Position of the general transcription factor TFIIF within the RNA polymerase II transcription preinitiation complex

Jesse Eichner¹, Hung-Ta Chen²,
Linda Warfield and Steven Hahn*

Division of Basic Sciences, Fred Hutchinson Cancer Research Center,
Seattle, WA, USA

The RNA polymerase (pol) II general transcription factor TFIIF functions at several steps in transcription initiation including preinitiation complex (PIC) formation and start site selection. We find that two structured TFIIF domains bind Pol II at separate locations far from the active site with the TFIIF dimerization domain on the Pol II lobe and the winged helix domain of the TFIIF small subunit Tfg2 above the Pol II protrusion where it may interact with upstream promoter DNA. Binding of the winged helix to the protrusion is PIC specific. Anchoring of these two structured TFIIF domains at separate sites locates an essential and unstructured region of Tfg2 near the Pol II active site cleft where it may interact with flexible regions of Pol II and the general factor TFIIB to promote initiation and start site selection. Consistent with this mechanism, mutations far from the enzyme active site, which alter the binding of either structured TFIIF domains to Pol II, have similar defects in transcription start site usage.

The EMBO Journal (2010) **29**, 706–716. doi:10.1038/emboj.2009.386; Published online 24 December 2009

Subject Categories: chromatin & transcription

Keywords: initiation; polymerase II; TFIIF; transcription

Introduction

The transcription preinitiation complex (PIC) is a key intermediate in transcription initiation by RNA polymerase II (Pol II) (Hahn, 2004; Thomas and Chiang, 2006). This complex contains promoter DNA, Pol II, and the general transcription factors TBP, TFIIA, TFIIB, TFIIF, TFIIE, TFIIH, and the elongation factor TFIIS (Guglielmi *et al.*, 2007; Kim *et al.*, 2007). The PIC is stabilized by many protein–protein and protein–DNA interactions among the general factors, Pol II, coactivator complexes, and promoter DNA (Hahn, 2004). Once formed, the PIC can transition to the open complex state, allowing Pol II to initiate RNA synthesis. After initia-

tion, Pol II escapes the promoter and transcription elongation is regulated by factors acting directly on Pol II, as well as an elaborate set of factors that couple elongation to chromatin modification and mRNA export (Sims *et al.*, 2004).

TFIIF functions in both transcription initiation and elongation and is one of three general factors, TFIIB, TFIIE, and TFIIF, that directly bind Pol II (Bushnell *et al.*, 1996). TFIIF facilitates Pol II incorporation into a PIC intermediate-containing promoter DNA, TBP, and TFIIB and is required for subsequent recruitment of TFIIE and TFIIH (Flores *et al.*, 1991; Maxon *et al.*, 1994). TFIIF has been proposed to make protein–protein interactions with other PIC components such as TFIIB, TFIIE, and the Pol II CTD phosphatase Fcp1 (Maxon *et al.*, 1994; Wang and Burton, 1995; Kamada *et al.*, 2003; Nguyen *et al.*, 2003; Chen *et al.*, 2007). TFIIF also has functions in open complex formation (Yan *et al.*, 1999), transcription start site selection (Ghazy *et al.*, 2004; Freire-Picos *et al.*, 2005; Chen *et al.*, 2007), promoter escape (Yan *et al.*, 1999), stabilization of an RNA–DNA hybrid within Pol II (Khapersky *et al.*, 2008), and, along with TFIIS, stimulates elongation by purified Pol II (Izban and Luse, 1992; Yan *et al.*, 1999; Zhang and Burton, 2004).

Mammalian and *Drosophila* TFIIF are composed of two subunits termed Rap74 and Rap30, originally discovered as Pol II-binding factors (Burton *et al.*, 1988) that correspond to the *Saccharomyces cerevisiae* TFIIF subunits Tfg1 and Tfg2 (Henry *et al.*, 1994). Tfg3/Taf14, a yeast-specific TFIIF subunit, is non-essential and is also a component of at least five other coactivator and chromatin remodelling complexes including TFIID and Swi/Snf (Henry *et al.*, 1994; Cairns *et al.*, 1996). Studies with the mammalian factor showed that TFIIF dimerizes using the N-terminus of both subunits and that these subunits intertwine to form a triple barrel fold with protruding loops and extensions (Gaiser *et al.*, 2000). The C-terminus of each TFIIF subunit contains a winged helix fold that is connected to the TFIIF dimerization domain by an unstructured region (Groft *et al.*, 1998; Kamada *et al.*, 2001).

An important unanswered problem concerns the mechanism by which TFIIF promotes transcription initiation. Two structural studies and one biochemical study have indicated different locations of TFIIF when bound to Pol II. EM experiments suggested that TFIIF bound near the Pol II clamp domain (termed ‘ β ’ pincer’ in bacterial Pol) and the active site cleft in a Pol II–TFIIF complex (Chung *et al.*, 2003). A preliminary crystal structure of a Tfg2-containing Pol II elongation complex also suggested that Tfg2 bound near the Pol II clamp (Kornberg, 2007). In contrast, biochemical studies showed that TFIIF in PICs bound near the Pol II domains termed lobe and protrusion (termed ‘ β pincer’ in bacterial Pol) located on the Rpb2 subunit directly above the Pol II cleft and far from the clamp (Chen *et al.*, 2007). Supporting this latter location, mutations in the Pol II lobe altered the affinity of Pol II–TFIIF binding and caused

*Corresponding author. Division of Basic Sciences, Fred Hutchinson Cancer Research Center, 1100 Fairview Ave N, Mailstop A1-162, PO Box 19024, Seattle, WA 98109, USA. Tel.: +1 206 667 5261; Fax: +1 206 667 6497; E-mail: shahn@fhcrc.org

¹Present address: Georgetown University School of Medicine, Washington, DC, USA

²Present address: Institute of Molecular Biology, Academia Sinica, Taiwan, ROC

Received: 13 October 2009; accepted: 1 December 2009; published online: 24 December 2009

upstream shifts in the transcription start site, the same phenotype observed with some mutations in the TFIIF dimerization domain. However, these studies raised the question of how TFIIF can influence transcription initiation and elongation from a position far removed from the enzyme active site.

To clarify both the location of TFIIF in the PIC and the mechanism of TFIIF in initiation, we have used a combination of biochemical probes, molecular modelling, and functional assays to locate the TFIIF dimerization domain, an essential unstructured TFIIF region, and the Tfg2 winged helix domain on Pol II. Our results also reveal a PIC-dependent localization of the Tfg2 winged helix near the upstream promoter DNA, and the positioning of an essential unstructured TFIIF linker near the Pol II cleft, and suggest a model for the function of TFIIF in initiation.

Results

Molecular modelling of yeast TFIIF domains

As no structure of any yeast TFIIF domain was available, we first generated molecular models of the yeast TFIIF dimerization domain and the Tfg1 and Tfg2 C-terminal winged helix domains to facilitate the positioning of biochemical probes on the surface of TFIIF. Previous protein sequence alignments between the N-terminal regions of the yeast and mammalian TFIIF subunits have been difficult because of low sequence similarity between these factors (Ghazy *et al*, 2004; Chen *et al*, 2007). For this reason, we used the sequence alignment tool HHPred that compares a profile-hidden Markov model of a multiple sequence alignment with a hidden Markov model derived from the protein structure database (Soding *et al*, 2005) (Supplementary Figure S1A and B). These results (summarized in Figure 1A and B) suggest that the region of

Tfg1 analogous to the human Rap74 dimerization region encompasses residues 92–416 with a yeast-specific insertion between residues 146 and 336. The dimerization region of Tfg2 is predicted to lie within residues 55–227 with another yeast-specific insert between residues 142–210. For this analysis and all work described below, the *S. mikatae* Tfg1 and *S. cerevisiae* Tfg2 genes were used, as *S. mikatae* Tfg1 completely complements Tfg1 function in *S. cerevisiae* and this DNA is much easier to manipulate in molecular cloning and mutagenesis (Chen *et al*, 2007; Yang *et al*, 2009). A similar analysis was used to generate alignments for the yeast TFIIF winged helix domains (Figure 1A and B). On the basis of these sequence alignments, structure models of the yeast TFIIF domains were generated using the program Modeller (Eswar *et al*, 2008) (see Materials and methods) (Supplementary Figure S2).

Identification of functionally important regions in TFIIF

To determine the optimal placement of biochemical probes to be used for positioning TFIIF in the PIC, we next identified functionally important regions in Tfg1 and Tfg2 using a genetic assay. For all our experiments, we focused on the two conserved TFIIF subunits, Tfg1 and Tfg2. A series of small internal deletions was generated in Tfg1 and Tfg2 and a plasmid shuffle assay was used to test *in vivo* function in growth assays on synthetic glucose plates at temperatures ranging from 18 to 36°C (summarized in Figure 1; Supplementary Table S1; Supplementary Figures S3 and S4).

The results from these functional assays were consistent with those derived from sequence analysis and structure modelling. Mutations that deleted segments from the conserved portion of the predicted dimerization domain

Figure 1 Summary of functionally important regions of conserved yeast TFIIF subunits. (**A**, **B**) The *S. mikatae* Tfg1 and *S. cerevisiae* Tfg2 subunits. Shown are the regions homologous to the human Rap74 and Rap30 dimerization domain, the yeast-specific dimerization domain insertions and the winged helix domains. Colours summarize the functionally important regions from Supplementary Table S1. Black, essential for yeast growth; orange, temperature/cold sensitive or slow growth when deleted; white, not required for normal growth. The numbers indicate the amino acid residues in which FeBABE was linked and * indicates the positions of FeBABE attachments that generate cleavage of Rpb1 and/or Rpb2 in PICs.

were all lethal, whereas small deletions within the yeast-specific insertions in the Tfg1 and Tfg2 dimerization domains were viable. The Tfg2 winged helix and adjacent linker sequence were essential for growth, consistent with mutations in the winged helix domain of human TFIIF that impairs transcription initiation (Tan *et al*, 1995). In contrast, deletion of the Tfg1 winged helix domain gave only a mild temperature-sensitive phenotype. Deletion of the N-terminal charged region of the Tfg1 linker (residues 415–452) gave temperature-sensitive cells whereas small deletions removing the remaining segments of the linker (residues 453–594) had no phenotype. In contrast, removing a large segment of this region resulted in a mild slow growth phenotype, perhaps indicating a redundant function encoded within the linker or the requirement for a minimal length linker.

A genetic assay was used to test for regions of TFIIF involved in transcription elongation (Supplementary Table S1; Supplementary Figures S3 and S4). Strains containing viable Tfg1 or Tfg2 deletions were tested for growth on 6-azauracil (6-AU), an assay in which a wide variety of mutants defective in transcription elongation show slow growth (Exinger and Lacroute, 1992). Surprisingly, none of the Tfg1 mutations had growth defects on 6-AU compared with growth without 6-AU. In contrast, strains containing deletion of Tfg2 residues 249–273 within the linker region grew slowly on 6-AU, suggesting that this portion of the linker has an elongation-specific function.

Position of the TFIIF dimerization domain on Pol II in the PIC

On the basis of the TFIIF structure model and functional analysis, the hydroxyl radical generating probe FeBABE (Datwyler and Meares, 2000) was linked at specific positions within the TFIIF dimerization domain, linker, and winged helix domains (summarized in Figure 1; Supplementary Table S2). These TFIIF-FeBABE derivatives were created by inserting single cysteine residues in either Tfg1 or Tfg2 lacking endogenous cysteines. Recombinant Tfg1 + Tfg2 was purified as described earlier (Chen *et al*, 2007) and linked to FeBABE at these cysteine residues. The TFIIF-FeBABE derivatives were all active in complementing yeast nuclear extracts depleted for TFIIF function using an *in vitro* transcription assay (Supplementary Figure S5). The TFIIF-FeBABE derivatives were incorporated into PICs (containing a Flag epitope tag on either Rpb1 or Rpb2) at an immobilized yeast *HIS4* promoter by supplementing yeast nuclear extracts with TFIIF-FeBABE (Chen *et al*, 2007). PICs were washed with transcription buffer and FeBABE was activated with H₂O₂ to produce hydroxyl radicals.

Protein cleavage was monitored on western blots, probed for the epitope tag located at either end of Rpb1 or Rpb2. For all the newly generated TFIIF derivatives reported here, hydroxyl radical cleavage was observed exclusively in Rpb2. Figure 2 shows representative cleavage data from the TFIIF-FeBABE derivatives. Rpb2 cleavage products that are reproducibly observed compared with a control reaction using a non-cysteine-containing TFIIF are marked with asterisks. Several other previously characterized Tfg1-FeBABE derivatives (Chen *et al*, 2007) (marked with asterisks in Supplementary Table S2) cleaved both Rpb1 and Rpb2, and these cleavage data are included in the following analysis. The position of protein cleavage in Rpb1 and Rpb2 was

calculated by comparing the size of these protein fragments with a ladder of *in vitro* translated Rpb1-Flag or Rpb2-Flag polypeptides as described earlier (Chen *et al*, 2007) (e.g. Supplementary Figure S6). Cleavage products from each TFIIF derivative are listed in Supplementary Table S2 and Supplementary Figures S3 and S4.

Figure 3A shows a summary of hydroxyl radical cleavage on Pol II derived from FeBABE linked to positions within the Tfg1- and Tfg2 dimerization domains. In this figure, 10 Rpb1 or Rpb2 residues on either side of the calculated cleavage sites are highlighted in blue (owing to the uncertainty in calculated cleavage site) and are shown on the surface of Pol II in the PIC structure model (Miller and Hahn, 2006; Kostreva *et al*, 2009). From this analysis, it is apparent that nearly all the cleavage is within the Pol II lobe domain with the remainder on the nearby Rpb1 jaw. Cleavage data from individual TFIIF-FeBABE derivatives were next analysed to manually dock the TFIIF dimerization domain on the Pol II surface. The deduced position of the TFIIF dimerization domain on the Pol II lobe is consistent with all the FeBABE cleavage data (Figure 3B; Tfg1, orange; Tfg2, red). Examples of data supporting this model are shown in Supplementary Figure S7. First, FeBABE linked to positions that are pointed away from Rpb2 on the dimerization domain model fail to cleave Pol II (Supplementary Figure S7A). Second, FeBABE linked to Tfg1 at one end of the dimerization domain (S359 and Y343) cleaves near the Pol II cleft (Supplementary Figure S7B; cleavage highlighted in pink). Third, FeBABE linked to the opposite end of the dimerization domain at Tfg2 S227 cleaves on the top surface of Rpb2 (Supplementary Figure S7C, cleavage highlighted in pink). In an earlier study, we found that the photoreactive non-natural amino acid BPA inserted at Rpb2 residues 278, 323, and 367 crosslinks to Tfg1 but not Tfg2 in the PIC (Chen *et al*, 2007). These results are in excellent agreement with the Pol II-TFIIF dimerization domain model (Supplementary Figure S8, blue residues).

Position of the Tfg2 winged helix domain in the PIC

We next examined TFIIF derivatives with FeBABE attached to the Tfg2 winged helix domain (Figure 1B; Supplementary Table S2B). In PICs, many of these derivatives cleaved Rpb2 within the protrusion domain, which lies above the Pol II cleft and is adjacent to the Pol II lobe (Figure 2B and C). No cleavage was observed in Rpb1. Figure 4A shows the positions of Pol II cleavage generated by the Tfg2 winged helix derivatives in which 10 residues on either side of the calculated cleavage sites are highlighted in blue. Cleavage was not observed on one surface of the protrusion domain corresponding to α -helices 2 and 3 at the N-terminus of Rpb2. From our method of analysis, cleavage within this N-terminal segment will generate products too small to observe on western blots using N-terminally tagged Rpb2 and too large to resolve from full-length Rpb2 when using C-terminally tagged Rpb2.

As was done for the TFIIF dimerization domain, cleavage data from individual TFIIF-FeBABE derivatives were analysed to locate a position and orientation of the winged helix domain that was consistent with the cleavage data. We found that positioning of the Tfg2 winged helix domain on top of the Pol II protrusion domain was consistent with our results (Figure 4B). In this model, FeBABE linked to Tfg2

Figure 2 TFIIF-FeBABE-dependent cleavage of Rpb2 in the PIC. PICs were formed at the *HIS4* promoter using nuclear extract supplemented with the indicated TFIIF-FeBABE derivatives, hydroxyl radical cleavage activated with H_2O_2 , and the resulting cleavage products visualized by western blot. *Indicates reproducibly observed cleavage products compared with a reaction using a non-cysteine-containing TFIIF. (A) Cleavage from FeBABE inserted in the TFIIF dimerization domain and Tfg1 linker regions. Left panel uses N-terminal Flag-tagged Rpb2. All other assays use C-terminal Flag-tagged Rpb2. (B, C) Cleavage from FeBABE inserted in the Tfg2 linker and winged helix regions.

positions that fail to generate Pol II cleavage are primarily pointed up and away from the surface of Rpb2 (Supplementary Figure S9A). Furthermore, FeBABE at residues on one face of the winged helix (M294, D301, R321) cleave between the lobe and protrusion domains (Supplementary Figure S9B). Finally, FeBABE at residues on the other face of the winged helix domain (D311, K348) generate cleavage products on the opposite side of the protrusion (Supplementary Figure S9C). The positioning of the winged helix domain is less precise than that of the dimerization domain because some of the protrusion is partially disordered and not visible in the Pol II structure, and some of the cleavage sites are predicted to be within these disordered regions. It is likely that some of this

disordered region becomes ordered when bound to Tfg2. The predicted position of the Tfg2 winged helix domain is in excellent agreement with the result that BPA inserted at Rpb2 residue Y57 crosslinks to Tfg2 (Supplementary Figure S8, green residue). In contrast, FeBABE positioned within the Tfg1 winged helix domain did not result in cleavage of Rpb1 or Rpb2.

Situated on the protrusion domain, the Tfg2 winged helix is within 20 Å of promoter DNA ~20 base pairs upstream of the TATA element. The winged helix motif is a commonly used DNA-binding domain, with helices 2 and 3 equivalent to helices 2 and 3 of the helix-turn-helix DNA-binding motif (Gajiwala and Burley, 2000). In the PIC model, Tfg2 helix 3 faces upstream promoter DNA and may interact with this

Figure 3 Model of the yeast TFIIF dimerization domain binding to Pol II in the PIC. (A) FeBABE cleavage data for all TFIIF-FeBABE derivatives linked to the TFIIF dimerization domain. The cleavage data are from Supplementary Table S2. Ten residues on either side of the calculated cleavage site in Rpb1 and Rpb2 are coloured blue and Pol II is shown in the PIC model (Kostrewa *et al*, 2009). TBP, green; TFIIB, yellow. (B) Model for the TFIIF dimerization domain binding to the Pol II lobe domain. On the basis of the cleavage data from individual TFIIF-FeBABE derivatives and manual docking, the indicated position of TFIIF can account for the observed Pol II cleavage patterns. Orange, Tfg1; Red, Tfg2.

DNA on modest bending of upstream DNA and/or movement of the flexible protrusion domain.

Position of the Tfg2 linker closely approaches the Pol cleft

The location of two structured TFIIF domains far from the Pol active site was unexpected, considering the essential role of TFIIF in transcription initiation. This either suggests that TFIIF promotes initiation by an indirect mechanism or that anchoring of TFIIF at two separate locations functions to position an unstructured region of TFIIF near the active site cleft. To test the latter model, FeBABE was positioned within the Tfg1 and Tfg2 unstructured linkers that connect the dimerization and winged helix domains (Figure 1A and B). These derivatives cleaved exclusively within Rpb2

Figure 4 Position of the Tfg2 winged helix domain in the PIC. (A) Combined cleavage data for all TFIIF-FeBABE derivatives linked to the Tfg2 winged helix domain. Ten residues on either side of the calculated cleavage sites in Rpb2 are coloured blue. One arrow points to α -helices 2 and 3 in the protrusion domain (Cramer *et al*, 2001), which cannot be assayed for cleavage because they are at the very N-terminus of Rpb2. The base pair highlighted in red and blue indicates the most upstream site of DNA melting in the human system (Miller and Hahn, 2006; Kostrewa *et al*, 2009). (B) Model for the position of the winged helix domain that best fits the cleavage data. α -helix 3 of the WH domain (the DNA recognition helix) is indicated.

(Supplementary Table S2; Supplementary Figures S3 and S4). Only two FeBABE positions within the Tfg1 linker (residues 480 and 541) cleaved Rpb2, and both of these Tfg1 derivatives cleave the Rpb2 protrusion domain underneath the predicted position of the Tfg2 winged helix (Figure 5A, purple residues). However, it is likely that the Tfg1 linker does not have a direct function in initiation as deletion of this linker region has no phenotype.

Similar experiments were conducted to map the position of the Tfg2 linker. FeBABE attached to Tfg2 residue 239 cleaved Rpb2 within the external 1 and 2 regions on the top of Rpb2 and far from the active site, consistent with our positioning of the TFIIF dimerization domain (Figure 5B, blue residues). Mutation of this Tfg2 region gives rise to a slow growth

Figure 5 Mapping the location of the Tfg1 and Tfg2 linkers on Pol II in the PIC. (A) FeBABE cleavage data from Supplementary Table S2A showing the positions of Rpb2 cleavage from FeBABE positioned at Tfg1 residues N480 and A541. Cleavage sites are highlighted in purple. (B) FeBABE cleavage data from Supplementary Table S2B showing the positions of Rpb2 cleavage from FeBABE positioned at Tfg2 Q239 (blue) and N274, A285 (violet). (C, D) PIC model with the combined deduced positions of the TFIIF dimerization domain, the Tfg2 winged helix domain, and the Tfg2 flexible linker represented as a dotted line.

phenotype. In contrast, insertion of FeBABE at Tfg2 residues 274 and 285 within the essential region of the linker cleaved Rpb2 underneath and on either side of the protrusion domain located at the edge of the Pol II cleft (Figure 5B, violet residues). On the basis of these results, the Tfg2 linker (residues 274–285) is the essential region of TFIIF closest to the Pol II active site. From this location, the linker may directly influence transcription initiation and start site selection by interaction with the Pol II protrusion and/or a general factor such as TFIIB that also resides in the cleft (Bushnell *et al*, 2004; Kostreva *et al*, 2009). Figure 5C and D shows the location of the Tfg2 linker deduced from these FeBABE cleavage results.

Comparison of TFIIF binding in the PIC and the Pol II–TFIIF complex

As mentioned above, two structural studies have suggested a different position for TFIIF binding on Pol II from that reported here (Chung *et al*, 2003; Kornberg, 2007). These studies used purified Pol II and TFIIF or Pol II and Tfg2, in contrast to our studies that mapped TFIIF binding in PICs. To test whether the position of TFIIF shifted in the purified Pol II–TFIIF complex compared with the PIC, we mixed purified Flag-tagged Pol II with purified TFIIF–FeBABE derivatives, activated hydroxyl radical production with H₂O₂, and mapped Rpb2 cleavage sites by western analysis.

FeBABE inserted within the Tfg1 dimerization domain produced an identical cleavage pattern in the PIC and in the purified Pol II–TFIIF complex (Figure 6A, lanes 1–8; Supplementary Table S3). The Rpb2 cleavage pattern showed

two additional cut sites compared with that in Figure 2, as the previous experiments used a Flag-tag at the N-terminus of Rpb2 and these experiments used Rpb2 with a C-terminal tag. However, these additional cleavage sites near the lobe domain are consistent with the above model for the positioning of the TFIIF dimerization domain on Pol II.

In contrast to these results, FeBABE positioned within the Tfg2 winged helix domain gave a different cleavage pattern compared with that observed in PICs (Figure 6B, lanes 9–16). These derivatives gave a series of four cleavage products, all of which were located in the Pol II lobe domain (Figure 6C). The faint western signals of Rpb2 observed in the PIC (Figure 6B, lanes 12, 14, 16) are comparable with the background signals observed when a non-cysteine TFIIF derivative is used in the cleavage reaction (Figure 6B, lane 10) and are in slightly different positions from cleavages observed in the purified Pol II–TFIIF complex (Figure 6B, lanes 11, 13, 15). From these results, we conclude that the position of the TFIIF dimerization domain in the Pol II–TFIIF complex and in the PIC is very similar, whereas the Tfg2 winged helix domain is positioned over the Pol II protrusion only in the PIC.

Function of the Tfg2 winged helix domain

To examine the function of the Tfg2 winged helix in more detail, we first tested whether the deletion of this domain results in reduced affinity of TFIIF for Pol II. Wild type, recombinant TFIIF (Tfg1 + Tfg2) or recombinant TFIIF lacking the Tfg2 winged helix domain (Δ 291–359) was incubated with Pol II attached to beads using Rpb2–Flag, and the resulting complexes were washed and analysed by western

Figure 6 Comparison of TFIIF binding to purified Pol II and to Pol II in PICs. TFIIF-FeBABE derivatives were mixed with purified Pol II containing a C-terminal Flag-tagged Rpb2 subunit, and hydroxyl radical production was activated and cleavage visualized by western blot. This was done in parallel with the same TFIIF derivatives incorporated in PICs. (A) FeBABE linked to the TFIIF dimerization domain. (B) FeBABE linked to the Tfg2 winged helix domain. (C) The Pol II-TFIIF dimerization domain model from Figure 3 with cleavage observed in reactions containing purified Pol II and the Tfg2 winged helix-FeBABE derivatives. The cleaved region, highlighted in blue, shows that the winged helix binds near the Pol II lobe in the Pol II-TFIIF complex. Cleavage data are from Supplementary Table S3.

blot. TFIIF lacking the Tfg2 winged helix domain bound Pol II approximately two-fold less well compared with wild-type TFIIF, consistent with a direct interaction between this domain and Rpb2 (Figure 7A, lanes 7–10).

We next examined the role of the Tfg2 winged helix domain in transcription initiation. If part of its function is to position the Tfg2 linker near the Pol II cleft, then mutation of this domain should have near identical defects in start site usage as mutations in the TFIIF dimerization domain. Nuclear extract was depleted of endogenous TFIIF and supplemented with either wild-type TFIIF or with TFIIF lacking the Tfg2 winged helix domain ($\Delta 291$ –359) (Figure 7B). *In vitro* transcription using these nuclear extracts showed that overall transcription was reduced approximately three-fold with the mutant TFIIF. Moreover, this reduction was accompanied by a major shift in transcription start site usage. *In vitro* transcription of *HIS4* normally initiates in two major clusters, with the more distal start site directing >70% of *HIS4* initiation in multi-round transcription. Strikingly, deletion of the Tfg2 winged helix domain results in equal usage of

both start sites. These results show that the Tfg2 winged helix, similar to the TFIIF dimerization domain, has an important function in transcription start site selection and productive initiation.

Mutations in the TFIIF dimerization domain, which alter the transcription start site

Mutations that shift the transcription start site to upstream positions and allow usage of a non-optimal sequence for initiation have been isolated in the TFIIF dimerization domain (Ghazy *et al*, 2004; Freire-Picos *et al*, 2005; Khaperskyy *et al*, 2008). To examine the molecular basis for these alterations, we determined the position of these mutations within the TFIIF–Pol II model (Figure 7C, green residues). Three of the four residues that shift the transcription start site when mutated are predicted to be buried within the dimerization domain (Tfg2 L59, Tfg1 W345, G358). This suggests that these residues do not lie at the interface of protein–protein contacts with either Pol II or another general factor, but are more likely to subtly alter the structure of the dimerization

Figure 7 Mutations in the TFIIF dimerization and winged helix domains lead to changes in transcription start site. **(A)** The Tfg2 winged helix contributes to Pol II–TFIIF binding. Pol II of 1 μ g attached to anti-Flag beads was mixed with the indicated amounts of purified TFIIF and Pol II-associated TFIIF recovered by IP was assayed by western blot. Variable amounts of purified TFIIF in lanes 1–5 were used to quantitate TFIIF-associated Pol II in lanes 7–10. The input levels of wild type (WT) and mutant TFIIF derivatives gave equivalent signals in a western blot assay (not shown). **(B)** Mutation of the Tfg2 winged helix alters start site usage. Nuclear extract depleted for TFIIF was supplemented with purified TFIIF and transcription at the *HIS4* promoter was visualized by primer extension. **(C)** Mutations in the TFIIF dimerization domain, which alter the transcription start site. The TFIIF dimerization domain-PIC model showing residues in green that when mutated shift the transcription start site upstream from the normal position (Ghazy *et al*, 2004; Freire-Picos *et al*, 2005). The Rpb9 subunit is shown in yellow and the two bound Zn atoms in Brown. **(D)** Primer extension analysis of *in vivo ADH1* mRNAs from selected Tfg1 and Tfg2 mutations.

domain. Several of these derivatives bind to Pol II with lower affinity (Ghazy *et al*, 2004), which is consistent with this hypothesis. In our PIC model, the TFIIF dimerization domain is situated adjacent to the Pol II subunit Rpb9 (Figure 7C) and we predict that Rpb9 and the TFIIF dimerization domain make direct protein–protein contacts that contribute to the affinity of TFIIF for Pol II. Finally, mutations in the Pol II lobe domain were found to alter the affinity of TFIIF for Pol II and to alter the transcription start site (Chen *et al*, 2007). Again, this agrees with the model that perturbing Pol II interactions with the TFIIF dimerization domain results in alteration of the transcription start site. The fourth Tfg1 residue that alters the transcription start site when mutated (E341) is predicted to be surface exposed and is located near the edge of the yeast-specific dimerization domain insert, mutation of which also influences start site selection (see below).

To probe whether the dimerization domain interacts with other general factors in the PIC, we inserted the site-specific photocrosslinker PEAS (Chen *et al*, 1994; Chen and Hahn, 2003) at Tfg2 residues R63 and H88, which are located within the dimerization domain and predicted to point up and away from Pol II (Supplementary Figure S6A). These derivatives showed no crosslinking in PICs aside from crosslinking to Tfg1 (not shown), suggesting that this surface of the dimerization domain does not make extensive contact with any other general factor in the PIC.

To test whether the previously isolated mutations in the dimerization domain were unique in their ability to alter the transcription start site, strains containing small internal deletions in Tfg1 or Tfg2 were assayed for transcription start site defects. Strains were grown in synthetic glucose media and RNA extracted and assayed for transcription start site changes at the *ADH1* gene, which is known to be sensitive to mutations in TFIIF (Ghazy *et al*, 2004). We found that deletion of two regions within the non-conserved insert of the Tfg1 dimerization region and the charged region of the Tfg1 linker produced a minor fraction of transcripts initiated from an upstream position (Figure 7D, lanes 2–4). No other Tfg1 mutations tested resulted in a start site shift. Strikingly, deletion of either half of the non-conserved insert within the Tfg2 dimerization region gave a strong shift in start site usage (Figure 7D, lanes 8–9). Our combined results show that alteration of the transcription start site can result from mutations that alter the Pol II binding of either structured TFIIF domain and are consistent with the model that these TFIIF domains function in part to anchor the Tfg2 linker region near the Pol II cleft.

Discussion

TFIIF has an essential function in multiple steps of transcription, but the mechanism of how TFIIF acts in PIC formation,

transcription start site selection, and initiation is not understood. In this work, we have defined and mapped the positions of functionally important TFIIF domains within the PIC. The positioning of these domains, combined with biochemical and genetic studies, suggests a model for TFIIF function in which two structured TFIIF domains are anchored to Pol II at separate locations far from the active site. The Tfg2 winged helix domain is in position to interact with promoter DNA upstream of the TATA and the binding of the TFIIF dimerization and winged helix domains positions a non-structured segment of Tfg2 near the Pol II active site cleft in which it may interact with Pol II and/or other general factors during initiation and start site selection.

Using TFIIF derivatives linked to the protein cleavage reagent FeBABE, we determined how the TFIIF dimerization domain binds to the surface of the Pol II lobe adjacent to the Pol II subunit Rpb9. Although this positioning differs from that predicted by two structural studies (Chung *et al*, 2003; Kornberg, 2007), our modelling of the dimerization domain in the PIC is strongly supported by genetic and biochemical results. First, our model is consistent with results obtained from insertion of the non-natural photoreactive residue BPA in the Pol II lobe, which crosslinks to Tfg1 in PICs (Chen *et al*, 2007). Second, mutations on the surface of the Pol II lobe alter both the binding of TFIIF to purified Pol II and transcription start site selection (Chen *et al*, 2007). Start site alterations are a phenotype associated with mutation of the TFIIF dimerization domain (Ghazy *et al*, 2004; Freire-Picos *et al*, 2005). Third, our model predicts that the TFIIF dimerization domain directly contacts Rpb9. Elimination of Rpb9 is known to alter the transcription start site (Hull *et al*, 1995; Sun *et al*, 1996), and Pol II lacking Rpb9 is defective in TFIIF binding as predicted by our model (Ziegler *et al*, 2003). Finally, BPA positioned on the Pol II clamp in PICs crosslinks to TFIIE (Chen *et al*, 2007) rather than TFIIF as predicted by the structure studies. Crosslinking of the Pol II-TFIIF complex and analysis by mass spectrometry have led to very similar conclusions by Rappsilber and colleagues (Chen *et al*, 2009).

Similar experiments show that the essential winged helix domain of Tfg2 is positioned over the Pol II protrusion domain. This flexible Pol II domain is located at the edge of the Pol II cleft directly over the enzyme active site and is partially disordered in the Pol II crystal structure. Unlike the TFIIF dimerization domain that binds the Pol II lobe in both the PIC and the Pol II-TFIIF complex, our results show that Tfg2 positioning on the protrusion domain is PIC-dependent. It is possible that positioning of the winged helix domain in the PIC is assisted by interaction with upstream promoter DNA. In contrast, the Tfg1 winged helix domain is not localized to Rpb1 or Rpb2. Consistent with this result, we found that deletion of the Tfg1 winged helix has a very mild phenotype.

The predicted position of the Tfg2 winged helix domain lies about 20 Å from the proposed path of upstream DNA in the PIC model (Kostrewa *et al*, 2009). However, only modest bending of upstream DNA and/or movement of the flexible protrusion domain would be required to allow interaction of the winged helix and upstream DNA. Helix 3 of the winged helix (the DNA recognition helix) faces upstream DNA in our model. Our modelling can explain previous results that biochemical probes in promoter DNA upstream and downstream from the TATA interact with TFIIF (Robert *et al*, 1998;

Kim *et al*, 2000; Miller and Hahn, 2006) and our new finding that the position of the winged helix is different in the PIC and the TFIIF-Pol II complex. Interaction of upstream promoter DNA with the Tfg2 winged helix may function to stabilize both the PIC and open complex states. This interaction may be particularly important at TATA-less promoters that presumably lack a TBP-induced DNA bend.

The finding that the structured TFIIF domains are located at separate sites on the surface of Pol II far from the active site is surprising as TFIIF has a central role in transcription initiation, start site selection, and stabilization of RNA-DNA hybrids. This suggests a model in which the structured TFIIF domains function to position another region of TFIIF near the Pol active site. To test this model, we mapped the location of the unstructured Tfg1 and Tfg2 linker regions that connect the dimerization domain to the winged helix domains. FeBABE attached to one segment of the Tfg1 linker (residues 480, 541) generated cleavage on the protrusion, but deletion of this Tfg1 linker segment had no effect in a genetic assay or in transcription start site selection. It therefore seems unlikely that this Tfg1-protrusion interaction influences initiation. Importantly, the closest approach of any essential TFIIF region to the Pol II active site is made by the Tfg2 linker residues Q239 and N274, in which FeBABE probes cleave along the top, bottom, and sides of the protrusion. This region of the Tfg2 linker is essential for viability and may influence initiation efficiency and start site selection by interaction with either nucleic acids, the Pol II protrusion, and/or with the TFIIB B-linker region located within the Pol II cleft (Bushnell *et al*, 2004; Chen and Hahn, 2004; Kostrewa *et al*, 2009).

Consistent with this model, mutations in either of the two structured TFIIF domains that anchor TFIIF to Pol II show similar defects in transcription start site selection with transcription initiating upstream of the normal location. Most previously isolated mutations in the TFIIF dimerization domain, which shift the transcription start site, are predicted by our model to be located in residues buried within the dimerization domain. This suggests that these mutations affect the structure of the dimerization domain and provide an explanation for their low affinity binding to Pol II. This is also consistent with additional mutations in Rpb9 and on the surface of the Pol II lobe, which affect the transcription start site and the interaction of TFIIF with Pol II. Similarly, deletion of the Tfg2 winged helix domain causes a major change in start site usage and decreases the efficiency of initiation *in vitro*.

Our combined results suggest two possible roles for TFIIF in promoting initiation and transcription start site selection. First, the Tfg2 winged helix is predicted to interact with upstream promoter DNA, stabilizing the PIC and open complex, and possibly contributing to structuring of the flexible protrusion domain that may indirectly alter initiation and start site selection. Second, the two TFIIF structured domains position the Tfg2 linker so that it may interact with flexible elements within the Pol II active sites that are involved in initiation and start site usage. A prime candidate for Tfg2 interaction is the TFIIB region that lies in the active site cleft and connects the TFIIB Zn ribbon and core domains. This region, originally termed the B-finger (Bushnell *et al*, 2004) and, more recently, the B-reader/B-linker (Kostrewa *et al*, 2009) is proposed to read the sequence upstream of the

transcription start site and to assist DNA strand opening during open complex formation. Interaction of Tfg2 with TFIIB could potentially position this TFIIB region in the optimal location for its role in initiation. Both of these mechanisms suggest that TFIIF–Pol II interaction leads to the ordering and positioning of flexible elements within Pol II and/or the general factors and explains how mutations in Pol II and TFIIF at locations far from the active site can influence initiation.

Materials and methods

Yeast strains, plasmids, and RNA analysis

TFG1 mutants were derivatives of plasmid pHCF404 (Chen *et al*, 2007) containing *S. mikatae TFG1* with a C-terminal 6 × Flag epitope tag and transcription driven by the *TFG1* promoter in the yeast single copy vector pRS315 (Sikorski and Hieter, 1989). *TFG2* mutants were derivatives of plasmid pSH829 containing *S. cerevisiae Tfg2* with a C-terminal 3 × Flag epitope tag and transcription driven by the *TFG2* promoter in pRS315. Internal deletions were constructed by *in vitro* mutagenesis and contained the protein sequence GSGSGS substituted for the deleted residues. Function of the internally deleted *TFG1* and *TFG2* genes was assayed in strain HCF401 (Chen *et al*, 2007) containing a chromosomal *TFG1* deletion and strain SHY736 containing a *TFG2* chromosomal deletion. Gene function was assayed by plasmid shuffle assay at 30°C and by growth of single colonies on glucose synthetic complete media. All assays were conducted in triplicate and were independently repeated. Expression levels of the mutant and wild-type epitope-tagged Tfg1 and Tfg2 proteins were assayed in a strain containing a wild-type non-tagged gene to allow expression levels of lethal mutations to be measured. All mutant proteins were expressed within approximately two-fold of wild-type levels. For tests of 6-AU sensitivity, strains containing a single copy of the *TFG1* or *TFG2* genes were transformed with a *URA3*-containing vector and strains were monitored for growth of single colonies on glucose synthetic complete plates containing 20, 50, and 200 µg/ml 6-AU at 30°C. Total RNA was extracted from strains grown at 30°C to A600 ~1 in synthetic glucose media and analysed by primer extension using the oligonucleotide: GTATTCCAAC TACCGTGGGATTCG (Ghazy *et al*, 2004).

Sequence alignment and structure modelling

HHpred (Soding *et al*, 2005) and Phyre (Kelley and Sternberg, 2009) were used for sequence and structure alignment of yeast and human TFIIF sequences. For Tfg1 alignment, residues 1–200 and 300–450 of *S. mikatae Tfg1* were separately aligned using HHpred. The sequence alignments of the TFIIF dimerization domain and winged helix domains were used with the program Modeller 9v5 (Esvar *et al*, 2008) using very thorough variable target function method and thorough MD optimization with simulated annealing repeated twice for each model. Thirty models were generated and examined using the programs ERRAT (Colovos and Yeates, 1993), verify3D (Luthy *et al*, 1992), and molprobit (Davis *et al*, 2007) and the top scoring models were used for analysis of protein cleavage results. Manual docking of TFIIF domains with Pol II in PIC models (Miller and Hahn, 2006; Kostrewa *et al*, 2009) was used to find positions of TFIIF consistent with the observed FeBABE cleavage results. Structure coordinates of the final PIC model containing TFIIF are available as Supplementary data and at: <http://labs.fhcr.org/hahn/>.

References

- Burton ZF, Killeen M, Sopta M, Ortolan LG, Greenblatt J (1988) RAP30/74: a general initiation factor that binds to RNA polymerase II. *Mol Cell Biol* **8**: 1602–1613
- Bushnell DA, Bamdad C, Kornberg RD (1996) A minimal set of RNA Pol II transcription protein interactions. *J Biol Chem* **271**: 20170–20174
- Bushnell DA, Westover KD, Davis RE, Kornberg RD (2004) Structural basis of transcription: an RNA polymerase II–TFIIF cocystal at 4.5 Å. *Science* **303**: 983–988

Recombinant TFIIF expression and linking to FeBABE

Tfg1 and Tfg2 derivatives lacking the endogenous cysteines or with the indicated cysteine substitutions were co expressed in *Escherichia coli* and purified as described (Chen *et al*, 2007). FeBABE was linked to the purified recombinant proteins and used in PIC formation and protein cleavage assays as described (Chen and Hahn, 2003, 2004; Chen *et al*, 2007). Cleavage sites were calculated using *in vitro* translated protein standards of Rpb1 or Rpb2 containing the relevant N-terminal or C-terminal Flag epitope tag (Chen and Hahn, 2003, 2004; Chen *et al*, 2007). For FeBABE cleavage assays using purified proteins, 210 ng purified Pol II (see below) was mixed with a four-fold molar excess of recombinant TFIIF and incubated in transcription buffer (100 mM potassium acetate, 20 mM HEPES, pH 7.6, 1 mM EDTA, 5 mM Mg acetate) for 30 min at 23°C. FeBABE cleavage was activated as described (Chen *et al*, 2007) and the cleavage reaction was quenched after 2 min at 23°C. Reactions were TCA precipitated and cleavage products visualized by western blot.

TFIIF–Pol II-binding assay

Pol II was purified from a strain derived from SHY815 containing an N-terminal HIS6-tagged RPB3 integrated in the chromosome and a C-terminal 3 × Flag-tagged Rpb2. Pol II was purified from whole cell extract using Ni Sepharose and Source Q chromatography (GE Healthcare). 1 µg of Pol II was bound to 20 µl anti-Flag Sepharose (Sigma) in transcription buffer containing 1 mg/ml BSA and 0.5% NP-40. Immobilized Pol II was mixed with 100 or 250 ng purified TFIIF and incubated in the above buffer for 1 h at 4°C. Beads were washed once in the above buffer and three times in transcription buffer with 550 mM potassium acetate (final concentration) with 0.5% NP-40. Beads were boiled in SDS–PAGE loading buffer and analysed by western blot.

In vitro transcription

In vitro transcription used yeast nuclear extracts and conditions as described (Ranish *et al*, 1999) except with the above transcription buffer lacking NP-40 and BSA. For TFIIF complementation studies, nuclear extract containing a C-terminal Flag-tagged Tfg1 was depleted by incubation of 300 µl extract with 75 µl packed anti-Flag beads for 1 h at 4°C. Beads were removed and the 1 h depletion was repeated with fresh anti-Flag beads to generate the final TFIIF-depleted extract.

Supplementary data

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

Acknowledgements

We thank James Fishburn for initial analysis of selected TFIIF deletions, Jeff Ranish for TFIIS antisera, Patrick Cramer, Juri Rappsilber, and Anass Jawhari for sharing unpublished results, members of the Hahn lab and Ted Young for valuable comments throughout the course of this work, and Beth Moorefield for comments on the paper. This work was supported by grant 5R01 GM053451 to SH.

Conflict of interest

The authors declare that they have no conflict of interest.

- Chen HT, Warfield L, Hahn S (2007) The positions of TFIIF and TFIIE in the RNA polymerase II transcription preinitiation complex. *Nat Struct Mol Biol* **14**: 696–703
- Chen Y, Ebright YW, Ebright RH (1994) Identification of the target of a transcription activator protein by protein-protein photocross-linking. *Science* **265**: 90–92
- Chen Z, Jawahri A, Fischer L, Buchen C, Tahir S, Kaminski 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 crosslinking and mass spectrometry. *EMBO J* **29**: 717–726
- Chung WH, Craighead JL, Chang WH, 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
- Colovos C, Yeates TO (1993) Verification of protein structures: patterns of nonbonded atomic interactions. *Protein Sci* **2**: 1511–1519
- Cramer P, Bushnell DA, Kornberg RD (2001) Structural basis of transcription: RNA polymerase II at 2.8 angstrom resolution. *Science* **292**: 1863–1876
- Datwyler SA, Meares CF (2000) Protein-protein interactions mapped by artificial proteases: where sigma factors bind to RNA polymerase. *Trends Biochem Sci* **25**: 408–414
- Davis IW, Leaver-Fay A, Chen VB, Block JN, Kapral GJ, Wang X, Murray LW, Arendall III WB, Snoeyink J, Richardson JS, Richardson DC (2007) MolProbity: all-atom contacts and structure validation for proteins and nucleic acids. *Nucleic Acids Res* **35** (Web Server issue): W375–W383
- Eswar N, Eramian D, Webb B, Shen MY, Sali A (2008) Protein structure modeling with MODELLER. *Methods Mol Biol* **426**: 145–159
- Exinger F, Lacroute F (1992) 6-Azauracil inhibition of GTP biosynthesis in *Saccharomyces cerevisiae*. *Curr Genet* **22**: 9–11
- Flores O, Lu H, Killeen M, Greenblatt J, Burton ZF, Reinberg D (1991) The small subunit of transcription factor IIF recruits RNA polymerase II into the preinitiation complex. *PNAS* **88**: 9999–10003
- Freire-Picos MA, Krishnamurthy S, Sun ZW, Hampsey M (2005) Evidence that the Tfg1/Tfg2 dimer interface of TFIIF lies near the active center of the RNA polymerase II initiation complex. *Nucleic Acids Res* **33**: 5045–5052
- Gaiser F, Tan S, Richmond TJ (2000) Novel dimerization fold of RAP30/RAP74 in human TFIIF at 1.7 Å resolution. *J Mol Biol* **302**: 1119–1127
- Gajiwala KS, Burley SK (2000) Winged helix proteins. *Curr Opin Struct Biol* **10**: 110–116
- Ghazy MA, Brodie SA, Ammerman ML, Ziegler LM, Ponticelli AS (2004) Amino acid substitutions in yeast TFIIF confer upstream shifts in transcription initiation and altered interaction with RNA polymerase II. *Mol Cell Biol* **24**: 10975–10985
- Groft CM, Uljon SN, Wang R, Werner MH (1998) Structural homology between the Rap30 DNA-binding domain and linker histone H5: implications for preinitiation complex assembly. *Proc Natl Acad Sci USA* **95**: 9117–9122
- Guglielmi B, Soutourina J, Esnault C, Werner M (2007) TFIIS elongation factor and Mediator act in conjunction during transcription initiation *in vivo*. *Proc Natl Acad Sci USA* **104**: 16062–16067
- Hahn S (2004) Structure and mechanism of the RNA polymerase II transcription machinery. *Nat Struct Mol Biol* **11**: 394–403
- Henry NL, Campbell AM, Feaver WJ, Poon D, Weil PA, Kornberg RD (1994) TFIIF-TAF-RNA polymerase II connection. *Genes Dev* **8**: 2868–2878
- Hull MW, McKune K, Woychik NA (1995) RNA polymerase II subunit RPB9 is required for accurate start site selection. *Genes Dev* **9**: 481–490
- Izban MG, Luse DS (1992) Factor-stimulated RNA polymerase II transcribes at physiological elongation rates on naked DNA but very poorly on chromatin templates. *J Biol Chem* **267**: 13647–13655
- Kamada K, De Angelis J, Roeder RG, Burley SK (2001) Crystal structure of the C-terminal domain of the RAP74 subunit of human transcription factor IIF. *Proc Natl Acad Sci USA* **98**: 3115–3120
- Kamada K, Roeder RG, Burley SK (2003) Molecular mechanism of recruitment of TFIIF- associating RNA polymerase C-terminal domain phosphatase (FCP1) by transcription factor IIF. *Proc Natl Acad Sci USA* **100**: 2296–2299
- Kelley LA, Sternberg MJ (2009) Protein structure prediction on the Web: a case study using the Phyre server. *Nat Protoc* **4**: 363–371
- Khaperskyy DA, Ammerman ML, Majovski RC, Ponticelli AS (2008) Functions of *Saccharomyces cerevisiae* TFIIF during transcription start site utilization. *Mol Cell Biol* **28**: 3757–3766
- Kim B, Nesvizhskii AI, Rani PG, Hahn S, Aebersold R, Ranish JA (2007) The transcription elongation factor TFIIS is a component of RNA polymerase II preinitiation complexes. *Proc Natl Acad Sci USA* **104**: 16068–16073
- Kim TK, Ebright RH, Reinberg D (2000) Mechanism of ATP-dependent promoter melting by transcription factor IIF. *Science* **288**: 1418–1422
- Kornberg RD (2007) The molecular basis of eukaryotic transcription. *Cell Death Differ* **14**: 1989–1997
- Kostrewa D, Zeller ME, Armache KJ, Seizl M, Leike K, Thomm M, Cramer P (2009) RNA polymerase II-TFIIF structure and mechanism of transcription initiation. *Nature* **462**: 323–330
- Luthy R, Bowie JU, Eisenberg D (1992) Assessment of protein models with three-dimensional profiles. *Nature* **356**: 83–85
- Maxon ME, Goodrich JA, Tjian R (1994) Transcription factor IIE binds preferentially to RNA polymerase IIa and requires TFIIF: a model for promoter clearance. *Genes Dev* **8**: 515–524
- Miller G, Hahn S (2006) A DNA-tethered cleavage probe reveals the path for promoter DNA in the yeast preinitiation complex. *Nat Struct Mol Biol* **13**: 603–610
- Nguyen BD, Abbott KL, Potempa K, Kobor MS, Archambault J, Greenblatt J, Legault P, Omichinski JG (2003) NMR structure of a complex containing the TFIIF subunit RAP74 and the RNA polymerase II carboxyl-terminal domain phosphatase FCP1. *Proc Natl Acad Sci USA* **100**: 5688–5693
- Ranish JA, Yudkovsky N, Hahn S (1999) Intermediates in formation and activity of the RNA polymerase II preinitiation complex: holoenzyme recruitment and a postrecruitment role for the TATA box and TFIIB. *Genes Dev* **13**: 49–63
- Robert F, Douziech M, Forget D, Egly JM, Greenblatt J, Burton ZF, Coulombe B (1998) Wrapping of promoter DNA around the RNA polymerase II initiation complex induced by TFIIF. *Mol Cell* **2**: 341–351
- Sikorski RS, Hieter P (1989) A system of shuttle vectors and yeast host strains designed for efficient manipulation of DNA in *Saccharomyces cerevisiae*. *Genetics* **122**: 19–27
- Sims III RJ, Belotserkovskaya R, Reinberg D (2004) Elongation by RNA polymerase II: the short and long of it. *Genes Dev* **18**: 2437–2468
- Soding J, Biegert A, Lupas AN (2005) The HHpred interactive server for protein homology detection and structure prediction. *Nucleic Acids Res* **33** (Web Server issue): W244–W248
- Sun ZW, Tessmer A, Hampsey M (1996) Functional interaction between TFIIB and the Rpb9 (Ssu73) subunit of RNA polymerase II in *Saccharomyces cerevisiae*. *Nucleic Acids Res* **24**: 2560–2566
- Tan S, Conaway RC, Conaway JW (1995) Dissection of transcription factor TFIIF functional domains required for initiation and elongation. *Proc Natl Acad Sci USA* **92**: 6042–6046
- Thomas MC, Chiang CM (2006) The general transcription machinery and general cofactors. *Crit Rev Biochem Mol Biol* **41**: 105–178
- Wang BQ, Burton ZF (1995) Functional domains of human RAP74 including a masked polymerase binding domain. *J Biol Chem* **270**: 27035–27044
- Yan Q, Moreland RJ, Weliky Conaway J, Conaway RC (1999) Dual roles for transcription factor IIF in promoter escape by RNA polymerase II. *J Biol Chem* **274**: 35668–35675
- Yang C, Khaperskyy DA, Hou M, Ponticelli AS (2009) Improved methods for expression and purification of *Saccharomyces cerevisiae* TFIIF and TFIIF; Identification of a functional *Escherichia coli* promoter and internal translation initiation within the N-terminal coding region of the TFIIF TFG1 subunit. *Protein Expr Purif* (e-pub ahead of print 7 October 2009)
- Zhang C, Burton ZF (2004) Transcription factors IIF and IIS and nucleoside triphosphate substrates as dynamic probes of the human RNA polymerase II mechanism. *J Mol Biol* **342**: 1085–1099
- Ziegler LM, Khaperskyy DA, Ammerman ML, Ponticelli AS (2003) Yeast RNA polymerase II lacking the Rpb9 subunit is impaired for interaction with transcription factor IIF. *J Biol Chem* **278**: 48950–48956