Function of the trigger loop in distinct steps of the transcription cycle

Dissertation zur Erlangung des Doktorgrades der Naturwissenschaften (Dr. rer. nat.)
der Naturwissenschaftlichen Fakultät III – Biologie und Vorklinische Medizin
der Universität Regensburg

vorgelegt von
Thomas Fouqueau
aus Paris, Frankreich

im Jahr 2013

Promotionsgesuch eingereicht am: 06.08.2013

Diese Arbeit wurde angeleitet von: Prof. Dr. Michael Thomm

Unterschrift:

Table of contents

I) Introduction

This work aims to have a better understanding of transcription machinery, using *Pyrococcus furiosus* (*Pfu*) as a model organism. This organism was isolated in 1986 from geothermally heated marine sediments collected at the beach of Porto Levante in Vulcano Island, Italy (Fiala and Stetter, 1986). *Pyrococcus* (literally "ball of fire") is a genus of Archaea, which represents one of the three domains of life (with Eukarya and Bacteria (Woese et al., 1990)). Archaea were originally seen as extremophiles that lived in harsh environments in terms of temperature, pH, salanity and pressure, such as hot springs and salt lakes, but they have been found in a broad range of mesophilic habitats including oceans (Adams, 1998; Delong, 1998), soils (Bintrimet al., 1997; Leininger et al., 2006) and human intestinal mucosa (Miller et al., 1982; Matarazzo et al., 2012).

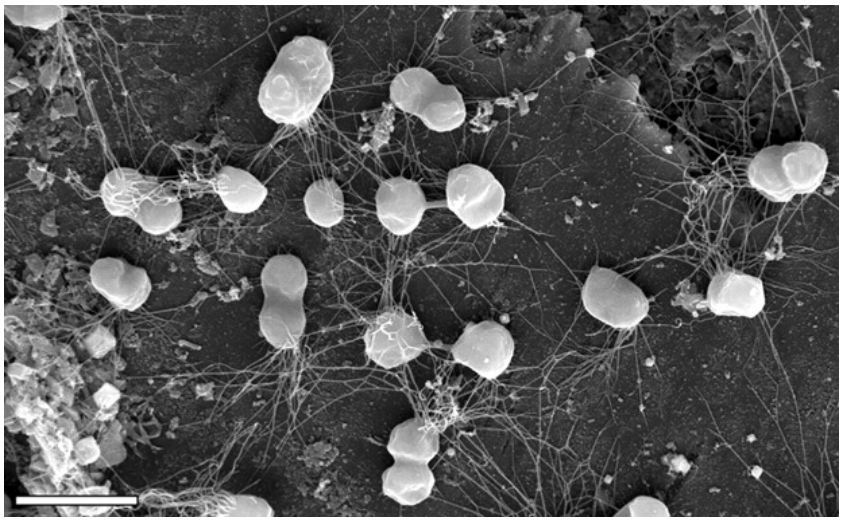

Figure 1. Growth of *P. furiosus* on the surface of sand grains from its natural habitat, visualized by scanning electron microscopy. Flagella attach the cells of the microcolony to the sand grain and to each other. Bar = 2 µm (Närther et al., 2006).

Archaea are very diverse organisms from morphological and metabolic point of view. They are single-celled organisms lacking nuclei and are therefore prokaryotes. Individual archaea range from 0.1 µm to over 15 µm in diameter, and some form aggregates or filaments up to 200 µm in length (Figure 1). Archaea are characterized by their unique ether-linked membrane lipids (Koga and Morii, 2007), and also by their unique enzymes such as specific DNA topoisomerases (Forterre et al., 2007) and DNA polymerases (Ishino et al., 1998). However, Archaea share some characteristics with the other two kingdoms. Thus, like bacteria, archaea usually has a single circular genome, their genes are grouped in operons and are regulated by bacteria-like transcription regulators (Bell et al., 1999). But it is with Eukarya that Archaea share most of its

information-processing systems (replication, transcription and translation). Indeed the majority of translation (Bell and Jackson, 1998), DNA replication (Kelman and White, 2005) and DNA repair factors (Kelman and White, 2005) are specifically shared between Archaea and Eukarya but are not present in Bacteria. Moreover, Archaeal RNA polymerases are closely related to their eukaryotic counterparts, in terms of both subunit composition and structure (Werner, 2008), and by the basal transcription factors required for initiation (Bell and Jackson, 2001). Thus, because of these similarities and because in archaea the number of factors involved are generally lower, archaeal systems act as simplified model systems for complex eukaryal processes. The investigations of transcription using Archaea provide therefore not only insights into the biology of archaeal cell, but allow better understanding of the experimentally limited eukaryotic transcription machinery too.

Study of archaeal transcription machinery using the hyperthermophilic organism *P. furiosus* has several advantages. The organism is capable heterotrophic growth on a wide range *of* substrates (starch, peptone, complex organic substrates, casein, and maltose). Under optimal conditions (100°C, pH 7), *P. furiosus* has a rapid doubling time of 37 minutes and can grow to high cell density (>10^{10} cells/ml). These characteristics are helpful for the isolation of endogenous RNAP (Fiala and Stetter, 1998). Although the organism is strictly anaerobic, the purification and the transcription activity of RNAP can be done under aerobic experimental conditions (Hethke et al., 1996). In addition, all basal transcription factors, transcription regulators and the eleven RNAP subunits can be individually purified from *E.coli* (Hausner et al., 1996; Goede et al., 2006). *P. furiosus* RNAP can be reconstituted from the individual RNAP subunits. This allows the design of RNAP substitution/deletion mutations that are potentially lethal *in vivo* and subsequent specific *in vitro* analysis (Naji et al., 2007; Naji et al., 2008). Since the sequence of the complete genome of *P.furiosus* is known (Robb et al., 2001), the identification and the characterization of transcription factors and regulators could be significantly improved. Moreover, by using a cryo-electron microscopy approach a relatively accurate prediction of *P.furiosus* RNAP architecture was obtained (Kusser et al., 2008). Thus, to over 15 years, the group of Prof. Dr. Michael Thomm contributes to the improvement of the knowledge of the transcription machinery by using the transcription system of *P. furiosus*. Recently, in addition to biochemical approaches and *in vitro* characterization, a genetic system was developed in this organism (Waege et al., 2010), allowing an enhancement of characterization of the transcription machinery of *P. furiosus* by *in vivo* data.

A. DNA-dependent RNA polymerase

All cells accomplish the transcription by one or more DNA dependent multisubunit RNAPs, which consist of 5-15 subunits and a molecular weight of up to 0.7 MDa (Cramer et al., 2008). Bacteria, archaea and chloroplast (PEP, plastid encoded polymerase) contain a single type of RNAP, while the eukaryotes contains three to five distinct types (RNAP I, II, III, IV and V) (Darst,

2001; Kanamaru and Tanaka, 2004; Cramer and Arnold, 2009; Grohmann et al., 2009a; Pikaard and Tucker, 2009; Ream et al., 2009). In addition to those enzymes, single-subunit RNAPs were also described in certain cells, like in mitochondria and chloroplast (NEP, nuclear encoded polymerase) (Gaspari et al., 2004; Kanamaru and Tanaka, 2004). Those enzymes are related to the single-subunit RNAPs from bacteriophages, such as T7, T3 or SP6, from which T7 RNAP is, structurally and functionally, best characterized (Steitz, 2009).

In eukaryotes, RNAP I synthesizes ribosomal RNAs (pre-rRNA 45S in yeast) which will form the major RNA sections of the ribosome. RNAP II synthesizes pre-messenger RNA (pre-mRNAs), small nuclear RNAs (snRNAs, ~125 nt) and small non-coding RNAs (microRNAs, ~22 nt), and RNAP III synthesizes transfer RNAs (tRNAs) and other small RNAs. Finally, RNAP IV and RNAP V, which are specific to the plants, are essential for the synthesis of small interfering RNAs (siRNAs) and other RNAs required for heterochromatin formation and gene silencing (Pikaard et al., 2008.; Wierzbicki et al., 2008;. Ream et al., 2009).

Archaea and bacteria contain only a single RNAP that catalyses the synthesis of all cellular RNAs (Darst, 2001; Grohmann et al., 2009a). Archaeal RNAP is, structurally and mechanistically, closely related to eukaryotic nuclear RNAP II (Langer et al., 1995). Figure 2A shows a comparison of the topology of the essential subunit and transcription factors in bacterial, archaeal and eukaryotic RNAPs (Werner and Grohmann, 2011). Sequence comparisons of the RNAP subunits lay that all multisubunit RNAPs derive from a common precursor enzyme (Huet et al., 1983). The bacterial RNAP has five subunits and any of the bacterial subunits has an archaeal/eukaryotic homologue (Sweetser et al., 1987; Ebright, 2000). The two largest RNAP subunits, β and β´ in bacteria, Rpb1 and Rpb2 in eukaryotes, and RpoA and RpoB (also known as Rpo1 and Rpo2) in archaea form about two-thirds of RNAP to form the catalytic centre and are derived from a common ancestor (Figure 2B) (Zhang et al., 1999; Cramer et al., 2001; Hirata et al., 2008b). In *P. furiosius,* and other archaea, the Rpb1 homologue is split into two subunits denoted RpoA´ and RpoA´´, respectively (Pühler et al., 1989). In Methanogenes and extreme Halophiles, the Rpb2 homologue is also split into two subunits (RpoB 'and RpoB''). The Rpb1/Rpb2 complex is anchored at one end into the Rpb3/Rpb11 heterodimer. Eukaryotic Rpb3/Rpb11 heterodimer (RpoD/L in archaea) together with Rpb10/Rpb12 (RpoN/P in archaea), as well as α-subunit homodimer in bacteria, form the assembly platform required for the efficient assembly and stability of RNAP (Werner et al., 2000; Werner and Weinzierl, 2002; Grohmann et al., 2009a). The smallest bacterial RNAP subunit ω, corresponding to the Rpb6 and RpoK in eukaryotes and archaea, respectively, also promots the RNAP assembly by latching the assembly platform (Minakhin et al., 2001).

The archaeal RpoH subunit lacks the N-terminal domain forming the lower jaw domain in eukaryotic homologue Rpb5. The C-terminal domain makes intricate contacts with the C terminus of the largest subunit (Rpb1 in eukaryotes, RpoA in archaea). Rpb8 and RpoG are located at the bottom of the RNAP between the assembly platform and the pore. Yeast Rpb8 is essential but its precise function remains unclear (Briand et al., 2001). In archaea, RpoG is present only in the

Crenarchaeota (Koonin et al., 2008; Kwapsiz et al., 2008; Korkhin et al., 2009). Recently, good indications of subunits and transcription factor homologies between the three nuclear RNAPs were also obtained by (Kuhn et al., 2007;. Carter and Drouin, 2009).

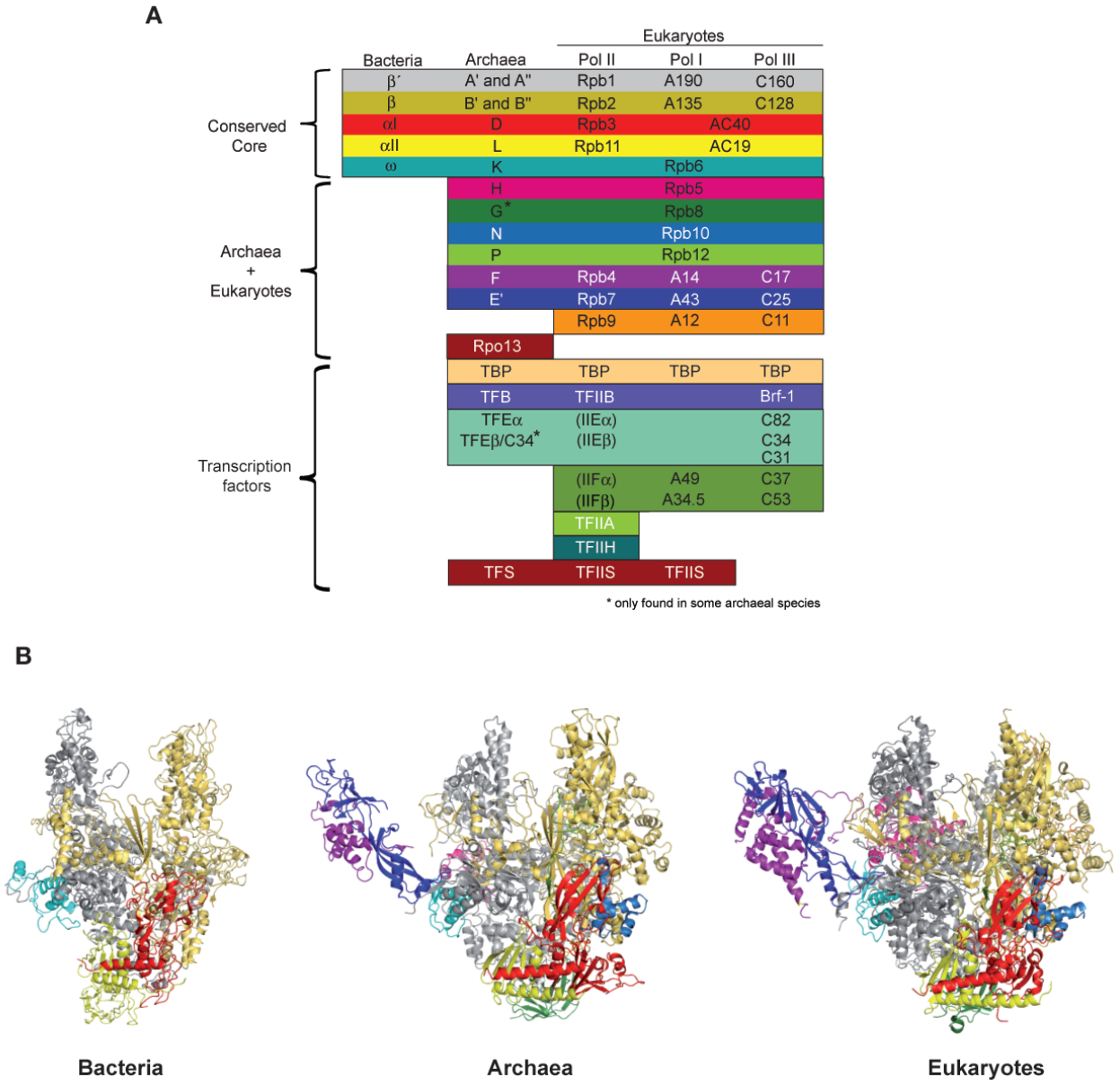

Figure 2: Composition and structure of multisubunit RNAPs. (**A**) Homology pattern of the subunits in the RNAPs of bacteria, archaea and eukaryotes (Werner, 2012). The specificity of the RNAP subunits is indicated in left. (**B**) Overall architecture of RNAPs from bacteria (*Thermus aquaticus* (1HQM) Minakhin et al., 2001), archaea (*Sulfolobus shibatae* (2Y0S) Wojtas et al., 2011) and eukaryotes (*Saccharomyces cerevisiae* (1Y1V) Kettenberger et al., 2004). The color code of the RNAP subunits is same as panel **A**.

Some subunits are specific for a single kingdom of life. Rpb9 is the only subunit found exclusively in eukaryotic RNAPs (Figure 2A). Rpb9 is related to the transcription factor TF(II)S, but with a loss of efficient RNA cleavage activity (Walmacq et al., 2009; Ruan et al., 2011), suggesting that Rpb9 was obtained through gene replication and alteration of catalytic C-ribbon. Rpo13 is the only archaea-specific RNAP subunit, and it is only present in a subset of archaeal genomes (Korkhin et al., 2009). Its function is unclear, but recent biochemical studies suggest that Rpo13 stabilize RNAP-DNA interaction by binding non-specifically to double strand DNA (Wojtas et al., 2012).

The previously mentioned subunits in archaea (RpoB, A ', A'', D, L, N, P, K, H and additional G, Rpo13 in Crenarchaeota) and their eukaryotic homologues in RNAP II (Rpb1, 2, 3, 5, 6, 8, 9, 10, 11 and 12) form the core part of the enzyme RNAP that resembles a crab claw (Cramer et al., 2001). The most pronounced difference between archaeal and eukaryotic enzymes and the bacterial one, is the presence of a stalk-like protrusion (RpoE/F and Rpb4/7 subcomplexes) (Cheetham and Steitz, 2000; Cramer et al., 2001; Hirata et al., 2008a; Grohmann and Werner, 2011). Indeed, the crystal structures of RNAP II and the archaeal RNAP, and also the ones of RNAP I and RNAP III, show the presence of the heterodimer forming the stalk above Rpb6/RpoK subunit (Armache et al., 2003; Bushnell and Kornberg, 2003; Jasiak et al., 2006; Kuhn et al., 2007; Korkhin et al., 2009)

B. Domains and structural elements of RNAPs

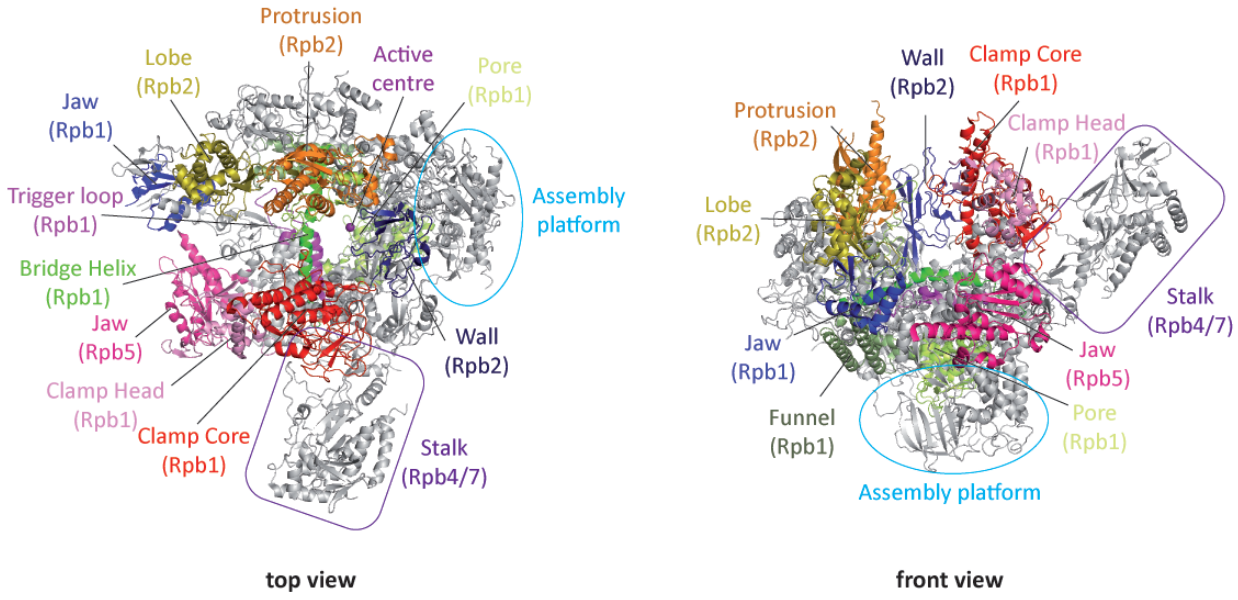

Figure 3. Structural elements of multisubunit RNAPs. Important domains and structural elements of multisubunit RNAPs are shown in RNAP II of *S. cerevisiae* (Kettenberger et al., 2004). The top view shows the active site (Metal ion A) at the centre of the enzyme. The Helix Bridge connects the two halves of the "crab claw", which each consist of mainly Rpb 1 (Clamp) and Rpb2 (Lobe and Protrusion) subunits domains. The front view shows the "Wall" and the position of the "Funnel", which forms the outer edge of the pore or the secondary channel.

1. Stalk (E/F subcomplex)

Archaeal and eukaryotic RNAPs (including RNAP IV and V) contain homologous subunits, which are not present in bacteria (Werner, 2008). E and F subunits (homologous to the eukaryotic Rpb4 and 7, respectively), form a stalk-like protrusion (Figure 3) which plays an important role during transcription initiation (Edwards et al., 1991; Armache et al., 2005; Grohmann et al., 2009). In archaea, those two subunits were shown to facilitate DNA melting and are required for the function of TFE (Werner and Weinzierl, 2005; Naji et a., 2007). During elongation, the E/F subunits interact with the nascent RNA emerging from the RNA exit channel of RNAP, and thus increase the processivity (Ujvári and Luse, 2006; Andrecka et al., 2009; Hirtreiter et al., 2010a). In addition, E/F may stabilize the elongation complex by inducing a conformational change in RNAP, such as the closure of the RNAP clamp (Armache et al., 2005). Recent studies on archaeal transcription termination showed that E/F significantly increases termination efficiency at weak termination signal (five dT stretch) (Hirtreiter et al., 2010a). *In vivo*, archaeal *rpo4* and eukaryotic *rpb4* genes are essential for survival, while archaeal *rpo7* and eukaryotic *rpb4* can be deleted with viability retained at moderate temperatures (Sheffer et al., 1999; Hirata et al., 2008a). Purified fractions of RNAP II of *S. cerevisiae* had substoichiometric amounts of Rpb4/7 that made its structural elucidation difficult for a long time (Cramer, 2004a). Reconstitution of the complete RNAP II from endogenous yeast core and recombinant Rpb4/7 allowed this obstacle to be overcomed structurally (Armache et al., 2003; Bushnell and Kornberg, 2003) and functionally (Edwards et al., 1991; Naji et al., 2007). The idea emerged that, in the yeast system, the stalk can assemble and disassemble during transcription cycle (Edwards et al., 1991). The relative ratio of RNAP II and Rpb4/7 in *S. cerevisiae* is dependent on the growth phase (Choder and Young, 1993). However, recent studies showed that E'/F on the archaeal RNAP from *Methanocaldococcus jannaschii* is stably incorporated into RNAP and that dynamic equilibrium with E'/F does not occur (Grohmann et al., 2009b).

2. Clamp domain

The high stability of RNAP elongation complexes prevents dissociation of RNAP from DNA and allows efficient transcription. This stability is mainly caused by the tight binding of the RNA/DNA hybrid to RNAP (Kireeva et al., 2000; Sidorenkov et al., 1998). In the elongation complex, the hybrid is nested in a highly complementary binding site, created by the closure of the mobile module called the “clamp” (Figure 3). The clamp is open in free RNAP and early transcription initiation complexes but a dramatic 30° rotation of the clamp occurs with the binding of the DNA template strand to three out of five “switch” regions (Gnatt et al., 2001). In the open state, the clamp allows promoter DNA to be loaded into and unwound in the active centre cleft. The binding of RNA/DNA hybrid to the folded switches stabilizes the closed state which accounts for the high

stability of initiation complexes and the high stability and processivity of elongation complexes (Cramer et al., 2001; Gnatt et al., 2001; Chakraborty et al., 2012).

Three loops that protrude from the clamp maintain the arrangement of the nucleic acids during the elongation (Figure 4A). The "rudder" is required for promoter opening in bacteria (Kuznedelov et al., 2002) and for transcription in archaea (Naji et al., 2008). The "lid" is important to stabilize the open promoter complex (Toulokhonov and Landick, 2006), in abortive transcription and serves as a wedge to facilitate RNA displacement by sterically blocking the formation of the overextended hybrid (Gnatt et al., 2001; Naji et al., 2008; Naryshkina et al., 2006). Finally, the double strand DNA is reformed at the back end of the transcript bubble by the "zipper" (Gnatt et al., 2001; Cramer et al., 2001). In bacteria, the zipper also contributes in promoter element (called "Z-element") recognition (Yuzenkova et al., 2011). In addition to these loops, the mobile part of the "flap loop" (flap tip) on top of the "wall" contributes in bubble maintenance and binds to nascent RNA hairpins that pause or terminate bacterial transcription (Figure 4B) (Toulokhonov and Landick, 2003; King et al., 2004). In Archaea and Eukaryotes, RNA hairpins do not affect transcription, probably because the flap tip is shorter in archaeal and eukaryotic RNAPs (Cramer, 2002). Moreover, unlike bacterial RNAP, eukaryotic RNAP II flap loop is not essential for transcription initiation (Palangat et al., 2011).

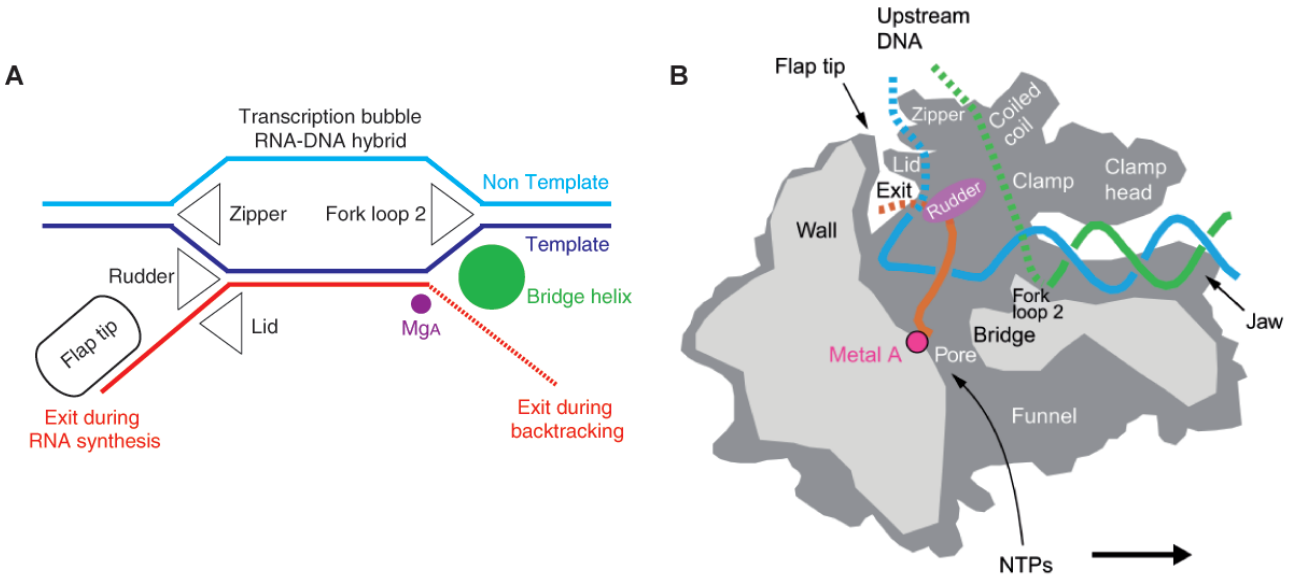

Figure 4. The RNAP elongation complex. (**A**) Schematic presentation of the arrangement of nucleic acids during RNA chain elongation. The DNA template and nontemplate strands are in blue and cyan, respectively, and the RNA is in red. The active site metal ion A is indicated by a pink sphere. Protein elements that are proposed to be involved in the maintenance of the arrangement of nucleic acids are indicated. (**B**) Cutaway view of the RNAP elongation complex. Cut surfaces are lightly shaded. During transcription, DNA enters the enzyme from the right (the polymerase moves to the right). Structural features that appear to be important for function are labeled. The DNA template and nontemplate strands are in blue and green, respectively. (Modified from Cramer, 2002).

The RNAP clamp coiled-coil motif was recently shown to be a binding site for several transcription factors pointing to its importance in transcription initiation and elongation (Figure 4B). The TF(II)B B-linker domain and bacterial σ2 domain, which are involved in promoter opening, were shown to bind to the clamp coiled-coil and the rudder (Kostrewa et al., 2010). Moreover, transcription initiation factor TF(II)E and universally conserved NusG/Spt5 elongation factor compete to bind on clamp coiled-coil motif (Grohmann et al., 2011; Grünberg et al., 2012; Martinez-Rucobo et al., 2011; Werner, 2012). The binding affinities of these factors are context dependent: TFE prevails over Spt4/5 in the initiation complex, whereas Spt4/5 prevails over TFE in the elongation complex. Thus, TFE prevents the inhibitory affect of Spt4/5 on transcription initiation and, during early elongation, Spt4/5 displaces TFE resulting in a high-processivity elongation complex.

3. Switch region

The “switch region” is located at the base of the clamp and serves as the hinge on which the clamp swings during clamp opening and clamp closure (Cramer et al., 2001; Cramer, 2002). Five segments of the switch region, termed “switch 1” through “switch 5”, undergo different conformations in open and closed clamp conformational states. It has been proposed that direct contacts between the switch region and DNA phosphates might coordinate clamp closure and DNA loading into the RNAP active centre (Gnatt et al., 2001; Vassylyev et al., 2007). In bacteria, this region is a target for several antibiotics that inhibit distinct steps of transcription initiation (Belogurov et al., 2009; Mukhopadhyay et al., 2008; Srivastava et al., 2011).

Switch 3 is a polypeptide loop which binds to each RNA base in a nascent transcript as it dissociates from the RNA/DNA hybrid (Kent et al., 2009). In archaea, it was shown to be crucial in transcript elongation, unlike bacteria, in which it is required to form stable complexes with nucleic acid scaffolds by controlling clamp closure (Santangelo and Reeve, 2010; Wiesler et al., 2012). This divergence is likely caused by the differences in charge and flexibility of archaeal and eukaryotic switch 3 loops (Santangelo and Reeve, 2010).

Recent studies on the bacterial switch region suggest that switch 1, 2, 4 and 5 contribute in thestart site melting mechanism (Wiesler et al., 2012). Indeed, a number of substitutions in the switch region affected transcription initiation. Analysis of switch 2 substitutions suggested that this region may be involved in start site selection, abortive initiation, promoter escape and transcript elongation (Majovski et al., 2005; Naji et al., 2007; Pupov et al., 2010). Furthermore, the invariant arginine (*Pfu* A´-R313; *Sce* Rpb1-R337; *Eco* ß´R339) of switch 2 was recently proposed to, in cooperation with switch 1, 4 and 5, undergo conformational changes that stabilize the DNA melting around the start site (Naji et al., 2007; Wiesler et al., 2012).

4. Active site

The catalytic cycle of RNAP (called nucleotide addition cycle) is driven by complex conformation changes that accompany NTP binding, catalysis, and RNAP translocation. When the NTP enters in the RNAP active site, via the secondary channel, a network of interactions between the incoming NTP and active site elements allow the proper positioning of the NTP and its incorporation into the nascent RNA. Recent studies identified two elements in the active centre of RNAP, the "Trigger loop" and the "Bridge helix", which appear to play key roles during the nucleotide addition cycle (Brueckner et al., 2009).

a. The trigger loop

The trigger loop (TL) is a polymorphous element of RNAP active site that is highly conserved among the three domains of life (Figure 5A). The TL is present in the largest subunit of eukaryotic RNAP II Rpb1 and the analogous β′ subunit of bacterial RNAP, and A subunit of archaeal RNAP (A″ in *Pfu* RNAP). In *E.coli* RNAP the TL contains a sequence insertion of 188 aa, called SI3. Structural and biochemical studies in yeast RNAP II and bacterial RNAPs, revealed the importance of the TL in substrates selection and catalysis. The conformational changes of the TL were proposed to link TL-NTP interaction with the substrate positioning and selection but also to be critical in translocation and proofreading (Kaplan et al., 2008; Brueckner et al., 2009; Huang et al., 2010; Yuzenkova et al., 2010; Yuzenkova and Zenkin, 2010; Zhang et al., 2010). Five distinct TL conformations have been observed: "open", "closed", "wedged", "trapped", and "locked" (Figure 5B) (Martinez-Rucobo and Cramer, 2013).

During nucleotide addition, in the absence of substrate, the TL adopts an "open" conformation in which its central part is unstructured (Kettenberger et al., 2004). Binding of an incoming NTP in the +1 site induces folding of the TL, resulting in extension of two helixes at the base of the TL and creating a closed, catalytically competent conformation of the active centre in which the NTP is properly aligned with the 3′-OH of the nascent RNA to facilitate catalysis (Vassylyev et al., 2007b; Wang et al., 2006). The "closed" TL forms a three-helix bundle with the Bridge helix (BH) that interacts with the substrate NTP and the template DNA base, resulting in the closure of the active site. Recent structural analysis on bacterial RNAP and yeast RNAP II proposed that TL residues Rpb1 Q1078, L1081 and L1085 (*Pfu* A′′ Q80, L83, and H87, respectively) contact the 2′-OH group, the base and the triphosphate moieties of the incoming NTP, respectively (Figure 5C), whereas the central part of BH contacts the template base (Vassylyev et al., 2007a; Wang et al., 2006; Yuzenkova et al., 2010; Zhang et al., 2010). However, many additional active centre residues make also essential interaction with the NTP substrate (Nudler, 2009; Cheung et al., 2011). The direct contact between TL residues with the substrate was proposed to link substrate positioning and

recognition, and to be critical for catalysis (Kaplan et al., 2008; Yuzenkova et al., 2010; Cheung et al., 2011).

Moreover, the TL was proposed to participate in translocation during the nucleotide addition cycle, and to be critical in intrinsic cleavage activity in bacteria (see below).

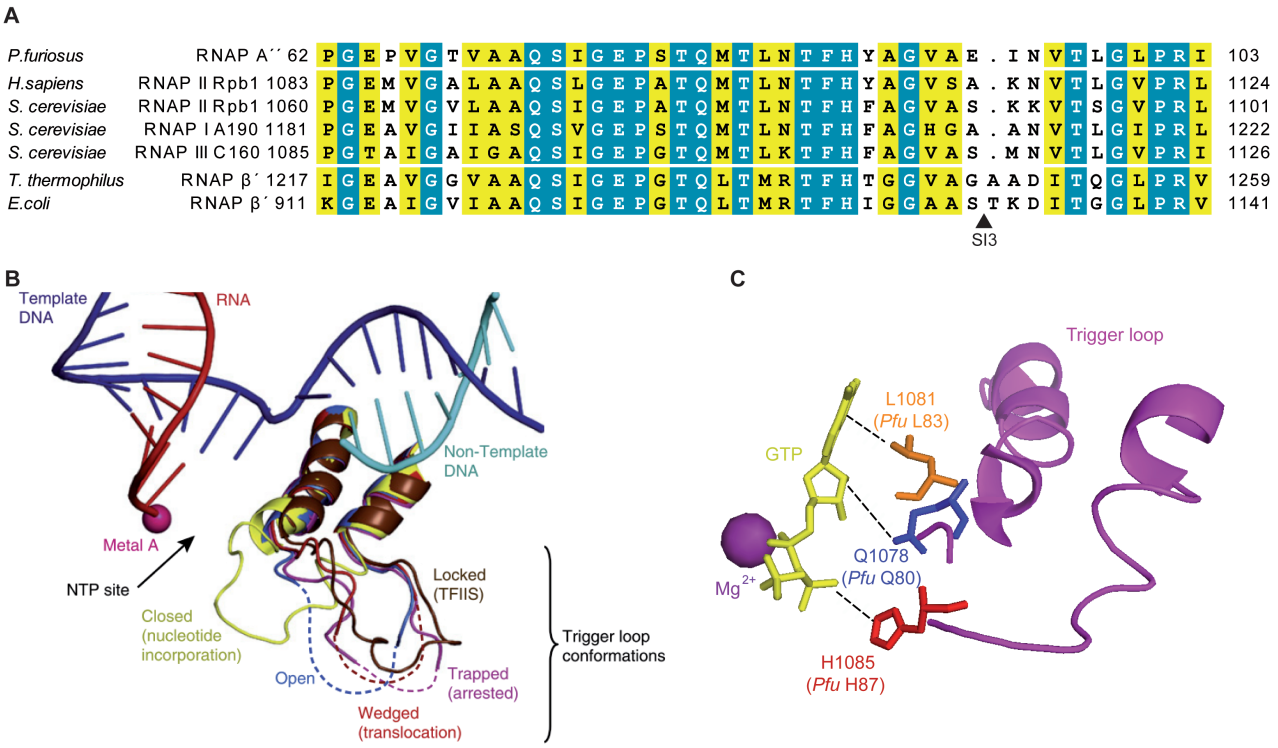

Figure 5. Conserved active site element: the trigger loop (TL). (**A**) The sequences alignment of the TL from archaeal RNAP (*P. furiosus*), eukaryotic RNAP II (*H. sapiens* and *S. cerevisiae*), RNAP I and RNAP III (*S. cerevisiae*) and bacterial RNAPs (*T. thermophilus* and *E. coli*). The black triangle indicates the position of insertion site of SI3 (188 aa) in the *E. coli* RNAP. (**B**) Comparison of TL conformations (Martinez-Rucobo and Cramer, 2013). Superposition of the five RNAP II TL conformations known structurally. "Open" TL in the post-translocation state (PDB 1Y1W (Kettenberger et al., 2004), blue), "closed" TL in the nucleotide incorporation state (PDB 2E2H (Wang et al., 2006), yellow), "wedged" TL in the translocation intermediate (PDB 2VUM (Brueckner and Cramer, 2008), red), "trapped" TL in the arrested complex (PDB 3PO2 (Cheung and Cramer, 2011), violet), and "locked" TL in the reactivation intermediate (PDB 3PO3 (Cheung and Cramer, 2011), brown). DNA template (blue), DNA non-template (cyan), RNA (red) and metal A (pink) are from the open state. (**C**) Closed TL forms a network of interactions with a nucleoside triphosphate (NTP) in the active centre (Wang et al., 2006). When correct NTP enter to the insertion site, TL invariant glutamine residue ((*Sce* Rpb1 Q1078; *Pfu* A´´ Q80), leucine reissue (*Sce* Rpb1 L1081; *Pfu* A´´ L83) and invariant histidine residue (*Sce* Rpb1 H1085; *Pfu* A´´ H87) were suggested to form a network of interaction with the base, the sugar 2´OH-group and β-phosphate of the NTP, respectively.

b. The bridge helix

The BH, an α-helix spanning the active site, is a highly conserved element which acts as pawl in a ratchet-like translocation mechanism to move DNA through the RNAPs (Bar-Nahum et al., 2005). The BH forms a stable three-helix bundle that is structurally flexible and isomerizes between a straight and a kinked conformation. Particularly, two flexible sites of BH, called N-terminal Hinge (BH-H_N) and C-terminal Hinge (BH-H_C), were shown to induce the kinked conformation (Weinzierl, 2011). Because BH kinking was observed on NTP-bound elongation complexes, and an increased BH kinking at the two hinges correlates directly with an increased rate of nucleotide addition, the BH dynamics were suggested to play a role in catalysis (Tan et al., 2008; Vassylyev et al., 2007a; Wang et al., 2006; Weinzierl, 2010). The BH N-terminus is tightly surrounded by other elements, such as experimentally uncharacterized "link domain" and the "F-loop". During nucleotide addition, the BH and F-loop form a gateway that contacts the link domain and the tip of the TL (Miropolkaya et al., 2009). Moreover, C-terminal BH influences the TL conformation and the BH-H_N may alter the position and conformation of the Link domain that is in direct physical contact with the nucleotide (Wienzierl, 2010). Taken together, because of its crucial role in translocation and in catalysis, and because of its contacts with other cleft loops, the BH is proposed to act as a central switchboard for catalysis and substrate movement coordination.

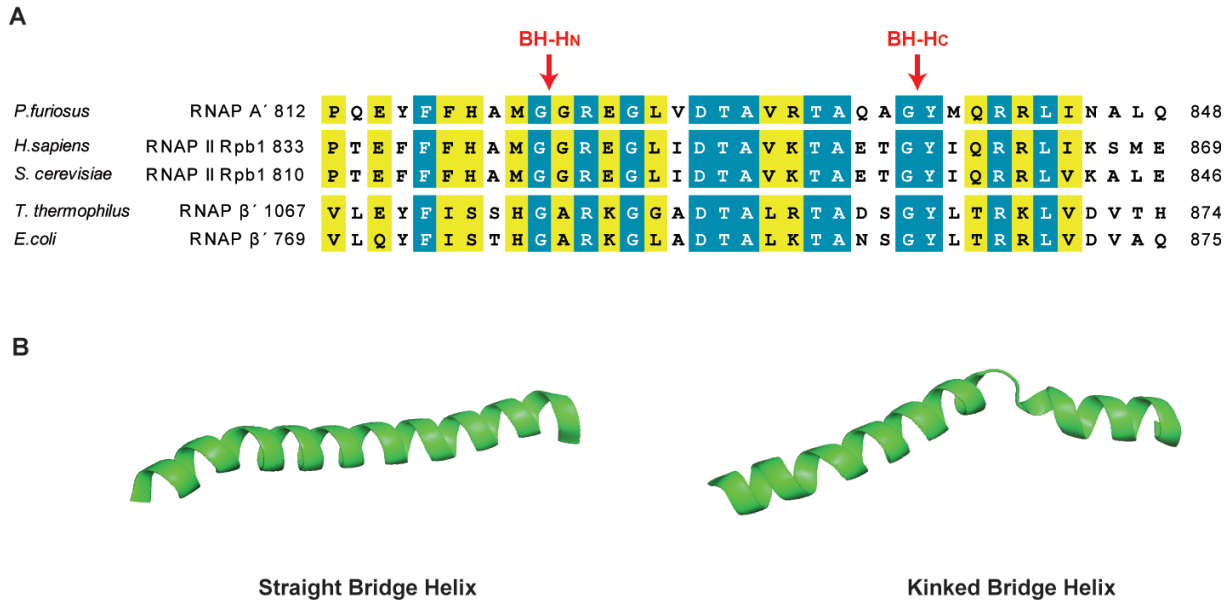

Figure 6. Conserved active site element: the bridge helix (BH). (**A**) The sequences alignment of the BH from archaeal RNAP (*P. furiosus*), eukaryotic RNAP II (*H. sapiens* and *S. cerevisiae*) and bacterial RNAPs (*T. thermophilus* and *E. coli*). The BH N-termus and C-termus Hinge sites are indicated with red arrows. (**B**) Comparison of two BH conformations from *T. thermophilus* RNAP. "Straight" state (PDB: 2O5I (Vassylyev et al., 2007a)) and kinked (PDB: 1IW7 (Vassylyev et al., 2002)) BH.

c. Further active site elements

Other conserved structural features of the RNAP active centre include “fork loops 1 and 2” (FL1 and FL2) and “F loop” (FL).

The FL1 is a small conserved segment of the larger fork domain, in the proximity of the active centre. The FL1, with the lid and the rudder, plays a key role in DNA/RNA strand separation. FL1 contacts the base pairs -6 and -7 in hybrid region, limiting strand separation (Westover et al., 2004a). Structural analysis suggested that FL1 conformation may fluctuate, engaging the single-strand DNA or RNA/DNA hybrid during transcription initiation or elongation respectively. Thus, after the formation of the nascent RNA (>8 nt), FL1 interacts with the rudder to lock the hybrid into a more stable interaction (Meyer et al., 2009).

The flexible FL2 directly interacts with an unpaired DNA residue in the non-template DNA strand, one nucleotide ahead from the active centre (the +2 site) and thus sterically preventing reannealing of the DNA strands (Andrecka et al., 2009; Cramer et al., 2001). This interaction also facilitates NTP sequestration through interaction with the adjacent segment of the fork subdomain I involved in the active centre of RNAP (Kireeva et al., 2011). Thus, FL2 may facilitate the non-catalytic (TL-independent) NTP incorporation in the active centre of RNAP and increase the rate of phosphodiester bond formation (Kennedy and Erie, 2011; Kireeva et al., 2011).

FL is located near the N-terminus of BH and directly contacts the closed TL in the NTP bound transcription elongation complex. Together with the BH, the FL forms a gateway that accommodates the folded TL during nucleotide addition. The FL may be required for the proper folding of the TL and may stabilize the closed conformation of the active centre during catalysis (Miropolskaya et al., 2009; Miropolskaya et al., 2010).

C. Transcription cycle

The synthesis of RNA from a DNA template is conserved among all RNAPs. The transcription cycle is divided into three distinct phases, initiation, elongation and termination, each of which is regulated by various factors (Figure 7). The structure and function of some factors are conserved across the three domains of life (NusG and Spt5), whereas other non-homologous factors show structural and/or functional similarities, suggesting that convergent evolution occurred to allow the same process (For example: Gre and TF(II)S, Sigma and TF(II)B).

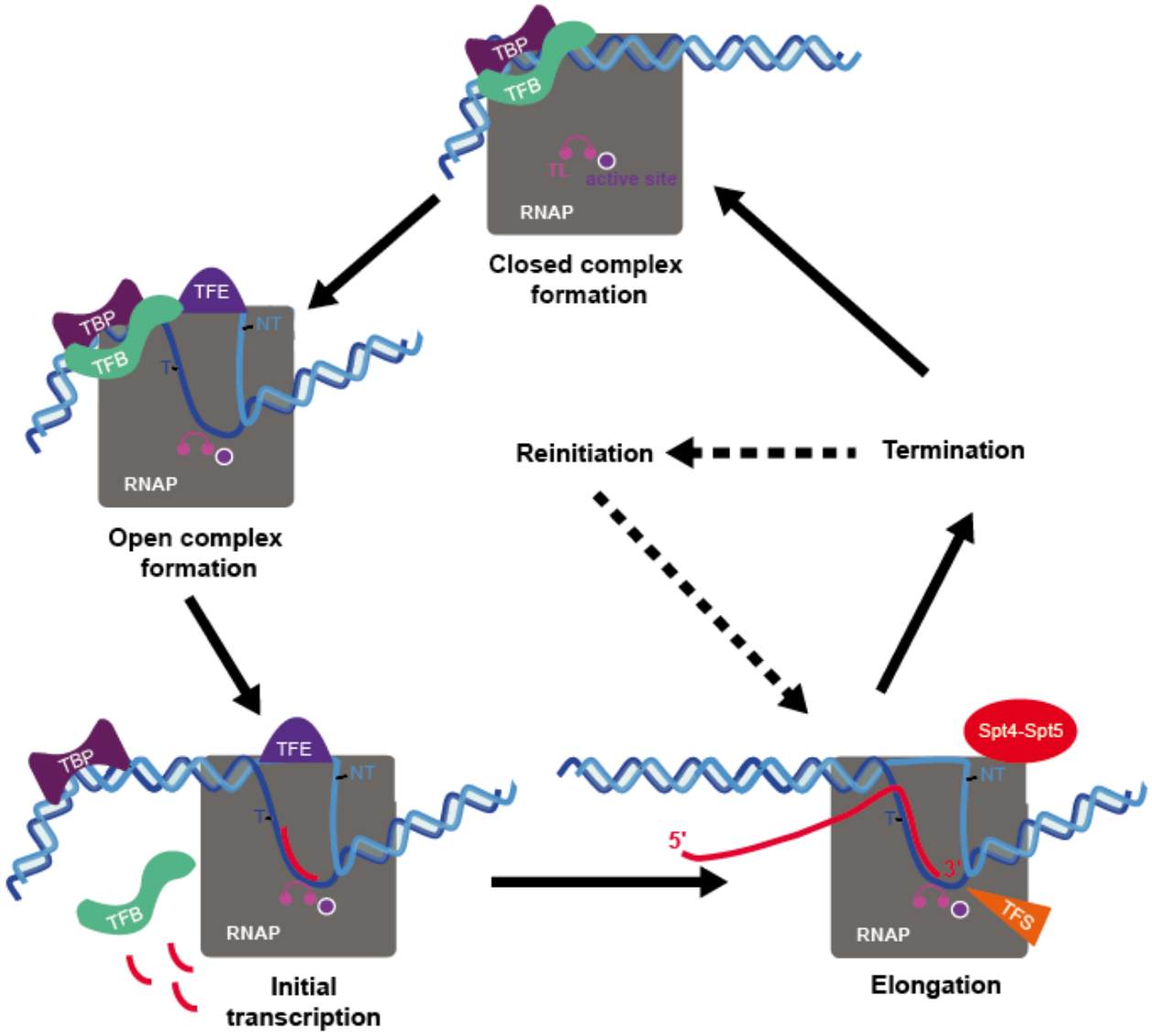

Figure 7. The archaeal transcription cycle. During initiation TBP and TFB assemble on the promoter and recruit RNAP. TFE stimulates DNA melting and the template strand loading into the active site during the next step of initiation. Spt4/5 and TFS associate with the elongation complex and stimulate processivity and proofreading, respectively. The DNA template (T) and non-template (NT) strands are in blue and cyan, respectively. RNA is in red and the active site is in purple.

1. Initiation of transcription

Promoter-directed transcription requires sequence-specific recruitment of RNAP to the promoter, initiation of RNA polymerization in a primer-independent manner and efficient escape from the promoter. This transcription phase is stimulated by evolutionarily unrelated basal initiation factors in all domains of life. However, as the molecular mechanisms of initiation are the same in all three domains and these non-homologues factors utilize the same RNAP-binding sites, they stimulate closely related mechanisms (Grohmann and Werner, 2011).

In bacteria, gene specific Sigma(σ)-factors interact with the core RNAP (ββ´ααω) to form holo-RNAP and enables specific binding of the enzyme to promoters (-10 and -35 elements). In addition to increasing RNAP sequence-specificity for promoters, it also facilitates DNA melting and template strand loading during the closed to open complex transition (Campbell et al., 2008, Murakami and Darst, 2003).

In eukaryotes, distinct general transcription factors (GTFs) are required to form, with the RNAP, the transcription initiation complex. The archaeal RNAP have identical but simplified set of minimal transcription initiation factors to eukaryotic RNAP II (Langer et al., 1995; Bartlett, 2005). Transcription initiation by RNAP II begins with assembly of polymerase and all five general initiation factors into a pre-initiation complex at the promoter and culminates in formation of an open complex and synthesis of the RNA transcript. In the first step, TBP (TATA-binding protein) subunit of TFIID complex (12 TAFs in yeast, TBP associated factors) binds specifically to the TATA box and induces bending of DNA by approximately 90°. TFIIA, by interacting with TBP, can stabilize this complex. In the second step, TFIIB functions as an adaptor by binding specifically to TATA-box-TBP complex and RNAP. TFIIF (Tetramer of two TFIIFα/RAP74 and two TFIIFβ/RAP30) strongly stabilizes this complex and recruits TFIIE (Dimer of TFIIEα and TFIIEβ in Metazoa, trimer of Tgf1, Tgf2 and Tgf3 in yeast) and TFIIH (10 subunits) into the complex. TFIIE binds on the clamp coiled-coil element and is required for open complex formation, that occurs by DNA melting generated by ATP-dependent DNA helicase activity (SSL2/XPB and RAD3/XPD) of TFIIH (Grünberg et al., 2012). The transcription initiation required phosphorylation of CTD provided by Kin28/CDK7 subunit of TFIIH, followed by promoter escape. TFIIF, E and H also suppress promoter-proximal pausing of RNAP (Dvir et al., 2001; Woychik and Hampsey, 2002).

In archaea, in contrast, there are only 3 GTFs named TBP, TFB and TFE, of which only TBP and TFB are essential for promoter-specific *in vitro* transcription initiation (Qureshi et al., 1995; Hausner et al., 1996). While the GTFs of RNAP II machinery are around 30 polypeptides with about 1560 kDa, the three archaeal proteins are only about 80 kDa. Moreover, the melting of the promoter DNA occurs without ATP hydrolysis, and there is so far no evidence of transcription cycle-dependent phosphorylation of archaeal RNAP (Hausner and Thomm, 2001).

Archaeal TBP, as eukaryotic TBP, has a symmetric saddle-shaped structure that is formed by two homologous domains (Figure 8A) (Nikolov et al., 1992). Eukaryotic TBP contains an amino-

terminal domain that is absent in archaeal TBP (Figure 8B). Archaea, however possess 6-10 acidic amino acids at the C-terminus which are not observed in eukaryotic TBPs (Bell and Jackson, 1998). The highly conserved core domain (saddle) is responsible for DNA binding, both upstream and downstream of the TATA box (Cox et al., 1997, Kosa et al., 1997).

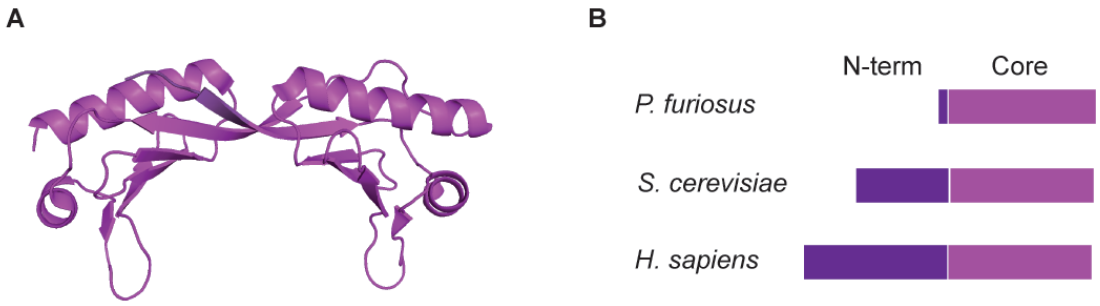

Figure 8. Structure and domain organization of TBP. (**A**) Structure of archaeal TBP from *M. jannaschii* (PDB: 2Z8U (Adachi et al., 2008). (**B**) TBP consists of an N-terminal domain (purple) which is absent in archaea and a highly conserved Core domain (magenta).

Archaeal TFB is a single polypeptide that is highly related to eukaryotic TFIIB (Figure 9A) (Ouzounis and Sander, 1992; Creti et al., 1993). It consists of N-terminal zinc-ribbon domain (Zn-ribbon) (Zhu et al., 1996), which interacts with the dock domain of RNAP, and the C-terminal core domain recognizes the BRE element of the promoter and ensures the correct orientation of the initiation complex (Bell and Jackson, 2000; Qureshi and Jackson, 1998; Lagrange et al., 1998). The highly flexible linker region that connects the TF(II)B domains (consisting of B-reader helix and B-linker) penetrates deep into the active centre of RNAP (Figure 9B). The B-reader is displaced by the growing RNA transcript (> 6 nt), whereas the B-linker is displaced by the rewinding of upstream DNA during TF(II)B release and promoter escape (Bushnell et al., 2004; Kostrewa et al., 2009). While eukaryotes have only one TFIIB, archaea encode mostly for two with Halophilic archaea even up to 6 TBPs and 7 TFBs (Werner, 2007). The additional copies of TFB often exhibit N- or C-terminal truncations or deviations in the functional areas (Werner, 2007). The assumption is that different TBP-TFB-sets, similar to the various σ factors in bacteria, recognize different subsets of promoters efficiently (Facciotti et al., 2007). In *Pfu,* the second TFB (TFB2) functions poorly in promoter-dependent transcription initiation, probably because of a truncation in B-finger/B-linker region (Figure 9A) (Micorescu et al., 2007).

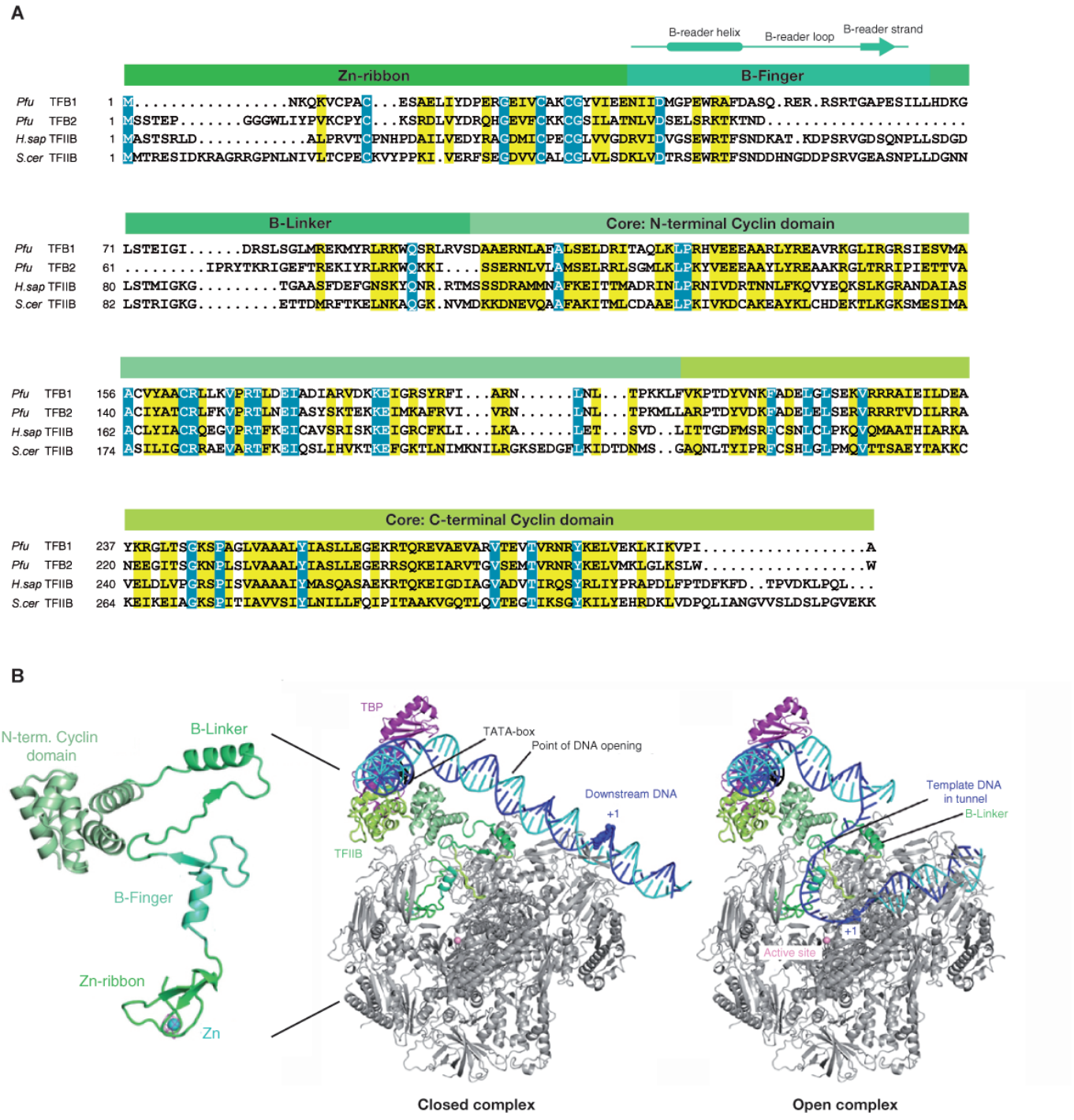

Figure 9. Structure and domain organization of TF(II)B. (**A**) The sequences alignment from archaeal TFB1 and TFB2 (*P. furiosus*) and eukaryotic TFIIB (*H. sapiens* and *S. cerevisiae*). The B-finger organization is indicated. (**B**) Structure of yeast TFIIB as observed in its complex with RNAP II and model of closed and open initiation complexes (Kostrewa et al., 2009). The DNA template and nontemplate strands are in blue and cyan, respectively. The TATA element is in black.

Archaeal TFE corresponds to the N-terminal part of the large subunit TFIIE-α (Figure 10A) (Kyrpides and Ouzounis, 1999). This consists of a winged-helix (WH) motif, which is a special form of the HTH motif (Brennan, 1993), typically found in transcription factors and DNA-binding proteins (Gajiwala and Burley, 2000). The structure of the *S. solfataricus* WH motif has been solved, due to its good preservation within the Archaea, this structure is likely to be applicable to the other archaeal TFEs (Meinhart et al., 2003). The preservation in TFIIE is lower, but sufficient to create homology models. A specific feature of the WH motif of TF(II)E is the extension of the canonical winged helix fold at the N and C termini, and the canonical three helices of the hydrophobic core. Located in the central part of TFE, there is also a conserved Zinc-binding motif and a predicted HTH motif at the non-crystallized C-terminus (Figure 10B) (Meinhart et al., 2003). TFE has a slight stimulatory effect on the transcription at limiting TBP concentrations or at weakly expressed promoters by stabilizing the open pre-initiation complex (Bell et al., 2001; Hanzelka et al., 2001). TFE binds to single stranded DNA (Grünberg et al., 2007), but the effect of TFE depends on the presence of E'/ F subcomplex. Indeed, TFE has, in the presence of E´ subunit, a stimulatory effect on promoter opening and on abortive transcription (Grünberg et al., 2007; Naji et al., 2007). Moreover, the RNAP clamp coiled coil domain and E´/F subcomplex were shown to be crucial for TFE binding and its effect on transcription activity (Ouhammouch et al., 2004; Naji et al., 2007; Grohmann et al., 2011). This suggests that, during transcription initiation, TFE is able to prevent binding of the elongation factor Spt4/5 on RNAP clamp coiled coil domain. Thus, by remaining associated with RNAP during early elongation, TFE can efficiently inhibit the inhibitory effect of Spt4/5 on transcription initiation (Grünberg et al., 2007; Grohmann et al., 2011; Werner, 2012). Recently, archaeal homologue of RPC34 (homologue of TFIIEβ subunit) was identified via computational search, but its function in transcription initiation has not yet been validated experimentally (Blombach et al., 2009).

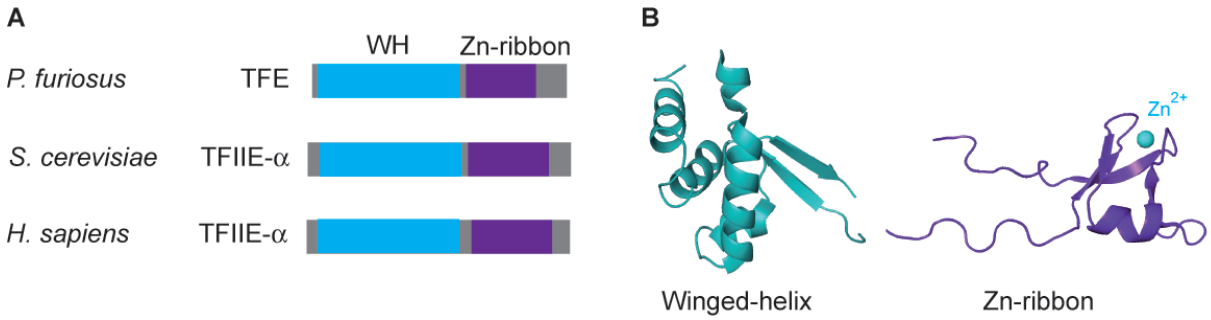

Figure 10. Structure and domain organization of TF(II)E. (**A**) TF(II)E consists of a highly conserved WH domain (cyan) and a Zn-ribbon domain (magenta). (**B**) Structure of archaeal TFE WH domain from *S. solfataricus* (PDB: 1Q1H (Meinhart et al., 2003) and eukaryotic TFIIE-α Zn-ribbon domain from *H. sapiens* (Okuda et al., 2004).

The pre-initiation complex (PIC) is obtained by the assembly of TBP, TF(II)B and archaeal RNAP/RNAP II to the promoter DNA (Bell and Jackson, 1998; Chen and Hahn, 2004; Kostrewa et al., 2009). Recent structural and biochemical studies have shown that the interaction between the B-linker domain of the TF(II)B and the coiled-coil region of clamp domain of RNAP is required for the DNA opening at the transcription start site (Kostrewa et al., 2009). The formation of the "open complex" allows the single-stranded DNA template to enter the active site (Wang et al., 1992; Cramer, 2004b). The RNAPs then enter then a non-productive RNA synthesis phase called the "abortive phase" during which small transcripts (3 to 9 nt) are repeatedly synthesized and released without the disengagement of the RNAP from the promoter (Kapanidis et al., 2006; Goldman et al., 2009). During this phase, the B-reader domain of TF(II)B clashes sterically with the growing RNA (> 5 nt). The B-linker domain is displaced by the rewinding of upstream DNA, promoting the disruption between the RNAP and TF(II)B interactions and, thus, leading to the "promoter escape" (Bushnell et al., 2004; Kostrewa et al., 2009).

2. Elongation

Following promoter escape, the RNA remains stably bound in the transcription elongation complex (TEC). The mechanism of RNA elongation was elucidated by structural studies of RNAP II elongation complex (Gnatt et al., 2001). The TEC is characterized by the transcription bubble that contains a short hybrid duplex formed between the template DNA strand and the RNA product (Kireeva et al., 2000). The "wall" domain of RNAP forces a 90° angle in the 8-9 bp DNA-RNA hybrid (Gnatt et al., 2001). During the formation of the TEC, the clamp swings over the cleft trapping the template and transcript, enhancing the stability of the TEC. The closure of the clamp is induced by the conformational changes of the five so-called "switches", located at the base of the clamp. In addition, cross-linking and structural studies of yeast RNAP II revealed that the two "jaws" grab the incoming DNA template and, thus serve to lend additional stability to the TEC and also to guide the DNA into the active site (Cramer et al., 2000; Korzheva et al., 2000; Wooddell and Burgess, 2000; Gnatt et al., 2001; Murakami et al., 2002). The 3′ end of the RNA is positioned above a pore (also known as secondary channel), through which nucleotides may enter and through which RNA may be extruded during backtracking. Additional structures of RNAP II elongation complex with the NTP substrate allowed nucleotide addition cycle mechanism to be elucidated (Kettenberger et al., 2004; Westover et al., 2004b; Wang et al., 2006; Brueckner and Cramer, 2008; Brueckner et al., 2009).

a. Nucleotide addition cycle

A conserved nucleotide addition cycle mechanism has been recently proposed for all three kingdoms of life (Brueckner et al., 2008; Brueckner and Cramer, 2009). Structural studies are consistent with a two-metal ion mechanism to catalyze nucleotide addition during elongation of a growing RNA (Steitz, 1998). The incoming NTP binds two Mg^{2+} ions designated as Mg_A and Mg_B in eukaryotes and archaea, and as Mg_I and Mg_{II} in bacteria (Cramer et al., 2001; Westover et al., 2004a). Mg_A is bound to the active site aspartate loop and binds the RNA 3´end, whereas Mg_B enters with the incoming NTP, bound to its triphosphate moiety (Cramer et al., 2001; Westover et al., 2004a). The incoming NTP first binds in the "pre-insertion site" in an open active centre conformation (Kettenberger et al., 2004; Vassylyev et al., 2007b). Catalysis is not permitted in this state, as the NTP triphosphate and Mg_B are too far from Mg_A (Brueckner et al., 2009). Folding of the TL then leads to closure of the active centre and shifts the NTP to the insertion site, where all the contacts required for catalysis occurs (Wang et al., 2006; Vassylyev et al., 2007b). Incorporation of the nucleotide in the 3´end of the RNA chain results in formation of a pyrophosphate ion (PPi). After NMP incorporation and PPi release, the enzyme move one nucleotide downstream the DNA to allow for addition of the next nucleotide. This process is likely accompanied by the TL relaxation and PPi release that may precede, or be coupled with, translocation (Brueckner and Cramer, 2008; Kireeva et al., 2010; Malinen et al., 2012). In the intermediate step of pre- to post-translocation, the TL was observed in a "wedged" conformation, bending the central BH. The conformational change of the central region of the BH is accompanied by the translocation of the downstream DNA until the next DNA template base reaches the pre-templating position above the bridge helix (Brueckner et al., 2009). In an archaeal reconstitution system, it has been shown that the kinking of the bridge helix is critical for catalysis (Tan et al., 2008).

The NAC mechanism is based on Brownian ratchet model that underlies TEC oscillation between pre-translocation and post-translocation states and directional polymerase movement (Landick, 2004; Bar-Nahum et al., 2005; Sousa, 2005). NTP binding temporarily stops the oscillation, acting like a pawl of a ratchet. Thus, after nucleotide incorporation, oscillation resumes around the next template position.

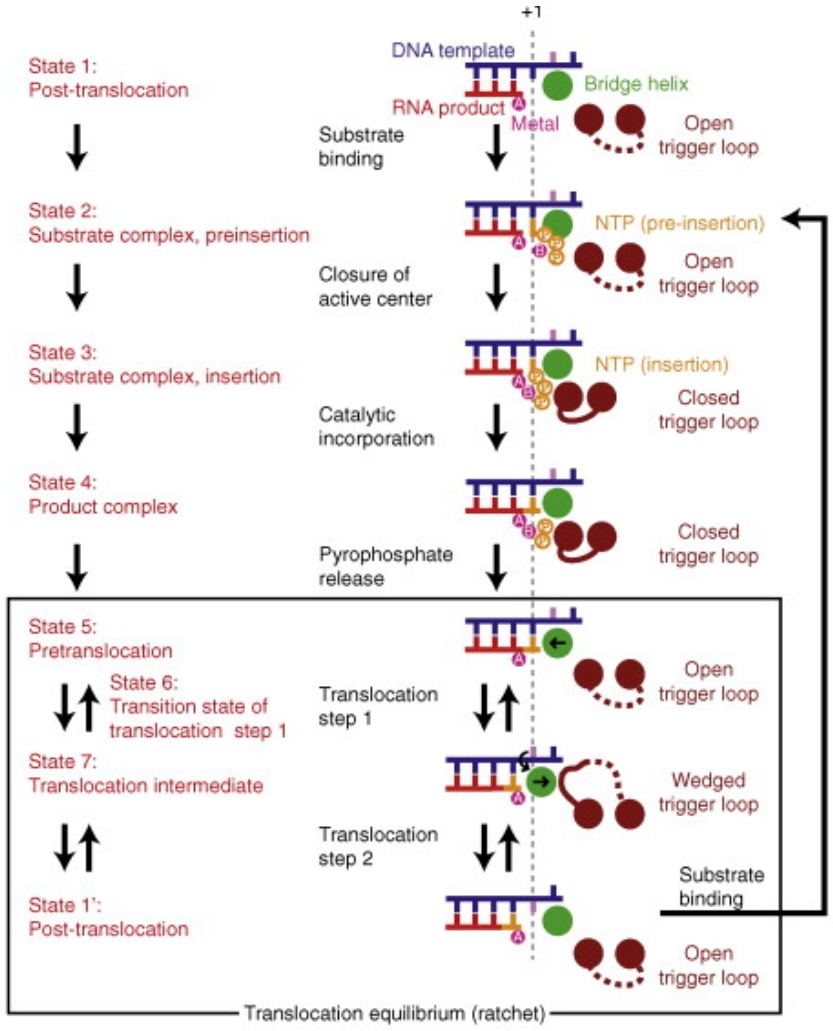

Figure 11. Schematic representation of the nucleotide addition cycle (Brueckner and Cramer, 2008; Brueckner et al., 2009). The phases of the nucleotide addition cycle are indicated in left. The vertical dashed line indicates the +1 site.

b. Nucleotide selection

During elongation, the RNAP selects the correct NTP substrate, discriminates rNTPs from dNTPs and selects the rNTP that is complementary to the DNA template base. Substrate selection was proposed to occur in two steps and to involve a conformation change of the TL from an open to a closed state (Vassylyev et al., 2007b; Kaplan et al., 2008; Kireeva et al., 2008; Yuzenkova et al., 2010; Zhang et al., 2010). The NTP substrate binds to an open active centre in a catalytically inactive pre-insertion state. The complement NTP (cNTP) base pairs with the DNA template base inducing folding of the TL which closes the active site and shifts the NTP to the insertion site (+1

site, also known as A site) (Westover et al., 2004; Wang et al., 2006; Vassylyev et al., 2007b; Brueckner and Cramer, 2008). Binding of non-complementary NTP (ncNTP) shifts the equilibrium between the closed and open conformations of the active site towards the open conformation, making release of the incorrect NTP more likely than nucleotide incorporation, thereby impairing misincorporation (Kireeva et al., 2008). In yeast and bacteria, mutations of several residues of the TL have been shown to affect this fidelity. In *E. coli*, β' G1136S mutation rendered the polymerase less discriminative against ncNTPs or complementary 2'dNTP (c2'dNTP), whereas the I1134V mutant enzyme was more accurate than the wild type (Bar-Nahum et al., 2005). In *T. aquaticus*, the affinity of a β' M1238A mutant of RNAP with ncNTPs was significantly higher, suggesting a critical role of this residue in ncNTPs discrimination (Yuzenkova et al., 2010). Similarly, genetically selected TL mutations H1085Y and E1103G were shown to compromise the fidelity of yeast RNAP II *in vitro* and *in vivo* (Kaplan et al., 2008; Kireeva et al., 2008). In addition, residues in close proximity to the TL were proposed also to be important for fidelity, as shown with the substitution mutant of the residue β D675 in *E. coli* RNAP (Holmes et al., 2006).

RNAPs discriminate cNTPs from c2'dNTPs by recognition of the 2'-OH group of the ribose. A conserved asparagine residue in the active centre (*E. coli* β' N458, *S. cerevisiae* Rpb1 N479, *Pfu* A' N456, respectively) was proposed to play a role in NTP/2'dNTP discrimination by contacting the ribose 2'-OH group (Kettenberger et al., 2003; Svetlov et al., 2004). Indeed, substitution of these residues leads to an increase in 2'dNTP incorporation (Svetlov et al., 2004; Wang et al., 2006). In addition, a conserved arginine (*E. coli* β' R425; *S. cerevisiae* Rpb1 R446; *Pfu* A' R423) was shown to contact the 2'-OH group in the crystal structures of *T. thermophilus* RNAP and yeast RNAP II elongation complexes (Wang et al., 2006; Vassylyev et al., 2007b; Cheung et al., 2011). The contact between TL residues and NTP was also proposed to be critical in dNTPs recognition. Yeast Rpb1 Q1078 and N1082 (*Pfu* A'' Q80 and N84, respectively) were shown to contact the ribose 2'-OH and 3'-OH groups, respectively, suggesting a possible contribution of these residues to NTP selection (Wang et al., 2006; Cheung et al., 2011).

c. Proofreading

During elongation phase, when a nucleotide misincorporation occurs, the mismatched 3'-nucleotide frays from the DNA template and induces elongation pausing and inhibition of RNA extension (Sydow *et al.,* 2009). Backtracking of the elongation complex by one nucleotide aligns the scissile phosphodiester bond with the active centre. The misincorporated nucleotide can then be removed via two modes: by intrinsic RNAP endonucleolytic activity or factor dependent cleavage activity (Gre factors in bacteria, TFIIS in eukaryote and its homologue TFS in archaea) (Erie et al., 1993; Thomas et al., 1998; Kettenberger et al., 2003; Laptenko et al., 2003; Lange and Hausner, 2004; Zenkin et al., 2006). Cleavage at the 3' end of the RNA transcript frees the NTP-binding site and creates a new 3'-OH group in the active site, allowing the transcription to resume.

In *T. aquaticus* RNAP, inhibition of intrinsic RNA cleavage was observed in the presence of the antibiotic Streptolydigin, which blocks TL in its inactive conformation, suggesting the critical role of the bacterial TL in RNA transcript hydrolysis (Temiakov *et al.,* 2005; Tuske *et al.,* 2005). In contrast, factor dependent RNA cleavage was proposed to occur independently of the TL (Rhoganian *et al.,* 2011). Recent co-crystallization of yeast RNAP II with TFIIS and bacterial RNAP with Gre factors showed the TL locked in inactive conformation by the acidic loop of these factors (Kettenberger et al., 2004; Wang et al., 2009; Tagami et al., 2010; Cheung and Cramer, 2011).

α. Intrinsic RNA cleavage

Intrinsic RNA cleavage activity in ternary elongation complexes was first observed in *E. coli* RNAP Independently of external factors (Surratt et al., 1991; Orlova et al., 1995). In eukaryotes, whereas RNAP I (Kuhn et al., 2007) and RNAP III (Thuillier et al., 1996; Alic et al., 2007) possess a strong intrinsic RNA cleavage activity, RNAP II possesses a very weak activity. The molecular basis for this difference can be explained by the lack of a catalytic residue (corresponding to E291 of TFIIS) on the C-ribbon domain of Rpb9 (Ruan et al., 2011). Since bacterial and archaeal RNAPs lack the Rpb9 homologue, the intrinsic cleavage was suggested to be carried out by the polymerization site itself (Rudd et al., 1994). Alkaline pH substantially stimulates RNA hydrolysis suggesting that increased deprotonation of water molecules in the active site may attack the phosphorous atom of the scissile phosphodiester bond (Orlova et al., 1995; Sosunov et al., 2003; Weilbaecher et al., 2003). Recently, structural and biochemical studies proposed a mechanism of intrinsic RNA cleavage (Figure 12) (Zenkin et al., 2006; Wang et al., 2009; Yuzenkova and Zenkin, 2010). When a nucleotide misincorporation occurs the RNAP pauses and the mismatched 3' NMP is frayed away from the DNA template, which prevents RNA extension by overlapping the tip of the closed TL and the NTP in the insertion site (Toulokhonov et al., 2007; Sydow et al., 2009). The polymerase then backtracks by one step, moving the 3′-terminal mismatched NMP from the fraying site to a proofreading site ("P" site) at position +2 (Sydow et al., 2009; Wang et al., 2009). In bacteria, the mismatched nucleotide was proposed to stabilize the Mg_{II} and/or the active water molecule, and thus stimulate transcript-assisted cleavage (Zenkin et al., 2006). Hydrolysis of a dinucleotide containing the mismatched NMP frees the NTP-binding site and restores the nucleophilic RNA 3′-OH group in the active site, allowing transcription to resume. The TL, especially the invariant histidine residue, was recently shown to be critical for intrinsic RNA cleavage in *T. aquaticus* RNAP (Yuzenkova and Zenkin, 2010). *Taq* β' H1242 was proposed to directly participate in penultimate phosphodiester bond reaction in two ways: (*i*) as a general base and (*ii*) by positioning the 3′ NMP moiety of the RNA for transcript-assisted catalysis. This conclusion contradicts an earlier study on *E. coli* RNAP (Zhang et al., 2009). However, these contradictory observations are probably due to the extensive noncomplementarity of the template and nontemplate DNA strands and a short RNA-DNA hybrid used in *E. coli* system, which may lock the TL in an inactive conformation. In yeast

RNAP II, the absence of the effect of α-amanitin on the intrinsic RNA cleavage argues against the involvement of the TL in an endonuclease reaction (Weilbaecher et al., 2003).

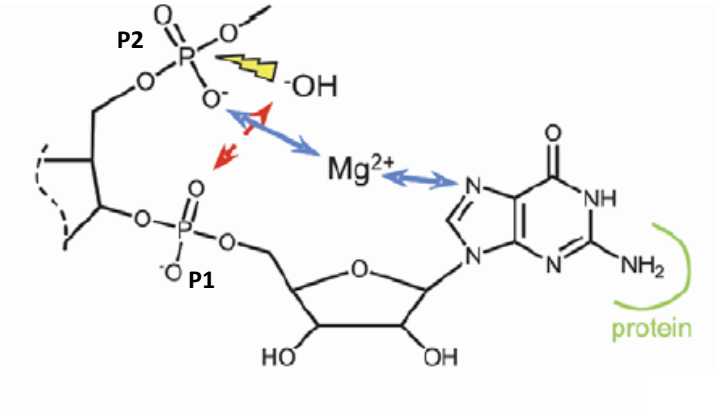

Figure 12. Transcript-assisted 3´ end nucleotide cleavage (P1: ultimate phosphodiester bond; P2: penultimate phosphodiester bond; ¯OH: active water molecule; red: hydrogen bonding, blue: coordination bonds) (Zenkin et al., 2006)

β. Factor-stimulated RNA cleavage

RNA cleavage by RNAP is dramatically stimulated by exogenous transcription factors. In bacteria, Gre factors (GreA and GreB) functionally correspond to the TF(II)S of Archaea and Eukaryote. The Gre factors consist of two domains: the N-terminal coiled-coil domain with two conserved acidic residues (Asp and Glu) on its tip, and the C-terminal barrel domain (Stebbins et al., 1995; Lamour et al., 2006; Laptenko et al., 2006; Symersky et al., 2006). The C-terminal domain of all Gre factors binds to the bacterial RNAP secondary channel. Whereas GreA induces the cleavage of 2-3 nt RNA fragments on short backtracked complexes, GreB induces the cleavage of 2 to 18 nt fragments in backtracked and arrested complexes (Borukhov et al., 1993; Feng et al., 1994). Their distinct functions derive from the difference of an essential basic patch on the surface of the N-terminal domain (Koulich et al., 1997; Kulish et al., 2000). Gre factors inhibit all known RNAP activities *in vitro* during transcription initiation and elongation (Hogan et al., 2002; Symersky et al., 2006). A recent crystal structure of the RNAP EC bound with Gfh1 (Gre factor homologue from the genus *Thermus*) from *T. thermophilus* revealed that the coiled-coil domain occludes completely the second channel and places the tip loop near the catalytic site (Tagami et al., 2010). Thus, Gfh1 blocks the access of the substrate NTP to the catalytic site and, in addition, traps the TL in an inactive, open conformation (distinct from the "locked" conformation observed in yeast RNAP II-TFIIS complexes, see below) (Figure 5B). Moreover, the Gre factors selectively bind on "ratcheted" RNAP to freeze the RNAP. The ratcheted state is characterized by the kinking of the BH and the expansion of the secondary channel that is probably required for Gfh1 binding (Tagami

et al., 2010; Sekine et al., 2012). Finally, the two conserved acidic residues on the tip stimulate the RNA cleavage by fixing the Mg_{II} and displacing the active water molecule in active site (Laptenko et al., 2003; Opalka et al., 2003; Sosunova et al., 2003).

Elongation factor TFIIS stimulates transcript RNA cleavage in RNAP II (Izban and Luse, 1992; Rudd et al., 1994). TFIIS is not essential for cell viability in yeast but its deletion generates sensitivity to oxidants like menadione and drugs like 6-azauracil (Nakanishi et al., 1995; Koyama et al., 2003; Koyama et al., 2007). TFIIS also interacts with RNAP I and is able to induce RNA cleavage (Schnapp et al., 1996). Recently, TFIIS was also detected as a RNAP III transcription factor (Ghavi-Helm et al., 2008). TFIIS of *S. cerevisiae* cleaves mismatched nucleotides faster than correctly paired RNA (Wang et al., 2009). Whereas TFIIS stimulates primarily dinucleotide cleavage in stalled complex, it can stimulate the cleavage of an oligonucleotide of up to 17 nt in a pronounced backtracked, arrested complex (Izban and Luse, 1992; Izban and Luse, 1993). In archaea, TFS also stimulates primarily dinucleotide cleavage (Hausner et al., 2000), but the pronounced backtracked complex was not observed (Spitalny and Thomm, 2003; Lange and Hausner, 2004). The structure of TFIIS was originally partially resolved by limited proteolysis and NMR. Based on these results TFIIS was divided into three domains (I, II and III) (Figure 13B) (Morin et al., 1996; Olmsted et al., 1998). The N-terminal domain I is poorly conserved among the TFIIS homologs and its structure has not been resolved yet (Labhart and Morgan, 1998). Domain II, III and the linker domains are fully sufficient for binding to the polymerase and for RNA cleavage stimulatory function of TFIIS (Awrey et al., 1998). Domain II forms a three-helix bundle, and domain III adopts a zinc-ribbon fold with a thin protruding β-hairpin. Domain II and the linker between domains II and III are required for RNAP II binding. Domain III inserts into the pore, placing the β-hairpin tip near the catalytic site, which stimulates RNA cleavage (Figure 13 and 14) (Cipres-Palacin and Kane, 1995; Awrey et al., 1998). Domain III is highly conserved between eukaryotes and archaea (Hausner et al., 2000). Recent crystal structure of yeast RNAP II-TFIIS complexes revealed that TFIIS binding is accompanied by large conformational change of RNAP II (Kettenberger et al., 2003; Kettenberger et al., 2004; Cheung and Cramer, 2010). TFIIS binds to the Rpb1 jaw domain through two conserved acidic residues (Asp and Glu) in β-hairpin of the domain III that are essential for the RNA cleavage activity (red triangles, Figure 13) (Jeon et al., 1994). These residues recruit and position Mg_B and position a water molecule in active centre, which acts as the nucleophile (Wang et al., 2009). Thus, a two metal ion mechanism, analogous to the bacterial Gre one, was proposed for TFIIS-stimulated RNA cleavage (Sosunov et al., 2003; Kettenberger et al., 2003; Wang et al., 2009; Sosunova et al., 2013). Moreover, domain III replaces the TL away from the backtracked RNA by trapping it in an inactive "locked" conformation, thus facilitating the cleavage reaction (Kettenberger et al., 2003; Kettenberger et al., 2004; Cheung et al., 2011).

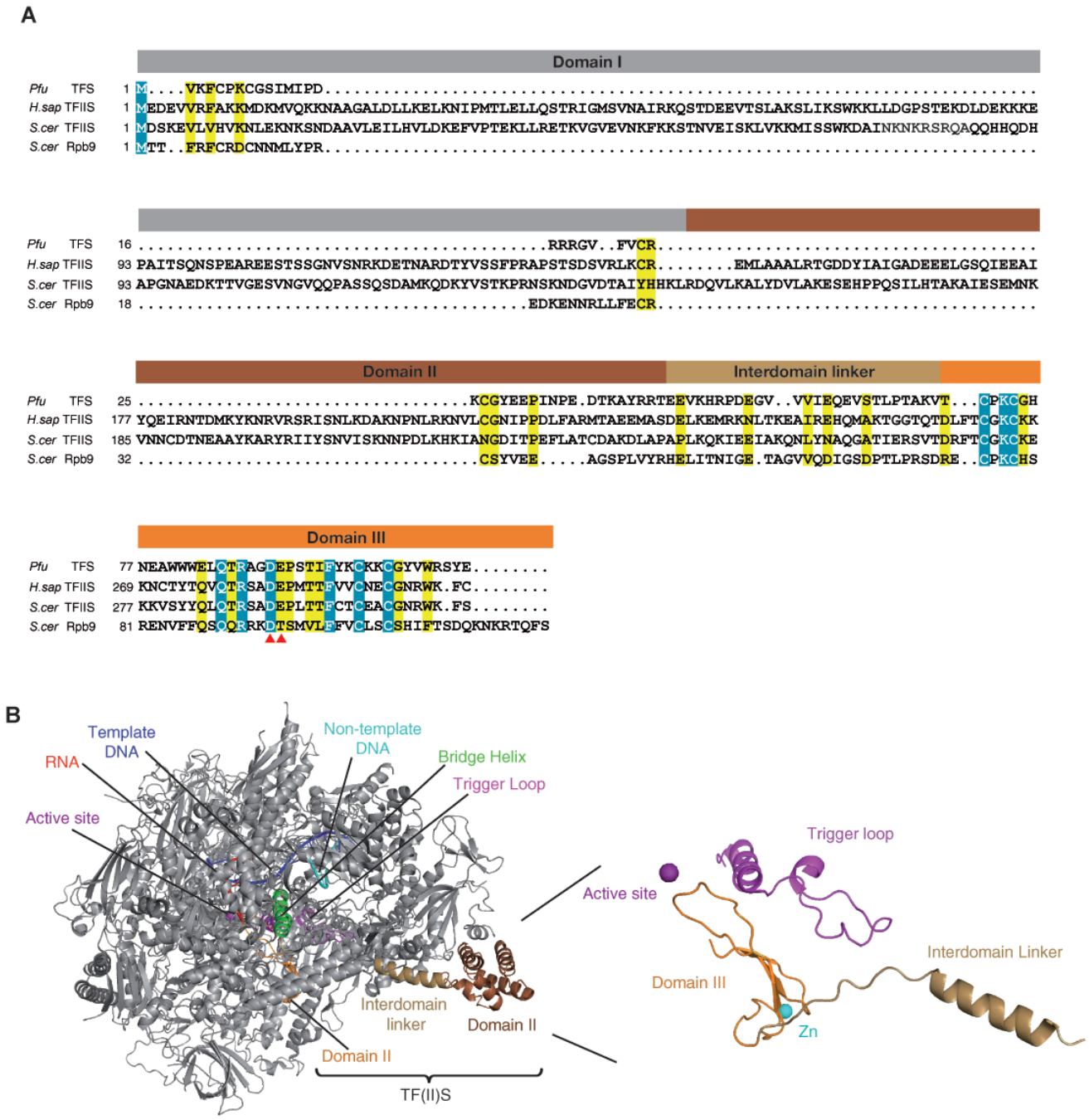

Figure 13. Structure and domain organization of TF(II)S. (**A**) The sequences alignment from archaeal TFS (*P. furiosus*), eukaryotic TFIIS (*H. sapiens* and *S. cerevisiae*) and Rpb9 subunit (*S. cerevisiae*). The catalytic residues (Asp and Glu) are indicated by red triangles. (**B**) Model of elongation complex with yeast TFIIS and structure of TFIIS and "locked" TL (PDB: 3PO3) (Cheung and Cramer, 2011).

Although the eukaryotic TFIIS and bacterial Gre factors lack sequence and structure similarities, both factors (*i*) bind to the secondary channel of RNAP, (*ii*) contain two conserved acidic residues at the tip of their catalytic domains, (*iii*) control RNAP activities by locking the TL in an inactive conformation and (*iv)* stimulate RNA cleavage via analogous two metal ion mechanisms. Thus, these similarities in their mechanisms of action indicate a universally conserved RNAP regulation mechanism of the cleavage factors.

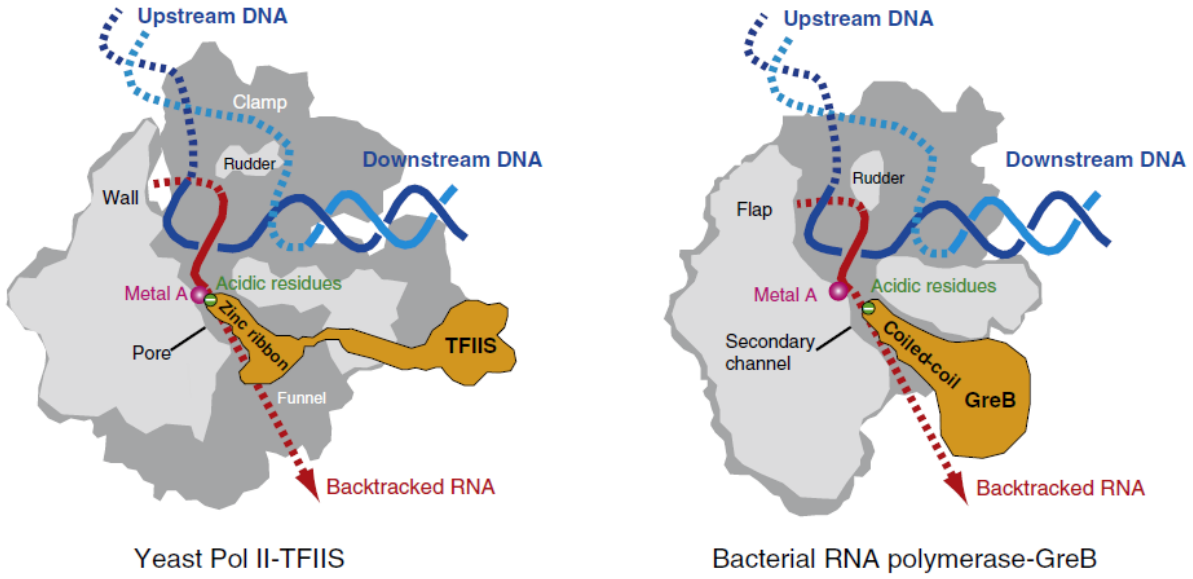

Figure 14. Schematic cutaway of the yeast RNAP II-TFIIS and bacterial RNAP-GreB complexes (Martinez-Rucobo and Cramer, 2012). TFIIS and Gre factors reach the active site of RNAPs with a hairpin that conserved the conserved acidic residues required for the simulation of RNA cleavage.

d. Processivity

Processivity refers to the property of the polymerase to remain associated on the transcribed DNA template during transcription elongation. Processivity is achieved by the high stability of the EC, which results from the tight binding on the DNA-RNA hybrid (Kireeva et al., 2000). In Archaea and Eukarya, the E'/F subunits complex was proposed to stimulate elongation and increase processivity by inducing the closure of RNAP clamp and by interacting with the nascent RNA (Armache et al., 2005; Hirtreiter et al., 2010a). In addition, the elongation factor Spt5 (NusG in bacteria), was shown to stimulate transcription processivity and elongation in archaeal and bacterial systems, respectively (Herbert et al., 2006; Herbert et al., 2010; Hirtreiter et al., 2010b), although it function has not yet been demonstrated in the eukaryotic system. Bacterial NusG and

archaeal Spt5 consist of an N-terminal NGN and C-terminal Kyprides, Ouzounis, Woese (KOW) domains. Eukaryotic Spt5 additionally contains an N-terminal acidic region and a set of short repeats at its C-terminus (CTR) (Figure). In archaea and Eukarya, Spt5 forms a heterodimer with Spt4 through its NGN domain (Werner, 2012). In addition, Spt5 NGN domain was shown to bind to the coiled-coil domain of the RNAP clamp, and serve as a recruitment platform for additional factors (Grohmann et al., 2011; Martinez-Rucobo et al., 2011). Recent structural and biochemical studies revealed that Spt5 stimulates the processivity by oscillating the active centre cleft with the NGN domain, locking the nucleic acids in the cleft and preventing their dissociation (Sevostyanova and Artsimovitch, 2010; Klein et al., 2011; Martinez-Rucobo et al., 2011).

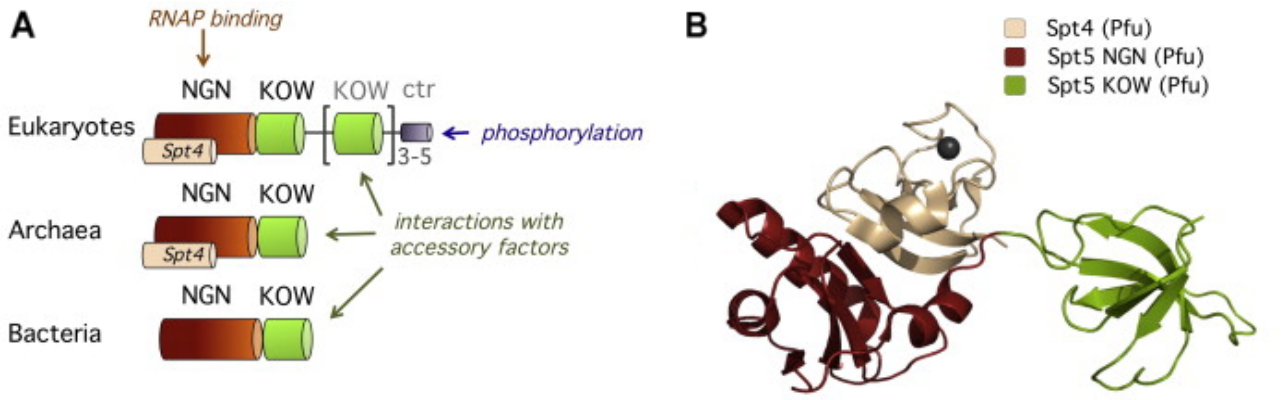

Figure 15. Structure and domain organization of Spt5-like factors (Werner, 2012)**.** (**A**) Spt4/5 and NusG consist of an NGN domain (red) and one or more KOW (*K*yrpidis, *O*uzounis, *W*oese) domains (green), only eukaryotic Spt5 contains two C-terminal repeats (ctr). In archaea and eukaryote, Spt5 forms a complex with Spt4 (wheat). (**B**) Structure of archaeal Spt4/5 complex from *P. furiosus* (PDB: 3P3B (Martinez-Rucobo, et al., 2011)).

3. Termination

Successful execution of the gene expression requires dissociation of the stable EC in order to release the RNA transcript and make RNAP available for the next round of transcription. Bacteria use two distinct strategies for transcription termination: intrinsic and factor-dependent termination (Santangelo and Artsimovitch, 2011). The intrinsic terminator signal is composed of a GC-rich RNA hairpin followed by a run of U residues (Gusarov and Nudler, 1999). Transcription of this poly-uracil sequence induces the pausing of the EC and the weak A-U bonds has a lower RNA-DNA duplex energy, allowing it to unwind and dissociate from the RNAP (Epshtein et al., 2007). However, many bacterial genomes lack intrinsic terminator signals, indicating the existence of a factor dependent termination mechanism. The Rho factor is an ATP-dependent helicase that binds to ribosome free RNAs, translocates to the EC and unwinds the RNA-DNA duplex (Epshtein et al.,

2010). Termination by eukaryotic RNAP II is more complex and involves both polyadenylation of the transcript and an exonuclease (Xrn2/Rat1) that stimulates termination mechanism reminiscent of bacterial Rho (West et al., 2004; Dengl and Cramer, 2009). Archaeal RNAP is able to terminate transcription independently of RNA secondary structures or factors (Spitalny and Thomm, 2008; Santangelo et al., 2009). The *in vitro* and *in vivo* assays published so far concluded that the underlying mechanism is related to those of eukaryotic RNAP III due to the oligo-dT sequence mediated termination and facilitated transcription re-initiation mechanisms (Braglia et al., 2005; Santangelo and Reeve 2006; Spitalny and Thomm 2008, Santangelo et al., 2009). However, the mechanism through which sequence-directed intrinsic termination in archaea occurs has not been clarified.

D. Aims of this thesis

While the role of the TL in NTP selection and catalysis during the elongation phase is well established, the function of the TL in initiation and terminations phases remains unclear. The scope of this work was to highlight the molecular functions of the TL by employing the recombinant RNAP transcription system of the euryarchaeote *Pyrococcus furiosus* (Werner and Weinzierl, 2002; Naji *et al.,* 2007). This system allowed us to introduce alanine substitutions and deletion mutations and thus to selectively analyze the TL and its key residues. We prepared six *Pfu* RNAP variants with TL mutations (Figure 16). TL mutations analyzed include L83A, H87A, ΔTLtip (A'' A89-N95; *Sc* Rpb1 A1087-K1093), where the mobile tip of the TL is deleted, and ΔTL (A'' T85-T97; *Sc* Rpb1 T1083-T1095), where the size of the deletion was designed based on available structural information in order not to influence enzyme stability, in contrast, to the bacterial TL deletion mutant that are likely to be destabilized (Yuzenkova et al., 2010; Zhang et al., 2010). In yeast, equivalent mutations were lethal (Kaplan et al., 2008; Kaplan et al., 2012). Furthermore, four supplementary mutants were analyzed in order to clarify their participation in NTP/2'dNTP discrimination: TL variants Q80A and Y88S in addition to A' R423A and A' N456A. Our results provide the first analysis of TL function in all three phases of the transcription cycle and unravel previously unknown TL functions.

II) Materials

A. Suppliers

1. Chemicals

Acrylamide, 2x	*Serva, Heidelberg*
Agarose NEEO Ultra	*Roth, Karlsruhe*
Ampicillin	*Sigma-Aldrich, Deisenhofen*
APS (ammonium persulfate)	*Serva, Heidelberg*
Bacto-agar	*Difco Laboratories, Michigan, USA*
Bacto-yeast extract	*Difco Laboratories, Michigan, USA*
Bacto-tryptone	*Difco Laboratories, Michigan, USA*
Bis-Acrylamide, 2x	*Serva, Heidelberg*
Blue Slick	*Serva, Heidelberg*
Bromophenol bleu	*Serva, Heidelberg*
Chloramphenicol	*Roth, Karlsruhe*
Chloroform	*Merck, Darmstadt*
Coomassie Brilliant Blue R250	*Serva, Heidelberg*
DNA ladders	*MBI Fermentas, Vilnius, Lithuania*
dNTPs mix	*MBI Fermentas, Vilnius, Lithuania*
DTT (dithiothreitol)	*Roth, Karlsruhe*
Dynabeads M280 Streptavidin (Dynal)	*Invitrogen, Karlsruhe*
EDTA	*Serva, Heidelberg*
Ethidium Bromide	*Serva, Heidelberg*
Formamid	*Merck, Darmstadt*
Formaldehyd 37%	*Merck, Darmstadt*
Glycerol	*Merck, Darmstadt*
GpU	*Sigma-Aldrich, Seelze*
Guanidine HCl	*Sigma-Aldrich, Seelze*
HEPES	*Roth, Karlsruhe*
Imidazol	*Sigma-Aldrich, Seelze*
IPTG (Isopropyl β-D-1-thiogalactopyranoside)	*Roth, Karlsruhe*
Isotopes ([α- 32 P]-UTP;[γ- 32 P]-ATP)	*Hartmann Analytic, Braunschweig*
Lysozyme	*Boeringer, Mannheim*
Magnesiumchlorid $MgCl_2$	*Merck, Darmstadt*
ß-Mercaptoethanol	*Roth, Karlsruhe*
NTPs	*New England Biolabs, Ipswich, USA*
Phenol/Chloroform/Isoamyl alcohol (25/24/1)	*Roth, Karlsruhe*

Piperidine	*Sigma-Aldrich, Seelze*
Potassium permanganate $KMnO_4$	*Merck, Darmstadt*
Protease Inhibitors Cocktail	*Roche Diagnostics, GmbH, Mannheim*
PMSF (Phenylmethylsulfonylfluorid)	*Roth, Karlsruhe*
Rotiphorese Gel 30 (37,5:1)	*Roth, Karlsruhe*
SDS (*Sodium dodecyl sulfate)*	*Roth, Karlsruhe*
Silver nitrate $AgNO_3$	*Roth, Karlsruhe*
Sodium chloride NaCl	*NORMAPUR, VWR Darmstadt*
TEMED (N,N,N',N'-Tetramethylethylenediamine)	*Roth, Karlsruhe*
Tris	*USB, Cleveland, USA*
Tween-20	*Thermo Fischer Scientific,Waltham, USA*
Urea	*Merck, Darmstadt*
Whatman-Paper	*3MM Maidstone, UK*
Xylencyanol FF	*Roth, Karlsruhe*

2. Enzymes and others proteins

BSA (special quality for molecular biology)	*Roche Diagnostics, GmbH, Mannheim*
Taq DNA-Polymerase	*MBI Fermentas, Vilnius, Lithuania*
Phusion High Fidelity DNA-Polymerase	*Finnzymes, Espoo, Finnland*
T4 DNA ligase	*New England Biolabs, Ipswich, USA*
T4 Polynukleotidkinase (PNK)	*New England Biolabs, Ipswich, USA*
Proteinstandards (# SM0661, #SM0671)	*MBI Fermentas, Vilnius, Lithuania*

3. Column chromatography

FPLC System

ÄKTA purifier 12	*GE healthcare, Chalfont St Giles, UK*

Ionenaustausch Chromatographie

MonoQ 5/50 GL	*GE healthcare, Chalfont St Giles, UK*
HiTrap SP HP 5 ml	*GE healthcare, Chalfont St Giles, UK*

Gelfiltration

HiLoad 16/60 Superdex 75	*GE healthcare, Chalfont St Giles, UK*

B. Genetic materials

1. Strains

Escherichia coli, DH5α	Taylor et al., 1993
Escherichia coli, BL21(DE3)-CodonPlus	*Stratagene, La Jolla, USA*

2. Plasmids

pET151/D-TOPO-A´	Souad Naji, Dissertation, 2006
pET151/D-TOPO-A´´	Souad Naji, Dissertation, 2006
pET151/D-TOPO-A´-R423A	This work
pET151/D-TOPO-A´-N456A	This work
pET151/D-TOPO-A´´-Q80A	This work
pET151/D-TOPO-A´´-L83A	Zeller Mirijam, Dissertation, 2009
pET151/D-TOPO-A´´-H87A	This work
pET151/D-TOPO-A´´-Y88S	This work
pET151/D-TOPO-A´´-ΔTLtip	This work
pET151/D-TOPO-A´´-ΔTL	This work
pUC19-*gdh*C20	Spitalny Patrizia, Dissertation, 2008
pUC19-*gdh*C9	Spitalny Patrizia, Dissertation, 2008
pUC17-TFS	Zeller Mirijam, Dissertation, 2009
pUC17-TFS-AA	This work

3. Primers for protein mutagenesis

A´´-F	AATATCCCGGAGCTAACTACGTGATAGACC
A´´-R	AGAGATTATGTGGTTTTGAACTTCCATTAG
A´´Q80A-F	TCCAATAGACTGGGCAGCAACAGTACCAAC
A´´Q80A-R	GAACCTTCAACCGCGATGACTCTCAACACT
A´´H87A-F	AGTCATCTGGGTTGAAGGTTCTCCAATAGA
A´´H87A-R	CTCAACACTTTCGCCTATGCAGGTGTTGCT
A´´Y88S-F	GAGAGTCATCTGGGTTGAAGGTTCTCCAAT
A´´Y88S-R	AACACTTTCCACTCTGCAGGTGTTGCTGAA
ΔTLtip-F	AACGTTACTTTAGGTTTGCCAAGA
ΔTLtip-R	ATAGTGGAAAGTGTTGAGAGTCAT
ΔTL-F	TTGGAGAGTCATCTGGGTTGAAGGTTCTCC
ΔTL-R	ACTTTAGGTTTGCCAAGAATTATAGAAATC
A´-F	CACCATGAAAAAAGTTATTGGAAGTATT

A´-R	CTATCACACCTTCGCCTTGTTATTCC
A´R423A-F	TATATCTCCATCCATTAGATGCCTCTCAAC
A´R423A-R	GTGCTGTTTAACGCACAGCCATCATTACAC
A´N456A-F	TACAGCCAAGTTGAGACGGAATGTTCTGTA
A´N456A-R	TGTCCACCATACGCTGCTGATTTTGATGGA
TFS-AA-F	TTGAAGTTCCCACCACCATGCTTCATTATG
TFS-AA-R	ACTAGGGCAGGAGCGGCGCCAAGTACAATA

4. Primers for promoter mutagenesis and oligonucleotides

M13-F	GCCAGGGTTTTCCCAGTCACGA
M13-R	GAGCGGATAACAATTTCACACAGG
MiniBub-F	CCGAAAGCTTTATATAGGCTATTGCCCTTTAATGTATCGTTAATGAGGTAATTTGGA
MiniBub-R	TCCAAATTACCTCATTAACGATACATTAAAGGGCAATAGCCTATATAAAGCTTTCGG
MiniStart-F	TTATATAGGCTATTGCCCAAAAATGTATTACTAATGAGGTAATTTGGAGCATATGGGG
MiniStart-R	CCCCATATGCTCCAAATTACCTCATTAGTAATACATTTTTGGGCAATAGCCTATATAA
EMSA-F	GTTTACCGAAAGCTTTATATAGGCTATTGCCCAAAAATGTATCGCCAATCACCTAATTTGG
EMSA-R	CCAAATTAGGTGATTGGCGATACATTTTTGGGCAATAGCCTATATAAAGCTTTCGGTAAAC
EC(U)-T	CCACCCTTACCTCCACCATATGGGAGATCCATTACAGCAGCCAAGCTCAAGTACTTACGCCTGGTCATTACTAGTACTGCCGG
EC(U)-NT	CCGGCAGTACTAGTAATGACCAGGCGTAAGTACTTGAGCTTGGCTGCTGTAATGGATCTCCCATATGGTGGAGGTAAGGGTGG
EC(A)-T	CCACCCTTACCTCCACCATATGGGAGATCCATTACAGCAGCCAAGCTCAAGTACGATCGCCTGGTCATTACTAGTACTGCCGG
EC(A)-NT	CCGGCAGTACTAGTAATGACCAGGCGATCGTACTTGAGCTTGGCTGCTGTAATGGATCTCCCATATGGTGGAGGTAAGGGTGG
ECTerm	CCACCCTTACCTCCACCATATGGGAGAAAAAAACAGCAGCCAAGCTCAAGTACTTACGCCTGGTCATTACTAGTACTGCCGG
ECTerm-NT	CCGGCAGTACTAGTAATGACCAGGCGTAAGTACTTGAGCTTGGCTGCTGTTTTTTTCTCCCATATGGTGGAGGTAAGGGTGG
RNA-EC	AUUUAGACCAGGCG
RNA-MEC(C)	AUUUAGACCAGGCGC
RNA-MEC(G)	AUUUAGACCAGGCGG

III) Methods

A. Cloning

Expression vectors containing *P. furiosus* RNAP subunits were provided by Souad Naji (2007): RpoL and RpoN were cloned into pET30 without His-tag, all other subunits were cloned in pET151/D-TOPO and have an N-terminal His_6-tag.

1. Gel purification of primers

10 µl of 100 µM primer solution, diluted in Tris-HCl pH 8, was mixed with 5 µl of 3xLoading buffer (98% Formamide, 10 mM EDTA, 0.1% bromophenol blue, 0.1% Xylene cyanol FF),denatured for 3 min at 95°C, and loaded to 12-15% PA / 6 M urea gel (750 V, 25 mA, 40 min). To visualize the nucleic acids, the gel was set on plastic wrap, placed on TLC Silica gel 60 F_{254} aluminium sheet (MERCK) and irradiated with UV light (UV-shadowing). The primer band was cut out and transferred into an Eppendorf tube. The primer DNA was eluted with 360 µl of TE´ buffer (10 mM Tris-HCl pH 8, 0.1 mM EDTA) and 40 µl of 3 M sodium acetate pH 5.2, overnight at 37°C on a shaker. The supernatant was transferred to a new Eppendorf tube and the DNA was precipitated with 20 µg of glycogen and 800 µl of 99% ethanol. After 30 min incubation at -80°C, the suspension was centrifuged for 30 minutes at 13 000 rpm. The DNA pellet was then washed with 700 µl of 70% ethanol, dried, and resuspended in 50 µl of 10 mM Tris-HCl pH 8. The DNA concentration of the primer solution was determined by measuring the absorption at 260 nm using the NanoDrop. The specific extinction coefficient of each oligo was determined using the website http://www.basic.northwestern.edu/biotools/oligocalc.html.

2. Sequence-specific mutagenesis of plasmids

The mutagenesis primers were designed with a phosphate group attached to the 5´end. The reactions was prepared in a total volume of 20 µl containing 1x HFbuffer, 0.2 µM dNTPs, 5 pmol forward and reverse primers, 1 ng of plasmid and 1.25 U of Phusion-Polymerase (Finnzymes). The PCR reaction was performed in following conditions: 1x 30 sec 98°C, 24x (10 sec 98°C, 30 sec 61-70°C, 15 sec/kb 72°C), 1x 5 min 72°C, 4°C hold. The PCR products were purified using the PCR purification kit (Qiagen).

3. Ligation of linear plasmid

The reactions was performed in a total volume of 50 µl containing 1x T4 Ligase buffer, 25 ng of PCR products, and 1 U of T4 DNA Ligase (NEB). The mixture was incubated for 1 h at 20°C, and 10 min at 65°C.

4. Transformation of *E. coli*

The *E. coli* strains DH5α and BL21(DE3)-CodonPlus were used as a host for standard sub-cloning and for the recombinant protein overexpression respectively. According to their transformation efficiency, 25-50 µl of cells per reaction were used. The cells were incubated for 15 min on ice with 1-3 µl of ligation mixture. The cells were subjected to water heat shock for 45-50 sec at 42°C, and incubated again on ice for 2 min. After the addition of 175-450 µl SOC medium (2% Bacto-tryptone, 0.5% Yeast extract, 10 mM NaCl, 2.5 mM KCl, 10 mM $MgCl_2$, 10 mM $MgSO_4$, 20 mM glucose), the cells were regenerated for 1 h at 37°C, on the shaker. 100 µl of suspension was plated on LB plates with the appropriate antibiotics, and grown overnight at 37°C.

B. Protein overexpression and purification

TBP was kindly provided by Winfried Hausner, TFB by Mirijam Zeller and Robert Reichelt, and TFE by Mirijam Zeller. The RNAP subunits were purified as described previously (Goede et al., 2006) by Wolfgang Forster (B, D, E´, F, H, L and P).

1. Protein overexpression

Four RNAP subunits (A ', A'', N and K) and TFS were expressed in *E. coli* BL21(DE3) CodonPlus. Expression strains were grown in the presence of 100 mg/l Ampicillin (A ', A'', K) or 50 mg/l Kanamycin (L, N) and 50 mg/l Chloramphenicol in LB medium (10 g/l NaCl, 10 g/l Bacto-tryptone, 5 g/l Yeast extract) to exponential phase (OD_{600} between 0.6 and 0.8) and then incubated on ice for 30 min. The recombinant protein expression was induced by addition of IPTG (final concentration 0.5 mM), and incubation at 18°C for 16-20 hours with shacking. The cells were harvested by centrifugation (10 min, 6000 rpm, 4°C, JA 10 rotor) and stored at -20°C.

2. Purification of recombinant *P. furiosus* RNAP subunit

a. Purification of the subunits from inclusion bodies (A', A'', and K)

Approximately 4 g of cell pellet were resuspended in 25 ml of "completed" Tris buffer (50 mM Tris-HCl pH 8, 5 mM 2-mercaptoethanol, 0.5 mM PMSF, 1 protease inhibitor cocktail tablet). A spatula (approx. 0.1 g) of Lysozyme was added and the suspension was incubated for 1 h on ice. After 3-4x 30 sec sonication (output: 30%), the sample was centrifuged (15 min 18 000 rpm, 4°C, SS34 rotor). The pellet was then resuspended in 25 ml IB buffer I (20 mM Tris-HCl pH 8, 2 M Urea, 500 mM NaCl, 0.1% Tween, 5 mM 2-mercaptoethanol, 0.5 mM PMSF) using a Potter (Wheaton), and sonicated and centrifuged as above (the supernatant obtained is called "supernatant 2"). The solubilization of the inclusion bodies (resuspension, sonication and centrifugation) was repeated again with the IB buffer I and buffer II (20 mM Tris-HCl pH 8, 500 mM NaCl, 0.1% Tween, 5 mM 2-mercaptoethanol, 0.5 mM PMSF), and the pellet was stored at -20°C.

α. Purification of A' and K subunits

The inclusion bodies were thawed on ice in 25 ml of binding buffer (20 mM Tris-HCl pH 8, 6 M Guanidine-HCl, 500 mM NaCl, 20 mM Imidazole, 5 mM 2-mercaptoethanol, 0.5 mM PMSF) and incubated for 1 h at RT with repeated vortexing until the complete solubilization. After centrifugation (20 min, 18 000 rpm, 4°C, SS34 rotor), the supernatant was filtered (0.23 µm). Applied slowly (0.5 ml/min) to the NiNTA column (HiTrap Chelating HP 5 ml) via the sample pump. The loaded sample was washed with 3 Column Volume (CV) Binding buffer and 3 CV Washing buffer (20 mM Tris-HCl pH 8, 6 M Urea, 500 mM NaCl, 20 mM Imidazole, 5 mM 2-mercaptoethanol, 0.5 mM PMSF). The urea was removed by a linear gradient of 10 CV Washing buffer to Refolding buffer (20 mM Tris-HCl pH 8, 500 mM NaCl, 20 mM Imidazole, 5 mM 2-mercaptoethanol, 0.5 mM PMSF), and another 10 CV Refolding buffer. The protein was then eluted in a 4 step gradient of: 2 CV Refolding buffer, 3 CV 5% Elution buffer (20 mM Tris-HCl pH 8, 500 mM NaCl, 500 mM Imidazole, 5 mM 2-mercaptoethanol, 0.5 mM PMSF), 5 CV 40% Elution buffer (proteins should elute in this step) and 5 CV 100% Elution buffer.

β. Purification of A'' subunit

The "supernatant 2" was filtered (0.45 µm) and applied slowly (0.5 ml/min) to the NiNTA column (HiTrap Chelating HP 5 ml) via the sample pump. The urea was removed by a linear gradient of 10 CV IB buffer I to Refolding buffer (20 mM Tris-HCl pH 8, 500 mM NaCl, 20 mM Imidazole, 5 mM 2-mercaptoethanol, 0.5 mM PMSF), and another 10 CV Refolding buffer. The protein was then eluted by 3 steps gradient of: 2 CV Refolding buffer, 15 CV linear gradients to

Elution buffer (20 mM Tris-HCl pH 8, 500 mM NaCl, 500 mM Imidazole, 5 mM 2-mercaptoethanol, 0.5 mM PMSF), 5 CV Elution buffer.

b. Purification of soluble subunit (N)

The cell pellet was resuspended in 20 ml of Binding buffer (20 mM Tris-HCl pH 8, 10 mM NaCl, 10% glycerol, 5 mM 2-mercaptoethanol, 0.5 mM PMSF, 1 protease inhibitor cocktail tablet). A spatula of Lysozyme was added and the suspension was incubated for 30 min on ice. After 3-4x 60 sec sonication (output: 30%), the sample was centrifuged (15 min 18 000 rpm, 4°C, SS34 rotor). The supernatant was incubated for 20 min at 90°C and then centrifuged (15 min 18 000 rpm, 4°C, SS34 rotor). The supernatant was filtered (0.45 µm) and applied slowly (0.5 ml/min) to the 1 ml MonoQ 5/50 GL column via the sample pump. The protein was eluted by a linear gradient of 20 CV Binding buffer to Elution buffer (20 mM Tris-HCl pH 8, 1 M NaCl, 10% glycerol, 5 mM 2-mercaptoethanol, 0.5 mM PMSF).

3. Purification of recombinant TFS

The cell pellet was resuspended in 20 ml of completed Tris buffer (20 mM Tris-HCl pH 8, 50 mM NaCl, 1 mM DTT). A spatula of Lysozyme was added and the suspension was incubated for 30 min on ice. After 4-5x 60 sec sonication (output: 35%), the sample was centrifuged (1 h, 20 000 rpm, 4°C, SS34 rotor). The supernatant is incubated for 15 min at 95°C and was centrifuged (20 min 20 000 rpm, 4°C, SS34 rotor). The supernatant was filtered (0.45 µm) and concentrated to 7 ml by ultracentrifugation (Millipore, MWCO 10.000). In two runs, the protein solution is separated through a Hiload 16/60 Superdex 75 column in gel filtration buffer (20 mM Tris-HCl pH 8, 1 M NaCl). TFS eluted around 74 to 80 ml (dyed in reddish fractions). The fractions were then pooled with Dilution buffer (20 mM Tris-HCl pH 8, 50 mM NaCl). The protein solution was applied to the 1 ml MonoQ 5/50 GL column and eluted by a linear gradient of 15 CV Dilution buffer to 1M NaCl, at a flow rate of 1 ml/min. TFS must elute after 10 ml.

4. Reconstitution of RNAP from *P. furiosus*

A mixture of RNAP subunits, each 2250 pmol, was loaded into a Dialysis cassette (0.5-3 ml, MWCO 3 500, Thermo Fisher Scientific) and denatured for 15 min in ice-cold TB-6 buffer (6 M urea, 40 mM HEPES-Na pH 7.3, 250 mM NaCl, 2.5 mM MgCl 2, 0.1 mM EDTA, 0.1 mM $ZnSO_4$, 10% Gycerol, 5 mM 2-mercaptoethanol, 0.5 mM PMSF) at moderate stirring. The dialysis cassette was then incubated in ice-cold TB-3 buffer (TB-6 with 3 M urea) for 15 min, and finally in ice-cold TB-0 buffer (TB-6 at without urea) for 1 hour. The reconstituted RNAP was incubated for 10 min at 70°C and precipitated proteins were removed by centrifugation (10 min, 13 000 rpm, room

temperature (RT)). The supernatant was slowly applied to the TB-0 equilibrated Superdex-200-10/300-GL column with a syringe. The polymerase eluted in a range between 8 and 10 ml, at a flow of 0.2 ml/min. The concentration of reconstituted RNAPs was compared by 4-20% gradient SDS gel and Silver staining.

C. DNA template preparation

1. Standard promoter-dependent transcription templates

gdh-C20, the standard transcription template, was cloned into pUC19 and amplified by PCR with M13 primers (297 bp). The reactions was prepared in a total volume of 50 µl containing 0.2 mM dNTPs, 2 mM $MgCl_2$, 15 ng of plasmid *gdh*-C20, 0.3 mM of each M13-F and M13-R primers, 2 U *Taq* DNA polymerase in 1x *Taq* buffer. The PCR reaction was performed in following conditions: 1X 3 min 95°C, 25X (5 sec 95°C, 30 sec 60°C, 30 sec 72°C), 5 min 72°C, 4°C . After purification of the PCR products with the PCR Purification Kit, the purity was verified on agarose gel (length 297 bp). The template *gdh*-C9 was obtained with the same protocol described above.

2. Pre-opened templates

Pre-opened templates contained mismatches at positions -9, -10 and -11 ("open upstream", O^U) or -2, -1 and +1 ("open start", O^S) relative to the transcription start site in *gdh-C20* were prepared. Mutations were introduced by fusion PCR using the MiniBub-F and MiniBub-R (O^U) or MiniStart-F and MiniStart-R (O^S) primers. Mutant and original DNA sequences were PCR amplified using one phosphorylated and one unphosphorylated primer. The PCR products were purified using the PCR Purification Kit and quantified with the NanoDrop. The phosphorylated strand was specifically digested with λ-exonuclease (Fermentas). The reactions was prepared in a total volume of 100 µl containing 1x reaction buffer, 40 nmol PCR product, 20 U of exonuclease and incubated for 1 h at 37 °C. Wild-type (WT) template strand and mutant nontemplate strand were mixed in equimolar amounts, re-annealed in 10 mM Tris-HCl pH 8 (5 min 95°C, 15 min 45°C, 15 min 40°C) and purified from agarose gels using the Qiagen gel extraction kit. For direct comparison, closed templates (C) were prepared in the same way, except that both template and nontemplate strand were derived from WT sequence. The templates were stored in 4°C.

3. $KMnO_4$-footprint template

The reactions was prepared in a total volume of 50 µl containing 0.2 mM dNTPs, 2 mM $MgCl_2$, 100 ng of plasmid *gdh*-C20, 0.5 mM of M13-F-bio and M13-R primers, 1 U *Taq* DNA polymerase in 1x *Taq* buffer. The PCR reaction was performed in following conditions: 1X 90 sec 95°C, 10X (5 sec

95°C, 30 sec 60°C, 1 min 72°C), 25X (5 sec 95°C, 30 sec 60°C, 1 min (+10 sec/cycle) 72°C), 4°C. Coupling of 100 µl PCR products with 50 µl Dynabeads M-280 Streptavidin (Invitrogen) was carried out by i) washing the beads 3 times with buffer A (10 mM Tris-HCl pH 7.5, 1 mM EDTA, 100 mM NaCl), ii) resuspension of the beads in buffer B (50 mM Tris-HCl pH 7.5, 1 mM EDTA, 100 mM NaCl), the addition of 50 µl PCR products, shaking for 30 min at RT, washing with 90 µl buffer B, addition of 50 µl PCR products, shaking for 30 min at RT, discarding the supernatant, iii) resuspension of the beads in 90 µl buffer C (10 mM Tris-HCl pH 7.5, 1 mM EDTA, 1 M NaCl), incubation for 1 min at RT, the resuspension in buffer C was repeated, and finally resuspension in 90 µl TE ' buffer. The templates were stored in 4°C.

4. Radioactively 5´end labeled EMSA template

The reactions was prepared in a total volume of 50 µl containing 0.2 mM dNTPs, 2 mM $MgCl_2$, 1 µl of genomic DNA, 0.5 mM of each EMSA-F and EMSA-R primers, 1 U *Taq* DNA polymerase in 1x *Taq* buffer. The PCR reaction was performed in following conditions: 1X 3 min 95°C, 30X (30 sec 95°C, 30 sec 58°C, 30 sec 72°C), 10 min 72°C, 4°C. After purification of the PCR products with the PCR Purification Kit, the purity was verified on agarose gel (length 97 bp) and quantified with the NanoDrop. The labeling reaction was prepared in a total volume of 50 µl containing at 12.7 pmol DNA Template, 20 U T4 PNK and 1.5 µl [γ-^{32}P] ATP with 0.55 MBq (220TBq / mmol) in 1 × PNK buffer. The mixture was then incubated for 30 min at 37°C, and 2 µl 0.2 M EDTA was added. The sample was mixed with 100 µl PCI and centrifuged for 10 min. The supernatant was transferred into a new Eppendorf tube with 20 µl of 10 M $LiCl_2$, 2 µg of Glycogen and 150 µl of ice cold 99% ethanol. Precipitation was carried out overnight at -20 ° C. The sample was centrifuged for 30 min, and the pellet was washed with 2 volumes of ice-cold 70% ethanol. After resuspension in 50 µl of 10 mM Tris-HCl pH 8, the template was purified using the Nucleotide Removal Kit (Qiagen) and resuspended in 50 µl of 10 mM Tris-HCl pH 8.

D. *In vitro* assays

1. *In vitro* promoter-dependent transcription assays

Transcription reaction contained 10 nM *gdh-C20* template, 100 nM TBP, 70 nM TFB and RNAP in a final volume of 25 µl were incubated in transcription buffer ((40 mM HEPES-Na pH 7.3, 250 mM NaCl, 2.5 mM $MgCl_2$, 0.1 mM, EDTA, 0.1 mM $ZnSO_4$; 10% Glycerol) containing NTPs (440 µM ATP, 440 µM GTP, 440 µM CTP, 2.7 µM UTP and [α-^{32}P]UTP at 0.15 MBq (110 tBq/mmol) for 10 min at 70°C. Transcription initiation assays were carried out with the template *gdh*-C9, where the first cytidine occurs at position +10, 40 µM priming RNA (Sigma), [α-^{32}P] complement NTP at 0.22 MBq (110 tBq/mmol). RNA was extracted with PCI and analyzed by electrophoresis in 8% (run off,

113 bp) or 28% (trinucleotides and short transcripts) Polyacrylamide (PA)/Urea gel. The transcription products were visualized with a FLA-5000 scanner (Fujifilm).

2. Band shift assays (EMSA)

Transcription initiation complex was assembled in transcription buffer (described above) containing 100 nM 5´end ^{32}P labeled template, 100 nM TBP, 70 nM TFB and RNAP in a final volume of 25 µl. After 10 min incubation at 37° C (70°C for endogenous RNAP(endoRNAP)), the reaction was cooled to 4 °C. 2.5 g of Heparin was added as a competitor and the samples were incubated at 37 °C (70 °C for endoRNAP) for additional 15 min. 17 µl of the reaction was loaded on a 4.5% PA / 2% Glycerol / TBE gel. TBE running buffer contained 2% of Glycerol.

3. $KMnO_4$-footprint assays

Transcription initiation complex was assembled in TB-0 buffer with 3 µl of immobilized template, 100 nM TBP, 70 nM TFB and RNAP in a final volume 25 µl, at 70°C for 5 min. 2.5 µl of 250 mM $KMnO_4$ (RT) were added, and the sample was incubated for at 70°C for additional 2 min. 1.5 µl of 2-mercaptoethanol (RT) and 20 µl of preheated (70 ° C) Stop buffer (125 mM EDTA, 1.25% SDS) were added. The sample was then transferred to ice. The supernatant was removed using the Magnetic particle separator (MPS) and the sample was resuspended with 18 µl water. 2 µl of Piperidine was added and the sample was incubated for 30 min at 90 ° C. The supernatant was transferred to 40 µl PCI using the MPS and centrifuged for 8 min. The supernatant (approx. 17 µl) was transferred to a new Eppendorf tube containing 30 µl of water. After the addition of 5 µl 3 M Sodium Acetate pH 5.2, 20 µg Glycogen and 125 µl of ice-cold 99% ethanol, the sample was incubated for 30 min at -80 ° C. After 30 min of centrifugation, the pellet was washed with 500 µl of ice-cold 70% ethanol and centrifuged for 10 min. The supernatant was discarded, and the pellet was dried by SpeedVac (2 min, 45 ° C) and resuspended in 10 µl 10 mM Tris-HCl pH 8. After addition of 5 µl 3x Protein loading buffer, the sample was denatured for 3 min at 95°C. The separation was carried out on a 6% PA/6 M Urea gel (Biorad Sequi-Gen®).

4. Bead-based RNA extension and TFS induced cleavage assays

The bead-based elongation complexes (ECs) containing complete complementary scaffolds were assembled and immobilized on Dynabeads M-280 Streptavidin essentially as described (Kireeva et al., 2000; Grünberg et al., 2010). The RNA was 5′ end labeled with [γ-^{32}P] ATP. For ECs assembly, RNAP was incubated with a hybrid of the DNA template strand annealed to the RNA in Elongation Buffer (EB: 20 mM HEPES [pH 7.6], 100 mM CH_3CO_2K [pH 7.6], 5 mM $Mg(CH_3COO)_2$, 0.1

mg BSA) for 5 min, then with the 3´-end-biotinylated non-template DNA strand for 5 min, and then with 2.5 g Heparin for 5min at 20°C. Beads were subsequently washed 3 times with EB pre-heated at 50°C. Beads were resuspended in EB. For 1 nt RNA extension assays, 100 µM of NTP (ATP, 2'dATP or UTP) were added, the mixture was incubated at 50°C. For the elongation and termination assays, 30 µM of NTPs mix were added. The mixture was incubated at 70°C for elongation assays and 70°C, 80 °C or 90°C for termination assays. For the TFS induced cleavage assays, 150 nM TFS were added and the mixture was incubated at 70°C. The reactions were stopped by transferring the sample into cold loading buffer containing Formamide. Samples were heated to 95°C and loaded on a 28% polyacrylamid gel containing 7 M Urea. For elongation assays and termination assays, the samples were loaded on a 15% polyacrylamid gel containing 7 M Urea. The radioactively 5′-labeled RNA products were visualized with a FLA-5000 scanner (Fujifilm). Gel bands were quantified using Aida Image analyzer (Raytest).

5. Bead-based RNA intrinsic cleavage assays.

For the intrinsic cleavage assays, the ECs was assembled in Cleavage Buffer (CB: 20 mM HEPES [pH 9], 100 mM CH_3CO_2K [pH 9], 10 mM $Mg(CH_3COO)_2$, 0.1 mg BSA), and then, incubated at 70°C. The reactions were stopped by transferring the sample into the ice. Beads were washed one time with CB and resuspended with CB with loading buffer containing Formamide. Samples were heated to 95°C and loaded on a 28% polyacrylamid gel containing 7 M Urea. The radioactively 5′-labeled RNA products were visualized with a FLA-5000 scanner (Fujifilm). Gel bands were quantified using Aida Image analyzer (Raytest).

6. Data analysis

The nucleotide (mis)incorporation rates obtained for various substrates and concentrations were fitted to the Michaelis-Menten equation; $k = k_{pol} \times [NTP] / (K_M + [NTP])$, were k_{pol} is the maximum NTP incorporation rate of the enzyme, [NTP] is the substrate concentration and K_M is the Michaelis constant. Kinetic data (Nucleotide additions, RNA hydrolysis) were fitted to a single exponential equation using non-linear regression in Sigmaplot (Systat Software Inc.).

IV) Results

A. TL mutants design

The high-resolution structure of yeast RNAP II and bacterial RNAP revealed that a highly conserved mobile element, the TL, is able to close the active site and stabilizes the NTP in the insertion site, allowing phosphodiester bond synthesis (Wang et al., 2006; Vassylyev et al., 2007b). In this conformation, the TL directly contacts the base and β-phosphate of the NTP through the homologues residue A'' L83 and A'' H87, respectively (Wang et al., 2006). Sequence alignment of the TL reveals a length variation between bacterial, eukaryotic and archaeal RNAPs. Indeed, bacterial TL contains insertions of one residue (*T. Thermophiles*) to 188 aa (SI3 in *E. coli*) insertions near the tip region suggesting a possible functional diversification of the TL in bacteria (Figure 16). In agreement with that, recent biochemical studies on bacterial and yeast systems revealed a different level of contribution of these specific residues in catalysis and fidelity during elongation (Kaplan et al., 2008; Yuzenkova et al., 2010; Zhang et al., 2010).

Figure 16. Sequence conservation of the TL in the three domains of life. Invariant (*blue*) and conserved (*yellow*) residues are shown in an alignment of TL sequences from human RNAP II (*Homo sapiens*), yeast RNAP II (*S. cerevisiae*), archaeal RNAP (*P. furiosus)* and bacterial RNAP (*T. thermophilus* and *E. coli*). Amino acid substitutions (all by alanine, except serine for Y88) introduced for this study in subunit A'' and deletion mutations are indicated by red square and red lines, respectively. The black triangle indicates the position of the insertion site of SI3 (188 aa) in the TL of *E. coli* RNAP. The α-helical and loop segments in the closed and open states of the TL are illustrated by thick and thin magenta lines, respectively.

Whereas biochemical analyses on eukaryotic RNAP II mutants are restricted to viable yeast strains, the closely related recombinant RNAP transcription system of *P. furiosus* allowed us to introduce alanine substitution and deletion mutations and thus to selectively analyze key residues of the TL. Six *Pfu* RNAP variants with TL mutations were prepared during the initial step of this work (Figure 16). The TL mutations analyzed include L83A, H87A, ΔTLtip (A'' A89-N95; *Sc* Rpb1 A1087-K1093), where the mobile tip of the TL is deleted, and ΔTL (A'' T85-T97; *Sc* Rpb1 T1083-T1095), where the size of the deletion was designed based on the structural information in order not to influence enzyme stability, in contrast to the bacterial TL deletion mutant, which are likely to be destabilizing (Yuzenkova et al, 2010; Zhang et al, 2010). Furthermore, four supplementary mutants were analyzed in order to clarify their participation

in NTP/2'dNTP discrimination: TL variants Q80A and Y88S (Figure 16) in addition to A' mutants R423A and N456A, that the homologues to residues in yeast that interact with the incoming NTP 2'-oH group (Cheung et al., 2010).

B. Reconstitution of TL mutants RNAPs and binding on the promoter

The mutant A'' with a His_6-tag at the N-terminus was purified as described in III) Methods. *In vitro* reconstitution of mutant enzymes did not show any significant differences with the WT enzyme suggesting that mutation within the TL does not affect the interaction between the RNAP subunits (Figure 17A).

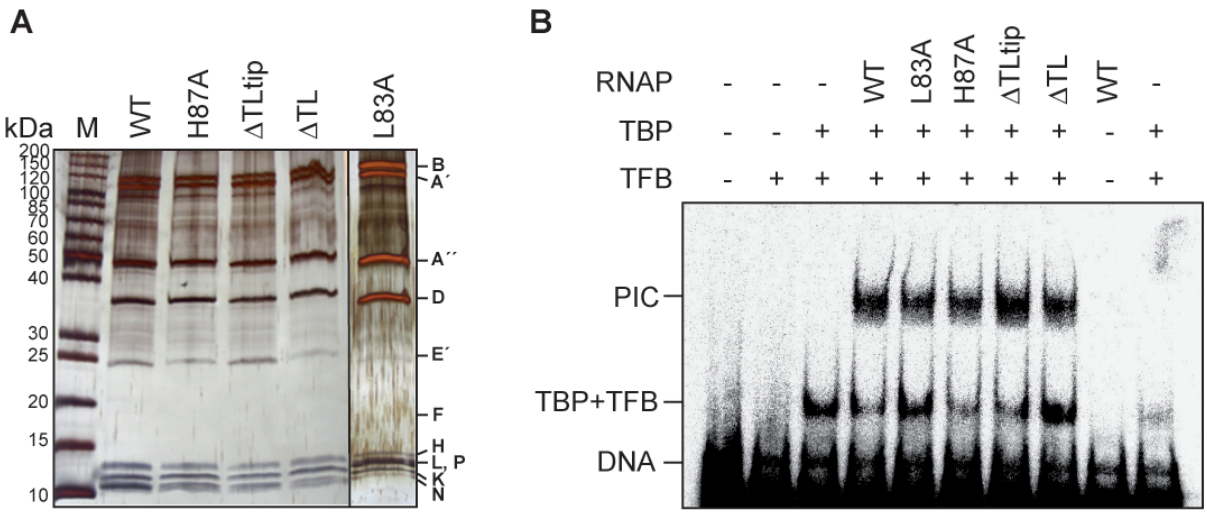

Figure 17. Reconstitution of mutant RNAPs. (**A**) Siver-staining of 4-20% SDS gradient gel with reconstituted WT and mutant RNAPs. (**B**) Band shift assay with 5´-end radioactively labeled template containing *gdh*-C20 promoter (Figure 18). PIC corresponds to the complex of TBP, TFB and RNAP.

In order to verify whether all TL mutant RNAPs are capable of being recruited to the DNA/TBP/TFB complex to form an initiation complex, electrophoretic mobility shift assays were performed (Figure 17B). Here, the initiation complex is composed of TBP, TFB and RNAP in the presence of radioactively 5'-end labeled DNA template containing the *gdh*-C20 promoter sequence. Native gel electrophoresis revealed that the mutant RNAPs did not show any significant defect in initiation complex formation, suggesting that the mutant RNAPs are capable to form a stable pre-initiation complex (PIC).

C. Function of the TL in transcription initiation

To analyze whether the TL RNAP mutants have an overall defect in promoter-dependent transcription, we subjected them to a promoter-specific transcription assay in which we employed the strong *P. furiosus* glutamate dehydrogenase (*gdh*-C20) promoter in the

presence of general transcription initiation factors (GTIFs) (Figure 18A). The 113 nt run off RNA products were synthesized by the WT RNAP and to a lesser extend by the mutant L83A, demonstrating that A'' L83 contributes, but is not essential for transcription (Figure 18B). In contrast, both TL deletion RNAP mutants as well as H87A displayed no transcription activity. These results show that the TL is critical for archaeal RNAP function, as it is the case for bacterial and eukaryotic RNAPs (Kaplan et al, 2008; Yuzenkova et al, 2010; Zhang et al, 2010).

To investigate whether the initiation phase of transcription requires the TL, we carried out an abortive transcription assay, using closed and pre-opened versions of the *gdh* promoter template in combination with a dinucleotide GpU priming RNA (RNA2) that was complementary to the +1 and +2 positions in the template DNA strand (Appendix Figure S1). By addition of the radioactive nucleotide [α-^{32}P]-UTP, we tested whether the mutant RNAPs were able to elongate the GpU RNA by one nucleotide. As shown in Appendix Figure S1, the L83A variant synthesized the trinucleotide product, but was less active than the WT enzyme. The TL deletion mutants and H87A were not active in this assay.

The initiation defect was apparently not due to a defect in DNA melting. Similar results were obtained on DNA templates pre-opened by three non-complementary nucleotides around the transcription start site +1, or around the putative start site of DNA melting (position -10) (Appendix Figure S1), indicating that promoter opening was not affected (Kostrewa *et al*, 2009). Transcription activity was also not increased by addition of TFE, which stimulates promoter opening and abortive transcription (Naji et al, 2007). Potassium permanganate footprint assays showed that mutations within the TL did not affect open complex formation, while TFE retains the stimulatory effect on all RNAP mutants during promoter opening (Figure 18C and D). Consequently, the TL is not required for stable open complex formation or TFE binding during initiation complex formation, but is essential for the synthesis of the first nucleotide bonds.

We next analyzed the capacity of WT RNAP and TL variants in the presence or absence of GTIFs to extend a priming RNA of 3 to 9 nt in length by one nucleotide (Figure 19A). These initiation scaffolds mimicked progressive steps in transcription initiation: RNAs up to 5 nt represent the phase of abortive initiation, whereas RNAs between 6 and 8 nt represent the phase of hybrid completion (Spitalny and Thomm, 2003; Sainsbury et al, 2013). At an RNA length of 9 to 11 nt, the bubble collapses and is accompanied by the beginning downstream migration of both the downstream (RNA9) and upstream (RNA10) edges of the RNAP (Spitalny and Thomm, 2003).

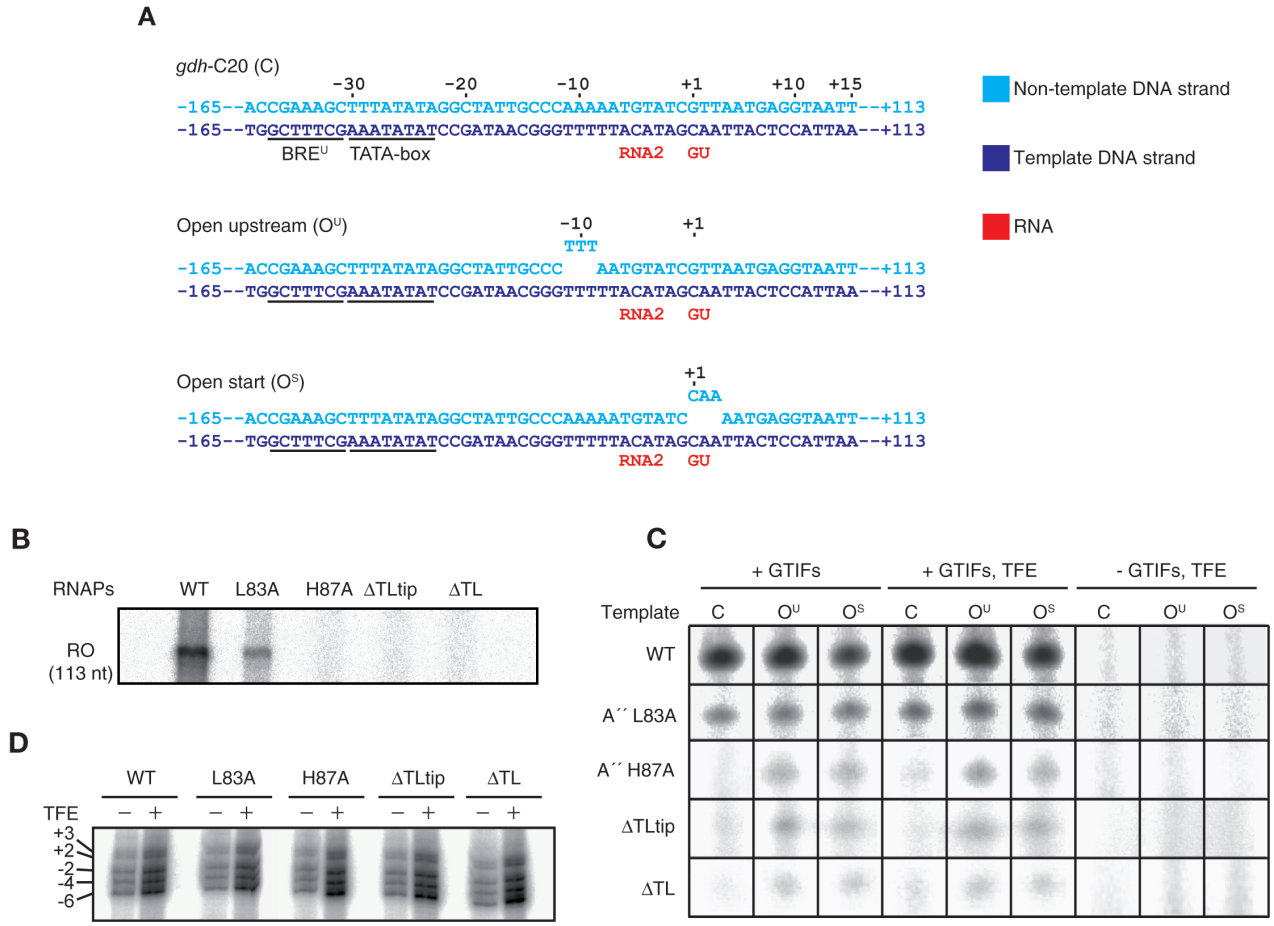

Figure 18. The TL is required for transcription initiation. (**A**) Schematic representation of closed and pre-opened templates containing the strong *gdh* promoter of *P. furiosus*. The BRE^{U} and the TATA-box are underlined. The first C residue in the transcribed region of the *gdh*-C20 template occurs at position +21 (Spitalny and Thomm, 2003). The templates "open upstream" and "open start" comprised a mismatch region from position –11 to -9 and +1 to +3 relative to the transcription start site, respectively (Kostrewa et al., 2009). (**B**) Run off (RO) *in vitro* transcription reactions were performed on the *gdh*-C20 template. (**C**) *In vitro* transcription reactions were performed on closed (C), "open upstream" (O^{U}) and "open start" (O^{S}) *gdh*-C20 templates in the presence (+) or absence (-) of GTIFs and TFE. The trinucleotide product after extension of RNA2 by 1 nt was detected. (**D**) Permanganate footprinting analysis was performed on the *gdh*-C20 promoter in the presence or absence of TFE.

The WT enzyme efficiently added one nucleotide in abortive initiation scaffolds, with the formation of a 5 nt RNA was most pronounced and the activity was fully dependent on GTIFs. On scaffolds mimicking the hybrid completion phase of initiation, NTP incorporation activity successively decreased with the length of the RNA, until at an RNA length of 8 nt, no nucleotide addition activity was observed. This suggests that the two successive RNA synthesis steps from RNA6 until RNA8 are necessary to maintain transcriptional activity of the initiating complex, presumably due to TFB rearrangements within the RNAP cleft that can only be achieved through a growing RNA chain, but not through a pre-synthesized RNA that binds the initiation complex (Figure 19C). In contrast, extension activity with RNA8 and RNA9 is restored and also independent of GTIFs (Figure 19B), indicating that RNAs longer than 6 nt begin to

displace TFB from the pre-initiation complexes. The TL mutant enzymes, excluding L83A, exhibited almost no activity on early abortive transcription scaffolds, whereas all mutant RNAPs were able to add the complementary nucleotide on the RNA8 and RNA9 assemblies, independently from the presence of GTIFs. These results show that the TL including the mobile tip element and A'' H87 are strictly required for transcription initiation.

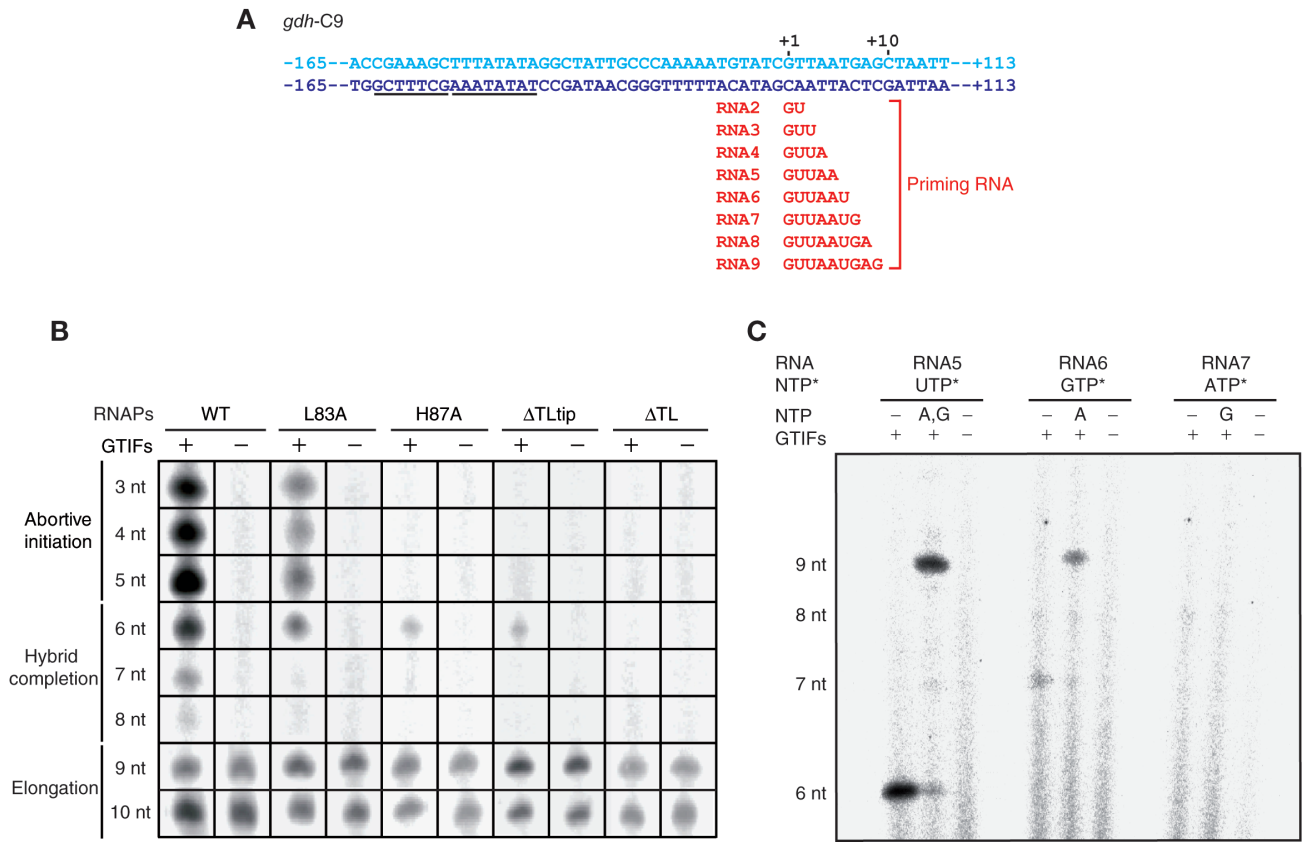

Figure 19. Two successive RNA synthesis steps from RNA6 to RNA8 are necessary for hybrid completion.
(**A**) Schematic representation of initiation templates containing the strong *gdh* promoter of *P. furiosus*. The first C residue in the transcribed region of the *gdh*-C9 template occurs at position +10 (Spitalny and Thomm, 2003). The sequence elements BRE^{U} and TATA-box are underlined. (**B**) *In vitro* transcription reactions were performed on *gdh*-C9 template in the presence or absence of GTIFs and 1 nt extension products of priming RNAs RNA2 to RNA9 (panel A) were analyzed on 28% polyacrylamid gels. (**C**) *In vitro* transcription reactions were performed on gdh-C9 template with WT RNAP in the presence or absence of GTIFs and extension products of priming RNAs RNA5 to RNA7 (panel A) were analyzed on a 28% polyacrylamid gels. Asterisk indicates radioactive nucleotide.

D. TL function in catalysis

1. Complement UTP addition

In order to explore the role of the archaeal TL during transcription elongation in more detail, we used an *in vitro* assembly of nucleic acids originally described by Kireeva et al (2000). Elongation scaffold EC(U) comprises a fully complementary 83 bp double-stranded DNA template and the radioactively labeled RNA14, which forms a 9 bp hybrid with the template DNA strand followed by a 5 nt 5' overhang that is non-complementary to the DNA (Figure 20A) (Grünberg et al., 2010). On this template, we determined the incorporation rate of a single complementary UTP (cUTP). The reaction of cUTP incorporation was slower for all mutant RNAPs in comparison with the WT RNAP (Figure 20B and C). As expected, among all mutants tested the ΔTL mutation had the strongest effect on the incorporation rate. However, as the WT enzyme was very fast in all assays at optimal reaction temperature (70°C) and because of the high amount of misincorporation products (RNA16 and RNA17), we could not measure the reaction rate (k_{pol}) and the substrate affinity (K_M) values. Thus a lower than optimal reaction temperature (50°C) was chosen that permitted determination of reaction kinetics. Moreover, since we also aimed at studying discrimination against a naturally occurring deoxyribonucleotide (2'dNTP), we used the template EC(A) (Figure 21A) which allowed us to determine the sequence-specific incorporation of one complementary ATP, one 2'dATP or of one non-complementary (nc)UTP.

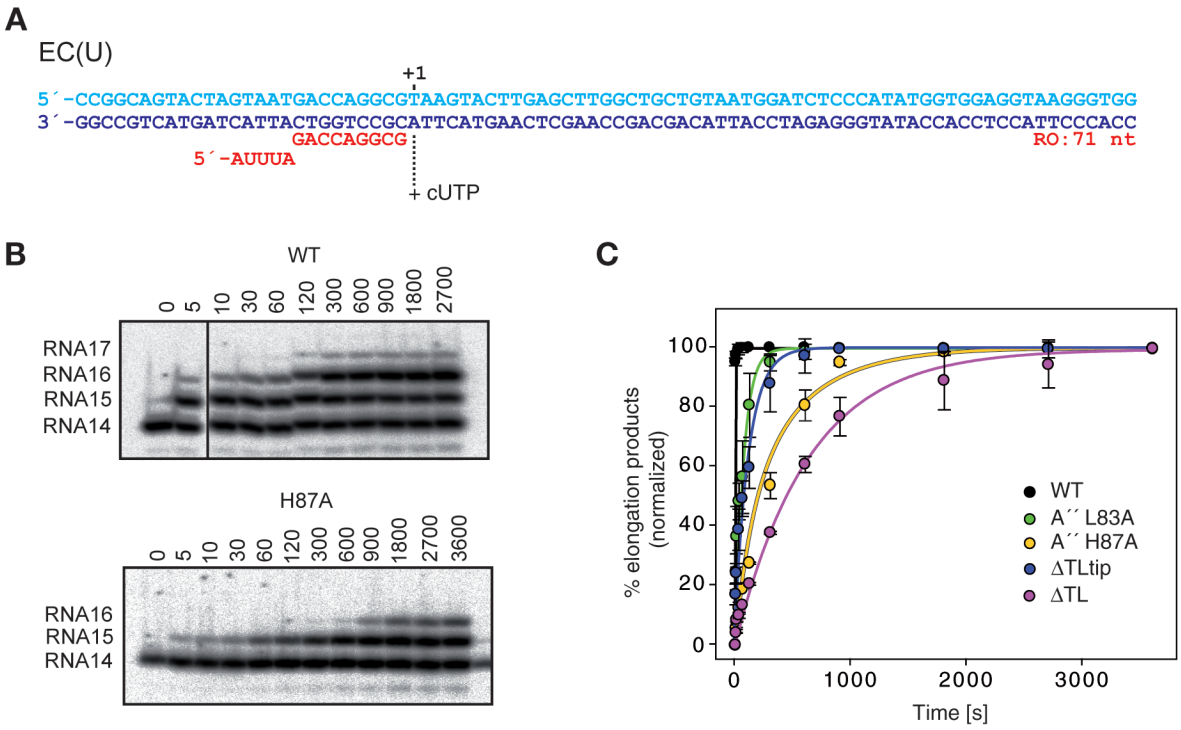

Figure 20. The role of the TL in catalysis with EC(U) template. (**A**) Elongation scaffold template EC(U) was employed for nucleotide incorporation assays. The 5' end of the RNA was labeled by ^{32}P. cUTP, complement UTP. (**B**) Representative gels of cUTP (100 µM) incorporation by WT and H87A RNAPs at 70°C are shown. (**C**) Kinetics of

cUTP (100 µM) incorporation on the EC(U) template. Solid lines represent a single exponential fit of single nucleotide addition time course data.

2. Complement ATP addition

Under these conditions, the WT enzyme added cATP at a k_{pol} of 13 s^{-1}, performing in the range of other RNAPs tested *in vitro* (Figure 21B and C; Table I) (Kaplan et al, 2008; Yuzenkova et al, 2010). The reaction with cATP was slower for all mutant RNAPs in comparison to WT RNAP (Appendix Figure S1). As previously observed, the ΔTL mutation had the strongest effect on the incorporation rate, resulting in an approximately 30-fold decrease, whereas the affinity to the substrate was decreased approximately 3-fold. Thus the archaeal TL is crucial for efficient NTP substrate binding and RNA chain elongation.

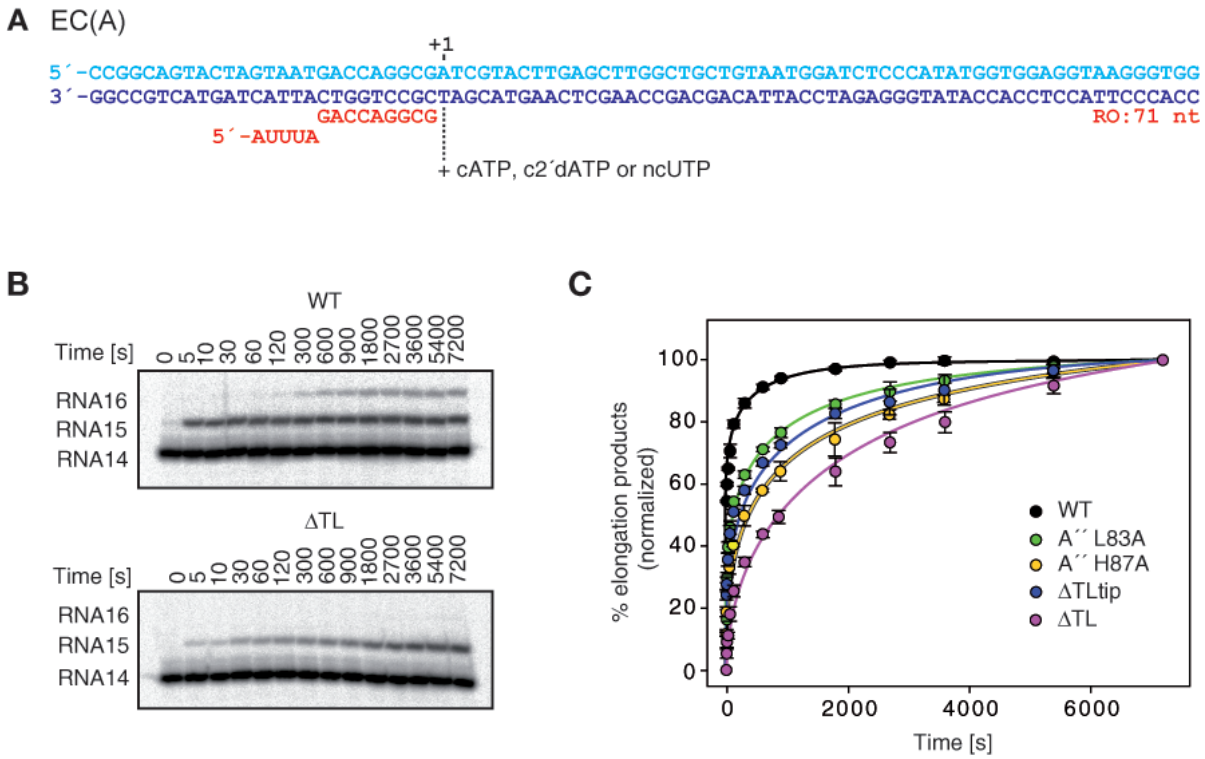

Figure 21. The role of the TL in catalysis with EC(A) template. (**A**) Elongation scaffold template EC(A) was employed for nucleotide incorporation and misincorporation assays. The 5' end of the RNA was labeled by ^{32}P. cATP, complement ATP. ncUTP, non-complement UTP. c2'dATP, 2'complement deoxy-ATP. (**B**) Representative gels of cATP (100 µM) incorporation by WT and ΔTL RNAPs at 50°C are shown. (**C**) Kinetics of cATP (100 µM) incorporation on the EC(A) template. Solid lines represent a single exponential fit of single nucleotide addition time course data.

Among the single point mutations examined, H87A resulted in the strongest defect, decreasing k_{pol} by about 10-fold and reducing substrate affinity by about 2-fold (Table I). The RNA synthesis rate with the L83A mutant decreased by 5-fold, while the substrate affinity was also reduced by around 2-fold. Interestingly, deletion of the mobile tip of the TL reduced the

incorporation rate, while increasing the affinity for the substrate (Table I). We conclude that the TL and specifically A''H87 are required for efficient catalysis and that A''L83 and A''H87 contribute to substrate binding, but not the centrally located mobile tip of the TL.

RNAP	cATP		ncUTP		c2'dATP	
	k_{pol} (s^{-1})	K_M (μM)	k_{pol} (s^{-1})	K_M (μM)	k_{pol} (s^{-1})	K_M (μM)
WT	13 ± 0.6	29 ± 2	2.1×10^{-3} ± 0.4	1128 ± 92	15×10^{-2} ± 0.2	228 ± 20
A" L83A	2.6 ± 0.2	53 ± 4	4.9×10^{-3} ± 0.5	566 ± 63	3.8×10^{-2} ± 0.4	170 ± 5
A" H87A	1.1 ± 0.5	68 ± 5	1.3×10^{-3} ± 0.3	737 ± 50	2.5×10^{-2} ±0.3	158 ± 16
ΔTLtip	2.1 ± 0.3	18 ± 1	1.6×10^{-3} ± 0.2	789 ± 60	5×10^{-2} ±0.1	184 ± 8
ΔTL	4.2×10^{-1} ± 0.3	88 ± 10	7.4×10^{-4} ± 0.5	405 ± 40	1.9×10^{-2} ±0.3	95 ± 5

Table I. K_M and k_{pol} for incorporation and misincorporation reactions by WT and mutant RNAPs. k_{pol} (reaction rate at saturating NTP concentration) and K_M were obtained by hyperbolic fitting of the kinetic data into the Michaelis-Menten equation as described in Materials and Methods. The values determined are shown with standard error (±).

E. TL function in NTP selection and transcription fidelity

The misincorporation reaction of ncUTP advanced slowly, k_{pol} was reduced by 3 to 4 orders of magnitude with all RNAPs, demonstrating the presence of an efficient discrimination mechanism for non-basepairing substrates in the archaeal enzyme (Table I; Appendix Figure S1). While the reaction rates with ncUTP were similar with mutant and WT enzymes, the K_M values were lower for all TL mutants. Accordingly, the discrimination rate for ncUTP was relatively high at approx. 240,000-fold for the WT RNAP, whereas the ability to recognize the correct NTP was reduced by approximately 25-, 40- and 90-fold for the H87A, L83A and ΔTL mutant RNAPs, respectively (Figure 22). Deletion of the TL tip resulted in only a moderate, 4-fold reduction on ncUTP discrimination. Thus, both A'' L83 and A'' H87 not only participate in catalysis and substrate binding, but also significantly contribute to the recognition of the correct NTP, while the mobile tip of the TL is dispensable for this function.

Incorporation and binding of 2'dATP by the WT enzyme was more efficient in comparison to the utilization of the non-cognate UTP as reflected by the lower, 680-fold discrimination rate (Figure 22C and D). As indicated in Table I, the ΔTL RNAP mutant most poorly distinguished between cATP and c2'dATP, exhibiting a discrimination rate that was reduced by almost 30-fold in comparison to the WT enzyme. The discrimination rate observed for the ΔTLtip mutant was only slightly reduced by around 1.5-fold, whereas the single point mutations L83A reduced discrimination by 3-fold, and H87A by approx. 7-fold.

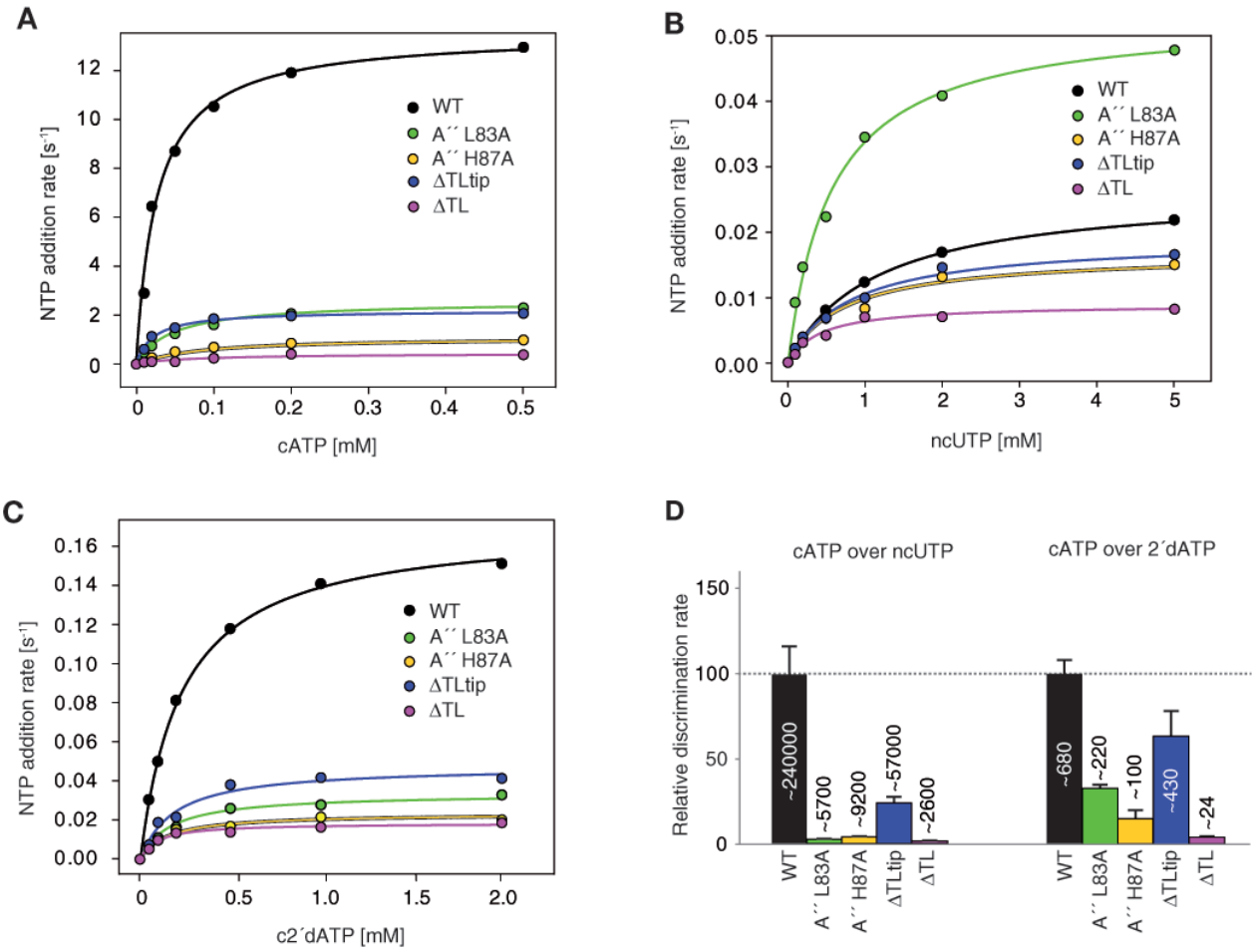

Figure 22. (Mis)incorporation dependence on the concentrations of (A) cATP, (B) ncUTP or (C) c2'dATP by WT, L83A, H87A, ΔTLtip and ΔTL RNAPs. The scaffold template EC(A) was incubated with the indicated substrate concentrations at 50°C. The reactions at each NTP concentration were performed in triplicates. The curves show the data fit into the Michaelis-Menten equation described in Materials and Methods. (**D**) Discrimination of cATP over c2'dATP and cATP over ncUTP by WT and mutants RNAPs. Discrimination ratios were calculated by taking the ratio of k_{pol}/K_M of cATP vs. 2'dATP or cATP vs. ncUTP. k_{pol} and K_M data are shown on Table 1.

F. TL function in NTP over 2'dNTP discrimination

The previous data indicate that the archaeal TL contributes to 2'dNTP discrimination, but also that additional residues might contribute. Structural studies of yeast RNAP II delineated an interaction network between Rpb1 residues R446 (*Pfu* A' R423) and N479 (*Pfu* A' N456) with the 2'-OH group of the ribose moiety (Wang et al, 2006; Cheung et al, 2011). In addition, a contact of the TL residue Q1078 (*Pfu* A'' Q80) with the 2'-OH group was observed recently (Figure 23A) (Cheung et al, 2011). Moreover, functional studies on yeast RNAP II showed that serine substitution of yeast TL residue Rpb1 F1086 (*Pfu* A'' Y88) could improve the 2'dNTP discrimination (Kaplan et al, 2008).

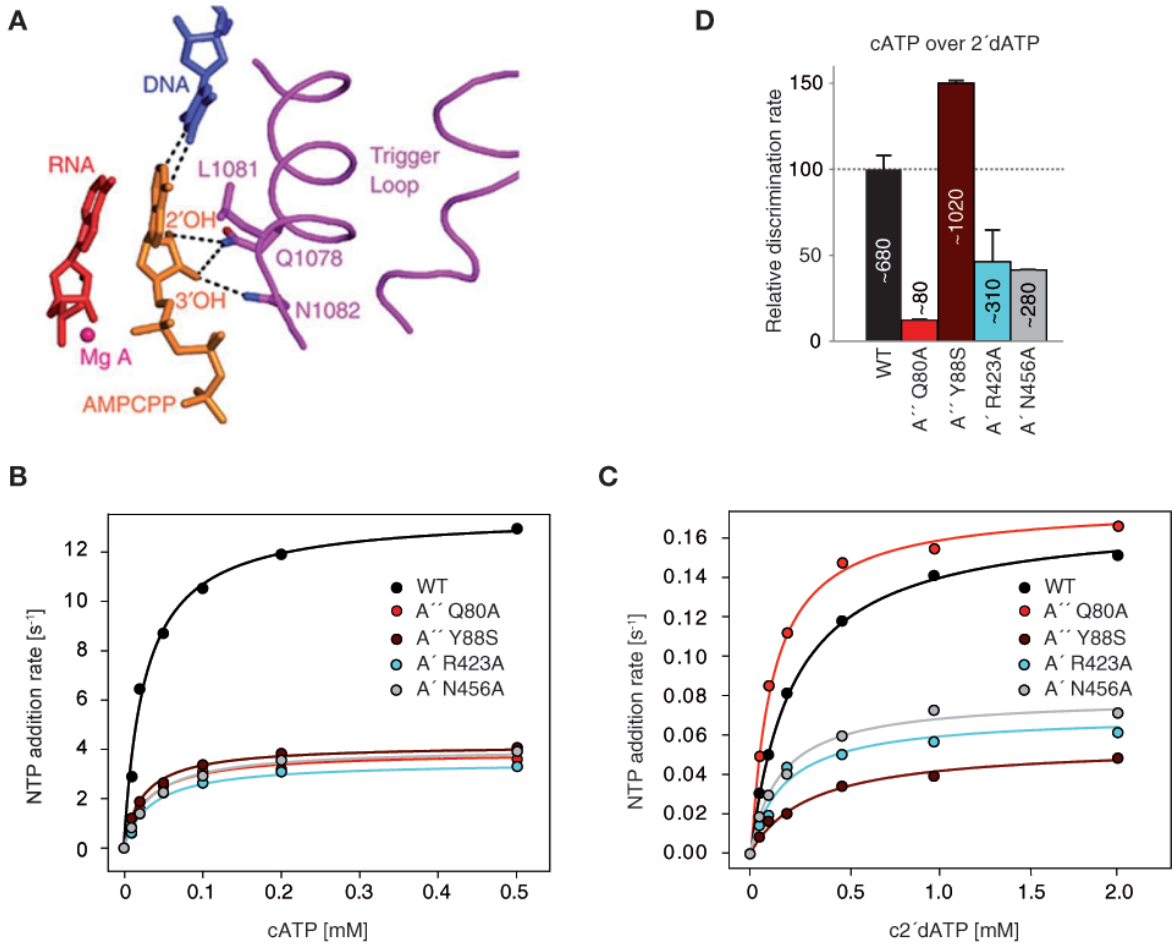

Figure 23. The role of the TL residue A'' Q80 in NTP 2´-OH group recognition. (**A**) Contact of AMPCPP with closed TL (Cheung et al., 2011). Incorporation dependence on concentrations of (**B**) cATP and (**C**) c2'dATP by WT, Q80A, Y88S, R423A and N456A RNAPs. The scaffold template EC(A) was incubated with the indicated substrate concentrations at 50°C. The reactions at each NTP concentration were performed in triplicates. The curves show the data fit into the Michaelis-Menten equation described in Materials and Methods. (**D**) Discrimination of cATP over c2'dATP by WT and mutants RNAPs. Discrimination ratios were calculated by taking the ratio of k_{pol}/K_M of cATP vs. 2'dATP. k_{pol} and K_M data are shown on Table 2.

These results prompted us to investigate the effect of analogous substitutions at homologues residues in the archaeal RNAP. In agreement with results from the eukaryotic system, mutation Y88S resulted in a phenotype that improved discrimination of cATP over 2'dATP (Figure 23 B, C and D; Table II). While the A' R423A and A' N456A substitutions had only a moderate effect on 2'dATP discrimination, we found that the A'' Q80A substitution indeed strongly decreased the discrimination against 2'dATP by approx. 8.5-fold (Figure 23B, C and D; Table II). These results show that a novel contact of a conserved glutamine residue in the TL with the NTP 2'-OH group contributes to selection of NTP over 2'dNTP substrates.

RNAP	cATP		c2'dATP	
	k_{pol} (s^{-1})	K_M (μM)	k_{pol} (s^{-1})	K_M (μM)
A" Q80A	3.8 ± 0.4	29 ± 2	1.7 x 10^{-1} ± 0.1	116 ± 9
A" Y88S	3.9 ± 0.3	26 ± 1	4.8 x 10^{-2} ± 0.5	325 ± 17
A' R423A	3.2 ± 0.2	30 ± 6	6.5 x 10^{-2} ±0.6	188 ± 9
A' N456A	3.8 ± 0.4	31 ± 2	7 x 10^{-2} ±0.4	161 ± 4

Table II K_M and k_{pol} for cATP and c2'dATP incorporation reactions by WT and mutant RNAPs. k_{pol} (reaction rate at saturating NTP concentration) and K_M were obtained as described in Table I and in Materials and Methods.

G. TL is not required for intrinsic RNA cleavage

In order to decipher the role of the archaeal TL during intrinsic RNA hydrolysis, we used the mismatched scaffold templates MEC(C) and MEC(G) in which the ultimate nucleotide at the 3′ end (CMP or GMP) of RNA15 is non-complementary to the corresponding template DNA base, while the upstream residues of template and RNA strands in the 9 bp hybrid region were complementary (Figure 24A). Such complexes are thought to be in the pre-translocated state, where the mismatched nucleotide is in position +1, either base-paired with the DNA template strand or frayed away from the template (Zenkin et al, 2006; Sydow et al, 2009). RNA cleavage requires backtracking of the complexes by one nucleotide, here referred to as +2 backtracked complexes (Zenkin et al, 2006; Sydow et al, 2009; Wang et al, 2009;).

Kinetic analysis of the intrinsic RNA cleavage reaction shows that neither the H87A mutation nor deletion mutants of the TL had a significant effect on the rate of RNA hydrolysis on either template in comparison to the WT activity (Figure 24B, C and E; Appendix Figures S2 and S3). This agrees with one study in the bacterial system (Zhang et al, 2010), but disagrees with another (Yuzenkova and Zenkin, 2010). The activity was higher on MEC(G) in comparison to MEC(C), with RNA degradation rates increased by approx. 3-fold (Figure 24E), consistent with earlier findings that describe faster RNA cleavage on purine than on pyrimidine mismatches (Sydow et al, 2009; Yuzenkova and Zenkin, 2010).

On MEC(C), the first cleavage product produced was RNA14, for both WT and mutant RNAPs (Figure 24C; Appendix Figure S2B), indicating that the 3'-terminal mismatched CMP may stimulate the cleavage reaction at the ultimate phosphodiester bond (Sosunov et al, 2003; Zenkin et al, 2006). Furthermore, this result shows that the TL does not influence the conformation of the ternary complex on MEC(C). In contrast, with MEC(G) the first and major cleavage product was RNA13, while RNA14 is a minor cleavage product arising later during the course of the experiment. Intriguingly, only the WT RNAP was able to produce the RNA14 cleavage product on the MEC(G) template, whereas cleavage with the H87A, ΔTLtip and ΔTL

RNAPs yielded RNA13 and smaller cleavage products only (Figure 24B; Appendix Figure S2A). These results indicate that A'' H87 and other residues in the TLtip region can influence the translocation state of the elongation complex.

Recently it was proposed that the TL of the *T. aquaticus* RNAP has a crucial role during substrate-assisted, intrinsic RNA cleavage activity at the penultimate phosphodiester bond by deprotonating the attacking water molecule with the invariant histidine (β' H1242) (Yuzenkova and Zenkin, 2010). To find out if the related archaeal TL residue A'' H87 participates in intrinsic cleavage by the same mechanism, we analyzed the Mg^{2+} and pH dependence of the dinucleotide hydrolysis reaction on the MEC(G) template with WT and mutant RNAPs. As observed with the corresponding bacterial RNAP mutants, the inflection point at approx. pH 7.5 observed in the pH profile of the WT RNAP shifts to a higher pH in the H87A and the ΔTL mutant RNAPs, while the pH profiles of the ΔTLtip RNAP mutant closely resembles the WT profile (Figure 24F). Unlike for bacteria, the A'' H87 mutant regained the RNA cleavage activity at high pH up to WT levels, indicating that other residues can substitute for the A'' H87 function. In addition, Mg^{2+} dependence was only slightly affected by the alanine substitution of A'' H87 or by the deletion of the TL (Figure 24G). This result indicates that, unlike in bacteria, A'' H87 is not essential to orient the 3´end NMP for NMP-assisted hydrolysis of the penultimate phosphodiester bond. Thus, A'' H87 is not essential, but catalytically participates in the intrinsic RNA hydrolysis reaction at the second phosphodiester bond.

In contrast to all other TL mutant enzymes, we found that mutant RNAP L83A weakly bound to the mismatched scaffolds and that the complexes formed dissociated easily (Figure 24B). Ternary complexes formed on the matched elongation template EC(U) exhibited a very low intrinsic cleavage activity for both WT and mutant RNAPs. This is consistent with the concept that intrinsic RNA cleavage by the RNAP is stimulated by the 3' terminal mismatched NMP (Sosunov et al, 2003; Zenkin et al, 2006; Sydow et al, 2009). Moreover, we find that the L83A RNAP was stable in the absence of a mismatched nucleotide at the 3'-end of the RNA. Thus, A'' L83 is required for stabilization of the RNAP on mismatched templates (Figure 24D). Together these data show that the TL is not required for intrinsic RNA cleavage by RNAP.

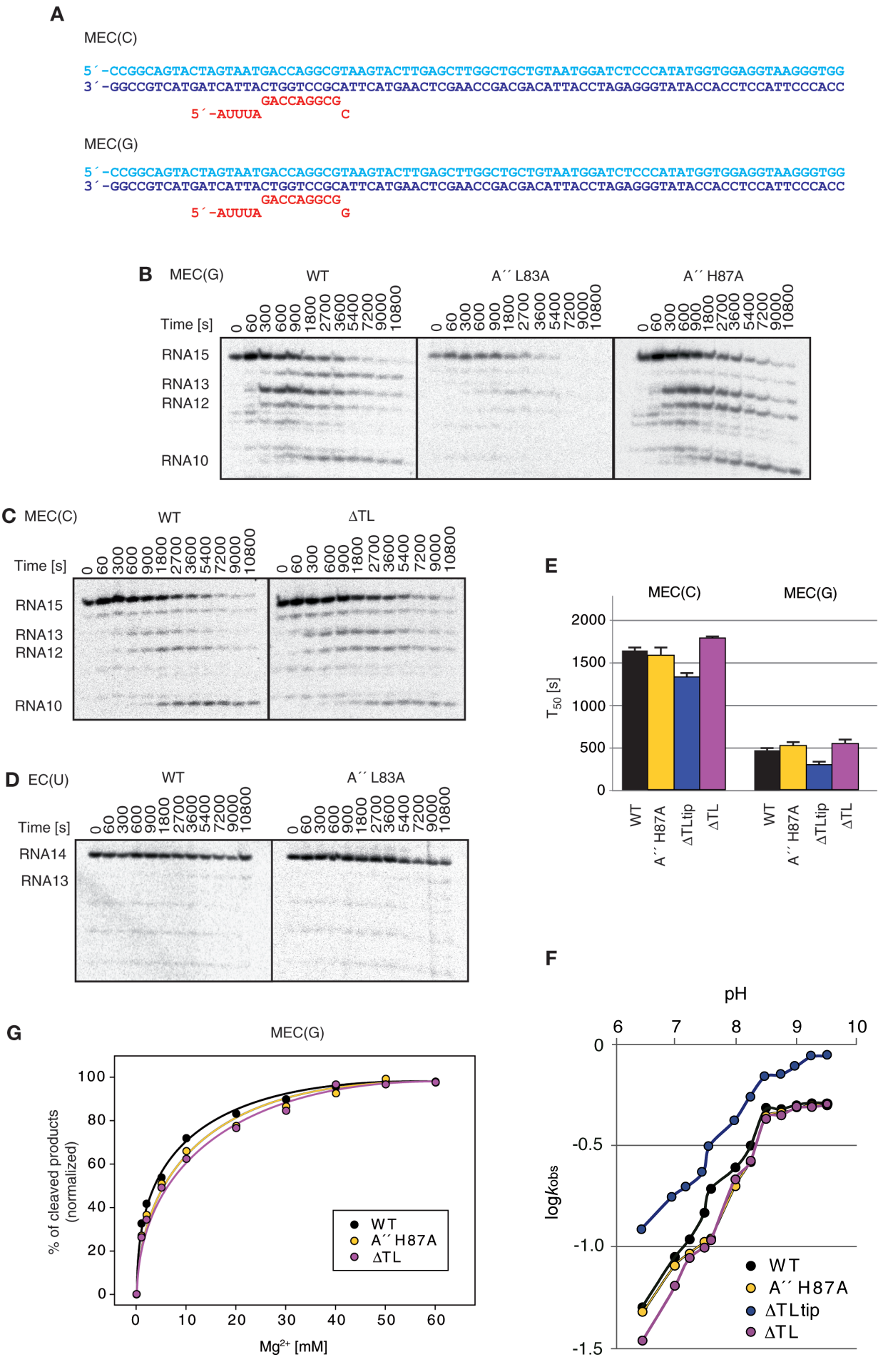
A
MEC(C)
5´-CCGGCAGTACTAGTAATGACCAGGCGTAAGTACTTGAGCTTGGCTGCTGTAATGGATCTCCCATATGGTGGAGGTAAGGGTGG
3´-GGCCGTCATGATCATTACTGGTCCGCATTCATGAACTCGAACCGACGACATTACCTAGAGGGTATACCACCTCCATTCCCACC
GACCAGGCG
5´-AUUUA C
MEC(G)
5´-CCGGCAGTACTAGTAATGACCAGGCGTAAGTACTTGAGCTTGGCTGCTGTAATGGATCTCCCATATGGTGGAGGTAAGGGTGG
3´-GGCCGTCATGATCATTACTGGTCCGCATTCATGAACTCGAACCGACGACATTACCTAGAGGGTATACCACCTCCATTCCCACC
GACCAGGCG
5´-AUUUA G
B MEC(G)
WT
A´´ L83A
A´´ H87A
Time [s]
0 60 300 600 900 1800 2700 3600 5400 7200 9000 10800
RNA15
RNA13
RNA12
RNA10
C MEC(C)
WT
ΔTL
Time [s]
0 60 300 600 900 1800 2700 3600 5400 7200 9000 10800
RNA15
RNA13
RNA12
RNA10
D EC(U)
WT
A´´ L83A
Time [s]
0 60 300 600 900 1800 2700 3600 5400 7200 9000 10800
RNA14
RNA13
E
MEC(C)
MEC(G)
T50 [s]
2000
1500
1000
500
0
WT
A´´ H87A
ΔTLtip
ΔTL
F
pH
6 7 8 9 10
0
-0.5
-1.0
-1.5
logkobs
WT
A´´ H87A
ΔTLtip
ΔTL
G
MEC(G)
% of cleaved products (normalized)
100
80
60
40
20
0
0 10 20 30 40 50 60
Mg2+ [mM]
WT
A´´ H87A
ΔTL

Figure 24. TL is not essential for intrinsic RNA cleavage activity. (**A**) Elongation scaffold templates employed for intrinsic cleavage assays. MEC(C) and MEC(G) contained one mismatched cytidine and guanosine at the 3'end of RNA, respectively. (**B)** and (**C**) Representative gels of intrinsic phosphodiester bond hydrolysis in MEC(G) and MEC(C) by WT, H87A and ΔTL RNAPs. (**E**) The times required to reach 50% RNA cleavage products (T_{50}) were obtained for MEC(C) and MEC(G) by single exponential fit of the intrinsic RNA cleavage product rate versus the reaction time (Supplementary Figure S3A and B). (**F**) pH profiles of second phosphodiester bond hydrolysis on the MEC(G) scaffold by WT, A'' H87A, ΔTLtip and ΔTL RNAPs. The observed cleavage rates (k_{obs}) were measured after a reaction time of 15 min. (**G**) Mg^{2+} dependences of the hydrolysis of the second phosphodiester bond in MEC(G).

H. TL is not required for TFS-stimulated RNA cleavage

The RNA cleavage activity in the presence of TFS occurred with WT and mutant RNAPs faster than without TFS (Figure 25C). Moreover TFS efficiently stimulated RNA hydrolysis in the absence of a 3'-mismatch. In line with earlier results, on both templates the first cleavage occurs predominantly at the penultimate phosphodiester bond (Figure 25A and B; Appendix Figure S4) (Wang et al, 2009; Grünberg et al, 2010). However, we note that on EC(U), the WT enzyme was also able to cleave the terminal 3'-phosphodiester bond, but the TL deletion mutants were not(Figure 25A, Appendix Figure S4A), indicating that the TL affects the translocation state on matched templates also in the presence of TFS. TFS also rescued the low stability of the ternary complex with the L83A mutant on MEC(G), demonstrating that binding of TFS itself exerts a stabilizing effect on this mutant in the presence of a 3'-terminal mismatch (Figure 25B; Appendix Figure S7). The slight decrease in RNA cleavage activity noted with the ΔTL and ΔTLtip mutants was also observed in the bacterial system and can be explained by a minor disarrangement of the active site caused by the deletion (Rhoganian et al, 2011). Moreover, the pH profiles of TFS-induced cleavage by WT and ΔTL RNAPs were similar to each other, and different from intrinsic RNA cleavage reaction (Figure 25D). The difference between the pH profiles of intrinsic and TFS-induced cleavage by WT and ΔTL RNAPs is consistent with previous observations in bacteria (Rhoganian et al, 2011), and can be attributed to stabilization of Mg_B and the attacking hydroxyl by the catalytic domain of TFS. Our results show that TFS-induced RNA cleavage proceeds without participation of the TL.

Figure 25. TL-independent TFS-induced RNA cleavage. Representative gels of TFS induced phosphodiester bond hydrolysis on (**A)** EC(U) and (**B**) MEC(G) scaffolds. Black asterisks indicate nonspecific RNA degradation products. The arrow indicates the RNA products cleaved at the terminal 3'-phosphodiester bond. (**C**) Kinetics of the TFS induced RNA cleavage reaction in MEC(G) scaffold template by WT and mutant RNAPs. Solid curves are the single exponential fits of the kinetics data. The reactions were performed in triplicates. pH profiles of second phosphodiester bond hydrolysis on the MEC(G) scaffold by WT and ΔTL RNAPs. The observed cleavage rates (k_{obs}) were measured after a reaction time of 5 min.

I. The TL functions in suppressing abnormal transcription termination

To investigate if the TL plays a role during transcription termination, we compared the termination ability of WT and ΔTL RNAPs. Elongation assays were performed as described above, but at increasing temperatures (70, 80 and 90°C) on template EC(U) and ECTerm, which contains a hepta-dT (T1 to T7) transcription termination sequence (Figure 26A). WT RNAP produced a run off transcript under all conditions tested (Figure 26B and C). Processivity improved with increased temperatures, apparent as fewer and less pronounced premature stop sites at dT. On both EC(U) and ECTerm templates, the overall amount of transcript slightly decreased at higher temperatures, while on the ECTerm template the termination efficiency improved, as reported previously (Figure 26B and C) (Santangelo et al, 2006; Spitalny and Thomm, 2008). Termination with WT RNAP occurred mainly at T5 and T7, whereas the ΔTL RNAP terminated transcription primarily at T-2, indicating that the TL contributes to anti-termination in T-rich sequences. At 90°C, the ΔTL enzyme was unable to elongate the scaffold RNA by more than 8 nt. At 70 to approx. 85°C premature stops were more prevalent with the TL deletion mutant in comparison to the WT enzyme (Appendix Figure S5). This effect was not reduced by increased NTP concentration (Appendix Figure S5).

Previously, the presence of the non-template DNA strand (NT) and RNAP subunits E' and F (homologues of eukaryotic Rpb7 and Rpb4, respectively) were characterized as processivity determinants (Hirtreiter et al, 2009). We therefore reconstituted WT and ΔTL enzymes lacking E'F and added either E'F, the NT or both during otherwise standard elongation complex assemblies. We found that processivity of the ΔTL enzyme was highly diminished in the absence of both NT and E'F so that elongation ceased at an RNA length of 30-35 nt, whereas the WT RNAP was still able to produce a 71 nt run off transcript under all conditions (Figure 26D). Adding back E'F allowed the ΔTL enzyme to slightly extend RNA synthesis until a length of about 35-40 nt. The run off product was produced by the ΔTL RNAP only in the presence of the NT.

However, neither WT nor mutant RNAPs were responsive to E'F addition in the presence of the NT (Figure 26D; Appendix Figure S5). We conclude that the ΔTL RNAP mutant exhibits a temperature-sensitive processivity defect and is prone to termination at T-rich sequences. These results show that the TL is required for suppressing termination at non-terminator sites.

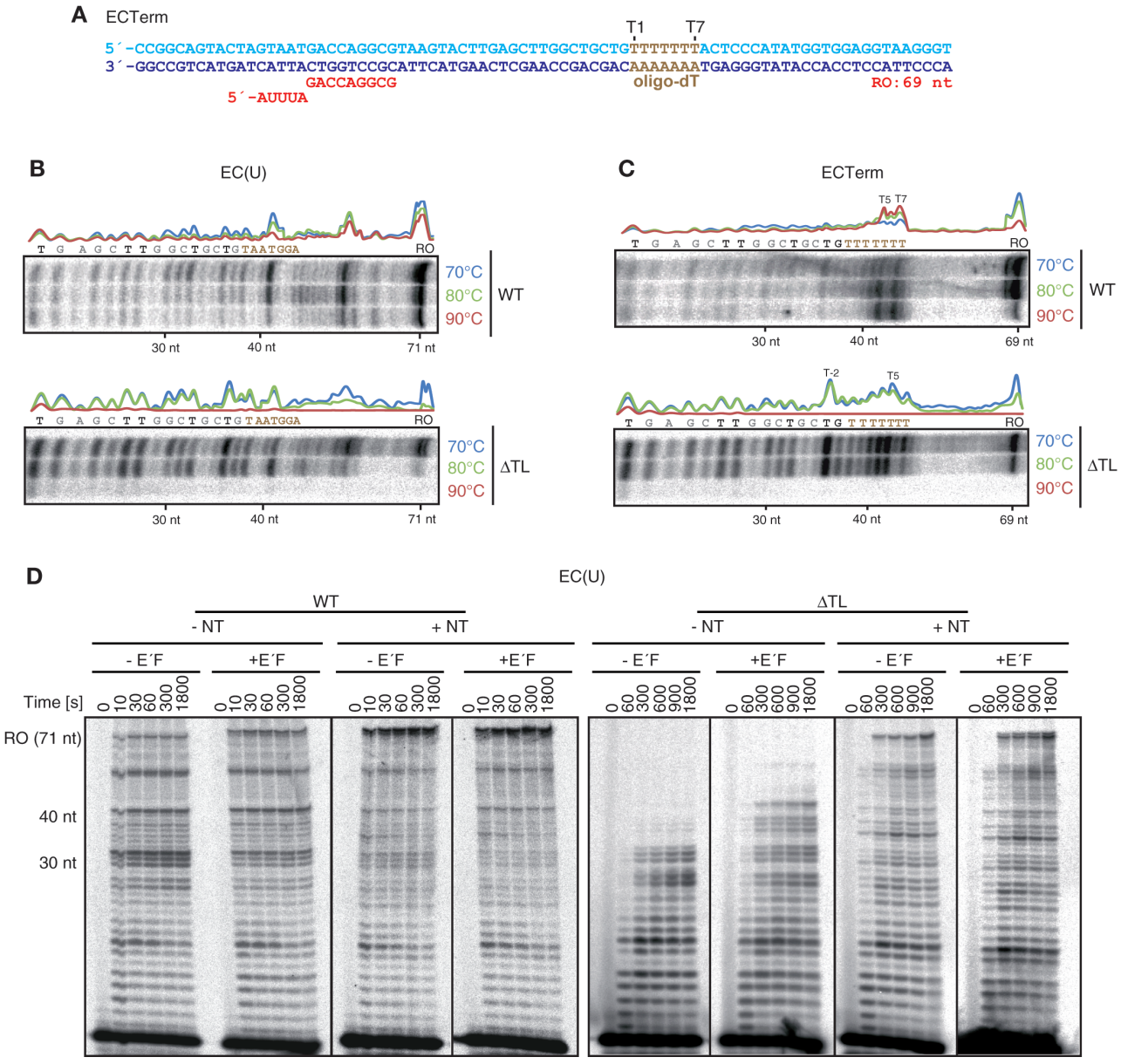

Figure 26. Implications of the TL in termination and processivity. (**A**) Elongation scaffold template ECTerm contains an oligo-dT sequence from positions +36 to +42. Transcription elongation assays with WT and ΔTL RNAPs were performed at at 70, 80 and 90°C, with (**B**) EC(U) and (**C**) ECTerm templates. The curves are based on quantification of elongation products after a reaction time of 5 min for WT RNAP and 15 min for ΔTL RNAP (arbitrary units). (**D**) Transcription elongation kinetics with EC(U) at 70°C in the presence or absence of the non-template DNA strand (NT) and/or E'F subunits by WT and ΔTL RNAPs.

V) Discussion

A. The essential role of the TL during transcription initiation

One of the striking findings of our analysis is the critical function of the TL for transcription initiation. Our reconstituted RNAPs were fully responsive to TFE, TFS and E'F in all assays, and transcription bubble formation was normal, indicating that the overall structural integrity was preserved despite the TL mutations. In co-crystal structures of RNAP II-TFIIB or bacterial RNAP-σ, the TL was not resolved or mainly flexible and therefore structural information of early transcribing complexes are not yet available (Vassylyev et al., 2002; Kostrewa et al., 2009; Cheung et al., 2011). However, the function of the TL during initiation must be catalytic, and formation of the first phosphodiester bonds cannot be compensated by other catalytic residues of the RNAP, or allosterically through TFB (Sainsbury et al., 2013). Given that in all promoter dependent assays the activity with the L83A mutant was reduced, while A'' H87 and the TLtip region of the TL were essential, NTP binding and appropriate active site closure, but not exact NTP positioning are the limiting factors for catalysis in initially transcribing complexes (ITC). Therefore, the TL is essential for capturing the first NTPs at the onset of transcription.

A recently described initially transcribing RNAP II-TFIIB complex from yeast showed that RNA extension beyond 6 and 7 nt requires RNA strand separation from the DNA to redirect the nascent transcript underneath the B-reader loop towards the RNA exit channel (Sainsbury et al., 2013). This is consistent with our finding that RNA6 and RNA7 are weakly extended by the WT enzyme and that maintenance of a productive ITC requires the consecutive steps of RNA synthesis between 6 and 8 nt.

B. The function of A'' L83 in transcription fidelity

Our single nucleotide addition experiments revealed, that A'' L83 contributes to fast catalysis and is a major discriminator against the wrong NTP. Among eukaryotic RNAPs the residue is conserved, whereas bacterial RNAPs also contain the non-polar, functionally equivalent residue methionine at the corresponding position. Substitution variants in eukaryotes were not yet assayed due to lethality of the mutants (Kaplan et al., 2012). According to Zenkin and colleagues works, the bacterial counterpart of the L83A mutant (*T. aquaticus* RNAP β' M1238A) was slow in catalysis, exhibited a decreased affinity for the cNTP and an increased affinity for the ncNTP (Yuzenkova et al., 2010). Crystal structures of elongation complexes from bacteria and yeast demonstrated that both, leucine, as well as methionine, stacks against the base of the incoming nucleotide, providing a structural basis for the role of this moiety in catalysis and fidelity (Wang et al., 2006; Vassylyev et al., 2007). Thus, these results suggest that the function of A'' L83 in nucleotide positioning and correct NTP selection is conserved among all three domains of life.

Interestingly, elongation complex stability of the L83A RNAP is drastically decreased on mismatched, but not on matched elongation scaffolds, and TFS is able to rescue this stability defect. Based on crystal structures of backtracked complexes in yeast, Wang and co-workers proposed that Rpb1 residues Q1078 to L1081 might contact the backtracked RNA (Wang et al., 2006). Notably, the destabilization of the L83A RNAP observed in our assays was apparently independent from the nature of the 3' end of the RNA, which defines the exact configuration of the mismatched complex (see below; Sydow et al., 2009; Wang et al., 2009). We thus propose that the interaction of A'' L83 with whichever mismatched 3' end of the RNA is required to avoid complex disassembly, probably by limiting the free motion of the 3' end NMP. Binding of TFS would impede RNA 3' end motion by directly contacting the scissile RNA bond, leading to ternary complex stabilization. Taken together, our results indicate that A'' L83 in the archaeal RNAP interacts with the nucleic acid base of either the incoming NTP in the insertion site or in the absence of NTPs with the base at the 3' end of the RNA, thus contributing to fidelity during transcription elongation and to stability of arrested ternary complexes once an erroneous NTP was incorporated.

C. Substrate binding and catalysis

Among all single point mutations tested, the H87A mutation had the strongest effect on the catalysis rate, in line with the conserved nature of this residue. In yeast, the alanine substitution of Rpb1 His1085 is lethal, but viability is restored in a double mutant with the superactive E1103G mutation (Kaplan et al., 2012). In bacteria, comparably strong defects on catalysis are only obtained with a double mutant, where the conserved histidine (*T. thermophilus* β' H1242/*E. coli* β' H936) and an arginine that is specific to the bacterial TL (*T. thermophilus* β' R1239/ *E. coli* R933) are substituted by alanine (Yuzenkova et al., 2010; Zhang et al., 2010). Thus, the critical function of A'' H87 during catalysis is confined to the archaea-eukaryotic type of RNAPs.

Deletion of only the mobile tip region of the TL results in a reduced catalytic rate, presumably due to incomplete final closure of the active site. However, this was the only mutant were we noted an increased affinity for the substrate, confirming that the substrate binding residues are not distorted in this variant. The improved substrate binding of the ΔTLtip RNAP might be the consequence of facilitated NTP access to a wider space in the active site caused by the deletion.

D. Discrimination against the wrong nucleotide

Mutagenesis of residues that were recently proposed to be involved in hydroxyl interactions (Cheung et al., 2011), showed that A' R423 and A' N456 moderately affect 2'dNTP discrimination, but the key residue is A'' Q80, for which the alanine substitution is lethal in yeast

(Kaplan et al., 2012). Available *in vitro* data confirmed that in *E. coli* (Svetlov et al., 2004) and yeast (Wang et al., 2006), the A' N456 homologue contributes to 2'-OH discrimination, but not the corresponding residue in the *T. aquaticus* RNAP (Yuzenkova et al., 2010). Crystal structure and gentle mutational analyses in yeast RNAP II suggested that the TL residue Q1078 may couple the 2'-OH and also 3'-OH recognition with TL folding. (Wang et al., 2006; Cheung et al., 2011; Kaplan et al., 2012). The fact that substitution of monoamine residue A'' Y88 by the non-aromatic serine improves 2'dNTP discrimination in the archaeal RNAP, as it is the case in yeast RNAP II (serine substitution of monoamine residue Rpb1 F1086) (Kaplan et al., 2008), supports the concept that the mechanism of NTP/2'dNTP distinction is conserved among the eukaryotic and the archaeal enzyme. Moreover, alanine substitution of A'' H87 results in imperfect discrimination against non-coding NTPs and 2'dNTPs, indicating that H87A causes weak closure of the active site preventing NTP recognition, as observed in other transcription systems (Kaplan et al., 2008; Huang et al., 2010;). Taken together, our results provide first evidence that A'' Q80 is required to recognize the 2'-OH group of the incoming NTP, and we propose that this control step requires complete active site closure through A'' H87.

E. TL-dependent and TL-independent RNA proofreading

The considerable differences observed between CMP and GMP mismatched RNA hydrolysis patterns indicate that the precise configuration of the ternary complexes depends on the type of mismatched NMP at the 3'-end, as described for similar assemblies in bacteria and eukaryotes (Sydow et al., 2009; Wang et al., 2009; Yuzenkova and Zenkin, 2010; Zhang et al., 2010). These conformational differences might contribute to the generally faster RNA hydrolysis rate on MEC(G) in comparison to MEC(C). This is expected as yeast RNAP II hydrolysis is faster on purine than on pyrimidine mismatched RNA templates (Sydow et al., 2009; Yuzenkova and Zenkin, 2010), revealing a significant functional divergence to the bacterial enzyme (Yuzenkova and Zenkin, 2010), neither A'' H87 nor the entire TL are essentially required for the RNAP intrinsic RNA hydrolysis activity.

Transcription fidelity is regarded to function as a Brownian ratchet mechanism (Bar-Nahum, 2005). The TL constantly swings against the nascent base pair located in the +1 position to test its stability. If it is stable, the RNAP translocates forward to position -1. If it is not stable, the 3' terminal nucleotide frays and the RNAP translocates backwards to form a +2-backtracked complex, which is able to cleave a dinucleotide. In our assays, dinucleotide cleavage is preferred in the WT enzyme, presumely through RNA self-assisted catalysis (Figure 4 B and C, Appendix Figure 5) (Zenkin et al., 2006). While the overall RNA cleavage rates, as well as cleavage at the penultimate phophosphodiester bond, were largely unaffected by the TL mutations, we found that the H87A and the TL deletion mutants were unable to induce RNA cleavage at the ultimate phosphodiester bond on MEC(G), which corroborates that the TL contributes to oscillation between pre- and post-translocated states.

Interestingly, the overall stability of the elongation complex with the L83A RNAP mutant was drastically decreased in the presence of a mismatched RNA 3' nucleotide and TFS was able to rescue this stability defect. Despite a reduced fidelity of the L83A mutant, we did not observe complex disassembly during elongation, indicating that binding of an NTP at the insertion site counteracts complex dissociation. Thus, the nascent RNA contacting residue A'' L83 prevents elongation complex disassembly during oscillation between the pre-, post- and reverse translocated states in mismatched complexes.

In crystal structures of frayed complexes, the TL is flexible, and the conserved histidine would overlap the frayed nucleotide if the TL was folded (Sydow et al., 2009), whereas in +2-backtracked complexes the residue is apparently directed away from the RNA (Wang et al., 2009). However, supported by the pH profile of the MEC(G) cleavage reaction, we propose that A'' H87 can participate as a proton donor in dinucleotide hydrolysis at the penultimate phosphodiester bond. However, our complementary experimental data on MEC(C) suggests that different catalytic residues and mechanisms are implicated in the RNA hydrolysis reactions depending on the conformation of the mismatched base. Altogether, the TL promotes transcription fidelity by two distinct ways: fit discrimination (error prevention) and mismatch recognition by stabilizing the enzyme in the backtracked conformation (error detection).

In contrast to the intrinsic RNA cleavage reaction, we found that TFS induced RNA hydrolysis virtually independent from TL mutagenesis. The structural basis for the non-involvement of the TL in factor induced RNA cleavage illustrates that the TL is immobilized in an inactive ("locked") conformation and folded away from the active site in the presence of TFIIS or Gre, respectively (Kettenberger et al., 2003; Tagami et al., 2010; Cheung and Cramer, 2011). On the matched template the WT but not the TL formed a minor cleavage product shortened by one nucleotide, indicating that the TL, also in the presence of TFS, encourages oscillation to the pre-translocated state, thereby favoring rapid elongation continuation over RNA cleavage in the absence of a mismatch.

F. Implications for the mechanism of transcription termination

In archaea, transcription termination occurs at single or multiple poly-T stretches and is more effective at elevated temperatures *in vitro* (Santangelo et al, 2006; Spitalny and Thomm, 2008; Santangelo and Reeve, 2009). In our analysis we found that the sensitivity to poly-T sequences in the archaeal enzyme is increased in the ΔTL enzyme, suggesting that the open conformation of the active site favors pre-mature transcription stops at termination sequences. Furthermore, early termination with the ΔTL RNAP is likely favored by the combined effects of the overall reduced processivity, substrate affinity and catalytic rate, which altogether has several implications for the mechanism of termination. We conclude that the TL prevents aberrant

termination at non-terminator sites, to ensure processive RNA synthesis. Thus, in the absence of the TL or upon its mutation, RNAP is more prone to pausing and termination at T-rich sites.

G. TL dynamics in the transcription cycle

Our key findings and previously observed TL contributions to the initiation and elongation phases of transcription are summarized in Figure 27. During initiation, A'' H87 and the tip region of the TL are essential for NTP binding to allow synthesis of the first nucleotide bonds and are required until hybrid completion (8 nt). Following bubble collapse (9 to 11 nt), the RNAP clears from the promoter to enter the productive elongation phase.

During elongation, ternary complexes in the post-translocated state bind the incoming NTP at the empty +1 site, which is accompanied by TL closure. The TL, and specifically A'' H87, is required for NTP binding and catalysis. A'' Q80 interacts with the 2'-OH group of the NTP in order to occlude 2'dNTPs. Base-stacking interactions by A'' L83 position the NTP and confer recognition of the complement substrate. Finally. The invariant histidine forms a stable interaction with the triphosphate moiety of the NTP, serves as proton donor to the β-phosphate during phosphoanhybride bond breakage (Castro et al., 2009; Huang et al., 2010; Carvalho et al., 2011). Pyrophosphate formation leads to the loss of interaction between the TL and the substrate, promoting the destabilization of the closed form of the TL (Da et al., 2012). The TL converts to the open conformation (pre-translocated state), which leads to the translocation of the nucleic acids. During translocation, A'' L83 forms a wedge next to the bridge helix which accompanies the movement of the nucleic acids (Brueckner and Cramer, 2008). Finally, translocation of the nucleic acids results in post-translocated state, in which the substrate binding site is free for binding the next incoming NTP, thus the NAC can be repeated (Brueckner et al., 2009).

Upon misincorporation, the RNA frays and the RNAP can reverse translocate (+2-backtracked conformation (Wang et al., 2009)), leading to dinucleotide cleavage through either TFS-stimulated RNA hydrolysis or intrinsic RNA cleavage that can be assisted by A'' H87. However, the TL stimulates forward translocation to the pre-translocated state that results in mononucleotide cleavage. Interestingly, A'' L83 is required to stabilize the backtracked complexes probably by interacting with the 3' NMP of the RNA.

The present work contributes to a deeper understanding of the role of the highly conserved TL in distinct phase of transcription by using the archaeal RNAP as model enzyme. Thus, we could demonstrate the functional conservation of the TL-dependent fit discrimination mechanism among the three domains of life, while conserved residues A''L83 and H87 have a distinctive contribution in catalysis and fidelity (NTP discrimination, intrinsic cleavage) than what was previously observed in bacterial RNAP. Because of the high sequence similarities and length conservation of the TL between *P. furiosus* RNAP and yeast RNAP II, our results on archaeal TL

functional contributions may be regarded as similar to the one of eukaryotic RNAP II in transcription initiation, elongation catalysis and fidelity.

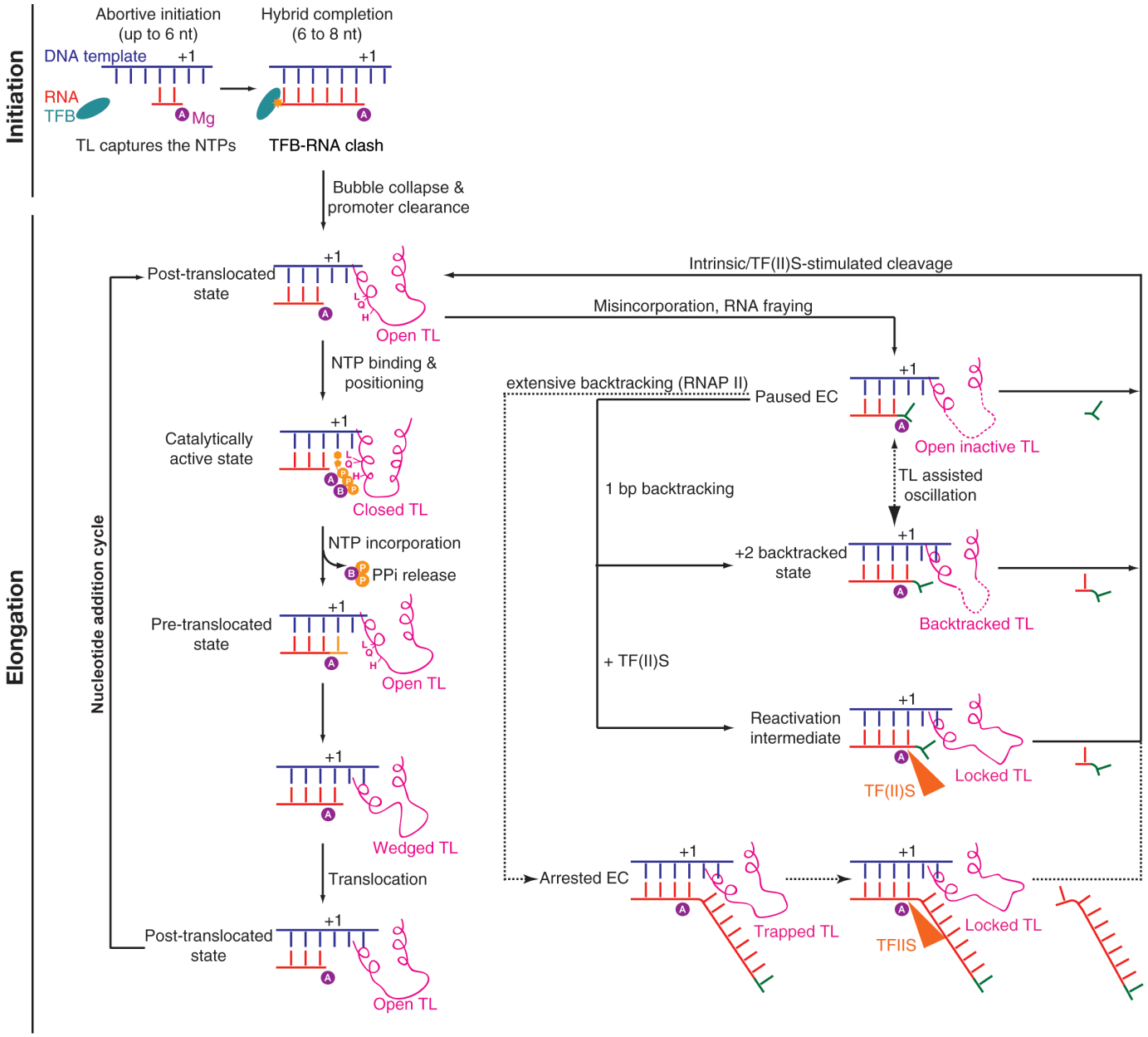

Figure 27. Schematic representations of TL dynamics in distinct transcription phases. The TL is essential for initial synthesis until a complete DNA-RNA hybrid formation (8 nt). The TL is important fidelity of transcription and catalysis during transcription elongation phase. Whereas L83 and H87 contribute to the recognition of the correct NTP, Q80 contributes to the recognition of the 2´OH-group of the NTP. When misincorporation occurs, the TL influences translocation but does not contribute in intrinsic or factor-stimulated RNA cleavage during proofreading.

VI) Bibliography

Adams M.W. (1998). The biochemical diversity of life near and above 100°C in marine environments. J Appl Microbiol. 85 Suppl 1:108S-117S.

Alic, N., Ayoub, N., Landrieux, E., Favry, E., Baudouin-Cornu, P., Riva, M., Carles, C. (2007). Selectivity and proofreading both contribute significantly to the fidelity of RNA polymerase III transcription. Proc Natl Acad Sci U S A. 104(25):10400-5.

Andrecka, J., Treutlein, B., Arcusa, M.A., Muschielok, A., Lewis, R., Cheung, A.C., Cramer, P., Michaelis, J. (2009). Nano positioning system reveals the course of upstream and nontemplate DNA within the RNA polymerase II elongation complex. Nucleic Acids Res. 37(17):5803-9.

Armache, K.J., Kettenberger, H., Cramer, P. (2003). Architecture of initiation-competent 12-subunit RNA polymerase II. Proc Natl Acad Sci U S A. 100(12):6964-8.

Armache, K.J., Mitterweger, S., Meinhart, A., Cramer, P. (2005). Structures of complete RNA polymerase II and its subcomplex, Rpb4/7. J Biol Chem. 280(8):7131-4.

Awrey, D.E., Shimasaki, N., Koth, C., Weilbaecher, R., Olmsted, V., Kazanis, S., Shan, X., Arellano, J., Arrowsmith, C.H., Kane, C.M., Edwards, A.M. (1998). Yeast transcript elongation factor (TFIIS), structure and function. II: RNA polymerase binding, transcript cleavage, and read-through. J Biol Chem. 273(35):22595-605.

Bar-Nahum, G., Epshtein, V., Ruckenstein, A.E., Rafikov, R., Mustaev, A., Nudler, E. (2005). A ratchet mechanism of transcription elongation and its control. Cell. 120(2):183-93.

Bartlett, M. S. (2005). Determinants of transcription initiation by archaeal RNA polymerase. Curr Opin Microbiol 8, 677-684.

Bell, S. D.,and Jackson, S. P. (1998). Transcription and translation in Archaea: a mosaic of eukaryal and bacterial features. Trends Microbiol 6, 222-228.

Bell, S. D., Kosa, P. L., Sigler, P. B., Jackson, S. P. (1999). Orientation of the transcription preinitiation complex in archaea. Proc Natl Acad Sci U S A 96, 13662-13667.

Bell, S. D., Brinkman, A. B., van der Oost, J., Jackson, S. P. (2001). The archaeal TFIIEalpha homologue facilitates transcription initiation by enhancing TATA-box recognition. EMBO Rep 2, 133-138.

Belogurov, G.A., Vassylyeva, M.N., Sevostyanova, A., Appleman, J.R., Xiang, A.X., Lira, R.,

Webber, S.E., Klyuyev, S., Nudler, E., Artsimovitch, I., Vassylyev, D.G. (2009). Transcription inactivation through local refolding of the RNA polymerase structure. Nature. 457(7227):332-5.

Bintrim SB, Donohue TJ, Handelsman J, Roberts GP, Goodman RM. (1997). Molecular phylogeny of Archaea from soil. Proc Natl Acad Sci U S A. 94(1):277-82.

Borukhov, S., Sagitov, V., Goldfarb, A. (1993). Transcript cleavage factors from *E. coli*. Cell. 72(3):459-66.

Blombach, F., Makarova, K.S., Marrero, J., Siebers, B., Koonin, E.V., van der Oost, J. (2009). Identification of an ortholog of the eukaryotic RNA polymerase III subunit RPC34 in Crenarchaeota and Thaumarchaeota suggests specialization of RNA polymerases for coding and non-coding RNAs in Archaea. Biol Direct. 4:39

Braglia, P., Percudani, R., Dieci, G. (2005). Sequence context effects on oligo(dT) termination signal recognition by *Saccharomyces cerevisiae* RNA polymerase III. J Biol Chem. 280(20):19551-62.

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

Briand, J.F., Navarro, F., Rematier, P., Boschiero, C., Labarre, S., Werner, M., Shpakovski, G.V., Thuriaux, P. (2001). Partners of Rpb8p, a small subunit shared by yeast RNA polymerases I, II and III. Mol Cell Biol. 21(17):6056-65.

Brueckner, F.,and Cramer, P. (2008). Structural basis of transcription inhibition by alpha-amanitin and implications for RNA polymerase II translocation. Nat Struct Mol Biol. 15(8):811-8.

Brueckner, F., Ortiz, J., Cramer, P. (2009). A movie of the RNA polymerase nucleotide addition cycle. Curr Opin Struct Biol. 19(3):294-9.

Bushnell, D.A.,and Kornberg, R.D. (2003). Complete, 12-subunit RNA polymerase II at 4.1-A resolution: implications for the initiation of transcription. Proc Natl Acad Sci U S A. 100(12):6969-73.

Bushnell, D. A., Westover, K. D., Davis, R. E., Kornberg, R. D. (2004). Structural basis of transcription: an RNA polymerase II-TFIIB cocrystal at 4.5 Angstroms. Science 303, 983-988.

Campbell, E. A., Westblade, L. F., Darst, S. A. (2008). Regulation of bacterial RNA polymerase sigma factor activity: a structural perspective. Curr Opin Microbiol 11, 121-127.

Carter, R.,and Drouin, G. (2009). Structural differentiation of the three eukaryotic RNA polymerases. Genomics. 94(6):388-96.

Carvalho, A.T.P., Fernandes, P.A., Ramos, M.J. (2011). The catalytic mechanism of RNA polymerase II. J. Chem. Theory Comput. 7:1177-1188.

Castro, C., Smidansky, E.D., Arnold, J.J., Maksimchuk, K.R., Moustafa, I., Uchida, A,, Götte, M., Konigsberg, W., Cameron, C.E. (2009). Nucleic acid polymerases use a general acid for nucleotidyl transfer. Nat. Struct. Mol. Biol. 16:212-218.

Chakraborty, A., Wang, D., Ebright, Y.W., Korlann, Y., Kortkhonjia, E., Kim, T., Chowdhury, S., Wigneshweraraj, S., Irschik, H., Jansen, R., Nixon, B.T., Knight, J., Weiss, S., Ebright, R.H. (2012). Opening and closing of the bacterial RNA polymerase clamp. Science. 337(6094):591-5.

Cheetham, G.M.,and Steitz, T.A. (2000). Insights into transcription: structure and function of single-subunit DNA-dependent RNA polymerases. Curr Opin Struct Biol. 10(1):117-23.

Cheung, A.C.,and Cramer, P. (2011). Structural basis of RNA polymerase II backtracking, arrest and reactivation. Nature. 471(7337):249-53.

Cheung A.C., Sainsbury S., Cramer P. (2011). Structural basis of initial RNA polymerase II transcription. EMBO J. 30(23), 4755-63.

Choder, M.,and Young, R.A. (1993). A portion of RNA polymerase II molecules has a component essential for stress responses and stress survival. Mol Cell Biol. 13(11):6984-91.

Cox, J. M., Hayward, M. M., Sanchez, J. F., Gegnas, L. D., van der Zee, S., Dennis, J. H., Sigler, P. B., Schepartz, A. (1997). Bidirectional binding of the TATA box binding protein to the TATA box. Proc Natl Acad Sci U S A 94, 13475-13480.

Cramer, P. (2002). Multisubunit RNA polymerases. Curr Opin Struct Biol. 12(1):89-97.

Cramer, P. (2004). Structure and function of RNA polymerase II. Adv Protein Chem. 67:1-42.

Cramer, P. (2004). RNA polymerase II structure: from core to functional complexes. Curr Opin Genet Dev. 14(2):218-26.

Cramer, P.,and Arnold, E. (2009). Proteins: how RNA polymerases work. Curr Opin Struct Bio. 19(6):680-2.

Cramer, P., Bushnell, D.A., Fu, J., Gnatt, A.L., Maier-Davis, B., Thompson, N.E., Burgess, R.R., Edwards, A.M., David, P.R., Kornberg, R.D. (2000). Architecture of RNA polymerase II and

implications for the transcription mechanism. Science. 288(5466):640-9.

Cramer, P., Bushnell, D.A., Kornberg, R.D. (2001). Structural basis of transcription: RNA polymerase II at 2.8 angstrom resolution. Science. 292(5523):1863-76.

Creti, R., Londei, P., Cammarano, P. (1993). Complete nucleotide sequence of an archaeal (*Pyrococcus woesei*) gene encoding a homolog of eukaryotic transcription factor IIB (TFIIB). Nucleic Acids Res 21, 2942.

Da, L.T., Wang, D., Huang, X. (2012). Dynamics of pyrophosphate ion release and its coupled trigger loop motion from closed to open state in RNA polymerase II. J. Am. Chem. Soc. 134:2399-2406.

Darst, S.A. (2001) Bacterial RNA polymerase. Curr Opin Struct Biol. (2):155-62.

Dengl, S.,and Cramer, P. (2009). Torpedo nuclease Rat1 is insufficient to terminate RNA polymerase II *in vitro*. J Biol Chem. 284(32):21270-9.

Dvir, A., Conaway, J. W., Conaway, R. C. (2001). Mechanism of transcription initiation and promoter escape by RNA polymerase II. Curr Opin Genet Dev 11, 209-214.

Ebright, R.H. (2000). RNA polymerase: structural similarities between bacterial RNA polymerase and eukaryotic RNA polymerase II. J Mol Biol. 304(5):687-98

Edwards, A.M., Kane, C.M., Young, R.A., Kornberg, R.D. (1991). Two dissociable subunits of yeast RNA polymerase II stimulate the initiation of transcription at a promoter *in vitro*. J Biol Chem. 266(1):71-5.

Epshtein, V., Cardinale, C.J., Ruckenstein, A.E., Borukhov, S., Nudler, E. (2007). An allosteric path to transcription termination. Mol Cell. 28(6):991-1001.

Epshtein, V., Dutta, D., Wade, J., Nudler, E. (2010). An allosteric mechanism of Rho-dependent transcription termination. Nature. 463(7278):245-9.

Erie, D.A., Hajiseyedjavadi, O., Young, M.C., von Hippel, P.H. (1993). Multiple RNA polymerase conformations and GreA: control of the fidelity of transcription. Science. 262(5135):867-73.

Facciotti, M.T., Reiss, D.J., Pan, M., Kaur, A., Vuthoori, M., Bonneau, R., Shannon, P., Srivastava, A., Donohoe, S.M., Hood, L.E., Baliga, N.S. (2007). General transcription factor specified global gene regulation in archaea. Proc Natl Acad Sci U S A. 104(11):4630-5.

Feng, G.H., Lee, D.N., Wang, D., Chan, C.L., Landick, R. (1994). GreA-induced transcript cleavage

in transcription complexes containing *Escherichia coli* RNA polymerase is controlled by multiple factors, including nascent transcript location and structure. J Biol Chem. 269(35):22282-94.

Fiala, G.,and Stetter, K. O. (1986). *Pyrococcus furiosus* sp. nov. represents a novel genus of marine heterotrophic archaebacteria growing optimally at 100°C". Archives of Microbiology 145: 56–61.

Forterre, P., Gribaldo, S., Gadelle, D. and Serre, M. C. (2007). Origin and evolution of DNA topoisomerases. Biochimie 89, 427–446.

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

Gaspari M., Larsson N.G., Gustafsson C.M. (2004). The transcription machinery in mammalian mitochondria. Biochim Biophys Acta. 1659(2-3):148-52.

Ghavi-Helm, Y., Michaut, M., Acker, J., Aude, J.C., Thuriaux, P., Werner, M., Soutourina, J. (2008). Genome-wide location analysis reveals a role of TFIIS in RNA polymerase III transcription. Genes Dev. 22(14):1934-47.

Gnatt, A.L., Cramer, P., Fu, J., Bushnell, D.A., Kornberg, R.D. (2001). Structural basis of transcription: an RNA polymerase II elongation complex at 3.3 A resolution. Science. 292(5523):1876-82.

Goede, B., Naji, S., von Kampen, O., Ilg, K., 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.

Goldman, S.R., Ebright, R.H., Nickels, B.E. (2009). Direct detection of abortive RNA transcripts *in vivo*. Science. 324(5929):927-8.

Grohmann, D., Hirtreiter, A., Werner, F. (2009) Molecular mechanisms of archaeal RNA polymerase. Biochem Soc Trans. Pt 1):12-7.

Grohmann, D., Hirtreiter, A., Werner, F. (2009) RNAP subunits F/E (RPB4/7) are stably associated with archaeal RNA polymerase: using fluorescence anisotropy to monitor RNAP assembly *in vitro*. Biochem J. 421(3):339-43.

Grohmann, D. and Werner, F. (2011). Recent advances in the understanding of archaeal transcription. Curr Opin Microbiol. 14(3):328-34.

Grohmann, D., Nagy, J., Chakraborty, A., Klose, D., Fielden, D., Ebright, R.H., Michaelis, J., Werner, F. (2011). The initiation factor TFE and the elongation factor Spt4/5 compete for the RNAP clamp during transcription initiation and elongation. Mol Cell. 43(2):263-74.

Grünberg, S., Bartlett, M. S., Naji, S., Thomm, M. (2007). Transcription factor E is a part of transcription elongation complexes. J Biol Chem 282, 35482-35490.

Grünberg, S., Warfield, L., Hahn, S. (2012). Architecture of the RNA polymerase II preinitiation complex and mechanism of ATP-dependent promoter opening. Nat Struct Mol Biol. 19(8):788-96.

Gusarov, I.,and Nudler, E. (1999). The mechanism of intrinsic transcription termination. Mol Cell. 3(4):495-504.

Hanzelka, B. L., Darcy, T. J., Reeve, J. N. (2001). TFE, an archaeal transcription factor in *Methanobacterium thermoautotrophicum* related to eucaryal transcription factor TFIIEalpha. J Bacteriol 183, 1813-1818

Hausner, W., Wettach, J., Hethke, C., Thomm, M. (1996). Two transcription factors related with the eucaryal transcription factors TATA-binding protein and transcription factor IIB direct promoter recognition by an archaeal RNA polymerase. J Biol Chem 271, 30144-30148.

Hausner, W., Lange, U., Musfeldt, M. (2000).Transcription factor S, a cleavage induction factor of the archaeal RNA polymerase. J Biol Chem. 275(17):12393-9.

Herbert, K.M., La Porta, A., Wong, B.J., Mooney, R.A., Neuman, K.C., Landick, R., Block, S.M. (2006). Sequence-resolved detection of pausing by single RNA polymerase molecules. Cell. 125(6):1083-94.

Herbert, K.M., Zhou, J., Mooney, R.A., Porta, A.L., Landick, R., Block, S.M. (2010). *E. coli* NusG inhibits backtracking and accelerates pause-free transcription by promoting forward translocation of RNA polymerase. J Mol Biol. 399(1):17-30.

Hethke, C., Geerling, A. C., Hausner, W., de Vos, W. M., Thomm, M. (1996). A cell-free transcription system for the hyperthermophilic archaeon *Pyrococcus furiosus*. Nucleic Acids Res 24, 2369-2376.

Hirata, A., Kanai, T., Santangelo, T.J., Tajiri, M., Manabe, K., Reeve, J.N., Imanaka, T., Murakami, K.S. (2008). 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(3):623-33.

Hirata, A., Klein, B.J., Murakami, K.S. (2008). The X-ray crystal structure of RNA polymerase from Archaea. Nature. 451(7180):851-4.

Hirtreiter, A., Grohmann, D., Werner, F. (2010). Molecular mechanisms of RNA polymerase--the F/E (RPB4/7) complex is required for high processivity *in vitro*. Nucleic Acids Res. 38(2):585-96.

Hirtreiter A., Damsma G.E., Cheung A.C., Klose D., Grohmann D., Vojnic E., Martin A.C., Cramer P., Werner F. (2010). Spt4/5 stimulates transcription elongation through the RNA polymerase clamp coiled-coil motif. Nucleic Acids Res. 38(12):4040-51.

Hogan, B.P., Hartsch, T., Erie, D.A. (2001). Transcript cleavage by *Thermus thermophilus* RNA polymerase. Effects of GreA and anti-GreA factors. J Biol Chem. 277(2):967-75.

Holmes, S.F., Santangelo, T.J., Cunningham, C.K., Roberts, J.W., Erie, D.A. (2006). Kinetic investigation of *Escherichia coli* RNA polymerase mutants that influence nucleotide discrimination and transcription fidelity. J Biol Chem. 281(27):18677-83

Huang, X., Wang, D., Weiss, D.R., Bushnell, D.A., Kornberg, R.D., Levitt, M. (2010). RNA polymerase II trigger loop residues stabilize and position the incoming nucleotide triphosphate in transcription. Proc Natl Acad Sci U S A. 107(36):15745-50.

Huet, J., Schnabel, R., Sentenac, A., Zillig, W. (1983). Archaebacteria and eukaryotes possess DNA-dependent RNA polymerases of a common type. EMBO J. 2(8):1291-4.

Ishino, Y., Komori, K., Cann, I. K. Koga, Y. (1998). A novel DNA polymerase family found in Archaea. J. Bacteriol. 180, 2232–2236.

Izban, M.G.,and Luse, D.S. (1992). The RNA polymerase II ternary complex cleaves the nascent transcript in a 3'----5' direction in the presence of elongation factor SII. Genes Dev. 6(7):1342-56.

Izban, M.G.,and Luse, D.S. (1993). The increment of SII-facilitated transcript cleavage varies dramatically between elongation competent and incompetent RNA polymerase II ternary complexes. J Biol Chem. 268(17):12874-85.

Jasiak, A.J., Armache, K.J., Martens, B., Jansen, R.P., Cramer, P. (2006). Structural biology of RNA polymerase III: subcomplex C17/25 X-ray structure and 11 subunits enzyme model. Mol Cell. 23(1):71-81.

Jeon, C., Yoon, H.,and Agarwal, K. (1994). The transcription factor TFIIS zinc ribbon dipeptide Asp-Glu is critical for stimulation of elongation and RNA cleavage by RNA polymerase II. Proc Natl Acad Sci U S A. 91(19):9106-10.

Kanamaru, K.,and Tanaka, K. (2004). Roles of chloroplast RNA polymerase sigma factors in chloroplast development and stress response in higher plants. Biosci Biotechnol Biochem. 68(11):2215-23.

Kapanidis, A.N., Margeat, E., Ho, S.O., Kortkhonjia, E., Weiss, S., Ebright, R.H. (2006). Initial transcription by RNA polymerase proceeds through a DNA-scrunching mechanism. Science. 314(5802):1144-7.

Kaplan, C.D., Larsson K.M., Kornberg R.D. (2008). The RNA polymerase II trigger loop functions in substrate selection and is directly targeted by alpha-amanitin. Mol Cell. 30(5), 547-56.

Kaplan, C.D., Jin, H., Zhang, I.L., Belyanin, A. (2012). Dissection of Pol II trigger loop function and Pol II activity-dependent control of start site selection *in vivo*. PLoS Genet. 8(4):e1002627.

Kelman, Z.,and White, M. F. (2005). Archaeal DNA replication and repair. Curr Opin Microbiol 8, 669-676.

Kennedy, S.R.,and Erie, D.A. (2011). Templated nucleoside triphosphate binding to a noncatalytic site on RNA polymerase regulates transcription. Proc Natl Acad Sci U S A. 108(15):6079-84.

Kent, T., Kashkina, E., Anikin, M., Temiakov, D. (2009). Maintenance of RNA-DNA hybrid length in bacterial RNA polymerases. J Biol Chem. 284(20):13497-504.

Kettenberger, H., Armache, K-J., Cramer, P. (2003). Architecture of the RNA polymerase II-TFIIS complex and implications for mRNA cleavage. Cell, 114, 347-57.

Kettenberger, H., Armache, K. J., Cramer, P. (2004). Complete RNA polymerase II elongation complex structure and its interactions with NTP and TFIIS. Mol Cell *16*, 955-965.

King, R.A., Markov, D., Sen, R., Severinov, K., Weisberg, R.A. (2004). A conserved zinc binding domain in the largest subunit of DNA-dependent RNA polymerase modulates intrinsic transcription termination and antitermination but does not stabilize the elongation complex. J Mol Biol. 342(4):1143-54.

Kireeva, M. L., Komissarova, N.,and Kashlev, M. (2000). Overextended RNA:DNA hybrid as a negative regulator of RNA polymerase II processivity. J. Mol. Biol. 299, 325-335.

Kireeva, M.L., Nedialkov, Y.A., Cremona, G.H., Purtov, Y.A., Lubkowska, L., Malagon, F., Burton, Z.F., Strathern, J.N., Kashlev, M. (2008) Transient reversal of RNA polymerase II active site closing controls fidelity of transcription elongation. Mol Cell. 30(5), 557-66.

Kireeva, M.L., Kashlev, M., Burton, Z.F. (2010). Translocation by multi-subunit RNA polymerases. Biochim Biophys Acta. 1799(5-6):389-401.

Kireeva, M.L., Domecq, C., Coulombe, B., Burton, Z.F., Kashlev, M. (2011). Interaction of RNA polymerase II fork loop 2 with downstream non-template DNA regulates transcription elongation. J Biol Chem. 286(35):30898-910.

Klein, B.J., Bose, D., Baker, K.J., Yusoff, Z.M., Zhang, X., Murakami, K.S. (2010). RNA polymerase and transcription elongation factor Spt4/5 complex structure. Proc Natl Acad Sci U S A. 108(2):546-50.

Koonin, E.V., Makarova, K.S., Elkins, J.G. (2007) Orthologs of the small RPB8 subunit of the eukaryotic RNA polymerases are conserved in hyperthermophilic Crenarchaeota and "Korarchaeota". Biol Direct. 2:38.

Korkhin, Y., Unligil, U.M., Littlefield, O., Nelson, P.J., Stuart, D.I., Sigler, P.B., Bell, S.D., Abrescia, N.G. (2009). Evolution of complex RNA polymerases: the complete archaeal RNA polymerase structure. PLoS Biol. 7(5):e1000102.

Korzheva, N., Mustaev, A., Kozlov, M., Malhotra, A., Nikiforov, V., Goldfarb, A., Darst, S.A. (2000). A structural model of transcription elongation. Science. 289(5479):619-25.

Kosa, P. F., Ghosh, G., DeDecker, B. S., Sigler, P. B. (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 U S A 94, 6042-6047.

Kostrewa, D., Zeller, M. E., Armache, K. J., Seizl, M., Leike, K., Thomm, M., Cramer, P. (2009). RNA polymerase II-TFIIB structure and mechanism of transcription initiation. Nature *462*, 323-330.

Koulich, D., Orlova, M., Malhotra, A., Sali, A., Darst, S.A., Borukhov, S. (1997). Domain organization of *Escherichia coli* transcript cleavage factors GreA and GreB. J Biol Chem. 272(11):7201-10.

Koyama, H., Ito, T., Nakanishi, T., Kawamura, N., Sekimizu, K. (2003). Transcription elongation factor S-II maintains transcriptional fidelity and confers oxidative stress resistance. Genes Cells. 8(10):779-88.

Koyama, H., Ito, T., Nakanishi, T., Sekimizu, K. (2007). Stimulation of RNA polymerase II transcript cleavage activity contributes to maintain transcriptional fidelity in yeast. Genes Cells. 2007 May;12(5):547-59.

Kuhn, C.D., Geiger, S.R., Baumli, S., Gartmann, M., Gerber, J., Jennebach, S., Mielke, T., Tschochner, H., Beckmann, R., Cramer, P. (2007). Functional architecture of RNA polymerase I. Cell. 131(7):1260-72.

Kulish, D., Lee, J., Lomakin, I., Nowicka, B., Das, A., Darst, S., Normet, K., Borukhov, S. (2000). The functional role of basic patch, a structural element of *Escherichia coli* transcript cleavage factors GreA and GreB. J Biol Chem. 275(17):12789-98.

Kusser, A. G., Bertero, M. G., Naji, S., Becker, T., Thomm, M., Beckmann, R., Cramer, P. (2008). Structure of an archaeal RNA polymerase. J Mol Biol 376, 303-307.

Kuznedelov, K., Korzheva, N., Mustaev, A., Severinov, K. (2002). Structure-based analysis of RNA polymerase function: the largest subunit's rudder contributes critically to elongation complex stability and is not involved in the maintenance of RNA-DNA hybrid length. EMBO J. 21(6):1369-78.

Kwapisz, M., Beckouët, F., Thuriaux, P. (2008). Early evolution of eukaryotic DNA-dependent RNA polymerases. Trends Genet. 24(5):211-5.

Kyrpides, N. C.,and Ouzounis, C. A. (1999). Transcription in archaea. Proc Natl Acad Sci U S A 96, 8545-8550.

Labhart, P.,and Morgan, G.T. (1998). Identification of novel genes encoding transcription elongation factor TFIIS (TCEA) in vertebrates: conservation of three distinct TFIIS isoforms in frog, mouse, and human. Genomics. 52(3):278-88.

Lagrange, T., Kapanidis, A. N., Tang, H., Reinberg, D., Ebright, R. H. (1998). New core promoter element in RNA polymerase II-dependent transcription: sequence-specific DNA binding by transcription factor IIB. Genes Dev 12, 34-44.

Lamour, V., Hogan, B.P., Erie, D.A., Darst, S.A. (2006). Crystal structure of *Thermus aquaticus* Gfh1, a Gre-factor paralog that inhibits rather than stimulates transcript cleavage. J Mol Biol. 356(1):179-88.

Landick, R. (2004). Active-site dynamics in RNA polymerases. Cell. 116(3):351-3.

Lange, U.,and Hausner, W. (2004). Transcriptional fidelity and proofreading in Archaea and implications for the mechanism of TFS-induced RNA cleavage. Mol Microbiol. 52(4):1133-43.

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(9):2251.

Langer, D., Hain, J., Thuriaux, P., Zillig, W. (1995). Transcription in archaea: similarity to that in eucarya. Proc Natl Acad Sci U S A 92, 5768-5772.

Laptenko, O., Kim, S.S., Lee, J., Starodubtseva, M., Cava, F., Berenguer, J., Kong, X.P., Borukhov, S. (2006). pH-dependent conformational switch activates the inhibitor of transcription elongation. EMBO J. 25(10):2131-41.

Laptenko, O., Lee, J., Lomakin, I., Borukhov, S. (2003). Transcript cleavage factors GreA and GreB act as transient catalytic components of RNA polymerase. EMBO J. 22(23):6322-34.

Larson, M.H., Zhou, J., Kaplan, C.D., Palangat, M., Kornberg, R.D., Landick, R., Block, S.M. (2012). Trigger loop dynamics mediate the balance between the transcriptional fidelity and speed of RNA polymerase II. Proc Natl Acad Sci U S A. 109(17):6555-60.

Leininger, S., Urich, T., Schloter, M., Schwark, L., Qi, J., Nicol, G.W., Prosser, J.I., Schuster, S.C., Schleper, C. (2006). Archaea predominate among ammonia-oxidizing prokaryotes in soils. Nature 442 (7104): 806–9.

Majovski, R.C., Khaperskyy, D.A., Ghazy, M.A., Ponticelli, A.S. (2005). A functional role for the switch 2 region of yeast RNA polymerase II in transcription start site utilization and abortive initiation. J Biol Chem. 280(41):34917-23.

Malinen, A.M., Turtola, M., Parthiban, M., Vainonen, L., Johnson, M.S., Belogurov, G.A. (2012). Active site opening and closure control translocation of multisubunit RNA polymerase. Nucleic Acids Res. 40(15):7442-51.

Martinez-Rucobo, F.W., Sainsbury, S., Cheung, A.C., Cramer, P. (2011). Architecture of the RNA polymerase-Spt4/5 complex and basis of universal transcription processivity. EMBO J. 30(7):1302-10.

Martinez-Rucobo, F.W., Cramer, P. (2013). Structural basis of transcription elongation. Biochim Biophys Acta. 1829(1):9-19.

Matarazzo F, Ribeiro AC, Faveri M, Taddei C, Martinez MB, Mayer MP. (2012). The domain

Archaea in human mucosal surfaces. Clin Microbiol Infect. 18(9):834-40.

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

Meyer, P.A., Ye, P., Suh, M.H., Zhang, M., Fu, J. (2009). Structure of the 12-subunit RNA polymerase II refined with the aid of anomalous diffraction data. J Biol Chem. 284(19):12933-9.

Miller, T. L., M. J. Wolin, E. C. de Macario, A. J. Macario. **(**1982). Isolation of *Methanobrevibacter smithii* from human feces. Appl. Environ. Microbiol. 43**:**227-232.

Micorescu, M., Grünberg, S., Franke, A., Cramer, P., Thomm, M., Bartlett, M. (2008). Archaeal transcription: function of an alternative transcription factor B from *Pyrococcus furiosus*. J Bacteriol. 190(1):157-67.

Minakhin, L**.**, Bhagat, S., Brunning, A., Campbell, E.A., Darst, S.A., Ebright, R.H., Severinov, K. (2001). Bacterial RNA polymerase subunit omega and eukaryotic RNA polymerase subunit RPB6 are sequence, structural, and functional homologs and promote RNA polymerase assembly**.** Proc Natl Acad Sci U S A. 98(3):892-7.

Miropolskaya, N., Artsimovitch, I., Klimasauskas, S., Nikiforov, V., Kulbachinskiy, A. (2009). Allosteric control of catalysis by the F loop of RNA polymerase. Proc Natl Acad Sci U S A. 106(45):18942-7

Miropolskaya, N., Nikiforov, V., Klimasauskas, S., Artsimovitch, I., Kulbachinskiy, A. (2010). Modulation of RNA polymerase activity through the trigger loop folding. Transcription. 1(2):89-94.

Morin, P.E., Awrey, D.E., Edwards, A.M., Arrowsmith, C.H. (1996). Elongation factor TFIIS contains three structural domains: solution structure of domain II. Proc Natl Acad Sci U S A. 93(20):10604-8.

Mukhopadhyay, J., Das, K., Ismail, S., Koppstein, D., Jang, M., Hudson, B., Sarafianos, S., Tuske, S., Patel, J., Jansen, R., Irschik, H., Arnold, E., Ebright, R.H. (2008). The RNA polymerase "switch region" is a target for inhibitors. Cell. 135(2):295-307.

Murakami, K. S.,and Darst, S. A. (2003). Bacterial RNA polymerases: the wholo story. Curr Opin Struct Biol 13, 31-39.

Murakami, K.S., Masuda, S., Campbell, E.A., Muzzin, O., Darst, S.A. (2002). Structural basis of transcription initiation: an RNA polymerase holoenzyme-DNA complex. Science. 296(5571):1285-90.

Naji, S., Bertero, M. G., Spitalny, P., Cramer, P., Thomm, M. (2008). Structure-function analysis of the RNA polymerase cleft loops elucidates initial transcription, DNA unwinding and RNA displacement. Nucleic Acids Res 36, 676-687.

Naji, S., Grünberg, S., 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.

Nakanishi, T., Shimoaraiso, M., Kubo, T., Natori, S. (1995). Structure-function relationship of yeast S-II in terms of stimulation of RNA polymerase II, arrest relief, and suppression of 6-azauracil sensitivity. J Biol Chem. 270(15):8991-5.

Naryshkina, T., Kuznedelov, K., Severinov, K. (2006). The role of the largest RNA polymerase subunit lid element in preventing the formation of extended RNA-DNA hybrid. J Mol Biol. 361(4):634-43.

Nikolov, D. B., Hu, S. H., Lin, J., Gasch, A., Hoffmann, A., Horikoshi, M., Chua, N. H., Roeder, R. G., Burley, S. K. (1992). Crystal structure of TFIID TATA-box binding protein. Nature 360, 40-46.

Nudler, E. (2009). RNA polymerase active center: the molecular engine of transcription. Annu Rev Biochem. 78:335-61.

Olmsted, V.K., Awrey, D.E., Koth, C., Shan, X., Morin, P.E., Kazanis, S., Edwards, A.M., Arrowsmith, C.H. (1998). Yeast transcript elongation factor (TFIIS), structure and function. I: NMR structural analysis of the minimal transcriptionally active region. J Biol Chem. 273(35):22589-94.

Opalka, N., Chlenov, M., Chacon, P., Rice, W.J., Wriggers, W., Darst, S.A. (2003). Structure and function of the transcription elongation factor GreB bound to bacterial RNA polymerase. Cell. 114(3):335-45.

Orlova, M., Newlands, J., Das, A., Goldfarb, A., Borukhov, S. (1995). Intrinsic transcript cleavage activity of RNA polymerase. Proc Natl Acad Sci U S A. 92(10):4596-600.

Ouhammouch, M., Werner, F., Weinzierl, R. O., Geiduschek, E. P. (2004). A fully recombinant system for activator-dependent archaeal transcription. J Biol Chem *279*, 51719-51721.

Ouzounis, C.,and Sander, C. (1992). TFIIB, an evolutionary link between the transcription machineries of archaebacteria and eukaryotes. Cell 71, 189-190.

Palangat, M., Grass, J.A., Langelier, M.F., Coulombe, B., Landick, R. (2011). The RPB2 flap loop of human RNA polymerase II is dispensable for transcription initiation and elongation. Mol Cell Biol. 31(16):3312-25.

Pikaard, C.S., Haag, J.R., Ream, T., Wierzbicki, A.T. (2008). Roles of RNA polymerase IV in gene silencing. Trends Plant Sci. 13(7):390-7.

Pikaard, C.S.,and Tucker, S. (2009). RNA-silencing enzymes Pol IV and Pol V in maize: more than one flavor? PLoS Genet. 5(11):e1000736.

Pühler, G., Leffers, H., Gropp, F., Palm, P., Klenk, H.P., Lottspeich, F., Garrett, R.A., Zillig, W. (1989). Archaebacterial DNA-dependent RNA polymerases testify to the evolution of the eukaryotic nuclear genome. Proc Natl Acad Sci U S A. 86(12):4569-73.

Pupov, D., Miropolskaya, N., Sevostyanova, A., Bass, I., Artsimovitch, I., Kulbachinskiy, A. (2010). Multiple roles of the RNA polymerase {beta}' SW2 region in transcription initiation, promoter escape, and RNA elongation. Nucleic Acids Res. 38(17):5784-96.

Qian, X., Jeon, C., Yoon, H., Agarwal, K., Weiss, M.A. (1993). Structure of a new nucleic-acid-binding motif in eukaryotic transcriptional elongation factor TFIIS. Nature. 365(6443):277-9.

Qureshi, S. A., Baumann, P., Rowlands, T., Khoo, B., Jackson, S. P. (1995). Cloning and functional analysis of the TATA binding protein from Sulfolobus shibatae. Nucleic Acids Res 23, 1775-1781.

Qureshi, S. A.,and Jackson, S. P. (1998). Sequence-specific DNA binding by the S. shibatae TFIIB homolog, TFB, and its effect on promoter strength. Mol Cell 1, 389-400.

Ream, T.S., Haag, J.R., Wierzbicki, A.T., Nicora, C.D., Norbeck, A.D., Zhu, J.K., Hagen, G., Guilfoyle, T.J., Pasa-Tolić, L., Pikaard, C.S. (2009). Subunit compositions of the RNA-silencing enzymes Pol IV and Pol V reveal their origins as specialized forms of RNA polymerase II. Mol Cell. 33(2):192-203.

Robb, F. T., Maeder, D. L., Brown, J. R., DiRuggiero, J., Stump, M. D., Yeh, R. K., Weiss, R. B., Dunn, D. M. (2001). Genomic sequence of hyperthermophile, *Pyrococcus furiosus*: implications for physiology and enzymology. Methods Enzymol 330, 134-157.

Roghanian, M., Yuzenkova, Y., Zenkin, N. (2011). Controlled interplay between trigger loop and Gre factor in the RNA polymerase active centre. Nucleic Acids Res. 39(10):4352-9.

Ruan, W., Lehmann, E., Thomm, M., Kostrewa, D., Cramer, P. (2011). Evolution of two modes of intrinsic RNA polymerase transcript cleavage. J Biol Chem. 286(21):18701-7.

Rudd, M.D., Izban, M.G., Luse, D.S. (1994). The active site of RNA polymerase II participates in transcript cleavage within arrested ternary complexes. Proc Natl Acad Sci U S A. 91(17):8057-61.

Rudd, M.D.,and Luse, D.S. (1996). Amanitin greatly reduces the rate of transcription by RNA polymerase II ternary complexes but fails to inhibit some transcript cleavage modes. J Biol Chem. 271(35):21549-58.

Sainsbury, S., Niesser, J., Cramer, P. (2013)Structure and function of the initially transcribing RNA polymerase II-TFIIB complex. Nature. 493:437-440.

Santangelo, T.J.,and Artsimovitch, I. (2011). Termination and antitermination: RNA polymerase runs a stop sign. Nat Rev Microbiol. 9(5):319-29.

Santangelo, T.J., Cubonová, L., Skinner, K.M., Reeve, J.N. (2009). Archaeal intrinsic transcription termination *in vivo*. J Bacteriol. 191(22):7102-8.

Santangelo, T.J.,and Reeve, J.N. (2006). Archaeal RNA polymerase is sensitive to intrinsic termination directed by transcribed and remote sequences. J Mol Biol. 355(2):196-210.

Santangelo, T.J.,and Reeve J.N. (2010). Deletion of switch 3 results in an archaeal RNA polymerase that is defective in transcript elongation. J Biol Chem. 285(31):23908-15.

Schnapp, G., Graveley, B.R., Grummt, I. (1996). TFIIS binds to mouse RNA polymerase I and stimulates transcript elongation and hydrolytic cleavage of nascent rRNA. Mol Gen Genet. 252(4):412-9.

Sekine, S., Tagami, S., Yokoyama, S. (2012). Structural basis of transcription by bacterial and eukaryotic RNA polymerases. Curr Opin Struct Biol. 22(1):110-8.

Sevostyanova, A.,and Artsimovitch, I. (2010). Functional analysis of *Thermus thermophilus* transcription factor NusG. Nucleic Acids Res. 38(21):7432-45.

Sheffer, A., Varon, M., Choder, M. (1999). Rpb7 can interact with RNA polymerase II and support transcription during some stresses independently of Rpb4. Mol Cell Biol. 19(4):2672-80.

Sidorenkov, I., Komissarova, N., Kashlev, M. (1998). Crucial role of the RNA:DNA hybrid in the processivity of transcription. Mol Cell. 2(1):55-64.

Sosunov, V., Sosunova, E., Mustaev, A., Bass, I., Nikiforov, V., Goldfarb, A. (2003). Unified two-metal mechanism of RNA synthesis and degradation by RNA polymerase. EMBO J. 22(9):2234-44.

Sosunova, E., Sosunov, V., Epshtein, V., Nikiforov, V., Mustaev, A. (2013). Control of transcriptional fidelity by active center tuning as derived from RNA polymerase endonuclease reaction. J Biol Chem. 288(9):6688-703.

Sosunova, E., Sosunov, V., Kozlov, M., Nikiforov, V., Goldfarb, A., Mustaev, A. (2003). Donation of catalytic residues to RNA polymerase active center by transcription factor Gre. Proc Natl Acad Sci U S A. 100(26):15469-74

Sousa, R. (2005). Machinations of a maxwellian demon. Cell. 120(2):155-6.

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, 30497-505.

Spitalny, P.,and Thomm, M. (2008). A polymerase III-like reinitiation mechanism is operating in regulation of histone expression in archaea. Mol Microbiol. 67(5):958-70.

Srivastava, A., Talaue, M., Liu, S., Degen, D., Ebright, R.Y., Sineva, E., Chakraborty, A., Druzhinin, S.Y., Chatterjee, S., Mukhopadhyay, J., Ebright, Y.W., Zozula, A., Shen, J., Sengupta, S., Niedfeldt, R.R., Xin, C., Kaneko, T., Irschik, H., Jansen, R., Donadio, S., Connell, N., Ebright, R.H. (2011). New target for inhibition of bacterial RNA polymerase: 'switch region'. Curr Opin Microbiol. 14(5):532-43.

Stebbins, C. E., Borukhov, S., Orlova, M., Polyakov, A., Goldfarb, A., Darst, S.A. (1995). Crystal structure of the GreA transcript cleavage factor from *Escherichia coli.* Nature. 373(6515):636-40.

Steitz, T.A. (1998). A mechanism for all polymerases. Nature. 391(6664):231-2.

Steitz, T.A. (2009). The structural changes of T7 RNA polymerase from transcription initiation to elongation. Curr Opin Struct Biol. 2009 Dec;19(6):683-90.

Surratt, C.K., Milan, S.C., Chamberlin, M.J. (1991). Spontaneous cleavage of RNA in ternary complexes of *Escherichia coli* RNA polymerase and its significance for the mechanism of transcription. Proc Natl Acad Sci U S A. 88(18):7983-7.

Svetlov, V., Vassylyev, D.G., Artsimovitch, I. (2004). Discrimination against deoxyribonucleotide

substrates by bacterial RNA polymerase. J. Biol. Chem. 279:38087-38090.

Sweetser, D., Nonet, M., Young, R.A. (1987). Prokaryotic and eukaryotic RNA polymerases have homologous core subunits. Proc Natl Acad Sci U S A. 84(5):1192-6.

Sydow, J.F., Brueckner, F., Cheung, A.C., Damsma, G.E., Dengl, S., Lehmann, E., Vassylyev, D., Cramer, P. (2009). Structural basis of transcription: mismatch-specific fidelity mechanisms and paused RNA polymerase II with frayed RNA. Mol Cell. 34(6):710-21.

Sydow, J. F.,and Cramer, P. (2009). RNA polymerase fidelity and transcriptional proofreading. Curr Opin Struct Biol *19*, 732-739.

Symersky, J., Perederina, A., Vassylyeva, M.N., Svetlov, V., Artsimovitch, I., Vassylyev, D.G. (2006). Regulation through the RNA polymerase secondary channel. Structural and functional variability of the coiled-coil transcription factors. J Biol Chem. 281(3):1309-12.

Tan, L., Wiesler, S., Trzaska, D., Carney, H.C., Weinzierl, R.O. (2008). Bridge helix and trigger loop perturbations generate superactive RNA polymerases. J Biol. 7(10):40.

Tagami, S., Sekine, S., Kumarevel, T., Hino, N., Murayama, Y., Kamegamori, S., Yamamoto, M., Sakamoto, K., Yokoyama, S. (2010). Crystal structure of bacterial RNA polymerase bound with a transcription inhibitor protein. Nature, 468(7326), 978-82.

Temiakov, D., Zenkin, N., Vassylyeva, M.N., Perederina, A., Tahirov, T.H., Kashkina, E., Savkina, M., Zorov, S., Nikiforov, V., Igarashi, N., Matsugaki, N., Wakatsuki, S., Severinov, K.,Vassylyev, D.G. (2005). Structural basis of transcription inhibition by antibiotic streptolydigin. Mol Cell. 19(5):655-66.

Thomas, M.S., Zou, C., Ishihama, A., Glass, R.E. (1997). The effect of a nested set of C-terminal substituted deletions on the function of the alpha subunit of *Escherichia coli* RNA polymerase. Int J Biochem Cell Biol. 29(12):1475-83.

Toulokhonov, I.,and Landick, R. (2003). The flap domain is required for pause RNA hairpin inhibition of catalysis by RNA polymerase and can modulate intrinsic termination. Mol Cell. 12(5):1125-36.

Thuillier, V., Brun, I., Sentenac, A., Werner, M. (1996). Mutations in the alpha-amanitin conserved domain of the largest subunit of yeast RNA polymerase III affect pausing, RNA cleavage and transcriptional transitions. EMBO J. 15(3):618-29.

Toulokhonov, I.,and Landick, R. (2006). The role of the lid element in transcription by *E. coli* RNA

polymerase. J Mol Biol. 361(4):644-58.

Toulokhonov, I., Zhang, J., Palangat, M., Landick, R. (2007). A central role of the RNA polymerase trigger loop in active-site rearrangement during transcriptional pausing. Mol Cell. 27(3):406-19.

Tuske, S., Sarafianos, S.G., Wang, X., Hudson, B., Sineva, E., Mukhopadhyay, J., Birktoft, J.J., Leroy, O., Ismail, S., Clark, A.D. Jr, Dharia, C., Napoli, A., Laptenko, O., Lee, J., Borukhov, S., Ebright, R.H., Arnold, E. (2005). Inhibition of bacterial RNA polymerase by streptolydigin: stabilization of a straight-bridge-helix active-center conformation. Cell. 122(4):541-52.

Ujvári, A.,and Luse, D.S. (2006). RNA emerging from the active site of RNA polymerase II interacts with the Rpb7 subunit. Nat Struct Mol Biol. 13(1):49-54.

Vassylyev, D.G., Sekine, S., Laptenko, O., Lee, J., Vassylyeva, M.N., Borukhov, S., Yokoyama, S. (2002). Crystal structure of a bacterial RNA polymerase holoenzyme at 2.6 A resolution. Nature. 417(6890):712-9.

Vassylyev, D.G., Vassylyeva, M.N., Perederina, A., Tahirov, T.H., Artsimovitch, I. (2007). Structural basis for transcription elongation by bacterial RNA polymerase. Nature. 448(7150):157-62.

Vassylyev, D.G., Vassylyeva, M.N., Zhang, J., Palangat, M., Artsimovitch, I., Landick, R. (2007). Structural basis for substrate loading in bacterial RNA polymerase. Nature. 448(7150), 163-8.

Vassylyeva, M.N., Svetlov, V., Dearborn, A.D., Klyuyev, S., Artsimovitch, I., Vassylyev, D.G. (2007). The carboxy-terminal coiled-coil of the RNA polymerase beta'-subunit is the main binding site for Gre factors. EMBO Rep. 8(11):1038-43.

Waege, I., Schmid, G., Thumann, S., Thomm, M., Hausner, W. (2010). A Shuttle Vector Based Transformation System for *Pyrococcus furiosus*. Appl Environ Microbiol. PMID: 20363792

Walmacq, C., Kireeva, M.L., Irvin, J., Nedialkov, Y., Lubkowska, L., Malagon, F., Strathern, J.N., Kashlev, M. (2009). Rpb9 subunit controls transcription fidelity by delaying NTP sequestration in RNA polymerase II. J Biol Chem. 284(29):19601-12.

Wang, D., Bushnell, D.A., Westover, K.D., Kaplan, C.D., Kornberg, R.D. (2006). Structural basis of transcription: role of the trigger loop in substrate specificity and catalysis. Cell, 127, 941-53.

Wang, D., Bushnell, D.A., Huang, X., Westover, K.D., Levitt, M., Kornberg, R.D. (2009). Structural basis of transcription: backtracked RNA polymerase II at 3.4 angstrom resolution. Science, 324(5931), 1203-6.

Wang, W., Carey, M., Gralla, J.D. (1992). Polymerase II promoter activation: closed complex formation and ATP-driven start site opening. Science. 255(5043):450-3.

Weilbaecher, R.G., Awrey, D.E., Edwards, A.M., Kane, C.M. (2003). Intrinsic transcript cleavage in yeast RNA polymerase II elongation complexes. J Biol Chem. 278(26):24189-99.

Weinzierl, R.O. (2010). The nucleotide addition cycle of RNA polymerase is controlled by two molecular hinges in the Bridge Helix domain. BMC Biol. 8:134.

Weinzierl, R.O. (2011). The Bridge Helix of RNA polymerase acts as a central nanomechanical switchboard for coordinating catalysis and substrate movement. Archaea. 2011:608385.

Werner, F. (2008). Structural evolution of multisubunit RNA polymerases. Trends Microbiol. 16(6):247-50

Werner, F. (2012). A nexus for gene expression-molecular mechanisms of Spt5 and NusG in the three domains of life. J Mol Biol. 417(1-2):13-27.

Werner, F.,and Grohmann, D. (2011). Evolution of multisubunit RNA polymerase in the three domains of life. Nat Rev Microbiol. 9(2):85-98.

Werner, F.,and Weinzierl, R. O. (2002). A recombinant RNA polymerase II-like enzyme capable of promoter-specific transcription. Mol Cell *10*, 635-646.

Werner, F.,and Weinzierl, R.O. (2005). Direct modulation of RNA polymerase core functions by basal transcription factors. Mol Cell Biol. 25(18):8344-55.

West, S., Gromak, N., Proudfoot, N.J. (2004). Human 5' --> 3' exonuclease Xrn2 promotes transcription termination at co-transcriptional cleavage sites. Nature. 432(7016):522-5.

Westover, K.D., Bushnell, D.A., Kornberg, R.D.(2004). Structural basis of transcription: separation of RNA from DNA by RNA polymerase II. Science. 303(5660):1014-6.

Westover, K.D., Bushnell, D.A., Kornberg, R.D.(2004). Structural basis of transcription: nucleotide selection by rotation in the RNA polymerase II active center. Cell. 119(4):481-9.

Wierzbicki, A.T, Haag, J.R., Pikaard, C.S. (2008). Noncoding transcription by RNA polymerase Pol IVb/Pol V mediates transcriptional silencing of overlapping and adjacent genes. Cell. 135(4):635-48.

Wiesler, S.C., Burrows, P.C., Buck, M. (2012). A dual switch controls bacterial enhancer-dependent transcription. Nucleic Acids Res. 40(21):10878-92

Woese, C. R., Kandler, O.,and Wheelis, M. L. (1990). Towards a natural system of organisms: proposal for the domains Archaea, Bacteria, and Eucarya. Proc Natl Acad Sci U S A 87, 4576-4579.

Wojtas, M.N., Mogni, M., Millet, O., Bell, S.D., Abrescia, N.G. (2012). Structural and functional analyses of the interaction of archaeal RNA polymerase with DNA. Nucleic Acids Res. 40(19):9941-52.

Wooddell, C.I.,and Burgess, R.R. (2000). Topology of yeast RNA polymerase II subunits in transcription elongation complexes studied by photoaffinity cross-linking. Biochemistry. 39(44):13405-21.

Woychik, N. A.,and Hampsey, M. (2002). The RNA polymerase II machinery: structure illuminates function. Cell 108, 453-463.

Yuzenkova, Y., Bochkareva A., Tadigotla V.R., Roghanian M., Zorov S., Severinov K., Zenkin N. (2010). Stepwise mechanism for transcription fidelity. BMC Biol. 8:54.

Yuzenkova, Y., Tadigotla, V.R., Severinov, K., Zenkin, N.(2011). A new basal promoter element recognized by RNA polymerase core enzyme. EMBO J. 30(18):3766-75

Yuzenkova, Y.,and Zenkin, N. (2010). Central role of the RNA polymerase trigger loop in intrinsic RNA hydrolysis. Proc Natl Acad Sci U S A.107,10878-83.

Zenkin, N., Yuzenkova, Y., Severinov, K. (2006). Transcript-assisted transcriptional proofreading. Science. 313(5786):518-20.

Zhang, G., Campbell, E.A., Minakhin, L., Richter, C., Severinov, K., Darst, S.A. (1999). Crystal structure of *Thermus aquaticus* core RNA polymerase at 3.3 A resolution. Cell. 98(6):811-24.

Zhang, J., Palangat, M., Landick, R. (2010). Role of the RNA polymerase trigger loop in catalysis and pausing. Nat Struct Mol Biol. 17(1):99-104.

Zhu, W., Zeng, Q., Colangelo, C. M., Lewis, M., Summers, M. F., Scott, R. A. (1996). The N-terminal domain of TFIIB from *Pyrococcus furiosus* forms a zinc ribbon. Nat Struct Biol 3, 122-124.

Erklärung

Hiermit erkläre ich, dass ich die vorliegende Arbeit selbstständig und ohne fremde Hilfe verfasst und nur die von mir angegebenen Quellen und Hilfsmittel verwendet habe. Diese Arbeit war bisher noch nicht Bestandteil eines Prüfungsverfahrens, andere Promotionsversuche wurden nicht unternommen.

Teilergebnisse dieser Arbeit sind bereits veröffentlicht worden:

Fouqueau, T., Zeller, M. E., Cheung, A. C., Cramer, P., Thomm, M. (2013). The RNA polymerase trigger loop functions in all three phases of the transcription cycle. Nucleic Acids Res. doi:10.1093/nar/gkt433 (in press).

Regensburg, im August 2013

VII) Appendix

A. Abbreviations

ATP	Adenosine Triphosphate
APS	Ammonium Persulfate
BH	Bridge Helix
bp	Base pair
BRE	B recognition element
BSA	Bovine Serum Albumin
CTD	C-terminal Domain
CTP	Cytidine Triphosphate
CV	Column Volume
DNA	Deoxyribonucleic Acid
dNTP	Deoxyribonucleoside Triphosphate
DTT	Dithiothreitol
EC	Elongation Complex
EDTA	Ethylenediaminetetraacetic acid
EMSA	Electrophoretic Mobility Shift Assay
gdh	Glutamate dehydrogenase
GTIFs	General Transcription Initiations Factors
GTP	Guanosine Triphosphate
HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
kDa	Kilodalton
MEC	Mismatched Elongation Complex
NAC	Nucleotide Addition Cycle
NMP	Nucleoside Monophosphate
nt	Nucleotide
NTP	Nucleoside Triphosphate
OD	Optical Density
PCR	Polymerase Chain Reaction
PIC	Pre-initiation Complex
PMSF	Phenylmethylsulfonylfluorid
PNK	Polynucleotide kinase
Pol I, II, III	RNA Polymerase I, II, III
RNA	Ribonucleic Acid
RNAP	RNA Polymerase
rpm	Rotation per minute
sec	Second
SDS	Sodium Dodecyl Sulfate
TEMED	Tetramethylethylenediamine

TBP	TATA-box Binding Protein
TEC	Transcription Elongation Complex
TL	Trigger Loop
TF(II)B	Transcription Factor (II)B
TF(II)E	Transcription Factor (II)E
TF(II)S	Transcription Factor (II)S
UV	Ultraviolet
WT	Wild Type

B. Supplemental figures

Figure S1

Representative gels of (**A**) cATP (100 µM) incorporation, (**B**) ncUTP (1 mM) misincorporation and (**C**) c2'dATP (100 µM) incorporation by WT and mutant RNAPs are shown, respectively. Schematic representation above the gel pictures describes the (mis)incorporation of (**A**) cATP, (**B**) ncUTP or (**C**) c2'dATP on the EC(A) scaffold template. The red asterisk indicates that the RNA is ^{32}P-labeled at
the 5' end.

Figure S2

Representative gels of intrinsic phosphodiester bond hydrolysis in (**A**) MEC(G) and (**B**) MEC(C) and scaffold templates by mutant RNAPs. Schematic representation above the gel pictures describes the reactions. The red asterisk indicates that the RNA is ^{32}P-labeled at the 5' end.

Figure S3

Kinetics of the intrinsic RNA cleavage reaction in (**A**) MEC(G) and (**B**) MEC(C) scaffold templates by WT, A'' H87A, ΔTLtip and ΔTL RNAPs. Solid curves are the single exponential fits of the kinetics data. The reactions were performed in triplicates.

Figure S4

Representative gels of TFS induced phosphodiester bond hydrolysis in (**A**) EC(U) and (**B**) MEC(G) scaffold templates by WT and mutant RNAPs are shown. Black asterisks indicate nonspecific RNA degradation products. Schematic representation above the gel pictures describes the reactions. The red asterisk indicates that the RNA is ^{32}P-labeled at the 5' end. The arrow indicates the RNA
products cleaved at the terminal 3'-phosphodiester bond.

Figure S5

Transcription elongation assays by WT and ΔTL RNAPs were performed at 70, 80 and 90°C, with (**A**) EC(U) and (**B**) ECTerm templates, in the presence of 100 µM NTPs. Samples were incubated for 5 min with WT RNAP and for 15 min with ΔTL RNAP. (**C**) Transcription elongation kinetics on the ECTerm scaffold template in the presence (+) or absence (-) of the non-template DNA strand (NT) and/or E'F subunits with WT and ΔTL RNAPs. (**D**) Transcription elongation was assayed between 80 to 94°C in 1°C increments on EC(U) and ECTerm scaffold templates. Samples were incubated for
5 min with WT RNAP and for 15 min with ΔTL RNAP.

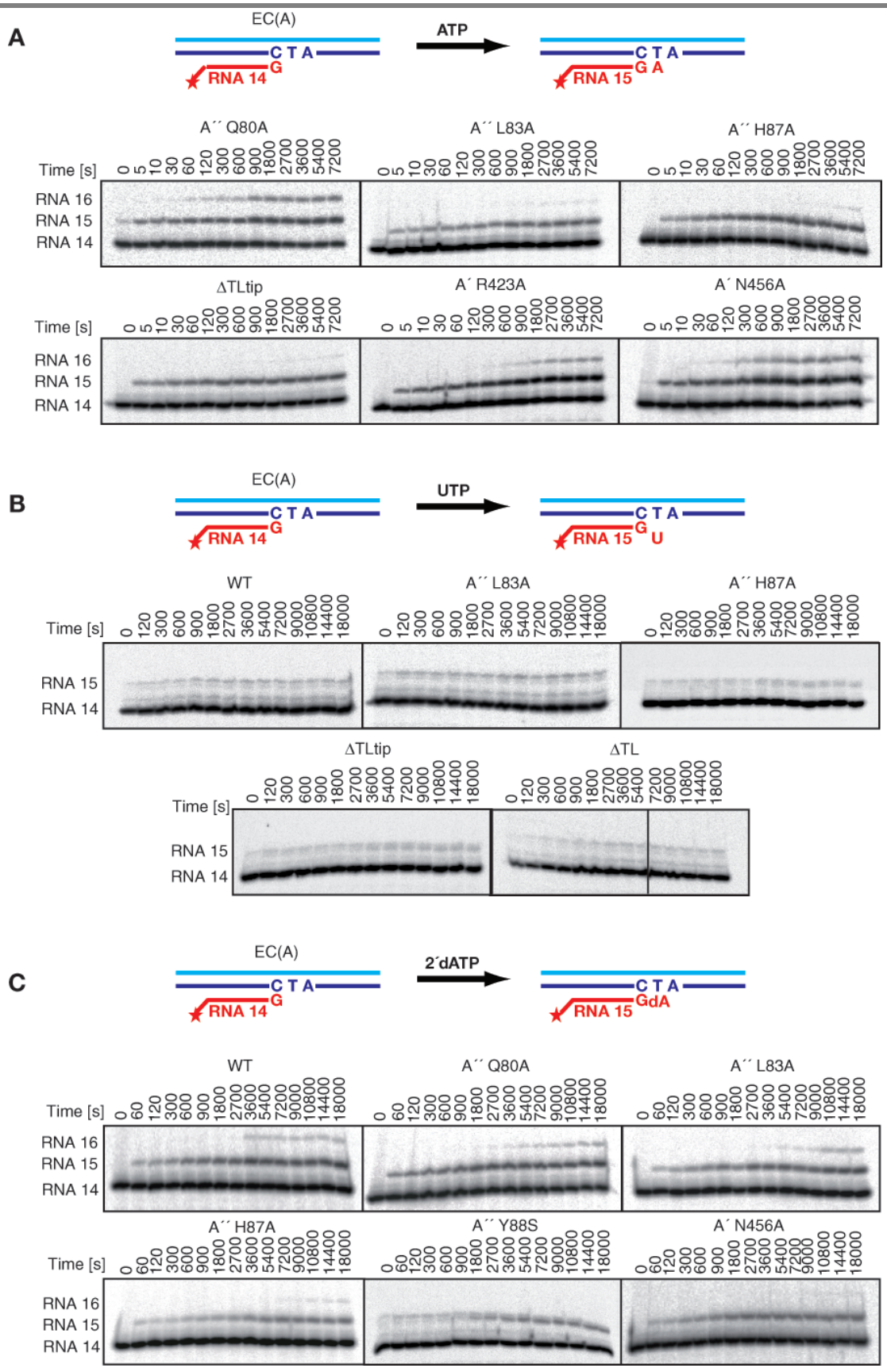
A
EC(A)
C T A
RNA 14
G
ATP
C T A
RNA 15
G A
A´´ Q80A
A´´ L83A
A´´ H87A
Time [s]
0 5 10 30 60 120 300 600 900 1800 2700 3600 5400 7200
RNA 16
RNA 15
RNA 14
ΔTLtip
A´ R423A
A´ N456A
B
EC(A)
C T A
RNA 14
G
UTP
C T A
RNA 15
G U
WT
A´´ L83A
A´´ H87A
Time [s]
0 120 300 600 900 1800 2700 3600 5400 7200 9000 10800 14400 18000
RNA 15
RNA 14
ΔTLtip
ΔTL
C
EC(A)
C T A
RNA 14
G
2´dATP
C T A
RNA 15
GdA
WT
A´´ Q80A
A´´ L83A
Time [s]
0 60 120 300 600 900 1800 2700 3600 5400 7200 9000 10800 14400 18000
RNA 16
RNA 15
RNA 14
A´´ H87A
A´´ Y88S
A´ N456A

Figure S1

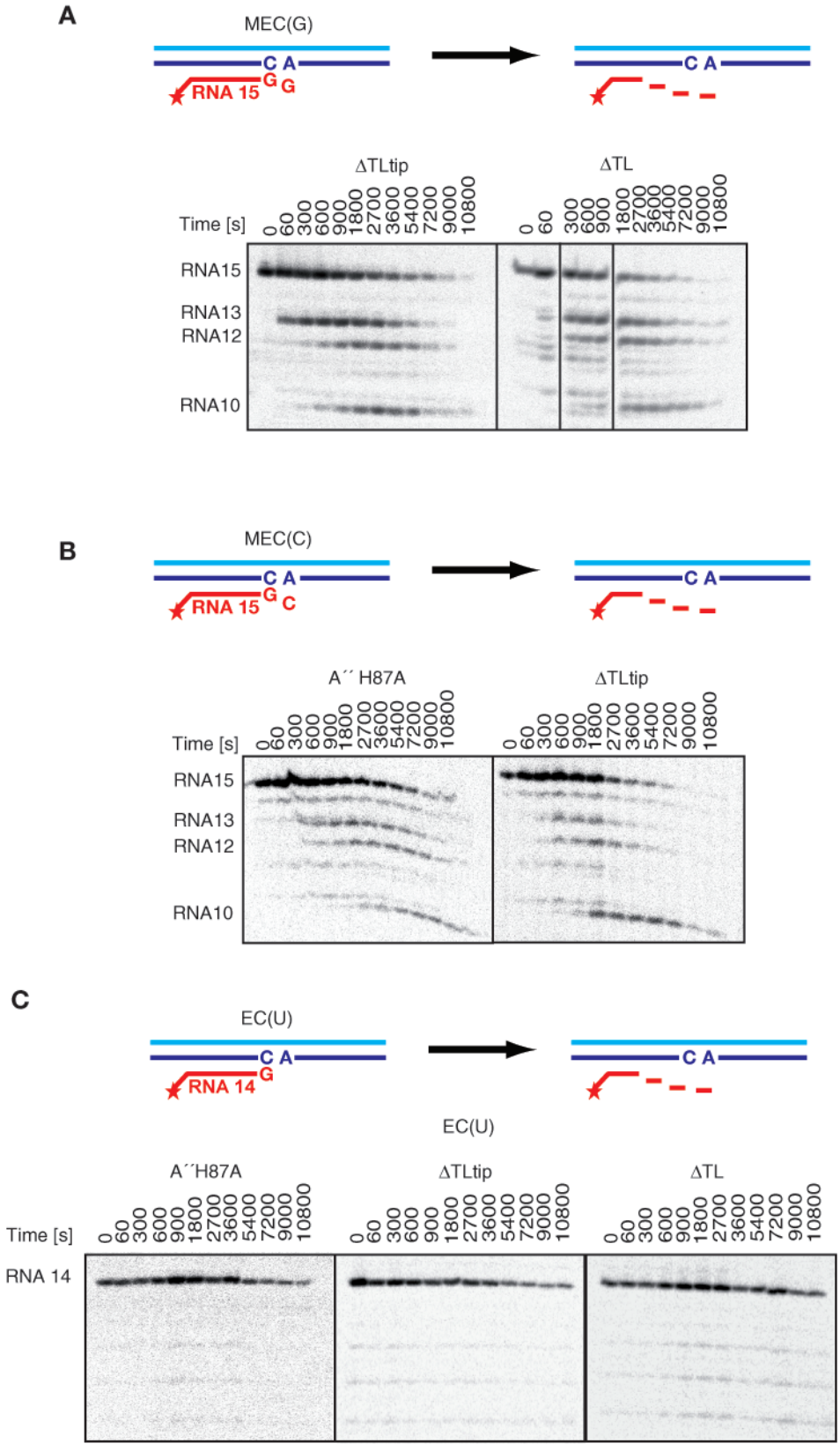
A
MEC(G)
C A
RNA 15 G G
C A
ΔTLtip
ΔTL
Time [s]
0 60 300 600 900 1800 2700 3600 5400 7200 9000 10800
0 60 300 600 900 1800 2700 3600 5400 7200 9000 10800
RNA15
RNA13
RNA12
RNA10
B
MEC(C)
C A
RNA 15 G C
C A
A´´ H87A
ΔTLtip
Time [s]
0 60 300 600 900 1800 2700 3600 5400 7200 9000 10800
0 60 300 600 900 1800 2700 3600 5400 7200 9000 10800
RNA15
RNA13
RNA12
RNA10
C
EC(U)
C A
RNA 14 G
C A
EC(U)
A´´H87A
ΔTLtip
ΔTL
Time [s]
0 60 300 600 900 1800 2700 3600 5400 7200 9000 10800
0 60 300 600 900 1800 2700 3600 5400 7200 9000 10800
0 60 300 600 900 1800 2700 3600 5400 7200 9000 10800
RNA 14

Figure S2

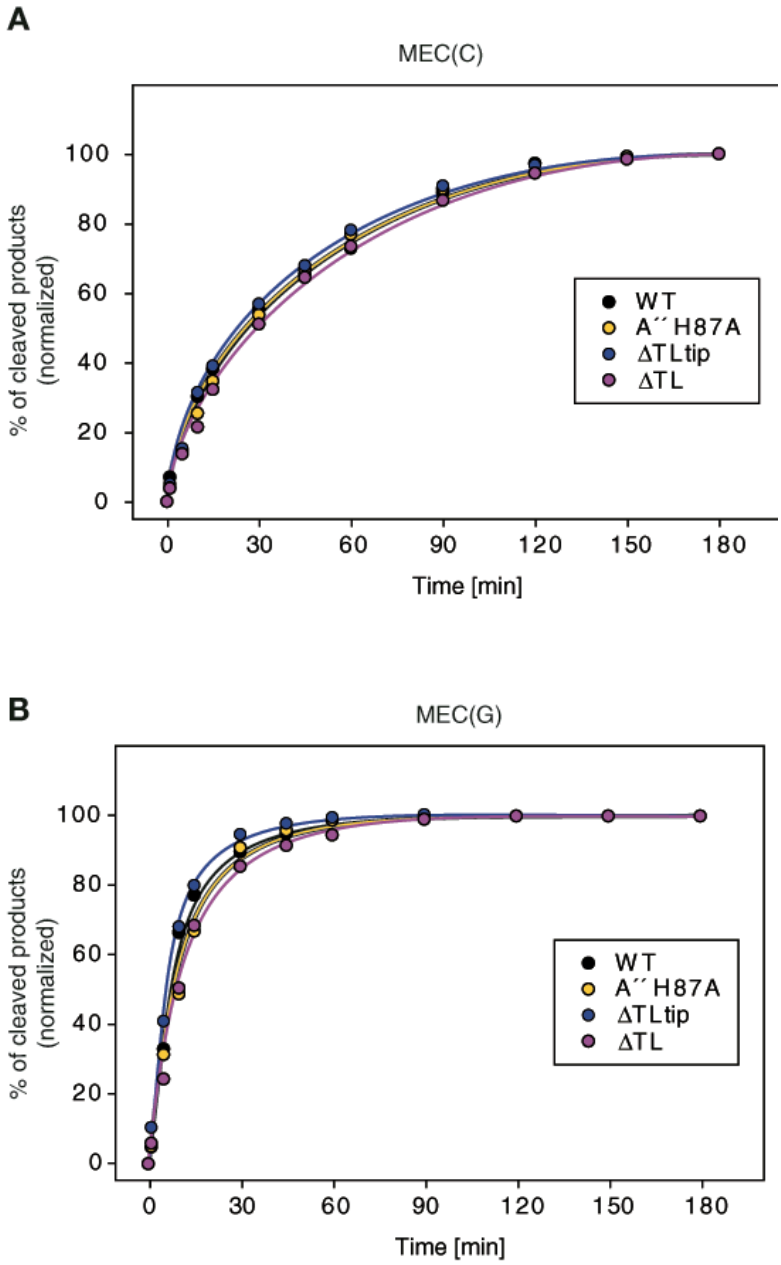
A
MEC(C)
% of cleaved products (normalized)
100
80
60
40
20
0
0
30
60
90
120
150
180
Time [min]
WT
A´´H87A
ΔTLtip
ΔTL
B
MEC(G)
% of cleaved products (normalized)
100
80
60
40
20
0
0
30
60
90
120
150
180
Time [min]
WT
A´´H87A
ΔTLtip
ΔTL

Figure S3

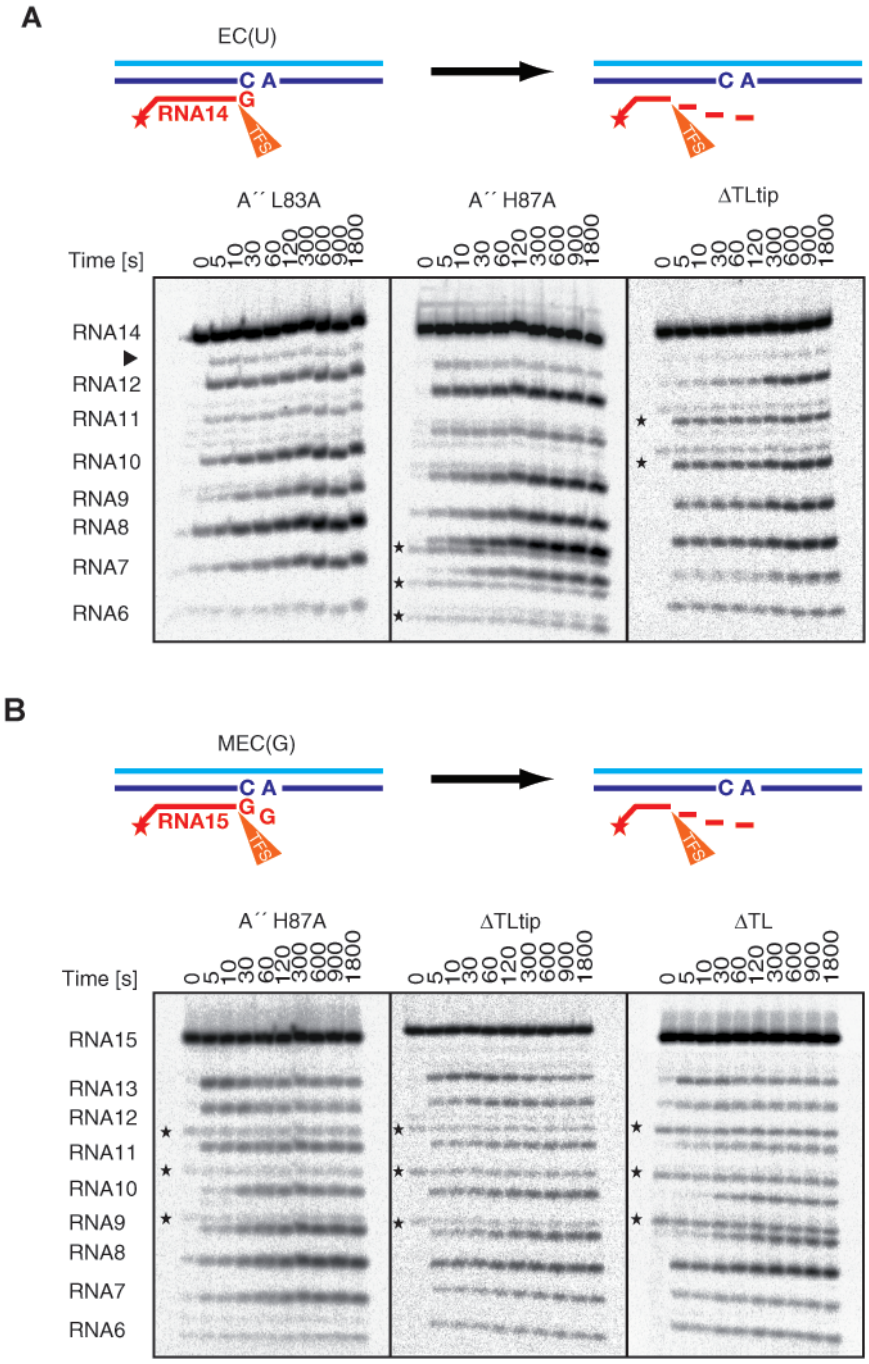
A
EC(U)
C A
G
RNA14
TFS
C A
TFS
A´´ L83A
A´´ H87A
ΔTLtip
Time [s]
0 5 10 30 60 120 300 600 900 1800
RNA14
RNA12
RNA11
RNA10
RNA9
RNA8
RNA7
RNA6
B
MEC(G)
C A
G G
RNA15
TFS
C A
TFS
A´´ H87A
ΔTLtip
ΔTL
Time [s]
0 5 10 30 60 120 300 600 900 1800
RNA15
RNA13
RNA12
RNA11
RNA10
RNA9
RNA8
RNA7
RNA6

Figure S4

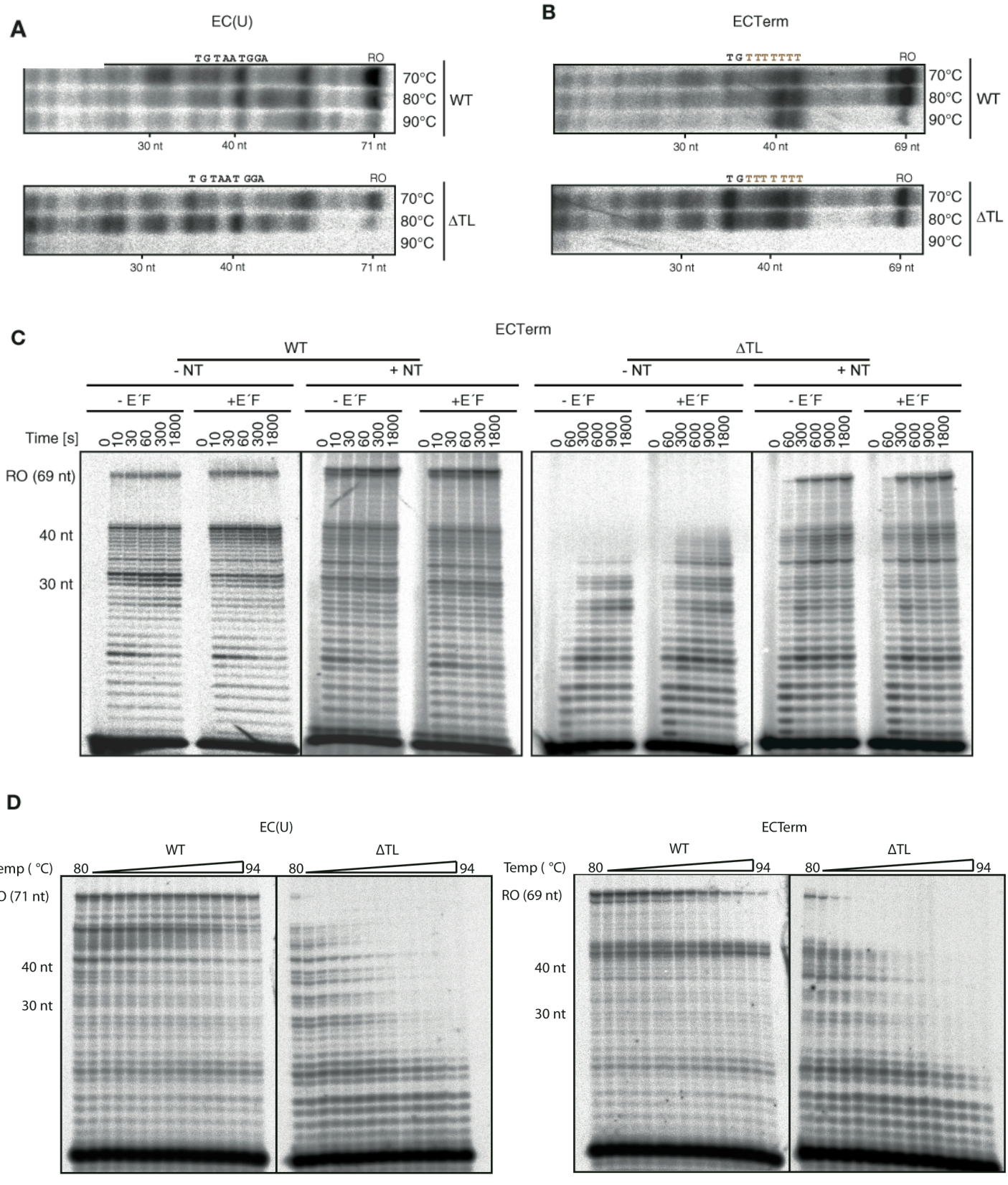
A
EC(U)
TG TAA TGGA
RO
70°C
80°C
90°C
WT
30 nt
40 nt
71 nt
T G TAAT GGA
ΔTL
B
ECTerm
TG TTT TTTT
69 nt
TG TTT T TTT
C
ECTerm
WT
ΔTL
- NT
+ NT
- E´F
+E´F
Time [s]
0 10 30 60 300 1800
0 60 300 600 900 1800
RO (69 nt)
40 nt
30 nt
D
EC(U)
ECTerm
Temp (°C)
80
94
RO (71 nt)
RO (69 nt)

Figure S5

Summary

RNA polymerases (RNAPs) carry out transcription in all living organisms. Whereas only one RNAP is present in bacteria and archaea, eukaryotes possess three to five specialized nuclear RNA polymerases (RNAP I, II, III, IV and V). To date, various RNAP structures have elucidated functional mechanisms of transcription, but biochemical analysis on eukaryotic RNAP II mutants are generally restricted to viable yeast strains. The active site of the RNAP catalyzes RNA chain growth by phosphodiester bond formation. The active site of all RNAPs contains two Mg^{2+} ions and an evolutionarily conserved mobile element, the trigger loop (TL) that functions in the elongation phase of transcription. Employing the reconstituted archaeal *in vitro* transcription system of *Pyrococcus furiosus,* we analyzed deletion mutants of the trigger loop and introduced alanine substitutions at key residues A'' Q80, A'' L83 and A'' H87. Furthermore, four supplementary mutants were analyzed in order to clarify their participation in NTP/2'dNTP discrimination.

The work of this thesis reveals that the archaeal TL is absolutely essential for transcription initiation, specially for capturing the incoming NTPs before a stable DNA–RNA hybrid is present in the active centre. Moreover, catalysis in initially transcribing complexes, active site closure and synthesis stimulation by the TL were crucial, as A'' H87 and the TLtip region were essential for initiation.

In vitro elongation assays also provided insights into TL function during the discrimination of correct NTPs from wrong dNTPs. Although A'' L83 and A'' H87 contribute to the recognition of the correct NTP and to the catalysis, A'' Q80 contributes to the recognition of the 2’OH-group of the NTP, indicating the critical role of the interaction between the TL and the NTP for the nucleotide incorporation fidelity.

Transcription fidelity also relies on proofreading, a post-incorporation mechanism that involves cleavage of a dinucleotide from the RNA 3’-end containing the misincorporated nucleotide. The TL is not required for the intrinsic cleavage activity of the enzyme but influences the translocation by being part of the Brownian ratchet that underlies translocation. The data shown here also suggests the critical role of the residue A'' L83 in backtracked complex stability. Proofreading is stimulated by extrinsic RNA cleavage factors such as TFS, and we found that the TL is also dispensable for TFS-stimulated cleavage, consistent with structures illustrating that the TL is in the locked conformation in the presence of the eukaryotic TFS counterpart, TFIIS.

Finally, our results revealed a function of the TL in transcription termination. We show that the sensitivity of the archaeal RNAP to poly-T sequences is increased on TL truncation, showing that the TL prevents aberrant termination at non-terminator sites, to ensure processive RNA synthesis.

Acknowledgements

First of all, I am very grateful to Professor Michael Thomm for giving me the chance to let me work on this exciting project, and also for his personal and scientific support throughout the time of my PhD. I also want to thank him for introducing me to the fascinating world of the RNA Polymerase and transcription.

I have been very fortunate to have a close collaboration with Professor Patrick Cramer, from the Gene center Munich, and I want to particularly thank him for his uncountable suggestions, discussions and interest for this work. Moreover, I would like to thank Dr. Alan Cheung for the professional discussions and helpful scientific expertise. Theirs constant advice and support were decisive in guiding this work.

Very special thanks to Dr. Mirijam Zeller for her enthusiasm and her constant support. I want to thank her especially for all encompassing help and support during my PhD but also later on, for her capacity for examining scientific issues and for her constructive criticism and discussion. Thank you for the great time together.

I am very thankful for all the present and former members of the Institute of Microbiology for their help and nice atmosphere in the lab. I thank Wolfgang Forster for his assistance in purification of *P. furiosus* RNAP subunits and for his permanent helps. Thanks to PD Dr. Winfried Hausner for his enthusiasm and for spreading his scientific knowledge. I would like to thank Robert Reichelt for his pleasant calmness, for the scientific discussions and for his great interest in science. Thank you for the great time together in and outside of the lab. I wish to give an especial thank to Johannes Danner and Jennifer Loose for the good times and for their cheerfulness. I also would like to thank Stefan Dexl for the good, but too short, time together in the lab.

I also want to thank Elisabeth Nagelfeld for her cheerfulness and for her constant encouragement and help.

I would like to extend my thanks also to the rest of the institute for the enjoyable time during my PhD and for letting me share in their charming Bavarian culture.

I am very thankful to Kathy Smollett for proofreading my PhD thesis.

Finally, I would like to thank my friends and family for their constant support and encouragement. Especially my parents, Marc Fouqueau and Rika Nishimura, my brother and sister for believing in me and for their moral and financial supports during my education and PhD. I would like to dedicate this thesis to them.

Erklärung

Hiermit erkläre ich, dass ich die vorliegende Arbeit selbstständig und ohne fremde Hilfe verfasst und nur die von mir angegebenen Quellen und Hilfsmittel verwendet habe. Diese Arbeit war bisher noch nicht Bestandteil eines Prüfungsverfahrens, andere Promotionsversuche wurden nicht unternommen.

Teilergebnisse dieser Arbeit sind bereits veröffentlicht worden:

Fouqueau, T., Zeller, M. E., Cheung, A. C., Cramer, P., Thomm, M. (2013). The RNA polymerase trigger loop functions in all three phases of the transcription cycle. Nucleic Acids Res. doi:10.1093/nar/gkt433 (in press).

Regensburg, im August 2013