The EMBO Journal Vol.16 No.24 pp.7468–7480, 1997

Three transitions in the RNA polymerase II transcription complex during initiation

Frank C.P.Holstege[1], Ulrike Fiedler and H.Th.Marc Timmers[2]

Laboratory for Physiological Chemistry, Utrecht University, PO Box 80042, 3508 TA Utrecht, The Netherlands

[1]Present address: Whitehead Institute for Biomedical Research, Massachusetts Institute of Technology, Nine Cambridge Center, Cambridge, MA 02142, USA

[2]Corresponding author
e-mail: H.T.M.TIMMERS@med.ruu.nl

We have analyzed transcription initiation by RNA polymerase II (pol II) in a highly efficient *in vitro* transcription system composed of essentially homogeneous protein preparations. The pol II complex was stalled on adenovirus major late promoter templates at defined positions, and the open region and RNA products of these complexes were examined. The first transition is formation of the open complex, which can be reversed by addition of ATPγS. The open region is no longer sensitive to ATPγS after formation of a four-nucleotide RNA, which constitutes the second transition. This indicates that the ATP-dependent DNA helicase activity of TFIIH is required to maintain the open region only during formation of the first three phosphodiester bonds. The downstream part of the transcription bubble expands in a continuous motion, but the initially opened region (–9/–2 on the non-template strand) recloses abruptly when transcription reaches register 11. This third transition is accompanied by a switch from abortive to productive RNA synthesis, which implies promoter clearance. Our findings provide a framework to analyze regulation of these specific transitions during transcription initiation by pol II.

Keywords: abortive transcription/DNA helicase/RNA polymerase II/TFIIH/transcription initiation

Introduction

Regulating the rate of transcription initiation by RNA polymerase is a major mechanism by which cells establish the proper expression patterns of their genes. Many proteins in eukaryotic cells are concerned with controlling the access of RNA polymerase II (pol II) to promoter regions of protein-encoding genes (for reviews, see Pirotta, 1995; Pazin and Kadonaga, 1997a,b). The establishment of *in vitro* transcription reactions and biochemical fractionations has led to the identification of the set of five basal transcription factors, which are minimally required to support accurate transcription initiation by pol II (for reviews, see Orphanides *et al.*, 1996; Roeder, 1996). Recent evidence indicates that pol II and the basal transcription factors are not floating around freely in the nucleus, but that they are pre-assembled into pol II holoenzyme complex(es) together with many other proteins (for reviews, see Koleske and Young, 1995; Halle and Meisterernst, 1996). The composition of these large complexes (>2000 kDa) differs depending on the isolation procedure (Kim *et al.*, 1994; Koleske and Young, 1994; Ossipow *et al.*, 1995; Chao *et al.*, 1996; Maldonado *et al.*, 1996), and this reflects a tendency to dissociate during purification. Analyses of the different holoenzyme preparations show in all cases tested that the complete set of five basal transcription factors (TBP, TFIIB, TFIIE, TFIIF and TFIIH) need to be present to support *in vitro* transcription by pol II (Kim *et al.*, 1994; Koleske and Young, 1994; Chao *et al.*, 1996; Maldonado *et al.*, 1996). This implies that the basal transcription factors perform unique roles in transcription initiation by pol II.

Biochemical analyses have established the pathway of formation of the pol II pre-initiation complex from its isolated components (for reviews, see Orphanides *et al.*, 1996; Roeder, 1996). In order to recognize promoter DNA, pol II requires the assistance of TBP (TATA-binding protein), TFIIB and TFIIF. In cellular extracts, TBP is found in complexes with other proteins. These TBP-associated factors (TAFs) play a role in transcriptional activation pathways and polymerase specificity, but for the basal pol II transcription reaction TBP alone suffices (for reviews, see Burley and Roeder, 1996; Verrijzer and Tjian, 1996). The TBP–TATA complex allows association of TFIIB, and together this forms the platform for promoter recognition by the TFIIF–pol II complex. Topological stress (Parvin and Sharp, 1993; Tyree *et al.*, 1993; Goodrich and Tjian, 1994; Holstege *et al.*, 1995) or a premelted region in promoter DNA (Pan and Greenblatt, 1994; Tantin and Carey, 1994; Holstege *et al.*, 1996) are required to observe transcription initiation with this so-called DBpolF complex, indicating that it has a very limited capacity to melt the DNA strands to allow initiation.

The basal transcription factors TFIIE and TFIIH are not essential for promoter recognition but are involved in promoter opening. TFIIE can stimulate transcription from supercoiled templates in the absence of TFIIH, and this is linked to the helical stability of promoter DNA (Holstege *et al.*, 1995). On relaxed DNA templates, TFIIE is essential for TFIIH recruitment to the pre-initiation complex (Flores *et al.*, 1992). Egly and co-workers found that two of the nine TFIIH subunits are ATP-dependent DNA helicases (Schaeffer *et al.*, 1993, 1994). It was shown in activated transcription assays by Gralla and co-workers that ATP hydrolysis is linked directly to promoter melting as determined by potassium permanganate sensitivity assays, which map single-stranded regions in DNA (Wang *et al.*, 1992). Using this assay and a highly purified basal transcription system, we obtained direct evidence that promoter opening requires the action of the ATP-dependent

DNA helicase(s) of TFIIH (Holstege *et al.*, 1996). With the adenovirus major late promoter (AdMLP), it was shown that a region predominantly upstream of the start site (–9/+2) becomes single stranded (Holstege *et al.*, 1996; F.C.P.Holstege and H.T.M.Timmers, in preparation). The formation of the –9/+2 open complex occurs prior to formation of the first phosphodiester bond, and the ATP dependence of this reaction agrees well with earlier studies showing that ATP hydrolysis is required at an early phase of transcription initiation (Sawadogo and Roeder, 1984; Luse and Jacob, 1987; Luse *et al.*, 1987; Conaway and Conaway, 1988). Further expansion of the transcription 'bubble' requires transcription initiation (Wang *et al.*, 1992; Holstege *et al.*, 1996).

The molecular events between open complex formation and promoter clearance by pol II have not been determined in precise detail. The strong protein–protein interactions which are responsible for assembly of the pre-initiation complex need to be broken during the promoter clearance step. This step also marks the transition to the productive elongation mode of the transcription complex. In an elegant set of experiments, Reinberg and co-workers determined the positions of basal factor release from the initiation complex (Zawel *et al.*, 1995). The role of ATP hydrolysis in promoter clearance is a controversial issue. Besides its ATP-dependent DNA helicase activity, TFIIH harbors a protein kinase specific for the C-terminal domain (CTD) of the largest subunit of pol II (for review, see Svejstrup *et al.*, 1996). Stimulation of promoter clearance by the CTD kinase of TFIIH may be promoter specific (Serizawa *et al.*, 1993; Akoulitchev *et al.*, 1995; Mäkelä *et al.*, 1995; Jiang *et al.*, 1996) and depends on the composition of the transcription system (Serizawa *et al.*, 1993; Li and Kornberg, 1994). Other studies have implicated the DNA helicase activity of TFIIH in promoter clearance (Goodrich and Tjian, 1994; Dvir *et al.*, 1996b).

In this study, we have used a highly purified and highly efficient *in vitro* system to examine the early phase of transcription initiation by pol II. DNA templates under control of the AdMLP were modified to stall pol II at defined positions. We have coupled the potassium permanganate sensitivity analysis of open regions in the DNA template to a direct analysis of the synthesized RNA products. This coupling allowed a careful description and provides a detailed model for the transcription initiation reaction by pol II.

Results

Our previous analyses of open complex formation indicate that prior to transcription initiation, the DNA strands of the AdMLP become single stranded from position –9 to +1 by the action of the ATP-dependent DNA helicase activity of TFIIH (Holstege *et al.*, 1996). Additional experiments showed that position +2 is also susceptible to permanganate (F.C.P.Holstege and H.T.M.Timmers, in preparation) and, therefore, the initially opened region of the AdMLP encompasses positions –9 to +2. Concomitant with formation of the first phosphodiester bond, we observed a large expansion of the open region to position +8, and we named this expansion the second step in promoter opening (Holstege *et al.*, 1996). When we analyzed opening of other pol II promoters (HIV, AdE4 and MMTV), expansions beyond the site of RNA synthesis were not observed (F.C.P.Holstege and H.T.M.Timmers, in preparation). In addition, the large expansion observed with the AdMLP was formed slowly and was increased by relatively high concentrations of CTP (>20 μM; F.C.P.Holstege and H.T.M.Timmers, unpublished data). Read-through has been observed during initiation of transcription by bacterial RNA polymerase (Carpousis and Gralla, 1980). Although other explanations are possible, we wondered whether the large expansion observed with the AdMLP is caused by pol II molecules transcribing beyond register 2 [we refer to register as the site of addition of the last nucleoside monophosphate (NMP) to the nascent RNA chain and we refer to position as the base pair of the DNA template relative to the transcription start site]. Bacterial RNA polymerase in the abortive transcription mode has a high rate of misincorporation (Metzger *et al.*, 1993). To investigate this hypothesis, we used the RNA chain terminator 3′-*O*-methyl-GTP (3′-OMeGTP) to block read-through. In addition, to stall pol II in specific registers, we constructed a series of AdMLP templates with G residues at varying positions (Figure 1).

Constructs to stall RNA polymerase II

AdML+2G : 5'-CCTC**AG**TCTCTTCCTCTAGAGTC

AdML+3G : 5'-CCTC**A**C**G**CTCTTCCTCTAGAGTC

AdML+4G : 5'-CCTC**A**CT**G**TCTTCCTCTAGAGTC

AdML+6G : 5'-CCTC**A**CTCT**G**TTCCTCTAGAGTC

AdML+9G : 5'-CCTC**A**CTCTCTT**G**CTCTAGAGTC

AdML+10G: 5'-CCTC**A**CTCTCTTC**G**TCTAGAGTC

AdML+11G: 5'-CCTC**A**CTCTCTTCC**G**CTAGAGTC

AdML+15G: 5'-CCTC**A**CTCTCTTCCTCTA**G**AGTC

Fig. 1. DNA sequence of the non-template strand of the pDNAdML+nG plasmids. The sequence of the transcription start and stall site of the different AdMLP templates used in this study is depicted. Indicated in bold are the natural transcription start site (A) and the first G residue in the RNA chain. The templates are identical outside the region shown.

Mutant AdMLP templates were labeled on the non-template strand and used in potassium permanganate sensitivity assays as described earlier (Holstege *et al.*, 1996; Holstege and Timmers, 1997). Briefly, initiation complexes were formed on AdMLP fragments by pre-incubation with homogenous preparations of TBP, TFIIB, TFIIE and TFIIF, which were expressed in bacteria or in insect cells with recombinant baculoviruses (TFIIF) (see also Holstege *et al.*, 1996). The pol II preparation was obtained from calf thymus extracts by a combination of ion-exchange and immunoaffinity chromatography. The TFIIH preparation was isolated from HeLa cells by an optimized protocol involving six purification steps. All protein preparations were free of contamination by other basal factor or pol II activities, as shown by leave-out experiments, and were devoid of nucleases, topoisomerases or co-factors like PC4 as determined by a number of assays (see Materials and methods and Holstege *et al.*,

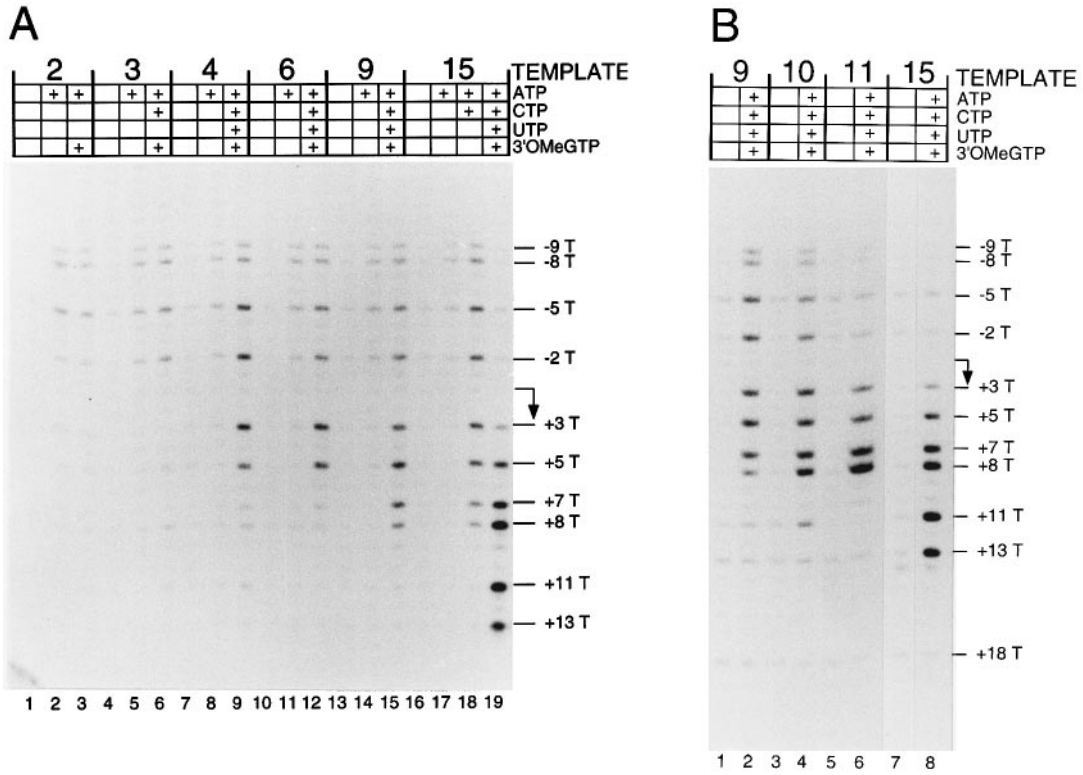

Fig. 2. The transcription bubble expands downstream continuously and recloses upstream discontinuously. (**A**) The results of permanganate sensitivity assays with pol II stalled at different positions. Transcription initiation complexes were assembled with AdML+2G (lanes 1–3), AdML+3G (lanes 4–6), AdML+4G (lanes 7–9), AdML+6G (lanes 10–12), AdML+9G (lanes 13–15) and AdML+15G (lanes 16–19) promoter fragments as described in Materials and methods. Reactions received either no nucleotides as a control (lanes 1, 4, 7, 10, 13 and 16), 60 μM ATP (lanes 2, 5, 8, 11, 14 and 17), 60 μM ATP and 120 μM 3′-OMeGTP (lane 3), 60 μM ATP, 10 μM CTP and 120 μM 3′-OMeGTP (lane 6), 60 μM ATP, 10 μM CTP, 10 μM UTP and 120 μM 3′-OMeGTP (lanes 9, 12, 15 and 19) or 60 μM ATP and 10 μM CTP (lane 18). After a 5 min incubation at 30°C, potassium permanganate was added for 90 s and the reaction was processed as described in Materials and methods. Positions of reactive thymidines are determined by co-migration of a G+A ladder of the AdML+11G fragment (data not shown) and are indicated to the right. The arrow marks the start site and direction of transcription. It should be noted that due to the T→G mutation no reactivity at the +3 position can be expected with the AdML+3G template in lane 6. (**B**) The initially melted open region (positions –9 to –2) recloses during translocation of pol II from register 9 to register 11. The open regions of transcription complexes stalled at the +9 (lane 2), +10 (lane 4), +11 (lane 6) and +15 (lane 8) positions were determined by permanganate sensitivity assays as described in (A). Reactions were reconstituted using AdML+9G (lanes 1 and 2), AdML+10G (lanes 3 and 4), AdML+11G (lanes 5 and 6) or AdML+15G (lanes 7 and 8) promoter fragments. Reactions received either no nucleotides (lanes 1, 3, 5 and 7) or 60 μM ATP, 10 μM CTP, 10 μM UTP and 120 μM 3′-OMeGTP (lanes 2, 4, 6 and 8). After a 5 min incubation at 30°C, potassium permanganate was added for 90 s and the reaction was processed as described in Materials and methods. The positions of reactive thymidines as determined with a co-migrating G+A ladder (data not shown) are indicated to the right. The arrow marks the start site and direction of transcription. It should be noted that due to the T→G mutation, no reactivity at the +11 position can be expected with the AdML+11G template in lane 6.

1996). The amounts of DNA template, basal factors and pol II used in the permanganate assay were titrated carefully in gel shift experiments to reach saturation of the template fragment. After pre-initiation complex formation, transcription was initiated by the addition of different combinations of ribonucleotides. In these assays, the efficiency of template usage for productive transcription is 25% at 4 fmol of template (see below).

The downstream part of the transcription bubble expands continuously whereas the upstream part recloses discontinuously

Figure 2 shows the results of the permanganate analysis using different mutant AdMLP templates, with the position of the first G residue indicated above the lanes. In all cases, addition of ATP to the reaction leads to increased sensitivity to permanganate of positions –9, –8, –5 and –2 (Figure 2A, lanes 2, 5, 8, 11, 14 and 17). Inclusion of 3′-OMeGTP with the AdML+2G fragment, which allows formation of only one phosphodiester bond, does not result in a downstream expansion of the open region (Figure 2A, compare lanes 2 and 3) as would be expected from our earlier study (Holstege *et al.*, 1996). When pol II proceeds to register 3, a small increase in the reactivity of the –2T is observed (Figure 2A, compare lanes 5 and 6). Stalling pol II in register 4 further increases the sensitivity of position –5 and –2, and new sensitivities at +3 and +5 are induced (Figure 2A, lane 9). No further expansions are observed with AdML+6G (Figure 2A, lane 12) but, as expected, with AdML+9G positions +7 and +8 become sensitive to permanganate (Figure 2A, lane 15). As a control, we included the AdMLP fragment with the first G residue at +15 in this experiment (Figure 2A, lane 16–19). In agreement with our earlier study (Holstege *et al.*, 1996), addition of ATP and CTP results in the large expansion to position +8 (Figure 2A, lane 18). This is well beyond the expected stall site in register 3, which specifies addition of a UMP residue (see Figure 1), and strongly suggests that our previously reported expansion upon formation of the first phosphodiester bond

is due to read-through by the pol II enzyme (see also below). Inclusion of ATP, CTP, UTP and 3′-OMeGTP in the AdML+15G reaction results in stalling of pol II at register 15 (Figure 2A, lane 19). This shows that the initially opened region from position –9 to –2 is no longer sensitive to permanganate and that the open region now encompasses position +3 to +13. These results show that in the pol II initiation complex, downstream melting of the template DNA keeps abreast with RNA synthesis. The downstream edge of the open region is within 1 bp from the register of RNA synthesis.

In the analysis of Figure 2A, the initially opened region (–9/–2), which is melted by the action of the TFIIH DNA helicase(s), recloses when pol II transcribes from register 9 to 15. To map the reclosure more precisely, two additional AdMLP constructs were analyzed. Figure 2B (compare lanes 2 and 6) shows that positions –9, –8, –5 and –2 are not reactive to permanganate when pol II is stalled in register 11, and this is similar to the reaction with the AdML+15G template (Figure 2B, lane 8). The AdML+10G fragment represents an intermediate situation (Figure 2B, compare lanes 2, 4 and 6). These results indicate that between register 9 and 11, the pol II initiation complex undergoes a conformational change, which results in reclosure of the –9/–2 region. Thus, while the downstream part of the open region expands in a continuous motion, the upstream part of the open region readopts the double-stranded conformation discontinuously.

Analysis of RNAs formed by stalled transcription complexes

To investigate whether pol II is stalled in the expected register, we analyzed the RNA products formed under conditions of the permanganate sensitivity assay. Initiation complexes were assembled on different AdMLP templates as previously, and transcription was initiated by the addition of 60 μM ATP, 10 μM [α-^{32}P]CTP, 10 μM UTP and 120 μM 3′-OMeGTP as indicated above the lanes. After 5 min of incubation, the reactions were stopped and analyzed. Figure 3A shows that in the absence of the RNA chain terminator, 3′-OMeGTP read-through products are observed with the AdML+4G, +6G, +9G, +10G, +11G and +15G templates (lanes 4–9). Longer exposure reveals the presence of read-through products in lanes 2 and 3 (data not shown). In several cases, the read-through products are one nucleoside longer than the expected product (4 nt in lane 4, 10 nt in lane 7, 11 nt in lane 8 and 15 nt in lane 9), which may indicate that these RNAs are elongated inefficiently due to mispairing of the last residue with the DNA template. The inclusion of 3′-OMeGTP in the reactions results in the production of a single RNA product in the cases of the AdML+10G, +11G and +15G templates (Figure 3A, lanes 14–16). While the expected RNA is the major species with the AdML+6G and +9G templates, products one nucleoside shorter are also observed (Figure 3A, lanes 12 and 13). With the AdML+3G and +4G templates, the expected product is formed with an efficiency similar to the one-nucleotide shorter product (Figure 3A, lanes 10 and 11). This suggests an inefficient incorporation of 3′-OMeGTP in these cases, which may lead to release of the nascent RNA from the pol II complex. For bacterial RNA polymerase, it is known that efficiency of NMP addition to a growing RNA chain is dependent on the specific NMP and on the sequence of the template (Levin and Chamberlin, 1987). Inclusion of 2 μg/ml α-amanitin in the reaction completely inhibits RNA synthesis, which verifies that RNAs are formed by pol II (data not shown). From the analysis of Figure 3A, we conclude that addition of 3′-OMeGTP to the reactions results in a more homogenous population of RNA products.

To investigate further the phenomenon of read-through by pol II, we tested the effects of CTP concentration and of incubation time on the formation of RNA products. Figure 3B shows that multiple RNA products can be detected when the CTP concentration is increased to 30 μM (lanes 1–5). Synthesis of such products is also dependent on a prolonged incubation. In addition, in permanganate sensitivity assays, we observed that increasing the CTP concentration in the assay results in increased sensitivities in position +3 to +8, whereas the –9/–2 region is not affected (data not shown). Together, these observations strongly support the hypothesis that expansion of the open region to position +8 in the presence of nucleotides, which should allow only formation of the first phosphodiester bond, is caused by complexes transcribing beyond register 2.

Whereas read-through products are detected easily in its absence, the RNA chain terminator suppresses the formation of RNAs longer than expected. The conclusion of pol II read-through has important implications for the interpretation of earlier studies (including our own) in which it was assumed that nucleotide deprivation stalled the pol II enzyme at a particular register. Even in the presence of a chain terminator, it is important to titrate carefully the concentration of ribonucleotides in the reaction. The RNA products which are shorter than expected may result from abortive transcription. Abortive RNAs are released rapidly from a transcription complex and their formation should not complicate the permanganate analysis.

Abortive RNA synthesis by pol II stops in register 11

The possibility of abortive RNA synthesis is supported by the observation in Figure 3A that shorter products are formed in much higher amounts than longer products. It should be noted that the 15 nt product contains six labeled CMP residues, whereas only one CMP is present in the 2 nt product of the AdML+3G reaction (see Figure 1). The primary criterion for abortive RNA synthesis is the accumulation of excess product over the amount of template present in the reaction (Jacob *et al.*, 1994; Gralla, 1996). To investigate whether formation of short RNAs fulfilled this criterion, a time course experiment was performed to measure formation of RNA products from the various AdMLP templates. Figure 4A shows the accumulation of products using 20 fmol of the AdML+3G or +15G templates in the reaction. Whereas the 15G product reaches a plateau very early after nucleotide addition (Figure 4A, lanes 9–15), the amount of both the 3G and 2 products increases linearly over the time of the assay. Product formation on the other AdMLP templates was analyzed similarly, and Figure 4B shows a graphic representation of this analysis. We found that after 1 min, the 11G and 15G products reach 70% of their maximal

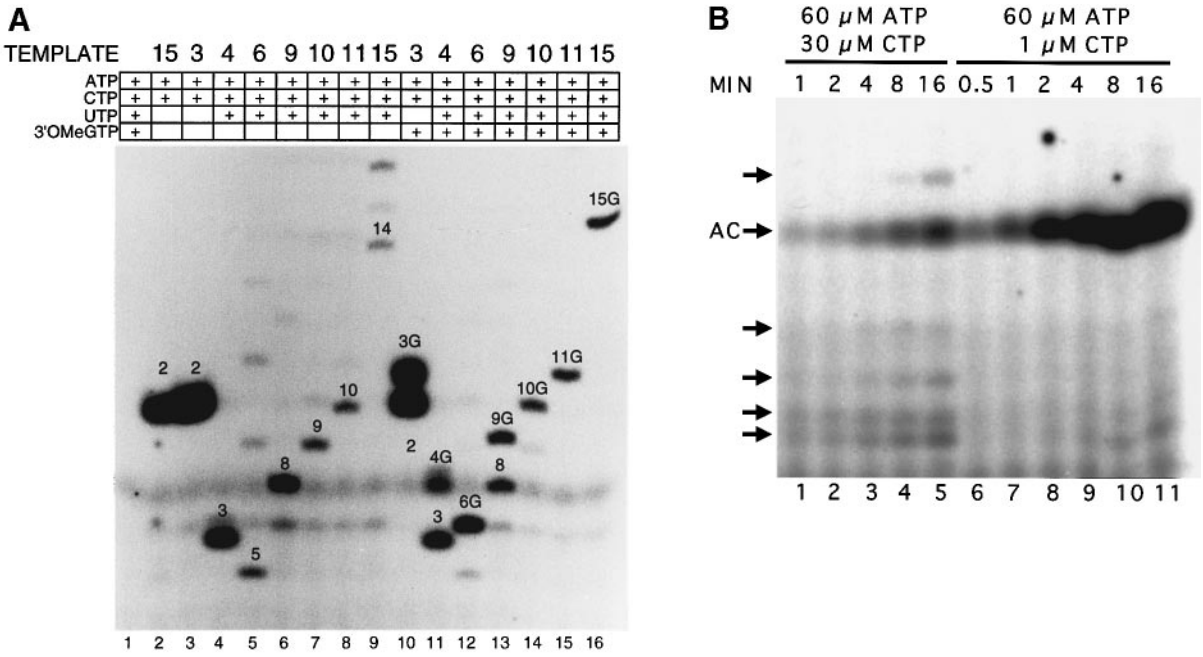

Fig. 3. Analysis of RNA products formed with mutant AdMLP templates. (**A**) Initiation complexes were assembled on mutant AdMLP templates (as indicated at the top of the figure) similarly to in the permanganate sensitivity assays. Nucleotide triphosphates were added, as indicated above the lanes, for 5 min. Subsequently, RNA products were processed and analyzed by denaturing polyacrylamide gel electrophoresis as described in Materials and methods. Reactions received either no promoter DNA fragment (lane 1), AdML+15G (lanes 2, 9 and 16), AdML+3G (lanes 3 and 10), AdML+4G (lanes 4 and 11), AdML+6G (lanes 5 and 12), AdML+9G (lanes 6 and 13), AdML+10G (lanes 7 and 14) or AdML+11G promoter fragment (lanes 8 and 15). All reactions were reconstituted with 20 fmol of DNA template and initiated with 60 μM ATP and 10 μM [α-^{32}P]CTP. In addition, reactions also received 10 μM UTP (lanes 4–9), 120 μM 3′-OMeGTP (lane 10) or 10 μM UTP and 120 μM 3′-OMeGTP (lanes 1 and 11–16). The size of each product (in nucleosides) is indicated by the number above the band, except for the two-nucleoside product in lane 10 for which the number is indicated below the band. Products which contain 3′-O-MeGMP are indicated by the letter G. Therefore, the product indicated by the number '3′ in lane 4 is ApCpU, whereas the product in lane 10 indicated by '3G' is ApCpG-3′-*O*-methyl. All reactions were treated with alkaline phosphatase to remove phosphate groups at the 5′-end of the RNAs. Besides reducing the background reactivity in each lane, this significantly decreases the migration rate of the smaller (five nucleosides or less) products. For these products, the migration rate initially increases with size. Substitution of a pyrimidine with guanosine results in a significant decrease in the migration rate (compare '3′ with '3G' in lanes 4 and 10). From RNA products longer than five nucleosides, the relative migration rate is as expected. The positions of the smaller products 2, 3G and 4G were verified by UV shadowing of co-migrating unlabeled marker RNA (see Materials and methods). The position of the three-residue product (lanes 4 and 11) is identical to the position of CpApC formed by the dinucleotide-primed reaction (Goodrich and Tjian, 1994; Holstege *et al.*, 1996). (**B**) Initiation complexes were assembled on the AdML+15G promoter as in (A). The reaction analyzed in lanes 1–5 received 60 μM ATP and 30 μM [α-^{32}P]CTP, and reactions of lanes 6–11 received 60 μM ATP and 1 μM [α-^{32}P]CTP. It should be noted that the specific activity of [α-^{32}P]CTP is 30 times higher in lanes 6–11 than in lanes 1–5. Reactions were stopped 0.5 (lane 6), 1 (lanes 1 and 7), 2 (lanes 2 and 8), 4 (lanes 3 and 9), 8 (lanes 4 and 10) or 16 min (lanes 5 and 11) after the addition of nucleotides. The position of the ApC product is indicated on the left. The arrows indicate the larger RNA products.

levels. After 4 min of incubation, no further increase is observed. In contrast, products shorter than 11 nt show a continuous accumulation over time. This is observed most dramatically with the 3G, 4G and 6G products, but the amount of 9G product also shows a linear increase with time. The 10G product represents an intermediate situation. At 15 min, 50% more 10G RNA is formed than 15G product. Quantitation of the products of Figure 4A showed that after 15 min ~20 fmol of RNA (2 and 3G RNA) was formed in the AdML+3G reaction, which contained 20 fmol of template.

In order to show conclusively that abortive transcription is involved in the synthesis of short RNA products, we tested whether less DNA template could be used in the assays. It was found that the minimal amount of AdML+15G template to sustain the level of 15G formation is 4 fmol of DNA fragment in the reaction (data not shown). The experiment in Figure 4B was repeated using 4 fmol of the different AdMLP templates to determine the amount of RNA product formed in 15 min reactions. The AdML+2G template could not be tested since the assay would require [α-^{32}P]3′-OMeGTP, which is not commercially available. Therefore, formation of the 2 nt product was measured by addition of ATP and CTP to the AdML+15G complex. Figure 5 shows that this results in formation of 2.2 molecules of RNA per DNA molecule. The AdML+3G reaction yields four RNA molecules (2 and 3G products) per template molecule. Formation of the 4G and 6G products is also catalytic, whereas one molecule of 9G RNA is produced per AdML+9G template molecule. In contrast, the AdML+11G and +15G templates yield 0.25 molecules of 11G and 15G RNA per DNA molecule, respectively. It is important to note that this analysis indicates that 25% of the templates is active in the 4 fmol transcription assays.

Taken together, these experiments show that pol II is capable of abortive synthesis of RNA molecules. The efficiency of abortive transcription is the highest with short RNAs (2 and 3 nt products), but can still be observed with products up to 10 nt in length. RNAs of 11 nt or longer cannot be synthesized in an abortive fashion. This indicates that after register 11 the pol II complex escapes the abortive cycle of RNA synthesis and is only capable of productive RNA synthesis.

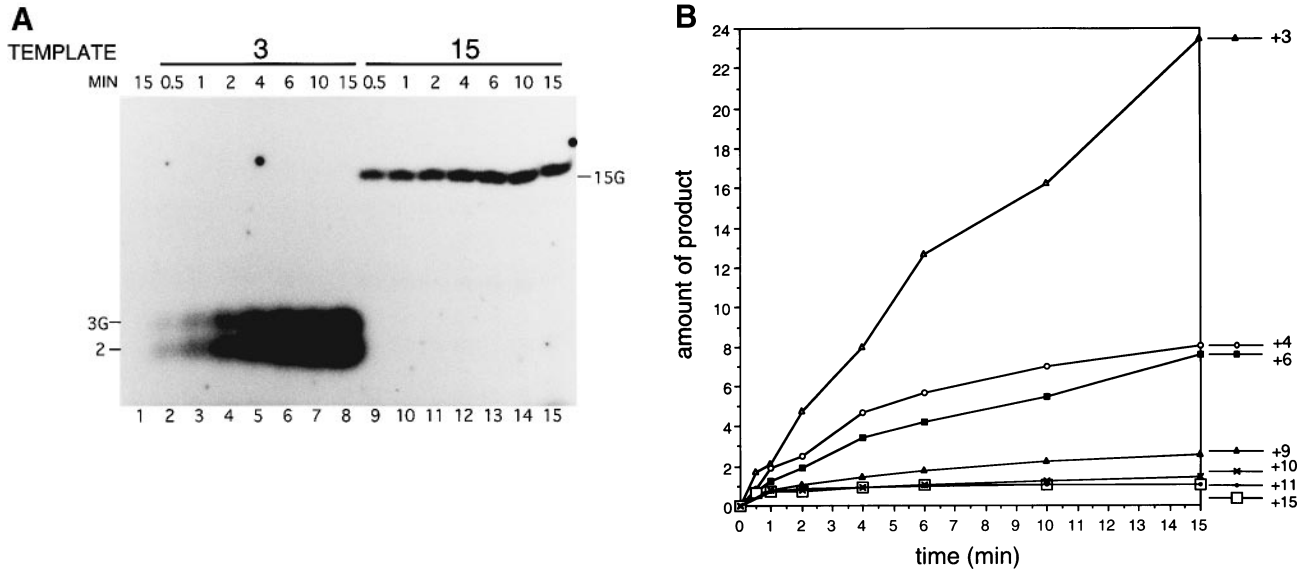

Fig. 4. Time course analysis of product formation with the different AdMLP templates. (**A**) Reactions were assembled as described in the legend of Figure 3A using 20 fmol of the AdML+3G (lanes 2–8) or AdML+15G promoter fragments (lanes 9–15). The reaction analyzed in lane 1 received no template DNA. Reactions received 60 μM ATP, 10 μM [α-^{32}P]CTP and 120 μM 3′-OMeGTP (lanes 2–8), or 60 μM ATP, 10 μM [α-^{32}P]CTP, 10 μM UTP and 120 μM 3′-OMeGTP (lanes 1 and 9–15). Reactions were stopped after 0.5 (lanes 2 and 9), 1 (lanes 3 and 10), 2 (lanes 4 and 11), 4 (lanes 5 and 12), 6 (lanes 6 and 13), 10 (lanes 7 and 14) and 15 min (lanes 1, 8 and 15) and RNA products were analyzed as described in Materials and methods. The positions of the products are indicated. (**B**) This figure is a graphic representation of time course experiments analyzing product formation with different AdMLP templates. Reactions were assembled and processed as described in the legend of Figure 2A. To the right, the AdMLP templates used are indicated. Only single bands were used for quantitation by phosphorimager analysis (except for the AdML+3G fragment which yielded the 2 and 3G bands). The values plotted on the *y*-axis were derived by dividing the results of each promoter by the number of cytosines incorporated in the RNA product. The values are relative to the amount of product from the AdML+15G fragment at 6 min. The relative amounts of 3G, 4G, 6G, 9G, 10G and 11G product after 15 min are 23.5, 8.0, 7.6, 2.5, 1.5 and 1.0, respectively.

Another characteristic of abortive transcription is that an abortively synthesized RNA is released from the transcription complex (Carpousis and Gralla, 1980). This implies that pol II in the abortive mode should be able to remove a 'dead-end' product like a 3G RNA from the transcription complex and can resume another cycle of transcription initiation. In contrast, pol II complexes in the productive elongation mode are unable to remove the 'dead-end' product. Therefore, productive pol II complexes cannot resume transcription. To test this, the following experimental approach was used. Initiation complexes were assembled on the different AdMLP templates and transcription was allowed in the presence of 3′-OMeGTP for 10 min. Subsequently, a 10-fold molar excess of GTP was added to compete for the 3′-OMeGTP chain terminator and transcription was allowed to proceed for 2 min. The RNA products were analyzed by a primer extension assay, which detects products of 63 nt and longer. Figure 6 shows that 3′-OMeGTP pre-incubation of complexes formed on the AdML+11G and +15G fragments blocks their ability to synthesize long RNA (lanes 13 and 15). Although inclusion of 3′-OMeGTP reduces the amount of synthesized RNA 2.5-fold with AdML+9G complexes, correctly initiated RNA can still be detected (Figure 6, lanes 9 and 10). As expected, the AdML+10G (Figure 6, lanes 11 and 12) represents an intermediate situation, and inclusion of 3′-OMeGTP inhibits 10-fold. In contrast, transcription from complexes formed with the AdML+2G, +3G, +4G and +6G fragments is not inhibited by 3′-OMeGTP in the pre-incubation (Figure 6, lanes 1 and 8). Interestingly, not only correctly initiated RNAs but also RNAs of shorter lengths are detected. Comparison with the co-migrated

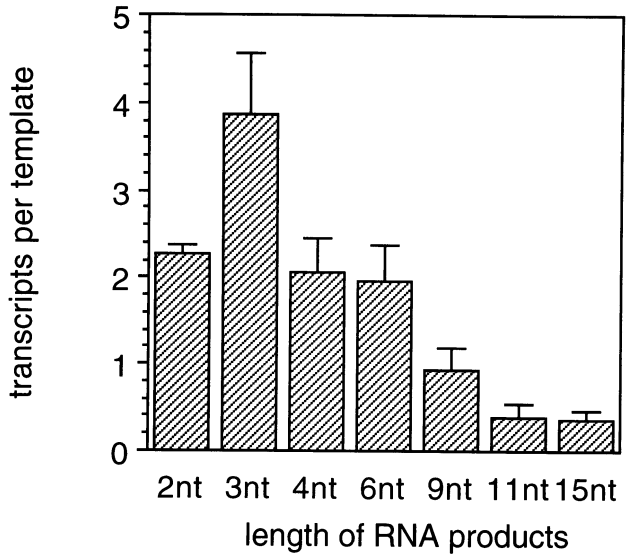

Fig. 5. Quantification of RNA products using a limiting amount of template. Reactions were assembled and processed as described in the legend of Figure 3A. The reactions received 4 fmol of template. The templates used were: AdML+15G for the 2 nt, AdML+3G for the 3 nt (which combines the amounts of the 2 and 3G products), AdML+4G for the 4 nt, AdML+6G for the 6 nt, AdML+9G for the 9 nt, AdML+11G for the 11 nt and AdML+15G for the 15 nt product. RNA products were quantified as described in the legend of Figure 4A. The bars indicate the number of RNA molecules formed per molecule of DNA template. Average values of three independent experiments are plotted.

DNA sequence shows that in each case the end of the shorter RNAs maps to the position of the introduced G residue. This suggests the following situation. When, for

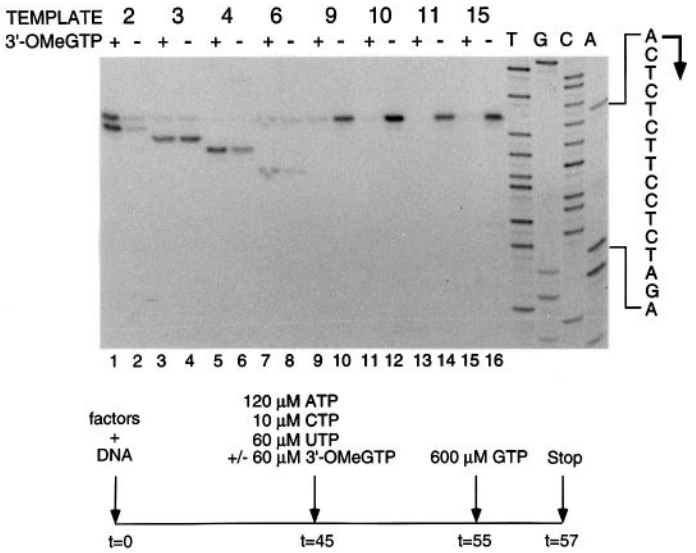

Fig. 6. Reinitiation of transcription by stalled transcription complexes. Transcription complexes were assembled on AdML+2G (lanes 1 and 2), AdML+3G (lanes 3 and 4), AdML+4G (lanes 5 and 6), AdML+6G (lanes 7 and 8), AdML+9G (lanes 9 and 10), AdML+10G (lanes 11 and 12), AdML+11G (lanes 13 and 14) or AdML+15G promoter fragments (lanes 15 and 16). Transcription was started by adding 120 μM ATP, 10 μM CTP and 60 μM UTP. As indicated, 60 μM 3′-OMeGTP was included in the reactions of the odd-numbered lanes and was omitted in the reactions of the even-numbered lanes. After 10 min, 600 μM GTP was added and the incubation was continued for 2 min. The reactions were stopped, processed and analyzed by primer extension as described in Materials and methods. Sequencing reactions using the AdML template and the oligonucleotide used for primer extension were co-migrated on the lanes labeled T, G, C and A. To the right is indicated a part of the AdML sequence and an arrow depicting the transcriptional start site.

example, the pol II complex is stalled in register 4, it can remove the 4G RNA and reinitiate transcription. It now has two options: (i) to start transcription at the normal start site or (ii) to start at register 4. The analysis of Figure 6 indicates that reinitiation in register 9 is not very favorable, since only RNAs initiated at position +1 are detected. One would expect that the concentrations of the nucleotides used for initiation are important for the choice between start sites. Indeed, it is observed that reducing the ATP concentration to 60 μM stimulates the production of shorter RNAs (data not shown). The converse experiment bears the complication that a further increase in ATP concentration will favor read-through during preincubation. In this experimental setup, competition for 3′-OMeGTP requires high concentrations of GTP (600 μM), which favor transcription initiation with a G residue. In conclusion, the experiment of Figure 6 shows that until register 10, transcription complexes can initiate a new cycle of transcription, which is the hallmark of abortive transcription. When transcription complexes reach register 11, this ability to reinitiate transcription is lost.

Stability of open complexes

Conaway and Conaway reported that ATPγS is a reversible inhibitor of transcription and acts at an early step in the initiation reaction (Conaway and Conaway, 1988; Dvir *et al.*, 1996b). We have described that this compound is unable to support promoter opening by the TFIIH DNA helicases (Holstege *et al.*, 1996). This suggests that ATPγS blocks the DNA helicase activity of TFIIH and can be used to assay the stability of open complexes. This was investigated by permanganate sensitivity assays of transcription complexes stalled in different registers. After formation of these complexes, ATPγS was added and the reduction in permanganate sensitivity was followed with time. Figure 7 shows this experiment for the initially opened region (Figure 7A) and for the AdML+4G complex (Figure 7B). The analysis of Figure 7 shows that the –9/–2 opening is rapidly reversed by the addition of ATPγS ($t_{1/2}$ = 45 s), whereas the open region in the AdML+4G complex is insensitive to ATPγS ($t_{1/2}$ > 10 min). Similarly, the ATPγS sensitivity of the AdML+2G and AdML+3G complexes was analyzed. The half-life of the AdML+2G complex is 90 s, whereas the AdML+3G complex has a half-life of 3–4 min (data not shown). These results indicate that the ATP-dependent activity of the DNA helicase(s) of TFIIH is no longer required when the pol II complex has reached register 4. These data are in agreement with the observations that ATPγS acts at an early step (Conaway and Conaway, 1988).

Discussion

In highly efficient transcription initiation assays we have stalled pol II transcription complexes in different registers. The analysis of the open regions in these stalled complexes was coupled with the analysis of formation of RNA products. Previously, we assumed that addition of ATP and CTP resulted in stalling of pol II in register 2 and reported that formation of the first phosphodiester bond is concomitant with an expansion to position +8 (Holstege *et al.*, 1996). Our current analysis indicates that the observed expansion was caused by pol II read-through, which can be suppressed by addition of an RNA chain terminator. Figure 2 shows that the downstream part of the open region expands in a continous motion within 1 bp of the register of NMP addition. Under the conditions of the permanganate assay, product formation was analyzed, and this indicated that pol II complexes were indeed stalled at the expected registers (Figure 3A). For RNA

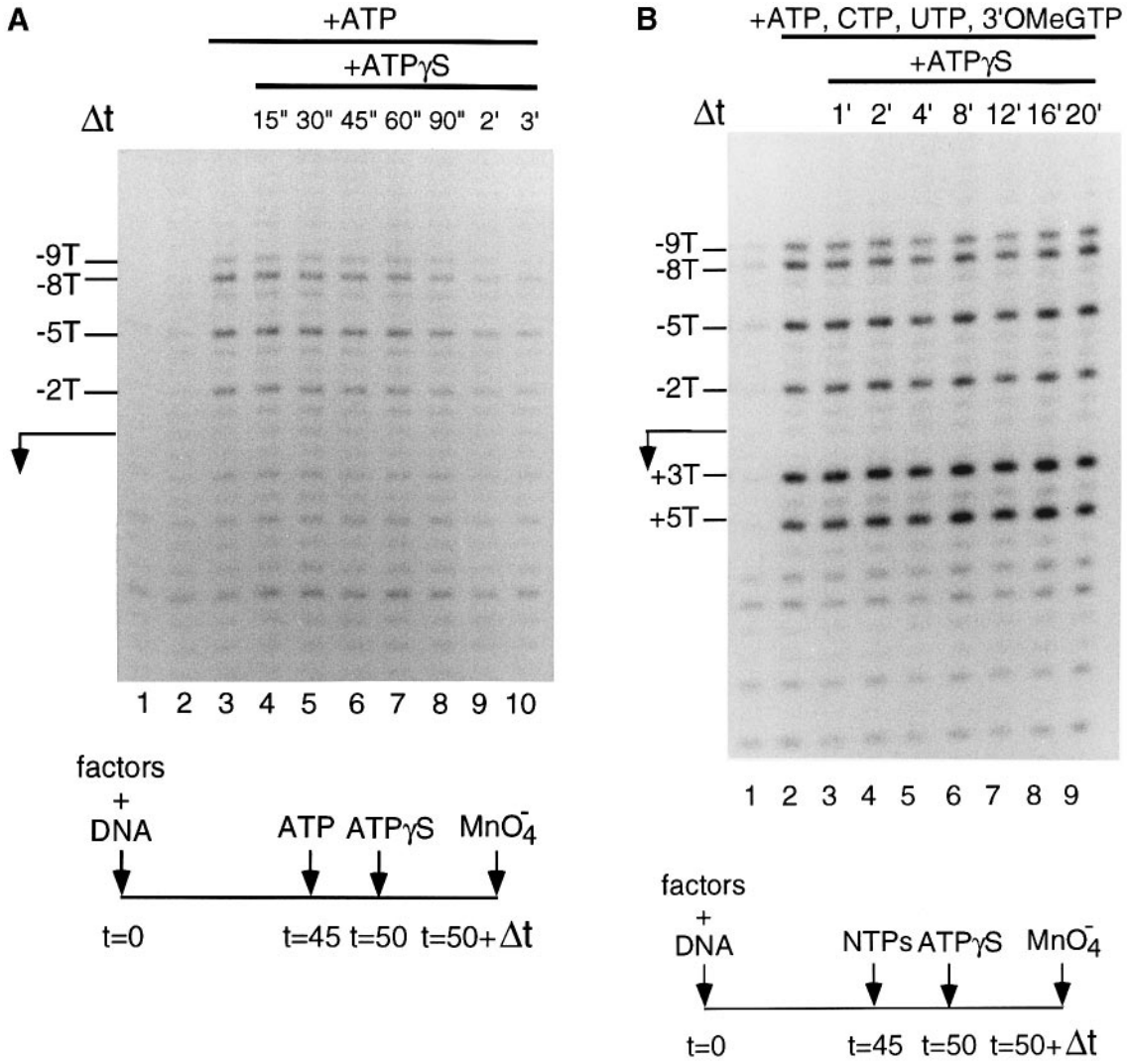

Fig. 7. After register 4, the open region of the transcription complex is insensitive to ATPγS inhibition. (**A**) Initiation complexes were assembled as described in the legend of Figure 1 and the effect of ATPγS was analyzed as depicted by the scheme indicated at the bottom. The reaction analyzed in lane 1 received no ATP. The reaction of lane 2 simultaneously received 10 μM ATP and 200 μM ATPγS. The reactions of lanes 3–10 received 10 μM ATP for 5 min and subsequently 200 μM ATPγS was added to the reactions of lanes 4–10 for the times indicated above the lanes. The reaction of lane 3 received only 10 μM ATP. Reactions were analyzed for permanganate sensitivity as described in Materials and methods. The positions of the reactive thymidines used for determination of the half-life are indicated to the left. Phosphorimager quantification of three independent experiments determined a half-life for the initially open complex of 45 s. (**B**) The ATPγS sensitivity of the open region of the AdML+4G complex was analyzed according to the scheme depicted at the bottom. The AdML+4G complex was formed by addition of 60 μM ATP, 10 μM CTP, 10 μM UTP and 120 μM 3′-OMeGTP to the assembled initiation complexes and a 5 min incubation (lanes 2– 9). To the reactions of lanes 3–9, ATPγS was added to 600 μM for the times indicated above the lanes. Reactions were analyzed for permanganate sensitivity as described in Materials and methods. The positions of the reactive thymidines used for determination of the half-life are indicated to the left. Phosphorimager quantification of three independent experiments determined a half-life for the open region of the AdML+4G complex of >10 min.

products of 6 nt and shorter, the assay yielded more than one RNA per template (Figure 5), which is formal proof of abortive transcription. The transcription complex is capable of abortive RNA synthesis until register 10, which marks the transition to productive transcription. This was verified by the reinitiation experiment of Figure 6, which indicates that at register 11 the pol II enzyme has cleared the promoter. The transcription 'bubble' achieves resistance to ATPγS in register 4 (Figure 7). This observation argues strongly against involvement of the ATP-dependent DNA helicases of TFIIH in promoter clearance.

In conclusion, our results indicate that at least three structural transitions occur within the pol II complex during initiation.

Three-step model for pol II transcription initiation

Our analysis provides a detailed model for transcription initiation by pol II, which is depicted in Figure 8. After assembly of the complete pre-initiation complex on the AdMLP, addition of ATP or dATP results in separation of the DNA strands from positions –9 to +2. This constitutes the first transition in the pol II initiation complex, and we showed previously that the DNA helicase activity of TFIIH is responsible for the formation of this open complex (Holstege *et al.*, 1996). We have used ATPγS, which acts as a reversible inhibitor of transcription (Conaway and Conaway, 1988), to determine the stability of the –9/+2 complex. This open complex decays with a short half-life of 45 s, which is very similar to the half-

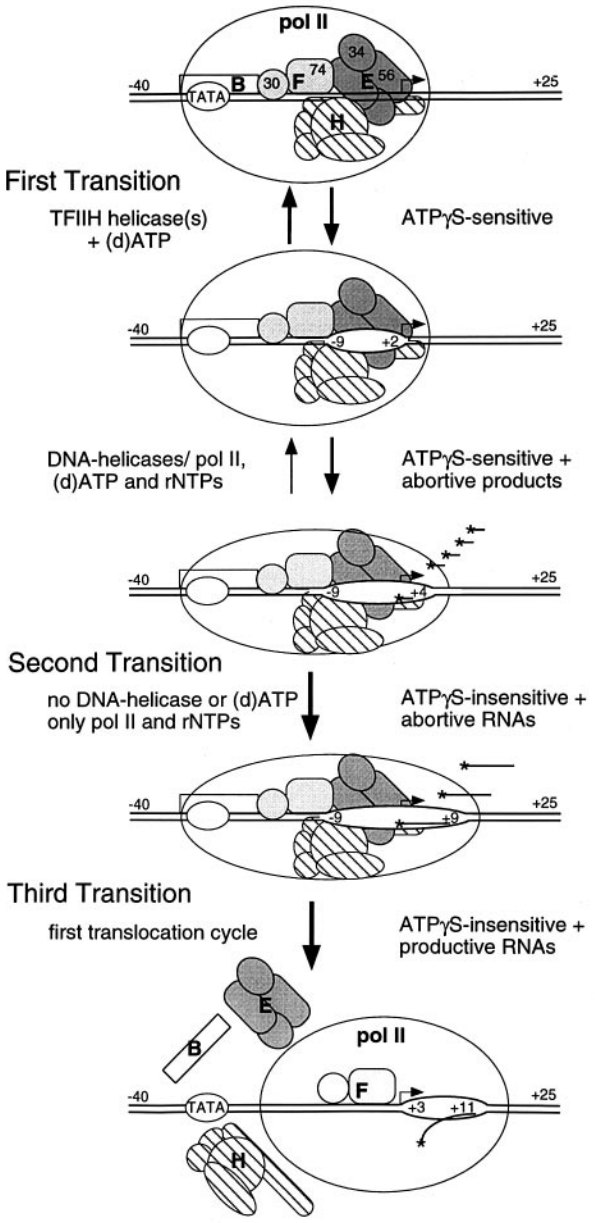

Fig. 8. Three transitions in the pol II transcription complex accompany transcription initiation. The figure depicts a model for the transcription initiation reaction of pol II on the AdMLP. The assembled initiation complex is depicted on the top of the figure, with the basal transcription factors and pol II as balloons. TBP is the unlabeled open balloon over the TATA box, whereas (subunits of) the other basal factors are indicated by a letter and by different shadings. It is important to note that for the sake of simplicity the DNA strands of the TATA box are drawn as straight lines while it is known from crystallographic studies that they are severely bent (Kim *et al.*, 1993a,b). The positions of the basal factors and pol II on promoter DNA are based on two- and three-dimensional crystallographic and protein–DNA cross-linking studies (for reviews, see Orphanides *et al.*, 1996; Roeder, 1996; Nikolov and Burley, 1997), but in most cases the positions are not exactly known. The numbers indicate positions of the DNA template and are relative to the transcription start site indicated by the arrow. The asterisks indicate the first residue of the RNA product. The three steps depict the structural transitions during initiation and are explained in the Discussion.

life of the activated pre-initiation complex (40 s) as determined by Conaway and co-workers (Conaway and Conaway, 1988; Dvir *et al.*, 1996b). This indicates that the complex is in dynamic equilibrium between open and closed conformation and that a continued ATP hydrolysis by TFIIH-associated DNA helicase(s) is required to sustain promoter opening. The –9/+2 complex has two options: fall back to the closed complex or initiate synthesis of an RNA chain by formation of the first phosphodiester bond. Synthesis of the ApC dinucleotide stabilizes the open complex ($t_{1/2}$ increases to 90 s). The high rate of abortive synthesis indicates that this product is not stably attached in the transcription complex. It is unclear whether the DNA strands remain melted when the ApC product is expelled from the pol II complex or whether the initiation complex undergoes a cycle of opening reactions. In either case, a collapse to a closed complex is no cause for abortive transcription. First, abortive transcription continues after register 4 when the complex has gained full resistance to ATPγS (Figure 7). Second, in the complete transcription system, we observed that ApC formation is more efficient using a heteroduplex fragment containing a pre-melted –4/+2 region than a normal AdMLP fragment as template (data not shown). The observation that the pol II complex in registers 2 and 3 is (moderately) sensitive to ATPγS may indicate that the ATP-dependent DNA helicases of TFIIH are involved in the expansion of the 'transcription bubble' until formation of the third phosphodiester bond. However, this is unlikely because transcription from an AdMLP fragment containing a heteroduplex region from positions –4 to +2 is fully independent of TFIIH and ATP hydrolysis (Holstege *et al.*, 1996; unpublished data). Therefore, we propose that until pol II reaches register 4 the open region is intrinsically unstable and easily readopts a double-stranded conformation. In this case, the DNA helicase activity of TFIIH is required for the next round of open complex formation.

The second transition in the initiation complex occurs at register 4. When pol II reaches this register, the probability of collapse of the open region is very low, and this results in resistance to ATPγS inhibition (Figure 7). This indicates that after register 4 the action of the TFIIH DNA helicases is not involved in the expansion of the open region. This is in agreement with the observations that HMG2 is unable to induce TFIIH or ATP dependence on transcription complexes at register 5 (Stelzer *et al.*, 1994). Our observations are also in agreement with experiments which showed that dinucleotide-primed synthesis of RNA of 9–13 nt is not inhibited by ATPγS (Dvir *et al.*, 1996b) and others which showed that initial transcription from the AdE4 promoter does not require ATP hydrolysis after opening (Jiang *et al.*, 1996). Also, Luse and co-workers have shown that a stable transcription complex is obtained after formation of a 4 nt product (Luse *et al.*, 1987). It is noteworthy that although the pol II complex is capable of abortive transcription until register 10, the efficiency of abortive synthesis decreases with RNA length. It is possible that di- and trinucleotides are expelled from the transcription complex more efficiently than longer products because they bind more weakly to the transcription complex. The reasons for this could be that short RNAs do not attach to the RNA-binding site in the pol II enzyme or that the short RNA–DNA hybrid is unstable in these cases, which is ~8 bp in an elongating polymerase (Landick, 1997).

The results from the permanganate sensitivity and

abortive transcription experiments indicate that a third structural transition occurs within the transcription complex when it goes from register 9 to register 11 (Figures 2B, 4 and 6). This transition can be described as promoter clearance since the complex in register 11 is unable to reinitiate transcription and the initially opened region (from –9 to –2) readopts the double-stranded conformation. It is important to note that others refer to promoter clearance as 'a step in which pol II leaves the promoter in a way that allows a new polymerase to enter and form a reinitiation complex' (Jiang *et al.*, 1996). Because association of a new pol II molecule is a slow step in our reconstituted transcription assay, we have been unable to test association of the next pol II molecule to register 11-stalled complexes. Several observations argue against an involvement of ATP hydrolysis in the promoter clearance step. First, transcription from supercoiled or pre-melted templates is completely independent of TFIIH and of ATP hydrolysis (Parvin and Sharp, 1993; Goodrich and Tjian, 1994; Parvin *et al.*, 1994; Tantin and Carey, 1994; Holstege *et al.*, 1996). Second, when we included in our assays the CTD kinase inhibitor H-8, we did not observe an effect on abortive transcription or on formation of the different open complexes (data not shown). Finally, complexes beyond register 3 are insensitive to ATPγS, which presumably blocks TFIIH DNA helicase activity. In addition, it was shown earlier that elongation of a 10 nt transcript does not require ATP hydrolysis (Sawadogo and Roeder, 1984). Contrary to these findings, several groups have provided evidence that promoter clearance is linked to ATP hydrolysis, suggesting an involvement of TFIIH (Goodrich and Tjian, 1994; Akoulitchev *et al.*, 1995; Dvir *et al.*, 1996a; Jiang *et al.*, 1996). Part of the discrepancy relates to the promoter analyzed. CTD kinase activity is required for promoter clearance from the AdE4 (Jiang *et al.*, 1996) and DHFR (Akoulitchev *et al.*, 1995) promoters. Previous analyses in purified transcription systems indicated that productive transcription and, thus, promoter clearance from a AdMLP template is independent of CTD phosphorylation (Serizawa *et al.*, 1993). Other studies analyzing AdMLP transcription suggested an involvement of TFIIH DNA helicase activity in the clearance step (Goodrich and Tjian, 1994; Dvir *et al.*, 1996a). However, these results are derived from dinucleotide-primed transcription reactions. We observed that in our system, these reactions are not a good representation of the transcription reaction with natural substrates (Holstege *et al.*, 1996). Overshadowing these considerations is the observation from Kornberg and colleagues that proteins present in crude systems can induce dependence on CTD kinase activity for transcription (Li and Kornberg, 1994). We also observed that the H-8 kinase inhibitor blocks transcription from AdMLP templates in crude extracts (data not shown). Apparently, these proteins are not basal transcription factors, because they are absent in a purified system (Li and Kornberg, 1994). Our observations support the model that a single ATP activation event, which results in formation of the open complex, can be sufficient for productive pol II transcription of AdMLP templates in a purified system.

In the model depicted in Figure 8, the positions of the basal factors and pol II are based on crystallographic and chemical cross-linking studies of (partially) assembled pre-initiation complexes (Robert *et al.*, 1996; for reviews, see Orphanides *et al.*, 1996; Roeder, 1996; Nikolov and Burley, 1997). A simultaneous release of TFIIB, TFIIE and TFIIH is depicted at the third transition (promoter clearance). This is just to indicate that we think that at this stage the interaction between these factors and pol II must be broken to allow translocation of the pol II enzyme and reclosure of the –9/+2 region. It is very possible that some factors dissociate prior to reclosure. In fact, Reinberg and colleagues have used template competition experiments to determine the release point for the basal factors (Zawel *et al.*, 1995). The use of a high concentration of nucleotides (1 mM) in this study gives some uncertainty to the exact stall sites, but their experiments indicate that TFIIB dissociates at an early stage. In addition, they have mapped TFIIE release between registers 1 and 10, TFIIF release between registers 10 and 68 and TFIIH release between registers 30 and 68.

Comparison of pol II and bacterial RNAP transcription initiation

Our detailed description of transcription initiation by pol II allows a comparison with the well-studied case of RNA polymerase (RNAP) from *Escherichia coli*. In the σ^{70} $\alpha_2\beta\beta'$ system, the σ^{70} factor mediates open complex formation by binding to the non-template strand, resulting in the formation of an open region of 12–15 bp also predominantly upstream of the transcription start (Siebenlist *et al.*, 1980). For most bacterial promoters, this is a slow step. By comparison, open complex formation in the pol II system is a fast step and involves an enzymatic activity, an ATP-dependent DNA helicase(s), to melt promoter DNA at similar positions (Holstege *et al.*, 1996). Abortive synthesis has been documented for many bacterial promoters, but in this system dinucleotide products are formed as a constant fraction of the total production of RNA transcripts (Carpousis and Gralla, 1980). Clearly, this is not the case for the pol II system as the ApC product is not formed in the reactions which stall pol II in registers 4–9 (Figure 3A). Formation of ApC in the AdML+3G complex is probably caused by an inefficient addition of 3′-OMeGTP. In the bacterial system, the efficiency of abortive synthesis decreases with the length of the products (Carpousis and Gralla, 1980), and the abortive stage of transcription is characterized by a lack of fidelity leading to read-through (Metzger *et al.*, 1993). Both characteristics are also observed in the pol II system (Figures 3 and 5). Formation of an ATPγS-insensitive open region at register 4 as the second transition of the pol II system obviously does not occur in the bacterial system. Promoter clearance by RNAP involves release by the σ^{70} factor from the transcription complex, which occurs around register 9 (Metzger *et al.*, 1993) and accompanies the switch to productive RNA synthesis. This is very similar to our proposal for pol II, although the basal factors TFIIB and TFIIE may be released prior to promoter clearance. Another important difference is that TBP remains associated with the promoter to allow reinitiation (Zawel *et al.*, 1995), whereas σ factor needs to bind core RNAP before reassociation with the promoter. We find that at its downstream edge the open region expands in a continuous movement, whereas at its upstream edge the open complex contracts discontinuously

(Figure 2). This is similar to what has been observed with open regions in bacterial RNAP complexes (Zaychikov *et al.*, 1995). It is important to stress that we analyzed movement of the open region and that our findings do not imply discontinuous translocation of the pol II moiety. Recent data from bacterial RNAP suggest a sliding clamp model for translocation (reviewed in Landick, 1997). We currently are employing different footprinting methods to investigate this issue for pol II complexes.

Implications for regulation of pol II transcription

The transition from abortive to productive transcription is also interesting with respect to transcription regulation. It is well documented that some promoters (like the *Drosophila hsp70* and mammalian c-*myc* and c-*fos* promoters) appear to be regulated by controlling the release of initiated pol II molecules which are stalled in promoter-proximal positions (for review, see Bentley, 1995). In addition, evidence has been obtained for several viral and cellular activators that transcription stimulation involves a stimulation of CTD phosphorylation of pol II. This occurs at promoter-proximal positions and results in a more processive form of pol II (Yankulov *et al.*, 1995; Parada and Roeder, 1996). Our definition of separate transitions in transcription initiation by pol II allows a further examination of the molecular mechanism that mediates transcriptional activation in these cases.

Materials and methods

Materials

Nucleotides were obtained from Pharmacia/LKB and Boehringer Mannheim. Radioactive nucleotides were obtained from Amersham or from NEN/DuPont. Restriction enzymes, Klenow fragment and calf intestine phosphatase were purchased from Boehringer Mannheim. $KMnO_4$ was from Janssen Chimica. Piperidine was from Fluka. Oligonucleotides and the T7 sequencing kit were from Pharmacia/LKB. ApC was purchased from Sigma. ApCp(2′-O-Me-G) and ApCpUp(2′-O-Me-G) were obtained from Eurogentec. The QIAEX II kit was supplied by QIAGEN, GmbH. The 8WG16 antibody was purchased from Promega. Quantification by phosphorimager analysis was performed on a Storm 820 gel scanner from Molecular Dynamics using ImageQuaNT software, version 4.2.

Purification of transcription factors

Pol II was affinity purified from calf thymus using the 8WG16 antibody as described previously (Holstege *et al.*, 1995). Recombinant histidine-tagged human TBP, recombinant human TFIIB and recombinant human TFIIE were expressed in *Escherichia coli* and purified as described previously (Timmers, 1994; Holstege *et al.*, 1995, 1996). Recombinant TFIIF (human RAP74/rat RAP30) was purified from baculovirus-infected insect cells and TFIIH was purified from HeLa cell extracts, as described previously (Holstege *et al.*, 1995). Preparations of TBP, TFIIB, TFIIE and TFIIF were homogenous, and preparations of TFIIH and pol II are highly purified as judged from silver staining of protein gels (see Holstege *et al.*, 1996). Leave-out transcription assays using a linearized AdMLP template showed that all protein preparations are free from contamination by other basal factors or pol II. In addition, preparations are negative for nuclease and topoisomerase activity, as determined in assays described earlier (Holstege *et al.*, 1996). Protein preparations are also negative for proteins that have a high affinity for heteroduplex DNA fragments, such as PC4 (S.Werten, F.W.M.Langen, R.van Schaik, H.Th.M.Timmers, M.Meisterernst and P.C.van der Vliet, in press).

Promoter DNA

The AdML promoter constructs were derived from pDNAdMLmut (depicted as AdML+15G in Figure 1) that encompasses the –50 to +10 Ad2 MLP with a *Bss*HII site introduced by site-directed mutagenesis (G→C) of the –15 position (Holstege *et al.*, 1996). The pDNAdML+2G,+3G, +4G, +6G, +9G, +10G and +11G plasmids were constructed by inserting the appropriate oligonucleotides into *Bss*HII–*Xba*I-digested pDNAdMLmut. All constructions were verified by DNA sequencing using the T7 sequencing kit (Pharmacia/LKB). For $KMnO_4$ sensitivity assays, *Eco*RI–*Hin*dIII-digested fragments were labeled on the *Eco*RI site using [α-^{32}P]dATP and Klenow fragment of *E.coli* DNA polymerase I using standard procedures (Sambrook *et al.*, 1989). For transcription assays, promoter DNA fragments were isolated from the appropriate pDNAdMLmut derivative by restriction enzyme digestion with *Pvu*II. The 372 bp promoter fragment was recovered by agarose gel electrophoresis and subsequent extraction using the QIAEX II kit following the manufacturers instructions. Concentrations of the isolated DNA fragments was determined by UV absorption using an UltrospecPlus spectrophotometer (Pharmacia/LKB).

$KMnO_4$ sensitivity assay

Radiolabeled probe (0.2–0.4 ng) was incubated at 30°C for 45 min with 12 ng of TBP, 25 ng of TFIIB, 80 ng of TFIIF, 40 ng of pol II, 30 ng of TFIIE and 160 ng of TFIIH in a 20 µl reaction that contained 12 mM HEPES–KOH (pH 7.9), 60 mM KCl, 5 mM $MgCl_2$, 3 mM $(NH_4)_2SO_4$, 0.6 mM EDTA, 0.7 mM dithiothreitol (DTT) and 90 µg/ml bovine serum albumin (BSA). Nucleotides (1 µl) were added for 5 min at the final concentrations specified in the figure legends. Two µl of 160 mM $KMnO_4$ was added for 2 min, after which the reactions were stopped by addition of 2 µl of 14.4 M β-mercaptoethanol. DNA was cleaved at modified residues by addition of cleavage mix (180 µl of 1.1 M piperidine, 1 mM EDTA, 1 mM EGTA and 25 µg/ml herring sperm DNA) and subsequent incubation for 30 min at 90°C. DNA was recovered and analyzed as described previously (Holstege and Timmers, 1997).

Transcription assay

Template DNA fragment (unless indicated otherwise: 5 ng of the 372 bp *Pvu*II fragment) was incubated at 30°C for 45 min with 12 ng of TBP, 25 ng of TFIIB, 80 ng of TFIIF, 40 ng of pol II, 30 ng of TFIIE and 160 ng of TFIIH in a 20 µl reaction that contained 12 mM HEPES–KOH (pH 7.9), 60 mM KCl, 5 mM $MgCl_2$, 3 mM $(NH_4)_2SO_4$, 0.6 mM EDTA, 0.7 mM DTT and 90 µg/ml BSA. After this assembly step, nucleotides were added to the concentrations indicated in the figure legends and incubated at 30°C for the appropriate time as indicated.

For analysis of short RNAs, transcription was stopped by heating the reaction mixture to 68°C for 3 min. After cooling on ice, 10 µl of the mixture was transferred to tubes containing 1 µl of 10× CIP buffer (500 mM Tris–HCl pH 8.5/1 mM EDTA) and 0.4 µl of calf intestine phosphatase (20 U/µl). After a 20 min incubation at 37°C, 3 µl of loading buffer (50% glycerol/200 mM EDTA/0.05% bromphenol blue) was added and 7 µl of the resulting mixture was analyzed by electrophoresis in 23% polyacrylamide (acrylamide:bisacrylamide = 19:2)–7 M urea gels. As suggested by S.Akoulitchev, treatment of the reaction mixtures with phosphatase greatly reduces the background signal due to non-incorporated [α-^{32}P]CTP (Akoulitchev *et al.*, 1995). In this gel system, RNA products of 6 nt or shorter do not run according to their molecular weight. RNA products were identified by co-migration of ApC, CpApC, ApCp(2′-O-Me-G) and ApCpUp(2′-O-Me-G) and UV shadowing of the gels using a 254 nm UV light source. The amount of RNA formed was determined by phosphorimager analysis using known concentrations of [α-^{32}P]CTP as a standard.

For primer extension analysis, transcription was stopped by addition of 180 µl of TENS (10 mM Tris–HCl pH 8.0, 1 mM EDTA, 200 mM NaCl, 0.1% SDS) and 5 µg of *E.coli* tRNA. The mixture was extracted with phenol–chloroform and chloroform. Then 75 fmol of radiolabeled M13 primer (5′-GTTGTAAAACGACGGCCAGT-3′) was added per reaction and nucleic acids were precipitated by addition of 0.1 vol. of 5 M NH_4Ac and 2 vols of 96% ethanol. The primer extension reaction was performed and analyzed as described previously (Holstege *et al.*, 1995). Correctly initiated transcripts yield primer extension products of 63 nt in this analysis, which were identified by co-migration of DNA sequencing reactions using the pAdMLmut and the M13 primer as the template–primer combination.

Acknowledgements

We thank J.-M.Egly and D.Reinberg for TFIIH antibodies and J.A.van der Knaap for isolation of TBP and TFIIB preparations. We acknowledge the contributions of R.Schiphof in providing HeLa cells and J.W.Borst in purification of pol II. We thank J.H.J.Hoeijmakers, S.Akoulitchev and W.van Driel for discussions. We also thank J.A.van der Knaap, A.J.M.Walhout and P.C.van der Vliet for critical reading of this manu-

script. In addition, we are grateful to P.C.van der Vliet for continued support and suggestions. U.F. was financially supported by an EMBO long-term fellowship. This work was also supported by the Netherlands Foundation for Chemical Research (SON).

References

Akoulitchev,S., Makela,T.P., Weinberg,R.A. and Reinberg,D. (1995) Requirement for TFIIH kinase activity in transcription by RNA polymerase II. *Nature*, **377**, 557–560.

Bentley,D. (1995) Regulation of transcriptional elongation by RNA polymerase II. *Curr. Opin. Genet. Dev.*, **5**, 210–216.

Burley,S.K. and Roeder,R.G. (1996) Biochemistry and structural biology of the transcription factor IID (TFIID). *Annu. Rev. Biochem.*, **65**, 769–799.

Carpousis,A.J. and Gralla,J.D. (1980) Cycling of ribonucleic acid polymerase to produce oligonucleotides during initiation *in vitro* at the lac UV5 promoter. *Biochemistry*, **19**, 3245–3253.

Chao,D.M., Gadbois,E.L., Murray,P.J., Anderson,S.F., Sonu,M.S., Parvin,J.D. and Young,R.A. (1996) A mammalian SRB protein associated with an RNA polymerase II holoenzyme. *Nature*, **380**, 82–85.

Conaway,R.C. and Conaway,J.W. (1988) ATP activates transcription initiation from promoters by RNA polymerase II in a reversible step prior to RNA synthesis. *J. Biol. Chem.*, **263**, 2962–2968.

Dvir,A., Conaway,R.C. and Conaway,J.W. (1996a) Promoter escape by RNA polymerase II. *J. Biol. Chem.*, **271**, 23352–23356.

Dvir,A., Garrett Pfeil,K., Chalut,C., Egly,J.-M., Conaway,J.W. and Conaway,R.C. (1996b) A role for ATP and TFIIH in activation of the RNA polymerase II preinitiation complex prior to transcription initiation. *J. Biol. Chem.*, **271**, 7245–7248.

Flores,O., Lu,H. and Reinberg,D. (1992) Factors involved in specific transcription by mammalian RNA polymerase: identification and characterization of TFIIH. *J. Biol. Chem.*, **267**, 2786–2793.

Goodrich,J.A. and Tjian,R. (1994) Transcription factors IIE and IIH and ATP hydrolysis direct promoter clearance by RNA polymerase II. *Cell*, **77**, 145–156.

Gralla,J.D. (1996) Global steps during initiation by RNA polymerase II. *Methods Enzymol*., **273**, 99–111.

Halle,J.-P. and Meisterernst,M. (1996) Gene expression: increasing evidence for a transcriptosome. *Trends Genet.*, **12**, 161–163.

Holstege,F.C.P. and Timmers,H.T.M. (1997) Analysis of open complex formation during RNA polymerase II transcription initiation using heteroduplex templates and potassium permanganate probing. *Methods*, **12**, 203–211.

Holstege,F.C.P., Tantin,D., Carey,M., van der Vliet,P.C. and Timmers,H.T.M. (1995) The requirement for the basal transcription factor IIE is determined by the helical stability of promoter DNA. *EMBO J*., **14**, 810–829.

Holstege,F.C.P., van der Vliet,P.C. and Timmers,H.T.M. (1996) Opening of an RNA polymerase II promoter occurs in two distinct steps and requires the basal transcription factors IIE and IIH. *EMBO J*., **15**, 1666–1677.

Jacob,G.A., Kitzmiller,J.A. and Luse,D.S. (1994) RNA polymerase II promoter strength *in vitro* may be reduced by defects at initiation or promoter clearance. *J. Biol. Chem.*, **269**, 3655–3663.

Jiang,Y., Yan,M. and Gralla,J.D. (1996) A three-step pathway of transcription initiation leading to promoter clearance at an activated RNA polymerase II promoter. *Mol. Cell. Biol*., **16**, 1614–1621.

Kim,J.L., Nikolov,D.B. and Burley,S.K. (1993a) Co-crystal structure of TBP recognizing the minor groove of a TATA element. *Nature*, **365**, 520–527.

Kim,Y., Geiger,J.H., Hahn,S. and Sigler,P.B. (1993b) Crystal structure of a yeast TBP/TATA-box complex. *Nature*, **365**, 512–520.

Kim,Y.J., Bjorklund,S., Li,Y., Sayre,M.H. and Kornberg,R.D. (1994) A multiprotein mediator of transcriptional activation and its interaction with the C-terminal repeat domain of RNA polymerase II. *Cell*, **77**, 599–608.

Koleske,A.J. and Young,R.A. (1994) An RNA polymerase II holoenzyme responsive to activators. *Nature*, **368**, 466–469.

Koleske,A.J. and Young,R.A. (1995) The RNA polymerase II holoenzyme and its implications for gene regulation. *Trends Biochem. Sci.*, **20**, 113–116.

Landick,R. (1997) RNA polymerase slides home: pause and termination site recognition. *Cell*, **88**, 741–744.

Levin,J.R. and Chamberlin,M.J. (1987) Mapping and characterization of transcriptional pause sites in the early genetic region of bacteriophage T7. *J. Mol. Biol*., **196**, 61–84.

Li,Y. and Kornberg,R.D. (1994) Interplay of positive and negative effectors in function of the C-terminal repeat domain of RNA polymerase II. *Proc. Natl Acad. Sci. USA*, **91**, 2362–2366.

Luse,D.S. and Jacob,G.A. (1987) Abortive initiation by RNA polymerase II *in vitro* at the adenovirus 2 major late promoter. *J. Biol. Chem.*, **262**, 14990–14997.

Luse,D.S., Kochel,T., Kueumpel,E.D., Coppola,J.A. and Cai,H. (1987) Transcription initiation by RNA polymerase II *in vitro*. *J. Biol. Chem.*, **262**, 289–297.

Mäkelä,T.P., Parvin,J.D., Kim,J., Huber,L.J., Sharp,P.A. and Weinberg,R.A. (1995) A kinase-deficient transcription factor TFIIH is functional in basal and activated transcription. *Proc. Natl. Acad. Sci. USA*, **92**, 5174–5178.

Maldonado,E. *et al*. (1996) A human RNA polymerase II complex associated with SRB and DNA-repair proteins. *Nature*, **381**, 86–89.

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

Nikolov,D.B. and Burley,S.K. (1997) RNA polymerase II transcription initiation: a structural view. *Proc. Natl Acad. Sci. USA*, **94**, 15–22.

Orphanides,G., Lagrange,T. and Reinberg,D. (1996) The general transcription factors of RNA polymerase II. *Genes Dev.*, **10**, 2657–2683.

Ossipow,V., Tassan,J.-P., Nigg,E.A. and Schibler,U. (1995) A mammalian RNA polymerase II holoenzyme containing all components required for promoter-specific transcription initiation. *Cell*, **83**, 137–146.

Pan,G. and Greenblatt,J. (1994) Initiation of transcription by RNA polymerase II is limited by melting of the promoter DNA in the region immediately upstream of the initiation site. *J. Biol. Chem.*, **269**, 30101–30104.

Parada,C.A. and Roeder,R.G. (1996) Enhanced processivity of RNA polymerase II triggered by Tat-induced phosphorylation of its carboxy-terminal domain. *Nature*, **384**, 375–378.

Parvin,J.D. and Sharp,P.A. (1993) DNA topology and a minimal set of basal factors for transcription by RNA polymerase II. *Cell*, **73**, 533–540.

Parvin,J.D., Shykind,B.M., Meyers,R.E., Kim,J. and Sharp,P.A. (1994) Multiple sets of basal factors initiate transcription by RNA polymerase II. *J. Biol. Chem.*, **269**, 18414–18421.

Pazin,M.J. and Kadonaga,J.T. (1997a) SWI2/SNF2m and related proteins: ATP-driven motors that disrupt protein–DNA interactions? *Cell*, **88**, 737–740.

Pazin,M.J. and Kadonaga,J.T. (1997b) What's up and down with histone deacetylation and transcription? *Cell*, **89**, 325–328.

Pirotta,V. (1995) Chromatin complexes regulating gene expression in *Drosophila*. *Curr. Opin. Genet. Dev.*, **5**, 466–472.

Robert,F., Forget,D., Li,J., Greenblatt,J. and Coulombe,B. (1996) Localization of subunits of transcription factors IIE and IIF immediately upstream of the transcriptional initiation site of the adenovirus major late promoter. *J. Biol. Chem.*, **271**, 8517–8520.

Roeder,R.G. (1996) The role of general initiation factors in transcription by RNA polymerase II. *Trends Biochem. Sci.*, **21**, 327–335.

Sambrook,J., Fritsch,E.F. and Maniatis,T. (1989) *Molecular Cloning: A Laboratory Manual*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

Sawadogo,M. and Roeder,R.G. (1984) Energy requirement for specific transcription initiation by the human RNA polymerase II system. *J. Biol. Chem.*, **259**, 5321–5326.

Schaeffer,L., Roy,R., Humbert,S., Moncollin,V., Vermeulen,W., Hoeijmakers,J.H.J., Chambon,P. and Egly,J.-M. (1993) DNA repair helicase: a component of BTF2 (TFIIH) basic transcription factor. *Science*, **260**, 58–63.

Schaeffer,L., Moncollin,V., Roy,R., Staub,A., Mezzina,M., Sarasin,A., Weeda,G., Hoeijmakers,J.H.J. and Egly,J.-M. (1994) The ERCC2/DNA repair protein is associated with the class II BTF2/TFIIH transcription factor. *EMBO J.*, **13**, 2388–2392.

Serizawa,H., Conaway,J.W. and Conaway,R.C. (1993) Phosphorylation of C-terminal domain of RNA polymerase II is not required in basal transcription. *Nature*, **363**, 371–374.

Siebenlist,U., Simpson,R.B. and Gilbert,W. (1980) *E.coli* RNA polymerase interacts homologously with two different promoters. *Cell*, **20**, 269–281.

Stelzer,G., Goppelt,A., Lottspeich,F. and Meisterernst,M. (1994) Repression of basal transcription by HMG2 is counterparted by TFIIH-associated factors in an ATP-dependent process. *Mol. Cell. Biol.*, **14**, 4712–4721.

Svejstrup,J.Q., Vichi,P. and Egly,J.-M. (1996) The multiple roles of transcription/repair factor TFIIH. *Trends Biochem. Sci.*, **21**, 346–350.

Tantin,D. and Carey,M. (1994) A heteroduplex template circumvents the energetic requirement for ATP during activated transcription by RNA polymerase II. *J. Biol. Chem.*, **269**, 17397–17400.

Timmers,H.T.M. (1994) Transcription initiation by RNA polymerase II does not require hydrolysis of the β–γ phosphoanhydride bond of ATP. *EMBO J.*, **13**, 391–399.

Tyree,C.M., George,C.P., Lira-De Vito,L.M., Wampler,S.L., Dahmus, M.E., Zawel,L. and Kadonaga,J.T. (1993) Identification of a minimal set of proteins that is sufficient for accurate initiation of transcription by RNA polymerase II. *Genes Dev.*, **7**, 1254–1265.

Verrijzer,C.P. and Tjian,R. (1996) TAFs mediate transcriptional activation and promoter selectivity. *Trends Biochem. Sci.*, **21**, 338–341.

Wang,W., Carey,M. and Gralla,J.D. (1992) Polymerase II promoter activation: closed complex formation and ATP-driven start site opening. *Science*, **255**, 450–453.

Yankulov,K., Yamashita,K., Roy,R., Egly,J.-M. and Bentley,D. (1995) The transcriptional elongation inhibitor 5,6-dichloro-1-beta-D-ribofuranosylbenzimidazole inhibits transcription factor IIH-associated protein kinase. *J. Biol. Chem.*, **270**, 23922–23925.

Zawel,L., Kumar,P. and Reinberg,D. (1995) Recycling of the general transcription factors during RNA polymerase II transcription. *Genes Dev.*, **9**, 1479–1490.

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 on July 18, 1997; revised on October 6, 1997