Molecular Cell
Article

RNA Polymerase I Contains a TFIIF-Related DNA-Binding Subcomplex

Sebastian R. Geiger,[1] Kristina Lorenzen,[2,3] Amelie Schreieck,[1] Patrizia Hanecker,[1] Dirk Kostrewa,[1] Albert J.R. Heck,[2,3] and Patrick Cramer[1,*]
[1]Gene Center and Department of Biochemistry, Center for Integrated Protein Science CIPSM, Ludwig-Maximilians-Universität München, Feodor-Lynen-Strasse 25, 81377 Munich, Germany
[2]Biomolecular Mass Spectrometry and Proteomics Group, Bijvoet Center for Biomolecular Research and Utrecht Institute for Pharmaceutical Sciences, Utrecht University
[3]Netherlands Proteomics Centre
Padualaan 8, 3584 CH Utrecht, The Netherlands
*Correspondence: cramer@lmb.uni-muenchen.de
DOI 10.1016/j.molcel.2010.07.028

SUMMARY

The eukaryotic RNA polymerases Pol I, II, and III use different promoters to transcribe different classes of genes. Promoter usage relies on initiation factors, including TFIIF and TFIIE, in the case of Pol II. Here, we show that the Pol I-specific subunits A49 and A34.5 form a subcomplex that binds DNA and is related to TFIIF and TFIIE. The N-terminal regions of A49 and A34.5 form a dimerization module that stimulates polymerase-intrinsic RNA cleavage and has a fold that resembles the TFIIF core. The C-terminal region of A49 forms a "tandem winged helix" (tWH) domain that binds DNA with a preference for the upstream promoter nontemplate strand and is predicted in TFIIE. Similar domains are predicted in Pol III-specific subunits. Thus, Pol I/III subunits that have no counterparts in Pol II are evolutionarily related to Pol II initiation factors and may have evolved to mediate promoter specificity and transcription processivity.

INTRODUCTION

Transcription in eukaryotic cells is performed by three related multisubunit RNA polymerases. Whereas Pol II synthesizes mRNAs and Pol III produces small RNAs such as tRNAs, Pol I transcribes the rRNA precursor. The structure of Pol II has been studied in detail (Cramer et al., 2008), but structural information is still limited for Pol I (De Carlo et al., 2003; Kuhn et al., 2007) and Pol III (Fernández-Tornero et al., 2007; Jasiak et al., 2006). All three polymerases contain a ten-subunit catalytic core of shared or homologous subunits and a peripheral heterodimeric subcomplex of conserved structure called A14/43, Rpb4/7, and C17/25 in Pol I, II, and III, respectively (Armache et al., 2005; Jasiak et al., 2006; Kuhn et al., 2007). In addition to these 12 related subunits, Pol I contains the specific subcomplex A49/34.5 (Kuhn et al., 2007), and Pol III contains the specific subcomplexes C37/53 (Kassavetis et al., 2010; Landrieux et al., 2006) and C82/34/31 (Wang and Roeder, 1997), which includes the stable C82/34 dimer (Lorenzen et al., 2007).

Because all three polymerases share a similar core structure and active center, their mechanism of DNA-dependent RNA elongation is similar and well understood at a structural level (Brueckner et al., 2009). In contrast, the structural basis for promoter-specific transcription initiation remains unclear. To recognize different promoters, the polymerases use different initiation factors. Pol II uses the initiation factors TFIIB, -D, -E, -F, and -H, which are apparently not related to Pol I initiation factors (Moss et al., 2007).

Here, we show that the Pol I subcomplex A49/34.5 consists of two distinct domains, and we provide the crystal structures of these domains. We demonstrate that the fold of the N-terminal A49/34.5 dimerization module is similar to that of TFIIF, whereas the A49 C-terminal domain contains a previously unobserved tWH domain that has DNA-binding activity and is predicted in TFIIE. Our results complete the structural information on Pol I subunits, demonstrate that Pol I, in contrast to Pol II, contains a DNA-binding surface domain, and show how the three RNA polymerases are evolutionarily related.

RESULTS

Pol I Subcomplex A49/34.5 Forms a Heterodimer with Two Domains

It was previously shown that the Pol I subunits A49 and A34.5 interact and can dissociate from Pol I (Huet et al., 1976; Kuhn et al., 2007). To investigate whether these two subunits form a heterodimer, we used native mass spectrometry (Benesch et al., 2007; Heck, 2008). We purified the complete Pol I as described (Kuhn et al., 2007) and obtained a mass of 593 kDa (Figure 1). This agrees well with the theoretical molecular weight of 592 kDa for Pol I and indicates that one copy of each Pol I subunit is present, including A49 and A34.5. Pol I preparations also gave rise to peaks with masses of 520 and 74 kDa, which correspond to the theoretical masses of Pol I lacking the A49/34.5 heterodimer (Pol IΔ, 518 kDa) and the free A49/34.5 heterodimer (74 kDa), respectively (Figures 1A and 1B). The stability of Pol I was dependent on the ionic strength. At 160 mM ammonium acetate, Pol I was largely dissociated into Pol IΔ and

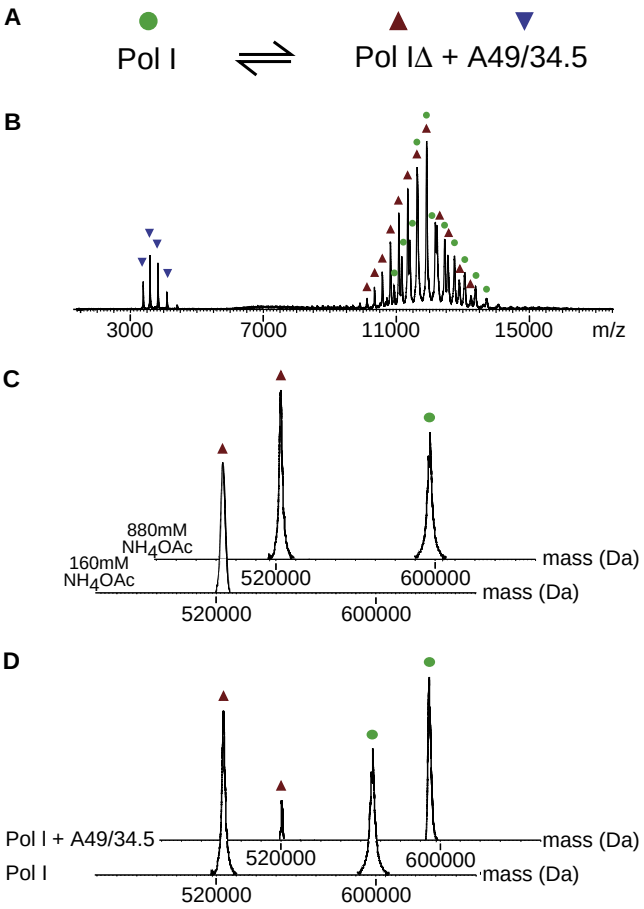

Figure 1. Native Mass Spectrometry of Pol I

(A) Dissociation of the A49/34.5 subcomplex from the complete 14 subunit Pol I generates the 12 subunit Pol IΔ.

(B) Native mass spectrum of Pol I, electrosprayed from aqueous ammonium acetate (880 mM). In the high m/z region, two different charge distributions are present, corresponding to Pol I (green dots) and Pol IΔ (brown triangles). The "free" A49/34.5 heterodimer is also detected in the low m/z region (blue triangles).

(C) Pol I stability depends on ionic strength. Two transformed zero-charge mass spectra are shown, recorded at 880 mM and 160 mM ammonium acetate.

(D) A49/34.5 can be associated with Pol IΔ. The lower spectrum was recorded with a sample containing Pol I in 880 mM ammonium acetate, whereas the upper spectrum was obtained after preincubation of the sample with a 5-fold molar excess of recombinant A49/34.5.

A49/34.5, whereas at 880 mM ammonium acetate, less dissociation was observed (Figure 1C). When an excess of recombinant A49/34.5 was added, higher relative amounts of complete Pol I were detected (Figure 1D). These results show that Pol I contains a single copy of each of its 14 subunits and that the A49/34.5 subcomplex forms a stable heterodimer.

To probe for flexible protein regions and to delineate domains, we subjected recombinant A49/34.5 (Kuhn et al., 2007) to limited proteolysis. The C-terminal tail of A34.5 and the central region of A49 were particularly sensitive to proteases (Figure 2A). Consistent with this finding, these regions are predicted to be largely unstructured (Biegert et al., 2006). Based on these results, we designed minimal variants for subunit coexpression and performed several rounds of solubility studies and variant optimization (Geiger et al., 2008). This revealed that the A49/34.5 heterodimer forms two structured domains connected by a flexible linker. The N-terminal regions of both subunits constitute a dimerization module, whereas the C-terminal region of A49 forms an independently folded domain (Figures 2A and 2B).

The A49/34.5 Dimerization Module Resembles TFIIF

We prepared the A49/34.5 dimerization module from various species and obtained crystals for the *Candida glabrata* variant $A49_{1-99}$/$A34.5_{25-143}$. The structure was determined by single wavelength anomalous diffraction at 2.9 Å resolution (Table 1). It revealed three interconnected β barrels (Figure 2C). A total of 17 interwoven β strands and three α helices contribute to a conserved hydrophobic core. Consistent with the structure, subunit dimerization is impaired by mutation of key core residues (Figure 2E) (Kuhn et al., 2007) or deletion of A49 strands β1, β1′, and β2 (data not shown). The structure resembles that of the TFIIF Rap74/30 dimerization module (Gaiser et al., 2000), with a DALI Z-score of 4.3 (Holm and Sander, 1995) and a root-mean-square deviation (rmsd) in 117 Cα atom positions of 2.6 Å. However, it lacks three elements of the Rap74/30 module, the β4-β5 hairpin, the Rap74 C-terminal helix α1, and the Rap30 loop α3-β5 (Figure 2D–2E).

Because of a pseudo 2-fold axis (Figure 2C), the A49/34.5 dimerization module may alternatively be superimposed with the TFIIF Rap74/30 module such that subunits A49 and A34.5 overlay with Rap30 and Rap74, respectively (Figure S1 available online). The similarity of the A49 and A34.5 dimerization domains is consistent with the formation of soluble homodimers after their expression in isolation (data not shown). Although native A49/34.5 is heterodimeric (Figure 1), heterotetramers are observed in the crystal lattice (Figures S2B and S2C). Tetramerization occurs via an unusual exchange of β strands, is an artifact of the high protein concentrations required for crystallization (Figure S2A), and may explain previous observations of TFIIF tetramers (Flores et al., 1990).

A49 Contains a Tandem Winged Helix Domain

The C-terminal domain of *S. cerevisiae* A49 could also be crystallized (Experimental Procedures), and its structure was determined by multiwavelength anomalous diffraction at 2.0 Å resolution (Table 1). The structure revealed a single domain (Figure 3A) with a fold that is not found in the database (Berman et al., 2000). The core of the domain consists of two tightly packed subdomains with a winged helix-turn-helix (winged helix, WH) fold (WH1, residues 254–321; WH2, residues 322–403) (Figure 3A). The WH fold is present in transcription factors and nucleic acid-binding proteins (Brennan, 1993; Gajiwala and Burley, 2000; Kenney, 2002) and consists of three α helices and three β strands in the order α1, β1, α2, α3, β2, and β3 (Figures 2A and 3A). Subdomains WH1 and WH2 additionally contain one or two short α helices, respectively, inserted in the β2-β3 loop (Figure 3A). In WH1, helix α1 is extended and surrounded with four additional N-terminal helices (α1*-α4*; Figure 3A). This N-terminal helical bundle and the β2-β3

insertions are flexible because they have higher B factors (Figure S3) and are partially disordered in an alternative crystal form (Table 1 and Figure S5C). The WH subdomains form an interface with conserved hydrophobic residues on their α1 helices that pack in an antiparallel fashion (Figures 3B and S4). Because the WH subdomains are packed into a single domain, we refer to this structure as tandem winged helix (tWH) domain.

The tWH Domain Binds Duplex DNA with a Conserved Surface

The tWH domain contains an extended positively charged surface (Figure 3C), suggesting that it binds nucleic acids. We investigated this using an electrophoretic mobility shift assay (EMSA; Experimental Procedures). A double-stranded (ds) DNA of random sequence caused a shifted band, indicating a tWH-dsDNA complex (Figure 4A, lanes 1 and 2). Similar results were obtained with dsDNA corresponding to parts of the Pol I promoter (Figure 4A, lanes 3, 4, 9, and 10). When duplexes of less than 30 base pairs were used, dsDNA binding was impaired (Figure 4A, lanes 5–8, and 11–14). The dimerization module did not bind DNA (Figure 4B, lane 2), providing a negative control. Consistent with these findings, a DNA affinity of A49 was noticed, but not investigated (Liljelund et al., 1992). These results unravel a DNA-binding function of A49/34.5 located in its tWH domain.

To predict the DNA-binding surface of the tWH domain, we superimposed the WH domain of the transcription factor OhrR-DNA complex (Hong et al., 2005) with both A49 WH subdomains. Superposition with the A49 WH1 subdomain resulted in a clash between DNA and protein, whereas superposition with the WH2 subdomain did not, suggesting that α3, β2, and β3 of WH2 bind DNA (Figure 4C). The putative DNA-binding surface is positively charged and conserved (Figures 3C and 3D). Consistent with these results, multiple point mutations on WH1 (Figure S5A) did not influence DNA binding (Figure 4B, lanes 4 and 5), whereas mutation of WH2 residues K356/S358, K359, R365, or K393 impaired binding (Figure 4B, lanes 6, 7, 9, and 10). Further consistent with the model, a polyethylene glycol molecule is bound in the crystal at the position of the modeled DNA, indicating affinity of this site for nucleic acids (Figure 4D). Thus, DNA binding of the tWH domain requires WH2 residues in α3 and β3, consistent with a typical WH-DNA interaction.

The tWH Domain Binds Single-Stranded Promoter DNA

To test whether the tWH domain could also bind single-stranded (ss) DNA, we used EMSA and 40 nucleotide long oligonucleotides corresponding to the template or nontemplate strand of the Pol I promoter. Indeed, the tWH domain bound ssDNAs with a sequence preference for the nontemplate promoter strand upstream of the transcription start site +1 (positions −110 to −20; Figure 4E). In contrast, the tWH domain did generally not bind ssDNAs with sequences of the template strand, the promoter core element (positions −25 to +8), or unrelated sequences (Figure 4E).

To investigate whether the tWH domain can simultaneously bind ssDNA and dsDNA, we labeled ssDNA and dsDNA with a red and green fluorescent dye, respectively. The tWH domain bound ssDNA in the presence of bound dsDNA (Figure 4F, lanes 2–4). Because the fluorescence intensities of the migrating protein-DNA complexes were similar, a competitive DNA binding is not likely. Moreover, a tWH variant defective in dsDNA binding still bound ssDNA (Figure 4F, lanes 5–7). However, a tWH variant lacking 12 C-terminal residues did not bind DNA at all (Figure 4F, lanes 8–10). Thus, the DNA-binding WH2 surface is required for dsDNA binding, whereas the basic C-terminal tail is required for both ssDNA and dsDNA binding. These results show that the tWH domain binds ssDNA with a preference for the nontemplate strand in the upstream promoter region and indicate that this activity is not mutually exclusive with its dsDNA-binding function.

A49/34.5 Domains Have Distinct Functions

To map protein regions that anchor A49/34.5 to the Pol I core, we tested A49/34.5 variants for binding to Pol IΔ. The dimerization module did not bind, but its extension by the A49 linker or by the A34.5 tail enabled binding (Figure 5A, lanes 1–4). The A49 linker is proteolytically sensitive (Figure 2A), and a purified linker variant (Figure 2B) was unstable (data not shown). The linker is predicted to contain five helices, two of which were observed in crystals. A part of the N-terminal linker helix (residues 100–106) was observed in a poorly diffracting crystal of the extended dimerization module variant $A49_{1–119}/34.5_{25–143}$ (Figure S3B and Table 1). The C-terminal linker helix (residues 172–181) was observed in the tWH crystal in one molecule in the asymmetric unit (Figure S3C) and in a different position in another crystal form (data not shown and Table 1). These results show that A49/34.5 is anchored on Pol I with its domain linker and suggest that the A49/34.5 domains maintain some mobility on the Pol I surface.

We previously showed that A49/34.5 is required for normal Pol I RNA cleavage activity and processivity on synthetic templates (Kuhn et al., 2007). To map regions required for these functions, A49/34.5 variants were added to Pol IΔ, and the activities were tested in vitro (Experimental Procedures). These experiments revealed that normal RNA cleavage requires the dimerization module and the A49 linker or the A34.5 tail (Figure 5B). Thus, the dimerization module mediates full RNA cleavage activity, but this requires anchoring the domain on Pol I. Processivity, as measured by the ability of Pol I to synthesize RNA to the end of the DNA template, required the tWH domain, including the basic C-terminal tail (Figure 5C). Taken together, our results reveal the functional architecture of yeast A49/34.5 (Figure 6) and enable homology modeling for the related mammalian subcomplex Paf53/49 (Figure S7).

DISCUSSION

Structure and DNA-Binding Function of A49/34.5

The main difference between Pol I and Pol II is the presence of the additional subcomplex A49/34.5 in Pol I. We show here that A49/34.5 forms a heterodimer that consists of two domains connected by a flexible linker. The N-terminal regions of both subunits form a dimerization module with a TFIIF-like fold, whereas the C-terminal region of A49 forms a tWH domain with two WH subdomains. The dimerization module mediates

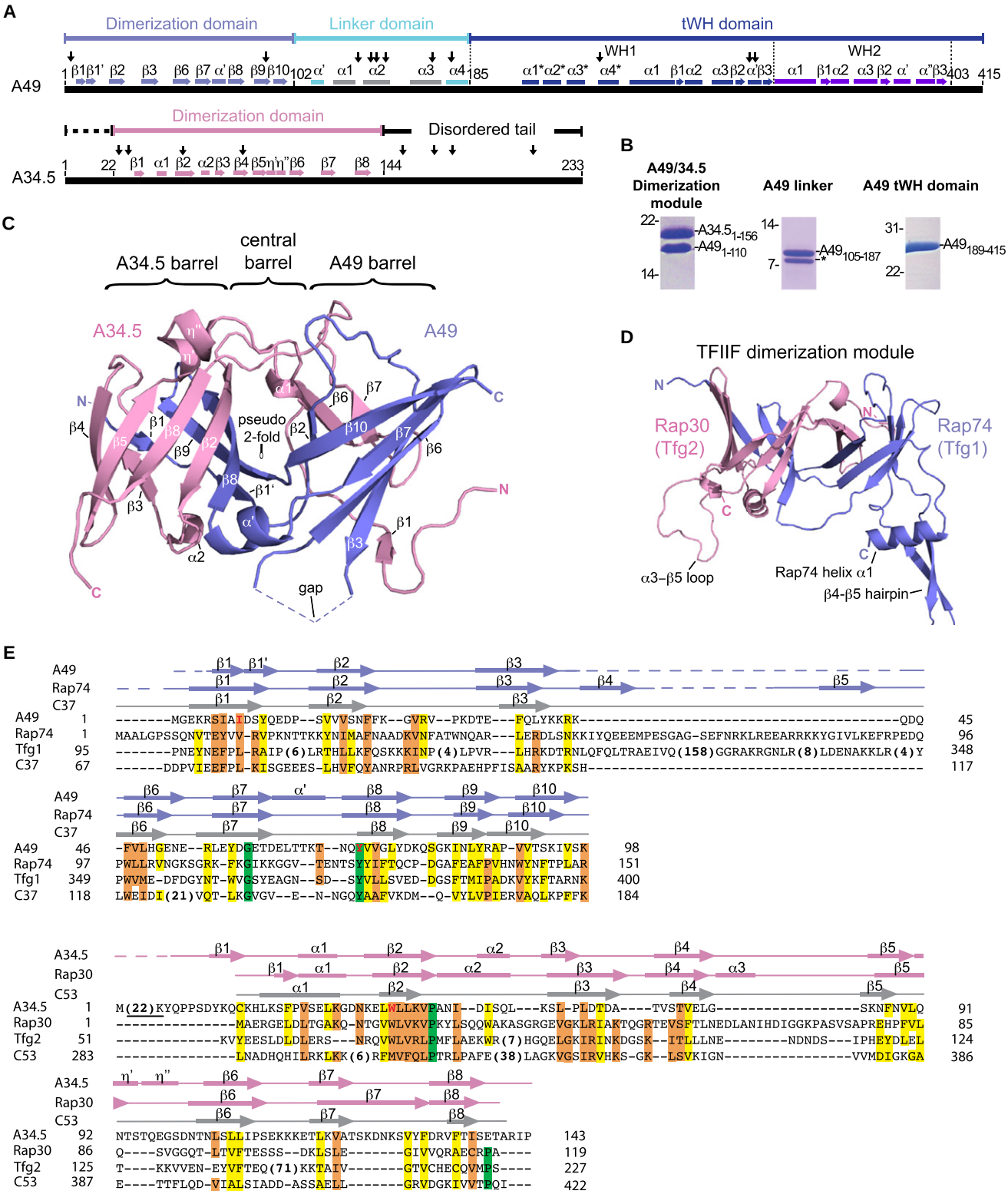

Figure 2. A49/34.5 Domain Organization and TFIIF-like Dimerization Module Structure

(A) Schematic representation of A49 and A34.5. The A49 N-terminal dimerization domain, interdomain linker, and C-terminal tandem winged helix (tWH) domain are shown in light blue, cyan, and dark blue, respectively. The A34.5 dimerization domain and C-terminal tail are in magenta and black, respectively. Secondary

Table 1. Diffraction Data and Refinement Statistics

A49/34.5 Variant	Dim. Module $A49_{1–99}/34.5_{25–143}$	Dim. Module (+ Linker Helix α′) $A49_{1–119}/34.5_{25–143}$	tWH Domain $A49_{171–403}$			tWH Domain $A49_{155–399}$
			Peak	Inflection	Remote	
Data Collection						
Space group	$P2_1$	$C222_1$	$P2_1$	$P2_1$	$P2_1$	$P4_3$
Unit cell axes (Å)	69.3, 131.9, 118.0	108.9, 221.8, 129.2	85.1, 78.1, 100.7			96.7, 96.7, 54.6
Unit cell β angle (°)	102.9	90	113.4	113.6	113.5	90
Wavelength (Å)	0.9792	0.9330	0.9790	0.9790	0.9180	0.9790
Resolution range (Å)	50–2.9 (3.0–2.9)[b]	80–4.0 (4.2–4.0)[b]	80–2.0 (2.08–2.0)[b]	50–2.8 (2.9–2.8)[b]	50–2.5 (2.6–2.5)[b]	80–2.35 (2.45–2.35)[b]
Unique reflections	88,325[a] (8,594)[b]	13,572[c] (1,801)[b]	159,335[a] (17,800)[b]	58,751[a] (5,918)[b]	82,833[a] (9,221)[b]	21,259[c] (2,466)[b]
Completeness (%)	97.8 (97.6)[b]	99.8 (99.5)[b]	99.0 (99.8)[b]	99.2 (99.5)[b]	99.3 (99.7)[b]	99.8 (99.8)[b]
Redundancy	4.9 (5.1)[b]	8.0 (10.7)[b]	3.6 (3.8)[b]	3.8 (3.8)[b]	3.8 (3.8)[b]	11.7 (12.5)[b]
Mosaicity (°)	0.19	0.34	0.16	0.17	0.30	0.09
R_{sym} (%)	3.7 (65.1)[b]	20.7 (79.2)[b]	6.6 (60.6)[b]	8.1 (69.7)[b]	7.7 (75.7)[b]	8.0 (59.7)[b]
I/σ(I)	24.2 (3.3)[b]	8.9 (3.7)[b]	16.8 (3.4)[b]	16.8 (3.4)[b]	18.3 (3.5)[b]	27.5 (6.3)[b]
Refinement						
Nonhydrogen atoms	13,794	6,539	9,698			3,587
Rmsd bonds	0.010	0.010	0.010			0.010
Rmsd angles	1.28	1.25	1.04			1.08
R_{cryst} (%)	19.4	28.7	18.9			18.8
R_{free} (%)	24.3	36.1	22.4			22.5
Preferred[d] (%)	95.2	83.0	98.5			99.0
Allowed[d] (%)	4.3	11.1	1.5			1.0
Disallowed[d] (%)	0.5	5.9	0.0			0.0

[a] Friedel pairs not merged.
[b] Values in parentheses are for highest-resolution shell.
[c] Friedel pairs merged.
[d] Ramachandron plot statistics from PROCHECK (Laskowski et al., 1993).

normal RNA cleavage activity of Pol I, whereas the tWH domain stimulates Pol I processivity and binds DNA. The tWH domain binds both ssDNA and dsDNA, and these interactions are apparently not mutually exclusive. This observation is consistent with a previously reported nonspecific dsDNA binding affinity of Pol I (Bric et al., 2004). The tWH domain binds preferentially to the nontemplate single strand in the upstream promoter region. These results suggest that the tWH domain is involved in promoter opening and/or stabilization of the early transcription bubble by trapping the nontemplate DNA strand. The transcription bubble was mapped for the protozoan amoebae *Acanthamoeba castellani* around the transcription start site (Kahl et al., 2000), but the location of the initial Pol I transcription bubble in yeast remains unknown.

structure elements are indicated as bars (α helices) or arrows (β strands). Dashed vertical lines delineate A49 tWH subdomains WH1 and WH2. Proteolytic cleavage sites are marked with vertical arrows.
(B) Preparation of recombinant A49/34.5 domains. SDS-PAGE analysis of the dimerization module, A49 linker, and A49 tWH domain. An asterisk indicates an N-terminal degradation product.
(C) Ribbon model of the dimerization module crystal structure, which is part of a heterotetramer, as shown in Figure S2. A49 and A34.5 are depicted in light blue and magenta, respectively. A pseudo 2-fold axis is indicated. Figures were prepared with PyMOL (DeLano Scientific).
(D) Structure of the TFIIF dimerization module (Gaiser et al., 2000). Rap74 and Rap30 are depicted in light blue and magenta, respectively. TFIIF-specific features are indicated.
(E) Alignment of dimerization module sequences from Pol I, II and III. Structure-based alignments of sequences of *C. glabrata* A49 and A34.5 with their counterparts in Pol II (*H. sapiens* Rap74 and Rap30, *S. cerevisiae* Tfg1 and Tfg2) and in Pol III (*S. cerevisiae* C37 and C53). Secondary structure elements observed in the A49/34.5 and TFIIF crystal structures are shown (bars, α helices; arrows, β strands; lines, loops; dashed lines, disordered regions). For C37 and C53, predicted secondary structure elements are in gray. As the hit with the highest score, HHpred (Biegert et al., 2006; Soding et al., 2005) identified similarity of the *S. cerevisiae* C53 C-terminal residues 300–421 to the N-terminal residues 13–118 of human Rap30 ($p = 0.0006$). For *S. pombe* C37, the hit with the highest score was detected between C37 residues 26–145 and Rap74 residues 10–151 ($p = 0.0037$). The β4-β5 hairpin is apparently lacking in C37. Conserved residues are highlighted according to decreasing conservation from green, through orange, to yellow. Residues located in the hydrophobic heterodimer interface are generally conserved between A49/34.5, Rap74/30, and C37/53. Insertions are marked with numbers in the sequence. Mutated residues that abolish proper folding (Kuhn et al., 2007) are highlighted in red. Residues not present in the crystallized variant are underlined.

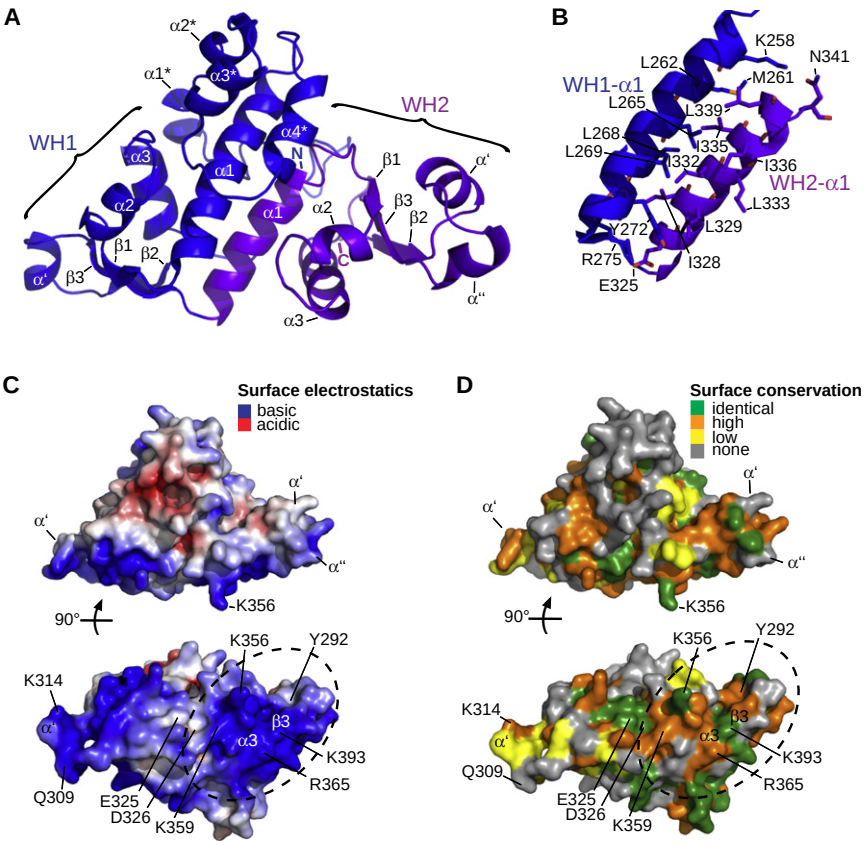

Figure 3. A Tandem Winged Helix Domain Structure in A49

(A) Ribbon diagram of the A49 tWH domain crystal structure. Subdomains WH1 and WH2 are displayed in blue and purple, respectively. For details, see Figure S3.

(B) The WH1-WH2 subdomain interface is formed by conserved residues.

(C) Surface charges. Red and blue indicate negative and positive charges, respectively, as calculated with APBS (Baker et al., 2001). A dashed circle indicates a positively charged conserved region. The top view is as in (A), and the bottom view is rotated around a horizontal axis by 90 degrees.

(D) Surface conservation. Decreasing conservation among seven different *Saccharomycotinae* species is indicated in green, orange, yellow, and gray. The views are as in (C).

Published results are consistent with a DNA-binding function of the tWH domain that may be relevant during transcription initiation and elongation in vivo. In *S. cerevisiae*, tWH residues 367–415 are responsible for a cold-sensitive phenotype, whereas A49 residues 1–119 in the dimerization domain and A34.5 are entirely dispensable in vivo (Beckouet et al., 2008). A yeast strain lacking residues 367–415 of the tWH domain was sensitive to 6-azauracil and mycophenolate and showed compromised Pol I processivity and a reduction of Pol I occupancy in the rDNA promoter and transcribed region (Beckouet et al., 2008). Mouse A49 is required for promoter-dependent transcription (Hanada et al., 1996). This was confirmed for *S. pombe* A49, and it was additionally shown that the region forming the tWH domain is responsible for a cold-sensitive phenotype (Nakagawa et al., 2003).

Pol I Resembles a Minimal Pol II-TFIIF-TFIIE Complex

Because TFIIF contains the dimerization module, but not the tWH domain, we searched for a tWH domain in another Pol II initiation factor. Two consecutive WH domains were found in TFIIE subunit β (Table S1): one that was structurally resolved (Okuda et al., 2000) and one that could be predicted (Figure S4) with HHPred (Soding et al., 2005). Modeling the arrangement of these two WH subdomains within a putative tWH structure revealed conserved hydrophobic residues in the subdomain interface and a putative conserved salt bridge between residues K86 and D147 (Figures 6 and S4). Consistent with these results, the mutation K86E impairs transcription (Tanaka et al., 2009). This suggests that TFIIEβ contains a tWH domain and that A49/34.5 combines features of TFIIF and TFIIE, which both contribute to Pol II promoter binding. TFIIF binds dsDNA (Groft et al., 1998), prevents nonspecific DNA binding to Pol II (Killeen and Greenblatt, 1992), and is required for initiation complex formation (Tan et al., 1995). TFIIE binds ssDNA (Okamoto et al., 1998) and dsDNA (Okuda et al., 2000; Tanaka et al., 2009). The ssDNA-binding activity requires TFIIEβ residues 257–277 (Okamoto et al., 1998), which correspond to the A49 C-terminal tail that is required for ssDNA binding to the A49 tWH domain in vitro (Figures 4E, 4F and S5D).

Pol III Contains TFIIF- and TFIIE-like Regions

There is also evidence that Pol III subunits resemble parts of TFIIF and TFIIE. The Pol III subcomplex C37/53 apparently forms a TFIIF-like dimerization module (Kuhn et al., 2007) and could be modeled with the A49/34.5 dimerization module structure (Figure 6). All β strands of a triple barrel were predicted (Figure 2E), and most hydrophobic core residues were conserved (Figure 2E). The related dimerization modules in A49/34.5, TFIIF, and C37/53 bind at similar locations on the Rpb2 side of Pol II (Chen et al., 2010; Fernández-Tornero et al., 2007; Kuhn et al., 2007). Unfortunately, we could not investigate this further because the structures could not be placed unambiguously into a previous Pol I electron microscopic density (Kuhn et al., 2007). Subunit C34 contains two consecutive WH subdomains (PDB 2DK8 and 2DK5) that could form a tWH domain, although the putative interface contains only three hydrophobic residues (V16, I23, and L94; Figure S4). Subunit C82 was also predicted to contain subsequent WH subdomains, but these include an additional N-terminal helix, as in the "extended WH domain" of TFIIEα (Meinhart et al., 2003). Thus, the C82/34 heterodimer (Lorenzen et al., 2007) may be distantly related to

the TFIIEα/β heterodimer. Consistent with these homologies, C37/53 and C82/34/31 function in transcription initiation (Brun et al., 1997; Kassavetis et al., 2010; Landrieux et al., 2006; Wang and Roeder, 1997). While this work was about to be completed, a bioinformatic analysis also suggested similarities between Pol III subunits and Pol II initiation factors (Carter and Drouin, 2010). These results indicate that Pol III is also related to a Pol II-TFIIF-TFIIE complex.

Polymerase Evolution and Promoter Specificity

Here, we provide evidence that Pol I- and Pol III-specific subcomplexes contain domains that are structurally related to parts of the Pol II initiation factors TFIIF and TFIIE (Figure 6). Our results and published data indicate that TFIIF- and TFIIE-related subcomplexes are involved in promoter binding. These subcomplexes, however, show substantial differences in their domain organization, suggesting that they evolved to allow for gene class-specific promoter recognition and apparently also for additional functions during elongation such as processivity.

EXPERIMENTAL PROCEDURES

Mass Spectrometry

A49/34.5, Pol I, and Pol IΔ were prepared as described (Kuhn et al., 2007; Fath et al., 2000). For native MS, the sample buffer was exchanged to a solution containing 160 or 880 mM ammonium acetate using centrifugal filter units (Millipore), and sample concentration was adjusted to 2 μM. MS was carried out on a Q-ToF I instrument (Lorenzen, 2007; van den Heuvel et al., 2006). The cone voltage was 150 V, and the needle voltage was 1.3 kV. The pressure in the source region was 10 mbar. Xenon was used as a collision gas with a pressure of 2×10^{-2} mbar (Lorenzen, 2007). Data were analyzed with MassLynx (Waters).

Recombinant Proteins

The genes for A49 and A34.5 were cloned sequentially into vector pET28b (Merck), resulting in a thrombin-cleavable N-terminal hexahistidine tag on A49. For bicistronic expression, a second ribosomal binding site was inserted before A34.5. The *S. cerevisiae* variant $A49_{1-110}/34.5_{1-156}$ was expressed for 18 hr at 18°C in *E. coli* BL21 (DE3) RIL cells (Agilent). Cells were harvested by centrifugation, resuspended in 50 ml buffer A (50 mM MES [pH 6.3], 300 mM NaCl, 10 mM β-mercaptoethanol, and 1 mM protease inhibitor mix containing 1 mM PMSF, 1 mM benzamidine, 200 μM pepstatin, and 60 μM leupeptin), and lysed by sonication. After centrifugation, the supernatant was loaded onto a 3 ml Ni-NTA column (QIAGEN) equilibrated with buffer A. The column was washed stepwise with 25 ml of buffer A containing 1 M NaCl and 15 ml of buffer A containing 10 mM imidazole. Protein was eluted with buffer A containing 200 mM imidazole. Eluted fractions were diluted 3-fold with buffer A lacking NaCl and incubated with thrombin (1 U protease/1 mg protein) for 16 hr at 4°C. A MonoQ column (GE Healthcare) was equilibrated with buffer B (50 mM MES [pH 6.3], 100 mM NaCl, and 5 mM DTT). Protein was obtained in the flowthrough, applied to a MonoS column (GE Healthcare), equilibrated with buffer B, and eluted with a linear gradient from 100 mM to 1 M NaCl. The sample was concentrated and applied to a Superose 12 column (GE Healthcare) equilibrated with 50 mM Tris [pH 7.5], 150 mM NaCl, and 5 mM DTT. Pooled peak fractions were concentrated to 20 mg/ml. Purification of other dimerization module variants was as above except $A49/34.5_{1-156}$, which was purified as described (Kuhn et al., 2007). The A49 linker variant $A49_{105-187}$ was expressed for 4 hr at 37°C in *E. coli* BL21 (DE3) RIL cells and purified as above. Purified linker was concentrated to ~10 mg/ml. A49 tWH domain variants were cloned, expressed, and purified essentially as above, except that the MonoQ step was skipped and gel filtration was carried out in 50 mM Tris [pH 7.0], 150 mM NaCl, and 5 mM DTT. Pooled peak fractions were concentrated to 16 mg/ml. Limited proteolysis was performed as described (Geiger et al., 2008; Hubbard, 1998), using trypsin and chymotrypsin as proteases.

Crystal Structure Determinations

Crystals for *C. glabrata* variant $A49_{1-99}/34.5_{25-143}$ were grown at 20°C in hanging drops, using as reservoir solution 22% PEG 3350 and 250 mM sodium fluoride. Crystals were cryoprotected by a stepwise transfer to reservoir solution containing 7%–20% PEG 400 and flash cooled by plunging into liquid nitrogen. Two additional methionines were introduced as nondisruptive point mutations (A49-V72M and A34.5-L55M). Selenomethionine-labeled variant was prepared as described (Budisa et al., 1995; Meinhart et al., 2003) and crystallized at 20°C using as reservoir solution 20% PEG 3350 and 50 mM Tris (pH 7.5). Crystals reached a size of 300 μm × 60 μm × 60 μm and were cryopreserved in reservoir solution containing 23% glycerol. SAD diffraction data were obtained at the Swiss Light Source and processed with XDS (Kabsch, 1993) (Table S1). Programs SHELXD/HKL2MAP (Pape and Schneider, 2004; Schneider and Sheldrick, 2002) detected 16 selenium sites, two sites in each of the eight heterodimers in the asymmetric unit. SHARP (de La Fortelle and Bricogne, 1997) was used for SAD phasing, and SOLOMON (Abrahams and Leslie, 1996) was used for density modification. NCS averaging was performed with DM (Cowtan, 1994). The model was built with COOT (Emsley and Cowtan, 2004) and refined with PHENIX (Afonine et al., 2005) and BUSTER (Blanc et al., 2004) (Table S1). Crystals for *C. glabrata* variant $A49_{1-119}/34.5_{25-143}$ were grown using as reservoir solution 20% PEG 3350 and 50 mM Tris (pH 7.0). Diffraction data were collected at the European Synchrotron Radiation Facility ESRF (Table S1). The structure was solved by molecular replacement with PHASER (McCoy et al., 2005), using the dimerization module structure as a search model. A NCS-averaged map was generated with PARROT (Cowtan, 2010). The model was manually extended and refined at 4 Å resolution (Table 1).

Crystals for *S. cerevisiae* $A49_{171-403}$ were grown at 4°C in hanging drops, using as reservoir solution 25% PEG 3350 and 100 mM Tris (pH 8.5). Crystals reached a size of 400 μm × 80 μm × 80 μm, were harvested in reservoir solution containing 10%–20% PEG 400, and were flash cooled. In addition to present methionine residues (M170, M256, and M397), two additional methionines were introduced (L178M and L261M), and selenomethionine-labeled protein was prepared as above and crystallized at 4°C using as reservoir solution 24% PEG 3350 and 50 mM Tris (pH 8.0). MAD data were obtained at ESRF and processed as above (Table 1), and the structure was determined essentially as above, except that MAD phasing was used. Twenty selenium sites were detected, four in each of five tWH domains in the asymmetric unit. The model was built automatically with ARP/wARP (Langer et al., 2008), manually adjusted, and refined (Table 1). Crystals for variant $A49_{155-399}$ were grown at 4°C in hanging drops, using as reservoir solution 100 mM Tris (pH 8.5), 26% PEG 4000, and 200 mM lithium sulfate monohydrate, reached a size of 80 μm × 30 μm × 30 μm, and were cryopreserved in 20% ethylene glycol. The structure was solved by molecular replacement using the A49 tWH structure as a search model and was refined (Table 1).

RNA Cleavage and Extension Assays

RNA cleavage assays were performed with a nucleic acid scaffold containing a three-nucleotide noncomplementary 3′ overhang (Kuhn et al., 2007), except that the reaction buffer contained 20 mM MES (pH 6.0). RNA extension assays were performed with a preannealed minimal nucleic acid scaffold as described (Kuhn et al., 2007).

Electrophoretic Mobility Shift Assay

EMSA assays were performed essentially as described (http://www.labs.fhcrc.org/hahn/methods) except that the buffer was 20 mM Tris (pH 8.0), 150 mM NaCl, 1 mM DTT, 40 μg/ml Heparin, 5 mM $MgCl_2$, and 4% glycerol. dsDNA was prepared by annealing equimolar amounts of complementary oligonucleotides (for sequences, see Table S2). For dual wavelength experiments, 6-FAM-labeled dsDNA and Cy5-labeled ssDNA were used. 1 pmol of DNA was incubated with 100 pmol of protein for 15 min at 4°C. DNA-protein complexes were resolved by gel electrophoresis (0.2 pmol per lane) in an 8%–0.08% acrylamid-bisacrylamid gel (25 mM Tris [pH 8.2], 19 mM glycine, 2.5% glycerol, and 0.5 mM DTT) at 150 V for 50 min and visualized with a Typhoon 9400 phosphoimager (GE Healthcare).

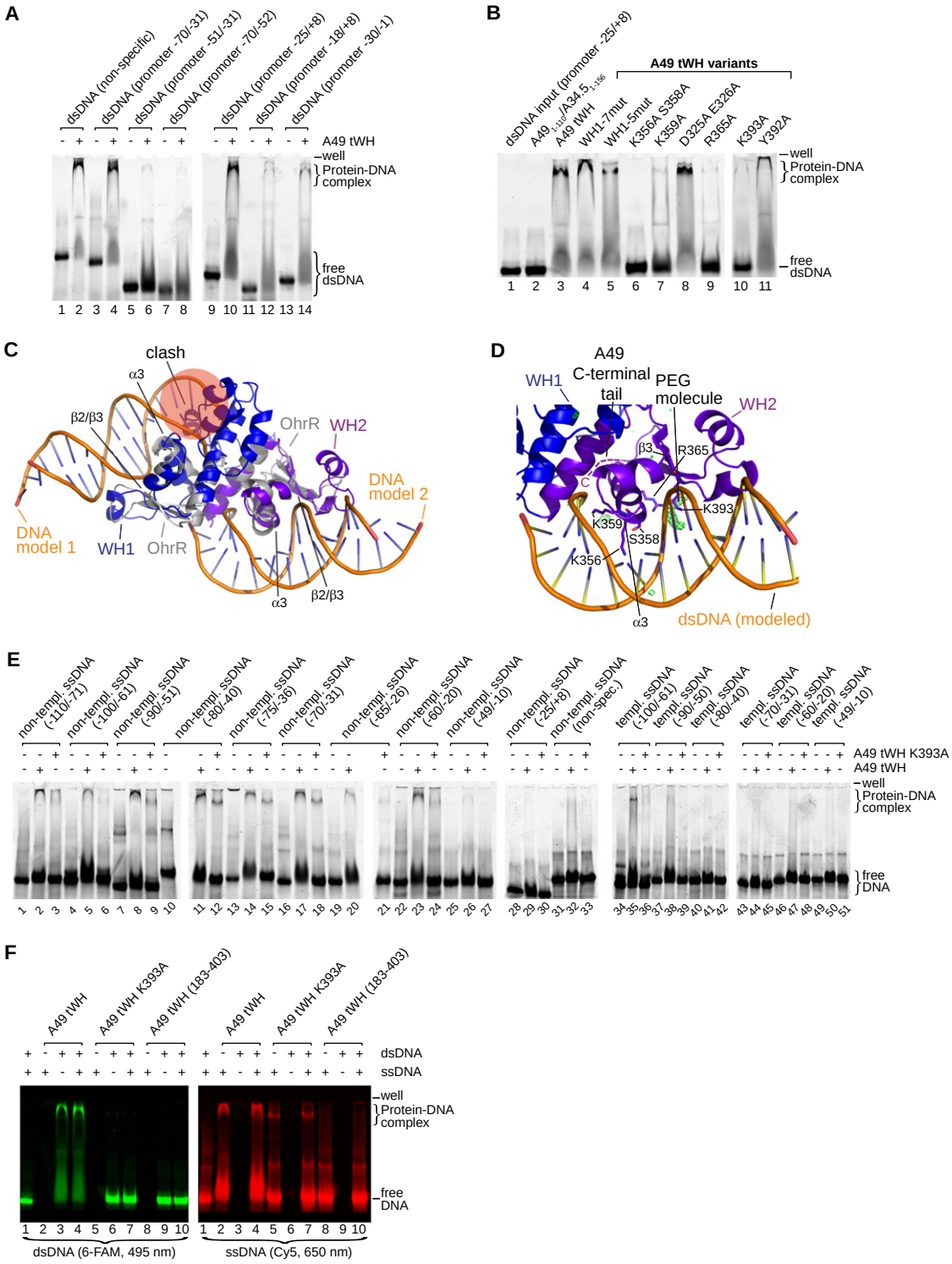

Figure 4. The A49 tWH Domain Binds DNA

(A) EMSA analysis of dsDNA binding. dsDNAs included a 47 base pair (bp) DNA of unrelated sequence (nonspecific) and regions of the Pol I promoter (Choe et al., 1992; Kulkens et al., 1991; Musters et al., 1989) as indicated by positions relative to the transcription start site +1. The latter included extended domain 2 of upstream promoter element (−70 to −31) and parts of it (−70 to −52, −51 to −31) and the core element domain 1 (−25 to +8) and parts of it (−18 to +8, −30 to −1).

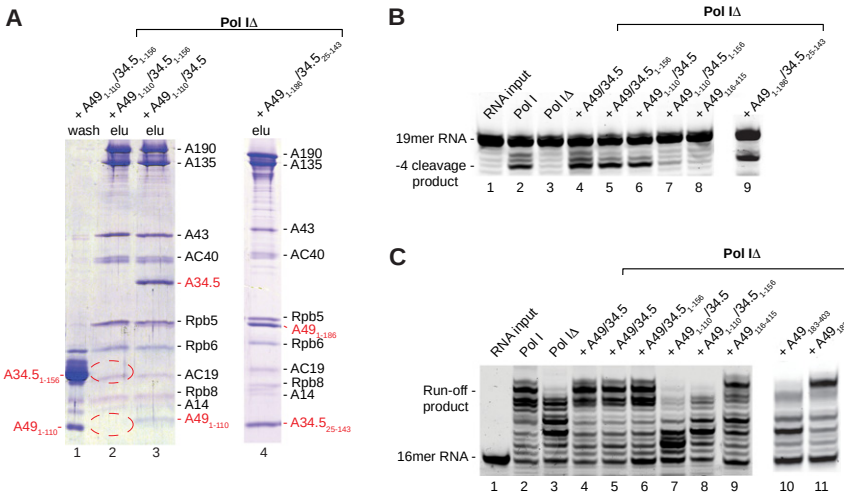

Figure 5. Different A49/34.5 Regions Mediate Pol I RNA Cleavage and Processivity

(A) The A49 linker or the A34.5 C-terminal tail is required to anchor A49/34.5 on Pol I. Wash and elution (elu) fractions of pull-down experiments were analyzed by SDS-PAGE. Pol IΔ subunits are indicated in black; A49 and A34.5 variants are in red. The dimerization module $A49_{1–110}$/$34.5_{1–156}$ did not bind Pol IΔ (lanes 1 and 2), but binding was enabled by its extension with the A49 linker ($A49_{1–186}$/$34.5_{25–143}$, lane 4) or with the A34.5 tail ($A49_{1–110}$/34.5, lane 3). Wash samples were TCA precipitated (lane 1). For controls, see Figure S6.

(B) The dimerization module stimulates RNA cleavage if it is anchored to Pol I. Pol I removes four nucleotides from the RNA (lane 2). This activity depends on A49/34.5, as Pol IΔ is defective in cleavage (lane 3). Cleavage can be restored by addition of A49/34.5 (lane 4) or its variants containing the dimerization module and either the A49 linker (lane 9) or the A34.5 tail (lane 6).

(C) Pol I processivity on a minimal DNA-RNA scaffold (Kuhn et al., 2007) requires the A49 tWH domain. Pol I elongates RNA (lane 1) by 12 nucleotides to the run-off product (lane 2), whereas Pol IΔ can not (lane 3). Processivity is restored by addition of recombinant A49/34.5 (lanes 4 and 5) or variants that contain the tWH domain (lanes 6, 9, and 11) and depends on the tWH C-terminal basic tail (compare lanes 10 and 11).

Protein Interaction Assay

15 μg of Pol IΔ was incubated with 5 μg of A49/34.5 variant in buffer C (5 mM HEPES [pH 7.8], 60 mM ammonium sulfate, 1 mM $MgCl_2$, 10 μM $ZnCl_2$, and 10 μM β-mercaptoethanol) for 25 min at 15°C. 40 μl of Ni-NTA beads (QIAGEN) equilibrated with buffer C was added, and the mixture was incubated for 1 hr at 4°C. Pol I bound to the beads via a hexahistidine tag on its A43 subunit. Beads were centrifuged and washed four times with buffer C containing 5 mM imidazole and 0.2% NP-40. Protein complexes were eluted with buffer C containing 200 mM imidazole and were incubated at 95°C for 5 min in SDS-PAGE loading buffer.

ACCESSION NUMBERS

Coordinates and structure factors have been deposited in the Protein Data Bank with accession codes 3NFG and 3NFF for the A49/34.5 dimerization module variants $A49_{1–99}$/$34.5_{25–143}$ and $A49_{1–119}$/$34.5_{25–143}$, respectively, and 3NFI and 3NFH for the A49 C-terminal tandem winged helix domain variants $A49_{171–403}$ and $A49_{155–399}$, respectively.

SUPPLEMENTAL INFORMATION

Supplemental Information includes seven figures and two tables and can be found with this article online at doi:10.1016/j.molcel.2010.07.028.

ACKNOWLEDGMENTS

We thank A. Kusser, S. Jennebach, C. Blattner, S. Benkert, M. Seizl, L. Wenzeck, E. Vojnic, F. Martinez, A. Schüller, and other members of the Cramer laboratory. We thank A. Jawhari, C. Kuhn, Y. Kraus, J. Söding, and L. Puchbauer (Gene Center Munich), as well as T. Hader (University of Regensburg). S.R.G. was supported by the Elitenetzwerk Bayern graduate program NanoBioTechnology. P.C. was supported by the Deutsche Forschungsgemeinschaft, the SFB646, the SFB TR5, the Nanosystems Initiative Munich NIM, the Jung-Stiftung, and the Fonds der chemischen Industrie. K.L. was supported by Roche (Penzberg). A.J.R.H. was supported by the Centre for Biomedical Genetics and the Netherlands Proteomics Centre, embedded in the Netherlands Genomics Initiative. We thank J. Basquin and K. Saldaña from the crystallization facility of the MPI for Biochemistry, Martinsried. Part of this work was conducted at beamline PXI of the Swiss Light Source at the Paul Scherrer Institute, Villigen, Switzerland. We thank the Structural Biology group at ESRF Grenoble, as well as the Protein Structure Factory at BESSY, Berlin.

Received: May 4, 2010
Revised: June 15, 2010
Accepted: June 29, 2010
Published: August 26, 2010

(B) Subdomain WH2 surface mutations impair dsDNA binding. The interaction of tWH variants with promoter dsDNA −25 to +8 (A) was analyzed by EMSA. In lane 2, the dimerization module was used as negative control. The tWH domain $A49_{183–415}$ and its variants were used in lane 3 and lanes 4–11, respectively (WH1-7mut, Q309A/F310A/R312A/S313A/K314A/D135A/R316A; WH1-5mut, S289A/E292A/Q309A/S313A/K314A; point mutant variants as indicated).

(C) Alternative superpositions of OhrR-DNA complex structure (PDB 1Z9C; Hong et al., 2005) onto A49 tWH subdomains WH1 (blue) or WH2 (purple). OhrR is depicted in gray and DNA in orange. For details, see Figure S5.

(D) Modeled interaction between A49 WH2 and dsDNA based on superposition in (C). Residues important for DNA binding are depicted. A polyethylene glycole molecule was observed in a difference electron density map (green mesh, contoured at 3.8σ).

(E) EMSA analysis of ssDNA binding. The A49 tWH domain and its variant K393A that is defective in dsDNA binding were used. The ssDNA sequences correspond to the template or nontemplate strand of the Pol I promoter as indicated by positions relative to +1.

(F) Binding of dsDNA and ssDNA is not mutually exclusive. 6-FAM-labeled dsDNA (−25 to +8) and Cy5-labeled ssDNA (−70 to −31) were visualized within the same gel at 495 nm (green, left) and 650 nm (red, right), respectively. tWH variants included wild-type $A49_{183–415}$ and its variant K393A, which is defective in dsDNA binding, and its variant $A49_{183–403}$, which lacks the C-terminal basic tail.

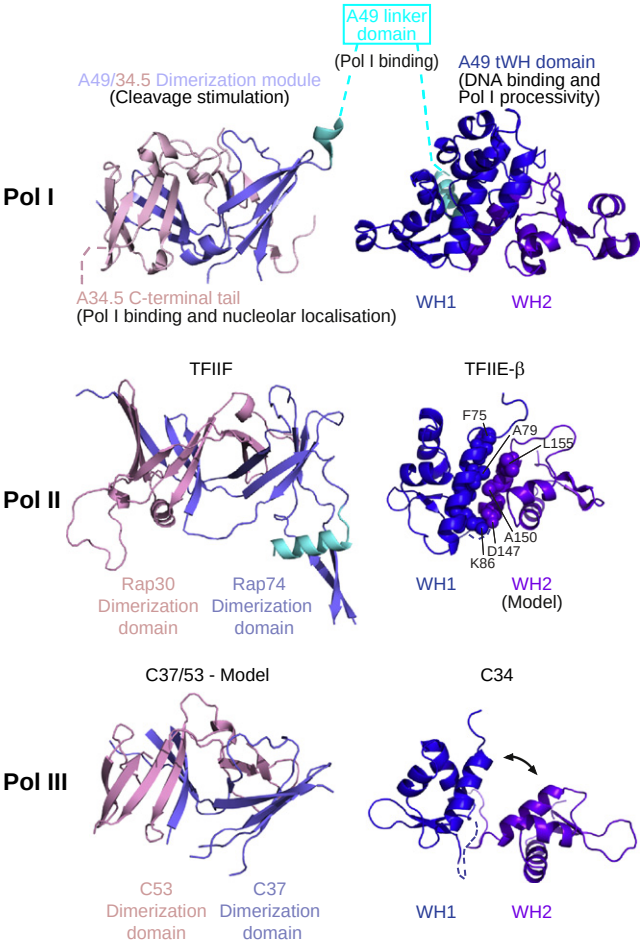

Figure 6. Evolutionary Relationship between Pol I/III-Specific Subunits and Pol II Factors

On the top, a model of the Pol I subcomplex A49/34.5 based on the structural information is shown. Colors are as follows: A49 dimerization domain, light blue; A34.5 dimerization domain, magenta; A49 linker, cyan; A49 WH1, blue; A49 WH2, purple. For functional assignments, compare text. The A34.5 C-terminal tail is important for nucleolar localization (Ushijima et al., 2008). In the middle, the A49/34.5-related domains in Pol II initiation factors are shown. The structure of the TFIIF Rap74/30 dimerization module (Gaiser et al., 2000) and the proposed TFIIEβ WH1 subdomain (Okuda et al., 2000) are known. The structure of the proposed TFIIEβ WH2 was modeled based on the A49 tWH structure. Some of the residues that potentially contribute to the tWH subdomain interface (F75, L78, A79, V82, A150, L151, and L155) are indicated with spheres. On the bottom, models for A49/34.5-related domains in Pol III-specific subunits are shown. The C37/53 dimerization module was modeled based on the A49/34.5 dimerization module structure with the use of a secondary structure-based alignment (Figure 2E). Structures for C34 WH1 and WH2 have been determined (PDB codes 2DK8 and 2DK5). Based on the modeling, it is unlikely that the WH1-WH2 interface is stable, and the two subdomains may move with respect to each other. For details, see Figure S4.

REFERENCES

Abrahams, J.P., and Leslie, A.G. (1996). Methods used in the structure determination of bovine mitochondrial F1 ATPase. Acta Crystallogr. D Biol. Crystallogr. *52*, 30–42.

Afonine, P.V., Grosse-Kunstleve, R.W., and Adams, P.D. (2005). A robust bulk-solvent correction and anisotropic scaling procedure. Acta Crystallogr. D Biol. Crystallogr. *61*, 850–855.

Armache, K.J., Mitterweger, S., Meinhart, A., and Cramer, P. (2005). Structures of complete RNA polymerase II and its subcomplex, Rpb4/7. J. Biol. Chem. *280*, 7131–7134.

Baker, N.A., Sept, D., Joseph, S., Holst, M.J., and McCammon, J.A. (2001). Electrostatics of nanosystems: application to microtubules and the ribosome. Proc. Natl. Acad. Sci. USA *98*, 10037–10041.

Beckouet, F., Labarre-Mariotte, S., Albert, B., Imazawa, Y., Werner, M., Gadal, O., Nogi, Y., and Thuriaux, P. (2008). Two RNA polymerase I subunits control the binding and release of Rrn3 during transcription. Mol. Cell. Biol. *28*, 1596–1605.

Benesch, J.L., Ruotolo, B.T., Simmons, D.A., and Robinson, C.V. (2007). Protein complexes in the gas phase: technology for structural genomics and proteomics. Chem. Rev. *107*, 3544–3567.

Berman, H.M., Westbrook, J., Feng, Z., Gilliland, G., Bhat, T.N., Weissig, H., Shindyalov, I.N., and Bourne, P.E. (2000). The Protein Data Bank. Nucleic Acids Res. *28*, 235–242.

Biegert, A., Mayer, C., Remmert, M., Soding, J., and Lupas, A.N. (2006). The MPI Bioinformatics Toolkit for protein sequence analysis. Nucleic Acids Res. *34*, W335–W339.

Blanc, E., Roversi, P., Vonrhein, C., Flensburg, C., Lea, S.M., and Bricogne, G. (2004). Refinement of severely incomplete structures with maximum likelihood in BUSTER-TNT. Acta Crystallogr. D Biol. Crystallogr. *60*, 2210–2221.

Brennan, R.G. (1993). The winged-helix DNA-binding motif: another helix-turn-helix takeoff. Cell *74*, 773–776.

Bric, A., Radebaugh, C.A., and Paule, M.R. (2004). Photocross-linking of the RNA polymerase I preinitiation and immediate postinitiation complexes: implications for promoter recruitment. J. Biol. Chem. *279*, 31259–31267.

Brueckner, F., Ortiz, J., and Cramer, P. (2009). A movie of the RNA polymerase nucleotide addition cycle. Curr. Opin. Struct. Biol. *19*, 294–299.

Brun, I., Sentenac, A., and Werner, M. (1997). Dual role of the C34 subunit of RNA polymerase III in transcription initiation. EMBO J. *16*, 5730–5741.

Budisa, N., Steipe, B., Demange, P., Eckerskorn, C., Kellermann, J., and Huber, R. (1995). High-level biosynthetic substitution of methionine in proteins by its analogs 2-aminohexanoic acid, selenomethionine, telluromethionine and ethionine in Escherichia coli. Eur. J. Biochem. *230*, 788–796.

Carter, R., and Drouin, G. (2010). The increase in the number of subunits in eukaryotic RNA polymerase III relative to RNA polymerase II is due to the permanent recruitment of general transcription factors. Mol. Biol. Evol. *27*, 1035–1043.

Chen, Z.A., Jawhari, A., Fischer, L., Buchen, C., Tahir, S., Kamenski, T., Rasmussen, M., Lariviere, L., Bukowski-Wills, J.C., Nilges, M., et al. (2010). Architecture of the RNA polymerase II-TFIIF complex revealed by cross-linking and mass spectrometry. EMBO J. *29*, 717–726.

Choe, S.Y., Schultz, M.C., and Reeder, R.H. (1992). In vitro definition of the yeast RNA polymerase I promoter. Nucleic Acids Res. *20*, 279–285.

Cowtan, K. (1994). Joint CCP4 and ESF-EACBM Newsletter on Protein Crystallography *31*, 34–38.

Cowtan, K. (2010). Recent developments in classical density modification. Acta Crystallogr. D Biol. Crystallogr. *66*, 470–478.

Cramer, P., Armache, K.J., Baumli, S., Benkert, S., Brueckner, F., Buchen, C., Damsma, G.E., Dengl, S., Geiger, S.R., Jasiak, A.J., et al. (2008). Structure of eukaryotic RNA polymerases. Annu. Rev. Biophys. *37*, 337–352.

De Carlo, S., Carles, C., Riva, M., and Schultz, P. (2003). Cryo-negative staining reveals conformational flexibility within yeast RNA polymerase I. J. Mol. Biol. *329*, 891–902.

de La Fortelle, E., and Bricogne, G. (1997). Maximum-likelihood heavy-atom parameter refinement for multiple isomorphous replacement and

multiwavelength anomalous diffraction methods. Methods Enzymol. *276*, 472–494.

Emsley, P., and Cowtan, K. (2004). Coot: model-building tools for molecular graphics. Acta Crystallogr. D Biol. Crystallogr. *60*, 2126–2132.

Fath, S., Milkereit, P., Podtelejnikov, A.V., Bischler, N., Schultz, P., Bier, M., Mann, M., and Tschochner, H. (2000). Association of yeast RNA polymerase I with a nucleolar substructure active in rRNA synthesis and processing. J. Cell Biol. *149*, 575–590.

Fernández-Tornero, C., Böttcher, B., Riva, M., Carles, C., Steuerwald, U., Ruigrok, R.W., Sentenac, A., Müller, C.W., and Schoehn, G. (2007). Insights into transcription initiation and termination from the electron microscopy structure of yeast RNA polymerase III. Mol. Cell *25*, 813–823.

Flores, O., Ha, I., and Reinberg, D. (1990). Factors involved in specific transcription by mammalian RNA polymerase II. Purification and subunit composition of transcription factor IIF. J. Biol. Chem. *265*, 5629–5634.

Gaiser, F., Tan, S., and Richmond, T.J. (2000). Novel dimerization fold of RAP30/RAP74 in human TFIIF at 1.7 A resolution. J. Mol. Biol. *302*, 1119–1127.

Gajiwala, K.S., and Burley, S.K. (2000). Winged helix proteins. Curr. Opin. Struct. Biol. *10*, 110–116.

Geiger, S.R., Kuhn, C.D., Leidig, C., Renkawitz, J., and Cramer, P. (2008). Crystallization of RNA polymerase I subcomplex A14/A43 by iterative prediction, probing and removal of flexible regions. Acta Crystallogr. Sect. F Struct. Biol. Cryst. Commun. *64*, 413–418.

Groft, C.M., Uljon, S.N., Wang, R., and Werner, M.H. (1998). Structural homology between the Rap30 DNA-binding domain and linker histone H5: implications for preinitiation complex assembly. Proc. Natl. Acad. Sci. USA *95*, 9117–9122.

Hanada, K., Song, C.Z., Yamamoto, K., Yano, K., Maeda, Y., Yamaguchi, K., and Muramatsu, M. (1996). RNA polymerase I associated factor 53 binds to the nucleolar transcription factor UBF and functions in specific rDNA transcription. EMBO J. *15*, 2217–2226.

Heck, A.J. (2008). Native mass spectrometry: a bridge between interactomics and structural biology. Nat. Methods *5*, 927–933.

Holm, L., and Sander, C. (1995). Dali: a network tool for protein structure comparison. Trends Biochem. Sci. *20*, 478–480.

Hong, M., Fuangthong, M., Helmann, J.D., and Brennan, R.G. (2005). Structure of an OhrR-ohrA operator complex reveals the DNA binding mechanism of the MarR family. Mol. Cell *20*, 131–141.

Hubbard, S.J. (1998). The structural aspects of limited proteolysis of native proteins. Biochim. Biophys. Acta *1382*, 191–206.

Huet, J., Wyers, F., Buhler, J.M., Sentenac, A., and Fromageot, P. (1976). Association of RNase H activity with yeast RNA polymerase A. Nature *261*, 431–433.

Jasiak, A.J., Armache, K.J., Martens, B., Jansen, R.P., and Cramer, P. (2006). Structural biology of RNA polymerase III: subcomplex C17/25 X-ray structure and 11 subunit enzyme model. Mol. Cell *23*, 71–81.

Kabsch, W. (1993). Automatic processing of rotation diffraction data from crystals of initially unknown symmetry and cell constants. J. Appl. Cryst. *26*, 795–800.

Kahl, B.F., Li, H., and Paule, M.R. (2000). DNA melting and promoter clearance by eukaryotic RNA polymerase I. J. Mol. Biol. *299*, 75–89.

Kassavetis, G.A., Prakash, P., and Shim, E. (2010). The C53/C37 subcomplex of RNA polymerase III lies near the active site and participates in promoter opening. J. Biol. Chem. *285*, 2695–2706.

Kenney, L.J. (2002). Structure/function relationships in OmpR and other winged-helix transcription factors. Curr. Opin. Microbiol. *5*, 135–141.

Killeen, M.T., and Greenblatt, J.F. (1992). The general transcription factor RAP30 binds to RNA polymerase II and prevents it from binding nonspecifically to DNA. Mol. Cell. Biol. *12*, 30–37.

Kuhn, C.D., Geiger, S.R., Baumli, S., Gartmann, M., Gerber, J., Jennebach, S., Mielke, T., Tschochner, H., Beckmann, R., and Cramer, P. (2007). Functional architecture of RNA polymerase I. Cell *131*, 1260–1272.

Kulkens, T., Riggs, D.L., Heck, J.D., Planta, R.J., and Nomura, M. (1991). The yeast RNA polymerase I promoter: ribosomal DNA sequences involved in transcription initiation and complex formation in vitro. Nucleic Acids Res. *19*, 5363–5370.

Landrieux, E., Alic, N., Ducrot, C., Acker, J., Riva, M., and Carles, C. (2006). A subcomplex of RNA polymerase III subunits involved in transcription termination and reinitiation. EMBO J. *25*, 118–128.

Langer, G., Cohen, S.X., Lamzin, V.S., and Perrakis, A. (2008). Automated macromolecular model building for X-ray crystallography using ARP/wARP version 7. Nat. Protoc. *3*, 1171–1179.

Laskowski, R.A., Moss, D.S., and Thornton, J.M. (1993). Main-chain bond lengths and bond angles in protein structures. J. Mol. Biol. *231*, 1049–1067.

Liljelund, P., Mariotte, S., Buhler, J.M., and Sentenac, A. (1992). Characterization and mutagenesis of the gene encoding the A49 subunit of RNA polymerase A in Saccharomyces cerevisiae. Proc. Natl. Acad. Sci. USA *89*, 9302–9305.

Lorenzen, K. (2007). Optimizing macromolecular tandem mass spectrometry of large non-covalent complexes using heavy collison gases. Int. J. Mass Spectrom. *268*, 198–206.

Lorenzen, K., Vannini, A., Cramer, P., and Heck, A.J. (2007). Structural biology of RNA polymerase III: mass spectrometry elucidates subcomplex architecture. Structure *15*, 1237–1245.

McCoy, A.J., Grosse-Kunstleve, R.W., Storoni, L.C., and Read, R.J. (2005). Likelihood-enhanced fast translation functions. Acta Crystallogr. D Biol. Crystallogr. *61*, 458–464.

Meinhart, A., Blobel, J., and Cramer, P. (2003). An extended winged helix domain in general transcription factor E/IIE alpha. J. Biol. Chem. *278*, 48267–48274.

Moss, T., Langlois, F., Gagnon-Kugler, T., and Stefanovsky, V. (2007). A housekeeper with power of attorney: the rRNA genes in ribosome biogenesis. Cell. Mol. Life Sci. *64*, 29–49.

Musters, W., Knol, J., Maas, P., Dekker, A.F., van Heerikhuizen, H., and Planta, R.J. (1989). Linker scanning of the yeast RNA polymerase I promoter. Nucleic Acids Res. *17*, 9661–9678.

Nakagawa, K., Hisatake, K., Imazawa, Y., Ishiguro, A., Matsumoto, M., Pape, L., Ishihama, A., and Nogi, Y. (2003). The fission yeast RPA51 is a functional homolog of the budding yeast A49 subunit of RNA polymerase I and required for maximizing transcription of ribosomal DNA. Genes Genet. Syst. *78*, 199–209.

Okamoto, T., Yamamoto, S., Watanabe, Y., Ohta, T., Hanaoka, F., Roeder, R.G., and Ohkuma, Y. (1998). Analysis of the role of TFIIE in transcriptional regulation through structure-function studies of the TFIIEbeta subunit. J. Biol. Chem. *273*, 19866–19876.

Okuda, M., Watanabe, Y., Okamura, H., Hanaoka, F., Ohkuma, Y., and Nishimura, Y. (2000). Structure of the central core domain of TFIIEbeta with a novel double-stranded DNA-binding surface. EMBO J. *19*, 1346–1356.

Pape, T., and Schneider, T.R. (2004). HKL2MAP: a graphical user interface for phasing with SHELX programs. Acta Crystallogr. D Biol. Crystallogr. *37*, 843–844.

Schneider, T.R., and Sheldrick, G.M. (2002). Substructure solution with SHELXD. Acta Crystallogr. D Biol. Crystallogr. *58*, 1772–1779.

Soding, J., Biegert, A., and Lupas, A.N. (2005). The HHpred interactive server for protein homology detection and structure prediction. Nucleic Acids Res. *33*, W244–W248.

Tan, S., Conaway, R.C., and Conaway, J.W. (1995). Dissection of transcription factor TFIIF functional domains required for initiation and elongation. Proc. Natl. Acad. Sci. USA *92*, 6042–6046.

Tanaka, A., Watanabe, T., Iida, Y., Hanaoka, F., and Ohkuma, Y. (2009). Central forkhead domain of human TFIIE beta plays a primary role in binding double-stranded DNA at transcription initiation. Genes Cells *14*, 395–405.

Ushijima, R., Matsuyama, T., Nagata, I., and Yamamoto, K. (2008). Nucleolar targeting of proteins by the tandem array of basic amino acid stretches identified in the RNA polymerase I-associated factor PAF49. Biochem. Biophys. Res. Commun. *369*, 1017–1021.

van den Heuvel, R.H., van Duijn, E., Mazon, H., Synowsky, S.A., Lorenzen, K., Versluis, C., Brouns, S.J., Langridge, D., van der Oost, J., Hoyes, J., and Heck, A.J. (2006). Improving the performance of a quadrupole time-of-flight instrument for macromolecular mass spectrometry. Anal. Chem. *78*, 7473–7483.

Wang, Z., and Roeder, R.G. (1997). Three human RNA polymerase III-specific subunits form a subcomplex with a selective function in specific transcription initiation. Genes Dev. *11*, 1315–1326.