High-Resolution Protein−DNA Contacts for the Yeast RNA Polymerase II General Transcription Machinery†

Bo-Shiun Chen,‡ Subhrangsu S. Mandal, and Michael Hampsey*

Department of Biochemistry, Division of Nucleic Acids Enzymology, Robert Wood Johnson Medical School, Piscataway, New Jersey 08854

Received May 18, 2004; Revised Manuscript Received July 31, 2004

ABSTRACT: We used site-specific protein−DNA photo-cross-linking to define contact points between *Saccharomyces cerevisiae* RNA polymerase II (RNAP II) and the general transcription factors TBP, TFIIB, and TFIIF on promoter DNA. We present three key findings: (i) the overall pattern of cross-link sites is remarkably similar between the yeast and the previously described human system, even though transcription initiates downstream of the DNA−TBP−TFIIB−RNAP II−TFIIF complex in the *S. cerevisiae* system; (ii) the yeast Rpb7 subunit of RNAP II forms strong and reproducible cross-links to both strands of promoter DNA; and (iii) a TFIIB arginine-78 to cysteine replacement (R78C), which shifts start site selection downstream of normal, does not affect TFIIB−DNA cross-links prior to promoter melting but instead affects downstream TFIIF−DNA interactions. These results support the premise that the overall structure of the RNAP II preinitiation complex is similar in all eukaryotes and imply that yeast RNAP II is able to scan template DNA downstream of the preinitiation complex for acceptable start sites. The novel Rpb7−DNA contact sites imply that either promoter DNA does not follow a straight path from TATA to the initiation site or the topology of Rpb7 within the DNA−TBP−TFIIB−RNAP II−TFIIF complex is different from that defined in the 12-subunit RNAP II X-ray structure. We discuss the implications of these results for the mechanism of preinitiation complex assembly and promoter melting.

Transcription initiation by RNA polymerase II (RNAP II)[1] requires assembly of a preinitiation complex composed of promoter DNA, RNAP II, and the general transcription factors TBP, TFIIB, TFIIE, TFIIF, and TFIIH (*1*−*3*). PIC assembly is nucleated by TBP binding to the TATA sequence, which induces a sharp bend in the DNA template. TFIIB then binds TBP and DNA, both upstream and downstream of TATA, forming a structure that serves as a platform for binding RNAP II and TFIIF. PIC assembly is completed by the binding of TFIIE and TFIIH.

Following PIC assembly, promoter DNA is melted in a manner dependent upon ATP hydrolysis and TFIIH helicase activity (*3*). Melting has been proposed to occur prior to placement of single-stranded template DNA into the active center cleft (*4*, *5*). In the presence of NTP substrates, the open complex is then able to initiate transcription. In metazoan systems, transcription typically starts at a single site located approximately ∼25 bp from the TATA box (*2*). Accordingly, RNAP II start sites appear to be defined simply by the architecture of the open complex: the position of the RNAP II catalytic center relative to the melted template defines the initiation site. Start site selection by RNAP II appears to be more complicated in *Saccharomyces cerevisiae*, where initiation often occurs at multiple sites within a window 40−120 bp from TATA (*2*). Nonetheless, hydroxyl radical footprinting of a yeast DNA−TBP−TFIIB complex demonstrated that ternary complex assembly occurs at the TATA box (*6*). Furthermore, promoter melting occurs at a fixed distance from TATA, regardless of the distance to start sites, suggesting that yeast RNAP II "scans" downstream template DNA for acceptable start sites (*7*). The mechanism by which RNAP II might scan promoter DNA for start sites is unknown.

TFIIB plays an important role in start site selection. Mutations in the yeast gene encoding TFIIB (*SUA7*) were identified on the basis of start site defects (*8*). Similar defects were conferred by *sua8* alleles of the *RPB1* gene (*9*), which encodes the largest subunit of RNAP II. Double *sua7 sua8* mutants are inviable, suggesting that interaction between TFIIB and Rpb1 underlies the mechanism of start site selection (*9*). Consistent with this premise, pairwise replacement of TFIIB and RNAP II from *S. cerevisiae* by their counterparts from *Schizosaccharomyces pombe* is both necessary and sufficient to shift start sites from the pattern characteristic of *S. cerevisiae* to that of *S. pombe* (*10*). TFIIB amino acid replacements that affect start site selection occur within a phylogenetically conserved region of the protein located just downstream of an N-terminal zinc ribbon (*11*−*17*). Replacements at similar positions in human TFIIB (*18*, *19*) and its archaeal counterpart, TFB (*20*), also affect start

† This research was supported by National Institutes of Health Grant GM39484 to M.H.

* To whom correspondence should be addressed. Phone: 732-235-5888. Fax: 732-235-5889. E-mail: michael.hampsey@umdnj.edu.

‡ Present address: Receptor Biology Unit, National Institute of Neurological Disorders and Stroke, National Institutes of Health, Bethesda, MD 20892.

[1] Abbreviations: RNAP II, RNA polymerase II; PIC, preinitiation complex; TBP, TATA binding protein; TFII, transcription factor of RNAP II; rTBP, recombinant TBP; rTFIIB, recombinant TFIIB; GTF, general transcription factor; NTP, nucleoside triphosphate; bp, base pair.

10.1021/bi048993r CCC: $27.50
Published on Web 09/14/2004

sites. The start site domain of TFIIB was recently shown to form a "B-finger" that projects into the active center of RNAP II (*21*) and is in proximity to the Rpb1 amino acid replacements encoded by the *sua8* alleles (*9*).

The *sua7-1* and *sua7-3* start site mutants encode, respectively, E62K and R78C amino acid replacements within the B-finger (*11*). These two replacements confer similar growth defects and identical effects on start site selection. E62 and R78 lie on opposite sides of the B-finger, presumably forming a stabilizing salt bridge (*21*). E62 and R78 replacements do not affect PIC assembly (*15*, *16*, *22*) but appear instead to affect the transition from abortive to productive initiation (*22*). Genetic analysis of the cold-sensitive growth phenotypes associated with E62K and R78C identified the largest subunit of TFIIF (Tfg1) and the Rpb2 and Rpb9 subunits of RNAP II as suppressors (*22*−*24*). The positions of the Rpb2 and Rpb9 suppressors within the X-ray structure of RNAP II suggest that the TFIIB defect is overcome by stabilizing RNAP II interaction with DNA downstream of the active center (*22*). We do not understand how TFIIF affects start site selection, either alone or in combination with TFIIB (*23*).

Crystal structures have been solved for a subset of RNAP II transcription complexes, including DNA−TBP−TFIIA (*25*), DNA−TBP−TFIIB (*26*, *27*), yeast RNAP II−TFIIB (*21*), the 10-subunit yeast RNAP II structure lacking Rpb4/Rpb7 (*28*), the complete 12-subunit RNAP II structure (*4*, *5*), and an RNAP II transcribing complex (*29*). Cryoelectron microscopy has also yielded a lower resolution image of yeast RNAP II−TFIIF (*30*). These and other structures provide extraordinary insight into the molecular basis of RNAP II transcription (*31*). Nonetheless, the molecular organization of RNAP II initiation complexes requires a comprehensive description of protein−protein and protein−DNA interactions. The contact sites between the human RNAP II machinery and promoter DNA were mapped by site-specific protein−DNA cross-linking (*32*−*35*). Similar experiments have mapped the structural organization of bacterial (*36*) and archaeal (*37*−*39*) initiation complexes. A description of site-specific DNA contacts for the yeast RNAP II machinery is lacking, despite the wealth of recent crystallographic images.

In this study we used site-specific protein−DNA cross-linking to map contact points on promoter DNA for a yeast DNA−TBP−TFIIB−RNAP II−TFIIF complex. Our results provide a comprehensive set of protein−DNA contact points that complement the recent X-ray and cryo-EM data for yeast RNAP II components. Moreover, our results define novel protein−DNA contacts that have implications for the mechanism of PIC assembly and promoter melting.

EXPERIMENTAL PROCEDURES

Yeast Strains. The yeast strains used in this study were derived from T16 (*MAT*α *cyc1-5000 cyc7-67 ura3-52 leu2-3,112 cyh2*) (*8*). YMH125 is isogenic to T16, except for the *sua7-3* allele encoding TFIIB R78C (*11*). Strains SHY407B (Rpb9-TAP) and SHY391A (Tfg1-TAP) were acquired from Steve Hahn (Fred Hutchinson Cancer Research Center).

Protein Purification. Recombinant yeast TBP was generously provided by Sungjoon Kim and Danny Reinberg (Robert Wood Johnson Medical School). The plasmid construct for expression of rTFIIB was a gift from Steve Buratowski (Harvard Medical School). The *sua7-3* allele encoding TFIIB R78C was amplified by the polymerase chain reaction, cloned into the pET-11d vector (Stratagene), and expressed in *Escherichia coli* strain BL21. The *sua7-3* sequence in pET-11d was confirmed by DNA sequence analysis. The normal and R78C forms of rTFIIB were purified as described previously (*15*). Briefly, BL21 transformants were grown in 2×YT medium supplemented with ampicillin (100 μg/mL) at 37 °C. When cell growth reached $A_{600nm} = 0.5$, rTFIIB expression was induced with IPTG (0.4 mM) for 4 h. Cell extracts were prepared and passed through an S-Sepharose column (Amersham Pharmacia Biotech); rTFIIB was eluted with 0.5 M potassium acetate. The full-length rTFIIB fraction also contained a truncation product that arises from internal translation initiation at methionine-119. To separate the full-length protein from the truncation product, the S-Sepharose eluate was applied to a hydroxyapatite column (Bio-Rad) equilibrated with 10 mM potassium phosphate (pH 7.6), 100 mM potassium acetate, 20% glycerol, 1 mM DTT, and 1 mM PMSF. rTFIIB was eluted with a gradient from 100 to 240 mM potassium phosphate in the same buffer. The final hydroxyapatite eluate was further purified by Mono S (HR 5/5) FPLC. Proteins were monitored by Coomassie Blue staining and confirmed by western blot using anti-TFIIB antibody.

RNAP II and TFIIF were purified from yeast strains SHY407B and SHY391A expressing TAP-tagged Rpb9 and Tfg1, respectively. Cells were grown at 30 °C in YPD medium, collected in late log phase, and lysed mechanically with glass beads in disruption buffer (0.2 M Tris−H_2SO_4, pH 7.9, 0.39 M ammonium sulfate, 10 mM magnesium sulfate, 20% glycerol, 1 mM EDTA). Proteins were purified as described previously (*40*). RNAP II lacking Rpb4 and Rpb7 was purified from SHY407B following resuspension of cell pellets in disruption buffer containing 50 mM $ZnCl_2$. RNAP II without the TAP-Rpb9 tag was purified by conventional and affinity (anti-CTD) chromatography; this protein was a generous gift from S. Kim and D. Reinberg.

Promoter DNA. Promoter DNA for protein−DNA cross-linking experiments was prepared as follows. Single-stranded DNA for primer extension was isolated using M13mp18-AdML-G-less and M13mp19-AdML-G-less constructs. These materials were made by PCR amplification of a 191 bp fragment of AdML-G-less promoter plasmid DNA [pG5MLT (A38T)] using synthetic primers that included flanking *Eco*RI and *Hin*dIII recognition sequences, digestion of amplified DNA, and ligation into the multiple cloning region of M13mp18 and M13mp19. pG5MLT (A38T) was constructed by site-directed mutagenesis of the standard pG5MLT promoter (*41*) to diminish initiation by yeast RNAP II at position +38. Phenyl azide-derivatized oligonucleotides were synthesized and purified essentially as described (*32*, *42*). These primers were then radioactively labeled, annealed to single-stranded template DNA, extended, and ligated according to standard procedures (*43*). Subsequent digestion with *Eco*RI and *Hin*dIII yielded DNA fragments that included the native AdML promoter (−50 to +10) followed by 117 bp of the G-less cassette.

Site-Specific Protein−DNA Photo-Cross-Linking. Reactions were performed as described previously (*32*) using chemically derivatized, radioactively labeled promoter DNA fragments and purified yeast transcription factors. Reaction

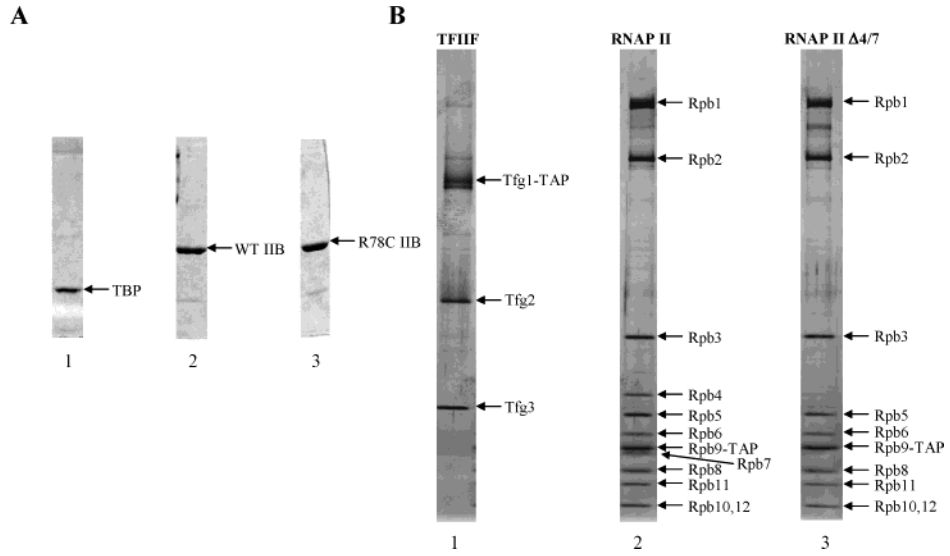

FIGURE 1: SDS−PAGE analysis of TBP, TFIIB, TFIIF, and RNA polymerase II. (A) SDS−10% polyacrylamide gel stained with Coomassie Blue R-250. Lanes: 1, rTBP; 2, wild-type rTFIIB; 3, R78C rTFIIB. (B) SDS−5−20% polyacrylamide gradient gel stained with silver nitrate. Lanes: 1, TAP-purified TFIIF using Tfg1-TAP; 2, TAP-purified RNAP II using Rpb9-TAP; 3, RNAP II lacking the Rpb4 and Rpb7 subunits.

mixtures (20 μL) contained 5.4 ng of DNA template (200 Bq/fmol), 5 ng of TBP, 15 ng of TFIIB (wild type or R78C), 15 ng of TFIIF, 60 ng of RNAP II, 20 mM Tris-HCl, pH 7.9, 20 mM potassium chloride, 40 mM potassium acetate, 5 mM magnesium chloride, 1 mM DTT, 25 mg/mL poly(dG-dC), and 100 μg/mL BSA. After 45 min at 28 °C, reaction mixtures were UV-irradiated for 2 min at 25 °C. Nuclease digestion was performed using 12 units of DNase I and 20 units of nuclease S1 with a protease inhibitor mix (Roche). Reactions were terminated by the addition of 0.2 volume of 10 M urea, 250 mM Tris-HCl, pH 6.8, 10% SDS, 25% β-mercaptoethanol, 50% glycerol, and 1% bromophenol blue, followed by heating for 2 min at 95 °C. Reaction products were analyzed by 5−16% gradient SDS−PAGE.

RESULTS

Site-Specific Protein−DNA Photo-Cross-Linking. We used site-specific, protein−DNA photo-cross-linking as described previously (*32*) to map precise positions of the general transcription factors within a DNA−TBP−TFIIB−RNAP II−TFIIF preinitiation complex. Briefly, promoter DNA fragments containing a phenyl azide photoactivatable cross-linking agent (probe) and an adjacent ^{32}P radiolabel were synthesized and incubated with purified yeast transcription factors to assemble a preinitiation complex (*32*, *42*). Following UV irradiation to cross-link proteins in physical proximity to the probe, complexes were extensively digested with nuclease, thereby converting cross-linked DNA to a ^{32}P tag. Labeled polypeptides were then identified by gel mobility assays and autoradiography. Our objectives were to answer the following questions: (i) Where do the yeast GTFs and RNAP II bind to promoter DNA relative to the transcription start site prior to promoter melting? (ii) How do the promoter contact points for these factors compare to the contact points for the human general transcription machinery? (iii) Does a mutation in yeast TFIIB that shifts start site selection in vivo affect protein−DNA contact points in vitro?

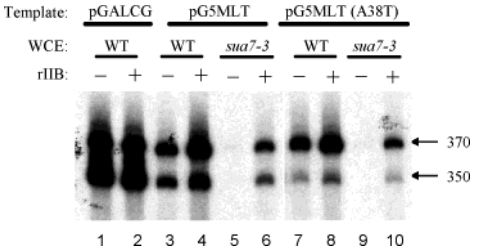

FIGURE 2: Yeast in vitro transcription system. Transcription from the pGALCG (*53*), pG5MLT (*41*), and pG5MLT (A38T) (Figure 3) templates initiates within the G-less cassette, yielding 350 and 370 nt products following RNase T1 digestion. Lanes: 1 and 2, wild-type (WT, strain T16) whole cell extract (WCE) using the pGALCG template with (+) or without (−) rTFIIB; 3−6, WT (T16) and *sua7-3* (YMH125) whole cell extracts using the pG5MLT template; 7−10, same as lanes 3−6, except using the pG5MLT (A38T) template, which includes the base pair substitution that diminishes initiation from the downstream (+38) start site that yields the 350 nt product.

Recombinant yeast TBP and TFIIB (rTBP and rTFIIB) were purified from *E. coli* as described previously (*15*). TFIIF and RNAP II were purified from yeast by tandem affinity purification using TAP-tagged derivatives of Tfg1 (TFIIF) and Rpb9 (RNAP II). Protein profiles are shown in Figure 1. The well-characterized adenovirus major late promoter (AdML) coupled to a G-less cassette (pG5MLT) was used as template DNA (*41*). The pG5MLT-G-less cassette includes a TATA box centered at position −30 and a single transcription start site at position +1. Although transcription initiates at +1 using a HeLa nuclear extract or purified human system, initiation occurs at two downstream sites located within the G-less cassette at positions +18 and +38 using the yeast system (Figures 2 and 3) (*22*, *44*). To minimize potential complications associated with dual start sites from the same promoter, we mutated the downstream start site from A → T, which significantly diminished initiation at position +38 relative to initiation at +18 (Figure 2, lane 1 vs lane 7). The *sua7-3* extract was defective for initiation from either promoter (lane 1; cf. lanes 5 and 9), and this defect was rescued by rTFIIB (lanes 6 and 10).

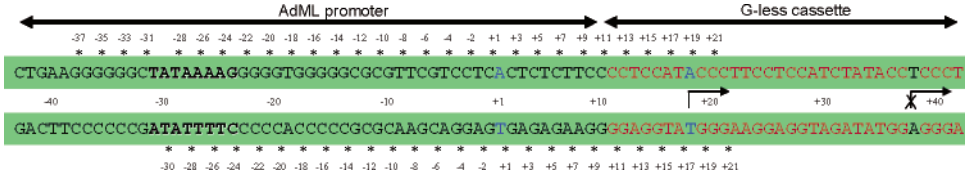

FIGURE 3: pG5MLT A38T template. Site-specific, phenyl azide photoactivatable DNA fragments that include an adjacent ^{32}P radiolabel span residues −37 to +21 on the nontemplate strand and residues −30 to +21 on the template strand. Transcription initiates at position +1 using a human transcription system but at positions +18 (arrow) and +38 using the yeast system; an A38T substitution was introduced to eliminate initiation at +38 (see Figure 2). Start sites at +1 and +18 are highlighted in blue; the altered start site at +38 is denoted in black. DNA sequences corresponding to the AdML promoter and G-less cassette are highlighted in black and red, respectively. Asterisks denote the positions of each phenyl azide site.

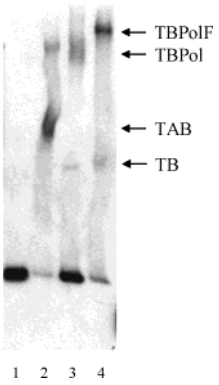

FIGURE 4: Assembly of derivatized template DNA into protein−DNA complexes. The electrophoretic mobility shift assay (EMSA) establishes that greater than 90% of the pG5MLT A38T template is assembled into a DNA−TBP−TFIIB−RNAP II−TFIIF complex. TFIIA is included here (lane 2) because the TB complex is not stable in the EMSA assay; TFIIA is dispensable for formation of the higher order TBPol and TBPolF complexes and was not included in any of the cross-linking experiments in this study (Figures 5 and 6). Key: TB, DNA−TBP−TFIIB; TAB, DNA−TBP−TFIIA−TFIIB; TBPol, DNA−TBP−TFIIB−RNAP II; TBPolF, DNA−TBP−TFIIB−RNAP II−TFIIF.

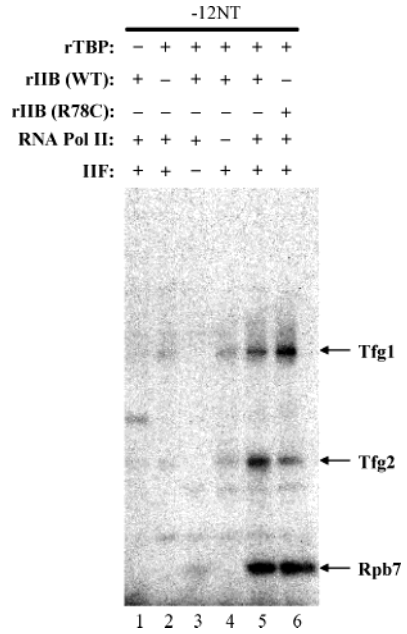

FIGURE 5: Protein−DNA photo-cross-linking is dependent upon the presence of TBP and TFIIB. Representative protein−DNA cross-linking data to the pG5MLT nontemplate strand position −12NT are depicted. Cross-linking to the Tfg1 and Tfg2 subunits of TFIIF and the Rpb7 subunit of RNAP II is dependent upon the presence of TBP and either the normal or R78C form of TFIIB.

We constructed 55 site-specific derivatized promoter DNA fragments, each including the photoactivatable cross-linking agent at a single, defined phosphate on the AdML-G-less A38T template. Site-specific derivatives span residues −37 to +21 on the nontemplate (top) strand and residues −30 to +21 on the template (bottom) strand (Figure 3). DNA−TBP−TFIIB−RNAP II−TFIIF complexes were assembled using either rTFIIB or the rTFIIB R78C derivative that shifts initiation downstream of normal. Randomly selected, derivatized DNA fragments were assayed by electrophoretic mobility shift assays to monitor assembly into higher order complexes. A representative assay is depicted in Figure 4. At the concentration of factors used in this study, essentially all DNA was incorporated into single homogeneous complexes. Control reactions established that the cross-link sites are position-dependent and dependent upon the presence of TBP and TFIIB, as shown in Figure 5 for nontemplate strand position −12. Accordingly, our results represent specific protein−DNA interactions.

The identities of cross-linked polypeptides were determined by denaturing polyacrylamide gel electrophoresis using purified GTFs and tagged derivatives of TFIIF and RNAP II subunits as markers. All experiments were done in the absence of ATP, TFIIE, and TFIIH, which are required for promoter melting; consequently, our data reflect contact points prior to open complex formation. Although the complexes characterized in this study are structurally homogeneous, they are not necessarily functionally homogeneous; i.e., in the presence of TFIIE, TFIIH, and ATP not all complexes would necessarily yield transcripts; therefore, we cannot be certain that specific protein−DNA cross-links defined here occur in initiation-competent complexes. Representative cross-linking data are shown in Figures 5 and 6A,B. Results are summarized schematically in Figure 7.

Cross-Link Sites between Yeast RNAP II and Promoter DNA. We found that three subunits of the 12-subunit RNAP II complex cross-linked with template DNA. Rpb1 and Rpb2 exhibited similar patterns, cross-linking with the nontemplate strand immediately upstream of the TATA box, beginning at position −37, and at several downstream sites between positions −12 and +18 (Figure 7). Similar, though less extensive, cross-links of Rpb1 and Rpb2 to the template strand were found. These results are comparable to those reported previously for the Rpb1 and Rpb2 subunits of human RNAP II using the AdML promoter (*32*, *35*). Given the dimensions of yeast RNAP II, estimated to be 140 × 136 × 100 Å (*45*), and the approximate distance of 190 Å between positions −37 and +18, it appears that RNAP II is able to bind sites on the template DNA that are separated

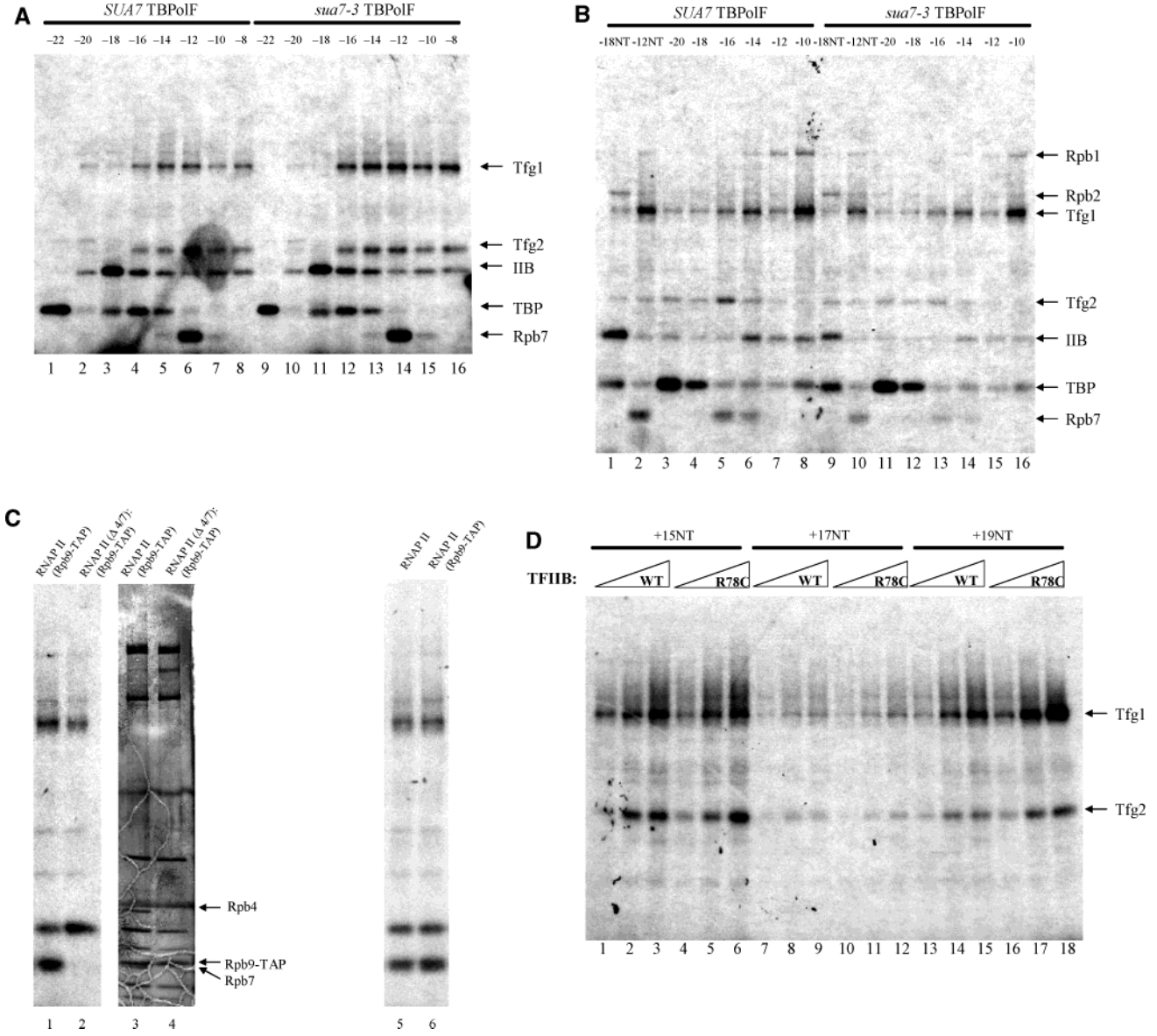

FIGURE 6: Protein−DNA photo-cross-linking results for the DNA−TBP−TFIIB−RNAP II-TFIIF promoter complex. (A) Representative protein−DNA cross-linking data to the pG5MLT nontemplate strand positions −22 to −8 using the normal or the R78C form of TFIIB (IIB). Tfg1 and Tfg2 denote the two largest subunits of TFIIF. (B) Representative cross-linking data to the template strand positions −20 to −10 and to nontemplate strand (NT) positions −18NT and −12NT using normal or R78C TFIIB. (C) Representative data showing cross-linking of Rpb7 to nontemplate strand position −12 using the complete 12-subunit RNAP II complex (lane 1) or 10-subunit complex lacking Rpb4 and Rpb7 (Δ4/7) (lane 2). Lanes 3 and 4 show silver nitrate stained TAP-purified RNAP II (lane 3) and RNAP II Δ4/7 (lane 4). Lanes 5 and 6 show cross-linking to nontemplate position −12 using either conventionally purified RNAP II (lane 5) or TAP-Rpb9 RNAP II (lane 6). TAP-purified RNAP II and RNAP II Δ4/7 are depicted in Figure 1. (D) Representative data showing enhanced TFIIF−DNA cross-linking associated with the R78C form of TFIIB to nontemplate (NT) strand positions +15 to +19. Photo-cross-linking experiments used three different concentrations (3.75, 7.5, or 15 ng) of normal TFIIB (WT) or its R78C derivative.

by distances greater than the dimensions of RNAP II. These results suggest that template DNA wraps around the yeast RNAP II complex, as suggested previously for human RNAP II (*32*, *34*).

The only other subunit of human RNAP II reported to cross-link with template DNA is Rpb5 (*32*). In our studies we saw no evidence for cross-linking of yeast Rpb5 with DNA. Instead, we found that the Rpb7 subunit cross-linked to DNA downstream of TATA at position −12 on the nontemplate strand and at positions −14 and −16 on the template strand (Figures 6A,B and 7). Three lines of evidence argue that the observed cross-links correspond to Rpb7. First, the apparent molecular mass of the cross-linking polypeptide is close to that of only Rpb7 or Rpb9-TAP (Figure 6C, lanes 1 and 3). Second, the cross-linking signal was lost when the experiment was repeated using the 10-subunit form of RNAP II that lacks the Rpb4 and Rpb7 subunits (Figure 6C, lanes 1 and 2), thereby defining dependence of the cross-link upon either Rpb4 and/or Rpb7. Third, there was no effect on the position of the cross-linking polypeptide when the experiment was repeated using RNAP II that lacks the Rpb9-TAP tag (Figure 6C, lanes 5 and 6), indicating that the cross-linking subunit is not Rpb9. We conclude that yeast Rpb7 directly contacts promoter DNA between the TATA box and start site.

GTF−Promoter DNA Cross-Links. As expected, TBP cross-linked with the nontemplate and template strands of promoter DNA at the TATA box (Figures 6A,B and 7).

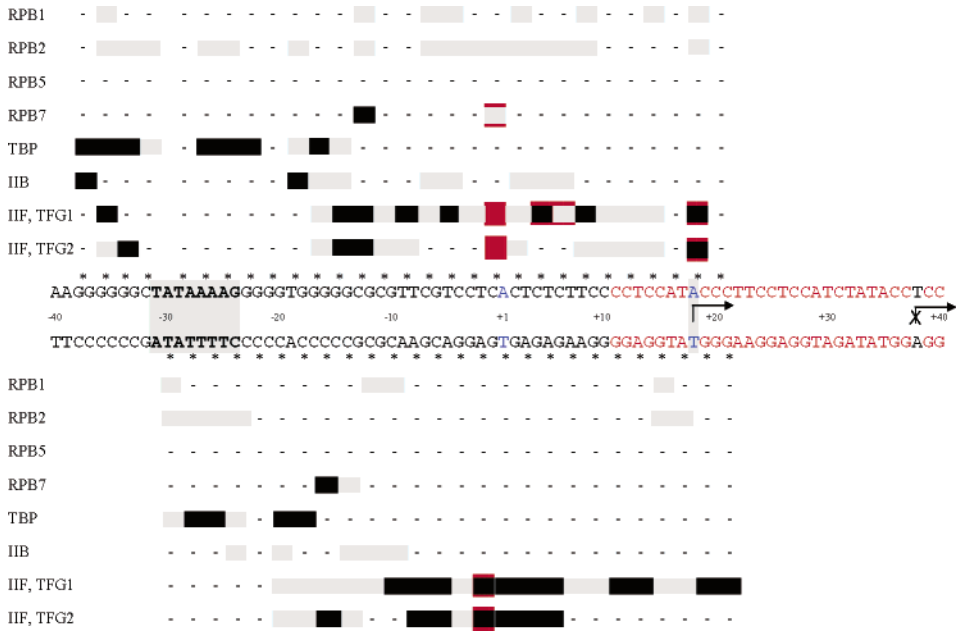

FIGURE 7: Summary of DNA−TBP−TFIIB−RNAP II−TFIIF photo-cross-linking results. The nontemplate (top) and template (bottom) strands of the AdML-G-less cassette are shown with AdML promoter DNA sequence denoted in black and the G-less cassette denoted in red. Phosphate positions analyzed in this study are indicated by asterisks. Sites exhibiting reproducible cross-links are indicated by solid bars; sites exhibiting less reproducible cross-links are indicated by shaded bars. The TATA element is highlighted by shading, and the transcription start site using the human system is denoted in blue at position +1. Transcription using the yeast system initiates within the G-less cassette (*44*) at position +18; normal initiation at +38 was diminished by the A38T substitution (Figure 2). Cross-linking results that are different using the R78C form of TFIIB are highlighted in red.

These interactions also extended immediately upstream (nontemplate strand) and downstream (both strands) of TATA. The cross-link sites for TFIIB were also similar to those of human TFIIB, contacting DNA upstream and downstream of TATA. These sites correspond to the specific contact points between promoter DNA and the C-terminal core domain of TFIIB observed in the DNA−TBP−TFIIB crystal structure (*26*) and are consistent with DNA−TBP−TFIIB footprinting data (*6*). Additional TFIIB cross-link sites were observed further downstream, including contact points flanking +1. These sites presumably represent contact points with the N-terminal domain of TFIIB that includes the zinc ribbon and B-finger.

TFIIF formed multiple DNA contacts, spanning promoter DNA from upstream of TATA to sites as far downstream as +18 on the nontemplate strand and +22 on the template strand (corresponding to the limits of the probes in this study). All of these contact sites are TBP-dependent. Extensive cross-linking between TFIIF and promoter DNA was reported previously for human TFIIF (*32*, *34*, *35*) and, in one case, included multiple contact points downstream of +1 (*34*). These results are generally consistent with the broad distribution of the Tfg1 and Tfg2 subunits of TFIIF across RNAP II revealed by cryoelectron microscopy of the yeast RNAP II−TFIIF complex (*30*). Moreover, much of the TFIIF density is associated with the Rpb4/Rpb7 subunits of RNAP II (*30*), which is likely to account for the common cross-links at positions −12 to −16 for TFIIF and Rpb7 (Figures 6A,B and 7).

Effect of the TFIIB R78C Replacement. The *sua7-3* mutant extract, containing the R78C form of TFIIB, is severely defective for transcription from the *CYC1* and AdML promoters (Figure 2) (*22*). Here we show that TFIIB R78C has no effect on the TFIIB−DNA interaction prior to promoter melting: cross-link sites to both DNA strands are identical for normal and R78C TFIIB (Figure 7). This result is consistent with our conclusion that R78C alters transcription after PIC assembly by affecting the competition between abortive initiation and promoter clearance (*22*). Similarly, contact with promoter DNA is not affected by an altered form of human TFIIB that affects start site selection (*19*). Interestingly, however, R78C enhanced the interaction of both subunits of TFIIF with the nontemplate strand at position +19 (Figure 6D, cf. lanes 16−18 with lanes 13−15; Figure 7); no effect of R78C on cross-linking of TFIIF to positions +15 or +17 was observed (Figure 6D, lanes 1−6 and 7−12). Quantitative effects of R78C on cross-linking of TFIIF also occurred at position +1 (Figure 7). Thus, R78C does not affect the direct interaction of TFIIB with DNA prior to promoter melting but instead affects DNA−TFIIF interactions. Whether this effect underlies the mechanism of altered start site selection associated with TFIIB and TFIIF mutations will be addressed by future cross-linking studies with promoter−open complexes.

DISCUSSION

The results of this study define a comprehensive set of protein−DNA contact points within the yeast DNA−TBP−TFIIB−RNAP II−TFIIF complex, thereby complementing the recent X-ray and cryo-EM structures of the yeast RNAP II transcription machinery. The key findings are that (i) the overall pattern of cross-link sites were remarkably similar between the yeast and human systems, supporting the premise that the architecture of the preinitiation complex is similar in all eukaryotes; (ii) in apparent contrast to human RNAP II, the yeast Rpb7 subunit forms strong and reproducible cross-links to both strands of promoter DNA at specific sites between the TATA box and initiator region; and (iii)

the TFIIB R78C replacement, which confers a marked effect on start site selection, has no effect on TFIIB−DNA cross-linking prior to promoter melting. Instead, R78C enhances specific TFIIF contacts with downstream DNA, an effect that underscores the expansive interaction of TFIIF with promoter DNA.

Our results have several important implications for the mechanism of transcription initiation. The recent X-ray structures of the complete 12-subunit yeast RNAP II complex revealed that the Rpb4/Rpb7 subcomplex, absent in the 10-subunit structures (*28*, *46*, *47*), is situated at the upstream face of RNAP II adjacent to the docking site for TFIIB (*4*, *5*). Rpb7 forms a wedge between the mobile "clamp" and flexible linker domain of Rpb1 such that the Rpb4/Rpb7 subcomplex appears to lock the clamp in a closed conformation. This structure would necessarily block entry of duplex DNA into the active site cleft, a surprising conclusion given that the Rpb4/Rpb7 subcomplex is required for transcription initiation but dispensable for elongation (*48*−*50*).

In the "clamp-open" form of RNAP II, double-stranded DNA could follow a straight path into the active site cleft (*28*). However, binding in the cleft in the clamp-closed state requires bending of the DNA by ∼90°, leading to the proposal that straight promoter DNA binds RNAP II, passing above the closed clamp, and that single-stranded template DNA descends into the active center of RNAP II upon melting and bending (*4*, *5*). However, a straight path for promoter DNA prior to melting is hard to reconcile with our observation that DNA interacts directly with Rpb7: the structure of the 12-subunit complex would require that promoter DNA bend significantly to enter the downstream cleft while making upstream contact with Rpb7.

We envision either of two scenarios to account for the robust cross-links between Rpb7 and promoter DNA. In one case, promoter DNA would not bind RNAP II as a straight duplex. Instead, the Rpb4/Rpb7 subcomplex might facilitate initiation by inducing a topological bend in the promoter as a prerequisite to promoter melting and entry of template DNA into the active center. This possibility is consistent with earlier cross-linking studies of human RNAP II, suggesting that promoter DNA wraps around RNAP II (*32*, *34*).

Alternatively, the topology of the Rpb4/7 subcomplex within the RNAP II holoenzyme does not reflect its topology in the DNA−TBP−TFIIB−RNAP II−TFIIF complex. Given the small interface between Rpb7 and core RNAP II (*4*, *5*), it seems plausible that the topology of the Rbp4/Rpb7 subcomplex might change during assembly of the preinitiation complex. If such a topological change involves dissociation of the Rpb7 "tip" that forms the wedge between the clamp and CTD linker, then the clamp could swing open to allow entry of the promoter DNA duplex into the active center cleft. This scenario might underlie the essential role of Rpb4/Rpb7 in transcription initiation (*49*).

Interaction between the Rpb7 subunit of RNAP II and DNA has also been detected by photoaffinity cross-linking during transcription elongation (*51*). Yeast Rpb7 cross-linked at low levels to positions −10 and −4, compared to robust cross-links at positions −12 to −16 in our study. [Rpb7 also cross-linked to position −29/−28; however, this cross-link was diminished in the presence of RNase H, implying that cross-linking was probably an artifact of the nontemplate strand failing to reanneal with the tailed template (*51*).] The Rpb7 subunit of RNAP II and its archaeal counterpart includes two nucleic acid binding motifs, leading to the proposal that Rpb7 binds the nascent transcript as it emerges from the RNA exit channel (*50*, *52*). If our Rpb7−DNA cross-link sites correspond to a binding site for DNA during PIC assembly and transcription initiation, then it is conceivable that the Rpb7−DNA interaction is replaced by Rpb7−RNA during the transition to elongation. This notion is in agreement with the weak interaction between Rpb7 and DNA in the elongation complex (*51*) and the dispensability of the Rpb4/Rpb7 complex for elongation in vitro (*49*).

The AdML-G-less cassette used in this study includes the same promoter used in earlier studies of human RNAP II (*32*, *35*). Accordingly, cross-linking results can be directly compared between the yeast and human systems. The most notable differences between the human and yeast cross-linking results are that human Rpb5, but not yeast Rpb5, binds DNA, whereas no cross-linking of human Rpb7 to DNA was reported (*32*, *35*). Perhaps these differences reflect differences in transcription mechanisms between the two systems. However, negative results from photochemical cross-linking experiments are not necessarily meaningful (*39*), and we caution against conclusions based on the absence of Rpb5 and Rpb7 cross-links in the respective studies.

The human and yeast studies are more notable for their similar overall cross-linking patterns. This is remarkable in that transcription initiation from the AdML-G-less cassette occurs at +1 with the human RNAP II machinery but at +18 and +38 using the yeast system (Figure 2) (*44*). This confirms that yeast RNAP II enters into a preinitiation complex at one site but initiates transcription further downstream. Although we have not yet determined cross-link sites for TFIIE and TFIIH, we anticipate that these factors will not affect the architecture of the DNA−RNAP II contacts, as TFIIE and TFIIH "slot" into the complex without altering the interactions of human RNAP II or the other general transcription factors with DNA (*32*, *35*). This scenario raises the question of whether promoter melting occurs at a defined distance from TATA or from the start site. Analysis of open complex formation at the *GAL1* and *GAL7* genes revealed that promoter melting begins at a fixed distance of ∼20 bp from the TATA elements of these two promoters, despite initiating 70 bp downstream at the *GAL1* promoter and 90 bp downstream at *GAL7* (*7*). These results led to the proposal that *S. cerevisiae* RNAP II scans downstream template DNA for acceptable start sites (*7*).

The structure of the RNAP II−TFIIB complex suggests a mechanism by which scanning might occur. The finger domain of TFIIB, which encompasses residues 55−88, protrudes into the RNA exit channel, extending beyond the active center (*21*). Kornberg and colleagues proposed that following strand separation the B-finger directly interacts with template DNA near the active center. Whether RNAP II initiates at the nucleotide positioned in the active center (+1) or proceeds downstream would depend on favorable interactions between TFIIB and the template sequence: in the absence of stable interactions, single-stranded template DNA would pass through the active center until more favorable interactions occur. TFIIB amino acid replacements that shift start site selection would do so either by directly affecting TFIIB−DNA interactions upstream of +1 or by

indirectly affecting interaction of the TFIIB "fingertip" with template DNA at the catalytic center (*21*).

Our observation that the TFIIB R78C replacement does not affect TFIIB−DNA cross-linking is consistent with this model. Prior to melting, promoter DNA would be unable to enter the active center of the clamp-closed form of RNAP II, and no effect of R78C on cross-linking of the B-finger to DNA would be expected. Instead, we found that R78C affected the interaction of TFIIF with DNA. Therefore, altered interaction of TFIIB within the RNA exit channel appears to affect Tfg1−DNA affinity prior to promoter melting and insertion of single-strand template DNA into the active center. We do not know how the TFIIB-mediated effect on TFIIF−DNA interaction correlates with start site changes, but it might be related to our isolation of altered forms of TFIIF as suppressors of the TFIIB E62K and R78C mutants (*22*, *23*). We are now interested in knowing how these and other suppressors of start site defects (*24*) affect protein−DNA cross-linking, both before and after promoter melting.

ACKNOWLEDGMENT

We are grateful to R. Ebright, D. Luse, D. Reinberg, and N. Woychik for valuable discussions and comments on the manuscript, to S. Kim and D. Reinberg for purified yeast TBP and RNAP II, to S. Hahn for yeast strains expressing TAP-tagged RNAP II and TFIIF, and to S. Buratowski for the rTFIIB expression plasmid.

REFERENCES

1. Orphanides, G., LaGrange, T., and Reinberg, D. (1996) The general initiation factors of RNA polymerase II, *Genes Dev. 10*, 2657−2683.
2. Hampsey, M. (1998) Molecular genetics of the RNA polymerase II general transcriptional machinery, *Microbiol. Mol. Biol. Rev. 62*, 465−503.
3. Hahn, S. (2004) Structure and mechanism of the RNA polymerase II transcription machinery, *Nat. Struct. Mol. Biol. 11*, 394−403.
4. Armache, K. J., Kettenberger, H., and Cramer, P. (2003) Architecture of initiation-competent 12-subunit RNA polymerase II, *Proc. Natl. Acad. Sci. U.S.A. 100*, 6964−6968.
5. Bushnell, D. A., and Kornberg, R. D. (2003) Complete, 12-subunit RNA polymerase II at 4.1-Å resolution: Implications for the initiation of transcription, *Proc. Natl. Acad. Sci. U.S.A. 100*, 6969−6973.
6. Lee, S., and Hahn, S. (1995) Model for binding of transcription factor TFIIB to the TBP-DNA complex, *Nature 376*, 609−612.
7. Giardina, C., and Lis, J. T. (1993) DNA melting on yeast RNA polymerase II promoters, *Science 261*, 759−762.
8. Pinto, I., Ware, D. E., and Hampsey, M. (1992) The yeast *SUA7* gene encodes a homolog of human transcription factor TFIIB and is required for normal start site selection *in vivo*, *Cell 68*, 977−988.
9. Berroteran, R. W., Ware, D. E., and Hampsey, M. (1994) The *sua8* suppressors of *Saccharomyces cerevisiae* encode replacements of conserved residues within the largest subunit of RNA polymerase II and affect transcription start site selection similarly to sua7 (TFIIB) mutations, *Mol. Cell. Biol. 14*, 226−237.
10. Li, Y., Flanagan, P. M., Tschochner, H., and Kornberg, R. D. (1994) RNA polymerase II initiation factor interactions and transcription start site selection, *Science 263*, 805−807.
11. Pinto, I., Wu, W.-H., Na, J. G., and Hampsey, M. (1994) Characterization of *sua7* mutations defines a domain of TFIIB involved in transcription start site selection in yeast, *J. Biol. Chem. 269*, 30569−30573.
12. Bangur, C. S., Pardee, T. S., and Ponticelli, A. S. (1997) Mutational analysis of the D1/E1 core helices and the conserved N-terminal region of yeast transcription factor IIB (TFIIB): identification of an N-terminal mutant that stabilizes TATA-binding protein-TFIIB-DNA complexes, *Mol. Cell. Biol. 17*, 6784−6793.
13. Pardee, T. S., Bangur, C. S., and Ponticelli, A. S. (1998) The N-terminal region of yeast TFIIB contains two adjacent functional domains involved in stable RNA polymerase II binding and transcription start site selection, *J. Biol. Chem. 273*, 17859−17864.
14. Faitar, S. L., Brodie, S. A., and Ponticelli, A. S. (2001) Promoter-specific shifts in transcription initiation conferred by yeast TFIIB mutations are determined by the sequence in the immediate vicinity of the start sites, *Mol. Cell. Biol. 21*, 4427−4440.
15. Cho, E. J., and Buratowski, S. (1999) Evidence that transcription factor IIB is required for a post-assembly step in transcription initiation, *J. Biol. Chem. 274*, 25807−25813.
16. Ranish, J. A., Yudkovsky, N., and 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.
17. Zhang, D. Y., Carson, D. J., and Ma, J. (2002) The role of TFIIB−RNA polymerase II interaction in start site selection in yeast cells, *Nucleic Acids Res. 30*, 3078−3085.
18. Hawkes, N. A., and Roberts, S. G. (1999) The role of human TFIIB in transcription start site selection in vitro and in vivo, *J. Biol. Chem. 274*, 14337−14343.
19. Fairley, J. A., Evans, R., Hawkes, N. A., and Roberts, S. G. (2002) Core promoter-dependent TFIIB conformation and a role for TFIIB conformation in transcription start site selection, *Mol. Cell. Biol. 22*, 6697−6705.
20. Bell, S. D., and Jackson, S. P. (2000) The role of transcription factor B in transcription initiation and promoter clearance in the archaeon *Sulfolobus acidocaldarius*, *J. Biol. Chem. 275*, 12934−12940.
21. Bushnell, D. A., Westover, K. D., Davis, R. E., and Kornberg, R. D. (2004) Structural basis of transcription: an RNA polymerase II-TFIIB cocrystal at 4.5 Angstroms, *Science 303*, 983−988.
22. Chen, B.-S., and Hampsey, M. (2004) Functional interaction between TFIIB and the Rpb2 subunit of RNA polymerase II: implications for the mechanism of start site selection, *Mol. Cell. Biol. 24*, 3983−3991.
23. Sun, Z. W., and Hampsey, M. (1995) Identification of the gene (SSU71/TFG1) encoding the largest subunit of transcription factor TFIIF as a suppressor of a TFIIB mutation in *Saccharomyces cerevisiae*, *Proc. Natl. Acad. Sci. U.S.A. 92*, 3127−3131.
24. Sun, Z. W., Tessmer, A., and 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.
25. Geiger, J. H., Hahn, S., Lee, S., and Sigler, P. B. (1996) Crystal structure of the yeast TFIIA/TBP/DNA complex, *Science 272*, 830−836.
26. Nikolov, D. B., Chen, H., Halay, E. D., Usheva, A. A., Hisatake, K., Lee, D. K., Roeder, R. G., and Burley, S. K. (1995) Crystal structure of a TFIIB-TBP-TATA-element ternary complex, *Nature 377*, 119−128.
27. Tsai, F. T., and Sigler, P. B. (2000) Structural basis of preinitiation complex assembly on human Pol II promoters, *EMBO J. 19*, 25−36.
28. Cramer, P., Bushnell, D. A., and Kornberg, R. D. (2001) Structural basis of transcription: RNA polymerase II at 2.8 angstrom resolution, *Science 292*, 1863−1876.
29. Westover, K. D., Bushnell, D. A., and Kornberg, R. D. (2004) Structural basis of transcription: separation of RNA from DNA by RNA polymerase II, *Science 303*, 1014−1016.
30. Chung, W. H., Craighead, J. L., Chang, W. H., Ezeokonkwo, C., Bareket-Samish, A., Kornberg, R. D., and Asturias, F. J. (2003) RNA polymerase II/TFIIF structure and conserved organization of the initiation complex, *Mol. Cell 12*, 1003−1013.
31. Woychik, N. A., and Hampsey, M. (2002) The RNA polymerase II machinery: structure illuminates function, *Cell 108*, 453−463.
32. Kim, T. K., Lagrange, T., Wang, Y. H., Griffith, J. D., Reinberg, D., and Ebright, R. H. (1997) Trajectory of DNA in the RNA polymerase II transcription preinitiation complex, *Proc. Natl. Acad. Sci. U.S.A. 94*, 12268−12273.
33. Forget, D., Robert, F., Grondin, G., Burton, Z. F., Greenblatt, J., and Coulombe, B. (1997) RAP74 induces promoter contacts by RNA polymerase II upstream and downstream of a DNA bend centered on the TATA box, *Proc. Natl. Acad. Sci. U.S.A. 94*, 7150−7155.
34. Robert, F., Douziech, M., Forget, D., Egly, J. M., Greenblatt, J., Burton, Z. F., and Coulombe, B. (1998) Wrapping of promoter

DNA around the RNA polymerase II initiation complex induced by TFIIF, *Mol. Cell 2*, 341−351.
35. Kim, T. K., Ebright, R. H., and Reinberg, D. (2000) Mechanism of ATP-dependent promoter melting by transcription factor IIH, *Science 288*, 1418−1422.
36. Naryshkin, N., Revyakin, A., Kim, Y., Mekler, V., and Ebright, R. H. (2000) Structural organization of the RNA polymerase-promoter open complex, *Cell 101*, 601−611.
37. Bartlett, M. S., Thomm, M., and Geiduschek, E. P. (2000) The orientation of DNA in an archaeal transcription initiation complex, *Nat. Struct. Biol. 7*, 782−785.
38. Bartlett, M. S., Thomm, M., and Geiduschek, E. P. (2004) Topography of the euryarchaeal transcription initiation complex, *J. Biol. Chem. 279*, 5894−5903.
39. Renfrow, M. B., Naryshkin, N., Lewis, L. M., Chen, H. T., Ebright, R. H., and Scott, R. A. (2004) Transcription factor B contacts promoter DNA near the transcription start site of the archaeal transcription initiation complex, *J. Biol. Chem. 279*, 2825−2831.
40. Puig, O., Caspary, F., Rigaut, G., Rutz, B., Bouveret, E., Bragado-Nilsson, E., Wilm, M., and Seraphin, B. (2001) The tandem affinity purification (tap) method: a general procedure of protein complex purification, *Methods 24*, 218−229.
41. Orphanides, G., LeRoy, G., Chang, C. H., Luse, D. S., and Reinberg, D. (1998) FACT, a factor that facilitates transcript elongation through nucleosomes, *Cell 92*, 105−116.
42. Lagrange, T., Kim, T. K., Orphanides, G., Ebright, Y. W., Ebright, R. H., and Reinberg, D. (1996) High-resolution mapping of nucleoprotein complexes by site-specific protein-DNA photocrosslinking: organization of the human TBP-TFIIA-TFIIB-DNA quaternary complex, *Proc. Natl. Acad. Sci. U.S.A. 93*, 10620−10625.
43. Sambrook, J., Fritsch, E. F., and Maniatis, T. (1989) *Molecular Cloning. A Laboratory Handbook*, 2nd ed., Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.
44. Lue, N. F., Flanagan, P. M., Sugimoto, K., and Kornberg, R. D. (1989) Initiation by yeast RNA polymerase II at the adenoviral major late promoter in vitro, *Science 246*, 661−664.
45. Darst, S. A., Edwards, A. M., Kubalek, E. W., and Kornberg, R. D. (1991) Three-dimensional structure of yeast RNA polymerase II at 16 Å resolution, *Cell 66*, 121−128.
46. Cramer, P., Bushnell, D. A., Fu, J., Gnatt, A. L., Maier-Davis, B., Thompson, N. E., Burgess, R. R., Edwards, A. M., David, P. R., and Kornberg, R. D. (2000) Architecture of RNA polymerase II and implications for the transcription mechanism, *Science 288*, 640−649.
47. Gnatt, A. L., Cramer, P., Fu, J., Bushnell, D. A., and Kornberg, R. D. (2001) Structural basis of transcription: an RNA polymerase II elongation complex at 3.3 Å resolution, *Science 292*, 1876−1882.
48. Pillai, B., Sampath, V., Sharma, N., and Sadhale, P. (2001) Rpb4, a non-essential subunit of core RNA polymerase II of *Saccharomyces cerevisiae* is important for activated transcription of a subset of genes, *J. Biol. Chem. 276*, 30641−30647.
49. Edwards, A. M., Kane, C. M., Young, R. A., and Kornberg, R. D. (1991) Two dissociable subunits of yeast RNA polymerase II stimulate the initiation of transcription at a promoter in vitro, *J. Biol. Chem. 266*, 71−75.
50. Orlicky, S. M., Tran, P. T., Sayre, M. H., and Edwards, A. M. (2001) Dissociable Rpb4-Rpb7 subassembly of RNA polymerase II binds to single-strand nucleic acid and mediates a post-recruitment step in transcription initiation, *J. Biol. Chem. 276*, 10097−10102.
51. Wooddell, C. I., and Burgess, R. R. (2000) Topology of yeast RNA polymerase II subunits in transcription elongation complexes studied by photoaffinity cross-linking, *Biochemistry 39*, 13405−13421.
52. Todone, F., Brick, P., Werner, F., Weinzierl, R. O., and Onesti, S. (2001) Structure of an archaeal homolog of the eukaryotic RNA polymerase II RPB4/RPB7 complex, *Mol. Cell 8*, 1137−1143.
53. Woontner, M., Wade, P. A., Bonner, J., and Jaehning, J. A. (1991) Transcriptional activation in an improved whole-cell extract from *Saccharomyces cerevisiae*, *Mol. Cell. Biol. 11*, 4555−4560.

BI048993R