Critical Reviews in Biochemistry and Molecular Biology, 2011; 46(1): 27–40

informa
healthcare

REVIEW ARTICLE

Archaeal RNA polymerase and transcription regulation

Sung-Hoon Jun, Matthew J. Reichlen, Momoko Tajiri, and Katsuhiko S. Murakami

Department of Biochemistry and Molecular Biology, The Pennsylvania State University, University Park, PA, USA

Abstract

To elucidate the mechanism of transcription by cellular RNA polymerases (RNAPs), high-resolution X-ray crystal structures together with structure-guided biochemical, biophysical, and genetics studies are essential. The recently solved X-ray crystal structures of archaeal RNAP allow a structural comparison of the transcription machinery among all three domains of life. The archaea were once thought of closely related to bacteria, but they are now considered to be more closely related to the eukaryote at the molecular level than bacteria. According to these structures, the archaeal transcription apparatus, which includes RNAP and general transcription factors (GTFs), is similar to the eukaryotic transcription machinery. Yet, the transcription regulators, activators and repressors, encoded by archaeal genomes are closely related to bacterial factors. Therefore, archaeal transcription appears to possess an intriguing hybrid of eukaryotic-type transcription apparatus and bacterial-like regulatory mechanisms. Elucidating the transcription mechanism in archaea, which possesses a combination of bacterial and eukaryotic transcription mechanisms that are commonly regarded as separate and mutually exclusive, can provide data that will bring basic transcription mechanisms across all life forms.

Keywords: *Archaea; RNA polymerase; transcription; X-ray crystallography; regulation*

Introduction

The last decade marked a revolution in understanding of the molecular detail of cellular RNA polymerases (RNAPs) from bacteria, archaea, and eukaryotes and their respective transcription mechanisms. The structure of bacterial *Thermus aquaticus* RNAP core enzyme, the first high-resolution X-ray crystal structure of any cellular RNAP revealed the structure–function relationship of the universally conserved core part of cellular RNAP (Zhang *et al.*, 1999; Darst, 2001). The structural studies of *T. aquaticus* holoenzyme and holoenzyme–fork junction promoter DNA complex have explained how the promoter recognition σ factor associates with the core enzyme and how the bacterial promoter DNA is recognized by holoenzyme to form the "transcription ready" promoter open complex (Murakami and Darst, 2003; Murakami *et al.*, 2002a; Murakami *et al.*, 2002b). The *Saccharomyces cerevisiae* RNAP II (Pol II) structures that include 10 subunits form (Cramer *et al.*, 2001), 12 subunits form (Armache *et al.*, 2005), complexed with a general transcription factor (GTF) TFIIB (Kostrewa *et al.*, 2009; Liu *et al.*, 2010) and the transcription elongation form (Gnatt *et al.*, 2001), have shown how the Pol II looks like and how it forms the promoter complex and how it transcribes RNA. These structures boosted our understanding of the structural basis of eukaryotic transcription. The structure of *Sulfolobus solfataricus* RNAP, the first X-ray crystal structure of archaeal RNAP solved recently, has completed the suite of cellular RNAPs of all three domains of life (Hirata *et al.*, 2008b).

Archaea was discovered as one of three branches of life, and since then, interest has grown for many reasons (Woese and Fox, 1977; Pace, 1997). The archaeal transcription system has been characterized as a hybrid of eukaryotic and bacterial transcription systems (Bell and Jackson, 1998). The archaeal basal transcription apparatus is very similar to that of eukaryote (Zillig *et al.*, 1978; Langer *et al.*, 1995), but its transcriptional regulatory factors are similar to those of bacteria (Brinkman *et al.*, 2003; Ouhammouch, 2004). In this review article, we will compare the architectural features of cellular

Address for Correspondence: Katsuhiko Murakami, Department of Biochemistry and Molecular Biology, The Pennsylvania State University, University Park, PA 16802, USA. E-mail: kum14@psu.edu

(Received 05 July 2010; revised 05 November 2010; accepted 05 November 2010)

ISSN 1040-9238 print/ISSN 1549-7798 online © 2011 Informa Healthcare USA, Inc.
DOI: 10.3109/10409238.2010.538662

RNAPs from three domains of life. We will then discuss the mechanism of transcription regulation in archaea, which are informed by high-resolution structure as well as biochemistry and genetics experiments.

Archaeal RNAP

The development of molecular biology has shown that molecular criteria like DNA sequence comparison of ribosomal RNA genes reveals more precise evolutionary relationships among the organisms than the classical morphological or cytological criteria. The classification of organisms into the "three domains"—bacteria, archaea, and eukaryote—is based on the molecular criteria and widely accepted (Woese *et al.*, 1990). Interestingly, archaea are so-called prokaryotic cells in terms of cytological features, but are thought to have common ancestry with eukaryotes considering their molecular features. Particularly, proteins involved in gene maintenance and expression are similar to those of eukaryotes. It has been shown that the basal transcription machinery including RNAP in archaea is closely related to the transcription machinery found in eukaryotes based on their subunit compositions (Rowlands *et al.*, 1994; Langer *et al.*, 1995) (Table 1). In the purified archaeal RNAP, the protein complex consists of 11–13 subunits depending on the species (Zillig *et al.*, 1978; Huet *et al.*, 1983) and it became clear, based on DNA cloning of each subunit coding gene, that the amino acid sequence similarities between archaeal and eukaryotic RNAPs are closer (Huet *et al.*, 1983; Langer *et al.*, 1995). Finally, in 2008, it became possible to compare the X-ray crystal structures of RNAP from the three domains of life and to begin to understand their evolutionary relationships (Zhang *et al.*, 1999; Cramer *et al.*, 2001; Hirata *et al.*, 2008b).

Table 1. Subunit composition of cellular RNAPs.

	Eukaryotic Pol II	Archaeal RNAP	Bacterial RNAP
Class I subunit	Rpb1	A′+A″	β′
	Rpb2	B	β
	Rpb3	D	αI
	Rpb6	K	ω
	Rpb11	L	αII
Class II subunit	Rpb4	F	
	Rpb5	H	
	Rpb7	E	
	Rpb8	G	
	Rpb10	N	
	Rpb12	P	
Class III subunit	Rpb9*Gdown1**	Rpo13**	δ**

*Found only in the eukaryotic Pol II. **Found only in certain species and these subunits are not orthologs

The X-ray crystal structures of archaeal RNAPs have been determined from two *Sulfolobus* sp., *S. solfataricus* (Hirata *et al.*, 2008b; Hirata and Murakami, 2009) and *S. shibatae* (Korkhin *et al.*, 2009) and the cryoelectron microscopy structure has been determined from *Pyrococcus furiosus* (Kusser *et al.*, 2008). Two X-ray crystal structures are very similar and are composed of 13 subunits with a molecular weight of about 380 kDa. The overall shape of *S. solfataricus* RNAP resembles a "crab claw" that includes 11 subunits with a protruding stalk that consists of E and F subunits (Figure 1). Based on a structural comparison among the archaeal, bacterial, and

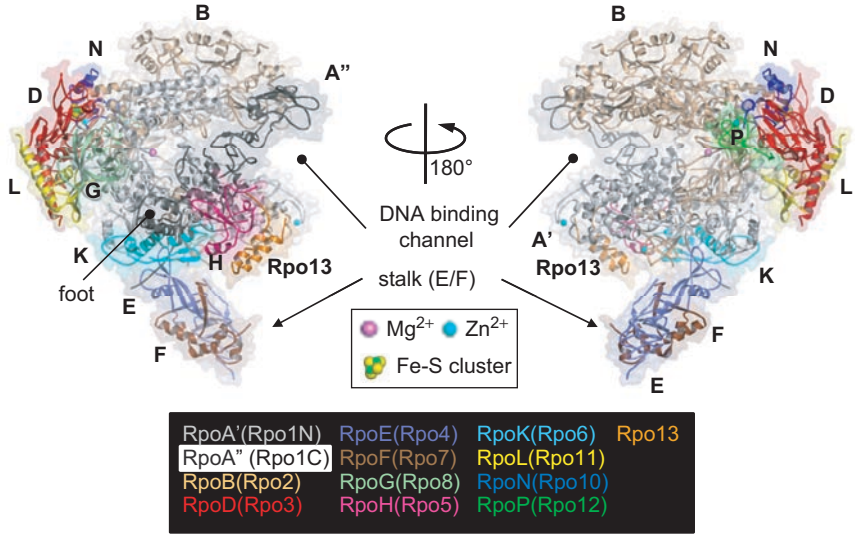

Figure 1. The 13-subunit *Sulfolobus solfataricus* RNAP structure (PDB: 3HKZ). The alpha-carbon backbone is shown as cartoon model, along with the transparent molecular surface. Each subunit is denoted by a unique color and labeled. Two sets of nomenclatures, one for traditional and the other is based on the eukaryotic terminology proposed by Korkhin *et al.*, are shown. Various structural features discussed in the text are also labeled.

eukaryotic RNAPs, it was determined that the dimensions of the double-stranded DNA binding channel and the architecture of active site of archaeal RNAP are highly conserved. However, in archaea, the largest subunit is divided into two polypeptides, A′ and A′′ subunits, which are encoded by separate genes in an operon (Langer *et al.*, 1995). Sequence alignments reveal that archaeal A′ and A′′ correspond to the N-terminal two-thirds and the C-terminal one-third of the Rpb1 subunit of Pol II (Table 1), respectively, and that the junction between A′ and A′′ is positioned at the "foot" domain (Cramer *et al.*, 2001). The C- and N-termini of A′ and A′′ form a four α-helix bundle domain, composed of one α-helix from C-terminus of A′ and three α-helices from N-terminus of A′′ (Figure 1) (Hirata *et al.*, 2008b).

Several archaeal RNAP subunits form stable subcomplexes. For examples, the D and L subunits form a heterodimer and function as a platform for the assembly of subunits (A′, A′′, and B) that comprise the active site (Goede *et al.*, 2006). The D/L subcomplex is located on the opposite side of the opening of the claw. Interestingly, the D subunit contains a 4Fe-4S cluster-binding motif that has been structurally characterized by the atomic resolution X-ray crystal structure of D/L subcomplex in addition to the entire *S. solfataricus* RNAP structure (Hirata *et al.*, 2008b; Hirata and Murakami, 2009). Three Cys residues (C183, C203, and C209) are ligands to the 3Fe-4S cluster and one additional C206 is positioned near the Fe-S cluster, suggesting that the cluster may exist as a 4Fe-4S *in vivo*. The Fe-S cluster is located ~45 Å from the enzyme active site, which suggests a structural role rather than a catalytic one. Furthermore, the *in vitro* site-directed mutagenesis study has shown that the Fe-S cluster plays a role in supporting the structural integrity of the D subunit and that it is essential in the formation of D/L subcomplex (Hirata *et al.*, 2008b). Interestingly, the 4Fe-4S cluster-binding motif is not conserved in all archaeal RNAPs. In fact, the amino acid residues holding the Fe-S cluster in the D subunit characterize a specific evolutionary lineage of archaea (Hirata and Murakami, 2009).

The E and F subunits form a stalk-like E/F heterodimer that binds to the core part of RNAP (Figure 1). It may modulate the position of the RNAP claw, thereby making the *S. solfataricus* RNAP adopt a closed clamp conformation. The E/F subcomplex is also known to be involved in both transcription initiation and elongation; it stimulates DNA melting and interacts with the newly synthesized RNA transcript through the RNA-binding motifs in the E and F subunits (Naji *et al.*, 2007; Werner, 2007). The RNA–protein interaction has been demonstrated by studies on its Pol II counterpart, Rpb7/Rpb4 subcomplex *in vitro* (Meka *et al.*, 2005; Ujvari and Luse, 2006). However, the interaction between RNA and this subcomplex has not been confirmed *in vivo*. The genetic study of hyperthermophilic archaeon, *Thermococcus kodakarensis*, has shown that the F subunit coding gene *rpoF* is not essential but *ΔrpoF* cells show a temperature sensitive phenotype (Hirata *et al.*, 2008a). RNAP preparations purified from *ΔrpoF* cells lacked subunit F and also subunit E and a transcription factor TFE that copurifies with RNAP from wild-type cells. But *in vitro*, this mutant RNAP exhibited no differences from wild-type RNAP in promoter-dependent transcription, abortive transcript synthesis, transcript elongation, or termination. The E/F subcomplex could be a binding platform of the TFE. This is consistent with TFE stimulation of archaeal RNAP activity requiring subunit E (Naji *et al.*, 2007; Ouhammouch *et al.*, 2004) and with reports of stimulatory transcription factor interactions with the homologous complexes in Pol I, II, and III. The extensions formed by A43 plus A14 in Pol I, and by C25 plus C17 in Pol III interact with polymerase-specific transcription initiation factors that recruit Pol I and Pol III to the appropriate promoters (Peyroche *et al.*, 2000; Kassavetis *et al.*, 2001). Given these observations, it seems most likely that the extensions formed by the archaeal subunits E plus F, and eukaryotic subunits A43 plus A14 (Pol I), Rpb4 plus Rpb7 (Pol II), and C25 plus C17 (Pol III) provide targets for some transcription factor binding, and so facilitate RNAP recruitment and transcription factor activation of the transcription machinery embodied in the core structures of these enzymes.

Structural comparison of RNAPs from three domains of life

The overall shape of bacterial, archaeal, and eukaryotic RNAPs resembles a crab claw (Figure 2). The largest 2–3 subunits form the most part of each claw-arm, generating a cleft for binding double-stranded DNA (Cramer *et al.*, 2001; Hirata *et al.*, 2008b; Zhang *et al.*, 1999). The enzyme active site is located at the bottom of the cleft that coordinates a catalytic metal Mg^{2+} by three invariant Asp residues found in the absolutely-conserved NADFDGD motif. The architecture around the cleft including the active site is highly conserved among all cellular RNAPs, which suggests that the catalytic mechanism of RNA synthesis from bacteria to human is conserved (Darst, 2001). But there are significant differences on the surfaces of these structures. The most significant difference is that archaeal RNAP and all three types of eukaryotic RNAPs have a protruding stalk-like structure that is absent from bacterial RNAP (Figure 2) (Armache *et al.*, 2005; Hirata *et al.*, 2008b; Kuhn *et al.*, 2007; Jasiak *et al.*, 2006).

Based on structural comparison, the subunits of cellular RNAPs can be characterized into three classes (Table 1). Class I subunits are conserved in all life forms, class II subunits are shared between archaeal and eukaryotic RNAPs, class III subunits are unique to each domain and except Rpb9, the others are found in certain species.

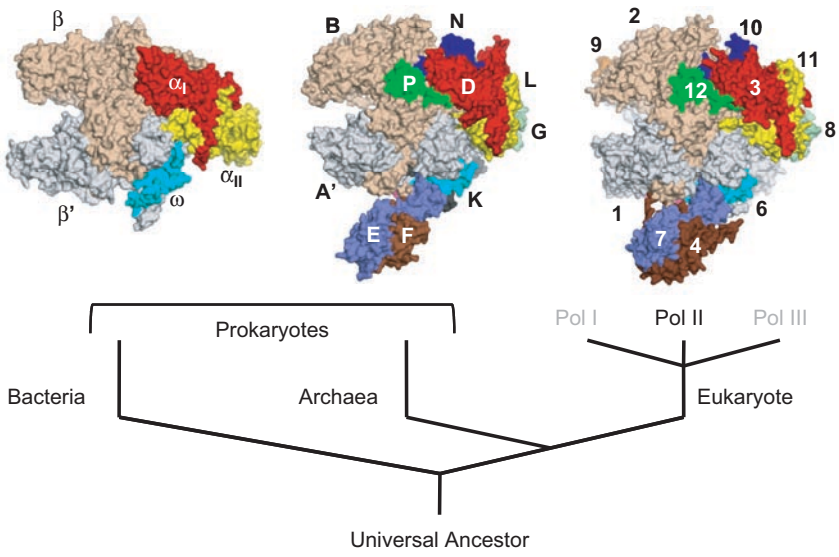

Figure 2. Surface representations of cellular RNAP structures from bacteria (left, *Thermus aquaticus* core enzyme), archaea (middle, *Sulfolobus solfataricus* RNAP), and eukaryote (right, *Saccharomyces cerevisiae* Pol II). Each subunit is denoted by a unique color and labeled. Orthologous subunits are depicted with the same color.

All five subunits of the bacterial RNAP core enzyme are class I indicating they all play essential roles in cellular RNAP functions. Structural similarities calculated by secondary structure matching algorithm have shown that the structures of class I subunits of eukaryote are more similar with archaeal subunits than those of bacteria (Hirata *et al.*, 2008b). Especially, the second largest Rpb2 subunit of *S. cerevisiae* Pol II shows about 90% structural similarity with the B subunit of *S. solfataricus* RNAP whereas 65% with the β subunit of *T. aquaticus* RNAP (Hirata *et al.*, 2008b). The differences in the structures of class I subunits between eukaryote and bacteria come from the extensive interactions with small subunits like class II and class III on the periphery. Class II subunits decorate the outside of class I subunits and are involved in the proper folding and assembly of RNAP (Werner and Weinzierl, 2002). The individual structures of class II subunits as well as their interactions with other subunits are highly conserved between archaea and eukaryote (Hirata *et al.*, 2008b). For E, N, and P subunits, the three smallest class II subunits in *S. solfataricus* RNAP, the structural similarities with the counterparts of *S. cerevisiae* Pol II are >90%. Although there is no class II subunit in bacteria RNAP, there are several motifs generated from class I subunits, which are located at the equivalent positions to the motifs from class II subunits of archaeal and eukaryotic RNAPs (Table 1). For example, the C-terminal tail of P subunit in archaeal RNAP (Rpb12 in Pol II) occupies the space between domain 2 of archaeal D subunit (Rpb3 in Pol II) and flap domain of B subunit (Rpb2 in Pol II) (Figure 3A and 3B). P subunit is essential subunit for the assembly of archaeal RNAP (Werner and Weinzierl, 2002). In bacterial RNAP, two loops in α_I subunit are present at the corresponding space of C-terminal tail of P subunit (Figure 3C). These loops likely help the assembly of α_I subunit and β subunit as C-terminal tail of P subunit does in archaeal RNAP.

Rpb9 of Pol II belongs to the class III subunit that is found only in Pol II (Table 1 and Figure 4C). Rpb9 plays an important role in maintaining transcriptional fidelity by mediating the intrinsic nuclease activity of Pol II (Nesser *et al.*, 2006) and it is also involved in the repair of DNA damages in the actively transcribed genes, termed transcription-coupled repair (Li and Smerdon, 2002). Although Rpb9 is highly conserved in eukaryotic organisms, yeast null mutants of *RPB9* have only a limited growth defect (Woychik *et al.*, 1991), whereas it is required for viability in higher organisms like *Drosophila* (Harrison *et al.*, 1992) showing again its intriguing position in evolution.

The δ subunit of bacterial RNAP, the Rpo13 of archaeal RNAP and Gdown1 of eukaryotic Pol II belong to the class III subunit. RNAP isolated from several Gram (+) bacteria, including *Bacillus subtilis*, contains an additional δ subunit (~20 kDa) in the core enzyme. The δ subunit consists with the N-terminal domain and negatively charged C-terminal tail. The NMR structure of the N-terminal domain has been determined (Motackova *et al.*, 2010), but its binding site on RNAP is unknown. *In-vitro* transcription assays have shown that the δ subunit either increases or decreases activities of transcription depending on promoters, and it may potentially influence the isomerization between the closed complex and the transcription-competent open complex (Lopez de Saro *et al.*, 1995).

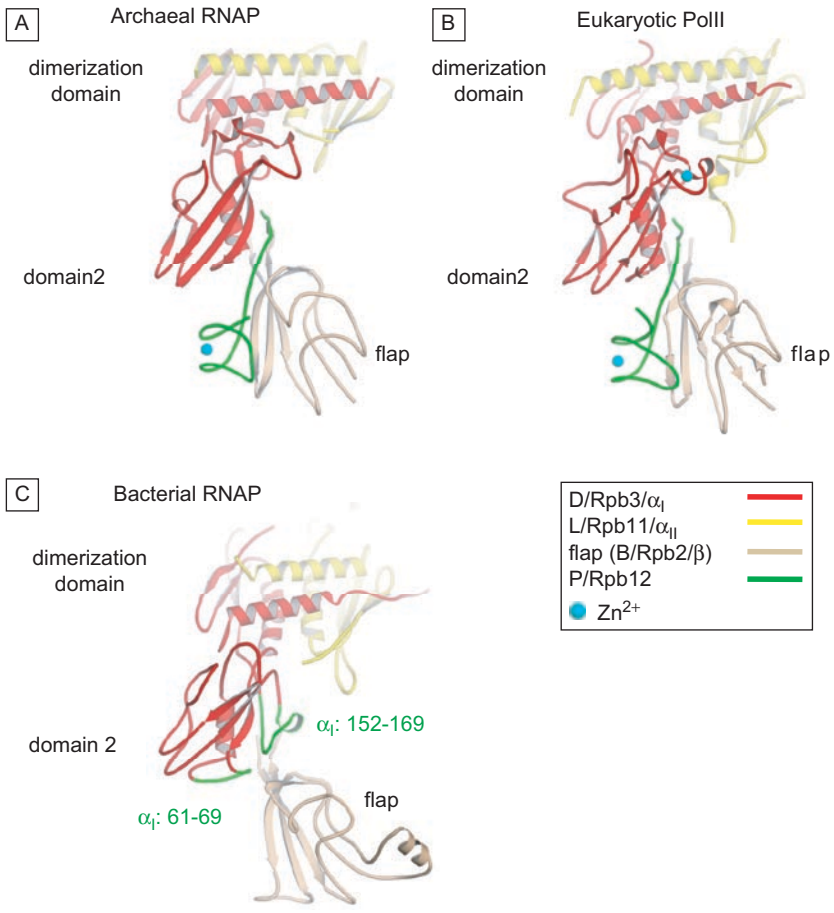

Figure 3. Interface of the D and P subunits and flap domain. (A) Ribbon representation of the archaeal domain 2 (D subunit), P subunit, and flap domain (B subunit) is shown (color coding of ribbons is indicated). A gap between domain 2 and flap domain is filled by P subunit, which forms a β-addition motif (in which a strand from one subunit is added to a β sheet of another). (B) Ribbon representation of the eukaryotic Pol II domain 2 (Rpb3), Rpb12, and flap domain (Rpb2) is shown. The orientation is the same as shown in the A. (C) Ribbon representation of the domain 2 (α_I subunit) and flap domain (β subunit) is shown. Bacterial α_I subunit residues 61–69 and 152–169 are present to the equivalent position of the C-terminal tail of P subunit, which buttresses α_I domain 2 and β flap interaction.

The Rpo13 has been identified as a new subunit of archaeal RNAP, which is only found in the order *Sulfolobales* and is located at a groove between the H subunit and the clamp head domain of the A′ subunit (Figures 1 and 4D) (Korkhin *et al.*, 2009). It has a helix-turn-helix motif and has been suggested to bind DNA. The role of Rpo13 in archaeal RNAP is an open question. (i) Rpo13 might have been introduced into the RNAP of last archaeal common ancestor to meet the evolutionary need of some subgroup of archaea. In this case its function would be a new one that does not exist in the subunits of eukaryotic RNAPs. (ii) Rpo13 might function as a "built-in" transcription factor-like A49 and A34.5 of eukaryotic Pol I and C37 and C53 of eukaryotic Pol III that are related to TFIIF (Kuhn *et al.*, 2007; Jasiak *et al.*, 2006). (iii) Rpo13 might play a similar role with the Rpb5 jaw domain of Pol II because they occupy similar locations in RNAPs (Figure 4C and 4D) and both were suggested to have DNA binding activities.

The Pol II isolated from metazoan cells contains an additional tightly associated polypeptide Gdown1, which is about 43 kDa molecular weight (Hu *et al.*, 2006). *In-vitro* transcription assay provided some evidences that the Gdown1 has a functional interaction with mediator complex for responding activator-dependent transcription. The structure of Gdown1 and its binding sites on the Pol II are unknown.

Comparison of transcription preinitiation complexes from three domains of life

When gene is expressed, the transcription machinery including GTFs and RNAP are recruited to promoter DNA to form preinitiation complex (PIC) (Hampsey, 1998; Hahn, 2004). The transcription-competent open complex is formed through the significant conformational changes of the PIC that include the unwinding of DNA around the

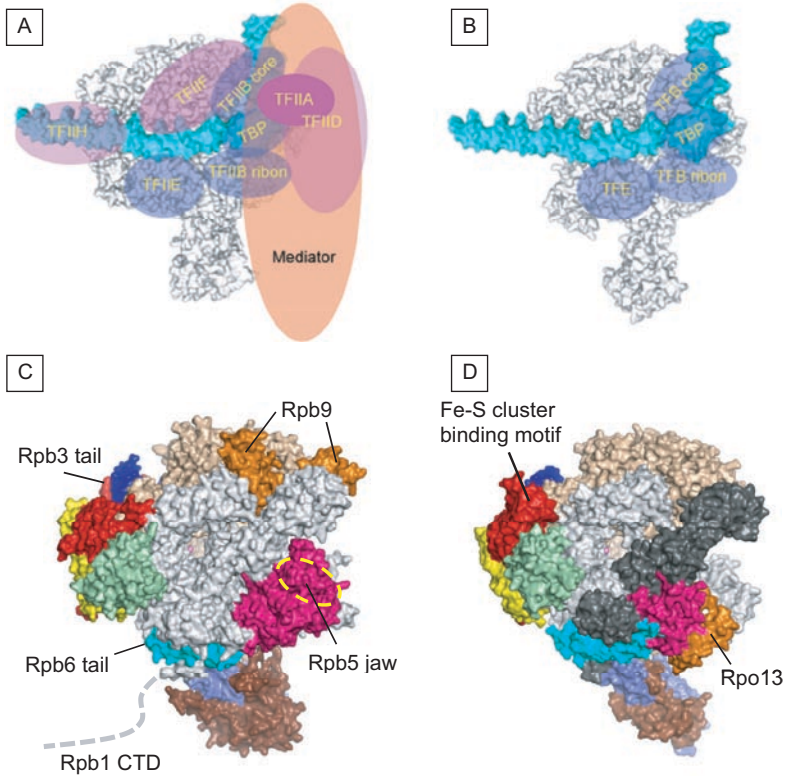

Figure 4. Comparison of the transcription machinery between eukaryote and archaea. (Top) Comparison of the (A) eukaryotic Pol II PIC and (B) the archaeal PIC. RNAP (gray) and promoter DNA (cyan) are represented as surface and general transcription factors are represented as ellipses. The common general transcription factors between eukaryote and archaea are colored transparent blue. (Bottom) Structural differences between (C) eukaryotic Pol II and (D) archaeal RNAP are highlighted and labeled. Subunit color-code is the same as in Figure 2.

transcription start site (TSS) and the positioning single-stranded template DNA near the RNAP active site. In bacteria, σ factor binds to core enzyme to form holoenzyme, which recognizes promoter DNA at around –35 and –10 from the TSS and makes a closed RNAP–promoter complex (Murakami and Darst, 2003). In the case of Pol II transcription, six GTFs, TFIIA, TFIIB, TFIID, TFIIE, TFIIF, and TFIIH, mediator complex and Pol II are recruited to the promoter DNA and form the PIC (Hampsey, 1998; Hahn, 2004) (Figure 4A). The archaeal PIC is similar to the eukaryotic Pol II transcription system and contains TBP, TFB, and TFE, which are orthologs of TBP, TFIIB, and TFIIEα, respectively, plus RNAP (Figure 4B) (Bell and Jackson, 1998). Only the structure of bacterial RNAP–promoter complex has been determined by X-ray crystallography (Murakami *et al.*, 2002a). However, the topological arrangements of archaeal transcription apparatus including GTFs (TBP and TFB) and RNAP on promoter DNA were well determined by high-resolution DNA photo-crosslink experiments (Renfrow *et al.*, 2004; Bartlett *et al.*, 2004), and the arrangements of these proteins on promoter DNA are in good agreements with their counterparts in the eukaryotic Pol II transcription system (Chen and Hahn, 2004; Kim *et al.*, 2000; Chen *et al.*, 2004; Forget *et al.*, 2004). Although archaeal PIC is much simpler than eukaryotic one, all the components of archaeal basal transcription machinery are highly related to the eukaryotic counterparts suggesting that the archaeal and eukaryotic transcription machines have come from the same origin. The tightly conserved, but much simpler and robust archaeal transcription system has been used for understanding the basic mechanism of Pol II transcription. For example, Cramer and Thomm groups have used archaeal *in-vitro* transcription system to characterize the structure and function relationship of yeast TFIIB in transcription initiation (Kostrewa *et al.*, 2009). Based on the newly determined Pol II–TFIIB complex X-ray crystal structure, they expected the DNA opening activity of the B-linker domain of TFIIB and have shown the same region of archaeal TFB has this DNA opening activity.

Compared with the Pol II transcription system, another advantage of the archaeal transcription system is that the active archaeal RNAP can be conveniently reconstituted from its individual subunits *in vitro* (Werner and Weinzierl, 2002; Naji *et al.*, 2007). With the reconstituted archaeal RNAP, it has been shown *in vitro* that the interaction between the B-linker domain of TFB and the coiled-coil region of clamp domain of RNAP is required for the DNA

opening activity of TFB, which was a nice complement to the observation from the X-ray crystal structure of Pol II–TFIIB complex (Kostrewa *et al.*, 2009).

Given that RNAP is the center of transcription machinery, structural comparisons between archaeal and eukaryotic RNAPs would give the insight for understanding the transcription machinery of two systems. The structural differences between archaeal and eukaryotic RNAPs can be regarded as simple additions of polypeptides to the archaeal RNAP rather than changes to the core structure (Hirata *et al.*, 2008b) (Figure 4C and 4D). Differences between archaeal and eukaryotic RNAPs found by X-ray crystal structures are Rpb9 and the N-terminal domain of Rpb5. In the X-ray crystal structure of *S. cerevisiae* Pol II, several regions are disordered that include the C-terminal hepta-peptide repeats of Rpb1 (Pol II CTD), Rpb3 C-terminal tail, the Rpb6 N-terminal tail, and the Rpb12 N-terminal region. Interestingly, these flexible regions do not exist in the archaeal RNAP amino acid sequence. Therefore, it is tempting to speculate that these flexible regions in Pol II have been gained during the evolution for interacting with the eukaryotic-specific transcription factors (Figure 4A). Further comparisons between archaeal and eukaryotic transcription apparatus could give the clues of how archaea and eukaryotes have developed their unique transcription system from their common ancestry during the evolution.

Transcription regulation in archaea

One striking feature of the archaea is that they possess a high degree of similarity to eukaryotes in the proteins utilized for gene expression. This includes proteins involved in coordinating DNA replication, translation, and transcription, and it is reflective of the common ancestry shared by the two groups (Woese *et al.*, 1990). The focus of second part of review will be to summarize the key proteins and the steps involved in establishing archaeal transcription initiation and its regulation.

General transcription factors in the archaeal PIC

Despite their capacity for *de novo* transcription without a primer, multisubunit cellular RNAPs are incapable of sequence-specific promoter recognition. These RNAPs require the aid of additional GTFs to properly position them at near the TSS before engaging in transcription. The first of the two required archaeal GTFs is an ortholog to the eukaryotic TBP, which is a highly conserved protein found in all known species of archaea and eukaryote, but not in bacteria. TBP orthologs are essential for the transcription of virtually all genes in these two domains (Thomm, 2007; Pugh, 2000). TBP binds promoter DNA along the minor groove of an AT-rich TATA-box sequence located ~25 bp upstream of the TSS and causes a sharp DNA bend (Kosa *et al.*, 1997; Kim and Burley, 1994; Nikolov *et al.*, 1996). The distortion of DNA backbone is thought to assist in the recruitment of other GTFs to the promoter. Archaeal TBPs and the C-terminal domain of the eukaryotic TBPs (TBP_c) are ~180 amino acids in length, and consist of two domains which are direct imperfect repeats. Archaeal TBPs and the eukaryotic TBP_c have ~30–40% amino acid sequence identity (Soppa, 1999) and have close structural similarity (Kosa *et al.*, 1997; Bell *et al.*, 1999b; Nikolov *et al.*, 1996; Kim and Burley, 1994). Although their structures are very similar, eukaryotic TBPs generally have an overall basic charge whereas archaeal TBPs tend to be acidic (Thomsen *et al.*, 2001). The eukaryotic TBPs have additional N-terminal domain that is absent in the archaeal molecule (Soppa, 1999; Thomsen *et al.*, 2001).

The second required archaeal GTF is TFB which is orthologous to eukaryotic TFIIB (Figure 5). TFB is highly conserved among different species (50–60% amino acid similarity), but sequence similarity between archaeal TFBs and eukaryal TFIIBs is only ~20–30% (Soppa, 1999). TFB interacts with a DNA sequences known as the B-factor recognition element (BRE), which is a purine rich (within nontemplate strand) sequence found immediately upstream of the TATA-box. The interaction between TFB and BRE appears to be necessary to determine direction of transcription (Bell *et al.*, 1999b; Littlefield *et al.*, 1999). The C-terminal region of TFB has two domains, each consisting of five or six α-helices (Figure 5). The X-ray crystal structure of DNA–TBP–TFB complex showed that both domains are involved in interaction with DNA and TBP (Korkhin *et al.*, 2001). The last two α-helices in the second repeat form a helix-turn-helix motif, which recognizes the DNA sequence of BRE (Littlefield *et al.*, 1999). The structure of archaeal TFB N-terminal domain including Zn^{2+} binding site has been determined by NMR (Zhu *et al.*, 1996). Later, X-ray structure of the yeast Pol II–TBP–TFIIB complex provided a nearly entire structure of TFIIB and positions of each TFIIB domain on the Pol II for understanding their functions (Kostrewa *et al.*, 2009). The N-terminal domain of TFB contains several motifs that play important roles in RNAP binding (B-ribbon and B-linker), DNA unwinding (B-linker), and TSS selection (B-reader). Previous *in vitro* transcription analyses showed that once 9–12 bases of the nascent transcripts are formed, TFB is released from the transcription initiation complex and recycled for the next round of transcription (Xie and Reeve, 2004; Spitalny and Thomm, 2003). TFB release from RNAP is also necessary to open the channel for nascent RNA exit from RNAP (Kostrewa *et al.*, 2009).

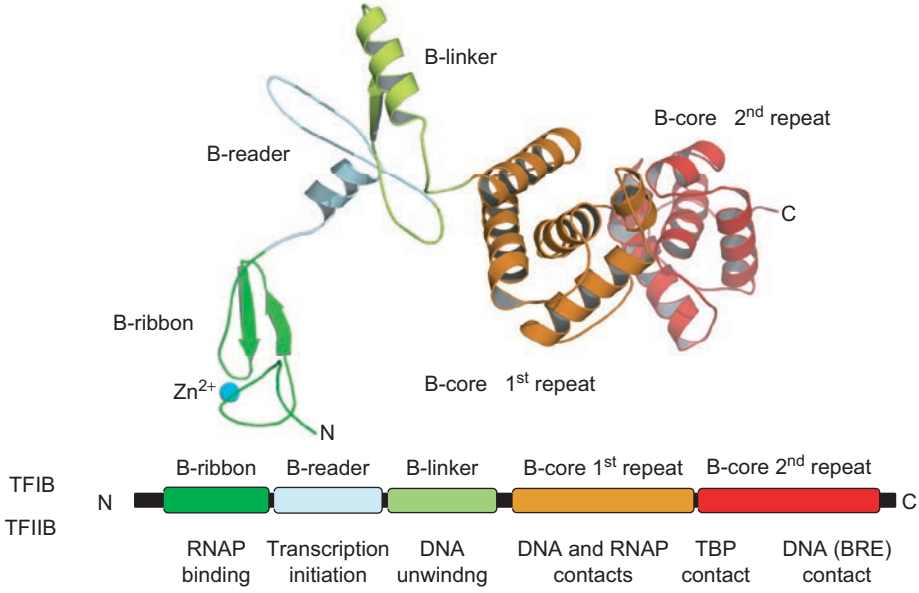

Figure 5. Schematic of TFB/TFIIB sequence architecture. The black bar represents the primary sequence. The conserved regions are labeled and their functions are indicated. The structure of TFIIB is also shown. Each conserved region has the same color as in the primary sequence.

In addition to having TBP and TFIIB orthologs, archaea possesses less well-characterized orthologs to the N-terminal domain of the α-subunit of eukaryotic TFIIE designated TFE (Figure 6) (Bell *et al.*, 2001; Thomm, 2007). Transcription from strong promoters is apparently unaffected by TFE (Bell *et al.*, 2001). It has been suggested that TFE is involved in stabilizing the RNAP open complex formation by enhancing DNA melting and DNA loading, and this activity is dependent upon the E subunit of RNAP (Naji *et al.*, 2007; Thomm *et al.*, 2009). Accordingly, TFE is copurified with RNAP but the mutant RNAP preparations purified from *ΔrpoF* cells lacked subunit F and also subunit E and TFE (Hirata *et al.*, 2008a). It has also been demonstrated that unlike TFIIE, TFE remains associated with RNAP during elongation (Grunberg *et al.*, 2007).

The TFS is the ortholog of C-terminal domain of the eukaryotic transcription elongation factor TFIIS, which plays a role in transcription proofreading by RNA hydrolysis (Langer and Zillig, 1993; Hausner *et al.*, 2000; Thomm, 2007). There have been no identified archaeal orthologs to eukaryotic TFIIA, TFIIF, or TFIIH nor have there been any identified archaeal homologs to any of the myriad of eukaryotic TBP-associated factors (TAF) which are found in TFIID (Thomm, 2007). Although no archaeal homologs to any eukaryotic TAFs have been identified, one protein, designated TBP-interacting protein 26 (TIP26) has been identified in *T. kodakarensis* KOD1 which can bind TBP and inhibit it from binding DNA (Matsuda *et al.*, 2001; Yamamoto *et al.*, 2006; Matsuda *et al.*, 1999). Although TIP26 homologs are not widespread in the archaea, its finding gives rise to the possibility that other proteins with analogous functions might exist in other archaeal species.

Transcription regulation in archaea by protein regulators and general transcription factors

In order to effectively survive in a competitive environment, organisms need to balance production of gene required for cell activities, while avoiding extraneous production of unnecessary proteins. The regulation of gene expression could potentially take place at any stage from transcription initiation to protein degradation. Regulation of transcription initiation is one of the important steps governing gene expression. Therefore, organisms have developed a variety of methods to achieve effective outcomes. Some archaeal transcription regulation mechanisms have been shown to be more closely related to bacterial than eukaryotic systems (Bell, 2005; Geiduschek and Ouhammouch, 2005), whereas others are more closely related to those in eukaryotes. Three examples of currently proposed models of transcription regulation in archaea are described below.

The first example is the utilization of bacterial-type transcription activators and repressors for transcription regulation. Although archaeal transcription machinery consists of a eukaryotic-type RNAP and GTFs, most archaeal genomes encode homologues to bacterial-type transcription regulators (Bell, 2005; Bell and Jackson, 2001; Geiduschek and Ouhammouch, 2005; Vierke *et al.*, 2003). The mechanism of repressing transcription in these cases entails binding to DNA in the promoter by a transcription

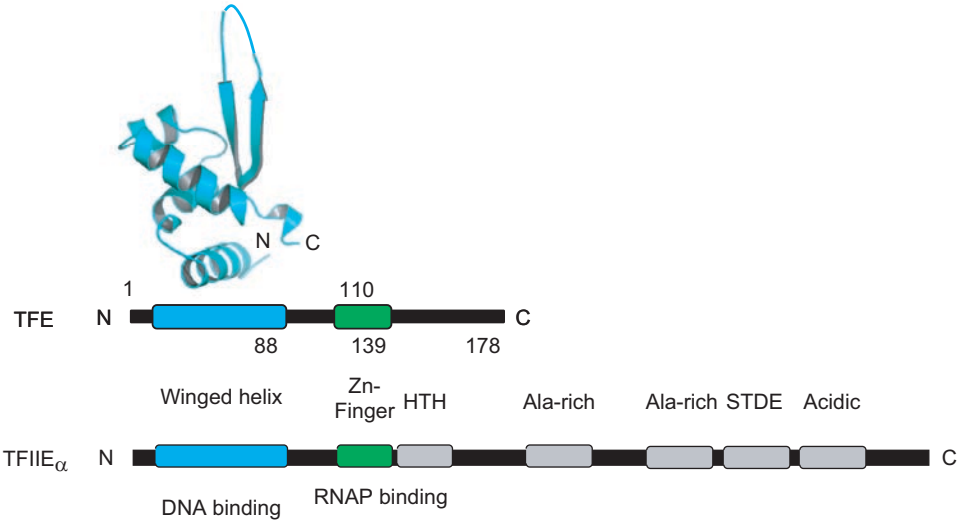

Figure 6. Schematic of TFE/TFIIEα sequence architecture. The black bar represents the primary sequence. The conserved regions are labeled and their functions are indicated. The X-ray crystal structure of the N-terminal winged helix domain of TFE (*Sulfolobus solfataricus* TFE, residues 1–88, PDB: 1Q1H) is also shown.

factor leading to the blocking of TFB and/or TBP bindings to DNA (e.g., *S. solfataricus* Lrs-14 (Fiorentino *et al.*, 2003)) or inhibiting RNAP recruitment to the promoter (e.g. *P. furiosus* LrpA (Brinkman *et al.*, 2002; Dahlke and Thomm, 2002), *Archaeoglobus fulgidus* MDR1 (Bell *et al.*, 1999a)). Some bacterial Lrp (Leucine-responsive regulator protein)-like factors in archaea are proposed to function as transcription activators, which enhance the formation of PIC (Ouhammouch, 2004). Examples of currently identified activators are *S. solfataricus* LysM (Brinkman *et al.*, 2002) and *Methanocaldococcus janaschii* Ptr2 (Brinkman *et al.*, 2003; Ouhammouch *et al.*, 2003).

The second example involves DNA packaging proteins, which is analogous to transcription regulation in eukaryotes; the nucleosome structure and histone modification play key roles in gene regulation. Eukaryotic histones possess positively charged N-terminal tails, and specific residues within the tails are covalently modified by acetylation, phospholylation, metylation, and ubiquitination (reviewed in Wu and Grunstein, 2000). Acetylation of histone tails, for example, allows the access of transcription factors to DNA to promote transcription initiation. Archaeal species from *Euryarchaeota* kingdom possess one to six different histones forming both heterodimers and homodimers. For example, two histone proteins in *Methanothermus fervidus*, HMfA and HMfB, exist as either a heterodimer (HMfA+HMfB) or a homodimer ($(HMfA)_2$ or $(HMfB)_2$) *in vivo* (Pereira and Reeve, 1998). Interestingly, alternative dimer forms have different affinities to DNA sequences. Therefore, combinations of histone proteins could potentially control accessibility of transcription factors to DNA promoters in the archaea (reviewed in Pereira and Reeve, 1998; Reeve, 2003). A recent *in vitro* transcription experiments revealed that the *M. jannaschii* histones inhibit transcription from promoter DNA and the *M. jannaschii* transcription activator Ptr2 competes for DNA binding with histones in effect counteracting the repressive effect of histones (Wilkinson *et al.*, 2010). Identification of archaeal histones raises the possibility of nucleosome structure presence in *Euryarchaeota* and a potential transcription regulation mechanism via its modification (Reeve, 2003). However, the archaeal histones lack the N- and C-terminal tails, which are the targets for modifications in eukaryotes. In addition, homologues to histone modification factors found in eukaryotes have not been identified in any archaeal genomes (Reeve, 2003).

Acetylation lowers binding affinity (Alba) is another double-stranded DNA binding protein identified in archaea, which can be reversibly acetylated and deacetylated by modification enzymes (Bell *et al.*, 2002). It has been proposed that the acetylation and deacetylation of Alba could influence transcription activity due to the lower DNA binding affinity exhibited by acetylated Alba compared with deacetylated form (Bell *et al.*, 2002). Also, repressed transcription activity *in vitro* on DNA templates in the presence of acetylated Alba had been reported (Bell *et al.*, 2002).

The third proposed model involves the multiplicity of GTFs that include TBP and TFB. Most higher eukaryotes are known to carry multiple orthologs of both TBP and TFIIB with divergent functional roles from these GTFs. These include at least four TBP orthologs known as TBP-related factors with a variety of different functional roles mostly during embryonic development (Crowley *et al.*, 1993; Dantonel *et al.*, 1999; Persengiev *et al.*, 2003), and two different TFIIB orthologs known as TFIIB-related

factors in the Pol III transcription system (Ferrari *et al.*, 2004; Paule and White, 2000; Saxena *et al.*, 2005). The presence of multiple TBP and TFIIB homologs is not unique to eukaryotes. Not long after the discovery of archaeal TBP and TFB, it was discovered that some archaeal species possessed more than one gene for TBP and/or TFB consistent with the possibility that archaeal species might utilize these alternative GTF orthologs to differentially regulate gene expression. The sequencing of an extrachromosomal megaplasmid in the extreme halophile, *Halobacterium* sp. NRC-1 revealed four open reading frames encoding for putative TBPs located on this megaplasmid (Ng *et al.*, 1998). Later, the complete genome sequencing of this organism revealed two additional TBP orthologs and seven different TFB orthologs (Ng *et al.*, 2000; Baliga *et al.*, 2000). Subsequent genome sequencing projects revealed multiple TFB orthologs (Kawarabayasi *et al.*, 1999; Thompson *et al.*, 1999; Lecompte *et al.*; 2001; Fukui *et al.*; 2005; She *et al.*, 2001; Kawarabayasi *et al.*, 1998; Facciotti *et al.*, 2007; Coker and DasSarma, 2007) or multiple TBP orthologs (Deppenmeier *et al.*, 2002; Galagan *et al.*, 2002; Maeder *et al.*, 2006) in several other archaeal species. It is now apparent that the presence of multiple archaeal GTF orthologs is not rare. The J. Craig Venter Institute currently lists the completed genomes for 42 archaeal species on its database, and 29 of these species have at least two identified TBPs or TFBs (Table 2). Among those species with multiple GTF orthologs, there is a tendency towards multiple TFBs in the halophiles, hyperthermophiles, and thermoacidophiles, whereas there is a tendency towards multiple TBPs in the methanogens. Whether or not this trend has any functional significance is unclear. However, in light of the important roles for TBP and TFB in DNA binding and DNA opening, respectively, this trend could be an indication that for the halophiles and hyperthermophiles the DNA opening step is the major target for gene regulation whereas for the methanogens, the DNA binding step is the primary target for regulation.

The majority of the experimental investigations into the function of multiple archaeal GTFs have been carried out in *Halobacterium* sp. NRC-1. This has been due to the large number of unique GTF homologs (six TBPs and seven TFBs which is the most for any archaeal species) (Table 2) (Ng *et al.*, 2000; Baliga *et al.*, 2000), and because of the availability of facile genetic tools. These investigations have largely concluded that the individual GTF isomers in *Halobacterium* sp. NRC-1 function to differentially regulate gene expression. Genetic analyses have established that at least 10 of its 13 GTF orthologs (four TBPs and six TFBs) can be individually deleted under standard laboratory growth conditions without hindering cell viability (Coker and DasSarma, 2007; Facciotti *et al.*, 2007) although no investigations into the consequences of deleting more than one GTF have been reported. One of the two TBPs that could not be deleted was *tbpE*, which is located on the main chromosome of *Halobacterium* sp. NRC-1. Proteomics analysis and quantitative reverse transcription-PCR established that the *tbpE* is more robustly expressed than the other *TBP* genes, and it was suggested that TBPe is the primary TBP utilized during growth. It was proposed that the additional TBPs provide added fitness based on defective growth of strains lacking the additional *tbp* genes (Goo *et al.*, 2003; Teufel *et al.*, 2008). It is unclear if one TFB ortholog functions as the primary TFB during growth, although *tfbB* is the only TFB that could not be knocked out (Coker and DasSarma, 2007; Facciotti *et al.*, 2007). In a global analysis, a preliminary TFB regulatory network was deduced based on integrated data from ChIP–chip analysis, transcriptomic data and *in vivo* protein–protein interactions between the various TBP and TFB orthologs (Facciotti *et al.*, 2007). In a separate report, one TBP, *tbpD* and one TFB, *tfbA* were shown to coordinately regulate roughly 10% of the *Halobacterim* sp. NRC-1 genome including several heat-shock response genes suggesting that preferential pairings of TBP–TFB may be exploited to direct gene expression. Mutant

Table 2 .Archaeal species with multiple annotated TBP or TFB genes.

Organism	Putative TBPs	Putative TFBs
Aeropyrum pernix K1	1	2
Candidatus methanoregula boonei 6A8	3	1
Haloarcula marismortui ATCC 43049	1	9
Halobacterium sp. NRC-1	6	7
Haloquadratum walsbyi DSM 16790	2	9
Hyperthermus butylicus DSM 5456	1	2
Metallosphaera sedula DSM 5348	1	2
Methanococcus maripaludis C5	2	1
Methanocorpusculum labreanum Z	2	1
Methanoculleus marisnigri JR1	2	2
Methanosarcina acetivorans C2A	3	1
Methanosarcina barkeri fusaro	2	1
Methanosarcina mazei Goe1	3	1
Methanospirillum hungatei JF-1	2	1
Natronomonas pharaonis sp	1	9
Picrophilus torridus DSM 9790	1	2
Pyrobaculum aerophilum IM2	1	3
Pyrobaculum arsenaticum DSM 13514	1	2
Pyrobaculum islandicum DSM 4184	1	2
Pyrococcus furiosus DSM 3638	1	2
Pyrococcus horikoshii shinkaj OT3	1	2
Sulfolobus acidocaldarius DSM 639	1	2
Sulfolobus solfataricus P2	1	2
Sulfolobus tokodaii strain 7	1	2
Thermococcus kodakarensis KOD1	1	2
Thermofilum pendens Hrk 5	1	2
Thermoplasma acidophilum DSM 1728	1	3
Thermoplasma volcanium GSS1	1	3

strains with either *tbpD* or *tfbA* deleted elicited defective growth in response to heat shock (Kaur *et al.*, 2006). Another TFB ortholog *tfbB* was shown to preferentially bind the heat-shock inducible promoter P_{hsp5} *in vitro* during incubation at 50 °C, but not during incubation at 37 °C. Furthermore, *tfbG* binding at P_{hsp5} was not detected at either temperature consistent with a specific role for *tfbB* in the regulation of *hsp5* during *Halobacterium* sp. NRC-1 heat-shock response (Lu *et al.*, 2008). Consistent with these observations for *Halobacterium salinarum,* TFB orthologs from *Haloferax volcanii, P. furiosus*, and *S. solfataricus* have been implicated in the heat-shock response or UV irradiation response (Thompson *et al.*, 1999; Shockley *et al.*, 2003; Paytubi and White, 2009). This suggests that the use of alternative TFB orthologs to direct the regulation of stress response genes may be a common feature in archaeal species.

A functional assessment of the two TFB orthologs in the hyperthermophile, *T. kodakarensis* suggested the possibility of functional redundancy for multiple TFBs (Santangelo *et al.*, 2007). It was determined that either of the two TFB orthologs can be individually deleted without hindering growth under optimal growth conditions. Furthermore, both TFB orthologs could support *in vitro* transcription equally well from several different promoters showing no difference in TSS selection (Hirata *et al.*, 2008a). In contrast, two TFB orthologs in *P. furiosus* (TFB1 and TFB2) showed substantial difference in their transcription activities *in vitro*. The TFB2, which lacks B-finger motif, is less active compared with TFB1 for all tested promoters (Micorescu *et al.*, 2008).

To date, there has been only one reported experimental investigation into the functional roles of multiple TBP orthologs in organisms with only a single TFB which was carried out in the mesophilic methanogen, *Methanosarcina acetivorans* (Reichlen *et al.*, 2010). As is the case for *Halobacterium* sp. NRC-1 *tbpE*, one of the *M. acetivorans* TBP orthologs, TBP1 appears to function in a dominant role to the two other TBPs, TBP2, and TBP3. However, these alternative TBPs did appear to be important for optimal growth when cells were cultured under nutrient limiting conditions or when forced to adjust their metabolic pathway from an energetically rich to an energetically poor substrate. These results suggested a possible role in for the alternative TBP orthologs transcriptional regulation specific to these conditions. However, specific targets for the alternative TBPs were not identified (Reichlen *et al.*, 2010).

Acknowledgements

We thank Rieko Yajima for critical reading of the manuscript. This work was supported by NIH grant GM087350.

Declaration of interest

The authors report no conflicts of interest.

References

Armache KJ, 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.

Baliga NS, Goo YA, Ng WV, Hood L, Daniels CJ and Dassarma S. 2000. Is gene expression in Halobacterium NRC-1 regulated by multiple TBP and TFB transcription factors? Mol Microbiol 36:1184–1185.

Bartlett MS, Thomm M and Geiduschek EP. 2004. Topography of the euryarchaeal transcription initiation complex. J Biol Chem 279:5894–5903.

Bell SD. 2005. Archaeal transcriptional regulation–variation on a bacterial theme? Trends Microbiol 13:262–265.

Bell SD, Botting CH, Wardleworth BN, Jackson SP and White MF. 2002. The interaction of Alba, a conserved archaeal chromatin protein, with Sir2 and its regulation by acetylation. Science 296:148–151.

Bell SD, Brinkman AB, van der Oost J and Jackson SP. 2001. The archaeal TFIIEalpha homologue facilitates transcription initiation by enhancing TATA-box recognition. EMBO Rep 2:133–138.

Bell SD, Cairns SS, Robson RL and Jackson SP. 1999a. Transcriptional regulation of an archaeal operon in vivo and in vitro. Mol Cell 4:971–982.

Bell SD and Jackson SP 1998. Transcription and translation in Archaea: a mosaic of eukaryal and bacterial features. Trends Microbiol 6:222–228.

Bell SD and Jackson SP 2001. Mechanism and regulation of transcription in archaea. Curr Opin Microbiol 4:208–213.

Bell SD, Kosa PL, Sigler PB and Jackson SP. 1999b. Orientation of the transcription preinitiation complex in archaea. Proc Natl Acad Sci USA 96:13662–13667.

Brinkman AB, Bell SD, Lebbink RJ, de Vos WM and Van der Oost J. 2002. The Sulfolobus solfataricus Lrp-like protein LysM regulates lysine biosynthesis in response to lysine availability. J Biol Chem 277:29537–29549.

Brinkman AB, Ettema TJ, de Vos WM and Van der Oost J. 2003. The Lrp family of transcriptional regulators. Mol Microbiol 48:287–294.

Chen BS, Mandal SS and Hampsey M. 2004. High-resolution protein-DNA contacts for the yeast RNA polymerase II general transcription machinery. Biochemistry 43 12741–12749.

Chen HT, Hahn S. 2004. Mapping the location of TFIIB within the RNA polymerase II transcription preinitiation complex: a model for the structure of the PIC. Cell 119:169–180.

Coker JA and DasSarma S. 2007. Genetic and transcriptomic analysis of transcription factor genes in the model halophilic Archaeon: coordinate action of TbpD and TfbA. BMC Genet 8:61.

Cramer P, Bushnell DA and Kornberg RD. 2001. Structural basis of transcription: RNA polymerase II at 2.8 angstrom resolution. Science 292:1863–1876.

Crowley TE, Hoey T, Liu JK, Jan YN, Jan LY and Tjian R. 1993. A new factor related to TATA-binding protein has highly restricted expression patterns in Drosophila. Nature 361:557–561.

Dahlke I and Thomm M. 2002. A *Pyrococcus* homolog of the leucine-responsive regulatory protein, LrpA, inhibits transcription by abrogating RNA polymerase recruitment. Nucleic Acids Res 30:701–710.

Dantonel JC, Wurtz JM, Poch O, Moras D and Tora L. 1999. The TBP-like factor: an alternative transcription factor in metazoa? Trends Biochem Sci 24:335–339.

Darst SA. 2001. Bacterial RNA polymerase. Curr Opin Struct Biol 11:155–162.

Deppenmeier U, Johann A, Hartsch T, Merkl R, Schmitz RA, Martinez-Arias R, Henne A, Wiezer A, Baumer S, Jacobi C, Bruggemann H, Lienard T, Christmann A, Bomeke M, Steckel S, Bhattacharyya A, Lykidis A, Overbeek R, Klenk HP, Gunsalus RP,

Fritz HJ and Gottschalk G. 2002. The genome of Methanosarcina mazei: evidence for lateral gene transfer between bacteria and archaea. J Mol Microbiol Biotechnol 4:453–461.

Facciotti MT, Reiss DJ, Pan M, Kaur A, Vuthoori M, Bonneau R, Shannon P, Srivastava A, Donohoe SM, Hood LE and Baliga NS. 2007. General transcription factor specified global gene regulation in archaea. Proc Natl Acad Sci USA 104:4630–4635.

Ferrari R, Rivetti C, Acker J and Dieci G. 2004. Distinct roles of transcription factors TFIIIB and TFIIIC in RNA polymerase III transcription reinitiation. Proc Natl Acad Sci USA 101:13442–13447.

Fiorentino G, Cannio R, Rossi M and Bartolucci S. 2003. Transcriptional regulation of the gene encoding an alcohol dehydrogenase in the archaeon *Sulfolobus solfataricus* involves multiple factors and control elements. J Bacteriol 185:3926–3934.

Forget D, Langelier MF, Therien C, Trinh V and Coulombe B. 2004. Photo-cross-linking of a purified preinitiation complex reveals central roles for the RNA polymerase II mobile clamp and TFIIE in initiation mechanisms. Mol Cell Biol 24:1122–1131.

Fukui T, Atomi H, Kanai T, Matsumi R, Fujiwara S and Imanaka T. 2005. Complete genome sequence of the hyperthermophilic archaeon *Thermococcus kodakaraensis* KOD1 and comparison with *Pyrococcus* genomes. Genome Res 15:352–363.

Galagan JE, Nusbaum C, Roy A, Endrizzi MG, Macdonald P, Fitzhugh W, Calvo S, Engels R, Smirnov S, Atnoor D, Brown A, Allen N, Naylor J, Stange-Thomann N, Dearellano K, Johnson R, Linton L, Mcewan P, Mckernan K, Talamas J, Tirrell A, Ye W, Zimmer A, Barber RD, Cann I, Graham DE, Grahame DA, Guss AM, Hedderich R, Ingram-Smith C, Kuettner HC, Krzycki JA, Leigh JA, Li W, Liu J, Mukhopadhyay B, Reeve JN, Smith K, Springer TA, Umayam LA, White O, White RH, Conway De Macario E, Ferry JG, Jarrell KF, Jing H, Macario AJ, Paulsen I, Pritchett M, Sowers KR, Swanson RV, Zinder SH, Lander E, Metcalf WW and Birren B. 2002. The genome of *M. acetivorans* reveals extensive metabolic and physiological diversity. Genome Res 12:532–542.

Geiduschek EP and Ouhammouch M. 2005. Archaeal transcription and its regulators. Mol Microbiol 56:1397–1407.

Gnatt AL, Cramer P, Fu J, Bushnell DA and Kornberg RD. 2001. Structural basis of transcription: an RNA polymerase II elongation complex at 3.3 A resolution. Science 292:1876–1882.

Goede B, Naji S, Von Kampen O, Ilg K and Thomm M. 2006. Protein-protein interactions in the archaeal transcriptional machinery: binding studies of isolated RNA polymerase subunits and transcription factors. J Biol Chem 281:30581–30592.

Goo YA, Yi EC, Baliga NS, Tao WA, Pan M, Aebersold R, Goodlett DR, Hood L and Ng WV. 2003. Proteomic analysis of an extreme halophilic archaeon, Halobacterium sp. NRC-1. Mol Cell Proteomics 2:506–524.

Grunberg S, Bartlett MS, Naji S and Thomm M. 2007. Transcription factor E is a part of transcription elongation complexes. J Biol Chem 282:35482–35490.

Hahn S. 2004. Structure and mechanism of the RNA polymerase II transcription machinery. Nat Struct Mol Biol 11:394–403.

Hampsey M. 1998. Molecular genetics of the RNA polymerase II general transcriptional machinery. Microbiol Mol Biol Rev 62:465–503.

Harrison DA, Mortin MA and Corces VG. 1992. The RNA polymerase II 15-kilodalton subunit is essential for viability in *Drosophila melanogaster*. Mol Cell Biol 12:928–935.

Hausner W, Lange U and Musfeldt M. 2000. Transcription factor S, a cleavage induction factor of the archaeal RNA polymerase. J Biol Chem 275:12393–12399.

Hirata A, Kanai T, Santangelo TJ, Tajiri M, Manabe K, Reeve JN, Imanaka T and Murakami KS. 2008a. Archaeal RNA polymerase subunits E and F are not required for transcription in vitro, but a *Thermococcus kodakarensis* mutant lacking subunit F is temperature-sensitive. Mol Microbiol 70:623–633.

Hirata A, Klein BJ and Murakami KS. 2008b. The X-ray crystal structure of RNA polymerase from Archaea. Nature 452:248.

Hirata A and Murakami KS. 2009. Archaeal RNA polymerase. Curr Opin Struct Biol 19:724–731.

Hu X, Malik S, Negroiu CC, Hubbard K, Velalar CN, Hampton B, Grosu D, Catalano J, Roeder RG and Gnatt A. 2006. A Mediator-responsive form of metazoan RNA polymerase II. Proc Natl Acad Sci USA 103:9506–9511.

Huet J, Schnabel R, Sentenac A and Zillig W. 1983. Archaebacteria and eukaryotes possess DNA-dependent RNA polymerases of a common type. Embo J 2:1291–1294.

Jasiak AJ, Armache KJ, Martens B, Jansen RP 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.

Kassavetis GA, Letts GA and Geiduschek EP. 2001. The RNA polymerase III transcription initiation factor TFIIIB participates in two steps of promoter opening. Embo J 20:2823–2834.

Kaur A, Pan M, Meislin M, Facciotti MT, El-Geweley R and Baliga NS. 2006. A systems view of haloarchaeal strategies to withstand stress from transition metals. Genome Res.

Kawarabayasi Y, Hino Y, Horikawa H, Yamazaki S, Haikawa Y, Jin-no K, Takahashi M, Sekine M, Baba S, Ankai A, Kosugi H, Hosoyama A, Fukui S, Nagai Y, Nishijima K, Nakazawa H, Takamiya M, Masuda S, Funahashi T, Tanaka T, Kudoh Y, Yamazaki J, Kushida N, Oguchi A, Kikuchi H et al. 1999. Complete genome sequence of an aerobic hyper-thermophilic crenarchaeon, Aeropyrum pernix K1. DNA Res 6:83–101, 145–152.

Kawarabayasi Y, Sawada M, Horikawa H, Haikawa Y, Hino Y, Yamamoto S, Sekine M, Baba S, Kosugi H, Hosoyama A, Nagai Y, Sakai M, Ogura K, Otsuka R, Nakazawa H, Takamiya M, Ohfuku Y, Funahashi T, Tanaka T, Kudoh Y, Yamazaki J, Kushida N, Oguchi A, Aoki K and Kikuchi H. 1998. Complete sequence and gene organization of the genome of a hyper-thermophilic archaebacterium, *Pyrococcus horikoshii* OT3 (supplement). DNA Res 5:147–155.

Kim JL and Burley SK. 1994. 1.9 A resolution refined structure of TBP recognizing the minor groove of TATAAAAG. Nat Struct Biol 1:638–653.

Kim TK, Ebright RH and Reinberg D. 2000. Mechanism of ATP-dependent promoter melting by transcription factor IIH. Science 28:1418–1422.

Korkhin Y, Littlefield O, Nelson PJ, Bell SD and Sigler PB. 2001. Preparation of components of archaeal transcription preinitiation complex. Methods Enzymol 334:227–239.

Korkhin Y, Unligil UM, Littlefield O, Nelson PJ, Stuart DI, Sigler PB, Bell SD and Abrescia NG. 2009. Evolution of complex RNA polymerases: the complete archaeal RNA polymerase structure. PLoS Biol 7, e102.

Kosa PF, Ghosh G, Dedecker BS and Sigler PB. 1997. The 2.1-A crystal structure of an archaeal preinitiation complex: TATA-box-binding protein/transcription factor (II)B core/TATA-box. Proc Natl Acad Sci USA 94:6042–6047.

Kostrewa D, Zeller ME, Armache KJ, Seizl M, Leike K, Thomm M and Cramer P. 2009. RNA polymerase II-TFIIB structure and mechanism of transcription initiation. Nature 462:323–330.

Kuhn CD, Geiger SR, 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.

Kusser AG, Bertero MG, Naji S, Becker T, Thomm M, Beckmann R and Cramer P. 2008. Structure of an archaeal RNA polymerase. J Mol Biol 376:303–307.

Langer D, Hain J, Thuriaux P and Zillig W. 1995. Transcription in archaea:similarity to that in eucarya. Proc Natl Acad Sci USA 92:5768–5772.

Langer D and Zillig W. 1993. Putative tfIIs gene of *Sulfolobus acidocaldarius* encoding an archaeal transcription elongation factor is situated directly downstream of the gene for a small subunit of DNA-dependent RNA polymerase. Nucleic Acids Res 21:2251.

Lecompte O, Ripp R, Puzos-Barbe V, Duprat S, Heilig R, Dietrich J, Thierry JC and Poch O. 2001. Genome evolution at the genus level: comparison of three complete genomes of hyperthermophilic archaea. Genome Res 1:981–993.

Li S and Smerdon MJ. 2002. Rpb4 and Rpb9 mediate subpathways of transcription-coupled DNA repair in *Saccharomyces cerevisiae*. Embo J 21:5921–5929.

Littlefield O, Korkhin Y and Sigler PB. 1999. The structural basis for the oriented assembly of a TBP/TFB/promoter complex. Proc Natl Acad Sci USA 96:13668–13673.

Liu X, Bushnell DA, Wang D, Calero G and Kornberg RD. 2010. Structure of an RNA polymerase II-TFIIB complex and the transcription initiation mechanism. Science 327:206–209.

Lopez de Saro FJ, Woody AY and Helmann JD. 1995. Structural analysis of the *Bacillus subtilis* delta factor: a protein polyanion which displaces RNA from RNA polymerase. J Mol Biol 252:189–202.

Lu Q, Han J, Zhou L, Coker JA, Dassarma P, Dassarma S and Xiang H. 2008. Dissection of the regulatory mechanism of a heat-shock responsive promoter in *Haloarchaea*: a new paradigm for general transcription factor directed archaeal gene regulation. Nucleic Acids Res.

Maeder DL, Anderson I, Brettin TS, Bruce DC, Gilna P, Han CS, Lapidus A, Metcalf WW, Saunders E, Tapia R and Sowers KR. 2006. The *Methanosarcina barkeri* genome: comparative analysis with *Methanosarcina acetivorans* and *Methanosarcina mazei* reveals extensive rearrangement within methanosarcinal genomes. J Bacteriol 188:7922–7931.

Matsuda T, Fujikawa M, Haruki M, Tang XF, Ezaki S, Imanaka T, Morikawa M and Kanaya S. 2001. Interaction of TIP26 from a hyperthermophilic archaeon with TFB/TBP/DNA ternary complex. Extremophiles 5:177–182.

Matsuda T, Morikawa M, Haruki M, Higashibata H, Imanaka T and Kanaya S. 1999. Isolation of TBP-interacting protein (TIP) from a hyperthermophilic archaeon that inhibits the binding of TBP to TATA-DNA. FEBS Lett 457:38–42.

Meka H, Werner F, Cordell SC, Onesti S and Brick P. 2005. Crystal structure and RNA binding of the Rpb4/Rpb7 subunits of human RNA polymerase II. Nucleic Acids Res 33:6435–6444.

Micorescu M, Grunberg S, Franke A, Cramer P, Thomm M and Bartlett M. 2008. Archaeal transcription:function of an alternative transcription factor B from Pyrococcus furiosus. J Bacteriol 190:157–167.

Motackova V, Sanderova H, Zidek L, Novacek J, Padrta P, Svenkova A, Korelusova J, Jonak J, Krasny L and Sklenar V. 2010. Solution structure of the N-terminal domain of *Bacillus subtilis* delta subunit of RNA polymerase and its classification based on structural homologs. Proteins 78:1807–1810.

Murakami KS and Darst SA. 2003. Bacterial RNA polymerases: the wholo story. Curr Opin Struct Biol, 13:31–39.

Murakami KS, Masuda S, Campbell EA, Muzzin O and Darst SA. 2002a. Structural basis of transcription initiation: an RNA polymerase holoenzyme-DNA complex. Science 296:1285–1290.

Murakami KS, Masuda S and Darst SA. 2002b. Structural basis of transcription initiation: RNA polymerase holoenzyme at 4 A resolution. Science 296:1280–124.

Naji S, Grunberg S and Thomm M. 2007. The RPB7 orthologue E' is required for transcriptional activity of a reconstituted archaeal core enzyme at low temperatures and stimulates open complex formation. J Biol Chem 282:11047–11057.

Nesser NK, Peterson DO and Hawley DK. 2006. RNA polymerase II subunit Rpb9 is important for transcriptional fidelity in vivo. Proc Natl Acad Sci USA 103:3268–3273.

Ng WV, Ciufo SA, Smith TM, Bumgarner RE, Baskin D, Faust J, Hall B, Loretz C, Seto J, Slagel J, Hood L and Dassarma S. 1998. Snapshot of a large dynamic replicon in a halophilic archaeon: megaplasmid or minichromosome? Genome Res 8:1131–1141.

Ng WV, Kennedy SP, Mahairas GG, Berquist B, Pan M, Shukla HD, Lasky SR, Baliga NS, Thorsson V, Sbrogna J, Swartzell S, Weir D, Hall J, Dahl TA, Welti R, Goo YA, Leithauser B, Keller K, Cruz R, Danson MJ, Hough DW, Maddocks DG, Jablonski PE, Krebs MP, Angevine CM, Dale H, Isenbarger TA, Peck RF, Pohlschroder M, Spudich JL, Jung KW, Alam M, Freitas T, Hou S, Daniels CJ, Dennis PP, Omer AD, Ebhardt H, Lowe TM, Liang P, Riley M, Hood L and Dassarma S. 2000. Genome sequence of Halobacterium species NRC-1. Proc Natl Acad Sci USA 97:12176–12181.

Nikolov DB, Chen H, Halay ED, Hoffman A, Roeder RG and Burley SK. 1996. Crystal structure of a human TATA box-binding protein/TATA element complex. Proc Natl Acad Sci USA 93:4862–4867.

Ouhammouch M. 2004. Transcriptional regulation in Archaea. Curr Opin Genet Dev 14:133–138.

Ouhammouch M, Dewhurst RE, Hausner W, Thomm M and Geiduschek EP. 2003. Activation of archaeal transcription by recruitment of the TATA-binding protein. Proc Natl Acad Sci USA 100:5097–5102.

Ouhammouch M, Werner F, Weinzierl RO and Geiduschek EP. 2004. A fully recombinant system for activator-dependent archaeal transcription. J Biol Chem 279:51719–51721.

Pace NR. 1997. A molecular view of microbial diversity and the biosphere. Science 276:734–740.

Paule MR and White RJ. 2000. Survey and summary: transcription by RNA polymerases I and III. Nucleic Acids Res 28:1283–1298.

Paytubi S and White MF. 2009. The crenarchaeal DNA damage-inducible transcription factor B paralogue TFB3 is a general activator of transcription. Mol Microbiol 72:1487–1499.

Pereira SL and Reeve JN. 1998. Histones and nucleosomes in Archaea and Eukarya: a comparative analysis. Extremophiles 2:141–148.

Persengiev SP, Zhu X, Dixit BL, Maston GA, Kittler EL and Green MR. 2003. TRF3, a TATA-box-binding protein-related factor, is vertebrate-specific and widely expressed. Proc Natl Acad Sci USA 100:14887–14891.

Peyroche G, Milkereit P, Bischler N, Tschochner H, Schultz P, Sentenac A, Carles C and Riva M. 2000. The recruitment of RNA polymerase I on rDNA is mediated by the interaction of the A43 subunit with Rrn3. Embo J 19:5473–5482.

Pugh BF. 2000. Control of gene expression through regulation of the TATA-binding protein. Gene 255:1–14.

Reeve JN. 2003. Archaeal chromatin and transcription. Mol Microbiol 48:587–598.

Reichlen MJ, Murakami KS and Ferry JG. 2010. Functional analysis of the three TATA binding protein homologs in *Methanosarcina acetivorans*. J Bacteriol 192:1511–1517.

Renfrow MB, Naryshkin N, Lewis LM, Chen HT, Ebright RH and Scott RA. 2004. Transcription factor B contacts promoter DNA near the transcription start site of the archaeal transcription initiation complex. J Biol Chem 279:2825–2831.

Rowlands T, Baumann P and Jackson SP. 1994. The TATA-binding protein: a general transcription factor in eukaryotes and archaebacteria. Science 264:1326–1329.

Santangelo TJ, Cubonova L, James CL and Reeve JN. 2007. TFB1 or TFB2 Is sufficient for *Thermococcus kodakaraensis* viability and for basal transcription in vitro. J Mol Biol 367 344–357.

Saxena A, Ma B, Schramm L and Hernandez N. 2005. Structure-function analysis of the human TFIIB-related factor II protein reveals an essential role for the C-terminal domain in RNA polymerase III transcription. Mol Cell Biol 25:9406–9418.

She Q, Singh RK, Confalonieri F, Zivanovic Y, Allard G, Awayez MJ, Chan-Weiher CC, Clausen IG, Curtis BA, de Moors A, Erauso G, Fletcher C, Gordon PM, Heikamp-De Jong I, Jeffries AC, Kozera CJ, Medina N, Peng X, Thi-Ngoc HP, Redder P, Schenk ME, Theriault C, Tolstrup N, Charlebois RL, Doolittle WF, Duguet M, Gaasterland T, Garrett RA, Ragan MA, Sensen CW and van der Oost J. 2001. The complete genome of the crenarchaeon *Sulfolobus solfataricus* P2. Proc Natl Acad Sci USA 98:7835–7840.

Shockley KR, Ward DE, Chhabra SR, Conners SB, Montero CI and Kelly RM. 2003. Heat shock response by the hyperthermophilic archaeon *Pyrococcus furiosus*. Appl Environ Microbiol 69:2365–2371.

Soppa J. 1999. Transcription initiation in Archaea: facts, factors and future aspects. Mol Microbiol 31:1295–1305.

Spitalny P and Thomm M. 2003. Analysis of the open region and of DNA-protein contacts of archaeal RNA polymerase transcription complexes during transition from initiation to elongation. J Biol Chem 278:30497–30505.

Teufel K, Bleiholder A, Griesbach T and Pfeifer F. 2008. Variations in the multiple tbp genes in different *Halobacterium salinarum* strains and their expression during growth. Arch Microbiol 190:309–318.

Thomm M.2007. Transcription: Mechanism and Regulation. In: Cavicchioli R. (ed.) Archaea: Molecular and Cellular Biology. Washington, DC: ASM Press.

Thomm M, Reich C, Grunberg S and Naji S. 2009. Mutational studies of archaeal RNA polymerase and analysis of hybrid RNA polymerases. Biochem Soc Trans 37:18–22.

Thompson DK, Palmer JR and Daniels CJ. 1999. Expression and heat-responsive regulation of a TFIIB homologue from the archaeon *Haloferax volcanii*. Mol Microbiol 33:1081–1092.

Thomsen J, de Biase A, Kaczanowski S, Macario AJ, Thomm M, Zielenkiewicz P, Maccoll R and Conway de Macario E. 2001. The basal transcription factors TBP and TFB from the mesophilic archaeon Methanosarcina mazeii: structure and conformational changes upon interaction with stress-gene promoters. J Mol Biol 309:589–603.

Ujvari A and Luse DS. 2006. RNA emerging from the active site of RNA polymerase II interacts with the Rpb7 subunit. Nat Struct Mol Biol 13:49–54.

Vierke G, Engelmann A, Hebbeln C and Thomm M. 2003. A novel archaeal transcriptional regulator of heat shock response. J Biol Chem 278:18–26.

Werner F. 2007. Structure and function of archaeal RNA polymerases. Mol Microbiol 65:1395–1404.

Werner F and Weinzierl RO. 2002. A recombinant RNA polymerase II-like enzyme capable of promoter-specific transcription. Mol Cell 10:635–646.

Wilkinson SP, Ouhammouch M and Geiduschek EP. 2010. Transcriptional activation in the context of repression mediated by archaeal histones. Proc Natl Acad Sci USA 107:6777–6781.

Woese CR and Fox GE. 1977. Phylogenetic structure of the prokaryotic domain: the primary kingdoms. Proc Natl Acad Sci USA 74:5088–5090.

Woese CR, Kandler O and Wheelis ML. 1990. Towards a natural system of organisms: proposal for the domains archaea, bacteria, and eucarya. Proc Natl Acad Sci USA 87:4576–4579.

Woychik NA, Lane WS and Young RA. 1991. Yeast RNA polymerase II subunit RPB9 is essential for growth at temperature extremes. J Biol Chem 266:19053–19055.

Wu J and Grunstein M. 2000. 25 years after the nucleosome model: chromatin modifications. Trends Biochem Sci 25:619–623.

Xie Y and Reeve JN. 2004. Transcription by *Methanothermobacter thermautotrophicus* RNA polymerase in vitro releases archaeal transcription factor B but not TATA-box binding protein from the template DNA. J Bacteriol 186:6306–6310.

Yamamoto T, Matsuda T, Inoue T, Matsumura H, Morikawa M, Kanaya S and Kai Y. 2006. Crystal structure of TBP-interacting protein (Tk-TIP26) and implications for its inhibition mechanism of the interaction between TBP and TATA-DNA. Protein Sci 15:152–161.

Zhang G, Campbell EA, Minakhin L, Richter C, Severinov K and Darst SA. 1999. Crystal structure of *Thermus aquaticus* core RNA polymerase at 3.3 A resolution. Cell 98:811–824.

Zhu W, Zeng Q, Colangelo CM, Lewis M, Summers MF and Scott RA. 1996. The N-terminal domain of TFIIB from *Pyrococcus furiosus* forms a zinc ribbon. Nat Struct Biol 3:122–124.

Zillig W, Stetter KO and Tobien M. 1978. DNA-dependent RNA polymerase from *Halobacterium halobium*. Eur J Biochem 91:193–199.

Editor: Michael M. Cox