The EMBO Journal (2012) 31, 3575–3587 | © 2012 European Molecular Biology Organization | All Rights Reserved 0261-4189/12
www.embojournal.org

THE EMBO JOURNAL

Regulation of mammalian transcription by Gdown1 through a novel steric crosstalk revealed by cryo-EM

Yi-Min Wu[1], Jen-Wei Chang[1], Chun-Hsiung Wang[1], Yen-Chen Lin[1], Pei-lun Wu[1], Shih-hsin Huang[1], Chia-Chi Chang[1], Xiaopeng Hu[2], Averell Gnatt[3,*] and Wei-hau Chang[1,4,5,*]

[1]Institute of Chemistry, Academia Sinica, Taipei, Taiwan, [2]School of Pharmaceutical Science, Sun Yet-Sen University, Guangzhou, China, [3]Department of Pharmacology and Experimental Therapeutics, University of Maryland School of Medicine, Baltimore, MD, USA, [4]Department of Biochemical Science and Technology, National Taiwan University, Taipei, Taiwan and [5]Genomic Research Center, Academia Sinica, Taipei, Taiwan

In mammals, a distinct RNA polymerase II form, RNAPII(G) contains a novel subunit Gdown1 (encoded by *POLR2M*), which represses gene activation, only to be reversed by the multisubunit Mediator co-activator. Here, we employed single-particle cryo-electron microscopy (cryo-EM) to disclose the architectures of RNAPII(G), RNAPII and RNAPII in complex with the transcription initiation factor TFIIF, all to ~19 Å. Difference analysis mapped Gdown1 mostly to the RNAPII Rpb5 shelf-Rpb1 jaw, supported by antibody labelling experiments. These structural features correlate with the moderate increase in the efficiency of RNA chain elongation by RNAP II(G). In addition, our updated RNAPII–TFIIF map showed that TFIIF tethers multiple regions surrounding the DNA-binding cleft, in agreement with cross-linking and biochemical mapping. Gdown1's binding sites overlap extensively with those of TFIIF, with Gdown1 sterically excluding TFIIF from RNAPII, herein demonstrated by competition assays using size exclusion chromatography. In summary, our work establishes a structural basis for Gdown1 impeding initiation at promoters, by obstruction of TFIIF, accounting for an additional dependent role of Mediator in activated transcription.

The EMBO Journal (2012) **31,** 3575–3587. doi:10.1038/emboj.2012.205; Published online 31 July 2012
Subject Categories: chromatin & transcription
Keywords: intrinsic disordered protein; mediator; RNA polymerase; single particle; TFIIF

*Corresponding authors. A Gnatt, Department of Pharmacology and Experimental Therapeutics, University of Maryland School of Medicine, Baltimore, MD 21201, USA. Tel.: +1 410 706 8239; Fax: +1 410 706 0032; E-mail: agnat001@umaryland.edu or W-H Chang, Institute of Chemistry, Academia Sinica, 128 Academia Road, Section 2, Nankang, Taipei 115, Taiwan. Tel.: +886 2 2789 8558; Fax: +886 2 2783 1237; E-mail: weihau@chem.sinica.edu.tw or weihau40@gmail.com

Received: 25 August 2011; accepted: 3 July 2012; published online: 31 July 2012

Introduction

RNA polymerase II (RNAPII) synthesizes all eukaryotic messenger RNA and constitutes the core of the protein encoding transcription machinery. Regulation of RNAPII transcription is crucial for cell growth and differentiation and is achieved by the concerted activities of a large number of proteins. As RNAPII is unable to recognize a promoter, promoter-specific initiation by RNAPII requires a set of conserved general transcription factors (GTFs), TFIIB, TFIID, TFIIF, TFIIE, and TFIIH to form a pre-initiation complex (PIC) at a promoter (Hahn, 2004). RNAPII was originally isolated as a multisubunit enzyme from mammalian cells (Roeder *et al*, 1976) but rigorous definition of its subunit composition (Rpb1–Rpb12) was facilitated by purification and characterization of its yeast counterpart (Young, 1991). The atomic coordinates of RNAPII were obtained by X-ray diffraction analysis of three-dimensional (3D) crystals grown from highly purified yeast protein (Cramer *et al*, 2001; Gnatt *et al*, 2001). The structure of RNAPII can be delineated as an assembly of distinct modules. A core module containing the active centre accounts for approximately half of the total mass of the enzyme (~500 kDa), and three additional modules surrounding the DNA-binding cleft are mobile: the jaw-lobe (Rpb1–Rpb2), shelf (Rpb5), and clamp (Rpb1). X-ray structures of RNAPII–TFIIB (Kostrewa *et al*, 2009; Liu *et al*, 2010) along with that of the TFIIB–TBP–TATA-element ternary complex (Nikolov *et al*, 1995) allow for the development of a model for the PIC. As large numbers of protein contacts are involved in PIC, its formation stands as a key point of regulation (Fuda *et al*, 2009).

Beyond PIC formation, an additional layer of regulation at the promoter requires co-activators for cell viability. Co-activator proteins convey signals from DNA-binding activators or repressors to RNAPII, allowing for up or down-regulation of gene expression. Among the co-activators, of critical importance is Mediator—an essential and conserved protein complex of ~30 polypeptides that supports gene activation by 'mediating' between the activator and the PIC (Conaway *et al*, 2005; Kornberg, 2005; Malik and Roeder, 2005, 2010). Remarkably, an *in-vitro* transcription assay containing purified mammalian proteins including GTFs, general co-activator PC4, and 12-subunit RNAPII, displayed an unregulated and unfettered high degree of transcription in the presence of DNA-binding activators with or without Mediator (Hu *et al*, 2006). As such, it appeared that transcription factor activation of RNAPII was not dependent upon Mediator in contrast to Mediator's established role as an essential transcriptional co-activator (Belakavadi and Fondell, 2006; Casamassimi and Napoli, 2007; Cai *et al*, 2009). This apparent discrepancy was resolved by the disclosure of a novel RNAPII isoform, RNAPII(G), containing the RNAPII-associated polypeptide, Gdown1 of 43 kDa, which suppresses activated transcription but is relieved only in the presence of Mediator (Hu *et al*, 2006).

Essentially, Gdown1 confers Mediator responsiveness upon RNAPII. Gdown1 is one of the many products of the GRINL1A complex transcription unit (Roginski *et al*, 2004) and is a novel RNAPII subunit (POLR2M) as it is resistant to dissociation from RNAPII by high salt and urea, and is found as a percent of native enzyme (Hu *et al*, 2006).

Considering the fundamental role of transcription, and the large number of interacting transcription proteins needed for effective transcription, it is important to derive a rudimentary understanding as to how Gdown1 could crosstalk with the transcription machinery. Here, by employing cryo-electron microscopy (cryo-EM) followed by single-particle analysis, we obtained the 3D structure of RNAPII(G) in an unstained state to ~19 Å and revealed the binding sites of Gdown1 on RNAPII. In addition, we obtained the 3D cryo-EM map of mammalian RNAPII–TFIIF and uncovered the densities of TFIIF on RNAPII and found TFIIF shared several sites with those of Gdown1. As such, the notion Gdown1 and TFIIF would exclude each other was suggested and confirmed by a gel-filtration competition assay. Our findings thus confer a steric mechanism underlying Gdown1 inhibits TFIIF function (Cheng *et al*, 2012; Jishage *et al*, 2012). Finally, the involvement of Mediator negating Gdown1 to restore transcription initiation is discussed.

Results

Biochemical characterization of bovine RNAPII and RNAPII(G)

After receiving the native bovine RNAPII and RNAPII(G) in ammonium-sulphate precipitant, the proteins were thawed and exchanged into physiological buffer conditions. At that stage, the RNAPII and RNAPII(G) enzymes were examined for their subunit composition on a SDS–PAGE stained by Coomassie Blue. As shown in Figure 1A, the Gdown1 in the native RNAPII(G) appears to be approximately stoichiometric when compared with the two largest RNAPII subunits, with the ratio Rpb1: Rpb2: Gdown1 ~0.81: 1: 0.74. Both forms of polymerase were tested for their activity in a nonspecific transcription elongation assay with tailed DNA template without the requirement of general transcription initiation factors. RNAPII and RNAPII(G) were active in generating early arrested RNA transcripts of 13–16 bases length and additional readthrough products of various lengths. Quantitation of early arrest or readthrough transcripts indicated a 1.5- to 2.5-fold increase in the amount of transcripts by RNAPII(G) compared with those of RNAPII (Figure 1B). This increase in activity of RNAPII(G) compared with RNAPII was also observed by others (Cheng *et al*, 2012; Jishage *et al*, 2012). We further analysed Gdown1's propensity as a disordered protein by rendering its sequence to folding analysis (Prilusky *et al*, 2005). Interestingly, the major folded region of Gdown1 appears to be in the N-terminal half, ranging from amino acid 55–113 (Figure 1C). To validate such prediction, recombinant Gdown1 proteins were subjected to limited trypsin proteolysis followed by mass spectroscopy. As anticipated, the cleavage mainly took place in the C-terminal region (Figure 1D).

Single-particle analysis of native bovine RNAPII and RNAPII(G) complexes in negative stain

RNAPII and RNAPII(G) display different behaviour in solution. As shown in an EM image (Figure 2A), RNAPII(G) predominantly formed monomers. By contrast, RNAPII mainly formed dimers (Figure 2D). Images of RNAPII dimeric particles or RNAPII(G) monomeric particles were aligned using the SPIDER (Frank *et al*, 1996) and clustered with XMIPP (Sorzano *et al*, 2004). 7689 RNAPII(G) particle images conferred a set of class averages resembling the 2D projections of yeast RNAPII X-ray structure (Figure 2B; Supplementary Figure 1A). By the common-line method (Penczek *et al*, 1996), those class averages were used to generate an initial model, which was used to guide the angular reconstruction (Penczek *et al*, 1994) of RNAPII(G) to obtain a volume with ~30 Å resolution (Supplementary Figure 1B). As the EM structure of RNAPII(G) was superimposed with the 12-subunit yeast RNAPII (Armache *et al*, 2005; PDB: 1WCM) (Figure 2C), good agreement was found, while Gdown1 density was virtually undetected. As to RNAPII dimers, alignment and clustering of selected dimeric particle images resulted in very few numbers of different classes, indicating the dimmers had preferred orientations. A class average of the dimer images (Figure 2E) carried the feature of yeast RNAPII dimers previously identified in 2D crystal (Darst *et al*, 1991; Asturias *et al*, 1998). To investigate the interface for dimerization of RNAPII, a dimer model was built from the 3D reconstruction of a negative-stained RNAPII(G) to suit the 2D dimer class average, by which the Rpb3 or Rpb4 subunit is suggested to participate in the dimer contact (Figure 2F). That Gdown1 can break the dimer formation suggests its binding sites on RNAPII include those interfacial areas. Interestingly, antibody against the Rpb3 subunit could induce a large fraction of RNAPII (~70%) to form monomers (Supplementary Figure 1C). As suggested by our low-resolution negative-stain EM study together with the folding analysis, Gdown1 may dwell on RNAPII in an extended form, leading us to pursue a higher quality and more detailed structure by cryo-EM. Furthermore, though native RNAPII and RNAPII(G) appear relatively stoichiometric, we were unable to achieve a complete separation of native RNAPII(G) so that some amounts, though very minimal, of 12-subunit RNAPII certainly contaminate our RNAPII(G) preparation. Therefore, for the cryo-EM study, we generated RNAPII(G) particles by reconstituting RNAPII with a three- to four-fold excess of recombinant Gdown1 (Hu *et al*, 2006).

Cryo-EM of bovine RNAPII elongation complex

For a rigorous assessment of Gdown1 on RNAPII, it was desired to directly compare a 13-subunit RNAPII(G) with a 12-subunit RNAPII, both derived from bovine. Therefore, a cryo-EM structure of bovine RNAPII (12-subunit) reconstructed from its monomer projections was pursued. However, the dimeric form of RNAPII as opposed to the monomeric form of RNAPII(G) was not suited for a direct comparison. In an attempt to generate monomers from the bovine RNAPII dimer, the screening of salt conditions was exhausted but none of them succeeded to dissociate the 12-subunit RNAPII dimers into monomers. Interestingly, as RNAPII was supplemented with a nucleic acid scaffold (Kettenberger *et al*, 2004; Chen *et al*, 2009) (Supplementary Figure 2A), RNAPII predominantly formed monomeric particles in cryo images (Supplementary Figure 2B). The 12-subunit bovine RNAPII reconstituted with dsDNA/RNA is herein designated as the RNAPII elongation complex. We first

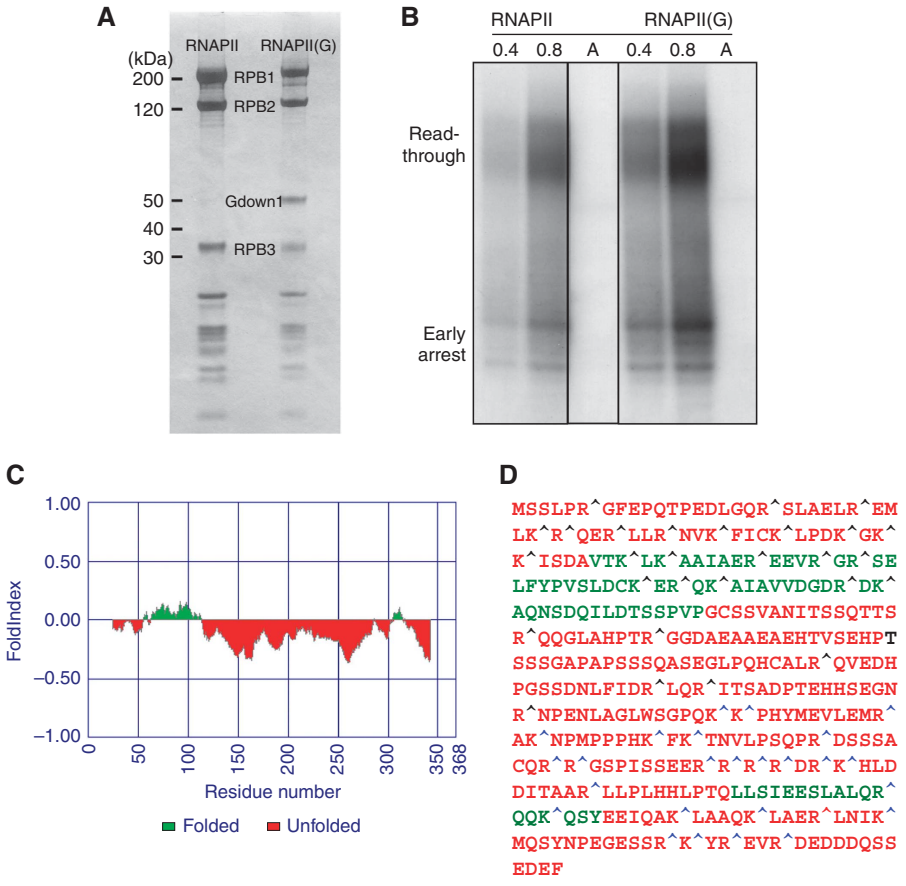

Figure 1 Biochemical and bioinformatics characterization of RNAPII(G). (**A**) Purification of native RNAPII and RNAPII(G). Both forms of calf thymus RNAPII are presented in the SDS–PAGE Coomassie stained gel, with Gdown1 and the RNAPII subunits Rpb1, Rpb2, and Rpb3 labelled. By dividing the integrated intensity over the respective molecular weight, the relative amounts of Rpb1, Rpb2, and Gdown1 in RNAPII(G) were determined to be 0.81: 1: 0.74. (**B**) Nonspecific transcription elongation assays. 0.4 and 0.8 μg of RNAPII (lanes 1–2) and RNAPII(G) (lanes 4–5) were used for the assays as previously described (Gnatt *et al*, 1997). RNA fragments from the early arrest or read-through are marked. As a control, α-amanitin, an RNAPII inhibitor, was added to RNAPII (lane 3) and RNAPII(G) (lane 6), respectively. All six lanes were from the same blot and only irrelevant lanes have been removed for the figure. Lanes 1 and 2 correspond to lanes 1 and 2 in the source gel; lane 3 corresponds to lane 4 in the source gel and lanes 4–6 correspond to lanes 6–8 in the source gel. The source data has been uploaded for full information. (**C**) Folding analysis of Gdown1. Program FoldIndex was used to evaluate the folding propensity of Gdown1. Two folded domains found are marked in green and unfolded region in red. (**D**) Limited proteolysis assay performed with trypsin. Single letter amino acid codes in the predicted folded region are denoted in green in contrast to the red colour for those in the predicted unfolded region. Cleavage sites of peptide fragments identified by mass spectrometry are labelled by blue carets while protected protease sites are denoted by black carets. Figure source data can be found with the Supplementary data.

generated a cryo-stained structure of bovine RNAPII elongation complex (Supplementary Figure 2C) by the angular reconstruction method using the negative-stained EM volume of RNAPII(G) as an initial model (Figure 2C). The cryo-stained structure of bovine RNAPII was consistent with that of the human RNAPII EM structure (EMD-1284) (Kostek *et al*, 2006) obtained by a similar approach (Supplementary Figure 2D). The cryo-stain reconstruction was employed to guide the angular parameters of ~20 000 unstained cryo-EM images of bovine RNAPII elongation complex to a resolution ~19 Å (Figure 3). As the resultant cryo-EM map of bovine RNAPII elongation complex was properly contoured according to the molecular mass of RNAPII, it gave no sporadic densities on the surface that did not belong to RNAPII but agreed nicely with that of the X-ray structure of the yeast RNAPII elongation complex filtered to the same resolution (Kettenberger *et al*, 2004; PDB: 1Y1W) (Figure 3). Since we did not inject any X-ray model of RNAPII for reconstructing the RNAPII cryo-EM images, the remarkable match between the cryo-EM maps of RNAPII with the X-ray structure strongly indicated that our reconstruction algorithm was reliable and could be applied to other RNAPII complexes in this study.

Cryo-EM of the 13 subunit bovine RNAPII–Gdown1 elongation complex

To assure a complete stoichiometric presence of Gdown1 in the reconstituted RNAPII(G) samples, we added four-fold recombinant human Gdown1 (rGdown1) to the 12-subunit bovine RNAPII to form the RNAPII–rGdown1 complex. RNAPII–rGdown1 was previously determined to be functionally equivalent to native bovine RNAPII(G) (Hu *et al*, 2006). The ratio of four for reconstitution was determined by titration of rGdown1 to RNAPII followed by negative-stain EM observation to assess the minimal amount of rGdown1 required to turn the majority of RNAPII dimmers into monomeric particles (Supplementary Figure 3B). Such ratio of Gdown1 to RNAPII was found to completely inhibit

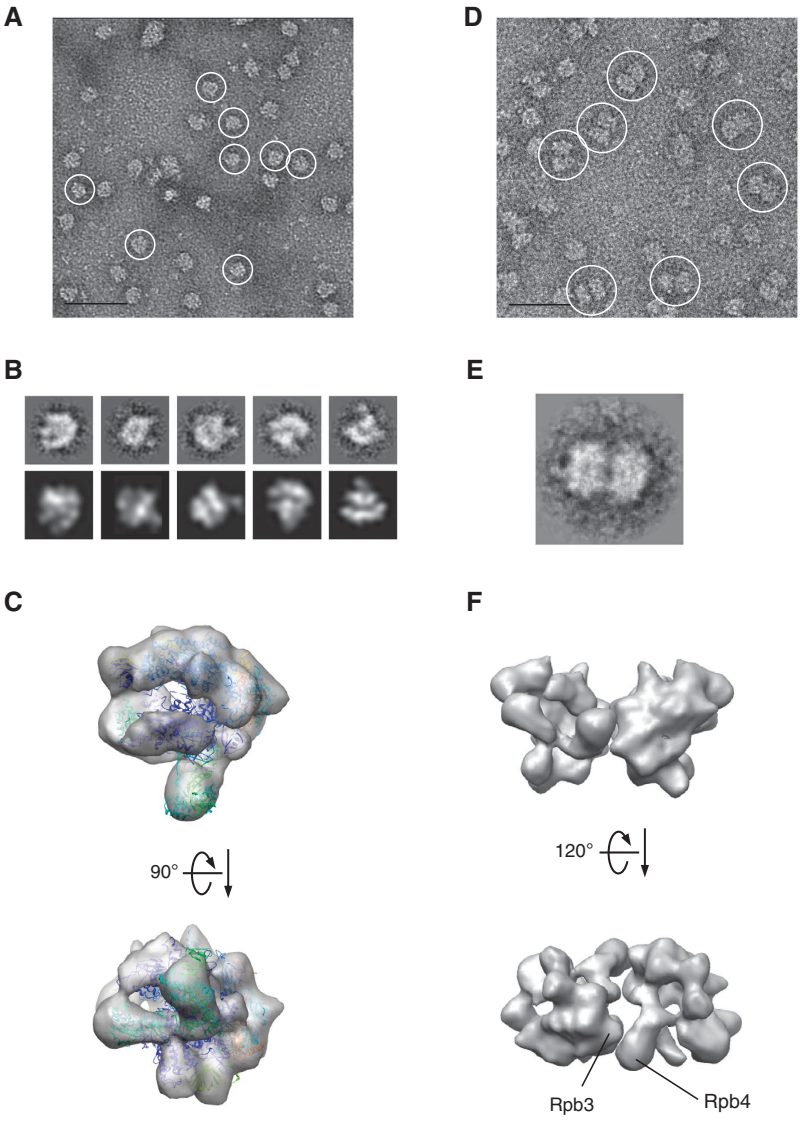

Figure 2 Image analyses of native bovine RNAPII complexes. (**A**) Native bovine RNAPII(G), enriched in monomers (white circles), were preserved in negative stain (2% uranyl acetate, 50 nm bar scale). (**B**) Upper row: class averages of bovine RNAPII(G) particles obtained after reference-free alignment and clustering reveal renowned RNAPII features such as groove and stalk (Rpb4/7). Lower row: re-projections from the 3D reconstruction of native RNAPII(G) that best match the class averages above. (**C**) EM structure of the native bovine RNAPII(G) superimposed with the yeast RNAPII X-ray structure (PDB: 1WCM). (**D**) Native bovine RNAPII, enriched in dimers (white circles), preserved in negative stain (50 nm scale bar scale). (**E**) A representative class average of RNAPII dimer. (**F**) 3D model of the RNAPII dimer built from reconstruction RNAPII monomers. An orientation best matches the class average in (**E**) shows subunit Rpb3 and Rpb4 likely participate the dimer interface.

promoter-specific transcription (Jishage *et al*, 2012). The RNAPII–rGdown1 complex is herein termed RNAPII–Gdown1. In addition, the RNAPII–Gdown1 complex was reconstituted with dsDNA/RNA to form the RNAPII–Gdown1 elongation complex, whose reconstruction would be used to compare with that of RNAPII elongation complex. The 3D cryo-EM reconstruction of RNAPII–Gdown1 (Supplementary Figure 3C) and RNAPII–Gdown1 elongation (Figure 4A) were both obtained to a resolution of 19 Å (Supplementary Figure 3E) from ~25 000 unstained particle images respectively via the same route as for the RNAPII elongation complex. By subtracting the cryo-EM volume of RNAPII elongation complex from that of RNAPII–Gdown1 elongation complex, a positive difference map was obtained and scored based on thresholding in units of standard deviation (σ). Those densities above 5σ spread on RNAPII extensively with the majority found in the vicinity of the DNA cleft (Figure 4B); they were estimated to account for a mass of ~40 kDa, fairly close to that of Gdown1. We divide those on the top of RNAPII into 'a' through 'e' according to where they dwell (Figure 4B). Remarkably, in region a, a bulky domain with mass of ~6 kDa above 10σ appeared on the Rpb5 shelf and connected to the Rpb1 jaw. We named this domain Gdown1-a. To verify the localization of Gdown1 on RNAPII, antibody labelling experiments were performed. By negative-stain EM, a polyclonal antibody against Gdown1 was located near the Rpb1 jaw where the Gdown1-a is (left panel in Figure 4D). To further locate the terminal regions of Gdown1, a glutathione S-transferase (GST) was fused to the N-terminus of Gdown1 and a monoclonal antibody against GST was found also near the Rpb1 jaw (right panel in Figure 4D). Since our structural results indicated Gdown1

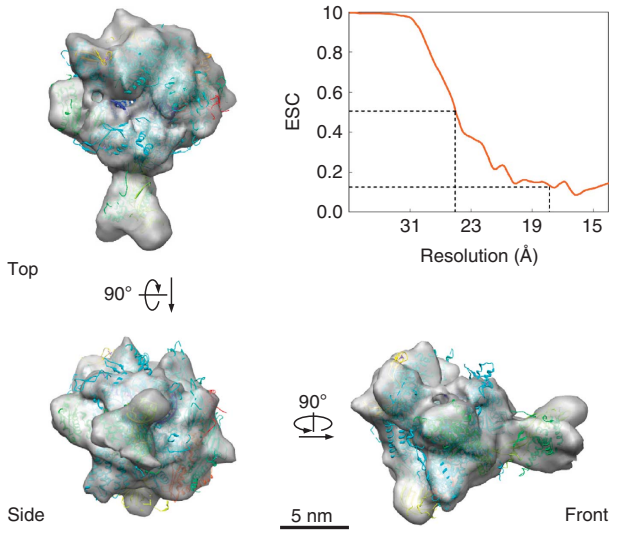

Figure 3 Cryo-EM reconstruction of the reconstituted bovine 12-subunit RNAPII elongation complex. Three views of the cryo-EM reconstruction of bovine RNAPII elongation complex at ~19 Å resolution (FSC 0.15). Superimposed with the EM envelop is the X-ray structure of yeast 12-subunit RNAPII (coloured ribbons) (PDB: 1Y1W). The threshold for rendering the EM reconstruction is chosen based on a molecular weight of ~520 kDa.

direct contacted the Rpb5 subunit of RNAPII, we tested if Gdown1 and Rpb5 would bind to each other with an *in vitro* pull-down assay. Indeed, recombinant Gdown1 and Rpb5 co-purified through two distinct affinity steps, supporting the notion that Gdown1 and Rpb5 would physically associate (Supplementary Figure 3F).

Cryo-EM of bovine RNAPII–TFIIF elongation complex

The structure of RNAPII–Gdown1 established herein has immediate functional implications. If common sites of association with RNAPII exist for other transcription factors, Gdown1 may poise a challenge for them to access RNAPII because Gdown1 binds to RNAPII so tightly as it does not dissociate either in high salt or in urea (Hu *et al*, 2006), in contrast to other transcription factors that readily dissociate from RNAPII in the presence of high salt (Cheng and Price, 2009).

Based on the reasons to be described, TFIIF was thought to be susceptible. In higher eukaryotes, TFIIF is composed of a large and a small subunit, RAP74 and RAP30, and their yeast homologues are dubbed Tfg1 and Tfg2 respectively. Yeast also contains a third TFIIF subunit Tfg3 not present in other eukaryotes (Henry *et al*, 1994; Kimura and Ishihama, 2004). In the previous cryo-EM study using the yeast proteins (Chung *et al*, 2003), the Tfg2 subunit was assigned in the DNA-binding cleft, mainly inferred from the crystal structures of bacteria RNAP holoenzyme (Murakami *et al*, 2002; Vassylyev *et al*, 2002), while the Tfg1 subunit was interpreted to reside largely alongside or on the RNAPII stalk, formed by subunits Rpb4 and Rpb7.

However, an alternative approach utilizing cleaving reagents mapped the residues of Tfg1 in the Tfg1/Tfg2 dimerization domain on the lobe/protrusion region of Rpb2 (Chen *et al*, 2007; Eichner *et al*, 2010). Recently, high-resolution cross-linking followed by mass spectroscopy allowed for positioning of almost all TFIIF residues on RNAPII (Chen *et al*, 2010), in which an extended track of Tfg1 was found on Rpb2, from the Rpb1 jaw-Rpb5 shelf to the Rpb2 lobe/protrusion. Providing that the interactions between RNAPII and TFIIF are largely conserved between yeast and mammals, RAP74—the mammalian homologue of Tfg1 would associate with mammalian RNAPII in a similar region.

To resolve the controversies as to the location of TFIIF on RNAPII, we re-investigated the cryo-EM structure of RNAPII–TFIIF but in the context of mammalian proteins by reconstituting bovine RNAPII with recombinant human TFIIF. The updated RNAPII–TFIIF cryo-EM map has direct relevance to mammals and would eliminate any imprecise inference from the yeast.

As the addition of recombinant TFIIF to bovine RNAPII did not turn the dimeric RNAPII into the monomeric form as effectively as Gdown1, we supplemented the RNAPII–TFIIF with nucleic acid scaffold, which is herein termed RNAPII–TFIIF elongation complex. The cryo-EM structure of the RNAPII–TFIIF elongation complex was reconstructed to a resolution of 19 Å (Supplementary Figure 4D) from ~25 000 monomeric particle images (Supplementary Figure 4B) using the similar approach for RNAPII(G). The 3D cryo-EM envelop of RNAPII–TFIIF elongation (Figure 5A) also matches well with that of the bovine RNAPII elongation cryo-EM structure. However, by subtracting RNAPII from RNAPII–TFIIF, densities attributed to TFIIF were revealed. Those scored above 5σ mainly appeared around the DNA cleft but not in the cleft and were distributed in regions labelled 'a' through 'e', used for describing Gdown1 (Figure 5B). Regions 'a' to 'c' accord with the localization of TFIIF on RNAPII by cross-linking experiments (Chen *et al*, 2010); for example, 'a' on the Rpb1 jaw domain corresponds to the charged domain of Tfg1 (RAP74), 'b' around the Rpb2 lobe area encircles the dimerization domain of Tfg1–Tfg2 (RAP74–RAP30) together with the N-terminus of Tfg1 (RAP74), and 'c' next to the Rpb2 protrusion corresponds to the linker of Tfg2 (RAP30). To further confirm the localization of RAP74 independently by EM, the same GST strategy was employed. The GST protein fused to the N terminus of RAP74 was detected on the Rpb2 side of RNAPII, near the jaw or lobe (Figure 5D), agreeing nicely with the finding by cross-linking (Chen *et al*, 2010). Importantly, as soon as the positive difference ascribed to TFIIF is overlaid with that of Gdown1 (Figure 5E), it is evident that most of Gdown1's sites clash with TFIIF. The conflict occurs on the jaw/shelf (region a in Figure 5E), where Gdown1-a meets with the charged region of RAP74 (Chen *et al*, 2010) and also likely with part of RAP30 (Wei *et al*, 2001; Le *et al*, 2005); near the lobe/protrusion (regions b and c), where Gdown1 clashes with the N-terminus of RAP74 and the RAP74–RAP30 dimerization domain (Chen *et al*, 2010); and on the clamp (region e), where Gdown1 can run into RAP30, according to the X-ray study of yeast RNAPII-Tfg2 (Kornberg, 2007). Such findings strongly suggest that Gdown1 and TFIIF would mutually exclude each other from accessing RNAPII.

Gdown1 blocks the interaction between RNAPII and TFIIF

To confirm that Gdown1 impedes TFIIF association with RNAPII, competition assays using size exclusion chromato-

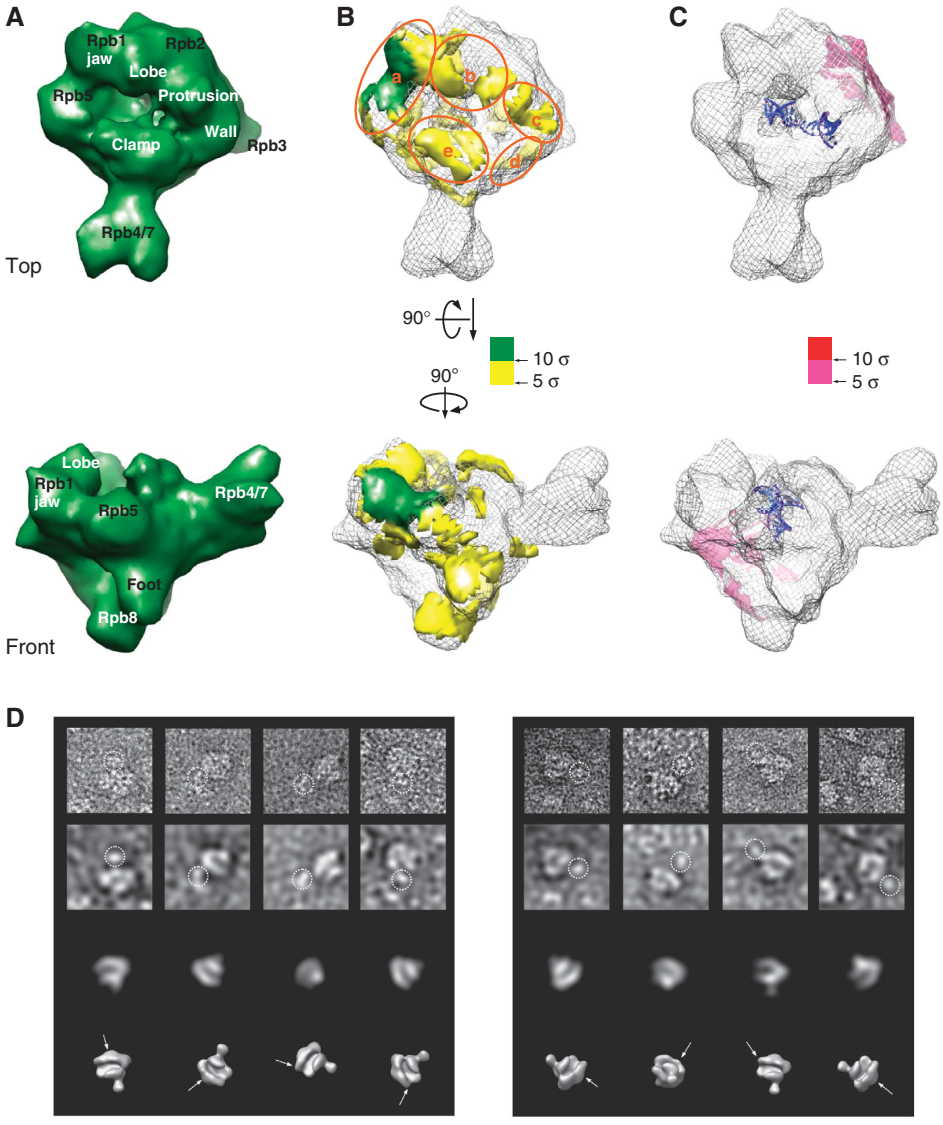

Figure 4 Cryo-EM reconstruction of reconstituted bovine RNAPII–Gdown1 and antibody labelling experiments. (**A**) Front and top views of the cryo-EM reconstruction bovine RNAPII–Gdown1 elongation at ~19 Å resolution (FRC 0.15) are depicted as solid deep green surface models. The threshold for rendering RNAPII–Gdown1 elongation is chosen based on a molecular weight of ~560 kDa. (**B**) A positive difference map was calculated between the bovine RNAPII–Gdown1 elongation and bovine RNAPII elongation complex (grey mesh) by using volumes that were both filtered to 15 Å, and shown in yellow hue above 5 σ. The most conspicuous density above 10 σ is in the gap between the Rpb5 shelf and the Rpb1 jaw is shown in deep green. (**C**) A negative difference map was calculated between the bovine RNAPII–Gdown1 elongation and bovine RNAPII elongation complex and shown in pink hue. The additional densities attributed to the RNAPII elongation complex are likely to be the result of conformational changes (outer densities). (**D**) The right panel presents RNAPII–Gdown1 EM analysis with a polyclonal antibody against Gdown1 (gift from Dr David Price, University of Iowa, USA). From top to bottom are raw negative-stained images, images filtered to 50 Å to reveal RNAPII gross features, matching projections of RNAPII at 50 Å, and the corresponding 3D models. The left panel shows a similar analysis employing a monoclonal antibody recognizing GST, fused to the N-terminus of Gdown1. Antibody densities are encircled in the top rows, and their location denoted with arrows in the lower 3D models.

graphy, namely gel filtration, were employed. To assess the migration behaviour of RNAPII, RNAP–TFIIF and TFIIF on a gel-filtration column, excess amount of recombinant human TFIIF (16 μg) were incubated with bovine RNAPII (10 μg, RNAPII to TFIIF ratio 1:8) and subjected to fractionation. As shown in Figure 6A, the silver-stained SDS–PAGE gel reveals the four largest subunits of RNAPII (Rpb1–Rpb3, Rpb5) and two subunits of TFIIF (RAP74 and RAP30), allowing for assignment of the molecular constituent in each fraction. Following these bands as markers, it appeared that RNAPII spread across many fractions (fractions 19–29). Nevertheless, RNAPII peaked in earlier fractions (fractions 20–22) together with the seemingly stoichiometric TFIIF co-migrating, indicating that the constituent was the RNAPII–TFIIF complex. The excess amount of free TFIIF (fractions 24–35) was partially resolved from the RNAPII–TFIIF complex (fractions 20–22) and ran slightly behind as anticipated. Likewise, an excess of Gdown1 (4 μg) was added to RNAPII (10 μg, RNAPII to Gdown1 ratio 1:4) to form the RNAPII–Gdown1 complex, resulting in a similar gel-filtration pattern; the RNAPII–Gdown1 complex appeared earlier (fractions 20–23) and the free Gdown1 followed (fractions 26–28) (Figure 6B). Interestingly, as Gdown1 was added to the pre-formed RNAPII–TFIIF complex, the TFIIF bands used to

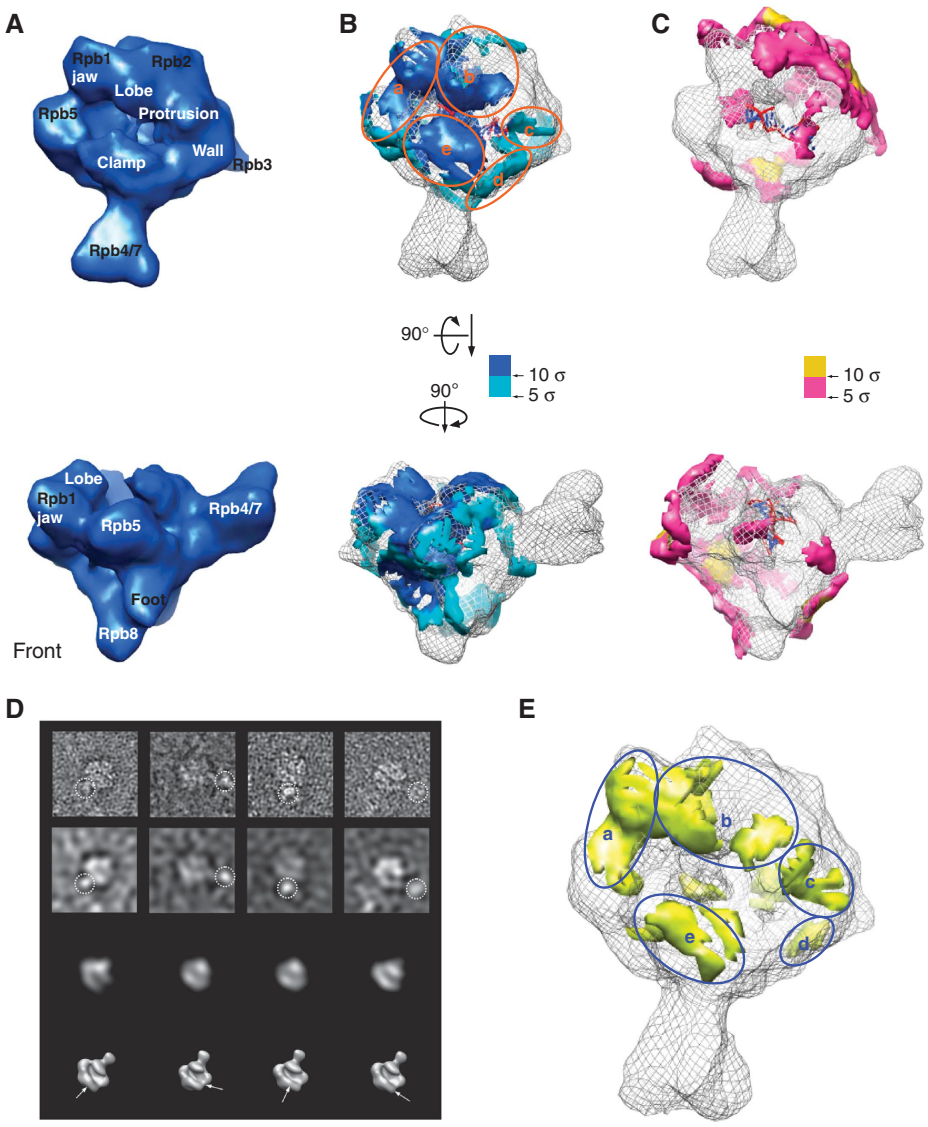

Figure 5 Cryo-EM reconstruction of RNAPII–TFIIF. (**A**) Top and front views of the cryo-EM reconstruction of bovine the RNAPII–TFIIF elongation complex at a resolution of ~19 Å (FRC 0.15). The threshold for rendering the complex is chosen based on the total mass of the complex (~620 kDa). (**B**) A positive difference map in blue hue is overlaid on the RNAPII elongation complex (grey mesh). The difference was obtained by subtracting RNAPII with nucleic acids from RNAPII–TFIIF with nucleic acid, both filtered to 15 Å. (**C**) A negative difference map in red hue is overlaid on the RNAPII elongation complex (grey mesh). (**D**) RNAPII–TFIIF images with a monoclonal antibody recognizing the GST fused to the N-terminus of TFIIF. From top to the bottom are raw negative-stained images, images filtered to 50 Å to reveal RNAPII gross features, matching projections of RNAPII at 50 Å, and the corresponding 3D models. (**E**) Common, overlapping regions attributed to both Gdown1 in Figure 4B (green) and regions attributed to TFIIF in Figure 5B (blue) are shown in yellow.

co-migrate with RNAPII in those earlier fractions (fractions 20–22) were retarded to the major positions of free TFIIF (after fraction 24) (Figure 6C). The replacement of TFIIF by Gdown1 in those earlier fractions indicates that Gdown1 displaces TFIIF from RNAPII, in keeping with the proposed steric exclusion mechanism. On the contrary, in the reverse experiment, the excess TFIIF (16 μg) employed to challenge the pre-formed RNAPII–Gdown1 failed to displace Gdown1 bound to RNAPII (Figure 6D).

Discussion

In this study, the single-particle cryo-EM technique was used to visualize the 3D structure of mammalian RNAPII and its isoform RNAPII(G) in the unstained state. Importantly, the Gdown1 domain 'a' of ~6 kDa was found situated on the Rpb5 shelf and connected to the adjacent Rpb1 jaw. It is noted that the interactions between Gdown1 and RNAPII in this region was recently reported by using cross-linking method (Jishage *et al*, 2012). Our observation of Gdown1 on RNAPII supports the idea that it is largely flexible or disordered because otherwise a bulky volume as large as the stalk of Rpb4–Rpb7 (45 kDa) would have been detected. Although at the current resolution, it is difficult to define the domain organization or orientation of Gdown1 on RNAPII, some questions along this line can be addressed: for instance, what would be the primary sequence that corresponds to Gdown1-a domain that seems to have a good tertiary structure? This sequence has to meet two

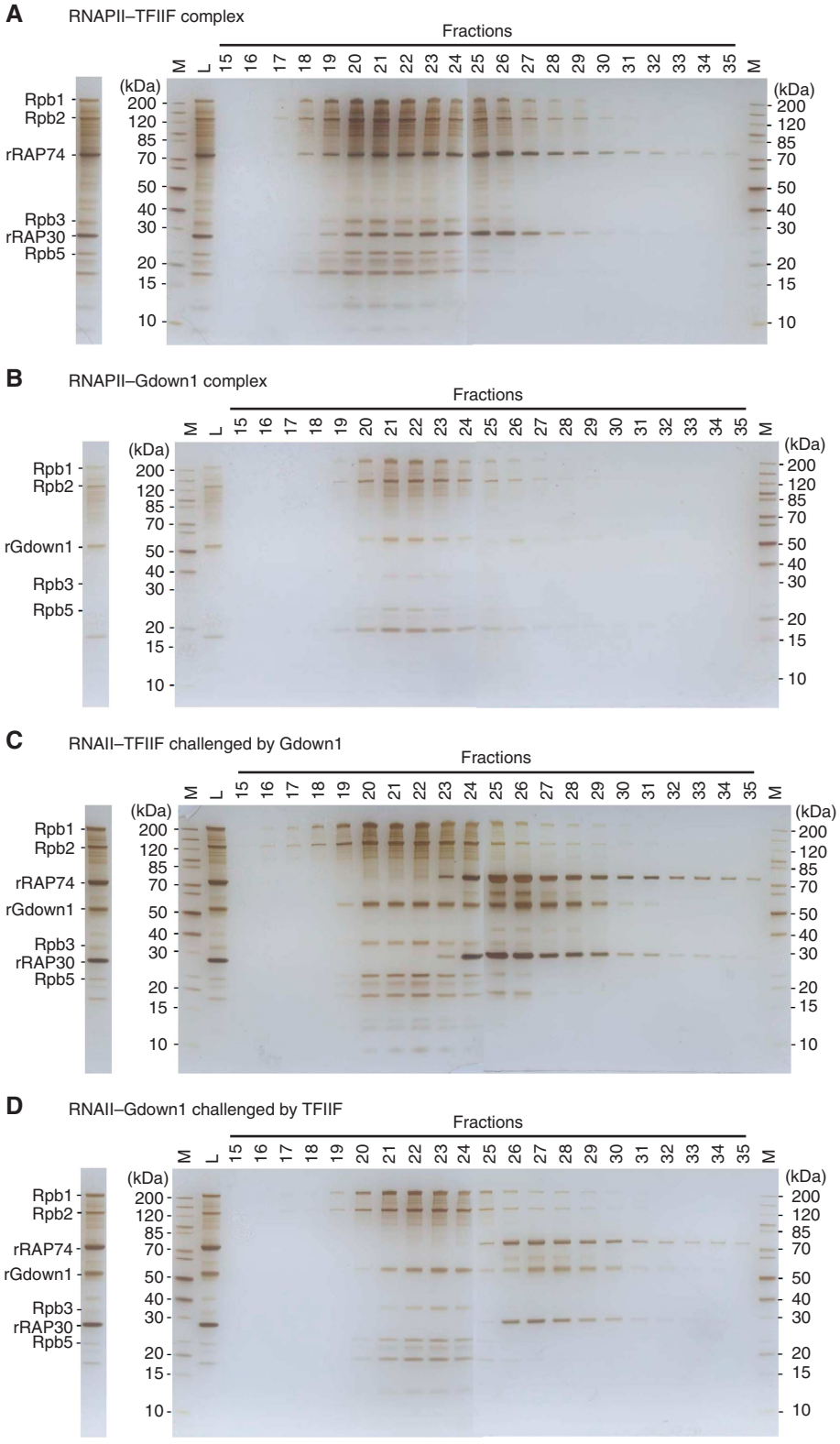

Figure 6 Competition assay using fractionation on a size exclusion column (Superose 6). (**A**) Major fractions of RNAPII–TFIIF complex (18–22) and major fractions of excess amount of free TFIIF (24–35) are visualized on SDS–PAGE with silver staining. To form RNAPII–TFIIF, eight-fold TFIIF was used. (**B**) Fractions of RNAPII–Gdown1 complex (20–23) and major fractions of free Gdown1 (26–28). To form RNAPII–Gdown1, four-fold Gdown1 was used (**C**) The RNAII–TFIIF complex challenged by Gdown1. The RNAPII-associated TFIIF (18–22) now disengages and becomes mostly free (24–35). RNAPII bound TFIIF are replaced by Gdown1 (19–22). (**D**) The RNAPII–Gdown1 complex challenged by TFIIF. Gdown1 remains mostly associated with RNAPII (21–24) and TFIIF remains free (25–35). Marked bands are major RNAPII subunits: Rpb1 (220 kDa), Rpb2 (133 kDa), Rpb3 (31 kDa), Rpb5 (25 kDa), TFIIF subunits: RAP74 (74 kDa), RAP30 (28 kDa), and rGdown1. In the gel, M stands for marker and L for load. The stained gel strip on the left of each panel represents ~2.5% of the column load.

criteria: (1) domain criterion; it forms a fold that resists proteolysis; (2) proximity criterion; it is close to the N-terminus because Gdown1-a is near the antibody recognizing the GST fused to the N-terminus (<40 Å including the GST dimensions); note that this condition would not be valid if Gdown1 were largely folded. As such, the best candidate would be the folding segment in the N-terminal half of Gdown1 (green island in Figure 1C). Strikingly, the total mass (~6.5 kDa) of those 59 amino acids virtually coincide with that of Gdown1-a. Of course, it awaits confirmation, perhaps best by high-resolution cross-linking experiments. Another question is which part of Gdown1 that causes RNAPII to form monomer? Our modelling analysis (Figure 2F) suggests that the Rpb3 subunit of RNAPII is situated at the dimer interface. In fact, Rpb3 was reported to cross-link with Gdown1 as well (Jishage *et al*, 2012). Indeed, a positive density ascribed to Gdown1 was detected on Rpb3 (Supplementary Figure 3D), and we reason that it corresponds to the sequence (277–358), which shows similarity to the C-terminal region of RAP30 (Jishage *et al*, 2012), whose yeast homologue region cross-links to Rpb3 (Chen *et al*, 2010).

Furthermore, we updated the cryo-EM structure of RNAPII–TFIIF using the mammalian proteins. The densities ascribed to TFIIF tether on RNAPII in the vicinity of DNA cleft in a massive manner and RAP74 is located on the Rpb2 side, not alongside the stalk of Rpb4–Rpb7 (Chung *et al*, 2003). Such assignment is entirely consistent with the results obtained by the cross-linking studies using homologous yeast proteins (Chen *et al*, 2010). As the densities of Gdown1 and those of TFIIF were overlaid, overlap of densities occurs on the regions of RNAPII as aforementioned. The comparison of RNAPII–TFIIF with RNAPII–Gdown1 revealed that TFIIF shared with Gdown1 not only a number of common sites but also similar flexible character.

That Gdown1 could exclude TFIIF was confirmed by gel filtration herein. Such exclusion has a profound impact on any process involving the RNAPII–TFIIF machinery. During transcription initiation, TFIIF is needed for the formation of the PIC complex to facilitate RNAPII initiation at a promoter. As such, Gdown1 exclusion of TFIIF would inhibit initiation, providing for a mechanism to explain the previous finding of Gdown1 repressing the promoter-specific transcription (Hu *et al*, 2006). Considering that TFIIF is at the heart of basal transcription, Gdown1 repression might be mitigated by a large excess of TFIIF, which was demonstrated by Roeder's lab (Jishage *et al*, 2012). However, in our hands two-fold excesses of TFIIF did not displace Gdown1, and furthermore we do not believe that extremely large excesses of TFIIF that might potentially dislodge Gdown1 are available *in vivo* for such a task. As such, it is important to consider the role played by Mediator in these events.

Previously, Mediator was shown to remove the Gdown1 restriction on activated transcription (Hu *et al*, 2006). Since elongation factors were absent from that study, Mediator necessarily acted during initiation. It is therefore likely that Mediator would be involved in the removal of Gdown1 from RNAPII at the stage of PIC formation through a mechanism coupled with Mediator function in enhancing the assembly of the PIC (Cantin *et al*, 2003; Baek *et al*, 2006; Takagi and Kornberg, 2006). At a promoter, Mediator could associate with RNAPII, and perhaps with RNAPII(G) as well (Jishage *et al*, 2012). Interestingly, a biochemical study on the head module of Mediator using yeast proteins indicated that the head module would not stably associate with RNAPII in the absence of TFIIF (Takagi *et al*, 2006). This line of thought was augmented by a recent cryo-EM study of human Mediator complex by Taatjes's team (Bernecky *et al*, 2011), showing the docking RNAPII on the Mediator appears to be dynamic, while TFIIF can promote the stability of the Mediator–RNAPII complex by specifying the orientation of RNAPII on the Mediator scaffold. Conversely, Mediator would promote the association of RNAPII and TFIIF. With this in mind and based on our finding of the TFIIF/Gdown1 steric overlap together with Gdown1's tight binding to RNAPII, we reason that the crosstalk of TFIIF with Gdown1 may be facilitated by Mediator during a three-way meeting, which could account for Mediator relief of Gdown1 repression at the promoter. To elaborate the molecular details, we introduce a scheme relating to the interplay among RNAPII, TFIIF, Mediator and Gdown1 according to the principles of protein association/dissociation and the stability of the Mediator–RNAPII–TFIIF complex (Figure 7). This would assure a highly regulated process with effective initiation for RNAPII(G) only in the presence of activators that can subsequently recruit Mediator to 'swap out' Gdown1 for TFIIF.

At the level of elongation, one would ask what and how Gdown1 might do since Gdown1 would take place of TFIIF, which is thought to remain associated with RNAPII after the promoter clearance and drastically increase the rate of transcription (Tan *et al*, 1994; Gu and Reines, 1995). First, the regulatory effect through steric interference by Gdown1 may not be limited to TFIIF but can be applicable to a number of RNAPII-associated elongation proteins other than TFIIF. For example, it is known that stalling of RNAPII can occur near the promoter, which depends on proteins of DSIF and NELF (Yamaguchi *et al*, 1999a, b). Such promoter proximal pausing is relieved by pTEFb phosphorylating Rpb1, the largest RNAPII subunit (Price, 2008). Can Gdown1 interfere with the binding of DSIF? To examine such, we compared our RNAPII–Gdown1 structure with the related one in which the RNAP is associated with Spt5, the homologue of DSIF (Klein *et al*, 2011), and also with the structure of RNAPII–TFIIS (Kettenberger *et al*, 2003) or RNAPII–CE (Suh *et al*, 2010), and therefore suggested that Gdown1 would affect DSIF but would not do so to TFIIS or CE. These conjectures have been confirmed by Price's *in-vitro* experiments (Cheng *et al*, 2012). Secondly, the observation of modest 1.5- to 2.5-fold increase in RNA production and the transcript length enhancement by Gdown1 (Cheng *et al*, 2012), though less efficient than TFIIF, prompts us to suspect that Gdown1 is a RNAPII processivity factor. Perhaps our EM structures presented herein would account for the structural basis of the enhancement of elongation and that the common enhancement of Gdown1 and TFIIF may be the result of their association with the common RNAPII regions by stabilizing the elongation prone conformation of RNAPII and restraining template release. To elaborate, consider that Gdown1-a on the Rpb1 Jaw-Rpb5 shelf can contribute to securing the incoming DNA and the C-terminal region has potential cryptic DNA-binding activity for upstream DNA (Price *et al*, 1989; Tan *et al*, 1994), while the entirety of Gdown1 can impact the relative positions and/or stability of the two RNAPII mobile elements: the jaw lobe and the clamp.

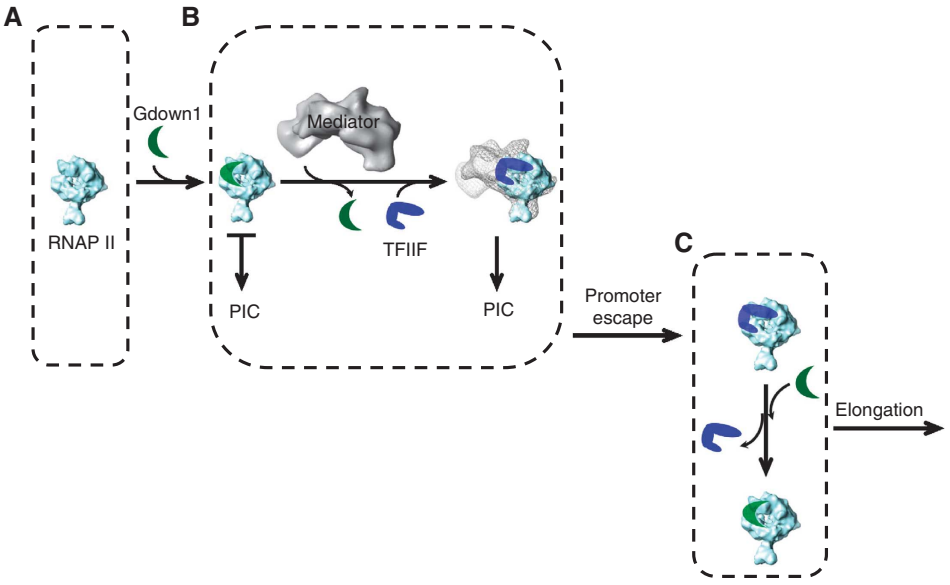

Figure 7 Steric model of Gdown1 repression and relief by Mediator. In this model, Mediator acts as a scaffold to promote the association of RNAPII with TFIIF to create an unfavourable circumstance for Gdown1 to dwell on RNAPII at the overlapping sites of TFIIF; henceforth destabilizing the association between RNAPII and Gdown1 and eventually yielding the swapping between Gdown1 and TFIIF on RNAPII to allow for transcription initiation to ensue (RNAPII is coloured in cyan, Gdown1 in green, TFIIF in blue, and Mediator in grey. PIC stands for 'pre-initiation complex'). (**A**) When RNAPII(G) is recruited to the promoter, the PIC would not be generated considering the exclusion of TFIIF on RNAPII and transcription would be prevented. (**B**) Steric hindrance of TFIIF binding to RNAPII by Gdown1 is relieved by the Mediator complex and Mediator swaps TFIIF for Gdown1 on RNAPII. (**C**) Once Gdown1 is removed, promoter escape ensues with the possibility of some RNAPII enzymes to re-associate with Gdown1, in accord with recent findings of RNAPII(G) downstream of promoters (Cheng *et al*, 2012) and others to remain with TFIIF. Considering that RNAPII(G) is substoichiometric, it is also possible that RNAPII initiates at promoters in the absence of Gdown1 and therefore unhindered by Gdown1.

By encompassing the DNA on all sides, akin to closing a door, which is slightly ajar, the opportunity for DNA to depart from the enzyme is greatly reduced. That Gdown1 can work as a processivity factor is in accordant with the observation of RNAPII elongation complexes entering productive elongation without loss of Gdown1 (Cheng *et al*, 2012).

Considering that a Gdown1 homologue does not exist in yeast, the question as to Gdown1's role in metazoans is raised. Perhaps for complex organism development there is much that can go wrong and Gdown1 presents an additional level of scrutiny requiring careful execution of gene expression. We also speculate that Gdown1 may be a master brake, limiting unwanted transcription of some genes by preferentially localized with proximity to their promoters (Cheng *et al*, 2012; Jishage *et al*, 2012) until Mediator removes Gdown1 and allows for TFIIF association. Another important need for Gdown1 could be in processivity, especially for extremely long genes such as the 2.5-Mb human dystophin gene which can require 16 h to transcribe (Tennyson *et al*, 1995). Gdown1 would secure and stabilize the elongation complex, yet be compatible with the TFIIS elongation factor (Wind and Reines, 2000) to aid in generating complete RNA transcripts. Such large genes are mostly metazoan specific, which perhaps explains a need for Gdown1 in higher organisms to generate some very extended RNA molecules.

In conclusion, our study has revealed a structural basis underlying the Gdown1 regulatory mechanism in part by restricting TFIIF association with RNAPII. In general, it is suggested that Gdown1 can restrict or permit RNAPII-binding proteins other than TFIIF that are involved in initiation and elongation.

Materials and methods

RNAPII and RNAPII(G) purification

RNAPII and RNAPII (G) were purified as previously described (Hu *et al*, 2006). Briefly, RNAPII was enriched by precipitating the homogenized and clarified calf thymus extract with polyethyleneimine (PEI); the pellet was dissolved, cleared by centrifugation, followed by Mono Q chromatography, and affinity chromatography using a monoclonal 8WG16 antibody column (Thompson *et al*, 1990). The eluted proteins were subjected to UNO-Q HPLC chromatography (Bio-Rad) to separate RNAPII and RNAPII(G). Both forms of RNAPII were stored as ammonium-sulphate precipitate under −80°C and shipped on dry ice via TNT service. For the subsequent EM and biochemical characterization, the frozen ammonium-sulphate precipitant was thawed immediately upon arrival, and dissolved in a buffer containing 50 mM Tris–Cl (pH = 7.5), 200 mM potassium acetate (KOAc), 5 mM $MgCl_2$, 5 mM DDT and 10% glycerol. The removal of ammonium-sulphate was achieved by buffer exchange with the dissolving buffer using an Amicon centrifugal filter unit (MWCO 100 kDa; Millipore). The proteins were further stored as aliquots (1 mg/ml) at −80°C.

RNAPII(G) nonspecific promoter assay

The transcription elongation assays were performed as previously described (Gnatt *et al*, 1997). Briefly, nonspecific initiation on a tailed template was employed. The tailed template was formed by annealing a nontemplate strand: 5′-AACACCAGCGAGCAAGGCGTT TCGGGGAAGAAAAA, and a template strand: TTTTTCTTCCCCGAA ACGCCTTGCTCGCTGGTGTTCCCCCCCCCCCC. 0.4 or 0.8 μg of polymerase, RNAPII or RNAPII(G), were employed, and α-amanitin, an RNAPII-specific inhibitor, was used as negative control. Quantification of bands was performed using the NIH free ImageJ software.

Expression of recombinant human Gdown1 and TFIIF

Recombinant Gdown1 expression plasmid, originally constructed in the pET151 vector (Hu *et al*, 2006), was subcloned into a pET21a vector using *Nde*I and *Not*I restriction sites. The Gdown1 protein was overexpressed in BL21 cells (Invitrogen) by IPTG induction (0.4 mM) and grown overnight at 20°C. The protein was purified by

using a nickel-NTA column with standard procedure as described by the manufacturer (Sigma). Further HPLC purification was depicted in Supplementary Figure 3A. The purified proteins were stored at −80°C. A bacterial expression plasmid encoding human TFIIF was constructed in the pACYCduet1 vector. To express TFIIF, *Escherichia coli* BL21 (DE3)-RIL cells (Merck) transformed with the TFIIF containing duet plasmid was grown at 37°C and the temperature was lowered to 30°C when OD reached 0.4 for induction with 0.84 mM IPTG. The cells were further grown for 3 h at 30°C before harvest. TFIIF was purified using a nickel-NTA column in the same manner as Gdown1 was, with slight modifications in the washing step by raising the concentration of NaCl to 300 mM and that of imidazole to 40 mM (see also Supplementary Figure 4A).

Limited proteolysis analysis of Gdown1

Purified recombinant Gdown1 was digested with trypsin (7500:1) at 30°C. After 40 min, the reaction was quenched by heating at 95°C for 10 min. The digestion products were subsequently analysed by LC mass spectrometry.

Size exclusion chromatography assays

To form RNAPII–Gdown1 or RNAPII–TFIIF complex, 10 μg of bovine RNAPII was mixed with individual recombinant proteins, rGdown1 or rTFIIF, and incubated for 60 min at 20°C, respectively. For rGdown1, 4 μg was used, representing a rGdown1 to RNAPII ratio of 4:1. As to the formation of RNAPII–TFIIF, about eight-fold recombinant TFIIF (16 μg) was used. The size exclusion chromatography was carried out using the AKTA purifier system (GE Health/Amersham Biosciences). A Superose 6 column PC 3.2/30 (GE Healthcare) was equilibrated with equilibrating buffer (50 mM Tris–HCl (pH 7.5 at 4°C), 100 mM NaCl, 1 mM TCEP, and 1% protease inhibitor cocktail). The protein sample (~30 μl) was loaded onto the column for fractionation with a total elution volume of 3.2 ml at a flow rate of 0.06 ml/min. The fractions were collected with 60 μl per fraction. The protein complex in each fraction was concentrated with nickel beads (Sepharose High Performance, GE Healthcare). In brief, 10 μl pre-equilibrated Ni-beads were added to individual fraction from (15–35) respectively, and incubated with mixing at 4°C overnight. The supernatant was discarded; the same volume of 2× SDS–PAGE sample buffer was added to the beads, boiled on a heating block for 10 min. The denatured protein samples were separated on a NuPAGE 4–12% Bis-Tris gel using MES running buffer and visualized with silver stain. In all, 5 μl of 10-fold diluted protein ladder was loaded onto an SDS–PAGE gel. To challenge RNAPII–Gdown1 with rTFIIF, RNAPII was first incubated with four-fold rGdown1 at 20°C and mixed at 550 r.p.m. for 60 min. After the first incubation period, eight-fold rTFIIF was added, and another 60 min of incubation was performed, followed by loading the protein samples onto a Superpose 6 PC 3.2/30 size exclusion column. To challenge RNAPII–TFIIF with rGdown1, RNAPII was first incubated with four-fold rGdown1, and eight-fold rTFIIF was added after the first incubation period for an additional incubation period before size exclusion chromatography was carried out.

EM sample preparation

In brief, a frozen aliquot of bovine RNAPII or RNAPII(G) (1 mg/ml) was freshly thawed and diluted with de-ionized water to a final concentration of 50 μg/ml for direct EM usage or for reconstitution. To reconstitute bovine RNAPII complexes, incubation of bovine RNAPII with additional proteins were performed at 20°C for 60 min before the EM sample was prepared. To form a RNAPII–Gdown1 complex, the molar ratio of recombinant Gdown1 to RNAPII was adjusted to 4:1. For RNAPII–elongation complexes, three-fold of nucleic acid scaffold containing the bubble DNA and a 35-nt RNA was used (Supplementary Figure 2A). For RNAPII–TFIIF, six-fold recombinant TFIIF together with three-fold nucleic acid scaffold was added. In this study, three different techniques for preserving RNAPII complexes were employed, including negative-stain, cryo-staining and unstained cryo. To make an EM specimen, 3 μl of protein solution was applied to Cu grid (300 mesh) or quantifoil grid coated with thin carbon film to enhance the Thon ring observation. The carbon film was made by evaporating a carbon rod onto a cleaved mica sheet with an evaporator (Edwards Auto 306), deposited on the grid to dry, and freshly discharged with amylamine in a glow-discharger (EM Science, EM-100) before applying protein solution. For the negative-stain technique, 2% uranyl acetate was used; the sample was either sandwiched by another layer of thin carbon (Tischendorf *et al*, 1974) or deep-stained (Stoops *et al*, 1991). For cryo-staining, the protein-adsorbed grids were first stained with 1% uranyl acetate and not allowed to dry; the sample was immediately blotted with a filter paper (Wattmann No. 1, smooth side) and plunged into liquid ethane with a customer built plunger in a cold room with 80% humidity (Golas *et al*, 2003; Ohi *et al*, 2004). The frozen EM samples were then stored under liquid nitrogen before EM. For unstained cryo preservation, the procedure was the same as cryo-staining except no staining agent was used.

Electron microscopy

Negative-stain preserved native bovine RNAPII or RNAPII(G) were collected on Kodak SO-163 films by a 200-kV HRTEM (JEOL 2011, LaB_6 filament, Cs = 1.0 mm) at magnification of ×40 000 with ~10 $e^-/Å^2$, defocus of ~2–3 μm. The micrographs were developed for 12 min at 25°C in full strength with the Kodak D-19 solution, dried and digitized on a Zeiss/Integraph flat-bed densitometer using a step size of 14-micron, corresponding to 3.5 Å on the object scale. For cryo-staining, Cu grids (300 mesh) were loaded with a layer of carbon film and glow discharged with amylamine; the specimens were observed using a Gatan 914 quick-load cryo-holder on a 200-kV field-emission microscope (JEOL 2010F, Cs = 3.0 mm) at magnification of ×52 000 and images of the reconstituted RNAPII complexes were collected on a 4K × 4K CCD (Gatan 895) with the doses controlled at ~10–20 $e^-/Å^2$ and defocus near ~2.5 μm. At such magnification, each CCD pixel corresponded to 2.9 Å. For cryo-EM imaging of unstained specimens, Quantifoil R2/2 grids (Quantifoil Micro Tools GmbH) were loaded with a very thin layer of carbon film and glow discharged with amylamine; images of the reconstituted RNAPII complexes were collected on a 4K × 4K CCD (Gatan 895) by the same 200 kV field-emission microscope with the magnification adjusted to ×78 000 and the doses ~10–20 $e^-/Å^2$ with defocus controlled in the range ~2.0–4.0 μm. At such magnification, each CCD pixel corresponded to 1.9 Å.

Immuno-electron microscopy

For antibody decoration against the GST-tagged Gdown1 in the RNAPII–Gdown1 complex or against the RAP74 subunit within the RNAPII–TFIIF complex, RNAPII complexes were incubated with two-fold antibody and then subjected to gel filtration for fractionating the RNAPII–Gdown1 antibody or the RNAPII–TFIIF antibody complex. Application of the polyclonal antibody against Gdown1 also followed the same procedure. To disclose the location of the antibody in complex, a 120-kV TEM (JEOL 1400, LaB_6 filament, Cs = 3.4 mm) and 4K × 4K CCD camera (Gatan 895) were used; the magnification used for visualizing antibody was ×104 000 and the defocus was ~1.5 μm.

Single-particle reconstruction of RNAPII complexes

To analyse RNAPII(G) preserved in negative-stain, a total of 7689 particle images were interactively selected by using EMAN BOXER (Ludtke *et al*, 1999). The images were pre-aligned on SPIDER (Frank *et al*, 1996), transferred to XMIPP for classification by the CL2D method (Sorzano *et al*, 2004) into 128 classes (Supplementary Figure 1A). Subsequently, representative class averages were selected and an initial model was derived by the common-line method (Penczek *et al*, 1996). To reconstruct a 3D structure of RNAPII(G) by back-projection, the initial model was injected for assigning the initial orientation parameters to the individual raw images by the projection matching method (Penczek *et al*, 1994), by which 10–20 cycles of angle refinement were used until convergence occurred.

Before we pursued cryo-EM of RNAPII complexes in the unstained state, reconstruction of each complex preserved by cryo-staining was generated. More precisely, the negative-stain reconstruction of RNAPII(G), lowpass filtered to 40 Å, was used as initial model for calculating the cryo-stained reconstruction of RNAPII–Gdown1 from a total of 13 000 cryo-stained images. The same procedure was repeated for obtaining the cryo-stained reconstruction of RNAPII in an elongation form also using 13 000 images, and for obtaining cryo-stained reconstruction of RNAPII–TFIIF, 13 000 images were employed as well. The resolution of reconstruction was estimated by the degree of similarity between the two reconstructions derived from each half-set using Fourier shell correlation (FSC) method and the reported resolution is based on FSC = ~0.15 (Kostek *et al*, 2006).

A total of 25 000 cryo-EM images of RNAPII–Gdown1 in the unstained state were used to calculate the cryo-EM reconstruction of RNAPII–Gdown1 by the back-projection method using the cryo-stained reconstruction of RNAPII–Gdown1 as the initial model. The same approach was used for obtaining the cryo-EM reconstruction of RNAPII in elongation form (~20 000 particles) and for that of RNAPII–TFIIF (25 000 particles) by using the respective cryo-stained reconstruction as the initial model. The resolution of the cryo-EM reconstruction was estimated by the FSC method as described. The difference mapping was performed using the RobEM program (R Ashmore and T Baker, unpublished data) and two thresholds (5σ and 10σ) were used to present the positive and negative difference maps.

Docking of the X-ray structure and volume rendering

UCSF Chimera (Goddard *et al*, 2007) was used to analyse and display the EM reconstructions. The PDB ribbons of yeast RNAPII (1WCM or 1Y1W) were first roughly aligned into the reconstruction volumes by visualization. Further optimization of the fitting was achieved by using 'Fit Model in Map' of Chimera. Each reconstruction was rendered with a threshold yielding a volume consistent with the molecular weight by assuming an average protein density of 1.21 $Å^3$ per dalton (Harpaz *et al*, 1994).

EMDataBank accession codes

The cryo-EM maps of bovine RNAPII elongation complex, RNAPII–Gdown1 complex, RNAPII–Gdown1 elongation complex and RNAPII–TFIIF elongation complex have been deposited in the EM database (http://emdatabank.org/) with accession codes EMD-5440 to EMD-5443.

Supplementary data

Supplementary data are available at *The EMBO Journal* Online (http://www.embojournal.org).

Acknowledgements

We acknowledge the Instrument Center of Academia Sinica (AS), Institute of Chemistry of AS, and YK. Hwu at Institute of Physics of AS for cryo-electron microscopes. W-HC thanks Yao-Yin Chuang and Yi-Yum Chen for assistance on TEM, Chen-Yu Li for helping trypsin digestion and Ting-Yi Wu for boxing cryo-EM images, and the Mass Spectroscopy Center at Institute of Chemistry of AS. W-HC is grateful for Dr Burton for TFIIF plasmids. W-HC and AG are indebted to Dr Price for anti-Gdown1 antibody and to Dr Price and Dr Roeder for communicating unpublished results of Gdown1. W-HC was supported by AS Thematic Grant (AS-99-TP-A03), AS Nanoscience programme, and National Science Council (NSC) of Taiwan (NSC99-2113-M-001-022-MY3 and NSC98-2113-M-001-020) for this work; Dr Y-M Wu and Y-C Lin have been supported by NSC postdoctoral fellowships (NSC98-2811-M-001-141, NSC96-2811-M-001-083 respectively). XH was supported by Natural Science Foundation of China (Nos. 30800169). AG was supported by the National Institutes of Health Grant (GM64474) and by the DOD Breast Cancer IDEA Award BC083495.

Author contributions: YM Wu did cryo-EM and image reconstruction; JW Chang did gel-filtration competition assays; JW Chang, SH Huang, and PL Wu did antibody EM; YC Lin compiled reconstruction algorithm; CH Wang and CC Chang did GST fusion and subcloning of recombinant TFIIF and Gdown1; X Hu did pull-down assay; X Hu and A Gnatt purified RNAPII and RNAPII(G). YM Wu and WH Chang designed the experiments and analysed the data; YM Wu, JW Chang, X Hu, A Gnatt, and WH Chang wrote the manuscript. Results and opinions herein represent those of the authors alone and not of the funding institutions unless otherwise specified.

Conflict of interest

The authors declare that they have no conflict of interest.

References

Armache KJ, Mitterweger S, Meinhart A, Cramer P (2005) Structures of complete RNA polymerase II and its subcomplex, Rpb4/7. *J Biol Chem* **280:** 7131–7134

Asturias FJ, Chang W, Li Y, Kornberg RD (1998) Electron crystallography of yeast RNA polymerase II preserved in vitreous ice. *Ultramicroscopy* **70:** 133–143

Baek HJ, Kang YK, Roeder RG (2006) Human Mediator enhances basal transcription by facilitating recruitment of transcription factor IIB during preinitiation complex assembly. *J Biol Chem* **281:** 15172–15181

Belakavadi M, Fondell JD (2006) Role of the mediator complex in nuclear hormone receptor signaling. *Rev Physiol Biochem Pharmacol* **156:** 23–43

Bernecky C, Grob P, Ebmeier CC, Nogales E, Taatjes DJ (2011) Molecular architecture of the human Mediator-RNA polymerase II-TFIIF assembly. *PLoS Biol* **9:** e1000603

Cai G, Imasaki T, Takagi Y, Asturias FJ (2009) Mediator structural conservation and implications for the regulation mechanism. *Structure* **17:** 559–567

Casamassimi A, Napoli C (2007) Mediator complexes and eukaryotic transcription regulation: an overview. *Biochimie* **17:** 1439–1446

Cantin GT, Stevens JL, Berk AJ (2003) Activation domain-mediator interactions promote transcription preinitiation complex assembly on promoter DNA. *Proc Natl Acad Sci USA* **100:** 12003–12008

Chen C-Y, Chang C-C, Yen C-F, Chiu MT-K, Chang W-H (2009) Mapping RNA exit channel on transcribing RNA polymerase II by FRET analysis. *Proc Natl Acad Sci USA* **106:** 127–132

Chen H-T, Warfield L, Hahn S (2007) The positions of TFIIF and TFIIE in the RNA polymerase II transcription initiation complex. *Nat Struct Mol Biol* **8:** 696–703

Chen ZA, Jawhari A, Fischer L, Buchen C, Tahir S, Kamenski T, Rasmussen M, Lariviere L, Bukowski-Wills J-C, Nilges M, Cramer P, Rappsilber J (2010) Architecture of the RNA polymerase II–TFIIF complex revealed by cross-linking and mass spectrometry. *EMBO J* **29:** 717–726

Cheng B, Li T, Rahl PB, Adamson TE, Loudas NB, Varzavand K, Gu J, Varzavand K, Cooper JJ, Hu X, Gnatt A, Young RA, Price DH (2012) Functional association of Gdown1 with RNA polymerase II poised on human genes. *Mol Cell* **45:** 38–50

Cheng B, Price DH (2009) Isolation and functional analysis of RNA polymerase II elongation complexes. *Materials Methods* **48:** 346–352

Chung WH, Craighead JL, Chang W-H, Ezeokonkwo C, Bareket-Samish A, Kornberg RD, Asturias FJ (2003) RNA polymerase II/TFIIF structure and conserved organization of the initiation complex. *Mol Cell* **12:** 1003–1013

Conaway RC, Sato S, Tomomori-Sato C, Yao T, Conaway JW (2005) The mammalian Mediator complex and its role in transcriptional regulation. *Trends Biochem Sci* **30:** 250–255

Cramer P, Bushnell DA, Kornberg RD (2001) Structural basis of transcription: RNA polymerase II at 2.8 angstrom resolution. *Science* **292:** 1863–1876

Darst SA, Kubalek EW, Edwards AM, Kornberg RD (1991) Two-dimensional and epitaxial crystallization of a mutant form of yeast RNA polymerase II. *J Mol Biol* **221:** 347–357

Eichner J, Chen HT, Warfield L, Hahn S (2010) Position of the general transcription factor TFIIF within the RNA polymerase II transcription preinitiation complex. *EMBO J* **29:** 706–716

Frank J, Radermacher M, Penczek P, Zhu Jun Li Y, Ladjadj M, Leith A (1996) SPIDER and WEB: processing and visualization of images in 3D electron microscopy and related fields. *J Struct Biol* **116:** 190–199

Fuda NJ, Ardehali MB, Lis JT (2009) Defining mechanisms that regulate RNA polymerase II transcription in vivo. *Nature* **461:** 186–192

Gnatt AL, Cramer P, Fu J, Bushnell DA, Kornberg RD (2001) Structural basis of transcription: an RNA polymerase II elongation complex at 3.3 Å resolution. *Science* **292:** 1876–1882

Gnatt AL, Fu J, Kornberg RD (1997) Formation and crystallization of yeast RNA polymerase II elongation complexes. *J Biol Chem* **272:** 30799–30805

Goddard TD, Huang CC, Ferrin TE (2007) Visualizing density maps with UCSF Chimera. *J Struct Biol* **157:** 281–287

Golas MM, Sander B, Will CL, Lührmann R, Stark H (2003) Molecular architecture of the multiprotein splicing factor SF3b. *Science* **300:** 980–984

Gu W, Reines D (1995) Identification of a decay in transcription potential that results in elongation factor dependence of RNA polymerase II. *J Biol Chem* **270:** 11238–11244

Hahn S (2004) Structure and mechanism of the RNA polymerase II transcription machinery. *Nat Struct Mol Biol* **11:** 394–403

Harpaz Y, Gerstein M, Chothia C (1994) Volume changes on protein folding. *Structure* **2:** 641–649

Henry NL, Campbell AM, Feaver WJ, Poon D, Weil PA, Kornberg RD (1994) TFIIF-TAF-RNA polymerase II connection. *Genes Dev* **8:** 2868–2878

Hu X, Malik S, Negroiu CC, Hubbard K, Velalar CN, Hampton B, Grosu D, Catalano J, Roeder RG, Gnatt A (2006) A Mediator-responsive form of metazoan RNA polymerase II. *Proc Natl Acad Sci USA* **103:** 9506–9511

Jishage M, Malik S, Wagner U, Uberheide B, Ishihama Y, Hu X, Chait BT, Gnatt A, Ren B, Roeder RG (2012) Transcriptional regulation by Pol II(G) involving mediator and competitive interactions of Gdown1 and TFIIF with Pol II. *Mol Cell* **45:** 51–63

Kettenberger H, Armache K-J, Cramer P (2003) Architecture of the RNA polymerase II-TFIIS complex and implications for mRNA cleavage. *Cell* **114:** 347–357

Kettenberger H, Armache K-J, Cramer P (2004) Complete RNA polymerase II elongation complex structure and its interactions with NTP and TFIIS. *Mol Cell* **16:** 955–965

Kimura M, Ishihama A (2004) Tfg3, a subunit of the general transcription factor TFIIF in *Schizosaccharomyces pombe*, functions under stress conditions. *Nucleic Acids Res* **32:** 6706–6715

Klein BJ, Bose D, Baker KJ, Yusoff ZM, Zhang X, Murakami KS (2011) RNA polymerase and transcription elongation factor Spt4/5 complex structure. *Proc Natl Acad Sci USA* **108:** 546–550

Kornberg RD (2005) Mediator and the mechanism of transcriptional activation. *Trends Biochem Sci* **30:** 235–239

Kornberg RD (2007) The molecular basis of eukaryotic transcription. *Proc Natl Acad Sci USA* **104:** 12955–12961

Kostek SA, Grob P, De Carlo S, Lipscomb JS, Garczarek F, Nogales E (2006) Molecular architecture and conformational flexibility of human RNA polymerase II. *Structure* **14:** 1691–1700

Kostrewa D, Zeller ME, Armache K-J, Seizl M, Leike K, Thomm M, Cramer P (2009) RNA polymerase II–TFIIB structure and mechanism of transcription initiation. *Nature* **462:** 323–330

Le TT, Zhang S, Hayashi N, Yasukawa M, Delgermaa L, Murakami S (2005) Mutational analysis of human RNA polymerase II subunit 5 (RPB5): the residues critical for interactions with TFIIF subunit RAP30 and hepatitis B virus X protein. *J Biochem* **138:** 215–224

Liu X, Bushnell DA, Wang D, Calero G, Kornberg RD (2010) Structure of an RNA polymerase II-TFIIB complex and the transcription initiation mechanism. *Science* **327:** 206–209

Ludtke SJ, Baldwin PR, Chiu W (1999) EMAN: semiautomated software for high-resolution single-particle reconstructions. *J Struct Biol* **128:** 82–97

Malik S, Roeder RG (2005) Dynamic regulation of RNAP II transcription by the mammalian Mediator complex. *Trends Biochem Sci* **30:** 256–263

Malik S, Roeder RG (2010) The metazoan Mediator co-activator complex as an integrative hub for transcriptional regulation. *Nat Rev Genet* **11:** 761–772

Murakami KS, Masuda S, Darst SA (2002) Structural basis of transcription initiation: RNA polymerase holoenzyme at 4 Å resolution. *Science* **296:** 1280–1284

Nikolov DB, Chen H, Halay ED, Usheva AA, Hisatake K, Lee DK, Roeder RG, Burley SK (1995) Crystal structure of a TFIIB-TBP-TATA-element ternary complex. *Nature* **377:** 119–128

Ohi M, Li Y, Cheng Y, Walz T (2004) Negative staining and image classification - powerful tools in modern electron microscopy. *Biol Proced Online* **6:** 23–34

Penczek PA, Grassucci RA, Frank J (1994) The ribosome at improved resolution: new techniques for merging and orientation refinement in 3D cryo-electron microscopy of biological particles. *Ultramicroscopy* **53:** 251–270

Penczek PA, Zhu J, Frank J (1996) A common-lines based method for determining orientations for $N > 3$ particle projections simultaneously. *Ultramicroscopy* **63:** 205–218

Price DH (2008) Poised polymerases: on your mark…get set…go! *Mol Cell* **30:** 7–10

Price DH, Sluder AE, Greenleaf AL (1989) Dynamic interaction between a Drosophila transcription factor and RNA polymerase II. *Mol Cell Biol* **9:** 1465–1475

Prilusky J, Felder CE, Zeev-Ben-Mordehai T, Rydberg EH, Man O, Beckmann JS, Silman I, Sussman JL (2005) FoldIndex: a simple tool to predict whether a given protein sequence is intrinsically unfolded. *Bioinformatics* **21:** 3435–3438

Roeder RG, Schwartz LB, Sklar VE (1976) Function, structure, and regulation of eukaryotic nuclear RNA polymerases. *Symp Soc Dev Biol* **34:** 29–52

Roginski RS, Mohan Raj BK, Birditt B, Rowen L (2004) The human GRINL1A gene defines a complex transcription unit, an unusual form of gene organization in eukaryotes. *Genomics* **84:** 265–276

Sorzano CO, Marabini R, Velázquez-Muriel J, Bilbao-Castro JR, Scheres SH, Carazo JM, Pascual-Montano A (2004) XMIPP: a new generation of an open-source image processing package for electron microscopy. *J Struct Biol* **148:** 194–204

Stoops JK, Momany C, Ernst SR, Oliver RM, Schroeter JP, Bretaudiere JP, Hackert ML (1991) Comparisons of the low-resolution structures of ornithine decarboxylase by electron microscopy and X-ray crystallography: the utility of methylamine tungstate stain and butvar support film in the study of macromolecules by transmission electron microscopy. *J Electron Microsc Tech* **18:** 157–166

Suh MH, Meyer PA, Gu M, Ye P, Zhang M, Kaplan CD, Lima CD, Fu J (2010) A dual interface determines the recognition of RNA polymerase II by RNA capping enzyme. *J Biol Chem* **285:** 34027–34038

Takagi Y, Calero G, Komori H, Brown JA, Ehrensberger AH, Hudmon A, Asturias F, Kornberg RD (2006) Head module control of mediator interactions. *Mol Cell* **23:** 355–364

Takagi Y, Kornberg RD (2006) Mediator as a general transcription factor. *J Biol Chem* **281:** 80–89

Tan S, Aso T, Conaway RC, Conaway JW (1994) Roles for both the RAP30 and RAP74 subunits of transcription factor IIF in transcription initiation and elongation by RNA polymerase II. *J Biol Chem* **269:** 25684–25691

Tennyson CN, Klamut HJ, Worton RG (1995) The human dystrophin gene requires 16 hours to be transcribed and is cotranscriptionally spliced. *Nat Genet* **2:** 184–190

Thompson NE, Aronson DB, Burgess RR (1990) Purification of eukaryotic RNA polymerase II by immunoaffinity chromatography. Elution of active enzyme with protein stabilizing agents from a polyol-responsive monoclonal antibody. *J Biol Chem* **265:** 7069–7077

Tischendorf GW, Zeichhardt H, Stoffler G (1974) Determination of the location of proteins L14, L17, L18, L19, L22, L23 on the surface of the 50S ribosomal subunit of *Escherichia coli* by immune electron microscopy. *Mol Gen Genet* **134:** 187–208

Vassylyev DG, Sekine S, Laptenko O, Lee J, Vassylyeva MN, Borukhov S, Yokoyama S (2002) Crystal structure of a bacterial RNA polymerase holoenzyme at 2.6 Å resolution. *Nature* **417:** 712–719

Wei W, Dorjsuren D, Lin Y, Qin W, Nomura T, Hayashi N, Murakami S (2001) Direct interaction between the subunit RAP30 of transcription factor IIF (TFIIF) and RNA polymerase subunit 5, which contributes to the association between TFIIF and RNA polymerase II. *J Biol Chem* **276:** 12266–12273

Wind M, Reines D (2000) Transcription elongation factor SII. *BioEssays* **4:** 327–336

Yamaguchi Y, Takagi T, Wada T, Yano K, Furuya A, Sugimoto S, Hasegawa J, Handa H (1999a) NELF, a multisubunit complex containing RD, cooperates with DSIF to repress RNA polymerase II elongation. *Cell* **97:** 41–51

Yamaguchi Y, Wada T, Watanabe D, Takagi T, Hasegawa J, Handa H (1999b) Structure and function of the human transcription elongation factor DSIF. *J Biol Chem* **274:** 8085–8092

Young RA (1991) RNA polymerase II. *Annu Rev Biochem* **60:** 689–715