

The Central Region of the *Drosophila* Co-repressor Groucho as a Regulatory Hub^{*S}

Received for publication, July 28, 2015, and in revised form, October 12, 2015. Published, JBC Papers in Press, October 19, 2015, DOI 10.1074/jbc.M115.681171

Pak N. Kwong[‡], Michael Chambers[‡], Ajay A. Vashisht[§], Wiam Turki-Judeh^{‡¶}, Tak Yu Yau[‡], James A. Wohlschlegel^{§¶}, and Albert J. Courey^{‡¶1}

From the Departments of [‡]Chemistry and Biochemistry and [§]Biological Chemistry and [¶]Molecular Biology Institute, UCLA, Los Angeles, California 90095

Background: The co-repressor Groucho has an essential, but disordered, central region.

Results: We identified over 160 central region-binding proteins, many of which, including components of the spliceosome, modulate Groucho-mediated repression.

Conclusion: Groucho regulates transcription by multiple mechanisms and may link the transcriptional and splicing machineries.

Significance: Its central region may serve as the hub of a regulatory network.

Groucho (Gro) is a *Drosophila* co-repressor that regulates the expression of a large number of genes, many of which are involved in developmental control. Previous studies have shown that its central region is essential for function even though its three domains are poorly conserved and intrinsically disordered. Using these disordered domains as affinity reagents, we have now identified multiple embryonic Gro-interacting proteins. The interactors include protein complexes involved in chromosome organization, mRNA processing, and signaling. Further investigation of the interacting proteins using a reporter assay showed that many of them modulate Gro-mediated repression either positively or negatively. The positive regulators include components of the spliceosomal subcomplex U1 small nuclear ribonucleoprotein (U1 snRNP). A co-immunoprecipitation experiment confirms this finding and suggests that a sizable fraction of nuclear U1 snRNP is associated with Gro. The use of RNA-seq to analyze the gene expression profile of cells subjected to knockdown of Gro or snRNP-U1-C (a component of U1 snRNP) showed a significant overlap between genes regulated by these two factors. Furthermore, comparison of our RNA-seq data with Gro and RNA polymerase II ChIP data led to a number of insights, including the finding that Gro-repressed genes are enriched for promoter-proximal RNA polymerase II. We conclude that the Gro central domains mediate multiple interactions required for repression, thus functioning as a regulatory hub. Furthermore, interactions with the spliceosome may contribute to repression by Gro.

Groucho (Gro)² is a conserved metazoan co-repressor that may be particularly critical for long range repression whereby repressors are able to establish large transcriptionally silent domains that can spread over many thousands of base pairs (1–3). Gro is essential in many developmental processes, including sex determination, neurogenesis, and pattern formation in *Drosophila* as well as myogenesis and hematopoiesis in vertebrates (2, 4, 5). Gro also has roles in multiple signal transduction pathways, including the Ras and Notch pathways (6–8). Furthermore, increased Gro activity correlates with the appearance of certain forms of cancer such as lung cancer (9, 10). Thus, understanding the mechanism of Gro-mediated repression should contribute to our understanding of long range repression and its role in development, signaling, and disease.

Sequence comparison of Gro family proteins reveals five domains (2, 10). The C-terminal WD repeat domain forms a β -propeller that interacts with the WRPW and eh1 motifs found in many Gro-dependent DNA-binding repressors (11). The N-terminal Q domain folds into a coiled coil structure that forms tetramers and perhaps higher order oligomers, and this self-association is required for robust repression (12–15). The central GP, CcN, and SP domains are believed to have essential functions even though their primary sequences are not well conserved. The GP domain interacts with the histone deacetylase Rpd3/HDAC1 (16, 17). Histone deacetylation is broadly associated with gene silencing, and treatment of flies with histone deacetylase inhibitors attenuates Gro-mediated repression (18). In addition, the GP domain is essential for nuclear localization because deletion of this domain prevents Gro nuclear uptake (19). The SP domain regulates Gro function negatively as its deletion leads to promiscuous repression and

* This work was supported by National Institutes of Health Grants GM44522 (to A. J. C.) and GM089778 (to J. A. W.). The authors declare that they have no conflict of interest with the contents of this article. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

^S This article contains supplemental Tables S1 and S2.

¹ To whom correspondence should be addressed: Dept. of Chemistry and Biochemistry, UCLA, 607 Charles E. Young Dr. East, Los Angeles, CA 90095-1569. Tel.: 310-825-2530; E-mail: courey@chem.ucla.edu.

² The abbreviations used are: Gro, Groucho; MudPIT, multidimensional protein identification technology; snRNP, small nuclear ribonucleoprotein; seq, sequencing; Pol II, RNA polymerase II; qPCR, quantitative PCR; dsRNA, double-stranded RNA; CKII, casein kinase II; ACF, ATP-dependent chromatin remodeling and assembly factor; Su(H), Suppressor of Hairless; Brk, Brinker; UAS, upstream activating sequence; EWS/FLI, Ewing sarcoma breakpoint region 1/Friend leukemia integration 1 transcription factor.

TABLE 1
PCR primers used in construction of plasmids encoding GST fusion proteins

Domain	Sequence
Q	ATTATTAGGATCCATGGATTACAAGGACGATGACGATAATATCCCTCACCGGTGCGCCACCCC ATTATTACTCGAGTCACTGCTGGGCGTGGATCTGTTCGCCA
GP	ATTATTAGGATCCATGGATTACAAGGACGATGACGATAAGTGCCAGGTGGACCACCTCAGCCGA ATTATTACTCGAGTCAAGCAATTGAGCAATCGCTCCTCGGC
CcN	ATTATTAGGATCCATGGATTACAAGGACGATGACGATAAGTTTCGCCGGCCGATCGTGAGAAGT ATTATTACTCGAGTCAATAGACACGTGCTCGCCGTGGGA
SP	ATTATTAGGATCCATGGATTACAAGGACGATGACGATAAGAGGTGCGCGATCGGGAAAGCTTGA ATTATTACTCGAGTCAACCCGTTAGGGCCGAGGGATGTGGA

developmental defects (19). Phosphorylation of the SP domain by Ras/MAPK signaling was shown to attenuate repression, providing a mechanism for regulating repression in response to environmental cues (20). Finally, the CcN domain is also targeted for phosphorylation by protein kinases and is required for repression by Gro (19, 21).

Sequence analysis of the Gro central domains strongly suggests that they are intrinsically disordered (19). Intrinsically disordered regions in proteins lack rigid three-dimensional structures under native conditions and can serve as hubs of large regulatory networks by mediating a wide array of highly specific protein interactions (22, 23). Increasing evidence suggests that intrinsically disordered domains have critical functions in transcriptional regulation (24, 25).

In this study, we set out to illuminate the mechanisms of Gro-mediated repression by identifying proteins that interact with the N-terminal Q domain and the three central domains. A proteomic screen revealed over 160 interacting proteins, many of which are components of protein complexes in a variety of functional categories such as chromatin remodeling and RNA processing. Perhaps most notably, the interactors included multiple components of the spliceosome, and a co-immunoprecipitation experiment suggests that a sizable fraction of U1 snRNP (a subcomplex of the spliceosome) is associated with Gro in embryonic nuclei.

As a means of systematically validating the functional significance of these interactions, we carried out a novel reporter assay using three different luciferase reporters that could be monitored simultaneously. These assays showed that many of the interacting proteins, including the protein components of U1 snRNP, are required for optimal Gro-mediated repression. Lastly, we compared the effects on the gene expression profile of Gro and U1 snRNP knockdown, finding a significant overlap in the regulated genes. Our results indicate that the central domains of Gro mediate multiple interactions required for repression and reveal a possible mechanism of Gro-mediated repression through an interaction with the spliceosome complex or subcomplexes. This reinforces previous studies suggesting that the spliceosome has roles in transcriptional regulation in addition to its roles in RNA processing (26–30).

Experimental Procedures

Plasmids—To generate plasmids for expression of glutathione *S*-transferase (GST) fusion proteins, sequences encoding the Gro domains were amplified by PCR and inserted between the BamHI and XhoI sites of pGEX4T (GE Healthcare). The Q domain included Gro amino acids 1–133, the GP domain included amino acids 134–194, the CcN domain included

amino acids 195–257, and the SP domain included amino acids 258–390. Sequences of PCR primers are provided in Table 1.

Plasmids used in the reporter assay were generated as follows. The red luciferase plasmid, G5DE5-pCBR, was generated by inserting the G5 DE5 enhancer region (14) into pCBR-Basic vector (Promega catalog number E1411) between the KpnI and XhoI sites. The green luciferase plasmid, DE5G5-pCBG68, was generated by inserting the luciferase gene using NcoI and SalI from pCBG68-basic vector (Promega catalog number E1431) into the DE5 G5 vector, which has UAS elements downstream of the reporter.³ Actin promoter-driven Dorsal (pPac DI), Twist (pPac Twi), and Gal4-Gro (pAct Gal4-Gro) plasmids have been described previously (14). The RpIII128 promoter-driven *Renilla* luciferase plasmid, RpIII128-Rluc, was obtained from Addgene (ID number 37380) (31).

Affinity Purification and Identification of Gro-interacting Proteins—Plasmids encoding the recombinant domains fused to GST or GST alone were transformed into BL21 cells. 250 ml of midlog cells were induced with 0.25 mM isopropyl 1-thio- β -D-galactopyranoside for an hour. Cells were pelleted at 4,000 \times g, resuspended in 25 ml of salty TE (0.15 M NaCl, 10 mM Tris, pH 8, 1 mM EDTA) with protease inhibitor (Life Technologies catalog number 88266), and incubated on ice for 30 min. Samples were incubated at 4 °C for 15 min after DTT and Triton X-100 were added to final concentrations of 5 mM and 1%, respectively. Cells were then disrupted through a microfluidizer (Microfluidics M110L) using standard conditions. The lysate was collected and centrifuged at 14,000 \times g for 10 min at 4 °C. Supernatant was collected, and 1 ml of glutathione-agarose resin (50% slurry) was added. After overnight incubation, the resin was washed with ice-cold PBS three times and stored at 4 °C.

Drosophila embryo nuclear extracts were prepared as described previously (32). To isolate Gro-interacting proteins, 20 μ g of glutathione bead-immobilized recombinant domains was mixed with nuclear extract containing 30 mg of protein (20 mg/ml) in 8 ml of HEMNK buffer (40 mM HEPES, pH 7.5, 5 mM MgCl₂, 0.2 mM EDTA, 1 mM DTT, 0.5% Nonidet P-40, 0.1 M KCl) at 4 °C overnight. Samples were washed six times for 15 min with 5 ml of HEMNK buffer. Proteins were first eluted with 5 ml of 2 M NaCl in HEMNK buffer and then with 2.5 ml of 2 M NaCl in HEMNK buffer for 20 min each. Eluted proteins were subjected to TCA precipitation prior to multidimensional protein identification technology (MudPIT) analysis. MudPIT analysis was performed as described previously (33). Peptide

³ A. J. Courey, unpublished data.

TABLE 2
RT-qPCR primers

Gene	Sequence
Gro	TTTATTACAACATGTTTCGAAATCATGC TTCGCTTTTGTGATGCGTTGCTAC
snRNP-U1-C	CTCAGGAACGGCATCAACGTT TATAATTAATTGTTTTCGCTATCGGG
Rpl32	CCCAAGGGTATCGACAACAGA CGATCTCGCCGAGTAAAC
U1 snRNA	ATACTTACCTGGCGTAGAGGTTAAC AACGCCATTCCCGGCTA

identifications were filtered using a false discovery rate cutoff of 0.05 as determined by the decoy database approach. Protein level false positive rates were less than 0.03 for all individual runs.

Supplemental Table S1C includes all the mass spectroscopy data for the two independent replicate screens carried out with each GST fusion protein and GST alone, and supplemental Tables S1, A and B includes selective data for 159 proteins that were detected in both replicates as well as three proteins (Histone H3, Caf1, and Bic) that were only detected in one replicate but for which other data confirm the significance of the interaction (supplemental Table S1B, notes 2 and 3). Ribosomal proteins were excluded from the lists in supplemental Table S1, A and B.

Gro Immunoprecipitation and Reverse Transcription-qPCR (RT-qPCR) Analysis of U1 snRNA—500 μ g of nuclear extract was incubated with 1.875 μ g of affinity-purified rabbit antibody against the Gro GP domain or rabbit IgG in a final volume of 250 μ l of HEMNK buffer overnight at 4 °C. 225 μ g of Protein A Dynabeads (Invitrogen catalog number 10001D) were incubated with the samples at 4 °C for 1 h. Samples were then washed with HEMNK buffer three times for 10 min each. For RT-qPCR, RNA was eluted in 10 μ l of water by heating to 80 °C for 2 min. Samples were treated with DNase I according to the manufacturer's protocol (Promega catalog number M6101). Reverse transcription was performed with 300 ng of random primer (Invitrogen catalog number 48190-011), and qPCR was performed using primers amplifying U1 snRNA (Table 2). Threshold cycle values were converted to percent input values by comparison with a standard curve generated from multiple serial dilutions of RNA isolated by TRIzol (Life Technologies catalog number 10296010) extraction from the input nuclear extract. Primer specificity was validated by melting curve analysis of the amplification products (data not shown).

For immunoblotting, samples were eluted in SDS-PAGE loading buffer. Proteins were detected with a mixture of mouse anti-Gro (Developmental Studies Hybridoma Bank; 1:650 dilution) and affinity-purified rabbit anti-GP domain (1:100 dilution) antibodies. Immunoblots were subsequently probed with goat anti-mouse 680 and goat anti-rabbit 800 IRDye-coupled secondary antibodies (LI-COR) and imaged with a LI-COR Odyssey imager.

Three-reporter Luciferase Assay—To guard against off-target effects, each candidate gene was knocked down with three non-overlapping dsRNAs when possible (the complete list of dsRNAs used is available upon request). Each dsRNA was tested in triplicate. dsRNA was synthesized by the *Drosophila* RNAi Screening Center and realiquoted into white flat bottom

96-well plates (USA Scientific catalog number CC7682-7968) at 150 ng/well in 10 μ l of water using a Beckman Coulter BioMek FX work station.

Transfections were carried out with Effectene reagent (Qiagen catalog number 301425). 6 μ g each of G5DE5-pCBR and DE5G5-pCBG68, 0.6 μ g of RpIII128-Rluc, 1 μ g of pPac D1, 0.3 μ g of pPac Twi, and 1.2 μ g of pAct Gal4-Gro were suspended in 600 μ l of buffer EC. 33 μ l of this mixture was added to 25 μ l of enhancer. After 2–3 min, 7.5 μ l of Effectene was added and mixed by pipetting up and down. 6 μ l of this mixture was immediately added into each well of a 96-well plate containing 150 ng of dsRNA. 4–8 min later, 100 μ l of S2 cells (diluted to 1×10^6 cell/ml) was added to each well. Cells were incubated at 24 °C for 2 days before assaying.

The luminescence signal was measured with a Molecular Devices LJM Analyst HT microplate reader using emission filters ET510/80m and E610LP (Chroma catalog numbers S-022658 and 138951). 50 μ l of D-luciferin (Chroma-Glo system, Promega catalog number E2980) was added to each well. Five minutes later, the reaction was stopped by the addition of 50 μ l of stop buffer containing coelenterazine (Dual-Luciferase system, Promega catalog number E1960). The luminescence signal was measured immediately without applying a filter.

To address the issue of signal overlap, raw signals were subjected to filter correction. The corrected red luminescence signal, R' , and green luminescence signal, G' , were calculated according to the following equations.

$$R' = \frac{\text{Lrf} - \text{Lgf} \times \left(\frac{\text{Grf}}{\text{Ggf}}\right)}{\left(\frac{\text{Rrf}}{R}\right) - \left(\frac{\text{Rgf}}{R}\right) \times \left(\frac{\text{Grf}}{\text{Ggf}}\right)} \quad (\text{Eq. 1})$$

$$G' = \frac{\text{Lgf} - R' \times \left(\frac{\text{Rgf}}{R}\right)}{\left(\frac{\text{Ggf}}{G}\right)} \quad (\text{Eq. 2})$$

Parameters were determined by expressing the individual luciferases and recording the luminescence signals with red and green filters and with no filter (data not shown). The ratio of green signal passed through the red filter, Grf/Ggf , was determined to be 0.0975; the ratio of red signal passed through the red filter, Rrf/R , was determined to be 0.42; the ratio of red signal passed through the green filter, Rgf/R , was determined to be 0; and the ratio of green signal passed through the green filter, Ggf/G , was determined to be 0.47. Lrf and Lgf are luminescence signals in which cells are co-transfected with both red and green luciferases. Lrf is the signal recorded with the red filter, and Lgf is the signal recorded with the green filter.

The signal from untransfected cells was then subtracted from the corrected data to eliminate background. Processed data were then normalized to the internal control *Renilla* luciferase. Finally, data were compared with the signal from cells in the same plate that were treated with control GFP dsRNA. A change in long or short range repression was considered significant if the p value was <0.1 . If multiple dsRNAs were tested for

Groucho/Spliceosome Interactions

a given gene (as was true in most cases; supplemental Table S2), then a change is only listed if the p value was <0.1 for at least two separate dsRNAs.

RNA-seq Library Preparation—Gro dsRNA was generated by PCR amplification of the first 800 nucleotides of the coding sequence using primers containing T7 promoters followed by *in vitro* transcription with T7 RNA polymerase. snRNP-U1-C dsRNA was generated by PCR and *in vitro* transcription of the snRNP-U1-C coding sequence with primers 5'-taatagcactcactatagggtactCAAAGTACTATTGCGACTACTGC and 5'-taatagcactcactatagggtactCTTGGGTCCGTTTCATGATTCC (lowercase letters represent the T7 promoter sequences). Transfection was carried out as described previously (34). RT-qPCR was used to determine the knockdown efficiency prior to RNA-seq library preparation. RT-qPCR primers targeted the 3'-UTRs of Gro and snRNP-U1-C. Rpl32 was used as a reference gene. The specificity of all primers was validated by melting curve analysis of the amplification products (data not shown). Sequences of the qPCR primers are listed in Table 2.

Total RNA was extracted with TRIzol according to the manufacturer's protocol. RNA integrity was determined with an Agilent 2100 Bioanalyzer using the RNA 6000 Nano kit (Agilent catalog number 5067-1511). Isolation of mRNA was carried out as follows. Streptavidin magnetic beads (Promega catalog number Z5481) were prepared in aliquots of 120 and 60 μl in $0.5\times$ SSC with 10 mM EDTA. 15 μg of total RNA was mixed with 1.5 μM biotinylated 15-mer poly(T) oligonucleotide in $0.5\times$ SSC with 10 mM EDTA. Samples were first incubated at 75 $^{\circ}\text{C}$ for 5 min followed by 15 $^{\circ}\text{C}$ for 10 min and 10 $^{\circ}\text{C}$ for 10 min. Samples were then incubated with 120 μl of magnetic beads at 4 $^{\circ}\text{C}$ for 2 h followed by 60 μl of magnetic beads at 4 $^{\circ}\text{C}$ for 30 min. The two aliquots of beads were combined and washed four times with 300 μl of ice-cold $0.1\times$ SSC containing 10 mM EDTA. mRNA was first eluted with 100 μl of water followed by 150 μl of water at 37 $^{\circ}\text{C}$ for 10 min each. Samples were precipitated with ethanol and stored at -80°C . Pulldown efficiency of mRNA and depletion efficiency of 18S rRNA were determined by RT-qPCR (data not shown).

The RNA-seq library was prepared according to the manufacturer's protocol (Epicenter, catalog numbers SSV21124 and RSBC10948). The concentration of the library was determined with Pico Green (Life Technologies catalog number Q32851) according to the manufacturer's directions. Fluorescence signal was measured using a TECAN M1000 fluorescence plate reader.

Bioinformatics—Alignment of paired end reads to the *Drosophila melanogaster* genome (assembly BDGP 5/dm3) was performed with Tophat2 (v2.0.9) (35) using default parameters. DESeq2 (v1.6.3) (36) was used for gene expression level normalization and differential expression significance testing. Histone modification and motif enrichment analysis were carried with i-cisTarget (37) using default parameters. Enriched gene ontology analysis was done with FlyMine (v31.0) (38) using default parameters.

Results

Identification of Gro-interacting Proteins—A previous study showed that deletion of the GP or CcN domain in the Gro central region led to a loss of Gro-mediated repression and to

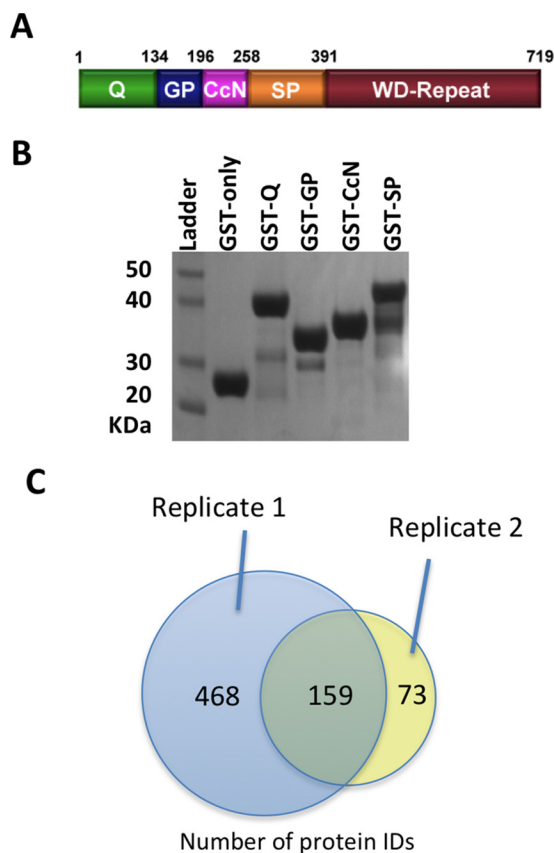


FIGURE 1. Purification of Gro-interacting proteins. A, schematic representation of Gro. The Q, GP, CcN, and SP domains were tagged with GST. B, the GST-tagged domains were expressed in *E. coli* and purified with glutathione-agarose beads. They were then resolved by 10% SDS-PAGE and visualized by Coomassie Blue staining. These proteins were then used as affinity reagents in the purification of Gro-interacting proteins from embryonic nuclear extracts that were subsequently identified by MudPIT (see Table 4 and supplemental Table S1). C, Venn diagram showing overlap between the non-ribosomal proteins identified in two replicate sets of affinity purification experiments. Fisher's exact test indicates that the overlap between the two sets is highly significant ($p < 2.2 \times 10^{-16}$).

lethality, whereas deletion of the SP domain led to reduced specificity of Gro-mediated repression and to reduced viability (19). To identify possible regulatory partners of these domains, we used them as affinity reagents to purify interacting proteins, which were then identified by mass spectrometry. The three central domains of Gro were expressed as GST-tagged proteins and purified from *Escherichia coli* lysates (Fig. 1, A and B). We also constructed a similarly tagged form of the N-terminal Q domain because previous studies suggested that, in addition to mediating Gro oligomerization, the Q domain engages in interactions with regulatory targets (39, 40).

The glutathione bead-immobilized GST-fused domains (or, as a negative control, immobilized unfused GST) were incubated with a *Drosophila* embryo nuclear extract. After extensive washing, interacting proteins were eluted with 2 M salt and analyzed by MudPIT (33) (supplemental Table S1C). Duplicate extract preparations and affinity purifications were carried out and analyzed on separate dates, and there was a high degree of overlap between the sets of proteins identified in these duplicate experiments (Fig. 1C). With three exceptions (see "Experimental Procedures"), only proteins that appeared in both rep-

TABLE 3
Enriched gene ontology groups of Gro-interacting proteins
 $p < 0.05$.

Enriched gene ontology	No. of genes
Gene expression	83
Chromosome organization	21
Chromatin modification	11
mRNA processing	53
Cell cycle	30
Cell differentiation	66
Developmental process	76
Neurogenesis	57
Anatomical structure development	74

licates were included in our list of Gro-interacting proteins (Fig. 1C and supplemental Table S1, A and B). Gene ontology analysis of this list of 162 proteins revealed a variety of functions, including regulation of gene expression, RNA processing, and developmental processes (Table 3).

89 of the 162 Gro-interacting proteins associated uniquely with one domain (in all but one case, the SP domain), whereas 32 interacted with two domains. In the case of 23 of the 32 proteins that interacted with two domains, one of these domains was the Q domain (supplemental Table S1A). This is consistent with the known role of the Q domain in homo-oligomerization (12–15). In accord with this role, chromatography using GST-Q as the affinity reagent resulted in the purification of some full-length endogenous Gro (supplemental Table S1C and data not shown). This could lead to the co-purification of Gro-interacting proteins that bind to regions outside the Q domain. Thus, 112 (89 plus 23) of the 162 detected interacting proteins can, in principal, be accounted for by the binding of Gro to a single central domain. However, at least 50 proteins (162 minus 112) are able to bind independently to two or three central domains. The ability to interact with multiple Gro domains could allow tighter binding or more versatile control of binding.

The list of interacting proteins (Table 4 and supplemental Table S1, A and B) contains multiple components of known multisubunit protein complexes. For example, we identified the α and β subunits of casein kinase II (CKII), a previously identified regulator of Gro activity (21). We also detected protein complexes involved in chromosome organization, including both components of the ATP-dependent chromatin remodeling and assembly factor (ACF), Acf1 and Iswi (41). Our proteomic screens also identified all the core protein components of the nucleosome (the core histones) as well as histone variant H2Av, consistent with previous studies demonstrating functional interactions between Gro and nucleosomes (42–44).

Perhaps most surprisingly, we discovered a number of components of the spliceosome among the group of Gro-interacting proteins, including all three proteins unique to U1 snRNP, components of U4/U6 snRNP, U2 snRNP, and the Sm complex (45, 46). To validate the interaction between Gro and U1 snRNP, *Drosophila* embryo nuclear extracts were subjected to immunoprecipitation using an affinity-purified antibody against the Gro GP domain or, as a negative control, rabbit IgG. An anti-Gro immunoblot of the immunoprecipitated material demonstrates the efficiency of the immunoprecipitation (Fig. 2A). RNA was extracted from the immunoprecipitates and ana-

TABLE 4
Representative Gro-interacting proteins

For the complete list, full gene/protein names, and UniProt identifiers, see supplemental Table S1.

Protein	Description
Chromosome organization	
Acf1	ACF chromatin remodeling complex
Iswi	ACF chromatin remodeling complex
Caf1	dNuRD chromatin remodeling complex
Nap1	Histone chaperone
Ball	H2A Thr-119 kinase
JIL-1	H3 Ser-10 kinase
Top1	Topoisomerase
Top2	Topoisomerase
Developmental process	
CKII α	CKII complex
CKII β	CKII complex
Nopp140	Negative regulator of CKII
Nito	Positive regulator of Wnt signaling pathway
Rm62	DEAD box helicase
Fmr1	Fragile X protein; interacting partner of Rm62
Vir	Involvement in sex determination
Snama	Involvement in eye morphogenesis
NonA	Involvement in visual perception
mRNA processing	
snRNP-U1-A	U1 snRNP complex
snRNP-U1-C	U1 snRNP complex
snRNP-U1-70K	U1 snRNP complex
U4-U6-60K	U4/U6 snRNP complex
CG7028	U4/U6 snRNP complex
Prp31	U4/U6 snRNP complex
Prp8	U5 snRNP complex
U2af38	U2 snRNP complex
U2af50	U2 snRNP complex
SF2	U2 snRNP complex
SmD2	Sm complex
SmD3	Sm complex
Nop60B	H/ACA ribonucleoprotein complex subunit
NHP2	H/ACA ribonucleoprotein complex subunit

lyzed by RT-qPCR with primers specific for U1 snRNA (a component of U1 snRNP). The results show that ~13% of the U1 snRNA in the nuclei of 0–12-h embryos is associated with Gro (Fig. 2B).

Functional Analysis of Gro-interacting Proteins—We next carried out functional assays to determine whether the interacting proteins are required for regulation of a Gro-responsive reporter gene. Previous studies established a reliable reporter assay for Gro function using a luciferase reporter containing Gal4 binding sites (UAS elements) as well as an artificial enhancer containing binding sites for the Dorsal and Twist activators (14, 16, 18, 47). Dorsal/Twist-activated transcription of this reporter is strongly repressed upon introduction of a Gal4-Gro fusion protein. By altering the position of UAS elements relative to the artificial enhancer, we were able to examine both short range and long range Gro-mediated repression simultaneously (Fig. 3, A and B). The reporter system relied on two variants of click beetle luciferase that use D-luciferin as a substrate and emit either red or green light (48). In addition, a plasmid encoding *Renilla* luciferase, which uses coelenterazine as a substrate, was used as an internal control for transfection efficiency, cell viability, and general effects on transcription and translation. We validated the three-reporter system using dsRNA against Dorsal, Gro, and Rpd3 (which is partially required for Gro-mediated repression (18)) (Fig. 3C). As predicted, Dorsal knockdown resulted in a complete loss of activation, Gro knockdown resulted in a complete loss of repression, and Rpd3 knockdown resulted in a partial loss of repression.

Each of the candidates from the screen for Gro-interacting proteins was knocked down by RNAi using up to three dsRNAs

Groucho/Spliceosome Interactions

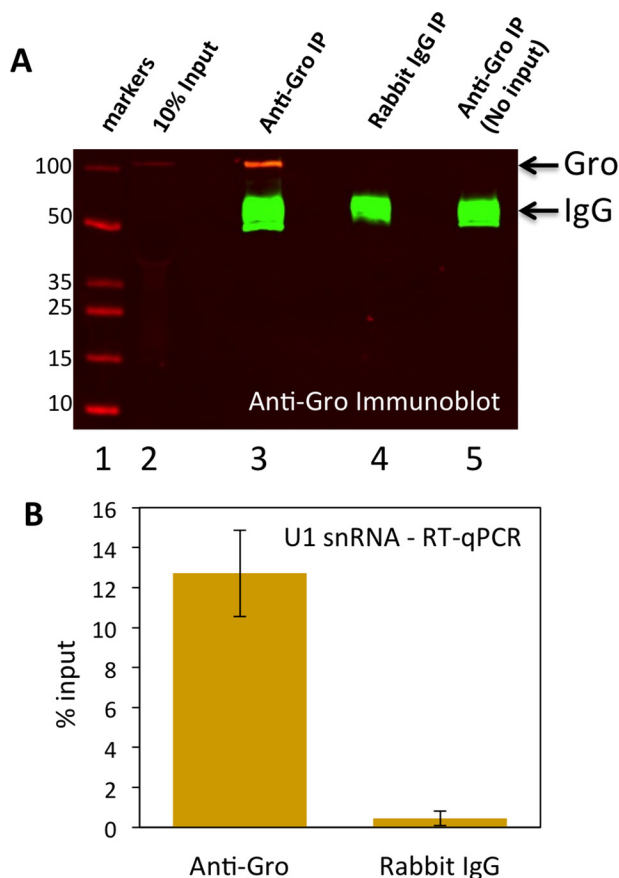


FIGURE 2. Validation of the interaction between Gro and U1 snRNP. *A*, 0–12-h *Drosophila* embryo nuclear extracts were subjected to immunoprecipitation using an affinity-purified polyclonal antibody directed against the Gro GP domain or, as a control, rabbit IgG. To assess immunoprecipitation efficiency and specificity, immunoprecipitates were subjected to SDS-PAGE and immunoblotting. The blot was probed with a mixture of the rabbit anti-GP domain antibody and a mouse monoclonal anti-Gro antibody and with IRDye-labeled secondary antibodies. The signal from the rabbit antibody was detected in the green channel of the IR imager, whereas the signal from the mouse antibody was detected in the red channel. Rabbit IgG heavy chain (*IgG*) and Gro bands are indicated with arrows on the right. The orange-yellow color of the Gro band is indicative of the overlap between the red and green signals. Lane 1, markers labeled in kDa; lane 2, 10% input; lane 3, anti-Gro immunoprecipitate (IP); lane 4, rabbit IgG immunoprecipitate; lane 5, mock anti-Gro immunoprecipitate from which input nuclear extract was omitted. *B*, RNA was extracted from immunoprecipitates prepared as described in *A*. The RNA from the immunoprecipitates as well as the RNA extracted from the input nuclear extracts was analyzed by RT-qPCR as described under “Experimental Procedures” to determine U1 snRNA levels. Error bars based on two independent biological replicates indicate S.D. A two-tailed *t* test gives $p = 0.016$.

per gene to guard against off-target effects. We excluded the histones from this analysis under the assumption that knockdown of these essential chromatin components would have pleiotropic deleterious effects on cell metabolism and because each histone is encoded by multiple genes, making efficient knockdown problematic. We therefore tested 157 genes in this S2 cell luciferase assay in most cases with multiple dsRNAs per gene (three if available), and each dsRNA was tested in triplicate. In total, we carried out ~1,300 assays (including controls) in a 96-well plate format using a partially automated approach (see “Experimental Procedures”).

A candidate was scored as a regulator of Gro-mediated repression if knockdown reproducibly resulted in either an

increase or a decrease in the level of repression (see “Experimental Procedures” for explanation of the statistical test of significance). 44 candidates met these criteria of which 28 interfered with optimal repression (*i.e.* repression increased upon knockdown; these were termed “negative regulators of Gro”), and 16 were required for optimal repression (*i.e.* repression decreased upon knockdown; these were termed “positive regulators of Gro”). We provide representative data for one negative regulator (Vir), one positive regulator (snRNP-U1-C), and one protein that is neither a positive nor a negative regulator (SR protein kinase, SRPK) (Fig. 3*D*); a list of all the positive and negative regulators (Table 5); and a separate list showing the quantitative effect of RNAi knockdown of each of the 44 regulators on repression by Gal4-Gro (supplemental Table S2). Of particular interest, four spliceosomal proteins, including two components of U1 snRNP, act as positive regulators of Gro, confirming the functional significance of the interaction between Gro and U1 snRNP. A few other noteworthy examples among the Gro regulators (Table 5 and supplemental Table S2) include both components of the CKII complex (CKII α and CKII β), which act as negative regulators, and the chromatin remodeling factor Acl1, which acts as a positive regulator (see “Discussion”).

Expression Profiling of Gro and snRNP-U1-C Knockdown Cells—snRNP-U1-C is one of the components of the U1 snRNP complex, which is responsible for 5' splice site recognition (46). In addition to its role in RNA processing, it has been shown to repress transcription of EWS/FLI-transactivated genes (30). Because our data indicated that snRNP-U1-C may also modulate Gro function, we examined the genome-wide role of snRNP-U1-C in Gro-mediated repression. Using RNA-seq, we compared the effects of snRNP-U1-C knockdown with those of Gro knockdown on the gene expression profile in S2 cells. Cells were treated with Gro or snRNP-U1-C dsRNA for 4 days, leading to 4-fold or greater knockdown of the Gro and snRNP-U1-C mRNA (Fig. 4*A*). The transcriptomes in wild-type and Gro knockdown S2 cells were quantitatively similar to those published previously (49, 50) (Fig. 4, *B* and *C*). We note that the genes differentially expressed in the snRNP-U1-C knockdown are enriched for genes containing introns as would be expected given the role of U1 snRNP in splicing. However, this set of genes also contains a number of intronless genes, consistent with the idea that snRNP-U1-C has roles in gene regulation apart from its role in splicing (Fig. 4*D*). We note that changes in the expression of an intronless gene could also reflect a requirement for the product of an intron-containing gene in the expression of the intronless gene.

98 genes were differentially expressed in both Gro and snRNP-U1-C knockdown cells (Fig. 4*E*) of which 36 were up-regulated in either case. These coordinately up-regulated targets included genes in various signaling pathways such as the Wnt, Notch, and Toll pathways (Table 6). Comparison with publically available ChIP-seq data on histone modification and transcription factor binding revealed that these coordinately regulated genes were most enriched for histone H3K36 methylation and the H3K36 methyltransferase ASH1 (Fig. 4*F*).

To determine whether the regulatory effects of knocking down Gro are likely to be direct, we compared our RNA-seq

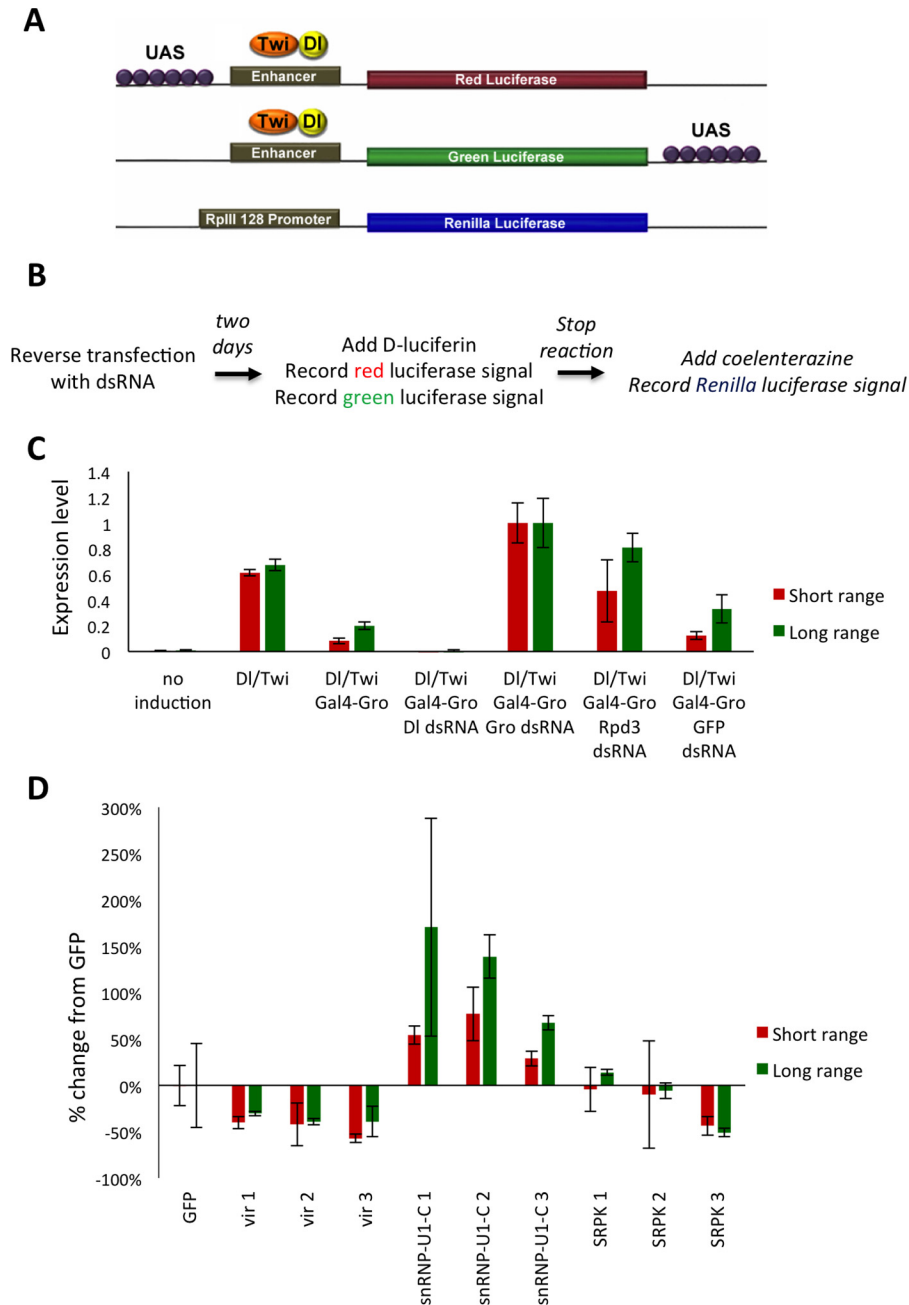


FIGURE 3. The three-reporter high throughput luciferase assay. *A*, schematic representation of the three reporters. Constructs are not drawn to scale. In the red luciferase reporter, the Gal4 binding sites (UAS elements) are immediately upstream of the enhancer, whereas in the green luciferase reporter, the UAS elements are about 2 kb downstream of the transcriptional start site. Expression is induced by the Dorsal (*Dl*) and Twist (*Twi*) activators and repressed by Gal4-Gro. The *Renilla* luciferase reporter under control of the class II promoter from the gene encoding RNA polymerase III subunit (RpIII128) was used as an internal control for transfection efficiency, cell viability, and general effects on transcription and translation. *B*, flow chart of the reporter assay. *C*, validation of the reporter assay. Co-transfection with Dorsal and Twist (*Dl/Twi*)-encoding plasmids activated both the red and green reporters, whereas addition of a plasmid encoding the Gal4-Gro fusion resulted in repression of the reporters. Dorsal, Gro (including Gal4-Gro), and the histone deacetylase Rpd3, which is partially required for Gro-mediated repression (18), were knocked down by RNAi. Data are normalized to the red and green signals from the Gro dsRNA sample. *Error bars* based on triplicate transfection assays represent S.D. *D*, representative results of the reporter assay. The luciferase reporter assay was carried out using three non-overlapping dsRNAs from the genes encoding *Vir*, snRNP-U1-C, and SR protein kinase (*SRPK*). The result of transfection with each dsRNA was compared with that of transfection with GFP dsRNA. *Error bars* based on triplicate transfection assays represent S.D.

data from Gro knockdown S2 cells with available S2 cell Gro ChIP data (49). Gro appears to bind many genes that it does not repress (Fig. 5A). This is consistent with observations made with numerous regulatory factors (51, 52) and suggests that binding, although required, is not sufficient for regulation. We observed an enrichment of Suppressor of Hairless (Su(H)) and Brinker (Brk) binding motifs within Gro ChIP-

seq peaks in the differentially expressed genes but not in the non-differentially expressed genes (Fig. 5B). Comparison of our RNA-seq data from Gro knockdown cells with available Pol II ChIP-chip data (53) also reveals an enrichment in Pol II pausing near the transcriptional start site in genes that are up-regulated upon Gro knockdown (*i.e.* genes that are repressed by Gro; Fig. 6).

Groucho/Spliceosome Interactions

TABLE 5

Positive and negative Gro regulators

See supplemental Table S2 for quantitative information on positive and negative regulation by these factors. See “Experimental Procedures” for an explanation of the test of statistical significance that genes had to pass to be included in this list.

Potential negative regulators of Gro ^a	Potential positive regulators of Gro ^b
CKII α , CKII β , Nopp140, fl(2)d, l(2)35Df, l(3)72Ab, Vir, NonA, Nito, X16, Nap1, JIL-1, Nop5, NHP2, FK506-bp1, CG3605, Prp31, Fmr1, Cdk12, CG6418, CG7372, CG7946, Srp68, Srp72, Ssrp, Pitsire, Pep, Nab2	snRNP-U1-C, snRNP-U1-70K, U2af50, U4-U6-60K, Rm62, Orc2, Smid, Acn, Acf1, Snama, CG1622, ZCHC8, CG4709, CG4806, Lat, Srp19

^a Negative regulators are defined as the products of those genes the knockdown of which led to increased repression by Gal4-Gro in the reporter assay.

^b Positive regulators are defined as the products of those genes the knockdown of which led to decreased repression by Gal4-Gro in the reporter assay.

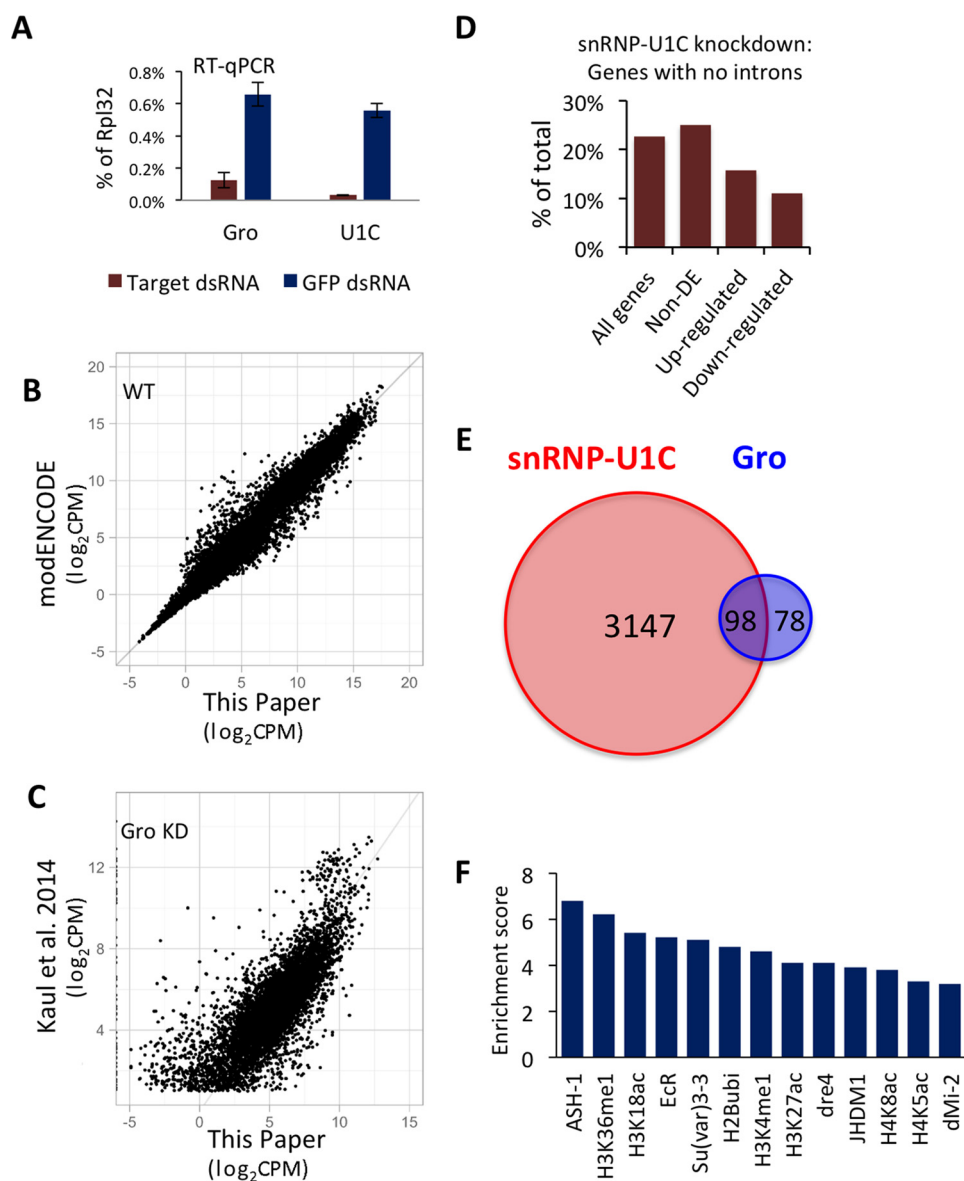


FIGURE 4. Genome-wide expression profiling reveals co-regulation of genes by Gro and snRNP-U1-C. *A*, expression of Gro and snRNP-U1-C mRNA after dsRNA treatment. RT-qPCR was performed after extraction of total RNA. Data were normalized to reference gene Rpl32. Error bars based on duplicate experiments represent S.D. *B*, comparison of transcriptomes from our wild-type S2 cell RNA-seq data and the modENCODE S2 cell RNA-seq data. *C*, comparison of transcriptomes from our Gro knockdown RNA-seq data and previously published Gro knockdown RNA-seq data (49). The transcripts that were detected at significant levels in only the previously published Gro knockdown study (represented by the points in contact with the vertical axis) correspond primarily to non-polyadenylated transcripts. In *B* and *C*, the scale on both axes is log₂(CPM) where CPM is counts per million sequence reads. *D*, based on RNA-seq analysis of wild-type and snRNP-U1-C knockdown cells, genes were categorized as non-differentially expressed upon knockdown (*Non-DE*; 12,028 genes), up-regulated upon knockdown (1,431 genes), and down-regulated upon knockdown (1,691 genes). The percentage of genes in each category with no introns is shown. Some *Drosophila* genes lack annotated transcripts, and thus it was not possible to determine their intron count. This results in a small numerical discrepancy between the number of differentially expressed genes included in this analysis and the number of snRNP-U1C differentially expressed genes shown in *E*. *E*, Venn diagram showing numbers of differentially expressed genes in Gro and snRNP-U1-C knockdown cells and the overlap between these sets. Fisher's exact test indicates that the overlap is highly significant ($p < 2.2 \times 10^{-16}$). *F*, enrichment of Gro/snRNP co-regulated genes for various features. Normalized enrichment scores were calculated using cumulative recovery curves (37). Scores above 2.5 are considered significant.

TABLE 6
Genes up-regulated upon knockdown of either Gro or snRNP-U1-C

Name	Function
Secreted Wg-interacting molecule	Wnt signaling pathway
Wnt oncogene analog 5	Wnt signaling pathway
E(spl)m2-BFM	Notch signaling
Spatzle	Toll signaling pathway
SH2 ankyrin repeat kinase	JNK cascade
Dawdle	SMAD protein signal transduction
CG33275	Rho protein signal transduction
Epac	Rap protein signal transduction
Boundary element-associated factor of 32kD	H3K9 methylation
Syncrrip	Dorsal/ventral axis specification
Fasciclin 1	Neuron recognition
Axotactin	Transmission of nerve impulse
Muscle-specific protein, 300 kDa	Skeletal muscle tissue development
Cheerio	Lamellocyte differentiation

Discussion

Previous studies showed that the disordered Gro central domains are essential for properly regulated transcriptional repression (2, 19). To shed light on the mechanism by which these domains function, we used them as affinity reagents to purify interacting proteins in *Drosophila* embryo nuclear extracts that were then identified by MudPIT. We identified over 160 interacting polypeptides, many of which associate with one another in a variety of multiprotein complexes. Several of these interacting proteins (e.g. the core histones and CKII) were previously characterized as Gro interactors, thus partially validating the screen. In addition, we validated the interaction between Gro and U1 snRNP by demonstrating the presence of U1 snRNA in an anti-Gro immunoprecipitate of embryonic nuclear extracts.

As a means of systematically validating interactions, we used a functional assay in *Drosophila* cells in which 157 of the interactors were each knocked down by RNAi to determine their requirement for Gal4-Gro-mediated repression of a luciferase reporter. In this way, we obtained evidence that 44 of the interactors have functional roles in Gro-mediated repression. 28 of these are required for repression, whereas 16 of them antagonize repression. The number 44 is probably an underestimate of the true number of functional interactors due to the artificiality of the reporter assay. For example, because we artificially recruit Gro to the reporter by tethering it to the Gal4 DNA binding domain, any interactions that work to help recruit Gro to the template will not be required. In addition, the reporters are introduced by transient transfection, and certain chromatin structures or modifications that contribute to Gro-mediated repression may not be reproduced in this context.

Gro Interactors Include Chromatin Remodelers, Protein Kinases, and Protein Complexes Involved in RNA Processing—Gro-mediated repression may be associated with changes in chromatin structure, including histone deacetylation and possibly increased nucleosome density (3, 18, 54). Consistent with this possibility, our proteomic screen identified a number of histone modifiers and ATP-dependent chromatin remodelers, including subunits of the ACF chromatin remodeling complexes (Acf1 and Iswi), the histone chaperone NAP1, and the histone kinases JIL-1 and Ball. Consistent with the idea that chromatin remodelers may be required for Gro-mediated repression by catalyzing changes in nucleosome density or

higher order chromatin structure, our reporter assay showed that Acf1 is required for optimal repression by Gro.

CKII is a heterotetrameric complex consisting of two copies of a catalytic subunit (CKII α) and two copies of a regulatory subunit (CKII β) (55, 56). A previous study showed that CKII phosphorylates Gro at multiple sites, including serines 239 and 253, to promote repression (21). We identified both the α and β subunits of CKII and the CKII negative regulator Nopp140 in our proteomic screen, but our findings are inconsistent with the view that CKII is a positive regulator of Gro and that Nopp140 acts by inhibiting CKII. This is because our reporter assays show that CKII α , CKII β , and Nopp140 are all negative regulators of Gro. However, our results are consistent with other findings showing that Gro phosphorylation can block repression (2). Furthermore, the effect we observed due to Nopp140 knockdown could reflect the role of this factor in processes other than CKII regulation (57).

In addition to several expected protein complexes, we also isolated many novel Gro-interacting proteins, one of which is the RNA helicase Rm62 (also known as p68). Rm62 is a DEAD box RNA helicase that has multiple functions, including roles in RNA processing, RNAi, and transcriptional regulation (58). Previous studies have shown a dual role for Rm62 in transcriptional regulation; its interaction with coactivator CBP/p300 may lead to gene activation (59), whereas its interaction with HDAC1 may lead to repression (60, 61). Our reporter assay confirms its function as a positive regulator of Gro-mediated repression as knocking down Rm62 resulted in attenuated Gro activity. Interestingly, Rm62 was also shown to be an essential splicing component through its action on the U1 snRNP (62, 63). The possible significance of the spliceosome in Gro-mediated repression is discussed below.

An Unanticipated Role for the Spliceosome in Gro-mediated Repression—One of the most surprising findings from our proteomic screen was the purification of a significant portion of the spliceosome complex, which suggests a potential role for the spliceosome in transcriptional regulation.

Pre-mRNA processing frequently occurs co-transcriptionally (64–66). Splicing factors are often recruited to nascent transcripts by the C-terminal domain of the Pol II large subunit and elongation factors (67, 68). In addition, there is evidence that co-activators are able to interact with splicing factors (27). The interaction between the transcriptional and splicing machinery may be functionally relevant because different promoters can yield transcripts that are subject to differential alternative splicing (69, 70). Although many studies have focused on the effect of transcription factors in splicing, there is also increasing evidence that promoter-proximal splicing elements can influence transcription (26, 28, 71).

U1 snRNP, a part of the spliceosome, consists of U1 snRNA, three U1 snRNP-specific proteins, and the seven-subunit Sm complex (46). Our list of 162 Gro-interacting proteins (supplemental Table S1, B and C) includes all three U1 snRNP-specific proteins (snRNP-U1C, snRNP-U1-70K, and snRNP-U1-A) as well as two subunits of the Sm complex (Sm-D2 and Sm-D3). We note that we also detected at least four other Sm complex subunits in one of the two replicate screens (Sm-B, Sm-F, Sm-D1, and Sm-G) (supplemental Table S1C). Additionally, we showed by co-immu-

Groucho/Spliceosome Interactions

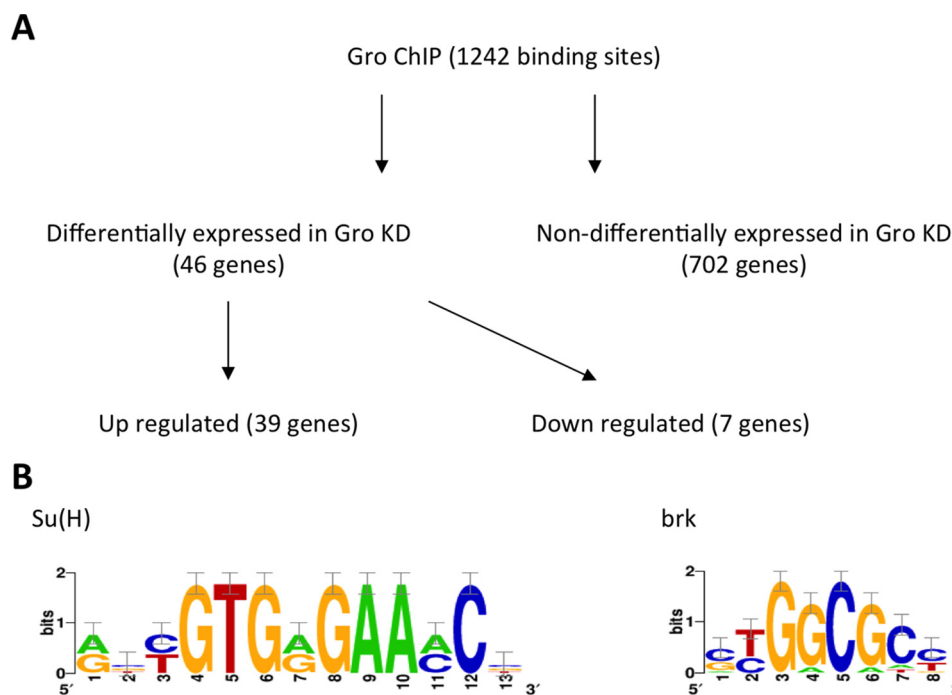


FIGURE 5. Gro binding regions in differentially expressed genes. A, S2 cell ChIP-seq data (49) identified 1,242 Gro binding sites, which map to 748 genes, 46 of which were differentially expressed when we knocked down Gro. Of the 46 differentially expressed genes, 39 were up-regulated and seven were down-regulated in response to Gro knockdown (KD). B, Gro binding regions in the 46 differentially expressed genes are significantly enriched for Su(H) and Brk binding sites.

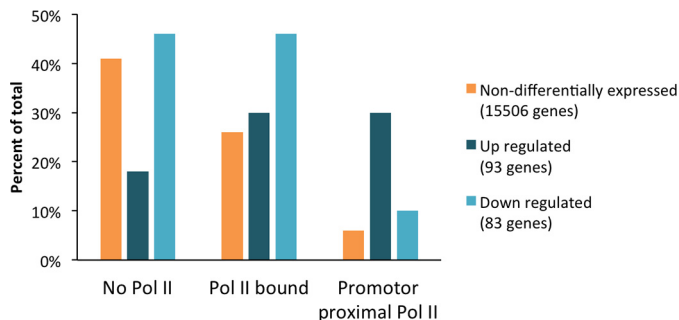


FIGURE 6. Gro-repressed genes are enriched for promoter-proximal Pol II. The percentage of non-differentially expressed genes and genes that are either up-regulated or down-regulated in Gro knockdown cells containing no Pol II bound or Pol II bound or enriched for promoter-proximal Pol II as ascertained by Pol II ChIP-chip analysis (53).

noprecipitation that ~13% of U1 snRNA, the RNA component of the U1 snRNP, is associated with Gro in embryonic nuclei. Thus, we have detected essentially the entire U1 snRNP in our proteomic screens for Gro-interacting proteins.

Data from our reporter assay suggest that the U1 snRNP complex is required for optimal Gro-mediated repression as snRNP-U1-C and snRNP-U1-70K knockdown attenuated repression. Consistent with our finding, it has been shown that snRNP-U1-C overexpression can decrease EWS/FLI-activated transcription (30). It is worth noting that the U1 snRNA is known to associate with transcription factor IIH and promote transcriptional initiation *in vitro* (29). Thus, the effect of the U1 snRNP complex in transcription regulation may be context-dependent.

Gro Recruitment Is Insufficient for Repression—The available S2 cell Gro ChIP-seq data (49) reveal 1,242 Gro binding sites in the S2 cell genome associated with 748 genes, whereas our RNA-seq analysis revealed that only 46 of these 748 genes are

differentially expressed in Gro knockdown S2 cells, implying that Gro binds to many genes that it does not regulate. The apparent contradiction could be explained by the absence of a required transcriptional activator in S2 cells to activate these genes upon Gro depletion. Regardless of the reason for the finding that Gro binds to many more genes than it regulates, this is a phenomenon that is common to many (perhaps most) eukaryotic gene-specific transcriptional regulators (51, 52). Gro ChIP-seq peaks associated with genes differentially expressed upon Gro knockdown are enriched for Su(H) and Brk binding motifs. This is in agreement with the known roles of Su(H) and Brk in the recruitment of Gro to target genes in the Notch and Dpp signaling pathways, respectively (72–74).

Genes that are up-regulated in Gro knockdown cells (and that are therefore candidate Gro repression targets) exhibit enrichment in Pol II pausing near the transcriptional start site. This finding is in agreement with the hypothesis that Pol II pausing is one mechanism to repress gene expression (75, 76). We note that our proteomic screen revealed the Pol II C-terminal domain kinase Cdk12 as a Gro-interacting protein (supplemental Table S1). By phosphorylating the C-terminal domain on Ser-2, Cdk12 may function to allow release of paused Pol II (77). Consistent with this idea, our reporter assay shows that Cdk12 functions to alleviate Gro-mediated repression (Table 5 and supplemental Table S2).

Genes that are differentially expressed in Gro and snRNP-U1-C knockdown cells are enriched for H3K36me1 as well as the H3K36 methyltransferase ASH1. Although H3K36me is involved in multiple functions, including transcriptional regulation, splicing, and DNA repair (78, 79), these findings suggest a previously unknown role for this histone mark in Gro-mediated repression.

The Gro Central Region as a Regulatory Hub of Repression Activity—In conclusion, our findings reinforce the idea that the Gro central domains, which are intrinsically disordered, are indispensable for repression (19). Previous studies from our laboratory and other laboratories show that the GP domain interacts with the histone deacetylase Rpd3/HDAC1, which may promote local histone deacetylation and alter nucleosome density (16, 18). The identification of the ACF chromatin remodeling complexes as a central region-interacting protein complex and our demonstration that knockdown of this protein attenuates Gro-mediated repression provide further support for the idea that regulation of chromatin structure is a critical aspect of Gro-mediated repression. Conversely, modulation of chromatin structure is likely not the only mechanism of Gro-mediated repression as histone deacetylase inhibitors and Rpd3 knockdown reduce, but do not abolish, Gro-mediated repression (16, 18) (Fig. 3C). Through a combination of proteomic screening, reporter assays, and genome-wide expression profiling, our results suggest a possible new mechanism of Gro-mediated repression involving the action of the spliceosome. Future experiments will focus on elucidating the underlying mechanisms by which these interacting partners act in Gro-mediated repression.

Author Contributions—A. J. C., P. N. K., and W. T.-J. conceived and planned the study, which was coordinated by A. J. C. P. N. K. and A. J. C. wrote the manuscript. P. N. K. and W. T.-J. conducted most of the experiments. T. Y. Y. assisted with the co-immunoprecipitation study. M. C. carried out the bioinformatics analysis. A. A. V. and J. A. W. carried out the MudPIT analysis. All authors reviewed and approved the manuscript.

Acknowledgments—We thank Robert Damoiseaux, Scientific Director of the UCLA Molecular Screening Shared Resource Center, for assistance in developing the mechanized multiple luciferase reporter screen. We are grateful to all the members of the Courey laboratory (past and present) for insight and advice.

References

- Courey, A. J., and Jia, S. (2001) Transcriptional repression: the long and the short of it. *Genes Dev.* **15**, 2786–2796
- Turki-Judeh, W., and Courey, A. J. (2012) Groucho: a corepressor with instructive roles in development. *Curr. Top. Dev. Biol.* **98**, 65–96
- Martinez, C. A., and Arnosti, D. N. (2008) Spreading of a corepressor linked to action of long-range repressor hairy. *Mol. Cell. Biol.* **28**, 2792–2802
- Chen, G., and Courey, A. J. (2000) Groucho/TLE family proteins and transcriptional repression. *Gene* **249**, 1–16
- Gasperowicz, M., and Otto, F. (2005) Mammalian Groucho homologs: redundancy or specificity? *J. Cell. Biochem.* **95**, 670–687
- Nagel, A. C., Krejci, A., Tenin, G., Bravo-Patiño, A., Bray, S., Maier, D., and Preiss, A. (2005) Hairless-mediated repression of notch target genes requires the combined activity of Groucho and CtBP corepressors. *Mol. Cell. Biol.* **25**, 10433–10441
- Blair, S. S. (2007) Wing vein patterning in *Drosophila* and the analysis of intercellular signaling. *Annu. Rev. Cell Dev. Biol.* **23**, 293–319
- Hasson, P., and Paroush, Z. (2006) Crosstalk between the EGFR and other signalling pathways at the level of the global transcriptional corepressor Groucho/TLE. *Br. J. Cancer* **94**, 771–775
- Allen, T., van Tuyl, M., Iyengar, P., Jothy, S., Post, M., Tsao, M. S., and Lobe, C. G. (2006) Grg1 acts as a lung-specific oncogene in a transgenic mouse model. *Cancer Res.* **66**, 1294–1301
- Buscarlet, M., and Stifani, S. (2007) The ‘Marx’ of Groucho on development and disease. *Trends Cell Biol.* **17**, 353–361
- Jennings, B. H., Pickles, L. M., Wainwright, S. M., Roe, S. M., Pearl, L. H., and Ish-Horowicz, D. (2006) Molecular recognition of transcriptional repressor motifs by the WD domain of the Groucho/TLE corepressor. *Mol. Cell* **22**, 645–655
- Song, H., Hasson, P., Paroush, Z., and Courey, A. J. (2004) Groucho oligomerization is required for repression *in vivo*. *Mol. Cell. Biol.* **24**, 4341–4350
- Chodaparambil, J. V., Pate, K. T., Hepler, M. R., Tsai, B. P., Muthurajan, U. M., Luger, K., Waterman, M. L., and Weis, W. I. (2014) Molecular functions of the TLE tetramerization domain in Wnt target gene repression. *EMBO J.* **33**, 719–731
- Chen, G., Nguyen, P. H., and Courey, A. J. (1998) A role for Groucho tetramerization in transcriptional repression. *Mol. Cell. Biol.* **18**, 7259–7268
- Pinto, M., and Lobe, C. G. (1996) Products of the grg (Groucho-related gene) family can dimerize through the amino-terminal Q domain. *J. Biol. Chem.* **271**, 33026–33031
- Chen, G., Fernandez, J., Mische, S., and Courey, A. J. (1999) A functional interaction between the histone deacetylase Rpd3 and the corepressor groucho in *Drosophila* development. *Genes Dev.* **13**, 2218–2230
- Choi, C. Y., Kim, Y. H., Kwon, H. J., and Kim, Y. (1999) The homeodomain protein NK-3 recruits Groucho and a histone deacetylase complex to repress transcription. *J. Biol. Chem.* **274**, 33194–33197
- Winkler, C. J., Ponce, A., and Courey, A. J. (2010) Groucho-mediated repression may result from a histone deacetylase-dependent increase in nucleosome density. *PLoS One* **5**, e10166
- Turki-Judeh, W., and Courey, A. J. (2012) The unconserved groucho central region is essential for viability and modulates target gene specificity. *PLoS One* **7**, e30610
- Hasson, P., Egoz, N., Winkler, C., Volohonsky, G., Jia, S., Dinur, T., Volk, T., Courey, A. J., and Paroush, Z. (2005) EGFR signaling attenuates Groucho-dependent repression to antagonize Notch transcriptional output. *Nat. Genet.* **37**, 101–105
- Nuthall, H. N., Joachim, K., and Stifani, S. (2004) Phosphorylation of serine 239 of Groucho/TLE1 by protein kinase CK2 is important for inhibition of neuronal differentiation. *Mol. Cell. Biol.* **24**, 8395–8407
- Dunker, A. K., Cortese, M. S., Romero, P., Iakoucheva, L. M., and Uversky, V. N. (2005) Flexible nets. The roles of intrinsic disorder in protein interaction networks. *FEBS J.* **272**, 5129–5148
- Uversky, V. N., and Dunker, A. K. (2010) Understanding protein non-folding. *Biochim. Biophys. Acta* **1804**, 1231–1264
- Tantos, A., Han, K. H., and Tompa, P. (2012) Intrinsic disorder in cell signaling and gene transcription. *Mol. Cell. Endocrinol.* **348**, 457–465
- Bondos, S. E., and Hsiao, H. C. (2012) Roles for intrinsic disorder and fuzziness in generating context-specific function in Ultrabithorax, a Hox transcription factor. *Adv. Exp. Med. Biol.* **725**, 86–105
- Damgaard, C. K., Kahns, S., Lykke-Andersen, S., Nielsen, A. L., Jensen, T. H., and Kjems, J. (2008) A 5' splice site enhances the recruitment of basal transcription initiation factors *in vivo*. *Mol. Cell* **29**, 271–278
- Monsalve, M., Wu, Z., Adelmant, G., Puigserver, P., Fan, M., and Spiegelman, B. M. (2000) Direct coupling of transcription and mRNA processing through the thermogenic coactivator PGC-1. *Mol. Cell* **6**, 307–316
- Furger, A., O'Sullivan, J. M., Binnie, A., Lee, B. A., and Proudfoot, N. J. (2002) Promoter proximal splice sites enhance transcription. *Genes Dev.* **16**, 2792–2799
- Kwek, K. Y., Murphy, S., Furger, A., Thomas, B., O'Gorman, W., Kimura, H., Proudfoot, N. J., and Akoulitchev, A. (2002) U1 snRNA associates with TFIIF and regulates transcriptional initiation. *Nat. Struct. Biol.* **9**, 800–805
- Knoop, L. L., and Baker, S. J. (2000) The splicing factor U1C represses EWS/FLI-mediated transactivation. *J. Biol. Chem.* **275**, 24865–24871
- Armknecht, S., Boutros, M., Kiger, A., Nybakken, K., Mathey-Prevot, B., and Perrimon, N. (2005) High-throughput RNA interference screens in *Drosophila* tissue culture cells. *Methods Enzymol.* **392**, 55–73
- Soeller, W. C., Poole, S. J., and Kornberg, T. (1988) *In vitro* transcription of the *Drosophila* engrailed gene. *Genes Dev.* **2**, 68–81
- Wohlschlegel, J. A. (2009) Identification of SUMO-conjugated proteins and their SUMO attachment sites using proteomic mass spectrometry.

- Methods Mol. Biol.* **497**, 33–49
34. Zhou, R., Mohr, S., Hannon, G. J., and Perrimon, N. (2013) Inducing RNAi in *Drosophila* cells by transfection with dsRNA. *Cold Spring Harb. Protoc.* **2013**, 461–463
 35. Kim, D., Pertea, G., Trapnell, C., Pimentel, H., Kelley, R., and Salzberg, S. L. (2013) TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. *Genome Biol.* **14**, R36
 36. Love, M. I., Huber, W., and Anders, S. (2014) Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* **15**, 550
 37. Herrmann, C., Van de Sande, B., Potier, D., and Aerts, S. (2012) i-cisTarget: an integrative genomics method for the prediction of regulatory features and cis-regulatory modules. *Nucleic Acids Res.* **40**, e114
 38. Lyne, R., Smith, R., Rutherford, K., Wakeling, M., Varley, A., Guillier, F., Janssens, H., Ji, W., McLaren, P., North, P., Rana, D., Riley, T., Sullivan, J., Watkins, X., Woodbridge, M., Lilley, K., Russell, S., Ashburner, M., Mizuguchi, K., and Micklem, G. (2007) FlyMine: an integrated database for *Drosophila* and *Anopheles* genomics. *Genome Biol.* **8**, R129
 39. Brantjes, H., Roose, J., van De Wetering, M., and Clevers, H. (2001) All Tcf/HMG box transcription factors interact with Groucho-related co-repressors. *Nucleic Acids Res.* **29**, 1410–1419
 40. Daniels, D. L., and Weis, W. I. (2005) β -Catenin directly displaces Groucho/TLE repressors from Tcf/Lef in Wnt-mediated transcription activation. *Nat. Struct. Mol. Biol.* **12**, 364–371
 41. Bouazoune, K., and Brehm, A. (2006) ATP-dependent chromatin remodeling complexes in *Drosophila*. *Chromosome Res.* **14**, 433–449
 42. Flores-Saaib, R. D., and Courey, A. J. (2000) Analysis of Groucho-histone interactions suggests mechanistic similarities between Groucho- and Tup1-mediated repression. *Nucleic Acids Res.* **28**, 4189–4196
 43. Sekiya, T., and Zaret, K. S. (2007) Repression by Groucho/TLE/Grg proteins: genomic site recruitment generates compacted chromatin *in vitro* and impairs activator binding *in vivo*. *Mol. Cell* **28**, 291–303
 44. Edmondson, D. G., Smith, M. M., and Roth, S. Y. (1996) Repression domain of the yeast global repressor Tup1 interacts directly with histones H3 and H4. *Genes Dev.* **10**, 1247–1259
 45. Patel, S. B., and Bellini, M. (2008) The assembly of a spliceosomal small nuclear ribonucleoprotein particle. *Nucleic Acids Res.* **36**, 6482–6493
 46. Will, C. L., and Lührmann, R. (2011) Spliceosome structure and function. *Cold Spring Harb. Perspect. Biol.* **3**, a003707
 47. Fisher, A. L., Ohsako, S., and Caudy, M. (1996) The WRPW motif of the hairy-related basic helix-loop-helix repressor proteins acts as a 4-amino-acid transcription repression and protein-protein interaction domain. *Mol. Cell Biol.* **16**, 2670–2677
 48. Thorne, N., Inglese, J., and Auld, D. S. (2010) Illuminating insights into firefly luciferase and other bioluminescent reporters used in chemical biology. *Chem. Biol.* **17**, 646–657
 49. Kaul, A., Schuster, E., and Jennings, B. H. (2014) The Groucho co-repressor is primarily recruited to local target sites in active chromatin to attenuate transcription. *PLoS Genet.* **10**, e1004595
 50. Celniker, S. E., Dillon, L. A., Gerstein, M. B., Gunsalus, K. C., Henikoff, S., Karpen, G. H., Kellis, M., Lai, E. C., Lieb, J. D., MacAlpine, D. M., Micklem, G., Piano, F., Snyder, M., Stein, L., White, K. P., Waterston, R. H., and modENCODE Consortium (2009) Unlocking the secrets of the genome. *Nature* **459**, 927–930
 51. Fisher, W. W., Li, J. J., Hammonds, A. S., Brown, J. B., Pfeiffer, B. D., Weiszmann, R., MacArthur, S., Thomas, S., Stamatoyannopoulos, J. A., Eisen, M. B., Bickel, P. J., Biggin, M. D., and Celniker, S. E. (2012) DNA regions bound at low occupancy by transcription factors do not drive patterned reporter gene expression in *Drosophila*. *Proc. Natl. Acad. Sci. U.S.A.* **109**, 21330–21335
 52. Walhout, A. J. (2011) What does biologically meaningful mean? A perspective on gene regulatory network validation. *Genome Biol.* **12**, 109
 53. Muse, G. W., Gilchrist, D. A., Nechaev, S., Shah, R., Parker, J. S., Grissom, S. F., Zeitlinger, J., and Adelman, K. (2007) RNA polymerase is poised for activation across the genome. *Nat. Genet.* **39**, 1507–1511
 54. Li, L. M., and Arnosti, D. N. (2011) Long- and short-range transcriptional repressors induce distinct chromatin states on repressed genes. *Curr. Biol.* **21**, 406–412
 55. Niefind, K., Guerra, B., Ermakowa, I., and Issinger, O. G. (2001) Crystal structure of human protein kinase CK2: insights into basic properties of the CK2 holoenzyme. *EMBO J.* **20**, 5320–5331
 56. Niefind, K., Raaf, J., and Issinger, O. G. (2009) Protein kinase CK2 in health and disease: protein kinase CK2: from structures to insights. *Cell. Mol. Life Sci.* **66**, 1800–1816
 57. Meier, U. T., and Blobel, G. (1992) Nopp140 shuttles on tracks between nucleolus and cytoplasm. *Cell* **70**, 127–138
 58. Fuller-Pace, F. V. (2006) DEXD/H box RNA helicases: multifunctional proteins with important roles in transcriptional regulation. *Nucleic Acids Res.* **34**, 4206–4215
 59. Rossow, K. L., and Janknecht, R. (2003) Synergism between p68 RNA helicase and the transcriptional coactivators CBP and p300. *Oncogene* **22**, 151–156
 60. Buszczak, M., and Spradling, A. C. (2006) The *Drosophila* P68 RNA helicase regulates transcriptional deactivation by promoting RNA release from chromatin. *Genes Dev.* **20**, 977–989
 61. Wilson, B. J., Bates, G. J., Nicol, S. M., Gregory, D. J., Perkins, N. D., and Fuller-Pace, F. V. (2004) The p68 and p72 DEAD box RNA helicases interact with HDAC1 and repress transcription in a promoter-specific manner. *BMC Mol. Biol.* **5**, 11
 62. Neubauer, G., King, A., Rappsilber, J., Calvio, C., Watson, M., Ajuh, P., Sleeman, J., Lamond, A., and Mann, M. (1998) Mass spectrometry and EST-database searching allows characterization of the multi-protein spliceosome complex. *Nat. Genet.* **20**, 46–50
 63. Liu, Z. R. (2002) p68 RNA helicase is an essential human splicing factor that acts at the U1 snRNA-5' splice site duplex. *Mol. Cell Biol.* **22**, 5443–5450
 64. Reed, R. (2003) Coupling transcription, splicing and mRNA export. *Curr. Opin. Cell Biol.* **15**, 326–331
 65. Han, J., Xiong, J., Wang, D., and Fu, X. D. (2011) Pre-mRNA splicing: where and when in the nucleus. *Trends Cell Biol.* **21**, 336–343
 66. Bentley, D. L. (2005) Rules of engagement: co-transcriptional recruitment of pre-mRNA processing factors. *Curr. Opin. Cell Biol.* **17**, 251–256
 67. Maniatis, T., and Reed, R. (2002) An extensive network of coupling among gene expression machines. *Nature* **416**, 499–506
 68. Bentley, D. (2002) The mRNA assembly line: transcription and processing machines in the same factory. *Curr. Opin. Cell Biol.* **14**, 336–342
 69. Cramer, P., Cáceres, J. F., Cazalla, D., Kadener, S., Muro, A. F., Baralle, F. E., and Kornblihtt, A. R. (1999) Coupling of transcription with alternative splicing: RNA pol II promoters modulate SF2/ASF and 9G8 effects on an exonic splicing enhancer. *Mol. Cell* **4**, 251–258
 70. Cramer, P., Pesce, C. G., Baralle, F. E., and Kornblihtt, A. R. (1997) Functional association between promoter structure and transcript alternative splicing. *Proc. Natl. Acad. Sci. U.S.A.* **94**, 11456–11460
 71. Merkhofer, E. C., Hu, P., and Johnson, T. L. (2014) Introduction to cotranscriptional RNA splicing. *Methods Mol. Biol.* **1126**, 83–96
 72. Hasson, P., Müller, B., Basler, K., and Paroush, Z. (2001) Brinker requires two corepressors for maximal and versatile repression in Dpp signalling. *EMBO J.* **20**, 5725–5736
 73. Paroush, Z., Finley, R. L., Jr., Kidd, T., Wainwright, S. M., Ingham, P. W., Brent, R., and Ish-Horowitz, D. (1994) Groucho is required for *Drosophila* neurogenesis, segmentation, and sex determination and interacts directly with hairy-related bHLH proteins. *Cell* **79**, 805–815
 74. Grbavec, D., and Stifani, S. (1996) Molecular interaction between TLE1 and the carboxyl-terminal domain of HES-1 containing the WRPW motif. *Biochem. Biophys. Res. Commun.* **223**, 701–705
 75. Adelman, K., and Lis, J. T. (2012) Promoter-proximal pausing of RNA polymerase II: emerging roles in metazoans. *Nat. Rev. Genet.* **13**, 720–731
 76. Gaertner, B., and Zeitlinger, J. (2014) RNA polymerase II pausing during development. *Development* **141**, 1179–1183
 77. Bartkowiak, B., Liu, P., Phatnani, H. P., Fuda, N. J., Cooper, J. J., Price, D. H., Adelman, K., Lis, J. T., and Greenleaf, A. L. (2010) CDK12 is a transcription elongation-associated CTD kinase, the metazoan ortholog of yeast Ctk1. *Genes Dev.* **24**, 2303–2316
 78. Lee, J. S., and Shilatifard, A. (2007) A site to remember: H3K36 methylation a mark for histone deacetylation. *Mutat. Res.* **618**, 130–134
 79. Wagner, E. J., and Carpenter, P. B. (2012) Understanding the language of Lys36 methylation at histone H3. *Nat. Rev. Mol. Cell Biol.* **13**, 115–126