Mapping of RNA Polymerase on Mammalian Genes in Cells and Nuclei

Jovan Mirkovitch* and James E. Darnell, Jr.

The Rockefeller University, New York, New York 10021

Submitted May 26, 1992; Accepted July 24, 1992

The assembly of an RNA polymerase II initiation complex at a promoter is associated with the melting of the DNA template to allow the polymerase to read the DNA sequence and synthesize the corresponding RNA. Using the specific single-stranded modifying reagent KMnO_4 and a new genomic sequencing technique, we have explored the melted regions of specific genes in genomic DNA of whole cells or of isolated nuclei. We have demonstrated for the first time in vivo the melting in the promoter proximal transcribed region that is associated with the presence of RNA polymerase II complexes. An interferon-inducible gene, ISG-54, exhibited KMnO_4 sensitivity over ~ 300 nucleotides downstream of the RNA initiation site in interferon-treated cells when the gene was actively transcribed but not in untreated cells where the gene was not transcribed. The extent of KMnO_4 modification was proportional to transcription levels. The KMnO_4 sensitivity was retained when nuclei were isolated from induced cells but was lost if the engaged polymerases were further allowed to elongate the nascent RNA chains ("run-on"). The sensitivity to KMnO_4 in isolated nuclei was retained if the run-on incubation was performed in the presence of α -amanitin, which blocks progress of engaged polymerases. A similar analysis identified an open sequence of only ~ 30 bases just downstream of the start site of the transthyretin (TTR) gene in nuclei isolated from mouse liver, a tissue where TTR is actively transcribed. This abrupt boundary of KMnO_4 sensitivity, which was removed completely by allowing engaged polymerases to elongate RNA chains, suggests that most polymerases transcribing this gene paused at about position +20. The possibility of mapping at the nucleotide level the position of actively transcribing RNA polymerases in whole cells or isolated nuclei opens new prospects in the study of transcription initiation and elongation.

INTRODUCTION

Transcription initiation by RNA polymerase II involves the participation of a number of general initiation factors. RNA polymerase II and these general factors assemble in a preinitiation complex at the promoter to form a structure that can initiate RNA synthesis in the presence of ribonucleotides (Saltzman and Weinman, 1989; Sawadogo and Sentenac, 1990; Roeder, 1991). In addition to the requirement for ubiquitous factors, transcription initiation almost always requires positive-acting factors that bind at promoters and/or remote enhancers (Johnson and McKnight, 1989; Mitchell and Tjian, 1989; Falvey and Schibler, 1991). These positive-

acting factors favor the assembly of the preinitiation complex and probably include newly recognized additional accessory proteins, named adaptors or coactivators, that facilitate transcriptional initiation (Martin, 1991; Roeder, 1991). After RNA synthesis is initiated, some general factors, such as TFIIF and TFIIS, interact with the elongating polymerase (Bengal *et al.*, 1991).

The last step in the assembly of the preinitiation complex consists in the formation of a so-called open complex, where the DNA at the initiation site is melted so that the polymerase can engage in RNA synthesis on addition of ribonucleotides. Such open complexes have been well described for bacterial promoters (Chamberlain, 1976; von Hippel *et al.*, 1984). A melted DNA conformation at eucaryotic promoters has been described in vitro for RNA polymerase I (Bateman and Paule, 1988), RNA polymerase III (Kassavetis *et al.*,

* Present address: Swiss Institute for Experimental Cancer Research, CH-1066 Epalinges s/Lausanne, Switzerland.

1990) and for a vaccinia virus promoter (Vos *et al.*, 1991). A conformational change of the RNA polymerase II complex on addition of ATP is believed to represent the formation of an open complex (Van Dyke *et al.*, 1989; Buratowski *et al.*, 1991), and the presence of unpaired DNA residues in a RNA polymerase II complex has been demonstrated recently in vitro (Wang *et al.*, 1992). DNA melting in vivo in association with RNA polymerase initiation has been described for bacterial promoters using KMnO_4 (Sasse-Dwight and Gralla, 1988, 1989, 1990). This single-stranded specific reagent is appropriate for such studies because it penetrates cells and preferentially modifies unpaired thymines. KMnO_4 has also been used in vitro to detect open complexes for eucaryotic RNA polymerase III (Kassavetis *et al.*, 1990) and vaccinia RNA polymerase (Vos *et al.*, 1991), RNA polymerase II (Wang *et al.*, 1992), as well as for the demonstration of T-antigen dependent melting at the SV40 origin of replication (Borowiec and Hurwitz, 1988).

Localization and measurement of the density of RNA polymerases along transcribed genes is usually performed by the run-on assay (Weber *et al.*, 1977). Nuclei are isolated from cells at 4°C and further incubated with radiolabeled ribonucleotides and Mg^{2+} at 30–37°C, so that the engaged polymerases that paused on cell lysis are allowed to elongate. The resulting labeled RNA is subsequently hybridized to different immobilized DNA probes to determine which regions of a gene are engaged in transcription. The intensity of the hybridization signal is proportional to the density of the polymerases along the gene fragment. In favorable cases where transcription occurs at a high rate, the run-on technique can determine transcription rate for segments only a few dozen nucleotides in length. In the best case described to date, it was possible to determine that many *Drosophila* genes contained polymerases paused at the 5' end of the transcription unit, an arrangement that could be common to eukaryotic gene expression (Rougvie and Lis, 1988, 1990).

We present here a new approach to map at the nucleotide level the position of RNA polymerase complexes in whole cells or nuclei. The melted regions associated with RNA polymerase II were mapped in vivo by KMnO_4 sensitivity on two mammalian genes using a newly described genomic sequencing technique (Mirkovitch and Darnell, 1991). The induction of transcription of the interferon-stimulated gene ISG54 resulted in a KMnO_4 sensitivity downstream of the RNA start site associated with a homogeneous distribution of polymerases along the first 282 transcribed nucleotides analyzed. The degree of KMnO_4 modification was proportional to transcriptional activity. This homogeneous distribution along the ISG54 transcribed sequence was retained when nuclei were isolated from transcribing cells. Incubation of nuclei in conditions allowing the extension of nascent RNA chains resulted in the loss of

KMnO_4 sensitivity as the polymerase complexes moved along the template. In contrast, the hepatocyte-expressed gene transthyretin (TTR) presented a paused polymerase close to the start site in isolated nuclei, even though the gene is actively expressed in intact tissue. The localization of RNA polymerases in whole cells at the nucleotide level should be of interest in the study of regulated transcription initiation, elongation, and termination.

MATERIALS AND METHODS

Cells and Nuclei

HeLaS3 cells were grown in suspension and treated with interferons exactly as described (Mirkovitch *et al.*, 1992). HeLa nuclei were isolated and digested with DNaseI as described (Mirkovitch *et al.*, 1992). Mouse liver and spleen nuclei were isolated and digested with DNaseI exactly as described (Mirkovitch and Darnell, 1991).

Run-On

About 10^8 nuclei were resuspended in 1 ml buffer RO (20 mM tris(hydroxymethyl)aminomethane [Tris]-HCl, pH 7.4, 40 mM KCl, 10 mM MgCl_2 , 1 mM dithiothreitol [DTT] and 20% glycerol). One milliliter of buffer RO was added containing ribonucleotides and/or α -amanitin, and polymerases were allowed to elongate for 12 min at 32°C. Final ribonucleotides were at 1 mM each, and α -amanitin was at 2 $\mu\text{g}/\text{ml}$. Reactions were stopped by sedimenting the nuclei and resuspending in 1 ml of wash buffer (Mirkovitch *et al.*, 1992) (15 mM Tris-HCl, pH 7.4, 0.2 mM spermine, 0.5 mM spermidine, 80 mM KCl, 2 mM K-EDTA, 1 mM DTT, 0.5 mM phenylmethylsulfonyl fluoride, 3 $\mu\text{g}/\text{ml}$ aprotinin, 0.5 $\mu\text{g}/\text{ml}$ leupeptin, and 1 $\mu\text{g}/\text{ml}$ pepstatin). Nuclei were then treated with DNaseI or KMnO_4 .

KMnO_4 Treatment of Cells, Nuclei, and DNA

HeLa cells grown in suspension were sedimented and resuspended in one-tenth volume of phosphate-buffered saline at either 4 or 37°C, and similar results were obtained at both temperatures. KMnO_4 from a 0.5 M freshly dissolved stock was added to the final concentration indicated in the figure legends and left to react for 1 min. Reactions were stopped by the addition of 2-mercaptoethanol and EDTA to final concentrations of 280 and 50 mM, respectively, and cells were lysed by the addition of one volume of solution 2 $\times$ SK (50 mM Tris-HCl, pH 8.0, 400 mM NaCl, 5 mM EDTA, 0.4% sodium dodecyl sulfate (SDS), and 0.2 mg/ml proteinase K). DNA was recovered after two phenol:ChCl₃:isoamylalcohol (25:24:1) extractions and precipitation with 2.5 volumes of ethanol at 4°C followed by centrifugation at $300 \times g$ for 5 min. Under these conditions, the manganese ions remain soluble and most of the cellular RNA does not sediment. Nuclei in wash buffer were treated with KMnO_4 on ice as for the cells, except that three volumes of 2 $\times$ SK buffer were added to lyse the nuclei. Naked genomic DNA or plasmid DNA were treated with KMnO_4 in TE (same as SK buffer without SDS) as for cells or nuclei, except that after quenching, the samples were extracted once with phenol mix and ethanol precipitated at –20°C. Dimethylsulfate and piperidine treatment of plasmid DNA were as described (Maxam and Gilbert, 1977).

Genomic Sequencing

The genomic sequencing of the TTR promoter (Mirkovitch and Darnell, 1991) and the ISG54 promoter (Mirkovitch *et al.*, 1992) were performed exactly as described. Briefly, purified genomic DNA was digested with restriction enzymes *Hind*III and either *Taq* I or *Msp* I (ISG54) or *Bgl* II and *Hae* III (TTR). The sequences corresponding to the promoters were isolated by hybridization with a biotinylated probe and immo-

bilized on streptavidin-agarose. The purified sequences were further recovered by alkali treatment. The KMnO_4 modification sites or DNaseI cleavage sites were mapped by primer extension using a high specific activity primer and run on 6% acrylamide sequencing gels. KMnO_4 -treated DNA samples were not treated with piperidine to cleave at the modified residues because Taq DNA polymerase does not seem able to polymerize easily through modified residues.

RESULTS

Single-Stranded DNA in the ISG54 Gene after Transcriptional Induction

Active transcription of the human ISG54 gene is dependent on interferon (IFN)- α treatment of cells, and induction occurs rapidly without new protein synthesis to high levels equal or greater than for actin or tubulin (Larner *et al.*, 1986; Reich *et al.*, 1987). Induced transcription requires an IFN-stimulated element (ISRE) that specifically binds at least three factors (Levy *et al.*, 1988), including ISGF-3, the positive activator of IFN-dependent transcription (Kessler *et al.*, 1988; Levy *et al.*, 1989) (Figure 1). ISG54 transcription is not induced by IFN- γ treatment alone; however, pretreatment of cells with IFN- γ results in about a fivefold superinduction of ISG54 transcription when the cells are further treated with IFN- α (Levy *et al.*, 1990). Using the single-stranded DNA cleaving reagent KMnO_4 , we analyzed this rapidly inducible single-copy gene in cells and isolated nuclei for the presence of melted DNA regions. HeLa cells grown in suspension were rapidly sedimented, resuspended in buffered saline, and treated with KMnO_4 for 1 min. Excess KMnO_4 was rapidly quenched by addition of 2-mercaptoethanol and EDTA. Cells were then immediately lysed with a SDS-proteinase K mix. After purification of the DNA, genomic sequencing was performed as previously described (Mirkovitch and Darnell, 1991; Mirkovitch *et al.*, 1992) to map the positions of KMnO_4 modification. Briefly, this procedure involves the purification of the sequence of interest from the bulk of the genomic DNA and subsequently a primer extension to reveal the modified residues. Even though the DNA template is subjected to a number of alkali treatments, the sites of KMnO_4 modification do not result in efficient strand scission. However, Taq DNA polymerase does not seem able to easily polymerize across modified residues and therefore a piperidine treatment that cleaves the sugar-phosphate backbone at KMnO_4 modified residues (Rubin and Schmid, 1980) is not necessary.

Figure 2 presents the sites of cleavage on the upper (coding) strand of the ISG54 promoter using primer 1 (at positions +156 to +131; see map in Figure 1). The cleavage pattern was identical for naked genomic DNA (Figure 2, lane 5) and for cells not expressing ISG54 (lanes 1 and 2). However, when transcription was induced by IFN- α treatment alone (lane 3), a number of residues appeared sensitive to KMnO_4 modification that had not been sensitive previously in uninduced cells or

Figure 1. Schematic representation of the ISG54 promoter region showing the positions of the primers used for genomic sequencing. Primer 1 was used for the upper strand and corresponded to positions +131 to +156. Primer 2 was used for the lower strand and corresponded to positions -169 to -140. The ISRE regulatory element is at -100 to -86 and the TATA box is at -30. The sequence of the promoter has been described in Levy *et al.* (1986).

naked genomic DNA. These residues corresponded to T's within ~ 115 bases that are downstream of the transcription initiation site (+1). The KMnO_4 sensitivity was directly proportional to transcriptional activity because superinduced cells (i.e., pretreated with IFN- γ before IFN- α , which increases transcription ~ 5 -fold) (Levy *et al.*, 1989) showed greater KMnO_4 cleavage that paralleled the increase in transcription (lane 4). In this sample, the signal from residues upstream of the start site were weaker in samples from transcribing cells as most templates had been modified by KMnO_4 close to the primer site.

In the upper strand, the transcription-associated sensitivity started at position +13, the first thymine residue downstream of the start site (Figure 1). The result presented in Figure 3 shows that on the other strand (lower, noncoding, or template strand), KMnO_4 sensitivity began at position +1. Therefore, the first transcribed nucleotide was in a single-stranded conformation. However, the first thymines upstream of the cap site at positions -4 (upper strand) and -9 (lower strand) were not sensitive. KMnO_4 sensitivity was found evenly distributed down to position +282 on the lower strand (*HindIII* site) that is in the intron of the ISG54 gene. A thymine residue at position -99 in the lower strand was found hypersensitive to KMnO_4 modification in all samples, including naked DNA (Figure 3). Sensitivity of a single isolated thymine residue may be indicative of a bent or other altered DNA structure (Borowiec *et al.*, 1987). This hypersensitive thymine maps to the ISRE, suggesting that this particular structure might be involved in the binding of transcription factors to this element.

As shown below, the single-stranded DNA conformation associated with transcription is indicative of the presence of transcribing RNA polymerase complexes that have to melt the DNA template to synthesize the corresponding RNA. These complexes seemed evenly distributed in the first 120 base pairs of the transcribed

Figure 2. KMnO₄ sensitivity of the ISG54 gene. ISG54 transcription was induced with different interferon treatments, and the promoter was probed for melted regions. IFN- α treatment was for 30 min (lane 3), IFN- γ for 18 h (lane 2), and superinduction of the gene was achieved by pretreatment with IFN- γ or 18 h followed by 30 min with IFN- α (lane 4). Cells or isolated nuclei were treated with 60 mM KMnO₄ on ice and the modified nucleotides in the upper strand were mapped by genomic sequencing using primer 1. Cells treated at 37°C with 15 mM KMnO₄ showed a similar sensitivity (data not shown). Nuclei were isolated from uninduced cells (lanes 6, 7, and 8) or superinduced cells (lanes 9, 10, and 11). Nuclei were incubated in run on conditions to allow engaged polymerases to elongate (lanes 7, 8,

sequence as most thymine residues appeared equally sensitive. The space occupied by a structure as large as RNA polymerase II is ~ 30 nucleotides as determined in vitro (Darst *et al.*, 1991; Linn and Luse, 1991; Szentirmay and Sawadogo, 1991). However, the stretch of DNA that would be rendered single-stranded by a single active polymerase is not known. We assume the KMnO₄ sensitive region of ~ 120 nucleotides represents several polymerase at different positions in different cells. In the absence of induced transcription, no sensitivity was found in uninduced cells, suggesting that there is no paused polymerase in the inactive ISG54 gene, unlike some other rapidly induced eucaryotic genes (Rougvie and Lis, 1988, 1990).

Transcription-Associated Melted Configuration Derives from Engaged RNA Polymerases

To examine the origin of the KMnO₄ sensitivity downstream of the cap site of ISG54, we analyzed conditions in which this sensitivity was lost or retained in isolated nuclei. The sensitivity was stable when nuclei were isolated from cells that had been induced to transcribe ISG54 (Figure 2, lane 9, and Figure 3, lane 1). In contrast, nuclei isolated from uninduced cells had the same relative sensitivity as naked DNA (Figure 2, lane 6). When nuclei from induced cells were incubated at 32°C in the presence of ribonucleotides and Mg²⁺, conditions that permit the engaged polymerases to elongate the nascent transcripts (run-on) (Weber *et al.*, 1977), the KMnO₄ sensitivity was lost, indicative of reannealing of the template after the passage of the transcription complex (Figure 2, lanes 10 and 11, and Figure 4, lane 2). However, if α -amanitin was included during the run-on incubation, the KMnO₄ sensitivity was retained (Figure 4, lane 3), demonstrating that the disappearance of KMnO₄ sensitivity is not merely due to the presence of ribonucleotides. The absence of ribonucleotides during the run-on also prevented the loss of KMnO₄ sensitivity (lane 5). Actinomycin D only partially prevented the disappearance of the sensitivity at the concentration used and resulted in the hypersensitivity of new residues, possibly due to some altered structure deriving from actinomycin D intercalation in only selected sites in the DNA template. Taken together, these results show that the single-stranded conformation found downstream of the cap site is indicative of the presence of active RNA polymerases. The polymerase ceases transcription on cell lysis, but resupply of ribonucleotides at 32°C allows continued elongation and hence reannealing of the template after passage of the polymerase complex.

10, and 11) in the presence of 0.06% sarkosyl (lanes 8 and 11). Purified control genomic DNA was treated with KMnO₄ in TE and then processed as the genomic samples (lane 5). A G-ladder was obtained by DMS-piperidine treatment of a ISG54 plasmid (lane 12).

ISG54 lower strand KMnO_4 sensitivity

Figure 3. Homogenous KMnO_4 sensitivity along the ISG54 lower strand. Nuclei isolated from superinduced cells were treated with KMnO_4 as in Figure 2, and the modified residues on the lower strand were mapped by genomic sequencing using primer 2.

We previously have described the sequences in the ISG54 promoter that are occupied by proteins in both whole cells and isolated nuclei (Mirkovitch *et al.*, 1992). Dimethyl sulfate (DMS) treatment of intact cells showed that the ISRE was constitutively occupied by proteins. However, induction of transcription increased the protection against DMS modification. Likewise, DNaseI footprints in isolated nuclei also showed a constitutive protection on the ISRE, and this protection was increased when the gene was induced, indicating a greater occupancy of the regulatory element by some protein(s) during increased transcription (Figure 5, lanes 2 and 3). In addition, DNaseI footprints revealed a weak but constitutive protection of the TATA box region in nuclei from uninduced cells (lane 2). The TATA box and sequences downstream also appeared more protected from DNaseI digestion on transcriptional activation (Mirkovitch *et al.*, 1992) (Figure 5, lanes 2 and 3). We were interested to test the stability of these footprints when polymerases were allowed to further elongate the nascent transcripts. DNaseI footprints appeared identical in samples whether polymerases were or were not allowed to elongate the nascent RNAs (Figure 5, lanes 4 and 5). Thus, the protein-DNA interactions at the ISRE and the TATA regulatory regions were stable after departure of the transcription complex and may possibly help in reinitiation on the chromatin template. This is in agreement with the *in vitro* data of Van Dyke *et al.* (1989) that demonstrated a stable association of TFIID (the TATA box binding protein) and USF, an upstream activating factor, on the adenovirus major late promoter during a transcription initiation cycle.

Paused RNA Polymerase Complex in TTR Gene in Mouse Liver Nuclei

To test for the generality of the structure of open complexes on active genes, we examined the mouse TTR gene in mouse liver nuclei. This gene is transcribed at its highest levels in adults in the liver and the choroid plexus, whereas its transcription is undetectable in most other tissues (Powell *et al.*, 1984). We have provided a detailed description of the DNA-protein interactions at the TTR promoter and enhancer in nuclei derived from liver or spleen (Mirkovitch and Darnell, 1991). To determine if this active chromatin structure was associated with engaged polymerases, we treated liver nuclei with KMnO_4 to map the single-stranded regions around the start site. Unlike the activated ISG54 gene where the KMnO_4 sensitivity was homogeneously distributed downstream of the cap site, only a subset of thymines around the cap site of the TTR gene in liver nuclei appeared sensitive to KMnO_4 modification (Figure 6, lane 5). From thymine residue +8 to thymine +38, all thymines and only thymines were sensitive to KMnO_4 modification, indicating a melted DNA conformation (all thymine residues, sensitive or not, are indicated by their position at the right of the figure). This sensitivity

**KMnO₄ sensitivity indicates the presence of
RNA polymerase complexes**

was not present in naked genomic DNA (lane 1) nor in spleen nuclei (lane 3), a tissue that does not transcribe TTR. In naked DNA, spleen nuclei, and in liver nuclei, the thymine residues flanking the +8 to +38 region presented a similar sensitivity (lane 1). Because the clearly melted region in liver nuclei stopped at +38, it seemed possible that a paused RNA might exist there. Run-on conditions in the presence of nucleotides removed the KMnO₄ sensitivity (lane 6). However, sarkosyl extraction of nuclei in the absence of ribonucleotides did not alter the single-stranded character of the template (lane 7), suggesting little displacement of RNA polymerase II under these conditions (Hawley and Roeder, 1985, 1987).

As described previously (Mirkovitch and Darnell, 1991), the TTR promoter presented a number of regions occupied by proteins in liver nuclei but not in spleen nuclei. These sites corresponded to the TATA region, an NF1 binding site at position -60 from the start site, two sites that can bind the hepatocyte-enriched factor HNF3, and one HNF4 binding site at position -230. The binding of different factors on the TTR promoter was determined in conditions that permitted elongation of the engaged polymerase. Figure 7 presents the DNaseI cleavage pattern of the upper strand of the TTR promoter in spleen nuclei (lane 1) and liver nuclei (lanes 2, 3, and 4) after various treatments. The characteristic footprints over the TATA region, at the NF1, HNF3, and -230 binding sites, were clearly protected from DNaseI digestion in untreated liver nuclei (lane 2) as well as in liver nuclei that were placed in run-on conditions (lane 3). This result indicates that chain elongation by the engaged polymerase complex was not associated with a loss of factors from the promoter and that at least under these conditions the site-specific DNA binding proteins are stably associated with the upstream regions of the TTR region.

DISCUSSION

Determining the position and density of transcription factors and polymerase II complexes in active genes provides an important approach for studying the molecular events associated with transcription. We show

Figure 4. KMnO₄ sensitivity indicates the presence of RNA polymerase complexes. Nuclei were isolated from superinduced HeLa cells and treated with 15 mM KMnO₄ after incubation in different conditions. Nuclei were either directly treated with KMnO₄ (lane 1) or incubated in run-on conditions with all four ribonucleotides (lane 2) to permit elongation of the nascent transcripts from engaged polymerases. If KMnO₄ sensitivity was lost in elongation conditions, it was retained in the presence of the RNA polymerase elongation inhibitor α -amanitin (lane 3) or in the absence of ribonucleotides (lane 5). The presence of actinomycin D, another inhibitor of elongation that intercalates in DNA, partially inhibited elongation as in the conditions used (~50 μ g of actinomycin for 1 mg of DNA) there may not have been enough inhibitor to efficiently inhibit all polymerases.

ISG54 promoter DNaseI sensitivity

here that the mapping of KMnO_4 -induced cleavage sites in whole cells or isolated nuclei can be used to determine at the single nucleotide level the position of engaged polymerases. In addition, the magnitude of the KMnO_4 cleavage is proportional to the transcribing activity and therefore measures the relative density of polymerase complexes. The homogeneous KMnO_4 sensitivity of the ISG54 transcribed sequence suggests that RNA polymerization proceeds at a constant rate along the first 282 nucleotides.

One of the final steps in the assembly of an RNA polymerization complex involves the formation of the so-called open complex, where the DNA is melted at the initiation site. The exact position of this melted region has been described for a number of RNA polymerase complexes in vitro. Studies on RNA polymerase I (Bateman and Paule, 1988), RNA polymerase III (Kassavetis *et al.*, 1990), vaccinia virus (Vos *et al.*, 1991), and RNA polymerase II (Wang *et al.*, 1992) have described melting only in the first 10 nucleotides or less upstream of the start site during the formation of an open complex. However, Zhang and Gralla (1989) have described in vivo KMnO_4 -sensitive region in the 30 nucleotides upstream of the SV40 major late promoter in CV-1 cells late during infection. In our studies, we did not observe any melted structure in vivo in the non-transcribed sequence (at least from position -4). However, the first transcribed nucleotide, a thymine in the lower (template) strand, was sensitive to KMnO_4 modification. It will be necessary to perform in vitro assembly experiments on the ISG54 promoter to determine the exact position of the melting in the open complex. As discussed for the *lac* promoter by Sasse-Dwight and Gralla (1989), the observation of such an open preinitiation complex in vivo may not be possible for some promoters. The detection of an open complex is possible only if the DNA remains melted for a substantial proportion of the initiation cycle. If there is no rate-limiting step for the beginning of RNA synthesis after the formation of the open complex, then the melted structure would appear only transiently and therefore may not be detected. A rate-limiting step is usually created in vitro by the absence of selected ribonucleotides. It will be of interest to analyze a number of eucaryotic promoters to determine which are open, pending some second event, and which are cleared more rapidly after initiation.

Surprisingly, the whole region downstream of the cap site of the ISG54 gene appeared protected from DNaseI

Figure 5. ISG54 promoter DNaseI sensitivity. Nuclei were isolated from uninduced cells (lane 2) or superinduced cell (lanes 3-6) and treated with 20 $\mu\text{g}/\text{ml}$ DNaseI for 1 min at 25°C. DNaseI cleavage sites were mapped by genomic sequencing on the lower strand using primer 2. Before DNaseI digestion, nuclei from transcriptionally active cells were incubated in run-on conditions (lane 5) in the absence of ribonucleotides (lane 4) or the presence of α -amanitin (lane 6).

Figure 6. A paused RNA polymerase just downstream of the TTR start site. KMnO_4 -induced modifications were mapped on the TTR upper strand on purified genomic DNA (lane 1), spleen nuclei (lanes 2 and 3), or liver nuclei (lanes 5 to 8) by genomic sequencing using the primer 1 (Mirkovitch and Darnell, 1991) as described. KMnO_4 modified residues specific to liver nuclei (lane 5) are apparent at thymine residues from position +8 to position +38. All thymine positions are indicated at the right of the figure. The sensitivity was lost if nuclei were incubated in run on conditions prior to KMnO_4 treatment (lane

when the transcription was induced (Figure 7, lane 2). This cannot originate from the presence of polymerase complexes because this DNaseI protection was retained whether or not the polymerases were allowed to elongate the nascent transcripts (Figure 7, compare lanes 2 and 4). However, a small molecule like DMS would methylate to the same extent the transcribed region in both uninduced and induced cells (Mirkovitch *et al.*, 1992). This suggests that induction of transcription transfers the transcribed template in another nuclear domain, possibly a "polymerization center" less accessible to DNaseI. Accessory polymerization cofactors of various types could be involved in such a more appropriate environment for polymerization.

In contrast to the ISG54 gene, a paused polymerase is present downstream of the cap site of the TTR gene in isolated nuclei. The origin of this pause site is unclear. The possibility that RNA polymerization complexes stop preferentially in this region during cell lysis for some unknown reason can not be ruled out. Alternatively, the mouse TTR gene may be similar to some *Drosophila* genes that were shown to have a polymerization pause site close to the cap site (Rougvie and Lis, 1988; 1990). In the *Drosophila* genes, however, sarkosyl or high salt extraction of nuclei is necessary to relieve the block to transcription, which is not the case for the TTR gene. The relevance of such polymerization block to the control of gene expression remains obscure. The block to transcription probably does not result in transcription termination because it is present in a tissue where the gene is expressed at its highest levels. This observation, as well as the work on *Drosophila*, suggests that there may be one or more rate-limiting step(s) after transcription initiation for some genes. Possibly the polymerase complex still interacts with the various transcription factors, and an additional event is necessary for liberating the polymerase complex from the promoter.

The paused polymerase on the TTR gene provides a maximal estimation for the region melted by an RNA polymerase II complex. If it is assumed that the polymerases are paused at the same position in all nuclei; the minimum length of the melted region created by an active RNA polymerase II transcribing complex should be ≥ 30 nucleotides because all thymines from position +8 to position +38 are unpaired. This value is in agreement with the result obtained for the SV40 major late promoter in infected cells (Zhang and Gralla, 1989) but larger than those observed in vitro for other RNA polymerases (Bateman and Paule, 1988; Kassavetis *et al.*, 1990; Vos *et al.*, 1991; Wang *et al.*, 1992). These data have been obtained on the upper strand that does not participate in DNA:RNA hybrids that could result

6). However, it was still present in the absence of ribonucleotides during run on (lane 8) or if the nuclei were simply extracted with 0.06% sarkosyl before KMnO_4 treatment.

Figure 7. TTR promoter DNaseI sensitivity. Spleen nuclei (lane 1), liver nuclei (lanes 2, 3, and 4), or naked plasmid DNA (lane 5) were treated with DNaseI and the cleavage sites mapped by genomic sequencing as described (Mirkovitch and Darnell, 1991). The liver-specific protected regions are marked at the left and correspond to the *in vitro* binding sites of the indicated proteins. These protections were stable if the nuclei were incubated in run-on conditions in the absence of ribonucleotides (lane 3) or the presence of ribonucleotides to allow elongation of the nascent transcripts from engaged polymerases (lane 4).

in protection to KMnO_4 modification and underestimation of the DNA "bubble." A melted region of 30 nucleotides, however, could be an overestimation because polymerases may not pause at the very same nucleotide in all nuclei.

The possibility of precisely mapping the distribution and density of active RNA polymerases on specific sequences in intact cells or isolated nuclei should help understand the complex mechanisms in transcription initiation, elongation, and termination. The ability to study the requirements for reinitiation in isolated nuclei

should provide a means to determine which components are stably associated with chromatin and which are necessary only transiently at each round of transcription initiation. The expression of some genes, such as *c-fos*, *c-myc*, and from the long terminal repeat of the human immuno-deficiency virus 1) (HIV LTR), have been shown to be regulated in some conditions at the level of transcription elongation (reviewed by Spencer and Groudine, 1990; Kerppola and Kane, 1991). Such regulatory mechanisms are generally difficult to demonstrate because careful run-on analyses with hybridization to very short probes are necessary. The assessment of polymerase density by KMnO_4 could help in rapidly characterizing such a regulatory level and could also precisely map the site of elongation block. In addition, the localization of the block by KMnO_4 sensitivity would also circumvent the use of artificial constructs in possibly irrelevant transfection systems.

ACKNOWLEDGMENTS

We thank Chris Schindler, Richard Pine, and anonymous referees for critical reading of the manuscript. We also thank P. Sorte (Hoffman-La Roche, Inc.) for the generous gift of α -interferon and D. Vapneek (Amgen) for providing γ -interferon. This work was supported by the National Institutes of Health grant CA-16006-19 and a training grant AI-07233-15.

REFERENCES

- Bateman, E., and Paule, M.R. (1988). Events during eukaryotic rRNA transcription initiation and elongation: conversion from the closed to the open promoter complex requires nucleotide substrates. *Mol. Cell. Biol.* 8, 1940-1946.
- Bengal, E., Flores, O., Krauskopf, A., Reinberg, D., and Aloni, Y. (1991). Role of the mammalian transcription factor IIF, IIS, and IIX during elongation by RNA polymerase II. *Mol. Cell. Biol.* 11, 1195-1206.
- Borowiec, J.A., and Hurwitz, J. (1988). Localized melting and structural changes in the SV40 origin of replication induced by T-antigen. *The EMBO J.* 7, 3149-3158.
- Borowiec, J.A., Zhang, L., Sasse-Dwight, S., and Gralla, J.D. (1987). DNA supercoiling promotes formation of a bent repression loop in lac DNA. *J. Mol. Biol.* 196, 101-111.
- Buratowski, S., Sopta, M., Greenblatt, J., and Sharp, P.A. (1991). RNA polymerase II-associated proteins are required for a DNA conformation change in the transcription initiation complex. *Proc. Natl. Acad. Sci. USA* 88, 7509-7513.
- Chamberlin, M. (1976). RNA polymerase: an overview. In: *RNA Polymerase*, ed. R. Losick, New York: Cold Spring Harbor Lab., 17-67.
- Darst, S.A., Edwards, A.M., Kubalek, E.W., and Kornberg, R. (1991). Three-dimensional structure of yeast RNA polymerase II at 16 Å resolution. *Cell* 66, 121-128.
- Falvey, E., and Schibler, U. (1991). How are the regulators regulated? *FASEB J.* 5, 309-314.
- Hawley, D.K., and Roeder, R.G. (1985). Separation and partial characterization of three functional steps in transcription initiation by human RNA polymerase II. *J. Biol. Chem.* 260, 8163-8172.
- Hawley, D.K., and Roeder, R.G. (1987). Functional steps in initiation and reinitiation from the major late promoter in a HeLa nuclear extract. *J. Biol. Chem.* 262, 3452-3461.

- Johnson, P.F., and McKnight, S.L. (1989). Eukaryotic transcriptional regulatory proteins. *Annu. Rev. Biochem.* 58, 799–839.
- Kassavetis, G.A., Braun, B.R., Nguyen, L.H., and Geiduschek, E.P. (1990). *S. cerevisiae* TFIIB is the transcription initiation factor proper of RNA polymerase III, while TFIIA and TFIIC are assembly factors. *Cell* 60, 235–245.
- Kerppola, T.K., and Kane, C.M. (1991). RNA polymerase: regulation of transcript elongation and termination. *FASEB J.* 5, 2833–2842.
- Kessler, D.S., Levy, D.E., and Darnell, J.E. (1988). Two interferon-induced nuclear factors bind a single promoter element in interferon-stimulated genes. *Proc. Natl. Acad. Sci. USA* 85, 8521–8525.
- Larner, A.C., Chaudhuri, A., and Darnell, J.E., Jr. (1986). Transcriptional induction by interferon. *J. Biol. Chem.* 261, 453–459.
- Levy, D.E., Kessler, D.S., Pine, R., Reich, N., and Darnell, J.E. (1988). Interferon-induced nuclear factors that bind a shared promoter element correlate with positive and negative transcription. *Genes Dev.* 2, 383–393.
- Levy, D.E., Kessler, D.S., Pine, R., and Darnell, J.E. (1989). Cytoplasmic activation of ISGF-3, the positive regulator of interferon- α -stimulated transcription, reconstituted in vitro. *Genes Dev.* 3, 1362–1371.
- Levy, D.E., Larner, A.C., Chaudhuri, A., Babiss, L.E., and Darnell, J.E., Jr. (1986). Interferon-stimulated transcription: isolation of an inducible gene and identification of its regulatory region. *Proc. Natl. Acad. Sci. USA* 83, 8929–8933.
- Levy, D.E., Lew, J.L., Decker, T., Kessler, D.S., and Darnell, J.E. (1990). Synergistic interaction between interferon- α and interferon- γ through induced synthesis of one subunit of the transcription factor ISGF-3. *EMBO J.* 9, 1105–1111.
- Linn, S.C., and Luse, D.S. (1991). RNA polymerase II elongation complexes paused after the synthesis of 15- or 35-base transcripts have different structures. *Mol. Cell. Biol.* 11, 1508–1522.
- Martin, K.J. (1991). The interactions of transcription factors and their adaptors, coactivators and accessory proteins. *Bioessays* 13, 499–503.
- Maxam, A.M., and Gilbert, W. (1977). A new method for sequencing DNA. *Proc. Natl. Acad. Sci. USA* 74, 560–564.
- Mirkovitch, J., and Darnell, J.E., Jr. (1991). Rapid in vivo footprinting technique identifies proteins bound to the TTR gene in the mouse liver. *Genes Dev.* 5, 83–93.
- Mirkovitch, J., Decker, T., and Darnell, J.E., Jr. (1992). Interferon induction of gene transcription analyzed by in vivo footprinting. *Mol. Cell. Biol.* 12, 1–9.
- Mitchell, P.J., and Tjian, R. (1989). Transcription regulation in mammalian cells by sequence-specific DNA binding proteins. *Science* 245, 371–378.
- Pfeifer, G.P., Steigerwald, S.D., Mueller, P.R., Wold, B., and Riggs, A.D. (1989). Genomic sequencing and methylation analysis by ligation-mediated PCR. *Science* 246, 810–813.
- Powell, D.J., Friedman, J.M., Oulette, A.J., Krauter, K.S., and Darnell, J.E., Jr. (1984). Transcriptional and post-transcriptional control of specific messenger RNAs in the adult and embryonic liver. *J. Mol. Biol.* 179, 21–35.
- Reich, N.B., Evans, B., Levy, D.E., Fahey, D., Knight, E., Jr., and Darnell, J.E., Jr. (1987). Interferon-induced transcription of a gene encoding a 15-kDa protein depends on an upstream enhancer element. *Proc. Natl. Acad. Sci. USA* 84, 6394–6398.
- Roeder, R.G. (1991). The complexities of eukaryotic transcription initiation: regulation of preinitiation complex assembly. *Trends Biochem. Sci.* 16, 402–408.
- Rougvie, A.E., and Lis, J.T. (1988). The RNA polymerase II molecule at the 5' end of the uninduced hsp70 gene of *D. melanogaster* is transcriptionally engaged. *Cell* 54, 789–804.
- Rougvie, A.E., and Lis, J.T. (1990). Postinitiation transcriptional control in *D. melanogaster*. *Mol. Cell. Biol.* 10, 7041–7045.
- Rubin, C.M., and Schmid, C.W. (1980). Pyrimidine-specific chemical reactions useful for DNA sequencing. *Nucleic Acids Res.* 8, 4613–4618.
- Saltzman, A.G., and Weinmann, R. (1989). Promoter specificity and modulation of RNA polymerase II transcription. *FASEB J.* 3, 1723–1733.
- Sasse-Dwight, S., and Gralla, J.D. (1988). Probing the *E. coli* *ginALC* upstream activation mechanism in vivo. *Proc. Natl. Acad. Sci. USA* 85, 8934–8938.
- Sasse-Dwight, S., and Gralla, J.D. (1989). KMnO_4 as a probe for lac promoter DNA melting and mechanism in vivo. *J. Biol. Chem.* 264, 8074–8081.
- Sasse-Dwight, S., and Gralla, J.D. (1990). Role of eucaryotic-type functional domains found in the prokaryotic enhancer receptor factor $\sigma 54$. *Cell* 62, 945–954.
- Sawadogo, M., and Sentenac, A. (1990). RNA polymerase B (II) and general transcription factors. *Annu. Rev. Biochem.* 59, 711–754.
- Spencer, C.A., and Groudine, M. (1990). Transcription elongation and eukaryotic gene regulation. *Oncogene* 5, 777–785.
- Szentirmay, M., and Sawadogo, M. (1991). Transcription factor requirement for multiple rounds of initiation by human RNA polymerase II. *Proc. Natl. Acad. Sci. USA* 88, 10691–10695.
- Van Dyke, M.W., Roeder, R.G., and Sawadogo, M. (1989). Physical analysis of transcription preinitiation complex assembly on a class II gene promoter. *Science* 241, 1335–1338.
- von Hippel, P.H., Bear, D.G., Morgan, W.D., and McSwiggen, J.A. (1984). Protein-nucleic acid interactions in transcription: a molecular analysis. *Annu. Rev. Biochem.* 53, 389–446.
- Vos, J.C., Saker, M., and Stunnenberg, H.G. (1991). Promoter melting by a stage-specific vaccinia virus transcription factor is independent of the presence of RNA polymerase. *Cell* 65, 105–113.
- 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.
- Weber, J., Jelinek, W., and Darnell, J.E., Jr. (1977). The definition of a large viral transcription unit late in Ad2 infection of HeLa cells: mapping of nascent RNA molecules labeled in isolated nuclei. *Cell* 10, 611–616.
- Zhang, L., and Gralla, J.D. (1989). In situ nucleoprotein structure at the SV40 major late promoter: melted and wrapped DNA flank the start site. *Genes Dev.* 3, 1814–1822.