Biochimica et Biophysica Acta 1829 (2013) 265–273

Contents lists available at SciVerse ScienceDirect

Biochimica et Biophysica Acta

journal homepage: www.elsevier.com/locate/bbagrm

Invited review

TFIIB-related factors in RNA polymerase I transcription ☆

Bruce A. Knutson *, Steven Hahn

Fred Hutchinson Cancer Research Center, Division of Basic Sciences, 1100 Fairview Ave. N, P.O. Box 19024, Mailstop A1-162, Seattle, WA 98109, USA

ARTICLE INFO

Article history:
Received 27 July 2012
Received in revised form 20 August 2012
Accepted 21 August 2012
Available online 30 August 2012

Keywords:
Rrn7
TAF1B
Pol I
TFIIB
Brf
rDNA transcription

ABSTRACT

Eukaryotic RNA polymerases (Pol) I, II, III and archaeal Pol use a related set of general transcription factors to recognize promoter sequences and recruit Pol to promoters and to function at key points in the transcription initiation mechanism. The TFIIB-like general transcription factors (GTFs) function during several important and conserved steps in the initiation pathway for Pols II, III, and archaeal Pol. Until recently, the mechanism of Pol I initiation seemed unique, since it appeared to lack a GTF paralogous to the TFIIB-like proteins. The surprising recent discovery of TFIIB-related Pol I general factors in yeast and humans highlights the evolutionary conservation of transcription initiation mechanisms for all eukaryotic and archaeal Pols. These findings reveal new roles for the function of the Pol I GTFs and insight into the function of TFIIB-related factors. Models for Pol I transcription initiation are reexamined in light of these recent findings. This article is part of a Special Issue entitled: Transcription by Odd Pols.

1. Introduction

RNA polymerase (Pol) I is the most specialized Pol as it exclusively transcribes the ribosomal DNA (rDNA) gene unit into a single rRNA transcript [1,2]. The large rRNA transcript or rRNA precursor is processed into the 18S, 28S, and 5.8S rRNAs that are key components of ribosomes [1,2]. Pol I transcription is robust and accounts for the majority of total RNA in cells [3], and its upregulation in human cells is a hallmark of cancer [2,4–6]. rDNA transcription takes place in a unique nuclear compartment called the nucleolus [7], where the majority of ribosome biogenesis occurs, producing over a million ribosomes per cell [3,8]. In most eukaryotes, the nucleolus contains hundreds of tandem ribosomal genes, located at one or more chromosomal loci, but only a subset of these genes is actively transcribed at any one time [1,2].

2. Pol I subunit composition

To synthesize rRNA, eukaryotic cells use a 14-subunit Pol I enzyme [9–11]. All three eukaryotic Pols (Pols I, II, III) contain 12 subunits that are either shared between the Pols, or are functionally and structurally related Pol-specific paralogs [9–11] (Table 1). Each Pol contains a core set of five shared subunits: Rpb5, 6, 8, 10, and 12 [9–12]. The Pol subunits AC40 and AC19 are shared between Pols I and III, and are paralogs to the Pol II subunits Rpb3 and Rpb11 [13–16]. Seven subunits are unique to Pol I, which include the two largest subunits that form the catalytic site (Rpa190 and Rpa135), the Rpb9/TFIIS/Rpc11-related paralog A12.2 [17], and the two Pol I stalk subunits A14 and A43 that are paralogous to Rpb4/7 and Rpc17/Rpc25 subunits from Pols II and III respectively [18–20].

The Pol I subunits A34.5 and A49 comprise a heterodimeric protein that contain domains structurally and functionally related to the Pol II general transcription factors (GTFs) TFIIF- and TFIIE [21–24]. A34.5 and the N-terminal domain of A49 dimerize with a triple barrel fold that is structurally related to the dimerization domain of the TFIIF subunits, Tfg1 and 2, and the Pol III TFIIF-like subunits, Rpc37 and Rpc53 [22]. A large unstructured linker connects the A49 dimerization domain to a tandem wing helix (tWH) domain that is structurally related to the tWH domains of the TFIIE subunit Tfa2 and the Pol III TFIIE-like subunit Rpc34. In terms of functional conservation, genetic evidence suggests that A34.5 and A49 perform TFIIF- and TFIIE-like functions in Pol I transcription elongation and PIC formation [22,23,25,26]. Thus, both Pols I and III evolved to stably incorporate TFIIE- and TFIIF-like GTFs, unlike the more loosely associated Pol II GTFs, TFIIE and TFIIF [21,27].

Given the sequence and structural similarity of each Pol, it is not surprising that the TFIIE and TFIIF subunits and their Pol I and III paralogs localize to evolutionary conserved positions on Pol. For Pols II and III, the dimerization domain of TFIIF and Rpc37/Rpc53 is positioned on the lobe domains of their respective Pols [28–30].

Abbreviations: Pol, RNA polymerase; GTF, general transcription factor; rRNA, ribosomal RNA; rDNA, ribosomal DNA; CE, core element; TSS, transcription start site; UAS, upstream activating sequence; UCE, upstream control element; UAF, upstream activating factor; UBF, upstream binding factor; TBP, TATA binding protein; CF, core factor; SL1, selectivity factor 1; tWH, tandem winged helix; HMG, high mobility group protein; ChIP, chromatin immunoprecipitation; PIC, preinitiation complex; CTD, C-terminal domain; Brf, TFIIB-related factor; RPG, ribosomal protein genes

☆ This article is part of a Special Issue entitled: Transcription by Odd Pols.

* Corresponding author. Tel.: +1 206 667 5746; fax: +1 206 667 6497.
E-mail address: bknutson@fhcrc.org (B.A. Knutson).

1874-9399/$ – see front matter © 2012 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.bbagrm.2012.08.003

Table 1
Yeast RNA polymerase I and paralogous counterparts.

	Pol I	Pol III	Pol II	Archaea	Bacteria
Core[1]	Rpa190	Rpc160	Rpb1	Rpo1	β'
	Rpa135	Rpc128	Rpb2	Rpo2	β
	AC40	AC40	Rpb3	Rpo3	α
	AC19	AC19	Rpb11	Rpo11	α
	Rpb6	Rpb6	Rpb6	Rpo6	ω
	Rpb5	Rpb5	Rpb5	Rpo5	
	Rpb8	Rpb8	Rpb8	Rpo8	
	Rpb10	Rpb10	Rpb10	Rpo10	
	Rpb12	Rpb12	Rpb12	Rpo12	
	Rpa12	Rpc11	Rpb9	TFS	
Stalk	Rpa14	Rpc17	Rpb4	Rpo4	
	Rpa43	Rpc25	Rpb7	Rpo7	
	A49-NT[2]	Rpc53	TFIIFα[4]		
GTFs	A34.5	Rpc37	TFIIFβ[5]		
	A49-CT[3]	Rpc34	TFIIEβ[6]	TFE	

[1]Shared Pol I subunits are outlined; [2]NT, N-terminal dimerization domain; [3]CT, C-terminal tWH domain; [4]TFIIFα, Tfg1; [5]TFIIFβ, Tfg2; [6]TFIIEβ, Tfa1

Likewise, previous EM studies and recent combined chemical crosslinking and mass spectrometry studies of Pol I show a similar position for the A49/A34.5 dimer on the Pol I lobe domain [31–33]. The crosslinking studies also showed that the A49 TFIIE-like tWH domain spans the Pol I active site cleft, bridging the clamp and protrusion domains [33], analogous to the positions observed for the TFIIE and Rpc34 tWH domains on Pol II and III [12,34,35]. Together, these findings suggest that each Pol transcription system follows a common structural, functional, and evolutionary framework.

3. Pol I transcription initiation factors

Transcription from rDNA repeats is controlled by a bipartite promoter with two essential cis-regulatory elements (Fig. 1). In yeast, Pol I promoters contain a core element (CE) that overlaps the transcription start site (TSS) and an upstream activating sequence (UAS) that is centered ~100 bp upstream of the TSS [1,2]. Yeast Pol I is recruited to the promoter via four different initiation factors including the UAS-binding upstream activating factor (UAF), TATA-binding protein (TBP), the CE-binding protein core factor (CF), and Rrn3. UAF consists of six subunits: UAF30, Rrn9, Rrn10, Rrn5, and histones H3 and H4 [36,37]. CF is composed of three subunits that include Rrn6, Rrn7, and Rrn11 [38,39].

Pol I preinitiation complex (PIC) formation in yeast begins with UAF binding to the UAS followed by TBP recruitment, which next facilitates recruitment of CF and Rrn3-bound Pol I to the promoter [40]. In vitro, UAF remains stably bound to the promoter of immobilized templates after initiation and operates as a reinitiation scaffold for subsequent rounds of transcription while CF, TBP, and Rrn3 are released [40]. In vivo, chromatin immunoprecipitation (ChIP) studies show that Pol I crosslinks to the promoter and rDNA coding regions as expected, whereas UAF, CF, and Rrn3 are exclusively promoter associated [41]. In the absence of Pol I and Rrn3, UAF, CF, and TBP remain stably bound to the promoter, suggesting that CF and TBP build a reinitiation scaffold in vivo [42].

Pol I GTFs are only weakly conserved between yeast and humans, but their promoters retain a similar bipartite structure. In humans, Pol I promoters contain an upstream control element positioned ~100 bp upstream of a TSS overlapping CE [1,2]. The human CE is targeted by a CF ortholog called selectivity factor 1 [1,2]. Three subunits of SL1 that include TAF1A, B, and C, that are orthologs of yeast Rrn6, 7, and 11, respectively [1,43,44]. Orthologous CF and SL1 components share ≤16% protein sequence identity. SL1 also contains three additional subunits that include TBP, the human-specific TAF1D subunit [45], and TAF12, a component of the conserved Pol II coactivators SAGA and TFIID [46,47]. The mechanistic roles of TAF1D and TAF12 subunits remain unclear.

Most components of yeast UAF beside the histone components also lack clear orthologs to the human Pol I factors. The UCE of human rDNA promoters is targeted by a dimer called the Upstream Binding Factor (UBF) [1,48]. UBF contains a series of tandem high mobility group (HMG) boxes [49] that bind and bend DNA, forming a loop that bridges the UCE and CE [50,51]. None of the UAF subunits share homology with UBF, but a UBF-related HMG box protein termed Hmo1 binds throughout the yeast rDNA locus [52,53], but it is still not clear whether UAF functions are replaced by UBF in humans or vice versa. A summary of yeast and human Pol I initiation factor orthology and homology is listed in Table 2.

Current PIC assembly models for mammalian rDNA promoters suggest that UBF binds to the promoters first and then facilitates SL1 recruitment, while other models suggest that UBF and SL1 bind together, or that SL1 facilitates UBF recruitment [2,54,55]. UBF and SL1 promoter binding is immediately followed by recruitment of Rrn3 bound Pol I [2,54,55]. After human Pol I initiation and promoter clearance, UBF and SL1 remain stably bound to the promoter, creating a reinitiation scaffold like those found in yeast for continued cycles of rDNA transcription [2,54,55].

4. Discovery of a Pol I TFIIB-related factor

Soon after the discovery of TFIIB [56–58], a TFIIB-related factor (BRF1) was discovered among the Pol III GTFs [59–61]. Brf1 is a subunit of the Pol III GTF TFIIIB that also contains subunits Bpd1 and TBP. Previous searches for a Pol I TFIIB-related factor focused on components of CF and SL1 since both interact with TBP and Pol I. However, no detectable TFIIB homology could be detected among the Pol I factors by traditional protein homology search methods. Given the conservation of the core Pol enzymes and GTFs, Pol I seemed unique in lacking a protein with homology to TFIIB. For almost two decades, it became accepted that either Pol I uses a TFIIB-like factor that has diverged beyond recognition by sequence homology, or that the Pol I machinery evolved a unique initiation mechanism distinct from Pol II and III [62–64].

Over the past decade, more sensitive sequence homology detection methods have been developed, such as HHpred [65], that are designed specifically to detect homology between distantly related proteins. Two groups independently reexamined the question of whether the Pol I transcription system required a TFIIB-related factor. Using HHpred, both groups discovered sequence and predicted structural homology between the yeast CF subunit Rrn7 and its human ortholog TAF1B with known TFIIB-like proteins in archaea and eukaryotes [63,64]. It was found that the N-terminus of Rrn7 and TAF1B shares homology with TFIIB, TFB, and the N-terminus of Brf [63,64]. Remarkably, the sequence identity between Rrn7 and other TFIIB-like proteins was very low, ranging from 8 to 13% [63,64]. Even more surprising, the protein sequence identity between Rrn7 and TAF1B was only 16% [63,64]. However, functional studies described below showed that these proteins share conserved function and structure. Together, these new findings uncovered the missing link between the Pol I transcription system and its Pol II and III counterparts, and shed new light on the Pol I initiation factor function.

5. Comparison of TFIIB family proteins

The overall domain architectures of TFIIB, TFB and Brf proteins are quite similar, but with several intriguing differences. All three TFIIB-related family members contain a zinc ribbon domain followed by a linker region that connects to tandem cyclin fold repeats [66,67] (Fig. 2A). Brf1 contains an extended C-terminal domain (CTD) immediately after the second cyclin repeat [66,68]. This unique Brf1 CTD folds into an elongated structure with three homology regions