The EMBO Journal vol.8 no.9 pp.2745 – 2754, 1989

A cinematographic view of *Escherichia coli* RNA polymerase translocation

Willi Metzger, Peter Schickor and Hermann Heumann

Abt. Molekulare Biologie der Genwirkungen, Max-Planck-Institut für Biochemie, D-8033 Martinsried b. München, FRG

Communicated by W.Zillig

A series of RNA synthesizing transcription complexes, initiated at the T7 A1 promoter and halted at specific base positions ranging from +12 to +40, were analyzed by footprinting techniques; exonuclease III was used to determine the position of the bound RNA polymerase on the DNA and hydroxyl radicals were used to visualize the protein–DNA contact sites within the protected areas. In the binding (open) complex without RNA there are two DNA-domains, differing in their protection pattern. The first, extending from position +18 to −13, termed 'melting domain', is fully protected, whereas the second, extending from −14 to −55, termed 'recognition domain', shows only partial protection. At this domain, RNA polymerase is attached to one side of the DNA only, as indicated by the 10-bp periodicity of the protection pattern. Our data show that the formation of a mature RNA transcribing complex is characterized by dissociation of the RNA polymerase from the recognition domain, whereby the size of the melting domain remains constant. This process is accomplished if the nascent RNA has reached a length of 11 bases. As the RNA reaches a length of 20 bases, the size of the melting domain decreases from ~30 to 23 bp. Further RNA synthesis leaves the protection pattern essentially unchanged. These data demonstrate that the formation of a mature RNA transcribing complex can be described by at least two transitions.

Key words: transcription/promoter

Introduction

Initiation of RNA synthesis, elongation and termination of the RNA chain are the three steps of transcription catalyzed by the DNA-dependent RNA polymerase (subunit composition: β', β, α_2, σ) which lead to RNA product formation (Chamberlin, 1976; von Hippel *et al.*, 1984). Initiation includes selection of the promoter sequence (von Hippel *et al.*, 1982; Ricchetti *et al.*, 1988), open complex formation with melting of the DNA and initiation of RNA synthesis. After synthesis of an RNA chain consisting of eight or nine bases (Hansen and McClure, 1980), the abortive transcription state (McClure *et al.*, 1978) enters the productive state (Carpousis and Gralla, 1980). This state is characterized by a stable complex (ternary complex) consisting of RNA, DNA and RNA polymerase without the initiating factor σ. The DNA is opened over a region of ~11–17 bp (Siebenlist, 1979; Gamper and Hearst, 1982; Kirkegaard *et al.*, 1983). In this open region ~12 bases (Hanna and Meares, 1983) of the 3′ end of the RNA are hybridized to the antisense strand of the DNA. This ternary complex moves downstream, incorporating the bases into the RNA chain according to the sequence in the antisense strand.

Dennis and Sylvester (1981) describe translocation of RNA polymerase during RNA synthesis as a rotation of the RNA polymerase. They assume that movement of the RNA polymerase along the DNA occurs in discrete steps. Gamper and Hearst (1982) propose that RNA polymerase moves along the DNA like a nut on a bolt, with the RNA polymerase bound to the two single-stranded regions in the DNA bubble.

Yager and von Hippel (1987) suggest that translocation is thermally driven, as in the case of one-dimensional diffusion of the RNA polymerase during promoter location (Ricchetti *et al.*, 1988). This implies that RNA polymerase has some degree of freedom in moving upstream and downstream during translocation. The direction of the process might be driven by the free energy gained by docking the following nucleoside triphosphate to the 3′ end of the RNA chain.

There is presently no unifying concept combining the different views of translocation. A lack of data due to experimental difficulties in tackling translocation problems might be the reason. The investigation of this process has been hampered due to the difficulties in bypassing the abortive transcription state and entering the productive state with a complex halted at a specific register of RNA synthesis.

This paper presents a strategy to overcome these difficulties. We prepared ternary complexes with RNA chain lengths of 11, 16, 20, 24, 25, 27, 29, 30, 33, 36 and 39 bases. These complexes were subjected to an exonuclease III and a hydroxyl radical treatment. The first probe provided information about the borders of the footprints of RNA polymerase on the DNA, the second probe about the contact points of RNA polymerase and DNA within the footprint. We illustrate the translocation of RNA polymerase during the initial phase of transcription by a sequence of single footprints. This cinematographic view includes a topological model of RNA polymerase and DNA during the transition from the binary to the ternary complex.

Results

Monitoring the movement of RNA polymerase during RNA synthesis required halting RNA synthesis at a specific base and subsequent determination of the position of RNA polymerase on the DNA by footprinting techniques.

Ternary complex formation with the T7 A1-promoter

The formation of a ternary complex in a specific register required a precise start and a precise stop of RNA synthesis bypassing the abortive transcription state.

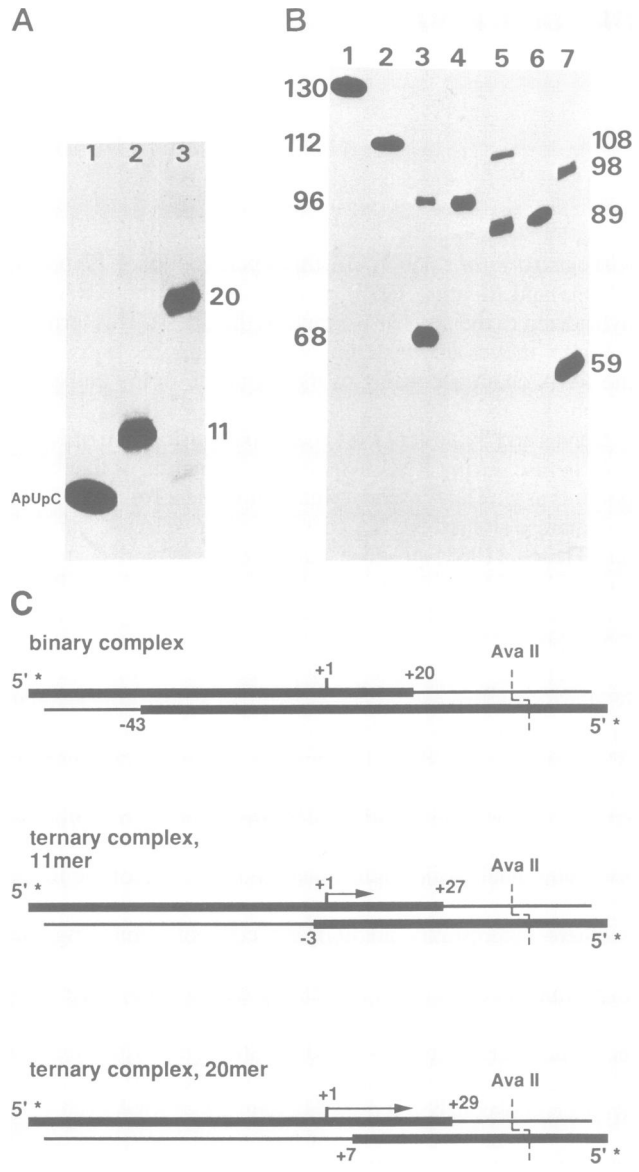
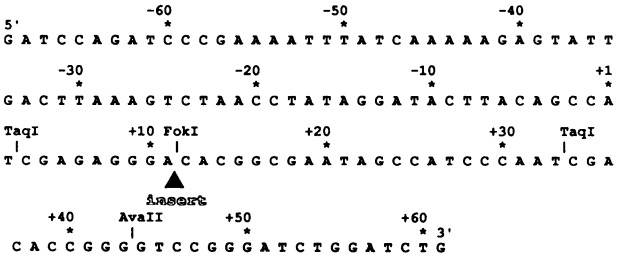


Fig. 1. Complexes obtained with the T7 A1 promoter carrying fragment (also see Figure 2). (**A**) RNA products of the ternary complexes stopped in the 11mer register by leaving out CTP and UTP in the assay (**lane 2**), and stopped in the 20mer register by leaving out UTP (**lane 3**). **Lane 1**: the primer ribo-oligonucleotide ApUp*C; * indicates the ^{32}P-label. (**B**) The exonuclease III footprints of the binary complex (**lane 5**), after cutting with *Ava*II, which removes the downstream positioned 5′ label (**lane 6**); of the ternary complex in the 11mer register (**lane 3**), after cutting with *Ava*II (**lane 4**); and of the 20mer register (**lane 7**). **Lanes 1** and **2** are size-markers. (**C**) A schematic representation of the exonuclease III footprints. Numbers refer to base positions, +1 indicates the start of RNA synthesis.

A precise start point for RNA synthesis was defined by using a ribotrinucleotide as a primer and a precise stop by leaving out one or two nucleoside triphosphates in the substrate mix. This artificial stop of RNA synthesis (Kinsella *et al.*, 1982), called stalling (Yager and von Hippel, 1987), was established by Levin *et al.* (1987). Using the A1 promoter of the phage T7 and ApUpC as a primer, RNA synthesis is stopped in the 20mer register if UTP is lacking and in the 11mer if UTP and CTP are lacking. The precise stop was verified by analysis of the length of the RNA chain bound to the complex (Figure 1A).



Insert	sequence	RNA-chain lengths (bold) with ATP,UTP:	ATP,UTP,CTP:
0	CGAGAGGGA	M0 **11**(C)	M0**20**(T)
5	CGAGAGGGAagagg	M5 **16**(C)	M5 **25**(T)
9	CGAGAGGGAagagagagg	M9 **20**(C)	M9 **29**(T)
13	CGAGAGGGAagagagagagagg	M13 **24**(C)	M13 **33**(T)
16	CGAGAGGGAagagaagagagagagg	M16 **27**(C)	M16 **36**(T)
19	CGAGAGGGAagagaagaagagagagagg	M19 **30**(C)	M19 **39**(T)

Fig. 2. Sequence of the 130 bp DNA fragment carrying the T7 A1 promoter (sense strand), inserted sequences and nomenclature of the different mutations. Restriction and insert sites are indicated. The lower cases represent the sequences inserted before the C on position +12. The whole sequence shows the inserts between the *Taq*I site at +2 and the *Fok*I site at +11 (see Materials and methods). Mx **y**(z) means: M = mutation; x = number of inserted nucleotides; **y** = RNA chain length in bases; z = stop inducing nucleotide (C or T at position +**1**).

Exonuclease III footprinting of RNA polymerase–DNA complexes

Open (binary) complex and RNA transcribing (ternary) complexes in the 11mer and 20mer register were subjected to exonuclease III treatments. The homogeneity and the length of the RNA products (Figure 1A) were checked. The borders of the RNA polymerase on the DNA were mapped by analyzing the lengths of the two 5′-labeled DNA strands left after exonuclease III digestion (Figure 1B). As expected, there are two bands from which the borders of the protected areas are determined. In order to attribute correctly the bands to the sense and antisense strands, the downstream 5′ label was removed by a restriction cut (*Ava*II). Figure 1C is a schematic display of the footprinting data of Figure 1B: during transition from the binary to the ternary complex in the 11mer register, the upstream border moves from position −43 to position −3, the downstream border from +20 to +27. This confirms DNase I footprinting data (see Table I) which showed previously that the size of the protected area decreases dramatically during transition from the specifically bound to the RNA transcribing state. The protection pattern changes further during the transition from the 11mer register complex to the 20mer register complex. The upstream boundary of the protected sequence moves nine bases in the direction of transcription as expected, whereas the downstream boundary moves only two bases. Taking the position of the 3′ end of the RNA chain as fixed, RNA polymerase contracts asymmetrically.

Exonuclease III footprints of ternary complexes in different registers

Since the technique of exonuclease III footprinting of stalled complexes proved to be successful for the 11mer and 20mer register, we extended the method to additional registers by mutagenesis of the downstream sequence of the A1 promoter. Between position +11 and +12 (a *Fok*I restriction site) additional sequences containing G and A bases were inserted (see Materials and methods and Figure 2). Each fragment with one of these insertions, shown in Figure 2, established two new registers applying the same experimental

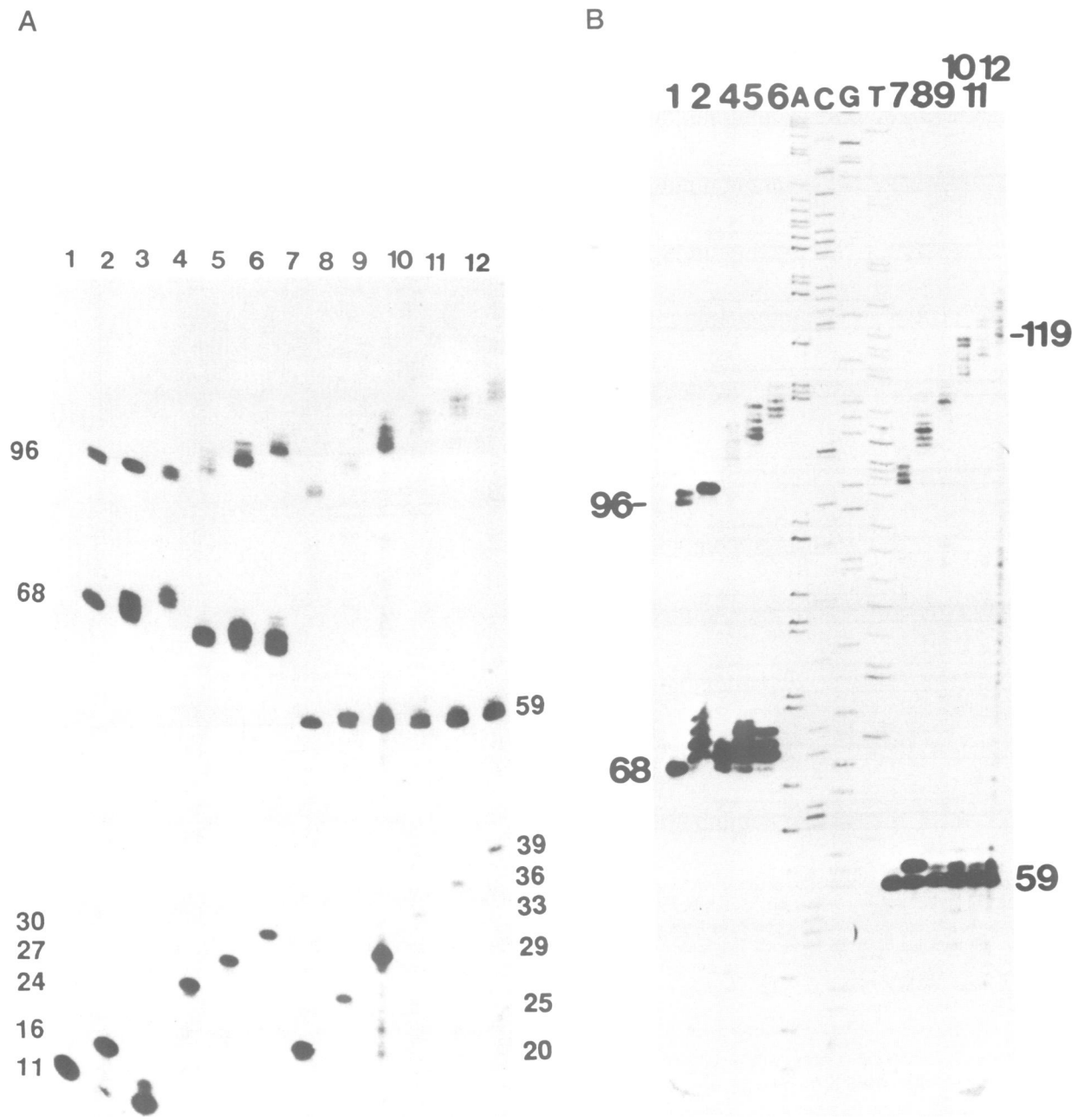


Fig. 3. Analysis of the RNA products and the exonuclease III footprints of the ternary complexes on an 8% acrylamide gel. (**A**) The RNA products (and exonuclease III footprints). **Lanes 1–6**: complexes stopped by lack of CTP and UTP [from left to right: M0 **11**(C), M5 **16**(C), M9 **20**(C), M13 **24**(C), M16 **27**(C), M19 **30**(C)]; **lanes 7–12**: complexes stopped by lack of UTP [from left to right: M0 **20**(T), M5 **25**(T), M9 **29**(T), M13 **33**(T), M16 **36**(T), M19 **39** (T)].

procedure developed for the wild-type promoter. Figure 3A shows the RNA products in these registers, and Figure 3B the corresponding exonuclease III footprints. Figure 4 is a graphic display of the footprints. We generated two sets of ternary complexes: the first set is halted before the first cytosine in the sense strand, the second set is halted before the first thymidine. Within each set of ternary complexes the DNA has a high degree of sequence homology: the sequence downstream of the halt-inducing base (C or T, which we term base position $+\bar{1}$), is identical within each set. The sequences upstream of $+\bar{1}$ are to a large extent homologous (see Figure 2). Using the $+\bar{1}$ position as a reference to arrange the different complexes, the following features of the footprints arise: the upstream-positioned border of the protected sequence moves downstream regularly, corresponding to the lengths of the transcripts. The distance between the $+\bar{1}$ position and the upstream border is constant (16 ± 1 bp). The protected area decreases from 30 bp in the 11mer register to ~23 bp in the 20mer register. In registers longer than 30 nucleotides the protected area increases slightly to 25 bp. This variation in the size of the protected area is caused by a change in only the downstream boundary of the RNA polymerase. These data suggest that there is a continuous transition of the ternary complex from the 11mer register to the 20mer register with a concomitant shortening of the protected area. We interpret

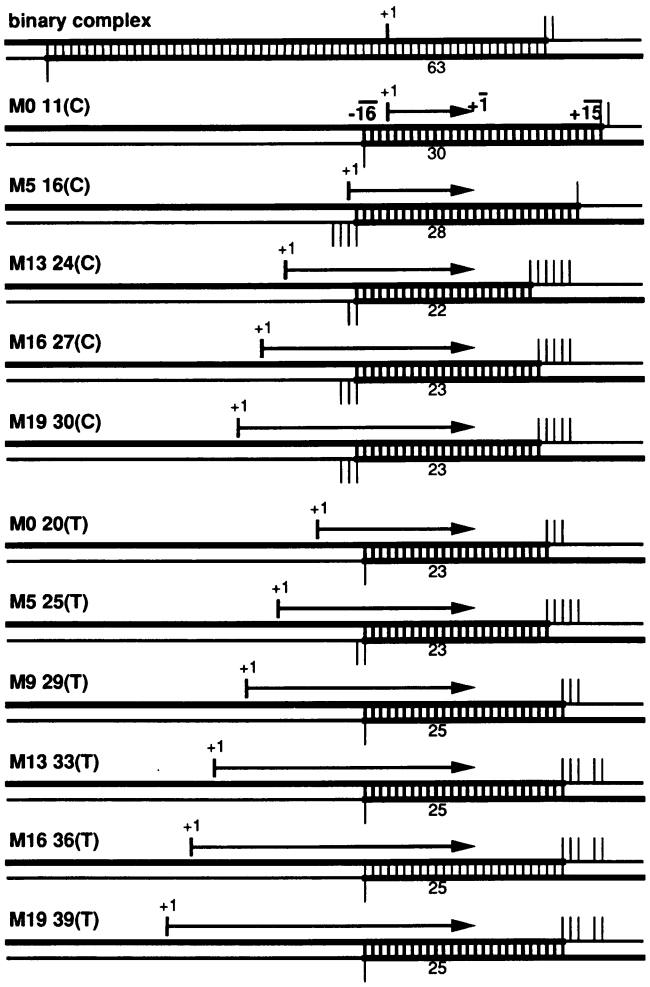


Fig. 4. Schematic representation of the exonuclease III footprints. The footprints are arranged in such a way that the 3′-ends of the RNA coincide. +1 marks the startpoint of transcription, $+\bar{1}$ refers to the position of the halt-inducing base. Each vertical line represents the position of 1 bp. The longer vertical lines represent the ends of the protected DNA areas obtained from Figure 3B. The numbers show the size of the protected areas in bp.

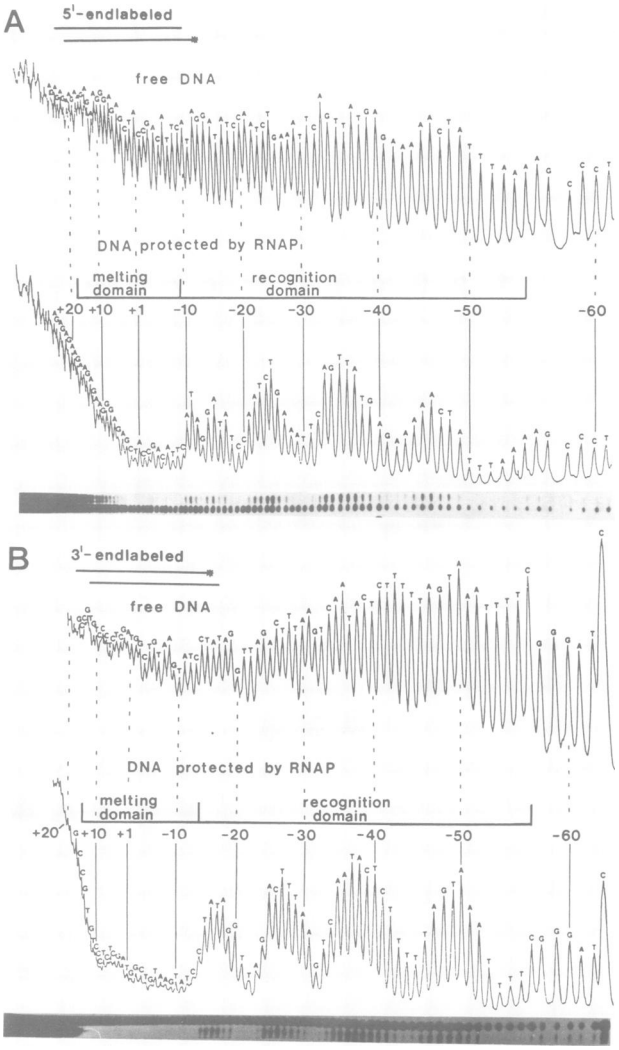


Fig. 5. Hydroxyl radical footprinting of the RNA polymerase bound to the T7 A1 promoter. The gel electrophoretic pattern and the corresponding densitrometric scan: (**A**) of the 5′ end labeled sense strand; and (**B**) of the 3′ end labeled antisense strand. (Note that the direction of transcription is from right to left.)

this transition as a maturation process of the ternary complex, which is completed when the RNA chain has reached a length of 20 bases. Another characteristic of the protection patterns is the clear definition of the upstream borders, in many cases within one base, whereas the downstream borders are less sharply defined and consist of more than one band in the gel, suggesting a variability in the positioning of the RNA polymerase with respect to the downstream border. It is unlikely that this is a result of a sequence-dependent digest of exonuclease III, since the sequence downstream of the $+\bar{1}$ position is the same within each set of ternary complexes.

Hydroxyl radical footprints of binary and ternary complexes

The exonuclease III probe was useful for monitoring the movement of the boundaries of the DNA areas protected by RNA polymerase during RNA synthesis, but this probe provides no information about the pattern of protein–DNA contacts within the protected area. This knowledge of the pattern would be useful, especially for discussion of the rotational model (Dennis and Sylvester, 1981). We have chosen hydroxyl radicals as a probe since this reagent cuts essentially non-sequence-specifically (Tullius and Dombroski, 1986) and therefore provides information about protein–DNA contacts without sequence-dependent disturbance.

We probed the open (binary) complex (Figure 5) and two RNA transcribing complexes, the M0 **11**(C) and the M0 **20**(T), halted in the 11mer and the 20mer register respectively (Figure 6). The first one represents a 'maturating' and the second a 'mature' ternary complex, as described by exonuclease III footprinting. Figure 7 summarizes the hydroxyl radical footprinting. In the open complex two areas can be discerned as suggested previously by Travers (1987). The first one reaches from base position +18 to position −10 with respect to the 5′ labeled sense strand. The depression of the bands within this region (melting domain) indicates full protection of the DNA. The second region (recognition domain), reaching from −11 to −52, displays a modulation of the intensity pattern with a periodicity of 10 bases, indicating that the RNA polymerase faces one side of the DNA (Figure 5).

The footprint of the RNA transcribing complexes (ternary complex) displays the same feature as does the open

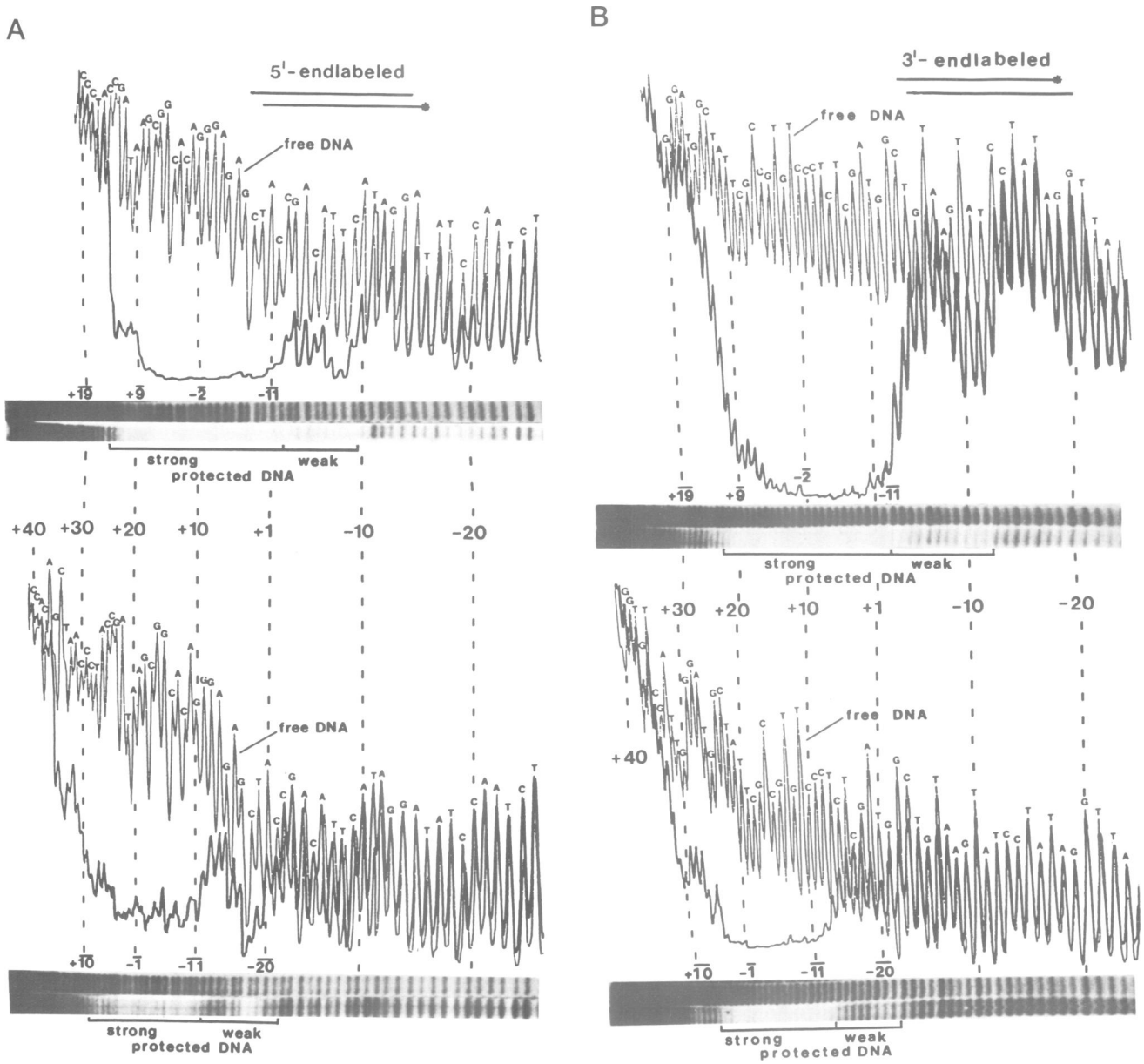


Fig. 6. Hydroxyl radical footprinting of the RNA polymerase in the 11mer register (upper panels) and the 20mer registers (lower panels). (**A**) 5′ end labeled sense strand; and (**B**) 3′ end labeled antisense strand.

complex. There are two domains: the melting domain and the recognition domain. However, the size of the recognition domain is drastically reduced. Only one of the four helical turns in the open complex is still protected in the RNA transcribing complex (Figure 7). Such a decrease in size was observed previously by DNase I footprinting (Carpousis and Gralla, 1985) and confirmed by our exonuclease III studies. The specific change of the pattern observed with the hydroxyl radical probe allowed us to identify the cause of this change. It indicates that the decrease in size is due to a dissociation of RNA polymerase in the upstream region of the recognition domain. The size of the melting domain ($\sim$27 bp) remains unchanged during transition from the open to the RNA transcribing state in the 11mer register. But the position of the melting domain is moved three bases further upstream if the protection patterns in the binary and ternary complexes are correlated with respect to position +1 and $+\bar{1}$,

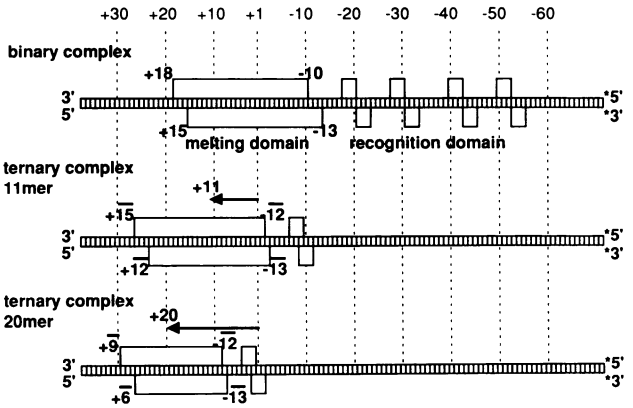


Fig. 7. A schematic representation of the hydroxy radical footprints of Figures 5 and 6.

the position of the halt inducing base. The size of the melting domain further decreases during transition to the 20mer state by ~6 bases. This change is due to a decrease in the melting domain at the downstream leading edge. This kind of asymmetric contraction of the footprint during transition of the RNA transcribing complex from the 11mer to the 20mer state was also observed with the exonuclease III probe.

Hydroxyl radical footprints were obtained with 5′ and 3′ labeled strands. Both show essentially the same pattern, but with a shift of 2–3 bases (Tullius and Dombroski, 1986), as expected for a protein attached to the cylindrical surface of the helically arranged DNA strands (Figures 5 and 6).

A size determination of the protected DNA is difficult since the transition from the protected to the unprotected area is not sharp. Therefore a comparison of the footprints obtained by exonuclease III and hydroxyl radicals is difficult. Notice that hydroxyl radical footprints determine both the upstream and downsteram border of protection by RNA polymerase on each strand of the DNA, whereas exonuclease III provides information about one border on each strand only. For a comparison of the borders obtained by the two different methods, the corresponding borders on the respective strands have to be compared (see Figures 1C and 7).

If the melting domain is considered, the downstream leading edge of the footprints obtained by exonuclease III and hydroxyl radicals agree. At the upstream leading edge the footprints obtained by exonuclease III are smaller than those obtained by hydroxyl radicals. In the binary complex, exonuclease III stops at −43, which is ~12 bases further downstream to the upstream leading edge determined by hydroxyl radicals. In the ternary complex the weakly protected areas (Figures 5 and 6) are not protected against exonuclease III. This indicates that exonuclease III can nibble into the area protected by RNA polymerase, which may be due to a weaker interaction of RNA polymerase with the DNA in this upstream region.

Backreaction of RNA chain formation

The RNA products were analyzed before and after exonuclease III digestion to examine the stability of the complexes.

There is one ternary complex in the 20mer register, the M9 **20**(C) complex (for nomenclature see Figure 2), which is not stable against exonuclease III digests (Figure 3A, lane 3). This complex forms the proper product of 20 bases (data not shown), but after exposure to exonuclease III the RNA chain is degraded. The RNA with the leading 5′ end radioactively labeled is progressively digested from the 3′ to the 5′ end until the 11- and 12mer register is reached. These RNA products are stably bound. Concomitantly, the exonuclease III footprint displays the characteristic pattern of an 11- or 12mer complex (Figure 3A, lane 3). This backreaction is reminiscent of the effect of pyrophosphate (Rozovskaya *et al.*, 1981 and 1982). Therefore, we studied the influence of pyrophosphate on the stability of the ternary complex in different registers (Figure 8). The reaction proceeds backwards to the 10mer but preferential stops can occur earlier. Exonuclease III also seems to be able to cause a backreaction of transcription like pyrophosphate. The shortest stably bound RNA product after pyrophosphate-induced backreaction is a 10mer, indicating a high energy barrier for further backreaction.

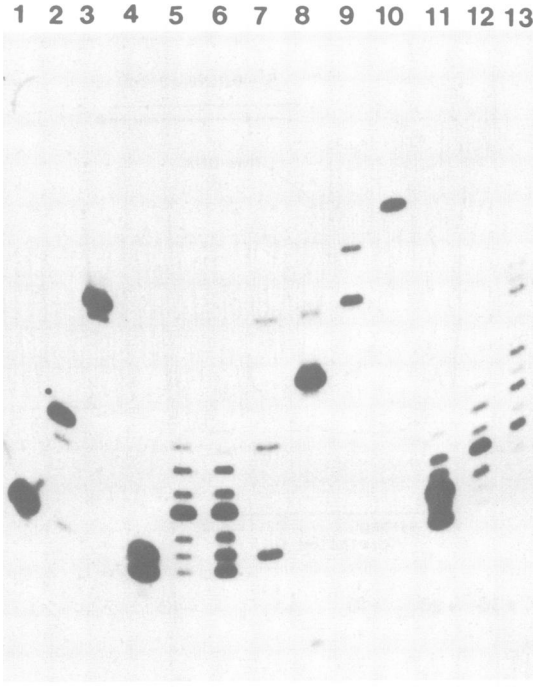


Fig. 8. Backreaction of RNA chain formation by addition of pyrophosphate. Ternary complexes were incubated with 10^{-2} M pyrophosphate for 40 min. The samples were applied to a non-denaturing gel, the complex bands were cut out, eluted and analyzed on a 20% sequencing gel. **Lanes 1, 2, 3, 8, 9** and **10** show the intact RNA products of chain lengths 11, 16, 24, 20, 25 and 33 bases respectively. **Lanes 4, 5, 6, 11, 12** and **13** show the corresponding degradation products bound to the complexes. **Lane 7** shows length markers of 11, 16 and 24 bases.

Discussion

The movement of RNA polymerase on the DNA during RNA synthesis was followed by analyzing a sequence of exonuclease III footprints on ternary complexes halted in different registers. The application of footprinting methods requires a 'freezing' of RNA synthesis in specific registers. This is an artificial interruption of a dynamic process. As a consequence, special attention must be given to the method inducing the interruption. Stalling of complexes is a method which halts RNA synthesis at a specific template base, where the corresponding nucleoside triphosphate is lacking in the substrate mix. A full restart of stalled complexes is possible by addition of the missing nucleoside triphosphate (data not shown; Levin and Chamberlin, 1987). Stalling of the complex is reversible and, therefore, preferable to other methods leading to an irreversible RNA chain termination. We used the T7 A1 promoter and derivatives of this promoter with sequences inserted upstream of the first halt inducing base. With ApUpC as a primer two sets of stalled complexes were generated, one halted before a deoxycytidine and another before a deoxythymidine (Figure 2).

The exonuclease and hydroxyl radical footprinting techniques provide complementary information about the boundaries and DNA contacts of RNA polymerase

The use of exonuclease III as a probe to monitor translocation has the advantage of a rather sharp transition from the protected to the unprotected area on the DNA, with the

Table I. Comparison of footprints of RNA polymerase–promoter complexes obtained by DNase I, exonuclease III and hydroxyl radicals

Promoter	Method	Complex	Protected area (approximately)	Authors
lac	DNase I	Binary	−50 to +20	Schmitz and Galas (1979)
lac-UV5	DNase I	Binary	−50 to +20	Carpousis and Gralla (1985)
lac-UV5	DNase I	Binary	−45 to +20	Straney and Crothers (1985)
lac-UV5	DNase I	Binary	− 45 to +20	Spassky (1986)
tac	DNase I	Binary	− to +20	Shi *et al.* (1988)
lac-UV5	DNase I	Binary	−42 to +20	Travers *et al.* (1983)
*tyr*T	DNase I	Binary	−65 to +20	Travers *et al.* (1983)
unc	DNase I	Binary	−75 to +20	Kanazawa *et al.* (1982)
A3	Exo III	Binary	−44 to −	Siebenlist *et al.* (1980)
lac-UV5	Exo III	Binary	−44(−24) to −	Straney and Crothers (1987a)
A1	Exo III	Binary	−43 to +20	Our results
A1	Hydrox. rad.	Binary	−55 to +18	Our results
lac-UV5	DNase I	Abortive	−50 to +20	Carpousis and Gralla (1985)
lac-UV5	DNase I	Abortive	−45 to +20	Spassky (1986)
lac-UV5	Exo III	Abortive	(−44)−24 to −	Straney and Crothers (1987a)
lac-UV5	DNase I	16mer	−5 to +25	Carpousis and Gralla (1985)
lac-UV5	DNase I	11mer	−15 to +20	Straney and Crothers (1985)
lac-UV5	Exo III	11mer	−6 to −	Straney and Crothers (1987a)
tac	DNase I	29mer	+7 to +37(+49)	Shi *et al.* (1988)
A1	Exo III	11mer	−3 to +27	Our results
A1	Exo III	20mer	+7 to +29	Our results
A1	Hydrox. rad.	11mer	(−11)−2 to +26	Our results
A1	Hydrox. rad.	11mer	(−1)+8 to +29	Our results

borders more clearly defined than in footprints obtained by other methods (Shalloway *et al.*, 1980).

It must be kept in mind, however, that (i) the probe can act in a sequence-dependent manner, which makes the interpretation of the protection pattern more difficult, and (ii) the probe itself can disturb the protein–DNA interaction.

DNase I is a probe with pronounced sequence-dependence. As a consequence, published footprints of RNA polymerase differ in size, as shown in Table I. A sequence-dependence of exonuclease III cannot be excluded but seems not to be very pronounced. Since the downstream sequences of the ternary complexes are identical within each set of the C-stopped and T-stopped ternary complexes, respectively, changes in this region should reflect changes in the accessibility of DNA to exonuclease III rather than a sequence-dependence of exonuclease III. The sequences at the upstream border differ in the different complexes, but the well defined border and the movement of this border in register, as expected, indicates that disturbance due to a sequence-dependent digest plays a minor role.

A possible pitfall in the interpretation of exonuclease III footprints is a partial disruption of the RNA polymerase–DNA interaction by exonuclease III. This becomes apparent at the upstream border of the footprints in the binary and ternary complexes. The footprints obtained by probing with hydroxyl radicals are 12 bases larger at the upstream leading edge in the binary complex and nine bases in the ternary complex (11mer and 20mer states). This indicates that areas of weaker protein–DNA interaction can be attacked by exonuclease III. For the *lac*-UV 5 promoter, a similar but even more pronounced attack by exonuclease III at the upstream border of the protected DNA was reported (Straney and Crothers, 1987a).

The results of our studies support the view that footprints obtained by hydroxyl radicals display a snapshot (Tullius, 1987) of the complex, whereas the footprint obtained by exonuclease III must be qualified as a lower limit. The high resolution of hydroxyl radical footprints allowed us to discern two binding domains, not visualized by other methods. The first domain, called the recognition domain, protects the recognition site around base position −35 and reaches from the upstream leading edge of the RNA polymerase to the 'Pribnow-box'. The RNA polymerase is attached to one side of the DNA in this region. We propose that the DNA is in the B-form, since the periodicity of the protection pattern is on average 10.5 bp. The second domain extending to position +18, including the 'Pribnow-box', comprises the melted DNA region. This region, called the melting domain, is fully protected by RNA polymerase.

The transcription bubble moves downstream in register with increasing RNA chain length

The position of the transcription bubble in the open and in the RNA-transcribing complex is well established: the DNA is opened from −8 to +3 in the binary complex (Siebenlist, 1979). An upper limit of the size of the transcription bubble in the ternary complex is 17 bases, as estimated by Gamper and Hearst (1982). The position of the transcription bubble in the ternary complex was determined by mapping the position of the DNA–RNA hybrid. It reaches 12 bases upstream from the $-\bar{1}$ position (Hanna and Meares, 1983). This is essentially confirmed by Shi *et al.* (1988). They propose, by DNA–DNA crosslinking, that at position $+\bar{1}$ or $+\bar{2}$ the first DNA bp is disrupted. The transcription bubble fits into the upstream leading part of the protected melting domain reaching from $-\bar{1}$ to $-\bar{12}$, referring to Hanna and Meares' data. The transcription bubble might extend, according to Gamper and Hearst, to $-\bar{17}$. In any case the footprinting pattern of the area comprising the transcription bubble ($\sim -\bar{1}$ to $-\bar{12}$) remains unchanged throughout transcription, as shown by probing with hydroxyl radicals (Figure 7). The same is valid for exonuclease

III footprints (Figure 4). The upstream border remains constant at position $-\overline{17}$ throughout RNA synthesis, i.e. the DNA re-anneals just at the end of the area protected by RNA polymerase. We conclude from all these findings that the transcription bubble moves in register with RNA synthesis.

The mature ternary complex with 20 or more bases in the RNA chain is formed via an intermediate with ~11 bases

In monitoring the change in size of the protected area from the binary to the mature ternary complex by exonuclease III, two states are suggested. One is the well known transition from the abortive to the productive state (Carpousis and Gralla, 1980). This transition occurs within synthesis of the first 10 or 11 bases of the RNA chain, with a concomitant decrease in the protected area from 63 to 30 bp. The size of the protected area decreases from 30 bp in the 11mer register to 23–25 bp in the mature ternary complex with an RNA chain length of 20 bases and longer, indicating a concomitant 'maturation' process of the ternary complex until the 20mer register is reached.

A simple topological model which explains the different features of the footprinting studies

Both transitions can be explained by the same kind of topological change of the RNA polymerase–DNA complex. We propose a dissociation of the recognition domain during transition from the binary to the ternary complex, whereas the melting domain remains bound. This can be realized by tilting the long axes of RNA polymerase and DNA around the $-\overline{16}$ position so that the angle between the axes increases. Only a minor increase in the angle of a few degrees would be required. This change of the topology of the RNA polymerase–DNA complex can be due to a lifting of the RNA polymerase (see Figure 9) or a bending of the DNA. The latter view is supported by the finding that an RNA polymerase induced bending of the promoter is also maintained in the ternary complex, probably with a larger bending angle (Heumann *et al.*, 1988b). We suggest that the topological change of the complex is triggered if the growing RNA chain reaches a certain critical length or if the σ-factor is released.

The change in the footprints during transition into the mature ternary complex can be explained by a further increase in the angle between RNA polymerase and the DNA axes. This resembles the opening of a pair of scissors with the joint of the scissors in the area upstream to the leading edge of the transcription bubble. This model explains why the upstream leading edge of the protected area is not affected, whereas the downstream leading border decreases in size. The opening angle of the scissors might depend on the length of the RNA. We speculate that RNA growing between the DNA and RNA polymerase acts like a wedge, increasing the angle until the RNA has reached a size of 20 bases.

The exonuclease III digests display a rather precisely determined footprint upstream and a rather fuzzy pattern downstream. We conclude from this finding that the upstream part with the transcription bubble is a defined arrangement of protein and DNA with a strong interacting domain, whereas the downstream part has a greater flexibility. Yager and von Hippel (1987) suggest that

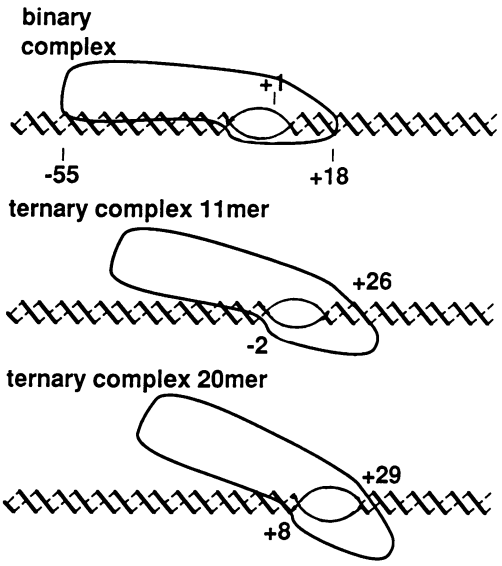


Fig. 9. A topological model of RNA polymerase and DNA derived from the footprinting studies describing the transition from binary to the ternary complex.

translocation of RNA polymerase might be thermally driven. We suggest that there is a thermally driven vibration of the RNA polymerase with respect to the DNA axis, with the fixpoint at the upstream border of the melting domain. This swing in the plain of the axes of RNA polymerase and DNA would explain the variation in the protected area downstream and the stability of the protected area upstream (Figure 6). This motion is in principle the same as the one we proposed to explain the change in the protection pattern during the transition from the binary to the ternary complex.

The backreaction of RNA synthesis: a means to understand the forward reaction

A model of translocation must also include an understanding of the backreaction of RNA synthesis. A backreaction can be induced by addition of pyrophosphate to a stalled complex. A similar effect is shown with exonuclease III. There is one ternary complex [M9 **20**(C)], which can be forced to go through a backreaction by addition of exonuclease III, indicating a low stability of this complex. Yager and von Hippel (1987) presented a concept which describes movement of the transcription bubble as the surmounting of an activation energy barrier. This concept is an energy-balance consideration of all the processes contributing or requiring energy during one translocational cycle; these processes include opening and closing of DNA and DNA–RNA bp within the transcriptional bubble. Assuming that this concept is also valid for the backreaction, our finding supports the idea that sequences upstream as well as downstream to the transcription bubble ($-\overline{9}$ position) contribute to the stability of the complex; the sequence of the M9 **20**(C) complex, which is not stable against exonuclease III digestion, has an identical sequence upstream of the $-\overline{9}$ position to that of the M0 **20**(T) complex, which is stable against exonuclease III. This shows that the sequence upstream of the $-\overline{9}$ position is not sufficient to obtain a stable complex. This conclusion is supported by comparing the sequences of the unstable M9 **20**(C) complex with the stable M13 **24**(C) complex. These complexes have identical

sequences downstream of the −9 position, indicating that this downstream sequence is also not sufficient to stabilize the ternary complex. A more detailed comparison of the M9 **20**(C) sequence with the sequence of stable complexes points out the base positions responsible for the stability, but a final decision as to which bases are important requires a systematic substitution of the sequences by site-specific mutagenesis.

Current views on the translocation process

Of the different aspects of the translocation process recently reviewed (Yager and von Hippel, 1987; Ricchetti and Buc, 1988, personal communication), the topological and some related mechanistic aspects of translocation will be discussed here.

Dennis and Sylvester (1981) propose that the entire RNA polymerase rotates, or at least the two proposed topologically equivalent DNA binding sites rotate in order to facilitate translocation of RNA polymerase. However, our data suggest that RNA polymerase does not change its orientation with respect to the direction of transcription, contradicting the rotational model.

Gamper and Hearst (1982) propose that RNA polymerase translocates along DNA like a nut on a bolt. This model is in line with all data available to date, including ours, but it has not yet been confirmed experimentally. This model, although very attractive, does not incorporate the fact that translocation is a directed process including an ordered sequence of reaction steps, such as binding of nucleoside triphosphate, incorporation of the following base in the RNA chain and translocation of RNA polymerase. The latter step requires partial dissociation of RNA polymerase from the DNA and subsequent rebinding with intermediate translocation of RNA polymerase. Straney and Crothers (1987a,b) address this problem for the start of RNA synthesis during the escape of RNA polymerase from the abortive to the productive state. They propose translocation of RNA polymerase via a stressed intermediate state, i.e. translocation, by an inchworm-like motion, either of RNA polymerase or DNA. This attractive model seems not to apply for the translocation of RNA polymerase in the productive state. A transient complex corresponding to the stressed intermediate could not be detected in the productive state.

Transition from the abortive to the productive state is characterized by formation of a stable ternary complex, σ factor release, dissociation of the recognition domain of the DNA from the RNA polymerase and rifampicin resistance. All of these processes must be finally triggered by the nascent RNA chain. The precise sequence and the triggering parameter, the length of the RNA chain or the stability of the RNA−DNA hybrid, which is itself dependent on length and base composition of the RNA, remain to be determined.

Materials and methods

Preparation of RNA polymerase

RNA polymerase was prepared according to Zillig *et al.* (1970). Only the fraction containing core enzyme was used. σ factor was isolated from the overproducing strain M 5219/pMRG 8 using the method of Gribskow and Burgess (1983). RNA polymerase was then reconstituted from core enzyme and σ as previously described (Heumann *et al.*, 1988a).

Construction and preparation of DNA fragments

A 130 bp DNA fragment carrying the T7 A1 promoter was prepared according to Heumann *et al.* (1987).

The promoter mutations were constructed as follows: two aliquots of the 130 bp wild-type fragment were cut with *Taq*I and *Fok*I respectively. A set of chemically synthesized oligonucleotides was inserted between the *Taq*I site at position +2 and the *Fok* I site at position +12. These fragments were designed to fulfill the following conditions: (i) the mutations should be introduced as far downstream as possible to avoid effects on the promoter efficiency described by Kammerer *et al.* (1986); (ii) each mutation should open two transcription registers by either initiating with ApUpC, GTP and ATP or with ApUpC, GTP, ATP and CTP; and (iii) possible mutations should cover a sufficient range of registers yet there should be overlaps between registers created with ApUpC, GTP and ATP or with ApUpC, GTP, ATP and CTP.

The mutations were ligated between the purified *Bam*HI−*Taq*I and *Fok*I−*Bam*HI fragments. Each of these mutated A1+ sequences, containing (G,A) inserts of from 5 to 19 nucleotides between positions +11 and +12 in the sense strand of the wild-type sequence, was then ligated into the vector pDS/Tol+ (Stueber and Bujard, 1982). After transformation into *Escherichia coli*, strain WK6, positive clones were isolated on chloramphenicol agar plates. The plasmids of single clones were isolated and screened. The inserted sequences containing a *Mbo*II recognition site allowed quick screening. The sequences of the mutated fragments were confirmed by recloning the fragments into the plasmid pMC5-8 (Stanssens *et al.*, in preparation) and sequencing according to Sanger *et al.* (1977).

The fragments were purified according to Heumann *et al.* (1987). They were 5′ end labeled with [γ-^{32}P]ATP (Amersham) or 3′ end labeled with [α-^{32}P]dATP (Amersham) according to Maniatis *et al.* (1982).

Preparation of the starter trinucleotide ApUpC

ApUpC was formed in an abortive transcription assay (McClure *et al.*, 1978). Binary complex of 0.3 nmol A1 fragment and 0.45 nmol RNA polymerase was established according to Heumann *et al.* (1986). Then 150 nmol purified ApU (Sigma), 1 mmol CTP (ICN) and $MgCl_2$ (6 mM) were added. The assay was incubated at 37°C for 5 h, mixed 1:1 with formamide (containing xylene cyanol as tracking dye) and applied to a 20% polyacrylamide gel (12.5 mM Tris−HCl, 10.5 mM borate, 0.12 mM EDTA, pH 8.6). The gel band containing ApUpC was detected by UV-shadowing at 254 nm, cut out and the ApUpC eluted by shaking the gel band in H_2O overnight. The eluate was concentrated in a Speedvac vacuum centrifuge. Radioactively labeled ApUp*C was prepared following the same procedure, but using 20 mCi of [α-^{32}P]CTP (Amersham) instead of cold CTP. The ApUp*C was detected on X-ray film (3 M).

Preparation of binary and ternary complexes

Open (binary) complexes were formed according to Heumann *et al.* (1986). The concentration for labeled or unlabeled DNA fragments varied between 0.1 and 0.3 mg/assay; the RNA polymerase concentrations were 2−7 mg/assay. The volume of a standard assay was 20 μl. Ternary complexes were formed in 8 mM Tris−HCl, 6 mM $MgCl_2$, pH 7.9. A 3-fold molar excess of ApUpC compared to the DNA fragment and labeled ApUp*C (50−1000 Bq) was added. The concentration of the triphosphates was 2×10^{-5} M. The transcription assays were incubated at 37°C for 40 min and then further processed by exonuclease III or hydroxyl radical footprinting.

Exonuclease III digests and hydroxyl radical footprinting

Before digestion with exonuclease III a 20-fold molar excess of heparin was added to binary and ternary complexes to destroy unspecific complexes. The exonuclease III digests were carried out in the same buffer as the transcription assay by adding 50−200 units of exonuclease III (BRL). Digests were incubated for 40 min. All available batches of exonuclease III contained a 5′ label cutting activity, which could be reduced by addition of *E.coli* tRNA (0.5 mg/ml). This activity could be an endonucleolytic one, as suggested by the correlation between the strengths of the gel bands representing the protected DNA strands and their accessibility to the endonuclease.

Hydroxyl radical footprinting was performed as described by Tullius and Dombroski (1986).

Purification and analysis of RNA products and exonuclease III digested DNA on acrylamide gels

All complexes after digestion with exonuclease III or after treatment with hydroxyl radicals showed clear bands upon application to non-denaturing acrylamide gels, thus indicating their stability. Therefore all binary and ternary complexes except binary ones treated with hydroxyl radicals were purified on non-denaturing gels as described previously (Fried and Crothers, 1981; Garner and Revzin, 1981; Heumann *et al.*, 1986). The complex bands were cut out and eluted by shaking overnight in 20% phenol−8 mM Tris,

pH 7.0. After extraction with chloroform the water phase was dried in a Speedvac centrifuge and then dissolved in 50% formamide containing xylene cyanol. After boiling for 3 min the samples were applied to sequencing gels (Maxam and Gilbert, 1977) (7 M urea, 100 mM Tris–HCl, pH 8.6, 84 mM borate, 1 mM EDTA). The gels contained either 20% (to analyze the RNA products) or 8% acrylamide (to analyze the exonuclease III digested DNA bands). The 20% gels were run at 40 W; the 8% gels were kept at a constant temperature of 50°C and run at 50 W.

Acknowledgements

We thank Drs J.Trent, M.Horne and D.Grogan for valuable discussion, and Ms G.Baer and Ms E.Rumpf for expert technical assistance. This work was supported by the Deutsche Forschungsgemeinschaft.

References

Amouyal,M. and Buc,H. (1987) *J. Mol. Biol.*, **195**, 795–808.
Carpousis,A.J. and Gralla,J.D. (1980) *Biochemistry*, **19**, 3245–3253.
Carpousis,A.J. and Gralla,J.D. (1985) *J. Mol. Biol.*, **183**, 165–177.
Chamberlin,M.J .(1976) In Chamberlin,M.J. and Losick,R. (eds), *RNA Polymerase*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, pp. 17–67.
Dennis,D. and Sylvester,J.E. (1981) *FEBS Lett.*, **124**, 135–139.
Fried,M.G. and Crothers,D.M. (1981) *Nucleic Acids Res.*, **9**, 6505–6525.
Gamper,H.B. and Hearst,J.E. (1982) *Cell*, **29**, 81–90.
Garner,M. and Revzin,A. (1981) *Nucleic Acids Res.*, **9**, 3047–3060.
Gribskow,M. and Burgess,R.R. (1983) *Gene*, **26**, 109–118.
Hanna,M.M. and Meares,C.F. (1983) *Proc. Natl. Acad. Sci. USA*, **80**, 4238–4242.
Hansen,U.M. and McClure,W.R. (1980) *J. Biol. Chem.*, **255**, 9564–9570.
Heumann,H., Metzger,W. and Niehörster,M. (1986) *Eur. J. Biochem.*, **158**, 575–579.
Heumann,H., Lederer,H., Kammerer,W., Palm,P., Metzger,W. and Baer,G. (1987) *Biochim. Biophys. Acta*, **909**, 126–132.
Heumann,H., Lederer,H., Baer,G., May,R.P., Kjems,J.K. and Crespi,H.L. (1988a) *J. Mol. Biol.*, **201**, 115–125.
Heumann,H., Ricchetti,M. and Werel,W. (1988b) *EMBO J.*, **7**, 4379–4381.
Kammerer,W., Deuschle,U., Gentz,R. and Bujard,H. (1986) *EMBO J.*, **5**, 2995–3000.
Kanazawa,H., Kazunori,M. and Masamitsu,F. (1982) *Biochem. Biophys. Res. Commun.*, **107**, 568–575.
Kinsella,L., Chin,Y., Su,J., Schulz,W. and Dennis,D. (1982) *Biochemistry*, **21**, 2719–2723.
Kirkegaard,K., Buc,H., Spassky,A. and Wang,J.C. (1983) *Proc. Natl. Acad. Sci. USA*, **80**, 2544–2548.
Levin,J.R. and Chamberlin,M.J. (1987) *J. Mol. Biol.*, **196**, 61–84.
Levin,J.R., Krummel,B. and Chamberlin,M.J. (1987) *J. Mol. Biol.*, **196**, 85–100.
Maniatis,T., Fritsch,E.F. and Sambrock,J. (1982) In *Molecular Cloning*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
McClure,W.R,. Cech,C.L. and Johnston,D.E. (1978) *J. Biol. Chem.*, **253**, 8941–8948.
Ricchetti,M., Metzger,W. and Heumann,H. (1988) *Proc. Natl. Acad. Sci. USA*, **85**, 4610–4614.
Rozovskaya,T.A., Chenchik,A.A. and Bibilashvili,R.S. (1981) *Mol. Biol.*, **15**, 498–510 (English transl.).
Rozovskaya,T.A., Chenchik,A.A. and Beabealashvilli,R.S. (1982) *FEBS Lett.*, **137**, 100–104.
Sanger,F., Nicklen,S. and Coulson,R. (1977) *Proc. Natl. Acad. Sci. USA*, **74**, 5463–5467.
Schmitz,A. and Galas,J.D. (1979) *Nucleic Acids Rs.*, **6**, 111–137.
Shalloway,D., Kleinberger,T. and Livingston,D.M. (1980) *Cell*, **20**, 411–422.
Shi,Y.B., Gamper,H., Van Houten,B. and Hearst,J.E. (1988) *J. Mol. Biol.*, **199**, 277–293.
Siebenlist,U. (1979) *Nature*, **279**, 651–652.
Siebenlist,U., Simpson,R.B. and Gilbert,W. (1980) *Cell*, **20**, 269–281.
Spassky,A. (1986) *J. Mol. Biol.*, **188**, 99–103.
Straney,D.C. and Crothers,D.M. (1985) *Cell*, **43**, 449–459.
Straney,D.C. and Crothers,D.M. (1987a) *J. Mol. Biol.*, **193**, 267–278.
Straney,D.C. and Crothers,D.M. (1987b) *J. Mol. Biol.*, **193**, 279–292.
Stueber,D. and Bujard,H. (1982) *EMBO J.*, **1**, 1399–1404.
Travers,A.A. (1987) *CRC*, **22**, 181–219.
Travers,A.A,. Lamond,A.I., Mace,H.A.F. and Berman,M.L. (1983) *Cell*, **35**, 265–273.
Tullius,T.D. (1987) *Trends Biochem. Sci.*, **12**, 297–300.
Tullius,T.D. and Dombroski,B.A. (1986) *Proc. Natl. Acad. Sci. USA*, **83**, 5469–5473.
von Hippel,P.H., Bear,D.G., Winter,R.B. and Berg,O.G. (1982) In Rodriguez,R.L. and Chamberlin,M.J. (eds), *Promoters Structure and Function*. Praeger, New York, pp. 3–31.
Yager,T.D. and von Hippel,P.H. (1987) In Neidhardt,F.C., Ingraham,J.L., Brooks Low,K., Magasanik,B., Schaechter,M. and Umberger,H.E. (eds), *Escherichia coli and Salmonella typhimurium*. American Society for Microbiology, Washington, pp. 1241–1275.
Zillig,W., Zechel,K. and Halbwachs,H.J. (1970a) *Hoppe Seyler's Z. Physiol. Chem.*, **351**, 221–224.

Received on March 30, 1989; revised on May 23, 1989