Published in final edited form as:

Mol Cell. 2017 November 02; 68(3): 491–503.e5. doi:10.1016/j.molcel.2017.09.031.

CBP regulates recruitment and release of promoter-proximal RNA polymerase II

Ann Boija[1,7], **Dig Bijay Mahat**[2,7], **Aman Zare**[3,4], **Per-Henrik Holmqvist**[1], **Philge Philip**[3,4], **David J. Meyers**[5], **Philip A. Cole**[5], **John T. Lis**[2,8], **Per Stenberg**[3,4,6,8], and **Mattias Mannervik**[1,8,9]

[1]Dept. of Molecular Biosciences, the Wenner-Gren Institute, Stockholm University, SE-10691 Stockholm, Sweden

[2]Dept. of Molecular Biology and Genetics, Cornell University, Ithaca, New York 14850, USA

[3]Dept. of Molecular Biology, Umeå University, SE-901 87 Umeå, Sweden

[4]Computational Life Science Cluster (CLiC), Umeå University, SE-901 87 Umeå, Sweden

[5]Dept. Pharmacology and Molecular Sciences, The Johns Hopkins University School of Medicine, 725 North Wolfe Street, Baltimore, MD 21205, USA

[6]Division of CBRN Defence and Security, FOI, Swedish Defence Research Agency, SE-906 21 Umeå, Sweden

Summary

Transcription activation involves RNA polymerase II (Pol II) recruitment and release from the promoter into productive elongation, but how specific chromatin regulators control these steps is unclear. Here we identify a novel activity of the histone acetyltransferase p300/CBP in regulating promoter-proximal paused Pol II. We find that *Drosophila* CBP inhibition results in “dribbling” of Pol II from the pause site to positions further downstream, but impedes transcription through the +1 nucleosome genome-wide. Promoters strongly occupied by CBP and GAGA-factor have high levels of paused Pol II, a unique chromatin signature and are highly expressed regardless of cell type. Interestingly, CBP activity is rate-limiting for Pol II recruitment to these highly-paused promoters through an interaction with TFIIB, but for transit into elongation by histone acetylation at other genes. Thus, CBP directly stimulates both Pol II recruitment and the ability to traverse the first nucleosome, thereby promoting transcription of most genes.

[7]These authors contributed equally.
[8]Senior authors: Mattias Mannervik, mattias.mannervik@.su.se.
Per Stenberg, per.stenberg@umu.se.
John T. Lis, johnlis@cornell.edu
[9]Lead contact

Author contributions

A.B., P.S. and M.M. designed the study. A.B. performed ChIP-qPCR, ChIP-seq, MNase-seq, and RNAi experiments, D.B.M. performed PRO-seq, and P.H.H. performed co-IP. A.Z., D.B.M., A.B. and P.P. analyzed data and together with J.T.L, P.S. and M.M interpreted the results. A.Z. prepared Figures with help from A.B. and D.B.M. D.J.M. and P.A.C. synthesized inhibitors. A.B., P.S. and M.M. wrote the manuscript, and D.B.M, A.Z. and J.T.L. helped with editing of the manuscript.

Introduction

Regulation of transcription occurs at two major steps in metazoans (reviewed by Core and Lis, 2008). The first step is the recruitment of RNA Polymerase II (Pol II) to the promoter, which is orchestrated by sequence-specific transcription factors that enable formation of a pre-initiation complex (PIC), consisting of Pol II and general transcription factors (GTFs). A successful formation of the PIC leads to rapid initiation and transcription of 20 to 60 nucleotides (Core et al., 2012). This brings Pol II to the second step of regulation where it pauses on many genes, an event mediated by the action of NELF and DSIF. The release of paused Pol II into productive elongation requires recruitment of P-TEFb kinase, which phosphorylates these pausing factors and the Pol II C-terminal domain (CTD) (reviewed in Zhou et al., 2012). Although pausing temporarily restrains Pol II from entering elongation, a majority of paused genes are highly expressed (reviewed in Adelman and Lis, 2012). In most instances, pausing may therefore be a rheostat-like mechanism to maintain and tune expression rather than to function as an on-off switch.

GAGA-factor (GAF, also known as Trithorax-like, Trl) is a sequence-specific transcription factor that is associated with 20% of expressed genes in *Drosophila,* and appears to have a role in both regulated steps (Li et al., 2013). GAF affects the first step by recruiting chromatin remodelers (Shimojima et al., 2003; Xiao et al., 2001) leading to local removal of nucleosomes (Fuda et al., 2015), and it promotes PIC formation by interacting with TFIID (Chopra et al., 2008). GAF also recruits NELF, which helps to establish Pol II pausing (Li et al., 2013).

Our understanding of how specific chromatin regulators are recruited to promoters and how they regulate the two major steps in transcription is incomplete. CREB-binding protein (CBP) and its paralog p300 are widely used transcriptional co-regulators with over 400 interaction partners (reviewed in Bedford et al., 2010). p300/CBP has histone acetyltransferase (HAT) activity and is known to acetylate lysine 18 and 27 of histone H3 and lysine 8 of histone H4, establishing an active chromatin state (Feller et al., 2015; Jin et al., 2011; Tie et al., 2009). p300/CBP is well known for its role at enhancers, and genome-wide mapping of CBP occupancy has been used to find novel enhancers both in mice and flies (Negre et al., 2011; Visel et al., 2009). However, p300/CBP occupies also other regulatory elements, including promoters (Philip et al., 2015; Wang et al., 2009). The mechanisms by which p300/CBP regulates transcription are not fully understood and a potential direct role for p300/CBP at promoters is a key unanswered question.

In *Drosophila*, there is only one homolog of p300 and CBP (known as *nejire)*, referred to as CBP in this manuscript. As in other organisms, a majority of *Drosophila* CBP peaks are found at enhancers (Philip et al., 2015). Herein, we analyzed the functions of this sole p300/CBP homolog at gene promoters in *Drosophila* S2 cells. We discovered that promoters co-occupied by CBP and GAF identify the most highly paused genes in the genome, that these promoters are enriched for a unique set of chromatin factors and modifications, and that transcription of these genes are the most CBP dependent. CBP is required for maintaining Pol II at the pause site, and facilitates efficient release of paused Pol II into productive elongation by overcoming the transcriptional barrier caused by the +1

nucleosome. CBP is also important for efficient Pol II recruitment to promoters by interacting with TFIIB. Our data suggest that CBP activity is rate-limiting for Pol II recruitment to promoters co-occupied by CBP and GAF, but for transit into elongation by histone acetylation at other genes. Thus, we identify a novel regulatory role for CBP at promoters.

Results

CBP occupies the promoters of virtually all expressed genes in S2 cells

To investigate the function of CBP in transcriptional control, we reanalyzed our previously described CBP ChIP-seq of *Drosophila* S2 cell chromatin (Philip et al., 2015) and determined the presence of CBP at expressed gene promoters (5400 genes based on RNA-seq, as defined in Cherbas et al., 2011). We plotted CBP occupancy and compared it to GAF (Fuda et al., 2015) and Pol II (Roy et al., 2010) data (Fig. 1A). This showed that CBP occupancy is higher than background (genomic mean) at virtually all promoters of expressed genes. Consistent with this notion, we found that the mean CBP occupancy was significantly higher at expressed than at non-expressed genes (Fig. 1B). It is possible that binding in the open chromatin of expressed promoters represent false positive “phantom peaks” described previously (Jain et al., 2015). However, there is a high concordance in the magnitude by which a CBP inhibitor affects transcription and the level of CBP occupancy (described in later sections, Fig. 3B), strongly suggesting functional CBP binding to most expressed promoters. We conclude that CBP can be found at the promoters of almost all genes expressed in S2 cells.

Strong CBP and GAF occupancy identifies a set of highly-paused promoters

To find factors that may be involved in recruiting CBP to promoters, we searched for sequence motifs in a +/− 100 bp window centered on the CBP summit at expressed promoters. We found that a $(GA)_n$ repeat is the most enriched motif in these sequences (E-value of 3.8e-038, Fig. S1A), and that its mean distance to the TSS is 177 bp upstream. We therefore plotted the promoter occupancy of the GAGA-motif binding factor GAF (Adkins et al., 2006), sorted on CBP binding strength. This shows that at promoters with the highest CBP enrichment, both CBP and GAF are present, whereas promoters with less CBP occupancy in most cases have very little or no GAF (Fig. 1A). Although we could not detect a direct interaction between GAF and CBP, we found that CBP occupancy was reduced at several loci upon GAF knock-down (Fig. S1B). Because GAF is known to be required to generate nucleosome depleted regions (Fuda et al., 2015) through its interaction with nucleosome remodelers like NURF (Xiao et al., 2001), GAF may be facilitating CBP occupancy indirectly through its opening of chromatin.

Whereas GAF is strongly associated with Pol II pausing (Duarte et al., 2016; Lee et al., 2008), a function for CBP in pausing has not been previously described. To examine the relation between CBP, GAF and pausing, we ranked the promoters of all 5400 expressed genes by the amount of CBP or GAF at the promoter (to account for differences in ChIP-seq enrichment), and plotted the CBP and GAF rank against the pausing index of the corresponding promoter (Fig. 1C). The pausing index (data from Kwak et al., 2013) is the

ratio of Pol II occupancy at the promoter versus the gene-body. This shows that there is a strong correlation between pausing, GAF, and CBP occupancy (Fig 1C). Because of this correlation, we sorted all expressed genes according to the combined CBP and GAF promoter rank, and divided them into 20 equally sized bins (Table S1). Bin 1 has the lowest enrichment of CBP + GAF and bin 20 the highest. We then plotted the average pausing index of these 20 bins (Fig. 1D). Bins 19 and 20 contain genes with significantly higher pausing index than genes in the other bins (Fig. 1D). We will refer to the 540 promoters in these bins as high CBP and GAF (HCG) promoters. Promoters in bin 15–18 are medium-paused and have medium CBP and GAF enrichment (MCG), whereas promoters in bins 1–14 are lowly-paused and have the least CBP and GAF (LCG, Fig. 1D and Fig. S1C). We performed ChIP-seq with another antibody raised against a different part of CBP (Lilja et al., 2007), which confirmed that HCG promoters are associated with the greatest amount of CBP (Fig. S1C).

Although the HCG class (labeled red in Fig. 1C and D) has the highest mean pausing index, not all highly-paused genes are found in this class. Genes with a high pausing index but low amounts of GAF and/or CBP are found in the other promoter bins. However, only 6 promoters with a high GAF rank have a high pausing index but are in the LCG class due to a low CBP rank (Fig. S1D). Thus, most paused, GAF-bound promoters also contain plenty of CBP. There are also about 430 genes that are paused but not expressed in S2 cells, but in this study we only examine expressed genes.

Together, our analyses show that promoters from genes bound by both CBP and GAF are highly paused, and indicate that CBP contributes to promoter-proximal pausing.

High CBP and GAF promoters contain more PIC components than other promoters

We expected promoters of highly-paused genes to contain more pause-related factors and polymerase than other gene promoters. We therefore plotted the average occupancy of the pause-inducing factor NELF (data from Gilchrist et al., 2010) and Pol II (data from Roy et al., 2010) at LCG, MCG and HCG promoters (+/− 500 bp around the TSS, Fig. 2A). This showed that HCG promoters contain more of both NELF and Pol II than promoters of other expressed genes, as expected (2.8 fold more NELF and 1.7 fold more Pol II at HCG vs LCG promoters, Fig. 2A). More Pol II and more pausing at these promoters relative to other promoters could be caused by inefficient release from pausing, more efficient Pol II recruitment, or a combination of both. We therefore plotted mean occupancy of the GTF TFIIA, which is involved in Pol II recruitment (data from Gilchrist et al., 2010), and found that the HCG promoters contain 2.0 fold higher levels of TFIIA than LCG genes (Fig. 2A). This result argues that efficient Pol II recruitment contributes to high pausing and Pol II occupancy. Our CBP inhibitor studies described below are consistent with efficient recruitment of Pol II to HCG promoters.

To confirm a higher concentration of PIC components at HCG promoters, we selected seven genes with similar levels of expression in S2 cells for further analysis. The genes *flw, Gαq,* and *CG14995* belong to the LCG class, *BTBD9* to MCG, whereas *Mrp4* and *rho* are in the HCG class. We compared these genes to the most extensively characterized paused gene, the *Drosophila hsp70* heat-shock gene, which is masked from our genome-wide dataset due to

its repetitive nature. ChIP-qPCR showed that HCG promoters (*Mrp4, rho)* and *hsp70* contain significantly more Pol II (10 fold), CBP (2.4 fold), GAF (10 fold) as well as the GTFs TBP (2.4 fold), TFIIA (2.6 fold), and TFIIB (3.2 fold) than the four other promoters ($p<0.05$, Fig. 2B, C and Fig. S2A). Together, these results show that HCG promoters strongly occupied by CBP and GAF contain substantially more of the general transcriptional machinery than promoters of other expressed genes. This indicates that Pol II is efficiently recruited to these promoters, and that the forces that act to recruit Pol II might also conspire to retain early elongating Pol II to the proximal pause region (Kwak et al., 2013).

High CBP and GAF promoters are enriched in specific chromatin factors and histone modifications

We extended our analysis in order to identify other factors and chromatin modifications enriched at HCG promoters. We plotted 40 previously mapped factors in S2 cells (data from Enderle et al., 2011; Roy et al., 2010) and found that Trithorax, Ash1 and Mi-2 are also enriched at HCG promoters (Fig. S2B), whereas the insulator proteins BEAF-32 and CP190 and the chromodomain protein MRG15 are depleted compared to other promoters (Fig. S2C).

Next, we plotted the mean levels of 25 previously mapped histone modifications (Roy et al., 2010) and found that HCG promoters are enriched in marks associated with enhancers, H3K4 monomethylation (H3K4me1, 2.5 fold), H3K18 and H3K27 acetylation (H3K18ac, 2.0 fold and H3K27ac, 1.9 fold), but depleted in H3K4me3, H3K9ac, and H4K16ac (1.5, 1.1, and 1.8 fold) relative to LCG promoters (Fig. 2D and E). Furthermore, the chromatin state of HCG promoters differs from most other promoters of expressed genes (using the 30 chromatin states model defined in the modENCODE project (Roy et al., 2010), Fig. S2D). Many HCG promoters are found in open chromatin (Fig. S2D), and are more DNAseI hypersensitive than promoters of other expressed genes (Fig. S2E). An example of the difference between HCG promoters and other promoters can be seen by comparing the *rho* locus (HCG) to *Gaq* (LCG, Fig. 2F). The *rho* locus is associated with more Pol II, TFIIA and H3K27ac, but less H4K16ac and H3K4me3, than *Gaq.*

We wanted to know if the characteristics of HCG promoters is mainly due to the presence of CBP or to GAF, and therefore plotted enrichment of all other factors and histone modifications against CBP and GAF occupancy levels separately (Fig. S3). However, these individual profiles are very similar to the ones obtained with the combined CBP+GAF rank, thereby preventing us from separating the role of CBP and GAF in the unique features of HCG promoters. Taken together, our results show that promoters of some highly-paused genes are strongly occupied by both CBP and GAF, and further enriched and depleted in specific chromatin factors and histone modifications relative to promoters of other expressed genes.

High CBP and GAF promoters are largely shared among cell types and strongly expressed during development

Genes with these HCG promoters are among the most highly expressed in S2 cells, despite their high pausing (Fig. S2F). Examination of gene expression at different developmental

stages showed that mRNAs from HCG genes are not the most abundant in early embryos (Fig. 2G). However, at later stages of embryo development, expression of other genes decreases whereas HCG genes remain strongly expressed, so that in late embryos, larvae, pupae and adults, they are the most highly expressed (Fig. 2G). Gene ontology analysis of the HCG class showed that they are enriched for the developmental processes imaginal disc development, epithelium development and gamete generation (Fig. S2G and Table S2). Indeed, many developmental transcription factors and signaling components, but also several house-keeping and metabolic genes are found within this class (Table S1). This combination of genes may explain their high mean expression throughout development.

Since many of the HCG genes are strongly expressed at different developmental stages, some of their features may be found in other cell types. We therefore used ChIP-seq data from early embryos to rank promoters according to CBP and GAF occupancy (Holmqvist et al., 2012; Roy et al., 2010). Whereas S2 cells are macrophage-like immune cells derived from late embryos (Schneider, 1972), early embryos consist of multipotent cells that are yet to enter terminal differentiation. We plotted the CBP+GAF ranking in S2 cells versus the ranking in early embryos and found that most of the promoters from the HCG class in S2 cells are also the most strongly occupied by CBP+GAF in embryos (Fig. 2H). This suggests that HCG promoters are largely shared between these very different cell types. Indeed, 318 of the 540 HCG promoters in S2 cells are among the 540 top CBP+GAF ranked promoters in early embryos (Fig. S2H).

Taken together, our results indicate that HCG genes are highly accessible for transcription in different cell types, allowing them to maintain a strong expression and be tuned by various developmental and metabolic signals.

CBP is required for transcription of thousands of genes and impacts High CBP and GAF genes most strongly

To investigate CBP's functions in transcription, we used a previously described inhibitor, C646, that selectively inhibits the catalytic activity of CBP (Bowers et al., 2010). To account for potential off-target effects of C646 we used C37 as a control, a compound very similar in structure to C646 but that shows no effect on CBP HAT activity (Bowers et al., 2010). C646 and related inhibitors have previously been shown to affect histone acetylation in *Drosophila* and mammalian cells within minutes (Crump et al., 2011; Dancy et al., 2012). To assess the immediate transcriptional response to CBP inhibition genome-wide, we performed precision run-on sequencing (PRO-seq) with S2 cells treated with C37 (control) or C646 for 10 min. This technique maps transcriptionally engaged Pol II with single nucleotide resolution (Kwak et al., 2013). The two biological replicates of each treatment showed strong correlation (Fig. S4). After normalization, we quantified the change in PRO-seq density upon CBP inhibition at all genes. We found 3790 genes with significant ($p<0.001$) down-regulation but only 42 genes with significant upregulation in gene-body PRO-seq density in C646-treated cells (Fig. 3A). However, all 42 up-regulated genes appear to represent false-positives, because the PRO-seq density increase occurs only in the promoter-proximal region. These reads were attributed to the gene-body because the genes are very short (Fig.

S4). We therefore conclude that the only direct effect of CBP inhibition is transcription down-regulation.

To determine if transcription down-regulation upon CBP inhibition is related to CBP levels, we plotted the change in gene-body PRO-seq density versus CBP enrichment (Fig. 3B). This shows an impressive correlation between the level of CBP occupancy at the promoter and the effect of CBP inhibition on transcription, arguing that C646 is indeed specific to CBP and that CBP promoter occupancy is functionally relevant to transcription. Note that even genes with the least CBP at their promoters are down-regulated upon CBP inhibition, which is consistent with occupancy of CBP at virtually all expressed genes (Fig. 1A and B). We then compared the PRO-seq change in LCG, MCG and HCG genes, and found that HCG genes are most strongly down-regulated after CBP inhibition (Fig. 3C). In summary, a large fraction of genes expressed in S2 cells require CBP for normal transcription, and those with strongest CBP promoter enrichment are the most CBP-dependent.

CBP inhibition reduces Pol II occupancy at High CBP and GAF promoters

Next, we wanted to investigate the step or steps during transcription where CBP plays a key role and assess if this role differs between HCG and other promoters. We started by measuring CBP and Pol II occupancy at promoters after CBP inhibition by ChIP-qPCR. We found that CBP occupancy is affected as early as 1 min after CBP inhibition at 6 out of 7 tested promoters, and is further decreased as the CBP inhibition progressed ($p<0.05$, Fig. 3D). Using an antibody against the Rpb3 subunit of Pol II, we observed that LCG and MCG promoters showed an increase in total Pol II occupancy at 10 and 60 min after CBP inhibition, followed by a decrease at 4 hours of treatment (Fig. 3E). By contrast, Pol II occupancy decreased much earlier after CBP inhibition at HCG promoters (*Mrp4, rho and hsp70*).

To confirm that these effects are specific to CBP, we knocked down CBP levels by RNAi and used another CBP inhibitor. Treating cells for 60 min with curcumin (that also inhibits CBP activity) resulted in a strong decrease of Pol II occupancy at HCG promoters, whereas LCG and MCG promoters were less strongly affected (Fig. S5A). Similarly, CBP RNAi had a significantly stronger effect on Pol II occupancy at HCG promoters than at promoters of LCG and MCG genes, whereas occupancy of CBP itself was reduced to a comparable extent at all tested promoters (Fig. S5B). These observations indicate that promoters differ in their response to CBP inhibition and that Pol II promoter occupancy at HCG genes is particularly CBP-dependent.

CBP is essential for the release of paused Pol II into productive elongation

The difference in Pol II occupancy between HCG and other promoters upon CBP inhibition at individual genes prompted us to look at Pol II occupancy at promoters genome-wide. We measured the PRO-seq signal in the promoter-proximal region (from TSS to 300 bp downstream) of all genes that show significant down-regulation in gene-body (n=3790). This identified two different responses to CBP inhibition - an increase in promoter-proximal Pol II at 2529 genes and a decrease at 1243 genes (Fig. 4A). Next, we plotted the PRO-seq fold change in this promoter-proximal region at LCG, MCG and HCG genes. Only HCG genes

showed decreased promoter-proximal Pol II after CBP inhibition, whereas LCG and MCG genes showed increased Pol II in this region (Fig. 4B). This result is in accord with the Pol II ChIP-qPCR shown in Fig. 3E. Taken together, both PRO-seq and ChIP-qPCR show that promoters of HCG genes respond differently than promoters of other expressed genes by failing to maintain high levels of promoter-proximal Pol II after CBP inhibition.

The increase in promoter-proximal Pol II upon CBP inhibition at LCG and MCG genes occurs mainly in the first 100 bp downstream of the TSS (Fig. 4C). Since all genes included in Fig. 4 are down-regulated in the gene body, the increase in promoter-proximal signal suggests that CBP has a function in releasing Pol II into productive elongation. At HCG genes (Fig. 4D), Pol II levels are decreased both in the promoter-proximal region and in the gene body, indicating that CBP has a crucial role in recruiting Pol II to these promoters.

We then compared the change in pausing index (calculated from the PRO-seq data) after CBP inhibition at these two classes of promoters. Genes with increased promoter-proximal Pol II are weakly paused in control cells, but more strongly paused after CBP inhibition as expected (Fig. 4E left). By contrast, genes with decreased PRO-seq reads in the promoter-proximal region have a high pausing index in control cells (Fig. 4E right). Surprisingly, despite reduced levels of promoter-proximal Pol II, the pausing index of these genes also increased upon CBP inhibition (Fig. 4E right). This indicates that the decreased PRO-seq signal in the gene body is of greater magnitude than in the promoter-proximal region, and that CBP is essential for the release of paused Pol II also at these genes. Thus, CBP may function to promote release of Pol II into productive elongation at all genes.

CBP regulates Pol II at High CBP and GAF promoters by controlling TFIIB promoter occupancy

Since CBP inhibition caused a reduction in Pol II occupancy at HCG promoters after 10 min of treatment, we performed a time-course study of Pol II occupancy after CBP inhibition using ChIP-qPCR. We used an antibody recognizing Pol II phosphorylated at Ser5 in the C-terminal domain (S5P), the form of Pol II that has initiated transcription. A statistically significant decrease in S5P Pol II occupancy was detected at HCG promoters after CBP inhibition, starting with a modest decline of approximately 20% at 3 min, 40% decline at 7 min and 80% decline after 60 min ($p<0.05$, Fig. 5A). The immediate effects of the inhibitor argue for a direct role of CBP activity on paused Pol II occupancy. Another CBP inhibitor, curcumin, showed a similar effect (Fig. S6A). GAF occupancy was not decreased after 10 min of CBP inhibitor treatment (Fig. 5B), indicating that reduced Pol II occupancy at HCG promoters upon CBP inhibition occurs while GAF is still bound. Interestingly, although TFIIB and TFIIF occupancy was reduced to approximately 50% after 10 min of CBP inhibition, TATA-binding protein (TBP) and TFIIA remained associated with HCG promoters (Fig. 5C). This indicates that CBP regulates Pol II promoter association at a post-TBP step that may involve an interaction with TFIIB. Consistent with this hypothesis, we found that TFIIB and CBP can be co-immunoprecipitated from untreated cells and that they remain associated after CBP inhibition (Fig. 5D). A CBP-TFIIB interaction in the presence of the inhibitor should result in a global reduction in TFIIB occupancy, since CBP binding to chromatin is strongly diminished upon inhibition (Fig. 3D). We performed TFIIB ChIP-seq

in cells treated with C37 control or C646 inhibitor for 10 min to test this hypothesis. Indeed, we found that TFIIB occupancy is dramatically reduced genome-wide after 10 min of CBP inhibition (Fig. 5E). The two biological replicates of each treatment were highly similar (Fig. S6B). Furthermore, the stronger CBP enrichment at the promoter, the more TFIIB was found in control-treated cells, but not after CBP inhibition (Fig. S6C). Based on these results, we propose that CBP recruits or stabilizes Pol II at promoters by regulating TFIIB occupancy, possibly through a direct interaction.

PRO-seq reveals a Pol II gatekeeper function for CBP

CBP inhibition not only affects the amount of Pol II at the promoter-proximal pause region but also the positioning of paused Pol II. A metagene profile of PRO-seq reads over all expressed promoter-proximal regions showed decreased amounts of Pol II over the canonical pause site, but “dribbling” of Pol II into positions downstream of the pause site after CBP inhibition (Fig. 6A). This dribbling of Pol II beyond the promoter-proximal pause site upon CBP-inhibition is readily apparent at the single gene level (Fig. 6B). To examine the prevalence of this effect, we generated a heatmap of the change in PRO-seq reads in 5 bp windows from the TSS to 600 bp downstream for all down-regulated genes, ranked by their level of down-regulation (Fig. 6C). This shows a decrease in PRO-seq signal over the canonical pause site (0–60bp downstream of the TSS) at virtually all of the genes, suggesting that CBP could be important for Pol II recruitment to most expressed genes and not only to HCG promoters. Importantly, essentially all genes also exhibit Pol II dribbling into the +60–180 bp region (increased PRO-seq signal, Fig. 6C). Since this effect is observed already by 10 min of inhibitor treatment, it appears to be a direct effect of CBP inhibition. We conclude that CBP is involved in maintaining Pol II at the canonical pause-site genome-wide.

CBP facilitates Pol II elongation by overcoming the +1 nucleosome barrier

In order to examine the role of nucleosomes in Pol II dribbling after CBP inhibition, we performed MNase-seq and compared the PRO-seq profiles to a nucleosome density map after CBP inhibition (Fig. 6D). Interestingly, the increased PRO-seq density correlates well with the position of the +1 nucleosome, indicating that Pol II dribbles from the pause site but has difficulty overcoming the transcriptional barrier caused by the first nucleosome when CBP is inhibited (compare Fig. 6C with D). We then plotted the PRO-seq ratio in CBP inhibited versus control cells as a function of distance from the average profile of nucleosome position in inhibitor-treated cells (Fig. 6E). This shows that the PRO-seq density in CBP inhibited cells is lower than in control cells in the pause region (around 100–200 bp upstream of the +1 nucleosome), but then increases until it reaches a maximum at the dyad axis of the +1 nucleosome, from where it decreases again (Fig. 6E). These results indicate that Pol II is released from the canonical pause site, but traverses the +1 nucleosome more slowly when CBP is inhibited.

Interestingly, the positioning of the nucleosomes did not change between control and inhibitor-treated cells, although occupancy was somewhat reduced for nucleosomes close to the TSS (Fig. 6F). A closer look at the MNase-seq change shows a reciprocal relationship between Pol II and nucleosomes (Fig. S7A). Nucleosome density was decreased where the

PRO-seq signal is stronger than in control, but slightly increased upstream and downstream of the peak of dribbled Pol II. This indicates that nucleosomes fill in the position vacated by paused Pol II, as previously suggested (Gilchrist et al., 2010; Gilchrist et al., 2008).

Our data shows that chromatin is not more compacted and that nucleosome position does not change after CBP inhibition, but that Pol II has difficulty traversing the +1 nucleosome. Since CBP is a HAT that targets H4K8, H3K18 and H3K27 (Feller et al., 2015; Jin et al., 2011; Tie et al., 2009), we investigated if the inhibition of CBP causes a reduction in acetylation of the +1 nucleosome. The level of H3K27ac in the promoter-proximal region, normalized for total H3 occupancy, was decreased upon CBP inhibitor treatment at both HCG and at other promoters (Fig. 6G). There was only a small effect on H3K27ac in the first minutes of treatment, but a substantial reduction after 10 min (Fig. 6G). These results show that the effects on histone acetylation occur concurrently with accumulation of Pol II upstream of the +1 nucleosome, and are consistent with a function for CBP-mediated histone acetylation in facilitating Pol II release into transcription elongation.

Taken together, our results reveal that CBP is required genome-wide both for recruitment of Pol II and for allowing Pol II to overcome the transcriptional barrier caused by the +1 nucleosome that resides downstream of the pause region. Upon CBP inhibition, release from Pol II pausing is unaffected but transcription through the +1 nucleosome is retarded, leading to dribbling of Pol II from the canonical pause site to some 50–100bp further downstream.

Discussion

Elucidating transcriptional regulatory mechanisms is of fundamental importance for understanding cell behavior in both normal and pathological conditions. In metazoan genomes, pausing of Pol II at a position close to the transcription start site is prevalent, occurring on a majority of active genes in a given cell type (Core et al., 2012). How Pol II pausing is regulated is not fully understood. Here we show that only 10 min of CBP inhibition results in dribbling of Pol II from the canonical pause site to positions further downstream. We propose that this effect can be explained by the combination of two different CBP activities, recruitment of Pol II to the promoter and facilitation of transcription beyond the first nucleosome (Fig. 7). In this scenario, CBP inhibition causes less Pol II recruitment, resulting in diminished Pol II amounts at the canonical pause site. The Pol II that has initiated transcription is released normally from the canonical pause, but accumulates upstream of the +1 nucleosome. Due to different rate-limiting steps at different types of promoters, the final outcome of CBP inhibition is either a decrease or an accumulation of Pol II in the promoter-proximal region (Fig. 7).

At HCG genes, CBP is critical for Pol II recruitment. A significant decrease in Pol II occupancy at HCG promoters can be observed already by 3 min whereas H3K27ac levels are not significantly different to that of control cells until 10 min of CBP inhibition. Thus, CBP has a direct effect on Pol II occupancy at HCG promoters that is at least partly independent of H3K27ac. This may involve an interaction with TFIIB. Although TFIIB can be acetylated, which stabilizes the interaction with TFIIF (Choi et al., 2003), we were unable to detect CBP-mediated TFIIB acetylation in S2 cells. However, the C-terminal zinc-finger

region in CBP is known to directly interact with TFIIB *in vitro* (Kwok et al., 1994), and our results support a model wherein CBP recruits TFIIB to promoters, possibly by this direct interaction. By recruiting TFIIB to promoters, CBP likely stimulates PIC assembly and Pol II recruitment, consistent with a highly dynamic TFIIB-promoter interaction (Zhang et al., 2016). Despite reduced Pol II levels in the promoter-proximal region after CBP inhibition, pausing is not completely eliminated and the pausing index increases, not decreases, indicating that CBP also modulates promoter escape. Thus, CBP promotes both recruitment and release of Pol II from these promoters.

At promoters of LCG and MCG genes, paused Pol II dribbles to more downstream positions, accumulates upstream of the +1 nucleosome, and fails to effectively enter productive elongation, causing an increase in total Pol II occupancy around the pause region upon CBP inhibition. Thus, CBP is rate-limiting for transcription of these genes at a post-recruitment step, likely by acetylation of the +1 nucleosome. This is consistent with single-cell analyses of transcription induction of a tandem gene array in mouse cells, where H3K27ac and p300/CBP were shown to be required for accumulation of elongating S2P Pol II (Stasevich et al., 2014). Together, these results demonstrate a critical function for CBP in efficient release of Pol II from the promoter-proximal pause into productive elongation.

Another explanation for the altered positioning of Pol II could be that CBP acts more directly to maintain Pol II at the pause site upstream of the first nucleosome, which may in turn determine how effectively a gene is activated. Dribbling of paused Pol II after CBP inhibition would shift its position into regions that are no longer efficiently targeted by activating mechanisms. The dribbled Pol II would then be slowed by the first, hypo-acetylated, nucleosome. In this scenario, CBP would stabilize or capture Pol II at the pause site rather than stimulate recruitment of Pol II to the promoter. This activity could also involve an interaction with TFIIB, and/or CBP-mediated acetylation of another protein involved in pausing. In mammalian cells, CBP can acetylate many proteins including the non-consensus lysines in the Pol II CTD (Schroder et al., 2013). We found that the *Drosophila* Pol II CTD is also acetylated in a CBP-dependent manner (Fig. S7). Another possibility is that CBP targets a pausing factor (e.g. NELF or DSIF) or core promoter factors.

C646 selectively inhibits p300/CBP over other acetyltransferases (Bowers et al., 2010). Although it is a competitive inhibitor that does not covalently modify p300/CBP (Bowers et al., 2010), covalent off-target reactivity with abundant cellular proteins has also been described (Shrimp et al., 2016). However, we find it highly unlikely that our observed effects on Pol II originate from merely off-target effects of C646. The short exposure (10 min) of the treated cells, the use of the similarly reactive compound C37 as a control, and the significant correlation between CBP occupancy and the effects of C646 (Figs 3–6) argue that CBP inhibition accounts for the transcriptional changes observed. Although C646 inhibits CBP's catalytic activity, we discovered that occupancy of CBP itself is rapidly affected by the CBP inhibitor, consistent with *in vitro* data showing that the curcumin inhibitor causes a conformational change that dissociates p300/CBP from a chromatin template (Black et al., 2006). We cannot therefore distinguish between CBP-mediated acetylation and a non-enzymatic CBP function in transcription.

Interestingly, typical features of active genes, such as H3K4me3 and H3K9ac are depleted in HCG promoters versus other expressed promoters. Instead, HCG promoters are associated with the enhancer features H3K4me1, H3K18ac, H3K27ac and enzymes mediating these modifications, Trithorax and CBP. One possibility is that looping between enhancers and HCG promoters is more stable than at other promoters, leading to detection of enhancer features at HCG promoters and contributing to efficient transcription from these genes. Another possibility is that depletion of active gene features and enrichment of enhancer marks at HCG promoters is orchestrated by the associated core promoter elements, mediating an alternative mode of transcription regulation. Differences in the chromatin landscape between promoters have previously been noted in *Drosophila* (Gilchrist and Adelman, 2012; Natsume-Kitatani and Mamitsuka, 2016), and recent studies have demonstrated unique chromatin features at different types of promoters also in human and yeast cells (Duttke et al., 2015; Kubik et al., 2015).

Our results demonstrate that CBP controls positioning of promoter-proximal paused Pol II at a majority of *Drosophila* genes by facilitating transcription through the first nucleosome and by stimulating Pol II recruitment. This global activator function for CBP is consistent with tethering experiments where a Gal4-CBP fusion protein strongly activated transcription from 24 different enhancer contexts (Stampfel et al., 2015). This suggests that CBP has critical roles not only in the previously described control of enhancer activity by H3K27 acetylation (reviewed in Holmqvist and Mannervik, 2013), but also in regulation of Pol II activity at the promoter.

STAR Methods

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Antibodies		
Rabbit polyclonal anti-CBP, affinity-purified	Holmqvist et al., 2012	N/A
Guinea-pig polyclonal anti-CBP, affinity purified	Lilja et al., 2007	N/A
Rabbit polyclonal anti-GAF (Trl)	Nakayama et al., 2007	N/A
Rabbit polyclonal anti-Rpb3	John Lis lab	N/A
Rabbit polyclonal anti-RNA polymerase II CTD Ser5 phosphorylation	Abcam	Cat#ab5131
Rabbit polyclonal anti-H3K4me3	Abcam	Cat#ab8580
Rabbit polyclonal anti-H3K9ac	Abcam	Cat#ab4441
Rabbit polyclonal anti-H3K27ac	Abcam	Cat#ab4729
Rabbit polyclonal anti-H3	Abcam	Cat#ab1791
Rabbit polyclonal anti-MRG15	This study	
Rabbit polyclonal anti-Ada2b	Lazlo Tora	Serum 1712
Rabbit polyclonal anti-TBP	Jim Kadonaga	N/A

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Rabbit polyclonal anti-TFIIA (large subunit)	Jim Kadonaga	N/A
Rabbit polyclonal anti-TFIIB	Jim Kadonaga	N/A
Rabbit polyclonal anti-TFIIF (dRAP30)	Jim Kadonaga	N/A
Rabbit polyclonal anti-HA tag	Abcam	Cat#ab9110
Bacterial and Virus Strains		
Biological Samples		
Chemicals, Peptides, and Recombinant Proteins		
C646 (CBP inhibitor)	Bowers et al., 2010	N/A
C37 (control)	Bowers et al., 2010	N/A
Curcumin	Tocris	Cat#2841
Critical Commercial Assays		
NEBNext Ultra II DNA Library Prep Kit	NEB	Cat#E7645
MEGAScript T7 Transcription kit	ThermoFisher	Cat#AM1334
Effectene Transfection Reagent	Qiagen	Cat#301425
5x HOT FIREPol EvaGreen qPCR Mix	Solis BioDyne	Cat# 08-25-00020
Deposited Data		
PRO-seq +/− CBP inhibition	This paper	GEO: GSE81649
TFIIB ChIP-seq and MNase-seq +/− CBP inhibition	This paper	GEO: GSE100614
CBP ChIP-seq	Philip et al., 2015	GEO: GSE64464
GAF ChIP-seq	Fuda et al., 2015	GEO: GSE40646
NELF and TFIIA ChIP-chip	Gilchrist et al., 2010	GEO: GSE20472
Chromatin factors ChIP-chip and ChIP-seq	Roy et al., 2010	modencode.org
Histone modifications ChIP-chip and ChIP-seq	Roy et al., 2010	modencode.org
Trx ChIP-seq	Enderle et al., 2011	GEO: GSE24521
RNA-seq	Cherbas et al., 2011	GEO: GSE15596
PRO-seq (Pol II pausing)	Kwak et al., 2013	GEO: GSE42117

REAGENT or RESOURCE	SOURCE	IDENTIFIER
CBP ChIP-seq *Drosophila* embryos	Holmqvist et al., 2012	GEO: GSE34221
Experimental Models: Cell Lines		
D. melanogaster: Cell line S2	Drosophila Genomics Resource Center	Cat#006
Experimental Models: Organisms/Strains		
Oligonucleotides		
Primers for ChIP-qPCR, RT-qPCR and dsRNA synthesis, see Table S3	This paper	N/A
Primers for *Drosophila* TFIIB ORF cloning, fwd: caccATGGCATCGACATCGAGACTG, rev: CATCTGTGGTAACTGATCAATGGGAGT		
Recombinant DNA		
pAWH-TFIIB (HA-tagged *Drosophila* TFIIB)	This paper	N/A
Software and Algorithms		
Bowtie2	Langmead and Salzberg, 2012	http://bowtie-bio.sourceforge.net/bowtie2/index.shtml
Samtools	Li et al., 2009	http://samtools.sourceforge.net/
BEDTools	Quinlan and Hall, 2010	http://bedtools.readthedocs.io/en/latest/
DESeq2	Love et al., 2014	https://bioconductor.org/packages/release/bioc/html/DESeq2.html
MEME	Machanick and Bailey, 2011	http://meme-suite.org/
R {GMD}	Zhao et al., 2011	https://CRAN.R-project.org/package=GMD
R (3.0.2)	R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing; 2014.	https://www.r-project.org/

REAGENT or RESOURCE	SOURCE	IDENTIFIER
STATISTICA 13	Dell Inc.	http://www.statsoft.com/Products/STATISTICA-Features
Other		

Contact for Reagent and Resource Sharing

Further information and requests for reagents should be directed to and will be fulfilled by the Lead Contact, Dr. Mattias Mannervik (mattias.mannervik@su.se).

Experimental Model and Subject Details

Drosophila S2 cells (DGRC catalog #006) were grown with Schneider's *Drosophila* medium supplemented with 10% FBS and 100 units/ml of penicillin and 100 μg/ml of streptomycin.

Method Details

Drug treatment of S2 cells—2×10^6 cells/ml of S2 cells were centrifuged and dissolved in FCS free media. 30 uM of CBP inhibitor (C646) or control drug (C37) in DMSO was added to the cells for the indicated time points (1, 3, 7, 10, 60, or 240 minutes). Alternatively, 400 μM of Curcumin (Tocris 2841) or DMSO (control) was added to the cells for the indicated time (10 or 60 minutes).

RNAi—GAF exonic sequence was PCR amplified from genomic DNA, whereas CBP and GFP exonic sequences were PCR amplified from cDNA using primers containing the T7 promoter plus the sequence complementary to the target gene (Table S3). dsRNA was produced using Ambion Megascript RNAi kit according to the manufacture's manual. Cells were centrifuged and washed twice in FCS free media. 37nM of dsRNA was added to the cells and shaken for 30s followed by 1 hour incubation. FCS containing media were added to a final concentration of 10% FCS. Cells were harvested after 4 days of GAF dsRNA treatment. For CBP dsRNA treatment, cells were treated a second time after 3 days and harvested after totally 6 days of treatment.

ChIP-qPCR and ChIP-seq—*Drosophila* S2 cells were grown to a density of $0.2–1 \times 10^7$ cells/ml and fixed in 1 % formaldehyde for 15 min at ambient temperature. The reaction was quenched by 0.16 M glycine pH 7.0 for 5 min and washed in PBS. Cells were sequentially washed with ChIP A (10 mM Hepes pH 7.6, 10 mM EDTA pH 8.0, 0.5 mM EGTA pH 8.0, 0.25 % Triton X100) and ChIP B (10 mM Hepes pH 7.6, 100 mM NaCl, 1 mM EDTA pH 8.0, 0.5 mM EGTA pH 8.0, 0.01 % Triton X100) for 10 min at 4°C followed by resuspension in Sonication buffer (50 mM Hepes, 140 mM NaCl, 1 mM EDTA, 1 % Triton, 0.1 % sodium deoxycholate, 0.1 % SDS, supplemented with proteinase inhibitor tablets, Roche) to a final concentration of $5–10 \times 10^7$ cells/ml. Nuclei were sonicated for 15 minutes using a Bioruptor (Diagenode), rotated for 10 minutes followed by centrifugation for 10 min at 13 000 rpm at 4°C.

A mix of Protein A and G Dynabeads (Invitrogen) blocked with BSA (1 mg/ml) and salmon sperm DNA (1 mg/ml, omitted from ChIP-seq) were mixed with indicated antibodies. Beads and antibodies were incubated for at least 2 hours followed by the addition of 0.5–1 × 10^7 cells. Chromatin and antibody bead complexes were formed during at least 2 hours followed by 5 minute washes with sonication buffer (50 mM Hepes, 140 mM NaCl, 1 mM EDTA, 1% Triton, 0.1% sodium deoxycholate, 0.1% SDS), WashA (as sonication buffer, but with 500 mM NaCl), WashB (20 mM Tris pH 8, 1 mM EDTA, 250 mM LiCl, 0.5% NP-40, 0.5% sodium deoxycholate) and TE.

Beads were resuspended in Elution buffer (50 mM Tris pH 8, 50 mM NaCl, 2 mM EDTA, 0.75% SDS, 20 μg/ml RNase A, 20 μg/ml glycogen) in a new tube. Cross-linking was reversed at 68°C for at least 4 hours and proteins removed by Proteinase K. DNA was purified with phenol-chloroform, ethanol precipitated and finally resuspended in 200 μl 0.1×TE.

ChIP samples were analyzed by qPCR or sequenced at SciLifeLab, Stockholm. 2 μl of DNA was used as template for qPCR, which was run in duplicates using 300nM primers and EvaGreen (Solis BioDyne) on a CFX96 Real-Time system (BioRad). Average Cq was calculated for each ChIP sample and compared to input. To account for the background of each individual ChIP, normalization was made to two intergenic sites devoid of known histone modifications and chromatin factors. ChIP values of histone modifications were further normalized to the total amount of histone H3. Primers are listed in Table S3. ChIP-seq libraries were prepared using the NEBNext Ultra II DNA Library Prep Kit (NEB) and paired-end sequenced on Illumina HiSeq2500, 2×125bp. The total number of mapped Mreads were 22.9 (C37 input rep1), 61.8 (C37 TFIIB rep1), 21.3 (C37 input rep2), 8.8 (C37 TFIIB rep2), 14.3 (C646 input rep1), 8.4 (C646 TFIIB rep1), 21.3 (C646 input rep2), and 7.1 (C646 TFIIB rep2). CBP ChIP-seq from *Drosophila* S2 cells (GEO accession: GSE64464) is described in (Philip et al., 2015).

Antibodies—The following antibodies were used in ChIP-qPCR; CBP (Holmqvist et al., 2012; Lilja et al., 2007; Lilja et al., 2003), GAF (kind gift from Susumu Hirose, National Institute of Genetics, Japan), Rpb3, Ser5 phosphorylated Pol II (Abcam, ab5131) H3K4me3 (Abcam, ab8580), H3K9ac (Abcam, ab4441), H3K27ac (Abcam, ab4729), H3 (Abcam, ab1791), Mrg15 (raised in rabbit), Ada2b (kind gift from Laszlo Tora, Université de Strasbourg), TBP, TFIIA, TFIIB, TFIIF (kind gifts from Jim Kadonaga, UCSD). The TFIIB antibody from Jim Kadonaga was used in ChIP-seq.

MNase-seq—1×10^8 S2 cells were treated with C646 CBP inhibitor or C37 control drug for 10 min and fixed in 1% formaldehyde for 2 minutes. The reaction was quenched by adding glycine to 125mM final concentration and incubated on ice. Cells were centrifuged at 3000 rpm for 5 minutes and resuspended in Nuclei buffer A (10mM Tris HCl pH 8, 300mM sucrose, 3mM $CaCl_2$, 2mM $MgAc_2$, 0.1% TritonX-100, 1x protease inhibitors (Roche)). Cells were homogenized in a Dounce homogenizer using 30 strokes followed by centrifugation at 1000 rpm. The pellet was washed in Nuclei buffer A followed by Digest buffer (15mM Tris HCl pH 8, 60mM KCl, 15mM NaCl, 1mM DTT, 250mM sucrose, 1mM $CaCl_2$) and finally resuspended in Digest buffer. 1×10^7nuclei were treated with 750 gel units

of MNase (NEB) for 30 minutes at room temperature to achieve about 80% mononucleosomes. The reaction was stopped by adding 0.1X 100mM EGTA and 3.5X Stop solution (143mM $NaHCO_3$/1.43% SDS/357mM NaCl/17.9mM EDTA). The crosslink was reversed by incubating at 65°C for 3.5 hours followed by a 30 minute incubation with Proteinase K. DNA was extracted using phenol-chloroform, precipitated and resuspended in RNase treated water followed by a incubation at 37°C for 30 minutes. Libraries were prepared using the NEBNext Ultra II DNA Library Prep Kit (NEB) and paired-end sequenced on Illumina HiSeq2500, 2×125bp. The total number of mapped Mreads were 20.3 (C37 rep1), 45.2 (C37 rep2), 34.3 (C646 rep1), and 42.4 (C646 rep2).

ChIP-seq and MNase-seq analysis—Reads were mapped to the Release 5 *Drosophila melanogaster* genome using Bowtie2 using default parameters. For the TFIIB ChIP-seq, read density was extracted and a $\log_2$-ratio between IP and the input was calculated. The data was median smoothed using a 100bp window and then subsampled to 10bp resolution. For the MNase-seq, the read density was extracted and subsampled to 10bp resolution, and an average of the two replicates was used for further analysis. The TFIIB data was normalized so that the average IP to input ratio is zero and the MNase data was normalized to the total number of mapped reads.

Isolation of nuclei, nuclear run-on and PRO-seq library preparation—Nuclei isolation and nuclear run-on were carried out essentially as described previously (Kwak et al., 2013; Love and Minton, 1985). In brief, S2 cells were treated with either C37 or C646 for 10 min followed by a PBS wash. Cells were resuspended in buffer L (10 mM Tris-HCl pH 7.5, 300mM sucrose, 10mM NaCl, 3 mM CaCl2, 2 mM MgCl2, 0.1% Triton X, 0.5 mM DTT, protease inhibitors cocktail (Roche), 4 u/ml RNase inhibitor (SUPERaseIN, Ambion) and immediately dounced 25 strokes with a tight fitting pestle. Lysed cells were mixed with an equal amount of buffer B (10 mM Tris-HCl pH 8, 2M sucrose, 10mM $NaCl_2$, 2 mM $MgCl_2$, 0.5 mM DTT, protease inhibitors cocktail (Roche), 4 u/ml RNase inhibitor (SUPERaseIN, Ambion) and loaded onto a buffer B sucrose pillow. The sample was spun at 23000g on a SW-41 rotor, the supernatant was removed, and the nuclei were washed once in storage buffer (50 mM Tris-HCl pH 8, 25% glycerol, 5mM $MgAc_2$, 0.1mM EDTA, 5mM DTT) and recovered by centrifuging at 1000g for 5 minutes. Isolated nuclei were resuspended in storage buffer and kept at −80°C. Nuclei from two biological replicates were prepared for each condition.

Nuclear run-on was carried out as described previously with some modifications (Kwak et al., 2013). Briefly, 10 million nuclei in 100 μl of storage buffer were mixed with 100 μl of 2x nuclear run-on buffer (10 mM Tris-HCl pH 8.0, 5 mM $MgCl_2$, 1 mM DTT, 300 mM KCl, 1% Sarkosyl, 50 uM biotin-11-A/C/G/UTP, 0.2 units/μl RNase inhibitor) and incubated at 30°C for three minutes. RNAs were isolated using TRIZOL LS and base-hydrolyzed with 200 nM final concentration of NaOH, generating an average size of RNA between 100–150 nucleotides. Nascent RNAs from run-on reactions, characterized by the addition of biotinylated nucleotide, were isolated with magnetic beads coated with streptavidin. After the ligation of 3′ adapter and a second biotin streptavidin affinity purification, the mRNA cap was removed and the 5′ adapter was ligated. Then a third biotin streptavidin affinity

purification was carried out and cDNA was generated by reverse transcription. The generated cDNA was amplified with 9 cycles of PCR using Illumina TruSeq small-RNA adaptors for sequencing.

Normalization—To analyze the effect of CBP inhibition on transcription, we identified genes with differential PRO-seq signal. However, to normalize the datasets from the two treatments we could not apply standard normalization techniques as they assume that only a minority of the data are changed between treatments (Landfors et al., 2011). Since CBP occupancy is higher than the genomic mean at virtually all expressed promoters (Fig. 1B), we expect it to affect expression of a large number of genes. We therefore normalized the PRO-seq datasets using the TSS proximal signal (TSS +100 bp) of the 5% (n=270) expressed genes with the least CBP signal. At these genes, we expect the smallest changes after CBP inhibition. After this normalization, we quantified the change in PRO-seq density upon CBP inhibition at all expressed genes.

Differential expression analyses of pause region and gene body—Significant change in PRO-seq density in the pause region and the body of gene was calculated using DESeq2 (Love et al., 2014) by analyzing the two biological replicates separately. PRO-seq count reads were calculated from the 100 bp pause region or gene body (500 bp downstream of TSS to 100 bp upstream of poly A site). Adjusted p-value cutoff of 0.001 and fold change of at least 1.25 was used in DESeq2. For all other analyzes of the PRO-seq data the reads from the two biological replicates were merged.

Data—Data for chromatin factors, chromatin states, histone modifications and expression is from (Cherbas et al., 2011; Enderle et al., 2011; Roy et al., 2010), GAF from (Fuda et al., 2015), Pol II pausing from (Kwak et al., 2013), NELF and TFIIA from (Gilchrist et al., 2010), CBP embryo data from (Holmqvist et al., 2012). Genome annotation 5.57 from Flybase was used (St Pierre et al., 2014). Promoters were defined as the +/− 500 bp region around the TSS. For genes with more than one annotated TSS, the promoter with highest CBP+GAF enrichment was selected. Expressed gene promoters were divided into 20 equally sized bins using CBP+GAF occupancy and grouped into 3 classes based on their mean pausing index. Since the confidence intervals of bins 19 and 20 overlap their mean pausing values, but not the mean pausing value of bin 18, they were denoted high CBP and GAF (HCG). Similarly, the confidence interval of bin 15 does not overlap the mean pausing index of bin 14, distinguishing MCG from LCG promoters.

To plot several ChIP datasets together they were linearly scaled so that zero corresponds to the genomic mean and 1 to the mean of the top 0.1% data points (to account for outliers with extremely high enrichment in some datasets). Zero then represents background levels of enrichment and 1 the strongest enrichment observed in the genome. For all plots, enrichment values of all factors were trimmed by their 99^{th} percentile in order to remove the extreme outliers.

Heatmaps—The heatmaps for CBP, GAF and Pol II were made in R, using Heatmap.3 function from GMD package. Prior to making these heatmaps, promoters lacking mappable ChIP-seq reads were excluded (n=254).

DNA motif discovery—CBP peaks at expressed promoters were defined as the position of the highest CBP enrichment in the promoter proximal region (TSS +/− 500bp). These positions were then extended 100bp in each direction and the corresponding sequences extracted using the getfasta function in BEDtools (Quinlan and Hall, 2010). Next, *de-novo* motif discovery was performed using the chip-meme module in the MEMEsuite (Machanick and Bailey, 2011). The parameters were set to find any number of motifs with minimum length of 6bp and maximum length of 15bp.

Cloning and expression of HA-tagged TFIIB—cDNA from S2 cells was prepared by isolating total RNA using Trizol (Invitrogen) according to manufactorers' protocol. RNA was DNase treated (Sigma) and used as template for cDNA synthesis (High capacity, Applied Biosystems). The resulting cDNA was PCR amplified with primers 5′ caccATGGCATCGACATCGAGACTG 3′ and 5′ CATCTGTGGTAACTGATCAATGGGAGT 3′ and cloned with the pENTR/D-TOPO cloning kit according to manufacturers instructions (Invitrogen). TFIIB was transferred from the resulting plasmid to the Actin5C-driven expression vector pAWH by Gateway LR Clonase II (Invitrogen) to introduce a C-terminal HA tag. 200ng of the expression plasmid was transfected to S2 cells by Effectene (QIAgen) according to manufacturers instructions, and harvested 2 days later.

Co-Immunoprecipitation of Micrococcal Nuclease solubilized protein extracts —S2 cells were washed in PBS, and resuspended in PBS + 0.1% Triton X-100. After 3 min mixing by rotation at room temperature, cells were collected at 2000g for 5 min and resuspended in 160µl MNase buffer (5% glycerol, 20mM Tris pH7.4, 60mM KCl, 15mM NaCl, 5mM CaCl2, 3mM MgCl2, 0.5% NP-40) to which 1mM DTT and 4.5U/ul Micrococcal Nuclease was freshly added. Samples were incubated at 25°C for 5 min. Reaction was stopped by the addition of 240µl MNase coIP dilution buffer (3.6mM Tris pH8.8, 12mM EDTA, 225mM NaCl, 60mM KCl, 1.2% NP-40, proteinase inhibitor tablets, Roche). Cells were shaken vigorously on a shaker at 4°C for 10 min, and centrifuged at 4°C, 10000g for 5 min. Supernatant was pre-cleared with 25µl of a mix of BSA blocked protein A and G Dynabeads (Invitrogen) at 4°C for 10 min. Resulting protein extract was added to 25µl BSA blocked protein A and G Dynabeads pre-incubated with anti-HA (Abcam, ab9110) or rabbit anti-CBP. After overnight incubation, the beads were washed twice with CoIP wash buffer (10mM Tris pH8.0, 140mM NaCl, 1.5mM MgCl2, 0.1% NP-40), once each with TEN (10mM Tris pH8.0, 1mM EDTA, 100mM NaCl) and 50mM Tris pH6.8. Beads were resuspended in 45µl Laemmli loading buffer, and 15µl was loaded on 5% or 12% SDS-PAGE gels for Western blot, and probed with guinea-pig anti-CBP (Lilja et al., 2003, 1:200) or anti-HA (Abcam, ab9110, 1:1000).

Quantification and Statistical Analysis

For all quantification analyses, sample size (N) and statistical tests performed are described in the corresponding figure legends.

Data and Software Availability

Accession Numbers—The accession numbers for data reported in this paper are GEO: GSE81649 (PRO-seq) and GEO: GSE100614 (TFIIB ChIP-seq and MNase-seq).

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We thank Susumu Hirose, Laszlo Tora and Jim Kadonaga for reagents, Karen Adelman for sharing data, and Stefan Åström for comments on the manuscript. This work was supported by grants from the Swedish Cancer Society and the Swedish Research Council (M.M.), the Knut and Alice Wallenberg (to EpiCoN, co-PI: P.S.), Kempe, Åke Wiberg, Magnus Bergvall and Carl Trygger foundations (P.S.), the NIH and FAMRI (D.M. and P.C.), and NIH GM25232 to J.T.L.

P.C. is a cofounder and advisor of Acylin Therapeutics.

References

Adelman K, Lis JT. Promoter-proximal pausing of RNA polymerase II: emerging roles in metazoans. Nat Rev Genet. 2012; 13:720–731. [PubMed: 22986266]

Adkins NL, Hagerman TA, Georgel P. GAGA protein: a multi-faceted transcription factor. Biochem Cell Biol. 2006; 84:559–567. [PubMed: 16936828]

Bedford DC, Kasper LH, Fukuyama T, Brindle PK. Target gene context influences the transcriptional requirement for the KAT3 family of CBP and p300 histone acetyltransferases. Epigenetics. 2010; 5:9–15. [PubMed: 20110770]

Black JC, Choi JE, Lombardo SR, Carey M. A mechanism for coordinating chromatin modification and preinitiation complex assembly. Mol Cell. 2006; 23:809–818. [PubMed: 16973433]

Bowers EM, Yan G, Mukherjee C, Orry A, Wang L, Holbert MA, Crump NT, Hazzalin CA, Liszczak G, Yuan H, et al. Virtual ligand screening of the p300/CBP histone acetyltransferase: identification of a selective small molecule inhibitor. Chem Biol. 2010; 17:471–482. [PubMed: 20534345]

Cherbas L, Willingham A, Zhang D, Yang L, Zou Y, Eads BD, Carlson JW, Landolin JM, Kapranov P, Dumais J, et al. The transcriptional diversity of 25 Drosophila cell lines. Genome Res. 2011; 21:301–314. [PubMed: 21177962]

Choi CH, Hiromura M, Usheva A. Transcription factor IIB acetylates itself to regulate transcription. Nature. 2003; 424:965–969. [PubMed: 12931194]

Chopra VS, Srinivasan A, Kumar RP, Mishra K, Basquin D, Docquier M, Seum C, Pauli D, Mishra RK. Transcriptional activation by GAGA factor is through its direct interaction with dmTAF3. Dev Biol. 2008; 317:660–670. [PubMed: 18367161]

Core LJ, Lis JT. Transcription regulation through promoter-proximal pausing of RNA polymerase II. Science. 2008; 319:1791–1792. [PubMed: 18369138]

Core LJ, Waterfall JJ, Gilchrist DA, Fargo DC, Kwak H, Adelman K, Lis JT. Defining the status of RNA polymerase at promoters. Cell Rep. 2012; 2:1025–1035. [PubMed: 23062713]

Crump NT, Hazzalin CA, Bowers EM, Alani RM, Cole PA, Mahadevan LC. Dynamic acetylation of all lysine-4 trimethylated histone H3 is evolutionarily conserved and mediated by p300/CBP. Proc Natl Acad Sci U S A. 2011; 108:7814–7819. [PubMed: 21518915]

Dancy BM, Cole PA. Protein lysine acetylation by p300/CBP. Chem Rev. 2015; 115:2419–2452. [PubMed: 25594381]

Dancy BM, Crump NT, Peterson DJ, Mukherjee C, Bowers EM, Ahn YH, Yoshida M, Zhang J, Mahadevan LC, Meyers DJ, et al. Live-cell studies of p300/CBP histone acetyltransferase activity and inhibition. Chembiochem. 2012; 13:2113–2121. [PubMed: 22961914]

Duarte FM, Fuda NJ, Mahat DB, Core LJ, Guertin MJ, Lis JT. Transcription factors GAF and HSF act at distinct regulatory steps to modulate stress-induced gene activation. Genes Dev. 2016; 30:1731–1746. [PubMed: 27492368]

Duttke SH, Lacadie SA, Ibrahim MM, Glass CK, Corcoran DL, Benner C, Heinz S, Kadonaga JT, Ohler U. Human promoters are intrinsically directional. Mol Cell. 2015; 57:674–684. [PubMed: 25639469]

Enderle D, Beisel C, Stadler MB, Gerstung M, Athri P, Paro R. Polycomb preferentially targets stalled promoters of coding and noncoding transcripts. Genome Res. 2011; 21:216–226. [PubMed: 21177970]

Feller C, Forne I, Imhof A, Becker PB. Global and specific responses of the histone acetylome to systematic perturbation. Mol Cell. 2015; 57:559–571. [PubMed: 25578876]

Fuda NJ, Guertin MJ, Sharma S, Danko CG, Martins AL, Siepel A, Lis JT. GAGA factor maintains nucleosome-free regions and has a role in RNA polymerase II recruitment to promoters. PLoS Genet. 2015; 11:e1005108. [PubMed: 25815464]

Gilchrist DA, Adelman K. Coupling polymerase pausing and chromatin landscapes for precise regulation of transcription. Biochim Biophys Acta. 2012; 1819:700–706. [PubMed: 22406341]

Gilchrist DA, Dos Santos G, Fargo DC, Xie B, Gao Y, Li L, Adelman K. Pausing of RNA polymerase II disrupts DNA-specified nucleosome organization to enable precise gene regulation. Cell. 2010; 143:540–551. [PubMed: 21074046]

Gilchrist DA, Nechaev S, Lee C, Ghosh SK, Collins JB, Li L, Gilmour DS, Adelman K. NELF-mediated stalling of Pol II can enhance gene expression by blocking promoter-proximal nucleosome assembly. Genes Dev. 2008; 22:1921–1933. [PubMed: 18628398]

Holmqvist PH, Boija A, Philip P, Crona F, Stenberg P, Mannervik M. Preferential genome targeting of the CBP co-activator by Rel and Smad proteins in early Drosophila melanogaster embryos. PLoS Genet. 2012; 8:e1002769. [PubMed: 22737084]

Holmqvist PH, Mannervik M. Genomic occupancy of the transcriptional co-activators p300 and CBP. Transcription. 2013; 4:18–23. [PubMed: 23131664]

Jain D, Baldi S, Zabel A, Straub T, Becker PB. Active promoters give rise to false positive 'Phantom Peaks' in ChIP-seq experiments. Nucleic Acids Res. 2015; 43:6959–6968. [PubMed: 26117547]

Jin Q, Yu LR, Wang L, Zhang Z, Kasper LH, Lee JE, Wang C, Brindle PK, Dent SY, Ge K. Distinct roles of GCN5/PCAF-mediated H3K9ac and CBP/p300-mediated H3K18/27ac in nuclear receptor transactivation. Embo J. 2011; 30:249–262. [PubMed: 21131905]

Kubik S, Bruzzone MJ, Jacquet P, Falcone JL, Rougemont J, Shore D. Nucleosome Stability Distinguishes Two Different Promoter Types at All Protein-Coding Genes in Yeast. Mol Cell. 2015; 60:422–434. [PubMed: 26545077]

Kwak H, Fuda NJ, Core LJ, Lis JT. Precise maps of RNA polymerase reveal how promoters direct initiation and pausing. Science. 2013; 339:950–953. [PubMed: 23430654]

Kwok RP, Lundblad JR, Chrivia JC, Richards JP, Bachinger HP, Brennan RG, Roberts SG, Green MR, Goodman RH. Nuclear protein CBP is a coactivator for the transcription factor CREB. Nature. 1994; 370:223–226. [PubMed: 7913207]

Landfors M, Philip P, Ryden P, Stenberg P. Normalization of high dimensional genomics data where the distribution of the altered variables is skewed. PLoS One. 2011; 6:e27942. [PubMed: 22132175]

Langmead B, Salzberg SL. Fast gapped-read alignment with Bowtie 2. Nat Methods. 2012; 9:357–359. [PubMed: 22388286]

Lee C, Li X, Hechmer A, Eisen M, Biggin MD, Venters BJ, Jiang C, Li J, Pugh BF, Gilmour DS. NELF and GAGA factor are linked to promoter-proximal pausing at many genes in Drosophila. Mol Cell Biol. 2008; 28:3290–3300. [PubMed: 18332113]

Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R. Genome Project Data Processing S. The Sequence Alignment/Map format and SAMtools. Bioinformatics. 2009; 25:2078–2079. [PubMed: 19505943]

Li J, Liu Y, Rhee HS, Ghosh SK, Bai L, Pugh BF, Gilmour DS. Kinetic competition between elongation rate and binding of NELF controls promoter-proximal pausing. Mol Cell. 2013; 50:711–722. [PubMed: 23746353]

Lilja T, Aihara H, Stabell M, Nibu Y, Mannervik M. The acetyltransferase activity of Drosophila CBP is dispensable for regulation of the Dpp pathway in the early embryo. Dev Biol. 2007; 305:650–658. [PubMed: 17336283]

Lilja T, Qi D, Stabell M, Mannervik M. The CBP coactivator functions both upstream and downstream of Dpp/Screw signaling in the early Drosophila embryo. Dev Biol. 2003; 262:294–302. [PubMed: 14550792]

Love JD, Minton KW. Screening of lambda library for differentially expressed genes using in vitro transcripts. Anal Biochem. 1985; 150:429–441. [PubMed: 2418709]

Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol. 2014; 15:550. [PubMed: 25516281]

Machanick P, Bailey TL. MEME-ChIP: motif analysis of large DNA datasets. Bioinformatics. 2011; 27:1696–1697. [PubMed: 21486936]

Nakayama T, Nishioka K, Dong YX, Shimojima T, Hirose S. Drosophila GAGA factor directs histone H3.3 replacement that prevents the heterochromatin spreading. Genes Dev. 2007; 21:552–561. [PubMed: 17344416]

Natsume-Kitatani Y, Mamitsuka H. Classification of Promoters Based on the Combination of Core Promoter Elements Exhibits Different Histone Modification Patterns. PLoS One. 2016; 11:e0151917. [PubMed: 27003446]

Negre N, Brown CD, Ma L, Bristow CA, Miller SW, Wagner U, Kheradpour P, Eaton ML, Loriaux P, Sealfon R, et al. A cis-regulatory map of the Drosophila genome. Nature. 2011; 471:527–531. [PubMed: 21430782]

Philip P, Boija A, Vaid R, Churcher AM, Meyers DJ, Cole PA, Mannervik M, Stenberg P. CBP binding outside of promoters and enhancers in. Epigenetics Chromatin. 2015; 8:48. [PubMed: 26604986]

Pugh BF, Venters BJ. Genomic Organization of Human Transcription Initiation Complexes. PLoS One. 2016; 11:e0149339. [PubMed: 26866362]

Quinlan AR, Hall IM. BEDTools: a flexible suite of utilities for comparing genomic features. Bioinformatics. 2010; 26:841–842. [PubMed: 20110278]

Roy S, Ernst J, Kharchenko PV, Kheradpour P, Negre N, Eaton ML, Landolin JM, Bristow CA, Ma L, Lin MF, et al. Identification of functional elements and regulatory circuits by Drosophila modENCODE. Science. 2010; 330:1787–1797. [PubMed: 21177974]

Schneider I. Cell lines derived from late embryonic stages of Drosophila melanogaster. J Embryol Exp Morphol. 1972; 27:353–365. [PubMed: 4625067]

Schroder S, Herker E, Itzen F, He D, Thomas S, Gilchrist DA, Kaehlcke K, Cho S, Pollard KS, Capra JA, et al. Acetylation of RNA polymerase II regulates growth-factor-induced gene transcription in mammalian cells. Mol Cell. 2013; 52:314–324. [PubMed: 24207025]

Shimojima T, Okada M, Nakayama T, Ueda H, Okawa K, Iwamatsu A, Handa H, Hirose S. Drosophila FACT contributes to Hox gene expression through physical and functional interactions with GAGA factor. Genes Dev. 2003; 17:1605–1616. [PubMed: 12815073]

Shrimp JH, Sorum AW, Garlick JM, Guasch L, Nicklaus MC, Meier JL. Characterizing the Covalent Targets of a Small Molecule Inhibitor of the Lysine Acetyltransferase P300. ACS Med Chem Lett. 2016; 7:151–155. [PubMed: 26985290]

St Pierre SE, Ponting L, Stefancsik R, McQuilton P. FlyBase 102--advanced approaches to interrogating FlyBase. Nucleic Acids Res. 2014; 42:D780–788. [PubMed: 24234449]

Stampfel G, Kazmar T, Frank O, Wienerroither S, Reiter F, Stark A. Transcriptional regulators form diverse groups with context-dependent regulatory functions. Nature. 2015; 528:147–151. [PubMed: 26550828]

Stasevich TJ, Hayashi-Takanaka Y, Sato Y, Maehara K, Ohkawa Y, Sakata-Sogawa K, Tokunaga M, Nagase T, Nozaki N, McNally JG, et al. Regulation of RNA polymerase II activation by histone acetylation in single living cells. Nature. 2014; 516:272–275. [PubMed: 25252976]

Tie F, Banerjee R, Stratton CA, Prasad-Sinha J, Stepanik V, Zlobin A, Diaz MO, Scacheri PC, Harte PJ. CBP-mediated acetylation of histone H3 lysine 27 antagonizes Drosophila Polycomb silencing. Development. 2009; 136:3131–3141. [PubMed: 19700617]

Visel A, Blow MJ, Li Z, Zhang T, Akiyama JA, Holt A, Plajzer-Frick I, Shoukry M, Wright C, Chen F, et al. ChIP-seq accurately predicts tissue-specific activity of enhancers. Nature. 2009; 457:854–858. [PubMed: 19212405]

Wang Z, Zang C, Cui K, Schones DE, Barski A, Peng W, Zhao K. Genome-wide mapping of HATs and HDACs reveals distinct functions in active and inactive genes. Cell. 2009; 138:1019–1031. [PubMed: 19698979]

Xiao H, Sandaltzopoulos R, Wang HM, Hamiche A, Ranallo R, Lee KM, Fu D, Wu C. Dual functions of largest NURF subunit NURF301 in nucleosome sliding and transcription factor interactions. Mol Cell. 2001; 8:531–543. [PubMed: 11583616]

Zhang Z, English BP, Grimm JB, Kazane SA, Hu W, Tsai A, Inouye C, You C, Piehler J, Schultz PG, et al. Rapid dynamics of general transcription factor TFIIB binding during preinitiation complex assembly revealed by single-molecule analysis. Genes Dev. 2016; 30:2106–2118. [PubMed: 27798851]

Zhao X, Valen E, Parker BJ, Sandelin A. Systematic clustering of transcription start site landscapes. PLoS One. 2011; 6:e23409. [PubMed: 21887249]

Zhou Q, Li T, Price DH. RNA polymerase II elongation control. Annu Rev Biochem. 2012; 81:119–143. [PubMed: 22404626]

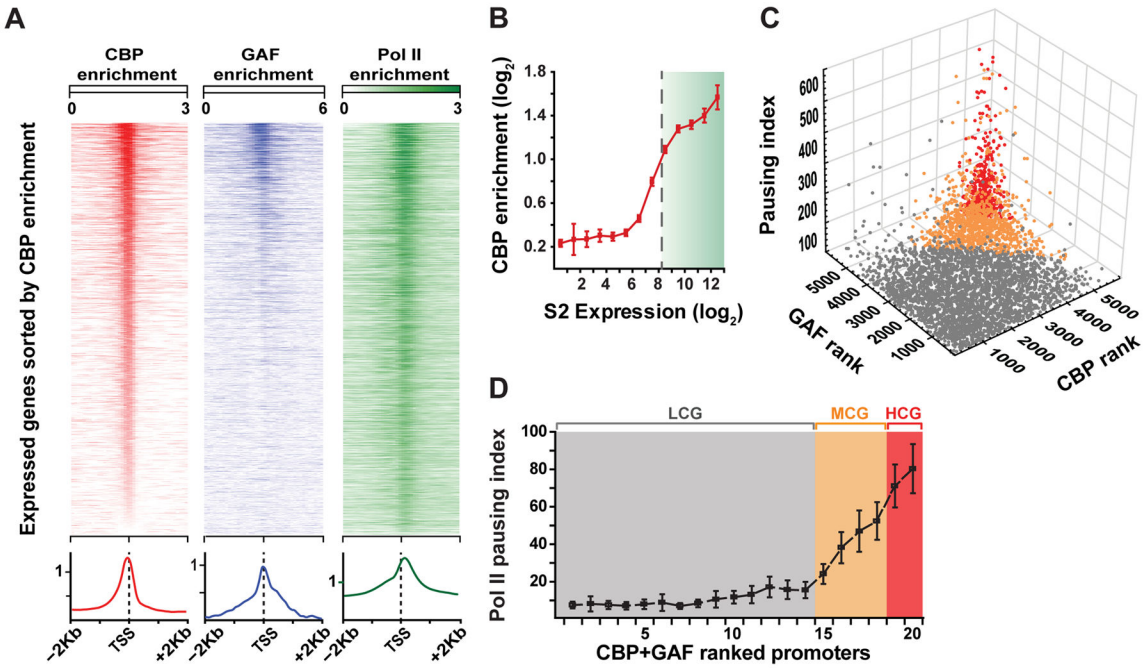

Figure 1. Promoters occupied by both CBP and GAF are highly paused

(**A**) CBP, GAF and Pol II occupancy in *Drosophila* S2 cells plotted around the TSS of all expressed genes (defined in (Cherbas et al., 2011), n=5400), sorted on CBP binding strength. (**B**) All *Drosophila* genes were binned according to expression value (log_2 RPKM) and the mean CBP occupancy calculated for each bin. Promoters of expressed genes (>300 RPKM (Cherbas et al., 2011), shaded green) show stronger CBP enrichment than promoters of non- or less-expressed genes. (**C**) The 5400 expressed genes in S2 cells (as defined in (Cherbas et al., 2011)) were ranked according to the amount of CBP or GAF at their promoter and plotted against their pausing index (Kwak et al., 2013). (**D**) Expressed genes were divided into 20 equally sized promoter bins based on the combined GAF + CBP rank order, and plotted against average Pol II pausing index (Kwak et al., 2013). Bins 19 and 20 have the highest CBP + GAF enrichment and the highest pausing index, (high CBP + GAF, HCG, colored red in panels C–D), bins 15–18 are medium-paused and have medium CBP and GAF enrichment (MCG, orange), and bins 1–14 are lowly-paused and have the least CBP + GAF (LCG, grey). Whiskers represent +/− 95% confidence interval. See also Figure S1 and Table S1.

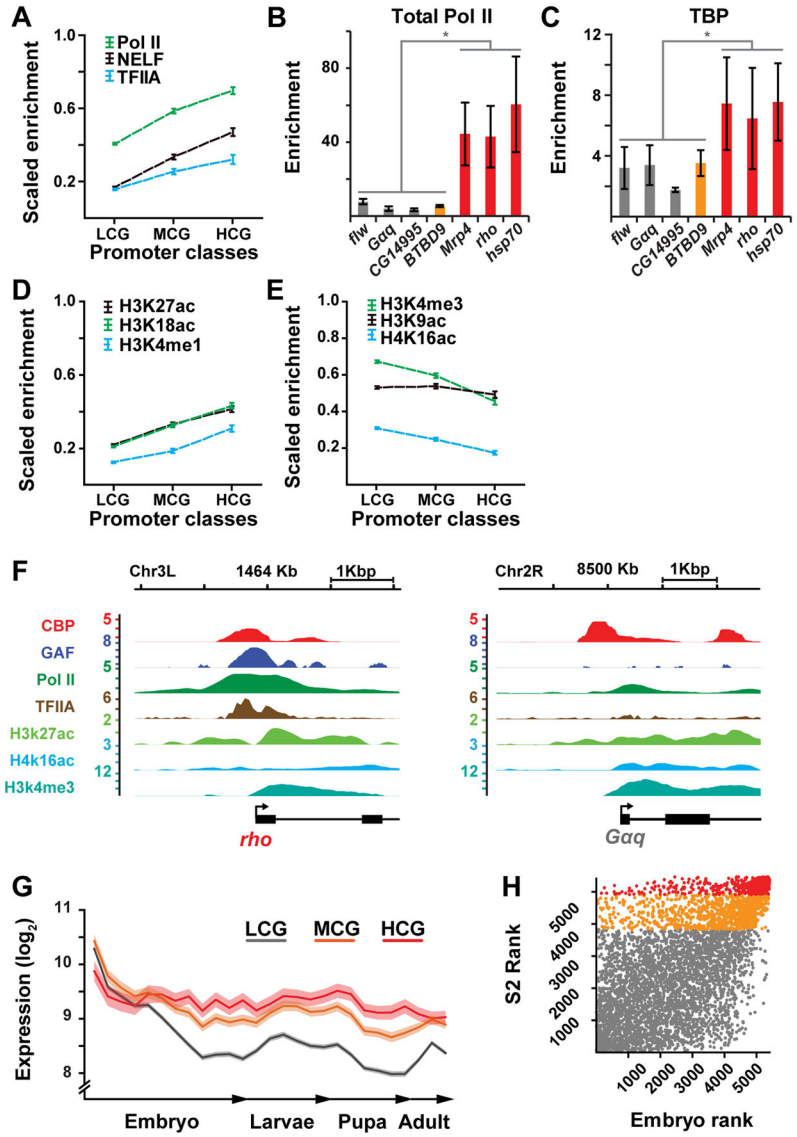

Figure 2. Unique chromatin features of promoters with high CBP and GAF enrichment

(**A**) Average Pol II, NELF and TFIIA (Gilchrist et al., 2010; Roy et al., 2010) scaled enrichment in the 5400 promoters of expressed genes with low (LCG), medium (MCG), and high CBP + GAF (HCG). (**B**–**C**) Enrichment (fold over 2 background loci) of Pol II (Rpb3 antibody) (**B**) and TBP (**C**) at LCG (*flw, Gaq, CG14995*), MCG (*BTBD9*), and HCG promoters (*Mrp4, rho, hsp70*) measured by ChIP-qPCR (n=2–5). Error bars represent S.E.M. HCG gene promoters contain significantly more Pol II and TBP than other promoters, two tailed t-test, $p<0.05$. (**D**–**E**) Average scaled enrichment of H3K27ac, H3K18ac, H3K4me1 (**D**), and H3K9ac, H3K4me3 and H4K16ac (**E**) (Roy et al., 2010) in the promoter classes. **(F)** ChIP-seq profiles of proteins and histone modifications at the *rho* (HCG) and *Gaq* (LCG) genes, illustrating the differences in enrichment between promoter classes. (**G**) Average expression of genes in the promoter classes across 27 developmental time points. Shades represent +/− 95% confidence interval. (**H**) Scatter plot of promoters

from all genes expressed in S2 cells (n=5400) ranked by combined CBP + GAF enrichment in S2 cells and in early (2–4h) embryos. HCG, MCG and LCG promoters according to the S2 ranking are color coded. See also Figures S2–S3 and Tables S1–S2.

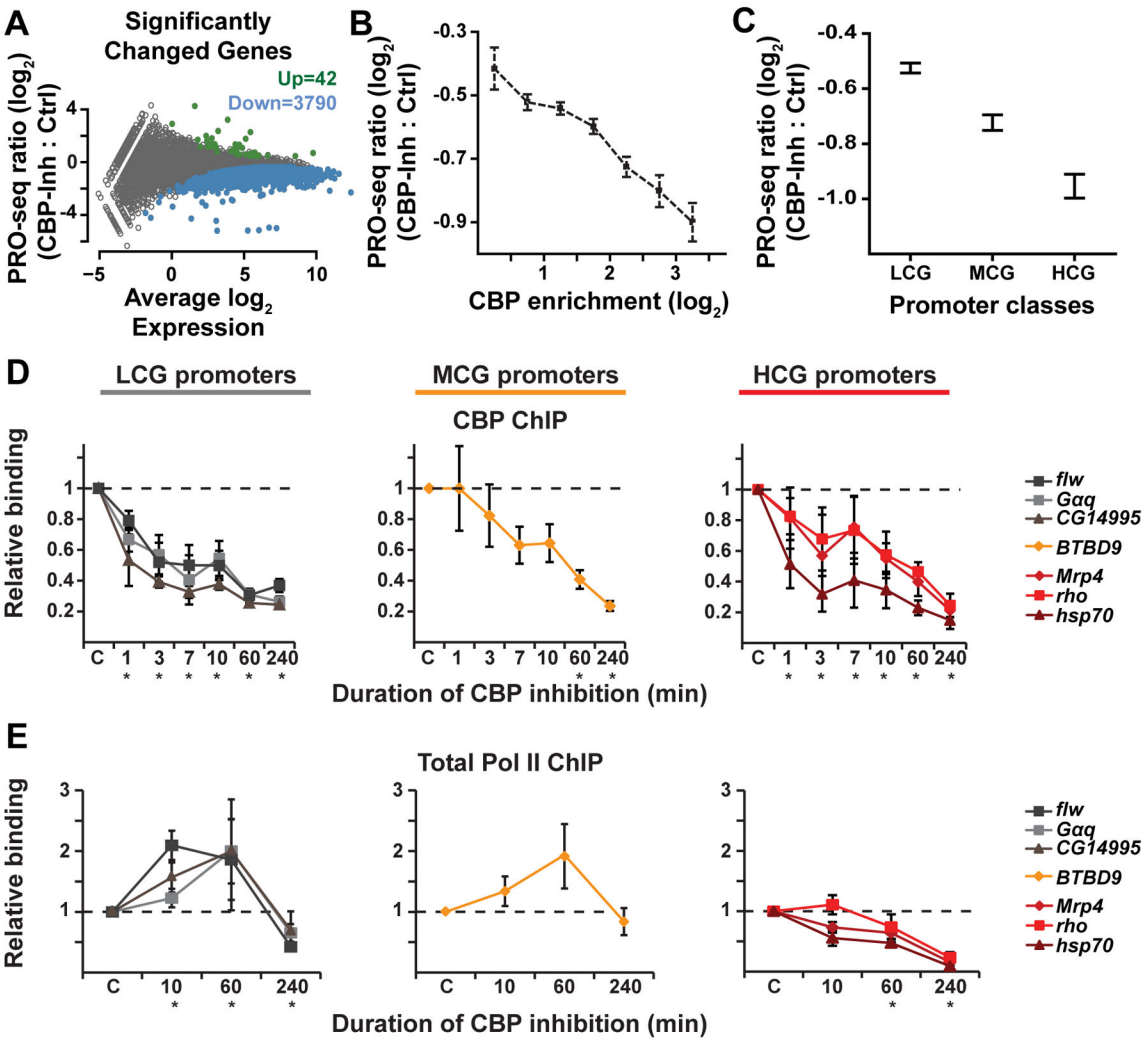

Figure 3. CBP inhibition affects Pol II occupancy and transcription

(**A**) Difference in gene-body (500 bp downstream of TSS to 100 bp upstream of poly A site) precision run-on sequencing (PRO-seq) reads between 10 min CBP-inhibitor (C646, CBP-Inh) and control drug (C37, Ctrl)-treated S2 cells after normalization. 3790 genes were down-regulated. (**B**) Average gene-body PRO-seq ratio between CBP inhibitor and control-treated cells plotted versus CBP enrichment at the promoter. (**C**) Average gene-body PRO-seq ratio between CBP inhibitor and control-treated cells plotted versus promoter class. (**D**) CBP and (**E**) total Pol II (Rpb3) ChIP-qPCR from S2 cells treated with C646 or control drug (C37) for the indicated time points. Occupancy at the indicated gene promoters is plotted relative to control-treated cells. Error bars represent S.E.M., n=2–5. A significant difference between C646- and control-treated cells for the genes as a group is indicated by an asterisk below the time points (two-tailed paired t-test, $p<0.05$). See also Figures S4–S5.

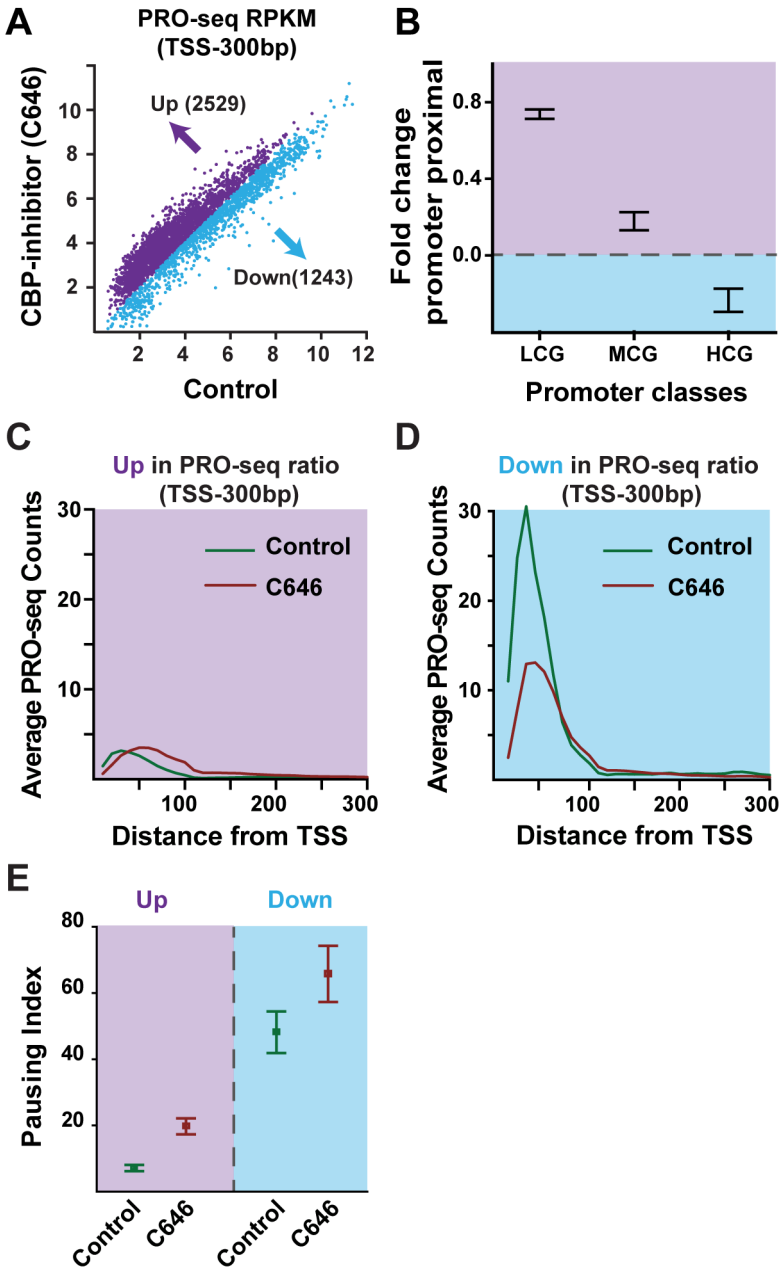

Figure 4. CBP regulates Pol II occupancy depending on promoter type, but is required for release into elongation genome-wide

Effects on promoters of genes (n=3790) with reduced gene body PRO-seq reads after 10 min of CBP inhibition (C646). (**A**) Scatterplot of PRO-seq counts in the promoter-proximal region (TSS-300 bp) after 10 min CBP inhibition (C646) compared to control (C37). 2529 genes had more PRO-seq reads in this region (shown in purle) and 1243 genes had less PRO-seq reads (shown in blue). (**B**) PRO-seq $\log_2$ fold change in the promoter-proximal region (TSS-300bp) after CBP inhibition in the promoter classes. (**C–D**) Metagene PRO-seq profiles (PRO-seq counts/10bp) for genes with decreased gene-body transcription after CBP inhibition but with increased (purple) (**C**) or decreased (blue) (**D**) promoter-proximal PRO-

seq reads. (**E**) Pausing Index (PRO-seq reads from TSS to 300bp region versus gene-body) in control and after 10 min of CBP inhibition. Increased pausing after inhibition suggests that CBP promotes release of Pol II into productive elongation.

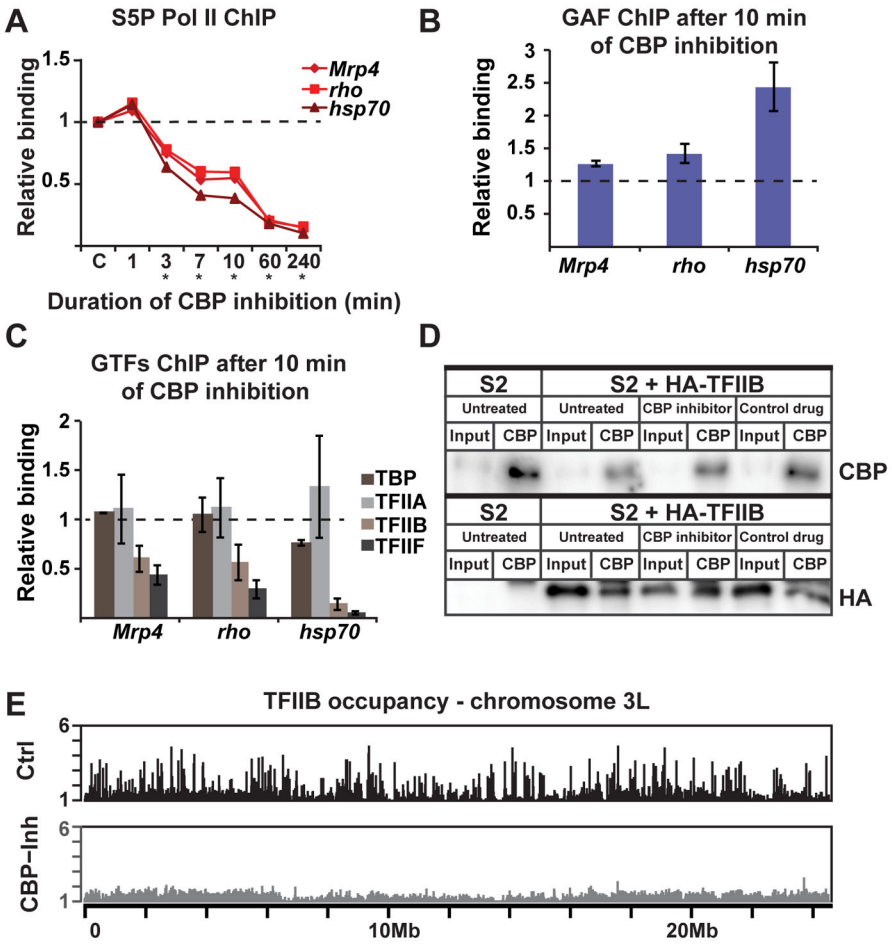

Figure 5. CBP recruits Pol II by promoting TFIIB occupancy

(**A**) Serine 5 phosphorylated (S5P) Pol II ChIP-qPCR after CBP inhibitor (C646) treatment for the indicated time points (n=2–5). Occupancy at HCG promoters (*Mrp4, rho* and *hsp70)* are plotted relative to control (C37)-treated cells. Error bars represent S.E.M., two-tailed paired t-test, *p<0.05 at the indicated time points for the genes as a group. (**B**) GAF ChIP from 10 min of C646-treated relative to control-treated S2 cells. (**C**) TBP, TFIIA, TFIIB and TFIIF ChIP after 10 min of drug treatment. (**D**) Co-immunoprecipitation from S2 cells either untransfected (S2) or transfected with HA-tagged TFIIB. Cells were untreated, treated with C646 or C37 control for 1 hour. 10% input and anti-CBP immunoprecipitated material was loaded on 5% or 12% SDS-PAGE gels. Top: CBP Western blot. Bottom: HA-tagged TFIIB blot. TFIIB interacts with CBP in the absence and presence of inhibitor. (**E**) TFIIB ChIP-seq in control (C37) and C646-treated cells (10 min). Genome browser views of entire chromosome 3L show reduced TFIIB occupancy after CBP inhibition. See also Figure S6.

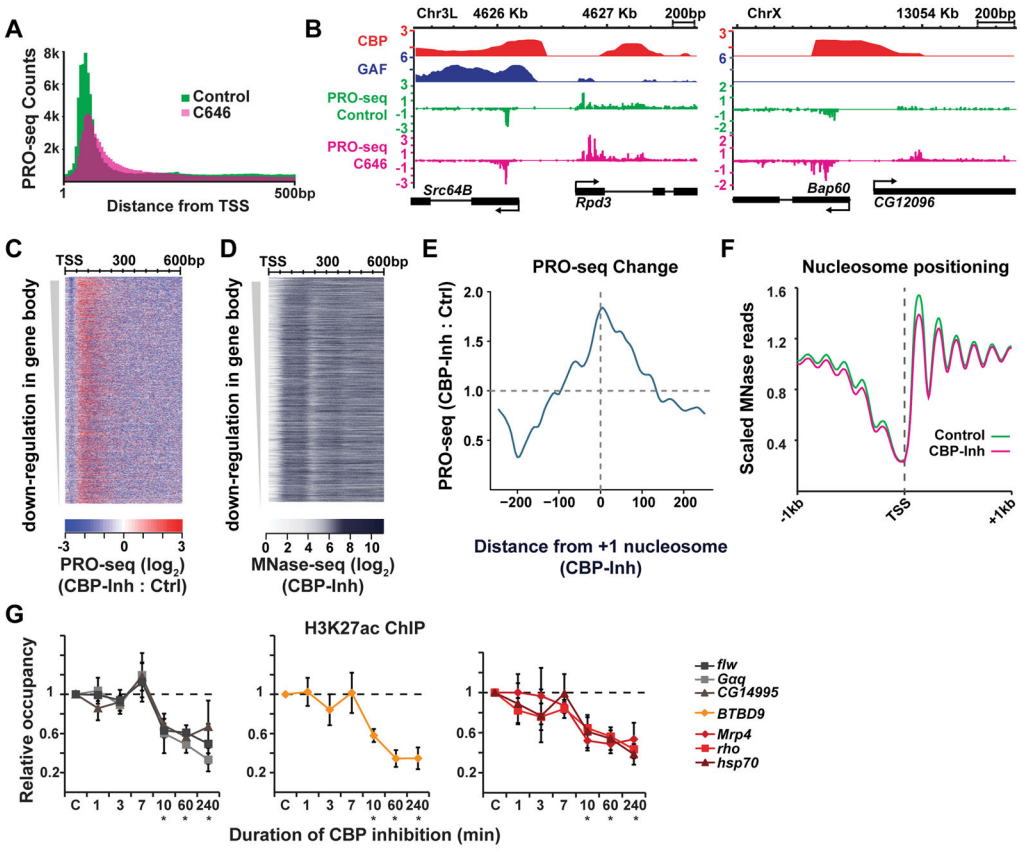

Figure 6. Promoter-proximal Pol II dribbles to more downstream positions after CBP inhibition but is retarded at the +1 nucleosome

(**A**) Metagene profiles of summed PRO-seq counts in all expressed genes, from TSS to 500 bp downstream in 5 bp bins, in control (C37) or CBP-inhibited (C646) S2 cells for 10 min. **(B)** CBP and GAF occupancy in untreated cells, as well as PRO-seq signal in control and C646-treated cells (in $\log_2$ scale) shown for the *Src64B, Rpd3, Bap60 and CG12096* genes. (**C**) Heatmap of $\log_2$ fold change in PRO-seq reads in 5 bp bins from the TSS to +600bp after CBP inhibition sorted on level of down-regulation in gene-body (n=3790). (**D**) MNase-seq after 10 min of CBP inhibition. (**E**) The ratio of PRO-seq change in CBP inhibited versus control cells plotted as distance from the dyad axis of the +1 nucleosome in inhibitor-treated cells. (**F**) Nucleosome position (from MNase-seq data) +/− 1 kb around the TSS of expressed genes (n=5400) in cells treated with control or CBP inhibitor for 10 min. (**G**) H3K27ac ChIP-qPCR after CBP inhibitor treatment for the indicated time points (n=2–4). Occupancy at promoters is plotted relative to control-treated cells (C). Error bars represent S.E.M., two-tailed paired t-test, *indicates $p<0.05$ at the indicated time points for the genes as a group. See also Figure S7.

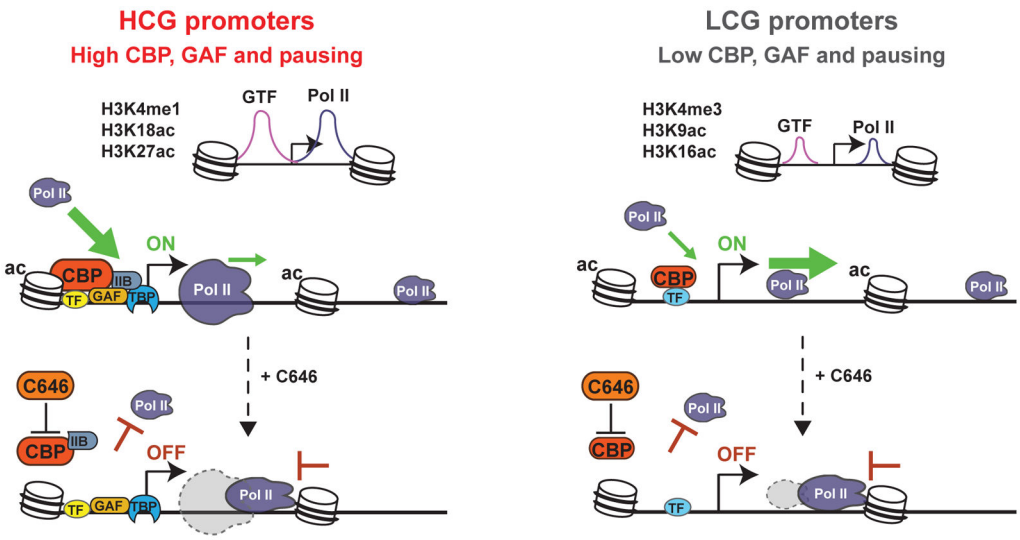

Figure 7. Model showing the functions of CBP on HCG and LCG promoters

(A) Genes with high CBP and GAF enrichment (HCG) are highly paused, enriched in enhancer chromatin marks, and contain high amounts of pre-initiation complex components. Release of Pol II into productive elongation is slow at these genes, and Pol II recruitment is heavily dependent on CBP (top). In CBP-inhibited cells, Pol II is released from the pause site and dribbles further downstream, but accumulates upstream of the +1 nucleosome that is hypo-acetylated. Moreover, CBP and TFIIB do not associate with chromatin, which affects Pol II recruitment, resulting in a decrease in promoter-proximal Pol II (bottom). Thus, CBP is rate-limiting for Pol II recruitment to these genes. (**B**) Low CBP and GAF (LCG) genes contain substantially less CBP, GAF and Pol II than HCG genes, are weakly paused and have promoters enriched in H3K4me3 and H4K16ac. Here, Pol II is efficiently released into elongation (top). After CBP inhibition, Pol II dribbles from the pause site but is retarded at the +1 nucleosome. The resulting accumulation of Pol II in the promoter-proximal region after CBP inhibition suggests that CBP is rate-limiting for release into productive elongation.