

Evidence that RNA polymerase II and not TFIIB is responsible for the difference in transcription initiation patterns between *Saccharomyces cerevisiae* and *Schizosaccharomyces pombe*

Chen Yang and Alfred S. Ponticelli*

Department of Biochemistry, School of Medicine and Biomedical Sciences, State University of New York, Buffalo, NY 14214-3000, USA

Received February 15, 2012; Revised March 26, 2012; Accepted March 27, 2012

ABSTRACT

The basal eukaryotic transcription machinery for protein coding genes is highly conserved from unicellular yeast to higher eukaryotes. Whereas TATA-containing promoters in human cells usually contain a single transcription start site (TSS) located ~30bp downstream of the TATA element, transcription in the yeast *Schizosaccharomyces pombe* and *Saccharomyces cerevisiae* typically initiates at multiple sites within a window ranging from 30 to 70bp or 40 to 200bp downstream of a TATA element, respectively. By exchanging highly purified factors between reconstituted *S. pombe* and *S. cerevisiae* transcription systems, we confirmed previous observations that the dual exchange of RNA polymerase II (RNAPII) and transcription factor IIB (TFIIB) confer the distinct initiation patterns between these yeast species. Surprisingly, however, further genetic and biochemical assays of TFIIB chimeras revealed that TFIIB and the proposed B-finger/reader domain do not play a role in determining the distinct initiation patterns between *S. pombe* and *S. cerevisiae*, but rather, these patterns are solely due to differences in RNAPII. These results are discussed within the context of a proposed model for the mechanistic coupling of the efficiency of early phosphodiester bond formation during productive TSS utilization and intrinsic elongation proficiency.

INTRODUCTION

The basal transcription machinery for eukaryotic protein coding genes is composed of the core enzyme RNA polymerase II (RNAPII) and the auxiliary general transcription factors (GTFs) TATA-binding protein (TBP), transcription factor IIB (TFIIB), TFIIF, TFIIE and TFIIH (1,2). The high degree of conservation of these factors between unicellular yeast and mammalian cells, combined with the rapid growth and ease of genetic manipulation of yeast, have led to the widespread use of the budding yeast *Saccharomyces cerevisiae* (*S. cerevisiae*, Sc) and the fission yeast *Schizosaccharomyces pombe* (*S. pombe*, Sp) as model organisms for studying the molecular mechanisms of eukaryotic transcription (3,4). Consistent with the conservation of the primary structures of RNAPII and the GTFs, hydroxyl radical and photo-cross-linking experiments have demonstrated that the overall architecture of yeast and human pre-initiation complexes (PICs) is conserved (5–8). Importantly, however, the transcription initiation patterns between humans and yeast are intriguingly distinct. While TATA-containing promoters in human cells (*Homo sapiens*, Hs) usually contain a single transcription start site (TSS) located ~30bp downstream of the TATA element, transcription in *S. pombe* and *S. cerevisiae* typically initiates at multiple sites within a window ranging from 30 to 70bp or 40 to 200bp downstream of a TATA element, respectively (9–12).

Genetic and biochemical studies in our lab and those of others have identified numerous mutations in RNAPII (Rpb1, Rpb2 and Rpb9 subunits) (13–18), TFIIB (19–24)

*To whom correspondence should be addressed. Tel: +1 716 829 2473; Fax: +1 716 829 2725; Email: asp@buffalo.edu

Present address:

Chen Yang, Department of Pharmacology and Toxicology and Massey Cancer Center, Virginia Commonwealth University School of Medicine, Richmond, VA 23298, USA.

and TFIIF (both Tfg1 and Tfg2 subunits) (25), which alter TSS utilization in *S. cerevisiae*. However, almost all of these mutations are in conserved residues, providing little insight as to factor(s) determining the distinct species-specific initiation patterns. One approach to identify the determining factor(s) is to sequentially swap a single or multiple potential factor(s) between two species to determine whether TSS utilization is changed accordingly. Because neither activator stimulation nor the presence of chromatin structure affect the distinct patterns of TSS utilization (26), the determining factor(s) must be a component(s) of the basal transcription machinery. TBP has long been known to be interchangeable between the yeast and human systems without altering TSS utilization (27), thereby leaving the remaining GTFs and RNAPII as potential candidates. Prior to the identification of the genes encoding the *S. pombe* GTFs, Kornberg and colleagues utilized crude extract fractions of GTFs and *in vitro* swap assays in order to identify the determining factor(s) underlying the distinct initiation patterns between *S. pombe* and *S. cerevisiae*. The results demonstrated that (i) the combination of ScTFIIE and ScTFIIH supported ~14% the transcriptional level compared to their *S. pombe* counterparts without changing the initiation pattern activity in the *S. pombe* system, (ii) ScTFIIF supported ~15% the transcriptional level as SpTFIIF without changing the initiation pattern activity in the *S. pombe* system and (iii) neither TFIIB nor RNAPII alone was interchangeable between the two systems. Significantly, however, the dual exchange of *S. pombe* TFIIB and RNAPII with their *S. cerevisiae* counterparts conferred the *S. pombe* initiation pattern in the *S. cerevisiae* system, demonstrating that TFIIB and RNAPII were solely responsible for species-specific TSS utilization (9).

TFIIB is typically viewed as a bridge between TATA–TBP and RNAPII within the PIC (2). Genetic (22), biochemical (22,28–30) and biophysical (31–35) analyses have identified four major domains in TFIIB that include (i) the N-terminal Zn-ribbon domain that binds to the dock domain of RNAPII, (ii) the highly conserved B-finger domain that projects through the RNA exit channel and into the RNAPII active center, (iii) the B-linker domain, which has been proposed to interact with the non-template strand and (iv) the C-terminal B-core domain (Figure 2A). The core domain comprises two direct repeats, the first (Core1) which contacts the tip of RNAPII Rpb2 wall domain, the DNA immediately downstream of TATA and TBP, and the second (Core2) which interacts with DNA immediately upstream of TATA and TBP, thereby contributing to the unidirectional formation of the TATA–TBP–TFIIB complex.

Although the above described swap assays identified TFIIB and RNAPII as the sole determinants for species-specific initiation patterns, it has remained unclear whether TFIIB or RNAPII alone, or their combination, determines the distinct initiation patterns. Specifically, because there are multiple interactions between TFIIB and RNAPII (34,35), the role of TFIIB in species-specific TSS utilization might simply reflect the inability of TFIIB and RNAPII from different species to functionally interact. Alternatively, TFIIB may contain

intrinsic species-specific domain(s) that directly confer distinct patterns of TSS utilization. In regard to the latter possibility, studies of the RNAPII–TFIIB structure have demonstrated that the B-finger, in which multiple mutations confer downstream TSS shifts in *S. cerevisiae* (19,22,24), is positioned near the upstream non-template sequence of the TSS in the RNAPII active center (29,33–35). The B-finger was initially reported to form a hairpin structure with the tip being designated as the B-fingertip. Interestingly, the B-fingertip in ScTFIIB contains four acidic residues, whereas the B-fingertip in SpTFIIB and HsTFIIB contain three and two acidic residues, respectively (Figure 2). The two aspartate residues in the HsTFIIB B-fingertip have been shown to be important for transcription efficiency (36,37) and it has been speculated that the amount of negative charge in the B-fingertip may play a direct role in the distinct patterns of TSS utilization downstream of a TATA element (1,33). In a more recent study of the RNAPII–TFIIB structure, it was reported that the residues previously assigned to the B-finger hairpin form a unidirectional extension comprising an α -helix and a mobile loop (35). Nevertheless, this region was referred to as the B-reader to underscore its proposed function during TSS selection/utilization (Figure 2). Interestingly, however, an *in vivo* study in *S. pombe* demonstrated that TSS utilization on the *ADHI* promoter was shifted upstream by over-expressing human TFIIB or an *S. pombe* TFIIB chimera with the C-terminal Core2 domain replaced by that of human TFIIB, indicating that TFIIB, specifically the second repeat of the core domain, determines the species-specific TSS utilization pattern between fission yeast and human cells (11,38). In contrast, an *in vivo* study in *S. cerevisiae* demonstrated that no ScTFIIB chimeras containing HsTFIIB replacements could confer upstream TSS shifts (39). Thus, it has remained unclear whether the differences in species-specific initiation patterns between humans, *S. pombe* and *S. cerevisiae* are globally governed by the same factors and functional domains.

In this study, we sought to clarify the roles of TFIIB and RNAPII in the distinct patterns of TSS utilization between *S. cerevisiae* and *S. pombe*. We report here the establishment of reconstituted transcription assays with highly purified RNAPII and GTFs for *S. cerevisiae* and *S. pombe* and results from swap assays confirming that the combination of RNAPII and TFIIB confer the distinct initiation patterns between these species. Surprisingly, however, the results from genetic and additional biochemical assays of TFIIB chimeras revealed that RNAPII, but not TFIIB as previously proposed, is solely responsible for the difference in TSS utilization patterns between *S. cerevisiae* and *S. pombe*.

MATERIALS AND METHODS

Plasmids

PCR products for the construction of plasmids containing SpTfg1(L/S), SpTfg2, SpTfg3, SpTFIIS and TFIIB chimeras were amplified with Pfu DNA polymerase

(Stratagene) and ligated into pJET1.2/blunt (Fermentas) (see Supplementary Table S1 for all primer sequences and end cloning restriction sites). The protein coding sequences for SpTFIIB, SpTfg1(L/S), SpTfg2, SpTfg3 and SpTFIIS were amplified from an *S. pombe* cDNA library that was generated using a first-strand cDNA synthesis kit (Fermentas) and mRNA isolated from *S. pombe* strain GP2. All coding sequences were confirmed by DNA sequencing (Health Research Inc., Roswell Park Cancer Institute). Plasmids pQE-ScIIB, pQE-HsIIB and their mutants were constructed in our previous work (21,40).

Plasmids pQE-ScIIB, pQE-SpIIB and pQE-HsIIB were used as starting templates for the construction of TFIIB chimeras. DNA fragments encoding TFIIB chimeras were generated using overlapping PCR to connect two or three PCR products coding for the separate TFIIB domains. Two PCR fragments with a 40–50 bp overlap in the ends were joined using 10 cycles of PCR without primers followed by 30 cycles of amplification with a primer pair designed to amplify only the ligated product. TFIIB chimera constructs cloned into pJET1.2/blunt were then subcloned into the pQE32* bacterial expression vector (21), the centromere-containing yeast vector pRS314 (p314) under control of the native ScTFIIB (*SUA7*) promoter, and into the high copy yeast vector pHX under control of the *S. cerevisiae* *ADH1* promoter (41).

The bacterial expression vector for the production of recombinant ScTFIIF (pDt/g1*g2) has been described previously (42). Recombinant SpTFIIF expression vectors included pDt/SpTfg1, pDt/SpTfg1–Tfg3 and pET15b–SpTfg2. The coding regions for SpTfg1 and SpTfg3 were inserted into pCOLADuet-1 (Novagen) using NcoI/BamHI and NdeI/XhoI, respectively, whereas the coding region for SpTfg2 was inserted into pET15b (Novagen) using NdeI/BamHI.

Plasmids pAdMLP/G- and pADH1/G-, which were utilized as templates in the reconstituted transcription assays, have been described previously (43,44).

Yeast strains, media and phenotypic analyses

Saccharomyces cerevisiae TFIIB (*SUA7*) plasmid shuffle strains FP153 {*MATa ura3-52 trp1-Δ63 sua7-Δ1* [+p316/ScIIB (*URA3*)]} and FP231 {*MATa ura3-52 trp1-Δ63 leu2::PET56 sua7-Δ1* [+p316/ScIIB (*URA3*)]} were described previously (21). For construction of strains containing the ScRpb2-HsWall loop substitution, overlapping PCR was used to replace the first loop of the ScRpb2 wall domain (residues 864–871, KKYGMSIT) with that from human Rpb2 (SKKGFQDE) with the co-introduction of a PdmI restriction site, a 3XHA tag at the C-terminus of Rpb2, and a KanMX G418-resistance marker immediately downstream of the coding region (see Supplementary Figure S6 for details). One-step gene replacement using G418-resistance was used to introduce the variant ScRpb2 into strain FP153 harboring plasmids p314-ScIIB or p314-ScIIB-HsBH2 instead of p316-ScIIB. The ScRpb2-HsWall loop replacement was confirmed by PCR amplification of genomic DNA followed by PdmI digestion and by immunoblotting using a 1:2000 dilution of mouse anti-HA antibody

(Sigma) (Supplementary Figure S6). *Schizosaccharomyces pombe* strain GP2 is an alias for the wild-type 972 (h-) strain from Urs Leopold. Rich (YPD), 5-FOA, Casamino Acids (CAA) and minimal (SD) media were prepared as described previously (25,41).

To determine the ability of TFIIB chimeras to support yeast cell growth, plasmid shuffle complementation assays were performed (21). Chimeric derivatives of p314-TFIIB or pHX-TFIIB were analyzed in strains FP153 or FP231, respectively. Transformants of strains FP153 or FP231 were spread on CAA medium lacking uracil and tryptophan, or on SD medium lacking uracil and leucine, respectively. The plates were incubated for 2–3 days at 30°C to select for cells harboring both the test plasmid (*TRP1*-containing or *LEU2*-containing) and the endogenous plasmid (*URA3*) containing the wild-type *S. cerevisiae* TFIIB gene. Cultures of the Ura+ Trp+ or Ura+ Leu+ transformants were subsequently grown to an OD₆₀₀ of 1.0, 5 μl of 10-fold serial dilutions were spotted on both the original selecting medium and on 5-FOA solid medium and the plates were incubated at room temperature for 2–4 days. 5-FOA, which is toxic to cells containing the *URA3* gene, selects for cells that have spontaneously lost the *URA3* plasmid containing wild-type ScTFIIB. The appearance of FOA-resistant colonies therefore reflects the ability of a TFIIB variant to support cell viability in the absence of its wild-type counterpart. FOA-resistant colonies were purified once by streaking on CAA medium lacking tryptophan or leucine and incubating the plates at room temperature for 3–4 days. The growth properties of FOA-resistant strains (confirmed to be Trp+ Ura- or Leu+ Ura-) were analyzed by growing the strains in liquid YPD at room temperature to an OD₆₀₀ of 1.0, spotting 5 μl of 10-fold serial dilutions on YPD plates, and incubating the plates for 3 days at 16, 23, 30 or 37°C. Plasmids rescued from FOA-resistant strains were re-sequenced to confirm that no recombination events had occurred in the TFIIB coding region between the wild-type ScTFIIB shuffle vector and the test vector harboring the TFIIB chimera.

Purification of RNAPII

Without specific mention, all protein purification procedures after cell harvest were performed at 4°C. *Saccharomyces cerevisiae* RNAPII was purified from strain BJ2168 as described (18). *Schizosaccharomyces pombe* RNAPII was purified from wild-type strain GP2 as described (45) with the following modifications. Cells (201 cultures) were harvested, suspended in 3 × lysis buffer (300 mM Tris–acetate, pH 7.9, 300 mM potassium acetate, 60% glycerol, 3 mM EDTA) containing 3 × protease inhibitors (1 × = 1 mM PMSF, 1 mM benzamidine, 20 μM bestatin, 10 μM chymostatin, 3 μM pepstatin A and 1 μM leupeptin). Cells were mechanically disrupted using glass beads and the supernatant and washes were pooled. Potassium acetate (1/9 volume of 5 M) was added with stirring for 15 min, the mixture was centrifuged at 20 000 × g for 15 min, and the supernatant was centrifuged again at 100 000 × g for 60 min. The cleared supernatant was adjusted to 40% saturation with ammonium sulfate

by the slow addition of solid (0.23 g/ml), neutralized by occasional addition of 1M KOH (10 µl/g of ammonium sulfate), and stirred for an additional 20 min (about 60 min total). The ammonium sulfate precipitate was recovered by centrifugation at 30 000 × *g* for 20 min, suspended with 100 ml buffer C-0 (20 mM HEPES–KOH, pH 7.9, 20% glycerol, 2 mM DTT, 1 mM EDTA/EGTA and 1 mM PMSF/benzamidine), and additional buffer C-0 was added until the conductivity was equal to that of buffer C containing 0.1 M ammonium sulfate (C-0.1). The suspension was applied to a 100 ml DE52 (Whatman) column equilibrated in buffer C-0.1, protein eluted with a linear gradient of buffer C-0.1 to C-0.6, and the presence of the hypo-phosphorylated form of RNAPII monitored by immunoblotting using 8WG16 antibody (Covance). Fractions containing hypo-phosphorylated SpRNAPII were pooled and bound to 8WG16-Sepharose for 2 h, the resin was washed and protein was eluted with 40% glycerol, 20 mM HEPES–KOH, pH 7.9, 2 mM DTT, 0.5 M ammonium sulfate, 0.2 mM EDTA, 0.2 mM EGTA and 1 mM PMSF/benzamidine.

Purification of TFIIB

Recombinant TFIIB proteins were induced and purified as described (21) with the following modifications. Cells (500 ml culture at A_{600} of 0.6–0.8) were induced with 1 mM IPTG for 3 h, harvested and frozen at -80°C . Frozen cell pellets were thawed on ice and lysed by re-suspending with 9 ml of lysis buffer (50 mM Tris–HCl, pH 8.0, 500 mM NaCl, 10% glycerol, 2 mM DTT and protease inhibitors) containing 2 mg/ml lysozyme and adding 1 ml of 10% Triton X-100 (in lysis buffer). RNase A and DNase I were added to final concentrations of 100 and 20 µg/ml, respectively, and the lysate was incubated for 40 min and centrifuged for 20 min at 15 000 × *g*. The supernatant was batch loaded onto 1 ml of Ni-NTA resin (Qiagen), and incubated for 90 min with constant rocking. The protein bound Ni-NTA resin was subsequently packed into a column, washed with 10 ml of Buffer W (20 mM Tris–HCl, pH 8.0, 10% glycerol, 0.5 M NaCl, 0.1% Nonidet P-40, 10 mM imidazole, 2 mM DTT and 1 mM PMSF/benzamidine), and protein was eluted with 10 ml of Buffer W containing 250 mM imidazole, collecting 1 ml fractions. Peak fractions of TFIIB were pooled (2–3 ml final volume), diluted with 4 volumes of Buffer T-0 (10 mM Tris–acetate, pH 7.9, 10% glycerol, 2 mM DTT, 1 mM EDTA and 1 mM PMSF/benzamidine; the number after the hyphen indicates the molar concentration of potassium acetate), and bound with 0.6 ml SP Sepharose resin (GE Healthcare) for 90 min with constant shaking. The protein bound resin was subsequently packed in a column, washed and protein step-eluted with Buffer T-0.1, -0.2, -0.3, -0.4 and -0.5. Peak fractions of TFIIB (T-0.4 and T-0.5 steps) were pooled and adjusted to a final concentration of 5 pmol/µl in T-0.2.

Purification of TFIIF

Purification of recombinant *S. cerevisiae* TFIIF (Tfg1–Tfg2 complexes) was performed as described previously (42). *Schizosaccharomyces pombe* TFIIF was purified

from *Escherichia coli* strain RosettaTM (DE3; Novagen) co-transformed with pET15b-His6SpTfg2 and pDT-SpTfg1. Cultures (500 ml LB containing 100 µg/ml ampicillin, 50 µg/ml kanamycin and 34 µg/ml chloramphenicol) were grown, induced, harvested and lysed as for TFIIB purification. Peak fractions of SpTFIIF from the Ni-NTA column were pooled, diluted with Buffer T-0 and bound with 0.6 ml DE52 resin (Whatman) for 90 min with constant shaking. The protein bound DE52 resin was packed in a column, washed and protein step-eluted with Buffer T-0.1, -0.2, -0.3, -0.4 and -0.5. Peak fractions of SpTFIIF (T-0.4 step) were pooled and adjusted to a final concentration of 2 pmol/µl in T-0.2.

Purification of other GTFs

Recombinant *S. cerevisiae* TFIIE was purified as described (18). *Saccharomyces cerevisiae* TFIH was purified from yeast strains YT062 or YT063 as described (42). Recombinant hexahistidine-tagged ScTBP, ScTFIIS and SpTFIIS were purified from induced BLR (DE3) strains (Novagen) transformed with pET15b based expression vectors using the same scheme for TFIIB purification.

Primer extension analysis

For analysis of *in vivo* RNA, 50 ml *S. cerevisiae* cultures were grown at 30°C in YPD, CAA-Ura-Trp or SD-Ura-Leu media, and harvested at OD_{600} of 1.0–1.5. Total RNA was isolated and mRNA 5' ends were mapped using 30 µg of total RNA by primer extension using AMV Reverse Transcriptase (Promega) and an *ADH1*-specific primer as described (25,46) (See Supplementary Table S1 for primer sequence). *In vitro* transcription products were extracted twice with phenol/chloroform and ethanol-precipitated before being analyzed by primer extension with a G-less cassette specific primer (See Supplementary Table S1 for primer sequence).

Reconstituted transcription assays

Reconstituted transcription assays were performed using pAdMLP/G- or pADH1/G-plasmid templates as described (18) with some modifications. Transcription reactions (30 µl) contained 50 mM HEPES–KOH, pH 7.6, 5–8% glycerol, 80 mM potassium acetate, 10 mM magnesium acetate, 5 mM EGTA, 2.5 mM DTT, 40 U Ribonuclease inhibitor (Fermentas), 30 mM creatine phosphate and 1.5 U/ml creatine kinase. Both the *S. cerevisiae* and the *S. pombe* reconstituted transcription reactions were supported by ScTBP (3 pmol), ScTFIIE (0.25 pmol) and ScTFIIF (0.25 pmol). The additional components included 3 pmol TFIIB, 0.4 pmol RNAPII, 2 pmol TFIIF and 0.12 pmol plasmid template as indicated in the figures. PICs were assembled by incubation at ambient temperature for 20 min and the standard transcription assays initiated by the addition of 1 µl of an NTP mixture to a final concentration of 1 mM ATP, 500 µM UTP, 200 µM GTP and 200 µM CTP, in accordance with the reported physiological levels of 1070 µM ATP, 530 µM GTP, 160 µM GTP and 220 µM CTP in *S. cerevisiae* (47). Reactions were terminated

after 20 min incubation at ambient temperature with the addition of 300 μ l of stop buffer (100 mM sodium acetate, pH 5.2, 10 mM EDTA, 1% SDS, 75 μ g/ml carrier tRNA) and were treated with proteinase K (170 μ g/ml final) for 20 min at 37°C. The products were extracted twice with phenol/chloroform, ethanol precipitated and the transcripts were analyzed by primer extension with a G-less primer as described in (46).

Direct isotope incorporation transcription assays were performed in a similar fashion except that a 2 μ l NTP mixture was added to a final concentration of 1 mM ATP, 500 μ M UTP, 20 μ M 3-*O*-methyl-GTP, 2 μ M CTP and 0.1 μ M α -³²P-CTP. Reactions were incubated for 20 min at ambient temperature and then terminated, treated with RNase T1, and analyzed by denaturing electrophoresis as described (18).

RESULTS

Swap assays between *S. cerevisiae* and *S. pombe* reconstituted transcription systems confirm that the dual exchange of TFIIB and RNAPII confers species-specific TSS utilization

A previous biochemical study utilizing crude fractions of *S. pombe* and *S. cerevisiae* GTFs indicated that TFIIB and RNAPII were solely responsible for species-specific TSS utilization (9). To confirm and expand upon those results, we established reconstituted transcription systems with highly purified factors that accurately reproduced the distinct patterns of *S. cerevisiae* and *S. pombe* TSS utilization. For *S. cerevisiae*, our recent improvements for production and purification of recombinant ScTFIIF and TAP-tagged native TFIIH complex enabled a reconstituted system of highly purified factors comprising recombinant ScTBP, ScTFIIB, ScTFIIE, ScTFIIF and native ScTFIIH and ScRNAPII (42). For *S. pombe*, reconstituted transcription systems comprising purified recombinant ScTBP, SpTFIIB, SpTFIIE, SpTFIIF and native SpTFIIH and SpRNAPII have been previously described (45,48), although the production and purification of recombinant SpTFIIF was reported to be problematic. Like ScTFIIF, SpTFIIF is composed of two major subunits (designated SpTfg1, SpTfg2) and a third non-essential subunit SpTfg3 (49). SpTfg1 and SpTfg2 were reported to be insoluble when expressed individually in *E. coli* (48) or when co-expressed in insect cells (45), whereas the addition of SpTfg3 in insect cell expression yielded a low amount of soluble heterotrimeric SpTFIIF (45). In an attempt to improve the production of recombinant SpTFIIF, we discovered that SpTfg1 exists in two forms due to alternative splicing and that the abundance of the transcript encoding the longer form (SpTfg1L) is over 20 times greater than that for the smaller form (SpTfg1S) *in vivo* (see Supplementary Figure S1A). In contrast to the production of soluble recombinant *S. cerevisiae* TFIIF using the vector pCOLADuet-1 (pDt) in the *E. coli* host strain Rosetta (DE3), co-expression of SpTfg1(L or S) and SpTfg2 in either the absence or presence of pET15b-SpTfg3 resulted in insoluble SpTFIIF (data not shown). Interestingly, however,

fully soluble and functional SpTFIIF complexes were obtained by co-transformation of pDt-SpTfg1 (L or S) and pET15b-His6SpTfg2, with SpTfg3 being dispensable for solubility and function (Supplementary Figure S1B and S1C).

Utilizing standard transcription assays and primer extension analysis of the reaction products, we initially determined that both the *S. cerevisiae* and *S. pombe* reconstituted systems could be supported by highly purified recombinant ScTBP, recombinant ScTFIIE and TAP-purified native ScTFIIH (42), and that species-specific transcription initiation on both the adenovirus major late promoter (Figure 1A) and the *S. cerevisiae ADHI* promoter (Figure 1B) was conferred by the species-specific combination of recombinant TFIIB, recombinant TFIIF and native RNAPII (Figure 1A, lanes 3 and 5 and Figure 1B, lanes 3 and 5). Consistent with previous *in vitro* results using nuclear extracts (21,22), the R64E ScTFIIB mutant, which confers downstream shifts in start site utilization *in vivo*, supported barely detectable levels of basal transcription (Figure 1A, lane 2 and Figure 1B, lane 2) whereas the ScTFIIF Tfg1 E346A mutant, which confers upstream shifts in start site utilization *in vivo*, supported increased levels of basal transcription and conferred upstream shifts in start site usage (Figure 1A, lane 1 and Figure 1B, lane 1). Interestingly, the homologous R60E SpTFIIB also conferred reduced levels of transcription and a downstream shift in start site utilization in the *S. pombe* system, demonstrating that the underlying mechanism of TFIIB-conferred downstream shifts is conserved between these yeast species (Figure 1A, lane 6 and Figure 1B, lane 6). In addition, the homologous SpTFIIF Tfg1 E147A mutant supported slightly increased transcription levels, but no significant upstream shifts were detected (Figure 1A, lane 7 and Figure 1B, lane 7).

To improve the sensitivity and reduce the background observed with the primer extension analysis, we subsequently utilized G-less cassette direct incorporation assays that involved transcription in the absence of GTP and treatment of the reaction products with RNase T1 prior to gel analysis (see 'Materials and Methods' section). In the absence of GTP, transcription initiation on the *ADHI* promoter in the *S. cerevisiae* system occurred primarily at the adenine at position +98, with initiation also occurring at positions +83A and +91A (Figure 1D, lane 1). In contrast, initiation in the *S. pombe* system occurred primarily at position +81A, at two additional sites not utilized by the *S. cerevisiae* system (+84A, +86A), as well as at the sites utilized by *S. cerevisiae* (Figure 1D, lane 8). When SpTFIIF was replaced by ScTFIIF in the *S. pombe* system, diminished levels of transcripts were observed with no change in the transcript pattern, indicating that TFIIF is not the determining factor for the *S. pombe* TSS utilization pattern (Figure 1D, lane 6). Consistent with previous observations, SpTFIIF could not efficiently substitute for ScTFIIF in the *S. cerevisiae* system (Figure 1D, lane 3) and neither RNAPII (Figure 1D, lanes 4 and 5) nor TFIIB (Figure 1D, lanes 2 and 7) by themselves were

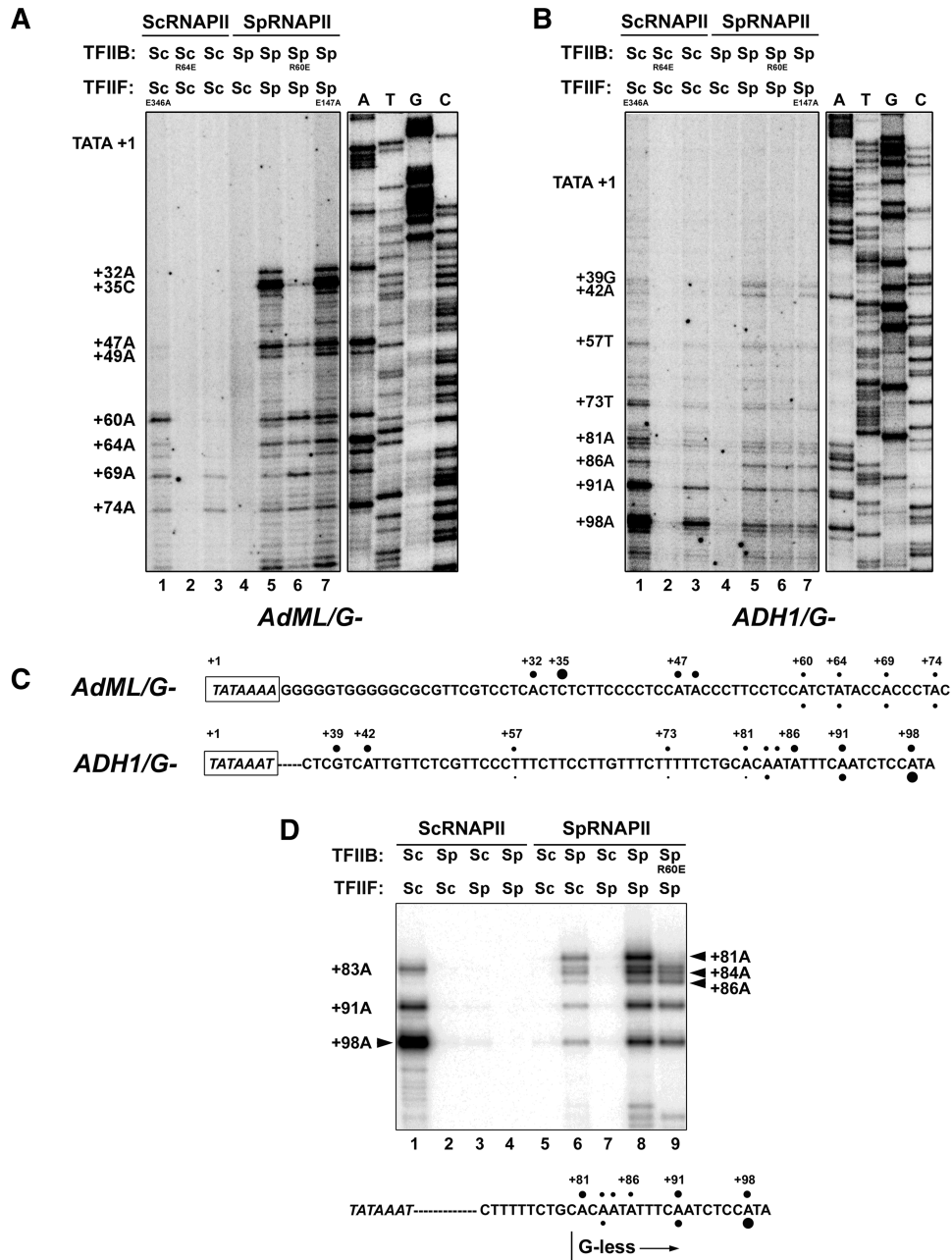


Figure 1. Swap assays between *S. cerevisiae* and *S. pombe* reconstituted transcription systems confirm that the dual exchange of TFIIB and RNAPII confers species-specific TSS utilization. (A and B) Standard transcription assays analyzed by primer extension. Pre-initiation complexes containing *S. cerevisiae* TBP, TFIIE, TFIIF and the indicated combinations of *S. cerevisiae* or *S. pombe* TFIIB, TFIIF and RNAPII were assembled on plasmid template pAdML/G- (A) or pADH1/G- (B) as described in ‘Materials and Methods’ section. Reactions were initiated by the addition of NTPs, terminated after 20 min incubation at ambient temperature, and the products were analyzed by primer extension utilizing a G-less cassette specific primer. The numbers flanking the panel indicate the positions of the minor and major TSSs, where +1 is defined as the first T in the TATA element. Shown in (C) is a depiction of the locations of the start sites in the *S. pombe* (above the sequence) or *S. cerevisiae* system (below the sequence), with the size of the circles indicative of the relative frequency of utilization. (D) Direct incorporation transcription assays. Pre-initiation complexes were assembled as in (B) and the reactions were initiated by the addition of GTP-lacking (G-) NTPs in the presence of α -³²P-CTP, terminated after 20 min incubation at ambient temperature, and the products treated with RNase T1 prior to denaturing polyacrylamide gel electrophoresis. The numbers flanking the panels indicate the positions of the TSSs under the GTP-lacking conditions, with the arrows highlighting the major characteristic sites for *S. cerevisiae* (left of panel, +98A) and for *S. pombe* (right of panel, +81A, +84A, +86A). Shown below the autoradiograph is a depiction of the locations of the start sites under GTP-lacking conditions in the *S. pombe* (above the sequence) or *S. cerevisiae* system (below the sequence), with the size of the circles indicative of the relative frequency of utilization.

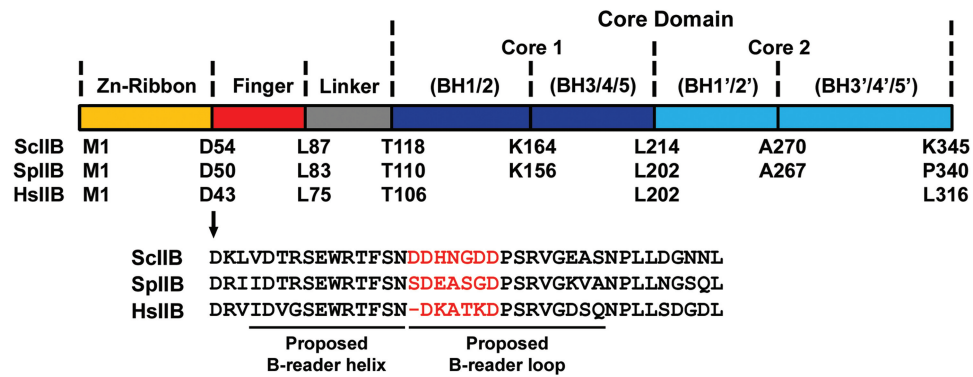


Figure 2. TFIIB structure and functional domains. Depicted at the top is the general structure of TFIIB and the positions of the functional domain boundaries for *S. cerevisiae* (Sc), *S. pombe* (Sp) and human (Hs) TFIIB that were used as junction points for the generation of chimeric variants. Shown below is the sequence of amino acids comprising the B-finger domains, with the proposed B-fingertip residues highlighted in red and the proposed B-reader helix and B-reader loop underlined.

functionally interchangeable between the two systems (9). Thus, these results obtained with highly purified factors confirm that the dual exchange of TFIIB and RNAPII confer the species-specific pattern of TSS utilization between *S. cerevisiae* and *S. pombe*.

Neither the B-finger nor the Zn-ribbon domain of TFIIB determines species-specific TSS utilization for *S. cerevisiae* *in vivo*

Although neither TFIIB nor RNAPII are functionally interchangeable between *S. cerevisiae* and *S. pombe* and the combination of TFIIB and RNAPII confers species-specific initiation patterns, it remains unclear whether both factors directly contribute to the mechanism of TSS utilization. As noted earlier, the role of TFIIB in species-specific TSS utilization might simply reflect the inability of TFIIB and RNAPII from different species to functionally interact. Alternatively, TFIIB may contain intrinsic species-specific domain(s) that directly confer distinct activities during TSS utilization. Because TFIIB has well-separated functional domains (Figure 2), we hypothesized that sequential replacement of one domain at a time with that of another species could identify any species-specific domain(s) of TFIIB responsible for conferring distinct initiation patterns. In contrast, if TFIIB chimeras were functionally compatible with different species without changing their respective distinct initiation patterns, it would strongly support the hypothesis that TFIIB does not play a direct mechanistic role in differential TSS utilization and that RNAPII itself is responsible for the species-specific initiation patterns.

To test this hypothesis, we first constructed TFIIB chimeras (under control of the *SUA7* promoter on a single-copy vector) that contained exchanges of the B-finger domain and determined their effects on *S. cerevisiae* cell growth and start site utilization *in vivo* using a plasmid-shuffle complementation assay (see 'Materials and Methods' section). As expected, neither wild-type SpIIB nor wild-type HsIIB could functionally substitute for ScIIB *in vivo* (Figure 3A-2 and A-3) and their presence did not confer any alteration in TSS

utilization in cells containing wild-type ScIIB (Figure 3C, lanes 2 and 3). Surprisingly, however, ScIIB chimeras containing the *S. pombe* or human B-finger domain were fully functional *in vivo* and did not confer any alterations to the *S. cerevisiae* TSS utilization pattern (Figure 3A-4, A-5, B-4*, B-5* and C, lanes 4, 5, 4* and 5*). Consistent with this result indicating functional conservation of the B-finger domains, an ScIIB mutant containing alanine substitutions for all four aspartate residues in the B-fingertip (Finger4D→A) was also indistinguishable from wild-type ScIIB (Figure 3A-10, B-10* and C, lanes 10 and 10*).

We next analyzed TFIIB chimeras containing exchanges of the Zn-ribbon domain. Neither SpIIB nor HsIIB chimeras containing the ScIIB Zn-ribbon were able to support cell growth in the absence of ScIIB (Figure 3A-6 and A-8) and their presence did not confer any alteration in TSS utilization (Figure 3C, lanes 6 and 8). For the reciprocal chimeras, ScIIB containing the HsIIB Zn-ribbon was also unable to substitute for ScIIB (Figure 3A-9), whereas ScIIB containing the SpIIB Zn-ribbon was nearly identical to wild-type ScIIB *in vivo* (Figure 3A-7, B-7* and C, lanes 7 and 7*). The results obtained for all of the wild-type and chimeric TFIIB constructs expressed from the ScIIB promoter on the single-copy vector were essentially identical to those obtained when they were over-expressed on the high-copy yeast vector pHX under control of the *S. cerevisiae* *ADHI* promoter (Supplementary Figure S2). Thus, taken together, these results demonstrate that neither the B-finger nor the Zn-ribbon domain of TFIIB is responsible for the distinct pattern of *S. cerevisiae* TSS utilization *in vivo*.

Neither the B-linker nor the Core domain of TFIIB determines species-specific TSS utilization for *S. cerevisiae* *in vivo*

We next addressed the potential role of the B-linker and Core domains in species-specific TSS utilization *in vivo*. When expressed at normal levels from a single-copy vector, an ScIIB chimera containing the SpIIB B-linker was fully functional *in vivo* (Figure 4A-4 and B-4*) while

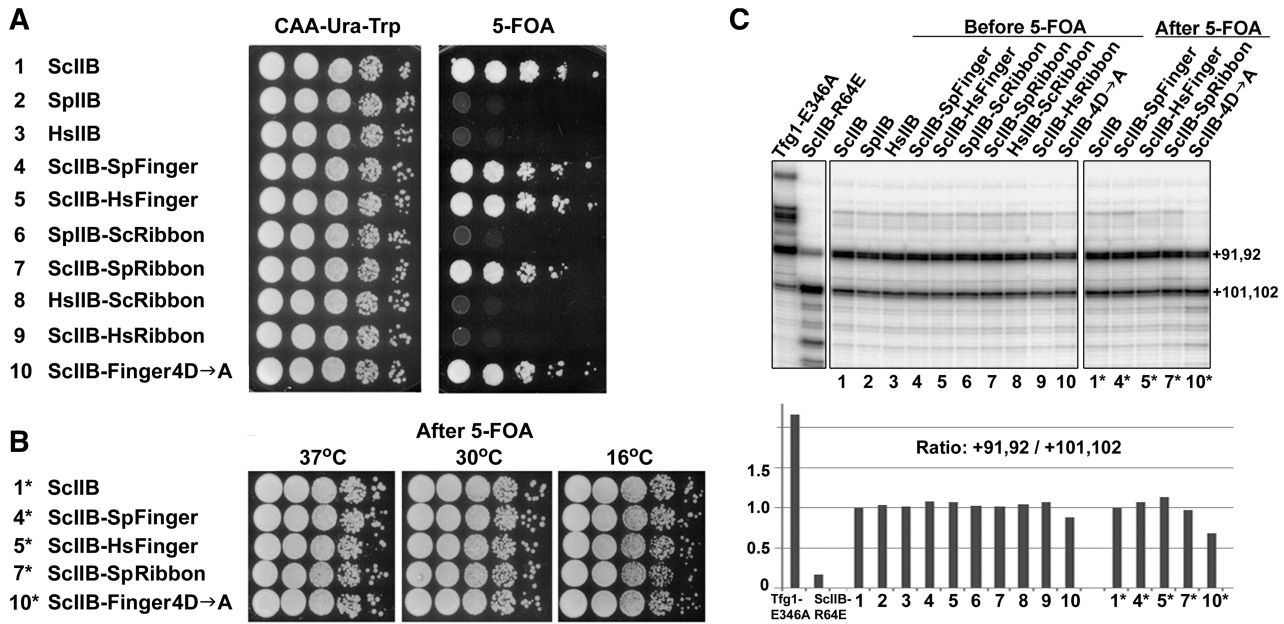


Figure 3. Neither the B-finger nor the Zn-ribbon domain of TFIIB determines species-specific TSS utilization for *S. cerevisiae* *in vivo*. (A) Plasmid shuffle complementation assay. *Saccharomyces cerevisiae* TFIIB shuffle strain FP153 was transformed with the indicated TFIIB-containing plasmid (*TRP1 CEN6* vector) and 10-fold serial dilutions of yeast cultures were spotted on CAA-Ura-Trp or 5-FOA plates and incubated for 3 days at ambient temperature. (B) Growth properties of *S. cerevisiae* strains containing chimeric B-finger or Zn-ribbon TFIIB variants. YPD cultures of the 5-FOA resistant derivatives from A (designated with an asterisk) were grown at ambient temperature and 10-fold serial dilutions spotted on YPD plates and incubated for 2–4 days at the indicated temperatures. (C) Primer extension analysis. Total RNA (30 μ g) from strains containing the indicated TFIIB variants were analyzed utilizing an *ADH1*-specific primer. The numbers to the right of the panel indicate the positions of the major TSSs, where +1 is defined as the A in the translation-initiating ATG. Shown below the gel is the quantitative analysis for the ratios of the major upper start sites (+91, +92) divided by the major lower sites (+101, +102), with the ratio in the wild-type TFIIB strain defined as 1.0. Mutant strains containing the TFIIB Tfg1-E346A or TFIIB R64E substitutions were analyzed as controls for TSS utilization closer to or more downstream from the TATA element, respectively.

ScIIB chimeras containing the entire SpIIIB Core domain or just the Core1 or Core2 subdomains supported cell viability but conferred reduced growth rates at all temperatures (Figure 4A-5, A-6, A-7 and B-5*, B-6*, B-7*). Quantitative analysis of TSS utilization on the endogenous *ADH1* promoter demonstrated that the ScIIB chimera containing the SpIIIB B-linker generated slightly more mRNA for the same amount of total RNA (Figure 4C, lane 4*), whereas the ScIIB chimeras containing the SpIIIB Core, Core1 or Core2 domain resulted in diminished levels of *ADH1* mRNA (Figure 4C, lanes 5*–7*). All four ScIIB chimeras conferred a modest downstream TSS shift to varying degrees, but importantly, none conferred utilization of more upstream start sites characteristic of the *S. pombe* initiation pattern (Figure 4C, lanes 4*–7*). Over-expression of these TFIIB chimeras gave similar results (Supplementary Figure S3) and analyses of TSS utilization on the *SNR20* and *HTB1* promoters indicated that these effects were not promoter specific (Supplementary Figure S4).

Although not directly related to the objective of determining the factor(s) responsible for the species-specific initiation patterns, we further investigated the basis for the reduced growth rates conferred by ScIIB chimeras containing the SpIIIB Core domains through the analysis of additional ScIIB chimeras (Supplementary Figure S5). ScIIB chimeras containing the first two helices

of SpIIIB Core1 (BH1/2) or Core2 (BH1'/2') were unable to support viability (Supplementary Figure S5A-3 and S5A-6), whereas those containing helices 3–5 from SpIIIB Core1 (BH3/4/5) or Core2 (BH3'/4'/5') were fully functional (Supplementary Figure S5A-4 and S5A-7). These results are consistent with the observation that Core1(BH1/2) or Core2(BH1'/2') contact RNAPII (34,35), whereas Core1(BH3/4/5) or Core2(BH3'/4'/5') physically interact the TATA-containing template-TBP complex (31,32). Because the second helix in Core1 (BH2) of human TFIIB has been reported to be incompatible with *S. cerevisiae* *in vivo* (39), ScIIB BH2 (residues KIVKDCAKE) was replaced partially or fully with that from SpIIIB (KIVADCAKE or KIVADTAKQ) or from HsIIIB (KIVVDRTNN or RNIVDRTNN). Consistent with previous observations (39), even partial replacement of ScIIB with HsIIIB BH2 resulted in *S. cerevisiae* lethality, whereas SpIIIB BH2 replacement conferred no alteration in growth or TSS utilization (Supplementary Figure S5C and S5D). In the RNAPII-TFIIB co-crystal structure, the ScIIB Core1 BH2 contacts the first loop of the Rpb2 wall domain (34,35), which is disordered in free RNAPII (50). Thus, to determine whether the interaction between TFIIB BH2 and the Rpb2 wall loop is species-specific, the ScRpb2 wall loop was replaced with the human Rpb2 wall loop (residues KKYGMSIT to SKKGFQDE) by directed PCR fragment recombination. The substitution

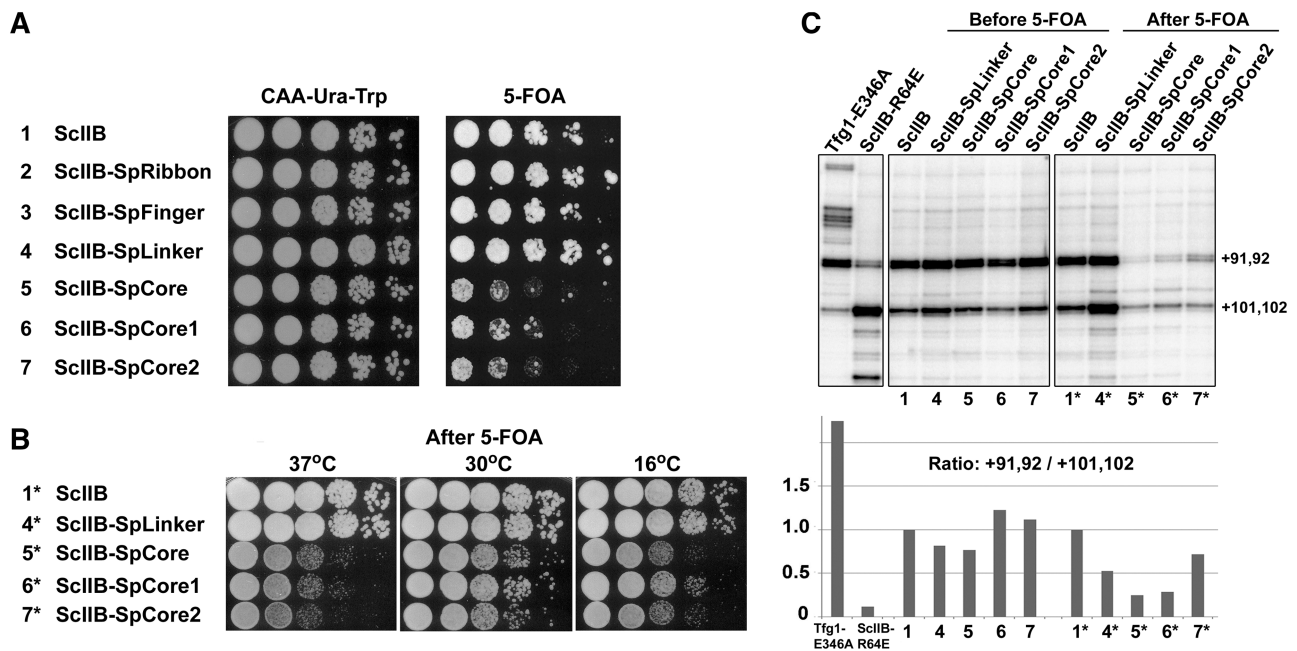


Figure 4. Neither the B-linker nor the Core domain of TFIIB determines species-specific TSS utilization for *S. cerevisiae* *in vivo*. (A) Plasmid shuffle complementation assay. *Saccharomyces cerevisiae* TFIIB shuffle strain FP153 was transformed with the indicated TFIIB-containing plasmid (*TRP1* *CEN6* vector) and 10-fold serial dilutions of yeast cultures were spotted on CAA-Ura-Trp or 5-FOA plates and incubated for 3 days at ambient temperature. (B) Growth properties of *S. cerevisiae* strains containing chimeric B-linker or Core domain TFIIB variants. YPD cultures of the 5-FOA resistant derivatives from A (designated with an asterisk) were grown at ambient temperature and 10-fold serial dilutions spotted on YPD plates and incubated for 2–4 days at the indicated temperatures. (C) Primer extension analysis. Total RNA (30 μ g) from strains containing the indicated TFIIB variants were analyzed utilizing an *ADH1*-specific primer. The numbers to the right of the panel indicate the positions of the major TSSs, where +1 is defined as the A in the translation-initiating ATG. Shown below the gel is the quantitative analysis for the ratios of the major upper start sites (+91, +92) divided by the major lower sites (+101, +102), with the ratio in the wild-type TFIIB strain defined as 1.0. Mutant strains containing the TFIIF Tfg1-E346A or TFIIB R64E substitutions were analyzed as controls for TSS utilization closer to or more downstream from the TATA element, respectively.

strain exhibited normal growth in the presence of wild-type ScIIB (Supplementary Figure S6C) although a slight downstream shift in TSS utilization was detected (Supplementary Figure S6D). Importantly, however, the presence of the human Rpb2 wall loop was unable to rescue the lethality conferred by ScIIB-HsBH2 (Supplementary Figure S6C).

Taken together, all of the ScIIB chimeras sequentially replaced with domains from *S. pombe* both in this and the preceding section supported *S. cerevisiae* cell viability without conferring an upstream TSS shift characteristic of the *S. pombe* initiation pattern. Thus, these *in vivo* results indicate that the difference between the *S. cerevisiae* and *S. pombe* TSS utilization patterns is not due to the function of a species-specific domain within TFIIB and strongly support the view that RNAPII itself is the determining factor.

Reconstituted transcription assays confirm that TFIIB is not the determinant for *S. cerevisiae* species-specific initiation patterns

To complement the *in vivo* analyses, the biochemical activities of TFIIB chimeras were directly determined by performing reconstituted *S. cerevisiae* and *S. pombe* transcription assays. Consistent with the *in vivo* results, the reconstituted transcription assays demonstrated that ScIIB chimeras containing the B-finger domain of SpIIB

or HsiIB, or containing the Zn-ribbon of SpIIB, conferred no change in TSS utilization on the *ADH1* promoter in the *S. cerevisiae* system (Figure 5A, lanes 4, 5 and 7). In addition, an ScIIB chimera containing the HsiIB Zn-ribbon, which was unable to support *S. cerevisiae* viability, nonetheless generated weak but above background transcript levels with no major TSS alteration *in vitro* (Figure 5A, lane 9). Similarly, an SpIIB chimera containing the ScIIB Zn-ribbon fully supported transcription in the *S. pombe* system with no TSS alteration (Figure 5B, lane 2), whereas an ScIIB chimera containing the SpIIB Zn-ribbon did not support detectable transcription (Figure 5B, lane 3).

Analyses of additional ScIIB chimeras containing substitutions of the B-linker or the Core domains provided further evidence that TFIIB is not responsible for the difference in the patterns of TSS utilization between *S. cerevisiae* and *S. pombe*. An ScIIB chimera containing the *S. pombe* Core2 domain supported modest levels of transcription in the *S. cerevisiae* system with initiation occurring at the preferred *S. cerevisiae* start sites, but did not support detectable levels of transcription in the *S. pombe* system (Figure 6A, lane 7 and Figure 6B, lane 7). Most importantly, however, three ScIIB chimeras that contained either the *S. pombe* B-linker, Core or Core1 domain supported transcription in both yeast reconstituted systems with their distinct TSS utilization

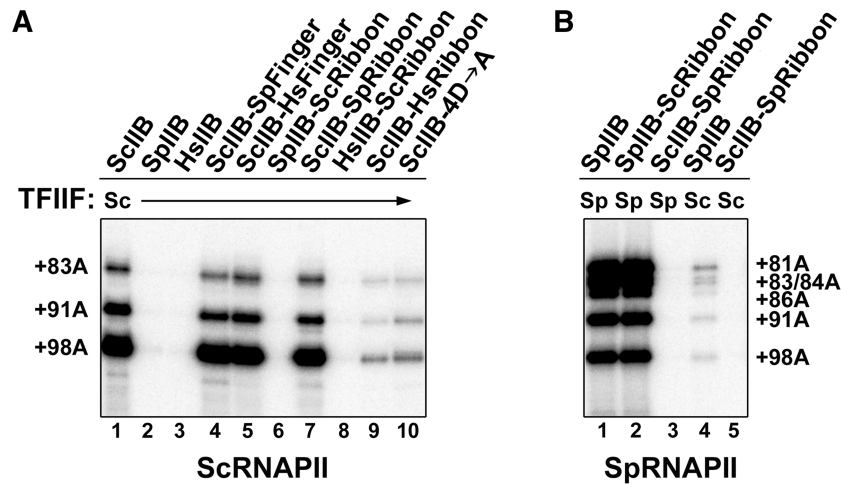


Figure 5. Reconstituted transcription assays confirm that neither the B-finger nor the Zn-ribbon domain of TFIIB determines species-specific TSS utilization for *S. cerevisiae*. Reconstituted *S. cerevisiae* (A) or *S. pombe* (B) direct isotope incorporation transcription assays were carried out as described in ‘Materials and Methods’ section. Reactions contained 3 pmol ScTBP, 3 pmol of the indicated TFIIB, 0.4 pmol RNAPII, 2 pmol TFIIF, 0.25 pmol ScTFIIE, 0.25 pmol ScTFIIH and 0.12 pmol plasmid template pADH1/G-. The numbers on the sides of the panel indicate the positions of the major TSSs within the G-less cassette, where +1 is defined as the first T in the *S. cerevisiae ADH1* TATA element.

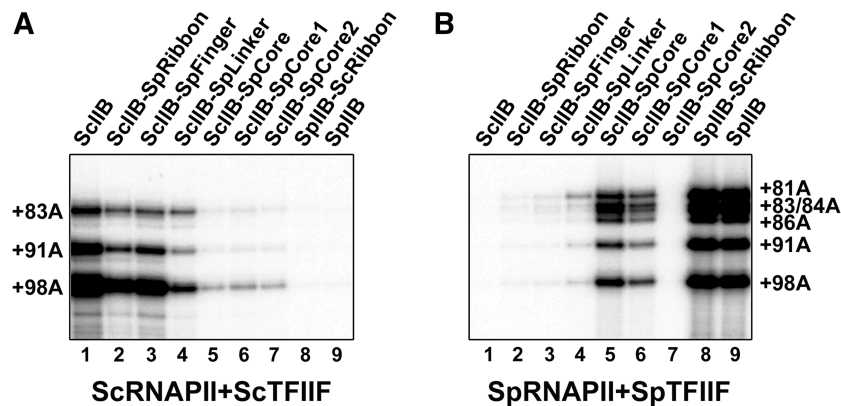


Figure 6. Reconstituted transcription assays confirm that neither the B-linker nor the Core domain of TFIIB determines species-specific TSS utilization for *S. cerevisiae*. Reconstituted *S. cerevisiae* (A) or *S. pombe* (B) direct isotope incorporation transcription assays were carried out as described in ‘Materials and Methods’ section. Reactions contained 3 pmol ScTBP, 3 pmol of the indicated TFIIB, 0.4 pmol RNAPII, 2 pmol TFIIF, 0.25 pmol ScTFIIE, 0.25 pmol ScTFIIH and 0.12 pmol plasmid template pADH1/G-. The numbers on the sides of the panel indicate the positions of the major TSSs within the G-less cassette, where +1 is defined as the first T in the *S. cerevisiae ADH1* TATA element.

patterns being retained (Figure 6A, lanes 4–6 and Figure 6B, lanes 4–6).

DISCUSSION

At the outset of this study, we utilized highly purified proteins to confirm previous observations that neither RNAPII nor TFIIB are by themselves functionally interchangeable between *S. cerevisiae* and *S. pombe* reconstituted transcription systems and that the species-specific combination of TFIIB and RNAPII confers the distinct species-specific pattern of TSS utilization (Figure 1). In order to better understand the mechanistic role of TFIIB within this TFIIB-RNAPII functional unit, we took advantage of the well-defined TFIIB functional domains in constructing TFIIB chimeras containing domains from different species and subsequently

analyzed their activities both *in vivo* and *in vitro*. Systematic replacement of ScTFIIB domains with those of SpTFIIB, which included the Zn-ribbon, B-finger, B-linker and Core domains, demonstrated that all were at least minimally compatible for supporting *S. cerevisiae* viability (Figures 3 and 4), and importantly, none of them conferred shifts to more upstream start sites characteristic of *S. pombe* either *in vivo* (Figures 3 and 4) or in reconstituted transcription assays (Figures 5 and 6).

The mammalian core promoter Initiator (Inr) element comprises the consensus sequence YYANT/AYY which contains the TSS (underlined) and can function to recruit the RNAPII machinery in either the absence or presence of a TATA element (51). In promoters containing a TATA element, the Inr is positioned 25–30 bp downstream and interacts with the TBP-associated factors (TAFs) TAF1 and TAF2 of the TFIID complex (51).

In contrast, preferred TSSs in *S. cerevisiae* are defined by the consensus sequence A(Ar_{ich})₅NYA(A/T)NN(Ar_{ich})₆ that can be positioned from 40 to 200 bp or greater downstream of a TATA element (52,53). Although the factor(s) recognizing this consensus motif in *S. cerevisiae* has not been identified, the B-finger domain of TFIIB has been viewed as a likely candidate in light of the fact that (i) B-finger mutations can alter start site utilization and (ii) molecular modeling of the TFIIB–RNAPII co-crystal structure with that of the DNA-containing RNAPII open complex has suggested that the B-finger is positioned adjacent to the non-template strand in the RNAPII active center (35). Indeed, this structural modeling led to the B-finger being renamed the ‘B-reader’ to underscore its proposed function as the likely candidate for Inr recognition (35). Importantly, however, the results from our study clearly do not support this hypothesis for B-finger function, because both the human and *S. pombe* B-fingers are fully compatible with the *S. cerevisiae* transcription machinery and do not alter the *S. cerevisiae* initiation patterns *in vivo* (Figure 3; Supplementary Figures S2 and S4) nor *in vitro* (Figure 5).

The results from this study strongly indicate that RNAPII is solely responsible for the difference in transcription initiation patterns between *S. cerevisiae* and *S. pombe*. Thus, how do we understand the effects of TFIIF, RNAPII and TFIIB B-finger mutations on these patterns in *S. cerevisiae* and what is the important underlying difference between *S. cerevisiae* and *S. pombe* RNAPII? First, it should be acknowledged that transcription initiation patterns are typically determined by primer extension analysis, and as such, these patterns not only reflect the recognition and productive initiation from a preferred sequence, but in addition, also reflect the proficiency of subsequent elongation to generate a non-aborted extended transcript that is capable of being detected by the oligonucleotide primer. In this regard, we have previously shown that amino acid substitutions in TFIIF that confer upstream shifts in start site utilization coordinately increase early phosphodiester bond formation, RNA–DNA hybrid stability and early elongation processivity (44). Moreover, substitutions in the switch 2 region of the Rpb1 subunit which confer downstream shifts similar to those by TFIIB B-finger substitutions coordinately confer diminished stability of the 3'-end of the RNA–DNA hybrid and increased abortive initiation (14). Although the effects of wild-type and mutant TFIIB B-finger on RNA–DNA hybrid stability in the polymerase active center remain to be determined, our combined studies strongly support the hypothesis that downstream and upstream shifts within the *S. cerevisiae* initiation window are a direct reflection of increased or diminished abortive initiation, respectively [see models presented in references (14,44)]. Significantly, several studies have provided clear support for a model for *S. cerevisiae* transcription initiation that involves PIC formation in the vicinity of the TATA element followed by a directional scanning/translocation process to downstream start sites (44,54,55). It has been proposed that the polymerase scans the downstream sequence for TSSs, and TFIIB mutations that shift initiation further downstream

have historically been referred to as mutations altering start site ‘selection’. Importantly, we refer to these shifts as an alteration in TSS utilization, which we suggest is a more accurate description of the underlying process. In light of the directional nature of the process, the polymerase is undoubtedly not being presented with all potential start sites at the same time and making a ‘selection’ between them. Rather, all evidence indicates that an encountered potential initiation site is either productively utilized or not, and in the case of TFIIB or RNAPII mutations that increase the frequency of abortive initiation, the initiating polymerase complex then continues further downstream with the opportunity to utilize a more distal site. In the case of TFIIF mutations that confer upstream shifts, suboptimal start sites encountered early in the process undergo diminished abortive initiation giving rise to an increase in mRNAs whose 5'-ends map closer to the TATA element, i.e. more productive utilization of upstream sub-optimal sites (44).

The conclusion that RNAPII is solely responsible for the difference in transcription initiation patterns between *S. cerevisiae* and *S. pombe* suggests an intrinsic difference between the respective polymerases. Although the specific molecular difference remains to be determined, it is tempting to speculate that the ability of *S. pombe* RNAPII to productively utilize start site sequences that are sub-optimal for *S. cerevisiae* RNAPII might be due to a difference in the *S. pombe* RNAPII active site that coordinately enhances early phosphodiester bond formation and, importantly, intrinsic elongation proficiency. In *S. cerevisiae*, only 4% of protein coding genes contain introns and the average gene length is ~1.5 kb. Significantly, however, ~40% of *S. pombe* protein coding genes contain introns and the average gene length is ~3 kb (56,57). Thus, the evolutionary divergence to utilize alternative splicing to generate protein diversity has been accompanied by the requirement to efficiently transcribe greater distances. Because our previous studies of *S. cerevisiae* TFIIF and the switch 2 region of the *S. cerevisiae* RNAPII active site have clearly demonstrated a direct correlation between the efficiency of early bond formation, the utilization of sub-optimal non-consensus start sites closer to the TATA element and more processive elongation (14,44), we propose that *S. pombe* RNAPII, and similarly human RNAPII, have evolved to be intrinsically more efficient in transcript elongation and that the structural changes underlying enhanced elongation also manifest in more efficient early phosphodiester bond formation during productive initiation. Further biochemical and structural studies of human, *S. pombe* and *S. cerevisiae* RNAPII are needed to test the proposed mechanistic coupling of initiation and elongation proficiency and the underlying structural differences between the polymerases.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online: Supplementary Table 1 and Supplementary Figures 1–6.

FUNDING

National Institutes of Health Public Health Service [GM51124 to A.S.P.]. Funding for open access charge: National Institutes of Health.

Conflict of interest statement. None declared.

REFERENCES

- Hahn,S. (2004) Structure and mechanism of the RNA polymerase II transcription machinery. *Nat. Struct. Mol. Biol.*, **11**, 394–403.
- Thomas,M.C. and Chiang,C.M. (2006) The general transcription machinery and general cofactors. *Crit. Rev. Biochem. Mol. Biol.*, **41**, 105–178.
- Kornberg,R.D. (2007) The molecular basis of eukaryotic transcription. *Proc. Natl Acad. Sci. USA*, **104**, 12955–12961.
- Mitsuzawa,H. and Ishihama,A. (2004) RNA polymerase II transcription apparatus in *Schizosaccharomyces pombe*. *Curr. Genet.*, **44**, 287–294.
- Chen,B.S., Mandal,S.S. and Hampsey,M. (2004) High-resolution protein-DNA contacts for the yeast RNA polymerase II general transcription machinery. *Biochemistry*, **43**, 12741–12749.
- Miller,G. and 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.
- 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.
- Kim,T.K., Ebright,R.H. and Reinberg,D. (2000) Mechanism of ATP-dependent promoter melting by transcription factor IIH. *Science*, **288**, 1418–1422.
- 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.
- Struhl,K. (1987) Promoters, activator proteins, and the mechanism of transcriptional initiation in yeast. *Cell*, **49**, 295–297.
- Choi,W.S., Yan,M., Nusinow,D. and Gralla,J.D. (2002) In vitro transcription and start site selection in *Schizosaccharomyces pombe*. *J. Mol. Biol.*, **319**, 1005–1013.
- Russell,P.R. (1983) Evolutionary divergence of the mRNA transcription initiation mechanism in yeast. *Nature*, **301**, 167–169.
- Berteroan,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.
- Majovski,R.C., Khapersky,D.A., Ghazy,M.A. and Ponticelli,A.S. (2005) A functional role for the switch 2 region of yeast RNA polymerase II in transcription start site utilization and abortive initiation. *J. Biol. Chem.*, **280**, 34917–34923.
- Chen,H.T., Warfield,L. and Hahn,S. (2007) The positions of TFIIF and TFIIE in the RNA polymerase II transcription preinitiation complex. *Nat. Struct. Mol. Biol.*, **14**, 696–703.
- Hull,M.W., McKune,K. and Woychik,N.A. (1995) RNA polymerase II subunit RPB9 is required for accurate start site selection. *Genes Dev.*, **9**, 481–490.
- 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.
- Ziegler,L.M., Khapersky,D.A., Ammerman,M.L. and Ponticelli,A.S. (2003) Yeast RNA polymerase II lacking the Rpb9 subunit is impaired for interaction with transcription factor IIF. *J. Biol. Chem.*, **278**, 48950–48956.
- 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.
- 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.
- 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.
- 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.
- 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.
- 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.
- Ghazy,M.A., Brodie,S.A., Ammerman,M.L., Ziegler,L.M. and Ponticelli,A.S. (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.
- Lorch,Y., LaPointe,J.W. and Kornberg,R.D. (1992) Initiation on chromatin templates in a yeast RNA polymerase II transcription system. *Genes Dev.*, **6**, 2282–2287.
- Buratowski,S., Hahn,S., Sharp,P.A. and Guarente,L. (1988) Function of a yeast TATA element-binding protein in a mammalian transcription system. *Nature*, **334**, 37–42.
- Chen,H.T. and Hahn,S. (2003) Binding of TFIIB to RNA polymerase II: Mapping the binding site for the TFIIB zinc ribbon domain within the preinitiation complex. *Mol. Cell*, **12**, 437–447.
- Chen,H.T. and Hahn,S. (2004) Mapping the location of TFIIB within the RNA polymerase II transcription preinitiation complex: a model for the structure of the PIC. *Cell*, **119**, 169–180.
- Buratowski,S. and Zhou,H. (1993) Functional domains of transcription factor TFIIB. *Proc. Natl Acad. Sci. USA*, **90**, 5633–5637.
- 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.
- Tsai,F.T. and Sigler,P.B. (2000) Structural basis of preinitiation complex assembly on human pol II promoters. *EMBO J.*, **19**, 25–36.
- 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.
- Liu,X., Bushnell,D.A., Wang,D., Calero,G. and Kornberg,R.D. (2010) Structure of an RNA polymerase II-TFIIB complex and the transcription initiation mechanism. *Science*, **327**, 206–209.
- Kostrewa,D., Zeller,M.E., Armache,K.J., Seizl,M., Leike,K., Thomm,M. and Cramer,P. (2009) RNA polymerase II-TFIIB structure and mechanism of transcription initiation. *Nature*, **462**, 323–330.
- Thompson,N.E., Glaser,B.T., Foley,K.M., Burton,Z.F. and Burgess,R.R. (2009) Minimal promoter systems reveal the importance of conserved residues in the B-finger of human transcription factor IIB. *J. Biol. Chem.*, **284**, 24754–24766.
- Tran,K. and Gralla,J.D. (2008) Control of the timing of promoter escape and RNA catalysis by the transcription factor IIB fingertip. *J. Biol. Chem.*, **283**, 15665–15671.
- Choi,W.S., Lin,Y.C. and Gralla,J.D. (2004) The *Schizosaccharomyces pombe* open promoter bubble: mammalian-like arrangement and properties. *J. Mol. Biol.*, **340**, 981–989.
- Shaw,S.P., Wingfield,J., Dorsey,M.J. and Ma,J. (1996) Identifying a species-specific region of yeast TFIIB in vivo. *Mol. Cell Biol.*, **16**, 3651–3657.

40. Bangur,C.S., Faitar,S.L., Folster,J.P. and Ponticelli,A.S. (1999) An interaction between the N-terminal region and the core domain of yeast TFIIB promotes the formation of TATA-binding protein-TFIIB-DNA complexes. *J. Biol. Chem.*, **274**, 23203–23209.
41. Ponticelli,A.S., Pardee,T.S. and Struhl,K. (1995) The glutamine-rich activation domains of human Sp1 do not stimulate transcription in *Saccharomyces cerevisiae*. *Mol. Cell Biol.*, **15**, 983–988.
42. Yang,C., Khapersky,D.A., Hou,M. and Ponticelli,A.S. (2010) Improved methods for expression and purification of *Saccharomyces cerevisiae* TFIIF and TFIH; 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.*, **70**, 172–178.
43. Pardee,T.S., Ghazy,M.A. and Ponticelli,A.S. (2003) Yeast and human RNA polymerase II elongation complexes: evidence for functional differences and postinitiation recruitment of factors. *Eukaryot. Cell*, **2**, 318–327.
44. Khapersky,D.A., Ammerman,M.L., Majovski,R.C. and Ponticelli,A.S. (2008) Functions of *Saccharomyces cerevisiae* TFIIF during transcription start site utilization. *Mol. Cell Biol.*, **28**, 3757–3766.
45. Spahr,H., Khorosjutina,O., Baraznenok,V., Linder,T., Samuelsen,C.O., Hermand,D., Makela,T.P., Holmberg,S. and Gustafsson,C.M. (2003) Mediator influences *Schizosaccharomyces pombe* RNA polymerase II-dependent transcription in vitro. *J. Biol. Chem.*, **278**, 51301–51306.
46. Ponticelli,A.S. and Struhl,K. (1990) Analysis of *Saccharomyces cerevisiae* his3 transcription in vitro: biochemical support for multiple mechanisms of transcription. *Mol. Cell Biol.*, **10**, 2832–2839.
47. Osorio,H., Carvalho,E., del Valle,M., Gunther Sillero,M.A., Moradas-Ferreira,P. and Sillero,A. (2003) H₂O₂, but not menadione, provokes a decrease in the ATP and an increase in the inosine levels in *Saccharomyces cerevisiae*. An experimental and theoretical approach. *Eur. J. Biochem.*, **270**, 1578–1589.
48. Tamayo,E., Bernal,G., Teno,U. and Maldonado,E. (2004) Mediator is required for activated transcription in a *Schizosaccharomyces pombe* in vitro system. *Eur. J. Biochem.*, **271**, 2561–2572.
49. Kimura,M. and Ishihama,A. (2004) Tfg3, a subunit of the general transcription factor TFIIF in *Schizosaccharomyces pombe*, functions under stress conditions. *Nucleic Acids Res.*, **32**, 6706–6715.
50. 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.
51. Juven-Gershon,T. and Kadonaga,J.T. (2010) Regulation of gene expression via the core promoter and the basal transcriptional machinery. *Dev. Biol.*, **339**, 225–229.
52. Zhang,Z. and Dietrich,F.S. (2005) Mapping of transcription start sites in *Saccharomyces cerevisiae* using 5' SAGE. *Nucleic Acids Res.*, **33**, 2838–2851.
53. Maicas,E. and Friesen,J.D. (1990) A sequence pattern that occurs at the transcription initiation region of yeast RNA polymerase II promoters. *Nucleic Acids Res.*, **18**, 3387–3393.
54. Giardina,C. and Lis,J.T. (1993) DNA melting on yeast RNA polymerase II promoters. *Science*, **261**, 759–762.
55. Kuehner,J.N. and Brow,D.A. (2006) Quantitative analysis of in vivo initiator selection by yeast RNA polymerase II supports a scanning model. *J. Biol. Chem.*, **281**, 14119–14128.
56. Fedorova,L. and Fedorov,A. (2003) Introns in gene evolution. *Genetica*, **118**, 123–131.
57. Wood,V., Gwilliam,R., Rajandream,M.A., Lyne,M., Lyne,R., Stewart,A., Sgouros,J., Peat,N., Hayles,J., Baker,S. *et al.* (2002) The genome sequence of *Schizosaccharomyces pombe*. *Nature*, **415**, 871–880.