Analyses of the Archaeal Transcription Cycle reveal a Mosaic of Eukaryotic RNA Polymerase II and III-like Features

Dissertation

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

vorgelegt von
Patrizia Spitalny
aus Karlsruhe

Januar 2008

Promotionsgesuch eingereicht: 15. Januar 2008
Diese Arbeit wurde angeleitet von: Prof. Dr. M. Thomm

Prüfungsausschuss:	Vorsitzender:	Prof. Dr. R. Wirth
	1. Gutachter und Prüfer:	Prof. Dr. M. Thomm
	2. Gutachter und Prüfer:	Prof. Dr. H. Tschochner
	3. Prüfer:	Prof. Dr. R. Sterner

Table of contents

I General Introduction

Transcription, the primary event in gene expression, plays a key role in the information processing pathways of all organisms. The synthesis of RNA from a DNA template is conserved among all DNA dependent RNA polymerases. The transcription cycle is divided into three major phases each of which is regulated by various factors and signal sequences. Starting with the promoter activation and initiation of RNA synthesis, a stable transcription complex is formed and as the nascent RNA is sufficiently long to stabilize this complex, the RNA Polymerase (RNAP) enters the elongation state. Finally the elongation ends when the RNA polymerase reaches one or more termination signals. The RNA is released and the RNAP starts subsequent rounds of transcription.

Within the last few years the transcription machineries of all domains of life have been studied extensively and many striking similarities especially between the archaeal RNA polymerase (RNAP) and the eukaryotic polymerase II (pol II) were elucidated (Bell and Jackson, 1998b; Thomm and Wich, 1988; Thomm, 1996). Although archaeal promoter structures as well as the sequences of their RNAP and of the transcription factors are closely related to their eukaryotic counterparts, the archaeal transcription machinery is vastly more simple than the eukaryotic pol II system. Archaea possess only one RNAP and the two transcription factors TBP and TFB suffice for promoter activation. This simplicity allowed a detailed analysis of mechanisms underlying different stages in the transcription cycle.

I.1 Initiation and elongation of archaeal transcription

Extensive studies during the last two decades provided detailed information on the mechanism of archaeal transcriptional initiation (Bartlett, 2005; Soppa, 1999).

Archaeal promoter activation is induced by the binding of the highly conserved transcription factor TBP to the TATA-box (Hausner *et al.*, 1991; Hausner *et al.*, 1996). The archaeal TATA box is an A-T rich eight-base-pair sequence element located around 25 bp upstream of the transcription start site. It has been identified as primary determinant of start site selection by different mutational analysis (Hain *et al.*, 1992; Hausner *et al.*, 1991; Reiter *et al.*, 1990). *In vivo* studies confirmed the essential role of the TATA element in archaeal promoter recognition (Palmer and Daniels, 1995). The saddle shaped TBP binds to the minor groove of the TATA-box with the DNA-binding region on the underside of the saddle and induces a DNA bending of about 65° (Kosa *et al.*, 1997; Littlefield *et al.*, 1999). The next step in

archaeal promoter activation is characterized by the binding of the TFIIB-related transcription factor TFB to the TBP-DNA complex. The C-terminal domain of TFB contacts TBP and contains a helix-turn-helix motif that mediates the sequence specific interaction with the transcription factor B recognition element (BRE) directly upstream of the TATA-box (Littlefield *et al.*, 1999). The contact with BRE is responsible for determining the orientation of the transcription complex (Bell *et al.*, 1999). By photocrosslinking experiments it has been shown that the N-terminal domain of TFB interacts with DNA around the transcription start site (Renfrow *et al.*, 2004). The N-terminal region of TFB also contains a zinc-ribbon that was shown to interact with the dock domain in subunit A' (Werner and Weinzierl, 2005) and with subunit K of the archaeal RNAP and may thereby have an important role in recruiting the RNAP (Magill *et al.*, 2001), while its B-finger was demonstrated to be involved in promoter opening (Micorescu *et al.*, 2007).

After the assembly of the TBP/TFB/DNA complex the RNA polymerase is positioned around the transcription initiation site (initiator element, INR; Hausner *et al.*, 1991; Thomm, 1996) by interaction of the RNAP dock domain with the TFB Zn-ribbon (Werner *et al.*, 2006). Upstream of the transcription start site the RNAP interacts with DNA around the transcription bubble via RNAP subunit B. The downstream contacts are mainly mediated by RNAP subunits A' and A'' and the front edge at around +18/+20 (Spitalny and Thomm, 2003) seems to be determined by subunit H (Bartlett *et al.*, 2004).

Although TBP and TFB are sufficient to recruit the RNAP for archaeal promoter-specific transcription initiation (Bell *et al.*, 1998; Hethke *et al.*, 1996; Qureshi *et al.*, 1997), the majority of archaeal genomes known so far contain a sequence for an additional transcription factor. It is homologous to the N-terminal region of the eukaryal TFIIE α-subunit (Aravind and Koonin, 1999; Bell and Jackson, 1998a; Kyrpides and Ouzounis, 1999) and therefore called TFE. In *in vitro* transcription assays it has been shown that the N-terminal part of the eukaryal TFIIEα is essential for basal and activated transcription (Ohkuma *et al.*, 1995). Archaeal TFE is not essential for basal *in vitro* transcription but it has a stimulatory effect on some promoters and under certain conditions (Bell *et al.*, 2001; Hanzelka *et al.*, 2001). Recently it could be demonstrated that TFE is stabilizing the transcription bubble (Naji *et al.*, 2007) and that it is also part of elongation complexes (Grünberg *et al.*, 2007).

During the assembly of the closed complex (Fig. 1A) the RNAP is only in weak contact to the DNA. The following conversion into the open complex is characterized by the separation of the DNA strands, accompanied by several conformational changes of the involved proteins and the DNA. The template strand is positioned into the active center and the RNAP-DNA

contact is stabilized by the B-finger and TFE (Werner and Weinzierl, 2005). The RNAP now enters the abortive state of transcription with repeated production of short transcripts (Fig. 1B). After synthesis of about 10 nucleotides the RNAP enters the elongation state. During promoter clearance and the transition from initiation to elongation the contact of the RNAP to the promotor bound transcription factors TBP and TFB is lost (Fig. 1C). Most likely TFB dissociates while TBP remains promoter bound (Xie and Reeve, 2004). Yet it could be demonstrated that on weak promoters also TBP dissociates (Geiduschek and Ouhammouch, 2005). TFE obviously remains attached to the mature elongation complex (Grünberg *et al.*, 2007).

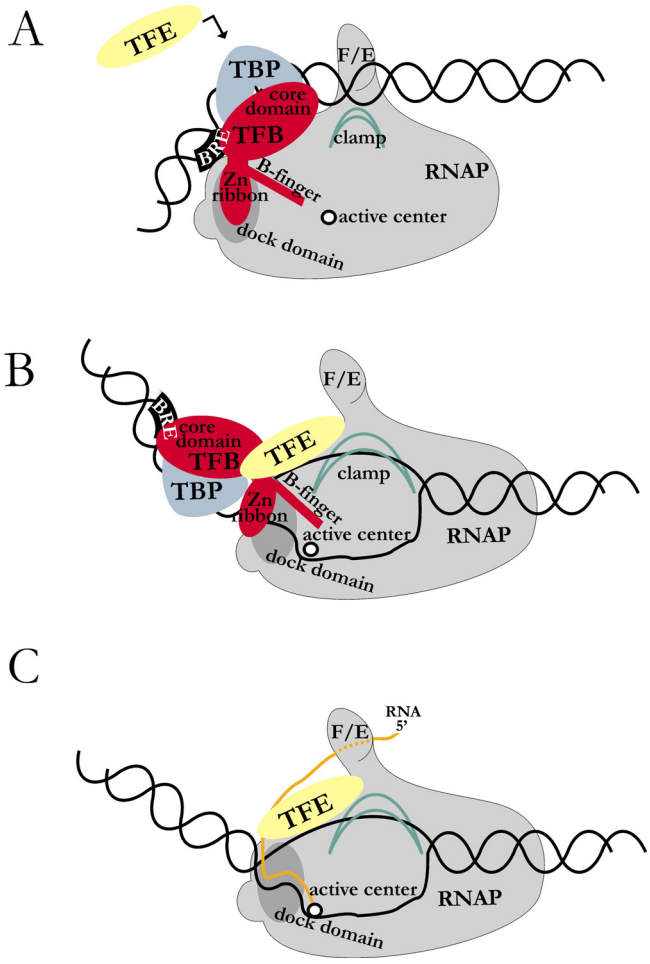

Figure 1**. Transcriptional initiation and elongation complex**. **A** The assembly of the preinitiation complex is mediated by the two transcription factors TBP (blue) and TFB (red). TFE (yellow) stimulates TBP binding under certain conditions. The RNAP (grey) is recruited to the promoter via interactions of the dock domain with the TFB zinc-ribbon. **B** Open complex formation is characterized by the melting of the DNA strands. The template strand comes into contact with the active center and the transcription bubble is stabilized by the B-finger and TFE. **C**. After synthesis of about 10 nucleotides RNAP looses contact to TBP and TFB while TFE remains associated with the elongation complex. The RNAP enters the elongation phase and synthesizes RNA in a synchronous and highly processive manner (modified after Werner *et al.*, 2006).

While the transition from initiation to elongation has extensively been studied in Bacteria and Eukarya (Kahl *et al.*, 2000; Kassavetis *et al.*, 1992; Metzger *et al.*, 1993; Samkurashvili and Luse, 1998; Schickor *et al.*, 1990), no data on that essential step in transcription had been avaliable for the archaeal system, until the detailed studies that are part of this work (Spitalny and Thomm, 2003) were published. In additionthe significance of RNAP structural elements for the early stages in the transcription cycle and the transition from initiation to elongation were elucidated by the analysis of mutated recombinant archaeal RNAPs (Naji et al., 2007) and will be discussed in this work.

Once the RNAP has reached the elongation phase the RNAP is readily synthesizing RNA molecules in a highly processive way. Yet, the RNAPs are no homogeneous population of elongating molecules. The elongation phase is sensitive to extrinsic and intrinsic signals leading to several intermediate states that include pausing, arrest and sliding of the transcription complexes (Fish and Kane, 2002).

In contrast to transcriptional initiation only few data on the elongation phase in Archaea are avaliable. This work characterizes an archaeal elongation complex stalled at position +20 in comparison to the eukaryal and bacterial counterparts (Spitalny and Thomm, 2003). Recent studies with mutant RNAPs on nucleic acid scaffolds revealed the function of several structural RNAP elements in transcription elongation (Naji et al, 2007).

Other analyses on archaeal elongation address TFS, a homologue of the polymerase II transcription factor TFIIS. TFS shows sequence similarity to the C-terminal domain of the eukaryotic transcription elongation factor TFIIS and to small subunits of all three eukaryotic RNA polymerases (Hausner *et al.*, 2000). TFS was shown to be a cleavage stimulatory factor similar to TFIIS (reviewed by Fish and Kane, 2002) and not a subunit of the archaeal RNAP (Hausner *et al.*, 2000). TFS acts on arrested or backtracked elongation complexes where the 3'- end of the nascent transcript is no longer located in the active center. By dinucleotide cleavage TFS generates a new 3'-end of the nascent RNA now positioned in the active center again (Lange and Hausner, 2004).

I.2 Termination of archaeal transcription

While in elongation state the RNAP is highly stable and synthesizes long RNA chains. Yet it destabilizes abruptly at certain termination signals. Archaeal termination has not been addressed by many studies so far. Based on early investigations (Muller *et al.*, 1985; Reiter *et al.*, 1988), an intrinsic termination mechanism in the archaeal system has been assumed.

Oligo-dT streches were shown to mediate transcript termination. A mutational study revealed 5'-TTTTAATTTT-3' as a termination signal for the $tRNA^{Val}$ gene of *Methanococcus vannielii* (Thomm et al., 1994). Deletion of two T residues from the 3'-end of the termination sequence significantly lowered termination efficiency and a deletion leaving only 5'-TTTTAA -5' completely abolished termination activity. Apart from the necessitiy of this octameric sequence the presence of tRNA secondary structures contributed significantly to the termination process. Deletion of the tRNA TΨC stem-loop structure resulted in reduced termination efficiency. Additionally it could be shown that a bacterial intrinsic terminator can replace the $tRNA^{Val}$ terminator completely. Recently, an *in vitro* single round system was established to study termination in a thermophilic archaeal transcription system based on the transcription system of *Methanobacterium thermoautotrophicum*. *M. t.* RNAP was demonstrated to terminate in response to several bacterial, phage and synthetic terminators (Santangelo and Reeve, 2006). In contrast to bacterial RNAPs the presence of a sequence capable of formation of a stem-loop structure is not essential for the termination of the archaeal RNAP. Although there are obviously no homologues of the bacterial termination factor rho encoded in archaeal genomes, archaeal transcription complexes are sensitive to disruption by the bacterial rho-factor (Santangelo and Reeve, 2006).

The present thesis not only contributes to the almost unknown field of archaeal termination, it also reveals the existance of a reinitiation mechanism in Archaea (Spitalny and Thomm, 2007)

I.3 Aim and outline of this thesis

Although considerable information on the archaeal transcription machinery is avaliable, yet many questions remain. Figure 2 demonstrates that the present thesis contributes to the understanding of mechanistic aspects that accompany all three major phases of the archaeal transcription cycle.

The chapter II "Analysis of the Open Region and of DNA-Protein Contacts of Archaeal RNA Polymerase Transcription Complexes during Transition from Initiation to Elongation" presents a detailed view on major conformational transitions that occur during early transcription. Footprinting analyses of stalled transcription complexes were conducted for positions +5 to +20 relative to the transcription start site. Exonuclease III was used to analyse the borders of the RNAP at the defined positions. The corresponding transcription bubbles as well as the RNA-DNA hybrid could be detected with potassium permanganate. The results

are discussed in the context of similar data avaliable for bacterial and eukaryotic RNAPs revealing a conserved mechanism between all DNA dependent RNA polymerases for the transition from transcriptional initiation to elongation.

In chapter III "Structure-function analysis of the RNA polymerase cleft loops elucidates initial transcription, DNA unwinding and RNA displacement" the influence of structural elements of the archaeal RNAP on different stages of the transcription cycle is analyzed. The avaliability of recombinant archaeal RNAPs from fully recombinant subunits provided the opportunity to selectively mutate structural elements known to have functional roles in bacterial and eukaryotic RNAPs. The close relationship to the eukaryotic pol II allowed the identification of four loop structures and of three essential amino acid residues that have been deleted or mutated, respectively. The resultant mutants showed defects at different stages of the transcription cycle and their impact on the dynamics of the transcription cycle is discussed.

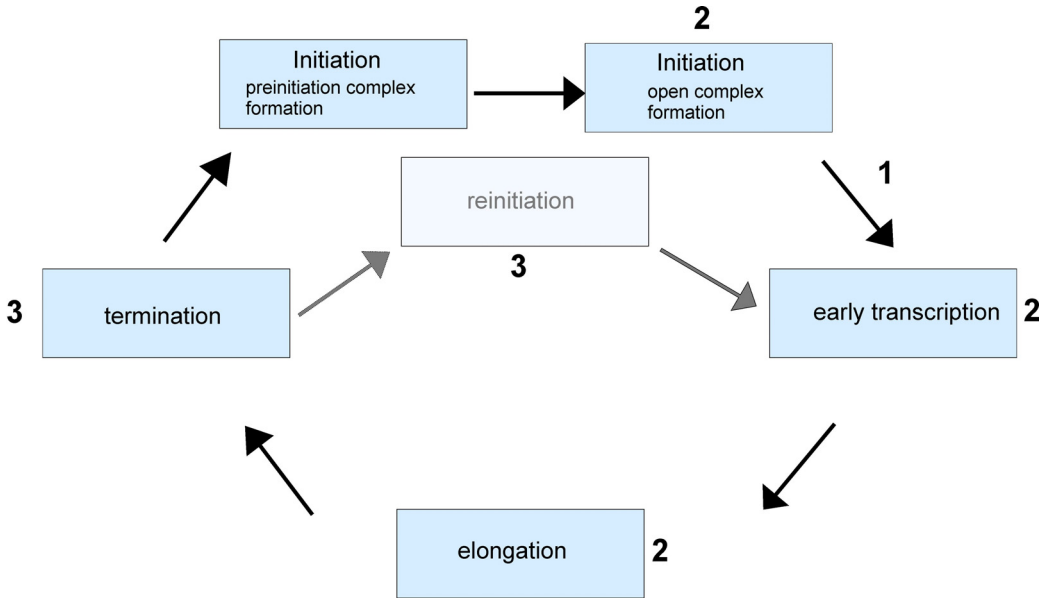

Figure 2. **The transcription cycle.** The different stages of the transcription cycle are depicted. The numbers inserted indicate what stages of the transcription cycle are analysed in the chapters of this thesis. 1: "Analysis of the Open Region and of DNA-Protein Contacts of Archaeal RNA Polymerase Transcription Complexes during Transition from Initiation to Elongation". 2: "Structure-function analysis of the RNA polymerase cleft loops elucidates initial transcription, DNA unwinding and RNA displacement". 3: "A polymerase III-like reinitiation mechanism is operating in regulation of histone expression in Archaea".

In chapter IV "A polymerase III-like reinitiation mechanism is operating in regulation of histone expression in Archaea" a complete archaeal histone gene with its adjacent four consecutive oligo-dT streches was used as a model system to address the question of termination mechanisms in hyperthermophilic Archaea. It could be demonstrated for the first

time that transcriptional recycling is present in the archaeal transcription system and that it plays an important role in transcriptional regulatory mechanisms. The results are discussed with respect to similarities in the termination and recycling processes of other DNA dependent RNA polymerases, especially to those of the eukaryotic polymerase III (pol III).

THE JOURNAL OF BIOLOGICAL CHEMISTY Vol. 278, No. 33, Issue of August 15, pp. 30497-30505, 2003

II Analysis of the Open Region and of DNA-Protein Contacts of Archaeal RNA Polymerase Transcription Complexes during Transition from Initiation to Elongation

Patrizia Spitalny* and Michael Thomm*[1]

Universität Kiel, Institut für Allgemeine Mikrobiologie, Am Botanischen Garten 1-9, D-24118 Kiel, Germany
* Present address: Universität Regensburg, Lehrstuhl für Mikrobiologie, Universitätsstr. 31, D-93053 Regensburg, Germany

Running Title: Stalled archaeal transcription complexes

Key words: Archaea, transcription, initiation, elongation, open complex, immobilized transcription complexes, transcription bubble, RNA-DNA hybrid, exonuclease III footprints

1 Corresponding author

E-mail: Michael.Thomm@Biologie.Uni-Regensburg.de
Phone: +49-941-943-3160
FAX: +49-941-943-2403

ABSTRACT

The archaeal transcriptional machinery is polII like but does not require ATP or TFIIH for open complex formation. We have used enzymatical and chemical probes to follow the movement of *Pyrococcus* RNA polymerase (RNAP) along *glutamate dehydrogenase* gene during transcription initiation and transition to elongation. RNAP was stalled between registers +5 and +20 using C-minus cassettes. The upstream edge of RNAP was in close contact with the archaeal transcription factors TBP/TFB in complexes stalled at position +5. Movement of the downstream edge of the RNAP was not detected by exonuclease III footprinting until register +8. A first structural transition characterized by movement of the upstream edge of RNAP was observed at registers +6/+7. A major transition was observed at registers +10/+11. In complexes stalled at these positions also the downstream edge of RNA polymerase started translocation and reclosure of the initially open complex occurred indicating promoter clearance. Between registers +11 and +20 both RNAP and transcription bubble moved synchronously with RNA synthesis. The distance of the catalytic center to the front edge of the exo III footprint was approximately 12 nucleotides in all registers. The size of the RNA-DNA hybrid in an

early archaeal elongation complex was estimated between 9 and 12 nucleotides. For complexes stalled between positions +10 and +20 the size of the transcription bubble was around 17 nucleotides. This study shows characteristic mechanistic properties of the archaeal system and also similarities to prokaryotic RNAP and polII.

INTRODUCTION

Transcription initiation requires formation of a preinitiation complex (PIC), melting of DNA, formation of the first phosphodiester bonds and promoter clearance involving movement of the open DNA region ("transcription bubble") and RNA polymerase. Finally, a stable ternary elongation complex is formed. These steps have been extensively studied during the last two decades in bacterial RNA polymerase and eukaryotic polymerase II (for reviews see 1, 2) and to less extent in eukaryotic RNA polymerase III (3, 4) and RNA polymerase I (5) systems. In Archaea, open complex formation at the *Methanococcus* $tRNA^{Val}$ (6) and at the 16S rRNA promoter of *Sulfolobus* (7) have been studied. The transition from initiation to elongation has not yet been investigated in Archaea.

In Bacteria, promoter isomerization from closed to open complex catalyzed by the predominant RNA polymerase holoenzyme ($\beta\beta\alpha_2\sigma^{70}$) occurs spontaneously in a temperature dependent manner (8, 9). By contrast, nuclear RNA polymerase II (polII; 10) and *Escherichia coli* RNA polymerase specific for promoters of genes involved in nitrogen metabolism ($\beta\beta'\alpha_2\sigma^{54}$; 11), require ATP hydrolysis for promoter melting. In the polII system promoter opening involves the helicase activity of TFIIH (12). Eukaryotic nuclear RNA polymerases I and III share with the σ^{70} containing *E. coli* RNA polymerase the ability to produce an open complex of 12-15 bp without ATP hydrolysis. In the polIII system the presence of the general transcription factor TFIIIB is required in addition for open complex formation (13).

Methods have been described to prepare ternary complexes stalled at different positions. Analyses of these transcription complexes by nuclease and chemical footprinting provided detailed insights into the basic mechanism of initiation of transcription in enteric bacteria and the eukaryotic polII system. PolII complexes were subjected to numerous structural alterations during formation of the first 30 phosphodiester bonds (14, 15, 16). In Bacteria, a discontinuous model of elongation (inchworming) was inferred from these studies (17, 18). The finding that movement of RNA polymerase along the DNA template was not synchronous with single-nucleotide additions was alternatively explained by transient backtracking of RNA polymerase (19). Goldfarb and coworkers provided evidence that the strength of the RNA-DNA hybrid is essential for maintaining stability of transcription

complexes by preventing backtracking of RNA polymerase (20). Irregular footprints observed earlier were interpreted by these authors in the light of their findings as reflections of mixed populations of transcription complexes in productive and backtracked states.

The archaeal transcription system is a simplified version of the eukaryotic polII machinery (21, 22). The archaeal TATA-box is recognized by an archaeal TATA-box binding protein (TBP). This interaction is stabilized by TFB, an homologue of general polII transcription factor TFIIB. This TBP-TFB promoter complex recruits the archaeal RNA polymerase which shows striking similarity in sequence and subunit composition to polII. With exception of TFE which is homologous to the alpha subunit of polII transcription factor TFIIE (23, 24) no other homologues of the basal eukaryotic transcriptional machinery were detected in archaeal genomes. Consistent with the lack of TFIIH in Archaea and in contrast to the striking general similarity to the polII system the archaeal RNA polymerase does not require hydrolysis of ATP for open complex formation at the $tRNA^{Val}$ promoter of *Methanococcus vannielii* (25). We recently have developed a cell-free transcription system for the hyperthermophile *Pyrococcus furiosus* (26). This highly purified system consisting of bacterially produced TBP and TFB, and RNA polymerase isolated from *Pyrococcus* cells was used for the characterization of the archaeal preinitiation complex (27), analysis of the trajectory of DNA in an archaeal transcription complex (28) and first studies on regulation of transcription in Archaea (29, 30). Here, we used immobilized templates to purify *Pyrococcus* ternary transcription complexes stalled in registers between +5 and +20. Analysis of these complexes by exonuclease III (exo III) and potassium permanganate ($KMnO_4$) footprinting provided a detailed view of the early steps of transcription in Archaea.

EXPERIMENTAL PROCEDURES

Reagents and Enzymes - Exonuclease III was purchased from NEB (New England Biolabs Inc.). Potassium permanganate was obtained from Merck (Darmstadt, Germany). [α -^{32}P] UTP and [γ -^{32}P] ATP were purchased from Hartmann Bioanalytics (Braunschweig, Germany).

Templates for in vitro transcription and footprinting reactions - Nine templates were constructed. All cytosine residues in the non template strand between the transcription start site and position +20 relative to the transcription start site were substituted by other bases using PCR and the plasmid pUC19 containing the *gdh (glutamate dehydrogenase)* gene from -95 to +163 from *Pyrococcus furiosus*.

The forward primer was complementary to sequences ~160 bp upstream of the transcription start site, and the reverse primer was partly complementary to sequences from positions -15 to +20 used to induce the point mutations. After hydrolysis of the amplified fragments with *EcoRI* the resultant DNA fragments contained the promoter and the mutated region downstream of the transcription start site of the *gdh* gene.The fragments were inserted between the *EcoRI* and *SmaI* (compatible to the blunt ends on one side of the fragment produced by PCR amplification) restriction sites of the vector pUC19. The resulting fragments were transformed into *E. coli* JM109. The resulting plasmids pUC19/*gdh*-C5, pUC19/*gdh*-C6, pUC19/*gdh*-C7, pUC19/*gdh*-C8, pUC19/*gdh*-C9, pUC19/*gdh*-C10, pUC19/*gdh*-C11, pUC19/*gdh*-C15 and pUC19/*gdh*-C20 were used to generate transcription templates by PCR with 263 to 278 bp in length. Oligonucleotides complementary to DNA sequences ~160 bp upstream and ~ 90 bp downstream of the transcription start site were used as primers. One primer was labelled with biotin and the resulting fragments were attached to streptavidin magnetic beads (Roche Molecular Biochemicals, Mannheim, Germany) according to the manufacturer's protocol.

Purification of Pyrococcus RNA polymerase - RNA polymerase from *Pyrococcus furiosus* was purified as described previously by (26).

Expression and purification of recombinant transcription factors - The transcription factor TBP from *P.woesei* was overproduced in *E. coli* as described previously by (27). The DNA sequences of *P. woesei* TBP and *P. furiosus* TBP show 100% identity. TFP from *P. furiosus* was expressed and purified as described previously for *P. woesei* (27).

Immobilized in vitro transcription assays - *In vitro* transcription assays were performed according to (26). A standard reaction mixture (25 µl) contained 60 ng immobilized template, 660 ng RNA polymerase, 150 ng of recombinant TBP, 120 ng recombinant TFB, 40 µM ATP and GTP, 2 µM UTP and 0.15 MBq [α -^{32}P] UTP (110 TBq/mmol). The transcription buffer contained HEPES 40 mM, EDTA 0.1 mM, dithiothreitol (DTT) 1 mM, KCl 300 mM and $MgCl_2$ 4 mM. Transcription reactions were performed for 3 minutes at 70°C.

To perform footprinting experiments the immobilized templates were labelled with [γ- ^{32}P] ATP. [α- ^{32}P] UTP was not added to the reaction.

Isolation of stalled ternary complexes - Ternary complexes stalled in *in vitro* transcription reactions at positions +5 to +11, +15 and +20 relative to the transcription start site were isolated at 20 °C by the use of a magnet, so DNA attached to magnetic beads could be located to one edge of the reaction tube and the supernatant could be removed. To remove TBP/TFB from promoter DNA, complexes were washed with transcription buffer containing 0.5% N-

lauroyl-sarcosine (NLS) and 40 µM GTP. Then, the isolated ternary complexes were resuspended in transcription buffer and either analyzed on a 28% denaturing polyacrylamide gel or supplemented with all four nucleotides (40 µM each) but no additional radioactivity in a total volume of 25 µl. During further incubation for 3 minutes at 70°C run-off transcripts were formed. Transcription reactions were stopped by the addition of loading buffer (98% formamide, 10 mM EDTA and 0.1% each bromphenol blue and xylene cyanol).

Exonuclease III footprinting - To perform footprinting experiments the immobilized DNA templates were labelled with [γ -^{32}P] ATP on the free 5'-end of either the coding or the RNA-like strand, depending on which strand was attached to the magnetic particle on the 5'-end. The *in vitro* transcription reaction was performed as described but no [α- ^{32}P] UTP was omitted. After the complexes had been stalled at positions +5 to +11, +15 and +20 relative to the transcription start site they were isolated as described. Then, they were resuspended in 25 µl of reaction buffer for exo III digestion (KCl 40 mM, $MgCl_2$ 2 mM, Tris-HCl pH 8.5, 100 mM and DTT 1 mM). After addition of 100 U exo III the samples were incubated at 37°C for 15 minutes. The reaction was stopped by the addition of loading buffer and the samples and sequencing reactions were loaded on a 6% denaturing sequencing gel.

$KMnO_4$ sensitivity assay - To perform $KMnO_4$ probing the immobilized DNA templates were labelled with [γ- ^{32}P] ATP on the free 5'-end of either the coding or the non coding strand, depending on which strand was attached to the magnetic particle on the 5'-end. The *in vitro* transcription reaction was performed as described but no [α- ^{32}P] UTP was added. After the complexes had been stalled at positions +5 to +11, +15 and +20 relative to the transcription start site they were isolated as described. The complexes were resuspended in 25 µl transcription buffer and 2.5 µl potassium permanganate (250 mM) were added. The samples were incubated for 5 minutes at 45 °C. The reactions were stopped by the addition of 1,7 µl 2-mercaptoethanol and 20 µl of stop mix (1,25% SDS, 125 mM EDTA). The supernatant was removed and the modified immobilized DNA was resuspended in 18 µl water and piperidine was added to a total volume of 20 µl. The DNA was subjected to cleavage by piperidine for 30 minutes at 90°C. Piperidine was removed by ethanol precipitation and the dried pellets were resuspended in loading buffer and loaded together with a sequence ladder onto a 6% denaturing sequencing gel.

To detect the open complex 2,5 µl $KMnO_4$ (250 mM) were added immediately after incubation of template DNA with TBP, TFB and RNA polymerase for 3 minutes at 70°C and the reaction was performed for another 3 minutes at 70°C. The reaction was stopped and subjected to piperidine treatment as described.

RESULTS

Stalled archaeal transcription complexes contain a homogenous population of RNA molecules- To investigate the transition between initiation and elongation we constructed a set of sequence variations of the *Pyrococcus gdh* gene sequence with their first C residue between position +6 and +21. RNA synthesis can be blocked at positions 5, 6, 7, 8, 9, 10, 11, 15 and 20 (Fig. 1) by omitting CTP from transcription reactions. The primers used for the construction of these gdh-C derivatives (Fig. 1) were biotinylated allowing rapid isolation of ternary transcription complexes by streptavidin coated magnetic beads (Experimental Procedures). Both in bacterial and eukaryotic systems read-through of RNA polymerase (RNAP) beyond the expected stall sites has been observed (31, 32).

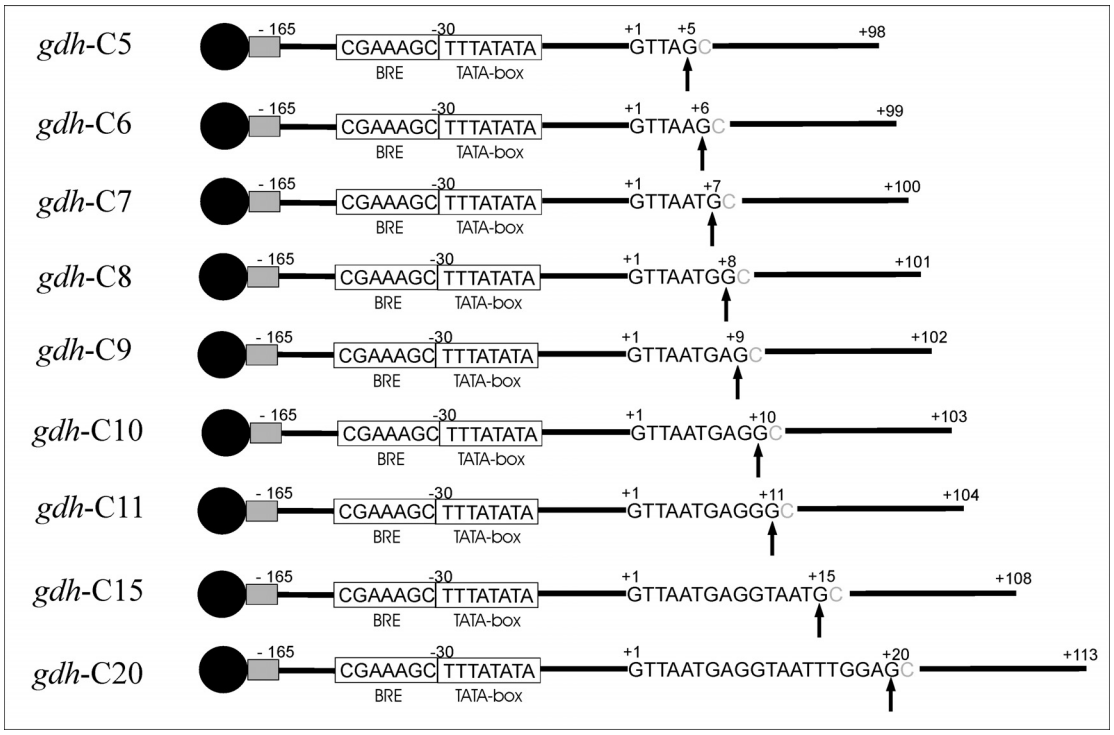

FIG. 1. **Templates for stalling and rapid isolation of archaeal ternary transcription complexes.** Schematic drawing of immobilized templates used for stalling the RNA polymerase at defined positions in *in vitro* transcription assays without CTP. The DNA sequences of the promoter and the mutated initiation region of the non template strand of the *gdh*-gene from *Pyrococcus furiosus* are depicted. The templates were constructed as described in Experimental Procedures. The TATA-box and the BRE are boxed. The nucleotides are numbered relative to the transcription start site and the stall positions are marked with arrows. The templates are immobilized on magnetic beads (black circle) using a biotin streptavidin linkage (gray box) on the 5'-end.

To establish the conditions for the synthesis of RNA products of correct size we analyzed first cell-free transcripts from the template containing the first C-residue at position +21 (gdh-C20; Fig. 2A). RNA products were labelled with [α-^{32}P] UTP. After short incubation times between 30 sec and 3 min an RNA product of 20 nt was synthesized as predominant product (data not

shown). After incubation times between 5 and 45 min, additional products of 21and 32 nt probably caused by missincorporation at positions 21 and 22 were observed. Products of the expected size were also transcribed from the other templates shown in Fig. 1 when transcription reactions were conducted for 3 min (Fig. 2A).

Stalled complexes are stable and transcriptionally competent - To investigate the stability of stalled ternary complexes, the various biotinylated templates (Fig. 1) were incubated for 3 min in transcription reactions and ternary complexes were purified by strepavidin coated magnetic beads. The RNA contained in these purified complexes was analyzed by polyacrylamide gel electrophoresis. In addition, the RNA released by the RNAP which was not bound to the magnetic beads was analyzed. The ratio of nascent RNA in ternary complexes to released RNA increased with the length of the RNA molecules synthesized (Table 1). When isolated complexes were incubated in transcription buffer supplemented with all nucleotides, labelled RNA associated with isolated complexes could be no longer detected (data not shown).

TABLE I: *Ratio of nascent to released RNA.* The amount of nascent RNA bound to isolated complexes and of released RNA from the supernatant of transcription reactions was analyzed on a 28% polyacrylamide gel and quantified by the use of a Fuji PhosphorImager for each stall position. The amount of nascent RNA increases with transcript length.

Register of stalled complex	C5	C6	C7	C8	C9	C10	C11	C15	C20
Nascent RNA (%)	19	11	11	18	20	21	43	73	75
Released RNA (%)	81	89	89	82	80	79	57	27	25

This finding suggests that the nascent RNA molecules were retained in functional ternary complexes which were elongated to run-off products after addition of nucleotides. To provide conclusive evidence that the isolated complexes were functionally active the RNA products released after addition of nucleotides were analyzed. Fig. 2B shows that run-off transcripts were synthesized under these conditions. Therefore, the isolated ternary transcription complexes are still functionally competent and seem suitable for subsequent analyses of footprints of the RNAP and of growth of transcription bubble at each of these stall sites. An additional analysis of the labelled RNA in isolated complexes stalled at each register between +5 and +20 (see Fig. 1) showed that RNA of the expected size was the major product in most cases (Fig. 2A, lanes 9, 11, 13, 15 and 17). Longer exposures of complexes stalled in register

+20 and +15 showed the existence of minor RNA products estimated to be 18 and 13 nt in length (Fig. 2A, lanes 15 and 17). At present it is unclear whether these shorter RNA products are caused by pausing of RNAP or whether they are due to hydrolysis of completed RNA from its 3´-end. All complexes stalled between +7 and +20 contained a 5 nt product suggesting the existence of DNA fragments in complex with RNAP paused at position +5. Further analysis of exo III and $KMnO_4$ footprints showed that the movement from +5 to position +6/+7 marks a significant transition in archaeal transcription which is probably a rate-limiting step (see below). However, these complexes stalled at +5 were not arrested since they could be chased after addition of nucleotides (Fig. 2A, lanes 6, 8, 10, 12, 14, 16 and 18).

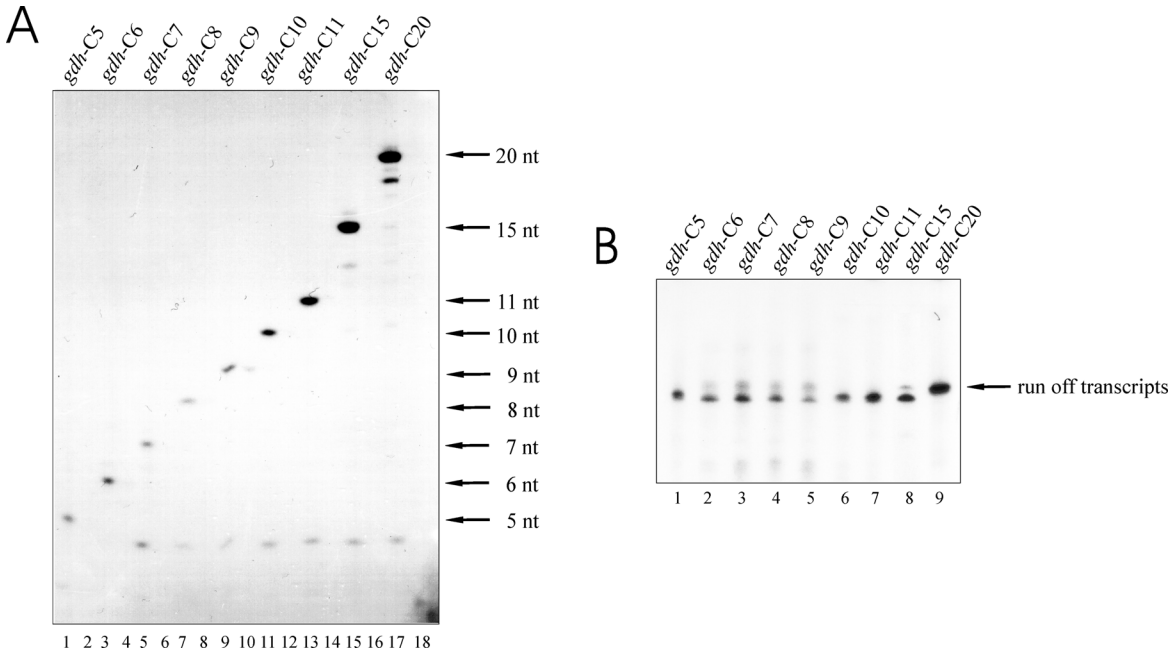

FIG. 2. **Isolated ternary complexes contain RNAs of correct size and are transcriptionally active.** Transcription experiments were performed as described in Experimental Procedures. The stalled complexes were washed in washing buffer to remove unincorporated nucleotides and released RNA. In **A** the isolated complexes were analyzed on a 28% polyacrylamide gel. Lanes 1, 3, 5, 7, 9, 11, 13, 15 and 17 show the RNA products of the isolated complexes stalled in the indicated registers. The higher mobility of the 5 nt RNA products in lanes 5, 7, 9, 11, 13, 15 and 17 is due to the last incorporated nucleotide in the nascent RNA being an A instead of a G. Minor products in lanes 15 and 17 could be detected after longer exposure. When the isolated complexes were chased by the addition of all NTPs (40 µM each) no RNA products could be detected in lanes 2, 4, 6, 8, 10, 12, 14, 16 and 18 indicating that all isolated complexes remained in a transcriptionally competent state. In **B** the run-off products in the supernatant ranging in length from 98 nt for *gdh*-C5 to 113 nt for *gdh*-C20 are shown. They were analyzed on a 6% polyacrylamide gel.

Interaction of stalled RNA polymerase with DNA probed by exonuclease III footprinting - We used exo III as a probe to identify the upstream and downstream boundaries of RNA polymerase at each of the stall sites. To define the upstream extent of the binding site, linear DNA was 5´-end labelled with ^{32}P on the template DNA strand, and the biotin label was

associated with the 5´-end of the complementary DNA strand. For analysis of the downstream extent of the RNAP binding site, the 5´-end of the RNA-like strand was labelled with ^{32}P, and the biotin label was attached to the 5´-end of the template DNA strand. Cell-free transcription reactions were conducted at 70 °C, the subsequent purification of transcription complexes at 20 °C. Since exo III was rapidly inactivated at 70 °C (data not shown), ternary complexes were incubated with exo III at 37 °C. At this temperature, initiation of transcription did not occur, but already formed isolated ternary complexes can be elongated by addition of a complete set of nucleotides (data not shown). Therefore, the complexes probed by exo III at 37 °C were transcriptionally active and competent.

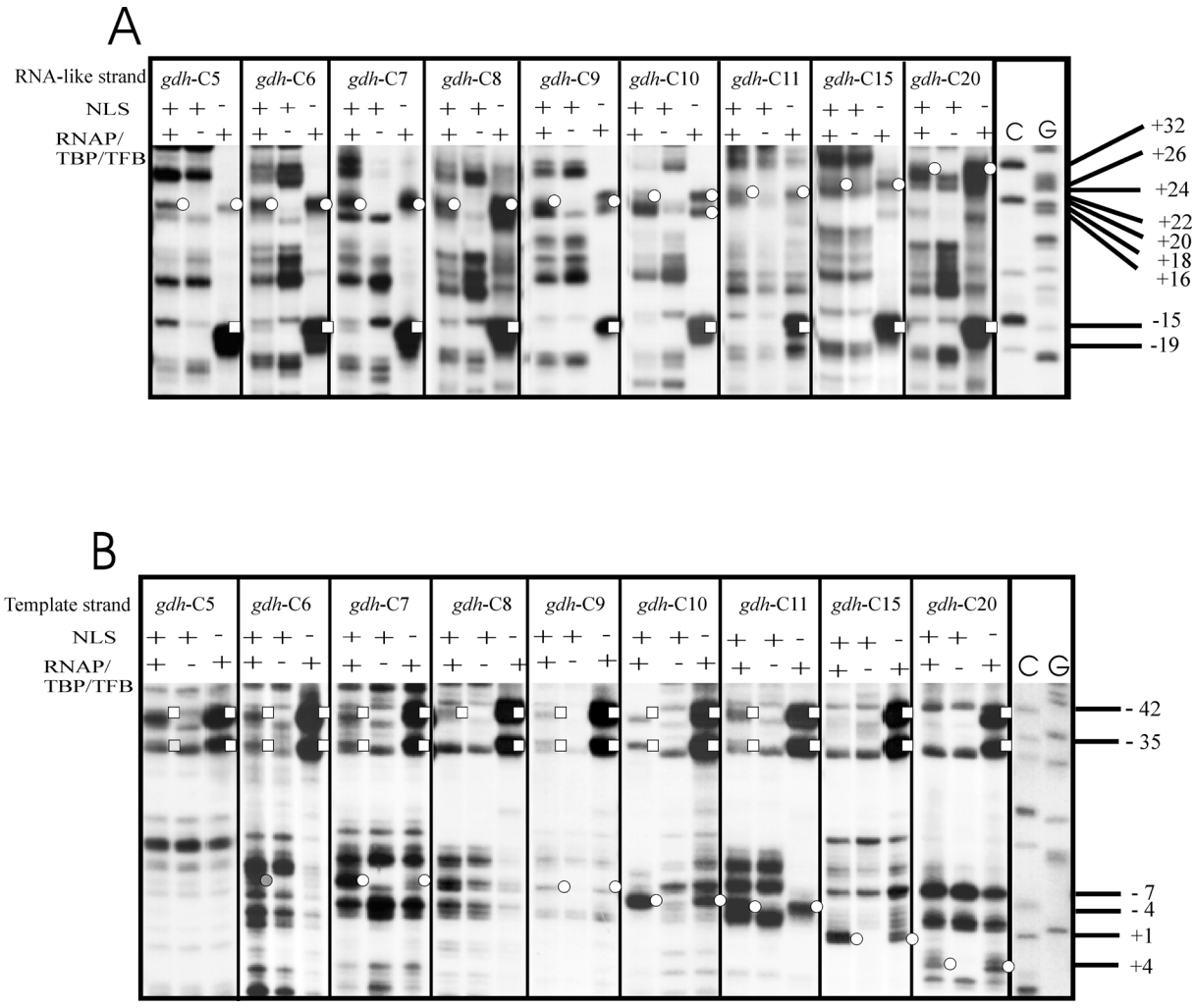

FIG. 3. **Exonuclease III footprints of stalled transcription complexes.** The complexes were stalled and subjected to treatment with exo III as described under "Experimental Procedures". The footprints on the RNA-like **A** and at the template strand **B** were analyzed on a 6% denaturing sequencing gel alongside a sequence ladder. The templates above refer to the registers in which the RNA polymerase was stalled. The anionic detergent N-lauroyl-sarcosine (NLS) was used to remove TBP and TFB from the DNA while the binding of RNAP remained stable. Using NLS produced a stronger background pattern in most cases (see control lanes RNAP/TBP/TFB -). The TBP/TFB footprints are marked with boxes. The downstream **A** and the upstream end **B** of the RNAP are indicated by circles. The positions relative to the transcription start site defined on a sequence ladder are indicated at right. The results are summarized in Fig. 5.

When the downstream boundary of the complex stalled at position +5 was analyzed, two distinct signals not present in the control reaction were identified (Fig. 3A, left panel). The strong diffuse band located between positions -19 to -15 corresponds approximately to the downstream end of the TBP/TFB footprint identified in the *Methanococcus* and *Sulfolobus* system by DNase I protection analyses (25, 33). The second signal at position +18 corresponds to the downstream edge of RNA polymerase identified in the *Methanococcus* and *Sulfolobus* system. After addition of N-lauroyl sarkosine (NLS) to the complexes the exo III stall site at position -19/-15 was no longer detected whereas the second signal at position +18 was not sensitive to NLS treatment (Fig. 3A, left lane in left panel). We therefore conclude that NLS removes TBP/TFB from the template whereas the archaeal RNA polymerase in ternary complexes remains associated with DNA like eukaryotic polII (14). Consecutive elongation of RNA from 5 to 8 nt did not cause movement of the downstream edge of RNAP (Fig. 3A and summary of footprinting data in Fig. 5). Between registers +9 and +20 the downstream edge of RNAP translocated approximately synchronously with RNA elongation. The downstream end of the RNAP footprint was located at position +20, +22, +24 +26 and +32 in registers +9, +10, +11 +15 and +20 (Fig. 3A and Fig. 5).

In register +5, a distinct upstream boundary of RNA polymerase could not be identified (Fig. 3B left lane). The two signals at positions -42 and -35 are almost identical with the upstream edges of the TPB/TFB and TBP DNaseI footprint at the *Pyrococcus gdh* promoter (27). In all archaeal systems investigated the upstream edge of RNAP could not be directly determined in preinitiation complexes. Addition of RNA polymerase to TBP/TFB promoter complexes caused extension of the protection patterns downstream but not upstream of the TBP/TFB binding site (33, 25). However, in complexes stalled between position +7 and +9 an upstream edge of RNAP could be identified at position -7 (Fig. 3B and Fig. 5). This finding indicates that a structural transition occurred in the early elongation complex stalled between +7 and +9. The upstream edge of complexes stalled at positions +10, +15 and +20 was located at position -4, +1 and +4 (Fig. 3B and Fig. 5). This finding indicates continuous movement of RNAP with the extension of the RNA chain between registers +10 to +20. To analyze the events during initiation and elongation in more detail, the open region and transcription bubble extension in stalled complexes were analyzed in addition.

Open complex, transcription bubble progression and RNA-DNA hybrid - To investigate open complex formation and transcription bubble extension we used potassium permanganate ($KMnO_4$) as a probe specific for thymidine (T)-residues in single stranded DNA. To

investigate the temperature dependence of open complex formation TBP/TFB were incubated individually or in combination with RNAP with end-labelled DNA-fragments containing the *gdh* promoter (Experimental Procedures) at 50 and 70 °C. Transcription reactions on linear templates were usually conducted at 70 °C and as expected T-residues in the region of the transcription start site were modified by $KMnO_4$ treatment (Fig. 4A and B, left panel and Fig. 5, upper panel) at 70°C. No $KMnO_4$ footprint was observed at 50 °C or when TBP/TFB alone were incubated at both temperatures with these templates (data not shown).

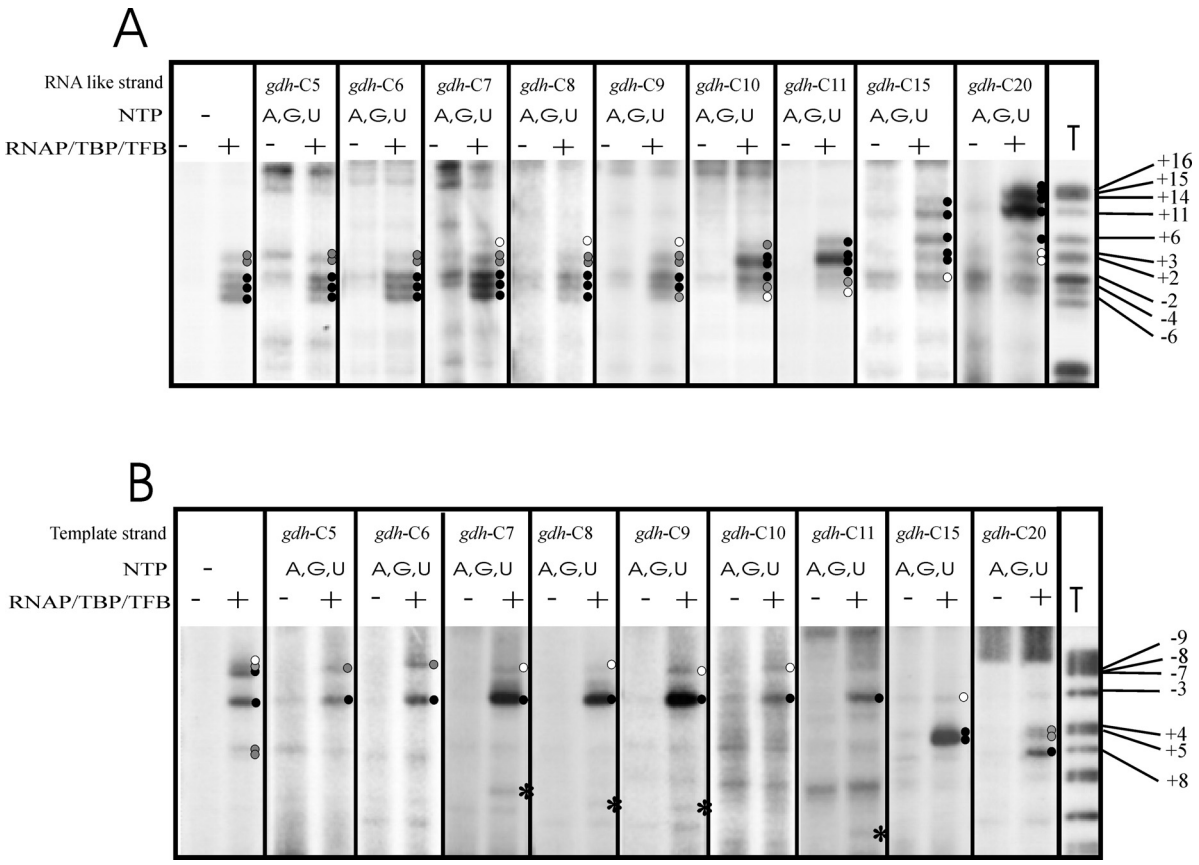

FIG. 4. **Mapping of the transcription bubble in stalled archaeal transcription complexes**. The single strand specific reagent $KMnO_4$ was applied as to stalled and isolated complexes as described under "Experimental Procedures". After piperidine treatment unpaired thymines on the RNA-like strand **A** and on the template strand **B** were analyzed on a 6% denaturing sequencing gel. The control lanes (RNAP/TBP/TFB -) show the patterns obtained on DNA without RNA polymerase. The templates above refer to the register in which the polymerase was stalled. Numbers beside the T sequence refer to the position relative to the transcription start site. Modified bases were indicated by black, gray and white circles referring to the intensity of modification. Modifications downstream of the stall position are marked by asterisks. Strand asymmetry of chemical modification within the estimated transcription bubble is consistent with an RNA-DNA hybrid. The results are summarized in Fig. 5.

These findings indicate that the RNA polymerase was required for strand separation at the promoter and that the open complex was not formed at 50 °C although ternary complexes can be elongated at temperatures down to 37 °C. Five T-residues at positions -6, -4, -2, +2 and +3 were modified on the RNA-like strand (Fig. 4A, left panel) and 2 residues at -7 and -3 were strongly modified on the coding strand (Fig. 4B, left panel). Five T-residues at positions -6, -

4, -2, +2 and +3 were modified on the RNA-like strand (Fig. 4A, left panel) and 2 residues at -7 and -3 were strongly modified on the coding strand (Fig. 4B, left panel). Additional T-residues with increased sensitivity to $KMnO_4$ were identified at positions -8, and -9 and at +4 and +5 on the coding DNA strand. These data indicate that the open complex extends from -9 to +5 at the *Pyrococcus gdh* promoter.

To investigate progression of the transcription bubble, transcription reactions with the templates shown in Fig. 1 were conducted at 70 °C, the ternary complexes isolated at 20 °C, and the $KMnO_4$ reactivity of T-residues in stalled complexes was analyzed at 45 °C. On the RNA-like DNA strand, the modification pattern of the complex stalled at position +5 was basically the same as in the preinitiation complex. On the coding DNA strand T-residues at -7 and -3 were modified. The T-residue at position +4 showed no sensitivity to $KMnO_4$ (Fig. 4B). This lack of reactivity of T-residues close to the 3´-end of nascent RNA was often observed on the coding strand (see Fig. 5 and below) of ternary transcription complexes. We conclude that these T-residues are hybridized with nascent RNA and thereby protected from modification with $KMnO_4$. This protection of T-residues at the coding strand was used for an estimation of the size of the RNA-DNA hybrid (see below). The finding that the T-residues at -8 and -9, in contrast to the T-residues at the same position in the open complex, were not sensitive to $KMnO_4$ in register +5 indicates reclosure of the open complex at these positions and movement of the upstream edge of the transcription bubble (see Fig. 5). In complexes stalled at positions +7, +8, +9 and +11 a modified T-residue two positions downstream of the stall site was detected on the template strand. By contrast, the T-residue 2 nucleotides downstream of the NMP addition site was not modified in complexes stalled at position +6 (see Fig. 4B). This finding indicates that the single-stranded region in the transcription complex can extend beyond the 3´-end of nascent RNA in complexes stalled in registers higher than 6. The complexes stalled at positions +7, +8, and + 9 showed very similar reactivity towards $KMnO_4$ on the coding DNA strand (Fig. 4) and most T-residues of the RNA-like strand, but the reactivity of the T-residue at position -6 was decreased indicating reclosure of the transcription bubble

A clear change in the modification pattern was observed on the RNA-like strand in registers +10 and +11 (Fig. 4A). The $KMnO_4$ reactivity at positions -6 and -4 was reduced, the reactivity of the T-residues at +2 and +3 dramatically increased and modification of the T-residue at position +6 was clearly increased. These findings indicate a major conformational change and movement of the transcription bubble in these registers.

Stalling RNAP at position +15 results in sensitivity of two T-residues towards $KMnO_4$ on the RNA-like strand at positions +14 and +11, which were not modified in register +11 (Fig. 4A). In addition, the T-residues at positions +6 and +3/+2 were modified. The T-residues at positions -4 and -6 showed no reactivity in this complex. On the coding strand in register +15 also strong changes in the modification pattern were detected (Fig. 4B). T-residues at position +4 and +5 showed strong reactivity, the signal at -3 was drastically reduced in intensity, modification of the T-residues at positions -7 and -8 could not be detected (Fig. 4B).

At register +20, again a very significant change of the modification pattern was observed. On the RNA-like strand, the T-residues at positions +16, +15, +14 and +11 showed reactivity towards $KMnO_4$, in addition, the T-residue at position +6 was sensitive (Fig. 4A). The T-residues at position +3/+2 showed reduced reactivity. On the coding DNA-strand, the reactivity of T-residues at -3 was eliminated, the sensitivity of T-residues at positions +4/+5 was reduced and a novel T-residue at +8 was modified (Fig. 4B). We estimate the open region at this stall site from +4 to +20, approximately 16 bp in length. An estimate of the extension of the transcription bubble in each register of transcription is given in Fig. 5.

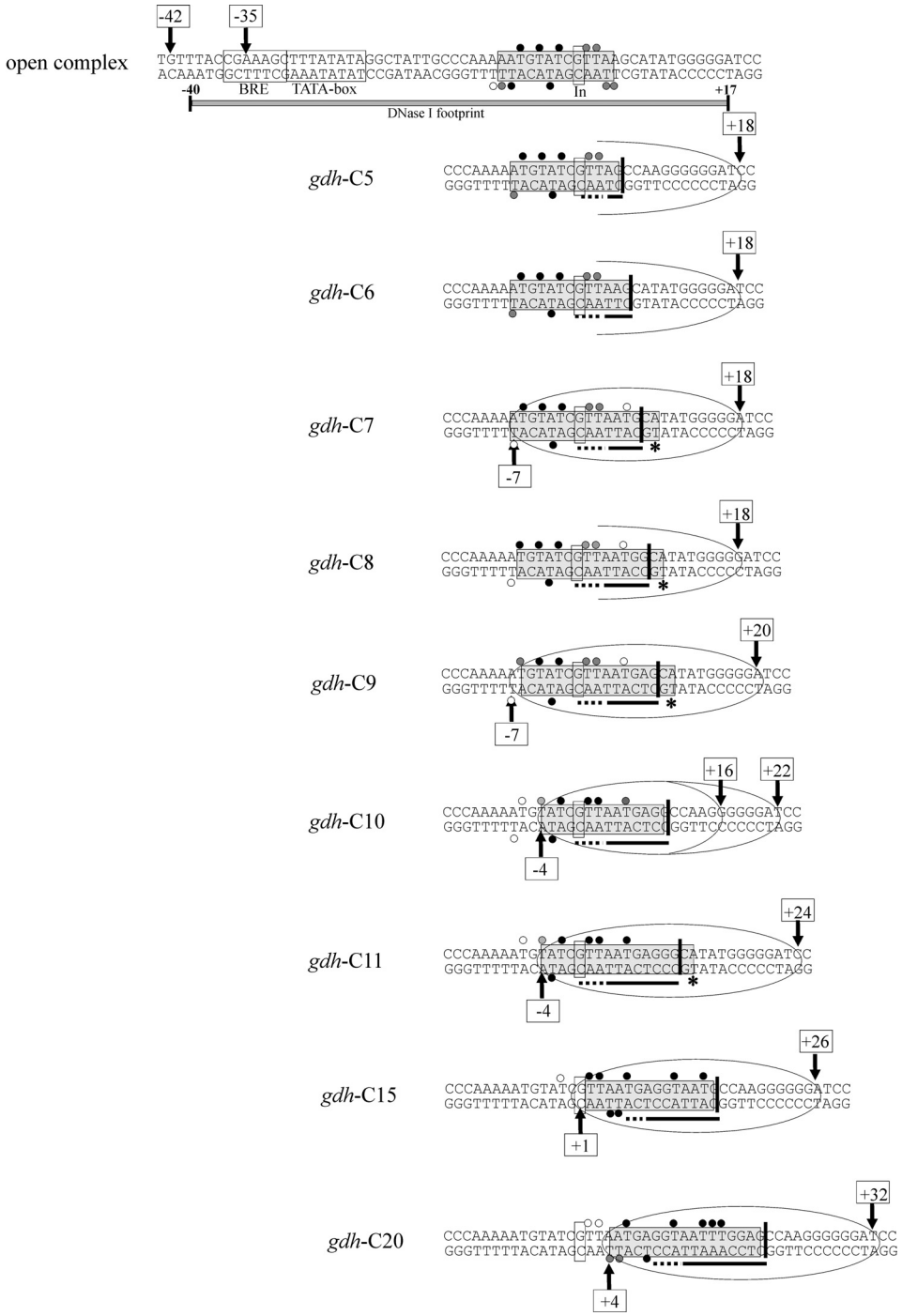

FIG. 5. **Map of exo III footprints, single stranded DNA regions and of the RNA-DNA hybrid in early archaeal elongation complexes.** The DNA sequences containing the transcription initiation region are shown for each stalled transcription complex and for the open complex the promoter is also depicted. The designation of templates on the left refer to the stall position. The TATA-box, the BRE and the transcription initiation site (In) are boxed. The black vertical lines indicate the stall positions. The limits of the RNA polymerase as defined by exo III footprinting were marked with arrows. The numbers refer to the nucleotide relative to the transcription start site. Black circles show the nucleotides with strong reactivity to $KMnO_4$, gray those with less and white circles those with weak reactivity to $KMnO_4$. Modifications downstream of the stall position are marked by asterisks. The gray box is the estimated extent of the transcription bubble. DNA strands were regarded as separated when the base in either one of the two strands was accessible to $KMnO_4$ modification. Strand asymmetry in chemical modification is due to an RNA-DNA hybrid. The portions of DNA supposed to be in a hybrid with RNA are underlined. Because of the sequence dependence of $KMnO_4$ the extent of the RNA-DNA hybrid can not be determined exactly (dotted line). For the same reason the size of the transcription bubble is a lower estimate in most cases.

The results presented here allowed also an estimate of the nascent RNA-DNA hybrid length. One striking example is the complex stalled at position +20. On the RNA-like strand, the T-residues at position +11, +14, +15 and +16 were clearly modified (Figs. 4A and 5). Therefore, the T-residues on the opposite strand at positions +12, +13 and +19 (see DNA sequence in Fig. 5) must be located within the melted DNA region. However, these T-residues on the coding strand showed no $KMnO_4$ signal (Fig. 4B). Considering this finding we suggest that this protection is due to an RNA-DNA hybrid of at least 9 nt. This is a minimal estimate as the next modified T-residue on the coding strand which is not protected by hydrogen bonding to adenine in RNA is located at position +8. Therefore, the length of the RNA-DNA hybrid may extend up to12 nucleotides (indicated by dotted lines in Fig. 5). In such a way, the length of the RNA hybrid was estimated in each register of transcription (summarized in Fig. 5).

DISCUSSION

Experimental design – We have investigated the movement of an archaeal RNAP and transcription bubble extension during transition from initiation to elongation using a series of complexes stalled between positions +5 and +20. The analysis of the limits of RNAP with exo III footprinting and of the melted DNA region with $KMnO_4$ footprinting was coupled with analyses of the formed RNA products. The templates did not contain C-residues up to the stall sites and therefore omitting CTP from transcription reactions was expected to cause stalling of RNAP at the desired positions. We have developed a stalling protocol involving short incubation times and rapid isolation of 5′-biotinylated templates (Fig. 1) by the use of streptavidin coated magnetic particles and a magnet that yielded ternary complexes containing RNAs of the correct size as major products. In registers between +7 and +20 always a second ternary complex was isolated that contained an RNA product of 5 nt (Fig. 2A) We assume that these complexes are paused at position +5. The existence of these complexes stalled close to the transcription start site complicated the interpretation of $KMnO_4$ footprinting data indicating reclosure of the open region at the upstream edge of the bubble during elongation but did not interfere with analyses of the upstream and downstream limits of RNAP and our analyses of extension of the transcription bubble at the downstream border. All the complexes isolated in this study were transcriptionally active and not arrested since addition of a complete set of NTPs resulted in elongation of these nascent RNAs in ternary complexes to run-off transcripts (Fig. 2A).

Three different conformations of RNAP and two distinct structural transitions were observed during early steps of archaeal elongation – The exo III footprinting data presented here and previous results suggest that the conformation of RNAP does not change during synthesis of the first five nucleotides. The exo III borders of the complex stalled at position +5 analyzed in this study (Fig. 3 and summary in Fig. 5) are basically the same as the limits of the PIC determined in the Pyrococcus and other archaeal systems by DNase I footprinting (27, 22, 25). One striking property shared between the complex stalled at position +5 and the PIC is that an upstream limit of the RNAP binding site cannot be defined. This finding suggests that the RNAP is in close contact with the transcription factors TPB/TFB assembled around the TATA-box/ BRE promoter elements (Fig. 5 top) in the PIC and in complexes stalled at position +5. In addition, also the downstream limit of the RNAP binding site is the same as in the PIC. These findings indicate that the RNAP does not move during synthesis of the first 5 nucleotides.

Our exo III footprinting data indicate that two distinct structural transitions occur between registers +6 and +20. The first transition was observed between registers +6 and +7. The RNAP seems to undergo a conformational change and /or to start translocation indicated by the presence of an RNAP induced exo III stop signal at position -7 (Fig. 3B and Fig. 5). Beyond register +6 an extension of the transcription bubble 2 nucleotides downstream of the NMP addition site was observed by $KMnO_4$ footprinting (see Fig. 4B and 5; T-residues labelled by an asterisk). Thus, two independent methods indicate that a structural transition occurs in complexes stalled at registers +6/+7. The conformation of these complexes is characterized in addition by an unchanged downstream edge of the exo III footprint that is located at position +18. Although we could not detect an upstream boundary of RNAP in register +8 (Figs. 3B and 5) the upstream edge of RNAP was consistently located at position -7 in registers +7 and +9. We therefore assume that the RNAP binding site in the first transition state extends from positions -7 to +18 over a DNA segment of 25 bp (Fig. 5).

The second clear structural transition occurs in complexes stalled at positions +10 and +11. Here, the downstream part of RNAP starts translocation and this movement continues synchronously with RNA elongation up to the stall position at +20 (Figs. 3A and 5). In each case the distance between the 3′-end of RNA and the downstream edge of RNAP was approximately 12 bp (Fig. 5). A somewhat longer but also constant distance has been found in active eukaryotic (34) and prokaryotic transcription complexes (35). Exo III borders very close to the site of NMP addition are characteristic for backtracking of RNAP and arrested complexes (34). Stalling *Pyrococcus* RNAP at position +10 produced beside the signal at

position +22 a second exo III pausing site at position +16 that was significantly closer to the 3´- end of the RNA. This signal at +16 is likely to be due to backtracking of RNA polymerase. This second complex at stall site +10 seems not to be arrested since all RNAs isolated in ternary complexes stalled at +10 could be chased by the addition of NTPs (Fig. 2A, lane 12 and Fig. 2B, lane 6). The finding that the distance between the 3´-end of the transcript to the leading edge of RNAP is constantly 12 bp supports our former conclusion that all isolated complexes were transcriptionally competent. The upstream end of complexes stalled between position +10 and +20 could also be clearly identified in each case and move also continuously with RNA elongation (Figs. 3B and 5). The complexes stalled at +11, +15 and +20 are characterized by coordinate movement of the active site and both the leading and trailing edges of RNAP.

Movement of transcription bubble and RNA-DNA hybrid – The conclusions inferred from exo III footprinting were confirmed and extended by analyses of transcription bubble extension in stalled complexes by $KMnO_4$ footprinting. The open region in the PIC was formed in a temperature-dependent manner and extended from position –9 to +5. Considering the limits of the method this is very similar to the open region in the PIC of a $tRNA^{Val}$ promotor of *Methanococcus* and an rRNA promoter of *Sulfolobus* which ranged from -11 to -1 and -12 to -1, respectively (25, 36). Up to register +6 no significant movement of the downstream edge of the bubble could be detected (Figs. 4 and 5). Since the three T-residues at positions -7 to -9 showed reduced activity towards $KMnO_4$ the process of reclosure of the open region at the upstream end was visible after synthesis of 5 and 6 nucleotides (Figs. 4B and 5). Only weak variations of the $KMnO_4$ sensitivities were observed between registers +7 and +9 (Figs. 4 and 5). The most significant change in these early registers was the extension of the length of the transcription bubble from 12 nt in register +5 to 17 in register +9 (Figs. 4 and 5). A more dramatic change of the $KMnO_4$ sensitivity patterns at positions +10 and +11 indicated a structural transition of the bubble at these stall sites. From analyses of the modifications patterns in complexes stalled at positions +10 +15 and +20 clearly movements of the bubble at the upstream and downstream edges could be inferred. This finding indicated that in early archaeal elongation complexes both the upstream and downstream edge of the bubble move synchronously with RNA synthesis. In complexes stalled at positions +11, +15 and +20 the open region encompasses 17, 14 and 17 nt in length. Since estimation of the bubble size by $KMnO_4$ footprinting depends on the presence of T-residues, the open region indicated by the

gray box in Fig. 5 is a lower estimate. We assume that the bubble size is at least 17 nt for complexes stalled between positions +10 and +20.

Analysis of the extent of the open region was complicated by the existence of complexes paused at +5 (visible in the lower part of Fig. 2A) which could cause additional $KMnO_4$ signals in the region of the transcription start site which were not part of the moving transcription bubble. However, careful inspection of the $KMnO_4$ modification patterns allowed clearly to define major transitions during translocation of the bubble. The formation of a hybrid between the growing RNA chain and the template DNA strand complicated an exact determination of the downstream limit of the bubble at the coding DNA strand. But, when the $KMnO_4$ modification patterns on both DNA strands and the weak $KMnO_4$ sensitive signals beyond the site of NMP addition (indicated by an asterisk in Fig. 4B) on the coding DNA strand were considered, it was possible to infer both the extent of the open region and the extension of the RNA-DNA hybrid.

The RNA-DNA hybrid grew continuously with RNA elongation in early registers of transcription. It was at least two nt in register +5, three in register +6, four in register +7, five in register +8, six in register +9, seven in register +12 and eight in register +11 (Fig. 5). When RNAP was stalled at position +15, the length of the RNA-DNA hybrid was at least 8, at stall site +20 at least 9 nt. The finding that the T-residue at positions +8 on the coding DNA strand was clearly modified and therefore not base paired with adenine in RNA in complexes stalled at position +20 indicates that the RNA-DNA hybrid encompasses not more than 12 bp (Fig. 5). We therefore conclude that the length of the RNA-DNA hybrid is between 8 and 12 bp during early elongation of archaeal transcription.

Comparison of mechanistic characteristics of archaeal, eukaryotic and bacterial RNAP – Although the basic mechanism of transcription and general structure of RNAP are highly conserved among the three kingdoms of life also distinct mechanistic and structural differences exist. The data described here provide first evidence for the dimensional parameters of a transcribing archaeal RNAP. As discussed in this paper the archaeal PIC and the *Pyrococcus* complex stalled at +5 are likely to extend over the DNA region from -42 to +18. A very similar DNA section extending form -55 to +18 is protected in the open complex formed by *E. coli* RNAP (37, 17). The upstream part from -55 to -14, designated as recognition domain is only partially protected in the *E. coli* open complex. After synthesis of 11 bp this recognition domain is completely dissociated from the DNA whereas the size of the second DNA-domain, the melted domain, remains constant (37). The recognition domain of

DNA bound by *E. coli* RNAP seems to be associated with the transcription factors TBP/TFB in the archaeal system. A major transition at registers +10/+11 was also observed in both systems. In *E. coli,* this transition is characterized by the dissociation of Sigma and the complex extends at this register from -3 to +27. The archaeal RNAP has also initiated promoter clearance at this register and extends over 28 bp from position -4 to +24. Thus, the overall dimension of the archaeal and bacterial complex stalled at +11 are very similar. A further contraction of *E. coli* RNAP binding site stalled at register +20 to 22 bp has been observed (37). By contrast, an RNAP binding site of 29 bp was found in archaeal complexes stalled at register +20. This exo III footprints of archaeal complexes stalled at +20 equal footprints of polII stalled between registers +20 and +23 which extend over 31-35 bp (38).
The distance of the catalytic center C to the front edge F of the footprint is constant in active and not retracted bacterial and polII complexes. The archaeal C-F values determined here were also constant at various registers but with 11-12 nucleotides shorter than in the bacterial (C-F= 18; (35)) and polII system (C-F=18-20; (34)).
In all domains of life a characteristic mechanistic similarity is the transition around register +10. At this point all RNAP seems to reach the elongation-commited state. We have not studied abortive products of the archaeal enzyme here but have clearly shown that complexes containing 5-9 nucleotides can be isolated and are fully elongatable. This is a common property of the archaeal enzyme and polII. By contrast *E. coli* RNAP which, at most promoters, is in an initiation state very similar to the open complex until position +10 and produces reiteratively abortive products in early registers without release and rebinding of RNAP.
From analyses of translocation of the transcription bubble three characteristic transitions have been postulated in the polII system (16, 15). The first transition is open complex formation. Similar to the archaeal system the eukaryotic open complex ranges from -9 to +2. By contrast the archaeal RNAP is able to catalyze DNA-strand separation in the absence of TFIIH helicase activity and ATP (25; Fig. 4). We have no evidence that the second transition in the polll initiation complex at register +4 characterized by insensitivity of the complex to ATPγS (15) occurs also in the archaeal system. In the polII system the region of the initially open complex readopts the double stranded conformation between registers +9 to +11 and this was described as the third transition. Reclosure of the most upstream part of the archaeal open complex was also observed in these registers. Both the archaeal RNAP and polII seem to start promoter clearance around register +10. In *E. coli* and polII complexes continuous opening of the downstream part of the open region and discontinuous reclosure of the upstream part have

been described (18, 15). The data shown here seem to indicate that failure to observe continuous reclosure of the upstream end may be not a mechanistic property of the elongation process but rather due to the presence of additional complexes stalled at an earlier register which might mask reclosure at the upstream edge (Fig. 2A).
The size of the RNA-DNA hybrid is with 9-12 bp in a similar range as in eukaryotic and prokaryotic elongation complexes analyzed by comparable methods (20, 39, 40).

Acknowledgements - We thank E. Zaychikov and H. Heumann (Max-Planck-Institut für Biochemie, Martinsried) for advising S. Francois (University of Kiel) in $KMnO_4$ footprinting at high temperatures. We would like to thank W. Hausner and U. Lange (Universität Kiel) for valuable advice and discussions. This work was supported by a grant of the Deutsche Forschungsgemeinschaft and of the Fonds der Chemischen Industrie to M. Thomm.

REFERENCES

1. Heumann, H., Zaychikov, E., Denissova, L., and Hermann, T. (1997) in *Nucleic Acids and Molecular Biology, Vol. 11*, (Eckstein, F., Lilley, D. M. J. eds) pp 151- 177, Springer-Verlag, Berlin Heidelberg.
2. Fiedler, U., and Timmers, H. Th. M. (2000) *BioEssays* **22**, 316-326
3. Kassavetis, G. A., Blanco, J.A., Johnsen, T.E., and Geiduschek, E. P. (1992) *J. Mol. Biol.* **226**, 47-58
4. Kassavetis, G. A., Letts, G. A., and Geiduschek, E. P. (1999) *EMBO J.* **18**, 5042-5051
5. Kahl, B. F., Li, H., and Paule, M.R. (2000) *J. Mol. Biol.* **299**, 75-89
6. Hausner, W., and Thomm, M. (2001) *J. Bacteriol.* **183**, 3025-3031
7. Bell, S. D., Jaxel, C., Nadal, M., Kosa, P. F., and Jackson, S. P. (1998) *Proc. Natl. Acad. Sci. U.S.A.* **95**, 15218-15222
8. deHaseth, P. L., and Helmann, J. D. (1995) *Mol. Microbiol.* **16**, 817-824
9. van Hippel, P.H., Yager, T.D., and S.C. Gill (1992) in *Transcription Regulation* (McKnight, S. and Yamamoto, K.eds) pp. 179-201, Cold Spring Harbor Laboratories, Press, Cold Spring Harbour
10. Jiang, Y., Smale, S.T., and Gralla, J.D. (1993) *J. Biol. Chem.* **268**, 6535-6540
11. Kustu, S., Santero, E., Keener, J., Popham, D., and Weiss, D. (1989) *J. Microbiol.* **53**, 367-376

12. Timmers, H. Th. M. (1994) *EMBO J.* **13**, 391-399

13. Kassavetis, G. A., Blanco, J. A., Johnsen, T. E., and Geiduschek, E. P. (1992) *J. Mol. Biol.* **226**, 47-58

14. Cai, H., and Luse, D. S. (1987) *J. Biol. Chem.* **262**, 298-304

15. Holstege, F. C. P., Fiedler, U., and Timmers, H. Th. M. (1997) *EMBO J.* **16,** 7468-7480

16. Yan, M., and Gralla, J. D. (1997) *EMBO J.* **16,** 7457-7467

17. Krummel, B., and Chamberlin, M. J. (1992) *J. Mol. Biol.* **225**, 239-250

18. Zaychikov, E., Denissova, L., and Heumann, H. (1995) *Proc. Natl. Acad. Sci. USA* **92,** 1739-1743

19. Komissarova, N., and Kashlev, M. (1997) *J. Biol. Chem.* **272,** 15329-15338

20. Nudler, E., Mustaev, A., Lukhtanov, E., Goldfarb, A. (1997) *Cell* **89**, 33-41

21. Thomm, M. (1996) *FEMS Microbiol. Rev.* **18,** 159-171

22. Bell, S. D., and Jackson, S.P. (2001) *Curr. Opin. Microbiol.* **4**, 208-213

23. Bell, S.D., Brinkman, A.B., van der Oost, J., Jackson, S.P. (2001) *EMBO reports* **21,** 133-138

24. Hanzelka, B. L., Darcy, T., and Reeve, J. N. (2001) *J. Bacteriol.* **183**, 1813-1818

25. Hausner, W., and Thomm, M. (2001) *J. Bacteriol.* **183**, 3025-3031

26. Hethke, C., Geerling, A.C., Hausner, w., de Vos, W. M., and Thomm, M. (1996) *Nucleic Acids Res.* **24**, 2369-2376

27. Hausner, W., Wettach, J., Hethke, C., and Thomm, M. (1996) *J. Biol. Chem.* **271,** 30144-30148

28. Bartlett, M. S., Thomm, M., and Geiduschek, E. P. (2000) *Nat. Struct. Biol.* **7**, 703-705

29. Dahlke, I., and Thomm, M. (2002) *Nucleic Acids Res.* **30**, 701-710

30. Vierke, G., Engelmann, A., Hebbeln, C., and Thomm, M. (2003) *J. Biol. Chem.* **278**, 18-26

31. Carpousis, A. J., and Gralla, J. D. (1980) *Biochemistry* **19,** 3245-3253

32. Fiedler, U., and Timmers, H. Th. M. (2001) *Nucleic Acids Res.* **29,** 2706-2714

33. Bell, S. D., and Jackson, S.P. (2000) *J. Biol. Chem.* **275**, 12934-12940

34. Samkurashivili, I., and Luse, D. S. (1998) *Mol. Cell. Biol.* **18,** 5343-5354

35. Nudler, E., Goldfarb, A., and Kashlev, M. (1994) *Science* **265**, 793-796

36. Bell, S. D., Jaxel, C., Nadal, M., Kosa, P. F., and Jackson, S. P. (1998) *Proc. Natl. Acad. Sci. U.S.A.* **95**, 15218-15222

37. Metzger, W., Schickor, P. and Heumann, H. (1989) *EMBO J.* **8**, 2745-2754

38. Samkurashvili, I., and Luse, D. S. (1996) *J. Biol. Chem.* **271**, 23495-23505

39. Sidorenkov, J., Komissarova, N., and Kashlev, M. (1998) *Mol. Cell* **2,** 55-64

40. Kahl, B. F., Li, H., and Paule, M.R. (2000) *J. Mol. Biol.* **299**, 75-89

NUCLEIC ACIDS RESEARCH doi:10.1093/nar/gkm1086 published online December 10, 2007

III Structure-function analysis of the RNA polymerase cleft loops elucidates initial transcription, DNA unwinding, and RNA displacement

Souad Naji,[1] Michela G. Bertero,[2] Patrizia Spitalny,[1] Patrick Cramer[2*] & Michael Thomm[1*]

[1]Lehrstuhl für Mikrobiologie und Archaeenzentrum, Universität Regensburg, Universitätsstr. 31, D-93053 Regensburg, Germany

[2]Gene Center Munich and Center for integrated Protein Science CiPS[M], Department of Chemistry and Biochemistry, Ludwig-Maximilians-Universität München, Feodor-Lynen-Strasse 25, 81377 München, Germany.

*Correspondence should be addressed to M.T. (michael.thomm@biologie.uni-regensburg.de) or P.C. (cramer@lmb.uni-muenchen.de).

Phone/Fax: 49-941-943-3160/49-941-943-2403 (Michael Thomm); 49-89-2180-76951/49-89-2180-76999 (Patrick Cramer)

Running Title: Functions of the RNA polymerase cleft loops

ABSTRACT

The active center clefts of RNA polymerase (RNAP) from the archaeon *Pyrococcus furiosus* (*Pfu*) and of yeast RNAP II are nearly identical, including four protruding loops, the lid, rudder, fork 1, and fork 2. Here we present a structure-function analysis of recombinant *Pfu* RNAP variants lacking these cleft loops, and analyze the function of each loop at different stages of the transcription cycle. All cleft loops except fork 1 were required for promoter-directed transcription and efficient elongation. Unprimed *de novo* transcription required fork 2, the lid was necessary for primed initial transcription. Analysis of templates containing a pre-melted bubble showed that rewinding of upstream DNA drives RNA separation from the template. During elongation, downstream DNA strand separation required template strand binding to an invariant arginine in switch 2, and apparently interaction of an invariant arginine in fork 2 with the nontemplate strand.

INTRODUCTION

Multisubunit RNA polymerases (RNAPs) catalyze RNA synthesis from a DNA template during gene transcription. The eukaryotic nucleus contains three RNA polymerases, called RNAP I, II, and III, whereas bacterial and archaeal cells contain only one RNA polymerase. Analysis of gene regulation requires a detailed structure-based understanding of the transcription mechanism. During the transcription cycle, the polymerases first assemble with initiation factors on promoter DNA (closed complex formation). The complex then unwinds the DNA double helix (open complex formation). The polymerase begins to synthesize short RNA oligonucleotides, which are often released (abortive transcription). When the RNA product reaches a certain length, the enzyme enters the elongation phase, characterized by a stable, processive elongation complex. The polymerase then elongates the RNA chain, unwinds downstream DNA, and rewinds upstream DNA. Finally the RNA transcript and the DNA are released during termination.

Detailed crystallographic structures are available for yeast RNAP II and bacterial RNAPs (1-5) and enable mechanistic studies of the transcription cycle by designing mutations. Four prominent loops were revealed above the active site in the polymerase cleft, named the rudder, lid, fork loop 1 and fork loop 2 (2,6). Whereas the rudder and lid protrude from the mobile clamp of the polymerase, the two fork loops are located on the opposite side of the cleft (Figure 1).

The rudder and lid were suggested to maintain the upstream end of the hybrid and the bubble (2,6-8). Functional roles of the rudder and lid were analyzed in the bacterial enzyme (9-11). Mutagenesis of the rudder showed that this element stabilizes the elongation complex but that it is not involved in maintaining the hybrid length (9).The lid was suggested to help separate RNA from DNA at the upstream end of the hybrid (2,3,6,7) but a mutant bacterial RNA polymerase lacking the lid could displace RNA normally (10). It was suggested that fork loop 2 blocks the path of the nontemplate strand before the active site, and thereby helps to separate the DNA strands at the downstream edge of the bubble (6,8). In yeast RNAP II, mutations in the proximity to fork loop 2 have been shown to lower the polymerization rate but no mutational *in vitro* studies on the rudder, lid, or fork loop 1 (12) have been reported.

Recently, recombinant forms of archaeal RNAPs became available, which enable rapid site-directed mutagenesis (13,14). The archaeal enzymes are closely related in sequence to eukaryotic RNAP II (15). All yeast RNAP II subunits have counterparts in the archaeal enzyme, except the small peripheral subunit Rpb8. In the *Pyrococcus furiosus* (*Pfu*) enzyme, 38 percent of the amino acid residues are identical with yeast RNAP II (16; 17). The

similarity of the archaeal and eukaryotic transcription machineries extends to protein-protein interactions of the polymerase subunits (13,16) and initiation factors that are required for promoter binding (18). The archaeal initiation factors TBP, TFB, and TFE have homologs in the eukaryotic RNAP II apparatus, named TBP, TFIIB, and TFIIE, respectively. Whereas the factors TBP and TFB are sufficient to bind and open promoter DNA in the *Pyrococcus* system at 70 °C (19,14) the eukaryotic machinery requires in addition TFIIF to bind the promoter, and TFIIE/TFIIH to open DNA. Despite these differences, the recent success in obtaining highly active recombinant *Pfu* RNAP opens up the possibility to rapidly prepare and functionally analyze mutant RNAP II-like enzymes (14).

Here, we analyzed the function of four recombinant archaeal RNA polymerase deletion mutant enzymes, each lacking one of the four cleft loops. Together with an analysis of three additional mutant enzymes carrying selected point mutations in fork loop 2 and switch region 2, another active center element that was thus far not studied by mutagenesis, our results unravel the functional significance of these elements at various stages of the transcription cycle. In addition, we have used different nucleic acid scaffolds to elucidate the initiation-elongation transition, one of the most dynamic and least understood aspects of the transcription cycle.

MATERIALS AND METHODS

Primer sequences

The sequence of primers used for mutagenesis and PCR are provided in the Supplementary Data.

Construction of subunit B (rpb2) and subunit A´ mutants by site-directed mutagenesis

The rudder and lid domains of subunit A´ and fork loop 1 and fork loop 2 of subunit B were deleted using a two-round, four primer technique. In round 1, two PCR products were generated containing the DNA region upstream and downstream from the deletion in separate reactions. Each PCR was performed using genomic DNA as template, end primers (FwdA and RevD) and a pair of primers flanking the internal sequence to be deleted (RevB and FwdC). The sequences of all primers are provided in supplementary material. The resulting products were analyzed by agarose gel electrophoresis and purified using a QIAquick spin Gel Extraction Kit. Purified DNA fragments were added to a second round of PCR. Fusion of the two intermediates was achieved as a result of overlapping complementary regions in the

products left and right to the deleted sequence formed in round 1. The products paired during the annealing phase of PCR round 2 and were amplified by the addition of primers complementary to the end of each single stranded DNA fragment (primers FwdA and RevD). For the generation of single point mutants, two complementary primers (B-R445A-RevB and B-R445A-FwdC for B-R445A; A´-R313A-RevB and A´-R313A-FwdC for A´-R313A; A´-K306A-RevB and A´-K306A-FwdC for A´-K306A) were used along with end primers (FwdA and RevD) to substitute arginine or lysine with alanine. After analysis of the PCR products by agarose gel electrophoresis, the resulting mutant DNA was purified and ligated into pET151/D-TOPO and transformed into *Escherichia coli*.

Overexpression of recombinant subunits

The subunits were expressed in BL21(DE3) Codon Plus TM by inducing exponentially grown cultures over night at 20 °C with IPTG. The his_6-tagged mutated subunits B and A´ were (like their wild type counterparts) highly insoluble and expressed as inclusion bodies in *E. coli* cells. They were solubilized in 6 M guanidine HCl and immobilized on a NiNTA column. The proteins were refolded on column by washing with a decreasing gradient of urea. The renatured subunits were eluted with imidazole and used for reconstitution of the RNAP as described previously (14).

Reconstitution of mutant and wild-type *Pyrococcus* RNAP

The RNAPs containing mutated and wild type components were reconstituted from 11 bacterially produced subunits after denaturation in TB buffer containing 6M urea and stepwise dialysis against TB buffer containing 3 M urea and no urea. The renatured RNAP assemblies were purified by Superdex 200 chromatography as described previously (14). The protein eluting as a homogenous peak from the Superdex 200 column was analyzed by SDS PAGE and in specific run-off transcription assays using the *Pyrococcus gdh* promoter as template. Fractions containing active RNAP were combined and used for transcriptional analyses.

Cell-free transcription reactions

Promoter-independent assay. Reactions were performed in a total volume of 100 µl containing 9 nM RNAP (or mutant derivative) 900 µM ATP, 90 µM UTP, 0.15 MBq(α-^{32}P)-

UTP (110 TBq/mmol) and 3 µg of poly[(dA-dT)] as template. Reactions were incubated for 30 min at 70 °C and counts insoluble in 5% TCA (w/v) were determined.

Promoter-directed assays. Specific i*n vitro* transcription reactions were essentially conducted as described by (20). *Xba*I digested plasmid pUC19 containing the *gdh* promoter region from -95 to +163 was incubated with 35 nM TBP, 30 nM TFB and 9 nM endogenous RNAP, recombinant RNAP or mutant derivative in 25 µl TB (40 mM Na-HEPES, pH 7.3, 250 mM NaCl, 2.5 mM $MgCl_2$, 0.1 mM EDTA, 5 mM β-mercaptoethanol, 0.1 mg/ml BSA). NTPs were added to 500 µM ATP, GTP, CTP and 10 µM UTP and 0.15 MBq(α-^{32}P)UTP (110 TBq/mmol). The reactions were assembled at 4 °C and started by transfer to 70 °C unless otherwise indicated. The 172 nt run-off transcripts and various abortive transcripts were analyzed in 6 %, 28 % or 20-28 % denaturing polyacrylamide gels as indicated and visualized by phosphoimaging. To construct the preopened bubble (Figure 5) the template and nontemplate strand (10 µM each) were incubated for 2 min at 92 °C and cooled down slowly over night to room temperature. This hybrid (10 µM) was used as template in cell-free transcription reactions. The RNA-DNA template strand hybrid (1a; Figure 6A) mimicking an elongation complex was assembled by heating 10 µM RNA and 5 µM template strand in a volume of 100 µl for 2 min at 92 °C and cooled down slowly overnight to room temperature. 0.5 µl of this mixture was used as template in the transcription assays shown in Figure 6C. Template EC3 (Figure 6A) was assembled by the same procedure and contained in addition to the hybrid RNA-DNA template strand 5 µM of the corresponding non template strand.

Abortive transcription reactions

300 ng double stranded template *gdh*-C15 (21); sequence of the promoter region is shown in Figure 4A) was incubated with transcription factors and *Pyrococcus* RNAP as described under cell-free transcription reactions. In the dinucleotide primed reactions shown in Figure 4 the reaction was started with 40 µM GpU and 10 µM UTP and 0.15 MBq(α-^{32}P)UTP (110 TBq/mmol).

Electrophoretic mobility shift assays

DNA fragments spanning the *gdh* promoter region from -60 to +37 were used as probes in assays with mutated and wild type RNAP as described previously (16). Binding reactions with the *gdh* promoter contained in addition TBP and TFB. The binding reactions (25 µl) contained 9 nM endogenous or 27 nM recombinant RNAP (or mutant derivatives), 100 nM

TBP, 100 nM TFB, 0.5 nM *gdh* promoter and 1 µg of poly[dI-dC)] as non specific competitor DNA.

Permanganate footprinting

Thymidine residues in the open complexes formed by wild type and mutant forms of RNAP in preinitiation complexes of the *gdh* promoter containing TBP and TFB were detected by treatment with potassium permanganate as described in the legend of Figure 3.

RESULTS

Design and preparation of Pfu RNA polymerase variants

Due to the high sequence conservation between the *Pfu* polymerase and yeast RNAP II, the *Pfu* lid, rudder, fork loop 1 and fork loop 2 were easily identified in a sequence alignment of the largest subunits (Figure 1). Whereas the loops themselves are not well conserved, the regions immediately preceeding and following the exposed loops are highly conserved, enabling unambiguous definition of the loop borders. The complete RNAP II elongation complex structure (8) guided the design of four *Pfu* RNAP deletion mutants that lack the lid (residues 224-236 in subunit A'), rudder (residues 281-300 in subunit A'), fork loop 1 (residues 413-420 in subunit B) and fork loop 2 (residues 442-452 in subunit B). The deletions were designed such that small stretches of amino acids were left intact and the ends resulting from truncation were connected without perturbation of the structure in the surrounding protein region (non-disruptive mutations).

In addition, we prepared three enzyme variants carrying alanine point mutations. One mutant modifies a conserved basic residue in fork loop 2, R445 in the B subunit, corresponding to Rpb2 residue R504 in RNAP II, and two mutants target basic residues in switch 2, K306 and R313 in the A' subunit, which correspond to Rpb1 residues K330 and R337, respectively, in RNAP II.. These two residues are highly conserved in the switch 2 element across the three domains of life (Figure 1A**,** lower panel).

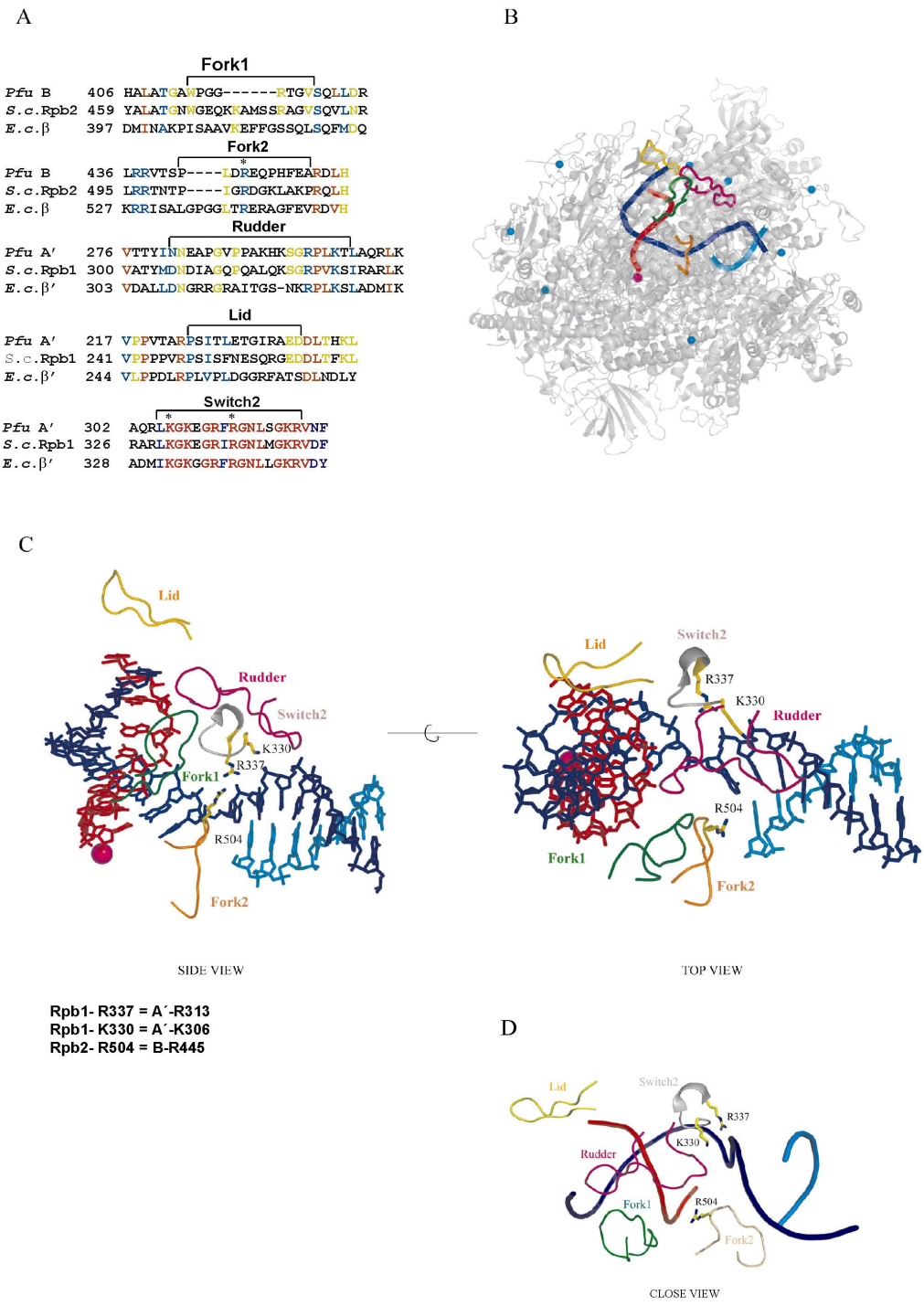

Figure 1. Design of the loop deletions and single point mutations. (**A**) Internal deletion of B and A´ subunits. Primary sequence alignment between the four loops fork1, fork 2, rudder and lid from *P. furiosus*, *S. cerevisiae* and *E. coli* (CULSTAL W). Invariant, conserved and weakly conserved residues are colored in red, blue and yellow, respectively. The extent of deletions and the single point mutations introduced in fork loop 2 and switch 2 are indicated and highlighted by asterisks. (**B**) Surface representation of the *S. cerevisiae* RNAP II elongation complex (8). Template DNA, nontemplate DNA, RNA, Mg^{2+} and Zn^{2+} ions are shown in blue, cyan, red, magenta and light blue, respectively. Lid, rudder, fork 1 and fork 2 are represented in yellow, magenta, green and okra, respectively. (**C**) Close up view of the RNAP II active site. Single point mutations are shown. The same color code as in (B) was used. The residues corresponding in the archaeal enzyme to the mutated putative Rpb1 and Rpb2 key residues are indicated. (**D**) Close up view of nucleic acids in the yeast RNAP II active center highlighting the position of single point mutations analyzed in this study, the colour code is like in (B).

Structural analyses of a yeast RNAP II elongation complex suggested that the basic amino acids R337 and K330 might be involved in pulling the template strand upwards in the active site and could be therefore critical for DNA strand separation in the active center (8). A close up view showing the position of the mutated basic amino acids relative to the nucleic acids in the active center is provided in Figure 1D. The four recombinant deletion enzymes, here referred to as Δlid, Δrudder, Δfork1 and Δfork2, and the three single point mutation enzymes, termed B-R445A, A´-K306A, and A´-R313A, were reconstituted as described (14). The seven purified variants showed a size-exclusion chromatography profile identical to the wild-type recombinant RNAP (not shown).

Fork 1 is the only cleft loop not required for promoter-dependent transcription

The functional properties of the *Pfu* RNAP variants were analyzed and compared to the properties of purified endogenous RNAP (20) and recombinant wild-type enzyme. In a non-specific transcription assay that uses poly-d(A-T) as a template, all mutant enzymes showed low activities close to background levels and also the activity of the reconstituted enzyme was low compared to that of the endogenous enzyme (data not shown). Therefore, the activity of the mutant enzymes was not be quantitatively determined in this assay and the mutant enzymes were rather analyzed in more sensitive assays measuring the synthesis of distinct RNA products.

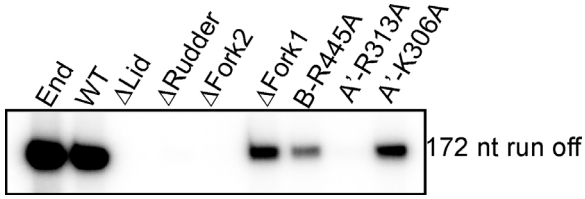

Figure 2. The cleft loops lid, rudder, fork 2 and the conserved arginine residue 313 in switch 2 are required for promoter-dependent initiation. Promoter-dependent assays. Equal amounts (100ng) of Superdex fractions of reconstituted RNAP mutants and of endogenous and reconstituted RNAP were incubated in standard transcription reactions in the presence of TBP (35 nM) and TFB (30 nM) with the *gdh* promoter (6.5 nM) as template. RNA products were analyzed on 6% polyacrylamide gels.

To analyze if the variants have an overall defect in promoter-dependent transcription, we subjected them to a promoter-specific transcription assay that uses as a template the strong *Pfu* glutamate dehydrogenase (*gdh*) promoter (20). The variants were incubated with DNA, TFB, and TBP, and reactions were initiated by addition of nucleoside triphosphates (NTPs). The run-off RNA product was efficiently synthesized by the endogenous and wild-type

recombinant RNAPs, providing a positive control (Figure 2). The Δfork1 RNAP was highly active, showing that fork loop 1 does not have an essential role in transcription. In contrast, Δlid, Δrudder and Δfork2 showed no transcription activity (Figure 2), pointing to an essential function of the lid, rudder and fork 2. Of the single point mutants, B-R445A and A´-K306A showed reduced but still significant activity, whereas A´-R313A was totally inactive. Taken together, our results confirm the functional importance of the rudder and lid, which was previously reported for the bacterial enzyme, and additionally show that fork 2 and switch 2 are required for normal polymerase function.

The rudder is required for open complex formation

The mutant variants were incubated at 37 °C with TBP, TFB, a DNA-fragment containing the *gdh* promoter, and competitor DNA, and were separated in native gels. All variants formed stable closed complexes, except the Δlid variant, which apparently formed large aggregates of unknown composition (not shown). However, the Δlid enzyme formed short transcripts in the presence of the TBP-TFB complex on a pre-melted template (Figure 5B, lane 12) indicating that its ability to interact with DNA and the promoter bound TBP-TFB complex was not generally impaired. In order to analyze the role of the mutated elements in open complex formation we used $KMnO_4$ footprinting, which identifies thymine residues in single-stranded DNA within melted regions. All cleft loop deletion mutants except Δrudder produced a footprint around the transcription start site at 70 °C, the switch 2 mutant R313A was also able to melt DNA in the promoter region (Figure 3). In particular, positions +3/+2, –2 and -4 were accessible to permanganate, indicating that the DNA strands are melted around the transcription start site (position +1). Promoter melting in the -4/+3 region is sufficient for specific initiation at 70 °C (14). These results indicate that the rudder of the archaeal enzyme has an important role in DNA strand separation and/or maintenance of melted DNA as in the bacterial counterpart (9).

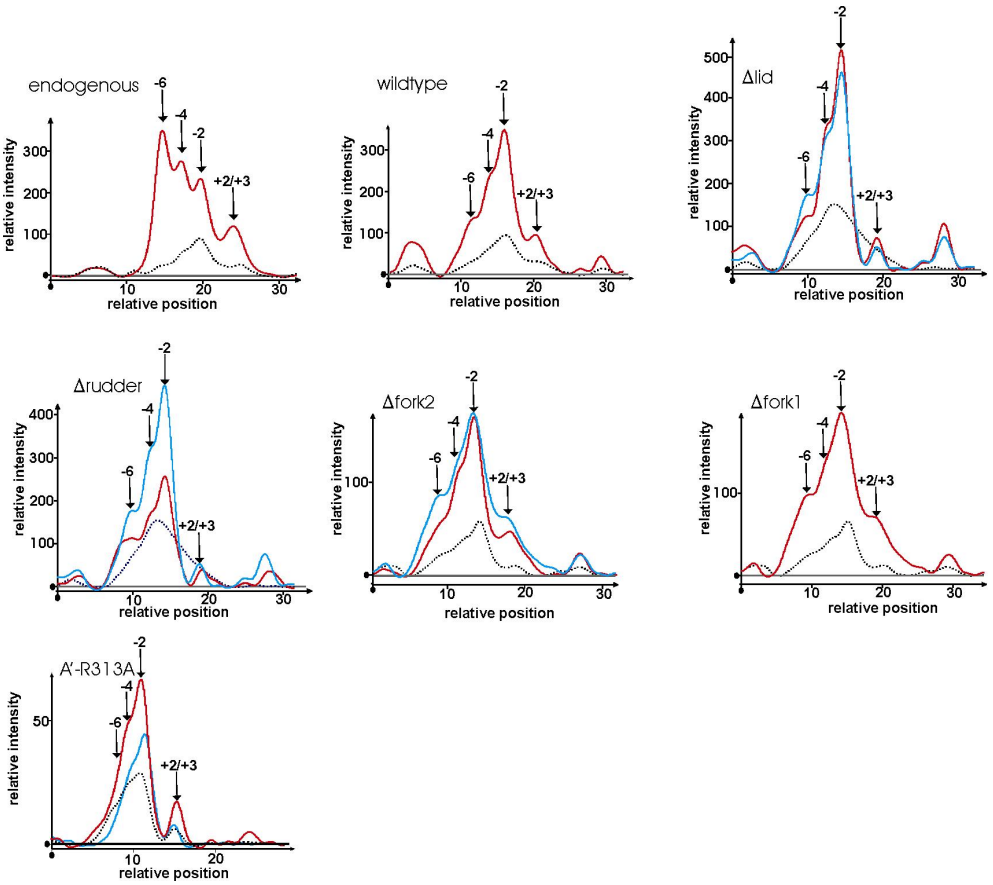

Figure 3. The cleft loops fork 1, fork 2 and lid and R-313 in switch 2 are not essential for the formation of open complexes. The various RNAP preparations and RNAP mutants as indicated in the Figure were incubated with the *gdh* promoter containing a 5´endlabelled template strand in the presence of TBP and TFB for 10 min at 70 °C to allow bubble opening and were probed with potassium permanganate for 5 min at 70 °C (21). Modified thymine residues in single stranded DNA were resolved after piperidine cleavage by electrophoresis in 6% sequencing gels. Phosphoimager traces of the $KMnO_4$ footprinting profiles of the endogenous recombinant and cleft loop polymerase mutants are shown. The dotted line represents the modification pattern in control reactions without RNAP, the red line the permanganate footprint. The blue line represents the signal obtained with WT enzyme in a control reaction on the same gel. The footprint of the endogenous enzyme extends from -6 to +3. This enzyme contains, in contrast to the reconstituted polymerase, nearly stoichiometric amounts of the subunits E´ and F (the archaeal Rpb7/Rpb4 homologues). The presence of the E' subunit in the endogenous enzyme is responsible for the extension of the permanganate footprint to the upstream end (14). The footprints of the reconstituted and mutant enzymes Δlid, Δfork 2 and Δfork 1 extend from +3 to -4. Note that formation of the footprint by the Δrudder enzyme is impaired, in particular at the downstream end.

The lid, rudder, and fork2 are essential for primed transcription

To investigate whether the initial phase of transcription is affected by our loop deletions and point mutations, we carried out an abortive transcription assay, using the *gdh* promoter as a template and the RNA dinucleotide 5'-GpU-3' as a primer (Figure 4A). By addition of the radioactive nucleotide [α-^{32}P]-UTP, we tested whether the polymerase variants were able to elongate the GpU primer to GpUpU. In the absence of TBP and TFB, none of the mutant enzymes was active (not shown). In the presence of TBP and TFB, endogenous and

recombinant *Pfu* RNAPs, the Δfork1 variant, and variants R445A and K306A readily synthesized the trinucleotide product (Figure 4B).

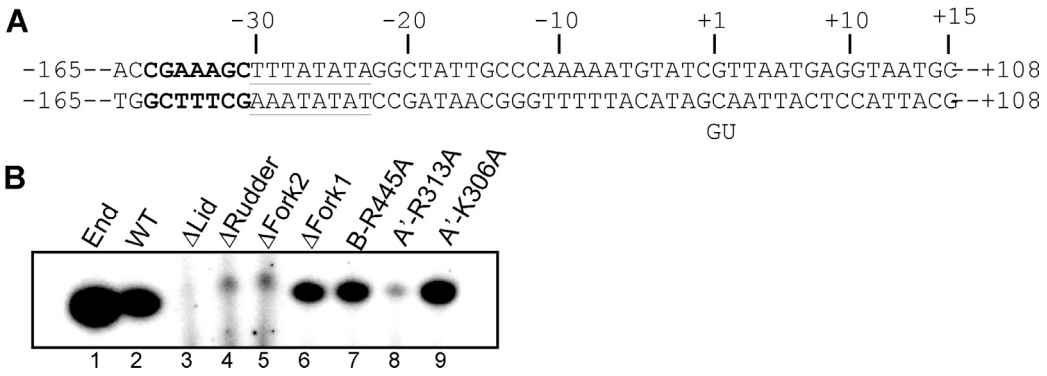

Figure 4. The loops rudder, lid and fork 2 and the conserved arginine residue 313 in switch 2 are required for synthesis of short abortive transcripts. (**A**) A template containing a modified sequence of the *gdh* promoter *gdh*-C15(-165 to +108; 21) was used to detect initial RNA synthesis. The TATA-box is underlined, the BRE element is indicated by bold letters and the transcription start site is indicated by +1. (**B**) Synthesis of a three nucleotide abortive transcript in dinucleotide primed reactions. Abortive RNA products were separated in a 28 % polyacrylamide gel.

The Δrudder, Δfork2, and R313A variants were highly defective, Δrudder and Δfork2 synthesized a product of different mobility indicating that they were also unable to perform the primed reaction. The Δlid enzyme however was totally inactive (Figure 4B, lane 3), suggesting that the lid plays an essential role in stabilizing the initially transcribing complex. To investigate whether increasing primer lengths would support abortive transcription by the Δlid enzyme, the reactions were performed in the presence of a tri-, tetra- or a pentanucleotide RNA primer. In contrast to the bacterial Δlid enzyme, which can synthesize longer transcripts in the presence of longer priming RNAs (11) the archaeal Δlid enzyme was unable to elongate tetra- and pentanucleotide RNAs (not shown).

Fork 2 is essential for unprimed de novo transcription

In a second approach to study initial RNA synthesis, we used a *gdh* template with a "pre-melted" mismatch bubble (bubble1) at positions –10 to +3, mimicking DNA in the open complex (Figure 5A). In the absence of transcription factors, the endogenous polymerase, the recombinant enzyme, the Δfork1 mutant, and the single point mutants R445A and K306A produced an 18 nt transcript as the major RNA product, compared to a full-length run-off product of 20 nt length (Figure 5B). This indicates that a pre-melted bubble can be bound by the polymerase and direct *de novo* RNA synthesis. It is not surprising that the lid is inactive in

this unprimed transcription assay, as it also fails to elongate RNA primers. In addition, the rudder, fork 2, and the arginine in switch 2 are essential for unprimed *de novo* transcription.
To investigate whether defects in *de novo* transcription may be compensated by initiation factors, we repeated the experiment in the presence of TBP and TFB. This showed that the lid, rudder, and the invariant arginine in switch 2 were not strictly required for *de novo* RNA synthesis in the context of an initiation complex. However, the Δfork 2 variant remained totally inactive, showing that this loop is essential for *de novo* transcription even within the initiation complex. The structural basis for this observation remains unknown.

DNA rewinding drives RNA displacement

A surprising result from the *de novo* transcription assays was that inclusion of TFB and TBP induced the accumulation of shorter, 11- and 12-mer RNA products, and strongly reduced the amount of 18-mer product (Figure 5B). The shorter products resulted from the presence of the TBP-TFB complex (Figure 5C), which apparently formed a barrier to further progression of the early transcribing complex. To investigate the nature of this barrier, we tested whether the RNA product was properly separated from the DNA template strand by treating the reaction products with RNase H, which specifically degrades RNA in a DNA-RNA hybrid. Indeed, the products were RNase H sensitive (Figure 5D), suggesting that RNA strand separation did not occur. To show that the failure of RNA displacement was due to the mismatch bubble design, we repeated the assay with a closed DNA. In this case, the longer RNA product was RNaseH resistant, showing that it was properly displaced from the template (Figure 5E). However, also on the closed template not displaced short RNAs were formed (Figure 5E, lane 4) indicating thatthe barrier is caused by the TBP/TFB complex independent of the presence of the mismatch bubble. The 1-nt upstream shift of the block in the closed template (Figure 5E) suggests an altered TBP/TFB interaction with the premelted bubble (unpublished data of Christine Richter and Winfried Hausner suggest that transcription initiates at the bubble template at +2 and at the closed template at +1) .

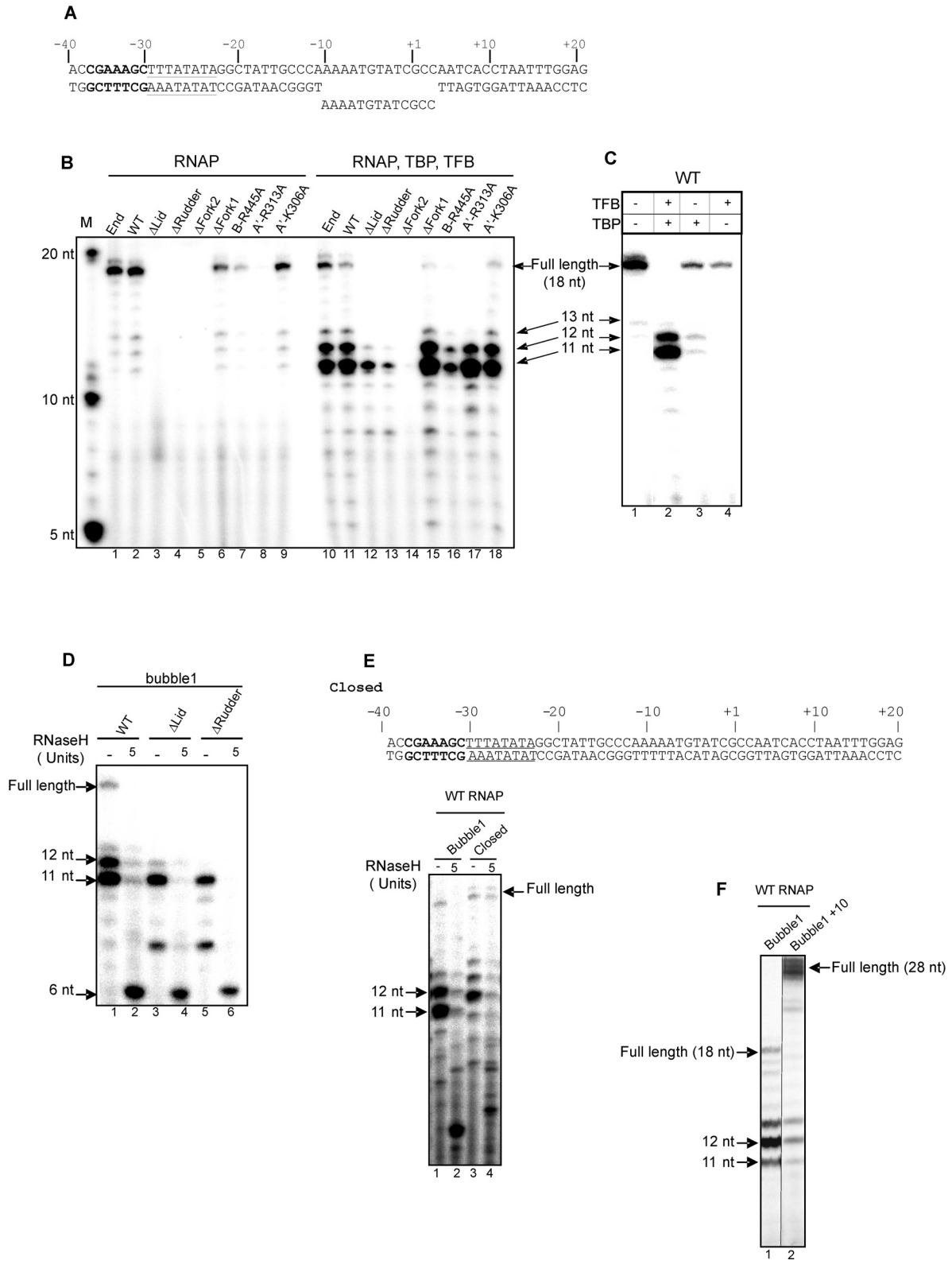

Figure 5. Initiation from a premelted mismatch bubble. (**A**) Template of the *gdh* promoter containing a mismatch in the region from -10 to +3 (bubble 1). The TATA box sequence is underlined the BRE sequence shown in bold letters. (**B**) Three loops lid, rudder and fork 2 and the conserved arginine residue 313 in switch 2 are required for the synthesis of full-length transcripts. The various RNAP fractions and mutants were preincubated in transcription reactions with the template shown in (A) for 30 min at 70 °C shown in the absence (left panel) and presence of TBP (35 nM) and TFB (30 nM) (right panel). Transcription was started by the addition of NTPs and the reactions were incubated for further 30 min at 70 °C. RNA products were analyzed on a 28 % polyacrylamide gel. The lane labeled M indicates RNA markers. (**C**) The +11 to +12 transcripts are

induced by the TBP-TFB complex. Transcripts from bubble 1 formed in the presence of the individual components indicated on top of the lanes were analyzed as in (B) Note that a weak barrier is also imposed in the presence of TBP alone (lane 3) and that TFB has an inhibitory effect on the synthesis of the full-length transcript (compare lanes 1 and 4) most likely by binding to free RNAP. (**D**) The 11 to 12 mer and full length transcripts are RNase H sensitive. Transcripts formed on bubble 1 were digested for 15 min at 37 °C with RNase H and analyzed as in (B). (**E**) RNA displacement requires upstream rewinding of the template strand. The double stranded closed template containing no mismatch (upper panel) and bubble 1 were transcribed in the presence of transcription factors and treated for 15 min at 37 °C with 5 units of RNase H as indicated. (**F**) The length of downstream duplex DNA affects promoter escape. Transcripts from bubble 1 and from a template with the same mismatch region but extended at the downstream end by 10 nt of duplex DNA were analyzed as in (B).

Our finding that the full-length transcript is RNase H sensitive on the mismatch bubble indicates that the mismatch bubble template prevents RNA displacement from the DNA template strand, and indicate that RNA displacement requires upstream rewinding of the DNA duplex, which cannot occur when the DNA strands are non-complementary, as in the mismatch bubble. Since the lid is located at the upstream end of the DNA-RNA hybrid, one may imagine that deletion of the lid would allow for continued growth of a persistent hybrid. However, the Δlid enzyme produced 11- and 12-mer and shorter RNA products but no full-length transcripts (Figure 5B), showing that the lid was not the cause of the observed elongation barrier. Taken together, these data are consistent with the idea that DNA-RNA strand separation is driven by successful competition of the DNA non-template strand with the RNA for the DNA template strand. When we used the mismatch bubble containing longer segments (e.g. 10 nt in Figure 5F**)** of downstream duplex DNA, the processivity of the WT enzyme was greatly increased and the arrest at +11/12 less pronounced (Figure 5F). This finding indicates that a certain length of DNA downstream duplex is required for efficient displacement of TFB/TBP, at least in our system.

Invariant arginines in fork 2 and switch 2 cooperate in DNA unwinding

After completion of the initiation-elongation transition, the polymerase alone can elongate the RNA chain. To investigate elongation by the mutant enzymes, we used nucleic acid scaffolds with a 83 nt template strand (EC-T) hybridized to 9 nt of a 14 nt RNA (Figure 6A), and assembly protocols first described by Kashlev and coworkers (22,23). In a first set of experiments, we analyzed elongation of the RNA hybridized to a template single strand (Figure 6A, template 1*a*). To investigate the effect of downstream duplex DNA a scaffold was used that was similar to synthetic templates used for analysis of elongation by bacterial RNAP (24). This duplex template contained a 12 nt heteroduplex region basepaired in part with the 13 nt RNA primer (EC3, see Figure 6A.). A lower incubation temperature of 60 °C was used to prevent DNA melting. The reconstituted wild-type polymerase was able to accumulate

larger amounts of incomplete transcripts but required a downstream DNA duplex for synthesis of a full-length transcript (Figure 6B, lanes 1 and 8). This finding indicates that downstream duplex DNA is required for processive RNA synthesis by the WT enzyme. The cleft loop mutants generally showed defects in these elongation assays, with the exception of fork1 (Figure 6B, lanes 5 and 12) and the mutant of residue K306 in switch 2 (data not shown). The Δlid, Δrudder and Δfork2 enzyme were unable to synthesize longer products on the single stranded template *a* (Figure 6B, lanes 2-4), B-R445A synthesized a weak ~40 nt product and A´-R313A a weak 20 nt product. These findings suggest that interaction of the cleft loops lid, rudder, fork2 and of R331A with nontemplate DNA are required for the formation of stable elongation complexes. In the presence of downstream duplex DNA the mutants rudder and lid retained the ability to synthesize RNA products up to a length of ~40 nt (Figure 6B, lanes 9 and 10). Previous structural studies had suggested that downstream DNA strand separation may involve binding and distortion of the template strand by switch 2 (8) and interference of fork 2 with the path of the non-template strand (6,8).

As predicted from this model, essentially no elongation activity was obtained for the Δfork2 enzyme (Figure 6B, lane 11). Strikingly, even a point mutation in fork 2 that mutates the invariant arginine R445 to an alanine almost abolished elongation activity on a template containing downstream duplex DNA (Figure 6B, lanes 13), consistent with a role of this arginine

in directing the nontemplate strand away from the active site. Also consistent with the proposed mechanism for downstream DNA separation, the switch 2 arginine mutation is totally inactive in elongation assays on template EC3 that contains a downstream DNA duplex **(**Figure 6B, lanes 14). Taken together, these results suggest that two invariant arginines, located in switch 2 and fork 2 on opposite sides of the cleft and incoming DNA play an important role in downstream DNA strand separation as suggested by structural studies. The role of the other adjacent amino acids still has to be investigated. The ability of the archaeal RNAP to synthesize >40 nt transcripts on a single stranded template (1a) was surprising since the bacterial enzyme forms only ~20 nt transcripts on similar templates (11,12). To investigate whether the archaeal RNAP forms extended DNA-RNA hybrids on a single stranded template the products of reactions with template 1a were digested with RNase H (Fig. 6c). To preclude posttranscriptional hybridization of RNA to the template during cooling down to 37 °C, the transcription reaction carried out at 60 °C were incubated at 70 °C with a thermostable RNase H. This revealed that transcripts from template 1a were generally RNase H sensitive, indicating the formation of extended RNA-DNA hybrids. The shorter

RNA products synthesized by *E. coli* polymerase on similar single stranded templates are also RNase H sensitive (9,11).

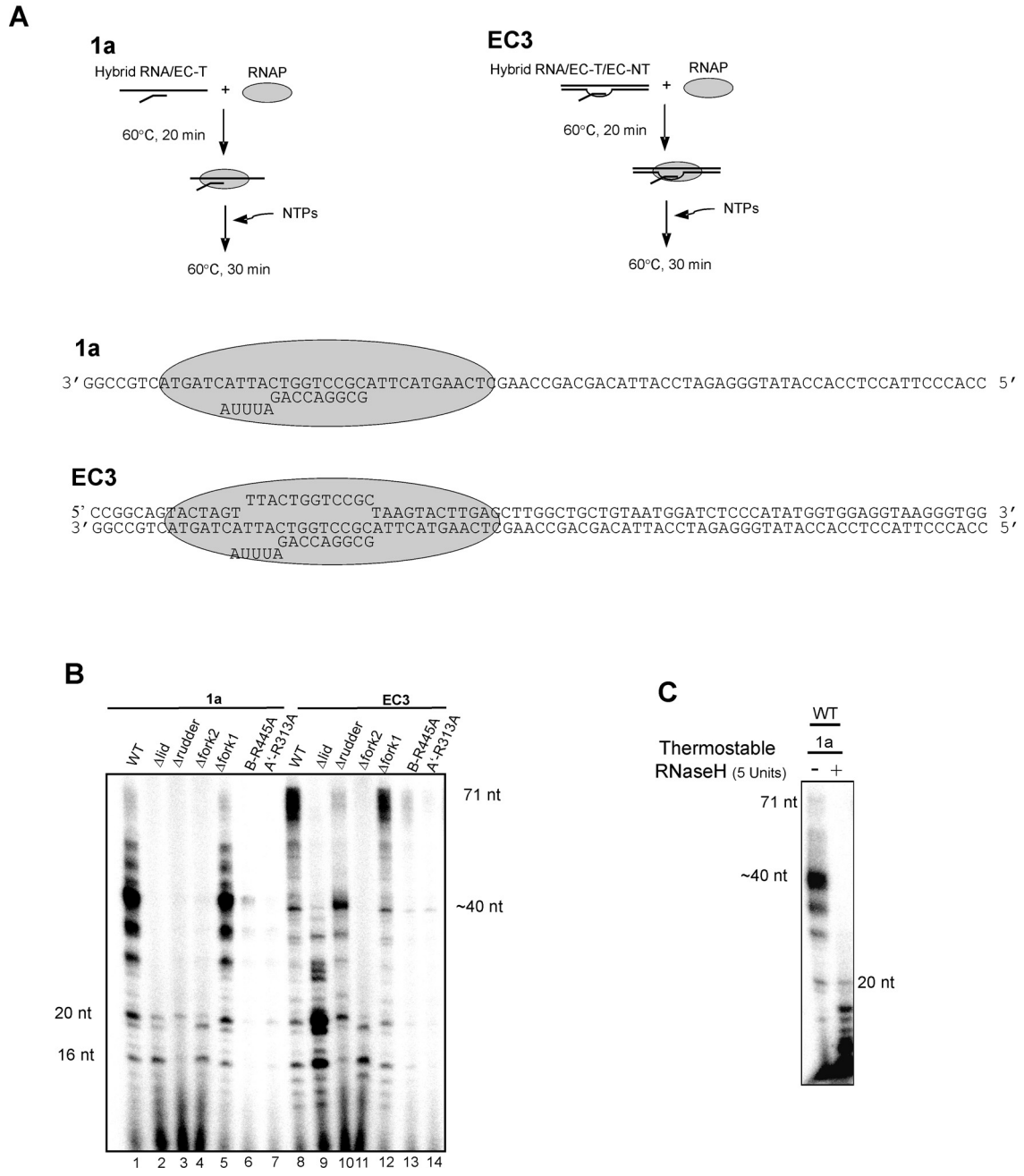

Figure 6. Lid, rudder, fork 2 and the conserved arginine residue 313 in switch 2 are required for normal elongation. (**A**) Experimental design and templates, template 1a contains the template strand hybridized with 9 nt of a 14 nt RNA primer, template ECR3 contains a 12 bp heteroduplex hybridized with the RNA primer. (**B**) Analysis of the effect of mutations on elongation. Mutants and wt enzymes were incubated with the templates indicated in the absence of TBP and TFB at 60 °C and RNA products synthesized were analyzed on 20-28 percent polyacrylamide gradient gels.. (**C**) The archaeal enzyme forms extended DNA–RNA hybrids on the single stranded template. Transcription reactions were conducted at 60 °C and digested with thermostable RNase H at 70 °C as indicated in the Figure.

DISCUSSION

Structural studies of RNAP II resulted in proposals for the functional roles of various polymerase elements that must be tested by mutational analysis. While such studies are now conveniently carried out in the bacterial system, only very few and non-lethal RNAP II mutations were so far introduced, because a reconstituted eukaryotic RNAP is not available (12). Thus a detailed structure-function analysis of the RNAP II system is lacking. Here we used variants of a recombinant RNAP II-like archaeal RNA polymerase to investigate the functional role of the cleft loops rudder, lid, fork 1, fork 2, and the switch 2 element during various stages of transcription. We tested the polymerase variants in several assays, including promoter-dependent transcription (Figure 2), open complex formation (Figure 3), elongation of short RNA primers (Figure 4), *de novo* transcription from a pre-opened bubble (Figure 5), and transcription elongation with various nucleic acid scaffolds (Figure 6). The results of this study are summarized in Table 1. All mutants were impaired in these assays to various degrees, with the exception of fork 1, which is the only loop not present in the bacterial enzyme.

Table 1. Comparison of the activities of the mutant enzymes relative to the reconstituted wt enzyme in various assays.

	Run off transcription (Figure 2)	Open complex (Figure 3)	Abortive transcription (Figure 4)	Full length transcript from open bubble (Figure 5b)	Elongation on double stranded scaffold (EC3) (Figure 6b) ≤ 40 nt	>40 nt
Δ lid	-	+	-	-	+	-
Δ rudder	-	-	-	-	+	(+)
Δ fork2	-	+	-	-	+	-
Δ fork1	+	+	+	+	+	+
B -R445A	+	+	+	(+)	(+)	(+)
A´-R313A	-	+	(+)	-	(+)	-
A´-K306A	+	+	+	+	+ (data not shown)	+ (data not shown)

(+) indicates strong impairment of activity, + good and – no activity .

Our analysis revealed that the rudder is important for stabilizing melted DNA in the open complex (Figure 3), consistent with formation of an inactive open complex by a bacterial rudder mutant (9). Initial transcription required the lid, rudder, fork 2, and switch 2 residue R313. The corresponding mutants were defective in extension of short RNA primers (Figure 4), and fork2 was incapable to synthesize longer RNAs from a pre-opened bubble (Figure 5). Our results are generally consistent with previous analysis of the bacterial rudder and lid, but

some differences were also observed. During RNA elongation on the single-stranded template the archaeal enzyme was able to synthesize transcripts > 40 nt (Figure 6C, lane 1) whereas the bacterial enzyme synthesized mainly 22-30 nt transcripts (11,10). The rudder was important for the maintenance of an actively elongating archaeal complex, whereas it seems less important in the bacterial system. The archaeal Δrudder enzyme was unable to reach the end of the template in preformed elongation complexes (Figure 6C), although the bacterial Δrudder enzyme is more processive and can produce full-length transcripts (9).

This work also provided insights into the crucial functions of fork loop 2. Fork 2 is strictly required for primed transcription and elongation, and a single point mutation of its conserved arginine residue (R445A) showed a severe defect in elongation effectivity (Figure 6C, lane 13). This arginine is highly conserved in fork 2 (ses Figure 1a) and is located directly at the junction of the two DNA strands close to the point of downstream DNA separation (8, 25). Several replacements in the eukaryotic fork 2, and mutations of sites interacting with fork 2 influence the elongation rate (reviewed by 12). Our data are consistent with an essential role of fork 2 in downstream DNA unwinding during elongation, most likely by interference with the path of the nontemplate strand as suggested by structural studies. However, the function of fork 2 is not restricted to downstream DNA unwinding. *De novo* RNA synthesis from a pre-melted bubble was totally impaired in the Δfork 2 variant (Figure 5B), suggesting that this loop is involved in stabilizing the NTPs during initial phosphodiester bond formation.

Unexpected insights into the transition from transcription initiation to elongation were obtained by the transcription assays with a pre-melted bubble (Figure 5). In the absence of initiation factors, run-off transcripts were synthesized in unprimed, *de novo* transcription. However, in the presence of initiation factors, RNA transcripts were generally limited to a length of 11-12 nucleotides which were not displaced independent of the presence of a preformed bubble (Figure 5E). RNase H probing revealed that persistent DNA-RNA hybrids were formed under these conditions, and that the mismatch bubble prevented RNA displacement of the full length transcript from the DNA template strand. Consistent with our findings, upstream re-closure of the transcription bubble begins when RNA has reached a length of 10-11 nt (21). The shorter RNA transcripts however were not induced by failure to extend the DNA-RNA hybrid past the lid, which normally stacks on the upstream end of the hybrid. Instead, they were due to the presence of the TBP-TFB complex. Synthesis of 11-12 nucleotide RNAs in our assay must involve displacement of the TFB finger domain located in the hybrid site (26) but does apparently not entirely release TFB. TFB may remain bound to the dock with its N-terminal ribbon domain (27, 26) but this alone cannot explain why RNA

synthesis stops prematurely, since the ribbon domain would not interfere with the DNA-RNA hybrid emerging from the cleft between the polymerase protrusion, wall, and clamp domains. Instead RNA synthesis may stop since the growing hybrid encounters the complex of TBP and the TFB core domain, situated above the cleft (27) when a length of 12-13 base pairs is reached. The existence of the 12-14 nt barrier on the duplex template (Figure 5E) that allows proper rewinding of upstream DNA indicates that the mismatch bubble is not the major cause of the first barrier. Furthermore, the ratio full-length to 11,12 nt transcripts is greatly increased when the length of duplex DNA downstream of the mismatch bubble was extended by 10 nt (Figure 5F). Taken together these results suggest that a minimal length of 27 nt of downstream duplex DNA are required for effective elongation past the barrier. Bubble reclosure at the upstream end of the open complex has been proposed as a key event in the promoter clearance transition and suggested to cause TFB-displacement (28) in the human RNAP II system, but is apparently not required for promoter escape and processive RNA synthesis in our system. The results with the preformed bubble described here are not specific for the archaeal enzyme. RNAP II from yeast shows similar properties on the pre-melted template analyzed in this study (Reich C., Naji S., Gerber J., Küsser A., Tschochner H, Cramer P, Thomm M, manuscript in preparation) and both RNAP II (29) and the archaeal enzyme can transcribe ~100 nt run-off transcripts with high effectivity when the premelted template contains longer segments of downstream duplex DNA (Figure 5F and Spitalny, P,. Naji, S., Thomm, M. unpublished data).

Finally, we have uncovered the essential role of switch 2, and in particular an invariant arginine in switch 2, A´-R313, at various points during transcription. This arginine was required for chain elongation from the minimal nucleic acid scaffolds (Figure 6, for transcription from a pre-melted bubble (Figure 5) and for primed transcription (Figure 4). Thus the interaction of this arginine with the DNA template strand backbone at position +2 is essential for correct template positioning in the active site. Further, these results provide insights into the structure of the open complex, since open complex formation apparently involves binding of the template in the active site at the location normally adopted during elongation, and thus leads to a complex prone to RNA chain initiation. Since mutation of R313 completely disabled RNA synthesis on templates containing downstream duplex DNA (Figure 6c), binding of the template strand to switch 2 is essential for DNA separation during elongation, most likely due to distortion of the incoming B-DNA duplex as suggested (8). Given the invariant nature of R313 in switch 2 in all three kingdoms of life, the same mechanisms and structural transitions will likely occur in all cellular RNA polymerases. The

mutation of R313 directly reports on the function of this residue, and not on the function of switch 2 in general, since mutation of another switch 2 residue, K306, which is not conserved and more distant from the template and active site, has essentially no effects in most assays.

More generally, our analysis showed that the recombinant archaeal RNA polymerase can be used for a structure-function analysis of aspects of a eukaryote-like transcription mechanism, including initiation and the initiation-elongation transition. Many more mutant polymerases and different nucleic acid scaffolds however must be analyzed before a satisfactory mechanistic understanding of the dynamic transcription cycle will emerge.

SUPPLEMENTARY DATA

Supplementary Data are avaliable at NAR online.

ACKNOWLEDGMENTS

This work was supported by grants of the Deutsche Forschungsgemeinschaft to Michael Thomm and Patrick Cramer and the Leibniz award of the DFG to Patrick Cramer. We thank Bernd Goede for bioinformatic work.

Conflict of interest statement. None declared.

REFERENCES

1. Cramer, P. *et al.* (2000) Architecture of RNA polymerase II and implications for the transcription mechanism. *Science* **288**, 640-649.
2. Cramer, P., Bushnell, D.A. & Kornberg, R.D. (2001) Structural basis of transcription: RNA polymerase II at 2.8 Ångstrom resolution. *Science* **292**, 1863-1876.
3. Zhang, G. *et al.* (1999) Crystal structure of *Thermus aquaticus* core RNA polymerase at 3.3 Å resolution. *Cell* **98,** 811-824.
4. Vassylyev, D.G. *et al.* (2002) Crystal structure of a bacterial RNA polymerase holoenzyme at 2.6 Å resolution. *Nature* **417**, 712-719.
5. 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***,* 7131-7134.
6. 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 Å resolution. *Science* **292**, 1876-1882.

7. 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**, 1014-1016.

8. 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.

9. 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**, 1369-1378.

10. Toulokhonov, I. & Landick, R. (2006) The role of the lid element in transcription by *E. coli* RNA polymerase. *J. Mol. Biol.* **361**, 644-658.

11. 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**, 634-643.

12. Trinh, V., Langelier, M.-F., Archambault, J. & Coulombe, B. (2006) Structural perspective on mutations affecting the function of multisubunit RNA polymerases. *Microbiol. Mol. Biol. Rev.* **70**, 12-36.

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

14. Naji, S., Grünberg, S. & Thomm, M. (2007) The RPB/ 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.

15. Pühler, G. *et al.* (1989) Archaebacterial DNA-dependent RNA polymerases testify to the evolution of the eukaryotic nuclear genome. *Proc. Natl. Acad. Sci. USA* **86**, 4569-4573.

16. 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**, 30581-30591.

17. Kusser, A.G., Bertero, M.G., Naji, S., Becker, T., Thomm, M., Beckmann, R. & Cramer, P. (2007) Structure of an archaeal RNA polymerase. *J.Mol.Biol.* doi:10.1016/j.jmb.2007.08.066

18. Bell, S.D. & Jackson, S.P. (2001) Mechanism and regulation of transcription in archaea. *Curr. Opin. Microbiol.* **4***,* 208-213.

19. Thomm, M. & Hausner, W. (2006) Transcriptional Mechanisms. in *Archaea: Evolution, Physiology and Molecular Biology* (eds. Garrett, R. & Klenk, H.-P.) 185-198 (Blackwell Publishing, Oxford, U.K.).

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

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

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

23. Kireeva, M.L., Komissarova, N., Waugh, D.S. & Kashlev, M. (2000) The 8-nucleotide-long RNA:DNA hybrid is a primary stability determinat of the RNA polymerase II elongation complex. *J. Biol. Chem.* **275**, 6530-6536.

24. Daube, S.S. & von Hippel, P.H. (1992) Functional transcription elongation complexes from synthetic RNA-DNA bubble duplexes. *Science* **258,** 1320-1324

25. 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,** 157-162

26. 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 Ångstroms. *Science* **303**, 983-988.

27. Chen, H.-T. & Hahn, S. (2003) Binding of TFIIB to RNA polymerase II: mapping the binding site fort he TFIIB zinc ribbon domain within the preinitiation complex. *Mol. Cell* **12**, 437-447.

28. Pal, M., Ponticelli, A.S. & Luse, D.S. (2005) The role of the transcription bubble and TFIIB in promoter clearance by RNA polymerase II. *Mol. Cell* **19**, 101-110.

29. Holstege, F.C.D., Tantin, D., Carey, M., van der Vliet, P.C. & Timmers, H.Th.M. (1995) The requirement for the basal transcription factor IIE is determined by the helical stability of promoter DNA. *EMBO J***, 14**, 810-819.

MOLECULAR MICROBIOLOGY doi: 10.1111/j.1365-2958.2007.06084.x published online: 20-12-2007

IV A polymerase III-like reinitiation mechanism is operating in regulation of histone expression in Archaea

Patrizia Spitalny and Michael Thomm

Department of Microbiology, University of Regensburg, Universitätsstraße 31, 93053 Regensburg, Germany

For correspondence. E-mail michael.thomm@biologie.uni-regensburg.de; Tel. (+49) 941 943 3160; Fax (+49) 941 943 2403.

Running title: Pol III-like termination in archaea

Key words: transcription, termination, RNA polymerase, Archaea, RNA polymerase III

Summary

An archaeal histone gene from the hyperthermophile *Pyrococcus furiosus* containing four consecutive putative oligo-dT terminator sequences was used as a model system to investigate termination signals and the mechanism of termination *in vitro*. The archaeal RNA polymerase terminated with high efficiency at the first terminator at 90 °C when it contained 5 to 6 T-residues, at 80 °C read through was significantly increased. A putative hairpin structure upstream of the first terminator had no effect on termination efficiency. Template competition experiments starting with RNA polymerase molecules engaged in ternary complexes revealed recyling of RNA polymerase from the terminator to the promoter of the same template. This facilitated reinitiation was dependent upon the presence of a terminator sequence suggesting that pausing at the terminator is required for recycling like in the RNA polymerase III system. Replacement of the sequences immediately downstream of the oligo-dT terminator by an AT-rich segment improved termination efficiency. Both AT-rich and GC-rich downstream sequences seemed to impair the facilitated reinitiation pathway. Our data suggest that recycling is dependent on a subtle interplay of pausing of RNA polymerase at the terminator and RNA polymerase translocation beyond the oligo-dT termination signal that is dramatically affected by downstream sequences.

Introduction

The mechanism of initiation of transcription in archaea is RNA polymerase II (polII) like (Bell and Jackson, 2001; Geiduschek and Ouhammouch, 2005; Thomm, 2007) and is dependent upon the general transcription factors TATA-binding protein (TBP) and TFB, both related in structure and function to eukaryotic TBP and TFIIB. Recent evidence suggests that a third factor, TFE that is homologous to the N-terminal part of the α subunit of TFIIE plays in addition a pivotal role in initiation (Bell *et al*., 2001; Hanzelka *et al*., 2001; Werner and Weinzierl, 2005, Naji *et al.,* 2007) and also in elongation by stabilizing the open complex and transcription bubble (Grünberg *et al.*, 2007). Archaeal transcriptional terminators were early described to contain oligo-T stretches (Reiter *et al.,* 1988; Brown *et al.,* 1989; Thomm *et al.,* 1994) that are also recognized by RNA polymerase III as terminator signals (Gunnery *et al.,* 1999; Geiduschek and Kassavetis, 1992; Braglia *et al.,* 2005) and a detailed study using a single-round *in vitro* system from a thermophilic archaeon has shown that these oligo-dT sequences without preceding RNA hairpin structures are sufficient to direct termination by an archaeal RNA polymerase (RNAP) *in vitro* (Santangelo and Reeve, 2006). Unlike in pol I and pol II transcripts the 3´ends of pol III and of archaeal transcripts are generated by transcriptional termination. Thus, both the sequences of terminators and the general mechanism of archaeal transcription termination seem to be pol III-like. However, the finding that rho-independent bacterial terminators and the bacterial rho-factor can mediate termination of transcription in an archaeal system (Santangelo and Reeve, 2006) show also some superficial similarities of archaeal termination to termination in the bacterial system.

With one exception (Thomm *et al.,* 1994) the complete transcription cycle involving initiation elongation and termination on an intact and complete archaeal gene has not yet been studied. Such an investigation is likely to lead to important insights because the mechanism of transcription reinitiation involving direct recycling of RNAP from terminator to promoter represents an important aspect of gene regulation in particular in the pol III but also in other transcription systems (Dieci and Sentenac, 1996; 2003). Furthermore, the sequences immediately downstream of terminator sequences have been shown to be in involved in pausing (Lee *et al.,* 1990; Palangat *et al.,* 2004) which is a precursor of transcriptional arrest and termination and the effects of downstream sequences on termination and recycling of RNAP have not yet been studied in the archaeal system.

Using a complete gene encoding an archaeal histone as template we demonstrate in a hyperthermophilic system that reinitiation has been established as regulatory mechanisms in the archaeal transcriptional machinery which is thought to be the evolutionary precursor of the

eukaryotic system. In addition, we show that GC-rich sequences downstream of the terminator inhibit recycling of RNAP from the terminator to promoter and are therefore likely to reduce the levels of gene expression at high temperatures.

Results

Transcriptional termination at 90°C

The histone gene *hpyA1* from *Pyrococcus furiosus* was chosen for termination experiments because it shows four consecutive oligo-dT sequences (T1 to T4) directly following the ORF (Fig. 1A) suspected to act as termination sites. The little information available on archaeal termination suggests mechanisms that are more closely related to pol III than to bacterial RNAP. But still archaeal termination is poorly little understood, especially in hyperthermophilic archaea. To investigate the termination events in hyperthermophilic archaea the histone encoding gene *hpyA1* was amplified from *Pyrococcus furiosus* genomic DNA and cloned into a plasmid vector. It shows a high transcription rate in *in vitro* transcription assays (Fig. 1B). It is a short protein enconding gene which can easily be transcribed completely in multiple round *in vitro* assays from the start point to the termination sites. Incubation of this template in linearized state at 80°C led to readthrough events at every termination site resulting in a predominant run off transcript (358b). In contrast, when incubated at 90°C, a temperature more close to the growth optimum of *Pyrococcus furiosus,* the first terminator was recognized very efficiently (Fig. 1B) leading to a predominant transcript of 250 nt (Fig. 1B. lanes 4-6). When truncated linearized templates (not containing a terminator) were used in *in vitro* transcription experiments template activity at 90°C was significantly lower (for *hpyA1* see Fig. 5B, right panel; for a comparison of the template activity of *gdh* promoter at 80 and 90 °C see Fig. 5B, left panel). Under our *in vitro* conditions purified RNA subjected to an incubation temperature of 90°C is degraded with a half –life time of ~20 min (Hethke, 1997; Hethke *et al.,* 1999). Therefore a specific terminator-dependent mechanism must exist allowing the high expression rate of the *hpyA1* gene *in vitro.*

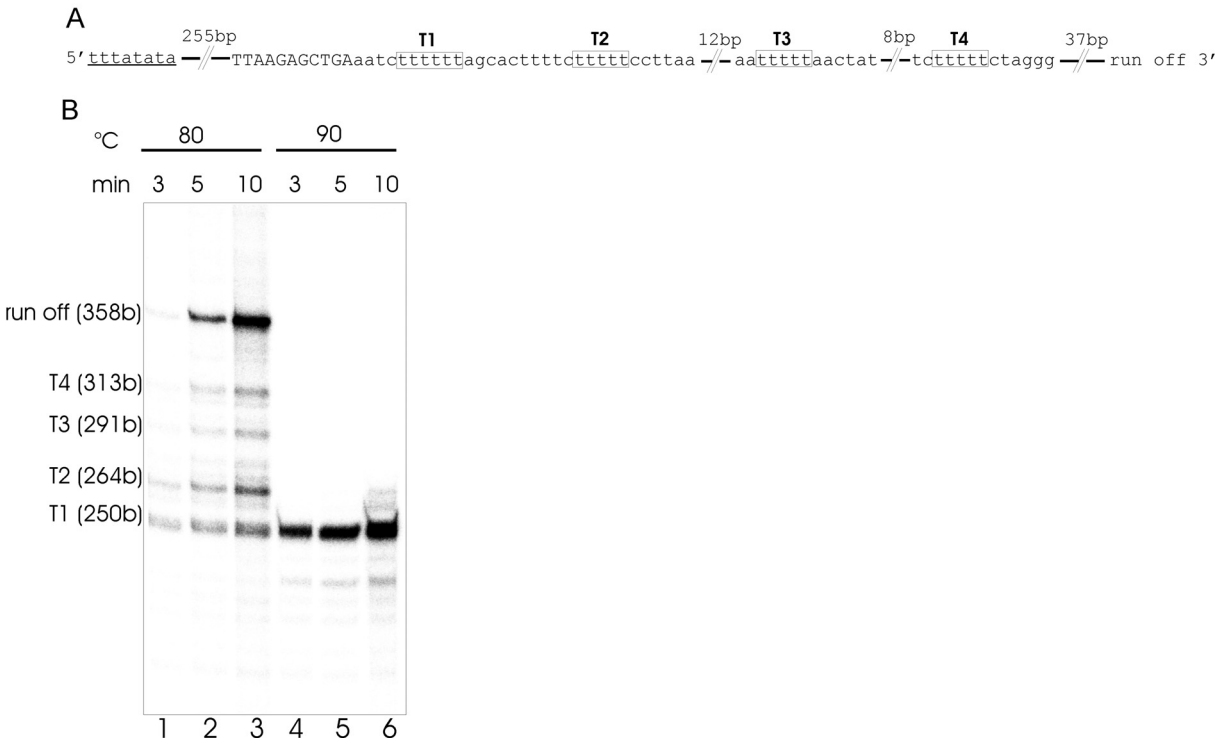

Fig. 1. Termination efficiency is increased at 90°C. In **A** the sequence of the RNA-like strand of *hpyA1* cloned into pUC19 is depicted from the TATA box (underlined) to the downstream end created by *PstI* digestion. The part of the sequence belonging to the ORF is written in capital letters. The main termination site (T1) and the tree backup terminators (T2-T4) are boxed. *In vitro* transcription of the template displayed in **A** at 80°C and 90°C, respectively, is shown in **B**. The lengths of the transcripts are indicated on the left. Transcription assays were performed with 46 nM RNAP, 238 nM TBP and 147 nM TFB.

The minimal termination signal is T_5

As shown in Figure 1 efficient archaeal termination at least at the *hpyA1* terminator is mediated by the archaeal RNAP alone without the need for additional factors. A simple run of T residues at the end of the *hpyA1* gene serves as termination signal for the archaeal RNA polymerase. In bacteria usually seven to nine T residues following a dyad symmetry element facilitate factor-independent termination. The minimal termination signal recognized by eukaryal polymerase III varies among different species (Bogenhagen and Brown, 1981; Cozarelli *et al.,* 1983; Allison and Hall, 1985; Hamada *et al.,* 2000). The T tract leads to extensive pausing of pol III (Matsuzaki *et al.,* 1994; Yin *et al.,* 1999) and termination efficiency tends to increase with the length of the T cluster (Allison and Hall, 1985).

To investigate the minimal termination signal in archaeal intrinsic termination leading to a pause state of elongation sufficiently long for termination to take place we introduced point mutations into the terminator region to stepwise alter the number of consecutive T residues (Figure 2A). The rate of termination efficiency at 90 °C was determined by the amount of readthrough transcripts terminated at the first backup terminator (T2). With 5 or more T

residues more than 80% of the transcription products were terminated at T1 (Figure 2B lanes 1 and 2). The mutations that leave 4 or less T residues (Figure 2B lanes 3 to 5) resulted in strongly decreased termination efficiency. The efficiency of duration of the pause at the T tract is partly determined by the avaliability of UTP in the reaction. By increasing the UTP concentration a slight decrease of termination efficiency at all terminators can be observed but still T1 acts as main termination site (data not shown).

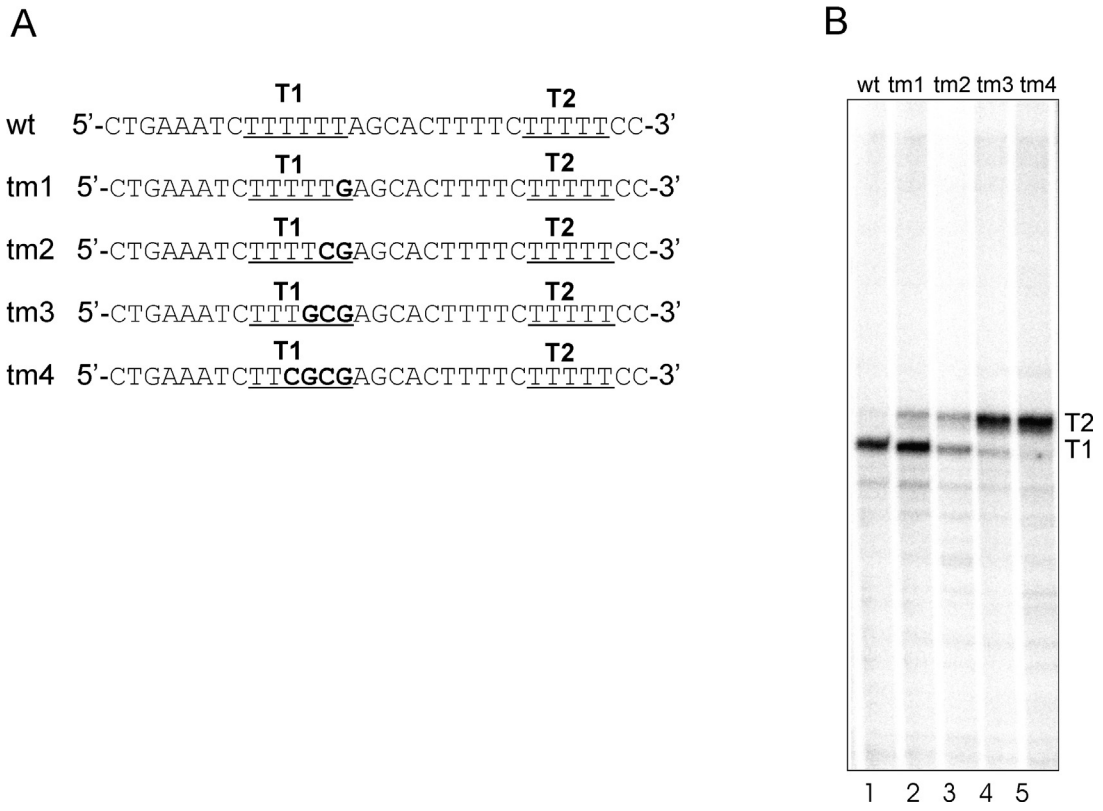

Fig. 2. The minimal termination signal is T_5. **A** Point mutations (bold) altering the length of the T-tract were introduced into the 3' flanking region of termination site T1. The effect of the different 3'mutations is shown in **B**. *In vitro* transcription experiments were performed at 90°C with the wildtype (wt) and mutant (tm1-tm4) templates. T1 and T2 indicate RNA products terminating at the main terminator and the first backup terminator, respectively. Transcription assays were performed with 46 nM RNAP, 238 nM TBP and 147 nM TFB.

Possible formation of a hairpin structure has no effect on termination

For bacterial termination it is well known that a GC-rich dyad symmetry element capable of stem-loop formation preceding the oligo-dT tract plays an essential role in transcriptional termination. For several pol III transcribed genes a palindromic sequence immediately upstream of the termination site has been described. Bogenhagen and Brown (1981) could show that these sequence elements have no influence on termination efficiency of the 5S RNA gene. In contrast in another study it has been shown that dyad symmetry elements

preceding the terminator stimulate pol III (Chu *et al.,* 1997). In a thermophilic *Methanothermobacter*-derived archaeal transcription system the presence of a sequence being capable of stem-loop formation was shown to contribute to termination efficiency (Santangelo and Reeve, 2006).

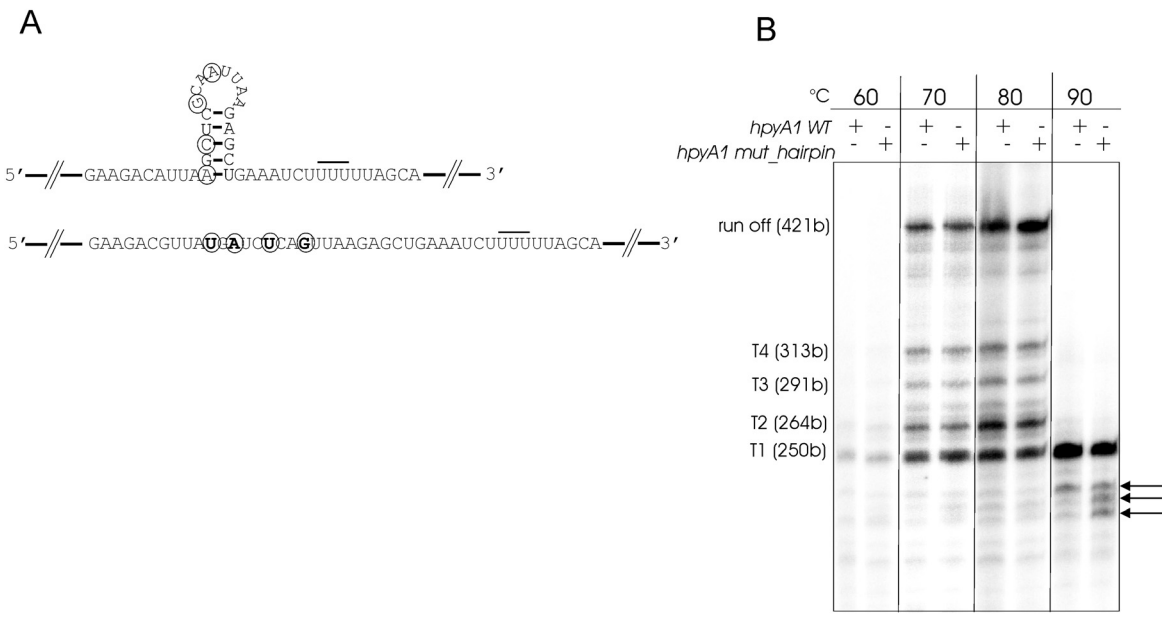

Fig. 3. A putative hairpin structure has no effect on termination. In **A** the 3' end of the RNA product from T1 is depicted. The palindromic sequence preceding the terminator is capable of hairpin formation. Point mutations introduced to eliminate the dyad symmetry are shown in bold and encircled in the sequence (lower panel), the mutated positions are encircled in the secondary structure. The termination site T1 is indicated by a bar on top of the consecutive U-residues. *In vitro* transcription products of the linearized wildtype (*hpyA1* WT) and mutant (*hpyA1 mut_hairpin*) templates are shown in **B**. The lengths of the RNA products are indicated on the left. Paused transcripts are indicated by arrows. *In vitro* transcription was performed at incubation temperatures as indicated above and contained 46 nM RNAP, 238 nM TBP and 147 nM TFB.

The *hpyA1* gene shows a palindromic sequence within the ORF located six nucleotides upstream of the T-stretch representing T1. It is unlikely that this stem loop consisting of a stem of 5 bp is stable in purified RNA at temperatures between 60 and 90 °C but a RNA secondary structure might be formed in transcribing ternary complexes. E.g. the phage λ tR2 terminator has been show to work in a *Thermus aquaticus* system at 65 °C (Naryshkina et al., 2006). To investigate the potential contribution of the palindromic sequence in the *hypA*1 gene to termination we introduced single point mutations that eliminated the dyad symmetry (Figure 3A; template *mut_hairpin*). At 90 °C, the major 250 nt RNA product from the *mut_hairpin* template was somewhat weaker than the transcript from the wt template (Figure 3B lanes 7 and 8). This result would be consistent with the findings of Chu *et al,* 1997, but may as well be an effect caused by the pausing of RNAP at the *mut_hairpin* template

upstream of the first termination site (Figure 3B, lane 8 see arrows). The high temperature conditions under which the transcription assays were performed in this study (80-90 °C) are likely to destabilize RNA secondary structures. Therefore, the temperature dependence of transcription from the wild type and mutated template *mut_hairpin* was analyzed also at temperatures ranging from 60 to 80 °C (60°C is the lowest incubation temperature allowing RNA synthesis in the *P. furiosus* system; Naji *et al.,* 2007). At 60°C, an evaluation of termination efficiency is difficult due to the low amounts of transcript obtained after an incubation time of 10 min at this temperature. No effect of the *mut_hairpin* mutation could be observed on the transcription efficiency at 70°C and 80 °C. Therefore, we conclude that the palindromic sequence preceding the T-tract seems to be no general requirement for archaeal transcription termination at the *Pyrococcus* histone gene *A*.

Competition experiments reveal that the RNAP is committed to reinitiate at the same template molecule

The limited information on archaeal termination suggests a pol III-like mechanism. It has been described (Wolffe *et al.,* 1986; Kassavetis *et al.,* 1989) that pol III transcribed genes are efficiently subjected to several rounds of transcription and the pol III transcription cycle is characterized by an increased reinitiation of pol III on the same template presumably without release of RNAP (Dieci and Sentenac, 2003). To investigate whether the archaeal RNAP is recycled from the terminator to the promoter competition experiments between two templates were performed. Both templates contained the *hpyA1* promoter and a C-minus cassette ranging from position +1 to +25 (Fig. 4A). This construct allows stalling of RNAP at +25 by conducting transcription assays in the absence of CTP. Promoter bound transcription factors can be removed by washing stalled complexes with N-lauroylsarcosine (NLS) and by purification of the immobilized ternary complexes which, in contrast to promoter bound initiation factors, are insensitive to NLS (Spitalny and Thomm, 2003). The second template used as competitor in the experiments described in Fig. 4C is a truncated version that did not contain a terminator (template 2 in Fig. 4A) and produced a run off transcript of 237 nt (13 nt shorter than the RNA terminated at T1). For analysis of single round transcription purified stalled complexes (step 2 of Fig. 4A) were chased by a full set of NTPs but in the absence of additional TBP/TFB. Having the same promoter sequences both templates were transcribed with similar efficiency at 80°C in single round transcription reactions (Fig. 4 B, SR). To analyze multiple rounds of transcription stalled RNAPs were washed with transcription buffer not containing NLS leaving TBP and TFB bound to the promoter (Spitalny and Thomm,

2003). Since no free RNAP is present in these assays newly initiated transcripts can be only formed under these conditions by RNAP molecules released from the terminator or end of the template or recycled from the terminator or the end of the DNA-fragment to the promoter. Analysis of transcripts synthesized in multiple round assays revealed that the ratio multiple round to single round transcripts was 3.1 for template 1 and 2.6 for template 2 (Fig. 4B).

We next analyzed transcript release from both templates by stalling RNAP at position +25, followed by isolation of ternary complexes (steps 1 and 2 in Fig. 4A). Next, 4NTPs but no competitor DNA was added to the stalled complexes and after incubation times ranging from 5 to 15 min ternary complexes were again purified and the elongation products in ternary immobilized complexes (p) and released RNA in the supernatant (s) were analyzed (Fig. 4C). As expected, transcript release was delayed at the terminator containing template (compare e.g. lanes 1 and 2 and 3 and 4 in the left and right panel of Fig. 4C). The tables (1) below the gel panels in Fig. 4C show the total amount of transcript in the pellet obtained by magnet particle separation following step 3 (Fig. 4A) containing the ternary complexes versus the amount of released transcripts found in the supernatant. While with template 2 after 10 minutes more than 80% of the run off transcript were found in the supernatant (Figure 4C, right panel, lanes 3 and 4) only about half of the transcribed RNA of template 1 was released after 10 minutes (Figure 4C left panel, lanes 3 and 4). But after 15 minutes incubation ~ 80 percent of RNAs were released from both templates. Table (2) below the gel panels in Fig. 4C documents the percentage of transcription products from T1, T2, T3 and T4 and of the run off product for each lane. The numbers are corrected for the varying amount of incorporated radioactivity for each transcription product which depends upon the occurrence of UMP in the transcript. As expected, the T1 signal decreased while the run off signal increased with time indicating a slow read through the termination signals at 80°C. Thus, the oligo–dT tracts lead either to termination or to a slow elongation through the terminator region.

The competition experiments were first performed at 80 °C as shown in Figure 4A. Initially transcribing polymerases were isolated while stalled at position +25. The promoter bound transcription factors were removed by washing with 0.5% NLS (see Spitalny and Thomm, 2003). For the following elongation by RNAPs purified in ternary complexes multiple rounds of transcription were allowed by adding a full set of NTPs and new transcription factors simultaneously with the alternative template 2. Thus, the starting conditions were the same for both competing templates. Figure 4D left panel shows the results of the competition experiments performed with template 2 as competitor DNA.

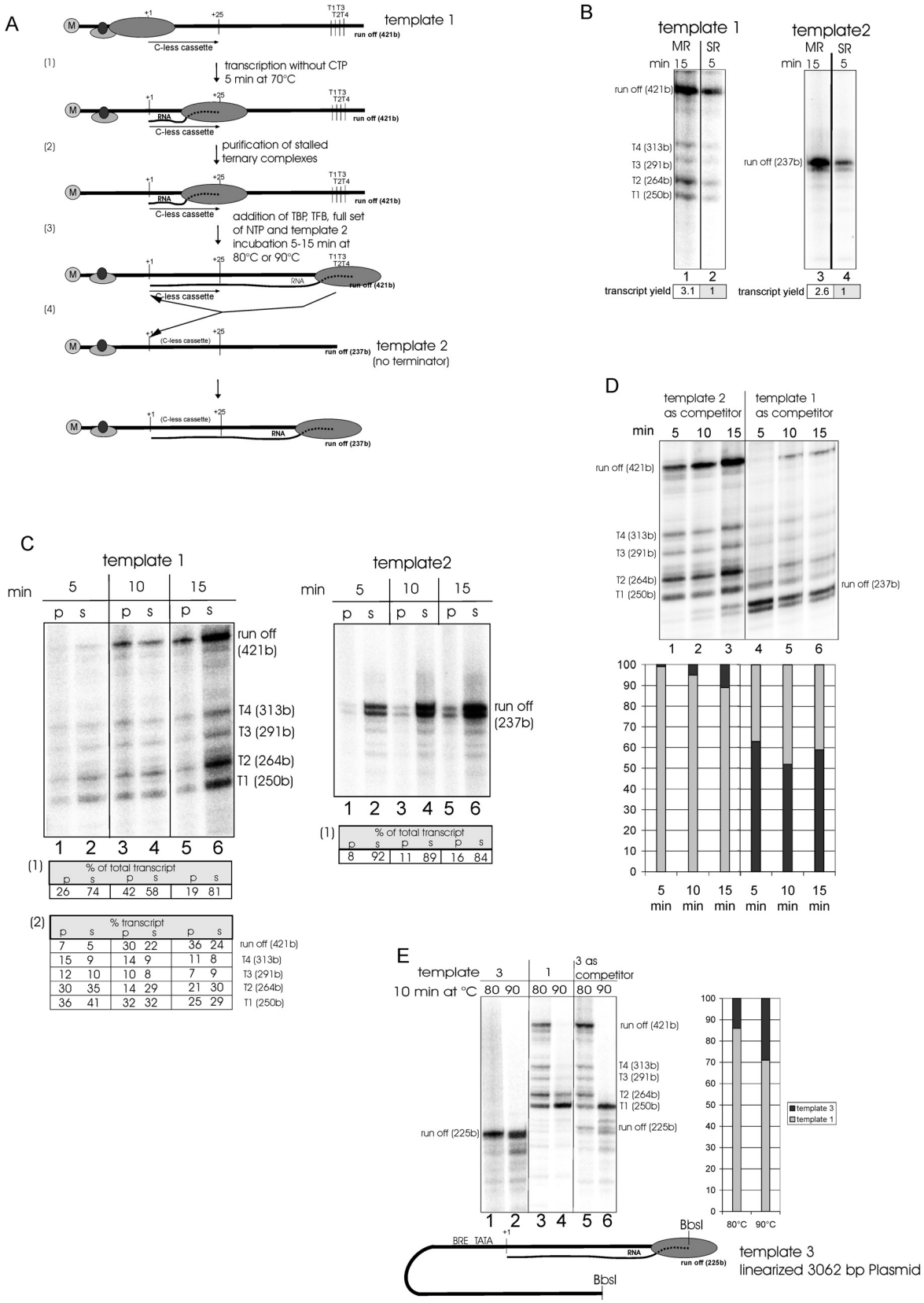

Fig. 4. Competition experiments reveal template commitment of an archaeal RNAP.
A. The *hpyA1-C25* containing the termination region (template 1) was incubated in transcription reaction without CTP to obtain stable transcription complexes stalled at position +25 (1). The complexes were isolated by magnetic attraction, the supernatant was removed and the complexes were washed with 0.5% NLS to remove all promoter bound transcription factors (2) (see Spitalny & Thomm, 2003). Then transcription buffer containing a full set of NTPs (440 µM each ATP, GTP, CTP, 2,7 µM UTP and 0,074 MBq [α-^{32}P]UTP (110 TBq/mmol)),

TBP and TFB (119 nM and 147 nM, respectively) and competitor template (template 2) was added (3). Template 2 is a truncated version of *hpyA1-C25* not containing a terminator. It produces run off transcripts ending 13 nucleotides before the first termination site. In the subsequent continued elongation at 80 or 90 °C (incubation time and temperature as indicated) the RNAP may choose either template to reinitiate transcription (4). For control reaction the same procedure was used with stalled complexes on template 2 and template 1 as competitor (Fig. 4D, right panel).
B. Single round (SR) transcription compared to multiple round transcription (MR) from the *hpyA*1 promoter at 80°C in the presence (template 1) and absence (template 2) of a linked terminator. The transcript yields were indicated below the lanes.
C. The terminator delays transcript release. The amount of RNA produced by isolated complexes after the indicated incubation time (procedure see A, steps 1-3) is shown. The RNA that is polymerase bound within the ternary complex (pellet, p) is shown in comparison to RNA that has already been released (supernatant, s). The tables below show the percentage of transcript after correction for the varying amount of radioactive UMP incorporated into RNAs of varying length. (1) shows the percentage of pelleted transcripts versus transcripts in the supernatant relative to the total of RNA (p + s) after 5, 10 and 15 min, respectively. In (2) the percentage of transcript in each band relative to the total RNA in each lane is shown.
D. The terminator directs RNAP to the same template. The experiment shown in the left panel was conducted with template 2 as competitor DNA (procedure as shown in A) and in the experiments shown in the right panel template 1 was used as competitor and template 2 to stall and isolate transcription complexes. Incubation time was as indicated. The diagram below demonstrates the relative amount of transcripts from template 1 (light grey) to transcripts from template 2 (dark grey).
E. Template commitment at 90 °C. Template 3 is a 3062 bp plasmid containing the *hypA1* gene linearized with *Bbs*I. Template 3 was not immobilized. This terminator less template 3 was used as competitor at 80 °C and at 90 °C. The diagram shows the amount of transcript from template 1 (light grey) and template 3 (dark grey), respectively. The incubation time for the reactions shown in lanes 1-6 was 10 min. The competition experiment was performed according to the scheme depicted in A, but template 3 was used as competitor.

Even after 15 minutes when more than 80% of the transcripts formed were released and RNAP molecules were free to choose a new initiation site on both templates (inferred from the data shown in Figure 4C) only about 10 percent of the total RNAs were transcribed from template 2 (Figure 4D lane 3 and corresponding diagramm below; the values were corrected for the varying amounts of UMP incorporated into the RNA products of different length). When RNAP ternary complexes were first formed on template 2 and template 1 was added as competitor a ratio of transcripts from either template of almost 50:50 was observed already after 10 minutes. The finding that preicubation with the template lacking the terminator leads to an equal distribution of transcripts from both templates excludes the possibility that the effect is being caused by the proximity of terminator and promoter of the same DNA molecule. Our results show a clear preference for the initially transcribed template in subsequent cycles of transcription in a terminator dependent manner and this can be best explained by recycling of the RNAP from terminator to promoter.

To further investigate this a template competition experiment was also performed at 90 °C using the terminator less template 3 that showed higher stability at 90 °C than the 335 nt template 2 used for the experiments shown in Fig. 4C and D. Template 3 (*hypA*1-*Bbs*I) consisted of a 3062 bp plasmid encoding the histone gene that was linearized by digestion with *Bbs*I and not immobilized (Fig. 4E). Run-off transcripts from template 3 were also transcribed with high activity at 90 °C (Fig. 4E, lane 2). When this template was used as

competitor in step 4 of Fig. 4A at 80 and 90 °C in each case transcription of template 1 was intrinsically favoured. At 80°C, only 14 percent total RNA was transcribed from template 3 added as competitor, at 90 °C only 30 percent. This finding shows that recycling on the first template occurs both at 80 and 90 °C in a terminator-dependent manner and that dissociation of RNAP from the first template is increased at 90°C.

Both terminator and associated promoter are necessary for a high transcription rate

In the pol III system (Dieci and Sentenac, 1996) facilitated recycling is dependent on the presence of a terminator sequence. As shown in Figure 4 this is also true for facilitated reinitiation in archaea. The *histone A1* gene was predicted to have a heat shock specific promoter (Gelfand *et al.,* 2000) and this provides a possible plausible explanation for the high transcription rate at 90°C reported in this work. In fact, run off transcription from the histone promoter on template 3 lacking the terminator occurred with similar activity at 80 and 90 °C (Fig. 5B, right panel) indicating this promoter shows high activity at 90 °C as predicted for a heat shock promoter. But the unusually strong increase ~2.4 fold increase of transcription at 90 °C compared to 80°C observed only on a template containing both promoter and terminator (Fig. 5A, lanes 1 and 2) is strictly dependent upon the presence of the terminator.
Can the presence of this terminator also activate the expression from additional *Pyrococcus* promoters? To investigate this, the histone terminators were ligated with 163 bp of a gene segment encoding the *Pyrococcus glutamate deydrogenase* (*gdh*) promoter. Run-off transcription assays from a template lacking the terminators revealed that this promoter was highly expressed at 80°C but showed significantly lower activity at 90 °C (Fig. 5B, left panel). At 80 °C, the construct containing the *gdh* promoter and the histone terminators had a similar template activity as the wt histone gene (Figure 5A lanes 1 and 3). At 90°C, termination efficiency at T1 linked with the *gdh* promoter was as high as at the wt histone gene, but the level of total transcripts was low compared to the levels of transcripts formed at the *hpyA1* template (Figure 5A, lanes 2 and 4). These findings suggest that the identified T-tracts can act independently as termination sites irrespective of the linked promoter from which initiation was started. The facilitated reinitiation pathway proposed here for the archaeal histone promoter is highly dependent on the presence of the terminator on the same template.

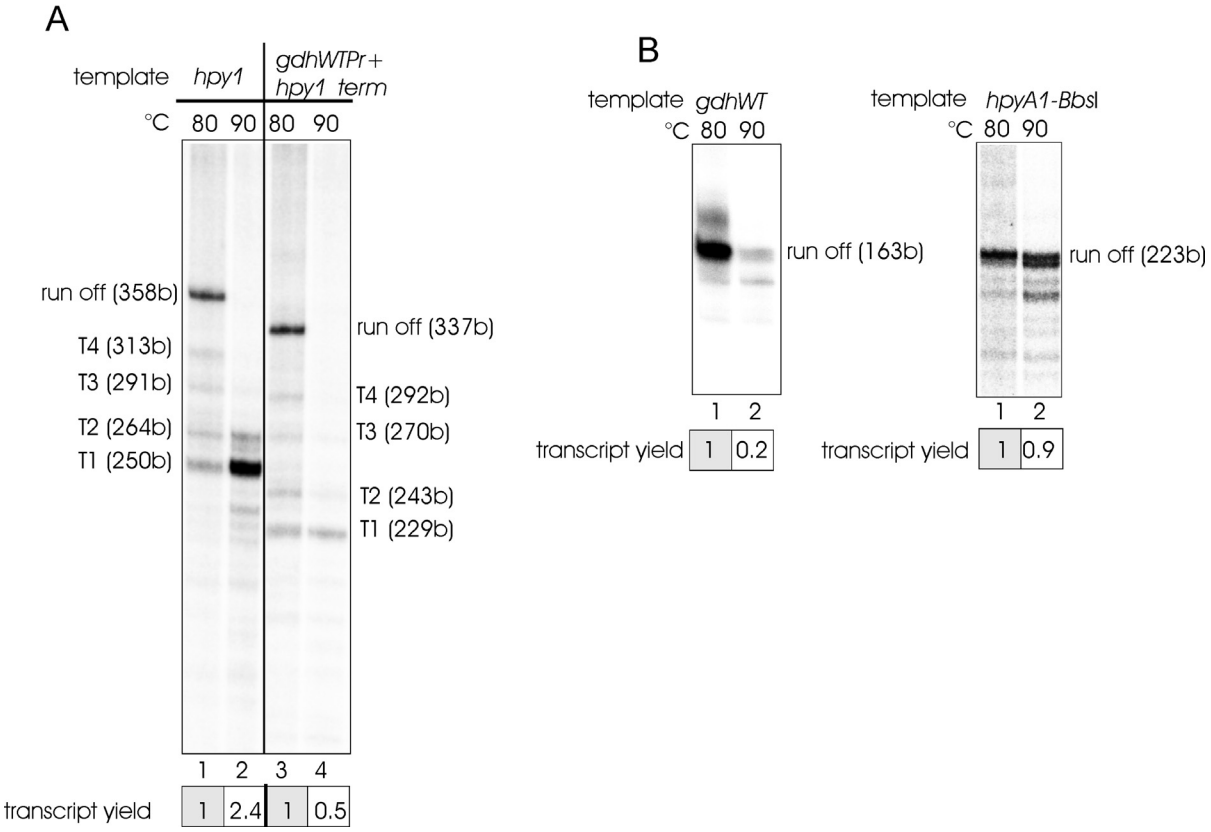

Fig. 5. Both terminator and associated promoter are necessary for high transcription rate at 90 °C.
A. The *hpyA1* terminator region was fused to the promoter region of *gdh* followed by 163 base pairs of *gdh* sequence. The wild type and the fused DNA template were incubated 10 min at the temperatures indicated on top of the lanes. The RNA products were analyzed on 6 percent PA gels. The transcriptional activity at 90 °C, relative to the activity at 80 °C which was defined as 1, is indicated below the lanes. The values for templates carrying the *hpyA*1 promoter were confirmed by four independent experiments. **B** shows run-off transcripts from a DNA fragment carrying the *gdh* promoter but lacking the *hypA1* terminator and from the truncated *hpyA1* template (template 3 of Fig. 4E) lacking the terminator region after restriction with *Bbs*I. Transcription assays were performed with 46 nM RNAP, 238 nM TBP and 147 nM TFB

The sequences immediately downstream of the main terminator affect termination efficiency and transcription rate

As described pausing of RNAP is a prerequisite to intrinsic termination. Palangat *et al.,* 2004 showed that the sequence downstream of an internal pausing site similar to a terminator sequence affects the paused conformation of RNA pol II.

To investigate wether the 3' flanking sequence to the oligo-dT tract has any influence on termination efficiency or transcription rate in archaea we constructed the mutated templates shown in Figure 6A. The 5 bp between T1 and T2 were mutated to consist of only AT- or GC-residues, respectively. *In vitro* transcription at 90°C revealed a strong effect of downstream sequences both on termination efficiency and transcription rate (Figure 6B). The diagramm in Figure 6C shows the varying termination efficiencies at T1 relative to the overall

transcript. While the termination efficiencies of the WT and GC templates do not differ significantly the AT mutant shows a much better termination at T1.

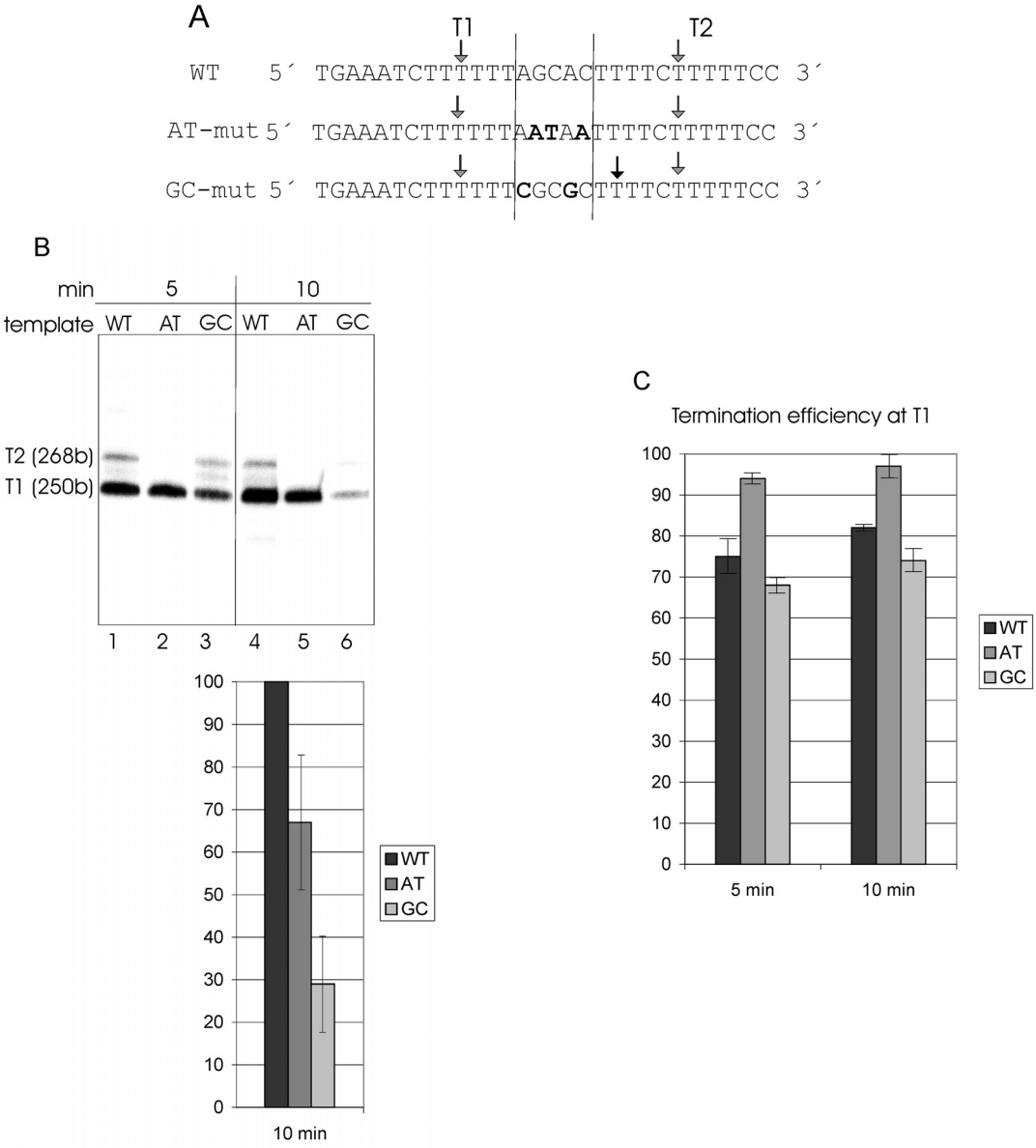

Fig. 6. The sequences immediately downstream of the main terminator affect termination efficiency and transcription levels.

A. The sequence between the main termination site (T1) and the first backup terminator (T2) was mutated to be either AT-rich or GC-rich. The substitutions are shown in bold type. The RNA release site identified by electrophoresis of RNA products form the wt template on 20% PA sequencing gel (data not shown) are indicated by grey arrows, an additional release site induced by GC-rich downstream sequences by a black arrow.

B shows the RNA products obtained from the linearized wildtype (wt) and mutated templates containing AT- and GC-rich downstream DNA after 5 and 10 min. *In vitro* transcription was performed at 90°C and incubation times were as indicated on top of the lanes. Transcription assays were performed with 46 nM RNAP, 119 nM TBP and 147 nM TFB. The diagram below displays the amount of total transcripts transcribed from the wt template and from the different mutant templates after 10 minutes.

C. The termination efficiency at T1 is shown for the wildtype and the mutant templates, respectively.

The level of overall transcripts is strongly affected by the mutation of the sequence downstream of the termination site (diagram Figure 6B). The highest amount of RNA is produced by the wiltdype DNA after 10 minutes of incubation (Figure 6B lane 4). The levels of transcripts formed at the AT mutant were reduced to around 70 percent compared with the wt template (diagram Fig. 6B). The most dramatic effect was observed with the GC mutant. After 5 min the levels of transcript were about 60 percent compared to the wt. Unexpectedly, after 10 min the transcript level was reduced to ~30 percent (Fig. 6B) of wt levels. This finding suggests that already after 5 min of incubation at 90°C further RNA synthesis is significantly impaired. Our finding that two single point mutations in the region downstream of T1 can abolish the observed activation of transcription at 90 °C (Fig.6 B, lines 4 and 6) demonstrates that the sequence within the terminator region is highly critical for recycling of RNAP. Clearly, RNA degradation is expected to occur with the same rate on transcripts from all templates shown in Fig. 6A, but this degradation is compensated on the WT template by recycling of RNAP from the terminator to the promoter. Once this mechanism is impaired by the GC mutation the rate of degradation is higher than the rate of RNA synthesis and therefore RNA levels are reduced to ~30 percent after 10 min of incubation. GC-rich DNA immediately downstream of the terminator sequence seems to abolish facilitated reinitiation activity.

Discussion

We selected a complete archaeal histone gene containing 4 consecutive downstream oligo-dT sequences as template for *in vitro* transcription experiments designed to investigate transcription termination signals and the mechanism of termination in a system operating at 90 °C. The results of this study provide evidence that termination is brought about by pol III-like terminator sequences and that the archaeal enzyme uses a pol III-like mechanism for termination.

Termination at high temperatures and terminator signals

The *Pyrococcus* histone gene *hpyA1* has been predicted as heat shock gene (Gelfand *et al.,* 2000) and in line with this prediction we find unusual high levels of *in vitro* transcription of the *hpyA1* template at 90°C *in vitro* (Figs. 1 and 5). *Pyrococcus* grows optimally between 90 and 100 °C and shows slower growth at 80 °C (Fiala and Stetter, 1986). Our finding that the termination efficiency at the first oligo-dT stretch (T1) *in vitro* is greatly enhanced at 90°C compared to 80 °C (Fig. 1) might simply reflect the adaptation of the *Pyrococcus* transcriptional machinery to growth at high temperatures. T1 is a weak terminator at 80 °C

(Fig. 1) and even the three additional back up terminators do not prevent read through to the end of the DNA-fragment at this temperature (Fig. 1, lanes 2 and 3). Therefore, the presence of these additional terminators can reduce read through during growth at lower temperature and might be an adaptation to the shallow water marine hydrothermal vent habitat of *P. furiosus* (Fiala and Stetter, 1986) that is characterized by rapidly fluctuating temperatures. The increase of the efficiency of termination with temperature could be a mechanism preventing read through and the expression of downstream genes upon heat shock and thereby contribute to a specific stimulation of expression of the *Pyrococcus* histone at elevated temperatures. The specific stimulation of histone gene expression upon heat shock seems to be physiologically important because binding of histones from hyperthermophilic archaea to DNA is likely to stabilize duplex DNA at elevated temperatures. 6 or 5 T-residues were sufficient for a high termination efficiency at 90 °C, 4 T residues lead to a efficiency of ~50 , three T-residues or less are not recognized as significant terminator signals (Fig. 2). By contrast, in the *M. thermoautotrophicus* (*M.t*) system deletion of one T-residue from the T6 stretch in the $t_{R'}$ terminator resulted in a ~fourfold reduced termination when termination efficiency of a construct containing this *E. coli* bacteriophage terminator was assayed *in vitro* (Santangelo and Reeve, 2006). This finding suggests that the signals directing efficient termination differ slightly among archaea like in eukaryotic pol III genes and systems from different species (Gunnery *et al.,* 1999). The minimal signal sufficing for termination of the archaeal enzyme at the *Pyrococcus* histone terminator resembles most the *Saccharomyces cerevisiae* system which requires also 5 or 6 T-residues as termination signal (Allison and Hall, 1985). A stem loop structure upstream of the histone terminator has no effect on termination at the *hpyA1* terminator (Fig. 3). By contrast, RNA hairpins seemed to contribute to termination efficiency in the *M. t*. system at the same incubation temperature (Santangelo and Reeve, 2006). Our findings indicate great similarities of the *Pyrococcus* termination to pol III termination at the *Xenopus* 5S gene. In both systems, a stretch of T-residues but no dyad symmetry sequences are required for termination (Bogenhagen and Brown, 1981).

The mechanism of termination is pol III-like

The findings that stem-loop structures are not necessary for termination and that addition of oligonucleotides complementary to the upstream half of the RNA hairpin stem which induce RNA release in the bacterial system (Yarnell and Roberts, 1999) have no effect on archaeal termination (Santangelo and Reeve, 2006) suggest that the mechanism of bacterial and archaeal intrinsic termination differ. In this study several lines of evidence indicate that the

mechanism of archaeal transcription is pol III-like. In the pol III system the oligo-dT stretches in terminators induce extensive pausing of RNAP (Matsuzaki *et al.,* 1994; Campell and Setzer, 1992) without formation of a stable stem-loop structure in the RNA which causes the bacterial RNA polymerase to pause and weakens its interactions with nascent RNA and template at rho-independent terminators (Reynolds and Chamberlin, 1992; Wang et al., 1997; Artsimovitch and Landick 1998; Wilson and von Hippel, 1995). We observed a significant delay of the release of RNA from template 1 containing the set of terminators compared to template 2 lacking terminator sequences (Fig. 4C). This finding suggests that the presence of the oligo-dT residues causes also extensive pausing of the *Pyrococcus* enzyme although pausing of *Pyrococcus* RNAP at these terminators has not been directly shown in this study.
The high *in vitro* transcription efficiency of pol III genes was shown to be mediated by a reinitiation mechanism bypassing most of the steps of the initial transcription cycle (Dieci and Sentenac, 1996). This reinitiation of pol III on the same template was dependent upon the presence of a functional terminator. We show here that the archaeal RNAP did not equally redistribute on both genes when a second template without functional terminator was added as competitor to stalled ternary complexes in multiple round transcription assays (Fig. 4D and E). The preferred transcription of the first template indicates rapid recycling of the archaeal enzyme from the terminator to the promoter of the same template. The finding that the expression of both templates is approximately equal when the template used for formation of ternary complexes lacks a terminator suggests that pausing of RNAP at the terminator is a prerequisite for recycling of the archaeal RNAP. Our finding that only ~10 percent of the competitor DNA is expressed 15 min after starting transcription from ternary complexes formed at the terminator containing template (Fig. 4D, lanes 5 and 6) suggests a mechanism involving reinitiation at the same gene without release of RNAP that was postulated to operate in the pol III system (Dieci and Sentenac, 2003).

Effects of downstream DNA on the reinitiation mechanism

The effects of downstream DNA on termination efficiency and pausing of RNAP have been studied in the bacterial (Lee *et al.,* 1990; Ederth *et al*., 2002) and human polymerase II system (Palangat *et al.,* 2004). In the pol III system initial (Bogenhagen and Brown, 1981) and later studies (Gunnery and Mathews, 1995; Gunnery *et al.,* 1999) revealed that the sequence context around oligo-dT terminator signals modulates termination efficiency. The dinucleotide CT immediately downstream of the 3´flank of terminators containing 5 T-residues (T_5) weakened termination efficiency, whereas an A or G residue following the T_5

track increased termination (Braglia *et al*., 2005). But on terminator sequences consisting of 6 T-residues like T1 downstream of the *hpyA1* gene analyzed in this study the weakening effect of the CT dinucleotide was lost (Braglia *et al*., 2005). In general, mutation lowering the duplex stability downstream of the T_5 track increased read through at the pol III terminator and thus weakened terminator efficiency (Braglia *et al*., 2005) By contrast, the mutation lowering the duplex stability immediately downstream of the *hpyA1* terminator increased termination efficiency (Fig.6B and C, AT mutant). In the bacterial system interstrand cross-linking of the DNA duplex downstream of the terminator decreased termination efficiency indicating that forward translocation of RNAP and melting of downstream DNA favour RNA release (Santangelo and Roberts, 2004). Considering the extreme temperature in the *Pyrococcus* system, the GC-content of downstream DNA is likely to have a more important effect on DNA melting and translocation of RNAP. According to the forward translocation model AT-rich sequences are likely to increase translocation and to reduce pausing at the terminator, GC-rich sequences probably favour pausing and inhibit downstream DNA melting and translocation of RNAP. The wt sequence downstream of *hpyA1* shows an intermediate GC-content favouring overall transcript synthesis most likely by allowing the reinitiation mechanism to occur. The AT-rich mutant sequence is likely to reduce pausing of RNAP at the terminator by stimulating downstream DNA opening and this is likely to impair reinitiation on the same template and to stimulate rapid release of RNA and therefore the termination efficiency is predicted to be higher, but the overall transcript synthesis predicted to be lower as observed (Fig. 6). The GC-rich downstream DNA mutant is predicted to favour RNAP pausing and to weaken downstream DNA opening and translocation. The finding that the GC rich mutant shows after 5 min of RNA synthesis similar termination efficiency as the wt is puzzling at first sigh. However, an additional weak termination site between T1 and T2 was utilized with higher efficiency by this mutant (indicated by an arrow in Fig. 6A) and more importantly, the overall RNA levels observed after 10 min of incubation in transcription reactions were only 30 percent of wt levels (Fig. 6B). This finding suggests that the predicted extended pausing and the observed pausing at an additional site induced by the GC-rich sequence impair the facilitated recycling of RNAP by an unknown mechanism. Because the paused RNAP molecules are not available for reinitiation and the synthesized RNA is rapidly degraded at 90 °C (Hethke *et al.,* 1999) the overall levels of RNA decrease with extended incubation time as observed (Fig. 6). Our findings indicate that a subtle balance between translocation of RNAP beyond the oligo-dT stretch and pausing at the terminator is important for termination and recycling in the archaeal system. The characteristics of this *Pyrococcus*

system, recycling and reduced stability of RNA at 90 °C, are responsible for the unexpected and to our knowledge unique effects of AT-rich and GC-rich downstream sequences on termination efficiency and the level of transcripts observed.

Experimental Procedures

DNA templates for in vitro transcription

Histone A1 gene (*hpyA1*) including promoter and terminator regions was amplified from *Pyrococcus furiosus* genomic DNA using the primers his-a1_F 5'-GGC AAT CTA TTT GGA A**T**T CGC TCT G-3' and his-a1_R 5'-GAT ATA CTT TAA TTT C**T**G **C**AG GCT C-3' containing a restriction site for *Eco*RI and *Pst*I, respectively. The fragment was inserted between the corresponding sites of pUC19. The resulting plasmids were transformed into *E. coli* JM109, amlified and purified. They were used as linearized templates after restriction with *Pst*I.

The templates with point mutations in the terminator region were constructed by PCR using two internal primers complementary to opposite strands of the plasmid pUC19 with *hpyA1*, both containing the desired point mutations. With one of the internal primers and either the M13 F or the M13 R primer two fragments with overlaping sequences in the region of the internal primers were produced. In a second PCR the overlapping sequences were fused and the complete fragments were amplified with the flanking M13 primers. The sequences of the point mutations were confirmed by sequencing and are noted in the corresponding figures.

To construct the template *gdhWTPr+hpyA1_term* the terminator region from *hpyA1* was amplified from *Pyrococcus furiosus* genomic DNA using the internal primer Hisa1_mH 5'-CAA GGC ACG CA**T** **C**TA GAA AGA C-3' and his-a1_R. After restriction of the PCR fragment with *Xba*I and *Pst*I the fragment was inserted between the corresponding restriction sites directly downstream of a *gdh* (glutamate dehydrogenase) gene segment in pUC19 containing the *gdh* sequence from -95 to +163 from *Pyrococcus furiosus.* The following steps were as described above. The mutations were confirmed by sequencing.

In vitro transcription assays

In vitro transcription assays were performed as described previously (Hethke et al. 1996). A standard transcription reaction mixture of 25 µl contained 250 ng linearized plasmid DNA, 46 nM RNAP, 238 nM or 119 nM recombinant TBP (as indicated in figure legends), 147 nM recombinant TFB, 440 µM each ATP, GTP, CTP, 2,7 µM UTP and 0,074 MBq [α-^{32}P]UTP

(110 TBq/mmol). The transcription buffer contained 40 mM HEPES, 0.1 mM EDTA, 1 mM DTT, 275 mM KCl, and 3 mM $MgCl_2$. Transcription reactions were performed 10 min at 90°C or as indicated in figures. Transcription reactions were stopped by the addition of loading buffer (98% formamide, 10 mM EDTA and 0.1% each bromphenol blue and xylene cyanol). Labeled transcripts were separated by electrophoresis on 6% polyacrylamide urea gels and visualized by phosphorimaging (FLA-5000, Fuji, Japan).

Immobilized in vitro transcription assays and competition experiments

To allow pausing and isolation of RNAP on the immobilized DNA template a C-minus cassette was introduced into the *hpyA1* gene by PCR. Two internal primers complementary to opposite strands of the *hpyA1* sequence from -6 to +23 relative to the transcription start site were used to substitute all C residues until position +25 (hpfA1-C25F 5'-CAA AAT GGA AAT GTG TTA TAA ATA AAA GG-3', hpfA1-C25R 5'-CCT TTT ATT TAT AAC ACA TTT CCA TTT TG-3'). The plasmids were constructed and transformed into *E.coli* as described above. The templates for immobilized *in vitro* transcription reaction were produced by PCR of pUC19 containing *hpyA1* with C-minus cassette (*hpyA1-C25*) by the use of the primers M13F and M13R. M13F was 5' modified with biotin and the resulting PCR fragments were immobilized on streptavidin magnetic beads (Roche Applied Science) according to the protocol of the manufacturer. To isolate transcription complexes the immobilized template *hpyA1-C25* was incubated in transcription reaction. A 25 µl reaction contained 30 ng immobilized template, 46 nM RNAP, 119 nM recombinant TBP, 147 nM recombinant TFB, 40 µM each ATP, GTP, 2 µM UTP and 0,074 MBq [α-^{32}P]UTP (110 TBq/mmol). Buffer conditions were as listed above. Transcription reactions were performed 5 minutes at 70°C. Transcription complexes paused at position +25 were isolated by magnet attraction at room temperature, washed with transcription buffer containing 0.5% N-lauroylsarcosine (NLS). Washing with NLS removes promoter bound transcription factors but stalled RNAP is retained in ternary complexes (Spitalny and Thomm, 2003). The isolated ternary complexes were resuspended in transcription buffer and supplemented with all four nucleotides (440 µM each ATP, GTP, CTP, 2.7 µM UTP and 0,074 MBq [α-^{32}P]UTP (110 TBq/mmol)), 119 nM recombinant TBP and 147 nM recombinant TFB to allow the stalled RNAPs to continue elongation (chase). Chase reaction was performed at 80 °C or 90 °C and the incubation time is indicated in the corresponding figures. For single round transcription TBP and TFB were omitted from the chase reactions (step 3 of Fig. 4D). Multiple round transcription in Fig. 4 B was performed by washing with transcription buffer not containing

NLS. These reactions contained no excess transcription factors but promoter bound transcription factors directed multiple rounds of transcription by RNAPs isolated by steps 1 and 2 of Fig. 4A.

For competition experiments immobilized *hpyA1-C25* templates lacking the terminator were constructed by PCR. The M13 F primer was biotinylated and the reverse primer was complementary to the region 12 -33 nucleotides before the terminator region. 30 ng of this template were added to the chase reaction of the isolated transcription complexes performed as described above.

Acknowledgments

This work was supported by a grant form the Deutsche Forschungsgemeinschaft.

References

Allison, D.S., and Hall, B.D. (1985) Effects of alterations in the 3' flanking sequence on *in vivo* and *in vitro* expression of the yeast *SUP*4-o tRNA^{Tyr} gene. *EMBO J* **4:** 2657-2664.

Artsimovitch, I., and Landick, R. (1998) Interaction of a nascent RNA structure with RNA polymerase is required for hairpin-dependent transcriptional pausing but not for transcript release. *Genes Dev.* **12:** 3110-3122.

Bell, S.D., and Jackson, S.P. (2001) Mechanism and regulation of transcription in archaea. *Current Opinion Microbio,* **4:** 208-213.

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

Bogenhagen, D.F., and Brown, D.D. (1981) Nucleotide sequences in Xenopus 5S DNA required for transcription termination. *Cell* **24:** 261-270.

Braglia, P., Rercudani, R., and Dieci, G. (2005) Sequence context effects on oligo(dT) termination signal recognition by *Saccharomyces cerevisiae* RNA polymerase III. *J Biol Chem* **280:** 19551-19562.

Brown, J.D., Daniels, C., and Reeve, J.N. (1989) Gene structure, organization and expression in archaebacteria. *Crit Rev Microbiol* **16:** 287-338.

Campbell, F.E., and Setzer, D.R. (1992) Transcription termination by RNA polymerase III: uncoupling of polymerase release from termination signal recognition. *Mol Cell Biol* **12:** 2260-2272.

Chu, W.-M., Ballard, R.E., and Schmid, C.W. (1997) Palindromic sequences preceding the terminator increase polymerase III template activity. *Nucl Acids Res* **25:** 2077-2082.

Cozzarelli, N.R., Gerrard, S.P., Schlissel, M., Brown, D.D., and Bogenhagen, D.F. (1983) Purified RNA polymerase III accurately and efficiently terminates transcription of 5S RNA genes. *Cell* **34:** 829-835.

Dieci, G., and Sentenac, A. (1996) Facilitated recycling pathway for RNA polymerase III. *Cell* **84:** 245-252.

Dieci, G., and Sentenac, A. (2003) Detours and shortcuts to transcription reinitiation. *Trends Biochem Sci* **28:** 202-209.

Ederth, J., Artisimovitch, I., Isaksson, L.A., and Landick, R. (2002). The downstream DNA jaw of bacterial RNA polymerase facilitates both transcriptional initiation and pausing. *J Biol Chem* **277:** 37456-37463.

Fiala, G., and Stetter, K.O. (1986) *Pyrococcus furiosus* sp. Nov. represents a novel genus of marine heterotrophic archaebacteria growing optimally at 100°C. *Arch Microbiol* **145**: 56-61.

Geiduschek, E.P., and Kassavetis, G.A. (1992) RNA polymerase III transcription complexes. In: Transcription Regulation (McKnight, S.L., and Yamamoti, K.R. (eds), Vol. **1**, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NW, pp. 247-280.

Geiduschek, E.P., and Ouhammouch, M. (2005) Archaeal transcription and its regulators. *Mol Microbiol* **56:** 1397-1407.

Gelfand, M.S., Koonin, E.V., and Mironov, A.A. (2000) Prediction of transcription regulatory sites in Archaea by a comparative genomic approach. *Nucl Acids Res* **28:** 695-705.

Grünberg, S., Bartlett, M.S., Naji, S., and Thomm, M (2007) Transcription factor E is a part of transcription elongation complexes. *J Biol Chem* in Press published on October 5, 2007 as doi:10.1074/jbs.M707371200.

Gunnery, S., and Mathews, M.B. (1995). Functional mRNA can be generated by RNA polymerase III. *Mol Cell Biol* **15:** 3597-3607.

Gunnery, S., Ma, Y., and Mathews, M.B. (1999) Termination sequence requirements vary among genes transcribed by RNA polymerase III. *J Mol Biol* **286:** 745-757.

Hamada, M., Sakulich, A.L., Boduru, S.B., and Marain, R.J. (2000) Transcription termination by RNA polymerase III in fission yeast. *J Biol Chem* **275:** 29076-29081.

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

Hethke, C., Bergerat, A. Hausner, W., Forterre, P., and Thomm, M. (1999) Cell-free transcription at 95°: Thermostability of transcriptional components and DNA topology requirements of *Pyrococcus* transcription. *Genetics* **152:**1325-1333.

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

Hethke, C. Aufstellung und Charakterisierung eines zellfreien Transkriptionssystems für das hyperthermophile Archaebakterium *Pyrococcus furiosus*. PhD thesis, University of Kiel.

Kassavetis, G.A., Riggs, D.L., Negri, R., Nguyen, L.H., and Geiduschek, E.P. (1989) Transcritpion factor IIIB generates extended DNA interactions in RNA polymerase III transcription complexes on tRNA genes. *Mol Cell Biol* **9:** 2551-2566.

Lee, D.N., Phung, L., Stewart, J., and Landick, R. (1990) Transcription pausing by *Escherichia coli* RNA polymerase is modulated by downstream DNA sequences. *J Biol Chem* **265:** 15145-15153.

Matsuzaki, H., Kassavetis, G.A., and Geiduscheck, E.P. (1994) Analysis of RNA chain elongation and termination by *Saccharomyces cerevisiae* RNA polymerase III. *J Mol Biol* **235:** 1173-1192.

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

Naryshkina, T., Kuznedelov, K., and 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**: 634-643

Palangat, M., Hittinger, C.T., and Landick, R. (2004) Downstream DNA selectively affects a paused conformation of human RNA polymerase II. *J Mol. Biol* **341:** 429-442.

Reiter, W.D., Palm, P., and Zillig, W. (1988) Transcription termination in the archaebacterium *Sulfolobus*: signal structures and linkage to transcription initiation. *Nucl Acids Res* **16:** 12445-12459.

Reynolds, R., and Chamberlin, M.J. (1992) Parameters affecting transcription termination by *Escherichia coli* RNA. *J Mol Biol* **224:** 53-63.

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:** 196-210.

Santangelo, T.J., and Roberts, J.W. (2004) Forward translocation is the natural pathway of RNA release at an intrinsic terminator. *Mol Cell* **14:** 117-126.

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

Thomm, M. (2007) Transcription: Mechanism and Regulation. In: Archaea Mollecular and Cellular Biology. R. Cavicchioli (ed.). ASM Press, Washington DC, U.S.A., Chapter *6,* 139-157.

Thomm, M., Hausner, W., and Hethke, C. (1994) Transcription factors and termination of transcription in *Methanococcus*. *System App. Microbiol* **16:** 148-155.

Wang, D., Severinov, K., and Landick, R. (1997) Preferential interaction of the *his* pause RNA hairpin with RNA polymerase β subunit residues 904-950 correlates with strong transcriptional pausing. *Proc Natl Acad Sci USA* **94:** 8433-8438.

Werner, F., and Weinzierl, R.O. **(**2005) Direct modulation of RNA polymerase core functions by basal transcription fabors. *Mol Cell Biol* **25**: 8344-8355.

Wilson, K.S., and von Hippel, P.H. (1995) Transcription termination at intrinsic terminators: the role of the RNA hairpin. *Proc Natl Acad Sci USA* **92:** 8793-8797.

Wolffe, A.P., Jordan, E., and Brown, D.D. (1986) A bacteriophage RNA polymerase transcribes through a *Xenopus* 5 S RNA gene transcription complex without disrupting it. *Cell* **44:** 381-389.

Yarnell, W.S., and Roberts, J.W. (1999) Mechanism of intrinsic transcription termination and antitermination. *Science* **284:** 611-615.

Yin, H., Artsimovitch, I., Landick, R., and Gelles, J. (1999) None-equilibrium mechanism of transcription termination from observations of single RNA polymerase molecules. *Proc Natl Acad Sci USA* **96:** 13124-13129.

V General Discussion

It is well known that Archaea, apart from having their unique characteristics, share many features with Bacteria and Eukarya. While morphological structures and metabolic proteins show similarities to those of Bacteria (Aravind and Koonin, 1999), the archaeal information processing pathways, including DNA replication, transcription and translation, are closely related to that of Eukarya (Bell and Jackson, 1998b; Edgell and Doolittle, 1997).
Several new mechanistic aspects of the archaeal transcription cycle were uncovered by the experimental approaches of the present thesis. The results are put into context with eukaryotic and bacterial transcription systems and many parallels between the mechanisms of transcription by RNA polymerases (RNAPs) from all three domains of life could be identified. The *in vitro* transcription system from the hyperthermophilic *Pyrococcus furiosus* (Hethke *et al.*, 1996) provided the basis to study archaeal transcription at its different steps in detail.

V.1 Initiation and elongation

This thesis contributed to a deeper understanding of initiation and the transition to elongation of archaeal transcription. The avaliability of an *in vitro* assembled recombinant form of the *Pyrococcus* RNA polymerase (RNAP; Naji *et al.*, 2007) allowed rapid site-directed mutagenesis and functional studies, independently of cell viability of RNAP mutants. The sequences of the archael RNAP are closely related to the eukaryotic RNA polymerase II (pol II; Langer *et al.*, 1995). High resolution structures of pol II (Bushnell and Kornberg, 2003; Kettenberger *et al.*, 2004) suggest structural elements important for initiation and elongation in eukaryotic transcription. The structural analyses revealed four prominent loops that are located in the polymerase cleft. These loops, named lid, rudder, fork1 and fork2, are in close contact with the transcription bubble and with the RNA-DNA hybrid. The functions of these loops have not yet been investigated in pol II, but due to the high degree of their conservation in pol II of yeast, in the bacterial RNAP and in the archaeal RNAP (see chapter III, figure 1) they could be identified and deleted in the recombinant archaeal enzyme. In addition three point mutants, one in fork2 and two in the structural element called switch2, were constructed (chapter III, Souad Naji). Experiments designed to test hypotheses that arose from the atomic structure of pol II were performed with the mutant archaeal enzymes. Specific transcription assays on the strong *gdh* (glutamate dehydrogenase) gene revealed significant transcriptional

defects in the mutants Δlid, Δrudder, Δfork2 and A'-R313A (chapter III, Souad Naji). In chapter III a stepwise analysis of the mutants in different phases of the transcription cycle demonstrated the functional importance of these structures.

The first step in the transcription cycle is closed complex formation, characterized by the assembly of transcription factors and recruitment of the RNAP. In *Pyrococcus* at incubation temperatures of 60°C or higher the closed complex converts into the open complex formation instantly and cannot be measured. Therefore closed complex formation was probed by band-shift assays performed at 37°C. It could be shown that obviously all mutant polymerases bind to the promoter platform (chapter III, Souad Naji).

The following open complex formation is a crucial step in transcription initiation and a plausible mechanism based on the atomic structure of pol II was suggested (Cramer *et al.*, 2001). The relevance of the mutated structures in open complex formation was analysed during this thesis by probing the initiation complexes with the single strand specific reagent $KMnO_4$. Though the lid element was shown to be important in stabilizing the open complex in the bacterial system (Toulokhonov and Landick, 2006), the experiments in this thesis led to the conclusion that all mutant enzymes, apart from Δrudder, opened DNA around the transcription initiation site. Structure-function models of eukaryotic (Westover *et al.*, 2004) and bacterial (Korzheva *et al.*, 2000) transcribing complexes show the rudder element in contact with DNA and nascent RNA at the upstream region of the DNA-RNA hybrid and suggest a role for the rudder element in formation and/or maintainence of the transcription bubble. The rudder element has been deleted in the bacterial RNAP (Kuznedelov *et al.*, 2002) and it could be demonstrated that in fact the transcription bubble was shortened in the downstream direction. Furthermore the rudder deletion dramatically destabilized nascent RNA in bacterial elongation complexes.

When the transcript reaches a certain length after initial RNA synthesis, the RNAP undergoes a transition to the elongation state. This transition is often referred to as promoter escape due to the disruption of all promoter specific contacts. It is a highly complex stage in early transcription, characterized by functional instability of the transcription complex and it has been addressed by extensive studies in both bacterial and eukaryotic transcription systems. However, many aspects still remain unclear. Chapter II of the present thesis elucidates mechanistic events that accompany promoter escape in the archaeal transcription system. Two

different footprinting assays have been used to follow the RNAP along the DNA in early transcription. The RNAP was stalled in registers +5 to +11, +15 and +20 relative to the transcription start site by using C-minus cassettes (chapter II, figure 1). The immobilization of the DNA template on magnetic beads allowed rapid isolation of ternary complexes that were still transcriptionally active (chapter II, figure 2). These complexes were incubated with exonuclease III to analyse the boundaries of the transcribing enzyme in the defined registers (chapter II, figure 3). The treatment with $KMnO_4$ showed apart from the size of the transcription bubble also the approximate extent of the RNA-DNA hybrid (chapter II, figure 4). Taken together, the analyses from chapter II identified for the first time the dimensional parameters of a transcribing archaeal RNAP and two distinct steps during promoter escape (chapter II, figure 5). It could be shown that the transition from initiation to elongation follows similar principles in all three domains of life. Promoter escape by pol II is described in a review by Dvir (2002). The critical point at which the early transcription complex reaches its full stability is determined by the formation of a 4 nt transcript (Kugel and Goodrich, 2002). It marks the first transition step referred to as “escape commitment”. The existence of this step in the pol II system is supported by the resolved structure of a pol II elongating complex which shows the RNA deeply buried in the pocket of the active site until position 3. The fourth nucleotide is already exposed and shows only limited interactions with the RNA polymerase (Gnatt *et al.*, 2001). At this point the polymerase domain called clamp can close over the RNA-DNA hybrid and stabilize the initiation complex. Chapter II gives no evidence that a similar step occurs in the archaeal system. The first transition that could be observed was between positions +6 and +7. At this stage an upstream end could be detected for the RNAP and an extension of the transcription bubble two nucleotides downstream of the stop position was observed. The ability of the RNAP to transcribe independently of the initiation factors at this stage indicates a transition critical for complex stability, as described for pol II at position +4. The downstream end of the RNAP remains stably at position +18 in registers +5 to +9, while the transcription bubble extends.

Around register +10 a second clear transition step occurs characterized by the forward translocation of the downstream end of the RNAP. The ratio of nascent RNA within the transcription complex to released transcripts increases significantly once the RNA reached the length of 10 nt (chapter II, table I) demonstrating a significantly lower potential for abortive transcription beyond this register. This transition between positions +9 and +11 was also documented for pol II (Holstege *et al.*, 1997). The archaeal transcription complexes showed a collapse in the upstream region of the transcription bubble, i.e. an abrupt reannealing of the

DNA strands, by the time the transcript was 10 nt in length. A central importance of reannealing for promoter escape by pol II was demonstrated by Pal *et al.* (2005). The bubble collaps in the pol II system occurs when the transcription bubble reaches a size of 17-18 bases and the RNA transcript has a length of at least 7 nt. It is assumed that the energy used to melt the DNA is retained within the transcription complex. The stability of the transcription complex decreases by the increase of the bubble and finally a critical state is reached. The collaps of the bubble is the consequence and the stored energy is used to reorganize a stable elongation complex. This may also be true for the archaeal system. Discontinuous extension of the transcription bubble relative to the RNAP during early transcription ends at position +9. The bubble reaches a maximum size of 18 bases and immediately afterwards it closes to 14 bases at position +10, while the enzyme translocates downstream.

Chapter III also provides interesting information on the transition from initiation to elongation. In figure 5 of chapter III two predominant transcripts of 11 and 12 nucleotides in length were obtained in transcription assays with the mismatch bubble dependent on the presence of the transcription factors TBP and TFB (Souad Naji). The structure of a complex between TFIIB and pol II revealed that TFIIB reaches into the pol II through the RNA exit channel. Its B-finger domain extends to the catalytic center and will clash with the nascent RNA once it reaches a length of 5-6 nt (Bushnell *et al.*, 2004). The reorganisation of the transcription complex following the bubble collaps, described by Pal *et al.*, 2005, includes TFIIB displacement thus the exit channel can be entered by the nascent RNA. Also in Archaea promoter escape and the dissociation of TFB are coupled as shown for *Methanothermobacter thermoautotrophicus* (Xie and Reeve, 2004). TFB may remain partly bound on the premelted templates and impede proper promoter escape. In addition the mismatch bubble in the experimental design of figure 5 in chapter III does not allow bubble collaps to occur so this can also be assumed to be the reason for the production of the 11-13 nt products. But figure 5E demonstrates that these products are also synthesized on a duplex template (Souad Naji). Therefore it is unlikely that the missing bubble collaps is responsible for the increased production of the 11-13 nt RNAs. Other transcription experiments during this thesis showed effective synthesis of run off transcripts on pre-melted templates that are about 100 bp long. Thus, *in vitro* transcription experiments were performed on templates that were extended by varying numbers of basepairs downstream of the mismatch bubble. It could be demonstrated that a minimal length of 27 nt downstream of the bubble allows transcription past the barrier (chapter III, figure 5F). The downstream DNA requirements for promoter escape in the archaeal system are consistent with analyses showing that synthesis of RNA of

15 nt or longer was inhibited at the AdML promoter when the downstream DNA was cleaved at position +39 (Dvir *et al.*, 1997). Cleavage further downstream at position +50 did not affect promoter escape.

The transition from initiation to elongation leads to an early elongation complex. In chapter II of this thesis an early elongation complex stalled at position +20 is characterized. The features characterizing an elongation complex that are acquired during early transcription, are similar in different RNAPs. In figure 1 early elongation complexes from the three domains of life are compared. In the present thesis it was shown that the archaeal elongation complex at position +20 covers about 29 bp of DNA. A similar extent of the RNAP binding site could be analysed for *E.coli* (Metzger *et al.*, 1989) and pol I (Kahl *et al.*, 2000) transcribing complexes. The dimensions of pol II elongation complexes vary, depending on the experimental approach, between 26 and 55 bp of protected DNA (Fiedler and Timmers, 2001; Kireeva *et al.*, 2000; Linn and Luse, 1991; Samkurashvili and Luse, 1998). But consistently with the data for the other RNA polymerases, several pol II complexes analysed by exonuclease III footprinting show a binding site that covers around 32 bp (Samkurashvili and Luse, 1998). For the archaeal transcription bubble 15 to 17 bp of DNA are melted, similar to the bubble of pol I. The bacterial RNAP and pol II have slightly shorter transcription bubbles. The archaeal RNA-DNA hybrid with a length of about 9 bp is similarly present in bacterial and eukaryotic transcription complexes. This is in line with the analyses of the pol II structure (Vassylyev *et al.*, 2007; Westover *et al.*, 2004). It was demonstrated for pol II that the loops lid and rudder are involved in separation of RNA from DNA once the transcript length has reached 8 nt. The further extension of RNA leads it beyond the lid that interacts with residues -8 to -10 of the RNA, thereby separating the strands and leading the RNA towards the exit path. The rudder is involved in stabilizing the transcription bubble upstream of the RNA-DNA hybrid.

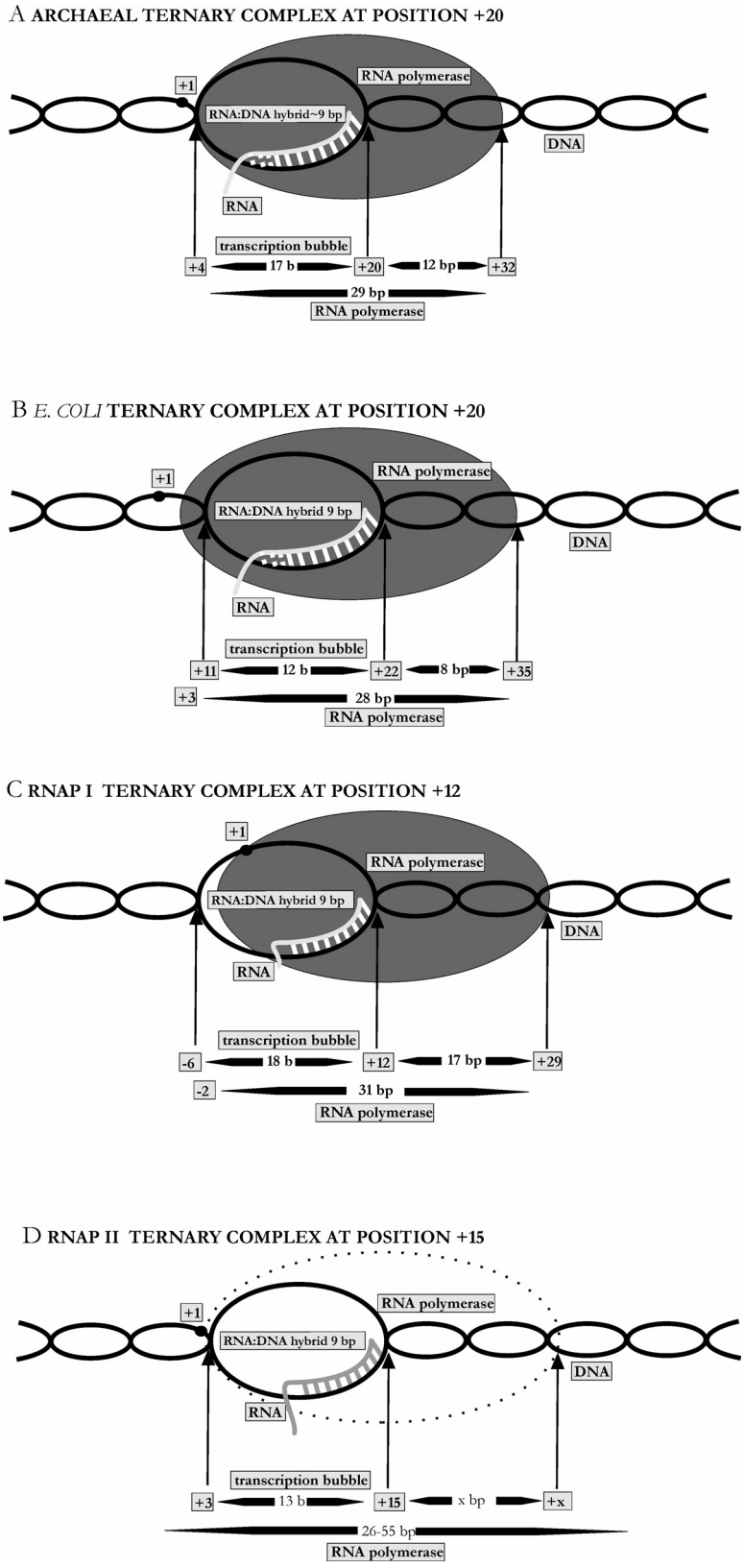

Figure 1. Comparison of archaeal (A), bacterial (B) and eukaryotic (C, D) ternary transcription complexes. The dimension of the RNAP, the size of the transcription bubble and the length of the RNA-DNA hybrid are depicted for each enzyme. The data are based on the following studies: Archaea: Spitalny and Thomm, 2003, *E.coli*: Zaychikov *et al.*, 1995, pol I: Kahl *et al.*, 2000 and the data for pol II are summarized from Fiedler and Timmers, 2001; Kireeva *et al.*, 2000; Linn and Luse, 1991; Samkurashvili and Luse, 1998.

Though being of different complexitiy, the RNAPs of all three domains of life share the basic transcriptional mechanisms. Chapter II and III of this thesis revealed several mechanistic and structural aspects, which serve to a better understanding of the fundamental processes that accompany initiation and elongation of archaeal transcription.

V.2 Termination

Termination of RNA synthesis is an essential process in the transcription cycle. It is required to prevent read throughs that may interfere with transcription of downstream genes and it leads to the maintainence of a pool of transcription competent RNAPs. Despite its importance for controlled gene expression, the mechanism of termination in Archaea is still unclear. The present thesis shows for the first time efficient intrinsic termination after transcription of a complete protein encoding gene in a hyperthermophilic archaeal system (chapter IV, figure 1). In addition a transcriptional recycling mechanism was revealed that was not yet known to exist in Archaea.

Intrinsic termination is well known in the bacterial system. In the fully nucleic acid-based termination a simple sequence encoding for a hairpin structure, followed by a uridine-rich region is sufficient to destabilize elongation complexes in Bacteria (reviewed in Platt, 1986). Most termination models propose that formation of the hairpin shortens the weak rU/dA hybrid (Komissarova *et al.*, 2002), leading to a partly rewinding of the DNA strands within the bubble region (Ryder and Roberts, 2003). The dissociation of the transcription complex follows. The oligo-dT sequence is needed to pause the elongation complex at the end of the T-tract, so the hairpin formation can take place (Gusarov and Nudler, 1999). A variant of this model is suggested by Yarnell and Roberts (1999). They assumed that termination is brought about by forward translocation of the RNAP and the transcription bubble in the absence of RNA synthesis This was confirmed by experiments of Santangelo and Roberts (2004). Hairpin formation drives the RNAP downstream while the bubble is rewound in the upstream region. The RNA-DNA hybrid is shortened significantly and thereby the entire complex is destabilized.

Though it has been documented that pol II shows some sensitivity to U-rich termination signals (Dedrick *et al.*, 1987; Komissarova *et al.*, 2002) and also in pol I termination an oligo-dT stretch is involved (Lang and Reeder, 1995), simple intrinsic termination in Eukaryotes is only known for pol III. Termination in pol III is induced by extensive pausing at an oligo-dT stretch (Matsuzaki *et al.*, 1994) and is apparently independent of additional factors. The few studies dealing with archaeal termination suggest an intrinsic termination at oligo-dT

sequences. In Bacteria the oligo-dT region is usually 8-10 bases long and has a dual role. U1-U5 in the nascent transcript are necessary for hairpin formation while U7 to U9 are responsible for pausing of the RNAP (Gusarov and Nudler, 1999). For archaeal termination in the hyperthermophilic system 5 T residues were sufficient to efficiently mediate termination (chapter IV, figure 2). Experiments to determine the precise position of termination, located the 3'end of RNA at the second, third and fourth T residue (data not shown). In addition in the archaeal intrinsic mechanism at the *hpyA1* terminator, formation of a hairpin is not required to terminate transcription (chapter IV, figure 3). These results show similarities with the findings of Bogenhagen and Brown, 1981, demonstrating that terminator preceding sequences capable of hairpin formation do not affect termination in the pol III system. Therefore intrinsic termination of *hpyA1* is pol III-like and clearly distinct from the bacterial process.

High levels of transcripts are produced by pol III. As demonstrated by Dieci and Sentenac, 1996, this is due to a termination dependent hyperprocessive reinitiation mechanism. It is characterized by the template commitment of the RNAP such that pol III obviously remains associated with the DNA template after the first round of transcription and rapidly reloads onto the same template in the following cycles. Transcription reinitiation in general significantly influences the level of RNA in living cells and has been described for all multisubunit RNAPs apart from the archaeal enzyme (Dieci and Sentenac, 2003). It is based on the assumption that some events during the first transcription cycle are responsible for accelerated transcription in subsequent cycles. These events can comprise modifications that are template and/or protein based (Dieci and Sentenac, 2003). The protein-based reinitiation mechanisms involve transcription proteins and their modifications during the transcription cycle.

Three different template based strategies of reinitiation are described by Dieci and Sentenac (2003) and depicted in figure 2. In basal reinitiation all time consuming steps of preinitiation complex formation in the first transcription cycle have to be repeated in subsequent rounds. In the PIC (preinitiation complex) assisted reinitiation one or more initiation factors remain bound to the promoter thus bypassing one or more assembly steps for RNAP recruitment. In the hyperprocessive reinitiation mechanism, so far only documented for pol III, the reinitiation process has gained its highest efficiency.

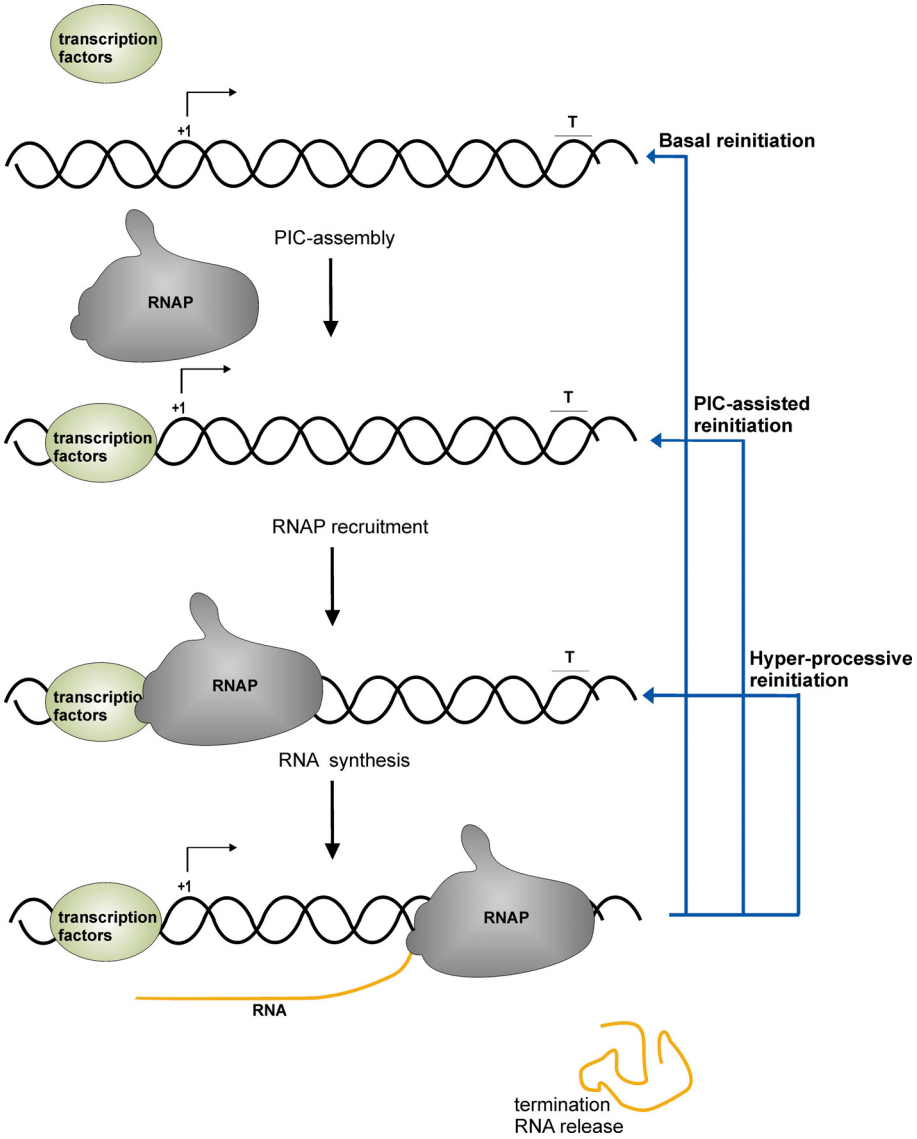

Figure 2. Template based reinitiation mechanisms after Dieci and Sentenac, 2003 (modified). Initiation of transcription requires the assembly of transcription factors in the promoter region to form a preinitiation complex (PIC). In a next step the RNA polymerase is recruited to the transcription start site (+1) and RNA synthesis is initiated. Transcription ends at a termination site (T) and the RNAP reinitiates subsequent rounds of transcription. Basal reinitiation requires all assembly steps of the initial cycle. PIC-assisted reinitiation occurs when one or more transcription factors remain promoter bound during multiple transcription cycles. In the hyper-processive reinitiation pathway the RNAP directs multiple rounds of transcription from the same template without being released.

The similarities between pol III termination and termination at the *hpyA1* terminator coupled with the detection of unusually high levels of RNA produced in *in vitro* transcription reactions at 90°C, suggested the existence of a similar reinitiation mechanism in Archaea. Competition experiments in chapter IV, figure 4 demonstrated that in multiple rounds of transcription the archaeal RNAP is committed to the template it initially transcribed. Similar to the mechanism described for pol III (Dieci and Sentenac, 1996), the recycling mechanism of the archaeal RNAP is clearly dependent on a coupling of the terminator with the associated promoter. Run

off transcription from a template harbouring the *hpyA1* promoter but lacking the terminator was equally high at 80°C at 90°C (chapter IV, figure 5B). When the terminator is linked with the associated promoter, a 2.4 fold increase of the transcript yield at 90°C can be observed (chapter IV, figure 5A).

As described by Dieci and Sentenac (1996), reinitiation pathways exist for the eukaryotic pol I-III as well as in the bacterial system. Pol I needs to be released at the terminator for repetitive polymerase recruitment due to the length of rRNA genes. PIC-assisted reinitiation is brought about by a rapid recruitment to a preformed preinitiation complex. In addition in the mammalian pol I system high frequency reinitiation was shown to be facilitated by the action of PTRF (Pol I and transcript release factor), a factor that dissociates transcription complexes paused by the termination factor TTF-I (Jansa *et al.*, 2001). PIC-assisted reinitiation is also known in the pol II system. While TFIIB and TFIIF dissociate, other components of the preinitiation complex remain at the promoter and favor a rapid reloading of the RNA polymerase (Yudkovsky *et al.*, 2000). The pol III system is known for the hyperprocessive reinitiation mechanism but also PIC-assisted reinitiation exists in pol III (Bieker *et al.*, 1985). PIC-based reinitiation in Bacteria is dependent on the σ-factor. The basal transcription cycle assumes release of the σ-factor after initiation but it could be shown that σ^{54} remains promoter bound after the transition to elongation (Tintut *et al.*, 1995). Other studies demonstrate that σ^{70} can be retained by transcription elongation complexes (Bar-Nahum and Nudler, 2001). Both mechanisms can contribute to regulatory pathways in transcription reinitiation.

Neither in prokaryotes nor in eukaryotes the intrinsic mechanism is the only way of terminating transcription. In Bacteria a forceful termination induced by the rho (ρ)-factor acts besides the intrinsic mechanism. The ρ-factor is composed of identical subunits building a ring shaped hexamer. It specifically interacts with the nascent RNA and uses physical forces to dissociate it from the transcription complex (Richardson, 2003). Although no homologous factors of ρ are present in Archaea, it has been shown that *Methanothermobacter thermoautotropicus* elongation complexes are sensitive to the ρ-factor (Santangelo and Reeve, 2006) and also eukaryotic pol II, but not pol I or pol III, elongation complexes could be disrupted by the ρ-factor (Lang *et al.*, 1998).

Eukaryotic pol I termination in yeast is dependent on the DNA-binding factor Reb1p. Reb1p binds specifically to the end of the transcription unit, causes pausing and stimulates transcript release over a T-stretch located just upstream (Lang and Reeder, 1995). In mammals

termination of pol I is mediated by TTF-I (transcription termination factor for pol I) and dissociation of the transcription complex is brought about by PTRF (Pol I and transcript release factor; Jansa *et al.*, 2001). For pol II termination two different models exist (Kornblihtt, 2004) having in common that pol II must transcribe through a poly(A) signal previous to termination. One model suggests that elongation factors travel with the polymerase, dissociate at the poly(A) signal and leave the enzyme in a termination competent state. The second model assumes that the nascent RNA is cleaved at the poly(A) signal resulting in an unprotected 5'end. It is degraded by 5'-3' exonucleases which run into the ternary transcription complex like torpedoes ("torpedo model") thereby destabilizing it thus the enzyme dissociates.
Although no homologues of any known termination factors have been found in archaea so far, the sensitivity to the ρ-factor and the similarities of the archaeal and eukaryotic transcription machineries, lead to the suggestion that there are additional termination processes present in archaea.

It has been demonstrated that certain subunits of eukaryotic RNA polymerases are involved in transcriptional termination and recycling. The pol III subunit C11 has been shown to be required for transcription reinitiation (Landrieux *et al.*, 2006). In pol I subunit A12 is homologous to Rpc11 and it seems to act as termination factor (Prescott *et al.*, 2004). The pol II subunit B9 is also homologous to Rpc11 and is required to release the RNA polymerase from arrest sites (Awrey *et al.*, 1997). No sequence similarities to these subunits can be found in the archaeal RNAP subunits. But all three subunits resemble TFIIS in sequence. The archaeal elongation factor TFS is related in sequence to TFIIS (Hausner *et al.*, 2000) and therefore it was suspected to be involved in the termination process. To analyse this, purified recombinant TFS from *Pyrococcus furiosus* was added in increasing concentrations to transcription reactions at 70°C and 90°C similar to those performed in chapter IV, figure 1, but no effect on termination or reinitiation could be observed (data not shown).

The high reinitiation activity described for pol III may also be due to a complex protein-protein interaction between the terminating pol III and promoter bound transcription factors (Ferrari *et al.*, 2004). This led to the idea that also in the archaeal termination-reinitiation process the transcription factors TBP or TFB may play a special role. Therefore experiments were performed similar to those described in chapter IV Figure 4A. The RNAP was stalled at position +25 and washed with NLS (N-lauroylsarcosine) to remove all promoter bound

factors. To perform single round transcription (SR) only NTPs, but no transcription factors were added for continued elongation. Addition of TBP and TFB to stalled complexes along with NTPs, allowed multiple rounds of RNA synthesis (MR). To test whether a single transcription factor influences transcript formation, either TBP or TFB were added to stalled complexes (Figure 3, step 3).

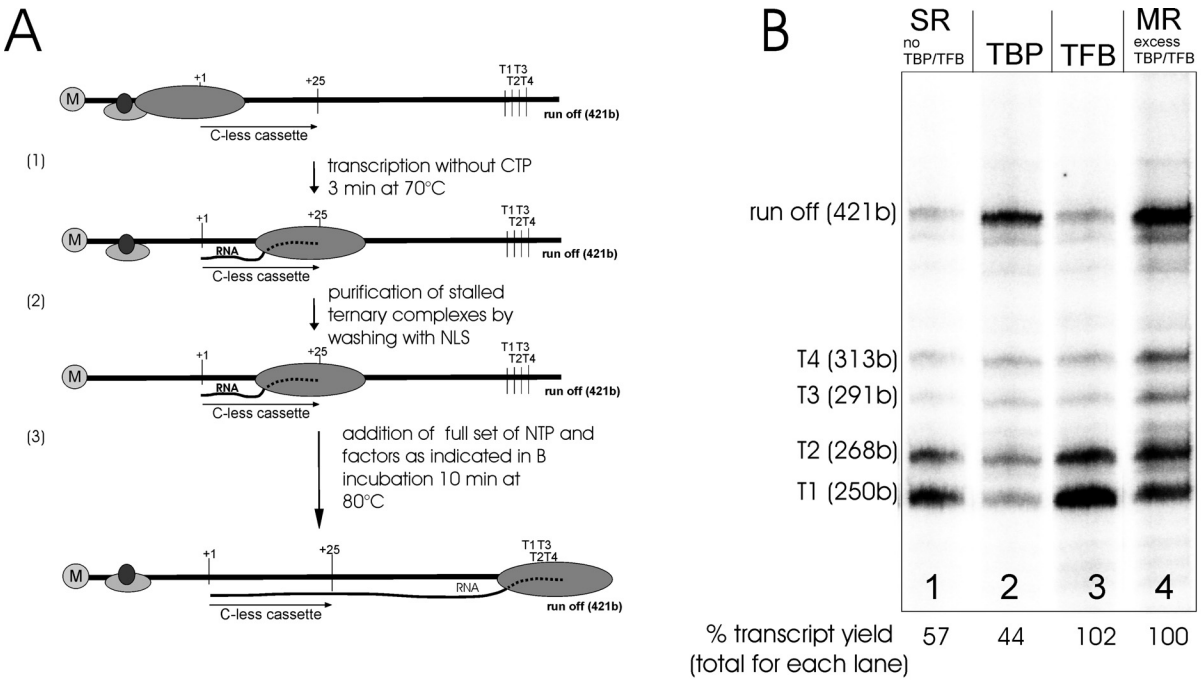

Figure 3. Is TFB involved in the reinitiation mechanism? **A.** The template *hpyA1* containing the termination region and a C-minus cassette was incubated in transcription reaction without CTP. Stable transcription complexes stalled at position +25 (1) were isolated by magnetic attraction, the supernatant was removed and the complexes were washed with 0.5% NLS to remove all promoter bound transcription factors (2). Then transcription buffer containing a full set of NTPs (440 µM each ATP, GTP, CTP, 2,7 µM UTP and 0,074 MBq [α-^{32}P]UTP (110 TBq/mmol)) and transcription factors as indicated in B, was added for continued elongation at 80°C (3). **B.** Single round transcription (SR) was performed without addition of transcription factors to the stalled complexes (lane 1). For multiple round transcription (MR) both TBP (119 nM) and TFB (147 nM) were added (lane 4). In lane 2 only TBP and in lane 3 only TFB was added to the stalled complexes. The lengths of the transcripts are indicated on the left. The total transcript yield compared to MR transcription is indicated below for each lane. The values are corrected for the varying amount of incorporated radioactivity depending on the length of each transcript.

First results showed that adding only TBP leads to transcriptional activity similar to that of single round transcription (figure 3B, see lanes 1 and 2), indicating that TBP alone does not allow reinitiation. In contrast, the experiments performed with only TFB showed a transcript level just as high as in multiple round transcription (figure 3B, see lanes 3 and 4). Additionally, termination efficiency at T1 was increased by incubation with TFB alone. These data have to be confirmed, but they suggest the following termination-reinitiation model

involving TFB. A high level of histone is required to maintain DNA stability especially at elevated temperatures. After transcription initiation the unusually AT-rich *hpyA1* promoter region might remain partly unwound due to thermodynamic instability of AT-rich sequences. After termination at the oligo-dT tract the RNAP possibly translocates beyond the termination region. The observation that TFB increases termination efficiency at the *hpyA1* terminator 1 (T1, figure 3, lane 3) suggests that TFB interacts with the RNAP paused in the termination process. It was shown that pausing at the sequences downstream of the terminator influences recycling efficiency (chapter IV, figure 6). The RNAP bound to downstream sequences might already have released RNA, making it distinguishable from paused transcription complexes and competent for interaction with TFB. A TFB-RNAP complex may be recruited to the promoter that could be still open at 90°C. TFB might contact the BRE contributing to the correct positioning around the transcription start site. For TFIIB it has been shown that it interacts with DNA independently of TBP (Lagrange *et al.*, 1998) so this can also be assumed for archaeal TFB.

A second copy of TFB exists in *Pyrococcus furiosus* (TFB2; Micorescu *et al.*, 2007) that lacks the B-finger motif. Recent analyses of TFB2 from *P. furiosus* showed that binding of TFB2 to the RNAP obviously is not impared (Micorescu *et al.*, 2007). It is able to form preinitiation complexes as efficiently as TFB but it shows a defect in stabilizing an open complex (Micorescu *et al.*, 2007) which may be due to the missing B-finger. Therefore a physiological significance of TFB2 in standard transcription initiation seems unlikely. Based on the first observations that TFB may somehow be involved in the termination-reinitiation pathway, a function for TFB2 in the cycling process of RNA synthesis under certain conditions could be suspected. To support a high histone expression *hpyA1* is under control of a heat shock-specific promoter (Gelfand *et al.*, 2000). Also the TFB2 level increases under heat shock conditions (Shockley *et al.*, 2003), leading to excess TFB2 relative to preinitiation complexes. Thus TFB2 might interact with the terminated RNAP, pausing downstream of the termination sequence and possibly guide it back to the promoter region. Under the special conditions of a promoter remaining unwound after the first round of transcription, the B-finger is not required to stabilize the premelted bubble. Normally the B-finger is inserted deeply into the active center of the RNAP (Bushnell *et al.*, 2004) thus slowing down promoter escape. Interaction with TFB2 lacking the B-finger, enables the RNAP to enter the elongation state more quickly and that again might contribute to a higher transcription efficiency. Therefore TFB2 may support an elevated histone expression under high temperature conditions and provide an additional regulatory element in gene expression.

The present work contributes to a deeper understanding of gene function by elucidating some essential mechanistic aspects of gene transcription. It created the base for further investigations especially concerning the termination of transcription that still remains obscure though being an important phase in gene expression with a high regulatory potential.

VI Summary

Novel mechanistic properties of the archaeal transcription system that reveal similarities especially to the eukaryotic RNA polymerase II and III systems were presented in thesis. The *in vitro* transcription system from the hyperthermophilic archaeon *Pyrococcus furiosus* provided the basis for the studies of the archaeal transcription cycle in the present work.

To obtain a detailed view on the dimensional parameters of the archaeal RNA polymerase during the transition from initiation to elongation, transcription complexes were paused in several registers from +5 to +20 by the use of C-minus cassettes. Exonuclease III footprinting analyses showed the RNA polymerase in close contact to the transcription factors TBP and TFB until a first structural transition in registers +6/+7 leads to a detectable upstream end of the RNAP. A second structural transition, observed in registers +10/+11, is characterized by bubble reclosure in the upstream part of the initially melted region and the first movement of the RNA polymerase downstream edge. RNA synthesis proceedes synchronously in early elongation between registers +11 and +20. The size of the transcription bubble in early elongation complexes is around 16 nucleotides and the RNA-DNA hybrid is about 9 bp in length. The RNA polymerase covers 26-29 bp of DNA and the distance of the catalytic center to the front edge of the RNA polymerase is approximately 12 bp.

Based on structural data on the eukaryotic RNA polymerase II, a structure-function analysis of *Pyrococcus furiosus* RNA polymerase mutants was performed. It elucidated the influence of structural polymerase elements on different stages of the transcription cycle. Recombinant archaeal RNA polymerases each carrying a deletion of one of four prominent cleft loops, named lid, rudder, fork1 and fork2, and three other RNAPs with point mutations were analysed. The Δrudder enzyme was demonstrated to be defective in open complex formation indicating an important role of the rudder loop in strand separation and/or maintainence of the transcription bubble. Experiments with templates containing a mismatch bubble resulted in predominant transcripts of 11 and 12 nucleotides. It could be demonstrated that a minimal length of 27 nt downstream of the bubble is required for promoter escape and transcription past the barrier.

The complete archaeal histone gene from *Pyrococcus furiosus* with its adjacent four consecutive oligo-dT streches was used as a model system to address the question of termination mechanisms in hyperthermophilic Archaea. At 90°C the archaeal RNA polymerase terminated with high efficiency at the first oligo-dT sequence when it contains at least 5 T residues. Possible hairpin formation has no influence on termination efficency.

Template competition experiments revealed the existence of a pol III-like termination-dependent reinitiation mechanism in the archaeal transcription system. Mutations of the sequences immediately downstream of the first termination signal dramatically affect the reinitiation mechanism.

VII Zusammenfassung

Die vorliegende Arbeit führt zu neuen Einblicken in den Mechanismus des archaeellen Transkriptionssystems und offenbart Ähnlichkeiten insbesondere zu den eukaryotischen RNA Polymerase II und III Systemen. Das *in vitro* Transkriptionssystem des hyperthermophilen Archaeons *Pyrococcus furiosus* bildete die Grundlage für die Untersuchung des archaeellen Transkriptionszyklus in dieser Arbeit.

Um ein detailiertes Bild von den strukturellen Eckdaten einer archaeellen RNA Polymerase während des Übergangs von der Initiation zur Elongation zu erhalten, wurden Transkripitonskomplexe in verschieden Registern zwischen +5 und +20 mit Hilfe von C-minus Kassetten pausiert. Die Analyse der pausierten Komplexe mit Exonuklease III zeigte, dass die RNA Polymerase in engem Kontakt zu den Transkriptionsfaktoren TBP und TFB steht, bis ein erster struktureller Übergang in den Registern +6/+7 ein Ablösen der RNAP von den Faktoren erkennen lässt. Das Schließen der Transkriptionsblase im stromaufwärts gelegenen Bereich der initial geöffneten Region und eine erste Vorwärtsbewegung des vorderen Endes der RNA Polymerase charakterisieren einen zweiten strukturellen Übergang in den Registern +10/+11. Zwischen den Registern +11 und +20 der frühen Elongationsphase läuft die RNA Synthese synchron. Im frühen Elongationskomplex hat die Transkriptionsblase eine Ausdehnung von etwa 16 Nukleotiden und das RNA-DNA Hybrid ist ungefähr 9 bp lang. Die RNA Polymerase Bindestelle umfasst 26-29 bp der DNA und der Abstand vom aktiven Zentrum zum stromabwärts gelegenen Ende der RNA Polymerase beträgt etwa 12 bp.

Eine auf Strukturdaten der eukaryotischen Polymerase II beruhende Struktur-Funktions-Analyse von *Pyrococcus furiosus* RNA Polymerase Mutanten beleuchtete den Einfluß struktureller Elemente der RNA Polymerase auf verschiedene Phasen im Transkriptions-zyklus. Rekombinante archaeelle RNA Polymerasen, von denen jede eine Deletion einer der vier aus der RNA Polymerase Spalte hervorstehende Schleifen, lid, rudder, fork1 und fork2, trug sowie drei andere RNA Polymerasen mit Punktmutationen wurden untersucht. Es konnte gezeigt werden, dass das Δrudder Enzym keinen offenen Komplex ausbilden konnte, was auf eine wichtige Rolle dieser Struktur in Strangtrennung und/oder Erhaltung der Transkriptions-blase hindeutet. Experimente mit einer Matrize, die einen vorgeöffneten Promotor enthält, brachten hauptsächlich vorzeitig abbrechende Transkripte von 11 und 12 Nukleotiden Länge hervor. Es konnte gezeigt werden, dass eine minimale Länge von 27 bp der DNA stromabwärts des Heteroduplexbereichs benötigt wird, damit die RNA Polymerase diese Barriere überwinden und Transkripte mit voller Länge synthetisieren kann.

Das komplette archaeelle Histon-Gen *hpyA1* von *Pyrococcus furiosus*, mit seinen nachfolgenden vier oligo-dT Sequenzen direkt stromabwärts des Gens, wurde als Modellsystem verwendet, um den Mechanismus der transkriptionellen Termination bei hyperthermophilen Archaeen zu untersuchen. Bei 90°C terminiert die archaeelle RNA Polymerase mit hoher Effizienz an der ersten oligo-dT-Sequenz wenn sie mindestens 5 T Reste enthält. Die mögliche Ausbildung einer Haarnadelstruktur hat keinen Einfluß auf die Terminationseffizienz. Die Existenz eines RNA polymerase III-ähnlichen terminations-abhängigen Reinitiations-mechanismus im archaeellen Transkriptionssystem konnte durch Kompetitionsexperimente mit pausierten Komplexen aufgedeckt werden. Mutationen der Sequenzen direkt stromabwärts der Terminationssignale haben einen entscheidenden Einfluß auf den postulierten Reinitiationsmechanismus.

VIII References

Aravind, L., and Koonin, E.V. (1999) DNA-binding proteins and evolution of transcription regulation in the archaea. *Nucleic Acids Res* **27**: 4658-4670.

Awrey, D.E., Weilbaecher, R.G., Hemming, S.A., Orlicky, S.M., Kane, C.M., and Edwards, A.M. (1997) Transcription elongation through DNA arrest sites. A multistep process involving both RNA polymerase II subunit RPB9 and TFIIS. *J Biol Chem* **272**: 14747-14754.

Bar-Nahum, G., and Nudler, E. (2001) Isolation and characterization of sigma(70)-retaining transcription elongation complexes from Escherichia coli. *Cell* **106**: 443-451.

Bartlett, M.S., Thomm, M., and Geiduschek, E.P. (2004) Topography of the euryarchaeal transcription initiation complex. *J Biol Chem* **279**: 5894-5903.

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. (1998a) Transcription in Archaea. *Cold Spring Harb Symp Quant Biol* **63**: 41-51.

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

Bell, S.D., Jaxel, C., Nadal, M., Kosa, P.F., and Jackson, S.P. (1998) Temperature, template topology, and factor requirements of archaeal transcription. *Proc Natl Acad Sci U S A* **95**: 15218-15222.

Bell, S.D., Kosa, P.L., Sigler, P.B., and 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., and Jackson, S.P. (2001) The archaeal TFIIEalpha homologue facilitates transcription initiation by enhancing TATA-box recognition. *EMBO Rep* **2**: 133-138.

Bieker, J.J., Martin, P.L., and Roeder, R.G. (1985) Formation of a rate-limiting intermediate in 5S RNA gene transcription. *Cell* **40**: 119-127.

Bogenhagen, D.F., and Brown, D.D. (1981) Nucleotide sequences in Xenopus 5S DNA required for transcription termination. *Cell* **24**: 261-270.

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**: 6969-6973.

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

Cramer, P., Bushnell, D.A., and Kornberg, R.D. (2001) Structural basis of transcription: RNA polymerase II at 2.8 angstrom resolution. *Science* **292**: 1863-1876.

Dedrick, R.L., Kane, C.M., and Chamberlin, M.J. (1987) Purified RNA polymerase II recognizes specific termination sites during transcription in vitro. *J Biol Chem* **262**: 9098-9108.

Dieci, G., and Sentenac, A. (1996) Facilitated recycling pathway for RNA polymerase III. *Cell* **84**: 245-252.

Dieci, G., and Sentenac, A. (2003) Detours and shortcuts to transcription reinitiation. *Trends Biochem Sci* **28**: 202-209.

Dvir, A., Tan, S., Conaway, J.W., and Conaway, R.C. (1997) Promoter escape by RNA polymerase II. Formation of an escape-competent transcriptional intermediate is a prerequisite for exit of polymerase from the promoter. *J Biol Chem* **272**: 28175-28178.

Dvir, A. (2002) Promoter escape by RNA polymerase II. *Biochim Biophys Acta* **1577**: 208-223.

Edgell, D.R., and Doolittle, W.F. (1997) Archaea and the origin(s) of DNA replication proteins. *Cell* **89**: 995-998.

Ferrari, R., Rivetti, C., Acker, J., and Dieci, G. (2004) Distinct roles of transcription factors TFIIIB and TFIIIC in RNA polymerase III transcription reinitiation. *Proc Natl Acad Sci U S A* **101**: 13442-13447.

Fiedler, U., and Timmers, H.T. (2001) Analysis of the open region of RNA polymerase II transcription complexes in the early phase of elongation. *Nucleic Acids Res* **29**: 2706-2714.

Fish, R.N., and Kane, C.M. (2002) Promoting elongation with transcript cleavage stimulatory factors. *Biochim Biophys Acta* **1577**: 287-307.

Geiduschek, E.P., and Ouhammouch, M. (2005) Archaeal transcription and its regulators. *Mol Microbiol* **56**: 1397-1407.

Gelfand, M.S., Koonin, E.V., and Mironov, A.A. (2000) Prediction of transcription regulatory sites in Archaea by a comparative genomic approach. *Nucleic Acids Res* **28**: 695-705.

Gnatt, A.L., Cramer, P., Fu, J., Bushnell, D.A., and Kornberg, R.D. (2001) Structural basis of transcription: an RNA polymerase II elongation complex at 3.3 A resolution. *Science* **292**: 1876-1882.

Grunberg, S., Bartlett, M.S., Naji, S., and Thomm, M. (2007) Transcription factor E is a part of transcription elongation complexes. *J Biol Chem*.

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

Hain, J., Reiter, W.D., Hudepohl, U., and Zillig, W. (1992) Elements of an archaeal promoter defined by mutational analysis. *Nucleic Acids Res* **20**: 5423-5428.

Hanzelka, B.L., Darcy, T.J., and 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., Frey, G., and Thomm, M. (1991) Control regions of an archaeal gene. A TATA box and an initiator element promote cell-free transcription of the tRNA(Val) gene of Methanococcus vannielii. *J Mol Biol* **222**: 495-508.

Hausner, W., Wettach, J., Hethke, C., and 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., and Musfeldt, M. (2000) Transcription factor S, a cleavage induction factor of the archaeal RNA polymerase. *J Biol Chem* **275**: 12393-12399.

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

Holstege, F.C., Fiedler, U., and Timmers, H.T. (1997) Three transitions in the RNA polymerase II transcription complex during initiation. *Embo J* **16**: 7468-7480.

Jansa, P., Burek, C., Sander, E.E., and Grummt, I. (2001) The transcript release factor PTRF augments ribosomal gene transcription by facilitating reinitiation of RNA polymerase I. *Nucleic Acids Res* **29**: 423-429.

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

Kassavetis, G.A., Blanco, J.A., Johnson, T.E., and Geiduschek, E.P. (1992) Formation of open and elongating transcription complexes by RNA polymerase III. *J Mol Biol* **226**: 47-58.

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

Kireeva, M.L., Komissarova, N., Waugh, D.S., and Kashlev, M. (2000) The 8-nucleotide-long RNA:DNA hybrid is a primary stability determinant of the RNA polymerase II elongation complex. *J Biol Chem* **275**: 6530-6536.

Komissarova, N., Becker, J., Solter, S., Kireeva, M., and Kashlev, M. (2002) Shortening of RNA:DNA hybrid in the elongation complex of RNA polymerase is a prerequisite for transcription termination. *Mol Cell* **10**: 1151-1162.

Kornblihtt, A.R. (2004) Shortcuts to the end. *Nat Struct Mol Biol* **11**: 1156-1157.

Korzheva, N., Mustaev, A., Kozlov, M., Malhotra, A., Nikiforov, V., Goldfarb, A., and Darst, S.A. (2000) A structural model of transcription elongation. *Science* **289**: 619-625.

Kosa, P.F., Ghosh, G., DeDecker, B.S., and 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.

Kugel, J.F., and Goodrich, J.A. (2002) Translocation after synthesis of a four-nucleotide RNA commits RNA polymerase II to promoter escape. *Mol Cell Biol* **22**: 762-773.

Kuznedelov, K., Korzheva, N., Mustaev, A., and 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**: 1369-1378.

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

Lagrange, T., Kapanidis, A.N., Tang, H., Reinberg, D., and 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.

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

Lang, W.H., and Reeder, R.H. (1995) Transcription termination of RNA polymerase I due to a T-rich element interacting with Reb1p. *Proc Natl Acad Sci U S A* **92**: 9781-9785.

Lang, W.H., Platt, T., and Reeder, R.H. (1998) Escherichia coli rho factor induces release of yeast RNA polymerase II but not polymerase I or III. *Proc Natl Acad Sci U S A* **95**: 4900-4905.

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**: 1133-1143.

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

Linn, S.C., and Luse, D.S. (1991) RNA polymerase II elongation complexes paused after the synthesis of 15- or 35-base transcripts have different structures. *Mol Cell Biol* **11**: 1508-1522.

Littlefield, O., Korkhin, Y., and Sigler, P.B. (1999) The structural basis for the oriented assembly of a TBP/TFB/promoter complex. *Proc Natl Acad Sci U S A* **96**: 13668-13673.

Magill, C.P., Jackson, S.P., and Bell, S.D. (2001) Identification of a conserved archaeal RNA polymerase subunit contacted by the basal transcription factor TFB. *J Biol Chem* **276**: 46693-46696.

Matsuzaki, H., Kassavetis, G.A., and Geiduschek, E.P. (1994) Analysis of RNA chain elongation and termination by Saccharomyces cerevisiae RNA polymerase III. *J Mol Biol* **235**: 1173-1192.

Metzger, W., Schickor, P., and Heumann, H. (1989) A cinematographic view of Escherichia coli RNA polymerase translocation. *Embo J* **8**: 2745-2754.

Metzger, W., Schickor, P., Meier, T., Werel, W., and Heumann, H. (1993) Nucleation of RNA chain formation by Escherichia coli DNA-dependent RNA polymerase. *J Mol Biol* **232**: 35-49.

Micorescu, M., Grunberg, S., Franke, A., Cramer, P., Thomm, M., and Bartlett, M. (2007) Archaeal transcription: function of an alternative transcription factor B from Pyrococcus furiosus. *J Bacteriol*.

Muller, B., Allmansberger, R., and Klein, A. (1985) Termination of a transcription unit comprising highly expressed genes in the archaebacterium Methanococcus voltae. *Nucleic Acids Res* **13**: 6439-6445.

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

Naji, S., Bertero, M.G., Spitalny, P., Cramer, P., and Thomm, M. (2007) Structure function analysis of the RNA polymerase cleft loops elucidates initial transcription, DNA unwinding and RNA displacement. *Nucleic Acids Res*.

Ohkuma, Y., Hashimoto, S., Wang, C.K., Horikoshi, M., and Roeder, R.G. (1995) Analysis of the role of TFIIE in basal transcription and TFIIH-mediated carboxy-terminal domain phosphorylation through structure-function studies of TFIIE-alpha. *Mol Cell Biol* **15**: 4856-4866.

Pal, M., Ponticelli, A.S., and Luse, D.S. (2005) The role of the transcription bubble and TFIIB in promoter clearance by RNA polymerase II. *Mol Cell* **19**: 101-110.

Palmer, J.R., and Daniels, C.J. (1995) In vivo definition of an archaeal promoter. *J Bacteriol* **177**: 1844-1849.

Platt, T. (1986) Transcription termination and the regulation of gene expression. *Annu Rev Biochem* **55**: 339-372.

Prescott, E.M., Osheim, Y.N., Jones, H.S., Alen, C.M., Roan, J.G., Reeder, R.H., Beyer, A.L., and Proudfoot, N.J. (2004) Transcriptional termination by RNA polymerase I requires the small subunit Rpa12p. *Proc Natl Acad Sci U S A* **101**: 6068-6073.

Qureshi, S.A., Bell, S.D., and Jackson, S.P. (1997) Factor requirements for transcription in the Archaeon Sulfolobus shibatae. *Embo J* **16**: 2927-2936.

Reiter, W.D., Palm, P., and Zillig, W. (1988) Transcription termination in the archaebacterium Sulfolobus: signal structures and linkage to transcription initiation. *Nucleic Acids Res* **16**: 2445-2459.

Reiter, W.D., Hudepohl, U., and Zillig, W. (1990) Mutational analysis of an archaebacterial promoter: essential role of a TATA box for transcription efficiency and start-site selection in vitro. *Proc Natl Acad Sci U S A* **87**: 9509-9513.

Renfrow, M.B., Naryshkin, N., Lewis, L.M., Chen, H.T., Ebright, R.H., and Scott, R.A. (2004) Transcription factor B contacts promoter DNA near the transcription start site of the archaeal transcription initiation complex. *J Biol Chem* **279**: 2825-2831.

Richardson, J.P. (2003) Loading Rho to terminate transcription. *Cell* **114**: 157-159.

Ryder, A.M., and Roberts, J.W. (2003) Role of the non-template strand of the elongation bubble in intrinsic transcription termination. *J Mol Biol* **334**: 205-213.

Samkurashvili, I., and Luse, D.S. (1998) Structural changes in the RNA polymerase II transcription complex during transition from initiation to elongation. *Mol Cell Biol* **18**: 5343-5354.

Santangelo, T.J., and Roberts, J.W. (2004) Forward translocation is the natural pathway of RNA release at an intrinsic terminator. *Mol Cell* **14**: 117-126.

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**: 196-210.

Schickor, P., Metzger, W., Werel, W., Lederer, H., and Heumann, H. (1990) Topography of intermediates in transcription initiation of E.coli. *Embo J* **9**: 2215-2220.

Shockley, K.R., Ward, D.E., Chhabra, S.R., Conners, S.B., Montero, C.I., and Kelly, R.M. (2003) Heat shock response by the hyperthermophilic archaeon Pyrococcus furiosus. *Appl Environ Microbiol* **69**: 2365-2371.

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

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

Spitalny, P., and Thomm, M. (2007) A polymerase III-like reinitiation mechanism is operating in regulation of histone expression in Archaea. *Mol Microbiol*.

Thomm, M., and Wich, G. (1988) An archaebacterial promoter element for stable RNA genes with homology to the TATA box of higher eukaryotes. *Nucleic Acids Res* **16**: 151-163.

Thomm, M. (1996) Archaeal transcription factors and their role in transcription initiation. *FEMS Microbiol Rev* **18**: 159-171.

Tintut, Y., Wang, J.T., and Gralla, J.D. (1995) A novel bacterial transcription cycle involving sigma 54. *Genes Dev* **9**: 2305-2313.

Toulokhonov, I., and Landick, R. (2006) The role of the lid element in transcription by E. coli RNA polymerase. *J Mol Biol* **361**: 644-658.

Trinh, V., Langelier, M.F., Archambault, J., and Coulombe, B. (2006) Structural perspective on mutations affecting the function of multisubunit RNA polymerases. *Microbiol Mol Biol Rev* **70**: 12-36.

Vassylyev, D.G., Vassylyeva, M.N., Perederina, A., Tahirov, T.H., and Artsimovitch, I. (2007) Structural basis for transcription elongation by bacterial RNA polymerase. *Nature* **448**: 157-162.

Werner, F., and Weinzierl, R.O. (2005) Direct modulation of RNA polymerase core functions by basal transcription factors. *Mol Cell Biol* **25**: 8344-8355.

Werner, F., Wiesler, S., Nottebaum, S., and Weinzierl, R.O. (2006) Modulation of RNA polymerase core functions by basal transcription factor TFB/TFIIB. *Biochem Soc Symp*: 49-58.

Westover, K.D., Bushnell, D.A., and Kornberg, R.D. (2004) Structural basis of transcription: separation of RNA from DNA by RNA polymerase II. *Science* **303**: 1014-1016.

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

Yarnell, W.S., and Roberts, J.W. (1999) Mechanism of intrinsic transcription termination and antitermination. *Science* **284**: 611-615.

Yudkovsky, N., Ranish, J.A., and Hahn, S. (2000) A transcription reinitiation intermediate that is stabilized by activator. *Nature* **408**: 225-229.

Zaychikov, E., Denissova, L., and Heumann, H. (1995) Translocation of the Escherichia coli transcription complex observed in the registers 11 to 20: "jumping" of RNA polymerase and asymmetric expansion and contraction of the "transcription bubble". *Proc Natl Acad Sci U S A* **92**: 1739-1743.

IX Appendix

IX.1 Danksagung

An erster Stelle gebührt mein herzlicher Dank Herrn Prof. Dr. M. Thomm, der mir die Durchführung dieser interessanten Doktorarbeit über einen langen Zeitraum ermöglicht hat. Unter seiner wissenschaftlichen Anleitung durfte ich genügend Freiraum für fachliche Entfaltung und kreatives Arbeiten genießen. Sein stetiges Interesse und seine große Unterstützung haben sehr zum Gelingen dieser Arbeit beigetragen.

Bei PD Dr. Winfried Hausner möchte ich mich für die vielen interessanten Diskussionen bedanken, zu denen er sich immer Zeit genommen hat und die mir viele Anregungen geboten haben.

Ein lieber Dank geht auch an Gudrun Vierke, die seit der Diplomarbeit an meiner Seite war und immer ein offenes Ohr für wissenschaftliche wie auch private Dinge hatte. Für die gute Arbeitsatmosphäre und die daraus erwachsene Freundschaft bedanke ich mich.

Sebastian Grünberg, der mich über die lange Zeit meiner Doktorarbeit in verschieden „Funktionen“ begleitet hat, danke ich vor allem für die vielen anregenden Gespräche zu wissenschaftlichen und nicht ganz so wissenschaftlichen Themen.

Für eine entspannte und heitere Atmosphäre im Arbeitsalltag bedanke ich mich bei meinen Labormitbewohnerinnen Mirijam Zeller und Christine Richter.

Auch den vielen nicht namentlich genannten Personen, die mich in Kiel und Regensburg während meiner Doktorarbeit begleitet haben, danke ich für eine unterhaltsame und fröhliche Zeit.

Nicht zuletzt danke ich meiner Familie, auf deren Unterstützung ich mich jederzeit verlassen konnte.

IX.2 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.

Die Ergebnisse dieser Arbeit sind bereits veröffentlicht worden:

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

2. Naji, S., Bertero, M. G., **Spitalny**,**P**, Cramer, P., and Thomm, M. (2007) Structure-function analysis of the RNA polymerase cleft loops elucidates initial transcription, DNA unwinding, and RNA displacement. Nucleic Acids Res. doi:10.1093/nar/gkm 1086 (in press).

3. **Spitalny, P**., and Thomm, M. (2007) A polymerase III-like reinitiation mechanism is operating in regulation of histone expression in archaea. Mol. Microbiol. doi: 10.1111/j.1365-2958.2007.06084.x (in press)

Veröffentlichung 2 entstand in Kooperation mit Prof. Dr. Patrick Cramer, Genzentrum der Ludwig-Maximilians-Universität München. Mein Anteil an dieser Veröffentlichung umfasst die Analyse der Ausbildung von offenen Transkriptionskomplexen durch die Mutanten der RNA Polymerase sowie Transkriptionsexperimente mit Heteroduplexmatrizen.

Das für Veröffentlichung 1 beschriebene immobilisierte Exonuclease III-Footprinting wird 2008 in der 3. Edition von DNA-Protein Interactions: Principles and Protocols, Humana Press, Totowa, New Jersey (Herausgeber: Benoît Leblanc und Tom Moss) als folgendes Kapitel erscheinen:
Spitalny, P. and Thomm, M.: Exonuclease III Footprinting on immobilized DNA Templates

Regensburg, Januar 2008