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

Received for publication, April 8, 2003, and in revised for, May 26, 2003
Published, JBC Papers in Press, June 3, 2003, DOI 10.1074/jbc.M303633200

Patrizia Spitalny† and Michael Thomm‡§

From the Universität Kiel, Institut für Allgemeine Mikrobiologie, Am Botanischen Garten 1-9, D-24118 Kiel, Germany

The archaeal transcriptional machinery is polymerase II (pol II)-like but does not require ATP or TFIIH for open complex formation. We have used enzymatic and chemical probes to follow the movement of Pyrococcus RNA polymerase (RNAP) along the 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 TATA box-binding protein/transcription factor B in complexes stalled at positions +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 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 ~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 pol II.

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 2 decades in bacterial RNA polymerase and eukaryotic polymerase II (for reviews see Refs. 1 and 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 tRNAVal (6) and at the 16 S 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 (β′αβ′γδσ) occurs spontaneously in a temperature-dependent manner (8, 9). By contrast, nuclear RNA polymerase II (pol II; see Ref. 10) and Escherichia coli RNA polymerase specific for promoters of genes involved in nitrogen metabolism (β′αβ′γδσ2; see Ref. 11) require ATP hydrolysis for promoter melting. In the pol II system promoter opening involves the helicase activity of TFIIH (12). Eukaryotic nuclear RNA polymerases I and III share with the σ70 containing E. coli RNA polymerase the ability to produce an open complex of 12–15 bp without ATP hydrolysis. In the pol II system the presence of the general transcription factor TFIIIB is required in addition to open complex formation (3).

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 pol II system. Pol II complexes were subjected to numerous structural alterations during formation of the first 30 phosphodiester bonds (14–16). 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 co-workers (20) provided evidence that the strength of the RNA-DNA hybrid is essential for maintaining stability of transcription complexes by preventing backtracking of RNA polymerase. 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 pol II machinery (21, 22). The archaeal TATA box is recognized by an archaeal TATA box-binding protein (TBP). This interaction is stabilized by transcription factor B (TFB), a homologue of general pol II transcription factor TFIIB. This TBP-TFB promoter complex recruits the archaeal RNA polymerase that shows striking similarity in sequence and subunit composition to pol II. With the exception of TFE, which is homologous to the α-subunit of pol II transcription factor TFIIE (23, 24), no other homologues of the basal eukaryotic transcriptional machinery were detected in archaeal genomes.
Stalled Archaeal Transcription Complexes

Consistent with the lack of TFIIH in Archaea and in contrast to the striking general similarity to the pol II system, the archaeal RNA polymerase does not require hydrolysis of ATP for open complex formation at the tRNA^Lys^ promoters of *Methanococcus vannielii* (6). We have developed recently a cell-free transcription system for the hyperthermophile *Pyrococcus furiosus* (28). This highly purified system consisting of bacteriophylylthylated TBP 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 (KMnO₄) footprinting provided a detailed view of the early steps of transcription in Archaea.

**EXPERIMENTAL PROCEDURES**

**Reagents and Enzymes**—Exonuclease III was purchased from New England Biolabs. Potassium permanganate was obtained from Merck. [α-³²P]UTP and [γ-³²P]ATP were purchased from Hartmann Bioanalytik, 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 *P. furiosus* was used to generate transcription templates by amplification (restriction sites of the vector pUC19. The resulting fragments were inserted between the EcoRI and SmalI (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 of 263–278 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 labeled with biotin, and the resulting fragments were attached to streptavidin magnetic beads (Roche Applied Science) according to the manufacturer’s protocol.

**Purification of Pyrococcus RNA Polymerase—**RNA polymerase from *P. furiosus* was purified as described previously (26). Expression and Purification of Recombinant Transcription Factors—The transcription factor TBP from *Pyrococcus woesei* was overproduced in *E. coli* as described previously (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).

**Im mobilized in Vitro Transcription Assays—**In vitro transcription assays were performed according to Ref. 26. A standard reaction mixture (25 μl) contained 60 ng of immobilized template, 660 ng of RNA polymerase, 150 ng of recombinant TBP, 120 ng of recombinant α-³²P ATP and GTP, 2 μM UTP, and 0.15 MBq [α-³²P]UTP (110 TBq/mmol). The transcription reaction buffer contained 40 mM HEPES, 0.1 mM EDTA, 1 mM dithiothreitol, 300 mM KCl, and 4 mM MgCl₂. Transcription reactions were performed for 3 min at 70 °C.

To perform footprinting experiments, the immobilized templates were labeled with [γ-³²P]ATP. [α-³²P]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 by ethanol precipitation. The supernatant was 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 min 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 labeled with [γ-³²P]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-³²P]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 exon III digestion (40 mM KCl, 2 mM MgCl₂, 100 mM Tris-HCl, pH 8.5, and 1 mM dithiothreitol). After addition of 100 units of exon III, the samples were incubated at 37 °C for 15 min. The reaction was stopped by the addition of loading buffer, and the samples and sequencing reactions were loaded on a 6% denaturing sequencing gel.

**KMnO₄ Sensitivity Assay—**To perform KMnO₄ probing, the immobilized DNA templates were labeled with [γ-³²P]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-³²P]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 of transcription buffer, and 2.5 μl of potassium permanganate (250 mM) were added. The samples were incubated for 5 min at 45 °C. The reactions were stopped by the addition of 1.7 μl of 2-mercaptoethanol and 20 μl of stop mix (1.25% SDS and 125 mM EDTA). The supernatant was removed, and the modified immobilized DNA was resuspended in 18 μl of water, and piperidine was added to a total volume of 20 μl. The DNA was subjected to cleavage by piperidine for 30 min 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.

**RESULTS**

**Stalled Archaeal Transcription Complexes Contain a Homogeneous 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 5–11, 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 (see under “Experimental Procedures”). Both in bacterial and eukaryotic systems read-through of RNAP beyond the expected stall sites has been observed (31, 32). 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 labeled with [α-³²P]UTP. After short incubation times between 30 s and 3 min, an RNA product of 20 nt was synthesized as the predominant product (data not shown). After incubation times between 5 and 45 min, additional products of 21 and 32 nt probably caused by misincorporation 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 streptavidin-coated magnetic beads. The RNA con-
tained in these purified complexes was analyzed by PAGE. 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 I). When isolated complexes were incubated in transcription buffer supplemented with all nucleotides, labeled RNA associated with isolated complexes could no longer be detected (data not shown). This finding suggests that the nascent RNA molecules were retained in functional ternary complexes that 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 labeled 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 KMnO₄ 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 because they could be chased after addition of nucleotides (Fig. 2A, lanes 6, 8, 10, 12, 14, 16, and 18).

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-labeled with ³²P 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 labeled with ³²P, and the biotin label was attached to the 5'-end of the template DNA strand. Cell-free transcription reactions were conducted at 70 °C and the subsequent purification of transcription complexes at 20 °C. Because 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.

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 cor-

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 P. furiosus are depicted. The templates were constructed as described under “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.
complexes were chased by the addition of all NTPs (40 in the nascent RNA being an A instead of a G. Minor products in lanes and 15 indicated registers. The higher mobility of the 5-nt RNA products show the RNA products of the isolated complexes stalled in the 17 transcriptionally active.

Transcription experiments analyzed on a 28% polyacrylamide gel. Rated nucleotides and released RNA. Stalled complexes were washed in washing buffer to remove unincorpo-

rated nucleotides and released RNA. 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. 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.

**TABLE 1**

| 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  |

responds approximately to the downstream end of the TBP/TFB footprint identified in the *Methanococcus* and *Sulfolobus* system by DNase I protection analyses (6, 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 NLS to the complexes, the exo III stall site at position −19 to −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 pol II (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 positions +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 up-

stream edges of the TBP/TFB and TBP DNase I 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 (6, 33). 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 ex-

tension 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 (KMnO₄) as a probe specific for thymidine (T) residues in single-stranded DNA. To investigate the temperature dependence of open complex formation TBP/TFB was incubated indi-

vidually or in combination with RNAP with end-labeled 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 KMnO₄ treatment (Fig. 4, A and B, left panel, and Fig. 5, upper panel) at 70 °C. No KMnO₄ footprint was observed at 50 °C or when TBP/TFB alone was incubated at both temperatures with these templates (data not shown). 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). Additional T residues with increased sensitivity to KMnO₄ 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, transcrip-

tion reactions with the templates shown in Fig. 1 were conducted at 70 °C, the ternary complexes isolated at 20 °C, and the KMnO₄ 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 position +7 and +9 were modified. The T residue at position +4 showed no sensitivity to KMnO4 (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 KMnO4. 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 complexes stalled at positions +7, +8, and +9 showed very similar reactivity toward KMnO4 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 KMnO4 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 toward KMnO4 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. 4. Mapping of the transcription bubble in stalled archaeal transcription complexes. The single strand-specific reagent KMnO₄ was applied 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.

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 KMnO₄ 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

(Fig. 4B). T residues at positions +4 and +5 showed strong reactivity; the signal at −3 was drastically reduced in intensity, and 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 toward KMnO₄; 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, −16 bp in length. An estimate of the extension of the transcription bubble in each register of transcription is given in Fig. 5.

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 positions +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 KMnO₄ 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 to 12 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).
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 cannot be determined exactly (dotted line). For the same reason the size of the transcription bubble is a lower estimate in most cases.
paused at position +5. The existence of these complexes stalled close to the transcription start site complicated the interpretation of KMnO₄ 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 because addition of a complete set of NTPs resulted in elongation of these nascent RNAs in ternary complexes to run-off transcripts (Figs. 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 (6, 22, 27). 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 KMnO₄ footprinting (see Figs. 4B and 5; T residues labeled 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 −12 bp (Fig. 5). A somewhat longer but also constant distance has been found in active eukaryotic (13) and prokaryotic transcription complexes (25). Exo III borders very close to the site of NMP addition are characteristic for backtracking of RNAP and arrested complexes (13). 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 because 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 KMnO₄ 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 tRNAval promoter of Methanococcus and an RNA promoter of Sulfolobus which ranged from −11 to −1 and −12 to −1, respectively (6, 7). Up to register +6 no significant movement of the downstream edge of the bubble could be detected (Figs. 4 and 5). Because the three T residues at positions −7 to −9 showed reduced activity toward KMnO₄, 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 KMnO₄ 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 KMnO₄ 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. Because estimation of the bubble size by KMnO₄ 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 KMnO₄ signals in the region of the transcription start site which were not part of the moving transcription bubble. However, careful inspection of the KMnO₄ modification patterns allowed clear definition of the 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 KMnO₄ modification patterns on both DNA strands and the weak KMnO₄-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 2 nt in register +5, 3 in register +6, 4 in register +7, 5 in register +8, 6 in register +9, 7 in register +12, and 8 in register +11 (Fig. 5). When RNAP was stalled at position +15, the length of the RNA-DNA hybrid was at least 8 and at stall site +20 at least 9 nt. The finding that the T residue at position +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 RNApol—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 RNApol. 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 from −55 to +18 is protected in the open complex formed by E. coli RNApol (17, 34). 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 (34). The recognition domain of DNA bound by E. coli RNApol 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 σ and the complex extends at this register from −3 to +27. The archaeal RNApol has also initiated promoter cleavage at this register and extends over 28 bp from position −4 to +24. Thus, the overall dimensions of the archaeal and bacterial complex stalled at +11 are very similar. A further contraction of E. coli RNApol-binding site stalled at register +20 to 22 bp has been observed (34). By contrast, an RNApol-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 pol II stalled between registers +20 and +23 which extend over 31–35 bp (35).

The distance of the catalytic center C to the front edge F of the footprint is constant in active and not retracted bacterial and pol II 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 (25)) and pol II system (C-F = 18–20 (13)).

In all domains of life a characteristic mechanistic similarity is the transition around register +10. At this point all RNApol seems to reach the elongation-committed 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 elongated. This is a common property of the archaeal enzyme and pol II. By contrast E. coli RNApol 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 RNApol.

From analyses of translocation of the transcription bubble three characteristic transitions have been postulated in the pol II system (15, 16). 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 RNApol is able to catalyze DNA strand separation in the absence of TFIIH helicase activity and ATP (6) (Fig. 4). We have no evidence that the second transition in the pol II initiation complex at register +4 characterized by insensitivity of the complex to ATPγS (15) occurs also in the archaeal system. In the pol II system the region of the initially open complex redadopts 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 RNApol and pol II seem to start promoter clearance around register +10. In E. coli and pol II complexes continuous opening of the downstream part of the open region and discontinuous reclosure of the upstream part have been described (15, 18). The data shown here seem to indicate that failure to observe continuous reclosure of the upstream end may not be 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 end (Fig. 2A).

The size of the RNA-DNA hybrid is within 9–12 bp in a similar range as in eukaryotic and prokaryotic elongation complexes analyzed by comparable methods (5, 20, 36).

Acknowledgments—We thank E. Zaychikov and H. Heumann (Max-Planck-Institut für Biochemie, Martinsried) for advising S. Francois (University of Kiel) in KMnO4 footprinting at high temperatures. We also thank W. Hausner and U. Lange (University Kiel) for valuable advice and discussions.

REFERENCES

1. Heumann, H., Zaychikov, E., Denissenova, L., and Hermann, T. (1997) Nucleic Acids Mol. Biol. 11, 151–177
2. Fiedler, U., and Timmers, H. Th. M. (2003) BioEssays 25, 316–326
3. Kassaevetis, G. A., Blance, J. A., Johnsen, T. E., and Geiduschek, E. P. (1992) J. Mol. Biol. 226, 47–56
4. Kassaevetis, 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 Timmers, 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 Hipel, P. H., Yager, T. D., and Gill, S. C. (1992) in Transcription Regulation (McKnight, S., and Yamamoto, K., eds) pp. 179–201, Cold Spring Harbor Laboratory, NY
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. 35, 367–376
12. Timmers, H. Th. M. (1994) EMBO J. 13, 391–399
13. Samkurai-stevi, I., and Luse, D. S. (1998) Mol. Cell. Biol. 18, 5343–5354
14. Cai, H., and Luse, D. S. (1997) J. Biol. Chem. 272, 298–304
15. Helie, 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, E., and Chamberlin, M. J. (1992) J. Mol. Biol. 225, 239–250
18. Zaychikov, E., Denissenova, L., and Heumann, H. (1999) Proc. Natl. Acad. Sci. U. S. A. 92, 1739–1743
19. Komissarova, N., and Kashlev, M. (1997) J. Biol. Chem. 272, 15329–15338
20. Nudler, E., Mustaev, A., Lakhmanov, 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 Rep. 21, 133–138
24. Hanzeoka, B. L., Darcy, T., and Reeves, J. N. (2001) J. Bacteriol. 183, 1813–1818
25. Nudler, E., Goldfarb, A., and Kashlev, M. (1994) Science 265, 793–796
26. Hethke, C., Geerling, A. C., Hausner, w., de Vos, W. M., and Thomm, M. (1996) Nucleic Acids Res. 24, 2389–2376
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. Dähne, I., and Thomm, M. (2002) Nucleic Acids Res. 30, 701–710
30. Vieker, G., Engelmann, A., Hebbeln, C., and Thomm, M. (2003) J. Biol. Chem. 278, 172–186
31. Carpacewski, 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, 12984–12940
34. Metager, W., Schickor, P., and Heumann, H. (1989) EMBO J. 8, 2745–2754
35. Samkurai-stevi, I., and Luse, D. S. (1996) J. Biol. Chem. 271, 24946–24955
36. Sidenlenkov, J., Komissarova, N., and Kashlev, M. (1998) Mol. Cell 2, 55–64
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

J. Biol. Chem. 2003, 278:30497-30505.
doi: 10.1074/jbc.M303633200 originally published online June 3, 2003

Access the most updated version of this article at doi: 10.1074/jbc.M303633200

Alerts:
- When this article is cited
- When a correction for this article is posted

Click here to choose from all of JBC's e-mail alerts

This article cites 35 references, 16 of which can be accessed free at http://www.jbc.org/content/278/33/30497.full.html#ref-list-1