

RESEARCH

Open Access



Drosophila architectural proteins M1BP and Opbp cooperatively form the active promoter of a ribosomal protein gene

Igor Osadchiy^{1†}, Anastasia Umnova^{2†}, Galina V. Pokholkova³, Anton Golovnin⁴, Vladimir A. Gvozdev⁵, Igor F. Zhimulev³, Pavel Georgiev^{2*} and Oksana Maksimenko^{1*}

Abstract

Background In *Drosophila*, architectural proteins are frequently found in promoters, including those of genes with extremely high expression levels, such as ribosomal protein genes (RPGs). The involvement of several of these proteins in gene regulation in *Drosophila* has been shown, but the exact mechanisms of their possible cooperative action have not been fully elucidated.

Results In this study we dissected the contribution of the architectural proteins Opbp and M1BP, which are co-localized at several RPG promoters near the transcription start site, to promoter functioning. We found that Opbp has two domains that directly interact with CP190, Putzig (Pzg), and Chromator (Chro) proteins, the cofactors which are required for the activation of housekeeping (*hk*) gene promoters. These domains have redundant functions *in vivo* and can tether the cofactors forming open chromatin regions when are artificially recruited to the “closed” chromatin. Additionally, we observed interactions between M1BP and the same cofactors. In the transgene assay, the transcription driven by the 192-bp part of *Rpl27A* RPG promoter is fully dependent on the presence of at least one Opbp or M1BP binding site and it is sufficient for the very high activity of this promoter integrated into the *hk* gene cluster and moderate expression outside the cluster, while presence of both sites even more facilitates transcription.

Conclusions This study demonstrates that different architectural proteins can work independently and in cooperation and fulfill partially redundant functions in the activation of RPG promoters.

Keywords Architectural C2H2 proteins, Zinc-finger proteins, Putzig, Chromator, CP190, Housekeeping genes, Z4, Chriz, Chro, Ribosomal protein genes

[†]Igor Osadchiy and Anastasia Umnova contributed equally to this work.

*Correspondence:

Pavel Georgiev
georgiev_p@mail.ru
Oksana Maksimenko
maksog@mail.ru

¹Center for Precision Genome Editing and Genetic Technologies for Biomedicine, Institute of Gene Biology, Russian Academy of Sciences, 34/5 Vavilov St., Moscow 119334, Russia

²Department of the Control of Genetic Processes, Institute of Gene Biology, Russian Academy of Sciences, 34/5 Vavilov St., Moscow 119334, Russia

³Laboratory of Molecular Cytogenetics, Institute of Molecular and Cellular Biology SB RAS, Novosibirsk 630090, Russia

⁴Department of *Drosophila* Molecular Genetics, Institute of Gene Biology, Russian Academy of Sciences, 34/5 Vavilov St., Moscow 119334, Russia

⁵NRC “Kurchatov Institute”-Institute of Molecular Genetics, Moscow 123182, Russia



Introduction

The promoter is the central component of the transcriptional apparatus that determine transcription initiation [1, 2]. It consists of various combinations of sequence motifs, the most common of which are the TATA box, initiator (Inr), and downstream core promoter element (DPE) [3, 4]. Several types of core promoters function via different mechanisms and have distinct biological properties [5, 6].

In *Drosophila*, transcription from a substantial fraction of core promoters is driven by the TFIID complex, which consists of the TATA-box binding protein (TBP) and 13–14 TBP-associated factors (TAFs) [4, 7]. TBP and several TAFs bind to specific core promoter motifs, recruit RNA polymerase II to the promoter, and initiate the transcription process [4, 8]. However, the TFIID complex is responsible only for the activity of genes regulated during development and a part of *hk* genes. Other genes are regulated by a homolog of TBP, named TBP-related factor 2 (TRF2) [5, 9]. TRF2 is common among bilateria [10] and lacks DNA-binding activity [9, 11, 12, 13]. Like TBP, TRF2 interacts with the basal transcription factors TFIIA and TFIIB [12, 14].

TRF2 is essential for transcription from *Drosophila* TATA-less promoters that contain either the TCT (polypyrimidine initiator) or DPE motifs [9, 15]. The TCT motif is a rare but biologically important core promoter motif in bilateria found in almost all RPG promoters [16]. In contrast to TBP associated with TAFs, TRF2 has been found in a complex with the remodeling complex NURF, chromatin protein Putzig (Pzg or Z4), and DNA-binding protein DREF [17, 18]. Because TRF2 does not directly bind to DNA, it has been suggested that DREF is responsible for recruiting TRF2 to some promoters that contain the specific DREF binding site [17, 19]. The Pzg protein interacts with NURF and thus can mediate the association of NURF with the TRF2 complex [18, 20, 21]. TRF2, DREF, Pzg, and NURF are predominantly found on promoters that regulate *hk* genes [18, 22].

Recently, it was found that Motif-1 Binding Protein (M1BP) plays a role in recruiting TRF2 to *hk* gene promoters that contain the TCT motif, and the M1BP peaks were detected in 37 out of 62 active RPG promoters [23]. M1BP contains the N-terminal ZAD domain required for homodimerization [24, 25] and the C-terminal cluster consisting of five zinc-finger domains of the C2H2 type (C2H2). M1BP specifically binds to a core promoter element called Motif 1 [26, 27], which has been found in over 2000 *Drosophila* promoters [28].

M1BP is involved in regulating both *hk* [23, 29] and developmental/inducible promoters [30, 31]. It was shown that associated with M1BP paused genes have well-positioned +1 nucleosome that presents a barrier to elongation of RNAP II [28, 30, 32, 33]. M1BP is also

involved in recruiting the homeotic Abd-A and Ubx proteins to promoters repressed by Polycomb Group (PcG) proteins [31], resulting in the reduction of PcG association and the release of paused Pol II. The C2H2 domains of M1BP have a high degree of homology with the C2H2 domains of the human ZKSCAN3 protein, which also binds to gene promoters [33]. As a result, M1BP and ZKSCAN3 bind to the same DNA motif.

Optix binding protein (Opbp) is another protein involved in transcriptional regulation. Opbp specifically binds to nearly 30 sites located exclusively in the promoters of *hk* genes (sites found at ten RPG promoters) [34]. It contains a cluster of five C2H2 domains that allow it to bind with very high specificity to a long consensus sequence. Interestingly, several *hk* gene promoters have motifs for both architectural proteins, Opbp and M1BP, in close proximity to each other and the transcription start site (TSS) (Additional File 1: Fig. S1).

Opbp and M1BP directly interact with the Centrosomal Protein 190 kDa, CP190 [29, 34], which preferentially binds near transcription start sites of active genes and is involved in the organization of open chromatin on promoters [29, 35, 36, 37, 38]. CP190 is involved in the recruitment to chromatin of the nucleosome remodeling factor (NURF), the Spt-Ada-Gcn5-acetyltransferase (SAGA) complex, the dimerization partner, RB-like, E2F, and multi-vulval class B (dREAM) complex, and the histone methyltransferase dMes4 [18, 39, 40, 41, 42]. The chromodomain protein, Chromator(Chro), interacts with Pzg and with JIL-1 histone H3S10 kinase and marks euchromatic interband regions of *Drosophila* polytene chromosomes [43, 44]. The CP190, Pzg, and Chro [40, 45] proteins co-immunoprecipitate from *Drosophila* extracts and are involved in promoter organization of the majority of *hk* genes [22, 35, 40, 46, 47]. The N-terminal region of CP190 is involved in multiple interactions with DNA binding architectural proteins and is sufficient for recruiting Pzg and Chro proteins to regulatory regions and chromatin opening [38]. It seems likely that CP190, Pzg and Chro are the main proteins involved in the recruitment of transcription complexes required for activation of *hk* promoters.

In this study, we demonstrate that M1BP and Opbp directly interact not only with CP190 but also with Pzg and Chro. Both proteins contain two regions at the N- and C-termini that interact with the cofactors. The M1BP and Opbp motifs are co-localized near the TCT initiators in the promoters of four RPGs. We examined the contribution of M1BP and Opbp to the activity of one of these promoters responsible for the expression of the *RpL27A* RPG. As a result, we found that these proteins can work independently and in cooperation since the presence of the motif for at least one of them is absolutely necessary for the promoter functioning and the presence of

both motifs even more potentiate transcription. We also observed a difference in expression depending on the transgene insertion locus, the mechanisms of which are not well understood and imply an interplay among the regulatory sequences in the locus [23, 44].

Results

Opbp and M1BP directly interact with the CP190, Chro, and Pzg proteins

The CP190, Pzg, and Chro proteins are associated with *hk* gene promoters and are essential for their activity [46]. The first goal of the work was to study in detail how the architectural proteins, Opbp and M1BP, can interact with these proteins.

To determine which of the proteins CP190, Pzg and Chro directly interact with Opbp and M1BP and to localize the interacting regions, we used the yeast two-hybrid assay (Y2H). We tested Opbp deleterious variants created based on the amino acid conservation among Drosophilidae and the biochemical properties of Opbp

protein regions (Additional File 1: Fig. S2; Additional File 2: Table S1). Previously, we found that the (30–114) amino acid (aa) region of Opbp is essential for interaction with CP190 [34]. In this assay, we used CP190 (245–524) and determined that there is a second CP190-interacting region outside the (1–114) aa region of Opbp. Next, we examined the interaction of Opbp with Pzg (1–520) and Chro and found that Opbp interacts with both proteins through the (64–105) and (492–562) regions, although Opbp Δ [64–105] Δ [544–562] also showed a significant decrease in the ability to interact with Pzg. Thus, Opbp contains two regions that simultaneously interact with the CP190, Pzg, and Chro proteins. (Fig. 1A and Additional File 1: Fig. S3).

Experimental evidence also shows that M1BP directly interacts with CP190 [29, 38]. In the Y2H assay (Fig. 1B and Additional File 1: Fig. S3) we showed that the (110–195) aa region and C-terminal (305–418) part of M1BP are essential for the interaction. Interestingly, the same regions of M1BP are required for interaction with Pzg,

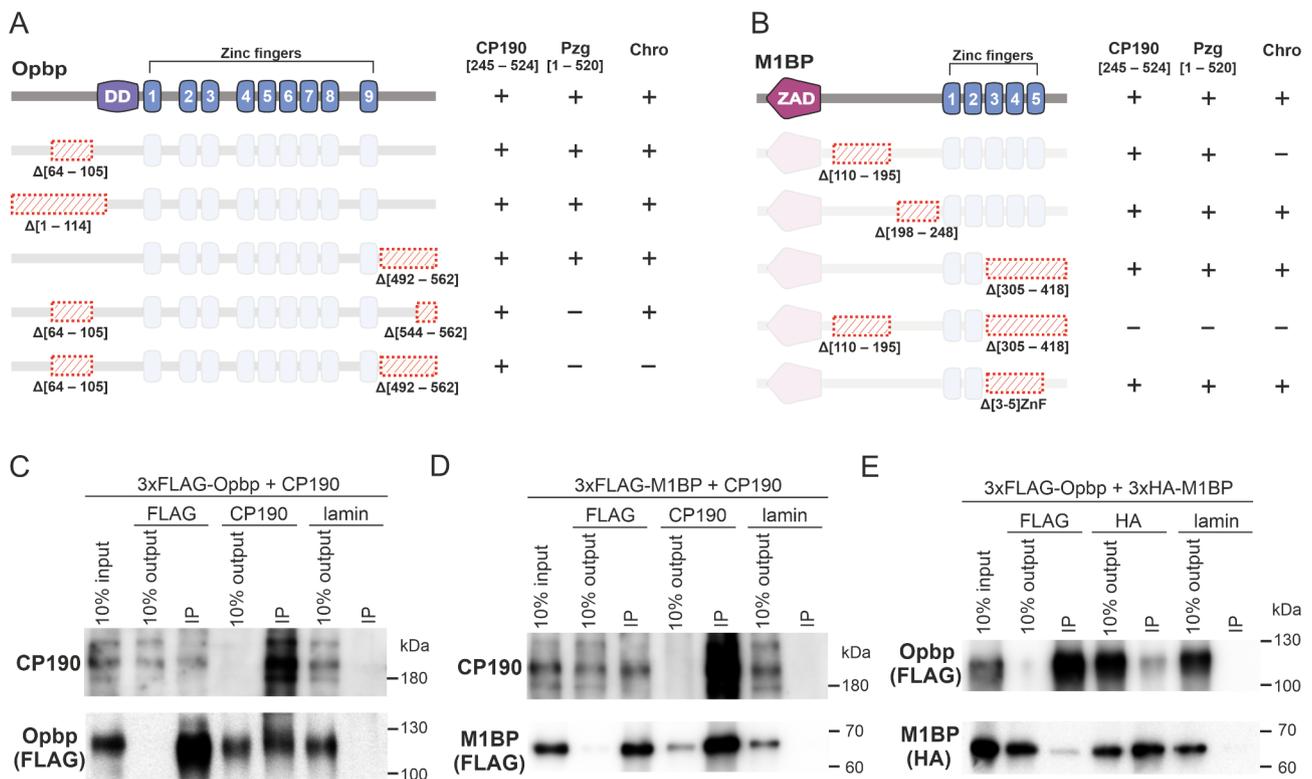


Fig. 1 Interactions of Opbp and M1BP with CP190, Pzg, and Chromator (Chro). **(A, B)** Schematic representations of Opbp **(A)** and M1BP **(B)**. Mapping of regions in Opbp and M1BP that interact with CP190[245–524], Pzg[1–520], and Chro in the yeast two-hybrid assay (Y2H). Protein domains of full-length Opbp **(A)** or M1BP **(B)** are indicated as solid line boxes, and red-dashed rectangles represent different deleted regions (amino acid residues shown below). “DD” - dimerization domain, “Zinc fingers” - C2H2-type zinc-finger domains, “ZAD” - Zinc-finger Associated Domain. The results are summarized in the right columns: “+” - presence and “-” - lack of yeast growth. See Fig. S3 for the yeast plates and the control used. **(C, D)** Confirmation of interactions between Opbp **(C)** and M1BP **(D)** with CP190 in the co-immunoprecipitation assay. Total extracts from *Drosophila* S2 cells co-transfected with 3xFLAG-Opbp **(C)** or 3xFLAG-M1BP **(D)**, and CP190 was immunoprecipitated with antibodies against FLAG epitope. The immunoprecipitates were analyzed by immunoblotting for CP190. **(E)** Interaction between Opbp and M1BP in the co-immunoprecipitation assay. Total extracts from *Drosophila* S2 cells co-transfected with 3xFLAG-Opbp and 3xHA-M1BP. “Input” refers to samples of the initial protein extract; “output” refers to the supernatant after the removal of the immunoprecipitated protein. Lamin is used as a control

while Chro interacts with (110–195) aa region only. Deletion of the C-terminal C2H2 zinc finger domains, fifth (M1BP Δ [5]Zf) or third to fifth (M1BP Δ [3, 4, 5] Zf), did not affect the interaction of M1BP with the CP190, Pzg, and Chro proteins.

The interactions found using the Y2H assay were confirmed in co-immunoprecipitation experiments using full-sized proteins. Opbp and M1BP were fused with the 3xFLAG epitope and expressed in *Drosophila* S2 cells. The FLAG antibodies co-immunoprecipitated Opbp or M1BP and CP190 from the lysates confirming that these proteins can effectively interact with the CP190 in vivo (Fig. 1C and D). We also confirmed interaction of Opbp and M1BP with Chro and Pzg using the lysates of S2 cells co-transfected with 3xFLAG-Opbp or 3xFLAG-M1BP and Chro or V5-Pzg (Additional File 1: Fig. S4).

Finally, we found out the interaction between Opbp and M1BP proteins using the lysate of S2 cells transfected with 3xFLAG-Opbp and 3xHA-M1BP (Fig. 1E).

Testing the ability of Opbp and M1BP to recruit promoter-associated proteins

To further explore the functional similarity of M1BP and Opbp in recruiting the promoter-associated proteins in vivo, we used the previously described model system based on polytene chromosomes of larvae salivary glands [37, 46] (Fig. 2A). In this system, 14 GAL4 binding sites (14xUAS) are inserted into the X chromosome 10A1-2 cytogenetic locus, which corresponds to a band in polytene chromosomes. Such bands are formed by compacted chromatin consisting of genes that are inactive in the salivary glands. In contrast, interbands on polytene chromosomes typically correspond to promoter regions of broadly expressed *hk* genes and have an “open” chromatin conformation [47, 48, 49]. It has been shown that Chro fused to the GAL4 DNA-binding domain can bind to GAL4 sites and induce the formation of an interband that separates the 10A1-2 band into two bands [46].

Here, different regions of the *opbp* gene encoding the N-terminal 1–290 aa (Opbp^N), C-terminal 440–562 aa (Opbp^C), or both regions (Opbp^{N+C}) were fused with the DNA-binding domain of the yeast GAL4 (GAL4 DBD) under the control of the *hsp70* promoter. These constructs were inserted into the *attP* site at 51C cytogenetic locus on the second chromosome using the ϕ C31-based integration system [50]. The 10A1-2 insertion was combined with each construct. To express the chimeric protein, flies were maintained at 29 °C from the embryonic to pupal stages. Earlier, it was shown that the GAL4 DBD alone is recruited to the 10A1-2 region but does not change the polytene chromosome organization and fails to recruit the promoter-associated proteins. Expression of Opbp-derived chimeric proteins resulted in a markedly decondensed zone at the edge of the 10A1-2 region,

which split the band in two (Fig. 2B). Thus, recruiting either the N- or C-terminal part of Opbp is sufficient for interband formation. On polytene chromosomes, the GAL4DBD-Opbp^N and GAL4DBD-Opbp^{N+C} proteins recruit CP190, Pzg, and Chro to the 10A1-2 site. In contrast, the C-terminal part of Opbp recruits only CP190.

Similarly, the ability of M1BP to form the open chromatin and recruit promoter-associated proteins to GAL4 sites was tested in the same model system. To prevent binding to endogenous M1BP sites, we used M1BP transgenes lacking the third to fifth (M1BP Δ [3–5]Zf) or only the fifth (M1BP Δ [5]Zf) C2H2 zinc-finger domain (Fig. 2C). According the Y2H results, deletion of these ZnF domains doesn't affect the interactions of M1BP with CP190, Chro and Pzg (Fig. 1B, Additional File 1: Fig. S3). Unexpectedly, although both chimeric proteins, GAL4DBD-M1BP Δ [3–5]Zf and GAL4DBD-M1BP Δ [5]Zf, bound the 10A1-2 region, neither of them formed the interband or recruited Chro, CP190, or Pzg. Thus, although M1BP interacts with these proteins in coimmunoprecipitation and in the Y2H assay, M1BP could not independently recruit chromatin opening factors in this model system, unlike Opbp.

Model systems for determining roles of the Opbp and M1BP motifs in promoter activity

In total, there are 31 regions specifically bound by Opbp, most of them (30/31) are located in *hk* gene promoters [34], and 16 of the them colocalizing with M1BP peaks (Additional File 1: Fig. S1). Notably, about half of the colocalized peaks display precise point binding near each other with peak summits for both proteins are less than 100 bp apart. The remaining half have a broader and more distributed binding pattern, colocalizing at a distance of ~400 bp on average between their peak summits. To determine roles of Opbp and M1BP in promoter activity, we chose the promoter that controls the *RpL27A* RPG, which is extremely highly expressed in all stages of *Drosophila* development. The 192-bp promoter used in this work contains the M1BP and Opbp motifs, which are located at 3 bp relative to each other and are 2 bp from a potential TCT initiator (Fig. 3A). Interestingly, the Opbp binding site and the TCT initiator are the most conservative sequences of the *RpL27A* promoter (Fig. 3B, Additional File 1: Fig. S5). As the reporter, we used the *firefly luciferase* (*Fluc*) gene and placed it downstream of the *RpL27A* promoter and between the UTRs of the *RpS28b* gene to facilitate its transcription (the *RpS28b* gene has short 5' UTR and contains the only intron in 3' UTR). In addition, we inserted a distinct DNA sequence between the *RpL27A* promoter and the 5' UTR of *RpS28b* for further RT-qPCR and qChIP analyses (Fig. 3A).

RPGs typically reside inside clusters of *hk* genes known as transcriptionally active A compartments on Hi-C

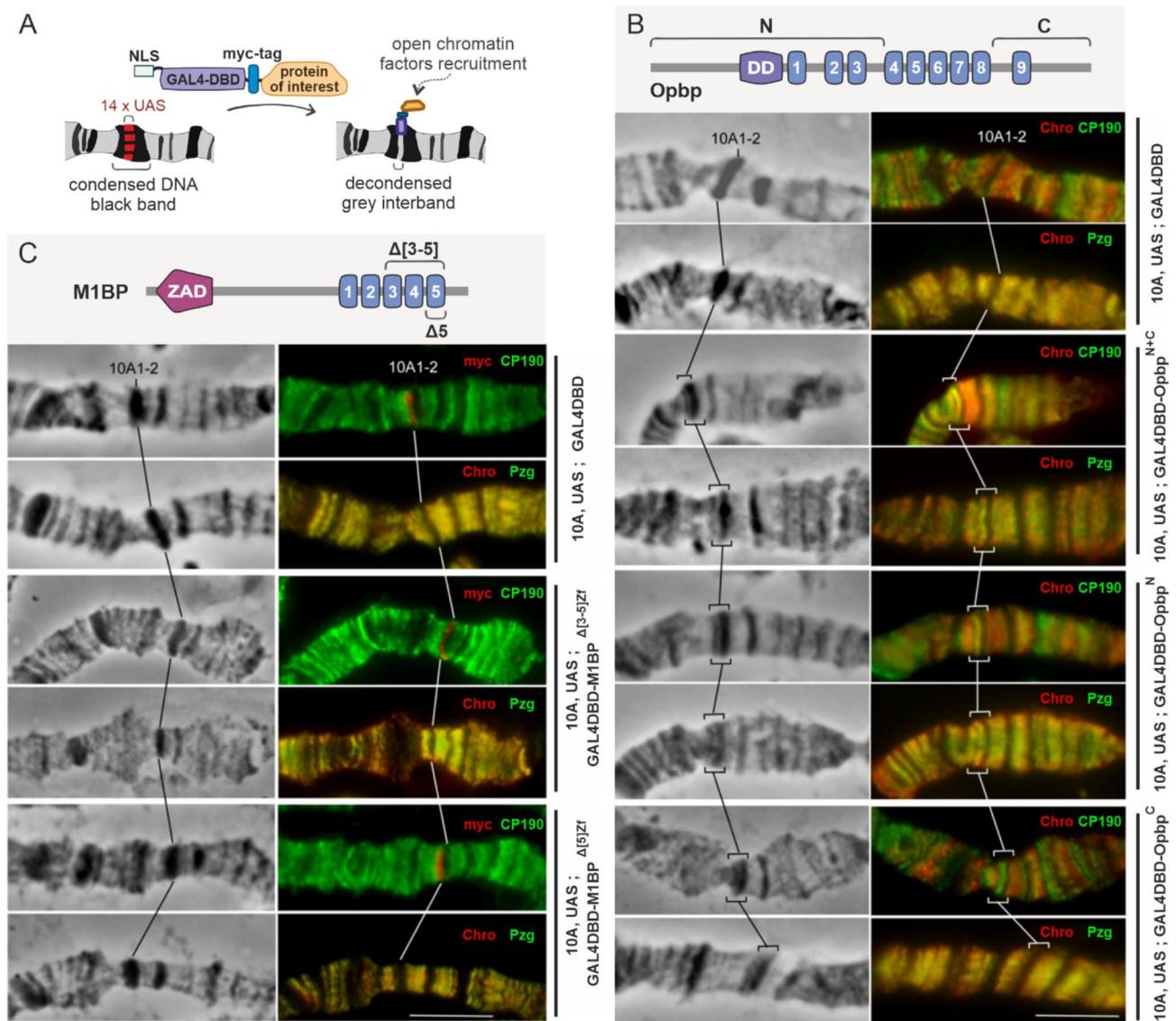


Fig. 2 Determination of the ability of Opbp and M1BP to recruit CP190, Pzg, and Chromator (Chro) proteins. **(A)** Schematic showing the model system used. The system consists of 2 parts. The fly line with an insertion of 14 GAL4 binding sites (14xUAS) into the 10A1-2 region of the X chromosome. And the fly line expressing GAL4 DNA-binding domain (GAL4DBD) fused with nuclear localization signal (NLS), myc-epitope, and protein of interest (domains of Opbp or M1BP variants). After the cross of these lines, the chimeric GAL4DBD-fused protein binds to 14xUAS binding sites at the 10A1-2 region. A band of condensed DNA at the 10A1-2 (black band) is expected to split in two parts separated by an "open" chromatin interband (grey interband) due to local tethering of open chromatin associated factors (CP190, Chro, or Pzg). The abilities of Opbp **(B)** and M1BP **(C)** to recruit CP190, Pzg, and Chromator (Chro) proteins and induce an "open" chromatin structure (an interband) were tested in this model system. **(B)** Scheme representing regions of the Opbp protein (**N** and **C**) used in the constructs (top panel). The left panel shows the polytene chromosomes in phase contrast. The right panel shows an overlay of the phase contrast and immunostaining with antibodies against CP190 (green), Chro (red), and Pzg (green). Thin lines denote the recruitment of CP190, Chro, Pzg, and a novel interband formed at the UAS sites in the 10A1-2 band. The recruitment of the Opbp regions fused with GAL4DBD (GAL4DBD-Opbp^N, GAL4DBD-Opbp^C, and GAL4DBD-Opbp^{N+C}) resulted in interband formation inside the band, and CP190, Chro, and Pzg proteins were detected in the decompacted area. **(C)** Schematic of the deletion variants of the M1BP protein (deletion of 3–5 ($\Delta 3-5$) or 5th ($\Delta 5$) zinc-finger domain) used in the model system (top panel). The left panel shows the polytene chromosomes in phase contrast. The right panel shows an overlay of the phase contrast and immunostaining with antibodies against myc (red), CP190 (green), Chro (red), and Pzg (green). Thin lines denote the recruitment of the myc-tagged GAL4DBD-fused M1BP variants (myc) to the 10A1-2 region. Despite the recruitment to the UAS sites, the GAL4DBD-M1BP variants did not induce the interband formation inside the 10A1-2 band. The absence of CP190, Chro, or Pzg protein recruitment was detected simultaneously with the presence of the signal for myc at the 10A1-2 band

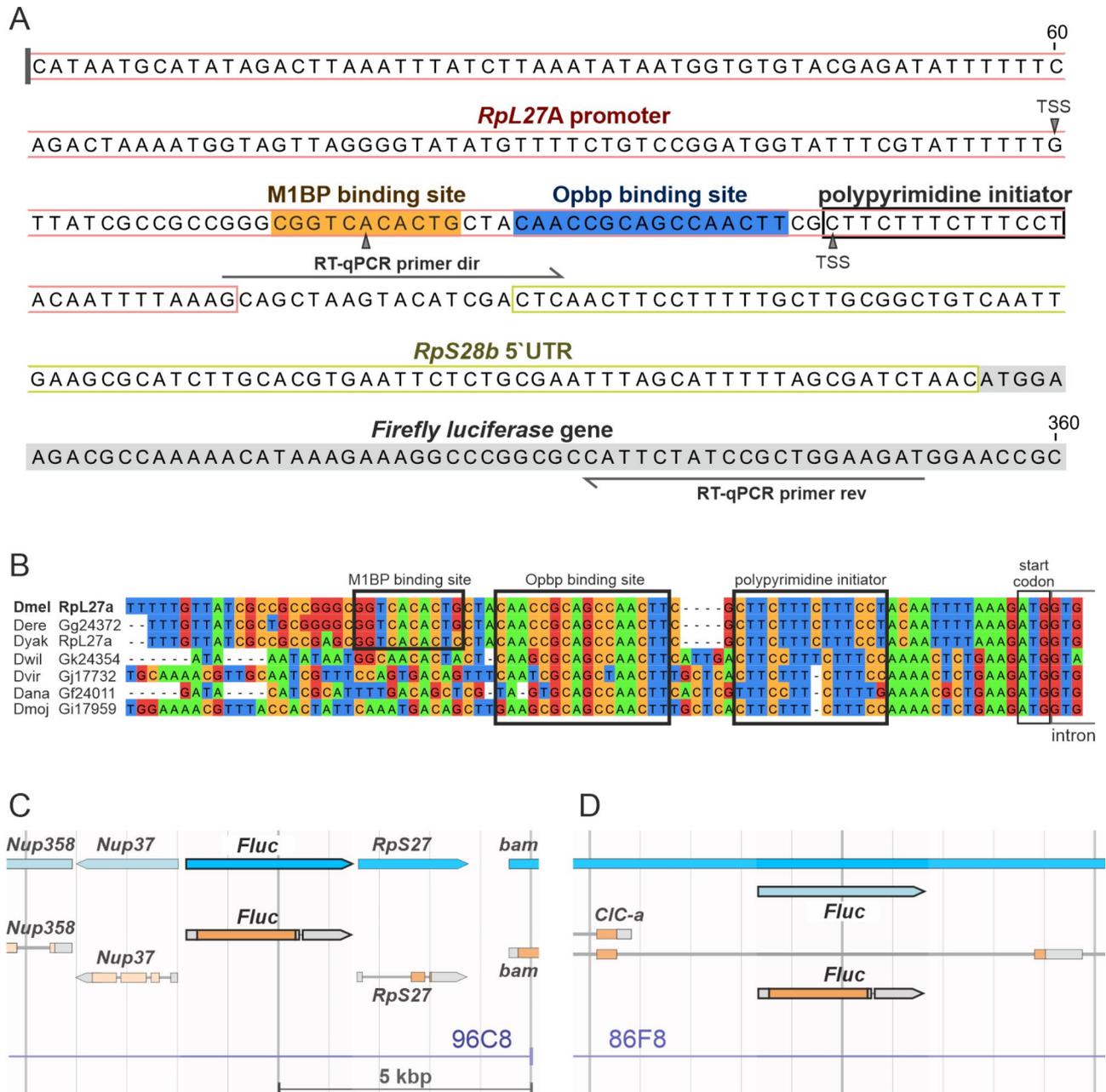


Fig. 3 Model system used to determine the role of cooperation between binding sites for M1BP and Opbp in the organization of ribosomal promoters. **(A)** The sequence of the model *RpL27A* promoter. The figure shows the region of the *RpL27A* promoter used, 5'UTR of the *RpS28b* gene, part of *Firefly luciferase* (*Fluc*) gene coding region, DNA motifs for M1BP and Opbp proteins, a potential TCT (polypyrimidine) initiator, transcription start sites (TSS) according to the known *RpL27A* isoforms, and regions of primers for further quantitative PCR analysis. **(B)** Part of the multiple sequence alignment (MSA) for *RpL27A* promoter sequences from Drosophilidae showing a high degree of conservation of the Opbp binding site, polypyrimidine initiator, and rapid divergence of the M1BP binding site (for the full MSA see Additional File 1: Fig. S5). **(C, D)** Schematic representation of the *Fluc* construct insertion loci, the 96C8 **(C)** and 86F8 **(D)** cytological loci (wider regions of insertion are depicted on Additional File 1: Fig. S6C, D)

maps [51]. According to a current model, transcription of these genes is boosted by the spatial proximity of their regulatory sequences, and thus, the expression of a gene depends on the expression of its neighbors [52]. For this reason, we used two different loci for the insertion of constructs providing *hk* and inducible gene surroundings. For the *hk* surrounding we created an attP landing

platform at the cytogenetic locus 96C8 for integrating the tested transgene between the *Nup37* and *RpS27* genes, located in a head-to-head orientation (Fig. 3C, Additional File 1: Fig. S6). The ribosomal protein *RpS27* gene is very highly expressed and similarly to *RpL27A* contains an Opbp-binding site in its promoter and a high M1BP ChIP-seq signal nearby (Additional File 1: Fig.

S1). Meanwhile, the *Nup37* gene is lowly/moderately expressed depending on the developmental stage. Despite the multiple difference in the gene expression levels, the distance between the annotated TSSs of the *Nup37* and *RpS27* promoters is only 203 bp. The *attP* site, inserted in between the genes using CRISPR/Cas9, was 52 bp away from the TSS of the *Nup37* promoter (Additional File 1: Fig. S6A, B, C, S7A). The integration of the *attP*-containing construct with the *mCherry* reporter gene under the control of the *Actin5C* gene promoter caused lethality in homozygous 96C8_ *attP* flies but the *loxP*-mediated deletion of the reporter restored normal viability.

To assess the role of the inducible gene surrounding in the expression of the transgenic reporter, we also used the landing platform in the genomic region 86F8 [50], which is located within a long intron of the *Chloride channel-a* (*Clc-a*) gene expressed in the stellate cells of Malpighian tubules. This region is associated with tissue-specific genes (Fig. 3D, Additional File 1: Fig. S6D, S7B). Inserting the construct with the model promoter ([MO] *Fluc*) into the 96C8 or 86F8 locus did not affect the viability of the homozygous transgene flies.

Testing the roles of Opbp and M1BP in activation of the model *RpL27A* gene promoter

To determine the contribution of the Opbp and M1BP proteins to activity of the *RpL27A* promoter, we created a series of transgenic lines with corresponding promoter mutations (Fig. 4A, Additional File 1: Fig. S8). These mutations included a mutation of either the Opbp ([MΔ] *Fluc*) or M1BP ([ΔO] *Fluc*) motif or mutation of both motifs ([ΔΔ] *Fluc*). We also deleted the upstream

promoter region ([Δ1-126 MO] *Fluc*) and isolated the shortened promoter from the nearby *Nup37* promoter using the SV40 polyadenylation signal ([SV40_Δ1-126 MO] *Fluc*). All these constructs were inserted at the 96C8 site using the ϕ C31-based integration system [50]. Alternatively, we inserted [MO] *Fluc*, [MΔ] *Fluc*, [ΔO] *Fluc*, and [ΔΔ] *Fluc* into the 86F8 locus. *Fluc* expression was studied in detail in homozygous adults (2–3 days of age) by directly assessing the enzyme activity in extracts for transgenes in both loci (Fig. 4A) and analyzing the amount of RNA by RT-qPCR at the late pupa stage for 98C8-based lines (Fig. 4B). The reporter behaved differently in the tested loci, with *Fluc* chemiluminescence at 86F8 reaching approximately 60% of that at 96C8. Of note, while [MO] *Fluc* was expressed at very high level, none of the tested constructs recapitulated the extremely high expression driven by the endogenous *RpL27A* promoter, maximally reaching nearly 10% (Additional File 1: Fig. S9A).

Then, we performed qChIP analysis of adult transgenic flies (Fig. 5B) to assess Opbp, M1BP and CP190 protein binding to a model promoter and to the control sites (Fig. 5A). The Opbp and M1BP proteins bind to the model *RpL27A* promoter at the 96C8 locus with approximately the same efficiency as to the native promoter and about 2–2.5-fold weaker at the 86F8 locus. Most likely, in the region of the *hk* cluster, Opbp and M1BP have special advantages (increased local concentration of proteins or more open chromatin) compared to the 86F8 locus associated with tissue-specific genes.

At the 96C8 locus, deletion of the M1BP motif ([ΔO] *Fluc*) moderately affects transcription, resulting in a

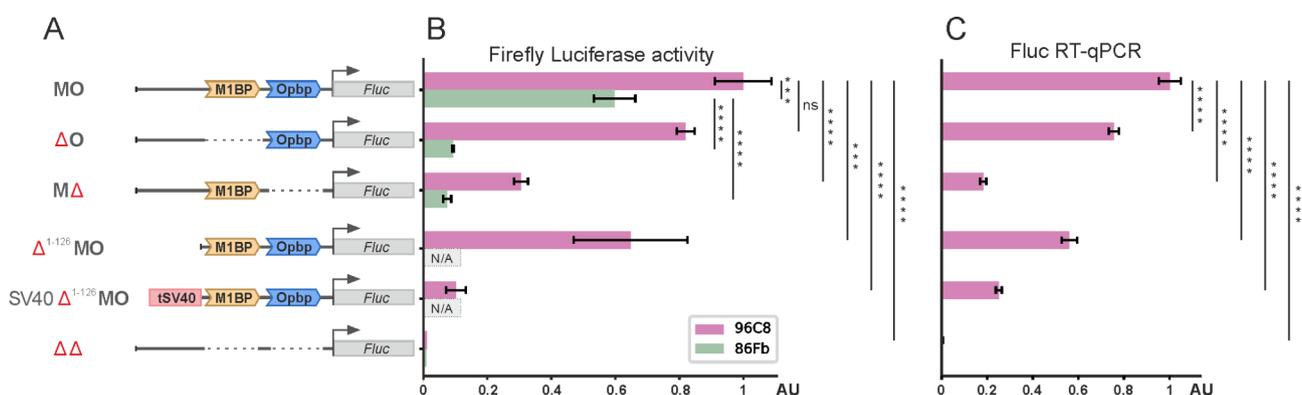


Fig. 4 Opbp and M1BP have similar functions and cooperatively drive transcription (A) Schematic of the promoter variants used to examine the role of the Opbp and M1BP motifs and the preceding 1–126-bp region (on the left). Yellow box - M1BP binding site, blue box - Opbp binding site, pink box - polyadenylation signal of the SV40 virus, gray rectangle with the arrow - reporter *Firefly luciferase* gene (*Fluc*). Deletions of the motifs are shown in dashed lines. (B) The histogram shows the *Fluc* activity in the homozygous transgenic fly lines' extracts. The order along the ordinate corresponds to that in (A). Error bars show the standard deviations of three replicates. Extracts from wild type flies were used as a negative control (data not shown). (C) Changes in expression levels of *Fluc* as measured by RT-qPCR with cDNAs synthesized from RNAs extracted from the transgenic fly lines. The analysis included all lines with construct insertion at the 96C8 genomic region. Transcript levels were determined by RT-qPCR with primers corresponding to *Fluc* and normalized relative to *Opbp* for the amount of input cDNA. The Histogram shows the changes of *Fluc* expression levels relative to expression in [MO] *Fluc* transgenic flies scaled as 1 AU (corresponds to 100%). Error bars show the standard deviations of three PCR measurements. The asterisks indicate the significance of the post-hoc Dunnett's test: *** - p -value < 0.01; **** - p -value < 0.0001; ns - not significant

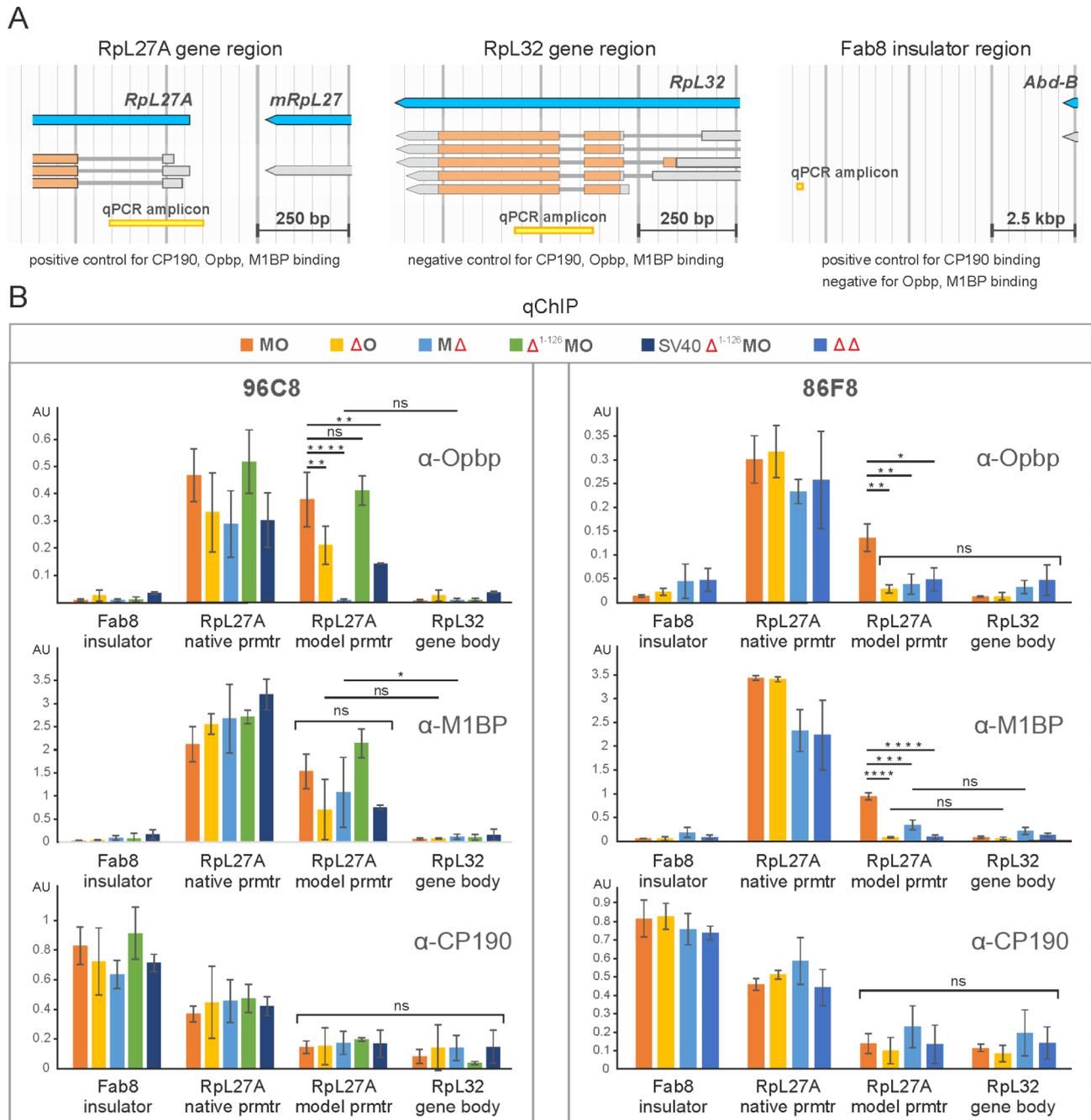


Fig. 5 M1BP and Opbp corroborate binding of each other. **(A)** Genomic regions used for controls in qChIP experiments. **(B)** Histograms show enrichments for Opbp, M1BP, and CP190 at the *RpL27A* promoter on chromatin isolated from transgenic fly lines. The results are presented as a percentage of input genomic DNA normalized on a control site in *RpS21* promoter. Error bars show standard deviations of PCR measurements from three independent experiments. The *RpL27A* gene promoter serves as a positive control for the binding of CP190, Opbp, M1BP. The *RpL32* gene body serves as a negative control for the binding of CP190, Opbp, M1BP. The Fab8 insulator is used as a positive control for the binding of CP190 and a negative control for the binding of Opbp and M1BP. Prmtr stands for "promoter". Error bars show standard deviations of three replicates. The asterisks indicate the significance of the post-hoc Dunnett's test: * - p -value < 0.05; *** - p -value < 0.01; **** - p -value < 0.0001; ns - not significant

roughly 20% decrease, while the deletion of the Opbp motif ([Δ M]Fluc) leads to a 3- and 5-fold decrease (Fig. 4A, B) in the chemiluminescence and RNA amount respectively. These results agree well with the qChIP data (Fig. 5B). Upon deletion of the M1BP site ([Δ O]

Fluc), the M1BP protein still binds to the *RpL27A* promoter but with a reduced efficiency. The Opbp protein also binds to the remaining motif about 2 times less than to the non-mutated promoter. Deletion of the Opbp motif in the model promoter ([Δ M]Fluc) prevents Opbp

binding and leads to a decrease in M1BP binding to its motif. Thus, Opbp and M1BP likely assist each other to more effectively bind to the promoter. This correlates with the observed interactions between proteins in co-immunoprecipitation.

At the 86F8 locus, deletion of either the M1BP or Opbp binding site leads to an even more extreme 6- or 8-fold reduction in *Fluc* expression, respectively (Fig. 4B). Moreover, the binding of both Opbp and M1BP to the model promoters [Δ O] and [Δ M] drops almost to background values, comparable to the deletion of both sites ([$\Delta\Delta$]Fluc line) (Fig. 5B).

As expected, simultaneous deletion of both motifs results in almost complete inactivation (reduction by 2000–3000 times) of the model *RpL27A* promoter, corroborating that Opbp and M1BP play a critical but partially redundant role in promoter activity. To confirm this, we assayed a truncated promoter at the 96C8 locus with deletion of the first 126 bp up to the M1BP binding site. The deletion decreased the *Fluc* reporter signal by approximately 25%, suggesting that the deleted part contributes to expression but is not essential. Interestingly, the insertion of the 260-bp SV40 polyadenylation signal further reduced the expression of the model *RpL27A* promoter by roughly 5 times relative to the intact promoter (Fig. 4A, B). Thus, the activity of the [Δ 1-126_MO] promoter heavily depends on the presence of the *Nup37* promoter nearby, which can compensate for the deleted part. The deletion of the upstream region in the [Δ 1-126_MO]Fluc Fly line did not result in sufficient changes in the binding efficiency of both proteins with the model promoter (Fig. 5B). However, the addition of the SV40

polyadenylation signal upstream of the shortened promoter resulted in an approximately 2-fold decrease in the binding of Opbp and worsening of M1BP binding to the model promoter. It seems likely that proteins bound to the 126-bp upstream sequence or the *Nup37* promoter facilitates the recruitment of M1BP and Opbp to their sites in the *RpL27A* promoter.

All homozygous fly lines with insertions into the 86F8 locus exhibited normal fertility and viability. However, insertions into 96C8 showed variations in these parameters depending on the specific model promoter mutation that was examined. In most cases, flies homozygous for the transgene displayed reduced fertility and increased pupal mortality. The [$\Delta\Delta$]Fluc line was a special case, with nearly 99% pupal mortality and complete sterility in homozygous crosses. Additionally, the homozygous [SV40_ Δ 1-126_MO]Fluc line demonstrated an extremely decreased fertility rate (Fig. 6B, Additional File 2: Table S2). Depending on the tested construct, the changes in viability and fertility rates are perhaps linked to the disturbance of the promoters of adjacent genes *Nup37* and *RpS27* or other genes in the cluster. To test this, we measured the expression of *Nup37*, *RpS27* (Fig. 6A) genes, and also potentially affected *RpL27A* and *RpS28b* used for the model construct, (Additional File 1: Fig. S10) using RT-qPCR. Interestingly, although the transcription rates of *Nup37* and *RpS27* were indeed similarly and correlatedly affected by the model promoter variants, no correlation is evident between the rate of *Fluc* expression and changes in *Nup37* or *RpS27* expression (Additional File 1: Fig. S11). We also observed no correlation between *Nup37* or *RpS27* expression and pupal mortality. *RpL27A*

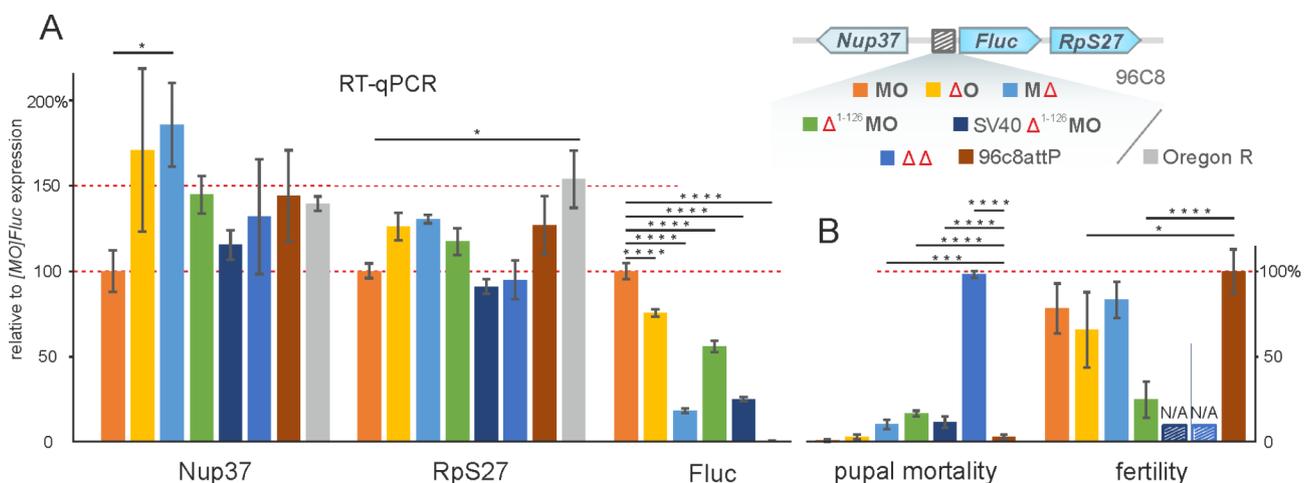


Fig. 6 Insertion of the model *RpL27A* promoter into the 96C8 changes the expression of the adjacent genes and fly viability. **(A)** Changes in the expression levels of *Fluc*, *RpS27* and *Nup37* as measured by RT-qPCR with cDNAs synthesized from RNAs extracted from the transgenic fly lines. The analysis included all lines with construct insertion at the 96C8 genomic region. Transcript levels were determined by RT-qPCR with primers corresponding to *Fluc* and normalized relative to *Opbp* for the amount of input cDNA. The Histogram shows the changes of *Fluc* expression levels relative to expression in [MO] *Fluc* transgenic flies scaled as 100%. Error bars show the standard deviations of three PCR measurements. Wild type flies (Oregon R) are used as a control. **(B)** Pupal mortality percentage was counted as ratio of dead pupae to all pupae multiplied by 100. Fertility rate was counted as a number of flies per vial scaled dividing by mean number of flies for the control line 96c8attP and multiplied by 100. Raw data is available in Additional File 2: Table S2

and *RpS28b* also demonstrated wild type expression levels in all transgenic lines (Additional File 1: Fig. S10).

Because CP190 is an external component of the promoter-bound complex, we tested its recruitment to the model *RpL27A* promoter (Fig. 5B). Unexpectedly, we detected CP190 at the endogenous *RpL27A* promoter but found only extremely reduced binding of CP190 (near the background level) with all variants of the model *RpL27A* promoter in all transgenic lines. The absence of CP190 on the *RpL27A* promoter in transgenes likely explains the relatively low expression of the reporter gene, compared to expression driven by the same promoter located at the native site of the genome.

Discussion

Promoters of *hk* genes are typically short, with 200–400 bp in length on average between their TSSs for genes in head-to-head orientation in *Drosophila* [53] that implies relatively simple organization of these promoters. In this study, we chose 192-bp part of *RpL27A* RPG promoter as a model for exploring the role of architectural proteins, considering very short 5' UTR (~12 bp between TCT initiator and the start codon), the presence of motifs for two architectural proteins immediately next to the TCT initiator and extremely high transcription rate. Although RPG promoters have been extensively studied in different eukaryotic species, a conventional model for their organization is lacking. Most of the *Drosophila* RPGs are expressed at extremely high levels in a coordinated manner to achieve the proper molar ratio [54], their promoters are TRF2-dependent [13] and have the TCT (polypyrimidine initiator) motif [16].

Previous studies have shown that M1BP directly interacts with TRF2 and is involved in recruiting the TRF2 complex to the RPG promoters [23]. In this study, we confirmed the previous finding that M1BP and Opbp directly interacts with CP190 [23, 29, 34], demonstrated that M1BP and Opbp co-immunoprecipitated with each other, interact with the Pzg and Chro proteins and localized two interacting regions in each protein. Interestingly, CP190, Pzg, and Chro also interact with each other and are frequently associated together on *hk* gene promoters [22, 29, 44, 45]. The TRF2 complex also contains Pzg [17]. Altogether, this supports the model in which *hk* promoters lacking core promoter elements are organized by DNA-binding transcription factors (TF) that recruit cofactors TRF2, CP190, Pzg, and Chro with corresponding complexes required for transcription (Fig. 7). Excessive protein interactions and functional redundancy of TFs make this system resistant to mutations that serves as a basis for rapid evolution of promoter sequences [55]. Interestingly, in *D. ananassae* (*Dana*) the M1BP motif in the *RpL27A* promoter has been mutated during the

evolution and is probably functionally replaced by the DREF motif that has emerged nearby (Additional File 1: Fig. S5).

Nevertheless, while mutagenesis of the TF motifs in the model *RpL27A* promoter supports this view, there is some discrepancy between the data obtained. In the *in vivo* polytene chromosomes model system the C-terminus of Opbp recruits only CP190 to the band and neither of the tested M1BP variants recruits any of the cofactor proteins despite the observed interactions in Y2H and CoIP. Several reasons could account for this, most likely it is a consequence of artificiality of this system, which doesn't reproduce the processes of TF binding, isomerization, chemical modification and so on. Since M1BP binds to a variety of *hk* and developmental gene promoters, interacts with many different transcription factors [28, 29, 30, 31] and participates both in activation and pausing, the other reason could be that M1BP by default tethers proteins that mask the surfaces for interaction with CP190, Pzg or Chro. It was shown [38] that, although M1BP frequently colocalizes with CP190 in polytene chromosomes, there are interbands bound by the M1BP protein only. In contrast, Opbp is an activator required for the activity of a small group of TRF2-dependent promoters of *hk* genes, including highly expressed RPGs [34], and so, effectively recruits CP190, Pzg, and Chro by N-terminus inducing formation of open chromatin.

We also observed quite unexpected results in qChIP experiments with our 192-bp *RpL27A* promoter model system (Fig. 5B). Opbp is likely localized on the promoter only through binding to a specific DNA motif. The incomplete disappearance of the signal for M1BP upon deletion of its site indicates that the M1BP protein can be attracted to the promoter region by recognition of additional low-affinity sites or with the aid of other DNA-binding proteins, particularly Opbp (Fig. 1E). Also surprisingly, CP190 does not effectively bind to the 192-bp *RpL27A* promoter (in both 96C8 and 86Fb loci). It was shown [27, 34] that CP190 affects transcription rate of both M1BP- and Opbp-dependent promoters. However, the presence of Opbp and M1BP in the promoter is not sufficient for efficient CP190 recruitment. Taking into account the 10-fold decrease in the transgene expression comparing with the *RpL27A* gene (Additional File 1: Fig. S9), additional regulatory elements are likely required for the effective recruitment of CP190 besides the 192-bp *RpL27A* promoter used in this study. Anyway, we suppose that reduced amount of recruited CP190 could be not detectable due to technical reasons (low resolution of the method) or its fast turnover at the active promoter. Unknown transcription factors are also associated with the upstream 126-bp region and are involved in the stabilization of Opbp and M1BP on the promoter and possibly

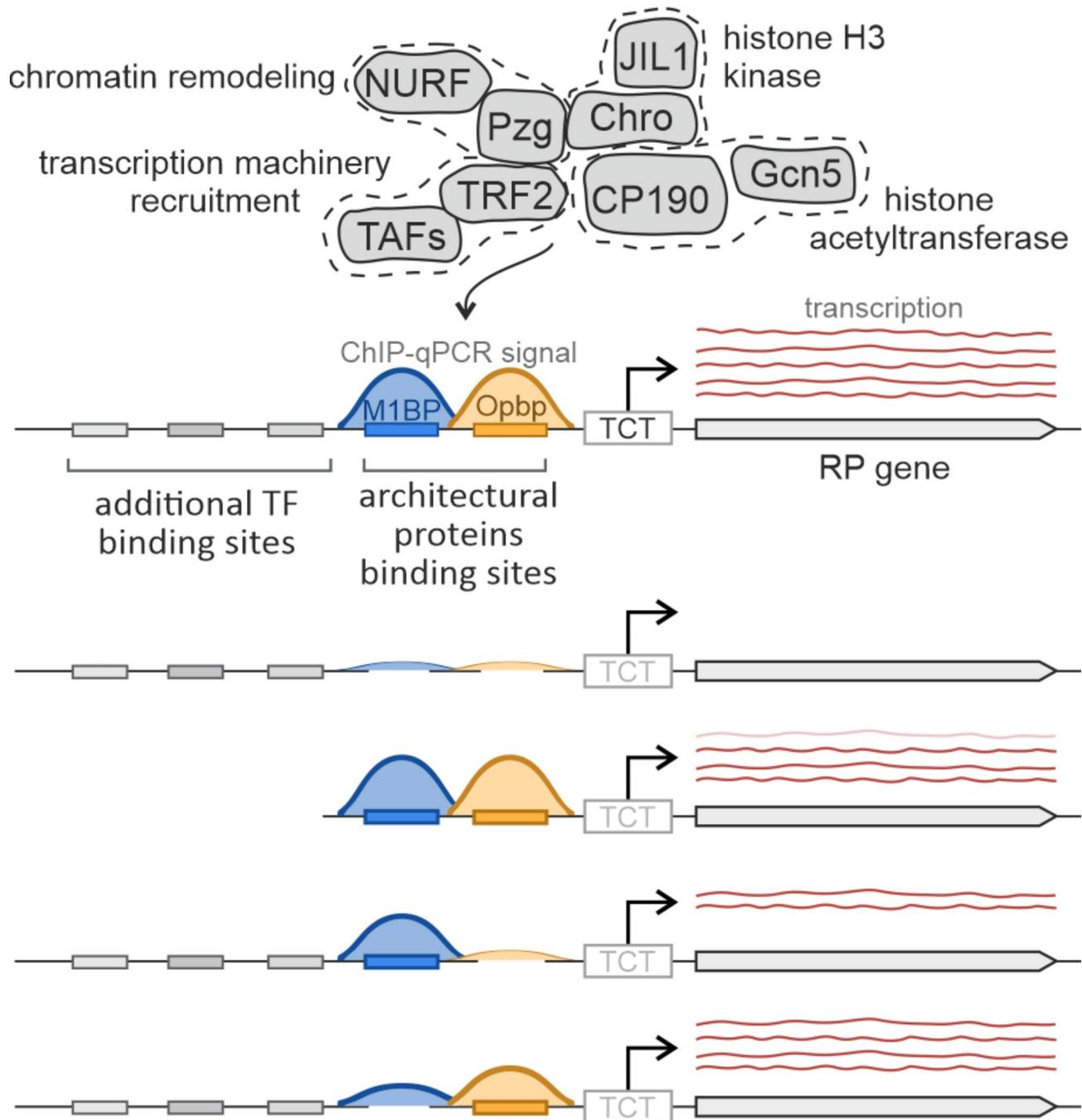


Fig. 7 Model of the ribosomal protein gene promoter organization. Architectural proteins (Opbp and M1BP) recruit the complexes for transcription (TRF2 and TAFs), chromatin remodeling (Pzg and NURF) and histone modification (CP190/Gcn5 and Chro/JIL1) to the core promoter

in the recruitment of transcription complexes. The functions of these sites in the stimulation of the *RpL27A* promoter can be replaced by the neighboring *Nup37* promoter.

Finally, we observed interesting effects depending on the transgene insertion locus. The model *RpL27A* promoter has very high activity in the cluster of *hk* genes in the 96C8 locus nearby the *RpS27* gene with a similar promoter organization. In addition, Opbp and M1BP

binding depends on the genomic environment and occurs more efficiently in the *hk* gene cluster. However, the promoter is also functional and has moderate activity in the 86F8 locus, which is located in the intron of the tissue-specific gene *CIC-a*. In the 96C8 platform, the *RpL27A* model promoter substitutes for *RpS27* and is located in a head-to-head orientation relative to the *Nup37* promoter. The distance between the transcription start sites of the model *RpL27A* and *Nup37* promoters is about

260 bp, which is close to the distance (203 bp) between the TSS of *RpS27* and *Nup37* in the WT. It seems likely that the *Nup37* promoter facilitates binding of the Opbp and M1BP proteins to the motifs in the truncated $\Delta 1$ -126 *RpL27A* promoter. At the same time, various mutations and deletions in the model *RpL27A* promoter, leading to its partial and complete inactivation, did not predictably affect transcription initiated by the *Nup37* and *RpS27* promoters, although the expression of these genes changes correlatedly (Additional File 1: Fig. S11).

There are several possible reasons for these effects. The promoters of *Nup37*, *Fluc* and *RpS27* might form a coordinated system in which *Nup37* and *Fluc* are coregulated due to the short distance between their TSSs, while *Fluc* and *RpS27*, with TSSs 3 kb apart, are likely brought together by architectural proteins M1BP and Opbp. When this is the case, strong expression driven by the model promoter attenuates the expression of neighboring genes (construct [MO]Fluc), probably via depletion of local transcription factors (TFs). Deletion of either M1BP or Opbp binding sites (constructs [M Δ]Fluc, [Δ O]Fluc) decrease *Fluc* transcription rate and the communication between the *Fluc* and *RpS27* promoters, leading to a release of TFs that predominantly activate *Nup37* nearby. On the contrary, the promoters of the [M Δ]Fluc, [Δ O]Fluc or [$\Delta 1$ -126_MO]Fluc constructs could recruit additional TFs and enhance the transcription of colocalized promoters. This could explain a decrease in the transcription of *Nup37* and *RpS27* in [$\Delta 1$ -126_MO]Fluc compared to [M Δ]Fluc, [Δ O]Fluc, and in [SV40_ $\Delta 1$ -126_MO]Fluc compared to [$\Delta 1$ -126_MO]Fluc. Moreover, the region of the attP site insertion is possibly a topology associated domain (TAD) border formed natively by the *Nup37*/*RpS27* promoters (Additional File 1: Fig. S7), and different variants of the model *RpL27A* promoter could affect other genes inside the TAD or more broadly in the gene locus. Disruption of the TAD could potentially explain the loss of fertility and extremely low survival rate for the [$\Delta\Delta$]Fluc construct. Further study is required to understand the interplay between architectural proteins, transcription and TAD organization.

Materials and methods

Fly crosses and generation of the Transgenic lines

Drosophila strains were grown at 25°C under standard culture conditions. The transgenic.

constructs were inserted into the 96C8, 86F8 or 51 C chromosome region using the ϕ C31-mediated site-specific integration system [50].

All constructs were based on an *attB*-contained integration vector with the *white* gene as a transgene marker. The resultant flies were crossed with γ^1w^{1118} flies, and the transgenic progeny were identified by the eye color.

The insertion of the *attP* site in between *Nup37* and *RpS27* genes (96C8attP fly line) was performed by the CRISPR/Cas9 technique (Additional File 1: Fig. S6). We used the fly CRISPR Optimal Target Finder tool (University of Wisconsin) to design a CRISPR target sequence [56]. The sgRNAs was cloned into the pCR vector based on pCFD4-U6:1_U6:3tandemgRNAs plasmid (Addgene#49411), using BbsI. The 5' and 3' flanking regions surrounding the CRISPR/Cas9 target site (homology arms for HDR) were cloned into the plasmid for homologous recombination surrounding the *mCherry* reporter flanked by *loxP* sites. As a source of Cas9 the helper plasmid was used (Addgene#62209). Plasmids mixture (concentration 300ng/ μ L) was injected in the γ^1w^{1118} embryos. Potential genome editing events were detected by mCherry fluorescence.

For Cre-loxP-mediated DNA fragment excision the recombinase-expressing line was used (γ^1w^1 ; *Kr^lf-1*/*Cyo*, *P[Cre w+]**DH1*; *MKRS/TM6B*). The Cre recombinase very efficiently induces excisions in the next generation. All excisions were confirmed by PCR analysis. Details of the crosses used for genetic analysis and the excision of functional elements are available upon request. The primers used are listed in Additional File 2: Table S3.

Plasmid construction

To construct the model system with *Fluc* under the control of *RpL27a* promoter the following fragments were PCR-amplified using the γ^1w^{1118} genomic DNA as a template: a 192-bp region of the *RpL27A* promoter, a 87-bp 5' UTR and 866-bp 3' UTR of *RpS28b*. For *Fluc* the pGL3-basic (Promega) plasmid was used as a template. All fragments were PCR-fused and the resulting 2826-bp model reporter fragment was digested with BamHI and HindIII, and cloned into the vector with *white*, *loxP* and an *attB* site. For the GAL4DBD chimeric constructs used in the polytene chromosomes model system the following fragments were PCR-amplified: Opbp^N (1-290), Opbp^C (440-562), Opbp^{N+C} ($\Delta 298$ –431), M1BP $\Delta [3-5]Zf$, M1BP $\Delta [5]Zf$; they were digested with AscI and NotI, and cloned into the p_attB-NLS-Gal4-DBD-myc vector, previously created for this system.

For constructs used in Y2H, the fragments of Opbp, M1BP, and Pzg were PCR amplified using corresponding cDNA as a template and in-frame cloned into either a GAL4 DNA-binding or activation domain containing vector (pGBT9 or pGAD424 respectively, Clontech). The deletion variants were generated using fusion PCR.

To express 3 \times FLAG-tagged proteins in the S2 cells, protein-coding sequences were subcloned into the pAc5.1 plasmid (Life Technologies, Carlsbad, CA, USA).

The primers used for PCR are listed in Additional File 2: Table S3.

Yeast two-hybrid assay

The yeast two-hybrid assay was performed as previously described [57]. The plasmids were transformed into *Saccharomyces cerevisiae* PJ69-4 A (*MATa trp1-901 leu2-3, 112 ura3-52 his3-200 gal4Δ gal80Δ LYS2::GAL1-HIS3 GAL2-ADE2 met2::GAL7-lacZ*) strain using LiAc/SS-DNA/PEG method [58], followed by plating on the medium lacking tryptophan and leucine. The plates were incubated at 30°C for 3 days and then streaked on selective medium lacking tryptophan, leucine, histidine (“SD-3”), and adenine (“SD-4”) and incubated at 30°C. Cell growth was assessed 3 days later. Each assay was repeated three times.

Co-immunoprecipitation assay

For Co-IP assay expression vectors were co-transfected into *Drosophila* S2 cells plated on 3 cm Petri dishes using Effectene Transfection Reagent (Qiagen) and HyClone SFX-insect medium (Thermo Fisher Scientific, Cytiva) as recommended by the manufacturer. After transfection, cells were incubated for 48 h and then collected by centrifugation at 700 g for 5 min, washed twice with 1×PBS, and resuspended in 300 µL of lysis buffer (10 mM HEPES, pH 7.9; 450 mM NaCl, 5 mM MgCl₂, 0.5% NP-40, 1 mM DTT, Protease Inhibitor Cocktail (Roche) and 1U/mL DNase I); incubated on ice for 45 min, with pipetting up and down three to four times to disrupt cell clumps; centrifuged at 15 000 g, 4 °C, for 15 min; and the supernatant was transferred to a new tube and diluted with four volumes of dilution buffer (10 mM HEPES, pH 7.9; 5 mM MgCl₂, 1 mM DTT, Protease Inhibitor Cocktail). The diluted lysate was again centrifuged at 15 000 g, 4 °C, for 15 min, and the supernatant was transferred to a new tube, with an aliquot (10% of total volume) of it being stored as input control. The lysate was then supplemented with 40 µL of anti-FLAG-conjugated Sepharose (Sigma) and incubated overnight at 4 °C on a rotary shaker. The beads were gently pelleted by centrifugation (700–1000 rpm at 4 °C, ~1 min), an aliquot of the supernatant was stored as output control, and the beads were washed with three portions of IP150 buffer (10 mM Tris-HCl, pH 7.5; 150 mM KCl, 10mM MgCl₂, 1 mM EDTA, 1 mM EGTA, 0.3 mM DTT, 0.1% NP-40, 10% glycerol), 10 min each, with pelleting between washes. The resulting immunoprecipitate was boiled with 1× Laemmli buffer (25 µL per sample) for 10 min, resolved by SDS-PAGE (20 µL per lane), and immunoblotted with appropriate antibodies.

Immunostaining of polytene chromosomes

Salivary glands were dissected from third-instar larvae reared at 29 °C. Polytene chromosome staining was performed as described previously [46]. The following primary antibodies were used: anti-CP190 (1:150),

anti-Chromator (1:600), anti-Pzg (1:15), anti-myc (1:150). 3–4 independent staining, and 4–5 samples of polytene chromosomes were performed with each Ovpb- and M1BP-expressing transgenic line.

Luciferase analysis

Fly lysate was prepared from 2- to 3-day-old adult males and assayed using the Firefly Luciferase Assay Kit (Biotium) following the manufacturer’s instructions. For each replicate, ten adult 2- to 3-day flies were collected, frozen in liquid nitrogen, and ground in 200 µl of Firefly Lysis Buffer (Biotium). The resulting lysate was 2 centrifuged at 2000 g for 5 min, and 20 µl of clear, fat-free middle phase was assayed using the Firefly Luciferase Assay Kit (Biotium) following the manufacturer’s instructions. Luciferase activity was estimated with a Clariostar reader (BMG, Germany). Analysis was performed in three independent biological replicates. The significance of changes in the expression level was estimated by two-sided independent Student’s t-test without equal variance assumption.

RT-PCR

Total RNA was isolated using the TRI reagent (Molecular Research Center, United States) according to the manufacturer’s instructions. RNA was treated with two units of Turbo DNase I (Ambion) for 30 min at 37 °C to eliminate genomic DNA. The synthesis of cDNA was performed using 2 µg of RNA, 200 U of RevertAid reverse transcriptase (Thermo Fisher Scientific, Fermentas), and 1 µM of oligo(dT) as a primer. The amounts of specific cDNA fragments were quantified by real-time PCR. At least three independent measurements were made for each RNA sample. Relative levels of mRNA expression were calculated in the linear amplification range by calibration to a standard genomic DNA curve to account for differences in primer efficiencies. Individual expression values were normalized with reference to *RpL32* and *ovpb* mRNA. Pearson correlation was used for statistical analysis of data for RT-qPCR analysis of *Nup37*, *RpS27* and *Fluc* expression. Data are expressed as mean ± SD. The primers used are listed in Additional File [lup2](#): Table [S3](#).

ChIP-qPCR

Chromatin was prepared from two- to three-day-old adult flies. A 1 g of adult flies was ground in a mortar in liquid nitrogen and resuspended in 20 mL of buffer A (15 mM HEPES-KOH, pH 7.6, 60 mM KCl, 15 mM NaCl, 13 mM EDTA, 0.1 mM EGTA, 0.15 mM spermine, 0.5 mM spermidine, 0.5% NP-40, 0.5 mM DTT) supplemented with 0.5 mM PMSF and Calbiochem Cocktail V. The suspension was then homogenized subsequently in a Potter and Dounce homogenizer with tight pestle, filtered through 100 µm Nylon Cell Strainer (Miltenyi

Biotec, United States), and cross-linked with 1% formaldehyde for 15 min at room temperature. Cross-linking was stopped by adding glycine to a final concentration of 125 mM. The nuclei were washed with three 10-mL portions of wash buffer (15 mM HEPES-KOH, pH 7.6, 60 mM KCl, 15 mM NaCl, 1 mM EDTA, 0.1 mM EGTA, 0.1% NP-40, protease inhibitors), one 5-mL portion of nuclear lysis basic buffer (15 mM HEPES, pH 7.6, 140 mM NaCl, 1 mM EDTA, 0.1 mM EGTA, 1% Triton X-100, 0.5 mM DTT, 0.1% sodium deoxycholate, protease inhibitors) and resuspended in 1 mL of nuclear lysis buffer (15 mM HEPES, pH 7.6, 140 mM NaCl, 1 mM EDTA, 0.1 mM EGTA, 1% Triton X-100, 0.5 mM DTT, 0.1% sodium deoxycholate, 0.5% SLS, 0.1% SDS, protease inhibitors). The suspension was sonicated in a Covaris ME220 focused-ultrasonicator (30 alternating 15-s ON and 45-s OFF intervals, peak power 75, duty % factor 25), and 50- μ L aliquots were used to test extent of sonication and to measure DNA concentration. Debris was removed by centrifugation at 14 000 g, 4 °C, for 10 min, and chromatin was pre-cleared with Protein A/G Magnetic beads (NEB). Corresponding antibodies were incubated for 1 h at room temperature with 20 μ L aliquots of Protein A (anti-Opbp(1:80), anti-M1BP(1:100) or G (anti-CP190 (1:40)) Magnetic beads (NEB) mixed with 200 μ L of PBST. Then antibodies-beads complexes were washed and equilibrated in nuclear lysis buffer. Chromatin samples containing 10–20 μ g of DNA equivalent in 200 μ L of nuclear lysis buffer (2 μ L aliquots of such pre-cleared chromatin being stored as input material) were incubated overnight, at 4 °C, with antibodies-beads complexes. After 3 rounds of washing with lysis buffer supplemented with 500 mM NaCl, and TE buffer (10 mM Tris-HCl, pH 8; 1 mM EDTA), the DNA was eluted with elution buffer (50 mM Tris-HCl, pH 8.0; 1 mM EDTA, 1% SDS), the cross-links were reversed, and the precipitated DNA was phenol-chloroform extracted with the PhaseLock Gel (VMR).

The enrichment of specific DNA fragments was analyzed by qPCR, using a StepOne Plus Thermal Cycler (Applied Biosystems). The primers used are listed in Additional File 2: Table S3.

Data analysis

Data analysis was performed using GraphPad Prism and Python. The one-way ANOVA (with Dunnett post-hoc) and Kruskal–Wallis (with Dunn post-hoc) test were used to determine statistically significant differences between the medians of independent groups depending on distribution and variances. Mann–Whitney test and t-test were used to determine statistically significant differences between two groups. Correlation was defined by Pearson correlation coefficients.

Antibodies

Antibodies	SOURCE	IDENTIFIER
Mouse anti-Myc	Invitrogen	46–0307
Mouse anti-Pzg	Prof H. Saumweber Miao Gan et al. 2011 https://doi.org/10.1007/s12038-011-9089-y	N/A
Rabbit anti-CP190	This lab	N/A
Rat anti-CP190	This lab	N/A
anti-Chro	A.A. Gorchakov	N/A
Rabbit anti-M1BP	This lab	N/A
Rat anti-Opbp	This lab, 2017 Zolotarev et al. https://doi.org/10.1093/nar/gkx840	N/A

Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s13072-025-00584-8>.

Supplementary Material 1
Supplementary Material 2
Supplementary Material 3
Supplementary Material 4

Acknowledgements

We are grateful to Farhod Hasanov for fly injections.

Author contributions

I.O., A.U., P.G., O.M. conceived and designed the project. I.O., A.G., G.V.P., A.U., O.M. performed all the experiments. V.A.G., I.F.Z., P.G., O.M., I.O., A.U. analyzed data. I.O., P.G., O.M. wrote the manuscript. The manuscript was edited and approved with contributions from all authors. All authors read and approved the final manuscript.

Funding

This work was supported by grants from the Russian Science Foundation № 19-74-30026-P.

Data availability

No datasets were generated or analysed during the current study.

Declarations

Ethics approval and consent to participate

Experiments involving laboratory animals were approved by the Human and Animal Ethics Committees of the Institute of Gene Biology (Moscow).

Consent for publication

All authors consent to publication.

Competing interests

The authors declare no competing interests.

Received: 30 November 2024 / Accepted: 5 April 2025

Published online: 16 April 2025

References

- Andersson R, Sandelin A. Determinants of enhancer and promoter activities of regulatory elements. *Nat Rev Genet.* 2020;21:71–87.
- Archuleta SR, Goodrich JA, Kugel JF. Mechanisms and functions of the RNA polymerase II general transcription machinery during the transcription cycle. *Biomolecules.* 2024;14:176.

3. Haberer V, Stark A. Eukaryotic core promoters and the functional basis of transcription initiation. *Nat Rev Mol Cell Biol.* 2018;19:621–37.
4. Vo Ngoc L, Kassavetis GA, Kadonaga JT. The RNA polymerase II core promoter in *Drosophila*. *Genetics.* 2019;212:13–24.
5. Mishal R, Luna-Arias JP. Role of the TATA-box binding protein (TBP) and associated family members in transcription regulation. *Gene.* 2022;833:146581.
6. Serebreni L, Stark A. Insights into gene regulation: from regulatory genomic elements to DNA-protein and protein-protein interactions. *Curr Opin Cell Biol.* 2021;70:58–66.
7. Chen H, Pugh BF. What do transcription factors interact with?? *J Mol Biol.* 2021;433:166883.
8. Chen X, Xu Y. Structural insights into assembly of transcription preinitiation complex. *Curr Opin Struct Biol.* 2022;75:102404.
9. Wang Y-L, Duttke SHC, Chen K, Johnston J, Kassavetis GA, Zeitlinger J, et al. TRF2, but not TBP, mediates the transcription of ribosomal protein genes. *Genes Dev.* 2014;28:1550–5.
10. Duttke SHC, Doolittle RF, Wang Y-L, Kadonaga JT. TRF2 and the evolution of the bilateria. *Genes Dev.* 2014;28:2071–6.
11. Dantanel J-C, Wurtz J-M, Poch O, Moras D, Tora L. The TBP-like factor: an alternative transcription factor in metazoa?? *Trends Biochem Sci.* 1999;24:335–9.
12. Rabenstein MD, Zhou S, Lis JT, Tjian R. TATA box-binding protein (TBP)-related factor 2 (TRF2), a third member of the TBP family. *Proc Natl Acad Sci USA.* 1999;96:4791–6.
13. Zehavi Y, Kedmi A, Ideses D, Juven-Gershon T. TRF2: transforming the view of general transcription factors. *Transcription.* 2015;6:1–6.
14. Teichmann M, Wang Z, Martinez E, Tjernberg A, Zhang D, Vollmer F, et al. Human TATA-binding protein-related factor-2 (hTRF2) stably associates with hTFIIA in HeLa cells. *Proc Natl Acad Sci USA.* 1999;96:13720–5.
15. Kedmi A, Zehavi Y, Glick Y, Orenstein Y, Ideses D, Wachtel C, et al. *Drosophila* TRF2 is a Preferential core promoter regulator. *Genes Dev.* 2014;28:2163–74.
16. Parry TJ, Theisen JWM, Hsu J-Y, Wang Y-L, Corcoran DL, Eustice M, et al. The TCT motif, a key component of an RNA polymerase II transcription system for the translational machinery. *Genes Dev.* 2010;24:2013–8.
17. Hochheimer A, Zhou S, Zheng S, Holmes MC, Tjian R. TRF2 associates with DREF and directs promoter-selective gene expression in *drosophila*. *Nature.* 2002;420:439–45.
18. Kwon SY, Grisan V, Jiang B, Herbert J, Badenhorst P. Genome-Wide mapping targets of the metazoan chromatin remodeling factor NURF reveals nucleosome remodeling at enhancers, core promoters and gene insulators. *PLoS Genet.* 2016;12:e1005969.
19. Zabidi MA, Arnold CD, Schernhuber K, Pagani M, Rath M, Frank O, et al. Enhancer-core-promoter specificity separates developmental and house-keeping gene regulation. *Nature.* 2015;518:556–9.
20. Kugler SJ, Gehring E-M, Walkkamm V, Krüger V, Nagel AC. The Putzig-NURF nucleosome remodeling complex is required for ecdysone receptor signaling and innate immunity in *Drosophila melanogaster*. *Genetics.* 2011;188:127–39.
21. Kugler SJ, Nagel AC. A Novel Pzq-NURF Complex Regulates Notch Target Gene Activity. Gonzalez-Gaitan M, editor. *MBoC.* 2010;21:3443–8.
22. Cubenas-Potts C, Rowley MJ, Lyu X, Li G, Lei EP, Corces VG. Different enhancer classes in *drosophila* bind distinct architectural proteins and mediate unique chromatin interactions and 3D architecture. *Nucleic Acids Res.* 2017;45:1714–30.
23. Baumann DG, Gilmour DS. A sequence-specific core promoter-binding transcription factor recruits TRF2 to coordinately transcribe ribosomal protein genes. *Nucleic Acids Res.* 2017;45:10481–91.
24. Bonchuk A, Boyko K, Fedotova A, Nikolaeva A, Lushchekina S, Khrustaleva A, et al. Structural basis of diversity and homodimerization specificity of zinc-finger-associated domains in *drosophila*. *Nucleic Acids Res.* 2021;49:2375–89.
25. Bonchuk AN, Boyko KM, Nikolaeva AY, Burtseva AD, Popov VO, Georgiev PG. Structural insights into highly similar Spatial organization of zinc-finger associated domains with a very low sequence similarity. *Structure.* 2022;30:1004–e10154.
26. Bailey TL, Boden M, Buske FA, Frith M, Grant CE, Clementi L, et al. MEME SUITE: tools for motif discovery and searching. *Nucleic Acids Res.* 2009;37:W202–8.
27. Ohler U, Liao G, Niemann H, Rubin GM. Computational analysis of core promoters in the *drosophila* genome. *Genome Biol.* 2002;3:research00871.
28. Li J, Gilmour DS. Distinct mechanisms of transcriptional pausing orchestrated by GAGA factor and M1BP, a novel transcription factor. *EMBO J.* 2013;32:1829–41.
29. Bag I, Chen S, Rosin LF, Chen Y, Liu C-Y, Yu G-Y, et al. M1BP cooperates with CP190 to activate transcription at TAD borders and promote chromatin insulator activity. *Nat Commun.* 2021;12:4170.
30. Raj A, Chimata AV, Singh A. Motif 1 binding protein suppresses wingless to promote eye fate in *drosophila*. *Sci Rep.* 2020;10:17221.
31. Zouaz A, Auradkar A, Delfini MC, Macchi M, Barthez M, Ela Akoa S, et al. The hox proteins Ubx and AbdA collaborate with the transcription pausing factor M1 BP to regulate gene transcription. *EMBO J.* 2017;36:2887–906.
32. Fuda NJ, Guertin MJ, Sharma S, Danko CG, Martins AL, Siepel A et al. GAGA Factor Maintains Nucleosome-Free Regions and Has a Role in RNA Polymerase II Recruitment to Promoters. Lieb JD, editor. *PLoS Genet.* 2015;11:e1005108.
33. Barthez M, Poplineau M, Elrefaey M, Caruso N, Graba Y, Saurin AJ. Human ZKSCAN3 and *drosophila* M1BP are functionally homologous transcription factors in autophagy regulation. *Sci Rep.* 2020;10:9653.
34. Zolotarev N, Maksimenko O, Kyrchanova O, Sokolinskaya E, Osadchiy I, Girardot C, et al. Opbp is a new architectural/insulator protein required for ribosomal gene expression. *Nucleic Acids Res.* 2017;45:12285–300.
35. Bartkuhn M, Straub T, Herold M, Herrmann M, Rathke C, Saumweber H, et al. Active promoters and insulators are marked by the centrosomal protein 190. *EMBO J.* 2009;28:877–88.
36. Kaushal A, Dorier J, Wang B, Mohana G, Taschner M, Cousin P, et al. Essential role of Cp190 in physical and regulatory boundary formation. *Sci Adv.* 2022;8:eabl8834.
37. Sabirov M, Kyrchanova O, Pokholkova GV, Bonchuk A, Klimenko N, Belova E, et al. Mechanism and functional role of the interaction between CP190 and the architectural protein Pita in *drosophila melanogaster*. *Epigenetics Chromatin.* 2021;14:16.
38. Golovnin A, Melnikova L, Babosha V, Pokholkova GV, Slovohtov I, Umnova A, et al. The N-Terminal part of *drosophila* CP190 is a platform for interaction with multiple architectural proteins. *IJMS.* 2023;24:15917.
39. Ali T, Kruger M, Bhuju S, Jarek M, Bartkuhn M, Renkawitz R. Chromatin binding of Gcn5 in *drosophila* is largely mediated by CP190. *Nucleic Acids Res.* 2017;45:2384–95.
40. Bohla D, Herold M, Panzer I, Buxa MK, Ali T, Demmers J, et al. A functional insulator screen identifies NURF and dREAM components to be required for enhancer-blocking. *PLoS ONE.* 2014;9:e107765.
41. Korenjak M, Kwon E, Morris RT, Anderssen E, Amzallag A, Ramaswamy S, et al. dREAM co-operates with insulator-binding proteins and regulates expression at divergently paired genes. *Nucleic Acids Res.* 2014;42:8939–53.
42. Lhoumaud P, Hennion M, Gamot A, Cuddapah S, Queille S, Liang J, et al. Insulators recruit histone methyltransferase dMse4 to regulate chromatin of flanking genes. *EMBO J.* 2014;33:1599–613.
43. Rath U, Ding Y, Deng H, Qi H, Bao X, Zhang W, et al. The chromodomain protein, chromator, interacts with JIL-1 kinase and regulates the structure of *Drosophila* polytene chromosomes. *J Cell Sci.* 2006;119:2332–41.
44. Gan M, Moebus S, Eggert H, Saumweber H. The Chriz–Z4 complex recruits JIL-1 to polytene chromosomes, a requirement for interband-specific phosphorylation of H3S10. *J Biosci.* 2011;36:425–38.
45. Melnikova LS, Kostyuchenko MV, Georgiev PG, Golovnin AK. The Chriz protein promotes the recruitment of the Z4 protein to the STAT-Dependent promoters. *Dokl Biochem Biophys.* 2020;490:29–33.
46. Pokholkova GV, Demakov SA, Andreenkov OV, Andreenkova NG, Volkova EI, Belyaeva ES, et al. Tethering of CHROMATOR and dCTCF proteins results in decompaction of condensed bands in the *drosophila melanogaster* polytene chromosomes but does not affect their transcription and replication timing. *PLoS ONE.* 2018;13:e0192634.
47. Zhimulev I, Vatolina T, Levitsky V, Tsukanov A. Developmental and housekeeping genes: two types of genetic organization in the *drosophila* genome. *IJMS.* 2024;25:4068.
48. Demakov SA, Vatolina TY, Babenko VN, Semeshin VF, Belyaeva ES, Zhimulev IF. Protein composition of interband regions in polytene and cell line chromosomes of *drosophila melanogaster*. *BMC Genomics.* 2011;12:566.
49. Vatolina TY, Boldyreva LV, Demakova OV, Demakov SA, Kokoza EB, Semeshin VF, et al. Identical functional organization of Nonpolytene and polytene chromosomes in *drosophila melanogaster*. *PLoS ONE.* 2011;6:e25960.
50. Bischof J, Maeda RK, Hediger M, Karch F, Basler K. An optimized transgenesis system for *drosophila* using germ-line-specific phiC31 integrases. *Proc Natl Acad Sci USA.* 2007;104:3312–7.
51. Lieberman-Aiden E, van Berkum NL, Williams L, Imakaev M, Ragozcy T, Telling A, et al. Comprehensive mapping of long-range interactions reveals folding principles of the human genome. *Science.* 2009;326:289–93.
52. Corrales M, Rosado A, Cortini R, Van Arensbergen J, Van Steensel B, Filion GJ. Clustering of *Drosophila* housekeeping promoters facilitates their expression. *Genome Res.* 2017;27:1153–61.

53. Yang L, Yu J. A comparative analysis of divergently-paired genes (DPGs) among drosophila and vertebrate genomes. *BMC Evol Biol.* 2009;9:55.
54. Petibon C, Malik Ghulam M, Catala M, Abou Elela S. Regulation of ribosomal protein genes: An ordered anarchy. *WIREs RNA.* 2021 [cited 2023 Jul 18];12. Available from: <https://onlinelibrary.wiley.com/doi/https://doi.org/10.1002/wrna.1632>
55. Ma X, Zhang K, Li X. Evolution of drosophila ribosomal protein gene core promoters. *Gene.* 2009;432:54–9.
56. Gratz SJ, Ukken FP, Rubinstein CD, Thiede G, Donohue LK, Cummings AM, et al. Highly specific and efficient CRISPR/Cas9-catalyzed homology-directed repair in drosophila. *Genetics.* 2014;196:961–71.
57. Fields S, Song O. A novel genetic system to detect protein-protein interactions. *Nature.* 1989;340:245–6.
58. Daniel Gietz R, Woods RA. Transformation of yeast by lithium acetate/single-stranded carrier DNA/polyethylene glycol method. *Methods in Enzymology.* Elsevier; 2002 [cited 2023 Aug 8]. pp. 87–96. Available from: <https://linkinghub.elsevier.com/retrieve/pii/S0076687902509575>

Publisher's note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.