Analysis of Termination and Pausing *in vivo* by RNA Polymerase II on Polyomavirus DNA

by

Richard Wen-chang Tseng

Department of Microbiology and Immunology

McGill University, Montreal, Quebec, Canada

A thesis submitted to the

Faculty of Graduate Studies and Research,

McGill University, in partial fulfillment of the

requirement for the degree of Doctor of Philosophy.

ISBN 0-315-44455-X

ABSTRACT

A novel S1 nuclease mapping method was devised to measure the efficiency and sites of transcription termination in vivo by RNA polymerase II on polyomavirus DNA. To determine termination efficiency, pulse-labeled nuclear RNA was hybridized with a DNA fragment which spans the transcription initiation region. Since polyomavirus DNA is circular, any RNA polymerases which do not terminate transcription will traverse the entire genome and pass through the initiation region again. Hybrids were treated with ribonuclease, bound to nitrocellulose filters, eluted from the filters by degrading single-stranded DNA with S1 nuclease, and analyzed by gel electrophoresis. The results show that about 70% of polymerases terminate transcription per genome traverse at 16 hours post infection; only 50% terminate at 28 hours. At 16 hours, the major transcription initiation site on the L DNA strand is at nucleotide 5050; but at 28 hours, the major initiation site is shifted upstream to nucleotide 5128. RNA polymerase terminates and/or pauses at many discrete sites which lie mostly downstream of the polyadenylation site. Some pausing and/or termination sites are also observed upstream of the polyadenylation site and are probably created by polymerases that traverse the genome more than once. Our experiments also show that the 3' end at nucleotide 1230 found by Triesman and Kamen (1981) in steady state nuclear RNA is created by a post-transcriptional RNA processing event and not by termination of transcription.

RESUME

Nous avons mis au point une nouvelle technique, ayant recours à la nucléase S1, afin de mesurer l'efficacité et de déterminer les sites de terminaison de la transcription par l'ARN polymérase II chez le virus du polyome.

Pour déterminer l'efficacité de la terminaison, nous avons hybridé l'ARN nucléaire "pulse-labeled" avec un fragment d'ADN incluant la région où s'effectue l'initiation de la transcription. Les hybrides ADN-ARN obtenus ont été traités avec la ribonucléase, fixés sur des filtres de nitrocellulose, traités avec la nucléase S1, et enfin analysés par électrophorèse sur gel. Les résultats obtenus nous démontrent qu'environ 70% des polymérases terminent la transcription du génome du virus du polyome 16 heures après l'infection; et seulement 50% terminent la transcription à 28 heures. A 16 heures, le site principal d'initiation de la transcription se situe au nucléotide 5050; mais à 28 heures ce site d'initiation est déplacé en amont au nucléotide 5128. L'ARN polymérase termine ou fait une pause à plusieurs sites précis qui sont localisés dans les 3 kilobases, en aval du site de polyadénylation. Nous avons également observé quelques sites de pause et/ou de terminaison en amont du site de polyadénylation; ces sites sont probablement créés par les polymérases qui ont fait le tour du génome plus d'une fois. Nos résultats expérimentaux démontrent aussi que l'extrémité 3' qui avait été localisée au nucléotide 1230 par Triesman et Kamen (1981) en analysant l'ARN nucléaire "steady state", est crée par coupure endonucléolytique et n'est donc pas dû à la terminaison de la transcription.

ACKNOWLEDGEMENTS

I would like to thank Michael McCleod, Maureen McNally, and Thu-Hang Tran for the construction of single-stranded DNA clones, and Susan Bolla and Daniel Tessier for their excellent technical assistance. I would like to thank Armen Kupelian, Michael Reedijk, and Bill Skarnes for their valuable suggestions. I would like to express my gratitude to Linda Duplessis for translating the abstract, to Fiona Lees for her excellent work typing this thesis, and to Anna Campana for her photographic work. I would like to thank Jacqueline Lanoix for her insightful discussion, corrections and critical reading of this thesis. Finally, I would like to thank my supervisor, Nicholas H. Acheson, for his advice, encouragement, and invaluable assistance throughout this investigation, and for editing this thesis.

CLAIM OF CONTRIBUTIONS TO KNOWLEDGE

1. I have developed an improved S1 nuclease mapping method for mapping *in vivo* pulse-labeled nuclear RNAs, which is potentially useful for studying unstable cellular or viral RNAs.

2. I have devised a strategy to measure the efficiency of transcription termination *in vivo* by RNA polymerase II on polyomavirus DNA. This strategy allows one to determine the proportion of RNA polymerases passing through the transcription initiation region which have not terminated during a previous round of transcription of a circular DNA.

3. I have determined that the efficiency of termination by RNA polymerase II on polyomavirus DNA is about 70% at 16 hours post infection, and 50% at 28 hours post infection.

4. I have used our improved S1 nuclease method to detect the sites of initiation of transcription by analyzing pulse-labeled RNA, thereby avoiding problems of interpretation that arise when steady-state RNA is used.

5. I have established that transcription of the L strand of the AT3 strain of polyomavirus DNA initiates near at least 7 different sites: nucleotides 5050, 5076, 5100, 5128, 5154, 5177 and 5203. In particular, the initiation site located near nucleotide 5076 (the 5' end of the late "leader" region) is a true initiation site as well as being a splicing site.

6. I have shown that at 16 hours post infection, the major initiation site of L strand RNA is at nucleotide 5050; however, at 28 hours post infection, the major site is shifted upstream, to nucleotide 5128.

7. I have shown that little or no processing (splicing or 3' end formation) is detected in nuclear RNA labeled for 1 or 2 minutes in vivo.

8. I have detected multiple 3' ends of pulse-labeled RNA, mostly downstream of the polyadenylation site of the L DNA strand; some 3' ends are also located upstream of the polyadenylation site within the coding region. These 3' ends may represent sites at which RNA polymerase II pauses or terminates transcription.

9. Our results show that the 3' end at nucleotide 1230 found in steady-state L-strand nuclear RNA is probably created by cleavage of a larger transcript and not by termination of transcription.

TABLE OF CONTENTS

LISTS OF FIGURES

LITERATURE REVIEW

Control of transcription termination in specific operons provides an important mechanism of genetic regulation in bacteria, and very likely in eukaryotes as well. Although a great deal is known about transcription termination in bacteria, much less is known in eukaryotes. This review describes what is known about termination signals, factors and the mechanism of termination in bacteria and in eukaryotes.

Termination of Transcription in Prokaryotes.

In prokaryotes, all types of RNA (ribosomal RNA, messenger RNA and transfer RNA) are transcribed by a single RNA polymerase. Efficient transcription of the bacterial genome depends on the ability of the cell to regulate the type and amount of mRNA produced from each of its thousands of genes. This not only requires the information to specify precise starting points for transcription, but stopping points as well. The ends of genes or operons (functional polycistronic units of genes) may be defined by the occurrence of an efficient termination site. This simple housekeeping function prevents transcriptional read-through by RNA polymerase into downstream genes or operons (Wu et al, 1980). In addition, a termination site can be utilized within a regulatory region to alter gene expression by controlling the ability of RNA polymerase to transcribe beyond that site (Adhya and Gottesman, 1978; Rosenberg and Court, 1979).

Thus, sites of transcription termination may serve several purposes in the cell: 1) as the stop signals at the ends of transcription units, they allow adjacent genes and/or operons to be transcribed and regulated independently; 2) as modulating elements in attenuation, they are regulated to permit the elongation or termination response to be sensitive to the

requirements of the cell for the products of genes located distal to the attenuator; 3) as conditional abortive elements in mutational polarity, they prevent wasteful depletion of cellular metabolites.

Transcription termination is a very complicated event involving the following: cessation of transcription, release of the completed RNA molecule and dissociation of the RNA polymerase from the DNA template. Analysis of the primary structure of termination regions has revealed common features that may be involved in the recognition of termination (Adhya and Gottesman, 1978). The stretch of DNA preceeding a transcription termination site is often GC-rich and usually contains a dyad symmetry element. As a result, the RNA transcripts can form a hairpin structure. These transcripts typically end in a series of uridine residues. In addition, the protein factor rho (Roberts, 1969) seems to be involved in termination *in vivo*. With the use of an *in vitro* transcription system, two different types of termination sites have been distinguished. Certain terminators can function efficiently with purified RNA polymerase alone and are called factor-independent terminators; others require in addition the protein factor rho; these are called factor-dependent terminators. This classification must be considered loosely, because the response at termination sites *in vitro* may be profoundly different from the response *in vivo*, complicating interpretations and predictions about their behavior.

1. Factor-independent Termination.

Terminators that can function efficiently with purified RNA polymerase holoenzyme and that do not require additional factors (such as rho protein) are called factor-independent terminators. In general, rho-independent termination sites consist of a GC-rich segment of dyad symmetry centered about 20 nucleotides before the stop site, followed by a stretch of 4-8

consecutive uridine residues in the transcript (Adhya and Gottesman, 1978). The terminated transcript ends within, or just distal to, the oligo (rU) sequence.

A current model proposes that this dyad symmetry is responsible for the formation of a hairpin structure in the nascent RNA, causing the polymerase to pause and disrupting the RNA-DNA hybrid helix (Rosenberg and Court, 1979; Farnham and Platt, 1980). The consecutive uridine residues near the 3' end of the RNA facilitate dissociation of the transcript from the template (Farnham and Platt, 1980; Martin and Tinoco, 1980) since RNA-DNA hybrids with rU-dA pairing are exceptionally unstable (Chamberlin, 1965; Riley et al, 1966).

Studies on transcription units which have been mutated at the GC-rich stem-loop structure region and the consecutive T-residues support this model. In general, mutations that decrease the predicted stability of the RNA hairpin structure also decrease termination efficiency, whereas mutations that increase the predicted stability of the hairpin structure increase termination efficiency (Rosenberg and Court, 1979; Zurawski and Yanofsky, 1980). Accordingly, mutations that reduce the number of terminal uridines also decrease the termination efficiency; only minimal termination is observed with a sequence of four uridines (Bertrand et al, 1977; Zurawski and Yanofsky, 1980). Using synthetic sites for transcription termination, Christie et al (1981) have shown that under the normal conditions of transcription in vitro, termination became more efficient with an increase in the length of the stem in the RNA hairpin or an increase in the number of consecutive uridine residues.

Incorporation of base analogs into the nascent transcript and the DNA template confirm the same findings. Base analogs, such as iodo-CTP and bromo-CTP, which form stronger GC base pairs in RNA stem-loop structures

increase termination efficiency. On the other hand, base analogs like inosine-5'-triphosphate (ITP), which can replace GTP and form I:C base pairs which are weaker than normal G:C base pairs, decrease termination efficiency. And base analogs such as bromo-UTP and allylamine-UTP, which form stronger base pairs with dA residues than normal UTP, elevate readthrough of termination sites (Neff and Chamberlin, 1978; Adhya et al, 1979; Farnham and Platt, 1980). Farnham and Platt (1982) have devised a method to incorporate deoxyribonucleotide base analogs in vitro into either strand of the tryptophan operon attenuator region. The results showed that base analogs which strengthen A:T or A:U base pairing enhance readthrough; but analogs which increase the strength of G:C base pairing in DNA have no effect on termination.

2. Factor-dependent Termination

Some in vivo termination signals are not recognized in vitro by purified RNA polymerase alone, but require additional factors to function efficiently. These terminators are called factor-dependent terminators. Rho was discovered by Roberts (1969) as an E. coli factor that facilitates termination of transcription at specific sites in vitro and in vivo on a wide variety of DNA templates from E. coli and various bacteriophages (Roberts, 1979; Adhya and Gottesman, 1978). Genetic studies on rho mutants confirm the requirement for rho as a transcription termination factor in vivo, and its activity is essential for cell growth (Richardson et al, 1975; Das et al, 1976).

Investigations of biochemical properties of rho have shown that it can bind to single-stranded RNA and has an RNA-dependent nucleotide triphosphatase (NTPase) activity that appears to be necessary for rho-dependent termination. The rho-NTPase activity has been studied

extensively on the uncoupled system, that is the NTPase activity stimulated by RNA alone in the absence of RNA transcription and termination of transcription. Synthetic polyribonucleotides containing cytosine are required for the rho-NTPase acitivity and chain length has to be greater than 50 nucleotides. Poly(dC) binds tightly to rho but does not stimulate the NTPase activity by itself. However, single-stranded poly(dC), when supplemented with small RNA fragments (poly[rC]), is capable of stimulating the rho-NTPase activity (Richardson, 1982). This suggests that rho-dependent termination may involve recognition of DNA sequence. Random copolymers of poly(rU, rC) which contain only 5% rC and poly(rA, rC) which contain as little as 10% rC are also efficient activators (Lowrey and Richardson, 1977a,b). Moreover, in vitro studies show that rho has a strong affinity for cytosine-containing RNA, especially with the highest affinity for the synthetic poly(rC). The binding of rho to RNA appears to correlate with the stimulation of the NTPase acitvity (Galluppi and Richardson, 1980). Therefore, rho may function by binding to either a specific site on DNA or to a site on the nascent mRNA, subsequently stimulating polymerase to terminate. In addition, rho does not bind to double-stranded RNA, while conditions that destabilize RNA secondary structure stimulate NTP hydrolysis (Richardson and Macy, 1981).

Using ATP analogs that have nonhydrolyzable β-γ bonds not only inhibits the rho-NTPase activity but also abolishes rho-dependent termination (Galluppi et al, 1976; Howard and deCrombrugghe, 1976). Genetic studies have shown that the rho mutants which are defective in termination and ATPase activity can be compensated by particular RNA polymerase mutations (Das et al, 1978; Guarente and Beckwith, 1978). This indicates the possible role of rho-RNA polymerase interaction which may link to termination-NTPase activity. RNA appears to be an essential

cofactor in both the termination-coupled and -uncoupled NTPase acitivity, and hydrolysis of ATP is coupled to termination of transcription. A possible role for the RNA-dependent ATPase activity is in translocation of a rho molecule along nascent RNA in the search for a RNA polymerase molecule paused at a termination site (Galluppi and Richardson, 1980). However, this hypothesis has not yet been proved.

Current studies of the rho-dependent termination site in the tryptophan operon, trp t', reveal several interesting features. Trp t' is located about 250 nucleotides past the last structural gene trpA, in a region that has few G-residues and is rich in AT-residues (Wu et al, 1981). In vitro, rho-dependent termination at trp t' is tightly coupled to the NTPase activity. Deletion of the trp t' region abolishes rho-dependent termination and rho NTPase activity. Transcription termination and rho NTPase activity are also diminished when ribonuclease T1 is present in the transcription reaction. These observations indicate that a specific RNA component within the trp t' transcript is required for NTPase activation and rho-dependent transcription termination (Sharp and Platt, 1984).

Unlike the rho-independent terminators, analysis of sequences of several rho-dependent terminators reveals very little sequence homology. In rho-dependent termination, conditions which favor reduced secondary structure in RNA are correlated with an increase in the NTPase activity and termination efficiency (Das et al, 1978; Richardson and Macy, 1981); in contrast, with rho-independent termination, enhancement of secondary structure in the RNA stem-loop results in increased termination (Rosenberg and Court, 1979). The trp t' region has a very low G content; this reduces the potential for the formation of stable secondary structures in the RNA transcript (Sharp and Platt, 1984). Therefore rho may require a region of RNA that is unstructured to catalyze termination of transcription.

3. Attenuation.

The expression of many prokaryotic amino acid biosynthetic operons is regulated by a phenomenon known as attenuation (Kolter and Yanofsky, 1982). Attenuation is the termination of transcription at the beginning of an operon, hence preventing distal gene expression. The sites at which this kind of termination occurs are called attenuators. They have the features of rho-independent terminators (GC-rich dyad symmetry followed by a stretch of U residues) and are located at the end of a 100-200 nucleotide leader RNA immediately preceeding the first structural gene of the operons (Kolter and Yanofsky, 1982). In vitro, the attenuator can terminate transcription efficiently with RNA polymerase alone; however, under in vivo conditions, factors can modulate the termination response and permit a certain fraction of the RNA polymerases to continue transcribing into the structural genes.

The tryptophan operon of E. coli has been studied extensively as a rho-independent terminator and as an attenuator (Bertrand et al, 1977; Oxender et al, 1979; Farnham and Platt, 1980; Yanofsky, 1981). The tryptophan attenuator is located within the transcribed 162 base pair leader region that precedes the structural genes of the operon. Transcription of this leader region is either terminated to give a 140-nucleotide leader transcript or allowed to proceed into the structural genes. This leader RNA codes for a 14-amino acid peptide containing two adjacent tryptophan residues (Lee et al, 1978). In the leader, there are four regions of base-pairing, capable of forming three stem and loop structures. Region 1 (nucleotides 54-68) which includes the tandem trp codons and the leader peptide translation stop codon, can form base-pairs with region 2 (nucleotides 76-91) which is immediately distal to the translation stop codon. Region 2 (nucleotides 74-85) is capable of base-pairing with region 3 (nucleotides 108-119). Region 3 (nucleotides

114-121) can also form base-pairing with region 4 (nucleotides 126-134). And region 4 is followed immediately by seven consecutive U residues (nucleotides 134-140) within which lies the 3' end of the leader RNA (Oxender et al, 1979).

Termination of transcription at the attenuator is regulated by the levels of charged and uncharged trp-tRNA, and translation of the leader RNA is an essential feature of this regulation. In the presence of abundant tryptophan, termination at the attenuator is more efficient. This is because the ribosome translating the newly transcribed leader RNA is able to synthesize the complete leader peptide. During this synthesis the ribosome masks regions 1 and 2 of the RNA and prevents the formation of stem-loop 1-2 or 2-3. Regions 3 and 4 are then free to form a stem-loop structure and signal the transcribing polymerase to terminate transcription at the stretch of U residues; the downstream trp operon structural genes will not be transcribed. Under conditions of tryptophan starvation, the ribosome stalls at the adjacent trp codons in the leader peptide coding region due to lack of charged trp-tRNA. Because only region 1 is masked, stem-loop 2-3 is then formed as regions 2 and 3 are transcribed. Formation of stem-loop 2-3 excludes the formation of stem-loop 3-4, which is required to signal transcription termination. Therefore, RNA polymerase is able to continue to transcribe the downstream structural genes, and hence lead to the production of tryptophan.

4. Antitermination.

In bacteriophage λ some protein factors have been shown to participate in transcriptional regulation as antitermination factors. Two major antitermination factors have been identified as the products of the N and Q genes (Greenblatt, 1981; Grayhack and Roberts, 1982), and their function is

critical for the lytic/lysogenic state in the life cycle of λ. There are several transcriptional terminators within each of the early operons of λ, such at tL1, tL2, tL3 of the leftward transcription and tR1, tR2 of the rightward transcription. Following λ infection, transcription by E. coli RNA polymerase is initiated at the P_L and P_R promoters, and is terminated at rho-dependent terminators tL1 and tR1 respectively. The leftward reading transcript codes for the N protein. This N protein functions as an antiterminator to prevent transcription termination of all the early operon terminators, thus allowing RNA polymerase to transcribe the λ genes that are essential for growth of the phage and that are located distal to the various terminators.

The N protein of λ stimulates antitermination at both rho-dependent and rho-independent terminators. For N protein to work, however, a variety of other phage and host functions appear to be required. One phage function is to provide nut (N utilization) sites in DNA or RNA that are recognized by N protein. Two nut sites, nutL and nutR, have been identified. They contain an identical region of dyad symmetry and a common sequence located 8-9 base pairs upstream of this dyad element. This common sequence has been called "box A" (Salstrom and Szybalski, 1978; deCrombrugghe et al, 1979; Friedmann and Olson, 1983). Molecular analysis of the nut sites has demonstrated that dyad symmetry and "box A" are crucial for the function of N-mediated antitermination. In addition, a number of host-encoded factors (NusA, NusB, NusC and NusE) are known to participate in the antitermination events (Friedman et al, 1981; Greenblatt and Li, 1981a,b).

It has been suggested that N protein functions directly on those RNA polymerases that pass through the nut site, altering them in such a way that they fail to terminate at distal transcription termination sites. The

nut site need not be located next to a promoter or a terminator implying that the events occurring at the nut site take place independently of transcription initiation or termination. The host protein factor NusA has been shown to bind tightly to N and is required for its action, perhaps by allowing N to bind to RNA polymerase (Greenblatt and Li, 1981b). The NusA protein also binds directly to RNA polymerase core enzyme (Greenblatt and Li, 1981a) and the "box A" sequence seems to be required for NusA protein to exert its effect in antitermination (Friedman and Olson, 1983). It is proposed that during transcription, RNA polymerase passes the "box A" element of the nut site. At this point the RNA polymerase undergoes a conformational change, which allows the binding of NusA and the subsequent binding of N protein to NusA. This modification renders the RNA polymerase unresponsive to termination signals.

Genetic studies of mutants have revealed that several host-encoded factors, the nus gene products, are also required in N-mediated antitermination. These include NusB, a 15.5 kd protein essential for bacterial growth (Strauch and Friedman, 1981); NusC, probably identical with the β subunit of RNA polymerase (Friedman et al, 1983); and NusE, identical to the ribosomal protein S10 (Friedman et al, 1981). The functions of these Nus proteins are still not understood.

Another λ phage protein, Q, has been shown to mediate antitermination at the λ 6S terminator in the presence of rho and NusA protein (Grayhack and Roberts, 1982). This protein seems to function in a manner similar to N protein, although there is no apparent sequence homology between the two proteins. Besides NusA protein, Q protein requires a qut (Q utililization) site for its antitermination function. The qut site is similar to nut sites, containing a "box A" sequence and a dyad symmetry region. Unlike N

protein, Q protein can exert its antitermination effect in a purified transcription system containing RNA polymerase, NusA protein, and a DNA template encoding the λ 6S transcript (Grayhack and Roberts, 1982).

Termination of Transcription in Eukaryotes.

The transcription process is much more complex in eukaryotes than it is in prokaryotes. Transcription takes place in the nucleus; the mature RNAs are transported to the cytoplasm after synthesis and processing events. Three different RNA polymerases have been identified, each one of which transcribes different classes of genes. They are designated RNA polymerases I, II, and III (for review, see Chambon, 1975; Roeder, 1976). RNA polymerase I is not affected by α-amanitin, a toxin of the poisonous mushroom Amanita phalloides, and is localized in the nucleolus (Roeder and Rutter, 1970). This enzyme is responsible for synthesis of ribosomal RNA (rRNA) (Roeder and Reeder, 1972). RNA polymerase II is completely inhibited by very low concentrations of α-amanitin (1 μg/ml) and transcribes protein coding genes which give rise to messenger RNA (mRNA). RNA polymerase III is inhibited by high concentrations of α-amanitin (100 μg/ml) and is responsible for the synthesis of transfer RNA (tRNA) and 5S RNA (Weinmann and Roeder, 1974). In recent years, cell-free systems have been developed that allow each of these 3 classes of RNA polymerase to transcribe each different class of genes accurately and specifically (Birkenmeier et al, 1978; Weil et al, 1979; Manley et al, 1980; Grummt, 1981). These systems provide a powerful tool for understanding the mechanism of transcription.

A. Termination of Transcription by RNA Polymerase III.

Termination of transcription by RNA polymerase III has been studied extensively on Xenopus 5S RNA. Repeating units of 5S ribosomal DNA of Xenopus laevis and Xenopus borealis have been isolated and sequenced (Fedoroff and Brown, 1978; Korn and Brown, 1978; Miller et al, 1978). Individual repeating units of cloned Xenopus 5S DNA can be faithfully transcribed either when injected into oocytes (Brown and Gurdon, 1978) or in oocyte nuclear extracts (Birkenmeier et al,1978). Using an oocyte nuclear extract for in vitro transcription, Bogenhagen and Brown (1981) have shown that termination of Xenopus 5S DNA transcription occurs within a cluster of four or more consecutive T residues in the 3' noncoding sequence. Studies of deletion mutants indicated that the sequences surrounding the T cluster can be varied widely without impairing termination. However, the nucleotide sequences adjacent to the T cluster influence the efficiency of termination. Efficient termination is observed whenever GC-rich sequences surround the T cluster, reminiscent of the prokaryotic rho-independent termination signal. However, in contrast to the termination site in prokaryotes, neither a dyad symmetry element near the end of the gene, nor a secondary structure in the transcript appear to be required in this case. Furthermore, purified RNA polymerase III from Xenopus laevis alone can terminate transcription at precisely the same site that is recognized in vivo and in complete transcription extracts (Cozzarelli et al, 1983). The purified polymerase also distinguishes between weak and strong terminator sequences with the same relative efficiency as the enzyme in complete extracts. Thus the information required for termination resides in the ability of RNA polymerase III to recognize a simple nucleotide sequence (such as a cluster of four or more consecutive T residues) at the end of the 5S gene.

B. Termination of Transcription by RNA Polymerase I.

In most eukaryotic species, 100-5,000 copies of ribosomal RNA (rRNA) genes per haploid genome are arranged in tandem arrays of repeated units (for review, see Long and Dawid, 1980). These genes are transcribed solely by RNA polymerase I (Roeder and Reeder, 1972). The primary rRNA transcripts are very large 35-47S (or 6-15 kb) precursor molecules which are processed into the mature 28S, 18S and 5.8S RNAs (Bowman et al, 1983) and incorporated into the ribosome.

Transcription of ribosomal DNA (rDNA) has been studied extensively (for review, see Sollner-Webb and Tower, 1986). It has been reported that the 3' end of the pre-rRNA of Xenopus laevis (Sollner-Webb and Reeder, 1979; Bakken et al, 1982) and Drosophila melanogaster (Mandal and Dawid, 1981) is coincident with the mature end of 28S RNA. The 3' end of the pre-rRNA of yeast (Veldman et al, 1980) and mouse (Kominami et al, 1982) appears to extend a few nucleotides into the 3' spacer. However, more recent studies have shown that RNA polymerase I does not stop transcription at the end of 28S RNA and that the correct 28S RNA terminus is generated by a processing event and not by termination (Grummt et al, 1985a, 1985b; Labhart and Reeder, 1986).

In nuclear run-off experiments using X. laevis rDNA, Labhart and Reeder (1986) have recently shown that transcription of the rRNA gene does not terminate at site T1 which was believed to be the termination site; instead, transcription continues past T1, into T2, another site located 235 nucleotides downstream of T1, and terminates at site T3 located 215 nucleotides upstream of the next rRNA start point. Site T3 seems to be the only site at which RNA polymerase I is released from the DNA template. Sites T1 and T2 are sites at which the primary transcript is rapidly processed to expose the mature 3' ends of 28S rRNA. A conserved 7 bp

element (GACTTGC), called the T3 box, located about 200 bp upstream of the next pre-rRNA initiation site is an essential part of the terminator for Xenopus RNA polymerase I (McStay and Reeder, 1986). In addition to its terminator activity, the T3 region can also act as an upstream element of the promoter for the adjacent gene and stimulate its function. Mutations within the T3 box eliminate the terminator activity and cause significant readthrough. The same mutations also eliminate the promoter stimulating function, suggesting that the T3 box is important for both activities.

In the mouse, Grummt et al (1985a) have shown that RNA polymerase I terminates transcription of rDNA 565 bp downstream of the 3' end of mature 28S rRNA, both in vivo and in vitro. This termination site is located immediately upstream of eight tandemly arranged repeated sequences. This repeated domain consists of 18 highly conserved nucleotides that contain a SalI restriction site (called the SalI box) surrounded by stretches of pyrimidines. In vivo, transcription terminates 25 bp upstream of the first SalI site (T1), within the pyrimidine-rich region. This site is also preferentially used in vitro at low template concentrations in a nuclear extract system. At increasing DNA concentrations a termination site (T2) upstream of the second SalI box is used. Competition experiments have suggested that the interaction of some factor (or factors) with the repetitive element in the 3' spacer is required for transcription termination by mouse RNA polymerase I. Exonuclease III protection experiments reveal the binding of a nuclear protein to the SalI box (Grummt et al, 1986b). Deletions, insertions, or point mutations within the SalI box eliminate or severely reduce transcription termination. Therefore, termination appears very specific and requires a repeated 18 bp SalI box that is present eight times in the spacer region downstream of the 3' end of pre-rRNA (Grummt et al, 1986b).

In addition, Grummt et al (1986a) have demonstrated the presence of a transcription termination sequence extending from -171 to -154 with respect to the pre-rRNA initiation site. This sequence (termed T0) exhibits striking structural and functional similarities with the conserved 18 bp SalI box. This upstream terminator (T0) is recognized by the same protein factor as the 3' terminal sites and is able to cause termination by RNA polymerase I in vitro; it also increases the efficiency of initiation of the adjacent promoter. Henderson and Sollner-Webb (1986) have obtained similar results for the mouse rRNA gene. An upstream promoter element also functions as a transcription terminator, although in this case, it is the upstream termination site that enhances promoter activity. The exact mechanism for this dual role is not yet understood.

C. Termination of Transcription by RNA Polymerase II.

There are many differences between prokaryotic and eukaryotic mRNA synthesis. In prokaryotes, the mRNA is polycistronic: that is, a single mRNA can code for several proteins. In eukaryotes, most mRNAs are monocistronic: that is, a single mRNA can code for only one protein. Eukaryotic mRNAs are capped at their 5' ends. This cap structure is a 7-methyl guanine which is added to the 5' terminal triphosphate shortly after initiation of transcription (Salditt-Georgieff et al, 1980). Messenger RNAs are internally spliced and methylated, and a 100-200 nucleotide long polyadenylic acid tail is added to their 3' ends by a poly-A synthetase. Once the fully processed eukaryotic mRNA is formed, it is then transported to the cytoplasm where translation takes place. In contrast, bacterial mRNA does not undergo these processing events, and begins to be translated even before transcription is completed.

Transcription termination appears to be a very important event in prokaryotic gene expression as previously described. It is not known whether transcription termination plays an important role in the regulation of eukaryotic gene expression. However, studies have shown that some eukaryotic RNA polymerases stop shortly downstream of the initiation site and therefore prevent the transcription of the entire mRNA. This phenomenon, which is called premature termination, resembles attenuation in prokaryotes. The production of prematurely terminated RNA was first discovered by researchers using the nucleoside analog 5,6-dichloro-1-beta-D-ribofurano-sylbenzimidazole (DRB) (Egyhazi, 1974). The DRB-resistant RNA molecules are about 100 to 400 nucleotides long (Tamm, 1977), their synthesis is sensitive to low concentrations (1 μg/ml) of α-amanitin and they have a cap structure at their 5' termini (Tamm et al, 1980). This indicates that the DRB-resistant RNA molecules are transcribed by RNA polymerase II. Premature termination has been found in HeLa cells (Segal et al, 1976), adenovirus (Fraser et al, 1978; Maderious and Chen-Kiang, 1984), SV40 (Laub et al, 1980), and polyomavirus (Montandon and Acheson, 1982).

Studies of transcription of SV40 late mRNA, adenovirus early and late mRNAs and mouse β-globin mRNA have revealed that RNA polymerase traverses the polyadenylation site (the 3' end of mature mRNA) and terminates up to several kilobases downstream (Ford and Hsu, 1978; Nevins and Darnell, 1978; Nevins et al, 1980; Hofer and Darnell, 1981). The 3' end of the mature mRNA is created by endonucleolytic cleavage of a longer precursor RNA subsequently followed by addition of polyadenylic acid. A highly conserved hexanucleotide AAUAAA is found 10-30 nucleotides upstream of most poly(A) sites (Proudfoot and Brownlee, 1976), and deletion of this sequence eliminates polyadenylation (Fitzgerald and Shenk, 1981). The following sections summarize what is known about termination of transcription by RNA polymerase II.

1. Yeast.

A spontaneous mutant cyc1-512 has a 38 base pair deletion in the 3' nontranslated region of the CYC1 locus in the yeast Saccharomyces cerevisiae. The deletion is located 130 base pairs downstream of the CYC1 translation termination codon, and about 15 base pairs upstream of the polyadenylation site. The sequence TATTTAT is found at both ends of the deleted region in CYC1; as a result of the deletion, only one copy of TATTTAT is left in cyc1-512 mutant. This mutant produces discrete sizes of CYC1 mRNAs ranging from 630 to 2400 nucleotides, while wild type CYC1 mRNA is approximately 630 nucleotides long including the poly(A) tail (Zaret and Sherman, 1982). This suggests that either 1) this region contains a termination signal, deletion of which causes RNA polymerase to read through the termination site and terminate further downstream; or 2) this region contains a polyadenylation signal, deletion of which leads to polyadenylation at aberrant sites.

It is known that AAUAAA is required for polyadenylation in higher eukaryotes (Fitzgerald and Shenk, 1981). An examination of the DNA sequence of the 3' non-translated region of CYC1 reveals no sequence related to AAUAAA in CYC1 mRNA. Analysis of a number of yeast genes that code for mRNAs indicates that most yeast mRNAs do not contain an AAUAAA-type sequence in the appropriate region. This suggests that signals other than AAUAAA may be used for polyadenylation in yeast. Analysis of steady-state polyadenylated [poly(A)$^+$] RNAs from total cellular RNA of cyc1-512 reveals that all the aberrantly sized RNAs are polyadenylated. In contrast to higher eukaryotes, yeast histone mRNA is polyadenylated (Fahrner et al, 1980). It appears that all or nearly all mRNAs become polyadenylated in yeast. This leads to the conclusion that in yeast polyadenylation may be tightly coupled to transcription termination (Zaret and Sherman, 1982).

Since only steady-state RNA was analyzed in this work, it is possible that non-polyadenylated RNA is degraded so rapidly that it cannot be detected in steady-state RNA. Therefore, one cannot rule out that the 38 bp region could be a processing signal for polyadenylation rather than a termination signal.

The mutant cyc1-512 not only produces several discrete sizes of CYC1 mRNAs, but also the combined steady-state levels of all the CYC1 mRNAs are approximately 10% of the normal amount of 630-nucleotide wild type CYC1 mRNA. Interestingly enough, there is a gene (OSM1) located downstream of the CYC1 locus that is transcribed from the opposite strand and terminates near the 3' end of CYC1. This gene normally produces a 1450-nucleotide transcript, but in cyc1-512, this transcript is absent and is replaced by a 2000-nucleotide transcript. The steady-state level of this 2000-nucleotide transcript in cyc1-512 is reduced by a factor of about five when compared to the amount of wild type 1450-nucleotide transcript. These results suggest that the 38 bp deletion affects the 3' ends of CYC1 and OSM1 transcripts so that they do not terminate and polyadenylate at the correct positions, and the transcription of both genes is now overlapping. The reduced amount of mRNA transcribed from these two genes in the mutant could be the result of collisions between RNA polymerases transcribing the same region in opposite directions. Alternatively, the addition of sequences beyond the normal 3' end of the mRNAs may affect the stability of these aberrant mRNAs. These results suggest that transcription termination may play an important role in the regulation of gene expression in yeast.

Since the 38-base pair region deleted in cyc1-512 shows functional importance, it is likely that some sequences within this region are important for transcription termination and/or polyadenylation. Comparison of 3' non-translated regions of CYC1 and 14 other yeast genes has revealed

that the sequences TAGT or TATGT, which are located in the center of the 38-base pair region, showed homology in 13 out of 14 yeast genes examined (Zaret and Sherman, 1982). This homologous sequence is located 10 to 40 nucleotides upstream of the known poly(A) site. A sequence TAG upstream of the TATGT, and a sequence TTT downstream from the TATGT are often seen in these genes, suggesting that this tripartite structure could form the signal for transcription termination and polyadenylation. The 3' nontranslated regions of yeast genes are extremely AT-rich, and no dyad symmetry structure can be formed in this region, resembling the rho-dependent terminators in prokaryotes (Platt, 1981). Zaret and Sherman (1982) therefore suggest that the region deleted in cyc1-512 may represent a eukaryotic counterpart to the rho-dependent bacterial terminator.

Recently, Henikoff and Cohen (1984) investigated sequences responsible for transcription termination in the yeast Saccharomyces cereviseae. They studied termination of transcription on a segment of Drosophila melanogaster DNA that complements an adenine-8 (ADE8) mutation in Saccharomyces cerevisiae and codes for the enzyme glycinamide ribonucleotide transformylase (Henikoff et al, 1981). The Drosophila ADE8 gene was cloned into the BamHI site of yeast plasmid YEp13 such that the 3' portion of the FLP gene of 2 μm plasmid DNA is oriented in the same direction as the ADE8 gene. This Drosophila ADE8 message in yeast is about 1 kilobase (kb) in length (Henikoff et al, 1981). Although the initiation sites of ADE8 mRNAs isolated from yeast and Drosophila cells differ considerably, the 3' end of these mRNAs are nearly the same.

Henikoff et al, (1983) have constructed deletions in the vicinity of the 3' end of ADE8 gene to determine the 3' boundary of a possible termination control signal. A fusion of ADE8 (about 1 kilobase) to about 0.5 kb of the FLP gene should increase the length of the message to about

1.5 kb if the putative ADE8 termination signal is deleted. Analysis of poly(A)$^+$ RNAs extracted from cells that carry various deleted plasmids has revealed that messages from deletions 712 and 753 are about 1.5 kb long indicating fusion transcripts, whereas those from deletions 782, 802 and 822 are about 1 kb long indicating unaltered termination. These numbers represent the number of nucleotides between the AUG start codon in ADE8 mRNA and the beginning of the deletion. The sequence AATAAA is at nucleotides 816-821. This sequence can be deleted without affecting the formation of 1 kb mRNA suggesting that in yeast this sequence is not required for polyadenylation. The termination control signal must be located between nucleotides 753 and 782. Two 8 base pair repeats, TTTTTATA, separated by a 17-nucleotide interval, may be important for this termination control. Further 3' end deletion analysis reveals that inserts which contain up to nucleotide 769 show fully normal termination, but deletion up to nucleotide 768 produces about 50% readthrough. The deletion 769 leaves TTTTTATA intact, while deletion 768 removes the last A-residue. Therefore, the 3' boundary of the transcription termination signal is the 3' boundary of TTTTTATA sequence.

To locate the 5' boundary of this signal, 5' end deletions from nucleotide 624 in the 3' direction were constructed (Henikoff and Cohen, 1984). Analysis of mRNA from these deletion plasmids shows that removal of sequences between 624 and 748 allows fully normal termination, while removal of sequences beyond this point shows increasing amounts of readthrough transcripts. From these results, the authors suggest that the termination signal is no more than 21 base pair long (from nucleotides 748 to 769), and part of the signal is the sequence TTTTTATA, which is seen in the 3' nontranslated region of other yeast genes. Sequences downstream of the 21 base pair signal sequence are also required for efficient

termination. Analysis of these downstream sequences within yeast mRNAs showed that there is not much homology among different genes; however, all such regions are AT-rich. Henikoff and Cohen (1984) came up with the same hypothesis as Zaret and Sherman (1982), that the mechanism for transcription termination in yeast may resemble bacterial rho-dependent termination (Platt, 1981).

2. Sea Urchin Histone Genes.

The histone gene repeat unit of the sea urchin Psammechinus miliaris has been isolated and cloned (Gross et al, 1976). This histone gene unit contains all 5 histone mRNA genes, H1, H2A, H2B, H3 and H4. Injection of cloned sea urchin histone genes into Xenopus oocyte nuclei produces high levels of H2A and H2B mRNAs indistinguishable from authentic sea urchin histone mRNAs, and low levels of H1, H3 and H4 mRNAs (Hentschel et al, 1980b). S1 nuclease mapping has shown that authentic 5' ends of all 5 sea urchin histone mRNAs are produced in injected frog oocyte nuclei; however, 5' ends of H4 mRNA are at low levels. Correct 3' ends are found for H1, H2A, H2B and H4 mRNAs, but not for H3 mRNAs. There is a deficiency in the production of the correct 3' terminus for the H3 mRNA, presumably as a result of poor processing.

Although histone genes are transcribed by RNA polymerase II, most histone mRNAs are not polyadenylated. The conserved AAUAAA motif found in polyadenylated mRNA is missing from histone mRNAs. Instead, a 23 base-pair conserved DNA sequence has been identified at the 3' ends of histone genes in many different species (Hentschel and Birnstiel, 1981). This sequence contains a 16 base-pair hyphenated inverted repeat and the 3' termini of histone mRNAs lies near an ACCA motif (Hentschel et al, 1980a). It has

been suggested that this homology sequence represents a regulatory signal either for transcription termination or maturation of a longer precursor (Busslinger et al, 1979).

To understand the possible function of this 23 base-pair conserved sequence a series of mutants were constructed in which deletion, insertion or point mutations were introduced into the H2A gene of the cloned sea urchin histone gene repeat unit h22 (Birchmeier et al, 1982, 1983). These mutated h22 DNAs were injected into Xenopus oocyte nuclei and histone mRNAs were analyzed. The results have shown that sequences immediately upstream of the conserved inverted repeat, as well as protein coding sequences of the H2A gene, can be removed without any serious effects on transcription initiation and 3' end formation. The inverted DNA repeat at the 3' end of the H2A gene is required for the generation of authentic 3' ends of H2A mRNA. A 12 base pair deletion which removes this palindrome and point mutations which destroy the symmetry of the inverted repeat prevent the formation of 3' termini. However, the presence of the terminal DNA repeat with the ACCA motif by itself is not sufficient for the generation of authentic 3' H2A mRNA termini; additional spacer sequences, about 80 base-pairs downstream of this region are also required. Transcription analysis of point mutations which prevent 3' end formation and pseudorevertants, which allow the production of H2A mRNA with authentic 3' ends, have revealed that an RNA hairpin structure rather than a DNA cruciform structure is essential for the formation of the 3' termini.

In vitro synthesized transcripts of the H2A gene with 3' extensions are rapidly processed into H2A mRNA with authentic 3' ends when injected into Xenopus laevis oocyte nuclei (Birchmeier et al, 1984). Injection into Xenopus oocyte nuclei of RNA synthesized in vitro from deletion mutants of the H2A gene has revealed that the inverted repeat and 200 nucleotides of

3' spacer sequence are required for the efficient processing of 3' termini. Injection of mutant H2A DNA from which the terminal inverted repeat is deleted does not produce any H2A RNA with authentic 3' termini due to lack of processing. Studies of a series of H2A spacer DNA deletion mutants which were processing-deficient have shown that the transcription has been terminated heterogeneusly within the first 130 to 230 nucleotides of H2A 3' spacer region. The results suggest that termination of transcription of sea urchin histone gene transcription in Xenopus oocyte nuclei occurs heterogeneously within the first 130 to 230 nucleotides of H2A spacer region, downstream of the terminal inverted DNA repeat of sea urchin histone genes. These transcripts are then rapidly processed into mature histone mRNAs.

Injection of the cloned sea urchin histone gene repeat h22 DNA into Xenopus oocyte nuclei results in faithful initiation of the histone mRNAs, but transcription termination or RNA processing to yield the H3 mRNA 3' ends is very inefficient, giving rise to heterogenous read-through transcripts. Some of these transcripts extend into the H2A gene and yield a bicistronic mRNA (Hentschel et al, 1980b). The simplest explanation of this finding is that the frog oocyte lacks a factor required for transcription termination of the H3 gene or processing of the 3' ends of transcripts. A 12 S protein component derived from a chromosomal salt wash fraction of sea urchin embryos was partially purified. Coinjection of this protein component with h22 DNA into Xenopus oocyte nuclei resulted in the formation of authentic 3' ends of H3 mRNA (Stunnenberg and Birnstiel, 1982). The generation of histone H3 mRNA 3' termini could also be enhanced by the injection of a small sea urchin RNA, U7 snRNA (about 60 nucleotides in length), into the oocyte one day prior to injection of the h22 sea urchin histone DNA. Furthermore, this 60 nucleotide RNA could be isolated

from the 12S protein component (Galli et al, 1983). Therefore, it appears that the formation of H3 mRNA 3' ends requires a small nuclear ribonucleoprotein (snRNP). Injection of 60-nucleotide small nuclear RNA into oocytes two days prior to injection of in vitro synthesized H3 RNA with 3' spacer extensions results in the formation of authentic 3' termini of H3 mRNA. The results suggest that the 3' ends of H3 mRNA are created by processing, not by termination (Birchmeier et al, 1984).

There is a sequence, CAAGAAAGA, in spacer DNA, beginning six nucleotides downstream of the 3' termini of histone mRNA, which is conserved in all the sea urchin histone genes (Hentschel and Birstiel, 1981). Analysis of the cDNA sequence of the 60 nucleotide U7 RNA has revealed that there are sequences near the 5' terminus of the U7 RNA which are complementary to the conserved sea urchin histone 3' inverted DNA repeat and the CAAGAAAGA motif (Strub et al, 1984). Studies of the processing of RNAs synthesized in vitro from linker scanning, deletion, and insertion mutants of the sea urchin H3 gene have reached several conclusions. In addition to the stem-loop structure of the mRNA, deletion of the CAAGAAAGA motif abolishes the synthesis of mature 3' termini of H3 mRNA. When the terminal ACCA sequence was replaced by a linker sequence, RNA processing was reduced but not completely abolished. Therefore, the ACCA sequence is required for maximal RNA processing but is not essential; the same is true of sequences within 50 nucleotides immediately downstream of the CAAGAAAGA motif. An insertion of 6 nucleotides between the 3' inverted repeat and CAAGAAAGA also abolished processing. Therefore, these two sequence blocks can be seen as a single signal for RNA processing, with a narrowly defined topology directing 3' maturation of histone mRNAs (Georgiev and Birnstiel, 1985).

3. Mouse β-major Globin Gene.

The mouse β-major globin gene has been isolated and sequenced (Konkel et al, 1978). The transcription unit that encodes β-major globin mRNA has been investigated in DMSO-induced murine erythroleukemia cells using mainly RNA labeled in vitro in isolated nuclei (Hofer and Darnell, 1981). Studies of RNA labeled in vitro with ^{32}P-UTP for 15 min in isolated HeLa cell nuclei have shown that there is no re-initiation of RNA chains, but elongation of already initiated chains proceeds for about 500 nucleotides or less (Weber et al, 1979); very little processing of previously labeled completed chains occurs (Blanchard et al, 1978). In vitro labeled nascent RNA was hybridized to cloned restriction fragments of the genomic β-globin DNA. Equimolar amounts of labeled nascent RNA hybridized to all the fragments encoding for the mRNA and to fragments extending at least 1400 nucleotides beyond the poly(A) addition site (Hofer and Darnell, 1981; Hofer et al, 1982). This result indicates that transcription of β-globin gene terminates within a region 1400 nucleotides downstream from the polyadenylation site. However, when in vivo pulse-labeled (20 min) nuclear RNA was used in the hybridization, much less RNA was hybridized to fragments containing sequences downstream of the poly(A) site. The result indicates that RNA synthesized from regions downstream of the poly(A) site is extremely unstable in vivo. Nevertheless, this finding reveals that there is substantial transcription past the poly(A) site in vivo (Hofer and Darnell, 1981).

In search of transcription termination sites, RNA labeled in isolated nuclei was hybridized to DNA fragments containing sequences downstream of the poly(A) site. The hybrids were treated with S1 nuclease and analyzed by polyacrylamide gel electrophoresis. There was no RNA species with a 3' end that would correspond to a single termination site; instead, RNA

molecules with many discrete 3' ends were detected. This result probably means that termination occurs at many sites throughout the region between 700 to 1500 nucleotides past the poly(A) site. Sequence analysis of this region reveals the presence of some inverted repeats and some homology among different globin genes (Citron et al, 1984).

To determine whether the termination region of the mouse β-globin gene functions to terminate transcription when moved to another genomic site, portions of the termination region have been inserted into the E1A transcription unit of adenovirus 5 (Falck-Pedersen et al, 1985). The site of insertion is at a unique XbaI site at 3.85 map units, within the 3' exon of the E1A region, and lies upstream of the E1A poly(A) site. The E1A promoter in this construction is also replaced by the promoter of adenovirus major late transcription unit (Logan and Shenk, 1984). This replacement leads to an increased accumulation of RNA from the E1A region late in virus infection with little effect early in infection. The mouse β-globin DNA fragments used to construct insertion mutants were fragment D (309 base pairs, containing the poly(A) site), E (443 base pairs, immediately downstream of fragment D), and F (809 base pairs, immediately downstream of fragment E). The putative termination region is located within fragment F. RNAs labeled either in isolated nuclei or in whole cells which were infected with recombinant virus were analyzed. In recombinant viruses which contain either the 5' portion (503 nucleotides) of the F region, inserted in either orientation, or the F+E+D fragment inserted opposite to its normal orientation, E1A transcripts were synthesized and polyadenylated at the normal E1A poly(A) site and contained the inserted globin sequences. Recombinant virus containing the D+E+F region of β-globin DNA in the normal orientation produces smaller E1A mRNAs. Analysis of the 3' ends of these mRNAs showed that the globin

poly(A) site was used exclusively. RNA labeled in nuclei isolated from cells infected with recombinant virus containing the mouse β-globin D+E+F region was hybridized to DNA fragments from the adenovirus E1A, E1B or globin D, E, or F regions. The results showed that transcription downstream of the globin F fragment is strongly reduced compared with transcription of upstream regions. In addition, transcription of the E1B gene is strongly reduced.

These results suggest that sequences between the mouse β-globin poly(A) site and the F region are required for transcription termination and that termination might only occur after an active polyadenylation site is utilized. During the late phase of infection, transcription of E1A DNA downstream from the globin F region is detected, and active transcription of the E1B gene is also resumed. The termination event in the late phase of infection is apparently not as efficient as during the early phase; this may account for the increased E1B transcription. It has previously been shown that the major late transcription unit is transcribed to the extreme righthand end of the adenovirus genome without an identifiable termination event during the late phase of infection (Fraser et al, 1979). This could be due to an antitermination event or to the deficiency of a termination factor during the late phase of adenovirus infection.

4. Other Higher Eukaryotic Genes and Viruses.

Studies of many viruses and eukaryotic mRNA coding genes have shown that RNA polymerase II transcribes across the polyadenylation site and terminates somewhere downstream of the poly(A) site. So far, the sequences at which transcription actually ceases are not known.

During the late phase of SV40 infection, RNA was labeled in vivo with ^{3}H-uridine for 1 min. Hybridization of the in vivo labeled RNA to SV40

L-strand DNA has shown that nearly equimolar amounts of labeled RNA hybridizes to the late mRNA coding region and to a region extending about 1000 nucleotides downstream of the poly(A) site (Ford and Hsu, 1978). This result indicates that transcription of L-strand DNA terminates beyond 1 kb past the poly(A) site.

In the case of polyomavirus, early in lytic infection E-strand RNA appears to be terminated and processed very rapidly and efficiently. By examining 15 min pulse labeled early nuclear RNA, Acheson and Mieville (1978) concluded that more that 80% of E-strand RNA terminated near the 3' end of early mRNA. But this result could be due to the rapid processing of 3' ends of early mRNA instead of transcription termination, because the labeling period was long compared to turnover and processing times. In contrast, transcription of the L DNA strand during the late phase of infection does not terminate efficiently, and can continue 3 or 4 times around the 5.3 kb circular DNA before termination occurs (Acheson, 1978). Studies of 1 or 2 min pulse-labeled nuclear late RNA (McNally, 1983; Tessier, D. and Acheson, N. H. personal communication) and late RNA labeled in isolated nuclei or viral transcription complexes (Skarnes, 1985) have revealed that only about half the amount of RNA polymerases was located at the DNA fragments downstream of the polyadenylation site when compared to the fragments upstream of the polyadenylation site. This suggests that the RNA polymerase may terminate transcription of the L DNA strand downstream of the poly(A) site with an efficiency of about 50%.

In adenovirus 2 transcription has been shown to proceed far beyond the polyadenylation sites of two early transcription units, E2 and E4 (Nevins et al, 1980). Late in adenovirus 2 infection, transcription of the major late transcription unit encompasses five poly(A) sites, and terminates within the terminal 650 nucleotides without reaching the end of the genome

(Nevins and Darnell, 1978; Fraser et al, 1979; Fraser and Hsu, 1980). A wide variety of mRNAs are processed from this primary transcript and fall into five groups (L1 to L5) which possess common polyadenylation sites. In contrast, during the early phase of infection, transcripts of the major late unit terminate near the middle of the genome, and only the mRNA of L1 family is produced (Shaw and Ziff, 1980; Nevins and Wilson, 1981). The reason for the difference in transcription pattern between early and late phase of infection is not clear. There are several possibilities: 1) after DNA replication, the viral chromatin structure may have changed, hence preventing termination in the middle of the genome; 2) some viral factors which inhibit termination (antiterminators) may be produced after DNA replication; or 3) during the late phase of infection the termination factors are depleted, reducing termination efficiency.

Using RNA labeled in vitro in isolated nuclei, LeMeur et al (1984) have analyzed the 3' ends of primary transcripts of the chicken ovalbumin gene. Hybridization of in vitro labeled nuclear RNA with DNA fragments covering the 3' region of the gene indicates that more than 90% of transcription terminates in a discrete region of 170 base pairs located 900 nucleotides downstream from the main polyadenylation site. Analysis of sequences within this 170 base pair region reveals two sequences analogous to a putative yeast termination signal, TTTTTATA (Zaret and Sherman, 1982). Therefore, the authors propose that this sequence may be involved in the process of transcription termination in the chicken ovalbumin gene.

The site of transcription termination in the mouse α-amylase gene Amy-2^a was determined by in vitro transcription in isolated nuclei of pancreatic cells (Hagenbuchle et al, 1984). Hybridization of in vitro elongated nascent transcripts to Amy-2^a restriction fragments indicates that over 90% of RNA polymerases terminate transcription in a region

between 2.5 and 4 kb downstream of the Amy-2^a polyadenylation site and that polymerases are gradually diluted out over the last 1.4 kb. S1 nuclease mapping of steady state nuclear RNA indicates that RNA polymerases terminate at multiple sites within this region. However, primary transcripts may be processed very rapidly into mature mRNAs; the 3' ends of primary transcripts may be very unstable. Therefore, 3' ends defined by S1 mapping of steady state nuclear RNA do not necessarily correspond to termination sites but may represent partially processed RNAs.

Rohrbaugh et al (1985) have localized the 3' termination region of the rabbit β1 globin gene in fetal liver by using in vitro labeled RNA in isolated nuclei. Their results indicate that transcription termination may occur 600 nucleotides downstream of the poly(A) site, within a 570-nucleotide region. The DNA segment immediately upstream of the region of declining transcription contains an inverted repeat and encodes a short RNA transcribed by RNA polymerase II from the opposite strand of the β1 globin gene. Therefore, the authors suggest that the inverted repeat and the opposite transcript may play a role in the termination of β1 globin gene transcription.

5. Conclusions.

In recent years, studies on transcription termination by RNA polymerase II in eukaryotic cells have accumulated, but very little is yet understood. In yeast consensus sequences such as TAG...TAGT...TTT and TTTTTATA, have been proposed to mediate formation of 3' ends (Zaret and Sherman, 1982; Henikoff and Cohen, 1984). Polyadenylation in yeast mRNA is apparently directly coupled with termination. However, it is still possible that these yeast sequences are involved in RNA processing rather than transcription termination.

Most histone mRNAs do not contain poly(A) tails. Studies of sea urchin histone genes injected into frog oocyte nuclei have revealed that the 3' ends of histone mRNAs are generated by RNA processing (Birchmeier et al, 1984). The conserved sequences, which include a stem-loop structure near the 3' end of the mRNA, and a 3' end ACCA motif, and a CAAGAAAGA nonamer 6 base pairs downstream of the 3' end, are required as a signal for processing of the 3' end. A sea urchin-specific small nuclear RNA (U7) has been shown to be required for the processing. However, the sites and role of transcription termination in these histone genes is not known.

Most mRNAs of higher eukaryotes contain a long poly(A) tail at their 3' ends. It is known that these 3' ends are created by endonucleolytic cleavage of the pre-mRNA, followed by the addition of polyadenylic acid to the newly exposed 3' end (Nevins and Darnell, 1978). The hexanucleotide AAUAAA as well as additional 3' flanking sequences, are required for the cleavage and polyadenylation reactions. Both in vivo and in vitro studies of transcription of mRNA genes have shown that RNA polymerases transcribe through the poly(A) site and terminate at mutiple sites from hundreds to thousands of nucleotides downstream of the poly(A) site. The nature of the termination signal is not yet understood. Many hypotheses have been proposed, some based on facts, some merely speculation. However, there is evidence indicating that transcription termination may be linked to polyadenylation or to recognition of a polyadenylation signal (Falck-Pedersen et al, 1985).

Lytic Cycle of Polyomavirus.

Polyomavirus is a member of the papovavirus family. It was originally isolated by accident when Gross (1953a) made extracts of the low-leukemia C3H mice and observed that a few mice inoculated with the extracts

developed salivary gland (parotid) adenocarcinomas while remaining free of any trace of leukemia. Gross (1953b) was able to separate the agent which causes salivary gland adenocarcinomas from the leukemia virus, hence he named this agent a parotid agent. Although parotid tumors are the most frequent consequence of inoculation by the parotid agent, a variety of other tumors including medullary adrenal tumors, epithelial thymic tumors, mammary gland carcinomas, renal carcinomas, liver hemangiomas, and subcutaneous fibrosarcomas also occasionally develop. Because the parotid agent was able to transform so many different cell types the name "polyoma virus" (poly=many oma=tumor) was proposed (Stewart et al, 1958) and now this is universally accepted.

The genome of polyomavirus consists of a circular, double-stranded DNA molecule (Dulbecco and Vogt, 1963; Weil and Vinograd, 1963). It has a molecular weight of 3.5×10^6 daltons and is enclosed in an icosahedral capsid 45 nm in diameter consisting of 72 capsomers and 420 structural subunits (Fenner et al, 1974). These capsomers are composed of a major capsid protein, VP1, and two minor capsid proteins, VP2 and VP3. The DNA within the virion is associated with four host histones and forms the viral chromatin (Freason and Crawford, 1972; Fey and Hirt, 1975). Polyomavirus DNA is comprised of 5292 base pairs (the A2 strain) or 5295 base pairs (the A3 strain) and has been completely sequenced (Friedmann, et al, 1979; Soeda et al, 1980).

The lytic infection of polyomavirus takes place in two phases, "early" and "late". The early phase lasts until the beginning of viral DNA replication, which occurs 12 hr after infection at 37°C. Three viral T antigens are synthesized during this period. The late phase begins with the onset of viral DNA replication and ends at cell death, which is about 48 hr after infection. The three viral structural proteins are synthesized

and progeny viral particles are assembled during the late phase. Viral DNA replication, RNA synthesis and virion assembly take place in the nucleus of the infected cell.

Upon infection, the virus particles absorb to the surface of the cell. The virions are then taken into the cytoplasm within pinocytic vesicles shortly after infection (Mattern et al, 1966). The viral capsid proteins are removed and viral chromatin is transported to the nucleus. The viral DNA in the nucleus can now undergo the subsequent steps in the viral lytic cycle.

Transcription of the polyomavirus genome is temporally regulated. During the early phase of the lytic infection, one of the two DNA strands (the E DNA strand) is transcribed predominantly. Transcription begins near the origin of viral DNA replication (around nucleotide 150) and proceeds counter-clockwise as shown in Fig. 1. The E strand RNA coding region spans approximately one half of the viral genome. The primary RNA transcript is spliced into three early mRNAs which encode for three early proteins known as large, middle and small T antigens (Soeda et al, 1979). The early mRNAs represent only 0.01-0.02% of the total cytoplasmic RNA (Weil et al, 1974; Acheson and Mieville, 1978). During the early phase, transcription of the viral E DNA strand is terminated efficiently, and E-strand RNA is rapidly transported from nucleus to cytoplasm. Viral mRNA appears in the cytoplasm within 6 min of its synthesis; more than half of the viral RNA synthesized in the nucleus is exported to the cytoplasm (Acheson, 1981). The primary transcript of the E DNA strand may extend beyond the polyadenylation site, which defines the 3' end of mature mRNA (Kamen et al, 1974; Acheson and Mieville, 1978).

Synthesis of polyomavirus T antigens during the early phase is followed by a mitotic host response. This comprises stimulation of overall

cellular RNA and protein synthesis, the activation of the cellular DNA-synthesizing apparatus and the duplication of the host cell chromatin (for review, see Acheson, 1981a). These changes in the host cell metabolism provide the materials for the efficient replication of viral DNA and production of viral progeny. Large T antigen has been shown to negatively regulate its own synthesis (Cogen, 1978; Kamen et al, 1980).

Replication of viral DNA can occur as early as 12 to 15 hours after infection, and marks the onset of the late phase of lytic infection. Large T antigen is required for the initiation of each round of viral DNA replication, but not for continued elongation. The requirement of large T antigen in viral DNA replication and the negative control on its own synthesis has been shown by experiments with temperature-sensitive mutants of polyomavirus affecting large T antigen (Franke and Eckhart, 1973). Viral DNA replication is semiconservative (Hirt, 1969) and bidirectional (Crawford et al, 1973; Griffin et al, 1974).

During the late phase of lytic infection, the opposite strand of DNA (the L DNA strand) is transcribed. Transcription is initiated in a region located at the other side of the DNA replication origin, with respect to the early initiation region, and proceeds in a clockwise direction (Fig. 1). The late coding region spans the right half of the viral genome and codes for the three viral capsid proteins, VP1, VP2 and VP3. The maximal rate of RNA synthesis occurs between 28 and 40 hours after infection, because of the increase of viral DNA templates which takes place. At this time, virus-specific RNA represents more than 10% of the RNA labeled in infected cells during a 5-30 min labeling period (Acheson et al, 1971). Late virus-specific mRNA accounts for 1-3% of total cytoplasmic RNA labeled during several hours (Buetti, 1974). Nearly all of these viral RNAs are transcribed from L DNA strand.

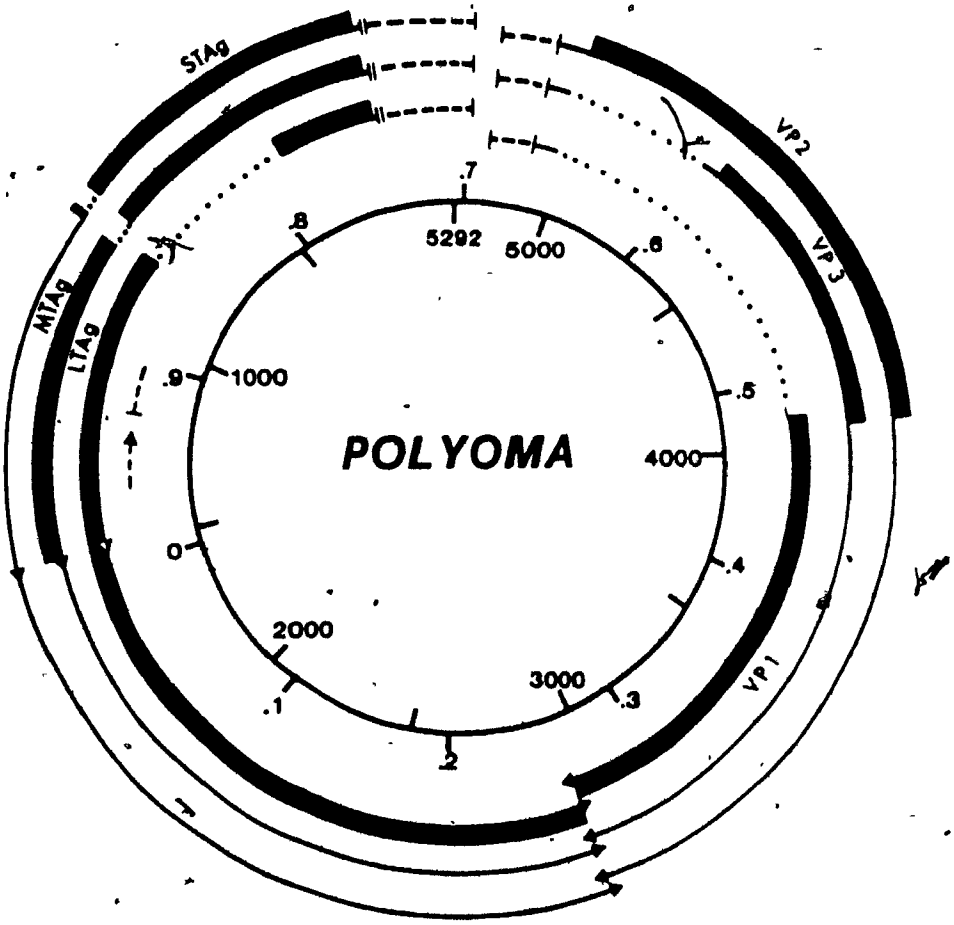

POLYOMA
STAg
MTAg
LTAg
VP2
VP3
VP1
5292
5000
4000
3000
2000
1000
0
.1
.2
.3
.4
.5
.6
.7
.8
.9

Figure 1: Physical map of polyomavirus genome.

The numbers on the inner side of the circle represent nucleotides numbered according to Soeda et al (1979, 1980). The numbers on the outer side of the circle represent map units, with the unique EcoRI site serving at the 0 reference point. The E-strand DNA is transcribed counter-clockwise and transcripts are processed into three early mRNAs which code for large tumor antigen (LTAg), middle tumor antigen (MTAg) and small tumor antigen (STAg). The L-strand DNA is transcribed clockwise and transcripts are processed into three late mRNAs which code for three viral capsid proteins, VP1, VP2 and VP3. Dashed lines represent regions where transcription initiates. The thickened portions on the outer lines represent protein coding regions on each mRNA. Dotted lines represent sequences removed from mRNA precursors by splicing. Arrowheads represent the 3' ends of mRNAs.

Termination and processing of the L-strand transcripts appear to be very inefficient. Only about 5% of total viral RNA synthesized becomes mRNA (Acheson, 1984). The bulk of late polyomavirus RNA is larger than the viral genome; this is caused by RNA polymerase II traversing the circular viral DNA genome several times before terminating transcription (Acheson, 1978). Acheson (1984) has confirmed that only 10 to 25% of nuclear viral RNA is polyadenylated and that it is polyadenylated within less than two min of its synthesis. Nuclear viral RNA which is polyadenylated exists as molecules of discrete lengths of about 2.2 + n(5.3) kilobases. These RNAs are a family of molecules which have the same 5' and 3' ends but differ in size by an integral number of viral genome lengths (5.3 kilobases) (Acheson, 1984). The mature mRNAs contain only a single mRNA body sequence (Buetti, 1974; Kamen et al, 1980), which is linked at its 5' end to one or more tandemly joined 57-nucleotide leader RNAs (Legon et al, 1979; Treisman, 1980). These multiple copies of the leader must arise by splicing between leader sequences present in the giant precursor RNAs. Thus, leader to leader splicing removes almost an entire genome-length of RNA from a precursor RNA molecule.

Studies with SV 40 have shown that as soon as viral capsid proteins are made, the empty viral capsid is assembled. The viral DNA-histone complexes subsequently interact with the empty capsids and form intact virions (Ozer, 1972; Ozer and Tegtmeyer, 1972). Although the exact mechanism of virus assembly is still not completely understood, there appears to be no specific mechanism for release of viral progeny from the infected cells, many virions remain cell-associated even after extensive cell death (Granboulan et al, 1963; Mattern et al, 1966). Assembly of polyomavirus virion is believed to be similar to that of SV40. The viral

particles begin to appear as early as 20-25 hours after infection, and can continue to appear up to 60-70 hours post-infection, the time at which cell death normally occurs. The lytic cycle is then complete.

Polyomavirus as a Model System to Study Transcription Termination.

At present, little is known about the sites and signals involved in the termination of transcription by RNA polymerase II. Polyomavirus lends itself particularly well to the study of transcription termination for the following reasons:

1. Its molecular biology has been extensively studied.
2. The nucleotide sequence of its small (5.3 kilobase) genome has been determined.
3. It is transcribed as a DNA-histone complex by RNA polymerase II in the nucleus of mouse cells; therefore it is analogous to cellular genes which code for mRNAs.
4. The virus has two distinct transcription units, one on each DNA strand.
5. During the early phase, transcripts of the E DNA strand are terminated and processed very efficiently; however, the amount of RNA synthesized is low.
6. During the late phase, transcripts of the L DNA strand are terminated and processed inefficiently, giving rise to giant transcripts which contain several tandem repeats of the entire DNA strand. Virus-specific RNA represents about 10% of total RNA synthesized in the cell and most of these viral RNAs are L-strand RNAs.

MATERIALS AND METHODS

Preparation of Primary Baby Mouse Kidney Cell Cultures (BMK).

Kidneys were removed from 10-day-old baby mice (Canadian Breeding Farms) and kept in phosphate-buffered saline lacking Mg^{++} and Ca^{++} (PBS). After removal of excess tissue, kidneys were chopped into very fine pieces with sterilized scissors and treated with 20 ml of 0.2% trypsin (Sigma) in PBS containing 0.03% penicillin (Sigma) and 0.01% streptomycin (Sigma) at room temperature for 30 minutes each time. The supernatant from the first trypsinization was discarded and the supernatants of the next 4 to 5 digestions were collected. Following centrifugation, the pellet was resuspended in Dulbecco's modified Eagle's medium (DMEM) (Flow Laboratories) supplemented with 5% heat treated newborn calf serum (NCS) or fetal bovine serum (FBS) (Flow Laboratories), 2.5 µg/ml fungizone (Squibb), 50 µg/ml gentamycin (Schering Canada Inc.) and 25 µg/ml vancomycin (Lilly). The suspension was plated onto 90 mm plastic tissue culture petri dishes (Nunc, Lux) at a volume of 10 ml. Kidneys from approximately 25 to 30 baby mice were prepared for 40 90-mm petri dishes. BMK cells were grown at 37°C in an atmosphere containing 5% CO_2. Within 3 to 4 days, the cells were nearly confluent (5×10^6 cells/petri dish) and were ready for infection. BMK cell cultures were mainly used for growing polyomavirus stocks and performing plaque assays.

Preparation of Mouse Fibroblast 3T6 Cell Cultures.

Swiss albino 3T6 cell line was purchased from the American Type Culture Collection (ATCC-CL96). These cells were maintained in DMEM with 10% FBS and antibiotics in 90 mm petri dishes. Upon reaching confluency ($2\text{-}3 \times 10^7$ cells per petri dish), the cells were trypsinized and replated

at a 1:10 dilution twice a week. After 3 days, the newly plated cells were nearly confluent (approximately 1 x 10^7 cells per petri dish), and were ready to be infected with virus. 3T6 cells were used for preparing polyomavirus RNAs.

Polyomavirus Stock.

The polyomavirus used in our laboratory is strain AT3 (Skarnes, Tessier and Acheson, in preparation). The DNA genome is comprised of 5309 base pairs. This strain resembles the polyomavirus A3 strain (Friedmann et al, 1979) in that it has a 11 bp deletion to the early side of the replication origin (nucleotides 44 to 54 of the A2 strain) (Griffin and Maddock, 1979) and an insertion of one nucleotide between nucleotides 2890 and 2891 of the A2 strain, which leads to a different C-terminal region in large-T antigen (Deininger et al, 1980; Soeda et al, 1980). However, strain AT3 has a 75-nucleotide direct repeat in tandem in the region of the transcriptional enhancer (nucleotides 5099-5174 and nucleotides 5175-5249), and a 52-nucleotide deletion (nucleotides 5207 to 5288 in the A2 strain). The 75-nucleotide repeat resembles that found in the MV strain (Ruley and Fried, 1983). Virus was plaque-purified and grown on primary BMK cells. Titers of virus stocks (4-10 x 10^8 pfu/ml) were determined by plaque assays on BMK cells. Virus stocks were stored at -20°C.

Labeling and Extraction of RNA.

3T6 cells were grown to subconfluency (about 1 x 10^7 cells) in 90 mm petri dishes and were infected with 0.4 ml of polyomavirus at a multiplicity of 20 to 40 pfu/cell. The plates were swirled every 15 minutes and kept in a 37°C, 5% CO_2 incubator for 2 hours. Following infection, 10 ml of DMEM containing gentamycin, fungizone, vancomycin (but

no FBS) was added to each plate. At 28 hours after infection, cells were labeled at 37°C with 500 µCi of [5-^{3}H] uridine (NEN, ICN, or Amersham; 25 to 30 Ci/mmol) in 1 ml of DMEM per 90 mm petri dish. Labeling was terminated by removing the radioactive medium and flooding monolayers with 10 ml of ice-cold TD buffer (25 mM Tris-HCl pH 7.5, 140 mM NaCl, 5 mM KCl, 0.3 mM Na_2HPO_4). Cells were then washed with 10 ml of ice-cold lytic buffer (10 mM triethanolamine-HCl pH 8.5, 10 mM NaCl, 1.5 mM $MgCl_2$) (Acheson, 1981). Monolayers were scraped in 1 ml of cold lytic buffer and cells were lysed by adding Nonidet P-40 (Shell) to a final concentration of 0.5%, and were gently disrupted with a Pasteur pipette.

Nuclei were sedimented at 3,000 rpm (Sorvall RC-5B centrifuge, HB-4 rotor) for 4 minutes at 2°C. The supernatant (cytoplasmic fraction) was transfered to another tube, and the nuclear pellet was resuspended in lytic buffer (1 ml per 90 mm petri dish of nuclei). One volume of extraction buffer (10 mM triethanolamine-HCl pH 8.5, 100 mM NaCl, 10 mM EDTA) containing 0.5 mM of aurintricarboxylic acid (ATA) (Sigma) was added to both cytoplasmic and nuclear portions. ATA was included to protect RNA from degradation by nucleases released during extraction procedures (Hallick et al, 1977). Sodium tri-isopropylnaphthalene sulfonate (TNS; Serva) (25%) was added to a final concentration of 2%. RNA was extracted three times in succession with an equal volume of phenol-chloroform-isoamyl alcohol (50:50:1) and once with chloroform-isoamyl alcohol (50:1). Nucleic acids were precipitated three times in succession from 0.2 M NaCl by adding 2.5 volumes of absolute ethanol, and then were treated with 10 µg/ml of RNase-free DNase I (Worthington Biochemical Corp., or BRL) at 0°C for 30 minutes in DNase buffer (10 mM sodium acetate pH 5.1, 100 mM NaCl, 2 mM $MnCl_2$). DNase was inactivated by adding EDTA (3 mM) and sodium dodecyl sulfate (SDS; BDH) (0.5%). RNA was then extracted with an equal volume of

phenol-chloroform-isoamyl alcohol (50:50:1) and ethanol precipitated three times (Acheson and Mieville, 1978). The RNA was stored at -20°C as the ethanol-precipitated form until further use.

TNS was purified after purchase. First, a 20% solution of TNS was prepared, then an equal volume of ethanol was added. The solution was kept at -20°C for one hour. Centrifugation was done at 10,000 rpm in Sorvall centrifuge to pellet the precipitate. The supernatant was then collected and lyophilized. A 25% TNS stock solution was prepared.

Glucosamine Treatment.

Cells were treated with glucosamine (Sigma) for 1 hour to reduce UTP pools and increase labeling efficiency when labeling times were to be 2 min or less (Scholtissek, 1971; Wertz, 1975). DMEM containing antibiotics with 20 mM glucosamine was neutralized with NaOH, then prewarmed and equilibrated in 37°C, 5% CO_2 incubator. One hour prior to labeling, infected 3T6 cell monolayers were treated with 20 mM glucosamine-DMEM medium. After 60 min incubation at 37°C, the medium was removed. Cells were washed 3 times with 5 ml DMEM which was prewarmed at 37°C. Cells were then labeled with [5-^{3}H] uridine as previously described.

Isolation of Recombinant Single-stranded DNA.

Restriction endonuclease fragments of polyomavirus DNA were cloned into single-stranded DNA bacteriophage vectors fd103 (Hermann et al, 1980) or M13 mp8 (Messing, 1983) by Nicholas Acheson, Michael McLeod, Maureen McNally (1984), Thu-Hang Tran, Daniel Tessier and myself in our laboratory. Single-stranded recombinant DNAs were prepared from polyethylene glycol (Sigma) precipitated phage preparations (Yamamoto et al, 1970) by hot phenol extraction. Circular single-stranded DNAs were checked for

integrity by agarose gel electrophoresis. Samples of double-stranded replicative form DNA from each preparation were tested for the presence of the correct fragment and absence of deletions by restriction endonuclease digestion.

DNA-RNA Hybridization and Nuclease Treatment.

^{3}H-labeled RNA was hybridized in solution with an excess of single-stranded recombinant bacteriophage DNA. Two methods were used.

A. Up to 20 µg of nuclear RNA were mixed with 4 µg of single-stranded DNA in 40 µl of 80% formamide, 0.4 M NaCl, 40 mM PIPES pH 6.4, 1 mM EDTA (Berk and Sharp, 1977; Favaloro et al, 1980). The mixture was incubated at 85°C or 65°C for 10 min, then transferred to 55°C for 3 hours for hybridization and cooled rapidly by adding 360 µl of ice-cold S1 buffer (30 mM sodium acetate pH 4.5, 0.3 M NaCl, 1 mM $ZnSO_4$, 5% glycerol). The hybrids were treated with S1 nuclease (Boehringer-Mannheim) at 45°C for 30 min (concentration of S1 nuclease varied in different experiments). One unit of S1 nuclease releases 1 µg acid-soluble deoxyribonucleotides from denatured DNA in 1 min at 37°C and pH 4.5 (Vogt, 1980). After S1 nuclease treatment, hybrids were extracted once with an equal volume of phenol-chloroform-isoamyl alcohol (50:50:1), and once with chloroform-isoamyl alcohol (50:1), and then ethanol precipitated. The pellet was resuspended in 200 µl of RNase buffer (10 mM Tris-HCl pH 7.5, 1 mM EDTA, 0.3 M NaCl) containing 0.1 to 0.5 µg/ml of RNase A (Sigma) and incubated at 30°C for 30 min. After ribonuclease digestion, proteinase K (Boehringer-Mannheim) was added at a concentration of 250 µg/ml and the sample was incubated at 37°C for 30 min. The hybrids were extracted with phenol-chloroform-isoamyl alcohol (50:50:1) and chloroform-isoamyl alcohol (50:1), and then ethanol precipitated in the presence of 10 µg of carrier tRNA.

The amount of RNA and DNA used in each experiment varied. In general, five times as much nuclear RNA as single-stranded DNA was mixed for hybridization. Transfer RNA was added to the hybridization mixture to bring the total amount of nucleic acids to 24 µg.

B. Up to 50 µg of total nuclear RNA was mixed with 10 µg of single-stranded DNA in 100 µl of 50 mM HEPES pH 7.0, 1 mM EDTA, 0.75 M NaCl. The mixture was incubated at 85°C for 10 min, then at 55°C for 10 to 30 min and cooled rapidly in an ice-water bath. Hybridization was complete within this period at this DNA concentration. Hybrids were diluted 2.5-fold to a final concentration of 0.3 M NaCl and digested with either ribonuclease T1 (200 to 800 units/ml) or ribonuclease A (0.1 to 0.5 µg/ml) at 30°C for 30 min. The NaCl concentration was adjusted to 0.75 M and 300 µl samples containing no more than 10 µg DNA were passed through 13 mm diameter nitrocellulose filters (BA-85, Schleicher and Schuell) clamped in a lucite manifold with one sheet of Whatman 3MM filter paper and two sheets of blotting paper to absorb the filtrate. Flow rates were about 1 ml per 30 seconds. Nitrocellulose filters were washed twice with 1 ml 50 mM HEPES pH 7.0, 1 mM EDTA, 0.75 M NaCl in the manifold. Nitrocellulose filters were then transferred to flat-bottomed polypropylene tubes and incubated with 333 units/ml of S1 nuclease in 0.4 ml of S1 buffer (30 mM sodium acetate pH 4.5, 300 mM NaCl, 1 mM $ZnSO_4$, 5% glycerol) at 37°C for 1 hour. The supernatant solution was subjected to phenol-chloroform-isoamyl alcohol (50:50:1) extraction and chloroform-isoamyl alcohol (50:1) extraction, and the hybrids were ethanol precipitated in the presence of carrier tRNA. Slight variations in these procedures are noted in Figure Legends.

Gel Electrophoresis and Fluorography.

DNA-RNA hybrids were subjected to electrophoresis on non-denaturing agarose, polyacrylamide or mixed agarose-polyacrylamide vertical slab gels (3 mm thick) in either TBE buffer (89 mM Tris-HCl pH 8.3, 89 mM sodium borate, 1 mM EDTA) or TAE buffer (20 mM Tris-HCl pH 7.8, 2.5 mM sodium acetate, 1 mM EDTA).

To prepare mixed 0.5% agarose-2.5% polyacrylamide gels, 50 ml of 1% agarose in water was equilibrated in a 45°C water bath and 50 ml of 5% polyacrylamide (acrylamide:bis-acrylamide 20:1) in 2X TBE buffer was equilibrated in a 37°C water bath. Both solutions were mixed together with 0.5 ml of 10% ammonium persulfate and 0.5 ml of 10% TEMED, and the gel was poured immediately.

Gels were prepared for fluorography by soaking in Enhance (New England Nuclear) according to the manufacturer's instructions. Briefly, gels were agitated gently in three gel volumes of Enhance for 1 hour at room temperature, then agitated in excess water for 90 min (during this period of time water was changed twice). Gels were transferred to a piece of Whatman 3MM filter paper, and dried with a Bio-Rad gel drier at 60°C under vacuum for 2.5 hours. The gels were then exposed to Kodak XAR-5 film at -70°C for periods of days to weeks.

Gel Slices.

In some experiments, certain lanes were removed from the untreated gel and were cut into 1 mm gel slices using a Bio-Rad gel slicer. Each slice was collected into a capped tube. Four hundred ul of Protosol (New England Nuclear) was added and slices were incubated in a 37°C water bath overnight. Four ml of Atomlight (New England Nuclear) or Universol (ICN Radiochemicals) scinitillation fluid were added and samples were counted in a LKB scinitillation spectrometer (background in tritium channel is about 10 cpm).

Cloning of Polyomavirus DNA into Plasmid pSP64.

Polyomavirus DNA was digested with XbaI, which cleaves twice at nucleotides 2477 and 2522. The 5264-nucleotide fragment was cloned into the XbaI site of the polylinker on the plasmid pSP64 (Melton et al, 1984), which was purchased from Promega Biotec. In order to reduce the chance of religation of the vector and decrease the background of antibiotic resistant colonies, the XbaI-digested pSP64 was treated with bacterial alkaline phosphotase (BAP) (BRL). The plasmid and insert DNAs were then mixed at a 1:2 ratio in 50 µl of ligation buffer (50 mM Tris-HCl pH 7.6, 10 mM dithiothreitol [DTT], 10 mM $MgCl_2$, 1 mM ATP, 10 µg/ml bovine serum albumin). Samples were incubated with T4 DNA ligase (10 units/ml) (New England Nuclear) in a 25°C water bath for one hour. Two volumes of ligation buffer were added and the mixture was transferred to a 15°C waterbath overnight. The ligation mixture was transformed into E. coli DH-1 competent cells, and ampicillin-resistant colonies were grown up. These transformants were screened for the polyomavirus DNA insert oriented such that the L-strand will be transcribed by SP6 RNA polymerase beginning at the SP6 promoter. The transformant containing the recombinant plasmid (pSP64-PyXbaI-L) was grown in chi broth (50 mM Tris-HCl pH 7.6, 25 gm/l yeast extract, 0.8 gm/l glucose, 0.1 gm/l diaminopimetic acid, 0.05 gm/l thymidine, 20 mM $MgCl_2$) in the presence of ampicillin. Recombinant plasmid DNA was purified by ethidium bromide-cesium chloride gradient centrifugation.

In Vitro Synthesis of Polyomavirus L-Strand RNA with SP6 RNA Polymerase.

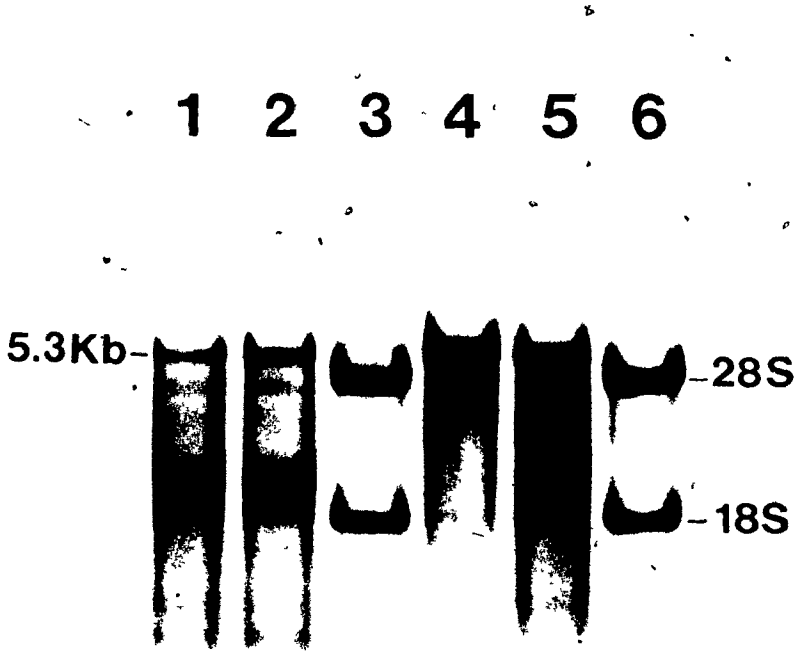
Recombinant plasmid pSP64-PyXbaI-L was linearized downstream of the 5.2 kb polyomavirus DNA insert via the SmaI restriction site of the polylinker. In vitro transcription was carried out by the procedure of

Melton et al (1984) with minor modifications. A 100 µl reaction mixture containing 40 mM Tris-HCl pH 7.5, 6 mM $MgCl_2$, 2 mM spermidine, 10 mM dithiothreitol, 100 units RNasin (Promega Biotec), 0.5 mM each of ATP, GTP, CTP and UTP, 100 µCi ^{3}H-CTP or ^{3}H-UTP (22 Ci/mmol, ICN Radiochemicals), 22.5 units of SP6 RNA polymerase (Promega Biotec) and 3 µg of SmaI-digested pSP64-PyXbaI-L DNA was incubated at 40°C for 30 min. An equal volume of transcription buffer containing 0.5 mM each of ATP, GTP, CTP and UTP was then added, and incubation continued for another 30 min at 40°C. The DNA template was degraded by adding 20 µg/ml of ribonuclease free DNase I (DPRF, Worthington Biochemicals Corp.) at 0°C for 30 min. To remove DNase activity, SDS (final concentration 0.5%) and EDTA (final concentration 3 mM) were added and RNA was extracted with equal volume of phenol-chloroform-isoamyl alcohol (50:50:1) and ethanol precipitated.

Purification of Full-length In Vitro Synthesized Polyomavirus L-strand RNA.

In vitro transcription of SmaI-digested pSP64-PyXbaI-L DNA should give a 5305 nucleotide run-off transcript. But many SP6 RNA polymerases terminated prematurely at several specific sites on the DNA template, when ^{3}H-UTP was present at 41 µM in the transcription mixture (Fig. 2, lanes 1, 2). Even in the reaction mixture containing 41 µM of ^{3}H-UTP plus 0.5 mM of unlabeled UTP, premature termination of transcription was observed (Fig. 2, lanes 4, 5). To obtain 5.3 kilobase full-length in vitro transcripts, RNA was suspended in 50 mM LiCl, 0.1% SDS, 10 mM Tris-HCl pH 7.4. The solution was heat-denatured at 95°C for 2 min and cooled rapidly in an ice-water bath. RNA was size-fractionated in a 15-30% sucrose gradient. After centrifugation at 55,000 rpm at 20°C for 150 min in a SW 60 Ti rotor, 9 fractions were collected from the bottom of the tube. Five µg of carrier tRNA was added to each fraction which was then ethanol precipitated.

A


1 2 3 4 5 6
5.3Kb
28S
18S

B

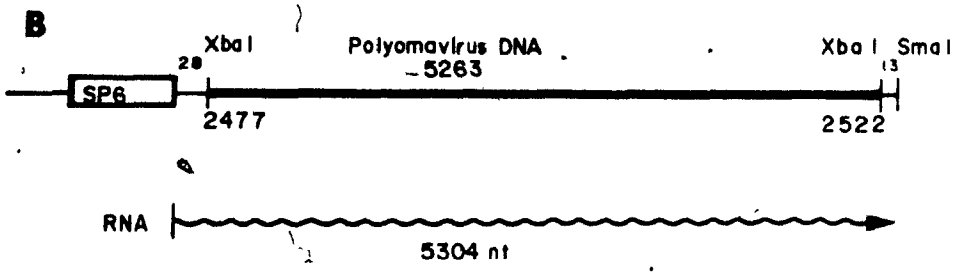

Xba I
Polyomavirus DNA
-5263
Xba I
Sma I
28
13
SP6
2477
2522
RNA
5304 nt

Figure 2: In vitro transcription of polyomavirus L-strand DNA (pSP64-PyXbaI-L) by SP6 RNA polymerase.

A. Each 10 µl of transcription mixture contained 1 µg of SmaI-digested pSP64-PyXbaI-L, 10 mM of DTT, 10 units of RNasin RNase inhibitor, 10 µCi of ^{3}H-UTP (~41 µM), 5 nanomoles each of ATP, GTP and CTP (0.5 mM). Two of the mixtures contained an additional 5 nanomoles of UTP (lanes 4 and 5). Either 1.5 units (lanes 1 and 4) or 6 units (lanes 2 and 5) of SP6 RNA polymerase were used in the reaction. Mixtures were incubated at 40°C for 1 hour. After DNase I digestion, RNA was phenol extracted and ethanol precipitated. Aliquots of each sample were taken and checked for radioactivity. Similar amounts of radioactivity (74,000-76,000 cpm) from each sample were glyoxal-denatured and analyzed in a 1% agarose slab gel as described in Materials and Methods. Markers were ^{3}H-uridine labeled cytoplasmic poly(A) RNAs (lanes 3 and 6). The size of full-length polyoamvirus run-off transcripts is about 5.3 kb.

B. Diagram showing SmaI digested pSP64-PyXbaI-L DNA. Open box is the SP6 promoter. Solid line corresponds to polyomavirus DNA cloned at the XbaI site of pSP64. Numbers above the line show the length of transcribed pSP64 and polyomavirus DNA. Numbers below the line show positions of the ends the of polyomavirus genome. The wavy line represents the run-off transcript. Number below the wavy line shows the size of the expected run-off transcript.

Aliquots of each fraction were glyoxal-denatured and analyzed by agarose gel electrophoresis and fluorography (Carmichael and McMaster, 1980) (Fig. 3). Fractions which contained full-length transcripts were pooled.

Glyoxal Denaturation.

A sample of RNA was suspended in 24 µl of 10 mM phosphate buffer (pH 6.9) containing 50% dimethyl sulfoxide (DMSO) and 1 M glyoxal. (To prepare 100 mM phosphate buffer pH 6.9, 40 ml of 0.1 M NaH_2PO_4 is mixed with 60 ml of 0.1 M Na_2HPO_4.) The glyoxal was purified by passing through a mixed bed resin [Bio-Rad AG501-X8] until the pH was neutral, and was stored at -20°C. The mixture was incubated at 50°C for 60 min (Carmichael and McMaster, 1980). The glyoxal-denatured RNA was then analyzed by running on a 1.1% agarose gel made in 10 mM phosphate buffer (pH 6.9). Gels were prepared for fluorography as previously described.

Oligo(dT)-cellulose Chromatography.

Cytoplasmic RNA (approximately 300 µg per sample) was suspended in 1 ml of 10 mM Tris-HCl, pH 7.4, 0.5 M NaCl, 0.1% SDS and was passed through a column containing 0.5 gm of oligodeoxy-thymidylic acid [oligo(dT)]-cellulose equilibrated with the same buffer. The unbound RNA was reloaded two more times and was collected. The column was washed with 15 ml of load buffer. The first 3 ml of this wash was collected and pooled with the flow-through as nonpolyadenylated RNA. Polyadenylated RNA was eluted with 5 ml of 10 mM Tris-HCl pH 7.4, 0.1% SDS (Aviv and Leder, 1972). Both nonpolyadenylated and polyadenylated RNA were ethanol precipitated.

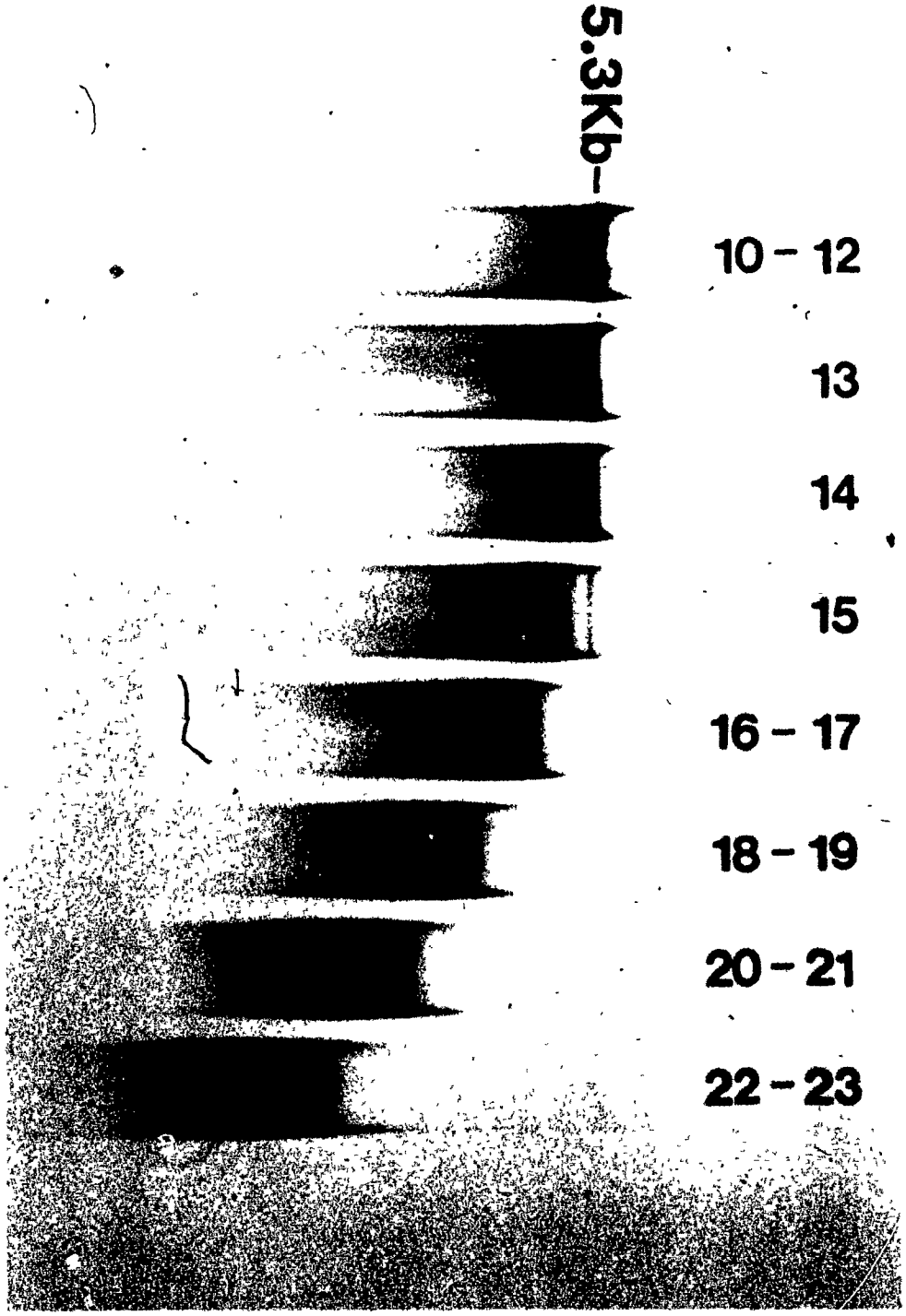

5.3Kb
10 – 12
13
14
15
16 – 17
18 – 19
20 – 21
22 – 23

Figure 3: Analysis of *in vitro* synthesized ^{3}H-polyomavirus L-strand RNA after sedimentation on a 15-30% linear sucrose gradient.

In vitro synthesized ^{3}H-polyomavirus RNA was suspended in 200 µl of 10 mM Tris-HCl pH 7.4, 50 mM LiCl, 0.1% sodium dodecyl sulfate, and was heat-denatured at 95°C for 2 minutes. RNA was layered on top of a 15-30% (wt/wt) linear sucrose gradient in the same buffer. Gradient was spun at 55,000 rpm in a Spinco SW-60 rotor at 20°C for 165 minutes. Fractions were collected from the tube bottom (8 drops per fraction), and aliquots were assayed for radioactivity. Ten µg of tRNA was added to each fraction, which was then ethanol precipitated twice in succession. Fractions in regions of interest were pooled. RNA was then resuspended in 10 mM Tris-HCl pH 7.4, 1 mM EDTA. Aliquots were taken and assayed for radioactivity. Similar amounts of radioactivity (10,000 cpm) wewe taken from each pool, glyoxal-denatured and analyzed on a 1.1% agarose slab gel as described in Materials and Methods. Above the fluorograph are the fraction numbers. The size of full-length run-off polyomavirus RNA is about 5.3 kb.

SECTION I

Development of a Technique for Mapping 5' and 3' Ends of *In Vivo* Pulse-labeled RNAs

INTRODUCTION

The aim of the work described in this section was to develop a method that would enable us 1) to define transcription initiation sites used at a given time, independently of stability or RNA processing; 2) to study the sequence and kinetics of RNA processing events (polyadenylation, splicing); and 3) to map the of 3' ends of RNAs created either by pausing of RNA polymerase on DNA template or by termination and release of RNA chains.

S1 nuclease is a single-strand-specific endonuclease isolated from *Aspergillus oryzae* (Vogt, 1973). It degrades both single-stranded DNA and RNA to yield all four nucleotide-5'-monophosphates. It does not degrade double-stranded DNA; nevertheless, supercoiled simian virus 40 (Beard *et al*, 1973) and polyomavirus (Germond *et al*, 1974) DNA are converted to unit length linear molecules by this enzyme. This could be due to weakly hydrogen-bonded or unpaired regions which are susceptible to S1 nuclease. Germond *et al* (1974) also demonstrated that S1 nuclease is capable of cutting DNA molecules containing single-stranded regions opposite to gaps. With these properties, S1 nuclease is used to detect the location of small deletions, insertions or any difference in base sequence between otherwise homologous DNAs (Shenk *et al*, 1975).

Berk and Sharp (1977) developed a method for mapping transcripts on viral genomes or on cloned DNA fragments. Unlabeled RNA is hybridized to ^{32}P-labeled viral DNA in 80% formamide at a temperature above the melting

temperature (Tm) of the DNA duplex, but below the Tm of the DNA-RNA hybrid. The DNA-RNA hybrids are treated with S1 nuclease under conditions which allow hydrolysis of single-stranded nucleic acid but do not produce strand breaks in hybrids. The size of protected DNA which is complementary to the RNA is then determined by electrophoresis. By performing this technique with a series of overlapping restriction fragments of DNA, 5' and 3' ends of transcripts can be accurately mapped.

Since both DNA strands of polyomavirus are transcribed, I decided to use single-stranded DNA clones to hybridize with ^{3}H-labeled RNAs. Restriction endonuclease fragments of polyomavirus DNA (Fig. 4) were cloned into the double-stranded replicative form (RF) DNA of single-stranded DNA bacteriophage fd103 (Hermann et al, 1980) or M13 mp8 (Messing). These phages replicate their DNA intracellularly as double-stranded DNA, but package only one DNA strand into the phage particle. By using these cloned single-stranded phage DNAs in hybridization, polyomavirus early RNAs will hybridize to the E DNA strand clones and late RNAs will hybridize to the L DNA strand clones only.

A major problem in pulse-labeling RNA in animal cells is the difficulty of getting a high specific activity of radioactive label in RNA. This is because large intracellular pools of RNA precursors cannot be diluted out or emptied. Scholtissek (1971) introduced a method for performing RNA pulse-chase experiments in primary chicken embryo fibroblasts. He demonstrated that, when labeled uridine was used as an RNA precursor, glucosamine could be used to rapidly deplete the UTP pool by forming UDP-N-acetyl-glucosamine, thus allowing the production of high specific activity pulse-labeled RNA.

I originally tried to apply Berk and Sharp's S1 nuclease mapping method directly to pulse-labeled nuclear RNA, by hybridizing labeled RNA to

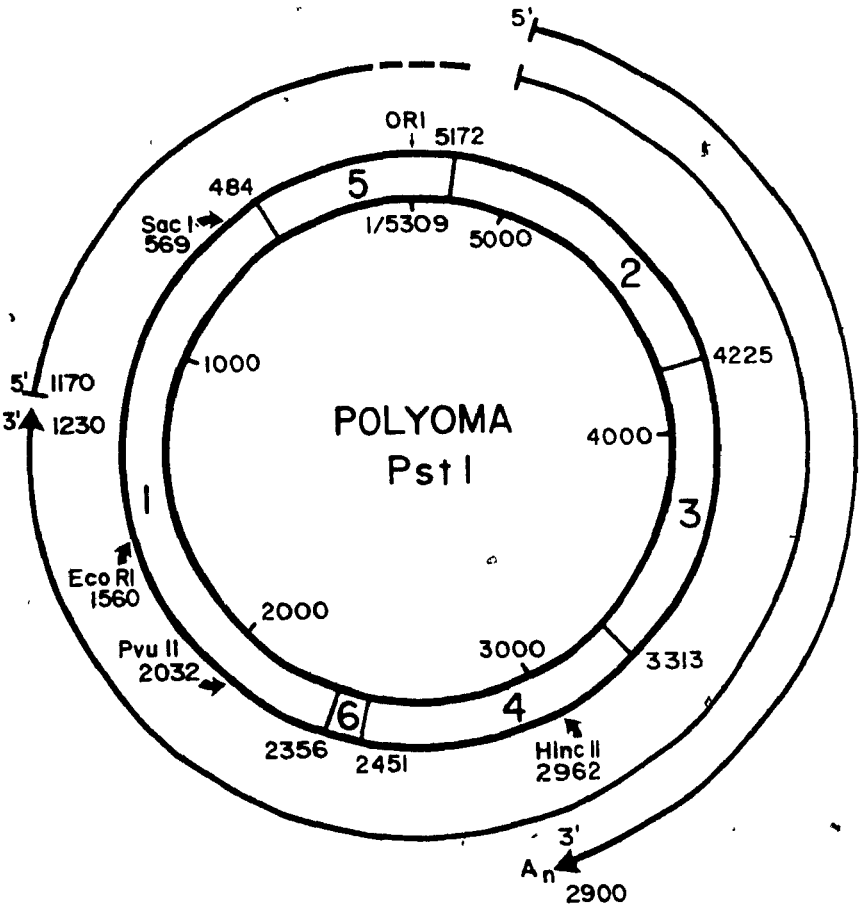

5'
ORI
5172
484
5
Sac I
569
1/5309
5000
2
4225
1000
5' 1170
3' 1230
POLYOMA
Pst I
4000
3
1
Eco RI
1560
2000
Pvu II
2032
3000
3313
6
4
2356
2451
Hinc II
2962
3'
An
2900

Figure 4: PstI restriction endonuclease map of polyomavirus DNA.

There are six PstI sites in polyomavirus DNA. Numbers inside the circle represent the polyomavirus DNA nucleotides (1-5309). Numbers outside the circle represent the PstI endonuclease cleavage sites. Replication origin is indicated as ORI. The cleavage sites of EcoRI (1560), HincII (2962), PvuII (2032), and SacI (569) are also indicated. All the PstI fragments (except fragment 6), HincII (2962) to PvuII (2032) fragment, EcoRI (1650) to PstI (484) fragment, PstI (2356) to SacI (569) fragment, and other restriction fragments not indicated in this map (such as HpaII fragments) were cloned into single-stranded DNA bacteriophage M13 or fd103. The outer circles represent L-strand transcripts; 5' and 3' ends are indicated. A 3' end in nuclear RNA at nucleotide 1230 and a 5' end at nucleotide 1170, as well as the L-strand polyadenylation site at nucleotide 2900, are also indicated.

unlabeled single-stranded DNAs. However, partially degraded labeled RNAs gave rise to unacceptably high backgrounds which obscured RNA-DNA hybrid bands in the gel. Therefore I had to modify Berk and Sharp's method. This section describes experiments which led to optimization of conditions for S1 nuclease mapping of pulse-labeled RNA.

RESULTS

Denaturation and Hybridization of Labeled RNA in 80% Formamide Solutions.

It has been shown that DNA-RNA and DNA-DNA association reactions proceed at much lower temperatures in aqueous formamide solutions than in aqueous solvents (Bonner et al, 1967; McConaughey et al, 1969). And DNA-RNA hybrids exhibit a significantly higher melting temperature than DNA-DNA duplexes in high concentrations of formamide (Casey and Davidson, 1977). Therefore, in solutions containing high formamide concentrations, DNA-RNA hybridization can be carried out under conditions which avoid reannealing of DNA-DNA duplexes.

Polyomavirus-infected 3T6 cells were labeled with ^{3}H-uridine for 150 min at 27 hr after infection (see Materials and Methods). The purified ^{3}H-labeled nuclear RNA, either alone or mixed with single-stranded fd103-polyomavirus PstI-1L clone (polyomavirus fragment PstI-1L cloned into single-stranded bacteriophage fd103, see Fig. 4), were suspended in 40 µl of hybridization buffer containing 80% formamide, 0.4 M NaCl, 40 mM PIPES pH 6.4 and 1 mM EDTA (formamide hybridization buffer). These samples were incubated at either 85°C or 65°C for 10 min to denature RNA-RNA hybrids (see Materials and Methods). Samples containing only RNA were cooled rapidly by adding 10 volumes of S1 nuclease buffer (30 mM sodium acetate pH 4.5, 0.3 M NaCl, 1 mM $ZnSO_4$, 5% glycerol). The remaining samples were transferred to 55°C (this temperature is above the Tm of DNA-DNA duplexes but below the Tm of DNA-RNA hybrids) for 3 hr, and then were cooled by adding 10 volumes of S1 buffer. All samples were treated with S1 nuclease (133 units/ml) and RNase A (2 µg/ml) as described in Materials and Methods. The results are shown in Figure 5. Incubation of the RNA alone at 85°C without subsequent incubation at 55°C led to degradation of almost all the

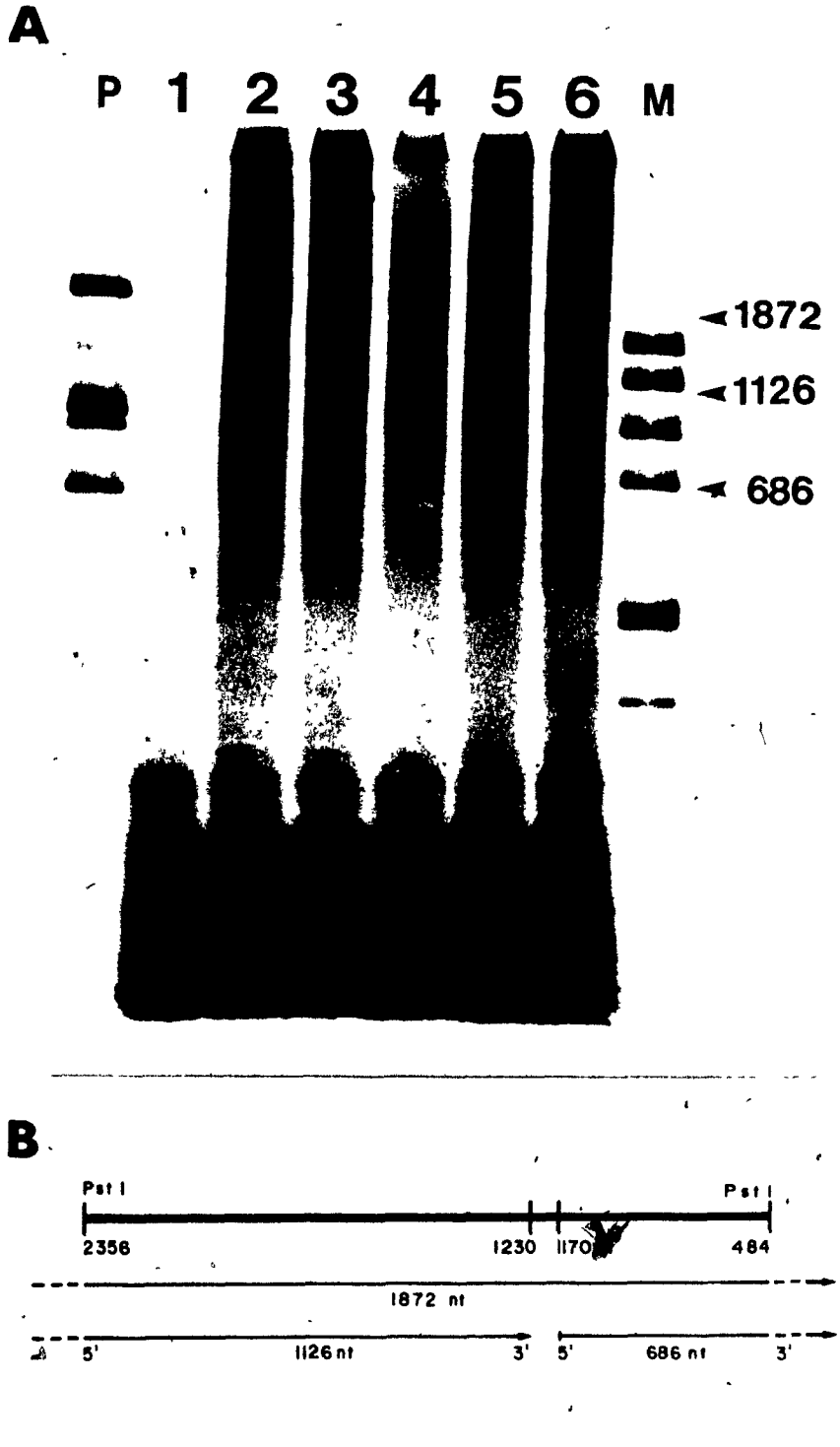

A
P 1 2 3 4 5 6 M
1872
1126
686
B
Pst I
Pst I
2356
1230
1170
484
1872 nt
5'
1126 nt
3'
5'
686 nt
3'

Figure 5: Effects of denaturation and reannealing in 80% formamide hybridization buffer on background radioactivity.

A. Twenty µg of nuclear RNA labeled for 150 minutes was suspended in 40 µl of hybridization buffer containing 80% formamide. Four µg of single-stranded fd103 bacteriophage DNA containing cloned polyomavirus PstI-1L was added to two of the samples (lanes 3 and 6). Half of the samples were heated at 85°C for 10 minutes; one of these was cooled rapidly by adding 360 µl of ice-cold S1 buffer (lane 1), the other two were transferred to 55°C for 3 hours before adding 360 µl of ice-cold S1 buffer (lanes 2 and 3). The other half of the samples were heated at 65°C for 10 minutes; one of these was cooled rapidly by adding ice-cold S1 buffer (lane 4), while the other two samples were transferred to 55°C for 3 hours before adding ice-cold S1 buffer (lanes 5 and 6). All samples were treated with 133 units/ml of S1 nuclease and 2 µg/ml of RNase A as described in Materials and Methods. Markers were double-stranded ^{3}H-polyomavirus DNA digested with PstI (lane P) or with MspI (lane M). Sizes and positions of expected DNA-RNA hybrids are indicated.

B. Diagram showing polyomavirus DNA fragment PstI-1L. Numbers below the line show positions of the ends of the fragment (nts 484 and 2356) and positions of 5' end (nt 1170) and 3' end (nt 1230) of steady-state RNA found by Treisman and Kamen (1981). Thin lines below the solid line show sizes of expected DNA-RNA hybrids.

high molecular weight RNAs by S1 nuclease and RNase A (lane 1), implying that denaturation is almost complete at 85°C. In contrast, RNA incubated at 65°C (lane 4) contained many high molecular weight RNAs which were resistant to S1 nuclease and RNase A digestion, and gave rise to a dark background in the fluorogram. Samples incubated at 85°C (lanes 2 & 3) or 65°C (lanes 5 & 6) for 10 min, then incubated at 55°C for 3 hr, showed the same background radioactivity upon electrophoresis and fluorography as samples incubated at 65°C for 10 min alone. These findings show that RNase and S1 nuclease resistant RNAs can be generated from nuclear RNA by incubating in annealing conditions, but are denatured by incubating at 85°C. Only the samples containing single-stranded DNA clone showed 3 bands of DNA-RNA hybrids (lanes 3 & 6). These 3 hybrid bands represent: 1) RNAs hybridized to the full-length polyomavirus fragment PstI-1L (1872 nucleotides); 2) RNAs which have 3' ends at nucleotide 1230, which generate a 1126 base pair hybrid; 3) RNAs which have 5' end at nucleotide 1176, which generate a 686 base pair hybrid (see Fig. 4).

Determination of the Cause of the Radioactive Background.

The radioactive background was present in all hybridization reactions containing ^{3}H-labeled nuclear RNA from infected cells alone or mixed with single-stranded DNA (Fig. 5), as well as in reactions containing ^{3}H-labeled nuclear RNA from uninfected 3T6 cells (data not shown). The background was resistant to S1 nuclease and RNase A digestions; therefore, it was probably produced by double-stranded RNA. RNase III is an endonuclease which specifically degrades double-stranded RNA (Robertson et al, 1968) and RNase H degrades the RNA moiety of DNA-RNA hybrids (Stein and Hausen, 1969; Stavrianopoulos and Chargaff, 1973). By treating the annealed RNA with either RNase III or RNase H after S1 nuclease and RNase A treatment, I attempted to identify the nature of the radioactive background.

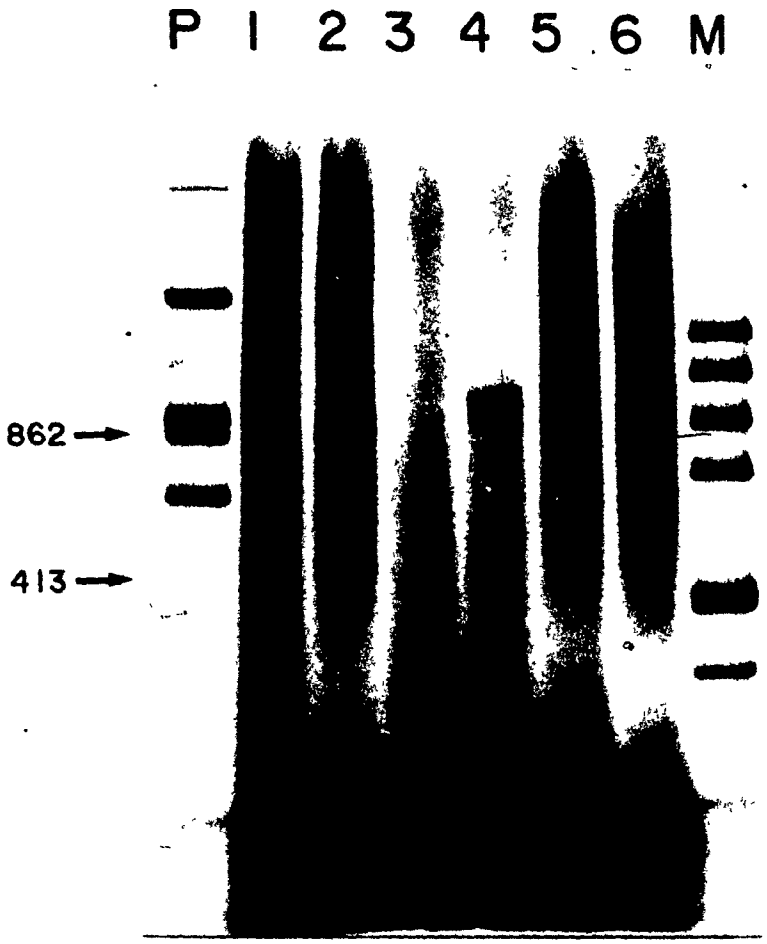
^{3}H-labeled Nuclear RNA either alone or with single-stranded fd103-polyomavirus fragment PstI-4L clone was incubated in the formamide hybridization buffer at 85°C for 10 min, then at 55°C for 3 hr. Samples were digested with S1 nuclease (100 units/ml) and RNase A (25 µg/ml) (Fig. 6, lane 1, RNA alone; lane 2, RNA with DNA). Two additional samples were treated with RNase III (100 units/ml) at 37°C for 1 hr (Fig. 6, lane 3, RNA alone; lane 4, RNA with DNA). The backgrounds of both samples were reduced. However, the DNA-RNA hybrids, for some unknown reason, migrated more slowly and the hybrid bands became more diffuse than those in the control. Two other samples were treated with RNase H (100 units/ml) at 37°C for 1 hr (Fig. 6, lane 5, RNA alone; lane 6, RNA with DNA). The background remained the same, but the intensity of the DNA-RNA hybrid bands in lane 6 was reduced. Since the background can be reduced by treatment with RNase III but not with RNase H, I concluded that double-stranded RNA is the most likely cause of the radioactive background.

Effect of Filter Binding on Recovery of DNA-RNA Hybrids and Background Radioactivity.

In order to reduce the background, I took advantage of the fact that single-stranded nucleic acids bind to nitrocellulose filters in high salt, whereas double-stranded nucleic acids do not (Nygaard and Hall, 1963; Gillespie and Spiegelman, 1965). Hybridization of RNA to single-stranded recombinant bacteriophage DNA, 7 to 11 kb long, will give rise to DNA-RNA hybrids joined to single-stranded bacteriophage DNA. The hybrids will therefore bind to nitrocellulose filters *via* their single-stranded DNA tails.

To test the filter binding method, a solution containing high salt and no formamide as the hybridization buffer was used (Padgett *et al*, 1983). I

A.



B.

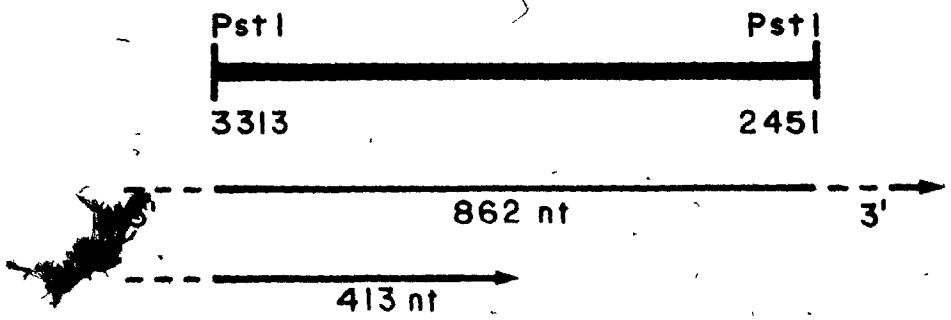


Figure 6: Effects of RNase III and RNase H on background radioactivity.

A. Three sets of samples containing either 15 µg of nuclear RNA labeled for 150 minutes alone (lanes 1, 3, 5) or RNA mixed with 3 µg of single-stranded fd103 bacteriophage DNA containing cloned polyomavirus PstI-4L were suspended in 40 µl of 80% formamide hybridization buffer. Samples were incubated at 85°C for 10 minutes, then transferred to 55°C for 3 hours before adding 360 µl of ice-cold S1 buffer. Hybrids were treated with S1 nuclease (100 units/ml) at 45°C for 30 minutes, and RNase A (2.5 µg/ml) at 30°C for 30 minutes followed by Proteinase K (250 µg/ml) at 37°C for 30 minutes, then ethanol precipitated as described in Materials and Methods. One set of samples was kept at this step and served as the control (lanes 1 and 2). One set was further treated with RNase III (100 units/ml) at 37°C for 1 hour (lanes 3 and 4); one other set was treated with RNase H (100 units/ml) at 37°C for 1 hour (lanes 5 and 6). Markers were double-stranded ^{3}H-polyomavirus DNA digested with PstI (lane P) or with MspI (lane M). Sizes and positions of expected DNA-RNA hybrids were indicated.

B. Diagram showing polyomavirus DNA fragment PstI-4L. Numbers below the solid line show positions of the ends of the fragment. The thin lines below the solid line show sizes of DNA-RNA hybrids corresponding to RNAs which read through this region or have 3' ends at nucleotide 2900.

first compared DNA-RNA hybridization in buffers with or without formamide. Results are shown in Figure 7. ^{3}H-labeled nuclear RNA and single-stranded bacteriophage DNA containing a polyomavirus DNA insert were suspended either in formamide hybridization buffer or in 50 mM HEPES pH 7.0, 0.75 M NaCl, 1 mM EDTA (high salt hybridization buffer). The hybridization mixtures were heated at 85°C for 10 min and transferred to 55°C for 2 hours. Hybrids were treated with S1 nuclease (100 units/ml) and then with RNase A (5 µg/ml) (Fig. 7, lanes 1 & 2). In both cases, the same amount of hybrids was observed (same intensity), indicating that hybridization was as efficient in the buffer without formamide as in the buffer with formamide under the conditions used. Another sample was treated first with RNase A (1 µg/ml), then with S1 nuclease (100 units/ml) (Fig. 7, lane 3). This resulted in the same amount of hybrid bands as the other samples; however, the background was higher due to the lower concentration of RNase A used.

To test whether filter-binding could reduce the radioactive background, ^{3}H-labeled nuclear RNAs were hybridized with single-stranded fd103-polyomavirus PstI-4L DNA in high salt hybridization buffer at 85°C for 10 min, then at 55°C for 2 hr. Hybrids were treated with RNase A (1 µg/ml) at 30°C for 30 min, followed by proteinase K (250 µg/ml) at 37°C for 30 min to inactivate the RNase. The hybridization mixture was divided into 3 equal parts. One part was extracted with phenol-chloroform-isoamyl alcohol (50:50:1) and then ethanol precipitated. The pellet was resuspended and digested with S1 nuclease (100 units/ml) at 45°C for 30 min, then analyzed on an agarose-acrylamide gel in Tris-borate buffer (Fig. 8, lane 2). NaCl was added to a final concentration of 0.75 M to the other two parts. Subsequently, they were passed through 13 mm nitrocellulose filters and were washed with 2 ml of the same buffer. One filter was treated with S1 nuclease to release DNA-RNA hybrids (Fig. 8, lane 3);

1 2 3 4

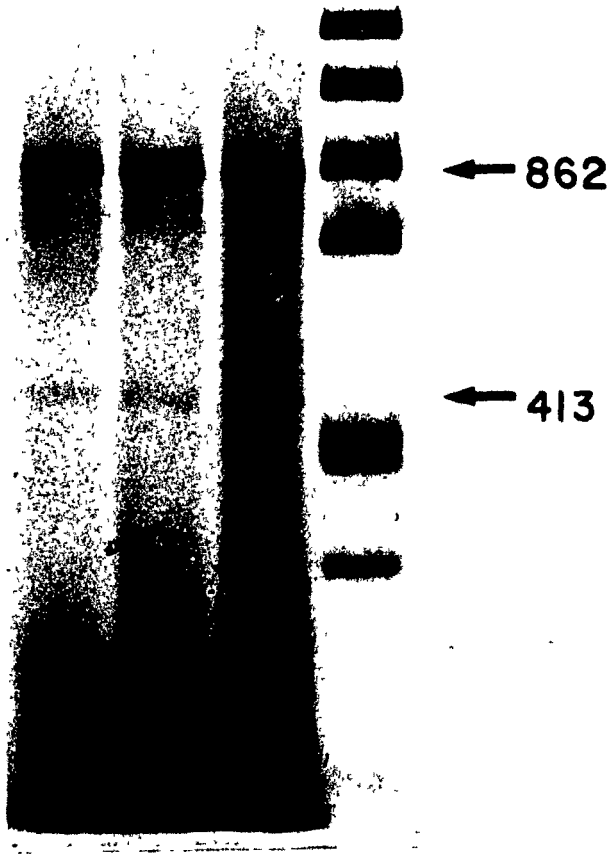
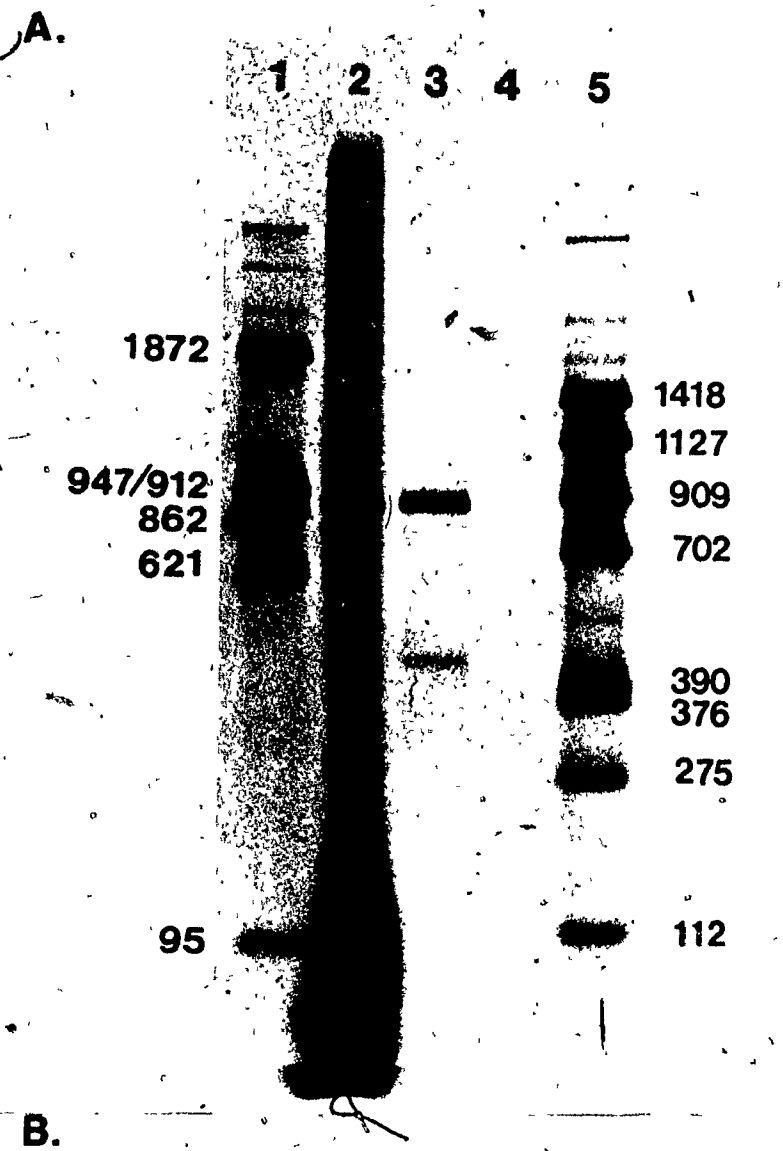


Figure 7: Comparison of hybridization in the presence or absence of formamide.

Fifteen µg of nuclear RNA labeled for 60 minutes was mixed with 3 µg of single-stranded fd103 bacteriophage DNA containing cloned polyomavirus PstI-4L. Mixtures were resuspended either in buffer which contains formamide (80% formamide, 0.4 M NaCl, 40 mM PIPES pH 6.4, 1 mM EDTA [lane 1]) or in buffer which contains no formamide (50 mM HEPES pH 7.0, 0.75 M NaCl, 1 mM EDTA [lanes 2 and 3]). All samples were heated at 85°C for 10 minutes, and then transferred to 55°C for 2 hours. Hybrids were treated with S1 nuclease (100 units/ml) and RNase A either at concentrations of 5 µg/ml (lanes 1 and 2) or 1 µg/ml (lane 3) as described in Materials and Methods. Markers were ^{3}H-polyomavirus DNA digested with MspI (lane 4). Sizes and positions of expected DNA-RNA hybrids are indicated.


A.
1
2
3
4
5
1872
947/912
862
621
95
1418
1127
909
702
390
376
275
112

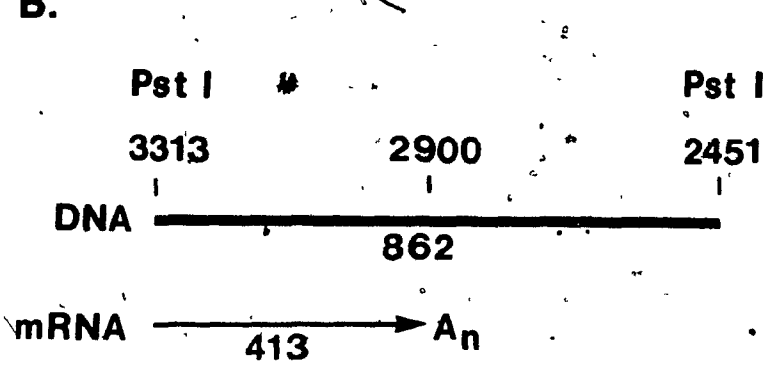

B.
Pst I
Pst I
3313
2900
2451
DNA
862
mRNA
413
An

Figure 8: Effect of filter binding on recovery of DNA-RNA hybrids and background radioactivity.

A. Nuclear RNA (37 μg, 1.7 x 10^6 cpm) from cells labeled with ^{3}H-uridine for 60 minutes was hybridized with 7.5 μg single-stranded bacteriophage fd103 DNA containing cloned polyomavirus PstI-4L. Hybrids were treated with RNAse A followed by proteinase K. An aliquot was phenol-chloroform extracted, ethanol precipitated, and digested with S1 nuclease (lane 2). Two other equal aliquots were bound to nitrocellulose filters (see Materials and Methods). One filter was incubated with S1 nuclease as above (lane 3); the other filter was washed with 1 mM EDTA (lane 4) before incubation with S1 nuclease. The 3 samples were analyzed by electrophoresis followed by fluorography as described in Materials and Methods. The major band in lanes 2 and 3 is the 862-nucleotide full-length DNA-RNA hybrid; the minor band is the 413-nucleotide hybrid formed with polyadenylated RNAs whose 3' ends are at nucleotide 2900 (see Fig. 8B). Markers were ^{3}H-polyomavirus DNA digested with PstI (lane 1) or HpaII (lane 5).

B. Diagram showing polyomavirus DNA fragment PstI-4L. Numbers above the line show positions of the ends of the fragment. Numbers below the lines show expected lengths of hybrids arising from hybridization with RNA which extends throughout the fragment or with RNA whose 3' end is at nucleotide 2900.

the other filter was further washed with 2 ml of 1 mM EDTA solution before S1 nuclease treatment (Fig. 8, lane 4). As shown on the autoradiogram, the filter binding step allowed retention of most of the hybrids and significantly reduced the background (Fig. 8, lane 3). However, most DNA-RNA hybrids were lost from the filter during washing with 1 mM EDTA.

Kinetics of Formation of DNA-RNA Hybrids in High Salt Hybridization Buffer.

Since single-stranded DNA is used in the hybridization experiments, DNA reannealing should not occur. Therefore denaturation prior to hybridization may not be necessary. Results of my experiments have shown that hybridization is as efficient in the high salt hybridization buffer as it is in the formamide hybridization buffer (see Figure 7). It was important to know the kinetics of hybridization in the high salt hybridization buffer, for use in future experiments.

Nuclear RNA labeled *in vivo* for 1 hr with ^{3}H-uridine and single-stranded fd103-polyomavirus PstI-4L DNA were suspended in the high salt hybridization buffer and were incubated at 85°C for 5 min. An aliquot was removed from the mixture and cooled rapidly in 9 volumes of ice-cold S1 buffer (Fig. 9, lane 1). The rest of the mixture was transferred to 55°C, aliquots were removed at 5, 30, 60 and 120 min intervals (Fig. 9, lanes 2 to 5 respectively), and cooled rapidly in S1 buffer. All samples were treated with S1 nuclease (100 units/ml) and RNase A (2 µg/ml). As shown in Figure 9, all samples showed the same intensity of hybrid bands, indicating that the hybridization was complete at time 0. This means that, at the DNA concentration used in this experiment (15 µg/200 µl), DNA-RNA hybridization is so rapid that it is essentially complete after 5 min at 85°C plus the time to cool the solution to 0°C.

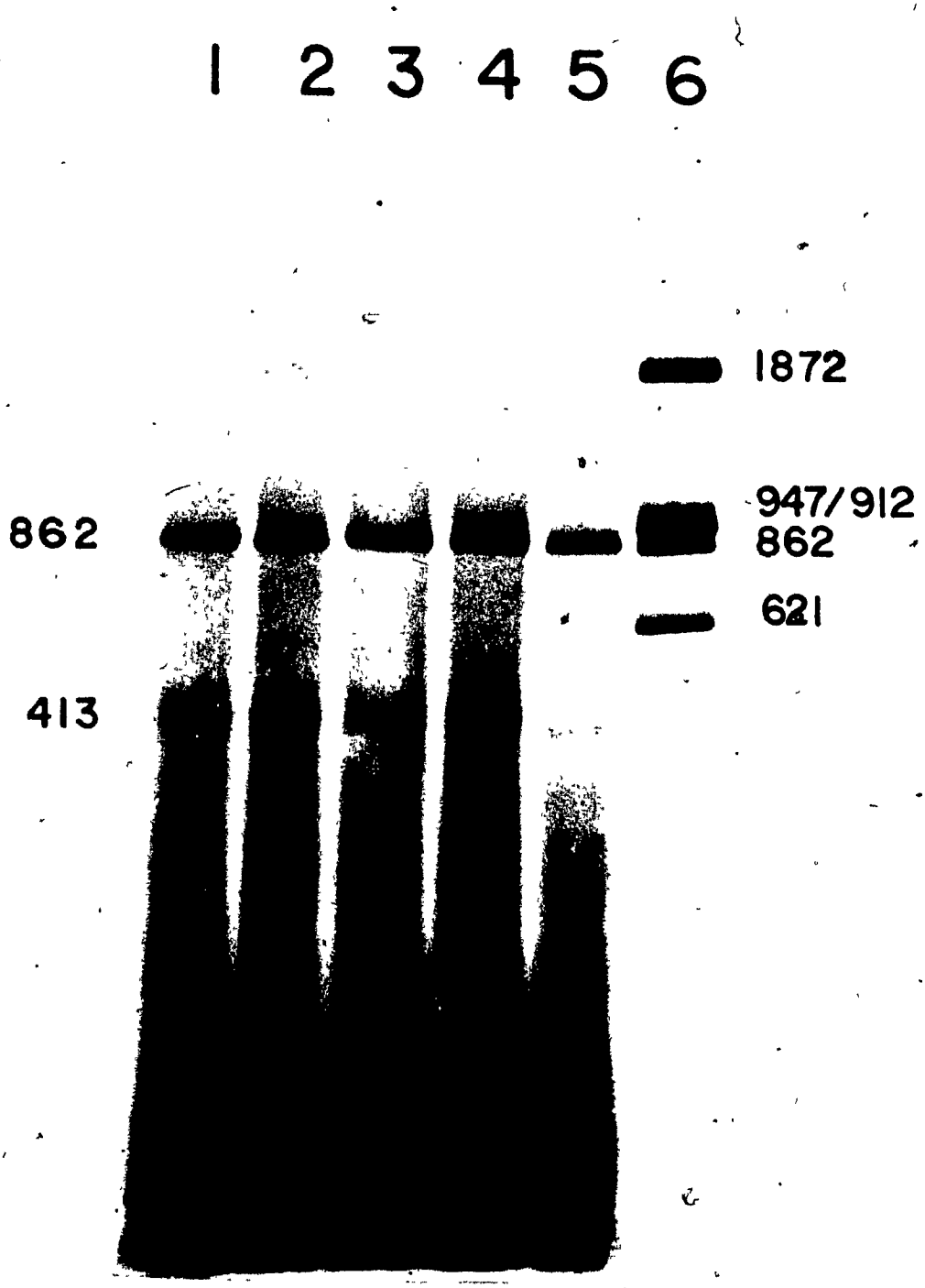

1 2 3 4 5 6
1872
947/912
862
621
862
413

Figure 9: Kinetics of formation of DNA-RNA hybrids in 50 mM HEPES, 0.75 M NaCl, 1 mM EDTA (high salt hybridization buffer).

Nuclear RNA (75 µg) labeled for 60 minutes at 28 hours post infection was hybridized with 15 µg of single-stranded fd103 bacteriophage DNA containing cloned polyomavirus PstI-4L. The hybridization mixture was incubated at 85°C for 5 minutes. Aliquots (40 µl) were removed before and at different times after incubation at 55°C and diluted into 360 µl of ice-cold S1 buffer. Hybrids were treated with 100 units/ml of S1 nuclease, 2 µg/ml RNase A and 125 µg/ml of proteinase K as described in Materials and Methods. Lane 1 is time 0 before hybridization was carried out at 55°C. Other lanes represent different times of incubation at 55°C: 5 minutes (lane 2), 30 minutes (lane 3), 60 minutes (lane 4) and 120 minutes (lane 5). Markers were ^{3}H-polyomavirus DNA digested with PstI (lane 6). Sizes and positions of expected DNA-RNA hybrids are indicated.

Determination of S1 Nuclease Concentration Required to Release DNA-RNA Hybrids from Nitrocellulose Filters.

S1 nuclease is used to release the DNA-RNA hybrids from the filters. To determine the optimum concentration of S1 nuclease, a titration experiment was performed. Since the RNase concentration may change the digestion pattern of S1 nuclease, two concentrations of RNase A were used. After hybridization, one set of samples was treated with 0.5 µg/ml of RNase A (Fig. 10, lanes 2 to 5), and the other set was treated with 5 µg/ml of RNase A (lanes 7 to 9). Subsequently, all samples were treated with proteinase K to inactivate the RNase, and hybrids were bound to filters. Filters were placed in flat-bottomed tubes containing 0.4 ml of S1 buffer and different amounts of S1 nuclease; 33 units/ml (lanes 2 & 7), 167 units/ml (lanes 3 & 8), 333 units/ml (lanes 4 & 9), or 667 units/ml (lanes 5 & 10). After incubation at 45°C for 30 min the supernatant was collected and phenol-chloroform extracted. The hybrids were then run on an agarose-polyacrylamide gel. As shown in Figure 10, 33 units/ml of S1 nuclease was not enough to release all hybrids treated with 0.5 µg/ml RNase from the filters (Fig. 10, lane 2); however, most of the hybrids treated with 5 µg/ml RNase were released but were not trimmed to the right size (Fig. 10, lane 7: note slower migration of the 1126 and 686 hybrid bands). Treatment with 5 µg/ml RNase A gave some extra bands in addition to the expected hybrid bands. Although the other three concentrations of S1 nuclease used gave rise to very similar results, i.e. they were all efficient in releasing all hybrids from the filters, 0.5 µg/ml RNase followed by 333 units/ml S1 nuclease (Fig. 10, lane 4) gave the best results.

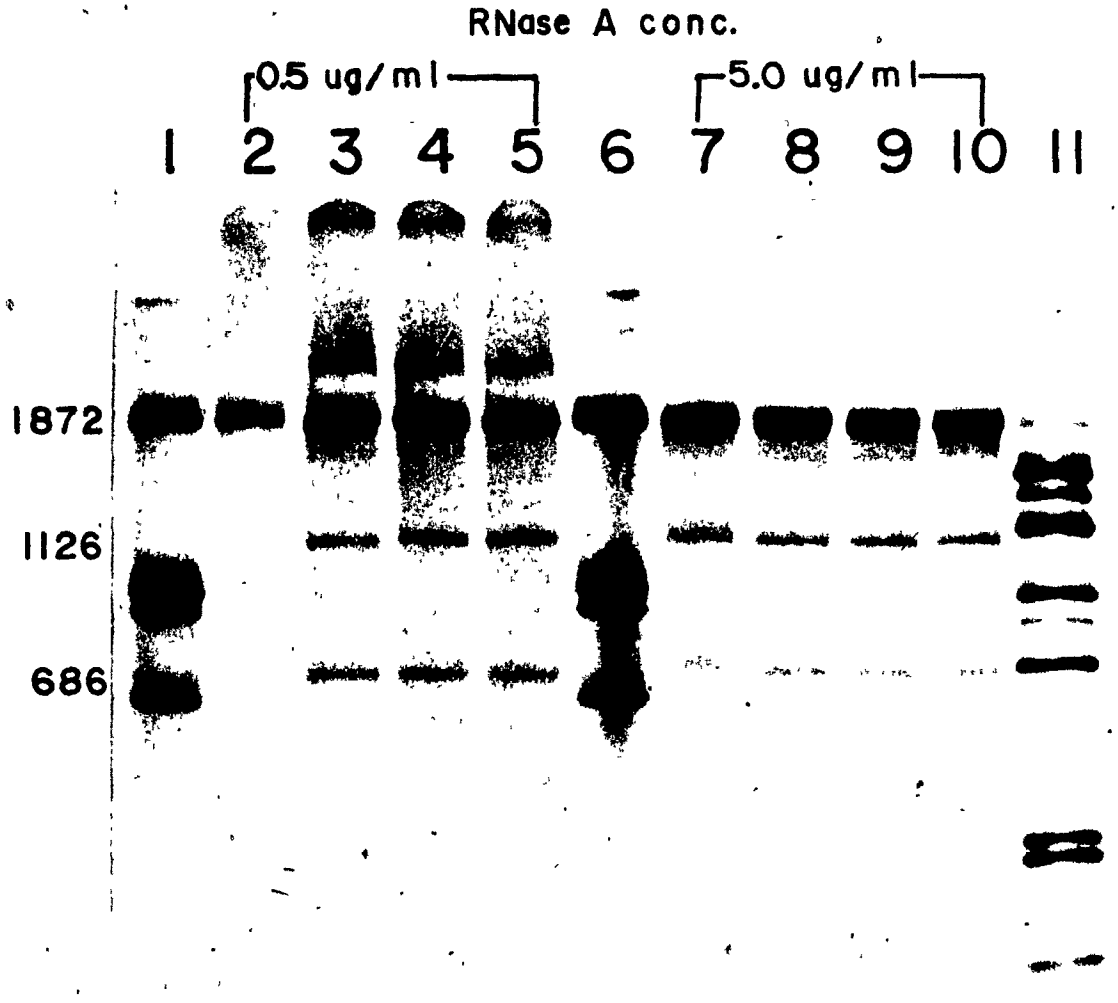

RNase A conc.
0.5 ug/ml
5.0 ug/ml
1 2 3 4 5 6 7 8 9 10 11
1872
1126
686

Figure 10: Effects of different concentrations of S1 nuclease on release of DNA-RNA hybrids from nitrocellulose filters.

Two samples of nuclear RNA (60 µg) labeled for 60 minutes were separately hybridized to 12 µg of single-stranded fd103 bacteriophage DNA containing cloned polyomavirus PstI-4L in 120 µl of 50 mM HEPES pH 7.0, 0.75 M NaCl, 1 mM EDTA buffer at 85°C for 5 minutes and then transferred to 55°C for 5 minutes. An equal volume of 50 mM HEPES pH 7.0, 1 mM EDTA was added to dilute the concentration of NaCl to 0.375 M. One sample was treated with 0.5 µg/ml of RNase A (lanes 2 to 5) and the other was treated with 5 µg/ml of RNase A (lanes 7 to 10) at 30°C for 30 minutes, both followed by proteinase K (125 µg/ml) treatment. NaCl was added to bring the concentration to 0.75 M. Each sample was divided into four equal portions and was passed through four 13 mm-nitrocellulose filters. Each filter was washed with 2 ml of loading buffer. The filters were treated with different concentrations of S1 nuclease: 33 units/ml (lanes 2 and 7), 167 units/ml (lanes 3 and 8), 333 units/ml (lanes 4 and 9), or 667 units/ml (lanes 5 and 10) at 45°C for 30 minutes. Hybrids were analyzed by gel electrophoresis. Markers were double-stranded ^{3}H-polyomavirus DNA digested with PstI (lanes 1 and 6) or MspI (lane 11). Sizes and positions of expected DNA-RNA hybrids are indicated by the numbers on the left hand panel (see Figure 5).

Effect of RNase Concentrations on DNA-RNA Hybrids.

In the filter binding method, unhybridized viral and cellular RNA is degraded by RNase before binding hybrids to the nitrocellulose filters. This treatment also trims any protruding single-stranded RNA tails from DNA-RNA duplexes. RNase A can nick the RNA portion of DNA-RNA hybrids non-specifically at certain sites (for example, see Skarnes, 1985). Nicking or gapping of RNA may allow S1 nuclease to degrade the DNA opposite the RNA breaks in DNA-RNA duplexes. Therefore, it is important to use a concentration of RNase which does not nick the RNA portion of DNA-RNA hybrids. Both RNase A and T1 are endonucleases. RNase A cleaves RNA after pyrimidine residues, and RNase T1 cleaves RNA after G residues. Therefore, RNase A and T1 were compared for their ribonuclease activity and nicking properties.

After hybridization, the hybrids were treated with different concentrations of RNase A or RNase T1 separately, and then were bound to nitrocellulose filters. The hybrids were released from the filters by incubation with 333 units/ml S1 nuclease and were analyzed on an agarose-polyacrylamide gel. Results are shown in Figure 11. When the hybrids were incubated at 45°C for 30 min with 10 µg/ml of RNase A, the full-length hybrids were completely degraded into many smaller hybrids (Fig. 11, lane 2). Incubation at 30°C for 30 min with 5 µg/ml of RNase A also led to degradation of some of the full-length hybrids and extra hybrid bands were seen (lane 3). When hybrids were treated with 0.5 µg/ml or 0.1 µg/ml of RNase A (lanes 4 & 5), or with 5000, 1000 or 100 units/ml of RNase T1 (lanes 6, 7 & 8), little or no degradation of DNA-RNA hybrids was noted. The background radioactivity was higher in the samples treated with 10 units/ml of RNase T1 (lane 9). Therefore, 0.1 to 0.5 µg/ml of RNase A or


RNase A
RNase T1
1 2 3 4 5 6 7 8 9 10
1872
1126
686

Figure 11: Effects of RNase concentrations on DNA-RNA hybrids.

Two hundred µg of ^{3}H-labeled nuclear NRA was hybridized to 40 µg of single-stranded fd103 bacteriophage DNA containing cloned polyomavirus PstI-1L in 400 µl of 50 mM HEPES pH 7.0, 0.75 NaCl, 1 mM EDTA at 85°C for 5 minutes followed by 55°C for 5 minutes. Hybrids were divided into 8 equal portions and were treated with different amounts of RNase A or T1 at 30°C (except for lane 2 at 45°C) for 30 minutes, followed by proteinase K digestion. The concentrations of RNase used were as follows: RNase A 10 µg/ml (lane 2), 5 µg/ml (lane 3), 0.5 µg/ml (lane 4), 0.1 µg ml (lane 5); and RNase T1 5,000 units/ml (lane 6), 1,000 units/ml (lane 7), 100 units/ml (lane 8), 10 units/ml (lane 9). Hybrids were bound to nitrocellulose filters and eluted by incubation with 333 units/ml of S1 nuclease at 37°C for 1 hour. Markers were double-stranded ^{3}H-polymavirus DNA digested either with PstI (lane 1) or with MspI (lane 10). Sizes and positions of expected DNA-RNA hybrids are indicated.

100 to 5000 units/ml of RNase T1 can be used to degrade unhybridized RNA followed by incubation with 333 units/ml S1 nuclease, without leading to significant non-specific degradation of DNA-RNA hybrids.

Determination of DNA Excess in Solution Hybridizations.

To ensure that hybridizations between ^{3}H-labeled RNA and unlabeled single-stranded DNA detects all RNA species present, it is important to establish that DNA is in molar excess over complementary RNA. To prove DNA excess, 5 µg of single-stranded fd103-polyomavirus PstI-2L DNA was separately hybridized to 5, 10, 25, 50, 100 or 200 µg of ^{3}H-labeled nuclear RNA from polyomavirus infected 3T6 cells. Hybridization mixtures or RNA alone were treated with RNase A (0.1 µg/µg RNA in 0.25 ml digestion buffer), followed by proteinase K. Hybrids were bound to nitrocellulose filters, which were washed and counted by scintillation spectrometry to measure bound ^{3}H radioactivity. Results are shown in Figure 12. Hybridization is linear with up to 25 µg of RNA, and does not reach a plateau even at an RNA concentration of 200 µg. In all subsequent experiments, hybridization was done with a RNA/DNA ratio of 5:1 to ensure DNA excess.

Binding of Single-stranded DNA to Nitrocellulose Filters.

In order to determine the optimal amount of single-stranded DNA to use in experiments, I have measured the binding of different amounts of single-stranded DNA to 13 mm diameter nitrocellulose filters. ^{3}H-labeled polyomavirus DNA (about 0.5 µg, 60,000 cpm) linearized by EcoRI was mixed with 102 µg of single-stranded fd103 DNA in 50 mM HEPES pH 7.0, 1 mM EDTA. After denaturing at 100°C for 10 min and rapid cooling in an ice-water bath, NaCl was added to a final concentration of 0.75 M. Aliquots of the

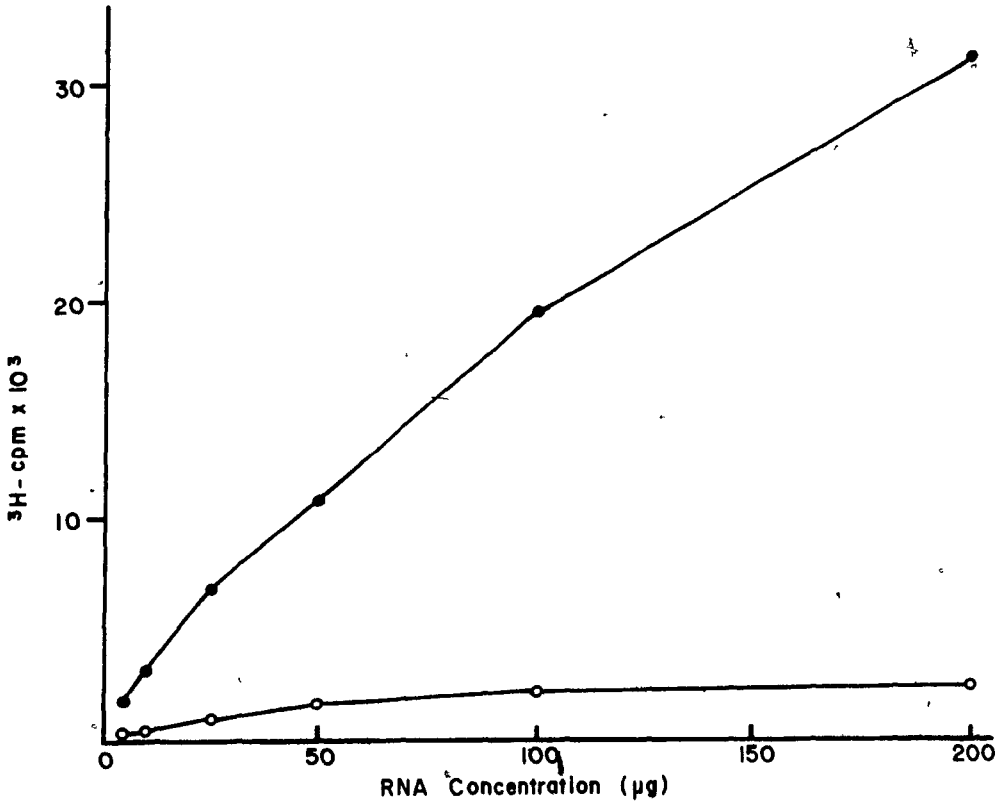

30
20
10
^{3}H-cpm x 10^3
0
50
100
150
200
RNA Concentration (µg)

Figure 12: Demonstration of DNA excess in solution hybridization of ^{3}H-labeled nuclear RNA with single-stranded bacteriophage-polyomavirus DNA.

The following amounts of ^{3}H-labeled nuclear RNA were alone or were hybridized with 5 µg of single-stranded fd103 containing polyomavirus PstI-2L fragment: 5, 10, 25, 50, 100, and 200 µg. Hybrids were treated with RNase A (0.1 µg per µg of RNA in 0.25 ml of digestion volume), followed by proteinase K. Hybrids were bound on nitrocellulose filters. The filters were dried; then scintillation fluid was added to each filter which was counted in an LKB scintillation spectrometer. (•) - DNA-RNA hybrids (RNA hybridized with DNA). (o) - RNA alone.

mixture containing 1, 5, 10, 15 or 20 µg of single-stranded DNA were filtered separately through nitrocellulose filters. The filters were washed with 3 x 1 ml of loading buffer (50 mM HEPES pH 7.0, 0.75 M NaCl, 1 mM EDTA) and then counted in an LKB scintillation spectrometer. Results are shown in Fig. 13. The binding of single-stranded DNA was quantitative with up to 15 µg of DNA per filter. Less radioactivity bound to filters which received 20 µg DNA compared to those which received 15 µg DNA. Therefore, the maximum binding capacity of the filters in this assay is approximately 15 µg single-stranded fd103 DNA. In most experiments which follow, I have used no more than 10 µg DNA per 13 mm filter.

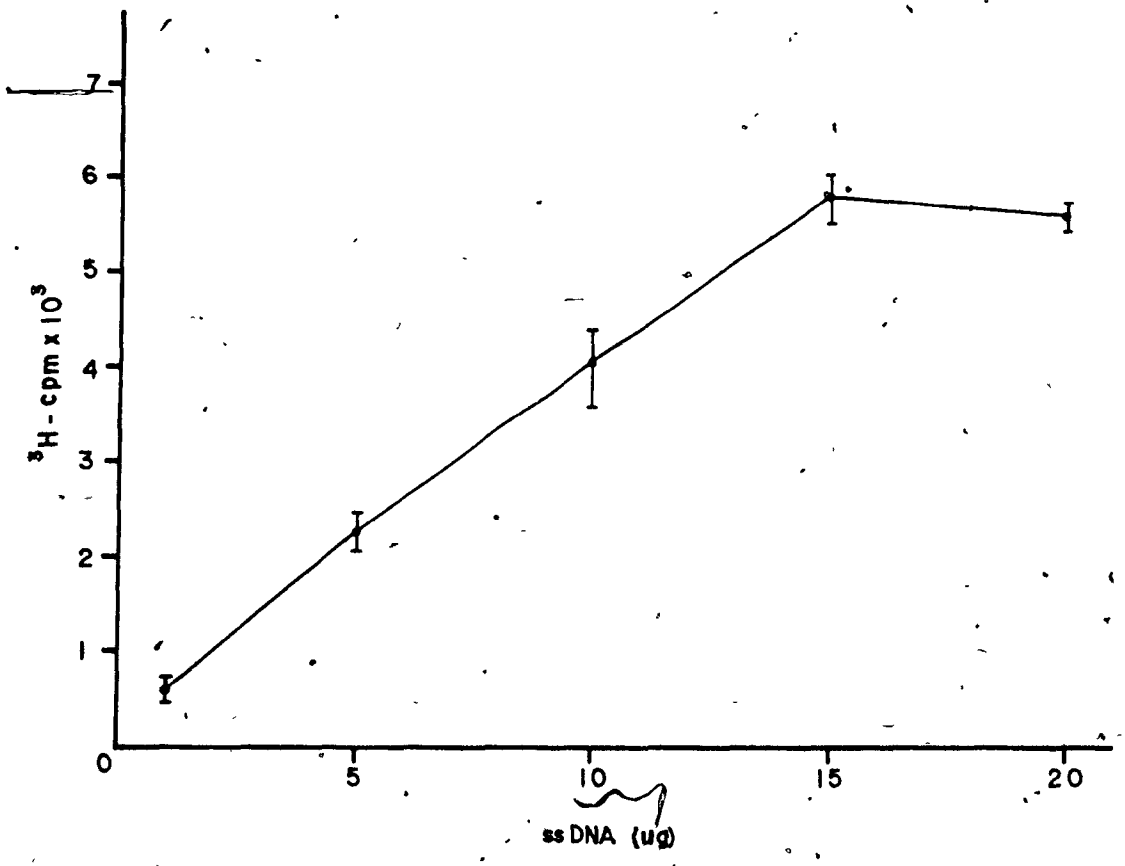

7
6
5
4
3
2
1
0
^{3}H - cpm x 10^3
5
10
15
20
ss DNA (ug)

Figure 13: Saturation curve of single-stranded DNA binding to 13 mm diameter nitrocellulose filters.

Single-stranded bacteriophage fd103 DNA (102 µg) was mixed with 60,000 cpm of EcoRI-digested ^{3}H-polyomavirus DNA in 50 mM HEPES pH 7.0. The mixture was incubated at 100°C for 10 minutes, then cooled rapidly in an ice-water bath. NaCl was added to a final concentration of 0.75 M. Aliquots corresponding to 1, 5, 10, 15 and 20 µg of ss DNA were bound on nitrocellulose filters (2 filters of each). The filters were washed with 2 ml of 50 mM HEPES pH 7.0, 0.75 M NaCl and then dried. Scintillation fluid was added to each filter, which was counted in an LKB scintillation spectrometer.

DISCUSSION

The thermal stability of nucleic acids affects the temperature of denaturation and hybridization. Several factors are known to affect the thermal stability; these are the concentration of the denaturing agent (e.g. formamide), the G + C content, the molarity of monovalent cations and the length of the nucleic acids. In general, the thermal stability of an RNA-RNA duplex is greater than that of a DNA-RNA hybrid, which, in turn, is greater than that of a DNA-DNA duplex (Casey and Davidson, 1977; Gray et al, 1981). The Tm (DNA) for a DNA duplex under varying experimental conditions can be estimated using the equation Tm(DNA) = 81.5 + 16.6 (log M) + 0.41 (% G + C) - 0.72 (% formamide); where M is the molarity of the monovalent cation, % G + C is the percentage of G + C, and (% formamide) is the percentage of formamide (Howley et al, 1979). A similar equation for RNA-RNA hybrids can be derived which accounts for the effect of (G + C) content, the molarity of monovalent cation (Gray et al, 1981) and the percentage of formamide (Cox et al, 1984). Thus, Tm(RNA) = 79.8 + 18.5 (log M) + 58.4 H_G + 11.8 $(H_G)^2$ - 0.35 (% formamide); where H_G is the (G + C) content represented as a fraction (Kallenbach, 1968).

Polyomavirus DNA contains 48% G + C. According to the above formulae, the Tm of polyomavirus double-stranded DNA in the formamide hybridization buffer containing 80% formamide and 0.4 M NaCl is 37°C and in the high salt hybridization buffer containing 0.75 M NaCl and no formamide is 99°C; the Tm of polyomavirus double-stranded RNA (self-annealing) in the formamide hybridization buffer is 75°C and in the high salt hybridization buffer is 108°C. Thus the Tm of polyomavirus DNA-RNA hybrids should be somewhere between these values. In the formamide hybridization buffer, the RNAs (both cellular and viral) were completely denatured at 85°C (Fig. 5).

Therefore hybridization was carried out at 55°C which is above the Tm of polyomavirus DNA. Since single-stranded DNA is used, theoretically the hybridization should be able to go to completion at any temperature below the Tm(RNA). In the high salt hybridization buffer, hybridization was complete within a 5 min incubation period at 85°C, suggesting that at 85°C DNA-RNA hybridization occurred and that the hybrids formed were stable (Fig. 9). This result can only be achieved by using single-stranded DNA in the hybridization reaction, because 85°C is still below the Tm of polyomavirus DNA in the high salt buffer.

When pulse-labeled nuclear RNA was hybridized with unlabeled single-stranded DNA in solution, a heavy nuclease-resistant radioactive background was produced. The background was so heavy that minor hybrid bands could only be seen with difficulty. The background was also produced in reactions containing labeled nuclear RNA alone (Fig. 5) and in reactions using labeled nuclear RNA from uninfected 3T6 cells (data not shown). This indicates that the background could be due to labeled cellular RNA. Since the background is resistant to digestion by both S1 nuclease and ribonuclease A, it is likely that it is due to labeled double-stranded RNA. The background was reduced when the hybridization mixture digested with RNase A and S1 nuclease was further treated with RNase III, an enzyme which specifically degrades double-stranded RNA (Robertson et al, 1968), but was not with RNase H, an enzyme that degrades the RNA moiety of DNA-RNA hybrids (Fig. 6). This result indicates that the background was indeed caused by double-stranded RNA. Interestingly, the hybrid bands became diffuse and migrated more slowly after RNase III treatment (Fig. 6, lane 4). The reason for this change is not known; it is possible that RNase III binds to the DNA-RNA hybrids and causes the shift of the hybrid band.

Single-stranded DNA binds efficiently to nitrocellulose filters in high salt, but double-stranded nucleic acids do not (Nygaard and Hall, 1963; Gillespie and Spiegelman, 1965). The DNA-RNA hybrids in our hybridization reaction contain long tails of bacteriophage fd103 or M13 single-stranded DNA. After hybridization the hybrids were treated with RNase to degrade unhybridized RNA, then bound to nitrocellulose filters and released from the filters by S1 nuclease treatment. Results showed that the filter binding step did significantly reduce the radioactive background while retaining most DNA-RNA hybrids (Fig. 8). When the filter was washed with 1 mM EDTA almost all the hybrids were lost, indicating that binding of single-stranded DNA to nitrocellulose is reversible in the absence of NaCl. Since I recovered all the hybrids from the filter in 0.75 M NaCl, I did not investigate the minimal salt concentration required for the binding of single-stranded DNA to nitrocellulose filters.

In the filter binding method hybrids are treated with ribonuclease before the filter binding step. Unhybridized RNA is degraded and therefore dose not bind to the filters. It is known that high concentrations of RNase can nick the RNA portion of DNA-RNA hybrids; in consequence, the hybrids could be degraded by S1 nuclease (Fig. 11, lane 2). Therefore, a concentration of RNase which degrades unhybridized RNA to a maximal extent without nicking the hybrids is required. As shown in Figure 11, 0.1 to 0.5 µg/ml of RNase A or 100 to 5000 units/ml RNase T1 gave similar results and did not nick the RNA portion of the hybrids when 333 units/ml of S1 nuclease was used.

When hybrids were treated with S1 nuclease in solution, 50 to 100 units/ml of S1 nuclease was enough to trim the hybrids to uniform size and gave rise to sharp hybrid bands (data not shown). To release hybrids from nitrocellulose filters, a higher concentration of S1 nuclease was required.

As shown in Figure 10, 333 to 667 units/ml of S1 nuclease produced sharp hybrid bands when 0.5 µg/ml RNase A was used. The requirements for higher concentrations of S1 nuclease to release the hybrids from the filters could be due to the limitation of contact between the enzyme and its filter-bound substrate.

In order to detect all the viral specific RNA in the hybridization mixture, a molar excess of DNA has to be used. As shown in Figure 12, the curve was linear when up to 25 µg of RNA was used, but did not reach a plateau even when 200 µg of RNA was used. This means that there is still a DNA excess when 200 µg of RNA was used in the hybridization. In this experiment, 0.1 µg of RNase A per µg of RNA was used to degrade the unhybridized RNA in 0.25 ml of digestion solution. The reason for using different amounts of RNase A to digest different amounts of RNA was to avoid under- or over-digestion when RNase concentration and digestion volume are fixed. Another way of doing this experiment would be to use a fixed RNase concentration but increasing digestion volume when more RNA is used in the hybridization. The highest amount of RNA used in this experiment (200 µg) was 40 times the lowest amount of RNA (5 µg), so the digestion volume would have to be increased 40-fold. Therefore I did not do the experiment this way. Digestion of high amounts of RNA with higher concentrations of RNase A may cause more degradation of RNA including hybridized RNA. This could be one of the reasons why with higher amounts of RNA the curve is not linear, even though the DNA is probably still in excess. The other reason could be that degraded RNA interferes with the binding of single-stranded DNA to the filter.

The binding capacity of 13 mm diameter nitrocellulose filters for single-stranded fd103 DNA is about 15 µg per filter (Fig. 13). More points between 10 and 50 µg of DNA would be required to determine the exact

binding capacity. Ten µg of DNA should not saturate the filter and should be completely retained by the filter under the experimental conditions used.

Taken together, in 50 mM HEPES pH 7.0, 0.75 M NaCl, 1 mM EDTA buffer, hybridization was complete within 5 min incubation at 85°C. The RNA to DNA ratio was 5 to 1 to ensure DNA excess and a maximum of 10 µg of single-stranded DNA was used to avoid saturating the filter. Therefore, 50 µg of nuclear RNA and 10 µg of DNA clone were used in each hybridization mixture. 0.1 to 0.5 µg of RNase A or 100 to 5000 units/ml of RNase T1 did not nick DNA-RNA hybrids, and 333 to 667 units/ml of S1 nuclease was used to release the hybrids from nitrocellulose filters and gave rise to sharp hybrid bands.

GUIDELINES REGARDING DOCTORAL THESIS CONTAINING QUOTATIONS FROM PUBLISHED OR SUBMITTED MANUSCRIPTS

The Candidate has the option, subject to the approval of the department, of including as part of the thesis the text of an original paper, or papers, suitable for submission to learned journals for publication. In this case the thesis must be still conform to all other requirements explained in Guidelines Concerning Thesis Preparation, (available at the thesis office). Additional material (experimental and design data as well as description of equipment) must be provided in sufficient detail to allow a clear and precise judgement to be made of the importance and originality of the research reported. Abstract, full introduction and conclusion must be included, and where more than one manuscript appears, connecting texts and common abstracts, introduction and conclusions are required. A mere collection of manuscripts is not acceptable; nor can reprints of published papers be accepted.

While the inclusion of manuscripts co-authored by the Candidate and others is not prohibited by McGill, the Candidate is warned to make an explicit statement on who contributed to such work and to what extent, and Supervisors and others will have to bear witness to the accuracy of such claims before the Oral Committee. It should also be noted that the task of the External Examiner is made much more difficult in such cases, and it is in the Candidate's interest to make authorship responsibilities perfectly clear.

In accordance with University regulations the above text has been quoted in full from Guidelines Concerning Thesis Preparation. One manuscript has been included as the body of this thesis, and appears forthwith. A common abstract, literature review, materials and methods,

and conclusions have been presented as required in the above guidelines. The specific contributions of the authors are presented in detail below in accordance with the same guidelines.

Manuscript Titled "**Use of a novel S1 nuclease RNA-mapping technique to measure efficiency of transcription termination on polyomavirus DNA.**" R.W. Tseng and N.H. Acheson. 1986. Mol. Cell. Biol. 6:1624-1632.

N.H. Acheson was responsible for the design of the formula to measure the termination efficiency and assisted in the preparation of the manuscript.

R.W. Tseng was responsible for all the experiments described in this paper.

This manuscript has been incorporated as Section II of this thesis. For the integrity of the thesis, the abstract has been omitted and materials and methods have been incorporated into the common materials and methods with a more detailed description. The method described in Section I was used to measure the efficiency of transcription termination, and Section II is the paper that describes this work.

SECTION II

Use of a Novel S1 Nuclease RNA Mapping Technique to Measure Efficiency of Transcription Termination on Polyomavirus DNA

INTRODUCTION

RNA polymerase II does not efficiently terminate transcription on the L strand of polyomavirus DNA during the late phase of productive infection (Acheson, 1978). As a result, some RNA polymerase molecules traverse the circular, 5.3 kb, double-stranded polyomavirus genome several times before dissociating and releasing their RNA chains. RNA molecules up to four times genome length have been detected in the nuclei of productively-infected cells (Acheson et al, 1971; Acheson, 1984). These multigenome-length RNAs undergo splicing within the nucleus to form 19S, 18S and 16S viral messenger RNAs (Kamen et al, 1980; Treisman, 1980). During this process, 57-nucleotide long "leaders" located 5.3 kb apart in multigenome-length transcripts are spliced together, producing mRNAs with multiple, tandem copies of the leader near their 5' ends (Legon et al, 1979; Acheson, 1981a). It has been postulated that the presence of multiple leaders may increase the ability of polyomavirus L-strand mRNAs to be translated (Treisman, 1980), thereby conferring a selective advantage to virus replication. In this light, the low termination efficiency which leads to production of multigenome-length mRNA precursors can be seen as a feature designed ultimately to enhance viral replication, although paradoxically it leads to removal by splicing and degradation in the nucleus of most of the viral RNA synthesized (Acheson, 1981a, 1984).

In contrast to the late phase, transcription appears to terminate efficiently during the early phase of polyomavirus infection. Transcription is restricted primarily to the E DNA strand during the early phase (Beard et al, 1976). No E-strand RNAs greater than genome size can be detected, and little transcription of regions of the E strand downstream of the polyadenylation site takes place (Acheson and Mieville, 1978). There is no E-strand region equivalent to the late leader sequence, and E-strand transcripts are efficiently exported from nucleus to cytoplasm (Acheson, 1981b). Thus none of the phenomena associated with multigenome-length transcripts are present during the early phase.

These observations suggest that transcription termination is regulated during the polyomavirus growth cycle, perhaps by a viral gene product. As a prelude to studies on the control of transcription termination in this system, we have devised an approach for measuring the efficiency of termination on the L DNA strand by RNA polymerase II. Since polyomavirus DNA is circular, RNA polymerase molecules which do not terminate during transcription of the L DNA strand will traverse the entire genome and will pass through the initiation region again. RNAs synthesized by polymerases which have traversed the genome at least once can be distinguished from RNAs synthesized by newly-initiated polymerases. When hybridized with a DNA fragment which spans the initiation sites, the former will give rise to full-length DNA-RNA hybrids; the latter will give rise to hybrids shorter than the DNA fragment used, because their 5' ends lie within the fragment used. Measurement of the relative amounts of these two classes of RNAs allows us to estimate the fraction of RNA polymerases which terminate per genome traverse.

To be able to carry out these experiments, we have modified Berk and Sharp's (1977) S1 nuclease RNA mapping method to allow analysis of *in vivo* pulse-labeled nuclear RNA. This modified method is described here in detail.

RESULTS

Transcription Initiation Region on the L Strand of Polyomavirus DNA.

Figure 14 shows the region where transcription of the polyomavirus L strand is initiated during the late phase of productive infection. Kamen and coworkers (Cowie et al, 1981; Treisman and Kamen, 1981) have shown that transcription initiates at multiple sites over a 100-nucleotide region between nucleotides 5050 and 5150 in the A2 strain of polyomavirus. Heiser and Eckhart (1982) found essentially identical results for the A3 strain, whose DNA sequence does not significantly differ from the A2 strain in this region (Deininger et al, 1979; Soeda et al, 1980; Tyndall et al, 1981). Our strain of polyomavirus, which we call AT3, is closely related to strain A3 but contains a 75-nucleotide direct repeat inserted immediately to the left of nucleotide 5175 (Fig. 14) (D.C. Tessier, W.G. Skarnes and N.H. Acheson, manuscript in preparation). As a result, nucleotides 5101-5175, which include the major initiation site at nucleotide 5128, are repeated and, as shown below, this upstream site (nucleotide 5203) is also used for transcription initiation but to a minor degree. The predominant transcription initiation sites in our strain are numbered 1 to 7 in Figure 14; these sites are nearly equally spaced about 25 nucleotides apart. Figure 14 also indicates the positions of the HpaII and PstI sites which define the ends of the DNA fragments cloned into single-stranded bacteriophage vectors for the purposes of this work.

Strategy of the Experiment.

Figure 15 outlines the strategy we used to distinguish RNAs made by recently-initiated RNA polymerases from RNAs made by polymerases which have completed at least one complete traverse of the polyomavirus genome before

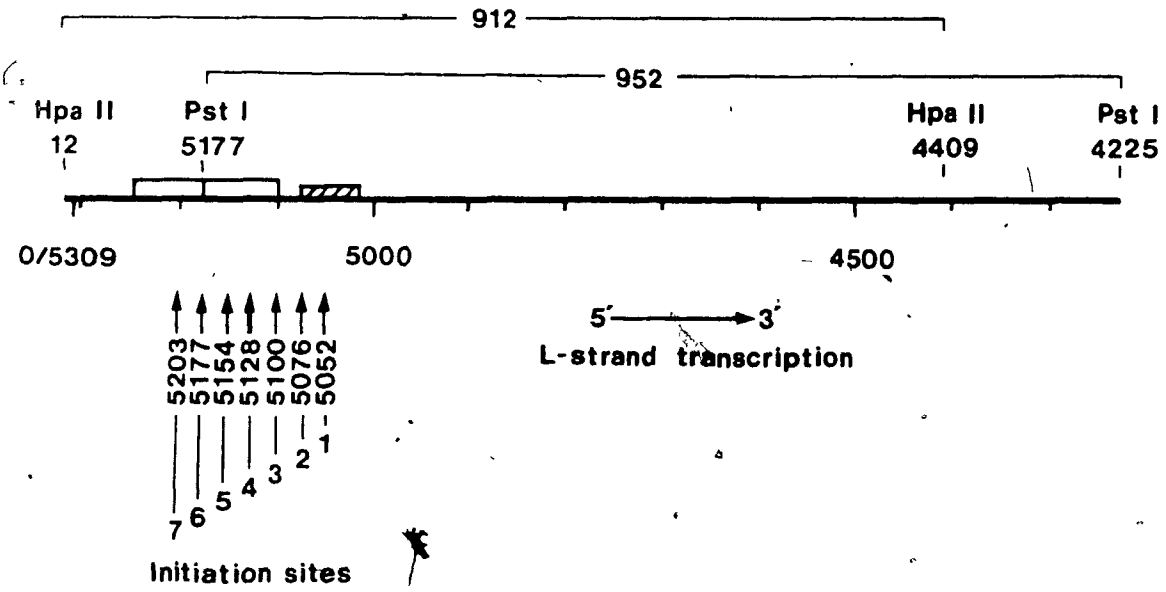
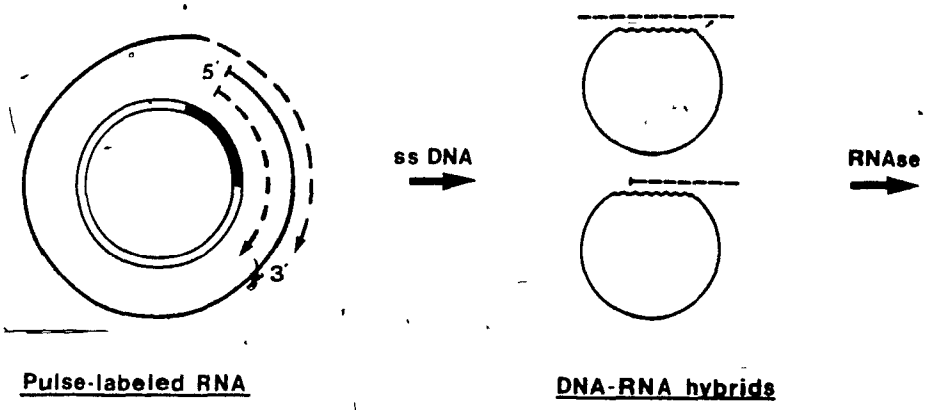

912
952
Hpa II
12
Pst I
5177
Hpa II
4409
Pst I
4225
0/5309
5000
4500
5203
5177
5154
5128
5100
5076
5052
7
6
5
4
3
2
1
Initiation sites
5'
3'
L-strand transcription

Figure 14: Region of polyomavirus DNA which contains initiation sites for L-strand transcription.

Polyomavirus strain AT3 differs from strain A3 (Deininger et al, 1979) principally by the presence of a 75-nucleotide direct repeat (nucleotides 5176-5250) and a 52-nucleotide deletion (after nucleotide 5288) between the DNA replication origin and the L-strand transcription start sites (D.C. Tessier, W.C. Skarnes and N.H. Acheson, manuscript in preparation). Nucleotide numbers are by convention the same as for strain A2 (Soeda et al, 1980; Griffin et al, 1981; Tyndall et al, 1981) between nucleotides 55 and 5175, but differ between nucleotides 5176-5309 and 1-54 because of duplications and deletions. The 75-nucleotide duplication is shown as two adjacent rectangular boxes. The 57-nucleotide late leader region, nucleotides 5020-5076 (Legon et al, 1979), is shown as a hatched rectangle. Vertical arrows below the line denote multiple start sites for L-strand transcription, which progresses in a rightward direction. Nucleotide numbers of initiation sites 1 through 7 are shown; these are estimated by comparison of our data (see Results) with those of Treisman (1980), Cowie et al (1981) and Heiser and Eckhart (1982). HpaII and PstI sites are shown above the line, along with the lengths of HpaII fragment 3 and PstI fragment 2.


5'
3'
ss DNA
RNAse
Pulse-labeled RNA
DNA-RNA hybrids

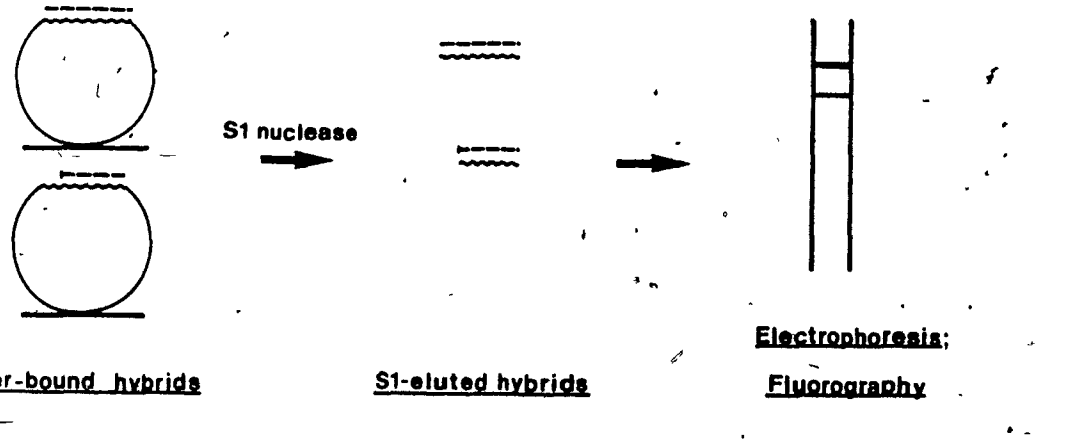

S1 nuclease
Filter-bound hybrids
S1-eluted hybrids
Electrophoresis;
Fluorography

Figure 15: Strategy of experiment to determine efficiency of transcription termination.

The double circle in the upper left panel depicts polyomavirus DNA; the filled arc represents a restriction fragment within which lie the initiation sites for L-strand transcription. The spiral lines outside the circle represent transcripts which have initiated at a site within this restriction fragment. The inner line represents RNA synthesized by recently-initiated RNA polymerase, which has passed through the fragment only once; the outer line represents RNA synthesized by RNA polymerase which has made a complete circuit of the DNA and has passed through the fragment a second time. Only the 3'-terminal 2 kb of these RNAs are labeled (dashed lined), due to very short exposure of infected cells to ^{3}H-uridine. These RNAs are hybridized with the complementary DNA strand of the restriction fragment cloned into a single-stranded bacteriophage vector (second panel). The labeled portion of RNAs made by recently initiated RNA polymerases will hybridize only with that portion of the DNA downstream of the initiation site; the labeled portion of RNAs made by RNA polymerases which have completed at least one circuit of the DNA will hybridize with the entire restriction fragment. For further description, see Results.

passing through the initiation region again. RNAs were labeled 28 hr after infection by incubating glucosamine-treated cells with high concentrations of [5-^{3}H]uridine for 1 or 2 min at 37°C (see Materials and Methods). During such a short labeling period, UTP pools are rapidly increasing in specific radioactivity; therefore, most of the incorporation of ^{3}H-uridine into RNA chains takes place in the last 20-30 seconds of the labeling period (Derman et al, 1976). RNA polymerases have been estimated to travel at a speed of about 5000 nucleotides per min (Greenberg and Penman, 1966). Thus the average length of RNA labeled at growing 3' ends should be about 2000 nucleotides, significantly less than a full genome traverse (5.3 kb). Derman et al (1976) showed that, in HeLa cells, the effective average length of the labeled portions of growing RNA chains did not change if the labeling period varied between 10 and 45 seconds. We have shown that the average length of the radioactive portions of polyomavirus RNAs labeled under these conditions is 2 kb (N.H. Acheson, unpublished results). Thus, every RNA chain which is growing through the initiation region at the time cells are harvested should be labeled only near its 3' end and, in particular, chains which initiated one or more genome traverses earlier will not contain radioactivity in their 5'-terminal portions. This is depicted in the upper left-hand panel in Figure 15. The short labeling time used also ensures that nascent RNAs which have undergone minimal processing are analyzed; this is covered more fully in the Discussion.

^{3}H-labeled nuclear RNA was hybridized in solution with circular, single-stranded bacteriophage DNA containing an insert of the L strand of the appropriate polyomavirus restriction fragment (2nd panel, Figure 15). Treatment of these DNA-RNA hybrids with ribonuclease A and S1 nuclease should degrade all unhybridized RNA and single-stranded RNA and DNA tails, allowing analysis of the remaining duplex structures by gel electrophoresis

as in the S1 nuclease mapping emthod of Berk and Sharp (1977). Although this method gave good results when labeled cytoplasmic RNA was hybridized, use of labeled nuclear RNA led to very high levels of ribonuclease-resistant background radioactivity throughout the area of interest in gel lanes (Fig. 8, lane 2). We found that this background was due to double-stranded RNA formed during reannealing of nuclear RNA. Denaturation of labeled nuclear RNA followed immediately by ribonuclease treatment reduced the background significantly; reannealing of denqatured nuclear RNA by itself brought back the background radioactivity. Treatment of hybridized nuclear RNA with double-strand specific ribonuclease III (Robertson et al, 1968) removed much of the background radioactivity without affecting DNA-RNA hybrids. Treatment with ribonuclease H (Stein and Hausen, 1969) removed DNA-RNA hybrid bands but did not affect the background (Fig. 6).

To get around this problem, we took advantage of the fact that the DNA-RNA hybrids formed in our hybridization reaction have long single-stranded DNA tails consisting of bacteriophage fd103 or M13 DNA. This allowed us to bind hybrids (as well as unhybridized DNA) to nitrocellulose filters in high salt after ribonuclease treatment (3rd panel, Figure 15). Degraded RNA and double-stranded RNA do not bind to nitrocellulose under the conditions we use. After washing, the filters are incubated with S1 nuclease, which releases the duplex DNA-RNA hybrids into solution and trims off any single-stranded DNA tails (4th panel, Figure 15). Figure 8 (lane 3) shows that, using this method, DNA-RNA hybrid bands were recovered quantitatively and most background radioactivity was removed. It should be noted that the DNA-RNA hybrids are not irreversibly bound to the nitrocellulose filter; washing filters with 1 mM EDTA before S1 nuclease treatment removes most of the hybrids (Fig. 8, lane 4).

By using this modified S1 nuclease mapping technique we were able to map 5' ends of newly-synthesized pulse-labeled RNAs. We could also quantitate levels of transcription through the initiation region due either to recently-initiated RNA polymerases, or to RNA polymerases which have made at least one complete circuit of the viral genome before passing through the initiation region again. The former gave rise to DNA-RNA hybrids shorter than the length of the restriction fragment used; the latter gave rise to full-length DNA-RNA hybrids (lower right-hand panel, Figure 15).

Analysis of Nuclear RNA: Mapping Sites of Transcription Initiation.

Figure 16 shows the pattern of DNA-RNA hybrid bands generated by hybridization of 2 min labeled nuclear RNA with either PstI fragment 2L (nucleotides 4225 to 5177, lane 2) or HpaII fragment 3L (nucleotides 4409 to 12, lane 4) (see Figure 14 for map positions). In both cases the most intense band was at the position of full-length DNA-RNA hybrids (952 nucleotides for PstI-2L; 912 nucleotides for HpaII-3L). Faster-migrating less intense bands showed a characteristic pattern of a major band followed by three closely-spaced minor bands. Size estimates derived from comparison with double-stranded DNA markers (Fig. 16) and DNA-RNA hybrids of known size (not shown) established that these four bands correspond to RNAs which have initiated at sites 1 through 4 (Fig. 14), with the major band being at site 4, nucleotide 5128. Two additional minor bands, corresponding to sites 5 and 7, can be seen in lane 4; they are not present in lane 2 as these initiation sites are upstream of the PstI fragment used.

Comparison of the patterns in lanes 2 and 4 confirms that bands 1-7 correspond to RNAs whose 5' ends lie within HpaII fragment 3. If the bands in lane 4 were generated by RNAs whose 3' ends lie within HpaII

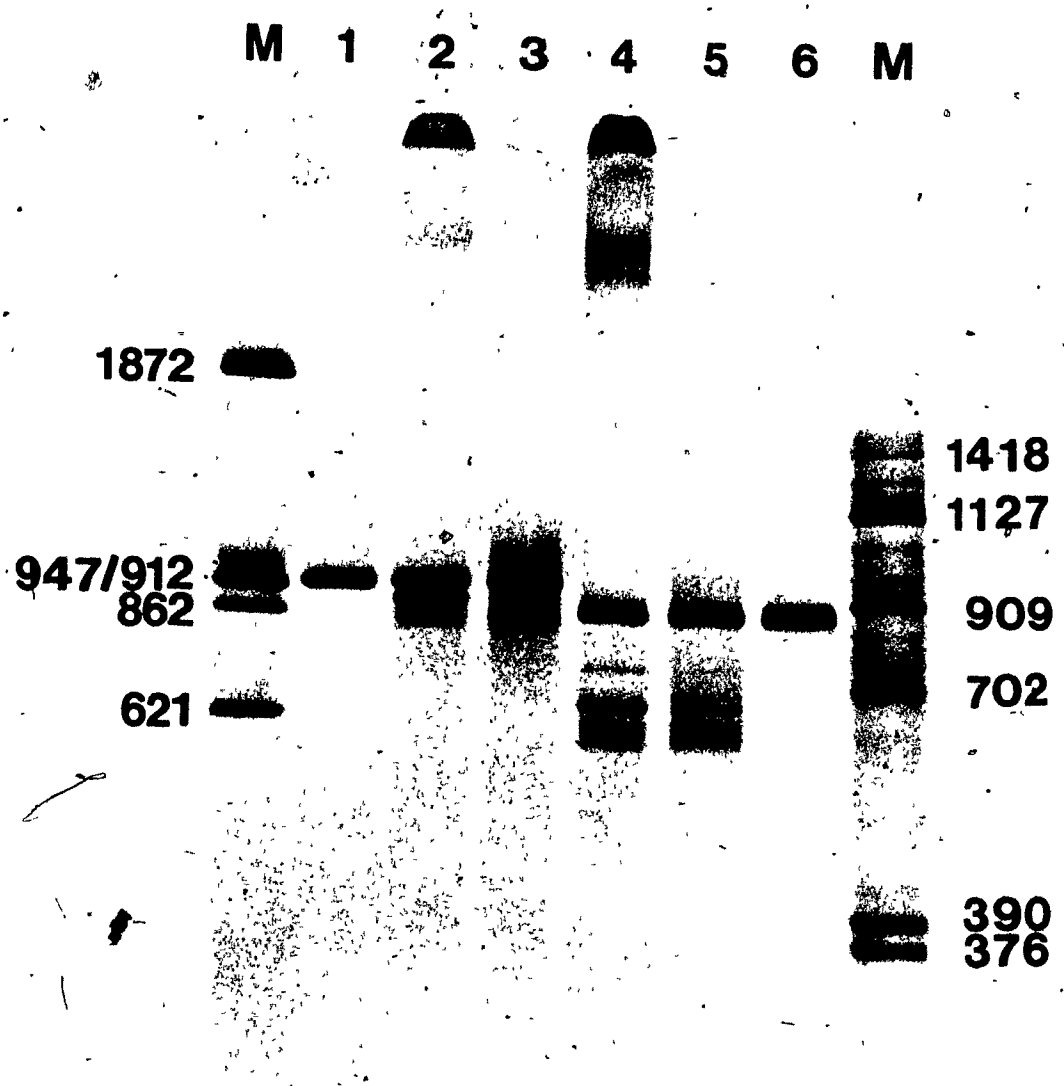

M
1
2
3
4
5
6
M
1872
947/912
862
621
1418
1127
909
702
390
376

Figure 16: Hybridization of ^{3}H-RNA synthesized *in vitro* or *in vivo* to polyomavirus DNA fragments PstI-2L or HpaII-3L.

^{3}H-uridine labeled RNAs described below were hybridized to 10 µg single-stranded fd103 DNA containing cloned polyomavirus PstI fragment 2L (lanes 1-3) or M13 mp8 DNA containing cloned polyomavirus HpaII fragment 3L (lanes 4-6) (see text and Fig. 14). Lanes 1 and 6: ^{3}H-labeled RNA (40×10^3 cpm) transcribed *in vitro* by SP6 RNA polymerase from a 5.25 kb XbaI fragment of polyomavirus DNA inserted into plasmid SP64 (see Materials and Methods); lanes 2 and 4: nuclear RNA (50 µg; 0.4×10^6 cpm) labeled for 2 minutes *in vivo* at 28 hours post infection; lanes 3 and 5: nuclear RNA (21 µg; 1×10^6 cpm) labeled for 60 minutes *in vivo* from 27 to 28 hours post infection. Hybrids were treated with 500 units/ml RNAse T1, bound to nitrocellulose filters, and eluted by incubation with S1 nuclease as described in Materials and Methods. Marker DNAs were as in Figure 8. Electrophoresis was at 50 V (20 mA) for 12 hours.

fragment 3, they would become smaller, not larger, when these RNAs were hybridized to PstI fragment 2 (Fig. 14). As a control, genome-length polyomavirus L-strand RNA was synthesized in vitro by using SP6 RNA polymerase (see Materials and Methods). When such ^{3}H-labeled RNA was hybridized with either PstI-2L or HpaII-3L single-stranded DNA, and hybrids were processed as outlined above, no bands shorter than full-length DNA-RNA hybrids were formed (Fig. 16, lanes 1 and 6). This important control confirms that bands 1-7 correspond to authentic 5' ends of labeled RNAs and are not artifacts caused by RNAse or S1 nuclease cleavage of longer DNA-RNA hybrids.

Lanes 3 and 5 of Figure 16 show the pattern of hybrids generated when 60 min labeled nuclear RNA was used. Although the overall patterns of less-than-full-length bands are similar to those seen with 2 min labeled RNA, the most prominent band now corresponds to site 2 (nucleotide 5076) rather than site 4 (nucleotide 5128). This result can be explained by leader-to-leader splicing. The 5' end (nucleotide 5076) of an internal leader can be spliced to the 3' end (nucleotide 5020) of the next upstream leader in multigenome-length RNAs (Treisman, 1980). Any RNAs in which this leader-to-leader splice has occurred will form a continuous hybrid with HpaII-3L only up to nucleotide 5076, the position of the splice junction (see below, Fig. 18B). This experiment shows that 60 min labeled nuclear RNA contains much unspliced RNA (full-length DNA-RNA hybrids) as well as some spliced RNA (band 2). Further separation of 60-min labeled nuclear RNA into size classes and into poly(A$^+$) and poly(A$^-$) fractions would be required to fully explain the distribution of DNA-RNA hybrid sizes seen in these lanes.

A higher-resolution pattern generated by hybridization of pulse-labeled RNA to DNA fragment PstI-2L is shown in Figure 17. This gel also

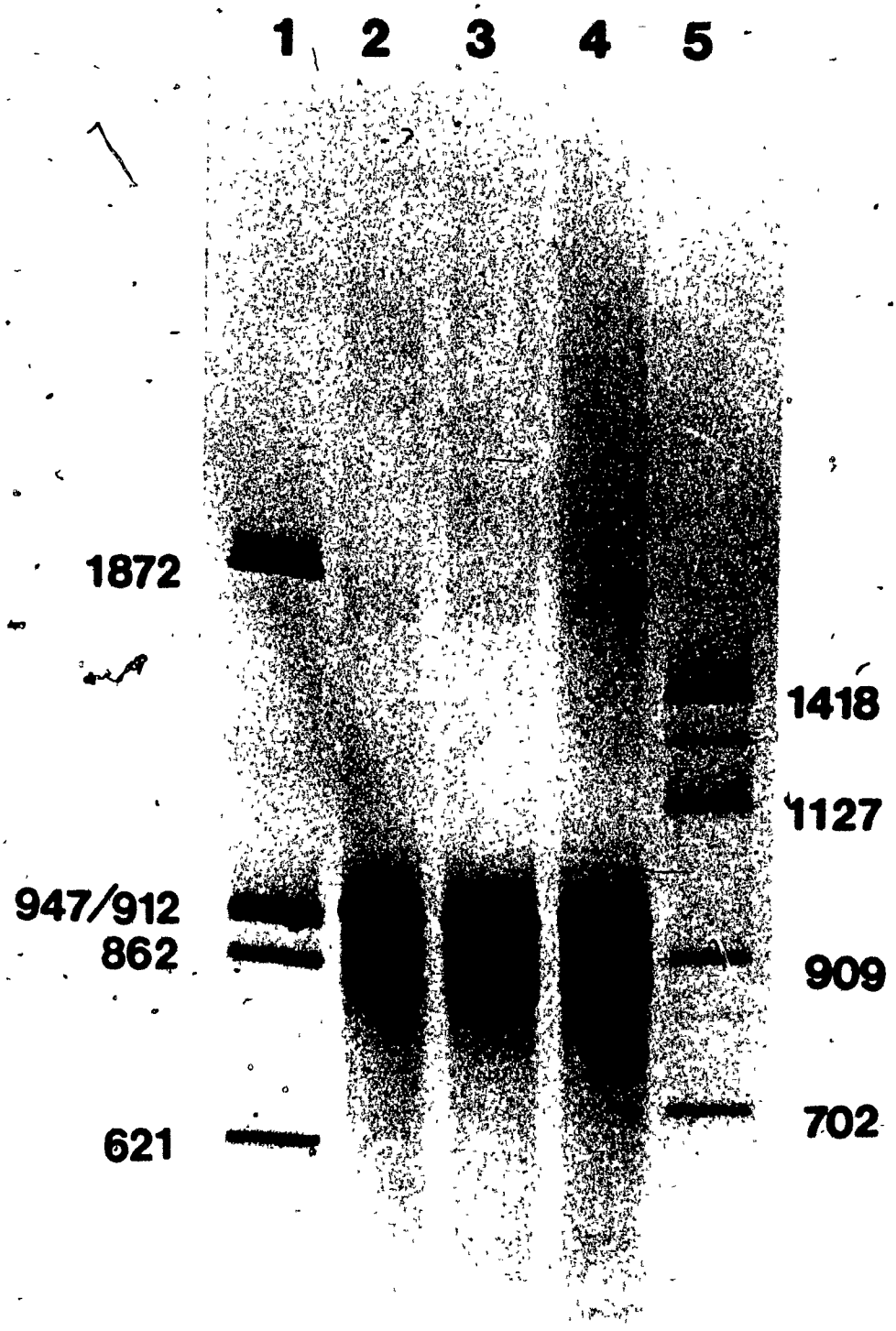

1
2
3
4
5
1872
947/912
862
621
1418
1127
909
702

Figure 17: Hybridization of ^{3}H-labeled nuclear RNA to polyomavirus DNA fragment PstI-2L.

Each of the following ^{3}H-uridine labeled RNA samples was hybridized with single-stranded bacteriophage M13 mp8 DNA containing cloned polyomavirus DNA fragment PstI-2L. Lane 2: nuclear RNA (160 µg, 0.73 x 10^{6} cpm) labeled for 1 minute at 28 hours post infection, hybridized with 30 µg DNA; lane 3: nuclear RNA (80 µg, 1.1 x 10^{6} cpm) labeled for 2 minutes at 28 hours post infection, hybridized with 20 ug DNA; lane 4: nuclear RNA (31 µg, 4.2 x 10^{6} cpm) labeled for 60 minutes from 27 to 28 hours post infection, hybridized with 10 µg DNA. Hybrids were treated as described in Materials and Methods. Marker DNAs were as in Figure 8. Electrophoresis was at 50 V (26 mA) for 16 hours.

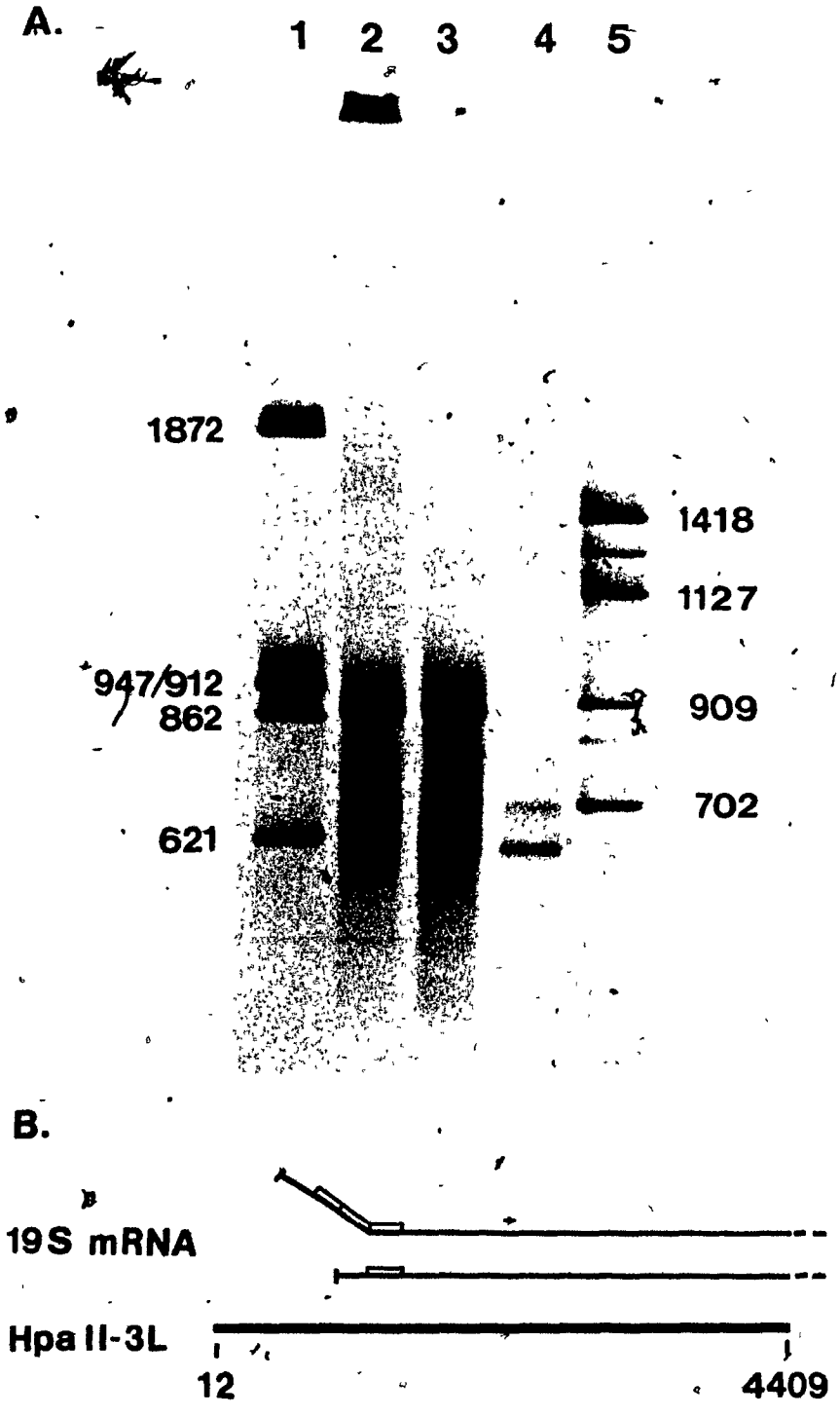

A.
1 2 3 4 5
1872
947/912
862
621
1418
1127
909
702
B.
19S mRNA
Hpa II-3L
12
4409

Figure 18: Hybridization of ^{3}H-labeled nuclear or cytoplasmic RNA to polyomavirus DNA fragment HpaII-3L.

A. Each of the following ^{3}H-uridine labeled RNA samples was hybridized with 10.5 µg single-stranded bacteriophage M13 mp8 DNA containing cloned polyomavirus DNA fragment HpaII-3L. Lane 2: nuclear RNA (57 µg, 0.33 x 10^6 cpm) labeled for 2 minutes at 28 hours post infection; lane 3: nuclear RNA (50 µg, 1.5 x 10^6 cpm) labeled for 60 minutes from 28 to 29 hours post infection; lane 4: poly(A$^+$) cytoplasmic RNA (0.79 x 10^6 cpm) labeled for 120 minutes from 27 to 29 hours post infection. Hybrids were treated as described in Materials and Methods. Marker DNAs were as in Figure 8.

B. Diagram showing hybridization between 19S L-strand mRNA and HpaII-3L DNA. Only unspliced RNA with a single leader sequence forms continuous hybrids up to the 5' end of the RNA; 5' ends of RNAs with multiple leaders will remain unpaired and will be digested by RNase treatment.

shows that the pattern generated with RNA labeled for 1 min (lane 2) or 2 min (lane 3) is identical, justifying the use of RNA labeled for 2 min, which generally has a specific radioactivity 3-5 times that of RNA labeled for 1 min.

Analysis of Cytoplasmic mRNA.

Figure 18A (lane 4) shows that hybridization of cytoplasmic polyadenylated mRNA with HpaII-3L DNA generated no full-length DNA-RNA hybrids. This was expected, since multigenome-length polyomavirus RNAs are confined to the nucleus (Acheson, 1984); all RNA sequences between the 5' end of a leader and the 3' end of the next upstream leader are removed by splicing before transport of mRNAs to the cytoplasm (Treisman, 1980).

Of the three classes of polyomavirus L-strand mRNAs (Kamen et al, 1980), only unspliced 19S mRNA will generate DNA-RNA hybrids which extend from the HpaII site at nucleotide 4409 up to the initiation region (Fig. 18B). 18S and 16S mRNAs are spliced between the distal end of the leader (nucleotide 5020) and the mRNA body (nucleotide 4707 or 4122) and therefore will not form continuous DNA-RNA hybrids with HpaII-3L. Thus, all bands seen in lane 4 must arise from 19S mRNA. Those mRNAs with a single, unspliced leader generate hybrid bands corresponding to mRNA 5' ends, some of which can be seen as minor bands in lane 4 of Figure 18A. 19S mRNAs with multiple, spliced leaders generate a single band corresponding to the 5' end of the leader sequence at nucleotide 5076 (major band in lane 4), because none of the leaders upstream of the most distal one will be able to hybridize to the single-stranded DNA clone, which contains only one copy of the leader sequence (Fig. 18B). Some 19S mRNAs with a single leader may also have 5' ends at this site (site 2 in Figure 14), as it is used

frequently as an initiation site (Figs. 16-18). Therefore, an undetermined fraction of the radioactivity in this band may arise from 19S mRNAs containing a single leader.

Lanes 2 and 3 of Figure 18 show the hybrid pattern formed by nuclear RNA labeled for 2 or 60 min at higher resolution than seen in Figure 16. In lane 2, bands corresponding to 5' ends at sites 1-5 and 7 can be distinguished, and site 4 is seen to be the most frequently used initiation site. In lane 3, all 7 bands can be seen (band 6 is faint), and the band corresponding to site 2 is predominant for reasons explained above.

Determination of Transcription Termination Frequency.

Gel lanes similar to lane 2 of Figure 18 were sliced and the radioactivity in each slice was measured. Figure 19 shows the profile of radioactivity from such a gel. As is already evident from the preceeding autoradiograms, the full-length band, which arises from transcription by RNA polymerases which have already made at least a complete genome traverse, contains more radioactivity than any of the shorter bands.

Seven independent nuclear RNA samples labeled for 1 or 2 min were hybridized with HpaII-3L DNA and DNA-RNA hybrids were analyzed by electrophoresis followed by slicing. Results are shown in Table I. To be able to compare molar amounts of hybrids of different lengths, results were normalized by dividing radioactivity associated with each band by the number of U residues calculated to be present in each RNA, then multiplying the result by 100. The sums of normalized values for all bands shorter than full-length DNA-RNA hybrids (column A) correspond to the relative number of RNA polymerase molecules which initiated and traveled at least as far as the end of HpaII-3L during the labeling period. The

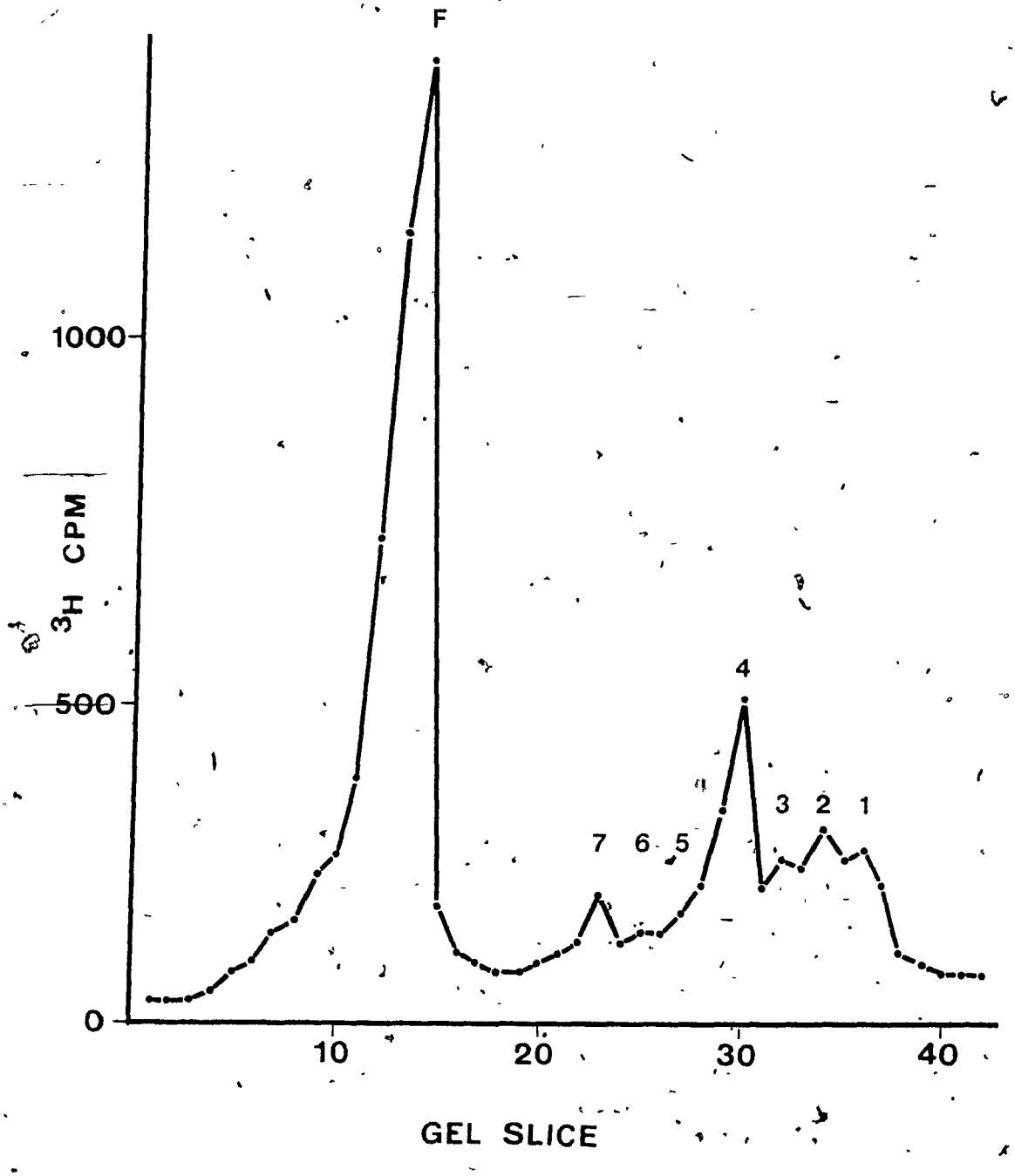

F
1000
500
0
^{3}H CPM
4
3
2
1
7
6
5
10
20
30
40
GEL SLICE

Figure 19: Radioactivity profile from sliced gel: nuclear RNA hybridized with DNA fragment HpaII-3L.

Nuclear RNA (101 µg, 600 x 10^3 cpm) labeled for 2 minutes at 28 hours post infection was hybridized with 20 µg single-stranded M13 mp8 DNA containing cloned polyomavirus HpaII fragment 3L. Hybrids were treated as described in Materials and Methods, except that 400 units/ml RNase T1 were used. After electrophoresis, the gel lane was sliced and slices were counted as described in Materials and Methods. Only that portion of the gel containing DNA-RNA hybrid bands is shown. F corresponds to full-length hybrids, and band numbers 1 to 7 correspond to initiation sites 1 to 7 shown in Figure 14.

TABLE I

Efficiency of termination per genome traverse by RNA polymerase II at 28 hr post infection by polyomavirus.[a]

Experiment	cpm per 100 U residues in DNA/RNA hybrids[b] A. Less-than-full-length	B. Full-length	Column A / (Column A + Column B) = Termination efficiency
1	1050	1180	0.47
2	1040	990	0.51
3	1750	1450	0.55
4	1220	1440	0.46
5	540	730	0.43
6	320	250	0.56
7	570	460	0.55

[a] Nuclear RNA labeled with ^{3}H-uridine for 1 minute (experiment 2) or 2 minutes (experiments 1, 3-7) at 28 hours post infection was hybridized with HpaII-3L. Hybrids, treated with RNAse and S1 nuclease, were subjected to electrophoresis and gel slices were counted as described in Experimental Procedures.

[b] cpm in full-length hybrid band (912 nucleotides) or in hybrid bands corresponding to RNA 5' ends at sites 1 through 7 (647-796 nucleotides) were divided by the number of U residues present in each RNA (determined from the known nucleotide sequence of polyomavirus DNA), then multiplied by 100. Normalized cpm in all less-than-full-length bands were summed.

normalized values for full-length DNA-RNA hybrids (column B) correspond to the relative number of RNA polymerase molecules which completed at least one genome traverse before passing through HpaII-3L again.

We have used these data to estimate the frequency of transcription termination per genome traverse of RNA polymerase by use of the model shown in Figure 20. To simplify the discussion, polyomavirus DNA is shown as an infinitely long linear tandem repeat, each repeat unit consisting of an entire 5.3 kb genome. This is formally equivalent to a circle from the point of view of a RNA polymerase molecule which is traversing the DNA unidirectionally, and is in fact the form in which multigenome transcripts are produced (Acheson, 1978). The left-hand end of this linear DNA molecule is defined as the site where transcription initiates, which we shall consider unique for the sake of argument.

RNA polymerase molecules initiate transcription and travel rightward, terminating with a fractional efficiency T ($0<T<1$) somewhere upstream of the next DNA repeat. For simplicity, Figure 20 shows there to be a single termination site, but termination could (and probably does) take place at multiple sites (R.W. Tseng and N.H. Acheson, manuscript in preparation); the number and location of termination sites is irrelevant to the present discussion. Downstream of the termination site, there will be only (1-T) times as many RNA polymerase molecules per unit length of DNA as began transcription. Thus for every polymerase molecule which transcribes through the initiation region (shown as a rectangle labled "a") in the first repeat, only (1-T) polymerases will transcribe through the initiation region (shown as a rectangle labeled "b_1") at the beginning of the second repeat. For example, if the termination efficiency T were 0.4 (40% of all polymerases terminate per genome traverse), then 1 - 0.4, or 0.6 (that is, 60%) of the polymerases which initiate would travel at least as far as the

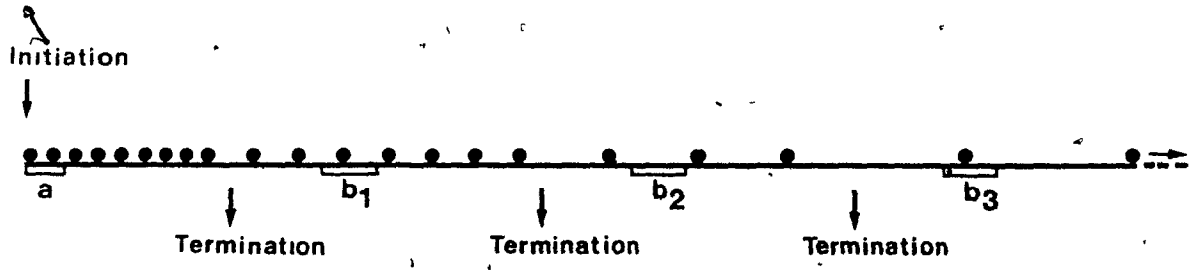

Initiation
a
b1
b2
b3
Termination
Termination
Termination

Figure 20: Model used for estimation of termination frequency per genome traverse by RNA polymerase.

Polyomavirus DNA is depicted as a linear molecule of n genome repeats, whose left-hand end is an initiation site for L-strand transcription. Transcription progresses from left to right on the L DNA strand. "a" designates the region between the initiation site and the end of the DNA fragment used for hybridization with pulse-labeled RNA (HpaII-3L in this paper). A fraction, T, of the RNA polymerases terminates somewhere between the distal end of this fragment and the distal end of the same fragment in the next downstream repeat. Similarly, an identical fraction, T, of the polymerases which have successfully transcribed beyond the end of b_1 terminates before passing the end of b_2 in the next genome repeat, and so on. The model is drawn assuming a termination frequency of 0.5 per genome traverse. See text for derivation of formula to calculate T from hybridization data.

end of the second initiation region. Similarly, (1-T) (1-T) of the polymerases which initiate will transcribe through the initiation region ("b_2") at the beginning of the third repeat unit, and so forth. The sum of all polymerases which will transcribe through the initiation regions in n repeats after having completed at least one genome traverse is therefore $(1-T) + (1-T)^2 + (1-T)^3 + \ldots + (1-T)^n$. This expression can be recognized as a binomial series, which conveniently reduces to the expression, $1/T - 1$. This formula allows us to calculate the termination efficiency, T, from the data in Table I. The molar ratio of full-length transcripts to less-than-full-length transcripts, or column B/column A, is equal to the ratio of polymerases transcribing downstream initiation regions to polymerases transcribing the first initiation region, or $1/T - 1$. A simple rearrangement gives T = column A/(column A + column B). Values for the termination efficiency calculated from this formula are given in the right-hand column in Table I. The average value for seven independent experiments was 0.50. Thus, approximately 50% of RNA polymerase molecules terminate transcription per traverse of the L strand of the polyomavirus genome during the late phase of productive infection.

DISCUSSION

We have devised a strategy to measure in vivo the efficiency of termination of transcription by RNA polymerase on a circular DNA molecule such as polyomavirus. Polymerases that have made at least one complete genome traverse will transcribe the entire length of a restriction fragment which contains the transcription initiation sites; in contrast, polymerases that have just begun transcription will transcribe only that part of the restriction fragment downstream of the initiation site. We can therefore distinguish between RNAs arising from these two classes of polymerase, which allows us to calculate termination frequency per genome traverse.

The interpretation of our data rests on two assumptions. First, we assume that all transcripts which extend through the initiation region to the end of the DNA fragment used have lifetimes significantly longer than the effective labeling time. If this were not the case, the amount of label in DNA-RNA hybrids might not be proportional to the number of polymerases transcribing that region, for some RNAs might be less stable than others. Although we have not rigorously proven this assumption, use of very short labeling times, whose effective length is on the order of 20-30 seconds (see Results), helps to ensure that this condition is met. Previous experiments have shown that a fraction (15-25%) of late polyomavirus RNAs is polyadenylated in the nucleus within 1-2 min of their synthesis, the remainder never being polyadenylated (Acheson, 1984). Little or no polyadenylation can be detected in samples labeled for 1 or 2 min and analyzed as described in this paper (data not shown). Transport of spliced late mRNAs to the cytoplasm begins approximately 15 min after their synthesis (Acheson, 1981b). No splicing can be detected in samples labeled for 1 or 2 min (Figures. 16-18). However, polyadenylated 3' ends and

splice junctions are readily detected by our S1 mapping technique in RNA samples labeled for as little as 10 min (data not shown). Thus, the assumption that we are analyzing unprocessed nascent RNAs in this paper appears justified.

We also assume that the likelihood that a given RNA polymerase molecule will terminate is independent of the number of genome traverses that polymerase has made. This is implicit in the formula developed from the model shown in Figure 20. This assumption seems reasonable, for it is difficult to imagine that an RNA polymerase molecule could "know" how long an RNA chain it carries, or how long a time it has been since it began transcription. If termination did vary as a function of the number of genome traverses, the termination frequency we have calculated could be interpreted as an average rate per genome traverse by RNA polymerase.

In order to carry out these experiments, we have modified Berk and Sharp's (1977) S1 nuclease mapping method so that in vivo-labeled RNAs, rather than unlabeled steady-state RNAs, can be mapped. Hybridization of labeled RNA to recombinant, single-stranded bacteriophage DNAs allows detection of RNAs transcribed from a specific region on a defined DNA strand. Furthermore, DNA-RNA hybrids can be bound to nitrocellulose filters via unhybridized single-stranded bacteriophage DNA tails (Padgett et al, 1983). This step removes double-stranded RNAs which are formed during reannealing of nuclear RNA and which give rise to background upon analysis of hybrids by electrophoresis. The bound DNA-RNA hybrids are subsequently released from the filters, and single-stranded DNA tails are trimmed by incubation with S1 nuclease.

This method should be of general use for detection of RNAs with unstable 5' or 3' ends which are present in pulse-labeled RNA but may be found in vanishingly small amounts in steady-state RNA. In particular,

unprocessed or partially processed primary transcripts can be analyzed. Transcription initiation sites can be mapped and quantitated independently of differential turnover of RNA. 3' ends of RNAs formed by termination of transcription can in principle be detected by this method, whereas analysis of steady-state RNA allows detection primarily of 3' ends of stable RNA species such as fully-processed mRNA. This technique should also enable the detection and mapping of unstable intermediates formed during 3'-end processing or splicing. A drawback is that pulse-labeling in vivo cannot provide RNAs of high specific radioactivity, which limits this technique to study of genes which are transcribed at very high rates, such as those of a number of viruses or highly reiterated cellular genes.

Our results show that about one in every two RNA polymerase molecules which initiate transcription of the L strand of polyomavirus DNA will terminate transcription per genome traverse. This agrees well with the termination frequency (60% per genome traverse) estimated by quantitation of the molar ratios of unspliced, polyadenylated L-strand RNAs which differ in length by integral multiples of the genome length (Acheson, 1984). These RNAs are found exclusively in the nucleus and are presumably precursors to mRNAs which contain multiple spliced leaders. That estimate was based on the assumptions that transcription termination downstream of a polyadenylation site precedes cleavage and polyadenylation at that site, and that the fraction of such RNAs which become polyadenylated is independent of the number of genome traverses made by the polymerase before terminating. Given these assumptions, the number of polyadenylated transcripts in each size class should reflect the number of transcripts which terminate downstream of the polyadenylation site during each genome traverse.

What is responsible for the low efficiency of termination on the polyomavirus L strand? Transcription termination signals encoded in viral DNA may be inherently weak; on the other hand, the machinery which recognizes such signals may be defective during the late phase in polyomavirus-infected cells. We have recently tried to test the former hypothesis by inserting a second polyadenylation signal into polyomavirus DNA (Lanoix et al, 1986). We reasoned that, since both polyadenylation and termination are inefficient in this system, they may be controled by the same signal elements. Neither the overall efficiency of polyadenylation nor the frequency of transcription termination on the L DNA strand (measured as described in this paper) was increased in cells infected with virus containing duplicated polyadenylation signals. We concluded that a viral gene product reduces the efficiency of both processes during the late phase of productive infection. It will be of interest to measure termination efficiency at different times during the late phase to see whether it is progressively reduced.

In the course of these experiments, we have detected seven distinct sites where L-strand transcription initiates in our strain of polyomavirus. Due to the length of DNA-RNA hybrids analyzed, the position of each initiation site is subject to an uncertainty of +/- 5 nucleotides. The positions shown in Figure 14 are based partly on more precise S1 nuclease mapping of the 5' ends of steady-state nuclear or mRNAs (Treisman, 1980; Cowie et al, 1981; Heiser and Eckhart, 1982). The major initiation site we find is near nucleotide 5128 (site 4); initiation is significantly less frequent and approximately equimolar near nucleotides 5050 (site 1), 5076 (site 2), and 5100 (site 3). Site 5 (nucleotide 5150) is used even less frequently. Sites 6 (nucleotide 5077) and 7 (nucleotide 5103) lie within the 75-nucleotide repeated sequence in our virus strain, and probably correspond to initiation within the upstream repeat at nucleotide sequences homologous to those at sites 3 and 4 (Fig. 14).

SECTION III

Changes in Major Initiation Sites and in Termination Efficiency for L-strand Transcription During the Late Phase of Lytic Infection

INTRODUCTION

During the early phase of lytic infection of mouse cells by polyomavirus, RNA is transcribed mostly from one strand (the E strand) of the viral DNA (Kamen et al, 1974). Transcription of the E strand is terminated efficiently during the early phase (Acheson and Mieville, 1978). There is a microheterogeneity in the initiation sites of E strand RNA; the major initiation sites are located between nucleotides 145 and 153, and a number of minor sites are located either downstream (nucleotide 300) or upstream (nucleotides 14 and 20) of the major initiation sites (Kamen et al, 1982). However, during the late phase of infection, major E strand initiation sites are shifted further upstream between nucleotides 5230 and 5260, which are minor initiation sites during the early phase of infection (Kamen et al, 1982; Fenton and Basilico, 1982). This shift could be due to repression of initiation at nucleotides 145-153 because of large T antigen binding to the early promoter in that area. At the same time, E strand transcripts complementary to all E DNA sequences can be detected; this implies that the termination of E DNA transcription is not very efficient during the late phase (Kamen et al, 1974; Beard et al, 1976; Kamen et al, 1982). Upstream start sites are also observed on the E strand of SV40 DNA during the late phase of infection (Ghosh and Lebowitz, 1981; Hansen et al, 1981).

Very little L-strand RNA can be detected during the early phase of the lytic infection. In contrast, most of the viral RNA synthesized during the late phase is copied from the L DNA strand. These L-strand RNAs are up to 4 times genome length and consist of tandem repeats of the entire viral genome (Acheson et al, 1971; Rosenthal et al, 1976; Birg et al, 1977; Acheson, 1978). This is due to inefficient termination of transcription by RNA polymerase (see Part II of this thesis). S1 nuclease mapping of the 5' ends of mRNA or steady state nuclear RNA has revealed that the initiation sites for L-strand RNAs are extremely heterogenous. These 5' ends are scattered over a 100 nucleotide stretch of the promoter region on the L DNA strand. The major initiation site is located at nucleotide 5128 (Treisman, 1980; Cowie et al, 1981; Heiser and Eckhart, 1982).

Using the S1 nuclease mapping technique that we have developed, we have found that at 16 hours after infection (that is, during the early part of the late phase), the major initiation site for L strand transcription is located at nucleotide 5050, instead of nucleotide 5128. Furthermore, we have found that termination of transcription is substantially more efficient at 16 hr after infection than at 28 hr after infection.

RESULTS

Hybridization of Nuclear RNA Labeled at 16 Hours or 18 Hours After Infection to Polyomavirus Fragment HpaII-3L.

Nuclear RNAs of polyomavirus-infected 3T6 cells, labeled with ^{3}H-uridine for 2 minutes at 16 hours or 18 hours after infection were hybridized to single-stranded M13 mp8 DNA containing cloned polyomavirus fragment HpaII-3L. Figure 21 (lane 3) shows the DNA-RNA hybrid bands of nuclear RNA labeled at 18 hours after infection. The pattern is similar to that of nuclear RNA labeled at 28 hours after infection (lane 1). The full-length hybrid is the most intense band. The hybrid band corresponding to initiation site 1 (the fastest migrating band) contains significantly more radioactivity than the bands corresponding to sites 2 and 3, and even slightly more than that corresponding to site 4, the major initiation site at 28 hours post infection. The pattern of DNA-RNA hybrids formed by nuclear RNA labeled for 2 minutes at 16 hours after infection is shown in Figure 22, (lane 2). Only four of the less-than-full-length hybrid bands, corresponding to initiation sites 1 through 4, as well as a full-length hybrid band, can be observed. In the four less-than-full-length hybrid bands, site 1 (at nucleotide 5050, within the leader) is the most intense band. About the same amount of radioactivity is present in hybrid bands corresponding to sites 2 to 4. These results indicate that site 1 is the major initiation site on the L-DNA strand at 16 hours after infection. There is a minor band which migrated faster than the band corresponding to site 1; this band was not usually seen in nuclear RNA labeled at 28 hours after infection, and could represent a distinct initiation site located downstream of the leader region, at approximately nucleotide 5020. As observed on this autoradiogram, sites 5 to 7 are rarely or never used at 16 hours post infection.

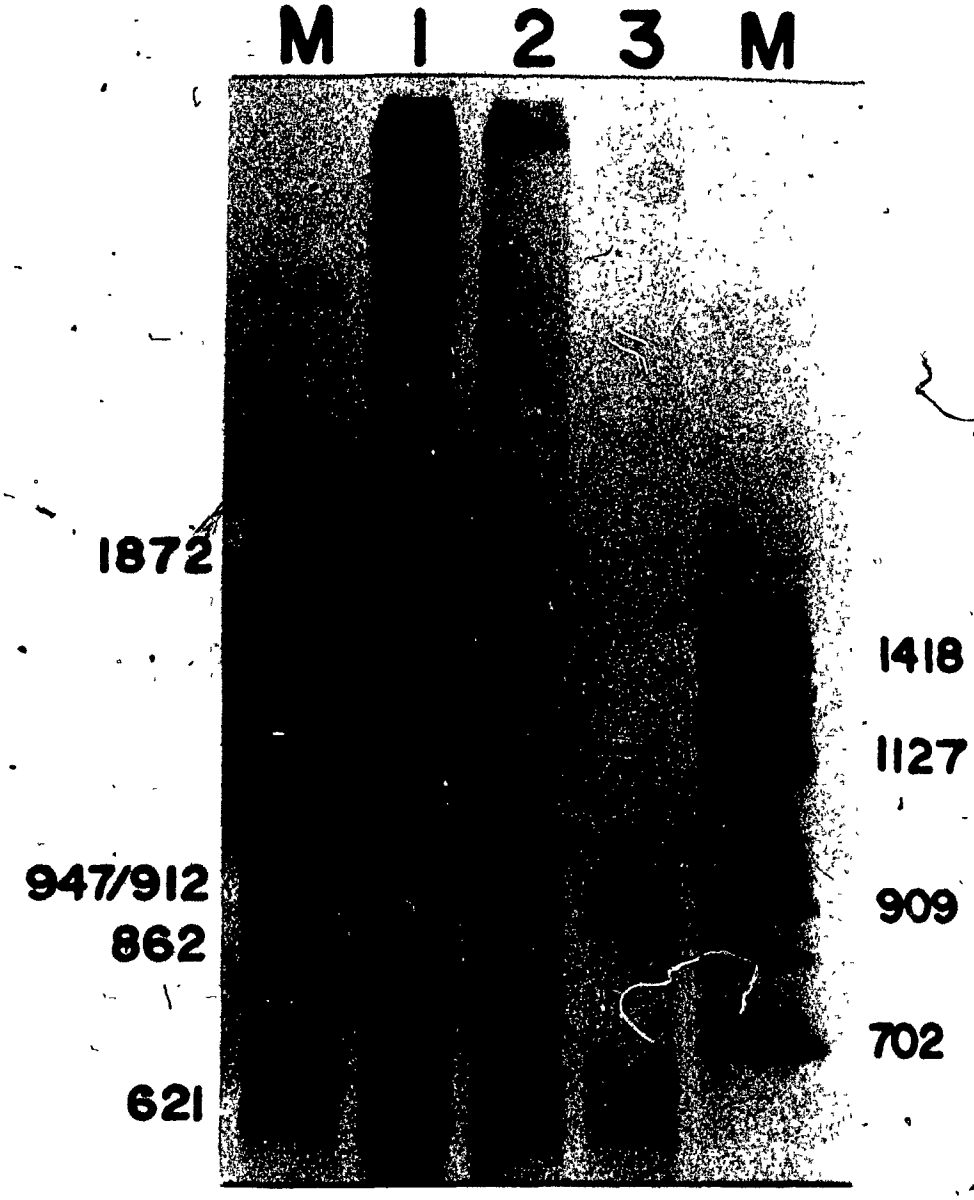

M 1 2 3 M
1872
947/912
862
621
1418
1127
909
702

Figure 21: Hybridization of nuclear RNA labeled with ^{3}H-uridine at 18 hours or 28 hours post infection to polyomavirus DNA fragment HpaII-3L.

Each of the following ^{3}H-uridine labeled RNA samples was hybridized with 10 µg (lanes 1 and 3) or 20 µg (lane 2) single-stranded bacteriophage M13 mp8 DNA containing cloned polyomavirus DNA fragment HpaII-3L. Lane 1: nuclear RNA (50 µg, 0.4 x 10^{6} cpm) labeled for 2 minutes at 28 hours post infection; lane 2: nuclear RNA (100 µg, 0.41 x 10^{6} cpm) (from cells infected with insertion mutant ins 4) labeled for 2 minutes at 28 hours post infection; lane 3: nuclear RNA (100 µg, 0.8 x 10^{6} cpm) labeled for 2 minutes at 18 hours post infection. Hybrids were treated as described in Materials and Methods. Marker DNAs were as in Figure 8. Electrophoresis was at 50V (16 mA) for 15 hours.

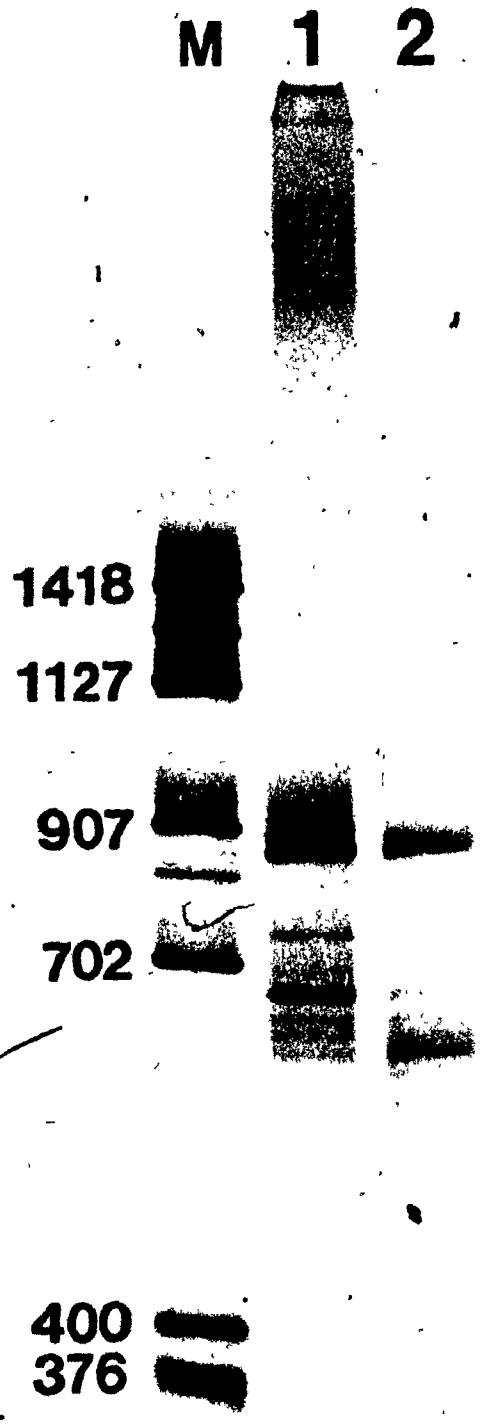

M
1
2
1418
1127
907
702
400
376

Figure 22: Hybridization of nuclear RNA labeled with ^{3}H-uridine at 16 hours or 28 hours post infection to polyomavirus DNA fragment HpaII-3L.

Each of the following ^{3}H-uridine labeled RNA samples was hybridized with 10 µg (lane 1) or 15 µg (lane 2) single-stranded bacteriophage M13 mp8 DNA containing cloned polyomavirus DNA fragment HpaII-3L. Lane 1: nuclear RNA (50 µg, 0.35 x 10^6 cpm) labeled for 2 minutes at 28 hours post infection; lane 2: nuclear RNA (300 µg, 4.25 x 10^6 cpm) labeled for 2 minutes at 16 hours post infection. Hybrids were treated as described in Materials and Methods. Markers were ^{3}H-polyomavirus DNA digested with MspI. Lengths of double-stranded DNA fragments are shown adjacent to each band. Electrophoresis was at 50 V (15 mA) for 15.5 hours.

Hybridization of Pulse-labeled Late Nuclear RNA from Mutant ins 4-infected Cells to Polyomavirus Fragment HpaII-3L.

Ins 4 is a viable mutant of polyomavirus constructed by Jacqueline Lanoix in our laboratory (Lanoix et al, 1986). A polyomavirus DNA fragment from nucleotides 2821 to 2964 was inserted into the HincII cleavage site between nucleotides 2964 and 2965 of wild-type polyomavirus DNA. This fragment contains polyadenylation signals (AATAAA) for both DNA strands and is inserted in the same orientation as in wild-type DNA. Therefore, this mutant contains two tandemly repeated functional polyadenylation signals. Both early RNAs during the late phase and late RNAs were found to be polyadenylated within the insert and at the normally used polyadenylation site (Lanoix et al, 1986).

RNA in ins 4-infected 3T6 cells was labeled with ^{3}H-uridine for 2 minutes at 28 hours after infection. Labeled nuclear RNA was hybridized to polyomavirus fragment HpaII-3L. The results are shown in Figure 21 (lane 2). The pattern of DNA-RNA hybrid bands is very similar to that shown by wild-type virus at 28 hours after infection (lane 1), i.e. initiation site 4 (at nucleotide 5128) is the major initiation site. However, the hybrid band corresponding to initiation site 1 is more intense than those corresponding to sites 2 and 3. This suggests that a greater number of RNA polymerase molecules initiated at site 1 than at sites 2 and 3, in contrast to the pattern seen with wild-type virus. This may be a result of slower growth of ins 4 virus, resulting in persistence of the transcription pattern normally seen at 18 hrs.

Termination Efficiency During the Early Part of the Late Phase.

Gel lanes similar to lane 3 of Figure 21 (18 hr post infection) and lane 2 of Figure 22 (16 hr post infection) were sliced and the

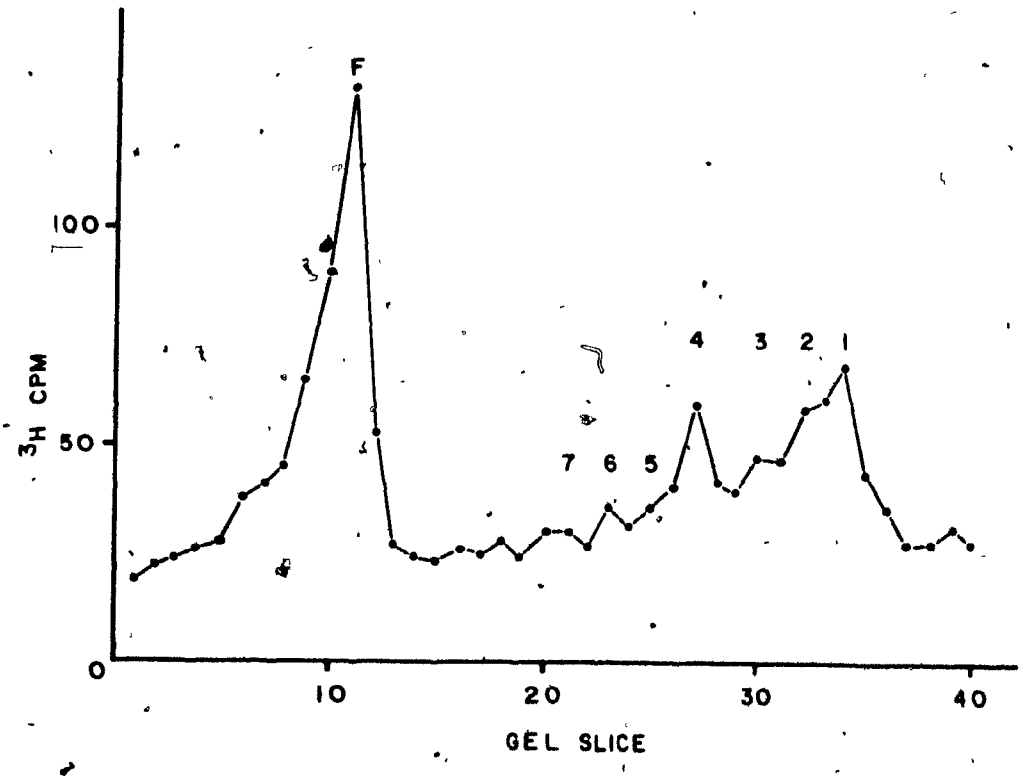

F
100
50
0
^{3}H CPM
4
3
2
1
7
6
5
10
20
30
40
GEL SLICE

Figure 23: Radioactivity profile from sliced gel: nuclear RNA labeled for 2 minutes at 18 hours post infection hybridized with polyomavirus DNA fragment HpaII-3L.

Nuclear RNA (130 µg, 1.1 x 10^6 cpm) labeled for 2 minutes at 18 hours post infection was hybridized with 10 µg of single-stranded M13 mp8 DNA containing cloned polyomavirus fragment HpaII-3L. Hybrids were treated as described in Materials and Methods. After electrophoresis the gel lane was sliced and gel slices were counted as described in Materials and Methods. Only that portion of the gel containing DNA-RNA hybrids bands is shown. F corresponds to full-length hybrids, and band nubmers 1 to 7 correspond to initiation sites 1 to 7 shown in Figure 14.

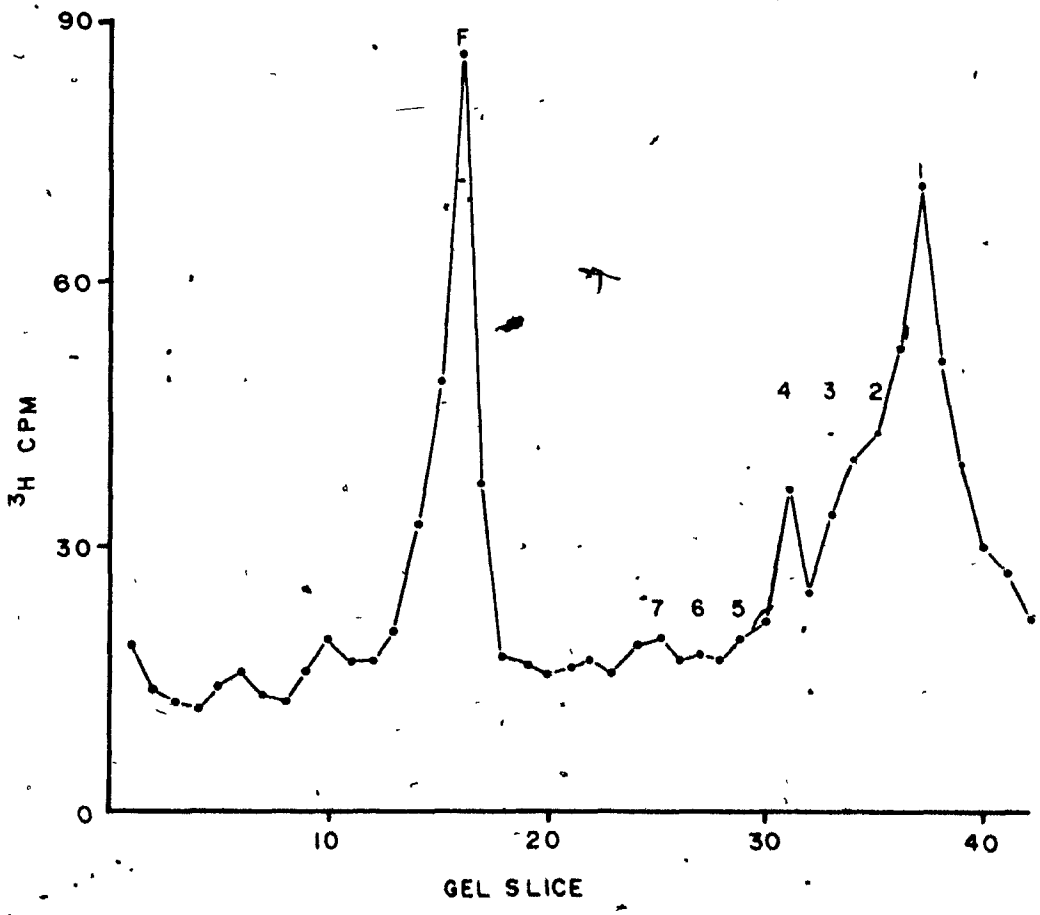

90
60
30
0
^{3}H CPM
F
I
4
3
2
7
6
5
10
20
30
40
GEL SLICE

Figure 24: Radioactivity profile from sliced gel: nuclear RNA labeled for 2 minutes at 16 hours post infection hybridized with polyomavirus DNA fragment HpaII-3L.

Nuclear RNA (300 µg, 4.35 x 10^6 cpm) labeled for 2 minutes at 16 hours post infection was hybridized with 15 µg of single-stranded M13 mp8 DNA containing cloned polyomavirus fragment HpaII-3L. Hybrids were treated as described in Materials and Methods. After electrophoresis the gel lane was sliced and gel slices were counted as described in Materials and Methods. Only that portion of the gel containing DNA-RNA hybrid bands is shwon. F corresponds to full-length hybrids, and band numbers 1 to 7 correspond to initiation sites 1 to 7 shown in Figure 14.

radioactivity of each slice was measured. Figure 23 shows the profile of radioactivity present in a sample isolated 18 hours after infection, and Figure 24 that in a sample isolated 16 hours after infection.

Table II shows the termination efficiency, calculated as explained in Section II for three independent nuclear RNA samples labeled at 16 hours after infection (experiments 1 to 3), and one nuclear RNA sample labeled at 18 hours after infection (experiment 4). At 16 hours after infection, the termination efficiency was calculated to be about 70%, meaning that seven out of ten RNA polymerase molecules transcribing the viral DNA stop transcription before completing a single cycle of the circular genome. At 18 hours after infection, the termination efficiency calculated for a single experiment was 60%. These results suggest that transcription termination is more efficient during the early part of the late phase than it is during the late part of late phase.

Termination Efficiency of Mutant ins 4 Late RNAs.

Figure 25 shows the profile of radioactivity in nuclear RNA labeled with ^{3}H-uridine for 2 minutes at 28 hours after infection of 3T6 cells with ins 4 and hybridized to polyomavirus fragment HpaII-3L. Table III shows the termination efficiency calculated from three independent experiments carried out with ins 4. The termination efficiency per genome traverse at 28 hours is about 48%, similar to that of wild-type polyomavirus at 28 hours, suggesting that duplication of the L-strand polyadenylation signal does not increase the efficiency of termination of L-strand transcription. Other experiments also showed that the overall efficiency of polyadenylation of L-strand RNA is not increased in cells infected with mutant ins 4 (Lanoix et al, 1986).

TABLE II

Efficiency of termination per genome traverse by RNA polymerase II at 16 to 18 hr post infection by polyomavirus.[a]

Experiment	cpm per 100 U residues in DNA-RNA hybrids[b] A. Less-than-full-length	B. Full-length	$\frac{\text{Column A}}{\text{Column A + Column B}}$ = Termination efficiency
1	220	100	0.69
2	180	70	0.72
3	430	150	0.74
4	220	140	0.61

[a] Nuclear RNA labeled with ^{3}H-uridine for 2 minutes at 16 hours (experiments 1-3) or 18 hours (experiment 4) post infection was hybridized with polyomavirus DNA fragment HpaII-3L. Hybrids, treated with RNase T1 and S1 nuclease, were subjected to electrophoresis and gel slices were counted as described in Materials and Methods.

[b] As described in Table I.

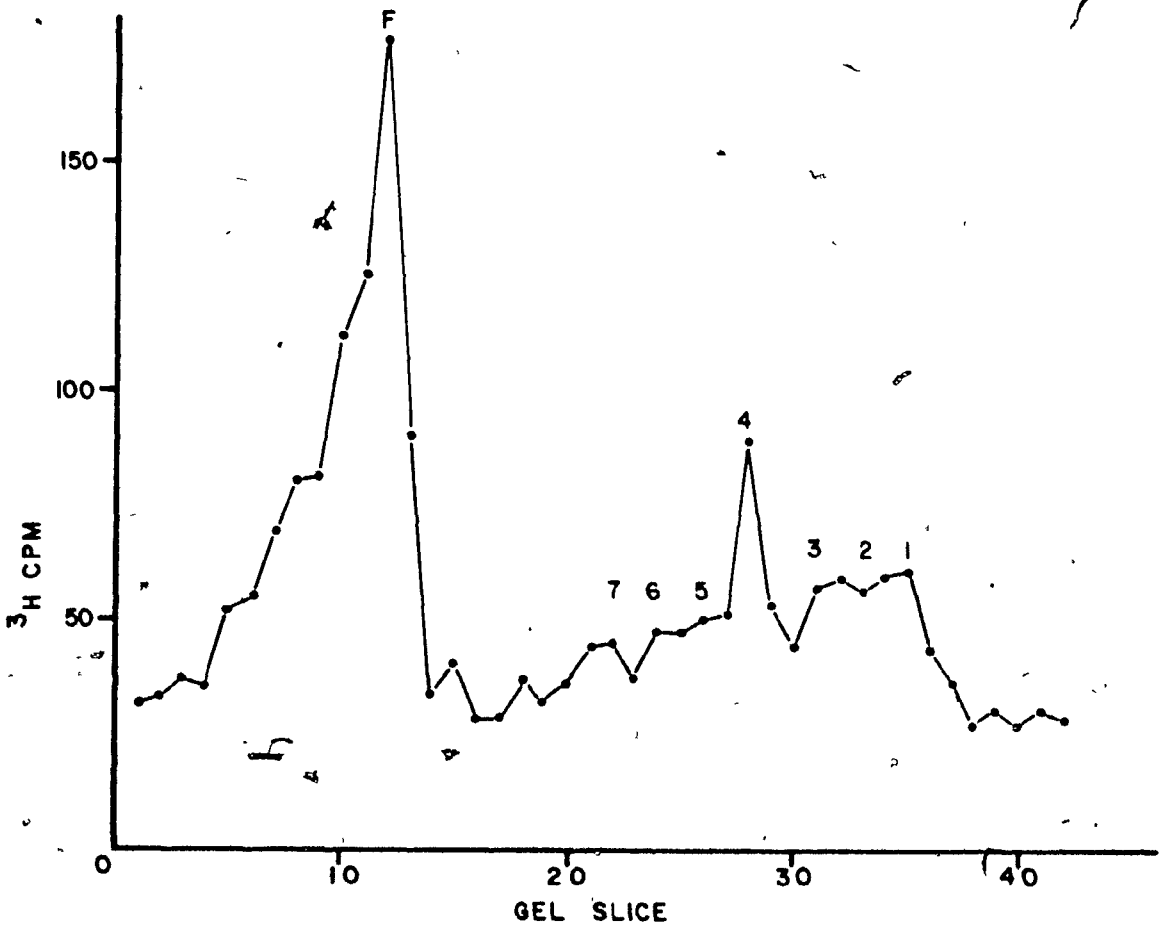

F
150
100
50
0
^{3}H CPM
4
3
2
1
7
6
5
10
20
30
40
GEL SLICE

Figure 25: Radioactivity profile from sliced gel: nuclear RNA from ins 4-infected cells labeled for 2 minutes at 28 hours post infection hybridized with polyomavirus DNA fragment HpaII-3L.

Nuclear RNA (120 µg, 0.4 x 10^6 cpm) from cells infected with mutant ins 4 (Lanoix et al, 1986) labeled for 2 minutes at 28 hours post infection was hybridized with 20 µg of single-stranded M13 mp8 DNA containing cloned polyomavirus HpaII-3L fragment. Hybrids were treated as described in Materials and Methods. After electrophoresis the gel lane was sliced and gel slices were counted as described in Materials and Methods. Only that portion of the gel containing DNA-RNA hybrid bands is shown. F corresponds to full-length DNA-RNA hybrids, and band numbers 1 to 7 correspond to initiation sites 1 to 7 shown in Figure 14.

TABLE III

Efficiency of termination per genome traverse by RNA polymerase II at 28 hr post infection by polyomavirus mutant ins 4.[a]

Experiment	cpm per 100 U residues in DNA-RNA hybrids[b] A. Less-than-full-length	B. Full-length	$\frac{\text{Column A}}{\text{Column A + Column B}}$ = Termination efficiency
1	210	250	0.46
2	340	420	0.45
3	760	630	0.55

[a] Nuclear RNA was labeled with ^{3}H-uridine for 2 minutes at 28 hours post infection of cells with polyomavirus mutant ins 4. RNA was hybridized with polyomavirus DNA HpaII-3L fragment. Hybrids, treated with RNase T1 and S1 nuclease, were subjected to electrophoresis and gel slices were counted as described in Materials and Methods.

[b] As described in Table I.

DISCUSSION

Preferential Usage of Initiation Sites at Different Stages of the Lytic Cycle.

Transcription initiation sites on the polyomavirus L-DNA strand have been studied extensively by S1 nuclease mapping of the 5' ends of mRNA or steady-state nuclear RNAs (Treisman, 1980; Treisman and Kamen, 1981; Heiser and Eckhart, 1982), or by decapping and enzymatically labeling 5' ends of mRNA (Cowie et al, 1981). These studies have revealed an extreme 5' end heterogeneity within a 100 nucleotide region (between nucleotides 5150-5050) immediately preceding the 57 nucleotide leader region (between nucleotides 5076-5020).

The results of our S1 mapping of 5' ends of pulse-labeled nuclear RNA isolated from cells infected with the AT3 strain of polyomavirus are in general agreement with these reports. However, due to the duplication of the region where initiation occurs (nucleotides 5099-5174), our virus strain has a higher degree of 5'-end heterogeneity, spreading over a 150 nucleotide region (nucleotides 5050-5202). Most RNA polymerases initiate transcription within an 80 nucleotide region, between nucleotides 5050 and 5128 (see Fig. 21 and Fig. 14 for the map of the initiation region). Within this region, there are four major initiation sites, namely: site 1 at nucleotide 5050; site 2 at nucleotide 5076; site 3 at nucleotide 5100 and site 4 at nucleotide 5128 (see Section II for more detailed description).

As shown in the results section, although all sites are used, there is a preferred major initiation site at different stages of lytic infection. At 28 hours after infection, the major initiation site is at nucleotide

5128 (site 4); the other 3 sites (sites 1 to 3) being used less frequently but at roughly equal frequencies. When 60 minute-labeled nuclear RNA was used, the DNA-RNA hybrid band corresponding to site 2, at nucleotide 5076, became the most intense band (Fig. 17, lane 4). This can be explained by the fact that there is little or no processing of nuclear RNA during a 1 or 2 min labeling period, whereas at 60 min most nuclear RNA is processed. The hybrid band corresponding to site 2 maps the 5' end of the last leader that is spliced to the VP1 mRNA coding sequences (Fig. 18). Since they only analysed steady-state nuclear RNA or mature mRNA, other investigators (Treisman, 1980; Treisman and Kamen, 1981; Heiser and Eckhart, 1982) could not distinguish between initiation at this site and splicing. When attempts were made to kinase the 5' ends of late RNAs, investigators could not measure the frequency at which this site was used because 5' ends located at nucleotide 5128 did not get kinased, showing that the kinase reaction can be selective (Cowie et al, 1981) and therefore inconclusive.

Although we found a substantial level of initiation within the leader region, at nucleotide 5050, previous authors (op. cit.) found only very low levels within this region. This discrepancy in results may be attributed to: 1) the difference in virus strain, although the sequences of the major initiation region are exactly the same; 2) the use of cytoplasmic poly (A)$^+$ RNA or steady-state nuclear RNA instead of nascent nuclear RNA. Some RNA precursors initiated at certain sites may never be processed into mature mRNA and therefore analysis of 5' ends of mRNAs present in the cytoplasmic poly (A)$^+$ fraction or in steady state nuclear RNA would not reveal all the initiation sites used during transcription, since these precursor transcripts might be rapidly degraded, not processed or not transported to the cytoplasm.

At 18 hr post infection, site 1 (nucleotide 5050) and site 4 (nucleotide 5128) were used to the same extent (Fig. 21, lane 3), in contrast to the preferential use of site 4 at 28 hr. Furthermore, at 16 hr post infection, there was a preferential usage of site 1 (nucleotide 5050), which became the major initiation site. This initiation site lies within the leader region. A similar shift of initiation site has been observed in both polyomavirus and SV40 E-strand RNA synthesized during the early phase, as compared to that of E-strand transcripts synthesized during the late phase of the lytic cycle (Fenton and Basilico, 1982; Ghosh and Lebowitz, 1981; Hansen et al, 1981). A possible explanation for the shift in initiation sites of E-strand transcription is that the binding of large T antigen to the nucleotide 94-152 region (Pomerantz et al, 1983) may prevent RNA polymerase from initiating at the major early initiation sites (between nucleotides 145 and 153) during the late phase.

The reason for the shift of initiation sites we have observed between the early and late parts of the late phase is not at all clear. We propose the following as possible mechanisms: 1) Both auxiliary elements for DNA replication are located between nucleotides 5097 and 5202 (Muller et al, 1983). During the early part of the late phase, this region may be occupied by proteins involved in DNA replication, which would force RNA polymerase to initiate more frequently at the more downstream site. As the late phase progresses, and more DNA templates become available for transcription, a smaller fraction of the DNA may be engaged in replication, and the upstream sites may therefore be used more frequently. 2) The same region also contains enhancer sequences required for both early and late gene expression (Mueller et al, 1984; Veldman et al, 1985); binding of large T antigen to this region could have the same effect of blocking access of RNA polymerase to these sites. 3) During the late part of the

late phase, increasing numbers of viral capsid proteins may bind to the viral DNA and cause changes in the chromatin conformation, shifting the major initiation site to the more upstream site. 4) An AT-rich region TAATTAAAAG (nucleotides 5158-5149) occurs in the DNA 30 nucleotides 5' to the initiation site at nucleotide 5128. This region is similar to the "Goldberg-Hogness" (TATA) box found at the same distance from most mRNAs capped termini (Gannon et al, 1979). The role of this sequence in genes that show high levels of heterogeneity at the mRNA start site is not known. Mutants that lack this sequence are fully viable (Cowie et al, 1981). It is possible that this sequence functions as a "TATA" box for polyomavirus late gene expression. This region may not be available when DNA replication is taking place during the early part of the late phase. Therefore, RNA polymerase is forced to use another AT-rich region TATTTTAAG (nucleotides 5085-5077) located further downstream, 35 nucleotides upstream of the initiation site at nucleotide 5050. Clearly further studies are required for a better understanding of this shift of initiation site.

Variations in Termination Efficiency at Different Stages of the Lytic Cycle.

During the early phase of the lytic cycle, most E-strand RNAs are terminated and processed very efficiently (Acheson and Mieville, 1978). However, during the late phase (28 hr post infection) termination and processing of L-strand RNAs are not efficient at all (Acheson, 1978, 1984). The results presented in this section show that the efficiency of transcription termination declines as the late phase progresses. During the early part of the late phase (at 16 hr post infection), the efficiency of termination is about 70%; it decreases to 60% at 18 hr, and further to 50% at 28 hr post infection.

An analogous phenomenon occurs with SV40. Processing of SV40 mRNAs in monkey cells is very rapid and efficient throughout the lytic cycle; very little unprocessed precursor being found in nuclei of infected cells (Chiu et al, 1978; Ford and Hsu, 1978). However, when SV40 DNA is injected into Xenopus oocyte nuclei, multigenomic length viral RNA molecules are synthesized, and processing of these giant molecules is very inefficient (Miller et al, 1982). Studies of the synthesis and degradation of heterogenous nuclear RNA in stage 6 oocytes of Xenopus laevis have shown that approximately 95% of this RNA turns over in the nucleus (Anderson and Smith, 1977). Diaz et al (1981) have shown that termination of histone gene transcription in lampbrush chromosomes is inefficient. Taken together, these results suggest that stage 6 oocytes of Xenopus laevis lack some processing factor(s). Thus, although the transcription rate is normal, most transcripts accumulate and turn over in the nucleus. Termination, splicing and polyadenylation of these transcripts are not efficient. Therefore, only a small portion of RNA precursors is fully processed to mature mRNA and transported to the cytoplasm.

This situation is very similar to that observed in polyomavirus-infected mouse cells during the late phase of infection. Transcription of viral DNA is still very active, but transcription termination and processing are very inefficient. Giant viral RNA molecules accumulate in the nuclei; only a small portion (5%) of these transcripts is processed and transported to the cytoplasm (Acheson, 1984). When an additional polyadenylation signal is inserted into the wild-type DNA (ins 4), the efficiency of polyadenylation and that of termination are not increased at 28 hr post-infection. This suggests that inefficient termination and polyadenylation (processing) are the result of limiting amounts of cellular factor(s) required for both these processes. However, we cannot rule out

other possibilities; for example, viral late gene products may act as antitermination factors. It would be interesting to know whether the efficiency of termination and that of polyadenylation are higher in ins 4-infected cells at 16 hr post infection compared to that at 28 hr. Howeyer, this experiment may be difficult to achieve, because of the lower titers of ins 4 virus stocks and the small amount of late viral RNAs present in infected cells during the early part of the late phase: in wild type-infected cells, about 0.05% of 2 min-labeled nuclear RNA hybridized to polyomavirus fragment HpaII-3L at 16 hr post infection (data not shown). Each lane on the gel described in the results section (Fig. 22, lane 2) contains 2 to 3 90 mm petri dishes worth of total nuclear RNA, and the autoradiograms have to be exposed for 4 to 5 weeks or longer, in order to obtain a reasonable signal. At 18 hr post infection, about 0.21% of 2 min labeled viral RNA in the infected cell nucleus hybridized to polyomavirus fragment HpaII-3L, and at 28 hr about 3%. In contrast, in ins 4-infected cells, only about 0.8% of 2 min-labeled nuclear RNA extracted at 28 hr post infection hybridized to fragment HpaII-3L.

As noted above, transcription of E-strand RNA is efficiently terminated during the early phase of the lytic cycle. However, there is evidence that during the late phase of infection, E-strand transcripts are complementary to the entire L DNA strand (Kamen et al, 1974; Beard et al, 1976; Kamen et al, 1982). The same results have been obtained in our laboratory with in vivo labeled RNA isolated 28 hr post infection (McNally, 1983) and with RNA labeled in vitro in isolated nuclei or in viral transcription complexes (Skarnes, 1985). In the case of E-strand RNA as well, inefficient termination during the late phase of the lytic cycle may be due to limiting amounts of termination and processing factors. However, more studies are required to prove this point. For example, it would be of

interest to study the kinetics of accumulation and processing of transcripts of a cellular gene during polyomavirus infection, especially during the late phase, to determine whether the processing of these cellular RNAs is efficient or not.

If termination is inefficient during the late phase of polyomavirus infection because of limiting amounts of termination factors, one would expect cellular genes to behave in the same fashion, competing for available factors. If, on the other hand, termination is inefficient because of an inherently weak termination signal on the L strand of polyomavirus, transcription of the cellular gene would be expected to be efficiently terminated.

SECTION IV

Identification of In Vivo Transcription Termination and/or Pausing Sites on the L-Strand of Polyomavirus DNA

INTRODUCTION

In higher eukaryotes, the mature 3' ends of mRNAs are created by endonucleolytic cleavage of precursor RNAs followed by the addition of a poly(A) tract to the newly exposed 3' ends (Nevins and Darnell Jr., 1978). The hexanucleotide, AAUAAA, is found 10 to 30 nucleotides upstream of polyadenylation sites of many mRNAs (Proudfoot and Brownlee, 1976). Removal or alteration of this AAUAAA motif prevents cleavage of the pre-mRNAs (Fitzgerald and Shenk, 1981; Montell et al, 1983; Wickens and Stephenson, 1984).

Not much is known about the sites at which RNA polymerase II terminates transcription of mRNA coding genes, nor is the nature of the sequences which signal transcription termination understood. There is evidence that RNA polymerases transcribe past polyadenylation sites and terminate up to several kilobases downstream of these sites (Ford and Hsu, 1978; Nevins et al, 1980; Hofer and Darnell Jr., 1981).

In vitro nuclear run-on experiments have revealed a nearly equimolar distribution of label in the spacer DNA immediately downstream of polyadenylation sites when compared to label in sequences upstream of these sites (Frayne et al, 1984; Mather et al, 1984). Other studies have shown a progressive reduction in the transcription of regions downstream of the polyadenylation site (Citron et al, 1984; Hagenbuchle et al, 1984; LeMeur et al, 1984; Rohrbaugh et al, 1985). S1 nuclease mapping of in vitro

labeled nuclear RNA (Citron et al, 1984), or steady state nuclear RNA (Hagenbuchle et al, 1984) have suggested that termination of transcription occurred at multiple sites within a region several hundred nucleotides long.

Falck-Pedersen et al (1985) inserted into the adenovirus-5 E1A gene segments of the mouse major β-globin gene, which contained a transcription termination region. Analysis of RNA labeled in isolated nuclei demonstrated that transcription terminated within the insert when it contained the mouse β-globin gene polyadenylation site plus 1.4 kb of DNA, including the region where heterogeneous transcription termination was observed. When inserted in the opposite orientation, this DNA fragment did not signal transcription termination. Insertion of the segment that contains the heterogeneous termination region alone in either orientation did not direct any termination, suggesting that termination of transcription might be linked to the polyadenylation event.

During the late phase of lytic infection, at 28 hr post infection, termination of transcription of polyomavirus L DNA strand by RNA polymerase II is inefficient. RNA polymerases can therefore traverse the viral genome several times without terminating, which gives rise to giant transcripts that contain tandem repeats of the viral genome (Acheson, 1978). We have shown that approximately 50% of RNA polymerases terminate transcription per genome traverse (This thesis: Tseng and Acheson, 1986). Hybridization of in vivo pulse-labeled nuclear RNA (McNally, 1983), or RNA labeled in vitro in isolated nuclei or in viral transcription complexes (Skarnes, 1985) to various restriction fragments of polyomavirus L DNA strand has revealed a lower level of hybridization to fragments downstream of the polyadenylation site when compared to the amount of radioactivity hybridized to upstream fragments.

Polyadenylation occurs within 1-2 min after RNA polymerase passes through the poly(A) site (Acheson, 1984). Transcripts downstream of polyadenylation sites are extremely unstable (Hofer and Darnell, 1981; Citron et al, 1984). Since processing of mRNA occurs so fast, unprocessed transcripts must only represent a very small proportion of total steady state RNAs. Therefore, in order to be able to detect 3' ends of RNAs which might correspond to termination sites, one must analyze nascent nuclear RNA labeled in vivo for a very brief period. Nuclear RNA labeled for 1 or 2 minutes (nascent RNA) should contain a mixture of processed and unprocessed RNA, and should allow us to map 3' ends that are created by termination. By use of our S1 nuclease mapping method (described in Section I; Tseng and Acheson, 1986), we have detected RNAs with a series of 3' ends which may represent sites where RNA polymerase II either pauses during elongation, or terminates transcription.

RESULTS

Transcription Does Not Terminate at a Single Major Site on Polyomavirus L-strand DNA.

3T6 cells infected with polyomavirus were labeled with ^{3}H-uridine for 60 min from 27 to 28 hr after infection. Nuclear RNA was hybridized to the L strand of polyomavirus PstI fragments 1 to 5, as well as to HincII-PvuII fragment 2L (nt 2032 to 2962) (see Figure 4). After ribonuclease treatment, filter binding and S1 nuclease treatment, DNA-RNA hybrids were analyzed by gel electrophoresis. The results are shown in Figure 26. The most prominent band present in all lanes represents full-length DNA-RNA hybrids (arrowheads). Lane 1 shows the results of hybridization of RNA with fragment PstI-1L. There are two other major hybrid bands besides full-length hybrids. These two bands correspond to RNAs with a 3' end at nucleotide 1230 (upper band), and a 5' end at nucleotide 1170 (lower band), first detected by Treisman and Kamen (1981) in steady-state nuclear RNA (see below, Figure 30). Lane 2 shows hybrids formed with fragment PstI-2L. Bands just below the full-length hybrid correspond to RNAs whose 5' ends are located at the transcription initiation sites contained within this fragment, and which are continuous with the remainder of the fragment (see Section II). The minor band about 480 nucleotides long corresponds to the 5' end of the body of VP3 mRNA which is spliced at nucleotide 4707 to the leader at nucleotide 5020 (see Figure 1).

Lane 3 shows the results of hybridization with fragment PstI-3L. Besides the full-length hybrid, the band about 810 nucleotides long corresponds to the 5' end of the body of VP1 mRNA which is spliced from nucleotide 4122 to the leader (see Figure 1). Lane 4 shows hybrids formed

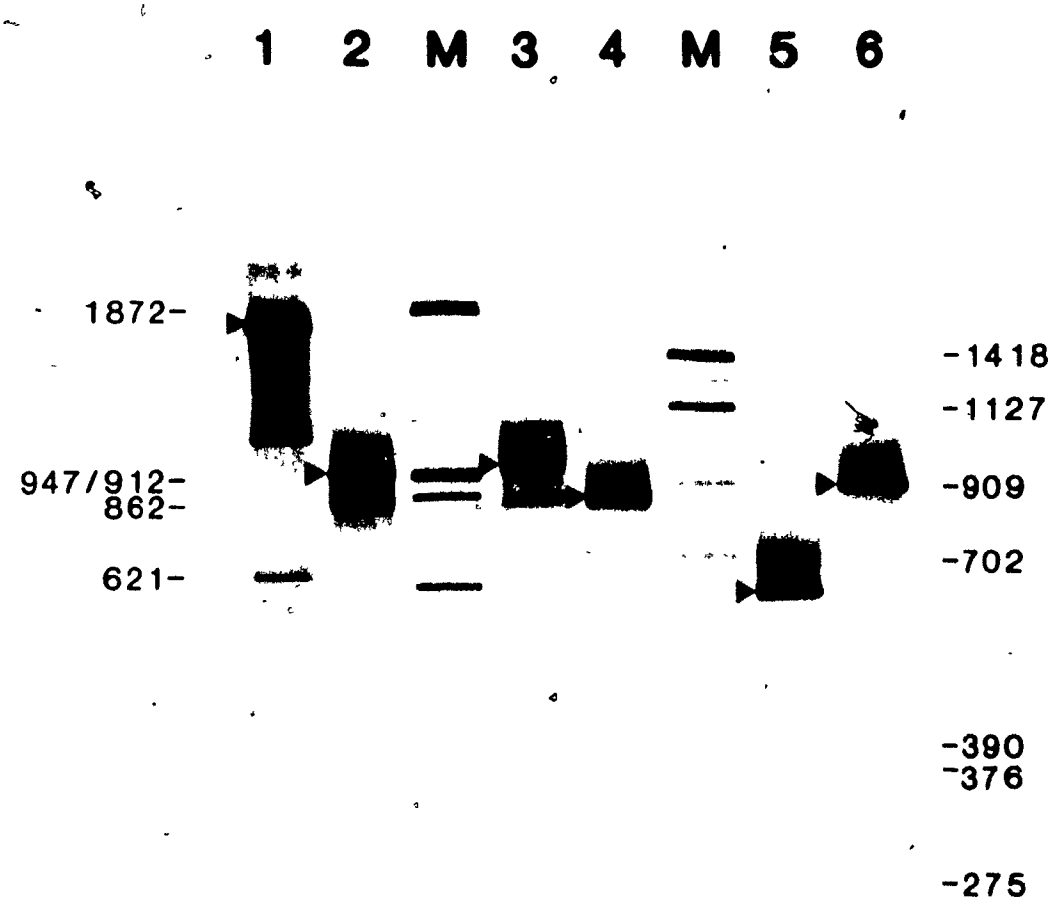

1 2 M 3 4 M 5 6
1872-
947/912-
862-
621-
-1418
-1127
-909
-702
-390
-376
-275

Figure 26: Hybridization of nuclear RNA labeled with ^{3}H-uridine for 60 min to polyomavirus DNA fragments.

Polyomavirus infected 3T6 cells were labeled with ^{3}H-uridine at 27 hr post-infection for 60 min. Total nuclear RNA (50 µg, 4.4 x 10^{6} cpm) was hybridized with 10 µg of single-stranded bacteriophage fd103 or M13 mp8 DNA containing cloned polyomavirus DNA fragment PstI-1L (lane 1), PstI-2L (lane 2), PstI-3L (lane 3), PstI-4L (lane 4), PstI-5L (lane 5), or HindII-PvuII-2L (lane 6) (see Figure 4 for the restriction map). Hybrids were treated with 200 units/ml of RNase T1, bound to nitrocellulose filters, and eluted by incubation with 333 units/ml of S1 nuclease as described in Materials and Methods. Position of full-length DNA-RNA hybrids are indicated by an arrowhead. Marker DNAs (lanes M) were ^{3}H-polyomavirus DNA digested with PstI or MspI as in Figure 8.

with fragment PstI-4L. This fragment contains the polyadenylation site for L-strand RNAs at nucleotide 2900 (Kamen et al, 1980). A hybrid band 410 nucleotides in length corresponds to RNAs with 3' ends at this site.

Lane 5 shows hybrids with fragment PstI-5L. Surprisingly, two hybrid bands migrated more slowly than the full-length hybrid (621 nucleotides); no hybrid larger than full length is expected after S1 nuclease treatment of DNA-RNA hybrids. We believe that these bands migrated more slowly because of secondary structure in the DNA-RNA hybrids (see below). A prominent minor hybrid band about 400 nucleotides long was also observed in lane 5. This band would correspond to an RNA whose 3' end is at nucleotide 80, located just downstream of a GC-rich palindrome between the two large T antigen binding sites A and B (see Figure 31B) and will be further discussed below. Lane 6 shows hybrids with HincII-PvuII fragment 2L. This fragment covers nearly 800 nucleotides downstream from the polyadenylation site and overlaps with fragments PstI-4L and PstI-1L. There were no major bands in addition to the full-length hybrid band.

Having established patterns of hybrids with 1 hr-labeled RNAs, I then analyzed late polyomavirus nuclear RNAs labeled for 2 min by using the same set of cloned DNA fragments (Fig. 27). Two minute-labeled RNA should contain mostly growing RNA chains and RNAs which have been terminated and released but not yet processed into mature mRNAs. Termination sites should be revealed by the presence of RNAs with specific 3' ends in pulse-labeled RNA. Again, the most prominent bands were full-length hybrids. In lane 1, the two smaller bands which were present in 1 hr-labeled RNA (Figure 26) were present at only low levels. This suggests that the 3' end at nucleotide 1230 and the 5' end at nucleotide 1170 are created by post-transcriptional cleavage of longer RNAs, and that very little cleavage takes place within 2 minutes of their synthesis. In lane 2, bands


1 2 M 3 4 M 5 6
1872-
947/912-
862-
621-
-1418
-1127
-909
-702
-390
-376
-275

Figure 27: Hybridization of nuclear RNA labeled with ^{3}H-uridine for 2 min to polyomavirus DNA fragments.

Polyomavirus infected 3T6 cells were labeled with ^{3}H-uridine for 2 min at 28 hr post infection. Total nuclear RNA (50 µg, 1.1 x 10^{6} cpm) was hybridized with 10 µg of single-stranded bacteriophage fd103 or M13 mp8 DNA containing cloned polyomavirus DNA fragment PstI-1L (lane 1), PstI-2L (lane 2), PstI-3L (lane 3), PstI-4L (lane 4), PstI-5L (lane 5), or HindII-PvuII-2L (lane 6) (see Figure 4 for the restriction map). Hybrids were treated with 200 units/ml of RNase T1, bound to nitrocellulose filters, and eluted by incubation with 333 units/ml of S1 nuclease as described in Materials and Methods. Position of full-length DNA-RNA hybrids are indicated by an arrowhead. Marker DNAs (lanes M) were ^{3}H-polyomavirus DNA digested with PstI or MspI as in Figure 8.

corresponding to RNAs whose 5' ends are at the transcription initiation sites were present. This is expected, as initiation of transcription should take place continuously, and therefore some newly-initiated RNAs should be synthesized during the 2-minute labeling period. On the other hand, bands which correspond to the splicing sites at nucleoides 4707 and 4122 (lanes 2 and 3) and to the polyadenylation site at nucleotide 2900 (lane 4), were not present in significant amounts. Since polyadenylation and splicing are post-transcriptional events, processed RNAs are not expected to be seen in briefly pulse-labeled RNA. Interestingly, the 400-nucleotide band in lane 5 which corresponds to an RNA with a 3' end at nucleotide 80 was as intense in 2 min-labeled RNA as in 1 hr-labeled RNA. Therefore, this 3' end may represent a strong pausing or termination site for RNA polymerase on polyomavirus DNA (see below). Finally, numerous minor bands were present in all lanes when 2 min-labeled RNA was analyzed, in contrast to the 1 hr-labeled RNA. These minor bands, if they represent 3' ends of RNAs, could reveal sites where RNA polymerase II either pauses or terminates transcription.

Kinetics of Processing of Viral RNAs.

The rate of appearance of viral RNAs whose 3' or 5' ends are created by processing events was examined more closely. Nuclear RNA was labeled with ^{3}H-uridine for 2, 10 or 60 minutes at 28 hr after infection and hybridized to PstI fragment 1L (Figure 28) or PstI fragment 4L (Figure 29). In both cases, the most prominent band was full-length hybrids, as noted previously. RNA labeled for 2 min gave rise to a series of discrete minor hybrid bands, which are reproducible in relative intensity and position in different experiments. As labeling time increased to 10 and 60 min, the intensity of bands corresponding to the major processed RNAs increased relative to that of the minor bands.

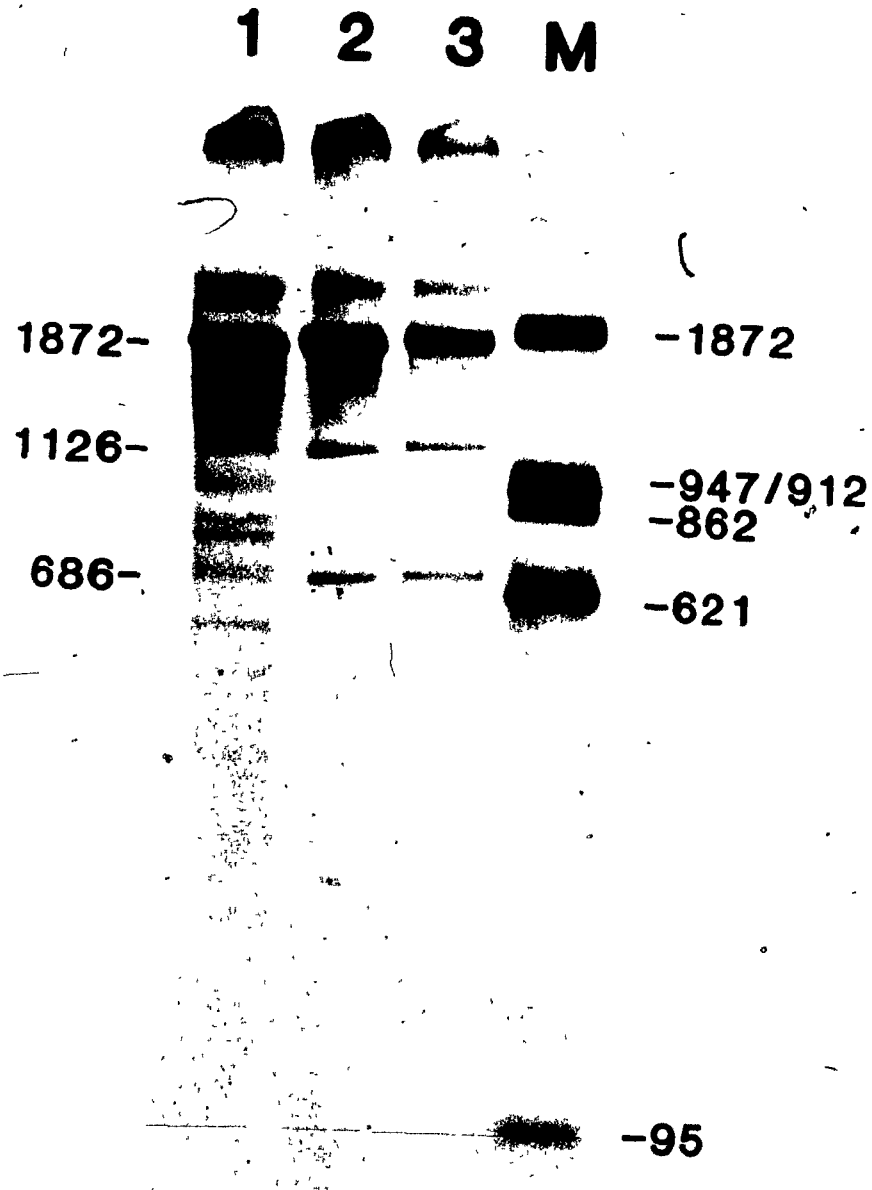

1
2
3
M
1872-
1126-
686-
-1872
-947/912
-862
-621
-95

Figure 28: Hybridization of ^{3}H-labeled nuclear RNA to polyomavirus DNA fragment PstI-1L.

Each of the following ^{3}H-uridine labeled RNA samples was hybridized with single-stranded bacteriophage fd103 DNA containing cloned polyomavirus DNA fragment PstI-1L. Lane 1: nuclear RNA (90 µg, 8.6 x 10^5 cpm) labeled for 2 min at 28 hr post-infection, hybridized with 18 µg DNA; lane 2: nuclear RNA (45 µg, 9.1 x 10^5 cpm) labeled for 10 min at 28 hr post infection, hybridized with 9 µg DNA; lane 3: nuclear RNA (20 µg, 9.5 x 10^5 cpm) labeled for 60 min from 27 to 28 hr post infection, hybridized with 4 µg DNA. Hybrids were treated with 1 µg/ml of RNase A, followed by 125 µg/ml of proteinase K, bound to nitrocellulose filters, and eluted by S1 nuclease as described in Materials and Methods. Sizes and positions of expected DNA-RNA hybrids are indicated. Marker DNAs were ^{3}H-polyomavirus DNA digested with PstI (lane M).

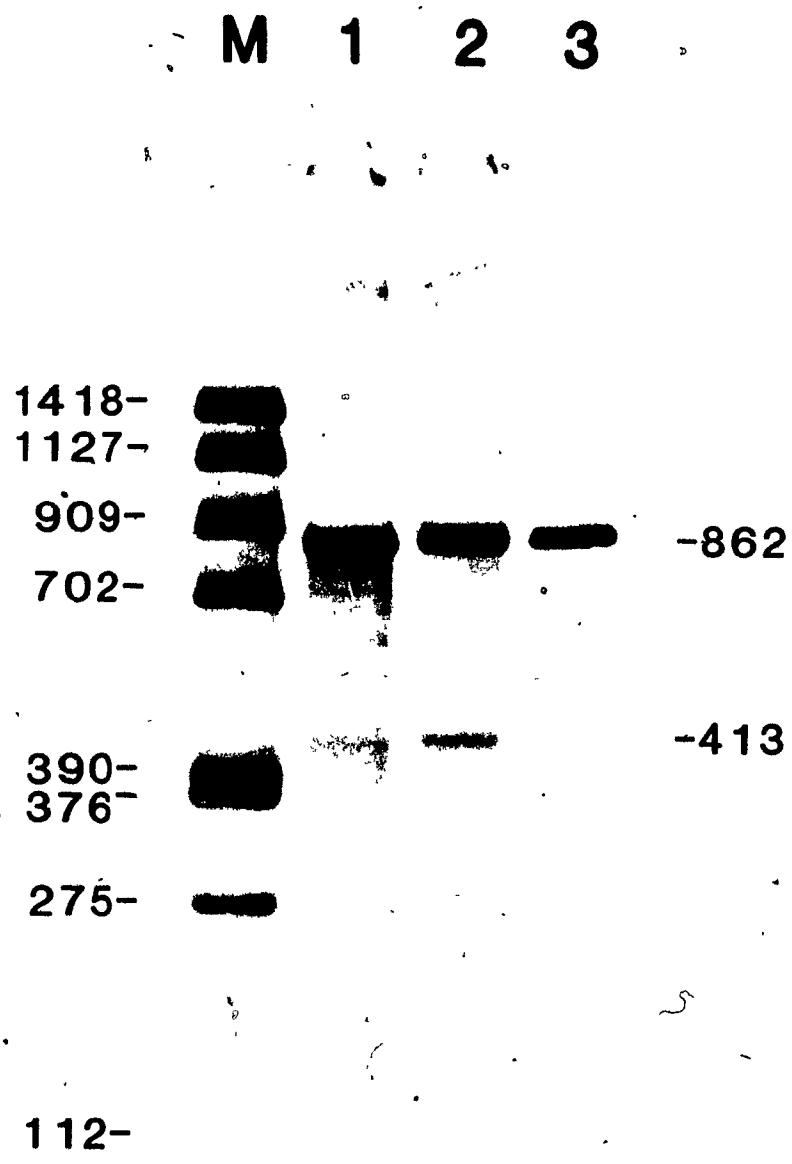

M 1 2 3
1418-
1127-
909-
702-
390-
376-
275-
112-
-862
-413

Figure 29: Hybridization of ^{3}H-labeled nuclear RNA to polyomavirus DNA fragment PstI-4L.

Each of the following ^{3}H-uridine labeled RNA samples was hybridized with single-stranded bacteriophage fd103 DNA containing cloned polyomavirus DNA fragment PstI-4L. Lane 1: nuclear RNA (90 µg, 8.6 x 10^{5} cpm) labeled for 2 min at 28 hours post infection, hybridized with 18 µg DNA; lane 2: nuclear RNA (45 µg, 9.1 x 10^{5} cpm) labeled for 10 min at 28 hours post infection, hybridized with 9 µg DNA; lane 3: nuclear RNA (20 µg, 9.5 x 10^{5} cpm) labeled for 60 min from 27 to 28 hours post infection, hybridized with 4 µg DNA. Hybrids were treated with 1 µg/ml of RNase A, followed by 125 µg/ml of proteinase K, bound to nitrocellulose filters, and eluted by S1 nuclease as described in Materials and Methods. Sizes and positions of expected DNA-RNA hybrids are indicated. Marker DNAs were ^{3}H-polyomavirus DNA digested with MspI (lane M).

Ideally, to compare relative intensities of these bands among different RNA samples, identical levels of S1 nuclease-resistant viral RNA should be loaded onto each gel lane. However, this was difficult to achieve in practice, since specific radioactivities of RNAs varied from sample to sample, and the proportion of labeled RNA which was virus-specific also varied. In the experiments shown in Figures 28 and 29, more labeled virus-specific RNA was present in samples labeled for 2 min (lane 1) than in samples labeled for 10 or 60 min (lanes 2 and 3). It can nevertheless be concluded that only very low levels of processed RNAs examined (RNAs with a 5' end at nt 1170 or a 3' end at nt 1230 in PstI fragment 1L, Figure 28; RNA with a polyadenylated 3' end at nt 2900 in PstI fragment 4L, Figure 29) were present in the 2 min-labeled RNAs. Therefore, the assumption that 2 min-labeled RNA contains mainly unprocessed and nascent RNA chains is largely justified, at least insofar as these particular processing events are concerned.

It is known that cleavage of mRNA precursors to form the polyadenylated 3' end of mature mRNA occurs within 1-2 min of their synthesis (Nevins and Darnell, 1978; Acheson, 1984). Since very little of the 2 min-labeled RNA has been cleaved at this site (Figure 29, lane 1), I conclude that the effective labeling time in this sample is 1 min or less (see Section II, Tseng and Acheson 1986). By 10 min of labeling, substantial levels of all three processed RNAs examined here have been made. Therefore, the cleavage of RNAs at nucleotides 1170 and 1230 must take place within a few minutes of the synthesis of their precursor RNA. Note that in neither case is a majority of the RNA processed even after 60 min (the intensity of full-length hybrids being much greater than that of the processed species, taking into account the relative lengths of RNAs in

the different bands). This suggests that some of the viral RNA synthesized is rapidly processed, and that much of it is never processed (see Acheson, 1981b).

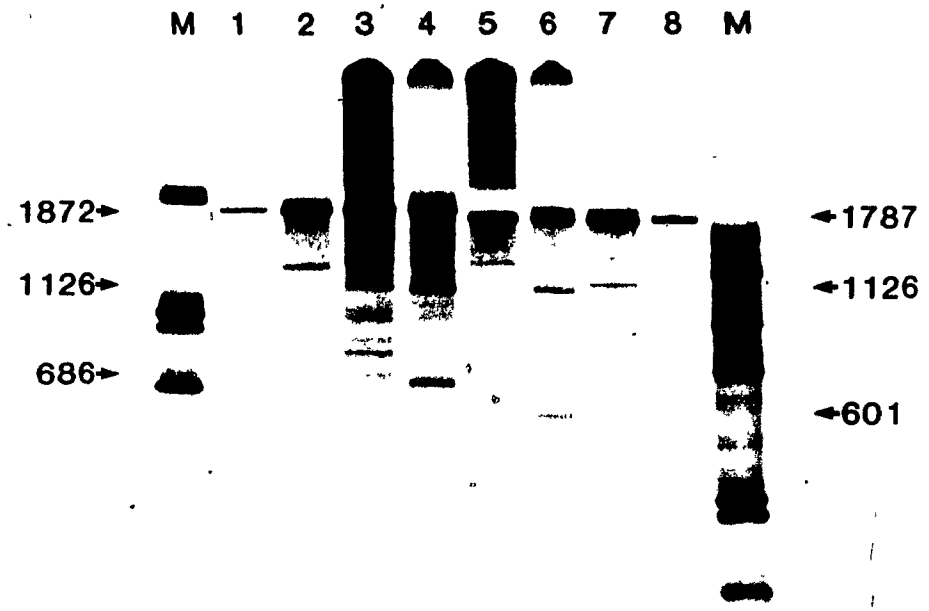
There is no evidence for the accumulation of an RNA species whose 5' end would be created by cleavage of a precursor at the polyadenylation site and whose 3' end would be located somewhere downstream of the end of PstI fragment 4L. Such an RNA would appear in Figure 29 as a band 450 nucleotides long, which should migrate somewhat more slowly than the 413-nucleotide band which corresponds to polyadenylated RNA. This suggests that the distal portion of the precursor RNA is very unstable after cleavage at the polyadenylation site, and that it is rapidly degraded.

Most of the Minor Bands Correspond to RNAs with Distinct 3' Ends.

It was important to determine whether the numerous minor bands detected in pulse-labeled RNA correspond to RNAs with specific 5' or 3' ends. If RNA polymerase either pauses during chain elongation or terminates transcription at specific sites on the template DNA, RNAs with specific 3' ends will be generated. On the other hand, rapid endonucleolytic cleavage at specific sites on precursor RNAs would give rise to a population of RNAs of which one-half would have 5' ends at those sites, and one-half would have 3' ends at those sites (see Figure 30B).

To determine whether the numerous minor bands detected in pulse-labeled RNA correspond to RNAs with specific 5' or 3' ends, I hybridized RNA with two cloned DNA fragments which differed only in their distal ends. PstI fragment 1 was cleaved with SacI and the resulting fragment was recloned into M13 mp8. This was carried out by Thu-Hang Tran in our laboratory. The resulting clone is 85 nucleotides shorter than PstI fragment 1 (see Figure 30B). RNAs whose 3' ends lie within this fragment

A.


M 1 2 3 4 5 6 7 8 M
1872
1126
686
1787
1126
601

B.

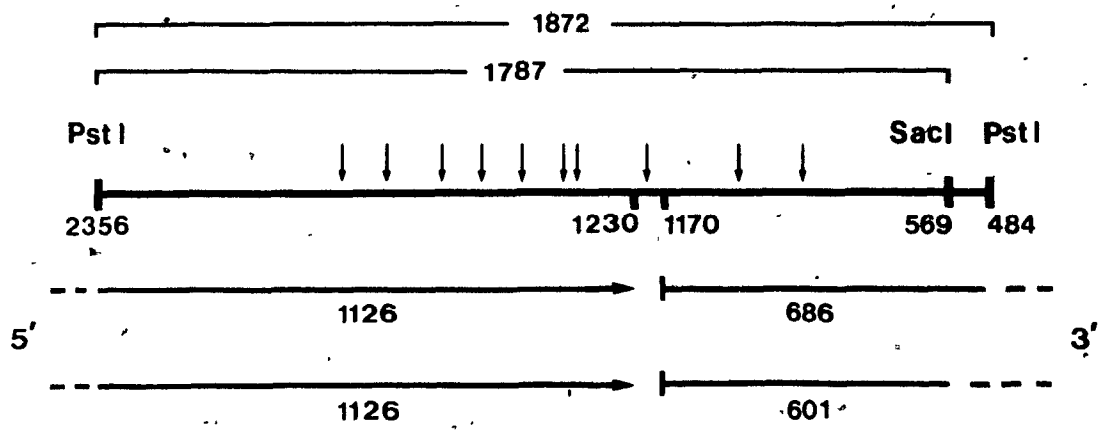

1872
1787
Pst I
Sac I
Pst I
2356
1230
1170
569
484
1126
686
5'
3'
1126
601

Figure 30. Hybridization of ^{3}H-RNA synthesized *in vitro* or *in vivo* to polyomavirus DNA fragments PstI-1L or PstI-1L-SacI subfragment.

A. ^{3}H-labeled RNAs described below were hybridized to 10 µg single-stranded M13 mp8 DNA containing cloned polyomavirus PstI fragment 1L (lanes 1-4) or PstI-1L-SacI subfragment (lanes 5-8) (see text). Lanes 1 and 8: ^{3}H-labeled RNA (4.5×10^3 cpm) transcribed *in vitro* by SP6 RNA polymerase from a 5.25 kb XbaI fragment of polyomavirus DNA inserted into plasmid SP64 (see Materials and Methods); lanes 2 and 7: ^{3}H-labeled RNA (90×10^3 cpm), same as lanes 1 and 8; lanes 3 and 5: nuclear RNA (50 µg, 0.45×10^6 cpm) labeled for 2 minutes *in vivo* at 28 hours post infection; lanes 4 and 6; nuclear RNA (21 µg, 1.3×10^6 cpm) labeled for 60 min *in vivo* from 27 to 28 hours post infection. Hybrids were treated with 400 units/ml RNase T1, bound to nitrocellulose filters, and eluted by incubation with S1 nuclease as described in Materials and Methods. Marker DNAs were as in Figure 8. Electrophoresis was at 50V (20 mA) for 10 hours.

B. Diagram showing DNA-RNA hybrids generated by S1 nuclease mapping using polyomavirus DNA fragment PstI-1L and PstI-1L-SacI subfragment. Solid line represents polyomavirus DNA fragment PstI-1L. SacI restriction site is indicated. Numbers below the line show the positions of PstI and SacI sites, and 5' end (nt 1170) and 3' end (nt 1230) found in steady-state RNA by Treisman and Kamen (1981). Numbers above the line show sizes of these two DNA fragments. Lines below the solid line show sizes of DNA-RNA hybrids which correspond to RNAs whose 5' ends at nt 1170 and 3' ends at nt 1230 when hybridized to these two DNA fragments. Arrows indicate positions of 3' ends of 2 min-labeled RNA.

will generate bands of equal length whether hybridized to the truncated PstI-SacI fragment or to the full-length PstI fragment 1L. On the other hand, RNAs whose 5' ends lie within this region will generate shorter hybrid bands when hybridized with the PstI-SacI fragment than when hybridized with the full-length PstI fragment 1L.

Nuclear RNA labeled in vivo for 2 or 60 min at 28 hr after infection, or polyomavirus L-strand RNA synthesized in vitro with SP6 RNA polymerase (see Materials and Methods), were hybridized to single-stranded DNAs corresponding to PstI fragment 1L or to the PstI-SacI subclone. The results are shown in Figure 30A. Full-length hybrids were slightly shorter in the samples hybridized to PstI-SacI fragment (lanes 5 to 8) than in the samples hybridized to PstI fragment 1L (lanes 1 to 4), as expected. When 60 min-labeled RNA was used (lanes 4 and 6), the band 1126 nucleotides long which corresponds to RNA with a 3' end at nucleotide 1230 migrated at the same position when hybridized with either DNA. The band 686 nucleotides long which corresponds to RNA with a 5' end at nucleotide 1170 was shifted, as expected, to a size of about 600 nucleotides when hybridized with the truncated fragment.

When 2 min-labeled RNA was used (lanes 3 and 5), most of the discrete minor bands migrated at the same positions regardless of which DNA fragment was used. These results demonstrate that most of the minor bands seen after hybridization with pulse-labeled RNA represent RNAs with 3' ends at numerous sites within this fragment. Therefore, these 3' ends represent either sites at which RNA polymerase pauses or sites at which the transcription is terminated and nascent RNA chains are released.

When in vitro synthesized RNA was hybridized to these DNA fragments, two other hybrid bands besides the full-length hybrid were detected (lanes 2 and 7). These two bands were present at very low levels, if at all, when

in vivo RNA was used. It is likely that they represent artifacts caused by ribonuclease or S1 nuclease, since only full-length run-off transcripts were used (see Materials and Methods). A nick could have been created around nucleotide 1770 which would generate hybrids 1282 (5' end) and 590 (3' end) nucleotides long when PstI fragment 1L was used (lane 2) or 1197 (5' end) and 590 (3' end) nucleotides long when the PstI-SacI fragment was used (lane 7). Lanes 1 and 8 show that when 20 fold-less in vitro RNA was used in hybridization, only the full-length hybrids could be seen.

A DNA-RNA Hybrid Band Which Migrates More Slowly Than Full-length Hybrids May be Caused by Internal Secondary Structure.

As noted above, a hybrid band which migrated more slowly than the full-length hybrid was observed in RNA samples hybridized to polyomavirus PstI fragment 5L (Fig. 31A, lane 1) (two slowly-migrating bands were observed in Figure 26, lane 5). We have verified the PstI-5L DNA insert in the double-stranded fd103 clone by cutting out the polyomavirus DNA insert with restriction endonuclease PstI, and found that the insert was correct (data not shown). The PstI-5L DNA fragment was also independently recloned into single-stranded bacteriophage M13 mp8. When nuclear RNA was hybridized to polyomavirus HpaII fragment 5L (nucleotides 399 to 9), which overlaps with PstI fragment 5L, an additional hybrid band also migrated more slowly than the full-length hybrid (390 bp) (data not shown). However, when hybridization was carried out in the presence of 80% formamide, 0.4 M NaCl at 55°C, instead of in 0% formamide, 0.75M NaCl at 85°C (see Materials and Methods), this slower migrating band disappeared (compare Fig. 31A, lanes 1 and 2). This result suggested that the slower migrating band represents the full-length DNA-RNA hybrid which contained secondary structure that altered its mobility. This secondary structure

A.


M
1
2
1872–
947/912–
862–
621–

B.

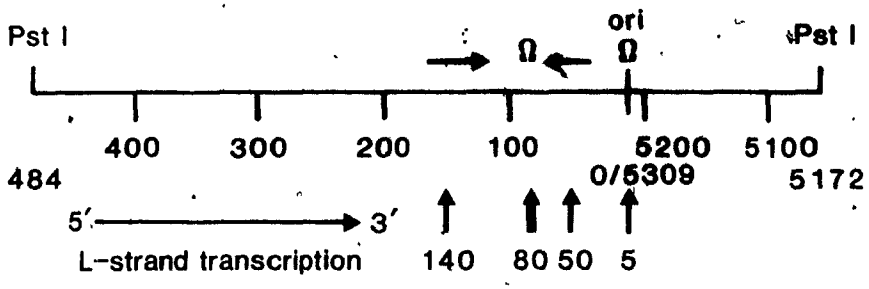

Pst I
ori
Ω
Ω
Pst I
400
300
200
100
5200
5100
484
0/5309
5172
5′
3′
L–strand transcription
140
80
50
5

Figure 31: Hybridization of ^{3}H-uridine labeled nuclear RNA to polyomavirus DNA fragment PstI-5L in the presence or absence of formamide.

A. ^{3}H-uridine labeled nuclear RNA (18 µg, 1.8 x 10^6 cpm) was hybridized with 10 µg of single-stranded bacteriophage M13 mp8 DNA containing polyomavirus DNA fragment PstI-5L in buffer containing no formamide, 0.75 M NaCl (lane 1) or in buffer containing 80% formamide, 0.4 M NaCl (lane 2) as described in Materials and Methods. Hybrids were treated with 200 units/ml of RNase T1, bound to nitrocellulose filters, and eluted by incubation with S1 nuclease as described in Materials and Methods. Arrowheads indicate positions of full-length DNA-RNA hybrids. Dots indicates prominent less-than-full=length hybrids, Marker DNAs were ^{3}H-polyomavirus DNA digested with PstI (lane M).

B. Diagram showing polyomavirus DNA fragment PstI-5L. The ori is the replication origin. (Ω) indicates position of GC-rich palindromes (see Figures 32 and 33). Arrows above the line indicate the large T antigen binding sites. Arrows below the line show positions and nucleotide numbers of the prominent less-than-full-length hybrids which correspond to 3' ends.

apparently does not form when hybridization is done in 80% formamide. In fact, sequence analysis reveals a possible GC-rich stem-loop structure between nucleotides 84 and 100 (Fig. 32).

A Possible Strong Pausing and/or Termination Site(s) on Fragment PstI 5L.

As mentioned earlier, nuclear RNA labeled for 2 min revealed a series of minor hybrid bands when hybridized to PstI fragment 5L (Fig. 27, lane 5) (These bands were more clearly visible in the original autoradiogram). When 60 min labeled nuclear RNA was hybridized, four minor hybrid bands were detected, comparable to those seen in the samples of 2 min RNA (Fig. 26, lane 5; Fig. 31A, lanes 1 and 2). The most prominent hybrid band in both the 2 min and 60 min labeled samples is 400 nucleotides in length and corresponds to an RNA whose 3' end is at nucleotide 80 (Fig. 32). This is just downstream of the GC-rich dyad symmetry element, and lies between large T-antigen binding sites A and B (Cowie and Kamen, 1984 and 1986). This band could represent a pausing and/or termination site for RNA polymerase. The other 3 minor hybrids have sizes of about 350, 435, and 480 base pairs, which would correspond to RNAs whose 3' ends are at nucleotides 140, 50 and 5. It is interesting to note that nucleotides 140 and 50 are located within large T-antigen binding sites A and C (Fig. 31B) (Cowie and Kamen, 1984). Nucleotide 5 is located at the origin of replication, within a 32 base pair GC-rich palindrome (Muller et al, 1983) (Fig. 33). These three sites could also be termination or pause sites.

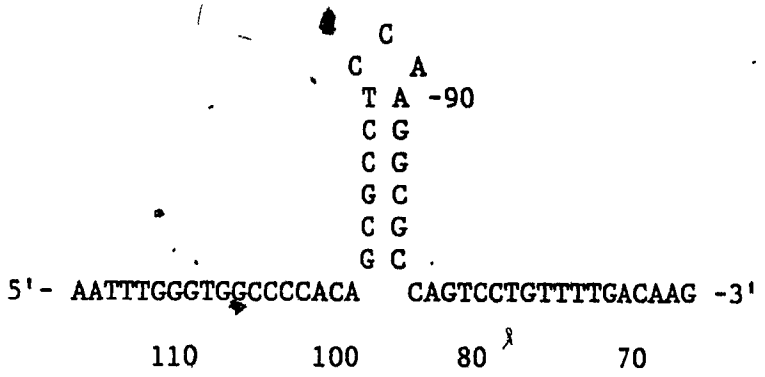


Figure 32: Nucleotide sequence of polyomavirus L DNA strand from nucleotides 117 to 67.

A GC-rich stem-loop structure can be formed between nucleotides 84 and 100. A putative strong pause and/or termination site at nucleotide 80 is located immediately downstream of the GC-rich palindrome.

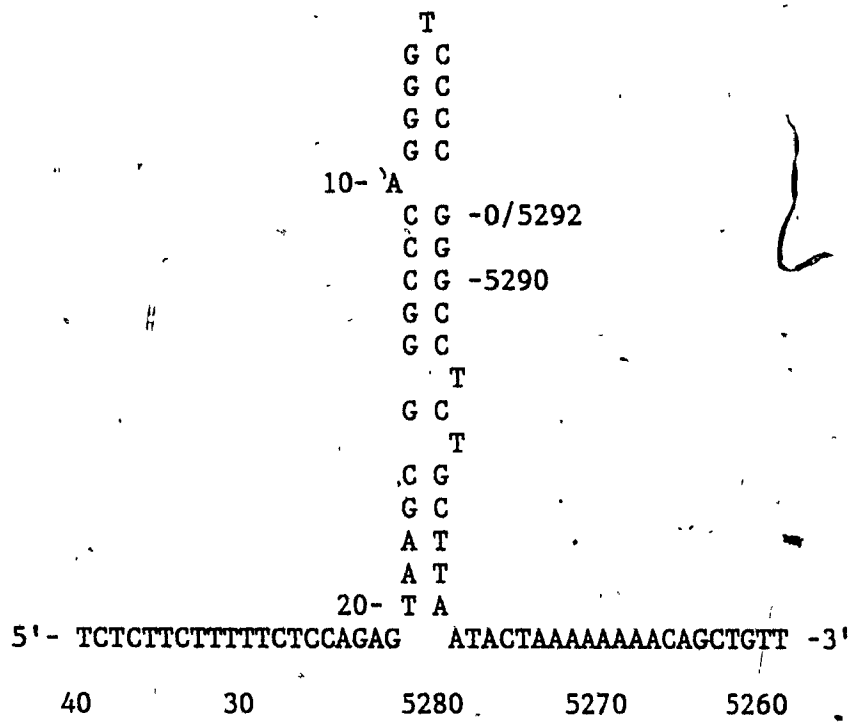


Figure 33: Nucleotide sequence of polyomavirus L DNA strand between nucleotides 5260 and 40.

This region is part of DNA replication origin (nucleotides 5265-90; Muller et al, 1983). A putative 32 base pair GC-rich palindrome can be drawn between nucleotides 21 and 5280.

DISCUSSION

RNA polymerase does not elongate RNA chains at a constant rate, but rather hesitates at specific sites *in vitro*. It is generally accepted that a transient pause is a prerequisite for termination *in vivo*. Termination of transcription is usually thought of as three separate events: 1) cessation of movement of RNA polymerase and of RNA chain elongation; 2) release of the completed RNA chain from the polymerase; 3) dissociation of RNA polymerase from the DNA template.

Analysis of 2 min pulse-labeled RNA with our S1 nuclease mapping method revealed a series of minor hybrid bands. Most of these hybrid bands correspond to the 3' ends of nascent RNAs and therefore represent termination and/or pause sites. This conclusion is based on the following observations: 1) no processing was detected in 2 min pulse-labeled nuclear RNA, suggesting that RNA labeled for this period of time represents nascent RNA; 2) the minor hybrid bands observed are not generated by ribonuclease or by S1 nuclease cleavage during our hybridization procedure since they were not detected in hybrids with RNA synthesized *in vitro*; 3) the experiment with the PstI-SacI subfragment of the PstI fragment 1L (Fig. 30) revealed no difference in the size of most hybrid bands, showing that they correspond to RNAs with 3' ends within the fragment.

It is impossible to determine the exact sizes of the hybrids by measuring their migration on a gel because DNA-RNA hybrids do not seem to migrate solely according to size. For example, the full-length hybrid between RNA and PstI fragment 3L (912 base pairs) migrated more slowly than the hybrid with the PstI fragment 2L (947 base pairs) (Figures 26 and 27). Apparent formation of secondary structure also can affect the migration of hybrids on a gel, as shown for the full-length hybrid between RNA and PstI

fragment 5L. In general, the best method for determining the size of DNA or RNA is to carry out electrophoresis on a denaturing gel. However, the RNA part of DNA-RNA hybrids is very sensitive to ribonucleases; nicking of the RNA would create artificial bands on the denaturing gel, and would further complicate the pattern. Therefore the most satisfactory method, for our purpose, was to determine the size of these DNA-RNA hybrids on a nondenaturing gel.

RNA labeled in vitro in isolated nuclei was analyzed in the reports on transcription of the mouse β-globin gene (Citron et al, 1984), the mouse α-amylase gene (Hagenbuchle et al, 1984), and the ovalbumin gene (LeMeur et al, 1984). This method of labeling offers several advantages: 1) the rate of elongation by RNA polymerase II is better controlled, e.g., in a 10 min labeling period, under certain conditions, the polymerase elongated nascent chains only about 500 nucleotides (Weber et al, 1977); 2) ^{32}P isotope can be used for labeling the nascent 3' ends of the RNA, allowing one to obtain a stronger signal and to use less RNA in each experiment. However, a major disadvantage of the technique concerns the interpretation of the results themselves. The multiple bands these authors observed may represent RNA polymerase pause sites created by the in vitro conditions and not sites of transcription termination. They may also represent extended pause sites downstream from sites where RNA polymerases paused prior to termination.

In vivo labeling of RNA also has its drawbacks: 1) It is difficult to control the rate of elongation by RNA polymerase under in vivo labeling conditions. 2) ^{3}H-uridine must be used for labeling, which means that a greater amount of labeled RNA is required in each experiment because: a) the maximum specific activity of ^{3}H-uridine is lower than that of ^{32}P labeled NTPs; b) the exogenously added uridine is diluted by internal UTP pools. 3) ^{3}H emits very weak β-particles; fluorography has to be used to

convert the energy of the β-particle to visible light, therefore fluorography has to be carried out for a much longer period. However, the in vivo labeling technique offers a unique advantage in that cells are maintained in the natural physical state during the labeling period. Therefore, one can be sure that 3' ends found on the pulse-labeled RNA represent the sites at which RNA polymerase II pauses or terminates transcription of the DNA template.

It is still not clear what causes RNA polymerase to pause. There is evidence that secondary structure such as a stem-loop structure on either the DNA or the RNA of the E. coli tryptophan operon causes RNA polymerase to pause and eventually release the RNA from the transcription complexes (Adhya et al, 1979; Oxender et al, 1979; Farnham and Platt, 1981). However, studies with several different templates have shown that pausing sites also occur in locations where RNA secondary structure is weak or absent (Kingston and Chamberlin, 1981; Lau et al, 1983; Morgan et al, 1983). In a cell-free eukaryotic system, it has been shown that RNA polymerase II pauses at the bacteriophage T7 (Kadesch and Chamberlin, 1982) and lambda 4S RNA terminator (Hatfield et al, 1983).

It is known that polyomavirus large T antigen binds specifically to multiple sites within the regulatory region on PstI fragment 5. Immuno-precipitation of protein-DNA complexes revealed two binding sites in the region upstream of early start sites for transcription. One is located between nucleotides 94 and 152, the other is located between nucleotides 5292 and 90 (Pomerantz et al, 1983). Using purified large T antigen, Cowie and Kamen (1984) revealed three closely spaced binding sites A (between nucleotides 25 and 75), B (between nucleotides 86 and 120) and C (between nucleotides 124 and 163). Two GC-rich palindromes are located within this region; one lies between nucleotides 84 and 100 (see Fig. 32), the other is

at the replication origin, between nucleotides 5280 and 21 (see Fig. 33). The four prominent minor hybrid bands located within this PstI fragment 5 would correspond to RNAs whose 3' ends are at nucleotides 5, 50, 80 and 140. All of them are located within the large T antigen binding region. This suggests that the binding of large T antigen to this region, together with the GC-rich palindrome, may cause RNA polymerase II to pause at these sites and may subsequently lead to termination of transcription.

In addition, in our laboratory, McNally (1983) has shown that there is a decrease in hybridization of in vivo pulse-labeled nuclear RNA to sequences downstream of the polyadenylation site, on the polyomavirus L DNA strand; she found an excess of hybridization to the fragment PstI-5L which is located upstream of the initiation sites for L strand RNA (see Figure 4). The same results were obtained by Tessier (Skarnes, Tessier, and Acheson, manuscript in preparation) using a solution hybridization technique, or with RNA labeled in vitro in isolated nuclei or on viral transcription complexes (Skarnes, 1985). Taken together, these results suggest that there is an excess of RNA polymerases stalling on fragment PstI-5L, implying the presence of strong pause sites within this fragment. In the light of these results, it is most likely that the 3' ends located at about nucleotides 5, 50, 80 and 140 within fragment PstI-5L correspond to those strong pausing sites.

In conclusion, it is difficult to distinguish between pausing and termination. We have shown that some of the 3' ends detected in my experiments correspond to pausing sites (above). Since we have demonstrated that about 50% of RNA polymerase molecules terminate transcription per traverse of the polyomavirus genome (Section II), it is reasonable to believe that most of the other 3' ends we detected represent the sites at which RNA polymerase II terminate transcription of the polyomavirus genome.

CONCLUSIONS

In prokaryotes, transcription termination plays an important role in gene expression, controlling the type and amount of genes to be expressed. It has been studied very extensively on E. coli, and two kinds of terminators have been found. One is the factor-independent terminator, which requires RNA polymerase only in vitro. It consists of a GC-rich segment of dyad symmetry and a stretch of dA residues in the template strand, which are transcribed as U residues in the RNA transcript. Termination occurs via formation of a stem-loop structure on the transcript, which causes RNA polymerase to pause and interrupts the RNA-DNA hybrid complex. The consecutive uridine residues then facilitate dissociation of the transcript from the template. The other terminators require a protein factor, such as rho, to function efficiently. Comparison of rho-dependent terminators reveal very little sequence homology, if any, between them. It is not known what recognition elements govern the behavior of rho-dependent termination sites.

In contrast, due to the complexity of the eukaryotic system, there is still very little known about termination, despite considerable efforts. However, it is recognized that termination also plays an important role in the control of eukaryotic gene expression, especially within gene clusters (Mather et al, 1984; Proudfoot, 1986; Grummt et al, 1986; Labhart and Reeder, 1986). Termination by RNA polymerase III and I are the most well-studied. For example, it is known that termination by RNA polymerase III requires only the polymerase itself. It can recognize the stretch of thymidine residues on the noncoding DNA strand and terminate transcription. RNA polymerase I requires protein factors which bind to specific sequences. Very little is known about termination of transcription by RNA polymerase

II despite a great deal of effort. It is known however that transcription extends beyond the polyadenylation site and terminates hundreds or thousands of nucleotides downstream of the poly (A) site. No control sequences or factors have yet been identified.

In this study, I have modified the S1 nuclease mapping method to allow me to map the 5' and 3' ends of in vivo pulse-labeled nuclear RNAs which represent mostly unprocessed RNA precurssors. This method can have a wide range of applications. For example, by pulse-labeling for different time periods, one can determine the kinetics of RNA processing, polyadenylation and also splicing, or determine which of these two processes occurs first; if there are more than one splicing sites which site is preferentially used.

Using this method, I have identified an initiation site of polyomavirus late RNA at nucleotide 5076 which is also a 5' splicing site of the late leader. The other initiation sites I mapped correspond to those reported by other investigators (Treisman et al, 1980; Treisman and Kamen, 1982; Heiser and Eckhart, 1982). I have also found a shift of the major initiation site from nucleotide 5050 at 16 hr post infection to a site further upstream at nucleotide 5128, at 28 hr post infection. Further investigation is required to understand the meaning of this change.

Calculation of the termination efficiency of polyomavirus L-strand RNA has revealed that termination is more efficient during the early part (16 hr post infection) than the late part (28 hr post infection) of the late phase. Perhaps a factor required for termination is present in limiting amounts during the late phase; it is also possible that some late gene product acts as an antiterminator, acting in a way similar to the N and Q proteins of bacteriophage λ, for example.

Mapping the 3' ends of pulse-labeled polyomavirus late RNAs has

revealed multiple stop sites instead of a single termination site. Similar results have been found for other eukaryotic mRNA genes, for example the mouse β-major globin gene (Citron et al, 1984), the mouse α-amylase gene (Hagenbuchle et al, 1984), and the chicken ovalbumin gene (LeMur et al, 1984). We have reason to believe that some of these sites represent pause sites and some termination sites. However, our method does not allow us to distinguish between pause sites and termination sites. Furthermore, the precise location of these sites cannot be accurately defined, due to the limitation of our gel system and possible secondary structure within the hybrids themselves. Therefore the sequences surrounding them cannot be determined.

Both termination and polyadenylation are inefficient during the late phase of polyomavirus lytic infection. This suggests that these two processes may be linked. The same observation is made when the termination region of the mouse β-globin is inserted into the E1A transcription unit of adenovirus 5 (Falck-Pedersen, 1985). Same polyadenylation factors may give a signal to the polymerase to stop transcription either by itself or with the help of other factors. In contrast, in the case of the adenovirus major late transcription unit, all polymerases pass five polyadenylaton sites and terminate within the terminal 650 nucleotides without reaching the end of the genome (Nevins and Darnell, 1978; Fraser et al, 1979; Fraser and Hsu, 1980). Therefore, there could be more than one type of mechanism in eukaryotic mRNA transcription termination.

It was generally believed that common function required common sequence, and that the key to understanding gene expression and regulation lies simply in determining the nucleic acid sequence. We have very little information about the nature of the signals. To carry out the function of such signals, all sorts of higher level structures of both nucleic acids

and proteins might be required, not just a simple DNA-protein interaction. Also, many different factors may be involved. The nature of transcription termination is far more complicated than we first believed, even in the simplest case. A reliable *in vitro* transcription system is needed to answer many of the unanswered questions.

LITERATURE CITED

Acheson, N.H. 1976. Transcription during productive infection with polyoma virus and simian virus 40. Cell 8:1-12.

Acheson, N.H. 1978. Polyoma virus giant RNAs contain tandem repeats of the nucleotide sequece of the entire viral genome. Proc. Natl. Acad. Sci. (USA). 75:4745-4748.

Acheson, N.H. 1981a. Lytic cycle of SV 40 and polyoma virus. In DNA Tumor Viruses, J. Tooze ed., (Cold Spring Harbor Laboratory, Cold Spring Harbor, New York), pp. 125-203.

Acheson, N.H. 1981b. Efficiency of processing of viral RNA during the early and late phases of productive infection by polyoma virus. J. Virol. 37:628-635.

Acheson, N.H. 1984. Kinetics and efficiency of polyadenylation of late polyomavirus nuclear RNA: generation of oligomeric polyadenylated RNAs and their processing into mRNA. Mol. Cell. Biol. 4:722-729.

Acheson, N.H., E. Buetti, K. Scherrer, and R. Weil. 1971. Transcription of the polyoma virus genome: synthesis and cleavage of giant late polyoma-specific RNA. Proc. Natl. Acad. Sci. (USA). 68:2231-2235.

Acheson, N.H., and F. Mieville. 1978. Extent of transcription of the E-strand of polyoma virus DNA during the early phase of productive infection. J. Virol. 28:885-894.

Adhya, S., and M. Gottesman. 1978. Control of transcription termination. Ann. Rev. Biochem. 47:967-996.

Adhya, S., P. Sarkar, D. Valenzuela, and U. Maitra. 1979. Termination of transcription by Escherichia coli RNA polymerase: influence of secondary structure of RNA transcripts on rho-independent and rho-dependent termination. Proc. Natl. Acad. Sci. (USA). 76:1631-1617.

Anderson, D.M., and L.D. Smith. 1977. Synthesis of heterogeneous nuclear RNA in full-grown oocytes of Xenopus laevis (Daudin). Cell 11:663-671.

Arrand, J.R., E. Soeda, J. Walsh, N. Smolar, and B.E. Griffin. 1980. Polyoma virus DNA: sequece from the late region that specifies the leader sequence for late mRNA and codes for VP2, VP3, and the N-terminus of VP1. J. Virol. 33:606-618.

Aviv, H., and P. Leder. 1972. Purification of biological active globin messenger RNA by chromatography on oligothymidylic acid-cellulose. Proc. Natl. Acad. Sci. (USA). 69:1408-1412.

Bakken, A., G. Morgan, B. Sollner-Webb, J. Roan, S. Busby, and R. H. Reeder. 1982. Mapping of transcription initiation and termination signals on Xenopus laevis ribosomal DNA. Proc. Natl. Acad. Sci. (USA). 79:56-60.

Beard, P., N.H. Acheson, and I.H. Maxwell. 1976. Strand-specific transcription of polyoma virus DNA early in productive infection and in transformed cells. J. Virol. 17:20-26.

Beard, P., J.F. Morrow, and P. Berg. 1973. Cleavage of circular, superhelical SV40 DNA to a linear duplex by S1 nuclease. J. Virol. 12:1303-1313.

Benjamin, T.L. 1966. Virus-specific RNA in cells productively infected or transformed by polyoma virus. J. Mol. Biol. 16:359-373.

Bertrand, K.P., L.J. Korn, F. Lee, and C. Yanofsky. 1977. The attenuator of the tryptophan operon of Escherichia coli: heterogeneous 3'-OH termini in vivo and deletion mapping of functions. J. Mol. Biol. 117:227-247.

Berk, A.J., and P.A. Sharp. 1977. Sizing and mapping of early adenovirus mRNAs by gel electrophoresis of S1 endonuclease- digested hybrids. Cell 12:721-732.

Birchmeier, C., W. Folk, and M.L. Birnstiel. 1983. The terminal RNA stem-loop structure and 80 bp of spacer DNA are required for the formation of 3' termini of sea urchin H2A mRNA. Cell 35:433-440.

Birchmeier, C., R. Grosschedl, and M.L. Birnstiel. 1982. Generation of authentic 3' termini of an H2A mRNA in vivo is dependent on a short inverted DNA repeat and on spacer sequences. Cell 28:739-745.

Birchmeier, C., D. Schumperli, G. Sconzo, and M.L. Birnstiel. 1984. 3' editing of mRNAs: sequence requirements and involvement of a 60-nucleotide RNA in maturation of histone mRNA precursors. Proc. Natl. Acad. Sci. (USA). 81: 1057-1061.

Birg, F., J. Favaloro, and R. Kamen. 1977. Analysis of polyoma virus nuclear RNA by mini-blot hybridization. Proc. Natl. Acad. Sci. (USA). 74:3138-3142.

Birkenmeier, E.H., D.D. Brown, and E. Jordan. 1978. A nuclear extract of Xenopus laevis oocytes that accurately transcribes 5S RNA genes. Cell 15:1077-1086.

Blanchard, J.M., J. Weber, W. Jelinek, and J.E. Darnell, Jr. 1978. In vitro RNA-RNA splicing in adenovirus 2 mRNA formation. Proc. Natl. Acad. Sci. (USA). 75:5344-5348.

Bogenhagen, D.F., and D.D. Brown. 1981. Nucleotide sequences in Xenopus 5S DNA required for transcription termination. Cell 24:261-270.

Bonner, J., G. Kung, and Bekhov. 1967. A method for the hybridization of nucleic acid molecules at low temperature. Biochemistry 6:3650-3653.

Bowman, L. H., W.E. Goldman, G.I. Goldberg, M. B. Hebert, and D. Schlessinger. 1983. Location of the initial cleavage sites in mouse pre-rRNA. Mol. Cell. Biol. 3:1501-1510.

Breathnach, R., and P. Chambon. 1981. Organization and expression of eukaryotic split genes coding for proteins. Ann. Rev. Biochem. 50:348-383.

Brown, D.D., and J.B. Gurdon. 1978. Cloned single repeating units of 5S DNA direct accurate transcription of 5S RNA when injected into Xenopus oocytes. Proc. Natl. Acad. Sci. (USA). 75:2849-2853.

Buetti, E. 1974. Characterization of late polyoma mRNA. J. Virol. 14:249-260.

Busslinger, M., R. Portmann, and M.L. Birnstiel. 1979. A regulatory sequence near the 3' end of sea urchin histone genes. Nucleic Acids Res. 6:1997-3008.

Carmichael, G.G., and G.K. McMaster. 1980. The analysis of nucleic acids in gel using glyoxal and acridine orange. Meth. Enzymol. 65:380-391.

Casey, J., and N. Davidson. 1977. Rates of formation and thermal stabilities of RNA:DNA and DNA:DNA duplexes at high concentrations of formamide. Nucleic Acids Res. 4:1539-1552.

Chamberlin, M. 1965. Comparative properties of DNA, RNA, and hybrid homopolymer pairs. Fed. Proc. 24:1446-1457.

Chambon, P. 1975. Eukaryotic nuclear RNA polymerases. Ann. Rev. Biochem. 44:613-638.

Chiu, N.H., M.F. Radonovich, M.M. Thoren, and N.P. Salzman. 1978. Selective degradation of newly-synthesized nonmessenger simian virus 40 transcripts. J. Virol. 28:590-599.

Christie, G.E., P.J. Farnham, and T. Platt. 1981. Rho-independent transcription termination: comparison of natural and synthetic terminator sites in vitro. Proc. Natl. Acad. Sci. (USA). 78:4180-4184.

Citron, B., E. Falck-Pedersen, M. Salditt-Georgieff, and J.E. Darnell, Jr. 1984. Transcription termination occurs within a 1000 base pair region downstream from the poly(A) site of the mouse β-globin (major) gene. Nucleic Acids Res. 22:8723-8731.

Cogen, B. 1978. Virus-specific early RNA in 3T6 cells infected by a ts mutant of polyoma virus. Virology 85:222-230.

Court, D., C. Brady, M. Rosenberg, D.L. Wulff, M. Behr, M. Mahoney, and S. Izumi. 1980. Control of transcription termination: a rho-dependent termination site in bacteriophage lamda. J. Mol. Biol. 138:231-254.

Cowie, A., P. Jat, and R. Kamen. 1982. Determination of sequences at the capped 5' ends of polyoma virus early region transcripts synthesized in vivo and in vitro demonstrates an unusual microheterogeneity. J. Mol. Biol. 159:225-255.

Cowie, A., and R. Kamen. 1984. Multiple binding sites for polyomavirus large T antigen within regulatory sequences of polyomavirus DNA. J. Virol. 52:750-760.

Cowie, A., and R. Kamen. 1986. Guanine nucleotide contacts within viral DNA sequences bound by polyomavirus large T antigen. J. Virol. 57:505-514.

Cowie, A., C. Tyndall, and R. Kamen. 1981. Sequences at the capped 5'-ends of polyoma virus late region mRNAs: an example of extreme terminal heterogeneity. Nucleic Acids Res. 23:6305-6322.

Cox, K.H., D.V. DeLéon, L.M. Angerer, and R.C. Angerer. 1984. Detection of mRNAs in sea urchin embroys by in situ hybridization using asymmetric RNA probes. Dev. Biol. 101:485-502.

Cozzarelli, N.R., S.P. Gerrard, M. Schlissel, D.D. Brown, and D.F. Bogenhagen. 1983. Purified RNA polymerase III accurately and efficiently terminates transcription of 5S RNA genes. Cell 34:829-835.

Crawford, L.V., C. Syrett, and A. Wilde. 1973. The replication of polyoma DNA. J. Gen. Virol. 21:515-521.

Das, A., D. Court, and S. Adhya. 1976. Isolation and characterization of conditional lethal mutants of Escherichia coli defective in transcription termination factor rho. Proc. Natl. Acad. Sci. (USA). 73:1959-1963.

Das, A., C. Merril, and S. Adhya. 1978. Interaction of RNA polymerase and rho in transcription termination: coupled ATPase. Proc. Natl. Acad. Sci. (USA). 75:4828-4832.

de Crombrugghe, B., M. Mudryj, R. DiLauro, and M. Gottesman. 1979. Specificity of the bacteriophage lambda N gene product (pN): nut sequences are necessary and sufficient for antitermination by pN. Cell 18:1145-1151.

Deininger, P., A. Esty, P. LaPorte, and T. Friedmann. 1979. Nucleotide sequence and genetic organization of the polyoma late region: features common to the polyoma early region and SV40. Cell 18:771-779.

Deininger, P., A. Esty, P. LaPorte, H. Hsu, and T. Friedmann. 1980. The nucleotide sequence and restriction enzyme sites of the polyoma genome. Nucleic Acids Res. 8:855-860.

Derman, E., S. Goldberg, and J.E. Darnell, Jr. 1976. hnRNA in HeLa cells: distribubution of transcript sizes estimated from nascent molecule profile. Cell 9:465-472.

Diaz, M.O., G. Barsacchi-Pilone, K.A. Mahon, and J.G. Gall. 1981. Transcripts from both strands of a satellite DNA occur on lampbrush chromosome loops of the newt Notophthalamus. Cell 24:649-659.

Dulbecco, R., and M. Vogt. 1963. Evidence for a ring structure of polyoma virus DNA. Proe. Natl. Acad. Sci. (USA). 50:236-242.

Egyhazi, E. 1974. A tentative initiation inhibition of chromosomal heterogeneous RNA synthesis. J. Mol. Biol. 84:173-183.

Fahrner, K., J. Yarger, and L. Hereford. 1980. Yeast histone mRNA is polyadenylated. Nucleic Acids Res. 8:5725-5737.

Falck-Pedersen, E., J. Logan, T. Shenk, and J.E. Darnell, Jr. 1985. Transcription termination within the E1A gene of adenovirus induced by insertion of the mouse β-major globin terminator element. Cell 40:897-905.

Favaloro, J., R. Treisman, and R. Kamen. 1980. Transcription maps of polyoma virus-specific RNA: analysis by two-dimensional nuclease S1 gel mapping. Meth. Enzymol. 65:718-749.

Farnham, P.J., and T. Platt. 1980. A model for transcription termination suggested by studies on the trp attenuator in vitro using base analogs. Cell 20:739-748.

Farnham, P.J., and T. Platt. 1981. Rho-independent termination dyad symmetry in DNA causes RNA polymerase to pause during transcription in vitro. Nucleic Acids Res. 9:563-577.

Farnham, P.J., and T. Platt. 1982. Effects of DNA base analogs on transcription termination of the tryptophan operon attenuator of Escherichia coli. Proc. Natl. Acad. Sci. (USA). 79:998-1002.

Fedoroff, N.V., and D.D. Brown. 1978. The nucleotide sequence of oocyte 5S DNA in Xenopus laevis. I. The AT-rich spacer. Cell 13:701-716.

Fenner, F., B.R. McAuslan, C.A. Mims, J. Sambrook, and D.O. White. 1974. The Biology of Animal Viruses. (New York and London, Academic Press, second edition).

Fenton, R.G., and C. Basilico. 1982. Changes in the topography of early region transcription during polyoma virus lytic infection. Proc. Natl. Acad. Sci. (USA). 79:7142-7146.

Fey, G., and B. Hirt. 1975. Fingerprints of polyoma virus proteins and mouse histones. Cold Spring Harbor Symp. Quant. Biol. 39:235-242.

Fitzgerald, M., and T. Shenk. 1981. The sequence 5'-AAUAAA-3' forms part of the recognition site for polyadenylation of late SV40 mRNA. Cell 24:251-260.

Flavell, A.J., A. Cowie, S. Legon, and R. Kamen. 1979. Multiple 5' terminal cap structures in late polyoma virus RNA. Cell 16:357-371.

Ford, J. P., and M. T. Hsu. 1978. Transcription pattern of in vivo-labeled late simian virus 40 RNA: equimolar transcription beyond the mRNA 3' terminus. J. Virol. 28:795-801.

Franke, B., and W. Eckhart. 1973. Polyoma gene function required for viral DNA synthesis. Virology 55:127-135.

Fraser, N.W., and M.T. Hsu. 1980. Mapping of the 3' terminus of the large late AD-2 transcript by electron microscopy. Virology 103:514-516.

Fraser, N.W., J.R. Nevins, E. Ziff, and J.E. Darnell, Jr. 1979. The major late adenovirus type 2 transcription unit: termination is downstream from the last poly(A) site. J. Mol. Biol. 129:643-656.

Fraser, N.W., P.B. Sehgal, and J.E. Darnell, Jr. 1978. DRB-induces premature termination of late adenovirus transcription. Nature 272:590-593.

Frayne, E.G., E.J. Leys, G.F. Crouse, A.G. Hook, and R.E. Kellems. 1984. Transcription of the mouse dihydrofolate reductase gene proceeds unabated through seven polyadenylation sites and terminates near a region of repeated DNA. Mol. Cell. Biol. 4:2921-2924.

Freason, P.M., and L.V. Crawford. 1972. Polyoma virus basic proteins. J. Gen. Virol. 14:141-155.

Friedman, D.I., and E.R. Olson. 1983. Evidence that a nucleotide sequence, "box A", is involved in the action of the Nus A protein. Cell 34:143-149.

Friedman, D.I., A.T. Schauer, M.R. Baumann, L.S. Baron, and S. L. Adhya. 1981. Evidence that ribosomal protein S10 participates in control of transcription termination. Proc. Natl. Acad. Sci. (USA). 78:1115-1118.

Friedman, D.I., A.T. Schauer, E.J. Mashni, E.R. Olson, and M.R. Baumann. 1983. E. coli factors involved in the action of the lambda gene N antitermination function. In Microbiology-1983, D. Schlessinger ed., (American Society for Microbiology, Washington, D. C.), pp. 39-42.

Friedmann, T., A. Esty, P. LaPorte, and P. Deininger. 1979. The nucleotide sequence and genome organization of the polyoma early region: extensive nucleotide and amino acid homology with SV40. Cell 17:715-724.

Galli, G., H. Hofstetter, H.G. Stunnenberg, and M.L. Birnstiel. 1983. Biochemical complementation with RNA in the Xenopus oocyte: a small RNA is required for the generation of 3' histone mRNA termini. Cell 34:823-828.

Galluppi, G., C. Lowery, and J.P. Richardson. 1976. Nucleoside triphosphate requirement for termination of RNA synthesis by rho factor. In RNA Polymerase, R. Losick and M. Chamberlin eds. (Cold Spring Harbor Laboratory, Cold Spring Harbor, New York), pp. 657-665.

Galluppi, G., J.P. Richardson. 1980. ATP-induced changes in the binding of RNA synthesis termination protein rho to RNA. J. Mol. Biol. 138:513-539.

Gannon, F., K. O'Hare, F. Perrin, J.P. LePennec, C. Benoist, M. Cochet, R. Breathnach, A. Royal, A. Garapin, B. Cami, and P. Chambon. 1979. Organisation and sequences at the 5' end of a cloned complete ovalbumin gene. Nature 278:428-434.

Gaudray, P., C. Tyndall, R. Kamen, and F. Cuzin. 1981. The high-affinity binding site on polyoma virus DNA for the viral large T protein. Nucleic Acids Res. 9:5697-5710.

Georgiev, O., and M.L. Birnstiel. 1985. The conserved CAAGAAAGA spacer sequence is an essential element for the formation of 3' termini of the sea urchin H3 histone mRNA by RNA processing. EMBO J. 4:481-489.

Germond, J.E., V.M. Vogt, and B. Hirt. 1974. Characterization of the single-strand-specific nuclease S1 activity on double-stranded supercoiled polyoma DNA. Eur. J. Biochem. 43:591-600.

Ghosh, P.K., and P. Lebowitz. 1981. Simian virus 40 early mRNA's contain multiple 5' termini upstream from a Hogness-Goldberg sequence; a shift in 5' termini during the lytic cycle is mediated by large T antigen. J. Virol. 40:224-240.

Gillespie, D., and S. Spiegelman. 1965. A quantitative assay for DNA-RNA hybrids with DNA immobilized on a membrane. J. Mol. Biol. 12:829-842.

Granboulan, N., P. Tounier, R. Wicker, and W. Bernhard. 1963. An electron-microscopic study of the development of SV40 virus. J. Cell Biol. 17:423-434.

Gray, D.M., J.J. Liu, R.L. Ratliff, and F.S. Allen. 1981. Sequence dependence of the circular dichroism of synthetic double-stranded RNAs. Biopolymers 20:1337-1382.

Grayhack, E.J., and J.W. Roberts. 1982. The phage lambda Q gene product: activity of a transcription antiterminator in vitro. Cell 30:637-648.

Greenberg, H., and S. Penman. 1966. Methylation and processing of ribosomal RNA in HeLa cells. J. Mol. Biol. 21:527-535.

Greenblatt, J. 1981. Regulation of transcription termination by the N gene protein of bacteriophage lambda. Cell 24:8-9.

Greenblatt, J., and J. Li. 1981a. Interaction of the sigma factor and the nusA gene protein of E. coli with RNA polymerase in the initiation-termination cycle of transcription. Cell 24:421-428.

Greenblatt, J., and J. Li. 1981b. The nusA gene protein of E coli: its identification and a demonstration that it interacts with the gene N transcription antitermination protein of bacteriophage lambda. J. Mol. Biol. 147:11-23.

Griffin, B.E., M. Fried, and A. Cowie. 1974. Polyoma DNA: a physical map. Proc. Natl. Acad. Sci. (USA). 71:2077-2081.

Griffin, B.E., and C. Maddock. 1979. New classes of viable deletion mutants in the early region of polyoma virus. J. Virol. 47:645-656.

Griffin, B.E., E. Soeda, B.G. Barrell, and R. Staden. 1981. Sequence and analysis of polyoma virus DNA. In DNA Tumor Viruses, J. Tooze ed., (Cold Spring Harbor Laboratory, Cold Spring Harbor, New York), pp. 843-891.

Gross, L. 1953a. A filterable agent, recovered from AK leukemic extracts, causing salivary gland carcinomas in C3H mice. Proc. Soc. Exp. Biol. Med. 83:414-421.

Gross, L. 1953b. Neck tumors, or leukemia, developing in adult C3H mice following inoculation, in early infancy, with filtered (Berkefeld N), or centrifuged (144,000 x g), Ak-leukemic extracts. Cancer 6:948-955.

Grummt, I. 1981. Specific transcription of mouse ribosomal DNA in a cell-free system that mimics control in vivo. Proc. Natl. Acad. Sci. (USA). 78:727-731.

Grummt, I., U. Maier, A. Ohrlein, N. Hassouna, and J.-P. Bachellerie. 1985a. Transcription of mouse rDNA terminates downstream of the 3' end of 28S RNA and involves interaction of factors with repeated sequences in the 3' spacer. Cell 43:801-810.

Grummt, I., H. Sorbaz, A. Hofmann, and E. Roth. 1985b. Spacer sequences downstream of the 28S RNA coding region are part of the mouse rDNA transcription unit. Nucleic Acids Res. 13:2293-2303.

Grummt, I., A. Kuhn, I. Bartsch, and H. Rosenbauer. 1986a. A transcription terminator located upstream of the mouse rDNA initiation site affects rRNA synthesis. Cell 47:901-911.

Grummt, I., H. Rosenbauer, I. Niedermeyer, U. Maier, and A. Ohrlein. 1986b. A repeated 18 bp sequence motif in the mouse rDNA spacer mediates binding of a nuclear factor and transcription termination. Cell 45:837-846.

Guarente, L.P., and J. Beckwith. 1978. Mutant RNA polymerse of Escherichia coli terminates transcription in strains making defective rho factor. Proc. Natl. Acad. Sci. (USA). 75:294-297.

Guarente, L.P., J. Beckwith, A.M. Wu, and T. Platt. 1979. A mutation distal to the messenger RNA endpoint reduces transcription terrmination in the tryptophan operon in Escherichia coli. J. Mol. Biol. 133:189-197.

Hagenbuchle, O., P.K. Wellauer, D.L. Cribbs, and U. Schibler. 1984. Termination of transcription in the mouse alpha-amylase gene Amy-2^a occurs at multiple sites downstream of the polyadenylation site. Cell 38:737-744.

Hallick, R.B., B.K. Shelm, P.W. Gray, and E.M. Orozco, Jr. 1977. Use of aurintricarboxylic acid as an inhibitor of nucleases during nucleic acid isolation. Nucleic Acids Res. 4:3055-3064.

Hansen, U., D.G. Tenen, D.M. Livingston, and P.A. Sharp. 1981. T antigen repression of SV40 early transcription from two promotors. Cell 27:603-612.

Hatfield, G.W., J.A. Sharp, and M. Rosenberg. 1983. Pausing and termination of human RNA polymerase II transcription at a procaryotic terminator. Mol. Cell. Biol. 3:1687-1693.

Heiser, W.C., and W. Eckhart. 1982. Polyoma virus early and late mRNA in productively infected mouse 3T6 cells. J. Virol. 44:175-18[illegible].

Henderson, S., and B. Sollner-Webb. 1986. A transcriptional terminator is a novel element of the promoter of the mouse ribosomal RNA gene. Cell 47;891-900.

Henikoff, S., and E. H. Cohen. 1984. Sequences responsible for transcription termination on a gene segment in Saccharomyces cerevisiae. Mol. Cell. Biol. 4:1515-1520.

Henikoff, S., J.D. Kelly, and E.H. Cohen. 1983. Transcription terminates in yeast distal to a control sequence. Cell 33:607-614.

Henikoff, S., K. Tatchell, B.D. Hall, and K.A. Nasmyth. 1981. Isolate of a gene from Drosophila by complementation in yeast. Nature 289:33-37.

Hentschel, C. C., and M. L. Birnstiel. 1981. The organization and expression of histone gene families. Cell 25:301-313.

Hentschel, C.C., J.C. Irminger, P. Bucher, and M.L. Birnstiel. 1980a. Sea urchin histone mRNA termini are located in gene regions downstream of putative regulatory sequences. Nature 285:147-151.

Hentschel, C.C., E. Probst, and M.L. Birnstiel. 1980b. Transcription fidelity of histone genes injected into Xenopus oocyte nuclei. Nature 288:100-102.

Herrmann, R., K. Neugebauer, E. Pirkl, H. Zentgraf, and H. Schaller. 1980. Conversion of bacteriophage fd into an efficient single-stranded DNA vector system. Mol. Gen. Genet. 177:231-242.

Hirt, B. 1967. Selective extraction of polyoma virus DNA from mouse cell culture. J. Mol. Biol. 26:365-369.

Hirt, B. 1969. Replicating molecules of polyoma virus DNA. J. Mol. Biol. 40:141-144.

Hofer, E., and J.E. Darnell, Jr. 1981. The primary transcription unit of the mouse β-major globin gene. Cell 23:585-593.

Hofer, E., R. Hofer-Warbinek, and J.E. Darnell, Jr. 1982. Globin RNA transcription: a possible termination site and demonstration of transcriptional control correlated with altered chromatin structure. Cell 29:887-893.

Holmes, W.M., T. Platt, and M. Rosenberg. 1983. Termination of transcription in E. coli. Cell 32:1029-1032.

Howard, B.H., and B. de Crombrugghe. 1976. ATPase activity required for termination of transcription by the E. coli protein factor rho. J. Biol. Chem. 251:2520-2524.

Howard, B.H., B. de Crombrugghe, and M. Rosenberg. 1977. Transcription in vitro of bacteriophage lamda 4S RNA: studies on termination and rho protein. Nucleic Acids Res. 4:927-939.

Howley, P.M., M.A. Isreal, M. Law, and M.A. Martin. 1979. A rapid method for detecting and mapping homology between heterologous DNAs. J. Biol. Chem. 254:4876-4883.

Jat, P., J. Roberts, A. Cowie, and R. Kamen. 1982. Comparison of the polyoma virus early and late promotors by transcrip- tion in vitro. Nucleic Acids Res. 10:871-887.

Kadesch, T.R., and M.J. Chamberlin. 1982. Studies of in vitro transcription by calf thymus RNA polymerase II using a novel duplex DNA template. J. Biol. Chem. 257:5286-5295.

Kallenbach, N.R. 1968. Theory of thermal transitions in low molecular weight RNA chains. J. Mol. Biol. 37:445-466.

Kamen, R., J. Favaloro, and J. Parker. 1980. Topography of the three late mRNA of polyoma virus which encode the virion proteins. J. Virol. 33:637-651.

Kamen, R., P. Jat, R. Treisman, and J. Favaloro. 1982. 5' termini of polyoma virus early region transcripts synthesized in vivo by wild-type virus and viable deletion mutants. J. Mol. Biol. 159:189-224.

Kamen, R., D.M. Lindstrom, H. Shure, and R.W. Old. 1974. Virus-specific RNA in cells productively infected or transformed by polyoma virus. Cold Spring Harbor Symp. Quant. Biol. 39:187-198.

Kamen, R., and H. Shure. 1976. Topology of polyoma virus messenger RNA molecules. Cell 7:361-371.

Kingston, R.E., and M.J. Chamberlin. 1981. Pausing and attenuation of in vitro transcription in the rrnB operon of E. coli. Cell 27:523-531.

Kolter, R., and C. Yanofsky. 1982. Attenuation in amino acid biosynthetic operons. Ann. Rev. Genet. 16:113-134.

Kominami, R., Y. Mishima, Y. Urano, M. Sakai, and M. Muramatsu. 1982. Cloning and determination of the transcription termination site of ribosomal RNA gene of the mouse. Nucleic Acids Res. 10:1963-1979.

Korn, L.J., and D.D. Brown. 1978. Nucleotide sequnce of Xenopus borealis oocyte 5S DNA: comparison of sequences that flank related eukaryotic genes. Cell 15:1145-1156.

Labhart, P., and R.H. Reeder. 1986. Characterization of three sites of RNA 3' end formation in the Xenopus ribosomal gene spacer. Cell 45:431-443.

Lai, C.J., R. Dhar, and G. Khoury. 1978. Mapping the spliced and unspliced late lytic SV40 RNAs. Cell 14:971-982.

Lanoix, J., R.W. Tseng, and N.H. Acheson. 1986. Duplication of functional polyadenylation signals in polyomavirus DNA does not alter the efficiency of polyadenylation or transcription termination. J. Virol. 58:733-742.

Lau, L.F., J.W. Roberts, and R. Wu. 1983. RNA polymerase pausing and transcript release at he lambda tR_1 terminator in vitro. J. Biol. Chem. 258:9391-9397.

Laub, O., E.B. Jakobovits, and Y. Aloni. 1980. 5,6-dichloro-1-β-D-ribofluranosylbenzimidazole enhances premature termination of late transcription of simian virus 40 DNA. Proc. Natl. Acad. Sci. (USA). 77:3297-3301.

Lee, F., K. Bertrand, G. Bennett, and C. Yanofsky. 1978. Comparison of the nucleotide sequences of the initial transcribed regions of the tryptophan operons of Escherichia coli and Salmonella typhimurium. J. Mol. Biol. 121:193-217.

Legon, S., A.J. Flavell, A. Cowie, and R. Kamen. 1979. Amplification in the leader sequence of late polyoma virus mRNAs. Cell 16:373-388.

LeMeur, M.A., B. Galliot, and P. Gerlinger. 1984. Termination of the ovalbumin gene transcription. EMBO J. 3:2779-2786.

Logan, J., and T. Shenk. 1984. Adenovirus tripartite leader sequence enhances translation of mRNAs late after infection. Proc. Natl. Acad. Sci. (USA). 81:3655-3659.

Long, E.O., and I.B. Dawid. 1980. Repeated genes in eucaryotes. Ann. Rev. Biochem. 49:727-764.

Lowery, C., and J.P. Richardson. 1977a. Characterization of the nucleotide triphosphate phosphohydrolase (ATPase) activity of RNA synthesis termination factor rho. I. Enzymatic properties and effects of inhibitors. J. Biol. Chem. 252:1375-1380.

Lowery, C., and J.P. Richardson. 1977b. Characterization of the nucleotide triphosphate phosphohydrolase (ATPase) activity of RNA synthesis termination factor rho. II. Influence of synthetic RNA homopolymers and random copolymers on the reaction. J. Biol. Chem. 252:1380-1385.

Maderions, A., and S. Chen-Kiang. 1984. Pausing and premature termination of human RNA polymerase II during transcription of adenovirus in vivo and in vitro. Proc. Natl. Acad. Sci. (USA). 81:5931-5935.

Mandal, R.K., and I.B. Dawid. 1981. The nucleotide sequence at the transcription termination site of ribosomal RNA in Drosophila melanogaster. Nucleic Acids Res. 9:1801-1811.

Manley, J.L., A. Fire, A. Cano, P.A. Sharp, and M.L. Gefter. 1980. DNA-dependent transcription of adenovirus genes in a whole cell extract. Proc. Natl. Acad. Sci. (USA). 77:3855-3859.

Martin, F.H., and I. Tinoco, Jr. 1980. DNA-RNA hybrid duplexes containing oligo (dA:rU) sequences are exceptionally unstable and may facilitate termination of transcription. Nucleic Acids Res. 8:2295-2350.

Mather, E.L., K.J. Nelson, J. Haimovich, and R.P. Perry. 1984. Mode of regulation of immunoglobulin μ- and Δ-chain expression varies during B-lymphocyte maturation. Cell 36:329-338.

Mattern, C.F., K.K. Takemoto, and W.A. Daniel. 1966. Replication of polyoma virus in mouse embryo cells: electron microscopic observations. Virology 30:242-256.

McConaughy, B.L., C.D. Laird, and B.J. McCarthy. 1969. Nucleic acid reassociation in formamide. Biochemistry 8:3289-3295.

McNally, M. 1984. Termination of transcription on the DNA of polyomavirus during the late phase of lytic infection. M. Sc. Thesis, McGill University, Montreal, Quebec.

McStay, B., and R.H. Reeder. 1986. A termination site for Xenopus RNA polymerase I also acts as an element of an adjacent promotor. Cell 47:913-920.

Melton, D.A., P.A. Krieg, M.R. Rebagliati, T. Maniatis, K. Zinn, and M.R. Green. 1984. Efficient in vitro synthesis of biologically active RNA and RNA hybridization probes from plasmids containing a bacteriophage SP6 promotor. Nucleic Acids Res. 12:7035-7056.

Messing, J. 1983. New M13 vectors for cloning. Meth. Enzymol. 101:20-78.

Miller, J.R., E.M. Cartwright, G.G. Brownlee, N.V. Fedoroff, and D.D. Brown. 1978. The nucleotide sequence of oocyte 5S DNA in Xenopus laevis. II. The GC-rich region. Cell 13:717-725.

Miller, K.G., and B. Sollner-Webb. 1981. Transcription of mouse rRNA polymerase I: in vitro and in vivo initiation and processing sites. Cell 27:165-174.

Miller, T.J., D.L. Stephens, and J.E. Mertz. 1982. Kinetics of accumulation and processing of simian virus 40 RNA in Xenopus laevis oocytes injected with simian virus 40 DNA. Mol. Cell. Biol. 2:1581-1594.

Montandon, P.E., and N.H. Acheson. 1982. Synthesis of prematurely terminated late transcripts of polyomavirus DNA is resistant to inhibition by 5,6-dichloro-1-beta-D-ribofluranosylbenzimidazole. J. Gen. Virol. 59:367-376.

Montell, C., E.F. Fisher, M.H. Caruthers, and A.J. Berk. 1983. Inhibition of RNA cleavage but not polyadenylation by a point mutation in mRNA 3' consensus sequence AAUAAA. Nature 305:600-629.

Morgan, W.D., D.G. Bear, and P.H. von Hippel. 1983. Rho-dependent termination of transcription. II. Kinetics of mRNA elongation during transcription from the bacteriophage lamda P_R promoter. J. Biol. Chem. 258:9565-9574.

Mott, J.E., J.L. Galloway, and T. Platt. 1985. Maturation of Escherichia coli tryptophan operon mRNA: evidence for 3' exonucleolytic processing after rho dependent termination. EMBO J. 4:1887-1891.

Mueller, C.R., A.M. Mes-Masson, M. Bouvier, and J.A. Hassell. 1984. Location of sequences in polyomavirus DNA that are required for early gene expression in vivo and in vitro. Mol. Cell. Biol. 4:2594-2609.

Muller, W.J., C.R. Mueller, A.M. Mes, and J.A. Hassell. 1983. Polyomavirus origin for DNA replication comprises multiple genetic elements. J. Virol. 47:586-599.

Neff, N.F., and M.J. Chamberlin. 1978. Termination of transcription by Escherichia coli RNA polymerase in vitro is affected by ribonucleoside triphosphate base analogs. J. Biol. Chem. 253:2455-2460.

Nevins, J.R., J.M. Blanchard, and J.E. Darnell, Jr. 1980. Transcription units of adenovirus type 2: termination of transcription beyong the poly(A) addition site in early regions 2 and 4. J. Mol. Biol. 144:377-386.

Nevins, J.R., and J.E. Darnell, Jr. 1978. Steps in the processing of Ad2 mRNA: poly(A)$^+$ nuclear sequences are conserved and poly(A) addition precedes splicing. Cell 15:1477-1493.

Nygaard, A.P., and B.D. Hall. 1963. A method for the detection of RNA-DNA complexes. Biochem. Biophy. Res. Comm. 12:98-104.

Oxender, D., G. Zurawski, and C. Yanofsky. 1979. Attenuation in the Escherichia coli tryptophan operon: role of RNA secondary structure involving the tryptophan codon region. Proc. Natl. Acad. Sci. (USA). 76:5524-5529.

Ozer, H.L. 1972. Synthesis and assembly of simian virus 40. I. Differential synthesis of intact virion and empty shells. J. Virol. 9:41-49.

Ozer, H.L, and P. Tegtmeyer. 1972. Synthesis and assembly of simian virus 40. II. Synthesis of the major capsid protein and its incorporation into viral particles. J. Virol. 9:52-60.

Padgett, R.A., S.F. Hardy, and P.A. Sharp. 1983. Splicing of adenovirus RNA in a cell-free transcription system. Proc. Natl. Acad. Sci. (USA). 80:5230-5234.

Penman, S. 1969. Preparation of purified nuclei and nucleoli from mammalian cells. In Fundamental Techniques in Virology, K. Habel and N. P. Salzman eds., (Academic Press Inc., New York), pp. 35-48.

Piper, P.W. 1979. Polyoma virus transcription early during productive infection of mouse 3T6 cells. J. Mol. Biol. 131:399-407.

Platt, T. 1981. Termination of transcription and its regulation in the tryptophan operon of E. coli. Cell 24:10-23.

Platt, T., and D.G. Bear. 1983. Role of RNA polymerase, rho factor, and ribosomes in transcription termination. In Gene Function in Prokaryotes, J. Beckwith, J. Davies, and J. Gallant eds., (Cold Spring Harbor Laboratory, Cold Spring Harbor, New York), pp. 123-161.

Pomerantz, B.J., and J.A. Hassell. 1984. Polyomavirus and simian virus 40 large T antigen binds to common DNA sequences. J. Virol. 49:925-937.

Pomerantz, B.J., C.R. Mueller, and J.A. Hassell. 1983. Polyomavirus large T antigen binds independently to multiple unique regions on the viral genome. J. Virol. 47:600-610.

Proudfoot, N.J., and G.G. Brownlee. 1976. 3' non-coding region sequences in eukaryotic mRNA. Nature 263:211-214.

Proudfoot, N.J. 1986. Transcription interference and termination between duplicated α-globin gene contructs suggests a novel mechanism for gene regulation. Nature 322:562-565.

Reeder, R.H., and R.G. Roeder. 1972. Ribosomal RNA synthesis in isolated nuclei. J. Mol. Biol. 67:433-442.

Richardson, J.P. 1982. Activation of rho protein ATPase requires simutaneous interaction of two kinds of nucleic acid binding sites. J. Biol. Chem. 257:5760-5766.

Richardson, J.P., C. Grimley, and C. Lowery. 1975. Transcription termination factor rho activity is altered in Escherichia coli with suA gene mutations. Proc. Natl. Acad. Sci. (USA). 72:1725-1729.

Richardson, J.P., and R. Conaway. 1980. RNA release activity of transcription termination protein rho is dependent on the hydrolysis of nucleotide triphosphate. Biochemistry 19:4293-4299.

Richardson, J.P., and M.R. Macy. 1981. RNA synthesis termination protein rho function: effects of conditions that destabilize RNA secondary structure. Biochemistry 20:1133-1139.

Riley, M., B. Maling, and M.J. Chamberlin. 1966. Physical and chemical characterization of two- and three-stranded adenine-thymine and adenine-uracil homopolymer complex. J. Mol. Biol. 20:359-389.

Roberts, J. 1969. Termination factor for RNA synthesis. Nature 224:1168-1174.

Roberts, J. 1976. Transcription termination and its control in E. coli. In RNA Polymerase, R. Losik and M. Chamberlin eds., (Cold Spring Harbor Laboratory, Cold Spring Harbor, New York), pp. 247-271.

Robertson, H.D., R.E. Webster, and N.D. Zinder. 1968. Purification and properties of ribonuclease III from Escherichia coli. J. Biol. Chem. 243:82-91.

Roeder, R.G. 1976. Eukaryotic RNA polymerases. In RNA polymerase, R. Losick and M. Chamberlin eds., (Cold Spring Harbor Laboratory, Cold Spring Harbor, New York), pp. 285-329.

Roeder, R.G., and W.J. Rutter. 1970. Specific nuclear and nucleoplasmic RNA polymerases. Proc. Natl. Acad. Sci. (USA), 65:675-682.

Rohrbaugh, M.L., J.E. Johnson III, M.D. James, and R.C. Hardison. 1985. Transcription unit of the rabbit β-1 globin gene. Mol. Cell. Biol. 5:147-160.

Rosenberg, M., and D. Court. 1979. Regulatory sequences involved in the promotion and termination of RNA transcription. Ann. Rev. Genet. 13:319-353.

Rosenthal, L.J., C. Saloman, and R. Weil. 1976. Isolation and characterization of poly(A)-containing intranuclear polyoma-specific "giant" RNAs. Nucleic Acid Res. 3:1167-1183.

Ruley, H.E., and M. Fried. 1983. Sequence repeats in a polyoma virus DNA region important for gene expression. J. Virol. 47:233-237.

Salditt-Georgieff, M., M. Harpold, S. Chen-Kiang, and J.E. Darnell, Jr. 1980. The addition of 5' cap structures occurs early in hn RNA synthesis and prematurely terminated molecules are capped. Cell 19:69-78.

Salstrom, J.S., and W. Szybalski. 1978. Coliphage lambda nutL⁻: a unique class of of mutants, defective in the site of gene N product utilizattion for antitermination of leftward transcription. J. Mol. Biol. 124:195-221.

Scholtissek, C. 1971. Detection of an unstable RNA in chicken fibroblasts after reduction of the UTP pool by glucosamine. Eur. J. Biochem. 24:358-365.

Sehgal, P.B., J.E. Darnell, Jr., and I. Tamm. 1976. The inhibition by DRB (5,6-dichloro-1-beta-D-ribofuranosyl benzimidazole) of hnRNA and mRNA production in HeLa cells. Cell 9:473-480.

Sharp, J.A., and T. Platt. 1984. Rho-dependent termination and concomitant NTPase activity requires a specific, intact RNA region. J. Biol. Chem. 259:2268-2273.

Shaw, A.B., and E.B. Ziff. 1980. Transcription from the adenovirus 2 major late promotor yield a single early family of 3' coterminal mRNA and five late families. Cell 22:905-916.

Shenk, T.E., C. Rhodes, P.W.J. Rigby, and P. Berg. 1975. Biochemical method for mapping mutational alterations in DNA with S1 nuclease: the location of deletions and temperature-sensitive mutations in Simian virus 40. Proc. Natl. Acad. Sci. (USA). 72:989-993.

Skarnes, W.C. 1985. In vitro study of polyomavirus transcription. M. Sc. thesis, McGill University, Montreal, Quebec.

Soeda, E., J.R. Arrand, N. Smolar, and B.E. Griffin. 1979. Sequence from early region of polyoma virus DNA containing viral replication origion and encoding small middle and (part of) large T antigens. Cell 17:357-370.

Soeda, E., J.R. Arrand, N. Smolar, J.E. Walsh, and B.E. Griffin. 1980. Coding potential and regulatory signals of the polyoma virus genome. Nature 283:445-453.

Sollner-Webb, B., and R.H. Reeder. 1979. The nucleotide sequence of the initiation and termination sites for ribosomal RNA transcription in X. laevis. Cell 18:485-499.

Sollner-Webb, B., and J. Tower. 1986. Transcription of cloned eukaryotic ribosomal RNA gene. Ann. Rev. Biochem. 55: 801-803.

Stavrianopoulos, J.G., and E. Chargaff. 1973. Purification and properties of ribonuclease H of calf thymus. Proc. Natl. Acad. Sci. (USA). 70:1959-1963.

Stein, H., and P. Hausen. 1969. Enzyme from calf thymus degrading the RNA moiety of DNA-RNA hybrids: effect on DNA-dependent RNA polymerase. Science 166:393-395.

Stewart, S.E., B.E. Eddy, and N. Borgese. 1958. Neoplasms in mice inoculated with a tumor agent carried in tissue culture. J. Natl. Cancer Inst. 20:1223-1244.

Strauch, M., and D.I. Friedman. 1981. Identification of the nusB gene product of Escherichia coli. Mol. Gen. Genet. 182:498-501.

Strub, K., G. Galli, M. Busslinger, and M.L. Birnstiel. 1984. The cDNA sequences of the sea urchin U7 small nuclear RNA suggest specific contacts between histone mRNA precursor and U7 RNA during RNA processing. EMBO J. 3:2801-2807.

Stunnenberg, H.G., and M.L. Birnstiel. 1982. Bioassay for components regulating eukaryotic gene expression: a chromosomal factor involved in the generation of histone mRNA 3' termini. Proc. Natl. Acad. Sci. (USA). 79: 6201-6204.

Tamm, I. 1977. Definition of subclasses of nucleoplasmic RNA. Proc. Natl. Acad. Sci. (USA). 74:5011-5015.

Tamm, I., T. Kikuchi, J. E. Darnell, Jr., and M. Salditt-Georgieff. 1980. Short capped mRNA precursor chains in HeLa cells: continued synthesis in the presence of 5,6-dichloro-1-beta-D-ribofuranosylbenzimidazole (DRB). Biochemistry 19:2743-2748.

Treisman, R. 1980. Characterization of polyoma late mRNA leader sequences by molecular cloning and DNA sequence analysis. Nucleic Acids Res. 8:4867-4888.

Treisman, R., and R. Kamen. 1981. Structure of polyoma virus late nuclear RNA. J. Mol. Biol. 148:273-301.

Tseng, R.W., and N.H. Acheson. 1986. Use of a novel S1 nuclease RNA-mapping technique to measure efficiency of transcription termination on polyomavirus DNA. Mol. Cell. Biol. 6:1624-1632.

Tyndall, C., G. LaMantia, C.M. Thacker, J. Favaloro, and R. Kamen. 1981. A region of the polyoma virus genome between the replication origin and late protein coding sequence is required in cis for both early gene expression and viral DNA replication. Nucleic Acids Res. 9:6231-6250.

Veldman, G.M., J. Klootwijk, P. de Jonge, R.J. Leer, and R.J. Planta. 1980. The transcription termination site of the ribosomal RNA operon in yeast. Nucleic Acids Res. 8: 5179-5192.

Veldman, G.M., S. Lupton, and R. Kamen. 1985. Polyomavirus enhancer contains multiple redundant sequence elements that activate both DNA replication and gene expression. Mol. Cell. Biol. 5:649-658.

Vogt, V.M. 1973. Purification and further properties of single-strand-specific nuclease from Aspergillus oryzae. Eur. J. Biochem. 33: 192-200.

Vogt, V.M. 1980. Purification and properties of S1 nuclease from Aspergillus. Meth. Enzymol. 65:248-255.

von Hippel, P.H., D.G. Bear, W.D. Morgan, and J.A. McSwiggen. 1984. Protein-nucleic acid interactions in transcription: a molecule analysis. Ann. Rev. Biochem. 53:389-446.

Weaver, R.F., and C. Weissmann. 1979. Mapping of RNA by a modification of the Berk-Sharp procedure: the 5' termini of 15S beta-globin mRNA precursor and mature 10S globin mRNA have identical map coordinates. Nucleic Acids Res. 7:1175-1193.

Weber, J., W. Jelinek, and J.E. Darnell, Jr. 1977. The definition of large viral transcription unit late in Ad2 infection of HeLa cells: mapping of nascent RNA molecules labeled in isolated nuclei. Cell 10:611-616.

Weil, P.A., D.S. Luse, J. Segall, and R.G. Roeder. 1979. Selective and accurate initiation at the Ad2 major late promoter in a soluble system dependent on purified RNA polymerase II and DNA. Cell 18:469-484.

Weil, R., C. Saloman, E. May, and P. May. 1974. A simplifying concept in tumor virology: virus specific "pleiotropic effectors". Cold Spring Harbor Symp. Quant. Biol. 39: 381-395.

Weil, R., and J. Vinograd. 1963. The cyclic helix and cyclic coil forms of polyoma virus DNA. Proc. Natl. Acad. Sci. (USA). 50:730-738.

Weinmann, R., and R.G. Roeder. 1974. Role of DNA-dependent RNA polymerase III in the transcription of tRNA and 5S RNA genes. Proc. Natl. Acad. Sci. (USA). 71:1790-1794.

Weintraub, H., A. Larson, and M. Grondine. 1981. Alpha-globin gene switching during the development of chicken embryos: expression and chromosome structure. Cell 24:333-344.

Wertz, G.W. 1975. Method for examining viral metabolism in cells in culture: metabolism of vesicular stomatitis virus RNA. J. Virol. 16:1340-1344.

Wicken, M., and P. Stephenson. 1984. Role of the conserved AAUAAA sequence: four AAUAAA point mutants prevent messenger RNA 3' end formation. Science 226:1045-1051.

Wu, A.M., A.B. Chapman, T. Platt, L.P. Guarente, and J. Beckwith. 1980. Deletions of distal sequence affect termination of transcription at the end of tryptophan operon in E. coli. Cell 19:829-836.

Wu, A.M., G.E. Christie, and T. Platt. 1981. Tandem termination sites in the tryptophan operon of Escherichia coli. Proc. Natl. Acad. Sci. (USA). 78:2913-2917.

Wu, A.M., and T. Platt. 1978. Transcription termination: nucleotide sequence at 3'end of tryptophan operon in Escherichia coli. Proc. Natl. Acad. Sci. (USA). 75:5442-5446.

Yamamoto, K.R., R.M. Alberts, R. Benzinger, L. Lawhorne, and G. Treiber. 1970. Rapid bacteriophage sedimentation in the presence of polyethylene glycol and its application to large-scale virus purification. Virology 40:734-744.

Yanofsky, C. 1981. Attenuation in the control of tryptophan operon expression. Nature 289:751-758.

Zaret, K.S., and F. Sherman. 1982. DNA sequence required for efficient transcription termination in yeast. Cell 28: 563-573.

Zuckermann, M., H. Manor, J. Parker, and R. Kamen. 1980. Electron microscopic demonstration of the presence of amplified sequences at the 5'-ends of the polyoma virus late mRNAs. Nucleic Acids Res. 8:1505-1519.

Zurawski, G., and C. Yanofsky. 1980. Escherichia coli tryptophan operon leader mutations which relieve transcription termination are cis-dominant to trp leader mutations which increase transcription termination. J. Mol. Biol. 142:123-129.