechanism of chromosomal boundary action: roadblock, sink, or loop? *Genetics* 187, 731–748.
- Gurudatta, B.V., Yang, J., Van Bortle, K., Donlin-Asp, P.G., and Corces, V.G. (2013). Dynamic changes in the genomic localization of DNA replication-related element binding factor during the cell cycle. *Cell Cycle* 12, 1605–1615.
- Hagège, H., Klous, P., Braem, C., Splinter, E., Dekker, J., Cathala, G., de Laat, W., and Forné, T. (2007). Quantitative analysis of chromosome conformation capture assays (3C-qPCR). *Nat. Protoc.* 2, 1722–1733.

- Handoko, L., Xu, H., Li, G., Ngan, C.Y., Chew, E., Schnapp, M., Lee, C.W., Ye, C., Ping, J.L., Mulawadi, F., et al. (2011). CTCF-mediated functional chromatin interactome in pluripotent cells. *Nat. Genet.* 43, 630–638.
- Hendrix, D.A., Hong, J.W., Zeitlinger, J., Rokhsar, D.S., and Levine, M.S. (2008). Promoter elements associated with RNA Pol II stalling in the *Drosophila* embryo. *Proc. Natl. Acad. Sci. USA* 105, 7762–7767.
- Hou, C., Li, L., Qin, Z.S., and Corces, V.G. (2012). Gene density, transcription, and insulators contribute to the partition of the *Drosophila* genome into physical domains. *Mol. Cell* 48, 471–484.
- Jee, J., Rozowsky, J., Yip, K.Y., Lochovsky, L., Bjornson, R., Zhong, G., Zhang, Z., Fu, Y., Wang, J., Weng, Z., et al. (2011). ACT: aggregation and correlation toolbox for analyses of genome tracks. *Bioinformatics* 27, 1152–1154.
- Jiang, N., Emberly, E., Cuvier, O., and Hart, C.M. (2009). Genome-wide mapping of boundary element-associated factor (BEAF) binding sites in *Drosophila melanogaster* links BEAF to transcription. *Mol. Cell. Biol.* 29, 3556–3568.
- Jin, F., Li, Y., Dixon, J.R., Selvaraj, S., Ye, Z., Lee, A.Y., Yen, C.A., Schmitt, A.D., Espinoza, C.A., and Ren, B. (2013). A high-resolution map of the three-dimensional chromatin interactome in human cells. *Nature* 503, 290–294.
- Kagey, M.H., Newman, J.J., Bilodeau, S., Zhan, Y., Orlando, D.A., van Berkum, N.L., Ebmeier, C.C., Goossens, J., Rahl, P.B., Levine, S.S., et al. (2010). Mediator and cohesin connect gene expression and chromatin architecture. *Nature* 467, 430–435.
- Lee, C., Li, X., Hechmer, A., Eisen, M., Biggin, M.D., Venters, B.J., Jiang, C., Li, J., Pugh, B.F., and Gilmour, D.S. (2008). NELF and GAGA factor are linked to promoter-proximal pausing at many genes in *Drosophila*. *Mol. Cell. Biol.* 28, 3290–3300.
- Li, H.B., Müller, M., Bahechar, I.A., Kyrchanova, O., Ohno, K., Georgiev, P., and Pirrotta, V. (2011). Insulators, not Polycomb response elements, are required for long-range interactions between Polycomb targets in *Drosophila melanogaster*. *Mol. Cell. Biol.* 31, 616–625.
- Lieberman-Aiden, E., van Berkum, N.L., Williams, L., Imakaev, M., Ragoczy, T., Telling, A., Amit, I., Lajoie, B.R., Sabo, P.J., Dorschner, M.O., et al. (2009). Comprehensive mapping of long-range interactions reveals folding principles of the human genome. *Science* 326, 289–293.
- Misteli, T. (2007). Beyond the sequence: cellular organization of genome function. *Cell* 128, 787–800.
- Nègre, N., Brown, C.D., Shah, P.K., Kheradpour, P., Morrison, C.A., Henikoff, J.G., Feng, X., Ahmad, K., Russell, S., White, R.A., et al. (2010). A comprehensive map of insulator elements for the *Drosophila* genome. *PLoS Genet.* 6, e1000814.
- 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.
- Phillips-Cremins, J.E., and Corces, V.G. (2013). Chromatin insulators: linking genome organization to cellular function. *Mol. Cell* 50, 461–474.
- Phillips-Cremins, J.E., Sauria, M.E., Sanyal, A., Gerasimova, T.I., Lajoie, B.R., Bell, J.S., 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.
- Sanyal, A., Baù, D., Martí-Renom, M.A., and Dekker, J. (2011). Chromatin globules: a common motif of higher order chromosome structure? *Curr. Opin. Cell Biol.* 23, 325–331.
- Sanyal, A., Lajoie, B.R., Jain, G., and Dekker, J. (2012). The long-range interaction landscape of gene promoters. *Nature* 489, 109–113.
- Schwartz, Y.B., Linder-Basso, D., Kharchenko, P.V., Tolstorukov, M.Y., Kim, M., Li, H.B., Gorchakov, A.A., Minoda, A., Shanower, G., Alekseyenko, A.A., et al. (2012). Nature and function of insulator protein binding sites in the *Drosophila* genome. *Genome Res.* 22, 2188–2198.
- Sexton, T., Yaffe, E., Kenigsberg, E., Bantignies, F., Leblanc, B., Hoichman, M., Parrinello, H., Tanay, A., and Cavalli, G. (2012). Three-dimensional folding and functional organization principles of the *Drosophila* genome. *Cell* 148, 458–472.
- Tanay, A., and Cavalli, G. (2013). Chromosomal domains: epigenetic contexts and functional implications of genomic compartmentalization. *Curr. Opin. Genet. Dev.* 23, 197–203.
- Vernimmen, D., De Gobbi, M., Sloane-Stanley, J.A., Wood, W.G., and Higgs, D.R. (2007). Long-range chromosomal interactions regulate the timing of the transition between poised and active gene expression. *EMBO J.* 26, 2041–2051.
- Vogelmann, J., Valeri, A., Guillou, E., Cuvier, O., and Nollmann, M. (2011). Roles of chromatin insulator proteins in higher-order chromatin organization and transcription regulation. *Nucleus* 2, 358–369.
- Wood, A.M., Van Bortle, K., Ramos, E., Takenaka, N., Rohrbaugh, M., Jones, B.C., Jones, K.C., and Corces, V.G. (2011). Regulation of chromatin organization and inducible gene expression by a *Drosophila* insulator. *Mol. Cell* 44, 29–38.
- Xiao, T., Wallace, J., and Felsenfeld, G. (2011). Specific sites in the C terminus of CTCF interact with the SA2 subunit of the cohesin complex and are required for cohesin-dependent insulation activity. *Mol. Cell. Biol.* 31, 2174–2183.