Published in final edited form as:

Protein Expr Purif. 2010 April ; 70(2): 172–178. doi:10.1016/j.pep.2009.09.021.

Improved methods for expression and purification of *S. cerevisiae* TFIIF and TFIH; Identification of a functional *E. coli* promoter and internal translation initiation within the N-terminal coding region of the TFIIF *TFG1* subunit

Chen Yang^a, Denys A. Khapersky^{a,b}, Min Hou^a, and Alfred S. Ponticelli^{a,*}

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

Abstract

The basal RNA polymerase II (RNAPII) transcription machinery is composed of RNAPII and the general transcription factors (TF) TATA binding protein (TBP), TFIIB, TFIIE, TFIIF and TFIH. Due to the powerful genetic and molecular approaches that can be utilized, the budding yeast *S. cerevisiae* has proven to be an invaluable model system for studies of the mechanisms of RNAPII transcription. Complementary biochemical studies of the *S. cerevisiae* basal transcription machinery, however, have been hampered by difficulties in the purification of TFIIF and TFIH, most notably due to the severe toxicity of the TFIIF Tfg1 subunit in *E. coli* and the complexity of the purification scheme for native TFIH. Here, we report the elimination of *TFG1*-associated toxicity in *E. coli*, the identification and removal of a functional *E. coli* promoter and internal translation initiation within the N-terminal coding region of *TFG1*, and the efficient production and two step purification of recombinant TFIIF complexes. We also report conditions for the efficient two step tandem affinity purification (TAP) of holo-TFIH, core TFIH and TFIK complexes from yeast whole cell extracts.

Keywords

S. cerevisiae; TFIIF; TFIH; transcription

Introduction

The general transcription factors TFIIF and TFIH play essential roles in transcription of eukaryotic protein encoding genes by RNA polymerase II (RNAPII). Studies in both mammalian cells and the budding yeast *S. cerevisiae* have demonstrated multiple roles for TFIIF during the transcription cycle that include (i) recruitment of RNAPII, TFIIE and TFIH into the preinitiation complex (PIC) [1,2]; (ii) wrapping of promoter DNA around RNAPII in the PIC [3]; (iii) modulating transcription start site utilization [4,5]; (iv) facilitating efficient promoter escape of RNAPII [6,7]; (v) increasing the rate of transcript elongation [8,9]; (v)

*Corresponding author: Mailing address: Department of Biochemistry, School of Medicine and, Biomedical Sciences, State University of New York, Buffalo, New York 14214-3000., Phone: (716) 829-2473. Fax: (716) 829-2725. asp@buffalo.edu.

^bCurrent location: Department of Microbiology and Immunology, Dalhousie University, Halifax, Nova Scotia, Canada B3H 3J5

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

stimulating the rescue of paused elongation complexes in conjunction with TFIIS [10]; and (vi) stimulating the RNAPII C-terminal domain (CTD) phosphatase Fcp1 to facilitate recycling of RNAPII back into the initiation-competent hypo-phosphorylated form [11]. Similarly, the TFIIF complex plays several important roles that include (i) unwinding of promoter DNA through multiple ATPase/helicase activities [12–19]; (ii) phosphorylation of the RNAPII C-terminal domain (CTD) during the transition from initiation to transcript elongation [20,21]; and (iii) participating in nucleotide excision repair of damaged DNA [22–24].

S. cerevisiae TFIIF comprises three protein subunits designated Tfg1, Tfg2, and Tfg3. The Tfg1 (calculated MW ~82 kDa) and Tfg2 (calculated MW ~47 kDa) subunits are homologous to the RAP74 and RAP30 subunits of mammalian TFIIF, respectively, and are essential for cell viability [25]. The Tfg3 subunit, also designated as Taf14, is associated with at least six yeast nuclear complexes, has no known homologues in higher eukaryotes, and is not required for viability [25–32]. The *S. cerevisiae* holo-TFIIF complex contains ten polypeptide subunits that include Ssl2, the core TFIIF sub-complex comprising Rad3, Tfb1, Tfb2, Ssl1, Tfb4 and Tfb5, and the TFIK sub-complex, which contains Ccl1, Tfb3 and Kin28 [12,33–35] (Fig. 3A).

Despite the powerful genetic approaches that can be utilized in studies with *S. cerevisiae*, comprehensive structure-function analyses of yeast TFIIF and TFIIF have been hindered by difficulties in the production of the recombinant proteins and/or purification of the native complexes from yeast cells. Although Holo-TFIIF has been purified from a strain containing a hexahistidine-tagged Tfb1 subunit, the purification scheme nonetheless involves five to six chromatographic steps to reach apparent homogeneity [36,37]. For TFIIF, native Tfg1-Tfg2-Tfg3 complexes can be isolated from yeast extract using tandem affinity purification (TAP), but additional purification is required due to contamination with RNAPII [38]. Importantly, the extreme toxicity of plasmids containing the Tfg1 subunit in *E. coli* has presented a major obstacle in the production of recombinant TFIIF. We previously reported a relatively inefficient approach for the production of functional recombinant Tfg1-Tfg2 complexes that involved a low-copy vector, induction of Tfg1, Tfg2 co-expression by infection with bacteriophage containing T7 RNA polymerase, and multi-step affinity purification [39]. In the work presented here, we report the elimination of *TFG1*-associated toxicity in *E. coli*, the identification of a functional *E. coli* promoter and internal translation initiation within the N-terminal coding region of *TFG1*, and two-step purification of recombinant Tfg1-Tfg2 complexes. We also report conditions for efficient TAP-purification of TFIIF complexes from yeast whole cell extracts.

Materials and methods

Expression plasmids and yeast strains

The initial dual-expression plasmid used in these studies (pDt/g1g2) was constructed using pCOLADuet-1 (Novagen) as the starting vector and contains the coding regions for the *S. cerevisiae* TFIIF Tfg1 and Tfg2 subunits (Tfg2p: myc-epitope tag at the C-terminus; Tfg1p: hexahistidine and myc-epitope tags at the N-terminus, FLAG and TAP tags at the C-terminus) under control of the T7 promoters in the adjacent MCS1 and MCS2 expression cassettes, respectively. The final dual-expression plasmid (pDt/g1*g2) was generated from pDt/g1g2 and contains the base substitutions indicated as mutant 2 (M2) in Figure 1C. *S. cerevisiae* strains YT062 and YT063 have been described previously [33] and contain TAP-tagged versions of the TFIIF subunits Tfb3 and Tfb4, respectively.

Primer extension analysis of RNA from *E. coli*

Overnight cultures of *E. coli* strain RosettaTM (DE3) (Novagen) harboring pDt/g1g2 or the subsequent M1 or M2 variants were grown at 37°C in LB media containing 50 g/ml kanamycin,

34 µg/ml chloramphenicol and then diluted to an initial OD₆₀₀ of 0.1 in 50 ml of LB media containing 50 µg/ml kanamycin, 34 µg/ml chloramphenicol. Cultures were grown at 37°C to an OD₆₀₀ of 0.6 and 10 ml was harvested immediately (–IPTG) or after 2 additional hrs. growth in the presence of 1 mM IPTG. Total RNA was isolated using hot acidic phenol essentially as described [40] and 30 µg was analyzed by primer extension as described [41] using the oligonucleotide primer 5'-dATATTCATTAGGATCTTCGTCC-3' (corresponding to *TFG1* positions +294 to +273, where +1 is defined as the A in the translation-initiating ATG).

Western blotting of Tfg1 proteins

For the analysis of Tfg1 proteins in *E. coli*, pre- and post-induction cultures (1.5 ml) were harvested by centrifugation, resuspended with 80 µl 2X SDS-PAGE sample buffer, and lysed by incubation at 95°C for 5 min. Samples were centrifuged at 14,000 × g for 3 min. and 5 µl of the supernatant resolved on an SDS-10% polyacrylamide gel, transferred to Immobilon-P PVDF membranes (Millipore), and probed with a mouse monoclonal anti-FLAG antibody (Sigma, 1:1000 dilution) followed by goat anti-mouse IgG coupled to horseradish peroxidase (Pierce, 1:1000 dilution). Immune complexes were visualized by the addition of SuperSignal ECL substrate (Pierce) and autoradiography. For the analysis of Tfg1p in purified recombinant Tfg1-Tfg2 complexes, 120 ng protein was analyzed using either the mouse monoclonal anti-FLAG antibody (Sigma, 1:1000 dilution) or a mouse monoclonal anti-His6 antibody (Covance, 1:1000 dilution) for detection of the Tfg1p C-terminus or N-terminus, respectively.

Expression and purification of recombinant yeast TFIIF

Overnight cultures of *E. coli* strain RosettaTM (DE3) harboring pDt/g1*g2 (inoculated from a freshly streaked colony) were grown at 37°C in LB media containing 50 µg/ml kanamycin, 34 µg/ml chloramphenicol and then diluted to an initial OD₆₀₀ of 0.1 in 1 liter of LB media containing 50 µg/ml kanamycin, 34 µg/ml chloramphenicol. The culture was grown at 30°C to an OD₆₀₀ of 0.6–0.9 and expression of the recombinant Tfg1 and Tfg2 proteins was induced with the addition of 1 ml 1M IPTG (1 mM final concentration). The culture was grown at 30°C for 3 hrs. and the cells were harvested by centrifugation and frozen at –80°C. Frozen cell pellets were thawed on ice, resuspended with 27 ml of ice-cold lysis buffer (50 mM Tris-HCl pH 8.0, 500 mM NaCl, 10% glycerol, 1 mM PMSF, 2 mM Benzamidine, 20 µM Bestatin, 3 µM Pepstatin A, 1 µM Leupeptin) containing 2 mg/ml lysozyme, and the suspension was incubated on ice for 30 min. with occasional mixing. Cells were lysed by the addition of 3 ml of 20% Triton X-100 (in lysis buffer), 300 µl of 1M MgCl₂, 300 µl of 10 mg/ml RNase A and 300 µl of 2 mg/ml DNase I. The suspension was mixed by inversion and incubated on ice for 30 min. with occasional mixing, after which another 300 µl of 2 mg/ml DNase I was added and incubation continued for an additional 30 min. The lysate was centrifuged at 15,000 × g for 20 min. at 4°C, and the cleared supernatant was adjusted to a final concentration of 1 mM imidazole and batch loaded onto 1 ml Ni-NTA resin (Invitrogen) for 90 min. at 4°C. The resin was subsequently packed into a column and washed with 30 ml of lysis buffer containing 1 mM imidazole, (omitting Bestatin, Pepstatin A, Leupeptin) and the protein was eluted with 10 ml of lysis buffer containing 250 mM imidazole (omitting Bestatin, Pepstatin A, Leupeptin), collecting 1 ml fractions. Peak fractions of Tfg1,2 protein were pooled (2–3 ml final), added to 400 µl of IgG Sepharose resin (GE Healthcare) pre-equilibrated with Buffer A (20 mM HEPES-KOH pH 7.6, 300 mM potassium acetate, 10% glycerol, 2 mM DTT, 0.5 mM EDTA, 0.01% NP-40, 1 mM PMSF, 2 mM Benzamidine), and incubated at 4°C overnight (or greater than 5 hrs.) with gentle rocking. The resin was harvested by centrifugation for 5 min. at 2500 × g, washed twice with 1.5 ml ice-cold Buffer A, and resuspended with 1.5 ml Buffer A containing 150 units AcTEV protease (Invitrogen). The mixture was incubated at 16°C for 3 hrs. with gentle rocking and then centrifuged as before, collecting the eluant supernatant. The resin was washed once with 0.5 ml Buffer A and the supernatant was pooled with the eluant supernatant. If necessary, the TEV eluant/wash could be concentrated using a Microcon

Ultracel YM-50 spin column (Millipore). Typical yield was 1–1.2 mg of TFIIF protein per 1 liter of starting culture.

Purification of TFIIF complexes from yeast extracts

Cultures (20 L) of yeast strains YT062 or YT063 were grown at 30°C in YPD media to an OD₆₀₀ of ~ 6.0, harvested by centrifugation, and the cell pellet (~160 g) resuspended in ½ volume (80 ml) of yeast lysis buffer (150 mM Tris-acetate pH 7.9, 100 mM potassium acetate, 20% glycerol, 1 mM DTT, 1 mM EDTA), and the suspension frozen at –80°C. The frozen suspension was thawed on ice, protease inhibitors were added (final concentrations: 2 mM Benzamidine, 1 mM PMSF, 3 µM Chymostatin, 2 µM Pepstatin A, 1 µM Leupeptin), and the cells were disrupted with a BioSpec Bead Beater, using an ice/water bath surrounded stainless steel vessel, 150 ml glass beads (0.5 mm), and 10 cycles of 1 min. beating and a 1.5 min. rest. The supernatant was harvested and the beads were rinsed once with 150 ml lysis buffer and pooled with the supernatant. Potassium acetate was added to a final concentration of 600 mM, the lysate was stirred for 20 min. at 4°C, and then centrifuged at 20,000 × g for 15 min. at 4°C. The supernatant was then centrifuged using a 45 Ti rotor at 100,000 × g for 60 min. at 4°C, or where indicated, adjusted to a final concentration of 0.33% polyethyleneimine (PEI) and stirred for 15 min. at 4°C prior to the high speed clarification. In the absence of PEI, three phases were observed after the high speed spin which included a top lipid phase, a middle clear lysate (harvested), and a loose pellet of precipitated material. In the presence of PEI, the pellet was solid, and the top phase was not observed. The clear lysate was harvested and filtered through Whatman 1 paper. Typical recoveries were approximately 270 ml for the untreated lysate (20–25 mg/ml protein) and 350 ml for the PEI treated lysate (16 mg/ml protein). The lysates were adjusted to a final concentration of 0.01% NP-40, protease inhibitors were again added, and 45 ml aliquots were added to 400 µl IgG Sepharose resin pre-equilibrated with Buffer A (20 mM HEPES-KOH pH 7.6, 300 mM potassium acetate, 10% glycerol, 2 mM DTT, 0.5 mM EDTA, 0.01% NP-40, 2 mM Benzamidine, 1 mM PMSF) and incubated at 4°C overnight (or greater than 2 hrs.) with gentle rocking. For each 45 ml sample, the resin was harvested by centrifugation for 3 min. at 2500 × g, washed once with 8 ml ice-cold Buffer A, and then washed three times with 1.2 ml ice-cold Buffer A in a microcentruge tube. The resin was then resuspended in 1.5 ml Buffer A containing 150 units AcTEV protease (Invitrogen), incubated at 16°C for 3 hrs. with gentle rocking, centrifuged for 2 min. at 2500 × g, and the eluant supernatant was collected. The resin was washed once with 1 ml Buffer A and the supernatant was pooled with the eluant. Each 2.5 ml pool was adjusted to 2 mM CaCl₂, raised to 10 ml total volume by the addition of 7.5 ml of calmodulin binding buffer (CBB; 20 mM HEPES-KOH pH 7.6, 300 mM potassium acetate, 10% glycerol, 2 mM CaCl₂, 2 mM DTT, 1 mM magnesium acetate, 1 mM imidazole, 0.01% NP-40, 2 mM Benzamidine, 1 mM PMSF), added to 200 µl CBB-equilibrated Calmodulin Affinity Resin (Stratagene), and incubated for 90 min. with gentle rocking at 4°C. The resin was harvested by centrifugation for 3 min. at 2500 × g, washed twice with 0.8 ml ice-cold CBB in a microcentruge tube, and the TFIIF complexes were eluted by the addition of 400 µl of calmodulin elution buffer (CEB; 20 mM HEPES-KOH pH 7.6, 300 mM potassium acetate, 10% glycerol, 3 mM EGTA, 2 mM DTT, 1 mM magnesium acetate, 0.01% NP-40, 2 mM Benzamidine, 1 mM PMSF) and incubation for 15 min. with gentle rocking at 4°C. The resin was harvested by centrifugation for 2 min. at 2500 × g, the eluant was harvested, and the elution procedure was repeated once more using 200 µl CEB. The pooled eluants (600 l) were then concentrated using a Microcon Ultracel YM-30 spin column (Millipore). Total yields from 20 liters of strain YT062 (Tfb3-TAP) were approximately 200 µg of a mixture of holo-TFIIF and TFIIFK for - PEI preps, and 90 µg of pure TFIIFK for + PEI preps. Total yields from 20 liters of strain YT063 (Tfb4-TAP) were approximately 100 µg of a mixture of holo-TFIIF and core TFIIF for - PEI preps, and 40 µg of pure core TFIIF for + PEI preps.

In vitro transcription and CTD kinase assays

Reconstituted *in vitro* transcription reactions were carried out using plasmid template pG5CG-as described [39], using 20 units of Ribolock RNase inhibitor (Fermentas) substituted for RNasin and the amount of TFIID indicated in the legend to Figure 4. CTD kinase reactions (15 μ l) contained 50 mM HEPES-KOH pH 7.6, 8% glycerol, 80 mM potassium acetate, 10 mM magnesium acetate, 5 mM EGTA, 2.5 mM DTT, 1 pmol (500 ng) purified *S. cerevisiae* RNAPII (IIA, hypo-phosphorylated form) and the amount of TFIID indicated in the legend to Figure 4. Reactions were initiated by the addition of dATP or ATP (400 μ M final) and incubated at ambient temperature for 10 min. Reactions were terminated by the addition of 15 μ l 2X SDS-PAGE sample buffer and 10 μ l was resolved on an SDS-8% polyacrylamide gel, transferred to Immobilon-P PVDF membranes (Millipore), and probed with mouse anti-phosphoSer5-CTD antibody (H14, Covance, 1:1000 dilution) followed by goat anti-mouse IgM coupled to horseradish peroxidase (Pierce, 1:1000 dilution). Immune complexes were visualized by the addition of SuperSignal ECL substrate (Pierce) and autoradiography.

Results

The N-terminal coding region of *S. cerevisiae* TFG1 contains a functional *E. coli* promoter and undergoes internal translation initiation at Met-114

We previously reported an approach for the production of functional recombinant Tfg1-Tfg2 complexes that involved a broad host range low-copy vector, induction of Tfg1, Tfg2 co-expression by infection with bacteriophage CE6 carrying T7 RNA polymerase, and multi-step affinity purification [39]. The low-copy vector was required due to the previously identified toxicity of the *TFG1* gene in *E. coli*, and induction by CE6 infection was required due to the inability to establish stable transformants of the co-expression vector in *E. coli* cells containing inducible T7 RNA polymerase, such as BLR (DE3). In addition, the purification involved the preparation and use of an anti-myc antibody column, making the approach somewhat cost ineffective in light of the yield of Tfg1-Tfg2 complexes obtained. To address this last issue, we sought to determine whether a TAP epitope tag could be substituted for the myc tag that was fused to the C-terminus of Tfg1. Using the plasmid pCOLADuet-1 as a co-expression vector for Tfg1 and Tfg2, we generated an initial construct (designated pDt/g1g2, see Materials and methods) and discovered that the addition of the C-terminal TAP tag significantly reduced the toxicity of the *TFG1*-containing plasmid in *E. coli*, suggesting that this toxicity involves the C-terminal region of Tfg1. More importantly, this result allowed for stable transformants of pDt/g1g2 to be established in host strains such as RosettaTM (DE3), which carries both IPTG-inducible T7 RNA polymerase as well as the plasmid pRARE, which supplies tRNAs for rarely used codons in *E. coli*. Using Rosetta as the host strain, we initially examined IPTG induction of Tfg1 by western blotting using antibody recognizing the FLAG epitope that was part of the C-terminal TAP tag used (see Materials and methods). The results demonstrated that full-length Tfg1 protein was induced, but that the vast majority of the Tfg1 detected was an N-terminal truncation product that, importantly, was present prior to IPTG induction (Fig. 1A, lanes 1, 4). This result suggested the potential presence of a cryptic functional *E. coli* promoter somewhere within the first 300 base pairs of the *TFG1* coding region. Using primer extension analysis, we identified an IPTG-independent transcript initiating within the N-terminal Tfg1 coding region (Fig. 1B, lane 1; Fig. 1C) and was able to eliminate the *E. coli* promoter generating this transcript by site-directed mutagenesis (mutant M1, Fig. 1B, lane 2; Fig. 1C). Significantly, the elimination of this *E. coli* promoter dramatically increased the growth rate of strains carrying the M1 plasmid (data not shown). Although the elimination of this transcript generated by the *E. coli* promoter abolished the production of the Tfg1 N-terminal truncation in the absence of IPTG, (Fig. 1A, lane 2), it did not abolish the generation of this product after IPTG induction (Fig. 1A, lane 5). This result suggested that, in addition to the presence of an *E. coli* promoter in the Tfg1 N-terminal coding region, this region might also contain an alternative

downstream ribosome binding site for internal translation initiation. Using site-directed mutagenesis, we confirmed that, indeed, internal translation initiation does occur at Tfg1 Met-114 to generate the prominent N-terminal truncation product (mutant M2, Fig. 1A, lane 6; Fig. 1C).

Induction and purification of recombinant *S. cerevisiae* Tfg1-Tfg2 complexes

Having eliminated both a cryptic *E. coli* promoter and the internal translation initiation within the N-terminal Tfg1 coding region, we proceeded to utilize the M2 mutant construct (designated pDt/g1*g2, Fig. 2A) for production and purification of Tfg1-Tfg2 complexes. Briefly, the approach included (i) IPTG induction of Tfg1 and Tfg2 in the host strain Rosetta™ (DE3); (ii) enrichment of Tfg1-Tfg2 complexes using Ni-NTA affinity chromatography; and (iii) purification of the Tfg1-Tfg2 complexes by binding to IgG-Sepharose beads and liberating by TEV protease cleavage (see Materials and methods for details). The complexes were sufficiently pure after the IgG-Sepharose step that further purification using calmodulin beads was unnecessary (Fig. 2B, lane 3). Western blot analysis using antibodies specific to the recombinant Tfg1 C-terminus (anti-FLAG) or N-terminus (anti-His6) confirmed that the single Tfg1 species present in the purified M2 preparation corresponded to full-length protein (Fig. 2C, lanes 2, 4). Importantly, *in vitro* transcription assays demonstrated that the TAP purified Tfg1-Tfg2 complexes were as functional, if not slightly more active, than Tfg1-Tfg2 complexes purified using the previously reported method (Fig. 4A).

Tandem affinity purification of TFIIF complexes from *S. cerevisiae* whole cell extracts

As noted earlier, purification of functional *S. cerevisiae* holo-TFIIF has previously been achieved using a strain containing a hexahistidine-tagged Tfb1 subunit and five to six chromatographic steps in the purification scheme to reach apparent homogeneity [36,37]. More recently, Kornberg and colleagues generated *S. cerevisiae* strains containing TAP-tagged versions of the TFIIF subunits Tfb3 and Tfb4 in order to utilize TAP purification to resolve an apparent discrepancy in the subunit composition of *S. cerevisiae* and human TFIIF complexes [33]. Using a TAP purification scheme that involved ammonium sulfate precipitation and 500 mM ammonium sulfate washing during the IgG-agarose chromatographic step, the authors successfully determined a revised subunit structure for yeast TFIIF with complete conservation to that of the human factor. Significantly, however, the stringent purification conditions used leads to the underrepresentation of some polypeptides, the dissociation of core TFIIF from TFIIFK, and complexes that do not support transcription in reconstituted assays (data not shown). Using these Tfb3-TAP and Tfb4-TAP strains, we were able to establish conditions for the isolation of functional TFIIF and TFIIFK complexes (see Materials and methods for details). Using the Tfb3-TAP strain, pure TFIIFK was isolated from polyethyleneimine (PEI) treated cell lysates (Fig. 3C), whereas a mixture of holo-TFIIF and TFIIFK was isolated from untreated cell lysates (Fig. 3B). Using the Tfb4-TAP strain, pure core TFIIF was isolated from PEI treated cell lysates (Fig. 3E), whereas a mixture of holo-TFIIF and core TFIIF was isolated from untreated cell lysates (Fig. 3D). Importantly, TFIIF complexes isolated from PEI lysates supported robust *in vitro* transcription in reconstituted assays (Fig. 4B), and exhibited CTD kinase activity comparable to that observed with the purified TFIIFK (Fig. 4D).

Discussion

In this study, we sought to develop improved approaches for (i) the expression and purification of recombinant *S. cerevisiae* TFIIF (Tfg1-Tfg2) complexes and (ii) the purification of functional native TFIIF complexes from yeast extracts. In regards to the expression and purification of recombinant Tfg1-Tfg2 complexes, a major obstacle has been the extreme

toxicity conferred by the *TFG1* gene in *E. coli*. As noted earlier, we previously reported an approach that partially circumvented the toxicity and allowed for the production of functional recombinant Tfg1-Tfg2 complexes through the use of broad host range low-copy vector, induction of Tfg1, Tfg2 co-expression by infection with bacteriophage CE6 carrying T7 RNA polymerase, and subsequent multi-step affinity purification [39]. In an attempt to simplify the purification scheme, we fortuitously discovered that the addition of a TAP tag to the C-terminus of Tfg1 significantly reduced the toxicity and allowed for stable transformants of our new co-expression plasmid to be established in host strains such as Rosetta™ (DE3), which carries both IPTG-inducible T7 RNA polymerase and the pRARE plasmid to provide rare tRNAs for improved translation in *E. coli*. Utilizing this new system in combination with western blotting and primer extension analysis, we subsequently identified both a cryptic *E. coli* promoter and an alternative translation initiation site (Met-114) in the N-terminal coding region of *TFG1* (Fig. 1). As a result of these two functional elements, a predominant N-terminal Tfg1 truncation product is generated both constitutively (cryptic promoter functioning in combination with the internal initiation at Met-114) as well as after IPTG induction of the T7 expression promoter (internal initiation of the full-length T7 generated transcript at Met-114). Importantly, elimination of both these elements by site-directed mutagenesis coordinately and completely eliminated the production of the N-terminal truncation product and the toxicity of the *TFG1* gene (Fig. 1 and data not shown). Since the initial addition of a C-terminal TAP tag partially eliminated the toxicity of a *TFG1* gene that was generating the N-terminal truncation product, the combined results strongly suggest that the toxicity of the *TFG1* gene in *E. coli* is due to the N-terminal truncation product that contains a wild-type C-terminus. Although the precise target and mechanism for this toxicity remains unclear, it is noteworthy that the RAP74 human homolog of Tfg1 has been reported to contain a C-terminal RNAPII binding domain that is masked by an intramolecular interaction with the RAP74 N-terminal region [42]. Thus, it is tempting to speculate that Tfg1 may also contain a C-terminal RNAPII binding domain that is masked by an intramolecular interaction with the N-terminal region, that this RNAPII binding domain has the potential to interact with *E. coli* polymerase, and that the combination of the cryptic *E. coli* promoter and Met-114 translation initiation constitutively generates an N-terminal truncation product with the potential to do so and confer extreme toxicity. In any event, the elimination of this toxicity and the significantly enhanced production and purification of Tfg1-Tfg2 complexes reported here should prove invaluable for future structure and function studies of *S. cerevisiae* TFIIF.

In the second part of this work, we sought to develop an improved approach for the purification of functional native TFIIF complexes from yeast extracts. As discussed earlier, TFIIF purification has previously been achieved using five to six traditional chromatographic steps to reach apparent homogeneity [36,37]. Although *S. cerevisiae* strains containing TAP-tagged versions of the TFIIF subunits Tfb3 and Tfb4 have been generated and utilized to examine the composition of yeast TFIIF complexes [33], the stringent purification conditions that were utilized, involving both ammonium sulfate precipitation and high salt column washes, leads to the underrepresentation of some polypeptides and complexes that do not support transcription in reconstituted assays. By using these strains, eliminating the ammonium sulfate precipitation step, and both reducing and changing the salt during the TAP purification scheme, we have established conditions for the two step isolation of TFIIF complexes that are functional in both reconstituted transcription and RNAPII CTD kinase assays (Figs. 3 and 4). As is the case for TFIIF, the improved approach for the isolation of functional TFIIF complexes should prove invaluable for future mechanistic studies of this essential and centrally important transcription factor.

Acknowledgments

The authors thank Yuichiro Takagi and Roger Kornberg for providing yeast strains YT062 and YT063. This work was supported by a Public Health Service grant (GM51124) from the National Institutes of Health to A.S.P.

References

1. Flores O, Lu H, Killeen M, Greenblatt J, Burton ZF, Reinberg D. The small subunit of transcription factor IIF recruits RNA polymerase II into the preinitiation complex. *Proc Natl Acad Sci U S A* 1991;88:9999–10003. [PubMed: 1946469]
2. Maxon ME, Goodrich JA, Tjian R. Transcription factor IIE binds preferentially to RNA polymerase IIa and recruits TFIIF: a model for promoter clearance. *Genes Dev* 1994;8:515–524. [PubMed: 7926747]
3. Robert F, Douziech M, Forget D, Egly JM, Greenblatt J, Burton ZF, Coulombe B. Wrapping of promoter DNA around the RNA polymerase II initiation complex induced by TFIIF. *Mol Cell* 1998;2:341–351. [PubMed: 9774972]
4. Ghazy MA, Brodie SA, Ammerman ML, Ziegler LM, Ponticelli AS. Amino acid substitutions in yeast TFIIF confer upstream shifts in transcription initiation and altered interaction with RNA polymerase II. *Mol Cell Biol* 2004;24:10975–10985. [PubMed: 15572698]
5. Sun ZW, Hampsey M. Identification of the gene (SSU71/TFG1) encoding the largest subunit of transcription factor TFIIF as a suppressor of a TFIIB mutation in *Saccharomyces cerevisiae*. *Proceedings of the National Academy of Sciences of the United States of America* 1995;92:3127–3131. [PubMed: 7724527]
6. Yan Q, Moreland RJ, Conaway JW, Conaway RC. Dual roles for transcription factor IIF in promoter escape by RNA polymerase II. *J Biol Chem* 1999;274:35668–35675. [PubMed: 10585446]
7. Khapersky DA, Ammerman ML, Majovski RC, Ponticelli AS. Functions of *Saccharomyces cerevisiae* TFIIF during transcription start site utilization. *Mol Cell Biol* 2008;28:3757–3766. [PubMed: 18362165]
8. Tan S, Aso T, Conaway RC, Conaway JW. Roles for both the RAP30 and RAP74 subunits of transcription factor IIF in transcription initiation and elongation by RNA polymerase II. *J Biol Chem* 1994;269:25684–25691. [PubMed: 7929273]
9. Izban MG, Luse DS. Factor-stimulated RNA polymerase II transcribes at physiological elongation rates on naked DNA but very poorly on chromatin templates. *J Biol Chem* 1992;267:13647–13655. [PubMed: 1618865]
10. Zhang C, Burton ZF. Transcription factors IIF and IIS and nucleoside triphosphate substrates as dynamic probes of the human RNA polymerase II mechanism. *J Mol Biol* 2004;342:1085–1099. [PubMed: 15351637]
11. Chambers RS, Wang BQ, Burton ZF, Dahmus ME. The activity of COOH-terminal domain phosphatase is regulated by a docking site on RNA polymerase II and by the general transcription factors IIF and IIB. *J Biol Chem* 1995;270:14962–14969. [PubMed: 7797476]
12. Svejstrup JQ, Vichi P, Egly JM. The multiple roles of transcription/repair factor TFIIF. *Trends Biochem Sci* 1996;21:346–350. [PubMed: 8870499]
13. Schaeffer L, Roy R, Humbert S, Moncollin V, Vermeulen W, Hoeijmakers JH, Chambon P, Egly JM. DNA repair helicase: a component of BTF2 (TFIIH) basic transcription factor. *Science* 1993;260:58–63. [PubMed: 8465201]
14. Guzder SN, Sung P, Bailly V, Prakash L, Prakash S. RAD25 is a DNA helicase required for DNA repair and RNA polymerase II transcription. *Nature* 1994;369:578–581. [PubMed: 8202161]
15. Sung P, Prakash L, Matson SW, Prakash S. RAD3 protein of *Saccharomyces cerevisiae* is a DNA helicase. *Proc Natl Acad Sci U S A* 1987;84:8951–8955. [PubMed: 2827162]
16. Sung P, Prakash L, Weber S, Prakash S. The RAD3 gene of *Saccharomyces cerevisiae* encodes a DNA-dependent ATPase. *Proc Natl Acad Sci U S A* 1987;84:6045–6049. [PubMed: 2957691]
17. Sung P, Bailly V, Weber C, Thompson LH, Prakash L, Prakash S. Human xeroderma pigmentosum group D gene encodes a DNA helicase. *Nature* 1993;365:852–855. [PubMed: 8413672]

18. Bootsma D, Hoeijmakers JH. DNA repair. Engagement with transcription. *Nature* 1993;363:114–115. [PubMed: 8483493]
19. Schaeffer L, Moncollin V, Roy R, Staub A, Mezzina M, Sarasin A, Weeda G, Hoeijmakers JH, Egly JM. The ERCC2/DNA repair protein is associated with the class II BTF2/TFIIH transcription factor. *Embo J* 1994;13:2388–2392. [PubMed: 8194528]
20. Feaver WJ, Svejstrup JQ, Henry NL, Kornberg RD. Relationship of CDK-activating kinase and RNA polymerase II CTD kinase TFIIH/TFIIK. *Cell* 1994;79:1103–1109. [PubMed: 8001136]
21. Lu H, Zawel L, Fisher L, Egly JM, Reinberg D. Human general transcription factor IIH phosphorylates the C-terminal domain of RNA polymerase II. *Nature* 1992;358:641–645. [PubMed: 1495560]
22. Svejstrup JQ, Wang Z, Feaver WJ, Wu X, Bushnell DA, Donahue TF, Friedberg EC, Kornberg RD. Different forms of TFIIH for transcription and DNA repair: holo-TFIIH and a nucleotide excision repairosome. *Cell* 1995;80:21–28. [PubMed: 7813015]
23. Feaver WJ, Huang W, Gileadi O, Myers L, Gustafsson CM, Kornberg RD, Friedberg EC. Subunit interactions in yeast transcription/repair factor TFIIH. Requirement for Tfb3 subunit in nucleotide excision repair. *J Biol Chem* 2000;275:5941–5946. [PubMed: 10681587]
24. Coin F, Oksenyich V, Mocquet V, Groh S, Blattner C, Egly JM. Nucleotide excision repair driven by the dissociation of CAK from TFIIH. *Mol Cell* 2008;31:9–20. [PubMed: 18614043]
25. Henry NL, Campbell AM, Feaver WJ, Poon D, Weil PA, Kornberg RD. TFIIH-TAF-RNA polymerase II connection. *Genes Dev* 1994;8:2868–2878. [PubMed: 7995524]
26. Cairns BR, Henry NL, Kornberg RD. TFG/TAF30/ANC1, a component of the yeast SWI/SNF complex that is similar to the leukemogenic proteins ENL and AF-9. *Mol Cell Biol* 1996;16:3308–3316. [PubMed: 8668146]
27. Treich I, Carlson M. Interaction of a Swi3 homolog with Sth1 provides evidence for a Swi/Snf-related complex with an essential function in *Saccharomyces cerevisiae*. *Molecular and cellular biology* 1997;17:1768–1775. [PubMed: 9121424]
28. Eberharter A, John S, Grant PA, Utley RT, Workman JL. Identification and analysis of yeast nucleosomal histone acetyltransferase complexes. *Methods* 1998;15:315–321. [PubMed: 9740719]
29. Steger DJ, Eberharter A, John S, Grant PA, Workman JL. Purified histone acetyltransferase complexes stimulate HIV-1 transcription from preassembled nucleosomal arrays. *Proceedings of the National Academy of Sciences of the United States of America* 1998;95:12924–12929. [PubMed: 9789016]
30. John S, Howe L, Tafrov ST, Grant PA, Sternglanz R, Workman JL. The something about silencing protein, Sas3, is the catalytic subunit of NuA3, a yTAF(II)30-containing HAT complex that interacts with the Spt16 subunit of the yeast CP (Cdc68/Pob3)-FACT complex. *Genes Dev* 2000;14:1196–1208. [PubMed: 10817755]
31. Shen X, Mizuguchi G, Hamiche A, Wu C. A chromatin remodelling complex involved in transcription and DNA processing. *Nature* 2000;406:541–544. [PubMed: 10952318]
32. Kabani M, Michot K, Boschiero C, Werner M. Anc1 interacts with the catalytic subunits of the general transcription factors TFIID and TFIIF, the chromatin remodeling complexes RSC and INO80, and the histone acetyltransferase complex NuA3. *Biochem Biophys Res Commun* 2005;332:398–403. [PubMed: 15896708]
33. Takagi Y, Komori H, Chang WH, Hudmon A, Erdjument-Bromage H, Tempst P, Kornberg RD. Revised subunit structure of yeast transcription factor IIH (TFIIH) and reconciliation with human TFIIH. *J Biol Chem* 2003;278:43897–43900. [PubMed: 14500720]
34. Ranish JA, Hahn S, Lu Y, Yi EC, Li XJ, Eng J, Aebersold R. Identification of TFB5, a new component of general transcription and DNA repair factor IIH. *Nat Genet* 2004;36:707–713. [PubMed: 15220919]
35. Giglia-Mari G, Coin F, Ranish JA, Hoogstraten D, Theil A, Wijgers N, Jaspers NG, Raams A, Argentini M, van der Spek PJ, Botta E, Stefanini M, Egly JM, Aebersold R, Hoeijmakers JH, Vermeulen W. A new, tenth subunit of TFIIH is responsible for the DNA repair syndrome trichothiodystrophy group A. *Nat Genet* 2004;36:714–719. [PubMed: 15220921]
36. Svejstrup JQ, Feaver WJ, LaPointe J, Kornberg RD. RNA polymerase transcription factor IIH holoenzyme from yeast. *J Biol Chem* 1994;269:28044–28048. [PubMed: 7961739]
37. Myers LC, Leuther K, Bushnell DA, Gustafsson CM, Kornberg RD. Yeast RNA polymerase II transcription reconstituted with purified proteins. *Methods* 1997;12:212–216. [PubMed: 9237165]

38. Rani PG, Ranish JA, Hahn S. RNA polymerase II (Pol II)-TFIIF and Pol II-mediator complexes: the major stable Pol II complexes and their activity in transcription initiation and reinitiation. *Mol Cell Biol* 2004;24:1709–1720. [PubMed: 14749386]
39. Ziegler LM, Khaperskyy DA, Ammerman ML, Ponticelli AS. Yeast RNA polymerase II lacking the Rpb9 subunit is impaired for interaction with transcription factor IIF. *J Biol Chem* 2003;278:48950–48956. [PubMed: 14522989]
40. Gerischer U, Durre P. mRNA analysis of the *adc* gene region of *Clostridium acetobutylicum* during the shift to solventogenesis. *J Bacteriol* 1992;174:426–433. [PubMed: 1370288]
41. Ponticelli AS, Struhl K. Analysis of *Saccharomyces cerevisiae his3* transcription *in vitro*: Biochemical support for multiple mechanisms of transcription. *Mol Cell Biol* 1990;10:2832–2839. [PubMed: 2188101]
42. Wang BQ, Burton ZF. Functional domains of human RAP74 including a masked polymerase binding domain. *J Biol Chem* 1995;270:27035–27044. [PubMed: 7592953]

Fig. 1.

The N-terminal coding region of *S. cerevisiae* *TFG1* contains a functional *E. coli* promoter and undergoes internal translation initiation at Met-114. (A) Western blot analysis. Bacterial lysates from *E. coli* strains harboring the wild-type Tfg1-Tfg2 co-expression vector (WT) or the indicated M1 or M2 site-directed Tfg1 variants were prepared prior to induction (– IPTG) or post induction (+ IPTG) and analyzed using an anti-FLAG (Tfg1 C-terminal) antibody. The positions of the full-length Tfg1 protein and the N-terminal truncation product due to internal translation initiation at Met-114 are indicated by the arrows. (B) Primer extension analysis. Total *E. coli* RNA (15 µg) was prepared from the indicated strains and analyzed as described in materials and methods. The transcript detected in lane 1 is generated by the cryptic *E. coli* promoter located in the Tfg1 N-terminal region; the major transcripts in lanes 4–6 are generated by the inducible T7 promoter. (C) DNA sequences from wild-type Tfg1 and the M1 and M2 site-directed variants. Shown are the sequences between Arg-55 and Met-114. The cryptic *E. coli* promoter is indicated by the shaded box, the generated transcript by the arrowhead, and the site-directed mutations introduced in the M1 and M2 variants are underlined.

Fig. 2.

Purification of recombinant *S. cerevisiae* Tfg1-Tfg2 complexes. (A) Plasmid pDt/g1*g2 used for the co-expression of recombinant yeast TFIIF. The g1* indicates the presence of the M2 variant Tfg1 sequence. (B) Purification of recombinant yeast TFIIF. Shown is an SDS-8% polyacrylamide gel stained with Gelcode Blue containing molecular mass standards (M); the eluant of wild-type yeast TFIIF from sequential IgG and CaM column chromatography (1 μ g, lane 1); the eluant of wild-type yeast TFIIF from sequential Ni-NTA and IgG column chromatography (1 μ g, lane 2); and the eluant of the M2 variant yeast TFIIF from sequential Ni-NTA and IgG column chromatography (500 ng, lane 3). The mobilities of the Tfg1 and Tfg2 subunits and their truncated fragments are indicated by the arrows. (C) Western blot

analysis. Eluants (120 ng) of wild-type yeast TFIIIF (lanes 1, 3) or the M2 variant yeast TFIIIF (lanes 2, 4) from sequential Ni-NTA and IgG column chromatography were analyzed using an anti-FLAG (Tfg1 C-terminal, lanes 1, 2) or anti-His6 (Tfg1 N-terminal, lanes 3, 4) antibody.

Fig. 3. Tandem affinity purification of TFIIH complexes from *S. cerevisiae* whole cell extracts. (A) Diagram for the composition of holo-TFIIH and sub-complexes. (B-E) Purification of yeast holo-TFIIH and sub-complexes. Shown are SDS-10% polyacrylamide gels stained with silver (BioRad) containing molecular mass standards and 0.4 – 1.0 μ g of the eluants from the IgG and calmodulin columns for the Tfb3-TAP (B,C) and Tfb4-TAP (D,E) strain lysates with or without polyethyleneimine treatment.

Fig. 4.

Functional assays of recombinant yeast TFIIF and native yeast TFIIH complexes. (A) *In vitro* transcriptional activity of recombinant TFIIF. Reconstituted transcription assays were carried out as described in materials and methods and contained no added TFIIF (lane 1), 250 ng of TAP purified TFIIF from pDt/g1*g2 (lane 2), or 250 ng of TFIIF purified as described previously (lane 3) [39]. (B) *In vitro* transcriptional activity of TFIIH complexes. Reconstituted transcription assays were carried out as described in materials and methods using 250 ng TAP purified TFIIF (M2 variant, designated as TFIIF*) and: no TFIIH (lane 2); 160 ng of a mixture of holo-TFIIH and TFIIK (lane 3, preparation from Tfb3-TAP strain – PEI); 50 ng TFIIK (lane 4, preparation from Tfb3-TAP strain + PEI); 160 ng of a mixture of holo-TFIIH and core TFIIH

(lane 5, preparation from Tfb4-TAP strain – PEI); or 130 ng core TFIIH (lane 6, preparation from Tfb4-TAP strain + PEI). (C) CTD kinase activities of TFIIH complexes. Kinase assays were performed as described in materials and methods and contained 500 ng purified *S. cerevisiae* RNAPII (IIA, hypo-phosphorylated form) and no TFIIH (lane 1), 160 ng TFIIH (lanes 2–4; Tfb3-TAP, –PEI) 36 ng TFIIK (lanes 5–7; Tfb3-TAP, +PEI) and 200 uM dATP or ATP as indicated. RNAPII CTD-ser5 phosphorylation was detected by western blotting using H14 antibody.