Published in final edited form as:

Methods. 2019 April 15; 159-160: 96–104. doi:10.1016/j.ymeth.2019.03.001.

In vitro assembly and proteomic analysis of RNA polymerase II complexes

Yoo Jin Joo¹, Scott B. Ficarro^{2,3}, Jarrod A. Marto^{2,3}, and Stephen Buratowski^{*,1}

¹Department of Biological Chemistry and Molecular Pharmacology, Harvard Medical School, Boston, MA

²Department of Cancer Biology Dana Farber Cancer Institute, Harvard Medical School, Boston, MA

³Blais Proteomics Center, Dana Farber Cancer Institute, Harvard Medical School, Boston, MA

Abstract

The RNA polymerase II (RNAPII) transcription cycle consists of multiple steps involving dozens of protein factors. Here we describe a useful approach to study the dynamics of initiation and early elongation, comprising an in vitro transcription system in which complexes are assembled on immobilized DNA templates and analyzed by quantitative mass spectrometry. This unbiased screening system allows quantitation of RNAPII complex components on either naked DNA or chromatin templates. In addition to transcription, the system reproduces co-transcriptional mRNA capping and multiple transcription-related histone modifications. In combination with other biochemical and genetic methods, this approach can provide insights into the mechanistic details of gene expression by RNAPII.

1. Introduction

Transcription is a key step in regulation of gene expression, fundamentally impacting almost all cellular processes. At various times, the transcription machinery interacts with RNA processing enzymes, histone modifiers, chromatin remodelers, and DNA repair factors [1–3]. These co-transcriptional processes can in turn affect concurrent and subsequent rounds of transcription, so it is therefore important to understand the cross-talk between these elements of gene expression. Full understanding requires defining the repertoire and dynamics of all the factors that participate in the initiation, elongation, and termination phases of the transcription cycle.

Multiple approaches have been used to identify factors involved in RNAPII transcription. Some transcription factor genes were first cloned based on mutant phenotypes in yeast [4–8], while a more complete set of RNAPII factors was defined by biochemical purification

*Corresponding author and lead contact, Tel: 1-617-432-0696, steveb@hms.harvard.edu.

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.

and reconstitution of transcription in vitro. Some factors were first discovered by direct physical associations with RNAPII, TATA-Binding Protein, or other key components of the gene expression machinery [9–12]. Details of interactions between factors and RNAPII have been visualized by X-ray crystallography [13], and more recently by cryo-electron microscopy and crosslink-based interaction mapping [14–17].

Although the basal transcription factors and RNAPII alone can assemble a functional pre-initiation complex (PIC) on a core promoter, maximal in vitro activity and transcription in vivo is additionally dependent on transcription activators and co-activators. Transcription activators bind specific regulatory promoter sequences that are usually located upstream from the core promoter, and so referred to as ‘upstream activation sequences (UASs)’ in yeast or ‘enhancers’ in higher eukaryotes. Activators have two major functions. First, they recruit histone acetyltransferases and chromatin remodelers to displace nucleosomes blocking the core promoter. Second, they recruit RNAPII via contacts with the Mediator co-activator complex or TFIID subunits. Therefore, a complex combination of DNA-protein and protein-protein interactions between DNA and transcription factors occurs during gene expression [18].

Post-translational modifications (PTMs) also have critical functions in RNAPII transcription. The best characterized is phosphorylation on the C-terminal domain (CTD) of the RNAPII subunit Rpb1. The CTD consists of many tandem repeats of the seven amino acid sequence (Y₁S₂P₃T₄S₅P₆S₇), and differential phosphorylation of residues during the transcription cycle creates a ‘CTD code’ by which specific CTD binding proteins are recruited at different stages [19]. CTD code ‘readers’ include enzymes needed for mRNA processing, termination, and histone modifications. Other phosphorylations promote elongation: P-TEFb/CDK9/Burl kinase phosphorylates DSIF/Spt5 to recruit Paf1C [20, 21] and also the Rpb1 linker domain to recruit Spt6 [22, 23]. Acetylations and methylations, particularly on histones, also regulate transcription [3, 24]. For example, bromodomains in the transcription factors TFIID, SAGA, p300/CBP, Swi/Snf, and the BRD proteins bind acetylated histones enriched near 5’ ends of genes, concentrating them near promoters.

We sought an in vitro system that would provide the following capabilities:

1. Identification and quantitation of transcription factors associated with the DNA template.
2. Time resolution to study dynamics of RNAPII complexes at specific stages of transcription.
3. The ability to monitor post-translational modifications.

Building on work from several labs [25, 26], we established an in vitro transcription system that meets these goals (Fig. 1). The system consists of the following parts:

1. Yeast nuclear extract, which contains a more complete set of factors than purified systems and therefore may better reproduce the in vivo environment.
2. A transcription template immobilized on magnetic beads, either as naked DNA or assembled into chromatin.

3. Reaction conditions that can enrich for PICs or elongation complexes in a time-resolved manner.
4. Analysis of template-associated proteins using multiplexed quantitative proteomic approaches and immunoblotting.

This in vitro system reproduces many aspects of the transcription cycle, including CTD phosphorylations, co-transcriptional mRNA capping, and several histone modifications. It nicely complements other approaches and can facilitate a better understanding of the molecular mechanisms of gene expression.

2. Preparation of components required for in vitro transcription on chromatin templates.

2.1. Yeast nuclear extract preparation.

Yeast nuclear extract was prepared as described previously with some modifications [18]. To minimize protein degradation and maintain robust transcription activity, a *S. cerevisiae* strain deleted for three protease genes (*PEP4*, *PRB1*, and *PRC1*) was used [27]. All buffers throughout the procedure contain 1 × protease inhibitors (1 mM PMSF, 1 µg/ml Aprotinin, 1 µg/ml Leupeptin, 1 µg/ml Pepstatin A, and 1 µg/ml Antipain) to further block proteolysis. Cells were grown in 8 liters of YPD medium containing 3% dextrose at 30°C to an optical density at 600 nm of 2. After Tris/DTT preincubation and wash, cells were resuspended in 40 ml YPD/1 M sorbitol supplemented with 15 mg of Zymolyase 100T (Seikagaku Corp.). Digestion was performed at 30°C with gentle shaking (120 rpm) until ~90% spheroplast formation, as monitored by phase contrast microscopy in 0.5% SDS. Spheroplasting usually takes around 1 hour, but varies between strains. Incomplete spheroplast formation leads to inefficient lysis and poor activity. However, spheroplast formation should be checked repeatedly to avoid overtreatment, which also reduces recovered transcription activity.

Spheroplasts were lysed by three slow passages through a motor-driven (Caframo) glass tissue grinder homogenizer (Wheaton/VWR 62400–802, 55 ml with 0.1 to 0.15 mm clearance smooth pestle). Cell debris and unlysed spheroplasts were removed by centrifuging twice at 3,000× g for 8 min and twice at 3,000× g for 5 min, recovering the viscous supernatant at each step. Nuclei were then pelleted by centrifugation at 20,000× g for 30 min. Nuclei were resuspended, and then lysed and extracted by adjusting the supernatant to 0.5 M ammonium sulfate (pH 7.6). After a spin at 144,000× g for 90 min with type 70Ti rotor (Beckman), the clear soluble fraction was adjusted to 75% saturation with ammonium sulfate, and the precipitates were pelleted. Then, the pellets were suspended and dialyzed against Dialysis Buffer (20 mM HEPES [pH 7.6], 10 mM MgSO₄, 75 mM ammonium sulfate, 1 mM EGTA, 20% glycerol, 3 mM DTT, 1 × protease inhibitors). Nuclear extracts with good transcription activity typically have 30–50 mg/ml protein concentration; lower concentration extracts are generally less active. Aliquots of the extract were flash frozen in liquid nitrogen and stored at –80°C. Typically, up to 400 mg of nuclear extract is recovered from 8 liters of yeast culture. Specific transcription activity was tested in vitro as described [28].

2.2. Purification of transcription activators from bacteria.

To enhance in vitro transcription, the recombinant transcription activators Gal4-VP16 and Gal4-Gcn4 were used as described [18, 29]. Proteins were expressed from plasmids pRJR-Gal4vp16 and pSH556 in *E. coli* expression strain BL21(DE3), and purified as described previously with several modifications [30, 31]. Bacterial cells in exponential growth stage were induced with 0.1 mM IPTG for 3 hours at 25°C. The hexahistidine-tagged proteins in soluble bacterial extracts were purified using Ni²⁺-NTA purification (Qiagen). The recombinant proteins were then further purified with a Mono Q (HR 5/5) column (GE Healthcare). Fractions from a linear gradient of 0.1–1.0 M NaCl in an Elution Buffer (20 mM HEPES-KOH [pH 7.5], 1 mM EDTA, 1 mM DTT, 20% glycerol, and 1 × protease inhibitors) were monitored by SDS-PAGE and Coomassie blue staining (Fig. 2C). Both proteins were dialyzed against 20 mM HEPES-KOH (pH 7.6), 500 mM K-acetate, 20% glycerol, 1mM EDTA, 10 μM ZnCl₂ and stored in aliquots at –80°C. The Gal4 binding domain binds zinc, so 10 μM ZnCl₂ was included in the growth media and buffers during purification to preserve activity.

2.3. Linear template DNA preparation.

Linear DNA was used as a template for the in vitro transcription and reconstitution reaction (Fig. 2A, B). For purification of PICs, a DNA fragment was amplified from plasmid pSH515 [32] containing a single Gal4 binding site upstream of the *HIS4* core promoter. For EC purification, DNA template was PCR amplified from pUC18-G5CYC1 G-[18], which contains five Gal4 binding sites and the *CYC1* core promoter upstream of an approximately 300 nucleotide G-less cassette. For both templates, the primer upstream of the promoter has a 5' biotin modification for immobilization on beads. After PCR amplification with Platinum Taq DNA polymerase (Invitrogen), primer dimers and small side products were removed using the DNA SizeSelector kit (Aline Biosciences). If multiple PCR products are generated, an additional gel extraction of the correct fragment is necessary.

3. In vitro chromatin reconstitution.

3.1. Expression and purification of yeast histones.

All four *S. cerevisiae* histones were purified from *E. coli* as described previously, with minor modifications [26]. Plasmids containing histone subunits (pRET-3A-H2A/F1038, pRET-3A-H2B/F1039, pRET-3A-H3/F1040, and pET28b-H4/F723) were transformed into bacterial expression strains. BL21(DE3)pLysS was used for H2A, H2B, and H3 expression, and BL21-CodonPlus(DE3)-RIL was used for H4 expression. For each histone, three liters of cells were cultured. Expression was induced by addition of 0.5 mM IPTG for 4 hours (H3 and H4) or 5 hours (H2A and H2B) at 37°C. Denatured histones were extracted from inclusion bodies with Unfolding Buffer (20 mM Tris-HCl [pH 7.5], 7 M guanidine hydrochloride, 10 mM DTT). The histones were further purified by passage through 5 ml Q Sepharose Fast Flow HR10/10 column (GE Healthcare), followed by binding and elution from a 5 ml SP Sepharose Fast Flow HR10/10 column (GE Healthcare) with buffer containing 600 mM NaCl, and finally dialyzed against 5 mM β-mercaptoethanol (β-ME) and 0.2 mM PMSF. After measuring protein concentration, aliquots were lyophilized and

stored at -80°C until needed for histone octamer reconstitution. The purity of the sample was confirmed using 16% SDS-PAGE and Colloidal Blue staining (Invitrogen).

3.2. Reconstitution of histone octamer.

Octamer reconstitution with the four yeast histones was performed essentially as described for *Xenopus* histones with some modification [33]. Equimolar amounts of H3 and H4 and 1.2-fold excess of H2A and H2B were mixed in Unfolding Buffer and assembled by dialyzing against Refolding Buffer (10 mM Tris-HCl [pH 7.5], 2 M NaCl, 1 mM EDTA, 5 mM β -ME, 0.2 mM PMSF). The mixture was concentrated to 0.5 ml with a Centrprep column (Millipore) that was pre-blocked with bovine serum albumen (BSA) before use. Refolded histones were separated on a Superose 12 (GE Healthcare) gel filtration column with 0.5 ml/min flow rate. Fractions were monitored by Colloidal Blue staining of a 16% SDS PAGE gel (Fig. 2D, left panel). Fractions contained an equimolar ratio of four histones were combined and concentrated with BSA-blocked Centricon 10 (Millipore). Octamer concentration was determined by measuring optical density at 276 nm. The sample was diluted to 2 M NaCl, 50% glycerol, aliquoted and stored in -20°C .

3.3. Purification of yeast histone chaperone, Nap1.

Yeast nucleosome assembly protein, Nap1, was expressed and purified from *E. coli*. The plasmid expressing GST tagged *S. cerevisiae* NAP1 (pGEX-6P-1-NAP1/F1034) was transformed into BL21-CodonPlus(DE3)-RIL cells. Protein expression was induced with 0.2 mM IPTG for 4 hours at 30°C . After cell collection and lysis, soluble GST-Nap1 was affinity purified using Glutathione Sepharose 4B beads (GE Healthcare) and eluted by digestion with 20 units Prescission protease (GE Healthcare) in 0.5 ml Prescission Protease Cleavage Buffer (50 mM Tris-HCl [pH 7.0], 150 mM NaCl, 1 mM EDTA, 1 mM DTT) for 16 hours at 4°C . The eluate was further purified by chromatography on a HR 5/10 Mono Q column (GE Healthcare). Fractions were assayed by 10% SDS PAGE and Coomassie brilliant blue staining (Fig. 2D, middle panel) and those with relatively pure Nap1 were combined and dialyzed against Buffer R (10 mM HEPES-KOH [pH7.5], 10 mM KCl, 1.5 mM MgCl_2 , 0.5 mM EGTA, 10 % Glycerol, 2.5 mM β -glycerophosphate, 0.2 mM PMSF, 1 mM DTT). After measuring protein concentration, Nap1 protein was aliquoted and stored at -80°C . In the assay for chromatin assembly (Fig. 2E), MNase digestion reactions did not produce the expected nucleosomal fragment sizes in the absence of Nap1 (not shown), and overall MNase digestion was strongly inhibited by nonspecific interactions between histones and DNA.

3.4. Purification of yeast Iws1a complex.

To produce properly spaced nucleosome arrays, *S. cerevisiae* Isw1a complex was purified from an Ioc3-TAP strain (YF2164) using tandem affinity purification modified as reported previously [34, 35]. The salt concentration in all buffers should be 300 mM KCl instead 150 mM NaCl. Briefly, total extract from 4 liters of yeast culture was prepared. For each purification, 0.8 ml IgG Sepharose beads (GE Healthcare) and 0.4 ml Calmodulin Sepharose 4B beads (GE Healthcare) were used. Sample purity was checked by silver staining of a 10% SDS-PAGE (Fig. 2D, right panel). After measuring protein concentration, the purified complex was aliquoted and stored at -80°C . ATP-dependent chromatin remodeling activity

was confirmed using partial MNase digestion assay, where the periodic cutting pattern indicative of positioned nucleosomes was dependent on Isw1a (Fig. 2E and data not shown).

3.5. Reconstitution of chromatin with linear DNA.

In vitro chromatin reconstitution was performed as described previously, with minor modifications [36]. Efficiency of assembly varies between different preparations of DNA template, so it is strongly recommended that optimal conditions be first determined by titrating the DNA:histone octamer ratio in 5% increments. Test tubes used in the reaction should be blocked with 2 mg/ml BSA and 0.1% NP-40 before use. In the 1× baseline reaction, the protein components (0.9 µg [8 pmol] histone octamer, 2 µg [30 pmol] Nap1, and 30 ng [135 fmol] Isw1a complex) were assembled in chromatin reconstitution buffer (10 mM HEPES-KOH [pH 7.5], 10 mM KCl, 5.5 mM MgCl₂, 0.5 mM EGTA, 10 % glycerol, and 2.5 mM β-glycerophosphate, 0.2 mM PMSF, 1 mM DTT) in a volume of 90 µl and preincubated for 30 min on ice. The reaction was then brought to 110 µl total reaction volume by adding the linear DNA template (centered at 660 ng [1.5 pmol]) together with 3 mM ATP, 700 ng creatine phosphokinase, and 30 mM phosphocreatine. The assembly reaction was allowed to proceed for 5 hours at 30°C. Nucleosome assembly was confirmed with a partial MNase digestion assay as described [37] or with a gel shift assay on 4% acrylamide gel as shown elsewhere [18]. Properly assembled chromatin will show a MNase ladder of ~145–165 bp nucleosomal fragments (Fig. 2E, showing the maximum four nucleosomes that can fit on this DNA fragment). We note that circular plasmid templates are more efficiently chromatinized than linear DNA with this protocol. Once optimal conditions are determined, the reaction can be scaled up as needed. For larger amounts of template, we prefer to do multiple 1× reactions rather than scaling up within a single test tube. Nucleosomal templates are assembled fresh just before use (the five hours 30°C incubation, followed by a 4°C storage, can be done overnight in a PCR machine).

4. In vitro assembly of RNAPII complexes on immobilized templates.

4.1. Isolation of pre-initiation complexes (PICs).

For RNAPII PIC analysis (Fig. 3A), complexes were assembled on immobilized *HIS4* template (Fig. 2A) as described previously, with minor modifications [18]. To couple DNA to magnetic beads, 660 ng of naked or chromatinized templates (the 1× reaction from Section 3.5) were mixed with an equal volume of 2× Binding Buffer (30 mM HEPES [pH 7.6], 300 mM NaCl, 8% PEG 8000, 20 mM EDTA, 0.04% NP40, 10% glycerol, 2.5 mM β-glycerophosphate, 0.2 mM PMSF, 1 mM DTT) containing 15 µl Dynabeads Streptavidin T1 slurry (10 µg/µl, Invitrogen). This mixture was incubated for one hour at 25°C with rotation. Coupled beads were magnetically purified, and the supernatant retained to test for efficient conjugation (assayed by ethanol precipitation and agarose gel electrophoresis to confirm depletion of free DNA). To reduce background binding, the template-coupled beads were blocked in 150 µl Transcription Buffer (20 mM HEPES-KOH [pH 7.6], 100 mM K-acetate, 1 mM EDTA, 5 mM Mg-acetate) supplemented with 60 mg/ml casein, 5 mg/ml polyvinylpyrrolidone, and 2.5 mM DTT for 30 min at 25°C. After washing three times in 400 µl Transcription Buffer, the immobilized templates were resuspended in 30 µl Transcription Buffer.

The 1× PIC assembly reaction described in this paragraph is sufficient for monitoring complexes by immunoblotting or in vitro transcription. However, for mass spectrometry analysis, we typically do a 6–10× scale-up. In the 1× reaction, 200–400 ng Gal4-VP16 or Gal4-Gcn4 activator (125–250 nM final concentration in the 60 µl reaction) was pre-incubated with 220 ng immobilized templates (10 µl immobilized template-beads mixture in Transcription Buffer) for 5 min at 25°C before the addition of other components. A “no activator” negative control reaction was performed in parallel with the same volume of the Gal4-VP16 dialysis buffer. The yeast nuclear extract supplies all other protein components of the RNAPII complex. Yeast nuclear extract (0.5–1 mg) and remaining components were mixed with template in a total reaction volume of 60 µl Transcription Buffer supplemented with 20 units RNasin (Promega), 1.32 µg creatine phosphokinase, 10 mM phospho-creatine, 0.03% NP-40, 2 µg *S. cerevisiae* tRNA (Sigma), 160 µM S-adenosyl-L-methionine (SAM, Sigma), 20 µM acetyl coenzyme A (Acetyl-CoA, Sigma), 1 µg Hae III-digested *E. coli* genomic DNA, and 1 × protease inhibitors. We find that the yeast tRNA and bacterial genomic DNA help reduce non-specific interactions of nucleic acid binding proteins or other basic proteins with the immobilized DNA template.

After a 45 min standing incubation at 25°C, the immobilized templates and bound proteins were isolated using a magnetic stand and washed three times with 600 µl of Wash Buffer (Transcription Buffer complemented with 0.05% NP40 and 2.5 mM DTT). The immobilized DNA templates and bound proteins were released from the beads by incubation with 60 units Pst I-HF (New England Biolabs) in 40 µl of Transcription Buffer for 30 min at 25°C with rotation. The supernatant was carefully decanted and cleaned up once again in the magnetic stand to remove residual magnetic beads.

For quantitative mass spectrometry analysis, the immobilized template reaction was scaled up 6-fold within a single test tube, producing a final elution volume of 240 µl. A 40 µl aliquot was used for silver staining and to confirm the RNAPII complex formation by immunoblotting for enrichment of RNAPII subunit Rpb1 (Fig. 3B). Note the lack of differential binding in the silver stained gel, demonstrating that background proteins dominate the eluates (also see below). After confirmation of activator-dependent RNAPII enrichment, the remaining elution fraction (200 µl) can be stored at –80°C until quantitative mass spectrometry analysis as described in Section 5.1.

Histone modifications influence transcription, so the chromatin template eluates were tested for histone acetylations and H3 K36 methylation. Yeast nuclear extracts apparently contain limiting levels of SAM and acetyl-CoA, as adding these cofactors produced higher levels of histone modifications (Fig. 3C). Testing the effect of chromatin on in vitro transcription showed that nucleosomes suppressed transcription as expected, and revealed that acetyl-CoA, but not SAM, stimulated transcription activation (Fig. 3D). This indicates that our in vitro system recapitulates transcription coactivation by histone acetyltransferases, consistent with previous in vivo studies [3]. Accordingly, all reactions here were performed with the added cofactors.

4.2. Purification of ECs.

EC purification on immobilized templates was performed using conditions slightly modified from the PIC protocol (Fig. 4A). Because ECs are transient and change rapidly, it is necessary to synchronize the *in vitro* transcription reaction to provide sufficient material for quantitative mass spectrometry. To this end, EC isolation was performed on a template carrying a 5×Gal4 binding site-*CYC1* promoter upstream of a G-less cassette (Fig. 2A). After PIC formation, synchronized transcription was initiated by addition of three NTPs (ATP, CTP, UTP) and the chain terminator (3'-O-me-GTP). This combination allows elongation until the ECs transiently stall at the end of the G-less cassette.

For a 1× experiment, the naked or chromatinized templates (220 ng) were immobilized onto 5 µl Dynabeads, blocked, and washed as described above for PICs. The immobilized templates were resuspended in 10 µl Transcription Buffer and Gal4-VP16 (400 ng) was pre-bound the immobilized templates for 5 min at 25°C. PICs were pre-assembled for 30 min at 25°C as in Section 4.1, and then transcription was initiated by adding 2.4 µl NTPs mix (final concentration of 400 µM each ATP, CTP, and UTP, and 40 µM 3'-Ome-GTP) at 25°C. We recently reported that CTD phosphorylation (Ser5P to Ser2P) and EC components change from early to late ECs as a function of time in our *in vitro* system, so the NTP incubation time can be varied between 1 to 16 min depending on the specific purpose of the experiment [38]. In the EC data analyzed below, the transcription reaction was performed for 2 minutes. The immobilized templates were recovered with a magnetic stand, and the reaction was stopped by two quick rinses of the pelleted beads in 400 µl Wash Buffer.

The transcription template has a unique Ssp I site located between the TATA box and transcription start site (Fig. 2A), so for the cleanest enrichment of ECs free of promoter-bound proteins, elution was performed with 60 units Ssp I-HF (New England Biolabs) in 40 µl Transcription Buffer for 30 min at 25°C with rotation. Otherwise, total DNA bound proteins (including PIC components remaining at the promoter) can be eluted using the Pst I-HF site located upstream from the activator binding sites (Fig. 2A). For quick analysis by SDS-PAGE and immunoblotting, it is also possible to simply boil the beads in loading buffer, although negative control background signals will depend on how much the analyzed factor binds non-specifically to the beads. As for quantitative mass spectrometry analysis of PICs, a 6-fold scaled up reaction was used for each EC quantitative mass spectrometry experiment, eluted in a final volume of 240 µl. An aliquot of the elution fraction (40 µl) was subjected to immunoblotting assay and silver staining (Fig. 4B). Total background proteins were dominant and equal throughout all samples, but RNAPII was significantly enriched by NTPs on the downstream DNA, as confirmed with an antibody against Rpb1 (8WG16). The remaining elution fraction (200 µl) was stored at -80°C until quantitative mass spectrometry, pending validation of RNAPII enrichment by immunoblot.

5. Quantitative mass spectrometry of the purified RNAPII complexes.

5.1. Preparation of samples for mass spectrometry

The elution fraction from the immobilized template binding assay (200 µl) was subjected to quantitative mass spectrometry as described previously, with modifications [38]. For trypsin

digestion, ammonium bicarbonate was added to a final concentration of 100 mM. To disrupt disulfides, proteins were reduced with 10 mM DTT at 56°C for 1 hour and cysteines were then alkylated with 22.5 mM iodoacetamide (IAA) for 45 min at 25°C in the dark. Digestion with trypsin (4 µg) was performed at 37°C for 16 hours. Digested peptides were purified on a C18 solid phase extraction cartridge (SOLA-RP, ThermoFisher Scientific). After the initial peptide desalting, a second clean-up was performed using paramagnetic beads (SP3 clean-up method) as described [39].

Isobaric tagging for quantitation of peptides can be accomplished using commercially available iTRAQ (Thermo Scientific) or TMT (Thermo Scientific) reagents. Our recent work has primarily used the 6-plex TMT reagents due to their additional multiplexing capacity. For TMT labeling, the peptides were reconstituted in 20 µl 50 mM triethylammonium bicarbonate (TEAB) [pH 8.0]. Each sample was incubated with 8 µL (1/5th unit, or 0.16 mg) of a unique TMT reagent (solubilized in 100% anhydrous acetonitrile) for 1 hour at 25°C.

To assess labeling efficiency, small aliquots of each reaction (5%) were combined and analyzed by tandem mass spectrometry with a 1-hour reverse phase gradient, with iTRAQ or TMT modification listed as a variable modification in the search program Mascot. If the ratio of fully labeled to total PSMs (peptide spectral matches) was <95%, labeling was deemed incomplete and peptides were subjected to another round of labeling as described above. If labeling was >95%, the remaining samples (stored during the interim at -80°C) were quenched with 5% hydroxylamine in 100 mM TEAB for 15 min at 25°C, combined and then desalted using SP3. Dried, desalted peptides were stored at -80°C until MS analysis.

5.2. Analyzing the samples with 3D LC-MS/MS

An online RP-SAX-RP (reverse phase-strong anion exchange-reverse phase) chromatography platform was used to fractionate peptides prior to MS analysis [40]. Peptides were dissolved in 120 µL of 20 mM ammonium formate [pH 10.0] and half samples were injected separately to provide technical replicates. After loading, peptides were fractionated by step elutions through the combination of first dimension high pH reverse phase column (150 µm I.D. capillary packed with 6 cm of 5 µm XBridge C18 material from Waters Corp.) and second dimension anion exchange column (150 µm I.D. capillary packed with 6 cm of 10 µm POROS10HQ; Applied Biosystems). Ten elution fractions of increasing organic and salt concentrations were each further resolved by a four hours acetonitrile gradient on the third dimension low pH reverse phase column (30 µm I.D. capillary packed with 50 cm of 5 µm Monitor C18; Column Engineering, Ontario, CA). The combination of organic and salt concentration gradients were optimized to produce similar peptide complexity in each fraction.

Peptide fractions were directly analyzed using a Q Exactive Hybrid Quadrupole-Orbitrap Mass Spectrometer (Thermo Scientific) operated in data-dependent mode. The 10 most abundant ions in each MS scan (m/z 300–2000, target=3E6, max fill time=500 ms, resolution=120K) were subjected to MS/MS (HCD fragmentation, 35% normalized collision energy, target=1E5, max fill time=100 ms, min AGC target=1.5E4, isolation window=1 Da,

fixed first mass= m/z 100). Dynamic exclusion was enabled, with a repeat count of 1 and an exclusion duration of 30 seconds.

5.3. Data analysis and presentation

Raw MS data files were directly accessed and converted to .mgf files as described [40]. The files were searched using Mascot (version 2.4.1 or later) against a forward-reversed yeast database from the *Saccharomyces* Genome Database (downloaded April 2010), with an appended cRAP (common repository of adventitious proteins) database of 752 entries. The mass tolerance values were set to 10 ppm and 25 mmu for precursor and product ions, respectively. After matching MS/MS patterns to predicted peptides, an Excel spreadsheet containing the Mascot search results filtered to 1% FDR was generated using Multiplierz software [41, 42]. Only peptides mapping uniquely to a given gene by our pep2gene algorithm were used for quantitation [43]. Reporter ion ratios were corrected for isotopic impurities, and then normalized for source protein variations. Normalization was performed using BSA (contained in the restriction enzyme used for elution from magnetic beads) or the total count of the reporter ion signals, which should be in equal abundance between samples. Corrected, normalized TMT reporter ion values were then used to calculate ratios between the various experimental conditions.

Taking examples from the experiments described above, multiple unique peptides assigned to RNAPII subunits were detected in both the PIC and EC purifications. Reporter ion ratios show that all peptides for Rpb1 were significantly enriched in PICs by Gal4-VP16 activator (red dots in Fig. 5A) and in ECs by NTPs (red dots in Fig. 5C), as were peptides for other RNAPII subunits (blue dots in Fig. 5A and C). To reduce data complexity and average out noise, reporter signal intensity values of all peptides matching a given protein can be summed before calculation of ratios. Scatter plots showing enrichment ratios for individual proteins (Fig. 5B and D for PICs and ECs, respectively) are much simpler than those showing individual peptides. Correlation coefficients between technical replicates show very good reproducibility of the mass spectrometry. The vast majority of identified proteins (gray dots, typically around one thousand) form a distribution around the scatter plot origin, indicating that these represent background binding that do not change with transcription conditions. It should be noted that the TMT reporter ion enrichment ratio for RNAPII and other transcription factors significantly underestimates their absolute enrichment as measured by immunoblotting. This signal compression is due to background TMT signal coming from unenriched peptides that coelute with the enriched peptide during chromatography [44]. Therefore, the Student t-test or mixed model analysis is used to determine the confidence level that a given protein is outside the background distribution (see Fig. 6B below).

The PIC purification experiment showed that Gal4-VP16 activator specifically increased binding of RNAPII subunits, all basal initiation factors (TBP, TFIIA, TFIIB, TFIIE, TFIIF, TFIIH, TFIIIS), and co-activators (Mediator, TAFs, SAGA, NuA4, and Swi/Snf) (Fig. 6A). Two biological replicates, each with two technical replicates, were used to test the reproducibility of activator-dependent enrichment. A “volcano” plot (Fig. 6B) shows the average fold-enrichment for each protein versus the Student t-test p-value for the six

replicate points falling within the distribution of the full protein dataset. Proteins above the horizontal dotted line fall below a 5% false discovery rate.

The scatter plots show that subunits within a complex vary in their absolute enrichment ratios, but clearly cluster (Fig. 6A). For example, it can be seen that the activator-induced increase in basal initiation factors is on average less than that for co-activators. The variations within a complex arise from different levels of background signals for different subunits (see Fig. 7C below), as well as the fact that some proteins exist in more than one complex. Indeed, information about shared subunits can be gleaned from the scatter plots. For example, there are five proteins (Taf5, Taf6, Taf9, Taf10, and Taf12) that are common to TFIID (TAF) and SAGA complexes. On average, SAGA-specific proteins show greater enrichment by transcription activator than TFIID/TAF-specific proteins (Fig. 6C). The shared subunits (TAFs/SAGA) show intermediate enrichment ratios, suggesting roughly equal amounts of SAGA and TFIID bound to the template (Fig. 6C). In contrast, the chromatin remodeling complex Swi/Snf shows activator-dependent recruitment while the related RSC complex does not (Fig. 6D). In this case, the shared subunits cluster with the RSC-specific subunits, indicating more bound RSC than Swi/Snf.

Chromatin significantly influenced the template-associated proteins. One set of proteins bound better to chromatin than naked DNA independently of transcription activator (“Enriched by chromatin” in Fig. 7A, see also[18]), while others preferred the naked templates (“Excluded by chromatin”). Absolute levels of all PIC components were reduced by nucleosomes, presumably reflecting the promoter occlusion that also occurs in vivo. However, the activator-induced enrichment ratios (“Enriched by an activator”) on chromatin are similar, or even increased, because of this lower basal association (Fig. 7B). This effect can also be seen by plotting the effects of activator versus those of chromatin (Fig. 7C), where those proteins that show the greatest chromatin suppression tend to be the most stimulated by activator. This analysis reveals that Gal4-VP16 activator has the greatest stimulatory effect on the co-activators SAGA and Swi/Snf.

In the experiment isolating EC proteins that appear on downstream DNA in the presence of NTPs (Fig. 4A), significant enrichment was seen for RNAPII subunits and the “core” transcription elongation factors Spt4/5 (DSIF), Paf1C, Spt6, Spn1, and Elf1 (Figs. 8A and B) [38]. An essential control reaction containing alpha-amanitin showed that enrichment is truly dependent on transcription and not just NTPs [28, 38]. Along with the core ECs, we also detected time-dependent enrichment of factors involved in co-transcriptional processes such as mRNA capping (Ceg1, Cet1, Abd1, Cbc2, and Sto1), histone modifications (Set1C/COMPASS, Set2, Set3, and Rph1), chromatin remodeling (FACT, Chd1, and ISWI), and phosphorylation (Bur1/CDK9). Detailed analysis of our EC experiments are discussed extensively elsewhere [38].

6. Concluding remarks

Work with purified factors has been essential for understanding RNAPII transcription. However, transcription mechanisms and dynamics also need to be studied in the more natural context of the full nuclear proteome. To this end, the combination of yeast nuclear

extract, immobilized DNA templates, and quantitative mass spectrometry provides an in vitro system for analyzing the composition of and transitions between initiation and elongation RNAPII complexes. We note that an increasing number of academic core facilities now offer multiplexed, isobaric-tagged mass spectrometry analysis. However, the immobilized template approach should also work with more common mass spectrometry methods that use unlabeled peptide counts to estimate protein abundance.

Basal transcription factors, activators and co-activators, nascent RNA interacting factors, histone modifying enzymes, and chromatin remodelers can be simultaneously monitored. Although experiment using yeast extracts are describe here, similar approaches using mammalian extracts have also proven fruitful [45]. Using chromatized templates, co-transcriptional histone modifications can be investigated. Future experiments can use pre-modified, mutated, or variant histones to determine their effects on transcription. Functions of individual factors can be probed using chemical inhibitors or extracts that have been genetically or biochemically depleted of those factors. By recapitulating many aspects of the in vivo transcription process in an in vitro system, we believe this system provides a comprehensive approach that complements other approaches.

Acknowledgements.

This work was supported by the US National Institutes of Health grants GM467498 and GM56663.

References

- [1]. Selth LA, Sigurdsson S, Sveistrup JQ, Transcript Elongation by RNA Polymerase II, *Annu Rev Biochem* 79 (2010) 271–93. [PubMed: 20367031]
- [2]. Bentley DL, Coupling mRNA processing with transcription in time and space, *Nat Rev Genet* 15(3) (2014) 163–75. [PubMed: 24514444]
- [3]. Li B, Carey M, Workman JL, The role of chromatin during transcription, *Cell* 128(4) (2007) 707–19. [PubMed: 17320508]
- [4]. Hampsey M, A review of phenotypes in *Saccharomyces cerevisiae*, *Yeast* 13(12) (1997) 1099–133. [PubMed: 9301019]
- [5]. Prather D, Krogan NJ, Emili A, Greenblatt JF, Winston F, Identification and characterization of Elf1, a conserved transcription elongation factor in *Saccharomyces cerevisiae*, *Mol Cell Biol* 25(22) (2005) 10122–35. [PubMed: 16260625]
- [6]. Scholes DT, Banerjee M, Bowen B, Curcio MJ, Multiple regulators of Ty1 transposition in *Saccharomyces cerevisiae* have conserved roles in genome maintenance, *Genetics* 159(4) (2001) 1449–65. [PubMed: 11779788]
- [7]. Winston F, Chaleff DT, Valent B, Fink GR, Mutations affecting Ty-mediated expression of the HIS4 gene of *Saccharomyces cerevisiae*, *Genetics* 107(2) (1984) 179–97. [PubMed: 6329902]
- [8]. Zhang L, Fletcher AG, Cheung V, Winston F, Stargell LA, Spn1 regulates the recruitment of Spt6 and the Swi/Snf complex during transcriptional activation by RNA polymerase II, *Mol Cell Biol* 28(4) (2008) 1393–403. [PubMed: 18086892]
- [9]. Dynlacht BD, Hoey T, Tjian R, Isolation of coactivators associated with the TATA-binding protein that mediate transcriptional activation, *Cell* 66(3) (1991) 563–76. [PubMed: 1907890]
- [10]. Spt6 M, Carthew RW, Greenblatt J, Isolation of three proteins that bind to mammalian RNA polymerase II, *J Biol Chem* 260(18) (1985) 10353–60. [PubMed: 3860504]
- [11]. Wade PA, Werel W, Fentzke RC, Thompson NE, Leykam JF, Burgess RR, Jaehning JA, Burton ZF, A novel collection of accessory factors associated with yeast RNA polymerase II, *Protein Expr Purif* 8(1) (1996) 85–90. [PubMed: 8812838]

- [12]. Krogan NJ, Kim M, Ahn SH, Zhong G, Kobor MS, Cagney G, Emili A, Shilatifard A, Buratowski S, Greenblatt JF, RNA polymerase II elongation factors of *Saccharomyces cerevisiae*: a targeted proteomics approach, *Mol Cell Biol* 22(20) (2002) 6979–92. [PubMed: 12242279]
- [13]. Hahn S, Structure and mechanism of the RNA polymerase II transcription machinery, *Nat Struct Mol Biol* 11(5) (2004) 394–403. [PubMed: 15114340]
- [14]. Ehara H, Yokoyama T, Shigematsu H, Yokoyama S, Shirouzu M, Sekine SI, Structure of the complete elongation complex of RNA polymerase II with basal factors, *Science* 357(6354) (2017) 921–924. [PubMed: 28775211]
- [15]. Louder RK, He Y, Lopez-Blanco JR, Fang J, Chacon P, Nogales E, Structure of promoter-bound TFIID and model of human pre-initiation complex assembly, *Nature* 531(7596) (2016) 604–9. [PubMed: 27007846]
- [16]. Schilbach S, Hantsche M, Tegunov D, Dienemann C, Wigge C, Urlaub H, Cramer P, Structures of transcription pre-initiation complex with TFIIF and Mediator, *Nature* 551(7679) (2017) 204–209. [PubMed: 29088706]
- [17]. Vos SM, Farnung L, Boehning M, Wigge C, Linden A, Urlaub H, Cramer P, Structure of activated transcription complex Pol II-DSIF-PAF-SPT6, *Nature* 560(7720) (2018) 607–612. [PubMed: 30135578]
- [18]. Sikorski TW, Joo YJ, Ficarro SB, Askenazi M, Buratowski S, Marto JA, Proteomic analysis demonstrates activator-and chromatin-specific recruitment to promoters, *J Biol Chem* 287(42) (2012) 35397–408. [PubMed: 22902623]
- [19]. Corden JL, RNA polymerase II C-terminal domain: Tethering transcription to transcript and template, *Chem Rev* 113(11) (2013) 8423–55. [PubMed: 24040939]
- [20]. Liu Y, Warfield L, Zhang C, Luo J, Allen J, Lang WH, Ranish J, Shokat KM, Hahn S, Phosphorylation of the transcription elongation factor Spt5 by yeast Bur1 kinase stimulates recruitment of the PAF complex, *Mol Cell Biol* 29(17) (2009) 4852–63. [PubMed: 19581288]
- [21]. Jonkers I, Lis JT, Getting up to speed with transcription elongation by RNA polymerase II, *Nat Rev Mol Cell Biol* 16(3) (2015) 167–77. [PubMed: 25693130]
- [22]. Sdano MA, Fulcher JM, Palani S, Chandrasekharan MB, Parnell TJ, Whitby FG, Formosa T, Hill CP, A novel SH2 recognition mechanism recruits Spt6 to the doubly phosphorylated RNA polymerase II linker at sites of transcription, *Elife* 6 (2017).
- [23]. Chun Y, Suh H, Joo Y, Batot G, Hill C, Formosa T, Buratowski S, Selective kinase inhibition shows that Bur1 (Cdk9) phosphorylates the Rpb1 linker. *Mol Cell Biol*, in revision (2019).
- [24]. Smolle M, Workman JL, Transcription-associated histone modifications and cryptic transcription, *Biochim Biophys Acta* 1829(1) (2013) 84–97. [PubMed: 22982198]
- [25]. Ranish JA, Yi EC, Leslie DM, Purvine SO, Goodlett DR, Eng J, Aebersold R, The study of macromolecular complexes by quantitative proteomics, *Nat Genet* 33(3) (2003) 349–55. [PubMed: 12590263]
- [26]. Vary JC Jr., Fazzio TG, Tsukiyama T, Assembly of yeast chromatin using ISWI complexes, *Methods Enzymol* 375 (2004) 88–102. [PubMed: 14870661]
- [27]. Inada T, Winstall E, Tarun SZ Jr., Yates JR 3rd, Schieltz D, Sachs AB, One-step affinity purification of the yeast ribosome and its associated proteins and mRNAs, *RNA* 8(7) (2002) 948–58. [PubMed: 12166649]
- [28]. Joo YJ, Ficarro SB, Soares LM, Chun Y, Marto JA, Buratowski S, Downstream promoter interactions of TFIID TAFs facilitate transcription reinitiation, *Genes Dev* 31(21) (2017) 2162–2174. [PubMed: 29203645]
- [29]. Sadowski I, Ma J, Triezenberg S, Ptashne M, GAL4-VP16 is an unusually potent transcriptional activator, *Nature* 335(6190) (1988) 563–4. [PubMed: 3047590]
- [30]. Cho EJ, Takagi T, Moore CR, Buratowski S, mRNA capping enzyme is recruited to the transcription complex by phosphorylation of the RNA polymerase II carboxy-terminal domain, *Genes Dev* 11(24) (1997) 3319–26. [PubMed: 9407025]
- [31]. Reeves WM, Hahn S, Activator-independent functions of the yeast mediator sin4 complex in preinitiation complex formation and transcription reinitiation, *Mol Cell Biol* 23(1) (2003) 349–58. [PubMed: 12482986]

- [32]. Ranish JA, Yudkovsky N, Hahn S, Intermediates in formation and activity of the RNA polymerase II preinitiation complex: holoenzyme recruitment and a postrecruitment role for the TATA box and TFIIB, *Genes Dev* 13(1) (1999) 49–63. [PubMed: 9887099]
- [33]. Luger K, Mader AW, Richmond RK, Sargent DF, Richmond TJ, Crystal structure of the nucleosome core particle at 2.8 Å resolution, *Nature* 389(6648) (1997) 251–60. [PubMed: 9305837]
- [34]. Puig O, Caspary F, Rigaut G, Rutz B, Bouveret E, Bragado-Nilsson E, Wilm M, Seraphin B, The tandem affinity purification (TAP) method: a general procedure of protein complex purification, *Methods* 24(3) (2001) 218–29. [PubMed: 11403571]
- [35]. Vary JC Jr., Gangaraju VK, Qin J, Landel CC, Kooperberg C, Bartholomew B, Tsukiyama T, Yeast Isw1p forms two separable complexes in vivo, *Mol Cell Biol* 23(1) (2003) 80–91. [PubMed: 12482963]
- [36]. Johnson A, Li G, Sikorski TW, Buratowski S, Woodcock CL, Moazed D, Reconstitution of heterochromatin-dependent transcriptional gene silencing, *Mol Cell* 35(6) (2009) 769–81. [PubMed: 19782027]
- [37]. Fyodorov DV, Kadonaga JT, Chromatin assembly in vitro with purified recombinant ACF and NAP-1, *Methods Enzymol* 371 (2003) 499–515. [PubMed: 14712724]
- [38]. Joo Y, Ficarro SB, Chun Y, Marto JA, B. S., In vitro analysis of RNA polymerase II elongation complex dynamics, *Genes Dev*, in press (2019).
- [39]. Hughes CS, Foehr S, Garfield DA, Furlong EE, Steinmetz LM, Krijgsvelde J, Ultrasensitive proteome analysis using paramagnetic bead technology, *Mol Syst Biol* 10 (2014) 757. [PubMed: 25358341]
- [40]. Ficarro SB, Zhang Y, Carrasco-Alfonso MJ, Garg B, Adelmant G, Webber JT, Luckey CJ, Marto JA, Online nanoflow multidimensional fractionation for high efficiency phosphopeptide analysis, *Mol Cell Proteomics* 10(11) (2011) O111 011064.
- [41]. Askenazi M, Parikh JR, Marto JA, mzAPI: a new strategy for efficiently sharing mass spectrometry data, *Nat Methods* 6(4) (2009) 240–1. [PubMed: 19333238]
- [42]. Parikh JR, Askenazi M, Ficarro SB, Cashorali T, Webber JT, Blank NC, Zhang Y, Marto JA, multiplierz: an extensible API based desktop environment for proteomics data analysis, *BMC Bioinformatics* 10 (2009) 364. [PubMed: 19874609]
- [43]. Askenazi M, Marto JA, Linial M, The complete peptide dictionary--a meta-proteomics resource, *Proteomics* 10(23) (2010) 4306–10. [PubMed: 21082763]
- [44]. Ting L, Rad R, Gygi SP, Haas W, MS3 eliminates ratio distortion in isobaric multiplexed quantitative proteomics, *Nat Methods* 8(11) (2011) 937–40. [PubMed: 21963607]
- [45]. Lin JJ, Lehmann LW, Bonora G, Sridharan R, Vashisht AA, Tran N, Plath K, Wohlschlegel JA, Carey M, Mediator coordinates PIC assembly with recruitment of CHD1, *Genes Dev* 25(20) (2011) 2198–209. [PubMed: 21979373]

Highlights:

- Transcription complexes are assembled from nuclear extract on immobilized templates
- Purified complexes are analyzed by quantitative mass spectrometry or immunoblotting
- Both initiation and elongation complexes can be isolated
- Transcription templates can be naked DNA or chromatin
- This system reproduces many co-transcriptional processes

Fig. 1.

Experimental workflow for purification and analysis of RNAPII complexes at different transcription stages.

Fig. 2. Components for preparation of chromatinized in vitro transcription templates.

(A) DNA templates used in this study. A *HIS4* core promoter (*HIS4p*) template was used for PIC purification. It contains a single Gal4 binding site upstream of the *HIS4* core promoter. For ECs purification, a template possessing five Gal4 binding sites upstream of the *CYC1* core promoter (*CYC1p*) and G-less cassette was used. (B) Purity of the linear DNA templates was examined in an ethidium bromide stained 1% agarose gel. (C) Purified recombinant transcription activators Gal4-VP16 and Gal4-Gcn4 are shown in a Coomassie-stained 10% SDS-PAGE gel. (D) Components for in vitro chromatin reconstitution are shown using colloidal blue (Histone octamer), Coomassie brilliant blue (Nap1), or silver stained (Isw1a) SDS-PAGE gel. (E) Chromatin reconstitution on the *CYC1p* template (660 ng) with the indicated component(s) was monitored by digestion with 2 or 4 units of MNase (Fermentas) for 10 min at 25°C. DNA fragmentation of nucleosomal units was examined by deproteinization and 1.3 % agarose gel stained with ethidium bromide.

Fig. 3. Purification of RNAPII PICs on *HIS4p* immobilized templates.

(A) Experimental workflow of RNAPII PIC purification. Note that final elution can be done with Pst I for cleanest results by mass spectrometry (3D LC-MSMS) or simply by boiling beads for western blot (WB). (B) Enrichment of RNAPII by activator was confirmed by immunoblotting with anti-Rpb1 (8WG16), while total purified proteins were detected by silver stained 10% SDS-PAGE. (C) Linear template DNA (220 ng *PMA1p*) was chromatinized and immobilized as described in the text. After incubation with 1 mg of yeast nuclear extract and indicated cofactors for 40 min at 25°C, total bound proteins were recovered by boiling. Histone and post-translational modifications on histones were detected by immunoblotting with specific antibodies against trimethylated H3K36 (H3K36me3, abcam, ab9050), acetylated H3 (H3ac, Millipore, 06–599), acetylated H4 (H4ac, Millipore, 06–866), and H3 (abcam, ab1791) separated by 15% SDS-PAGE. (D) Plasmid DNA containing G-less cassette and *CYC1* core promoter (pUC18-G5CYC1 G-/SB649), either naked or chromatinized, was transcribed in vitro, with or without the indicated cofactors, for 45 min at 25°C with Gal4-VP16 or Gal4-Gcn4. RNase T1 digested transcripts were recovered, separated by 6% UREA PAGE, and detected using ³²P autoradiography.

Fig. 4. Purification of RNAPII ECs on *CYC1p* immobilized templates.
(A) Experimental workflow for EC purification. (B) NTP-dependent enrichment of RNAPII at the downstream end of templates was confirmed by immunoblotting with anti-Rpb1 (8WG16), while total purified proteins were detected by silver stained 10% SDS-PAGE.

Fig. 5. Analysis and presentation of quantitative mass spectrometry results.

(A and B) RNAPII PICs were purified as described in Fig. 3A. Proteins eluted by Pst I digestion were subjected to quantitative mass spectrometry using TMT reagents. A scatter plot presents log₂ scaled ratio of activator-dependent binding (+Gal4-VP16/-Gal4-VP16) for unique peptides (A) and proteins (B) from two technical replicates. Spots of the unique peptides and protein for Rpb1 were marked as red. Spots of the unique peptides and protein for other RNAPII subunits were marked as blue. (C and D) RNAPII ECs were purified as described in Fig. 4A. Proteins eluted by Ssp I digestion were subjected to quantitative mass spectrometry using TMT reagents. A scatter plot presents log₂ scaled ratio of transcription-

dependent binding (+NTPs/-NTPs) for unique peptides (C) and proteins (D) from two technical replicates.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Fig. 6. Identification of proteins recruited to immobilized templates by transcription activator. (A) A scatter plot shows log₂ ratios of activator-dependent binding (+Gal4-VP16/-Gal4-VP16) as in Fig. 5B. Various transcription-related complexes are marked with different colors. (B) A volcano plot shows reproducibility and statistical significance for the quantitative mass spectrometry from two biological replicates times two technical replicates each. Each dot represents a single protein. The same color code as in panel A was used for different complexes. A p-value below 0.05 (above dashed line) signifies a less than 5% probability of being a false positive for enrichment. (C and D) Scatter plots compare log₂ ratios of activator-dependent protein association for specific and shared subunits from (C) SAGA and TFIID/TAFs or (D) RSC and Swi/Snf complexes, marked with different colors.

Fig. 7. Chromatin effects on PICs assembly.

(A) RNAPII PICs were purified on naked and chromatinized templates. Proteins bound to each template were recovered by Pst I digestion and subjected to quantitative mass spectrometry. A heat-map generated with Multiple Experiment Viewer shows differential binding of all identified proteins (796 total), with each row representing one identified protein. Proteins clustered into four groups (enriched by chromatin, enriched by activator, excluded by chromatin, and unchanged). (B) A scatter plot presents log₂ ratio of activator-dependent binding (+Gal4-VP16/-Gal4-VP16) on naked template (x-axis) versus chromatin template (y-axis). Each dot represents one protein. Factors of interest are marked with different colors representing a complex or category. (C) A scatter plot compares log₂ ratio of activator-dependent binding (+Gal4-VP16/-Gal4-VP16) on chromatin template (x-axis) versus log₂ ratio of chromatin effect (on naked template/on chromatin template). Each dot represents one protein.

Fig. 8. Analysis of proteins identified in elongating RNAPII complexes.

(A) A scatter plot presents log₂ ratio of proteins enriched on a Ssp I -released fragment (+NTPs/-NTPs), from two technical replicates as in Fig. 5D. Each dot represents one protein, color-coded as indicated. Key elongation factors are labeled. (B) A volcano plot shows reproducibility and statistical significance for the quantitative mass spectrometry from three biological replicates times two technical replicates each. Each dot represents a single protein labeled as in panel (A).