nature structural & molecular biology

ARTICLES

https://doi.org/10.1038/s41594-018-0118-5

# Architecture of Pol II(G) and molecular mechanism of transcription regulation by Gdown1

Miki Jishage[1], Xiaodi Yu[2,5], Yi Shi[3,6], Sai J. Ganesan[4], Wei-Yi Chen[1,7], Andrej Sali[4], Brian T. Chait[3], Francisco J. Asturias[2,8] and Robert G. Roeder[1*]

**Tight binding of Gdown1 represses RNA polymerase II (Pol II) function in a manner that is reversed by Mediator, but the structural basis of these processes is unclear. Although Gdown1 is intrinsically disordered, its Pol II interacting domains were localized and shown to occlude transcription factor IIF (TFIIF) and transcription factor IIB (TFIIB) binding by perfect positioning on their Pol II interaction sites. Robust binding of Gdown1 to Pol II is established by cooperative interactions of a strong Pol II binding region and two weaker binding modulatory regions, thus providing a mechanism both for tight Pol II binding and transcription inhibition and for its reversal. In support of a physiological function for Gdown1 in transcription repression, Gdown1 co-localizes with Pol II in transcriptionally silent nuclei of early *Drosophila* embryos but re-localizes to the cytoplasm during zygotic genome activation. Our study reveals a self-inactivation through Gdown1 binding as a unique mode of repression in Pol II function.**

In eukaryotes, Pol II is the central machine responsible for transcription of genes that specify messenger RNAs, small nuclear RNAs, and microRNAs. In response to signals that result in expression of specific sets of genes, Pol II is recruited to cognate gene promoters through the concerted functions of transcription activators, co-activators that include the Mediator, and the general transcription factors (GTFs)[1]. The GTFs (TFIIB, -D, -E, -F, and -H) play essential roles in transcription through interactions with Pol II at core promoters during preinitiation complex (PIC) formation and function (initiation)[2]. The large (25–30 subunit) Mediator complex acts principally in the regulation of transcription through joint enhancer-bound activator and Pol II interactions that facilitate both Pol II recruitment and enhancer-promoter interactions[3–5]. However, the Mediator has multiple functions in transcription that are not fully understood[3,4].

While conventional Pol II is composed of 12 subunits (POLR2A to POLR2L), Pol II(G) is a newly discovered form of Pol II containing a stoichiometric, tightly associated Gdown1 subunit (POLR2M)[6]. Although Pol II and Mediator are highly conserved from yeast to human[7], Gdown1 is metazoan-specific[6]; and its ubiquitous expression[8,9] suggests a global role in transcriptional regulation. However, the biological roles of Gdown1 remain largely unknown. In vitro studies revealed that in the absence of Mediator, Pol II(G), unlike Pol II, inhibits transcription, suggesting a primary function of Gdown1 in transcription repression. This effect is attenuated in the presence of Mediator[6], indicating that Mediator somehow reverses the inhibitory effect of Gdown1. Thus, the formation of Pol II(G) elicits an enhanced dependence on Mediator for transcription. Although the underlying mechanism is unclear, the Mediator-dependent Pol II(G)-mediated transcription appears to differ from Pol II-mediated transcription.

In vitro studies also revealed that Gdown1 prevents TFIIF from associating with Pol II, which in turns leads to inhibition of PIC assembly[10]. TFIIF plays a critical role in transcription initiation[2,7] by stabilizing the PIC through interactions with other GTFs (including TFIIB and TFIIE) as well as promoter DNA[11]. However, the fact that Pol II interactions with Gdown1 and TFIIF are mutually exclusive indicates that for successful PIC formation, either Gdown1 must be dissociated from Pol II or, alternatively, the inhibitory effect of Gdown1 on TFIIF binding to Pol II must somehow be neutralized (for example, by structural changes). Our previous studies were unable to deduce the fate of Gdown1 when Pol II(G)-mediated repression was reversed due to the extremely strong binding to Pol II[10].

Here, we examine in detail both physical and functional interactions between Pol II and Gdown1. We map Gdown1 locations on Pol II by cryo-electron microscopy (cryo-EM), chemical cross-linking with mass spectrometry readout (CX-MS), and integrative modeling approaches, and further establish Gdown1 Pol II interaction regions by biochemical analyses. We find that Gdown1 locations on Pol II perfectly overlap contact sites for TFIIF and TFIIB. Moreover, we identify mutations in Gdown1 that bypass the Mediator requirement for reversal of the repression. Analyses of these mutants show that the primary Pol II binding region(s) of Gdown1 are regulated by two essential regulatory regions that further stabilize Gdown1–Pol II interactions. An integrative modeling approach uniquely reveals the location of these regulatory regions, providing insights into the molecular mechanisms underlying both the robust binding of Gdown1 to Pol II and the Mediator-facilitated dissociation of Gdown1 from the Pol II binding site. This model is consistent with the further demonstration of an essential role of Gdown1 in *Drosophila melanogaster* early embryonic development,

[1]Laboratory of Biochemistry and Molecular Biology, The Rockefeller University, New York, NY, USA. [2]Department of Integrative Structural and Computational Biology, The Scripps Research Institute, San Diego, CA, USA. [3]Laboratory of Mass Spectrometry and Gaseous Ion Chemistry, The Rockefeller University, New York, NY, USA. [4]Department of Bioengineering and Therapeutic Sciences, Department of Pharmaceutical Chemistry, California Institute for Quantitative Biosciences, University of California, San Francisco, San Francisco, CA, USA. [5]Present address: Pfizer Inc., Groton, CT, USA. [6]Present address: Department of Cell Biology, University of Pittsburgh School of Medicine, Pittsburgh, PA, USA. [7]Present address: Institute of Biochemistry and Molecular Biology, National Yang-Ming University, Taipei, Taiwan. [8]Present address: Department of Biochemistry and Molecular Genetics, University of Colorado School of Medicine, Aurora, CO, USA. *e-mail: roeder@mail.rockefeller.edu

where Gdown1 co-localization with Pol II in nuclei is inversely correlated with active gene transcription.

## Results

**Molecular architecture of Pol II(G) based on cryo-EM and CX-MS analyses.** Gdown1 and TFIIF associate with Pol II in a mutually exclusive manner, which prevents efficient formation of PIC and leads ultimately to inhibition of transcription[10]. To further elucidate the molecular mechanism of the inhibition, we first set out to visualize Gdown1 location on Pol II by single-particle cryo-EM. Human Pol II(G) was purified from nuclear extracts of cells expressing flag-tagged Gdown1 (ref. [10]), and Pol II(G) cryo-EM images were computationally screened and clustered to obtain a cryo-EM map of Pol II(G) at ~4 Å resolution (Table 1, Methods, Supplementary Fig. 1). All major Pol II domains were well resolved in the Pol II(G) cryo-EM map, with the exception of the clamp and RPB4 and RPB7 subunits (Fig. 1a and Supplementary Fig. 1d,e) that appeared blurred (possibly due to high mobility). An atomic model of human Pol II was obtained by real-space refinement of a model derived from relevant portions of the published model of bovine Pol II[12]. To identify Gdown1 density in the Pol II(G) map, we calculated a difference map between the Pol II(G) cryo-EM map and a 4 Å molecular map derived from corresponding portions of the bovine Pol II atomic model. Three major densities were observed: (1) around RPB3 and RPB10; (2) at the RPB2 protrusion and wall domains; and (3) at the RPB1 dock domain (Fig. 1a).

To identify which Gdown1 residues might correspond to Gdown1 densities in the Pol II(G) cryo-EM map, we performed CX-MS analyses[13,14] with purified Pol II(G) (Methods and Supplementary Fig. 1a), and obtained 40 cross-links (Methods and Supplementary Table 1) that were used to build maps of Gdown1–Pol II inter-molecular interactions (Fig. 1b) and Gdown1 intra-molecular interactions (Fig. 1c). Inter-molecular cross-links were detected in Pol II subunits RPB1, RPB2, RPB3, and RPB10. With the exception of RPB1 dock contacts, these cross-links were in accord with cryo-EM data.

**Table 1 | Cryo-EM data collection, refinement, and validation statistics**

| | Pol II(G) (EMD-7997, PDB 6DRD) |
|---|---|
| **Data collection and processing** | |
| Magnification | 22,500 |
| Voltage (kV) | 300 |
| Electron exposure (e–/Å$^2$) | 40 |
| Defocus range (μm) | –1.5 to –2.8 |
| Pixel size (Å) | 1.31 |
| Symmetry imposed | *C*1 |
| Initial particle images (no.) | 201,527 |
| Final particle images (no.) | 141,619 |
| Map resolution (Å) | 3.7 |
| FSC threshold | 0.143 |
| Map resolution range (Å) | 3.5 to 7.5 |
| **Refinement** | |
| Initial model used | PDB 5FLM |
| Model resolution (Å) | 4 |
| FSC threshold | 0.5 |
| Model resolution range (Å) | |
| Map sharpening *B* factor (Å$^2$) | –156 |
| Model composition | |
| Non-hydrogen atoms | - |
| Protein residues | 3,349 |
| Ligands | 5 |
| *B* factors (Å$^2$) | |
| Protein | 59.53 |
| Ligand | 146 |
| R.m.s. deviations | |
| Bond lengths (Å) | 0.01 |
| Bond angles (°) | 1.11 |
| Validation | |
| MolProbity score | 1.99 |
| Clashscore | 7.47 |
| Poor rotamers (%) | 0.68 |
| Ramachandran plot | |
| Favored (%) | 88.64 |
| Allowed (%) | 11.21 |
| Disallowed (%) | 0.15 |

**Gdown1 interaction sites overlap TFIIF and TFIIB contact sites on Pol II.** Relative to effects of Gdown1 on its function, TFIIF is composed of Rap74 and Rap30 subunits that form a hetero-dimer whose functional domains, dimerization domain, and WH domain are connected by the Rap30 linker domain[7] (Fig. 2a). Strikingly, Gdown1 densities on Pol II overlapped perfectly with the interaction site of the Rap30 linker domain on RPB2 (ref. [11]) (Fig. 2b) and also with the interaction site of the TFIIF dimerization domain on the RPB2 lobe[15–17] (Supplementary Fig. 2a). In addition, four lysines in the Gdown1 C-terminus cross-linked to the RPB2 protrusion near the binding site of the Rap30 WH domain[18] (Figs. 1b and 2c). This perfect overlap unequivocally confirms, and further details, the molecular basis of Gdown1 inhibition of TFIIF binding to Pol II[10].

Our cryo-EM analyses also revealed Gdown1 density overlaps with the Pol II interaction sites (RPB1 dock and RPB2 wall domains) of the TFIIB B-core (Fig. 2a and Supplementary Fig. 2b) and B-ribbon domains (Fig. 2d)[19,20]. Consistent with these results, we observed cross-linking of the Gdown1 C-terminal region to residues (K820, K821) in the RPB2 wall domain (Fig. 1b and Fig. 2d); and although a corresponding density was not detected, the Gdown1 C-terminal region was also cross-linked to residues (K213, K331) in the RPB1 clamp domain (Fig. 1b and Supplementary Fig. 2c) that is recognized by the TFIIB B-linker[19,20]. These striking coincidences of Gdown1 C-terminal region and TFIIB interaction sites on Pol II prompted us to examine whether Gdown1 could also inhibit TFIIB binding to Pol II. To this end, we performed a Pol II binding assay with $^{35}$S-labeled TFIIB and an in vitro transcription assay with a premelted promoter template that retains the TFIIB requirement for transcription but partially bypasses the normal TFIIF requirement[21] and, correspondingly, decreases Gdown1 transcriptional inhibition through TFIIF[22]. Notably, the results show that Gdown1 indeed interferes with the interaction of TFIIB with Pol II and inhibits transcription (Supplementary Fig. 2d,e). Since the Pol II clamp loop rudder extends into the active site, directly interacts with DNA and facilitates separation of the RNA transcript from DNA[23], these results also suggest a hindrance to the Gdown1 C-terminal interaction with this domain once RNA synthesis has begun.

**Identification of Gdown1 domains that stabilize the interaction of Pol II binding region I with Pol II.** We next determined Pol II binding and functional domains of Gdown1 by biochemical studies. GST-pulldown (GST: glutathione-S-transferase) assays showed robust binding of the Gdown1 C-terminal fragment (181–368) to Pol II, but only weak binding of the N-terminal fragment (and only

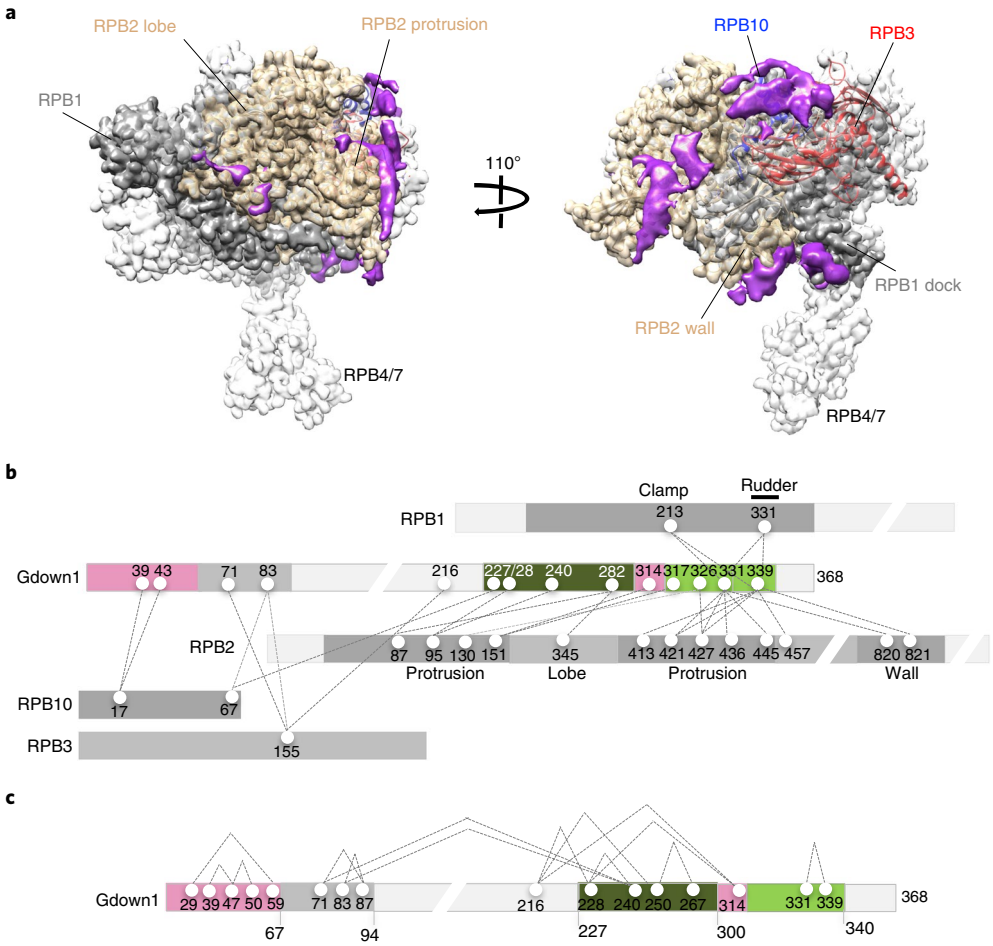


**Fig. 1 | Molecular architecture of Pol II(G) by cryo-EM and CX-MS analyses. a**, Cryo-EM structure of Pol II(G). Non-Pol II (Gdown1) density is shown in purple. An atomic model of human Pol II was obtained by real-space refinement of a model derived from relevant portions of the published model of bovine Pol II (PDB 5FLM). To identify Gdown1 density in the Pol II(G) map, the difference between all stable portions of the Pol II(G) cryo-EM map was calculated. **b**, Diagram of cross-links between Pol II subunits and Gdown1. Purified Pol II(G) was subjected to cross-linking with the amine-specific cross-linker, disuccinimidyl suberate (DSS), followed by high-resolution MS. **c**, Intra-molecular cross-links map of Gdown1.

when fused C-terminally to GST) (Supplementary Fig. 3a,b). In complementary in vitro transcription assays (Supplementary Fig. 3c), the C-terminal fragment, but not the N-terminal fragment, showed a dose-dependent inhibition that was significantly less than that observed for full-length, suggesting that while the N-terminal region (analyzed further below) cannot act alone it somehow contributes to the intrinsic inhibitory activity of the C-terminal region. In a further analysis of the C-terminus based in part on conserved amino acids, we generated various C-terminal deletion mutants and tested their effects on basal transcription and Pol II binding (Supplementary Fig. 3a,d–f). Fragments containing C-terminal deletions to residue 314 maintained basal transcription inhibitory activity (Supplementary Fig. 3d), whereas fragments with further deletions either lost inhibitory activity while maintaining significant (but not full) Pol II binding (fragments 1–298 and 1–269) or lost both Pol II binding and inhibitory activity (fragment 1–226) (Supplementary Fig. 3e,f). Based on these results, high-sequence conservation (Fig. 3c), and further results in Fig. 3d (discussed below), we identify residues 299–314 as a C-terminal transcription inhibitory region (C-TIR) important for Gdown1 inhibitory activity but not for general Pol II binding. Based on several Gdown1 deletion and binding studies (Supplementary Fig. 3f–j), we define Gdown1(227–298) as Pol II binding region I. Gdown1(315–340) is defined as Pol II binding region II based on cross-linking results (Fig. 1b). Although its Pol II binding activity is significantly weaker than that of Pol II binding region I (Supplementary Fig. 3g, also see below), the C-terminal region(315–368) that includes Pol II binding region II is clearly involved in the inhibitory activity (Supplementary Fig. 3d,e).

In a further analysis of the C-TIR region by mutation of conserved residues, full-length Gdown1 with L303A and L304A mutations showed a near-complete loss of transcriptional inhibitory activity (Fig. 3d), while retaining full Pol II binding activity (Fig. 3e). These results indicate that the two hydrophobic residues, and thus C-TIR, are critical for Gdown1 inhibitory activity and, further, that the loss of Gdown1 activity is not due to an overall loss of Pol II binding. Therefore, we postulated that C-TIR might have qualitative and/or quantitative effects on one or both of the flanking Pol II binding regions. In this regard, the C-TIR L303/4A mutation significantly reduced the binding to Pol II of the Gdown1(227–368) and (194–368) fragments containing Pol II binding regions I and II (Supplementary Fig. 3h,j, and Fig. 3g, lane 9)—consistent with its effect on Gdown1 function and a role in altering specific Pol II interactions or functional properties of Pol II binding region I or II (note that the severe effect of the L303/304 mutation on binding of the Gdown1(194–368) fragment (Fig. 3g, lane 9) relative to the full-length Gdown1 (Fig. 3e, lane 6) reflects the loss of N-terminal sequences that, like C-TIR, also stabilize binding of the C-terminal

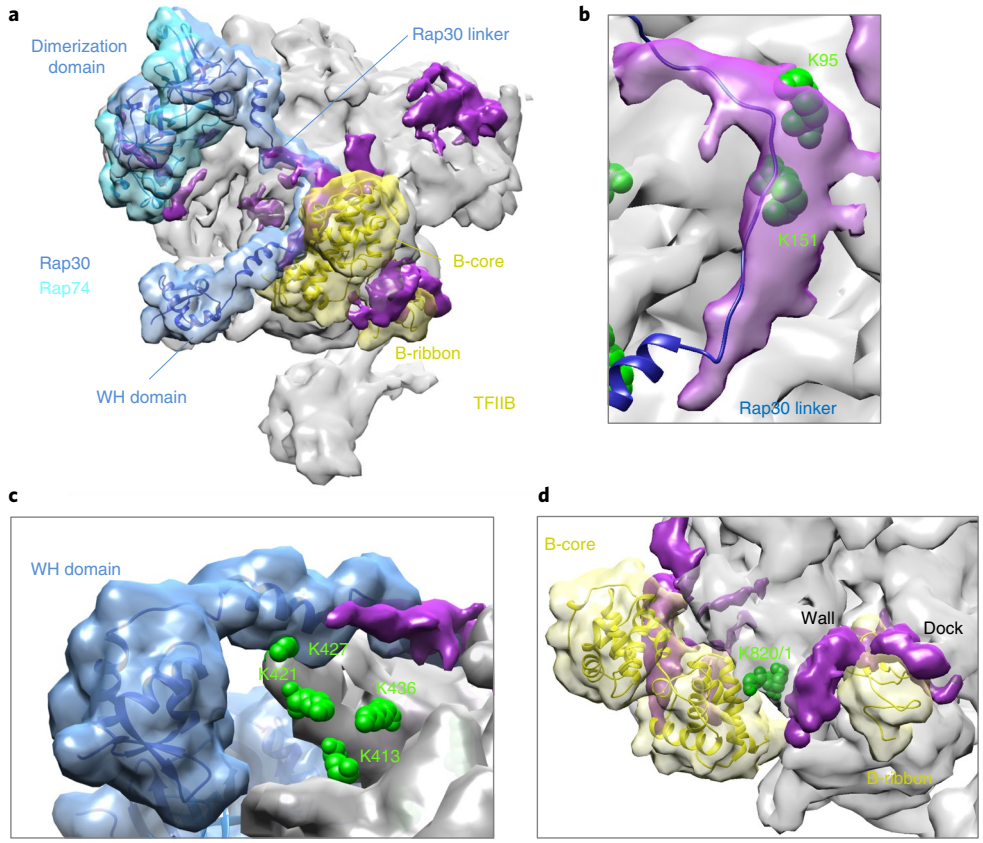


**Fig. 2 | Gdown1 interaction sites overlap TFIIF and TFIIB contact sites on Pol II. a**, Positions of TFIIF (Rap74 and Rap30) and TFIIB relative to Gdown1 density (shown in purple) on Pol II. Rap74, Rap30, and TFIIB are shown in light blue, dark blue and yellow, respectively. **b**, Overlap of Gdown1 density on the RPB2 protrusion domain with the Rap 30 linker domain. Two RBP2 lysines (K95 and K151) that cross-link to the C-terminal region of Gdown1 are shown in green. The Rap30 linker domain is shown in dark blue. **c**, Position of the Rap 30 WH domain (shown in dark blue) relative to Gdown1 density (shown in purple). Four RPB2 lysines that cross-link to the C-terminal region of Gdown1 are shown in green. **d**, Positions of TFIIB B-core and B-ribbon domains (shown in yellow) relative to Gdown1 density. Two RPB2 lysines (K820, K821) that cross-link to the C-terminal region of Gdown1 are shown in green.

domains). Although the exact role of C-TIR in modulating or stabilizing the Pol II interactions of the Gdown1 C-terminal domains is unclear, the extensive cross-linking of these domains (Fig. 1b) suggests that it may facilitate their direct cooperative binding to Pol II.

In further consideration of Pol II binding region I, Gdown1 residues K228 and K240 were found to cross-link to K95 in the RPB2 protrusion (Fig. 1b and Fig. 3b). The protrusion domain surrounding Gdown1 density, where Rap30 linker also binds (Fig. 2b), is enriched in hydrophobic amino acids (shown in orange in Fig. 3b), and, reciprocally, there are several highly conserved hydrophobic amino acids in the cross-linked Gdown1 region (227–248) (Fig. 3a), suggesting the potential involvement of hydrophobic interactions in the Gdown1–Pol II interaction at this site. In support of this idea for Pol II binding domain I, Gdown1(194–298) fragments in which selected hydrophobic residues were changed to alanine (Fig. 3a) lost Pol II binding (except for the P244A mutant) (Fig. 3f). Surprisingly, the same mutations failed to elicit any obvious deficit in Pol II binding when analyzed in the context of the large C-terminal Gdown1(194–368) fragment (Fig. 3g), although this may simply reflect compensation by Pol II binding region II and/or the C-TIR for general Pol II binding and does not eliminate a role for the indicated binding region I hydrophobic interactions in the inhibitory functions of Gdown1. The detrimental effect of the L303/4A mutation on binding of the Gdown1(194–368) fragment to Pol II (Fig. 3g, lane 9) is consistent with the results discussed earlier and again indicative of a stabilizing effect of C-TIR on Pol II binding region I.

**C-TIR acts cooperatively with N-terminal TIR to stabilize the interaction of Pol II binding region I with Pol II.** In a further analysis of the large 180-residue Gdown1 N-terminal fragment that alone does not repress transcription but nonetheless contributes significantly to Gdown1 inhibitory activity (Supplementary Fig. 3c), a small N-terminal deletion significantly reduced transcription inhibition by Gdown1 (Fig. 4b) without affecting Gdown1 binding to Pol II (Supplementary Fig. 4a). However, an electrophoretic mobility-shift assay (EMSA) demonstrated that the normal Gdown1 inhibition of PIC assembly was significantly reduced (Supplementary Fig. 4b), suggesting that the Gdown1 N-terminus somehow affects Gdown1–Pol II interactions.

To investigate the mechanism, we sought to identify N-terminal amino acids involved in Gdown1–Pol II interactions. Notably, the Gdown1 cryo-EM density near RPB3 (Fig. 4c), where Gdown1 K39 and K43 were cross-linked to K17 in RPB10 (Fig. 1b and Fig. 4e), overlapped with an acidic amino acid-enriched region in RPB3 (Fig. 4c,d). The presence of highly conserved arginine or lysine residues in the Gdown1 N-terminal region (Fig. 4e) thus suggested a salt-bridge interaction. In support of this prediction, the binding to Pol II of a GST-fused Gdown1(1–96) fragment in which these basic amino acids were mutated to aspartic acid was almost completely lost, whereas mutations of highly conserved acidic amino acids to arginine stabilized the Pol II interaction (Fig. 4f), indicating that these amino acids are important for Pol II interactions. Therefore, we defined this region (1–67) as an N-terminal transcription inhibition region (N-TIR). Although N-TIR is located more than 160

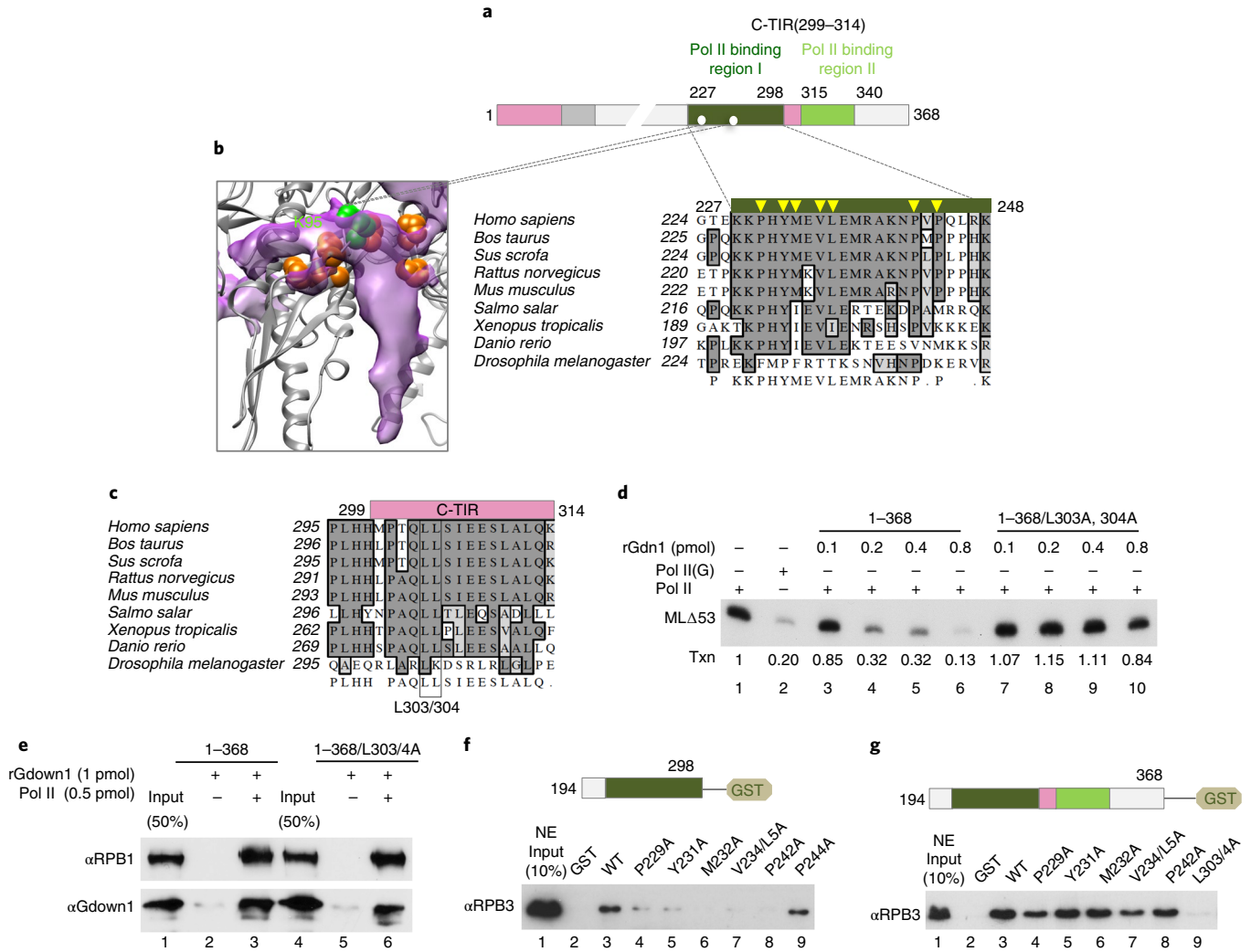


**Fig. 3 | C-TIR stabilizes the interaction of Pol II binding region I with Pol II. a**, Schematic of Gdown1 functional domains. Pol II binding region I (227–298), C-TIR (299–314), and Pol II binding region II (315–340) are shown in dark green, pink, and light green, respectively. Cross-linked residues (K228 and K240) in Pol II binding region I are shown in white circles. Alignment of binding region I sequences from various species. Amino acid residues that were changed to alanine by site-directed mutagenesis are shown in yellow triangles. **b**, RPB2 protrusion domain surrounding Gdown1 density (shown in purple). K95 (shown in green) in the RPB2 protrusion domain (shown in gray ribbon) cross-links to K228 and K240 in the Gdown1 Pol II binding region I. Hydrophobic amino acids in the protrusion domain are shown in orange. **c**, Alignment of TIR sequences from various species. Amino acid residues that were changed to alanine by site-directed mutagenesis are boxed in black. **d**, In vitro transcription assay with Gdown1 WT (1–368) and L303/4A mutant (1–368/L). Reactions contained 10 ng of TBP, 10 ng of TFIIB, 25 ng of TFIIF, 10 ng of TFIIEα, 5 ng of TFIIEβ, 50 ng of Pol II or Pol II(G) and indicated amounts of His-tagged full-length Gdown1(1–368) or L303/4A mutant. Reactions were incubated at 30 °C for 1 h and purified RNA products resolved by PAGE. **e**, Pol II interaction assays. Full-length Gdown1 or L303/4A mutant fragments were incubated with Pol II and subjected to immunoprecipitation with anti-CTD antibody (8W16). Bound proteins were analyzed by immunoblot. **f**,**g**, Pol II binding assay. Immobilized GST or indicated C-terminally fused GST-Gdown1 proteins with or without indicated mutations were incubated with HeLa nuclear extract in Buffer C containing 0.3 M KCl and 0.1% NP40. After washing, bound proteins were analyzed for bound Pol II by immunoblot. Results shown in panels **d**–**g** are representative of a minimum of three independent experiments. Uncropped gel and blot images are shown in Supplementary Dataset 1.

amino acids away from Pol II binding region I (Fig. 4a), cross-links between the two regions (Fig. 1c) indicate their close apposition in Pol II-bound Gdown1. A further analysis showed that K29/30D mutations in a Gdown1(1–244) fragment eliminated its binding to Pol II (Fig. 4g), thereby explaining the loss of transcription inhibition on the N-terminal deletion (Fig. 4a).

Although roles for N-TIR and C-TIR in stabilizing the Pol II binding region I interaction in Gdown1 repression became evident (Fig. 4a), localizations of these regions on Pol II (especially C-TIR, which did not show any significant binding to Pol II), remained unknown. To determine the locations, we performed cross-linking-based integrative structure modeling[24,25] entailing a four-step procedure (Methods, Supplementary Fig. 5, Fig. 6, and Supplementary Note). This analysis indicated an N-TIR localization near RPB10 and RPB3 where Gdown1 density was detected (Fig. 4h), consistent with the biochemical results (Fig. 4f). Although corresponding cryo-EM density was not observed, the N-terminal region (65–94) was localized to RPB10 and RPB3 and mapped adjacent to Pol II binding region I. Thus, N-TIR interacts with Pol II binding region I through the (65–94) region (Fig. 4h), explaining how N-TIR could stabilize the Gdown1–Pol II interaction. Pol II binding region I covers the RPB2 protrusion surface where two distinct densities overlap well, and Pol II binding region II also localizes to the RPB2 protrusion. The location of C-TIR was identified as an overlapping

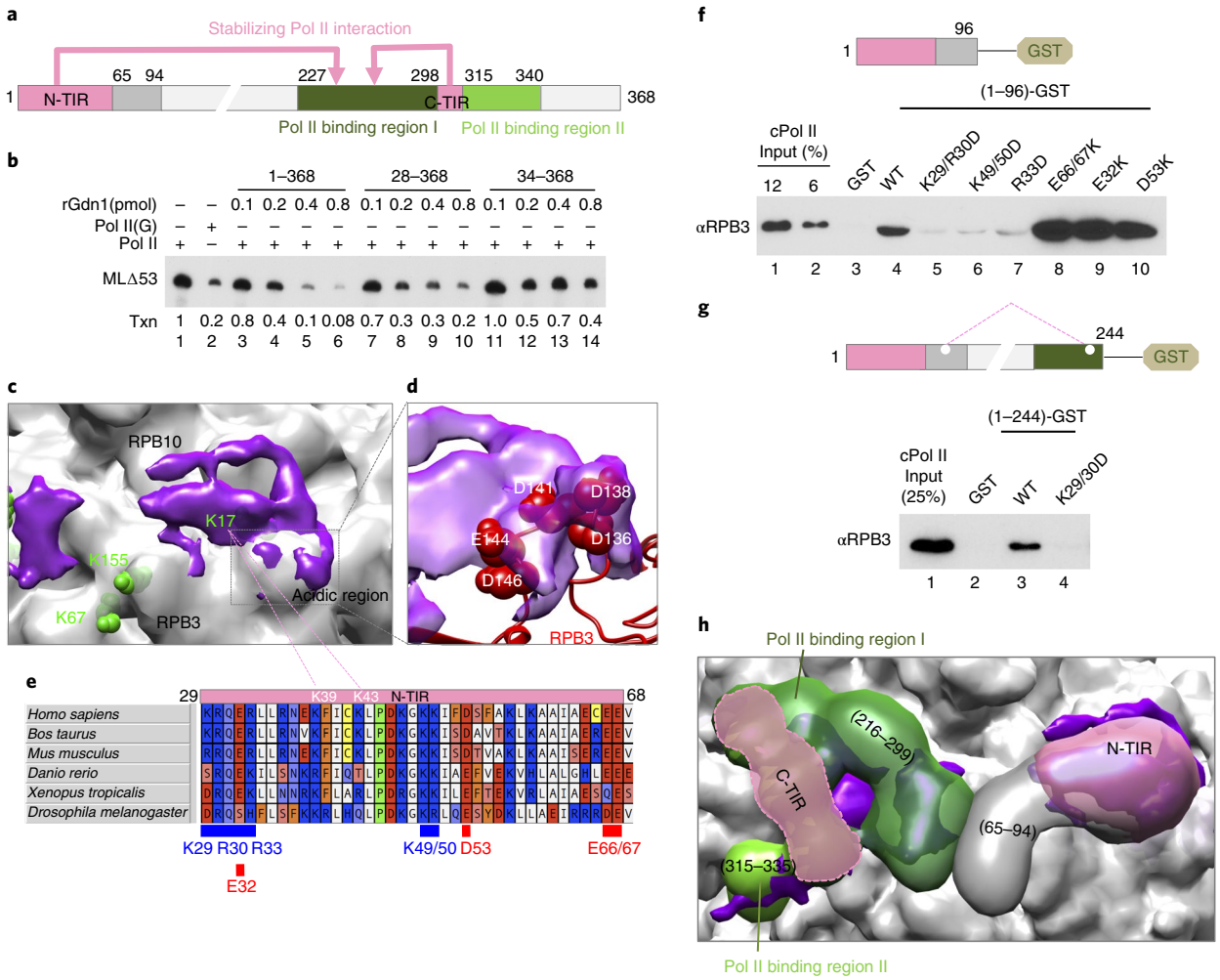


**Fig. 4 | N-TIR and C-TIR stabilize the interaction of Pol II binding region I with Pol II. a**, Schematic of Gdown1 functional domains. **b**, In vitro transcription assay with Gdown1 WT (1–368) and N-terminal deletion mutants. Assay is as described in Fig. 3d. **c**, Position of Gdown1 density on RPB10 and RPB3. Cross-linked RPB3/10 lysines (Fig. 1b) are shown in green. **d**, Overlap of Gdown1 density (purple) with the acidic region residues (red) in RPB3. **e**, Schematic of cross-linked N-TIR residues and alignment of N-TIR sequences (28–68) from various species. Residues changed by site-directed mutagenesis from basic to acidic or from acidic to basic are shown in red or blue, respectively. **f**,**g**, Binding of WT and mutant Gdown1 N-terminal fragments to Pol II. Immobilized GST fusion protein were incubated with purified Pol II in buffer C containing 0.1 M KCl and 0.1% NP40. After washing, bound Pol II was monitored by immunoblot. **h**, Location of Gdown1 functional domains determined by integrative structure modeling. Results shown in panels **b**, **f**, and **g** are representative of a minimum of three independent experiments. Uncropped gel and blot images are shown in Supplementary Dataset 1.

region between (216–314) and (300–335) (Supplementary Fig. 4c), whereas about two-thirds of the region does not overlap with the density (Fig. 4h). This may explain why C-TIR does not interact with Pol II. Overall, the modeling study revealed that both TIRs are located near Pol II binding regions, thus furnishing evidence for stabilization of Pol II binding region I by TIRs.

**Gdown1 nuclear co-localization with Pol II is inversely co-related with transcriptional activation in early fly embryos.** Whereas in vitro studies clearly show that Pol II(G) has unique properties relative to Pol II, little is known about the biological roles of Pol II(G) and why higher metazoans need two distinct forms of Pol II to regulate gene transcription. Beyond earlier indications of Gdown1 functions in mammalian cells[10], we sought to establish in vivo functions of Gdown1 in the context of global gene regulation in *Drosophila* embryogenesis. Initial studies established that Pol II(G), evidenced by reciprocal co-immunoprecipitation of Pol II and Gdown1, is present both in *Drosophila* embryo nuclear extracts (Fig. 5a) and in derived chromatographic fractions that are distinct from those containing Pol II (data not shown). Gdown1 is present at all life-cycle stages and was detected in both nuclear (more abundant) and cytoplasmic (less abundant) fractions at the embryonic stage (Fig. 5b) but appeared predominantly cytoplasmic at the adult stage (Fig. 5b,c). Strikingly, Gdown1 co-localizes with Pol II in nuclei at the transcriptionally silent syncytial blastoderm stage (Fig. 5d), but is detected only in the cytoplasm and not in the nuclei that retain Pol II at the later cellular blastoderm stage at which global transcription is initiated (Fig. 5e). Moreover, the pole cells that are transcriptionally silent retain nuclear Gdown1 at stage 5 (Fig. 5f). These data suggest a role for Gdown1, through formation of Pol II(G), in transcriptional repression.

We also demonstrated that whereas homozygous fly embryos carrying *Gdown1* alleles with P element insertions grow to first-instar larva, adult flies are not obtained. This requirement for Gdown1 was further confirmed by siRNA-mediated Gdown1 knockdown in the embryo. In related genetic studies, maternal Gdown1 knockout embryos were not obtained, further indicating that Gdown1 plays a critical role in early embryonic development in the fly.

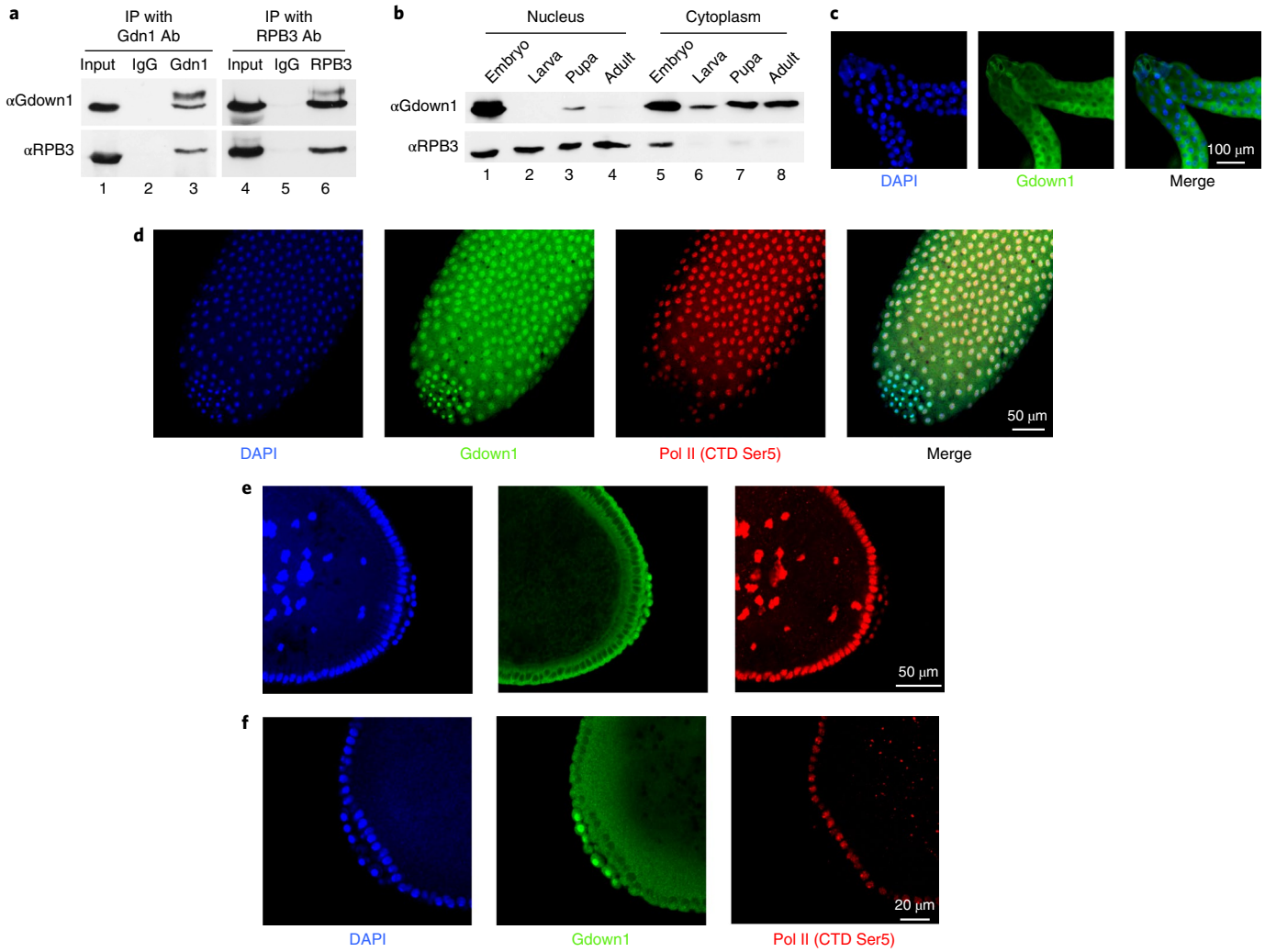


**Fig. 5 | Gdown1 nuclear co-localization with Pol II is inversely co-related to transcriptional activation in early *Drosophila* embryos. a**, Immunoprecipitation with *Drosophila* embryo nuclear extracts. Anti-Gdown1 (lane 3) or anti-RPB3 (lane 6) or IgG (lane 2 and 5) was incubated with fly embryo nuclear extract in Buffer C containing 0.3 M KCl and 0.1% NP40. After washing, bound proteins (Gdown1 and RPB3) was detected by immunoblot. **b**, Gdown1 cellular localization during *Drosophila* development. Nuclear fractions or cytoplasm fractions were prepared from each development stages. Gdown1 and RPB3 were detected by immunoblot. Results shown in panels **a** and **b** are representative of three independent experiments. Uncropped blot images are shown in Supplementary Dataset 1. **c**, Gdown1 immunostaining in salivary gland. **d**–**f**, Immunostaining of Gdown1 and Pol II (CTD ser5) at a syncytial blastoderm stage (**d**) and a cellular blastoderm stage (**e** and **f**).

Although we have not yet been able determine which embryonic gene expression events are influenced by Gdown1 knockout, these studies have provided strong evidence that Gdown1 has essential functions in vivo.

## Discussion

The present study has identified key Pol II binding domains (I and II) and modulatory domains (N-TIR and C-TIR) in metazoan-specific Gdown1 (Fig. 4a), which effects a unique mode of repression of Pol II function. Although Pol II interactions of individual Gdown1 functional domains are not sufficient to inhibit transcription, their inter-connected (cooperative) interactions with Pol II establish robust binding. At the same time, weak interactions of individual domains could potentially facilitate Gdown1 dissociation by Mediator and subsequent transcription.

The C-terminal Pol II binding regions (I and II) are the primary cause of the inhibitory activity of Gdown1 by preventing both TFIIF and TFIIB from binding to Pol II. Among the functional domains in TFIIF, the Rap30 linker has been shown to be essential both for growth in yeast[15] and for transcription initiation in vitro[26]. Moreover, recent cryo-EM studies of the PIC[18] and an initially transcribing complex[27] showed that the Rap30 linker makes contacts with multiple sites, including the RPB2 external domain, the RPB2 protrusion, TFIIB, TBP, and a downstream promoter region. Since these sites are essential for PIC formation, it was proposed that the linker interactions in the PIC might position other essential TFIIF domains[18]. Hence, the Gdown1 C-terminal interaction with the RPB2 protrusion domain could explain the mutually exclusive interactions of Pol II with TFIIF and Gdown1.

Notably, whereas the Gdown1(227–298) region interaction with the Pol II protrusion domain is stable, the Gdown1(315–368) region interaction appears quite fluid as evidenced by cross-linking to multiple domains in RPB2 and RPB1. Although deletion of this region did not result in a complete loss of Gdown1 inhibitory activity in the in vitro transcription assay, its interactions with the Pol II wall and rudder domain were found to inhibit TFIIB binding to Pol II and could be critical in a cellular context. Since TFIIB plays a pivotal role in PIC formation by direct binding both to Pol II and

promoter DNA[7], the Gdown1 interaction could thus further impede Pol II incorporation into the PIC. Also, the Gdown1 interactions with the Pol II clamp loop may interfere with DNA loading to the cleft. Taken together, a role of the Gdown1 C-terminal Pol II binding region in transcription repression is to prevent TFIIF and TFIIB from binding to Pol II, which contributes to the maintenance of Pol II an inactive state in the absence of Mediator.

Whereas Pol II binding region I is the major site for Gdown1 binding to Pol II, the interaction itself does not suffice for the Gdown1 transcription inhibition activity. As mutations in both TIRs can bypass the Mediator requirement for transcriptional activation, TIRs play essential roles in the inhibitory activity. Notably, whereas C-TIR, which consists of only 15 amino acids, does not directly bind to Pol II, point mutations in C-TIR cause a complete loss of repression. Although it remains unclear how C-TIR stabilizes the Gdown1 C-terminus interaction with Pol II, the modeling study predicts that C-TIR(300–314) is localized on the RPB2 protrusion parallel to the Gdown1(216–299) region. Since C-TIR probably possesses substantial secondary structure[28], it may directly associate with the Gdown1(216–299) region to stabilize the C-terminal Pol II interaction. The study also indicates that half of the Gdown1(300–314) region does not overlap with the density on the RBP2 protrusion, which may explain the absence of an independent C-TIR binding activity.

While C-TIR is located near Pol II binding region I, N-TIR, which is positioned at RPB3 and RPB10, is far from this region. However, the Gdown1(65–94) region, which is connected to N-TIR, directly contacts to the Pol II binding region I. This may explain how N-TIR deletion leads to the loss of transcription inhibition. Although N-TIR and C-TIR both show little independent binding to Pol II, their joint stabilization of the Gdown1 C-terminal interaction may account for the robust binding of Gdown1 to Pol II that is resistant to dissociation even by 2 M urea[6].

Since the mutations in N-TIR and C-TIR can relieve the normal Mediator requirement for transcription by Pol II(G), it is possible that Mediator interactions with Pol II disrupt N-TIR and/or C-TIR interactions with Pol II. In this regard, our cryo-EM study revealed at least two Mediator interacting sites on RPB3 and the RPB1 dock domain. While it remains unknown which region of Gdown1 interacts with the RPB1 dock domain, which is a contact site for MED18 in the Mediator head module[27], N-TIR is located around the RPB3 site where the Mediator tail module is likely to be located[29]. Interestingly, the location of Mediator tail domain subunit Med2 in the yeast PIC complex has been mapped near RPB3 (ref. [29]), where, more intriguingly, N-TIR also interacts (Supplementary Fig. 4d). This result raises the possibility that a Mediator tail domain interaction with RPB3 may lead to a concomitant N-TIR dissociation and de-stabilization of the Gdown1–Pol II interaction, which could be at least part of the mechanism by which Mediator reverses Gdown1-mediated transcription inhibition in vitro. In relation to biological relevance, we observed a joint Gdown1 and Pol II nuclear localization in transcriptionally silent embryos and selective re-localization of Gdown1 to the cytoplasm during zygotic genome activation. These results, the Mediator-dependent reversal of Gdown1 inhibition of Pol II[6], and Pol II(G) localization upstream of transcription start sites[10] suggest a biological role for Gdown1 that maintains Pol II in an inactive, potentially poised, state until an appropriate activation signal is generated.

## Online content

Any methods, additional references, Nature Research reporting summaries, source data, statements of data availability and associated accession codes are available at https://doi.org/10.1038/s41594-018-0118-5.

Received: 3 March 2018; Accepted: 24 July 2018;
Published online: 6 September 2018

## References

1. Roeder, R. G. Lasker basic medical research award. The eukaryotic transcriptional machinery: complexities and mechanisms unforeseen. *Nat. Med.* **9**, 1239–1244 (2003).
2. Grunberg, S. & Hahn, S. Structural insights into transcription initiation by RNA polymerase II. *Trends. Biochem. Sci.* **38**, 603–611 (2013).
3. Malik, S. & Roeder, R. G. The metazoan Mediator co-activator complex as an integrative hub for transcriptional regulation. *Nat. Rev. Genet.* **11**, 761–772 (2010).
4. Allen, B. L. & Taatjes, D. J. The mediator complex: a central integrator of transcription. *Nat. Rev. Mol. Cell Biol.* **16**, 155–166 (2015).
5. Malik, S. & Roeder, R. G. Mediator: a drawbridge across the enhancer-promoter divide. *Mol. Cell* **64**, 433–434 (2016).
6. Hu, X. et al. A mediator-responsive form of metazoan RNA polymerase II. *Proc. Natl. Acad. Sci., USA* **103**, 9506–9511 (2006).
7. Sainsbury, S., Bernecky, C. & Cramer, P. Structural basis of transcription initiation by RNA polymerase II. *Nat. Rev. Mol. Cell Biol.* **16**, 129–143 (2015).
8. Fagerberg, L. et al. Analysis of the human tissue-specific expression by genome-wide integration of transcriptomics and antibody-based proteomics. *Mol. Cell. Proteomics.* **13**, 397–406 (2014).
9. Yue, F. et al. A comparative encyclopedia of DNA elements in the mouse genome. *Nature* **515**, 355–364 (2014).
10. Jishage, M. et al. Transcriptional regulation by Pol II(G) involving mediator and competitive interactions of Gdown1 and TFIIF with Pol II. *Mol. Cell* **45**, 51–63 (2012).
11. He, Y., Fang, J., Taatjes, D. J. & Nogales, E. Structural visualization of key steps in human transcription initiation. *Nature* **495**, 481–486 (2013).
12. Bernecky, C., Herzog, F., Baumeister, W., Plitzko, J. M. & Cramer, P. Structure of transcribing mammalian RNA polymerase II. *Nature* **529**, 551–554 (2016).
13. Leitner, A. et al. Probing native protein structures by chemical cross-linking, mass spectrometry, and bioinformatics. *Mol. Cell. Proteomics.* **9**, 1634–1649 (2010).
14. Shi, Y. et al. A strategy for dissecting the architectures of native macromolecular assemblies. *Nat. Methods* **12**, 1135–1138 (2015).
15. Eichner, J., Chen, H. T., Warfield, L. & Hahn, S. Position of the general transcription factor TFIIF within the RNA polymerase II transcription preinitiation complex. *EMBO J.* **29**, 706–716 (2010).
16. Chen, Z. A. et al. Architecture of the RNA polymerase II-TFIIF complex revealed by cross-linking and mass spectrometry. *EMBO J.* **29**, 717–726 (2010).
17. Muhlbacher, W. et al. Conserved architecture of the core RNA polymerase II initiation complex. *Nat. Commun.* **5**, 4310 (2014).
18. He, Y. et al. Near-atomic resolution visualization of human transcription promoter opening. *Nature* **533**, 359–365 (2016).
19. Sainsbury, S., Niesser, J. & Cramer, P. Structure and function of the initially transcribing RNA polymerase II–TFIIB complex. *Nature* **493**, 437–440 (2013).
20. Kostrewa, D. et al. RNA polymerase II-TFIIB structure and mechanism of transcription initiation. *Nature* **462**, 323–330 (2009).
21. Pan, G. & Greenblatt, J. Initiation of transcription by RNA polymerase II is limited by melting of the promoter DNA in the region immediately upstream of the initiation site. *J. Biol. Chem.* **269**, 30101–30104 (1994).
22. Mullen Davis, M. A., Guo, J., Price, D. H. & Luse, D. S. Functional interactions of the RNA polymerase II-interacting proteins Gdown1 and TFIIF. *J. Biol. Chem.* **289** (2014).
23. Gnatt, A. L., Cramer, P., Fu, J., Bushnell, D. A. & Kornberg, R. D. Structural basis of transcription: an RNA polymerase II elongation complex at 3.3 A resolution. *Science* **292**, 1876–1882 (2001).
24. Sali, A. et al. Outcome of the first wwPDB hybrid/integrative methods task force workshop. *Structure* **23**, 1156–1167 (2015).
25. Russel, D. et al. Putting the pieces together: integrative modeling platform software for structure determination of macromolecular assemblies. *PLoS Biol.* **10**, e1001244 (2012).
26. Tan, S., Conaway, R. C. & Conaway, J. W. Dissection of transcription factor TFIIF functional domains required for initiation and elongation. *Proc. Natl. Acad. Sci., USA* **92**, 6042–6046 (1995).
27. Plaschka, C. et al. Architecture of the RNA polymerase II–Mediator core initiation complex. *Nature* **518**, 376–380 (2015).
28. Wu, Y. M. et al. Regulation of mammalian transcription by Gdown1 through a novel steric crosstalk revealed by cryo-EM. *EMBO J.* **31**, 3575–3587 (2012).
29. Robinson, P. J. et al. Structure of a complete mediator-RNA polymerase II pre-initiation complex. *Cell* **166** (2016).

## Acknowledgements

This study was supported by National Institutes of Health Grants CA202245 and DK071900 (R.G.R.), GM067167 (F.J.A.), GM103314 and GM109824 (B.T.C.), GM067167 (F.J.A.), GM083960 and GM109824 (A.S.), and Tri-I Stem Cell Initiative Grant 2014-21 (R.G.R.). We thank Z. Xue for preparation of nuclear extracts and S. Malik for discussions and critical reading of the manuscript.

## Author contributions

M.J. developed the purification procedure and purified Gdown1, Pol II, and Pol II(G) for biochemical assays, cryo-EM and CX-MS analyses, and performed biochemical studies. Cryo-EM experiments were designed by X.Y. and F.J.A., and performed by X.Y. Y.S. and B.C. designed and performed CX-MS analysis. S.G. and A.S. designed and performed integrative structure modeling. W.C. performed immunostaining and confocal imaging. M.J. and R.G.R. wrote the manuscript with input from all co-authors.

## Competing interests

The authors declare no competing interests.

## Additional information

**Supplementary information** is available for this paper at https://doi.org/10.1038/s41594-018-0118-5.

**Reprints and permissions information** is available at www.nature.com/reprints.

**Correspondence and requests for materials** should be addressed to R.G.R.

**Publisher's note:** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

## Methods

**Cell line.** HeLa S cell line that stably expresses FLAG-tagged human Gdown1 (ref. [10]) was used for Pol II(G) purification. The cell line was examined and tested negative for mycoplasma contamination.

**Purification of Pol II and Pol II(G).** Nuclear extracts were prepared from FLAG-tagged Gdown1 stable cell line[10] and dialyzed in TGE buffer (20 mM Tris-HCl pH 7.9 at 4 °C, 25% glycerol, 0.1 mM EDTA, 2 mM DTT, 0.5 mM PMSF) containing 0.1 M ammonium sulfate. The dialyzed nuclear extracts were fractionated on a DE52 column. The 0.3 M ammonium sulfate fraction was subjected to anion exchange chromatography (Hi-trapQ) and eluted with a linear gradient from 0.1 M to 1 M KCl in buffer B. The eluate was dialyzed in buffer C containing 0.3 M KCl, subjected to M2-agarose affinity chromatography, and eluted with 3X FLAG peptide.

**Preparation of recombinant Gdown1 proteins.** Human Gdown1 was cloned into pET21b (C-terminal 6× His tag). After transformation, *Escherichia coli* BL21(DE3) RIL cells were grown at 37 °C to $A_{600\,nm}$ of 0.6 before protein expression with 1 mM IPTG for 4 h at 25 °C. Subsequent steps were completed at 4 °C unless otherwise noted. Cells were lysed by sonication in Lysis buffer (50 mM HEPES pH 7.6, 500 mM NaCl, 10 mM imidazole, 1 mM PMSF, 10% glycerol,1 mM DTT). Cleared lysate was subjected to affinity chromatography using Ni-NTA agarose (Qiagen), and excess chaperone was removed by washing the resin with Lysis buffer containing 5 mM ATP and denatured *E. coli* protein, at room temperature. Protein was eluted with Lysis buffer supplemented with 250 mM imidazole. The eluate was exchanged into buffer B (20 mM HEPES pH 7.9, 100 mM KCl, 0.2 mM EDTA, 20% glycerol, 1 mM DTT). Diluted protein was subjected to anion exchange chromatography (Hi-trapQ) and eluted with a linear gradient from 0.1 M to 1 M KCl in buffer B. Gdown1-containing fractions were pooled and dialyzed in buffer B.

**Antibodies.** Antibodies against Gdown1 were raised in rabbits as described elsewhere[10]. Antibodies against RPB3 were purchased from Bethyl Laboratories. Antibodies against full-length *Drosophila* Gdown1 were validated by western blot, immunoprecipitation, and immunostaining with purified recombinant *Drosophila* Gdown1 protein, purified *Drosophila* Pol II(G) from *Drosophila* embryo nuclear extracts, and Gdown1 null *Drosophila* embryos.

**In vitro transcription and electrophoretic mobility-shift assays.** Pol II and GTFs were purified as described previously[30]. Bovine Pol II and Pol II(G) were purified as described previously[6]. In vitro transcription with purified components was performed as described[30]. EMSAs were performed as described[31].

**Protein interaction assays.** For the GST-pulldown assays, approximately 5 µg of each GST protein was immobilized on glutathione-Sepharose beads and incubated for 2 h at 4 °C with HeLa nuclear extract or the $^{35}$S-labeled proteins that were expressed in the TNT Quick Coupled Transcription/Translation system (Promega). After washing with buffer C containing 0.1 M or 0.3 M KCl and 0.1% NP40, bound proteins were analyzed by SDS-PAGE and were subjected to immunoblot or autoradiography.

**Chemical cross-linking and mass spectrometry.** Five µg of natively isolated Pol II(G) was chemically cross-linked with 0.5 mM disuccinimidyl suberate (Creative Molecules) for 30 min at 25 °C with constant agitation. The reaction was then quenched with 50 mM ammonia bicarbonate. After disulfide reduction and cysteine alkylation, the cross-linked Pol II(G) complexes were separated on a SDS-PAGE gradient gel (4–12%), which was briefly stained by GelCode Blue Stain Reagent (Thermo Fisher) to enable the visualization of the cross-linked complexes. Efficiently cross-linked materials that correspond to gel regions of >250 kDa were excised and digested in-gel using trypsin[14,32–35]. After peptide extractions, the cross-linked peptides were separated into six fractions using a self-packed pH10 C18 reverse resin, acidified and analyzed by liquid chromatography–mass spectrometry. For cross-link identifications, the purified peptides were dissolved in sample loading buffer (5% MeOH, 0.2% FA) and analyzed by an Orbitrap Q Exactive Plus mass spectrometer (Thermo Fisher). The instrument was coupled online to an EASY-nLC 1000 System (Thermo Fisher) for chromatographic separation of peptides. Peptide mixtures were loaded onto an Easy-Spray column (C18, 3 mm particle size, 200 Å pore size, and 50 µm × 15 cm, Thermo Fisher). The top eight most abundant ions (with charge stage of 4–7) were selected for fragmentation by higher-energy collisional dissociation. The raw data were searched by pLink. All spectra were manually verified as previously described[14,32,34–36]. The cross-link data was visualized and analyzed by the CX-Circos software.

**Cryo-EM analysis.** For grid preparation and data acquisition, 3.5 µl of 0.2 mg ml$^{-1}$ purified Pol II(G) complex was applied to the glow-discharged C-flat R2/2 grid coated with a home-made continuous thin layer of carbon. In order to overcome an orientation bias, the grid was treated with 5 µl of 0.1% (w/v) poly-L-lysine hydrobromide (Polysciences) prior applying the sample. Excess sample was manually blotted and vitrified in liquid ethane. The entire procedure was carried out at 4 °C and 98% relative humidity. Cryo grids were loaded into a Titan Krios transmission electron microscope operating at 300 keV. Images were automatically acquired with Leginon[37] at a nominal magnification of 22,500× (1.31 Å per pixel at the specimen level) using a total dose of ~40 electrons per Å$^2$ and a nominal defocus range of 1.5–2.8 µm. A total of 2,710 images were acquired using a Gatan K2 Summit direct electron detector, operated in electron counting mode. Each image was acquired as a 35 frame dose fractionated movie over a 7 s exposure time.

**Electron microscopy data processing.** Dose-fractioned movies were aligned using dosefgpu_driftcorr[38] with a frame offset of 7 and a B factor of 1,000 pixels$^2$. The Pol II(G) particles were automatically picked from the motion-corrected images using FindEM in the Appion pipeline[39], via the templates derived from 5,000 picked particles using DoGpicker. Particles were then extracted using Relion 1.4 (ref. [40]) with a box size of 224 pixels. CTF parameters were estimated using the programs CTFFIND4 (ref. [41]), A total of 201,527 particles were extracted using a box size of 224 pixels$^2$. The two- and three-dimensional (2D and 3D) classification and refinement were performed with Relion 1.4. Two rounds of 2D classification and one round of 3D classification were performed to select the homogenous particles. After 3D classification, one set of 141,619 particles were then submitted to particle-based motion correction and radiation damage weighting, and followed by 3D auto-refinement. All 3D classification and 3D refinements were started from a 60 Å low-pass filtered version of the X-ray crystal structure of bovine RNA polymerase II (PDB 5FLM)[12]. The refinement resulted in an overall structure at a resolution of 4.0 Å based on the gold-standard FSC = 0.143 criteria[42]. Prior to visualization, all density maps were corrected for the modulation transfer function (MTF) of the detector, and then sharpened by applying a negative B-factor that was estimated using automated procedures[43].

In order to improve the quality of the reconstructed map at the N-terminal region of Gdown1, we used the signal-subtracted focus classification and refinement in RELION[44]. We subtracted projections from the Pol II of the reconstruction in the experimental particle images using the relative orientation of each experimental image from the last 3D auto-refinement run of the overall polished particles. The subtracted experimental images were then subjected to the 3D classification with a soft binary mask around the N-terminal region of Gdown1 and no alignments. Thirty-five thousand and twenty-six signal-subtracted particles from the 3D class with the strongest Gdown1 signals were selected, and the corresponding experimental polished particles were subjected to 3D auto-refinement by continuing with the last angular and translational search parameters from the 3D auto-refinement step of the overall Pol II(G). This approach produced a better quality map around the N-terminal region of Gdown1. The same strategy was also applied for the C-terminal region of Gdown1, but did not further improve the map that indicates a dynamic interaction between the C-terminal region and Pol II.

To build the Pol II atomic model, the cryo-EM structure of bovine Pol II (PDB 5FLM) served as the reference. The starting model was placed in the density by rigid-body fitting in UCSF Chimera[45]. Further model adjustments were done manually using Coot[46]. Consistent with previous observations, the apo Pol II preserves a flexible clamp. We were not able to model the clamp part of the Pol II(G). Focus refinement using a soft mask around the C-terminal region of Gdown1 improved the local density, and we were able to trace the main chain. The map is not good enough to assign the sequence. Refinement of the Pol II and Gdown1 models against the Pol II(G)1 cryo-EM map was done by using the real space refinement function implemented in Phenix[47]. Final models were validated using Molprobity[48] and the FSC of the final model versus the Pol II(G)1 map (Supplementary Fig. 1). Integrative structural modeling is described in the Supplementary Note.

**Preparation of *Drosophila* nuclear extracts.** Embryos (0–12 h old) were collected from a mass population of *Drosophila melanogaster*. The nuclear extracts were made as previously described[49].

**Immunostaining and confocal imaging.** Embryos were washed, dechorionated, and fixed as previously described[50]. After devitellinization in methanol, embryos were rinsed with PBST (PBS containing 0.2% Tween 20) three times, and incubated with blocking solution (Roche) for 1 h at room temperature. Embryos were then incubated with primary antibody overnight at 4 °C. After washing with PBST three times, embryos were incubated with secondary antibody for 1 h at room temperature, washed with PBST three times, and mounted in Vectashield (Vector Laboratories) for microscopy. Primary antibody was diluted in blocking solution as follows: anti-dGdown1 (Roeder lab) 1/200, anti-CTD phosphor-ser5 (Clone H14, BioLegend) 1/100. Confocal images of immuno-stained embryos were obtained using a LSM 780 laser scanning confocal microscope (Zeiss).

**Reporting Summary.** Further details on research design can be found in the Nature Research Reporting Summary linked to this article.

## Data availability

The cryo-EM density maps of Pol II(G) have been deposited in the Electron Microscopy Data Bank under accession EMD-7997. Additional source data are available from the corresponding author upon reasonable request.

## References


30. Malik, S. & Roeder, R. G. Isolation and functional characterization of the TRAP/mediator complex. *Methods Enzymol.* **364**, 257–284 (2003).
31. Malik, S., Guermah, M. & Roeder, R. G. A dynamic model for PC4 coactivator function in RNA polymerase II transcription. *Proc. Natl Acad. Sci., USA* **95**, 2192–2197 (1998).
32. Algret, R. et al. Molecular architecture and function of the SEA complex, a modulator of the TORC1 pathway. *Mol. Cell. Proteomics.* **13**, 2855–2870 (2014).
33. Fernandez-Martinez, J. et al. Structure and function of the nuclear pore complex cytoplasmic mRNA export platform. *Cell* **167**, 1215–1228.e1225 (2016).
34. Shi, Y. et al. Structural characterization by cross-linking reveals the detailed architecture of a coatomer-related heptameric module from the nuclear pore complex. *Mol. Cell. Proteomics.* **13**, 2927–2943 (2014).
35. Sun, J. et al. The architecture of a eukaryotic replisome. *Nat. Struct. Mol. Biol.* **22**, 976–982 (2015).
36. Cevher, M. A. et al. Reconstitution of active human core Mediator complex reveals a critical role of the MED14 subunit. *Nat. Struct. Mol. Biol.* **21**, 1028–1034 (2014).
37. Suloway, C. et al. Automated molecular microscopy: the new Leginon system. *J. Struct. Biol.* **151**, 41–60 (2005).
38. Li, X. et al. Electron counting and beam-induced motion correction enable near-atomic-resolution single-particle cryo-EM. *Nat. Methods* **10**, 584–590 (2013).
39. Lander, G. C. et al. Appion: an integrated, database-driven pipeline to facilitate EM image processing. *J. Struct. Biol.* **166**, 95–102 (2009).
40. Scheres, S. H. RELION: implementation of a Bayesian approach to cryo-EM structure determination. *J. Struct. Biol.* **180**, 519–530 (2012).
41. Rohou, A. & Grigorieff, N. CTFFIND4: Fast and accurate defocus estimation from electron micrographs. *J. Struct. Biol.* **192**, 216–221 (2015).
42. Scheres, S. H. & Chen, S. Prevention of overfitting in cryo-EM structure determination. *Nat. Methods* **9**, 853–854 (2012).
43. Rosenthal, P. B. & Henderson, R. Optimal determination of particle orientation, absolute hand, and contrast loss in single-particle electron cryomicroscopy. *J. Mol. Biol.* **333**, 721–745 (2003).
44. Bai, X. C., Rajendra, E., Yang, G., Shi, Y. & Scheres, S. H. Sampling the conformational space of the catalytic subunit of human gamma-secretase. *eLife* **4**, e11182 (2015).
45. Pettersen, E. F. et al. UCSF Chimera: a visualization system for exploratory research and analysis. *J. Comput. Chem.* **25**, 1605–1612 (2004).
46. Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. Features and development of Coot. *Acta Crystallogr. D. Biol. Crystallogr.* **66**, 486–501 (2010).
47. Afonine, P. V. et al. Towards automated crystallographic structure refinement with phenix.refine. *Acta Crystallogr. D. Biol. Crystallogr.* **68**, 352–367 (2012).
48. Chen, V. B. et al. MolProbity: all-atom structure validation for macromolecular crystallography. *Acta Crystallogr. D. Biol. Crystallogr.* **66**, 12–21 (2010).
49. Ueda, H., Sonoda, S., Brown, J. L., Scott, M. P. & Wu, C. A sequence-specific DNA-binding protein that activates fushi tarazu segmentation gene expression. *Genes Dev.* **4**, 624–635 (1990).
50. Kanesaki, T., Hirose, S., Grosshans, J. & Fuse, N. Heterotrimeric G protein signaling governs the cortical stability during apical constriction in Drosophila gastrulation. *Mech. Dev.* **130**, 132–142 (2013).

nature research

Corresponding author(s): Robert Roeder

# Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see Authors & Referees and the Editorial Policy Checklist.

## Statistical parameters

When statistical analyses are reported, confirm that the following items are present in the relevant location (e.g. figure legend, table legend, main text, or Methods section).

| n/a | Confirmed | |
|---|---|---|
| ☐ | ☒ | The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement |
| ☐ | ☒ | An indication of whether measurements were taken from distinct samples or whether the same sample was measured repeatedly |
| ☐ | ☒ | The statistical test(s) used AND whether they are one- or two-sided<br>*Only common tests should be described solely by name; describe more complex techniques in the Methods section.* |
| ☐ | ☒ | A description of all covariates tested |
| ☐ | ☒ | A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons |
| ☐ | ☒ | A full description of the statistics including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals) |
| ☐ | ☒ | For null hypothesis testing, the test statistic (e.g. *F*, *t*, *r*) with confidence intervals, effect sizes, degrees of freedom and *P* value noted<br>*Give P values as exact values whenever suitable.* |
| ☐ | ☒ | For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings |
| ☐ | ☒ | For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes |
| ☐ | ☒ | Estimates of effect sizes (e.g. Cohen's *d*, Pearson's *r*), indicating how they were calculated |
| ☐ | ☒ | Clearly defined error bars<br>*State explicitly what error bars represent (e.g. SD, SE, CI)* |

*Our web collection on statistics for biologists may be useful.*

## Software and code

Policy information about availability of computer code

| | |
|---|---|
| Data collection | Integrative Modeling Platform (IMP 2.8), MODELLER 9.3, HHPred, PSIPRED, DISOPRED, CTFFIND4, Relion 1.4 |
| Data analysis | UCSF Chimera, CX-Circos, matplotLib, GNUPLOT 5.0, Coot |

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors/reviewers upon request. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research guidelines for submitting code & software for further information.

## Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

All data generated or analyzed during this study are included in this article and its supplementary information files.

# Field-specific reporting

Please select the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☒ Life sciences ☐ Behavioural & social sciences

For a reference copy of the document with all sections, see nature.com/authors/policies/ReportingSummary-flat.pdf

# Life sciences

## Study design

All studies must disclose on these points even when the disclosure is negative.

| | |
|---|---|
| Sample size | The sample size for integrative structure determination was determined by counting 5,000,000 modelled structures from 100 independent runs and 1,693 good-scoring structures. |
| Data exclusions | No data were excluded. |
| Replication | Experimental findings were reliably reproduced. |
| Randomization | n/a |
| Blinding | n/a |

## Materials & experimental systems

Policy information about availability of materials

| n/a | Involved in the study |
|---|---|
| ☒ | ☐ Unique materials |
| ☐ | ☒ Antibodies |
| ☐ | ☒ Eukaryotic cell lines |
| ☒ | ☐ Research animals |
| ☒ | ☐ Human research participants |

### Antibodies

| | |
|---|---|
| Antibodies used | Antibodies against full-length drosophila Gdown1 was raised in rabbits (Covance). Antibodies against human Gdown1 was generated by genomic antibody technology (Strategic Diagnostics), and have been published in Mol Cell 45, 51-63, doi:10.1016/j.molcel.2011.12.014 (2012). Antibodies against RPB3, was purchased from Bethyl Laboratories (Cat#A303-771A). Antibodies against CTD phosphor-ser5 was purchased from BioLegend (Clone H14, Cat#920304). |
| Validation | Antibodies against full-length drosophila Gdown1 was validated by western blot, immunoprecipitation, and immunostaining with purified recombinant drosophila Gdown1 protein, purified drosophila Pol II(G) from drosophila embryo nuclear extracts, and Gdown1 null drosophila embryos. |

### Eukaryotic cell lines

Policy information about cell lines

| | |
|---|---|
| Cell line source(s) | FLAG-tagged Gdown1 stable HeLa cell line is described in detail in Mol Cell 45, 51-63, doi:10.1016/j.molcel.2011.12.014 (2012). |
| Authentication | The cell line was generated in our laboratory, and verified by visual inspection of morphology and flag-tagged Gdown1 expression. |
| Mycoplasma contamination | The cellline was examined and tested negative for mycoplasma contamination. |
| Commonly misidentified lines (See ICLAC register) | No commonly misidentified cell lines were used. |

# Method-specific reporting

| n/a | Involved in the study |
|---|---|
| ☒ | ☐ ChIP-seq |
| ☒ | ☐ Flow cytometry |
| ☒ | ☐ Magnetic resonance imaging |