***In vivo* evidence for back and forth oscillations of the transcription elongation complex**

**Francine Toulmé, Martine Guérin,
Nathalie Robichon, Marc Leng
and A.Rachid Rahmouni¹**

Centre de Biophysique Moléculaire, CNRS, rue Charles Sadron,
45071 Orléans cedex 2, France

¹Corresponding author
e-mail: rahmouni@cns-orleans.fr

We have used a combination of DNA and RNA footprinting experiments to analyze the structural rearrangements experienced by a transcription elongation complex that was halted *in vivo* by a protein readblock. We show that the complex readblocked within an (ATC/TAG)_n sequence is in a dynamic equilibrium between upstream- and downstream-translocated conformers. By increasing the strength of the putative RNA–DNA hybrid, the ternary complex is readily trapped in the downstream-translocated conformation, where the melted DNA region is limited to 8 bp. The shift of the equilibrium towards the downstream location is also achieved by introducing within the 5' end of the message an RNA sequence that can pair with a segment of the transcript in the vicinity of the halted ternary complex. Our results demonstrate that within certain template DNA sequences, the back and forth oscillations of the ternary complex actually occur in a multipolymerase system and inside the cell. Furthermore, the *cis*-acting effect of the upstream RNA sequence underscores an important phenomenon in gene regulation where a transcript may regulate its own elongation.

Keywords: *in situ* probing/lateral oscillations/RNA polymerase/transcription elongation complex

Introduction

The elongation of RNA chain during transcription takes place within the ternary complex that comprises the RNA polymerase with its accessory factors, the DNA template and the nascent transcript. These basic components of the ternary complex constitute the key intermediates in transcriptional regulation at the level of elongation in all living organisms (for reviews see Richardson and Greenblatt, 1996; Uptain *et al.*, 1997). Over the past few years, a variety of enzymatic and chemical probes have been used to investigate the structure of the ternary complex. Most of the work was performed *in vitro* with *Escherichia coli* RNA polymerase artificially halted by NTP withholding. The analyses of a large number of ternary complexes halted at successive positions on the template revealed two modes of translocation of the polymerase. At most template positions, the enzyme

translocates monotonously (step-by-step) in synchrony with the nucleotide addition cycle. However, at certain short DNA segments, such as those preceding pause or termination sites, a discontinuous (inchworm-like) translocation was observed (Krummel and Chamberlin, 1992; Lee and Landick, 1992; Nudler *et al.*, 1994, 1995; Wang *et al.*, 1995; Zaychikov *et al.*, 1995). The discontinuous movement is characterized by a substantial contraction of the RNA polymerase footprint on the DNA in such a way that the downstream edge of the polymerase appears to remain fixed for multiple rounds of nucleotide addition while the upstream edge slides in concert with RNA growth. The contraction of the footprint is also associated with a progressive conversion of the halted ternary complex into an arrested state from which resumption of RNA synthesis requires factor-induced (GreA or GreB in *E.coli* and SII in eukaryotes) endonucleolytic cleavage of the 3'-proximal part of the transcript (Lee *et al.*, 1994; Gu *et al.*, 1996).

These observations were first interpreted in terms of internal flexibility of RNA polymerase, where a mobile catalytic center is supposed to carry out multiple rounds of RNA synthesis within a fixed enzyme (Chan and Landick, 1994; Nudler *et al.*, 1994, 1995; Chamberlin, 1995). However, more recent re-evaluations of these results now suggest that the altered footprints observed at some template positions could be explained by a backward translocation of the whole RNA polymerase along the DNA and RNA chains (Reeder and Hawley, 1996; Komissarova and Kashlev, 1997a,b; Nudler *et al.*, 1997). In this view (Landick, 1997; Gelles and Landick, 1998; von Hippel, 1998; Nudler, 1999), the transcribing RNA polymerase invariably covers a DNA segment of ~35 bp, among which 12–16 bp are unwound forming the transcription bubble. Successive melting and reannealing of base pairs at either end of the bubble maintain the two DNA strands separated as the polymerase translocates monotonously along the template adding new nucleotides into the RNA. Within this melted region, the non-template strand is in a single-stranded configuration, whereas the template strand forms a constant 8–9 bp heteroduplex with the 3'-proximal part of the transcript. When the ternary complex is halted at positions where the RNA–DNA hybrid is weak, the polymerase slides backward to re-form a more stable heteroduplex instead of forming the next phosphodiester bond. This upstream movement shifts the transcription bubble, the RNA–DNA hybrid and the catalytic center away from the 3' end of the transcript which becomes extruded out of the polymerase and thus abnormally close to the front end of the enzyme's footprint on the DNA. The upstream shift of the polymerase relocates the catalytic center at an internal position within the transcript. Thus, unless the sliding process is reversed, endonucleolytic cleavage at this new position is necessary to re-align the 3' end of the RNA with the active center of the enzyme.

Most of the features of this new model have been drawn from analyses of arrested ternary complexes of both *E. coli* RNA polymerase and eukaryotic RNA polymerase II, where the loss of catalytic activity during arrest was shown to coincide with the appearance of upstream-translocated forms of the complex (Reeder and Hawley, 1996; Komissarova and Kashlev, 1997a; Nudler *et al.*, 1997; Samkurashvili and Luse, 1998). Further studies of such complexes suggested that the RNA–DNA hybrid plays a crucial role in the lateral stability of the ternary complex (Komissarova and Kashlev, 1997b, 1998; Nudler *et al.*, 1997; Sidorenkov *et al.*, 1998). These results have led to new ideas of transcript elongation in which actively transcribing RNA polymerase is viewed as an oscillating enzyme that has the option to move back or forth at each template position. The decision as to which of these two movements occurs is dictated by the relative stability of the DNA–DNA or RNA–DNA base pair that can be formed at the downstream branching point of the transcription bubble (Guajardo and Sousa, 1997; Landick, 1997; Gelles and Landick, 1998; von Hippel, 1998). In light of these new ideas, it is now widely believed that the lateral oscillations of RNA polymerase are at the basis of many mechanisms that regulate transcript elongation, including pausing, termination and proofreading. However, all the available experimental data have been obtained with isolated RNA polymerases transcribing short DNA fragments. Therefore, it is still unclear whether these dynamic rearrangements of the ternary complex could actually occur within the cellular context where the presence of accessory factors, physiological concentrations of NTPs, DNA supercoiling, and possibly the anchoring of the RNA polymerase and/or the DNA template to structural components of the cell (Rahmouni and Wells, 1992; Iborra *et al.*, 1996; Cook, 1999 and references therein), may interfere with such processes.

In a previous study (Guérin *et al.*, 1996) we performed the mapping of an *E. coli* transcription elongation complex that was halted *in vivo* by a protein readblock while transcribing a long trinucleotide repeating sequence (ATC/TAG)_n. In the present study, we used a combination of DNA and RNA footprinting experiments to analyze the structural rearrangements experienced by such a readblocked complex inside the cell. Our results provide evidence that the ternary complex halted within the (ATC/TAG)_n sequence is in a dynamic equilibrium between upstream- and downstream-translocated conformers. We show that the readblocked ternary complex can be readily stabilized in the downstream-translocated conformation either by increasing the strength of the putative RNA–DNA hybrid (changing the DNA sequence within the complex) or by introducing within the 5' part of the transcript an 'antisense' RNA sequence that can form base pairs with the nascent RNA in the vicinity of the complex.

Results

Rationale

We have previously reported the design of an *E. coli* system in which a transcription elongation complex is halted *in vivo* by a protein readblock for structural analyses by *in situ* probing experiments (Guérin *et al.*, 1996). In this approach, the RNA polymerase that initiates

Fig. 1. Schematic representation of the relevant parts of the plasmids used in this study. The promoter and the 22 bp *lac* operator motif are shown as boxes. The repeats are represented by the sequence of the non-template strands with the Cs at position -6 relative to the operator motif underlined. The restriction sites used for the mapping and sequencing analyses are indicated.

transcription from a constitutive promoter within a plasmid is readblocked at a downstream position by the *lac* repressor bound to its operator motif. The ternary complex was halted within a trinucleotide repeating sequence (ATC/TAG)_n where the structural status of each base residue could be analyzed by at least one of the following probes: osmium tetroxide (OsO₄), chloroacetaldehyde (CAA), dimethylsulfate (DMS) and UV light. These *in situ* analyses provided the picture of a ternary complex in which the DNA helix is melted over 12–16 bp with the catalytic center of the polymerase (last nucleotide added in the RNA) located at <6 bp from the repressor binding site. Furthermore, among other important structural features revealed by this study, a close analysis of the OsO₄ and CAA reactivity patterns of the two DNA strands led us to suggest that the ternary complex halted within the (ATC/TAG)_n sequence is probably in a dynamic equilibrium between multiple laterally translocated forms (Guérin *et al.*, 1996).

Here, we have explored this possibility by extending our *in situ* investigations with new plasmid constructs that were intended to stabilize the readblocked ternary complex in one conformation. The relevant parts of these plasmids are shown in Figure 1. As for pATC21, which was used in our previous work, all the constructs have in common the presence of an (ATC/TAG)_n sequence derivative inserted downstream from the HisR promoter (at position +40 relative to the transcription start site in the vector) in the orientation that gives rise to (AUC)_n-containing RNAs. The repeats are immediately flanked at their downstream side by a 22 bp *lac* operator sequence. For clarity, we will designate the base residues within the repeats by negative numbers that indicate their positions relative to the upstream edge of the operator motif.

A *cis*-acting 'antisense' RNA sequence stabilizes the halted ternary complex in the downstream-translocated conformation

Recent *in vitro* studies have shown that the pairing of antisense oligonucleotides to a segment of the transcript immediately behind the RNA polymerase prevents the backward translocation of the ternary complex and therefore its conversion into an arrested state (Reeder and Hawley, 1996; Komissarova and Kashlev, 1997a,b). To find out whether the ternary complex in our *in vivo*

Fig. 2. Primer extension analyses of the *in situ* CAA modifications on the non-template strands of pATC21 and pATC21H. The horizontal arrow indicates the direction of transcription from the promoter to the operator (Op). The densitometer scans of the minus IPTG lanes are shown in the bottom panel. The upstream and downstream margins (C-22 and C-6) of the relevant halted ternary complex are indicated by arrows. Note the presence of a second elongation complex proximal to the promoter, shown by a bracket. Owing to the detection method (primer extension), this complex is more clearly seen when analyzing the opposite DNA strand (see Figure 3 in Guérin *et al.*, 1996). The densitometer scans were obtained on a Molecular Dynamics PhosphorImager using IMAGEQUANT software Version 3.3 for data processing. This was also used for the images shown in Figure 4.

system can be stabilized in the downstream-translocated conformation by a similar mechanism, we conducted a comparative study between pATC21 and its derivative pATC21H. As depicted in Figure 1, pATC21H contains an additional (ATC/TAG)₂₀ repeat inserted in the opposite orientation at the *Hpa*I site (start site of transcription). Thus, the HisR-promoted transcript in pATC21H harbors an (AUG)₂₀ sequence at its 5' end that can possibly fold back and hybridize to a segment of the (AUC)_n RNA in the vicinity of the halted ternary complex (see Figure 7C).

As a first step to test this hypothesis, we used the single-strand-specific probe CAA to characterize the transcription bubble of the halted ternary complexes within the two plasmids. *Escherichia coli* cells bearing the plasmids and grown in the presence or the absence of the inducer isopropylthio β -D-galactopyranoside (IPTG) were treated with CAA as previously described (Guérin *et al.*, 1996). After plasmid DNA extraction, the modified bases on the non-template strand were revealed by their ability to terminate the extension of a ³²P end-labeled primer with Klenow fragment of DNA polymerase I. In the presence of IPTG, only background reactivities were detected, reflecting the overall transcriptional activity (Figure 2, top panel). In the absence of the inducer, both plasmids showed a cluster of reactive base residues (mostly Cs and As) located within the repeat immediately upstream from the operator motif (Figure 2, top panel). The modified bases mark the DNA region of the halted ternary complex that is accessible to the probe in an open configuration. In agreement with our previous estimation (Guérin *et al.*,

1996), in both plasmids the melted region involves 12–16 bp with the downstream margin of the bubble at position –6 relative to the operator motif. However, a more quantitative analysis of the results by phosphorimager scanning reveals a significant difference between the two reactivity patterns (Figure 2, bottom panel). For pATC21, as we previously noted, the CAA reactivities at the upstream part of the bubble and at the downstream margin are relatively low as compared with the internal positions. Obviously, this behavior is more strongly marked in pATC21H where a contraction of the overall reactivity pattern is observed (see the densitometer scans). In effect, the base residues at the upstream part of the bubble are even less reactive, whereas the reactivity of the internal positions as well as the downstream margin increases significantly. Although we have no direct proof that the 5' terminal (AUG)₂₀ RNA motif is actually hybridizing to its complementary sequence behind the polymerase, this shrinking of the CAA footprint in pATC21H strongly suggests that the *cis*-acting motif mediates a shift of the dynamic equilibrium towards the downstream-translocated form of the complex.

We next performed S1 nuclease protection experiments to map the 3' ends of the truncated transcripts resulting from the two halted ternary complexes. Cellular RNAs isolated from *E. coli* cells harboring the plasmids were hybridized to a 3' end-labeled single-stranded DNA probe and digested with S1 nuclease under appropriate conditions (see Materials and methods). The template strand of the *Bam*HI–*Pst*I fragment (Figure 1), which is the same for

Fig. 3. S1 mapping of the 3' ends of the short RNA transcripts produced *in vivo* from the plasmids pATC21 and pATC21H. Py and Pu denote the Maxam and Gilbert sequencing ladders obtained with the DNA probe (template strand of the *Bam*HI–*Pst*I fragment). Note that positions –6 and –9 shown by the arrows in the S1 mapping migrate at the level of their following Ts on the sequencing ladder (T-5 and T-8, respectively). This is because in the Maxam and Gilbert chemical cleavage, the T residues no longer exist within the DNA fragment. This is also the case for Figure 5A.

both plasmids, was used as the DNA probe. The data in Figure 3 clearly show that the RNAs extracted from cells grown in the absence of IPTG contain short transcripts that have been elongated up to a region between positions –6 and –9. Again, the S1 protection patterns reveal a remarkable difference between the RNA transcripts obtained with the two constructs. For pATC21, the cluster of RNA 3' ends is distributed in nearly equal amounts between positions –6 and –9. In contrast, most of the truncated transcripts produced from pATC21H have their 3' end at position –6. Thus, these results corroborate our conclusion with the CAA footprints. They imply that the halted ternary complex in pATC21 is in equilibrium between at least two laterally translocated forms and that in pATC21H the equilibrium is shifted towards the downstream location.

To characterize further the halted ternary complexes within the two plasmids, we analyzed the guanine residues within the template strand by DMS probing experiments. After growth in the presence or the absence of IPTG, the *E. coli* cells containing the plasmids were treated with DMS, and the DNA was extracted. Subsequently, the *Bam*HI–*Pst*I DNA fragments were selectively labeled at

the 3' end of the template strand and the modified bases revealed by hot piperidine cleavage. In agreement with our previous report (Guérin *et al.*, 1996), the results in Figure 4 show that in the absence of IPTG the G residue at position –6 is hypermethylated in both plasmids. The extent of hypermethylation, however, is drastically reduced in pATC21H (see densitometer scans). Hence, these DMS probing data are consistent with the experiments described above. They emphasize the structural differences between the ternary complexes within the two plasmids.

We have previously ascribed the hypermethylation of the G-6 residue to its presence in a highly exposed configuration within the catalytic center of the polymerase (Guérin *et al.*, 1996). Indeed, it seems conceivable that the template base at the catalytic center should be blocked in the orientation facing the NTP substrate for proper incorporation via Watson–Crick pairing. As will be discussed below, our present results extend this interpretation by revealing the temporal character of the process. That is, once the incoming CMP residue is added into the transcript, the template G residue is no longer blocked in the highly exposed configuration.

Increasing the strength of the putative RNA–DNA hybrid traps the halted ternary complex in the downstream-translocated conformation

To gather more evidence for the structural rearrangements in our *in vivo* system, we sought to control the lateral stability of the halted ternary complex by increasing the rC–dG content of the putative RNA–DNA hybrid. In order to introduce the sequence changes inside the repeat, we constructed plasmid series that have shorter inserts made out of oligonucleotides instead of (ATC/TAG)_n polymers. As shown in Figure 1, the pATC6b construct contains a run of four C–G bp (between positions –6 and –9) within the 18 bp (ATC/TAG) repetition. The structural properties of the halted ternary complex in pATC6b were compared with those obtained with the control plasmid pATC6a which harbors a perfect (ATC/TAG)₆ stretch (Figure 1). The two plasmids were first analyzed by *in situ* CAA probing experiments. The reactivity patterns of the non-template strands, shown in Figure 5A, offer a revealing picture of the effect of the C–G run on the apparent features of the transcription bubble. As for pATC21, the CAA reactivities in pATC6a are spread over a region of ~14 base residues located upstream from position –6. In contrast, the CAA modifications in pATC6b are stronger and are limited to a segment of at most eight base residues (between positions –6 and –13). Obviously, these results indicate that the halted ternary complex in pATC6b is stably positioned at the downstream location. Thus, the presence of the four consecutive rC–dG bp within the putative RNA–DNA hybrid seems to exert a strong effect on the lateral stability of the ternary complex. The same conclusion was reached by the nuclease S1 protection experiments shown in Figure 5B. Total RNAs extracted from cells harboring the plasmids were hybridized to the template strand of the corresponding *Bam*HI–*Pst*I DNA fragment and digested with S1 nuclease. Whereas the short RNAs produced from pATC6a showed a cluster of 3' ends (between positions –6 and –9) that is similar to the one obtained for pATC21, only a single band

Fig. 4. *In situ* DMS modification patterns of the template strands in pATC21 and pATC21H. The direction of transcription (horizontal arrow) and the *lac* operator motif (box) are depicted. The densitometer scans of the minus IPTG lanes shown in the bottom panel were obtained as in Figure 2. The position of the G-6 residue is indicated.

corresponding to a 3' end at position -6 was observed for pATC6b (Figure 5B).

We next compared the methylation status of the G residues within the template strands (Figure 6). As expected, the G residue at position -6 in the control plasmid pATC6a was heavily methylated in the absence of IPTG. However, this IPTG-dependent hypermethylation was not detected in pATC6b. The same methylation pattern was observed either in the absence or the presence of the inducer (Figure 6). As is generally seen in the Maxam and Gilbert sequencing ladders, the higher methylation of the G-6 residue as compared with the surrounding Gs in pATC6b is inherent in its position at the 5' side of the (G)₄ stretch. To show this phenomenon *in vivo*, we constructed an additional control plasmid (pATC6c) in which the (G)₄ stretch is located downstream from position -6 (see Figure 1). Indeed, in this construct, the G residue on the 5' side of the stretch (position -3) also has an inherently high reactivity with DMS (Figure 6). By means of CAA footprinting and S1 mapping experiments (results not shown), we found that the halted ternary complex in pATC6c has the same conformational behavior as in pATC6a. This is consistent with the IPTG-dependent hypermethylation of the G-6 residue shown in Figure 6 (pATC6c). Together, these results provide a strong indication that the hypermethylation of the G-6 residue is linked to the dynamic rearrangements of the halted ternary complex.

Discussion

In this study, we have used DNA and RNA footprinting experiments to analyze the conformational

behavior of a transcription elongation complex that was halted *in vivo* by a protein readblock. Our results indicate that the ternary complex halted within an (ATC/TAG)_n sequence is distributed between upstream- and downstream-translocated conformers. The different forms of the complex are apparently in an active equilibrium that can be shifted towards the downstream conformation either by increasing the strength of the putative RNA-DNA hybrid or by the pairing of upstream RNA sequences to the nascent transcript in the vicinity of the complex. Thus, within the framework of the current models for transcript elongation (see Introduction), these footprinting data provide a dynamic view of the structural rearrangements experienced by the ternary complex inside the cell.

In this view (see the model in Figure 7), the transcribing RNA polymerase elongates the transcript until the operator-bound repressor becomes an extreme physical barrier. Within this downstream-translocated complex, the catalytic center of the enzyme (last nucleotide added in the RNA) is at position -6 relative to the upstream edge of the 22 bp operator sequence. We have previously shown that the *lac* repressor binds one side of the DNA double helix with a footprint covering the full operator motif (Guérin *et al.*, 1996). The ability of the polymerase to extend the transcript 5-6 bp closer to the repressor footprint was then taken to suggest a structural flexibility between the catalytic site and the front end of the enzyme. An alternative interpretation, however, is that the halted polymerase and the repressor occupy opposite sides on the helix with the 5-6 bp distance being the half-turn angular separation between the two individual footprints. Therefore, we have no obvious reason to believe that the

Fig. 5. (A) Primer extension analyses of the *in situ* CAA modifications on the non-template strands in pATC6a and pATC6b. Relevant segments of the repeats upstream from position -6 are depicted on the side with the dots representing the observed CAA modifications. (B) S1 mapping of the 3' ends of the short RNAs produced *in vivo* from the plasmids pATC6a and pATC6b.

polymerase within this downstream-translocated complex is in a constrained configuration.

As revealed by the CAA reactivity pattern in pATC6b, the single-stranded transcription bubble of the complex involves at most 8 bp with the downstream margin of the bubble within the catalytic center of the polymerase. The eight nucleotides (nt) of the template strand in the bubble are presumably paired with the 3' proximal part of the transcript which is actually the size of the RNA-DNA hybrid found in many *in vitro* studies (Nudler *et al.*, 1997; Komissarova and Kashlev, 1998; Sidorenkov *et al.*, 1998). When this RNA-DNA hybrid is relatively strong, such as in pATC6b (five rC-dG bp in the 8 bp heteroduplex), the complex remains stably positioned at this downstream location (Figure 7B). However, in the case of a weaker hybrid such as in pATC21, pATC6a and pATC6c (three rC-dG bp in the 8 bp heteroduplex), the halted ternary complex experiences a reverse translocation with an apparent shortening of the 3'-proximal part of the transcript by 3 nt from position -6 to position -9 (Figure 7A). Most *in vitro* studies assign the transcript cleavage function to the catalytic center of RNA polymerase, an activity that is simply stimulated by GreA, GreB or SII factors (Rudd *et al.*, 1994; Orlova *et al.*, 1995). Thus, from our S1 mapping data, we infer that the upstream shift of the complex relocates the catalytic site of the enzyme at position -9 where transcript cleavage occurs. Presumably, the backward translocation of the halted complex is limited

to 3 bp by the presence of a second polymerase further upstream which is in line on the template (bracket in Figure 2). In the presence of NTP substrates inside the bacteria, the polymerase re-elongates the transcript from the new 3' end up to position -6 (the downstream location). In this regard, the detection of small amounts of RNAs with 3' ends at positions -7 and -8 (see Figures 3 and 5B) may reflect either the presence of the elongation intermediates or the fact that the upstream translocation and the concomitant transcript cleavage actually occur in a step-by-step process.

The DMS probing experiments show that the hypermethylation of the G-6 residue on the template strand is associated with the structural rearrangements of the complex. The most reasonable interpretation of this result is that the N7 position of the template G residue becomes highly exposed to DMS attack during the incorporation of the opposite rC residue on the transcript. Since the incorporation of an NMP residue is a relatively fast process, its detection by DMS probing (hypermethylation) strongly suggests that the ternary complex halted within the (ATC/TAG)_n sequence is repeatedly switching between the upstream and downstream locations with the accompanying transcript cleavage and re-synthesis. The apparent accessibility of a large region of the non-template strand to CAA modification in pATC21 and its contraction in pATC21H provide additional support for this back and forth oscillation of the complex. Obviously, the modifica-

Fig. 6. *In situ* DMS modification patterns of the template strands in pATC6a, pATC6b and pATC6c. The symbols are the same as in Figure 4.

tion pattern in pATC21 results from the superimposition of the footprints of the different forms of the complex that are in equilibrium (Figure 7A). The shift of the equilibrium induced by the *cis*-acting sequence in pATC21H increases the concentration of the downstream-translocated form and, thus, leads to stronger modifications within a somewhat reduced region (Figure 7C). However, the fact that the apparent melted region in pATC21H is still larger than the one detected in pATC6b (8 nt) indicates that the effect of the upstream (AUG)_n RNA sequence is not sufficient to trap the ternary complex completely in the downstream location. In pATC21H, we observed a decrease of the G-6 hypermethylation, but not its complete disappearance (Figure 4), and we detected a low amount of RNAs with the 3' end at position -9 (Figure 3). These results suggest that the putative pairing of the upstream (AUG)_n RNA motif probably acts by lowering the frequency of lateral oscillations of the complex.

In this connection, our finding that the extent of the single-stranded transcription bubble is limited to 8 bp when the RNA polymerase is trapped in one position has important implications for the current models of lateral oscillations of the ternary complex. Indeed, this estimation differs significantly from the usual 12–16 bp of melted region found in many *in vitro* and *in vivo* studies of halted ternary complexes (see for example Kainz and Roberts, 1992; Lee and Landick, 1992; Zaychikov *et al.*, 1995).

Although we cannot exclude the possibility that the extent of the melted region is dependent upon the sequence context, our data point out the attractive hypothesis that the size of the transcription bubble is very often overestimated due to the presence of multiple laterally translocated isoforms during the analyses. If this is true, then the size of the melted region in each elongation intermediate should be limited to the 8 bp that are required to accommodate the RNA–DNA heteroduplex. The ‘zip-ping’ model of transcript elongation suggests that the lateral oscillations of the ternary complex should be controlled by the direct competition between the DNA–DNA or RNA–DNA base pairs that can be formed at the downstream branching point of the transcription bubble (Komissarova and Kashlev, 1997b; Nudler *et al.*, 1997; Sidorenkov *et al.*, 1998). In the perspective where the melted DNA region has the same size as the RNA–DNA hybrid in each elongation intermediate, similar competition at the upstream branching point of the bubble should also contribute to the lateral stability of the complex.

Although the dynamic rearrangements described in this study are those experienced by a ternary complex that was halted artificially, we believe that they probably reflect many mechanistic aspects of transcript elongation *in vivo*. To transcribe the cellular DNA efficiently, RNA polymerase has to overcome certain transcriptional impediments. As suggested by the *in vitro* studies, these can be template DNA sequences that induce lateral oscillations of the ternary complex causing a slowdown in RNA chain synthesis and ultimately leading the polymerase to pause, arrest or terminate transcription. In the light of our results, these lateral movements of the complex and their regulatory consequences are likely to occur at the gene level (multipolymerase system and inside the cell). Moreover, the modulation of the lateral stability of the complex by a *cis*- or a *trans*-acting RNA sequence adds a new mechanism by which the cell may regulate gene expression.

Chromosomal proteins, as well as site-specific binding proteins, constitute another class of blocks to transcript elongation *in vivo*. Our present data show that within certain template DNA sequences, such impediments induce a backward sliding of RNA polymerase. By measuring the expression of the downstream CAT gene, we found that a reasonable amount of polymerases bypass the readblock (results not shown). This is also seen by the amount of full-length probe in the S1 mapping experiments. Preliminary results in our laboratory indicate that the bypass is probably enhanced by the structural rearrangements of the complex via the transcript cleavage and re-synthesis process. Hence, as previously suggested from *in vitro* studies (Reines and Mote, 1993), factor-stimulated transcript cleavage within protein-readblocked ternary complexes may play a crucial role in the processivity of RNA polymerase when transcribing DNA in the cellular context.

Materials and methods

Enzymes, chemicals and oligonucleotides

The restriction enzymes and phage T4 polynucleotide kinase were obtained from New England Biolabs or Eurogentec Belgium. Klenow fragment of DNA polymerase I was purchased from Amersham. S1

Fig. 7. A model of the structural rearrangements experienced by the transcription elongation complex that was halted *in vivo* by the protein readblock. The direction of transcription is from left to right. The RNA polymerase is represented by the open rectangles and the bound lac repressor by the filled box. (A) Halting of the polymerase within an $(ATC/TAG)_n$ sequence leads to lateral oscillations of the ternary complex between positions -6 and -9 (6 and 9 bp distance between the catalytic center of the enzyme and the upstream edge of the operator motif, respectively). The oscillations are accompanied by the cleavage of the 3' proximal part of the transcript (3 nt) and its resynthesis. Since the ternary complex is in equilibrium between the upstream and downstream locations, the apparent size of the transcription bubble results from the superimposition of the footprints (filled rectangles) of the two forms of the complex. (B) When the RNA–DNA hybrid is relatively strong, the complex is trapped at the downstream location (position -6) where the size of the transcription bubble is limited to 8 bp. (C) The equilibrium is shifted towards the downstream location by the presence of the 5' terminal $(AUG)_n$ RNA sequence that can pair with a segment of the nascent transcript in the vicinity of the halted polymerase.

nuclease and T4 DNA ligase were from Gibco-BRL. DMS and most chemicals, including antibiotics, were obtained from Sigma. CAA was bought from Fluka chemie and double distilled before use (boiling point $78-80^\circ\text{C}$). Unlabeled dNTPs and ddNTPs were purchased from Boehringer Mannheim, whereas $[\alpha\text{-}^{32}\text{P}]\text{dATP}$ and $[\gamma\text{-}^{32}\text{P}]\text{ATP}$ were from Amersham. All the oligonucleotides were synthesized by Eurogentec Belgium.

Plasmids and bacteria

The original plasmid vector having the CAT gene under the control of the constitutive HisR promoter is a derivative of pKK232-8; its construction was described previously (Guérin *et al.*, 1996; Figueroa-Bossi *et al.*, 1998). The plasmid pATC21H was derived from pATC21 (Guérin *et al.*, 1996) by inserting an $(ATC/TAG)_{20}$ polymer into the *HpaI* site. The orientation of the insert was checked by dideoxy sequencing. The construction of pATC6a, pATC6b and pATC6c was carried out exactly as previously reported for pATC21, except that the inserts were made out of oligonucleotides instead of polymers. Briefly, the 18mer double-stranded oligonucleotides were first ligated to a 22mer autocomplementary *lac* operator oligonucleotide. The desired ligation product was purified on a 12% polyacrylamide gel and inserted into the filled-in *SaI* site of the vector (position $+40$ relative to the transcription start site). The correct constructs were selected by dideoxy sequencing analyses. As for their construction, the plasmids were maintained in *E. coli* strain SU1675 (Figueroa-Bossi *et al.*, 1998). To have a high yield of intracellular *lac* repressor for the *in vivo* experiments, the cells harboring the plasmids were co-transformed with a pACYC177 derivative (pAC177IQ) in which a 1100 bp *lac* IQ gene was inserted between the *Bam*HI and *Ban*I sites. The two plasmids were maintained within the cell by a double selection (ampicillin 100 $\mu\text{g}/\text{ml}$ and kanamycin 30 $\mu\text{g}/\text{ml}$). In general, the isolation of the DNA fragments, the ligation reactions and the transformations into the *E. coli* cells were performed according to standard procedures (Sambrook *et al.*, 1989).

In situ CAA and DMS modifications and DNA analyses

Cells containing the plasmids were grown at 37°C in 10 ml of M9 minimal medium supplemented with casamino acids (4 mg/ml) and the appropriate antibiotics. At an optical density at 600 nm of 0.8, cells were exposed to IPTG (1 mM final concentration) for 15 min when required. The incubations with the chemical probes, as well as the subsequent DNA extractions by the alkaline lysis procedure, were performed exactly as previously reported (Guérin *et al.*, 1996). The DNA modifications were analyzed either by primer extension (CAA) or by direct detection after 3' end labeling and hot piperidine cleavage (DMS). The primer extension method, including its use for the dideoxy sequencing, has already been described (Rahmouni and Wells, 1992). The extension products were analyzed on a 6% polyacrylamide gel containing 7 M urea in parallel with dideoxy sequencing ladders. For the direct detection, the plasmids extracted from the cells were digested with *Bam*HI and *Pst*I, and labeled by filling in the 5' overhangs with the Klenow fragment of DNA polymerase I. The desired 3' end-labeled fragment was purified from an 8% polyacrylamide gel and treated with 1 M piperidine for 20 min at 90°C . After lyophilization, the samples were analyzed on an 8% denaturing polyacrylamide gel in parallel with Maxam and Gilbert sequencing ladders.

RNA extraction and S1 nuclease mapping

In this study, we wished to obtain highly accurate determinations of the RNA 3' ends. Thus, we used a new extraction method in which the bacteria are quickly lysed and phenol extracted to avoid the RNA truncations by the RNases. Cells bearing the plasmids were grown in 5 ml of Luria–Bertani medium with the appropriate antibiotics up to an OD at 600 nm of 0.6–0.7 and induced with 1 mM IPTG for 15 min when desired. One milliliter of the culture was mixed with an equal volume of a hot solution containing 20 mM sodium acetate pH 5.2, 0.3 M sucrose and 2% SDS, and heated in a boiling water bath for 30 s. The hot samples were immediately extracted with 2 ml of phenol. After

a second extraction with phenol/chloroform, the nucleic acids were precipitated with ethanol and resuspended in sterile water. The single-stranded DNA probes labeled at the 3' end were prepared from the appropriate plasmids and purified by gel electrophoresis as previously described (Guérin et al., 1996). Twenty micrograms of total RNAs (half of the RNA extracted from 1 ml culture) were mixed with the DNA probe (10 000 c.p.m.) in 50 µl of 40 mM HEPES pH 6.5, 0.3 M NaCl, 1 mM EDTA, 1 mM dithiothreitol and 0.1% Triton. The samples were first heated to 95°C for 10 min and immediately put at 55°C, then allowed to reach 30°C slowly. The hybridization process was continued at this temperature for 12 h. After addition of 300 µl of S1 buffer [30 mM sodium acetate pH 5.2, 1 mM zinc sulfate; 3 µg/ml denatured herring DNA and 5% glycerol] containing 240 U of S1 nuclease, the samples were incubated for 15 min at 12°C. The digestion was arrested by addition of EDTA and phenol/chloroform extraction. The protected hybrids were precipitated with ethanol and analyzed on an 8% denaturing polyacrylamide gel in parallel with sequencing ladders.

Acknowledgements

This work was supported in part by la Ligue Contre le Cancer, L'Association de la Recherche sur le Cancer and l'ANRS.

References

- Chamberlin, M.J. (1995) New models for the mechanism of transcription elongation and its regulation. *Harvey Lect.*, **88**, 1–21.
- Chan, C.L. and Landick, R. (1994) New perspectives on RNA chain elongation and termination by *E. coli* RNA polymerase. In Conaway, R.C. and Conaway, J.W. (eds), *Transcription: Mechanisms and Regulation*. Raven Press, New York, NY, pp. 297–321.
- Cook, P.R. (1999) The organization of replication and transcription. *Science*, **284**, 1790–1795.
- Figueroa-Bossi, N., Guérin, M., Rahmouni, A.R., Leng, M. and Bossi, L. (1998) The supercoiling sensitivity of a bacterial tRNA promoter parallels its responsiveness to stringent control. *EMBO J.*, **17**, 2359–2367.
- Gelles, J. and Landick, R. (1998) RNA polymerase as a molecular motor. *Cell*, **93**, 13–16.
- Gu, W., Wind, M. and Reines, D. (1996) Increased accommodation of nascent RNA in a product site on RNA polymerase II during arrest. *Proc. Natl Acad. Sci. USA*, **93**, 6935–6940.
- Guajardo, R. and Sousa, R. (1997) A model for the mechanism of polymerase translocation. *J. Mol. Biol.*, **265**, 8–19.
- Guérin, M., Leng, M. and Rahmouni, A.R. (1996) High resolution mapping of *E. coli* transcription elongation complex *in situ* reveals protein interactions with the non-transcribed strand. *EMBO J.*, **15**, 5397–5407.
- Iborra, F.J., Pombo, A., McManus, J., Jackson, D.A. and Cook, P.R. (1996) The topology of transcription by immobilized polymerases. *Exp. Cell Res.*, **229**, 167–173.
- Kainz, M. and Roberts, J. (1992) Structure of transcription elongation complex *in vivo*. *Science*, **255**, 838–841.
- Komissarova, N. and Kashlev, M. (1997a) Transcriptional arrest: *Escherichia coli* RNA polymerase translocates backward, leaving the 3' end of the RNA intact and extruded. *Proc. Natl Acad. Sci. USA*, **94**, 1755–1760.
- Komissarova, N. and Kashlev, M. (1997b) RNA polymerase switches between inactivated and activated states translocating back and forth along the DNA and the RNA. *J. Biol. Chem.*, **272**, 15329–15338.
- Komissarova, N. and Kashlev, M. (1998) Functional topography of nascent RNA in elongation intermediates of RNA polymerase. *Proc. Natl Acad. Sci. USA*, **95**, 14699–14704.
- Krummel, B. and Chamberlin, M.J. (1992) Structural analysis of ternary complexes of *Escherichia coli* RNA polymerase. Deoxyribonuclease I footprinting of defined complexes. *J. Mol. Biol.*, **225**, 239–250.
- Landick, R. (1997) RNA polymerase slides home: Pause and termination site recognition. *Cell*, **88**, 741–744.
- Lee, D.N. and Landick, R. (1992) Structure of RNA and DNA chains in paused transcription complexes containing *Escherichia coli* RNA polymerase. *J. Mol. Biol.*, **228**, 759–777.
- Lee, D.N., Feng, G.H. and Landick, R. (1994) GreA-induced transcript cleavage is accompanied by reverse translocation to a different transcription complex conformation. *J. Biol. Chem.*, **269**, 22295–22303.
- Nudler, E. (1999) Transcription elongation: Structural basis and mechanisms. *J. Mol. Biol.*, **288**, 1–12.
- Nudler, E., Goldfarb, A. and Kashlev, M. (1994) Discontinuous mechanism of transcription elongation. *Science*, **265**, 793–796.
- Nudler, E., Kashlev, M., Nikiforov, V. and Goldfarb, A. (1995) Coupling between transcription termination and RNA polymerase inchworming. *Cell*, **81**, 351–357.
- Nudler, E., Mustaev, A., Lukhtanov, E. and Goldfarb, A. (1997) The RNA–DNA hybrid maintains the register of transcription by preventing backtracking of RNA polymerase. *Cell*, **89**, 33–41.
- Orlova, M., Newlands, J., Das, A., Goldfarb, A. and Borukhov, S. (1995) Intrinsic transcript cleavage activity of RNA polymerase. *Proc. Natl Acad. Sci. USA*, **92**, 4596–4600.
- Rahmouni, A.R. and Wells, R.D. (1992) Direct evidence for the effect of transcription on local DNA supercoiling *in vivo*. *J. Mol. Biol.*, **223**, 131–144.
- Reeder, T.C. and Hawley, D.K. (1996) Promoter proximal sequences modulate RNA polymerase II elongation by a novel mechanism. *Cell*, **87**, 767–777.
- Reines, D. and Mote, J., Jr (1993) Elongation factor SII-dependent transcription by RNA polymerase II through a sequence-specific DNA-binding protein. *Proc. Natl Acad. Sci. USA*, **90**, 1917–1921.
- Richardson, J.P. and Greenblatt, J. (1996) Control of RNA chain elongation and termination. In Neidhardt, F.C. (ed.), *Escherichia coli and Salmonella typhimurium*, 2nd edn. American Society for Microbiology, Washington, DC, pp. 822–848.
- Rudd, M.D., Izban, M.G. and Luse, D.S. (1994) The active site of RNA polymerase II participates in transcript cleavage within arrested ternary complexes. *Proc. Natl Acad. Sci. USA*, **91**, 8057–8061.
- Sambrook, J., Fritsch, E.F. and Maniatis, T. (1989) *Molecular Cloning: A Laboratory Manual*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
- 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.
- Sidorenkov, I., Komissarova, N. and Kashlev, M. (1998) Crucial role of the RNA:DNA hybrid in the processivity of transcription. *Mol. Cell*, **2**, 55–64.
- Uptain, S.M., Kane, C.M. and Chamberlin, M.J. (1997) Basic mechanisms of transcription elongation and its regulation. *Annu. Rev. Biochem.*, **66**, 117–172.
- von Hippel, P.H. (1998) An integrated model of the transcription complex in elongation, termination and editing. *Science*, **281**, 660–665.
- Wang, D., Meier, I., Chan, C.L., Feng, G., Lee, D.N. and Landick, R. (1995) Discontinuous movements of DNA and RNA in RNA polymerase accompany formation of a paused transcription complex. *Cell*, **81**, 341–350.
- 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. USA*, **92**, 1739–1743.

Received June 3, 1999; revised and accepted July 27, 1999