ATTENUATION MAY REGULATE GENE EXPRESSION IN ANIMAL VIRUSES AND CELLS

Authors: **Yosef Aloni**
Department of Genetics
Weizmann Institute of Science
Rehovot, Israel

Nissim Hay
Hooper Foundation
University of California
San Francisco, California

Referee: Guiseppe Attardi
Department of Biology
California Institute of Technology
Pasadena, California

I. INTRODUCTION

One of the most important advances in understanding the regulation of gene expression came with the concept of Jacob and Monod.[1] This concept proposed that the regulation of gene expression is achieved by controlling the frequency of the initiation of transcription. However, while many genes are in fact controlled in this manner, it is becoming apparent that transcription of many other genes is also regulated by transcription termination.[2-9] This is especially true in the intragenic termination of transcription found in bacteria and their viruses which have been termed attenuation.[2-5] While most control mechanisms at the initiation level are qualitative (all or none), the regulation by attenuation is quantitative. In eukaryotes there is an accumulating body of evidence which suggests that transcriptional differences between cells of different tissues, developmental or cell-cycle stages, or hormonal states may be more quantitative than qualitative (no large changes over two or three orders of magnitude). This led to the speculation that attenuation of transcription may play a major role in the regulation of eukaryotic gene expression.[10]

In the present communication we review the available information which indicates that a mechanism resembling attenuation in prokaryotes operates in eukaryotes. We shall deal primarily with studies carried out with the small DNA tumor virus SV40 which, in this respect, is the best-studied system to date. SV40 provides several unique advantages as a model system for the study of transcription and posttranscription regulation in eukaryotes, including the following:

1. The viral genome is a small circular molecule that contains genetic information for only six proteins, and there may be additional small polypeptides. The DNA can be obtained in large quantities, which is imperative for many experiments in molecular biology.
2. The same RNA polymerase (polymerase II) transcribes both viral and cellular RNA.
3. The viral and cellular RNAs undergo similar posttranscriptional modification (e.g., capping at the 5′ terminus, polyadenylation at the 3′ terminus, internal methylation, and splicing).
4. A number of mutants and hybrid viruses are available for study.

5. Transcriptional complexes are easy to obtain.
6. The entire nucleotide sequence of the genome of this virus and the localization of the major and minor initiation sites for transcription have been determined.[11-18]

A variety of reviews of different aspects of SV40 are listed in Tooze[11] and Lebowitz and Weissman.[12] Here we shall review only that part which has relevance to the studies of attenuation. There are also a variety of reviews concerning attenuation in prokaryotes,[2-4,8,9] and here we shall summarize only the basic principles of this mechanism in the tryptophan operon in *Escherichia coli*.

In the tryptophan operon, attenuation is a mechanism of gene expression which couples transcription and translation. Whereas the mechanism of transcription initiation in prokaryotes and eukaryotes share similar regulatory elements,[2,3,8,19-23] the mechanism of translation initiation in prokaryotes and eukaryotes and selection of the AUG start codon differ in some principle features.[24-28] It should, therefore, be assumed that if eukaryotes have inherited the attenuation mechanism from their prokaryotic ancestors,[29] they had to develop a new mechanism for coupling transcription and translation. There are a variety of reviews concerning translation in eukaryotes and prokaryotes,[28,30,31] and here we shall relate only to some aspects of the mechanism of mRNA recognition by eukaryotic ribosomes during initiation of protein synthesis.

II. INTRODUCTION TO THE SV40 SYSTEM

The SV40 genome is comprised of early and late genes that are localized in symmetrical halves of the viral DNA. The segment between 0.67 and 0.17 on the map is transcribed in counterclockwise direction prior to the onset of viral DNA replication, and codes for the early viral proteins. The second segment (from 0.67 to 0.17) is transcribed in abundance after initiation of viral DNA replication in a clockwise direction. It encodes the information for the late proteins (see Figure 1).

The transcription of SV40 early RNA starts at a principal initiation site 21 bp downstream from a TATA box.[21,32,33] The efficient initiation of early transcription is also dependent on enhancer sequences composed of 72 bp tandem repeats, 100 nucleotides upstream of the TATA box.[34] There are two alternative splices in SV40 early mRNA (see Figure 1); one splice removed only 66 nucleotides. As a result, a continuous reading frame encoding small t antigen, together with a termination codon, is present, and therefore, the codons of the messenger downstream from small t termination codon are not translated.[35,36] The other form of early messenger RNA has a longer internal region removed, including the termination codon for small t. The codons from the initial portion of small t are joined in phase to a large open reading frame to produce the message for large T.[35,36] The sequence of the SV40 early region contains an alternative reading frame of 91 continuous codons beginning with four AUGs located at the 3′ terminal portion of the early region (0.21 to 0.16 map units). Although some of the 91 codons in the second open reading frame lie in the untranslated 3′ end of the large T messenger RNA, it has been suggested that this alternative reading frame might be used at some time in SV40 infection.[11,12,37]

A specific decrease in transcription of the SV40 early transcription unit has been observed in late periods of infection.[11,38-40] Late in infection when there is a very large increase in total SV40 RNA in infected cells, there is only 1/20th as much RNA complementary to the early transcription unit as to the late transcription unit.[39-41] This is due to autoregulation of transcription by T antigen.[38-40] T antigen has been shown to bind three identical sites within a 100 to 150-base pair region that overlaps the early promoter.[42-44] This binding is necessary for DNA replication[45] and independent for late transcription,[46] but it also reduces considerably the early transcription of SV40 RNA.

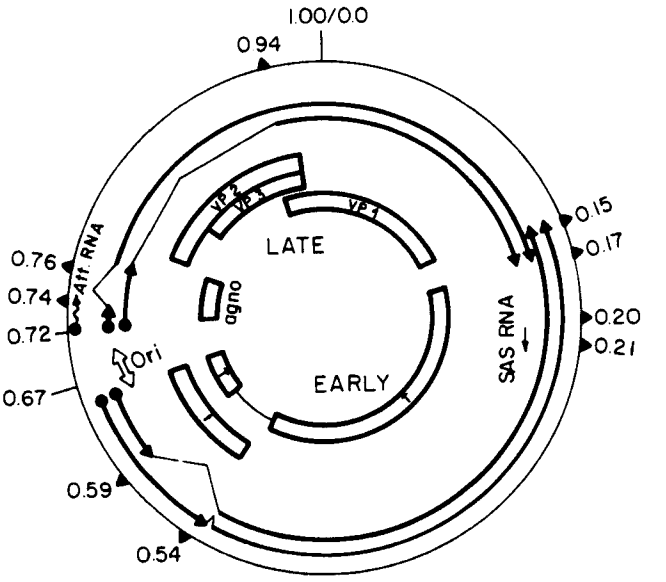

FIGURE 1. Map of SV40 RNAs and proteins. Solid straight lines mark the extent of the early and late mRNAs, while the thin nonstraight lines indicate spliced-out introns. Arrows indicate direction of transcription. The 5′ end of early and major late RNA are indicated by closed circles. The map units of initiation, splicing, and polyadenylation sites are indicated. Proteins are indicated by the framed areas. Ori is the origin of replication.

In contrast to early promoter, the promoter of late RNA transcription lacks a TATA box, and late SV40 mRNAs have heterogeneous initiation sites. There are two major classes of SV40 late mRNA, the 16S and 19S (Figure 1) present in a relative abundance of approximately 5:1.[11,12] Each class of RNA contains multiple forms of leader sequences,[47,48] all of which are transcribed from the nuclease-sensitive stretch of viralDNA[49,50] corresponding to the gap in the viral minichromosome.[51-54] The 16S mRNA family is transcribed to form the capsid protein VP1 and has most of the coding region for VP2 and VP3 removed by splicing, while 19S mRNA directs the synthesis of both VP2 and VP3[11,12] (Figure 1). The structure of the leader sequence attached to both 16S and 19S mRNA is complex and, with a single exception, there appears to be no leader that is identical in both 16S and 19S mRNA.[12,47,48] The principal sources of the heterogeneity of SV40 late mRNA are the multiple 5′ end of the RNA,[12,48] the presence of internal splices within 16S mRNA leader segments, and the use of three alternative 3′ ends in 19S leader segments. In addition, a minor portion of SV40 late mRNA has a tandem reduplication of the segment of the leader.[55,56] Presumably, this arises as a result of splicing within RNA molecules which contain transcripts of more than one copy of the viral genome.[56]

The predominant type of leader sequence in the 16S mRNA is 202 nucleotides in length. It contains an initiation condon, 61 sense codons, and the termination condon of the leader protein-agnoprotein[47] (see Figure 2). Therefore, 16S mRNA is the RNA from which agnoprotein is translated.[57] The AUG of agnoprotein is the first AUG proximal to the cap in the 16S mRNA and is surrounded by consensus-favorable nucleotides.[26,28,29] A question, therefore arises how the major capsid protein VP1 is translated, since upstream in the messenger RNA there is the translatable coding frame of agnoprotein. In addition to the principal late leader of 16S mRNA, where the 5′ end lies at nucleotide 243 (according the numbering system of Reddy et al.[17]), there are multiple forms of the late leader of 16S mRNA whose 5′ ends lie upstream to that of the principal late leader. All of these additional

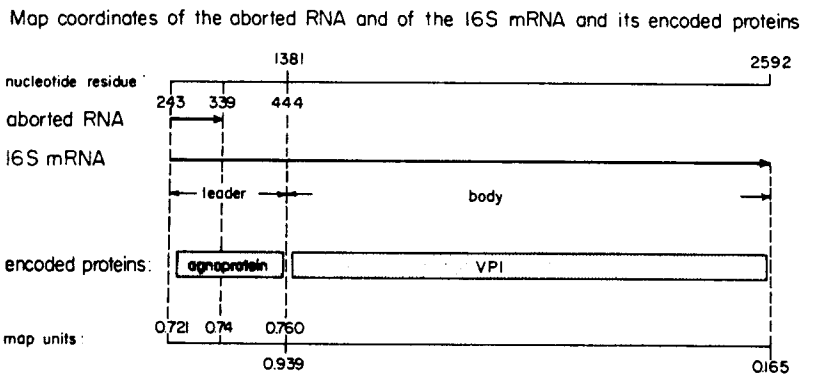

FIGURE 2. Map coordinates of the attenuated RNA and of the 16S mRNA and its encoded proteins. Nucleotide residue numbers refer to the wild-type SV40 sequence by Reddy et al.[17] (From Hay, N., Skolnik-David, H., and Aloni, Y., *Cell*, 29, 183, 1982. With permission.)

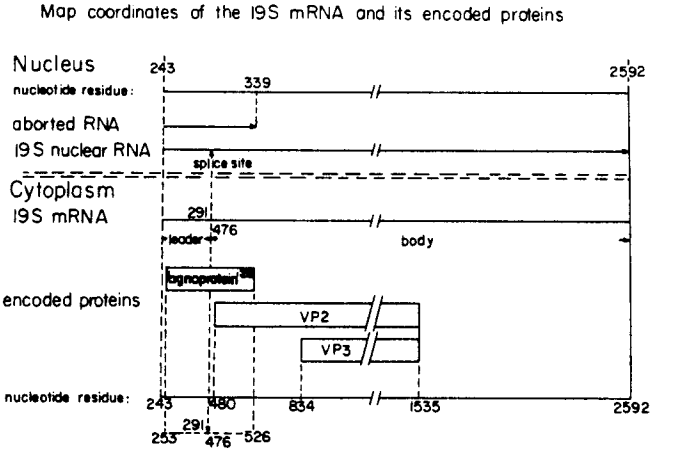

FIGURE 3. Map coordinates of the attenuated RNA, of the 19S nuclear RNA, and of the 19S mRNA and its encoded proteins. Residue numbers are from Reddy et al.[17] (From Aloni, Y. and Hay, N., *Mol. Biol. Rep.*, 9, 91, 1983. With permission.)

leaders of SV40 late mRNA have internal splices within them. These splices remove all AUG triplets from the leader.[57]

The 19S mRNA family contains only leader sequences that represent transcripts of continguous segments of SV40 DNA, without internal splices. However, unlike the 16S mRNA, there are three alternative 3′ ends for 19S leaders. One of these corresponds to the 3′ end of the 16S leaders, the second occurs in sequences upstream from the DNA encoding the 5′ end of the major 16S leader, and the third, which is the most abundant, lies 43 nucleotides downstream from the 5′ end of the major 16S leader, removing most of the coding region and one of the two AUG triplets in the 16S leader.[48] Consequently, in the most abundant leader of 19S mRNA there is an open reading frame for the synthesis of 29 amino acids protein. This coding frame shares the first 13 amino acids with the agnoprotein coding frame and overlaps with the AUG of VP2 (see Figure 3). The translation product of this conding frame has not yet been identified.

In addition to polyA-containing mRNA, a small RNA species arises in relatively large quantities, late in SV40 lytic infection.[58] This small RNA is called SV40-associated small

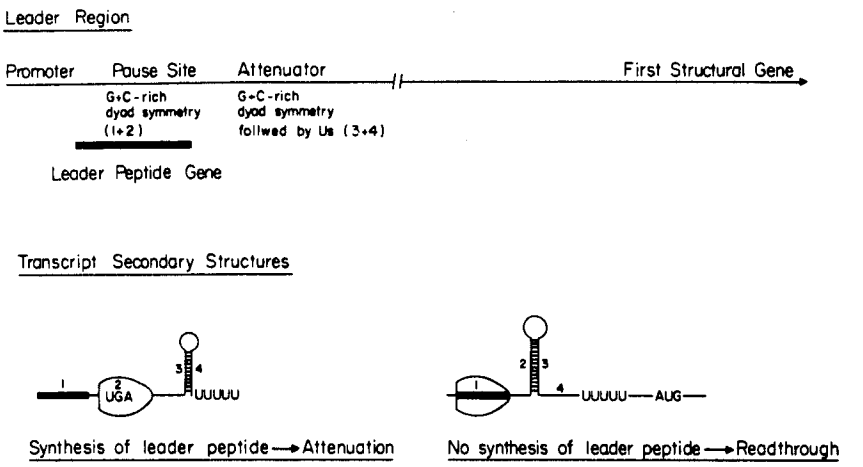

FIGURE 4. Model of attenuation in amino acid biosynthetic operon. (From Kolter, R. and Yanofsky, C., *Ann. Rev. Genet.*, 16, 113, 1982. With permission.)

RNA (SAS-RNA), and it was suggested that it is a processing product from the 3′ untranslated portion of the late SV40 primary transcripts[59,60] (see Figure 1). The mechanism by which this RNA is produced and its function are not known.

III. A MODEL OF ATTENUATION IN AMINO ACID BIOSYNTHETIC OPERONS OF BACTERIA

Attenuation of transcription in bacteria was first found and well studied in amino acid operons.[4,9] All amino acid operons which are regulated by attenuation have a leader region which contains certain consistent features (see Figure 4). Near the beginning of the leader message there is a ribosome binding site. Within the leader message there is a translation initiation codon, AUG, followed by a coding frame for a small peptide that has several codons of the amino acid of the operon and a stop codon. The RNA in this region can form extensive secondary structures. There are three regions of dyad symmetry (inverted repeats), which are capable of forming two mutually exclusive secondary structures. One of the structures consists of a G-C-rich stem-and-loop followed by oligo U. This conformation causes transcription termination, which prevents transcription from continuing into the subsequent coding sequences. The alternative structure permits transcription of the structural genes and does not cause termination. The formation of the two alternative structures is regulated by the translation of the small peptide encoded in the leader transcript. The three regions of dyad symmetry are found in four segments of the transcript. Segments 1 and 2 form one stem-and-loop, while 3 and 4 form another (the terminator). 1 + 2 and 3 + 4 structures do not overlap and can exist together. However, segments 2 and 3 can, also, form a stem-and-loop which, due to overlap, cannot coexist with structures 1 + 2 and 3 + 4. Movement of the ribosomes during translation of the leader peptide along the leader transcript regulates which of the secondary structures is formed.

The position of the translating ribosome is dependent on the concentration of the specific amino acid in the cell. Thus, for example, when there is an excess amount of tryptophan in the cell, the ribosome moves up to the UGA terminator codon. The UGA codon is located on segment 2 and the ribosome, therefore, prevents segment 2 from base pairing with segment 3. This leads to formation of 3 + 4 which signals transcription termination. In contrast, when there is a shortage of tryptophan in the cell the translating ribosome stalls at the tryptophan codons of the leader peptide that are located on segment 1. In this case, the

ribosome prevents segment 1 from base pairing with segment 2. This leads to formation of 2 + 3. When 2 + 3 is formed it prevents terminator (3 + 4) formation, as segment 3 is no longer available for base pairing and the RNA polymerase will transcribe the structural gene.

Before discussing whether a mechanism of attenuation which couples transcription and translation could operate in eukaryotes, we shall briefly summarize the main differences between translation initiation in eukaryotes and prokaryotes.

IV. A MECHANISM OF mRNA RECOGNITION BY EUKARYOTIC RIBOSOMES DURING INITIATION OF PROTEIN SYNTHESIS

The fundamental components of the translation machinery are the same for prokaryotes and eukaryotes. However, there are striking dissimilarites in the process of translation initiation, between prokaryotes and eukaryotes. The major dissimilarities are connected with the fact that prokaryotic mRNA is structurally and functionally polycistronic, and eukaryotic mRNA is, in general, functionally monocistronic. In eukaryotic mRNA, the first AUG from the 5′ end of the RNA is the functional AUG. A striking feature of the eukaryotic mRNA is that there are usually no other AUG triplets preceding the functional initiator codon. The first AUG rule holds for ~90% of eukaryotic mRNAs, including many with very long 5′ leaders.[28,61-66] Based on this and additional experimental observations, a model for initiation of translation in eukaryotic cells was suggested. According to the simplest version of this model, which is called a "scanning model", the 40S ribosomal subunit binds first at or near the 5′ end of the mRNA molecule, and then migrates on the mRNA until it encounters the first AUG triplet, at which point a 60S subunit would join and the first peptide bond would be synthesized.[24,25] In this version of the model the initiator codon is defined merely by its position, i.e., closest to the 5′ end of the mRNA. However, there are ~10% of the mRNA species in which ribosome initiates at an AUG triplet that is not closest to the 5′ end of the mRNA.[26,28] Inspection of the nucleotides sequence flanking the apparently non-functional upstream AUG triplets led to a modified version of the scanning model.[26,28,29] According to the modified version of the scanning model, if the first AUG codon encountered by the migrating 40S ribosomal subunit occurs in the optimal sequence context ($\frac{A}{G}$NNAUGG), 40S subunit will uniquely initiate at that site; however if the first AUG triplet occurs in a less favorable context (e.g., GNNAUG $\frac{C}{U}$ or $\frac{C}{U}$ NNAUGG) some 40S subunits will stop and initiate there while some will bypass that site and initiate further downstream.[26,28,29] In addition, the 40S subunit can bypass the first AUG if it is base paired with other sequences.[67,68]

V. ATTENUATION IN THE CONTROL OF SV40 GENE EXPRESSION

A. Introduction

Promoters and terminators are potential sites at which gene expression can be controlled.[3,69] In prokaryotes RNA polymerases, with and without accessory factors, specifically recognize promoter and terminator sequences. In eukaryotes, although sequences which constitute promoters have been partially characterized,[70-73] the mechanism of termination of RNA polymerase II transcription is not known. In prokaryotes, transcription termination sites are located within as well as at the end of the operons. The internal termination sites cause premature termination of the transcripts and quantitatively regulate the level of gene expression, by selectively reducing the transcription of distal portions of the operons. This mechanism of regulation has been termed attenuation.[4] Pause site, within the operon, sets the rate of transcription below that attainable by polymerase-promoter interaction alone and, therefore, considered as a mechanism of attenuation.[4] Attenuation is not a complete termination process. It can be overcome by regulatory proteins, as in bacteriophage λ,[74] or in response to changes in physiological conditions as in the case of amino acid operons in

bacteria.[4] The sequence at the termination is GC-rich and often possesses dyad symmetry. The 3′ terminus of the transcript typically contains a series of uridine residues. The role of these and other features, such as secondary structure in the transcript, distal DNA sequences, protein factors involved in termination and readthrough, and RNA-RNA or RNA-DNA interaction, is not fully understood (for reviews see References 4, 8, 9, 69, 74-77).

In eukaryotes, an abundant population of promoter-proximal RNA chains has been observed and studied, mainly in whole nuclear RNA,[78,79] in adenovirus type 2,[80-82] and in simian virus 40.[83,84] On the basis of these results it has been suggested that a premature termination process resembling attenuation in prokaryotes occurs in eukaryotes.[82] Moreover, these studies have shown that the adenosine analog 5,6-dichloro-1-β-D-ribofuranosylbenzimidazole (DRB) enhances premature termination, but its mode of action is not understood.[78,80,84,85]

Although the use of the term attenuation in eukaryotes closely follows the original usage in prokaryotic systems, there was no experimental basis to suggest that the mechanism of attenuation in eukaryotes is similar to that in prokaryotes. Furthermore, there is no experimental evidence that DRB enhances premature termination at the attenuation site.

The determination of the complete nucleotide sequences of SV40 and the localization of the major and minor initiation sites for late transcription[18,36,47,86,87] provide means for elucidating the nucleotide sequences involved in the attenuation mechanism. Below is a review of results indicating that a mechanism resembling attenuation in prokaryotes regulates SV40 late gene expression.

B. Rationale for a Mechanism of Attenuation in SV40

The product of encapsidation, the virion, is composed of SV40 DNA and cellular histones surrounded by a capsid consisting of viral proteins VP_1, VP_2, and VP_3 in the proportion 420/36/96, respectively, per capsid.[88,89] The simplest encapsidation mechanism requires self-assembly of viral proteins to form capsids. We assume that a controlled production of the structural proteins VP_1, VP_2, and VP_3 is critical for an efficient self-assembly process.[90] This is true for other viruses as well. We have suggested that attenuation and modulation of RNA secondary structure are the mechanisms that quantitatively regulate the production of the capsid proteins.[91,92] We describe below experiments that support this suggestion.

C. The Promoter-Proximal SV40 Nascent RNA Terminates in Vitro at a Transcription Termination Signal

Experiments performed in the last few years have shown an abundant population of promoter-proximal nascent RNA chains on SV40 transcriptional complexes (VTC), minichromosomes, and isolated nuclei.[91-98] These three systems are capable of in vitro elongating the in vivo preinitiated viral RNA. We adopted these systems in order to verify whether these short nascent viral RNAs are of discrete lengths and whether DRB enhances premature termination at physiological sites.

In the following experiment VTC were prepared from untreated and DRB-treated cultures and incubated in vitro for 5 min in the presence of α-^{32}P-UTP. The briefly elongated RNAs were purified and analyzed by gel electrophoresis. Figure 5 lanes A and B show one major band in a position corresponding to a length of 93 to 95 nucleotides as well as other minor bands in the RNA preparations of both untreated and DRB-treated cultures.[91] Longer viral RNA is absent in the preparation of the DRB-treated culture, as shown by the fact that the region above the major band has no radioactivity. Identical bands were obtained when the viral RNA was purified from minichromosomes, nuclear matrices, and isolated nuclei.[93,95,97]

To locate the RNA in the major band on the physical map of SV40 DNA, the RNA was eluted from the gel and hybridized to two sets of restriction fragments. The map locations of the restriction fragments of the two sets are shown in part E of Figure 5. Lane C of

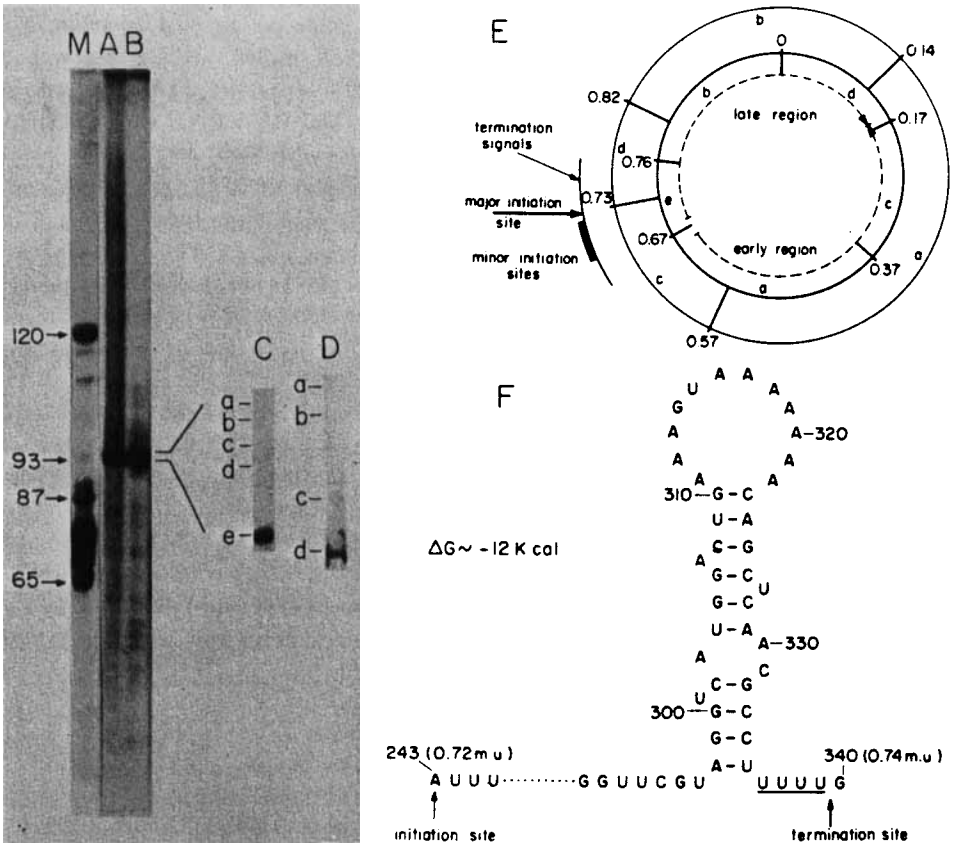

FIGURE 5. Size analysis of viral RNA of untreated and DRB-treated cultures and the predicted secondary structure of the attenuator region. VTC were extracted from untreated and DRB-treated infected cultures by the Sarkosyl extraction method[185] and incubated in vitro in the presence of α-^{32}P-UTP for 5 min, and ^{32}P-RNAs were extracted. The labeled RNAs were denatured and subjected to electrophoresis on a 12% acrylamide gel in 7 *M* urea. (Lane A) ^{32}P-RNA of untreated cultures; (lane B) ^{32}P-RNA of DRB-treated cultures; (lane M) length markers of *E. coli* RNA. The RNA in the major bands in lanes A and B (93 to 95 nucleotides in length) was eluted from the gel by electrophoresis[210] and hybridized to two sets of restriction fragments obtained by digestion of SV40 DNA with Eco. R1. Hpa 1 and Bgl 1 (lane C) and with Taq 1, Hpa 11, and Bam H1 (lane D). The map location of the fragments in lane C are shown in the inner circle and those of lane D in the outer circle in (E). (F) Secondary structure at the 3′ end of the attenuated RNA. The ΔG was calcaulted as described.[211] Nucleotide residue numbers refer to the wild-type SV40 sequence of Reddy et al.[17] (From Hay, N., Skolnik-David, H., and Aloni, Y., *Cell*, 29, 183, 1982. With permission.)

Figure 5 shows that hybridization occurred in one set solely with fragment e (0.67 to 0.76 map units), and lane D of Figure 5 shows that hybridization occurred in the second set solely with fragment d (0.73 to 0.82 map units). On the basis of these results, we concluded that the labeled RNA in the major band originates from a region of the genome spanning 0.73 to 0.76 map units (nucleotides 267 to 440).[17]

More direct evidence for the position of the 93 to 95 nucleotide RNA was obtained by the experiment illustrated in Figure 6. The 93 to 95 nucleotide RNA was recovered from a gel similar to that shown in lane A of Figure 5 and was hybridized with the "late" strand of a fragment spanning from the Hpa II site (nucleotide 267) to the Bam H1 site. The RNA-DNA hybrid was treated with RNase to remove the nonhybridized portion of the RNA, and the length of the RNase-resistant portion was determined, after denaturation, by gel electrophoresis.[91] Figure 6 shows that the RNase treatment reduced the length of the RNA from

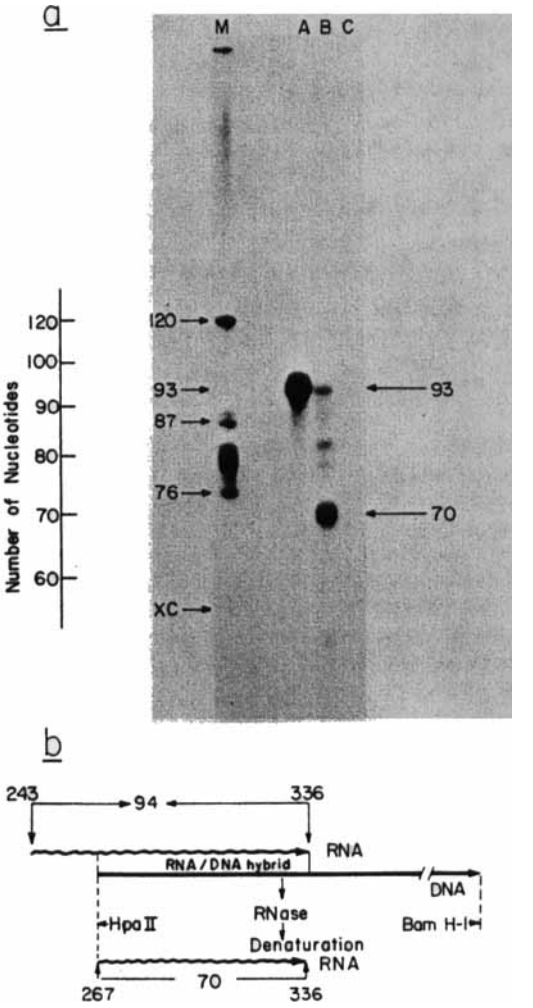

FIGURE 6. Mapping of the 93 to 95 nucleotide labeled RNA by RNAase digestion of RNA-DNA hybrid. (a) Shows an autoradiograph of 12% acrylamide gel in 7 *M* urea as in Figure 5; (lane A) the 93 to 95 nucleotide labeled RNA annealed with the L strand of the 0.73 to 0.14 map unit DNA fragment with no RNAase treatment; (lane B) same as in lane A, but with RNase treatment; (lane M) length markers of *E. coli* RNA; (b) diagram showing the plan of the experiment and the deduced topography of the labeled RNA with the lengths of the RNA before and after RNase treatment. Nucleotide residue numbers refer to the wild-type SV40 sequence of Reddy et al.[17] (From Hay, N., Skolnik-David, H., and Aloni, Y., *Cell*, 29, 183, 1982. With permission.)

93 to 95 nucleotides to 69 to 71 nucleotides. The minor bands may result from incomplete digestion of the RNA. We concluded that the 5′ end of the original RNA is at nucleotide 243 and the 3′ end at nucleotides 335 to 337.[17] We have designated this RNA, attenuated RNA. Further evidence confirming the position of the attenuated RNA came from a fingerprint analysis of RNase T1 digest of this RNA, followed by nearest-neighbor analysis of each spot.[94] The latter analysis also allowed for the identification of the in vivo pause sites.

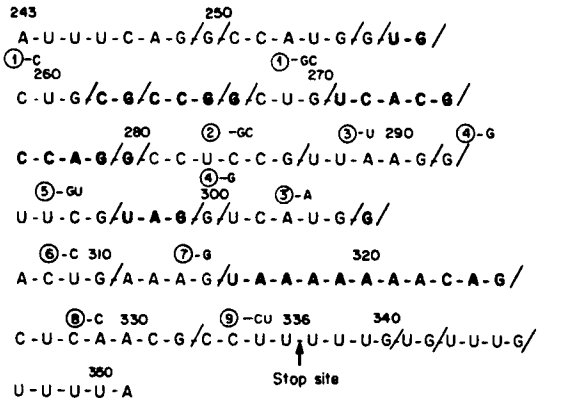

FIGURE 7. The oligonucleotides obtained after the RNase T1 digestion and the labeled nucleotides in each oligomer obtained following the RNase T2 digestion. Nucleotide residue numbers refer to wild-type SV 40 sequence of Reddy et al.[17] The slash indicates the RNase T1 cleavage site (following G). The nucleotides above each oligonucleotide are those which are labeled following the RNase T2 digestion. The shaded oligonucleotides are not labeled in the present analysis. (From Skolnik-David, H. and Aloni, Y., *EMBO J.*, 2, 179, 1983. With permission.)

D. The Location of the in Vivo Pause and in Vitro Attenuation Sites: GC-Rich Dyad Symmetry Region in the Process of Pausing and Attenuation

When the labeled RNA present in the major band is digested with RNase T1 and fingerprinted,[99] an oligonucleotide pattern characteristic of a transcript from the region between nucleotides 243 and 335 to 337 can be predicted. However, in an actual experiment, only those oligonucleotides which were elongated in vitro should produce radioactive spots, because the in vivo synthesized oligonucleotides are unlabeled. The junctions between labeled and unlabeled oligonucleotides define the in vivo pause sites of RNA polymerase molecules.

Figure 7 shows the base sequence spanning the major initiation site (nucleotide 243) and nucleotide 351, the predicted RNase T1 cleavage sites (following G residues), and the predicted labeled nucleotides obtained after the RNase T2 digestion of each of the oligonucleotides. Note that several oligonucleotides are not expected to be labeled. Ten consecutive oligonucleotides spanning nucleotide 259 to nucleotide 336 were identified.[94] It was evident that spot 1 contained the two oligonucleotides CUG (nucleotides 259 to 261 and 268 to 270) because the ratio of radioactivity between CMP and GMP in spot 1 was 2:1. It should be noted that some polymerase molecules may have initiated elongation 13 nucleotides downstream from the transcription initiation site, at nucleotide 256, but we were unable to resolve it, because nucleotides 256 to 258 could not be identified by our analyses.

In none of our fingerprint analyses have we ever observed an additional spot that could correspond to the oligonucleotide CCAUG (nucleotides 251 to 255) (see Figure 7). It appears, therefore, that in vivo the first pause site of active RNA polymerase molecules in VTC is located 13 to 16 nucleotides downstream from the transcription initiation site. It is noteworthy that spot 2 consistently contained about two to three times more radioactivity than spot 1, in spite of the fact that spot 2 has only two labeled phosphates as compared to three in spot 1 (see Figure 7). Based on this observation we concluded that the highest concentration of RNA polymerase molecules on VTC is in the vicinity of the oligonucleotide CCUCCG (281 to 286) of spot 2.[94] This could be the location of an in vivo pause site. If this is an exclusive pause site, then the occurrence of RNA polymerase molecules upstream from it could reflect

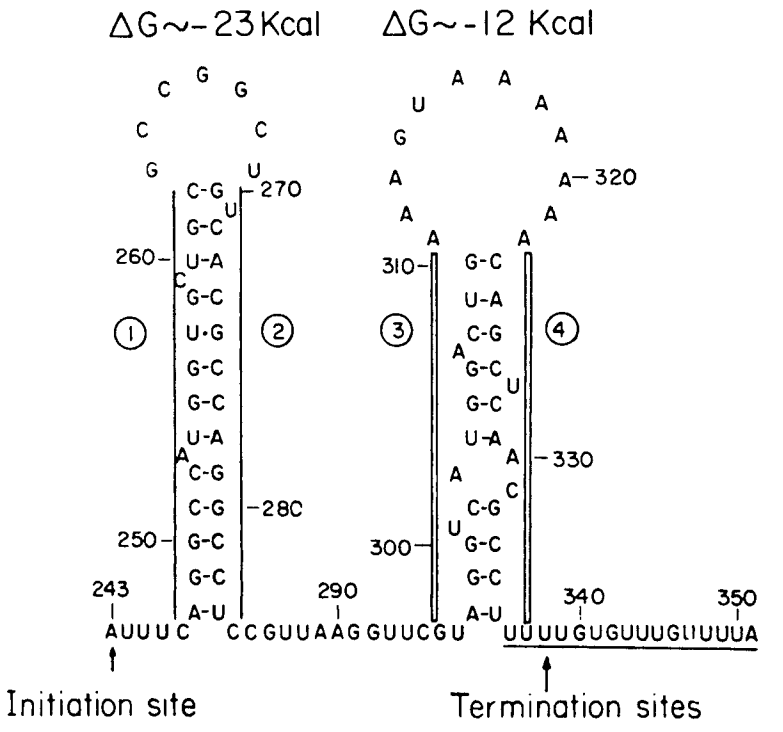

FIGURE 8. The "attenuation" conformation of the attenuated RNA. The ΔG was calculated as described.[211] Nucleotide residue number refers to the wild-type SV40 sequence of Reddy et al.[17] (Taken from Hay, N., Skolnik-David, H., and Aloni, Y., *Cell*, 29, 183, 1982. With permission.)

a situation in which two or more RNA polymerase molecules have initiated transcription on the same VTC and they are physically blocked by the enzyme present at the pause site. Alternatively, it is possible that RNA polymerase molecules also pause in vivo at the open promoter complex.[100,101]

The precise 3′ end of the RNA in the major band (the in vitro attenuation site) maps at nucleotide 336. This conclusion was based on the 1:1 ratio of radioactivity between CMP and UMP in spot 9[94] (see Figure 7). The RNA in the major band is, therefore, 94 nucleotides long (nucleotides 243 to 336). This confirms the above estimations.

SV40 DNA sequences spanning nucleotides 243 and 336 possess dyad symmetries sufficient for the formation of stable hairpin conformation. This conformation is designated "attenuation". The pair of inverted repeat sequences 1 + 2; 3 + 4 of the "attenuation" conformation is shown in Figure 8.

The oligonucleotide of spot 2, which we have suggested above to be the position of a pause site of RNA polymerase molecules in vivo, is at the end of a GC-rich region of dyad symmetry (1 + 2 in Figure 8). The major in vitro attenuation site is at the second uridine residue that follows a GC-rich region of dyad symmetry (3 + 4 in Figure 8). Based on these observations, we concluded that the eukaryotic RNA polymerase II, when transcribing SV40 DNA, responds to signals for pausing and attenuation, similar to those in prokaryotes.[4,8,9,69,74,76,77]

It has been suggested that in the *trp* leader region the function of the early (1 + 2) hairpin is to retard the polymerase in order to permit the translating ribosome to catch up and remain coupled to the transcription apparatus.[102,103] This, of course, does not apply to the SV40 system where transcription and translation are uncoupled processes. However, there are

several other potential functions for the pause in the first opening of the DNA helix and at the 1 + 2 hairpin structure. One of them is to help mediate interactions between RNA polymerase and other proteins such as "attenuator" and "antiattenuator" factors. Another function is to allow capping and methylation of the cap at the 5′ end of the RNA transcripts. In HeLa cells and adenovirus 2-infected cells, the prematurely terminated transcripts (i.e., attenuated RNAs) are already capped and methylated,[79,82] but the RNA at the pause sites has not been analyzed. In this regard it is interesting to note that Lycan and Danna[104] have found in nascent SV40 RNA, in addition to capped ends, some unprocessed pppA ends, and undermethylated caps.

E. The Attenuator RNA is Not Associated with the Nuclear Matrix: Association of Nascent RNA with the Nuclear Matrix Pertains to the Antiattenuation Mechanism

The structural framework of the nucleus, known as the nuclear matrix, is obtained when nuclei are depleted of their membranes, soluble molecules and chromatin by subsequent treatments with detergents, nucleases, and high salt.[105-107] Several reports have demonstrated the association of newly synthesized RNA with the nuclear matrix as well as RNA processing products,[108-116] presumably, via two proteins of Mr 41,500 and 43,000.[117]

We addressed the question of whether the nuclear matrix is actively involved in the attenuation mechanism in SV40.[97] Isolated nuclei were prepared from SV40-infected cells, pretreated with or without DRB or proflavine, and were incubated in the presence of α-^{32}P-UTP for 5 min. Nuclei, DNase plus salt-soluble fractions, and nuclear matrices were prepared and ^{32}P-labeled RNAs were extracted. One portion of each ^{32}P-RNA preparation was hybridized to restriction fragments as outlined in Figure 5. The remaining portion of each preparation was used to select the viral RNA by hybridization to, and elution from, SV40 DNA on filters. The purified RNAs were then subjected to gel electrophoresis.

Figure 9 shows that total nuclear viral RNA hybridized with all the restriction fragments, but with a relative enrichment of hybridization with the promoter-proximal fragment e and its adjacent fragment b. The hybridization with fragment e is more pronounced with the RNA derived from the DNase and salt-soluble fraction as compared with that of the matrix fraction. These results indicate that the soluble fraction is enriched with the promoter-proximal viral RNA species. The gel electrophoresis analyses in Figure 9b show that the enrichment of the soluble fraction with the promoter-proximal viral RNA species is, at least in part, due to the existence of the 94 nucleotide attenuator RNA. The attenuator RNA is found among total nuclear RNA and in the soluble fraction but is not detected in the matrix fraction.

The enrichment in the accumulation of the attenuator RNA is even more pronounced in preparation of ^{32}P-RNA synthesized in nuclei from the DRB pretreated cells. Figure 9a shows that in all three preparations (nuclear, matrix, and soluble) there is an almost exclusive signal with fragment e. However, the gel analysis (Figure 9b) shows that the attenuator RNA is detectable only in the nuclear and soluble preparations and not in the matrix fraction. These observations suggest that the matrix-bound viral RNA is heterogeneous in its length and it, presumably, spans the entire length of fragment e. It can also be seen in Figures 9a and 9b, that when the infected cells are pretreated with proflavine, the in vitro synthesized ^{32}P-labeled RNA does not hybridize with restriction fragment e and there is no detectable 94 nucleotide band on the gel. This is in agreement with our observation indicating that pretreatment of the infected cells with the intercalating drug, proflavine, causes read-through.[91] Although the attenuator RNA is not bound to the matrix, the matrix could well be involved in the attenuation mechanism.

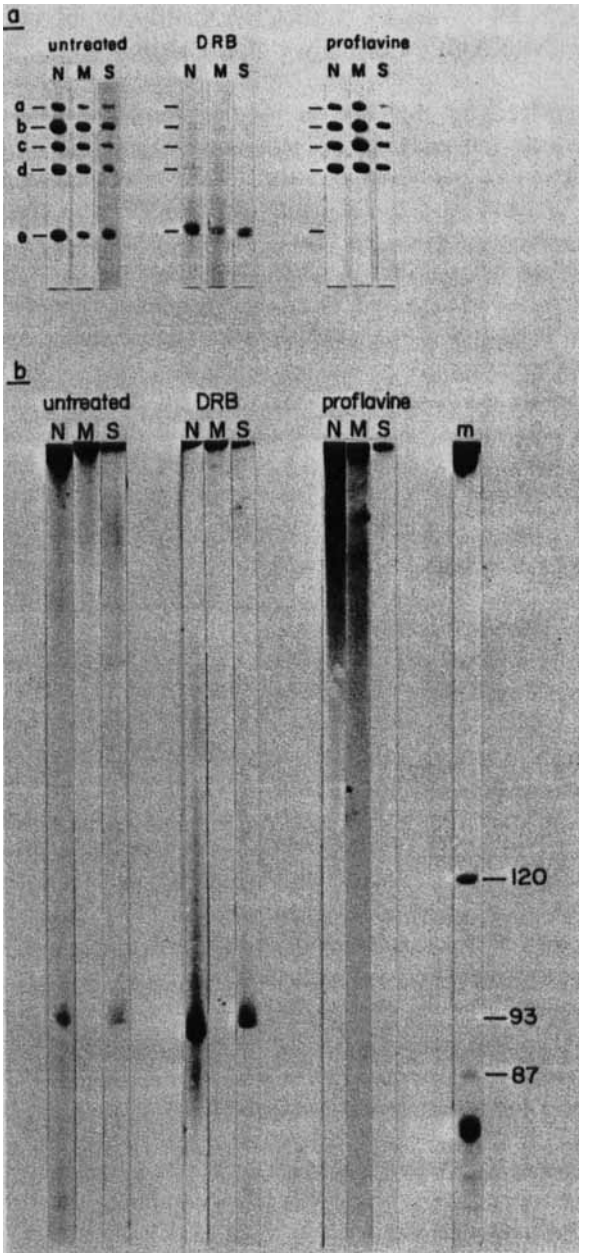

FIGURE 9. The attenuated RNA is not associated with the nuclear matrix. (a) Isolated nuclei (N) prepared from an equal number of untreated, DRB-treated (75 μ*M*, 40 min), or proflavin-treated (80 μM, 2 min) SV40 infected cells were incubated with α-^{32}P-UTP as described in Figure 5. The nuclear matrix fraction (M) and a DNase plus salt-soluble fraction (S) were prepared and ^{32}P-labeled RNA was extracted and hybridized to the five DNA restriction fragments as described in Figure 5; (b) an equal portion of each RNA sample obtained as described in (a) was hybridized to, and eluted from, SV40 DNA filters and the SV40-specific RNA molecules were anlayzed by gel electrophoresis in 12% acrylamide — 7 *M* urea gels as in Figure 5; (m) ^{32}P-labeled tRNA markers. (From Abulafia, R., Ben-Ze'ev, A., Hay, N., and Aloni, Y., *J. Mol. Biol.*, 172, 467, 1984. With permission.)

VI. ATTENUATION IN SV40 AS A MECHANISM OF TRANSCRIPTION TERMINATION BY RNA POLYMERASE II

We shall now diverge from dealing with the mechanism of attenuation and show that attenuation in SV40 may serve to study the mechanism of transcription termination by RNA polymerase II. In the previous experiments the attenuator RNA was found in isolated nuclei, nuclear matrices, or VTC after pulse labeling with (α-^{32}P) UTP for a maximum of 5 min. These experiments were done using limited concentrations of UTP (0.2 to 0.3 μM). When the same experiments were repeated but incubation was carried out for longer times, or when a high concentration of UTP was used, or when pulse chase conditions were employed, there was a significant reduction in the production of the attenuator RNA, and almost complete readthrough of transcription through the attenuation site.

In order to verify whether the production of the 94- to 96-nt attenuator RNA is a result of pausing of the RNA polymerase, which is enhanced by limited UTP concentration or is a result of transcription termination, minichromosomes leached from isolated nuclei[118] and nuclei isolated from infected cells with a detergent-free hypotonic buffer were used.[98] Using these two experimental systems, it was found that efficient transcription termination occurs at low ionic strength but not at high ionic strength.

A. Introduction to the Minichromosome System

SV40 DNA is found in infected cells in the form of a minichromosome. It possesses a beaded structure composed of cellular histones and supercoiled viral DNA in a molecular complex which is very similar to that of cellular chromatin.[119-121] This has made SV40 an attractive model system in which to study the organization and expression of eukaryotic chromatin. The chromatin structure of the 5'-flanking region of several actively transcribed cellular and viral genes is known to be different from the bulk of chromatin. Such regions are generally hypersensitive to DNase I.[49,122-128] Consistent with this is the observation that ~25% of the SV40 minichromosomes display, precisely within the sensitive region, a stretch of DNA not contained within a typical nucleosome structure.[51,53,54] Transcription occurs on the SV40 minichromosomes,[84,129] and the late promoters map at the exposed region.[11,12,47,130] The possible role of the exposed region in directing the specificity of transcription initiation of the late genes is suggested by the observation that E. coli RNA polymerase and the eukaryotic polymerase II initiate transcription primarily within this region in vitro.[52,131] It is assumed that the higher-order structure of the actively transcribed minichromosome is involved at other levels of regulation during transcription.[127,132]

B. RNA Synthesis Directed by SV40 Minichromosomes Is a Salt-Sensitive Reaction

At 48 hr postinfection nuclei were prepared from SV40-infected cells and leached in hypotonic buffer to yield transcriptionally active viral minichromosomes.[118,133] Figure 10 illustrates the kinetics of RNA synthesis at 30°C under various salt concentrations. Incubation of minichromosomes for increasing lengths of time in a reaction mixture for RNA synthesis at moderate salt concentrations (30 mM ammonium sulfate or 100 mM NaCl) led to incorporation of labeled nucleotide with linear kinetics for only 5 min. On the other hand, incubating the minichromosomes at high salt concentrations (300 mM ammonium sulfate or 500 mM NaCl) results in continued (^{32}P)UMP incorporation. This effect is more pronounced with ammonium sulfate than with NaCl since incorporation proceeded linearly for as long as 60 min in the presence of the former salt.

C. Elongating Viral RNA Is Associated with the Minichromosomes While Prematurely Terminated RNA Detaches from the Template

To determine the relationship between the viral template responsible for transcription and

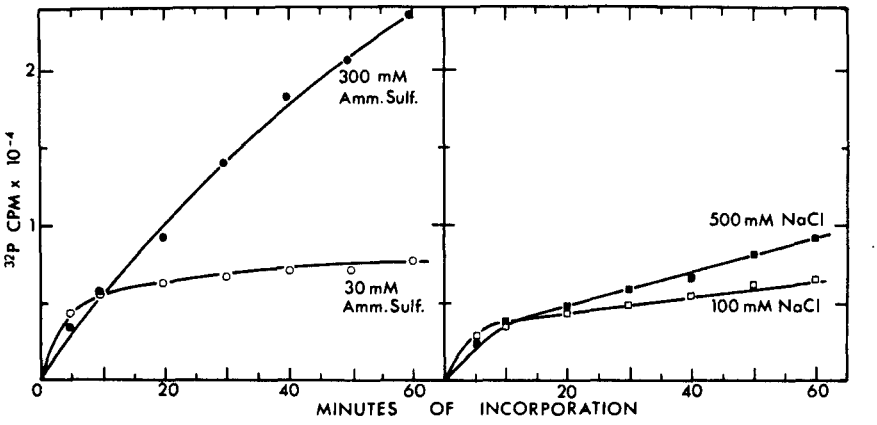

FIGURE 10. Kinetics of (^{32}P)UMP incorporation directed by nuclear extracts from SV40-infected cells. Nuclei were prepared by hypotonic lysis of the cells and the nuclear extracts obtained were supplemented with 2 m*M* $MnCl_2$, (α-^{32}P)UTP (40 μCi/100$\mu\ell$ final volume), 1 m*M* ATP, GTP, and CTP, and salt as indicated. Aliquots were removed and spotted on strips on Whatman 3 MM paper and precipitated with trichloroacetic acid. (From Pfeiffer, P., Hay, N., Pruzan, R., Jakobovits, E. B., and Aloni, Y., *EMBO J.*, 2, 185, 1983. With permission.)

the RNA made at the various salt concentrations, SV40-infected cells were labeled with (^{3}H)thymidine before extraction of the nuclei. The labeled minichromosomes were incubated for synthesis of ^{32}P-labeled RNA either at moderate or at high salts for varying times. The reactions were stopped by the addition of EDTA to 10 m*M* and the samples were loaded onto sucrose gradients for separation of the major viral nucleoprotein pools.[118]

As shown in Figure 11, the extracts incubated at moderate salts (100 m*M* NaCl; 30 m*M* ammonium sulfate) contained two major peaks of viral nucleoprotein, of which the 250S peak comprised mature and immature virions and the 75S peak the minichromosome pool of replicating and transcribing molecules.[118]

RNA synthesis for 5 min revealed two major populations, of which one was bound to the 75S structures reflecting attachment of RNA to the minichromosome template and a second population was present at the top of the gradient in a nonbound state. Increasing the time of incubation to 20 or 60 min led to almost total release of RNA from the 75S peak and to its accumulation at the top of the gradient as a template-free fraction. This was in contrast to the results obtained at high salt concentrations (500 m*M* NaCl; 300 m*M* ammonium sulfate) where no major fraction of unbound RNA was obtained. Rather, increasing the time of synthesis led to accumulation of RNA into the minichromosome peak as a species attached to its template. The fraction of RNA running ahead of the minichromosome peak suggests the presence of long RNA chains. As expected from the kinetics of incorporation (see Figure 10), the number of counts synthesized at high salt increased continuously with increasing time of synthesis, while synthesis at low salt rapidly plateaued.

It should be noted that the minichromosome peak displayed in the sucrose gradient profile at high salt was shifted to a 55S position in the gradient. This, as previously reported, is due to the loss of histone H1 from these molecules which leads to a reduction in the compaction of the minichromosome.[134,135] Minichromosome-associated and unbound RNAs were purified and hybridized to southern blots of the five restriction fragments shown in Figure 5. Figure 12 shows that the RNA associated with the minichromosome after 60 min incubation at low salt (100 m*M* NaCl) hybridized primarily with fragment e and, to a lesser extent, with fragment b which lies immediately downstream, reflecting poor growth of RNA chains. Labeled RNA found at the top of the gradient, on the contrary, hybridized exclusively with fragment e. The situation is even more pronounced at high salt (300 m*M* ammonium

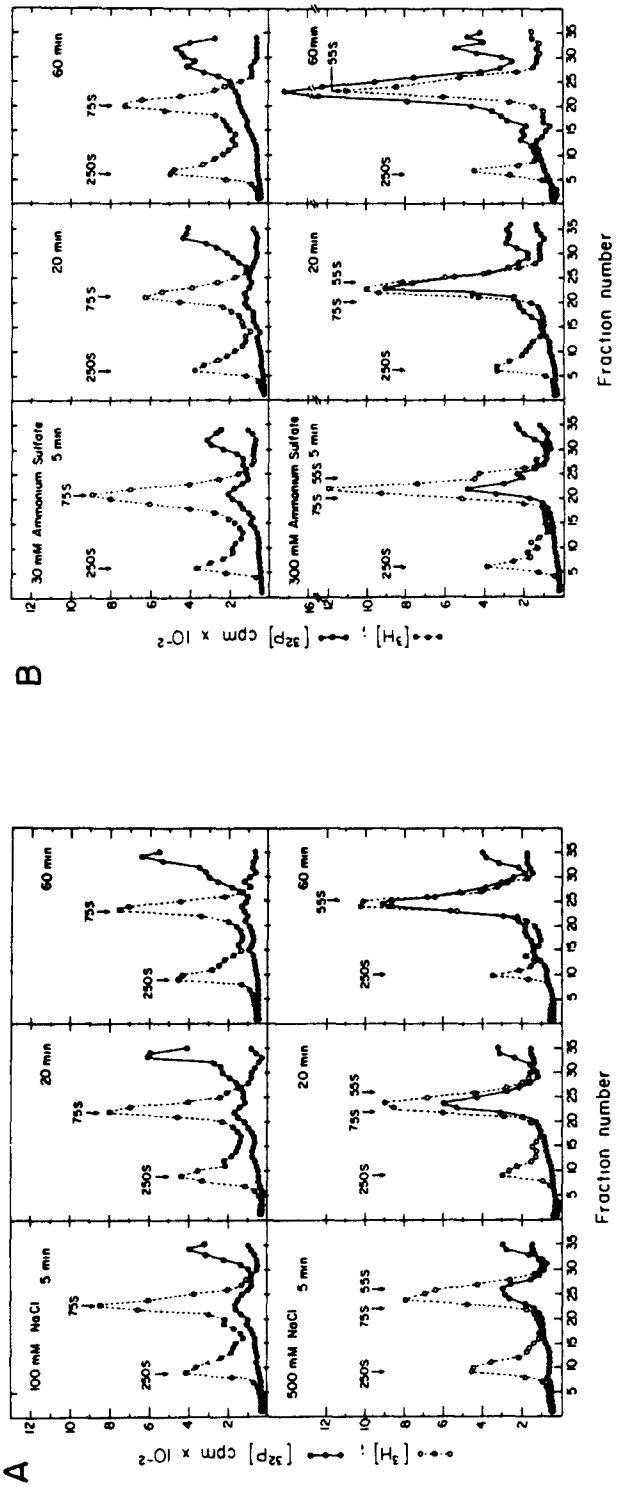

FIGURE 11. Sucrose gradient analysis of transcription mixtures programmed by nuclear extracts from SV40-infected cells. Nuclei were prepared by hypnotic leaching of 21 plates of CV-1 cells at 46 hr postinfection in 15 mℓ buffer 'H'[115] and the nuclear extract was concentrated by dialysis against PEG 20,000 and equilibrated overnight against buffer 'H'. Transcription mixtures (300 μℓ) containing 240 μℓ nuclear extracts were adjusted to the desired salt concentration ([A] NaCl; [B] ammonium sulfate) and supplemented with NTPs, $MnCl_2$, and 100 μCi (α-^{32}P)UTP as described.[95] The evolution of the acid-insoluble radioactivity upon incubation at 32°C was followed (see Figure 10) and after 5-, 20-, and 60-min total RNA was prepared from half the sample for hybridization analysis,[95] while the other half was chilled on ice and layered onto precooled 5 to 30% sucrose gradients in buffer 'H' supplemented with 100 m*M* NaCl. After centrifugation for 105 min in a Beckman SW41 rotor at 4°C and 32,000 rpm, the gradients were collected from the bottom and the acid-precipitable radioactivity was determined on 40 μℓ aliquots of each fraction spotted on strips of Whatman 3 MM® paper. The fractions containing the *minichromosome*-associated and free RNA (top of the gradient) were pooled and the RNA extracted for further analysis. (From Pfeiffer, P., Hay, N., Pruzan, R., Jakobovits, E. B., and Aloni, Y., *EMBO J.*, 2, 185, 1983. With permission.)

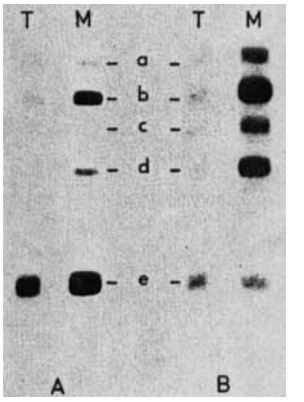

FIGURE 12. Hybridization pattern of the minichromosome-associated and free RNA to southern blots of the SV40 DNA restriction fragments depicted in Figure 5. RNA synthesized by nuclear extracts at 100 m*M* NaCl (A) or 300 m*M* ammonium sulfate (B). T: free RNA remaining at the top of the gradient; M: minichromosome-associated RNA. (From Pfeiffer, P., Hay, N., Pruzan, R., Jakobovits, E. B., and Aloni, Y., *EMBO J.*, 2, 185, 1983. With permission.)

sulfate), where minichromosome-associated RNA hybridized to the "late" fragments (b, d, and e) as well as to the "early" fragments (a and c). This result does not discriminate between hybridization to the "late" or "early" strands but indicates, on the other hand, distribution of RNA polymerase molecules over almost all the viral genome. In contrast, RNA found at the top of the gradient hybridized primarily with the promoter-proximal e fragment. This indicates that SV40-specific RNA found detached from the template both after low and high salt incubations does not originate from random degradation or accidental release of growing RNA chains, but rather points to the existence of a well-defined class of RNA molecules prematurely terminated and released from their template. Analysis by gel electrophoresis indicated that the RNA found at the top of the sucrose gradient contained the 95-nt attenuator RNA.[95]

D. The Effect of Ionic Strength on the Production of the Attenuator RNA in Isolated Nuclei

The mechanism of transcription termination at the attenuation site was studied using isolated nuclei, since it is considered to be a more complete system in comparison to the minichromosome system.

Figure 13 shows the results of experiments carried out in order to verify the effect of salt concentration on the production of the attenuator RNA in isolated nuclei. The incubation mixture contained salt of either high (100 m*M* $[NH_4]_2SO_4$) (A) or low (100 m*M* NaCl) (B) ionic strength. In the presence of 100 m*M* $(NH_4)_2SO_4$ following a 5-min pulse with 0.3 μM (α-^{32}P) UTP, a major intense band of 95 nt is revealed (Figure 13A, P). However, when the "pulse" was followed by a "chase" the majority of the radioactivity in the 95-nt band disappeared and appeared at the upper part of the gel (Figure 13A, C_1). Only a small fraction (~10%) of the 95-nt at RNA was unchasable. These results indicate that during the pulse the majority of the RNA polymerase molecules paused but did not terminate transcription at the attenuation site. A major band of 95 nt was also produced following a 5-min "pulse" in the presence of 100 m*M* NaCl (Figure 13B, P). However, in contrast to incubation with ammonium sulfate, "chase" for 10 or 30 min under the present conditions of low ionic strength had almost no effect on the intensity of the 95-nt band (Figure 13B, C_1 and C_2).

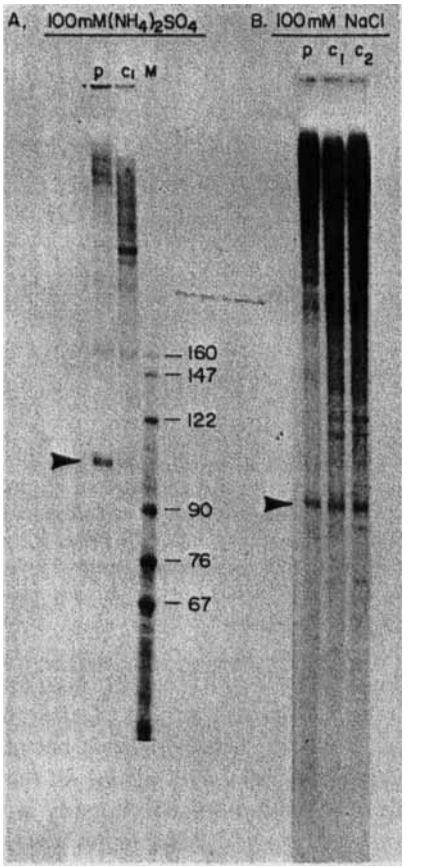

FIGURE 13. Size analyses of viral RNAs produced in isolated nuclei incubated in the presence of salt of high ([A] 100 m*M* $[NH_4]_2SO_4$) and low ([B] 100 m*M* NaCl) ionic strength under "pulse" and "pulse-chase" conditions. (P) Nuclei were incubated in a transcription reaction mixture containing 100 μCi (0.3 μ*M*) of (α-^{32}P) UTP for 5 min ("pulse"); (C_1) nuclei were incubated as in (P), but following the "pulse" 20 μ*M* of unlabeled UTP were added and transcription continued for 10 min ("chase"); (C_2) as in (C_1), but transcription in the presence of unlabeled UTP continued for 30 min. The labeled RNAs produced were purified and analyzed by gel electrophoresis. The arrows point to the position of the 95 nt attenuator RNA; (M) size markers obtained by end labeling of the Hpa II restriction fragments of pBR322 DNA. (From Hay, N. and Aloni, Y., *Nucl. Acids Res.*, 12, 1401, 1984. With permission.)

These results indicate a real and efficient transcription termination rather than pausing of the RNA polymerase molecules at the attenuation site.

The results of Figure 14 provide additional support for a real transcription termination at the attenuation site under conditions of low ionic strength. In this experiment, isolated nuclei were incubated in the presence of 10 μ*M* (α-^{32}P) UTP for 20 min. It is evident that the 95-nt attenuator RNA was produced when the incubation mixture contained 100 m*M* NaCl (Figure 14A) but not when the incubation mixture contained 100 m*M* $(NH_4)_2SO_4$ (Figure 14B). The ~64-nt band was identified to be the SAS-RNA,[58-60] which is produced from the 3′ end of the late primary transcript.[213]

Figure 15 shows that readthrough which leads to the disappearance of the 95-nt band during the "chase" or prolonged in vitro incubation depends on the ionic strength and is

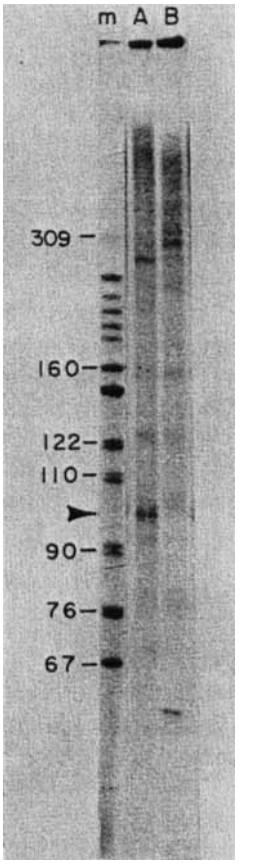

FIGURE 14. Size analyses of viral RNAs produced in isolated nuclei, incubated in the presence of salt of high and low ionic strengths for a long duration. Nuclei were incubated in a transcription reaction mixture containing 10 μM (α-^{32}P) UTP for 20 min in the presence of salts of low (100 mM NaCl) in (A) or high (100 mM $(NH_4)_2SO_4$) in (B) ionic strengths. The labeled RNAs produced were purified and analyzed by gel electrophoresis. The arrow points to the position of the 95 nt attenuator RNA. (m) Size markers as in Figure 13. (From Hay, N. and Aloni, Y., *Nucl. Acids Res.*, 12, 1401, 1984. With permission.)

not specific for ammonium sulfate. It is evident that the 95-nt band also disappeared and there was almost a complete readthrough when the incubation mixture contained either 0.2 or 0.3 M NaCl but not 0.1 M NaCl. (The ionic strength of 0.1 M $[NH_4]_2SO_4$ is equivalent to 0.25 M of NaCl.) In the following experiments the standard transcription reaction mixture contained 0.1 M NaCl and 10 μM UTP.

E. Possible Involvement of Cellular Factor in Transcription Termination at the Attenuation Site

The effect of high salt concentrations in the incubation mixture on the production of the attenuator RNA raised the possibility that a salt-soluble factor is involved in transcription termination at the attenuation site. To verify this possibility, nuclei of uninfected cells were washed with 300 mM NaCl, precipitated by low speed centrifugation, and the soluble supernatant fraction was concentrated by vacuum dialysis against a buffer containing 100 mM NaCl. Two dialysis tubes were used, one with a M.W. cutoff of 10 k and the second

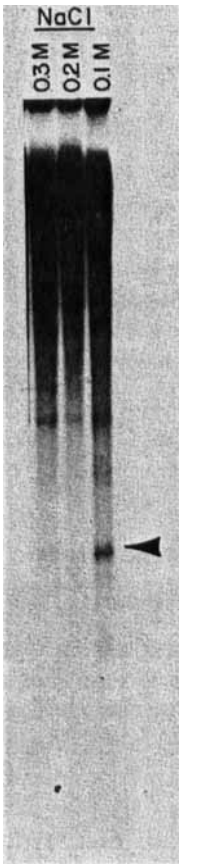

FIGURE 15. Size analyses of viral RNAs produced in isolated nuclei in the presence of various salt concentrations. Nuclei were incubated in a transcription reaction mixture containing 10 μM (α-^{32}P) UTP for 15 min in the presence of various salt concentrations, as indicated. The labeled RNAs produced were purified and analyzed by gel electrophoresis. The arrow points to the position of the 95 nt attenuator RNA. (From Hay, N. and Aloni, Y., *Nucl. Acids Res.*, 12, 1401, 1984. With permission.)

with a M.W. cutoff of 75 k. These fractions were designated "10 kd HSF" and "75 kd HSF", respectively. In order to determine whether these high salt-soluble fractions contain a transcription termination factor, nuclei of SV40-infected cells were washed with 300 mM NaCl and divided into four fractions. The first contained the high salt prewashed nuclei (Figure 16a), the second contained the same as the first but, in addition, the "10 kd HSF" (Figure 16b), the third contained the same as the first but, in addition, the "75 kd HSF" (Figure 16c), and the fourth contained the same as the third but the "75 kd HSF" was heated at 80°C for 10 min before adding it to the high salt prewashed nuclei (Figure 16d). Following preincubation for 5 min at 20°C the various components of the standard transcription reaction mixture were added and incubation was allowed to proceed for 15 min. Figure 16a shows that washing the nuclei with high salt (300 mM NaCl) abolished the production of the attenuator RNA (see Figures 13, 14, and 15). Figure 16 (b and c) shows that transcription termination at the attenuation site was restored upon addition of the two "HSF". The results of Figure 16 show that the activity of the "HSF" was inactivated by heating.

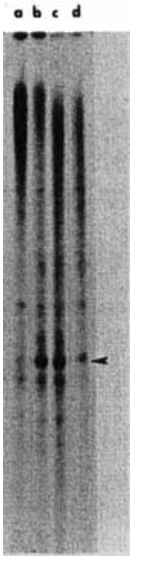

FIGURE 16. Size analyses of viral RNAs produced in isolated nuclei with and without the presence of "HSF". Isolated nuclei were washed with 300 m*M* NaCl and preincubated for 5 min either without "HSF" (a), with "10 kd HSF" (b), or with "75 kd HSF" (c). (For details see Hay and Aloni[98]); (d) is the same as (c) but the "75 kid HSF" was heated at 80°C for 10 min before addition to the standard transcription reaction mixture. "HSF" extracted from 1×10^8 nuclei was added to 3×10^7 nuclei. Incubation was then carried out for 15 min at standard reaction mixture. The labeled RNAs produced were purified and analyzed by gel electrophoresis. The arrow points to the position of the 95 nt attenuator RNA. (From Hay, N. and Aloni, Y., *Nucl. Acids Res.*, 12, 1401, 1984. With permission.)

The interpretation of the above results is that a nuclear factor or a multimer of it having a M.W. higher than 75 k and which is heat labile is involved in transcription termination by RNA polymerase II. It is worth noting that Leer et al.[136] have found a high salt-soluble factor which is involved in transcription termination of a ribosomal RNA, i.e., a polymerase I transcript, and it has been suggested that a high salt-soluble factor is involved in transcription termination of histone mRNA, i.e., a polymerase II transcript.[137,138]

F. The Secondary Structure of the Attenuator RNA Is Involved in Transcriptrion Termination

The resemblance of the possible secondary structure of the attenuator RNA (see Figure 8) to the termination signal of the prokaryotic polymerase[96] and the observation that the intercalating drug, proflavine, abolishes attenuation[91] (see Figure 9) led us to suggest that the secondary structure of the attenuator RNA is involved in pausing and transcription termination of the RNA polymerase II transcribing SV40. In the following experiment this suggestion was studied more directly, by replacing GTP with its analog ITP in the standard transcription reaction mixture. In bacteria and bacteriophages, ITP was shown to abolish transcription termination by destabilizing RNA stem-and-loop structure.[139-142]

As shown in Figure 17, there is a significant increase in readthrough transcripts when ITP replaced GTP in the standard reaction mixture (Figure 17B and 17A, respectively). Incorporation of ITP into RNA can reduce the stability of either RNA-DNA or RNA-RNA

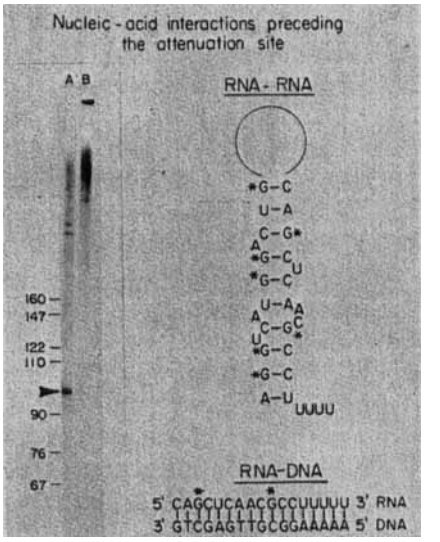

FIGURE 17. Size analyses of viral RNAs produced in isolated nuclei incubated in the presence of GTP or ITP in the transcription reaction mixture. (A) A standard reaction mixture; (B) ITP (400 μM replaced GTP in the standard transcription reaction mixture. Transcription was allowed to proceed for 15 min. The labeled RNAs produced were purified and analyzed by gel electrophoresis. The arrow points to the position of the 95 nt attenuator RNA. Possible RNA-RNA and RNA-DNA interactions at the 3′ end of the attenuator RNA (see Figure 5F) are also represented. The (*) indicates a G-residue which was replaced by an I-residue during transcription in the presence of ITP. The position of the size markers are shown on the left side of the autoradiogram. (From Hay, N. and Aloni, Y., *Nucl. Acids Res.*, 12, 1401, 1984. With permission.)

hybrids. In the case of SV40 attenuator RNA, it is evident that incorporation of ITP has a much higher effect on the stability of the RNA-RNA as compared to the RNA-DNA hybrid.

As illustrated in the schematic base pairing in Figure 17, it is evident that in the RNA-DNA hybrid the incorporation of ITP can destabilize only two of the eight G-C bonds immediately preceding the attenuation site. In contrast, in the RNA-RNA hybrid, the incorporation of ITP can destabilize all the G-C bonds, thus eliminating the entire RNA secondary structure. This may suggest that it is the RNA stem-and-loop structure which is essential for transcription termination.

High ionic strength can also stabilize RNA secondary structure and, therefore, lead to an increased production of the attenuator RNA. Indeed, during a short in vitro incubation the production of the attenuator RNA is enhanced when incubation is performed at high salt concentrations as compared to low salt concentrations (see Figure 13). However, as discussed above, the production of the attenuator RNA under these conditions is due to pausing of the RNA polymerase at the attenuation site rather than transcription termination.

G. Conclusions

The existence of a premature transcription termination process, i.e., attenuation, during SV40 late transcription provides a good system for studying the mechanism of transcription termination by the eukryotic RNA polymerase II. The work described above indicates that a nuclear factor and the RNA secondary structure are essential regulatory elements involved

in transcription termination. We suggest that interactions similar to those involved in prokaryotic transcription termination[143] operate in eukaryotes. We further suggest that, similar to the mechanism of transcription termination in prokaryotes, the intramolecular hairpin structure in the RNA transcript (resulting from dyad symmetry in the template) impedes the progress of the RNA polymerase and leads it to pause at the start of the U-residues which follow it. Whether the RNA polymerase pauses or terminates transcription probably depends on a delicate balance between the rate of formation of the RNA hairpin and the stability of the rU-dA hybrid region.[143] This balance can be affected by the in vitro conditions. Indeed, we have observed in the present studies that a combination of low UTP concentration and high ionic strength enhances the pause of the RNA polymerase at the terminator. Transcription termination, is, however, inefficient. The most efficient transcription termination was found to occur with a combination of limited UTP and low ionic strength. We, therefore, speculate that high ionic strength in the transcription reaction mixture has two effects: (1) it stabilizes the RNA hairpin structure which enhances pausing of the RNA polymerase; and (2) it removes a nuclear factor which is involved in transcription termination and, therefore, reduces the efficiency of the termination process. At low ionic strength transcription termination is efficient due to the formation of the RNA hairpin structure, the exceptionally unstable rU-dA interaction, and the presence of a termination factor. The termination factor apparently brings the enzyme to a complete stop and leads to the release of the RNA transcript from the template. The increased termination observed in the present study, at low UTP concentration, is, presumably, due to the slower elongation rate which increases the duration of the pause. As a result there could be a better chance for the termination factor to interact with the RNA polymerase and to lead it to terminate transcription. That the concentrations of the ribonucleotides may influence the extent of termination has also been observed in prokaryotes.[144]

Support for the conclusion that the eukaryotic polymerase II can respond to the prokaryotic terminator is provided by the observation that in a cell-free eukaryotic system, the bacteriophage lambda 4S RNA terminator caused human polymerase II to pause on the template and to partially terminate transcription of transcripts initiated at the adenovirus 2 major late promoters.[145] The observation that only partial termination occurred and that the polymerase failed to respond to the lambda t_{int} and t_{R1} terminators may indicate the lack of a termination factor in the in vitro system and/or the involvement of a specific sequence in the termination process.

Similar terminators were shown to occur at the 3′ end of other polymerase II transcripts, such as in the minute virus of mice (MVM), an autonomous parvovirus (see below), and at the 3′ end of human U_1 RNA.[146] It is worth noting that a termination site has been identified for the mouse β-major globin gene that is transcribed by RNA polymerase II. This site is located about 1000 bases downstream from the poly(A) site, but the nature of the terminator is still unknown.[147] The mammalian mitochondiral polymerase appears to respond to a similar termination signal at the 3′ end of the 16S rRNA;[148] likewise, polymerase III responds to a similar terminator at the 3′ end of the VA RNA in Ad.[149] Moreover, based on the DNA sequence of several viruses we have suggested that there is a potential of forming terminators at promoter proximal regions.[96] Terminators which contain only the hairpin but not the run of Us or only the run of Us were, also, reported. Birchmeier et al.[150] have shown that the 3′ terminus of histone mRNA may be defined by dyad symmetry 30 to 40 bases downstream from the protein stop codon. These sequences are necessary for generating the 3′ terminus, although downstream sequences are, also, involved. AT-rich dyad symmetries with no U-residues following them were, also, found at regions involved in transcription termination of various yeast genes.[151] This terminator resembles the rho-dependent terminator in bacteria. The run of Us was shown as an essential element for RNA polymerase III transcription termination.[152] Based on these observations we suggest that there are several mechanisms

for transcription termination in eukaryotes. The SV40 attenuator resembles the rho-independent terminator of prokaryotes. Nevertheless, the finding that a cellular factor is involved in transcription termination at the SV40 attenuator is not surprising, because it is possible that in prokaryotes the rho factor actually is needed to recognize these signals in the bacterium and it enhances termination in vitro.[2] It is interesting to note that in order to see the effects of the rho factor and the cellular factor in vitro, it is necessary to use conditions of relatively low ionic strength.

It has been previously shown that a cellular factor with an estimated M.W. of 50 to 100 k is involved in transcription termination of rRNA in Tetrahymena-a polymerase I transcript.[136] Similarly, a cellular factor with estimated M.W. of 200 to 250 k was needed to generate the 3′ end of histone mRNA, a polymerase II transcript.[137,138] Both termination factors were extracted from nuclear fractions with high salt buffers. Similar to the rho factor these eukaryotic transcription termination factor(s) could be active in multimeric form.[2] Our factor as well as the others are still crude, and further purification and characterization are needed to determine whether there are more than one eukaryotic transcription termination factors. In this regard it is interesting to note that based on idiosyncrasies in termination efficiencies of histone gene transcription in xenopous oocyte and HeLa cells, Bendig and Hentschel have suggested the presence of termination factors which are relative species or even tissue specific.[153]

We would like to mention the possibility that the SV40 attenuator is functioning also as a terminator when the RNA polymerase reaches it again during transcription of the circular genome. A dual function of the SV40 terminator is supported by the observation that, following injection of SV40 DNA into xenopous oocytes, RNA species corresponding in lengths to the 94- to 98-nt attenuator RNA[214] and to a genome length transcript accumulate.[154]

We shall now proceed with our discussion of the attenuation mechanism in SV40 with the suggestion that the leader protein (agnoprotein) has a transactive negative effect on SV40 late transcription.

VII. THE LEADER PROTEIN (AGNOPROTEIN) ENHANCES PREMATURE TERMINATION

The late leader region of SV40 encodes a 7900-Mr (61 amino acid) protein termed agnoprotein[17] (see Figure 2). The agnoprotein is highly basic with a turnover rate of ~2 to 3 hr[155] (see below). It has a high affinity for both single- and double-stranded DNA.[155] Agnoprotein is not essential for virus growth in tissue culture, although an altered phenotype is observed in deletion mutants which lack its coding region. These mutants are characterized by small plaques as well as slower growth.[155-158] The exact role of the agnoprotein in the virus life cycle is unknown. However, three functions have been suggested: (1) it interacts with the major capsid protein VP_1 in the process of encapsidation;[158,159] (2) it has a transacting positive effect on late transcription;[60] and (3) we have postulated that the leader protein in SV40, like the leader peptide in the biosynthetic operons in bacteria, has a transactive negative effect on late transcription.[91,92] As an approach to defining this function we have determined the following: (1) its subcellular distribution in SV40-infected cells and (2) the correlation between its synthesis, at various times after infection, and the level of premature termination using wild-type virus and an insertion mutant Δ79 (a generous gift of Dr. G. Khoury). Δ79 Is a viable insertion mutant in which two nucleotides were inserted at the Hpa II site[155] (see Figure 20). Consequently, whereas the RNA secondary structure transcribed from the promoter-proximal region is maintained, this mutant is unable to synthesize an authentic leader protein. The results and conclusions of these experiments are as follows: first, the predominant localization of the agnoprotein is in the cytosol. However, a minor detectable fraction of agnoprotein can be found associated with the nuclear matrix and in

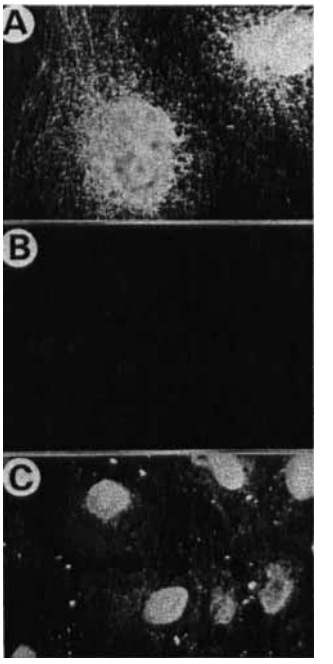

FIGURE 18. Indirect immunofluorescent staining of SV40-infected cells with antibodies against agnoprotein and T-antigen. (A) SV40-infected BSC-1 cells were stained with antibody against agnoprotein, that was raised by injecting agnoprotein from SDS gels into a rabbit. The second antibody was rhodaminated goat antirabbit antibody. Before staining, the cells were fixed with 3.7% formaldehyde in PBS and permeabilized with 1% Triton X-100® in PBS for 10 min; (B) uninfected BSC-1 cells stained as above; (C) SV40-infected BSC-1 cells as in (A) were stained with anti-T antibody and rhodaminated goat antirabbit antibody as in (A). (From Aloni, Y., Hay, N., Skolnik-David, H., Pfeiffer, P., Abulafia, R., Pruzan, R., Ben-Asher, E., Jakobovits, E. B., Laub, O., and Ben-Ze'ev, A., *Developments in Molecular Virology IV*, Kohn A. and Fuchs, P., Eds., Martinus Nijhoff Publishers, Boston, 1983, 1. With permission.)

association with the viral minichromosome.[215] An indirect immunofluorescence method with antibodies raised against purified agnoprotein or synthetic peptides[216] confirmed the biochemical fractionation procedure. In Figure 18 it can be seen that rabbit antiserum against the agnoprotein showed an intensive granular fluorescence (Figure 18A). The same immune antiserum was nonreactive with mock-infected cells (Figure 18B). Figure 18C shows a similar experiment with antibodies raised against SV40 T-Ag. It is evident that the pattern of fluorescence in Figure 18A indicates the dominant localization to be the perinuclear and cytoplasmic regions with a granular appearance. Detectable fluorescence was also observed in the nucleus. In contrast, Figure 18C shows an absolute nuclear localization of the SV40 large T-Ag in the same set of cells. Similar results were obtained by Nomura et al.[161] Second, Figure 19 shows the time course of appearance of the agnoprotein. It is interesting to note that the synthesis of the structural proteins VP_1, VP_2, and VP_3 is not increased from 42 to 50 hr postinfection, rather, on the contrary, it appears that there is even a somewhat reduced synthesis of these viral proteins. In contrast, the agnoprotein is synthesized at the highest

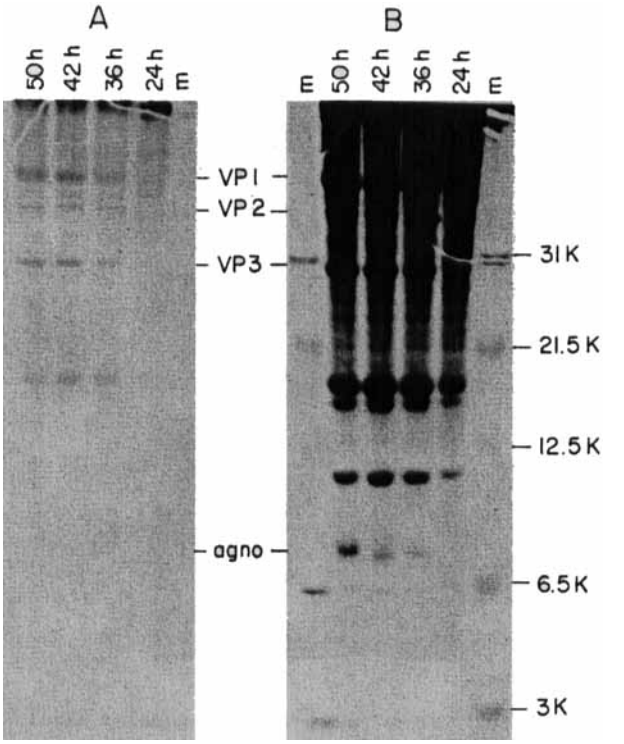

FIGURE 19. Time course of the synthesis of the leader protein (agnoprotein) of SV40. BSC-1 cells (0.5×10^6 cells) were infected with WT SV40. At various times after infection cells were labeled for 3 hr with ^{14}C-arginine, collected in 0.1 mℓ sample buffer (10 m*M* sodium phosphate, pH 7.2, 7 *M* urea, 1% SDS, 1% β-mercaptoethanol, 0.01% bromophenol blue), lysed, and analyzed on a 15% polyacrylamide gel containing 0.1% SDS in 6 *M* urea. The gel was fluorographed and dried. The autoradiogram was obtained by exposing the dried gel to Kodak XAR-V® film at −80°C. (m) Designates molecular weight markers. (From Aloni, Y., Hay, N., Skolnik-David, H., Pfeiffer, P., Abulafia, R., Pruzan, R., Ben-Asher, E., Jakobovits, E. B., Laub, O., and Ben-Ze'ev, A., *Developments in Molecular Virology IV*, Kohn, A. and Fuchs, P., Eds., Martinus Nijhoff Publishers, Boston, 1983, 1. With permission.)

rate at the latest time after infection (50 hr). Pulse-chase experiments performed at 50 hr postinfection indicate a half life of 2-3 hr for the agnoprotein (see below). The short half-life of the agnoprotein and its accumulation at the latest time after infection indicate a potential in vivo regulatory function for the leader protein which is most pronounced at the latest time after infection. Indeed, a direct correlation between the accumulation of the agnoprotein and the production of the 95 nucleotide attenuated RNA in isolated nuclei was observed. Moreover, cells infected with the viable insertion mutant Δ79 showed no agnoprotein and almost no premature termination.[217] Based on these results we conclude that the function of the agnoprotein is to stabilize the viral RNA conformation that leads to premature termination. In addition, a potential salt-sensitive rho-like factor may mediate the release of the attenuated RNA from its template.[98]

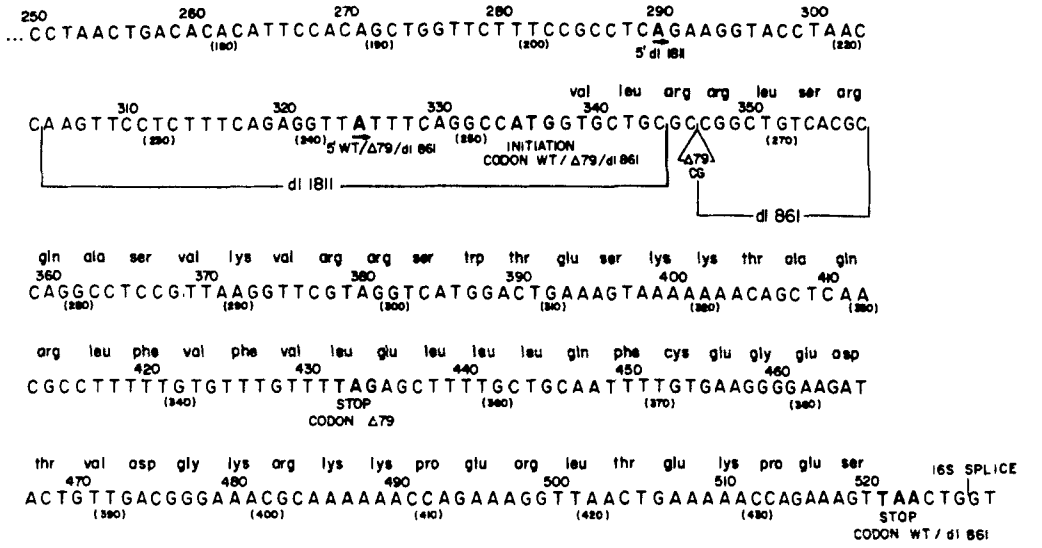

FIGURE 20. Nucleotide sequence of the major late leader region of SV40 mRNAs. Nucleotide residues are numbered according to Tooze[11] (upper numbers). The underlined numbers in brackets are according to Reddy et al.[17] The bracket lines indicate the deletion regions of mutants d11811 (not discussed) and d1861. The position of insertion of CG in the DNA of mutant Δ79 is also indicated. The arrows show the position of major initiation sites of transcription of W.T. and mutant late RNA. Initiation codons and stop codons of W.T. agnoprotein, d1861 agnoprotein, and Δ79 agnoprotein are designated.

VIII. d1861 SYNTHESIZES AGNOPROTEIN SHORTENED BY FOUR AMINO ACIDS

A. Introduction

d1861 Is a viable deletion mutant of SV40 that has been only partially characterized.[60,156] From nuclease S_1 analyses it was estimated that d1861 is missing 16 to 25 nucleotides downstream from the major initiation site (residue 325)[11] within the late leader region. Based on the site and extent of the deletion two predictions were made: (1) the major initiation site for late transcription is shifted upstream[47,48,60] and (2) d1861 is incapable of synthesizing the agnoprotein, encoded by the major leader region of SV40 late RNA.[60]

The suggestion that agnoprotein has transactive positive effect on late transcription was based on mixed infection experiments involving WT SV40 and d1861.[60] Given that it contradicts our suggestion, we decided to investigate d1861 more carefully.[162] The results obtained show that in comparison to WT, d1861 is missing 12 nucleotides downstream and in-phase with the AUG start codon of agnoprotein (residues 347 to 358) (see Figure 20). Furthermore, in spite of the deletion, the major initiation site for late transcription is the same as that in WT SV40 (residue 325)[11] and d1861 is still capable of synthesizing agnoprotein.[162] However, in comparison with WT agnoprotein, d1861 agnoprotein is shorter by four amino acids, it is produced in a lesser amount, and it has a shorter turnover rate ($t_{1/2}$ ~ 1 hr). Below is a description of representative experiments.

B. Does d1861 Actually Synthesize Agnoprotein?

In order to determine whether the agnoprotein encoded by d1861 is actually produced in virus-infected cells, the following experiment was performed. Cells were infected with WT virus, d1861, and the mutant Δ79 in which two nucleotides were inserted at the Hpa II site.[155] Δ79 Encodes for a 32-amino acid polypeptide (see Figure 20). At 45 hr postinfection the cells were labeled for 2 hr with ^{14}C-arginine and protein analyses were carried out by

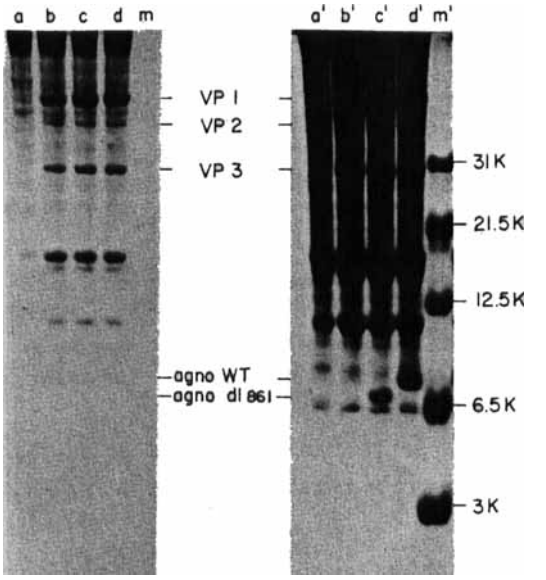

FIGURE 21. Protein analysis of cells infected with SV40 W.T., SV40 mutant Δ79, SV40 mutant d1861, and uninfected cells. At 45 h.p.i. of BSC-1 cells with SV40, the cells were labeled with ^{14}C arginine for 2 hr and protein analysis was done as described in Hay et al.[162] (Lanes a and a') Proteins of mock infected cells. (lanes b and b') proteins of Δ79 infected cells; (lanes c and c') proteins of d1861 infected cells; (lanes d and d') proteins of W.T. infected cells; (lanes m and m') molecular weight markers. (A) Exposure of the gel for 12 hr; (B) exposure of the gel for 48 hr. (From Hay, N., Kessler, M., and Aloni, Y., *Virology*, 137, 160, 1984. With permission.)

polyacrylamide gel electrophoresis. The gel was exposed to X-ray films for short and long durations (Figure 21). Following a short exposure, only the major viral proteins VP_1, VP_2, and VP_3 are apparent. Moreover, it is evident that the three structural proteins were synthesized to about the same extent in cells infected with Δ79 (b), d1861 (c), or WT (d). Agnoprotein was not revealed in either of the infected cells. However, following a long exposure of the gel to an X-ray film the agnoprotein was recognized in cells infected with WT (d') as well as with d1861 (c'), although in the latter, as expected, the agnoprotein is definitely smaller. The 32-amino acid polypeptide was not recognized in Δ79-infected cells even following a shorter pulse with ^{14}C-arginine and longer exposure (data not shown), indicating that if this polypeptide is synthesized it is extremely unstable. Taking into account that d1861 agnoprotein contains 2 arg less than WT agnoprotein, it appears that four to five times less d1861 agnoprotein was labeled with ^{14}C-arginine during 2 hr as compared with WT agnoprotein. This could result either from a slower rate of synthesis, from a shorter half-life time, or from both. These possibilities were studied next.

C. The Rate of Synthesis and the $t_{1/2}$ of d1861 Agnoprotein

In order to determine the rate of synthesis and the turnover rate of d1861 agnoprotein and to compare these parameters with WT agnoprotein, cells were infected with d1861 or WT and were starved for arginine and then pulse-labeled for 10 min with ^{14}C-arginine. Following the pulse, the cells were washed and fresh medium containing cold arginine was added for

a chase period. Immediately after the pulse and at various times during the chase, equal samples were taken for protein gel analysis. The gel was then exposed to X-ray film for a short duration for the analysis of WT agnoprotein (Figure 22A) and for a long duration for the analysis of the d1861 agnoprotein (Figure 22B). The bands corresponding to the two agnoproteins were also scanned. Figure 22C represents the calculated average area under the corresponding peaks. It is evident from Figure 22C that during a 10-min pulse there is already about five times more label in WT agnoprotein than in d1861 agnoprotein. Moreover, it is apparent that whereas the $t_{1/2}$ of WT agnoprotein is about 3 hr, that of d1861 is only 1 hr. The finding that d1861 agnoprotein is synthesized at a slower rate than WT agnoprotein was somewhat surprising, because it could indicate a lower presence of RNA containing the full length of the agnoprotein coding frame, i.e., contiguous major leader of 16S mRNA. Remember, that the major initiation site for late transcription was found to be at the same residue for both WT and d1861 and that the structural proteins VP_1, VP_2, and VP_3 are equally produced in cells infected with the two viruses.[162] Alternatively, it is possible that the AUG initiation codon of d1861 agnoprotein is not as available for ribosome binding as that of the WT agnoprotein. These two possibilities are not mutually exclusive.

IX. A MODEL FOR QUANTITATIVE REGULATION OF SV40 GENE EXPRESSION BY ATTENUATION AND mRNA MODULATION IN A FEEDBACK CONTROL MECHANISM

A. Regulation of Synthesis of VP_1

We have shown that in SV40 viral RNA initiated at nucleotide 243 terminates in vitro 95 nucleotides downstream, at a typical prokaryotic transcription termination structure, suggesting that an attenuation mechanism, resembling attenuation in prokaryotes, may regulate SV40 late transcription. The same 95 nucleotides comprise the 5′ end of a 202 nucleotide leader (nucleotides 243 to 444) that is spliced to the body (nucleotides 1381 to 2592) of the 16S mRNA that encodes for the structural protein VP_1 (see Figure 2).

The structural relationship of the 16S leader to the leader sequences of prokaryotic mRNAs is striking and raises the possibility that, like them, it, also, participates in an attenuation mechanism by influencing the expression of its downstream structural gene VP_1. The 16S leader sequence, like those of prokaryotes, encodes information for a small protein.

We have noticed two pairs of inverted complementary repeat sequences in the attenuated RNA and at the 5′ end of the major 16S mRNA.[91] These two pairs are designated 1 + 2 and 3 + 4 in Figure 23. In the nucleus, conformation 3 + 4 serves as a typical transcription termination structure that has been implicated in the process of premature termination.[4,8,9,69,74,76,77] When conformation 1 + 2 is first produced during transcription, the production of the competing conformation (2 + 3) (Figure 23B) is prevented. Thus, the base pairing that occurs between 1 and 2 appears to be most important in the stabilization of the base pairing in the terminator region 3 + 4. A striking observation is that in conformation 1 + 2, 3 + 4, the agnogene AUG and the adjacent four nucleotides are sequestered in the stem 1 + 2, leaving only five nucleotides for ribosome binding to the 16S mRNA in the cytoplasm. A second AUG (nucleotides 303 to 305) is sequestered in the stem of 3 + 4. We predict that when conformation 1 + 2, 3 + 4 prevails at the 5′ end of the 16S mRNA, the AUG initiation codon for translation of VP_1 (nucleotides 1423 to 1426) is the first to be encountered by the ribosome, and, thus, VP_1 is synthesized. Conversely, when conformation 2 + 3 prevails at the 5′ end of the 16S mRNA, the agnogene AUG is accessible for ribosome binding and the agnoprotein is translated. We speculate that modulation between the two conformations is a fundamental element in a feedback control mechanism that regulates the amount of 16S mRNA synthesized in the nucleus and the amounts of VP_1 produced in the cytoplasm. For such control at the attenuator site, a deficiency or surplus

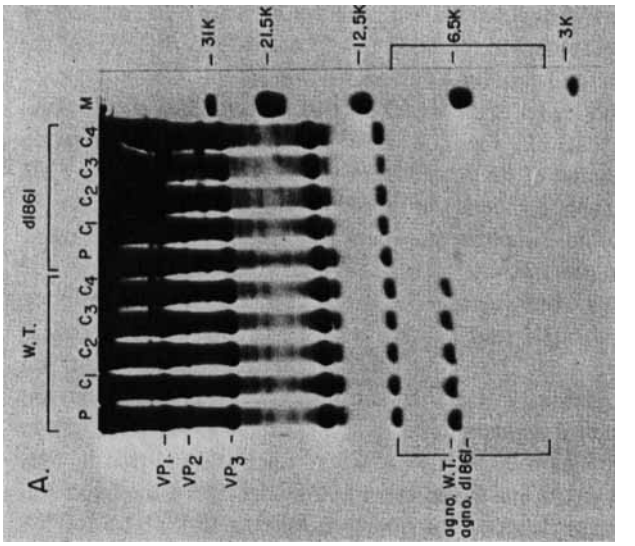

FIGURE 22. The half-life of W.T. and d1861 agnoprotein. At 48 hr after infection with SV40 W.T. and d1861 the cells were starved for arginine for 3 hr. Then 20 μCi/mℓ of ^{14}C-arginine was added for 10-min pulse (P). Following the pulse the cells were washed and medium-containing cold arginine was added for chase periods. After the pulse and during the chase period, samples were taken for protein analysis as described in Hay et al.[162] (P) 10-min pulse, (C_1) $1^1/_2$-hr chase, (C_2) 2-hr chase, (C_3) 3-hr chase. (A) The dried gel was exposed for 15 days to Kodak XAR-V® film at −80°C; (B) longer exposure of the region indicated by brackets in (A). The positions of W.T. and d1861 agnoproteins are designated; (C) graphical representation in which each point in the curve represents result of scanning of the band of agnoproteins after pulse and after chase periods. ●—●, W. T. agnoprotein; ○—○, d1861 agnoprotein. (From Hay, N., Kessler, M., and Aloni, Y., *Virology*, 137, 160, 1984. With permission.)

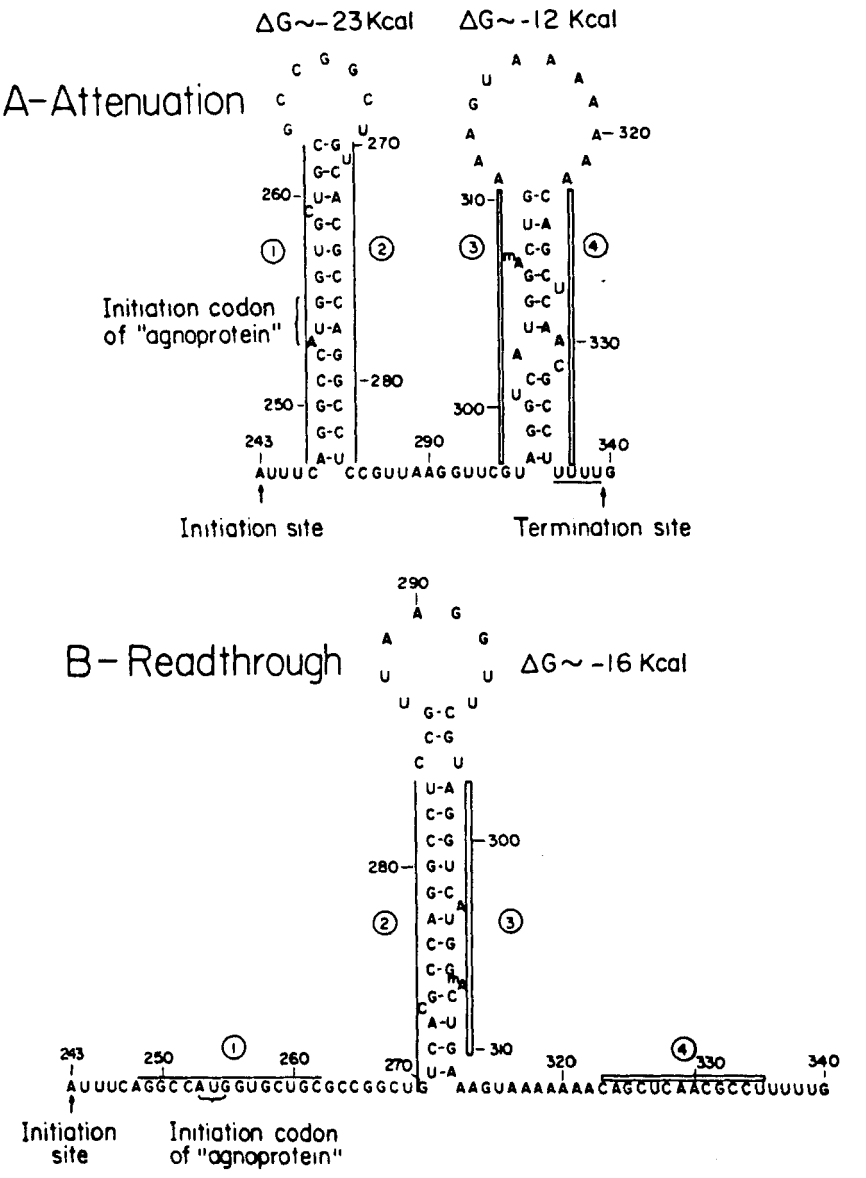

FIGURE 23. Schema of alternative conformation in the attenuated RNA and at the 5′ end of the 16S mRNA. (A) Attenuation conformation showing sequestration of initiation codon of agnoprotein in the stem 1 + 2, and a typical termination signal in 3 + 4; (B) readthrough conformation in which the initiation codon of agnoprotein is available for ribosome attachment. For details see text. The ΔG were calculated as described by Timoco et al.[211] (From Hay, N., Skolnik-David, H., and Aloni, Y., *Cell*, 29, 183, 1982. With permission.)

of VP_1 must be sensed and communicated to the transcribing RNA polymerase molecule. Our prediction is that the concentration of the agnoprotein in the cell is the distinctive feature communicated to the transcribing polymerase in the regulation of termination at the attenuator. According to this suggestion, the amount of agnoprotein in the cell determines, at least in part, the equilibrium between the alternative conformations of the 95 nucleotides at the 5′ end of the transcript. The suggestion that the agnoprotein is a molecule that transfers information from the cytoplasm to the nucleus is supported by the observations that it can

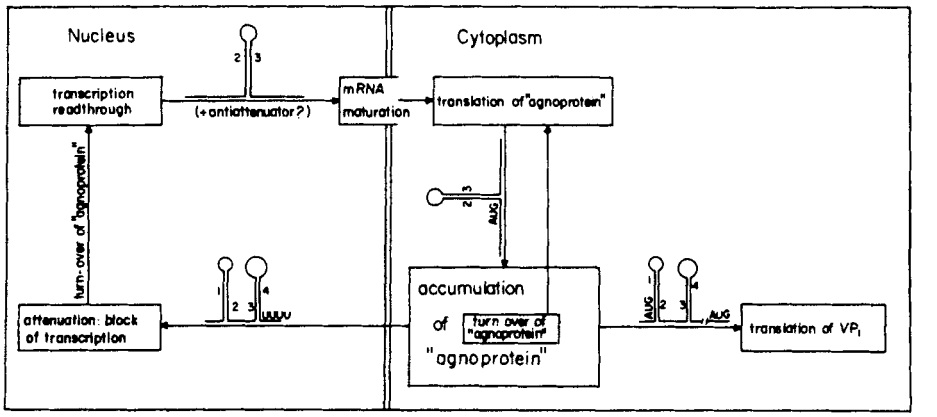

FIGURE 24. A model for quantitative regulation of the synthesis of VP_1 by attenuation and mRNA modulation in a feedback control mechanism. Image 1 + 2, 3 + 4 is (A) of Figure 23, and image 2 + 3 is (B) of Figure 23. For details see text and Hay et al.[91,92]

be isolated from both the cytoplasm and the nucleus (see Figure 18), that it can be found in association with SV40 minichromosomes,[91,161] and that it is a nucleic-acid-binding protein.[155,161] Moreover, the short half-life (2 to 3 hr) of the agnoprotein[155] (see above) is suitable for that of a regulatory protein.

The details of the model are illustrated in Figure 24.[91] During transcription of the 5′ end of the transcripts, initiating at nucleotide 243, a certain fraction, arising from the readthrough 2 + 3 conformation, is synthesized into a complete primary transcript. We speculate that an antiattenuator factor is needed to facilitate the 2 + 3 conformation. The primary RNA transcripts are processed to mature 16S poly(A)$^+$ mRNA and are transported to the cytoplasm. In this conformation, the accessible agnogene AUG is encountered by the ribosome, and 61 triplets are translated to the agnoprotein. Once a critical concentration of the agnoprotein is attained, it serves as a repressor of its own synthesis by binding to its mRNA and stabilizing conformation 1 + 2, 3 + 4 (see Figure 25). As a result of this RNA folding, the ribosome encounters the initiation codon for the translation of VP_1, and the structural protein is synthesized. At the same time, the agnoprotein is transported to the nucleus and stabilizes conformation 1 + 2, 3 + 4, leading to premature termination (Figure 24). Consequently, when the VP_1 is synthesized in the cytoplasm, transcription in the nucleus is aborted. Reinitiation of the process occurs once the concentration of the agnoprotein is reduced below a critical level.

B. Regulation of Synthesis of VP_2

The most abundant group of 19S mRNA (the mRNA of VP_2 and VP_3) is characterized by a gap of 184 nucleotides which extends exclusively from residue 292 to 475 (see Figure 3). Thus, as a result of the splicing process 152 of the 202 nucleotides of the 16S leader are missing, consequently, the 19S mRNA cannot encode for the agnoprotein. Instead, in the leader of 19S mRNA there is an open reading frame for the synthesis of a 29-amino acid protein.[48] This protein shares the first 13 amino acids with the agnoprotein and overlaps with the AUG of VP_2 (Figure 26). Hence, a small overlapping region of the SV40 genome can potentially code for the synthesis of two different proteins in two reading frames (see Figure 3), a situation that finds precedent in the SV40 genome in the C terminus of VP_2/VP_3 and the N terminus of VP_1. We have designated this postulated,[48,92] but as yet undiscovered, protein "agnoprotein-29".[92] As a result of the splicing process the nucleotides comprising the 3 + 4 stem-and-loop structure of Figure 23 are also removed. As a result

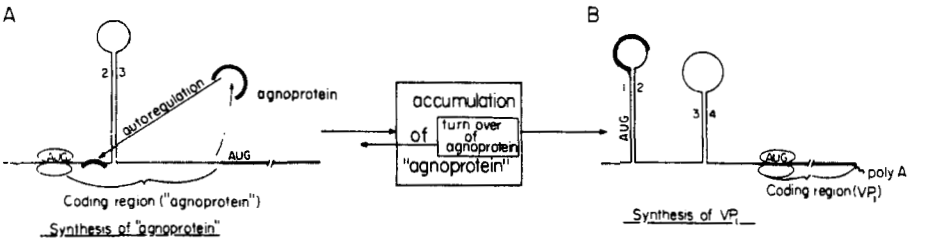

FIGURE 25. Agnoprotein regulates the conformation at the 5′ end of the 16S mRNA. For details see text and Hay et al.[91,92]

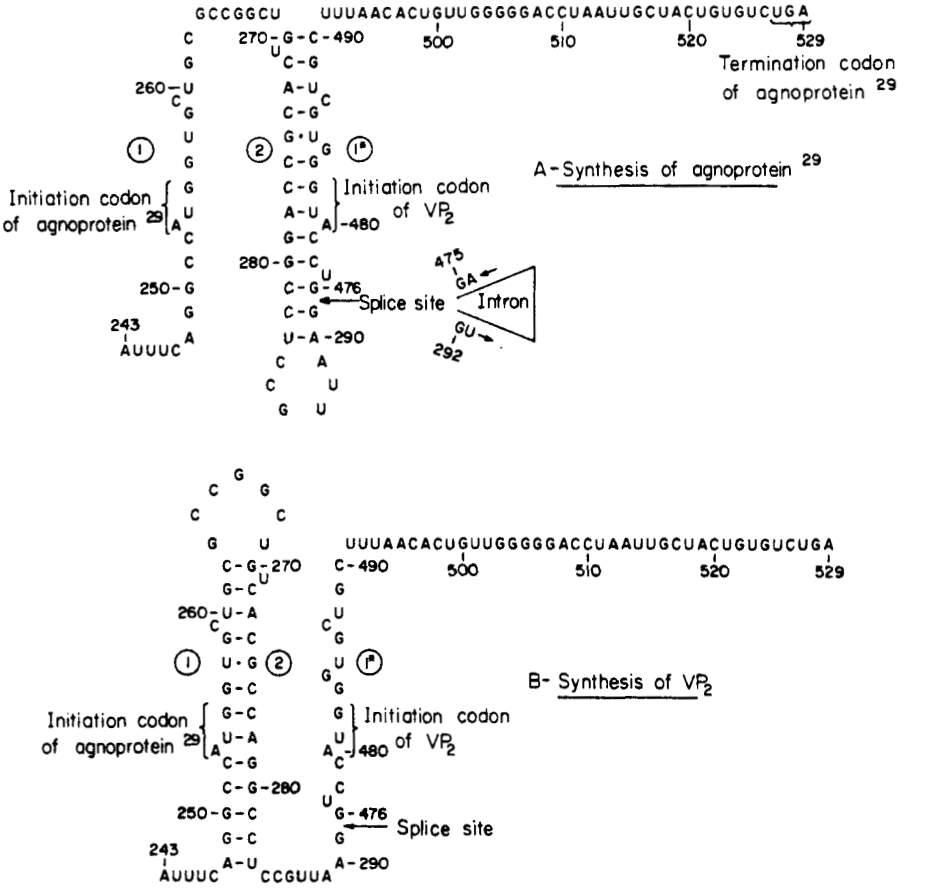

FIGURE 26. Schema of alternative conformation at the 5′ end of the 19S mrNA. (A) Conformation showing sequestration of initiation codon of VP_2 while the initiation codon of "agnoprotein-29" is available for ribosome attachment. The sequence shows the entire reading frame of "agnoprotein-29" (B) conformation showing sequestration of initiation codon of "agnoprotein-29"; while the initiation codon of VP_2 is available for ribosome attachment. For details see, also, Aloni and Hay.[92] Residue numbers are according to Reddy et al.[17]

in the 19S mRNA it is not possible to form the same alternative conformations as were suggested for the 16S mRNA in the regulation of the synthesis of VP_1. Since these alternative conformations (see Figure 23) are fundamental elements in our model[91] substitute sequences must exist if the synthesis of VP_2 and VP_3 are regulated by a feedback mechanism similar to that regulating the synthesis of VP_1. Indeed, 17 out 18 nucleotides preceding and including

the AUG of the agnoprotein are identical to the nucleotides preceding the AUG corresponding to the initiation codon for VP_2[163] (see strand 1 and 1* in Figure 26). Consequently, two alternative conformations can be predicted for the sequences at the 5′ end of the 19S mRNA (see Figure 26). In conformation A, the agnogene-29 AUG is available for ribosome binding and the VP_2 AUG is sequestered in the stem 2 + 1*. Conversely, in conformation B the agnogene-29 AUG is sequestered in the stem 1 + 2 and the VP_2 AUG is available for ribosome binding.

Similar to the regulation of synthesis of VP_1[91] we speculate that modulation between the two alternative conformations is a fundamental element in a feedback control mechanism that regulates the amount of 19S mRNA synthesized in the nucleus and the amount of VP_2 produced in the cytoplasm. At the level of transcription the regulation between attenuation and readthrough is based on the same alternative conformations as suggested for the synthesis of 16S mRNA (see Figure 23). This is because the 16S and 19S mRNAs are spliced from the same primary transcript. That is, differential splicing occurs after the decision to attenuate or not.[11,12] Again, as in our original model,[91] the prediction is that the concentration of the agnoprotein, either that of the 16S or 19S or both, is the distinctive factor communicated to the transcribing polymerase in the regulation of termination at the attenuator. According to this suggestion, the amount of agnoprotein in the cell determines, at least in part, the equilibrium between the alternative conformations of the 95 nucleotides at the 5′ end of the primary transcript and at the 5′ end of the 19S mRNA in the cytoplasm.

C. Regulation of Synthesis of VP_3

Because of the uncertainty about the nature of the VP_3-mRNA,[11,12] it is difficult to speculate whether the synthesis of VP_3 is regulated by a similar feedback mechanism as that which regulates the synthesis of VP_1 and VP_2. The current notion is that VP_3 is synthesized, at least in part, from a 19S mRNA with a leader that is different from that of the VP_2-mRNA.[11,12] Ghosh et al.[47] suggested that in the 19S species lacking any gap or containing a small gap between residues 444 and 476, base pairing can occur between sequences at and adjacent to the VP_2 AUG and a proximate segment of RNA which is retained in these 19S RNA species (see Figure 27). This base pairing may make the AUG of the VP_2 translation initiator at residue 480 to 482 inaccessible to ribosomes. It is interesting to note that in such a situation the AUG of the agnoprotein translation initiator would be sequestered in the 1 + 2 stem (see Figure 27). Thus, as discussed for the regulation of synthesis of VP_1 and VP_2, the stabilization of conformation 1 + 2 is a prerequisite for the synthesis of VP_3. It is noteworthy that the 19S RNA species lacking any gap or containing a small gap between residues 444 and 476 would encode information for the same 61 amino acid agnoprotein as that encoded in the leader of the 16S mRNA. Since the VP_3 mRNA retains the same sequences involved in the modulation of the secondary structure at the 5′ end of the 16S mRNA (see Figure 23), the production of VP_3 could be regulated by the same feedback mechanism that regulates the synthesis of VP_1. However, an additional base pairing should be formed in order to make the AUG of VP_2 translation initiator inaccessible to ribosomes.*

D. Antiattenuation

We must bear in mind that the models presented here and in our previous communications,[91,92] however complex, are still not complete. Thus, we can postulate, as in bacteria and their viruses, the existence of a positive control element that suppresses transcription termination, namely, a molecule that acts as an antiattenuator leading to the extension of the aborted RNA beyond the termination site.[218] Of the several possible mechanisms sug-

* We would like to note that the model of the role of the agnoprotein is open now to direct tests by in vitro and in vivo transcription and translation assays.

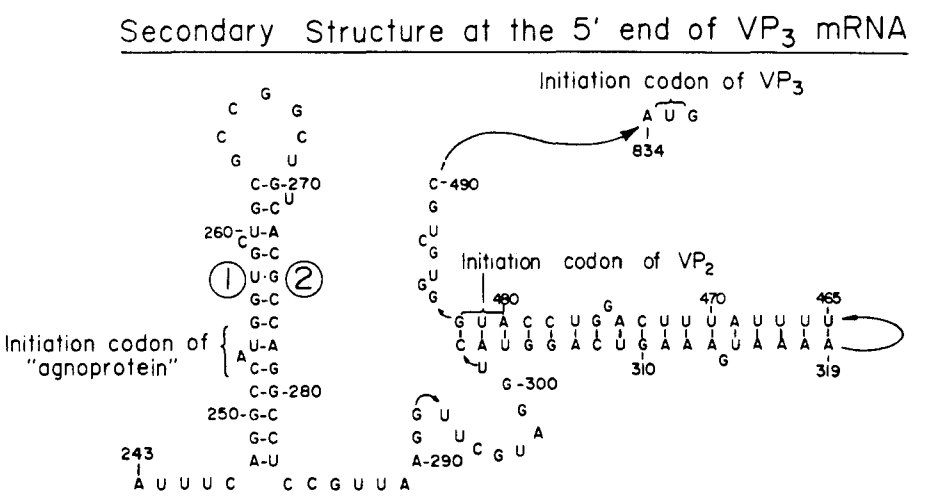

FIGURE 27. Schema of a conformation at the 5' end of VP_3 mRNA. The conformation shows sequestration of initiation condons of agnoprotein and VP_2. The initiation codon of VP_3 is available for ribosome attachment. The base-pairing showing sequestration of initiation codon of VP_2 is from Ghosh et al.[47] For details see, also, Aloni and Hay.[92] Residue numbers are according to Reddy et al.[17]

gested previously by which antiattenuation is achieved in prokaryotic systems,[4,5,9,69,74,76,77] we prefer the following two: (1) the antiattenuator interacts with a subunit of the RNA polymerase, enabling it to transcribe through the termination signal; (2) the antiattenuator binds to the nascent RNA affecting the secondary structure and preventing transcription termination. This might be accomplished by the movement of the nascent RNA under the action of the antiattenuator into a cellular domain where constraints influence RNA secondary structure formation. Relevant to this proposition is our observation that attenuated RNA is not attached to the nuclear matrix while readthrough transcripts are attached.[97]

E. Agnoprotein Fulfills a Role Which in the Attenuation Mechanism in Amino Acid Biosynthetic Operons of Bacteria is Achieved by a Successful Synthesis of the Leader Peptide

The similarities between our model and the attenuation model in amino acid biosynthetic operons of bacteria is striking.[4,9] If SV40 has inherited the attenuation control mechanism from its prokaryotic ancestors, then it had to devise mechanisms to overcome the nuclear membrane barrier which uncouples transcription and translation in eukaryotes. It appears that the modification that the eukaryotic virus has made is using the leader protein itself instead of using the translation process of the leader peptide per se.[4,9] In the prokaryotic systems successful synthesis of the leader protein and not the final product is the distinctive feature which signals premature termination via stabilization of the 3 + 4 RNA conformation.[4,9] In the eukaryotic system translation has no direct effect on transcription, because the two processes occur in separate cellular compartments. Instead, in SV40, the leader protein itself is transported from the cytoplasm to the nucleus where it might stabilize the 1 + 2, 3 + 4 RNA conformation which can lead to transcription termination (see Figure 24). An additional striking similarity between the prokaryotic and eukaryotic systems is the location of the pause sites. In both systems, it is located at the 1 + 2 stem-and-loop structure. Figure 28 summarizes the similarities shared by the prokaryotic and eukaryotic systems. It should be mentioned that other modifications or supplements that relate to the different natures of the prokaryotic and eukaryotic RNA polymerase are conceivable.

As an approach to determine whether attenuation is a specialized mechanism of the SV40

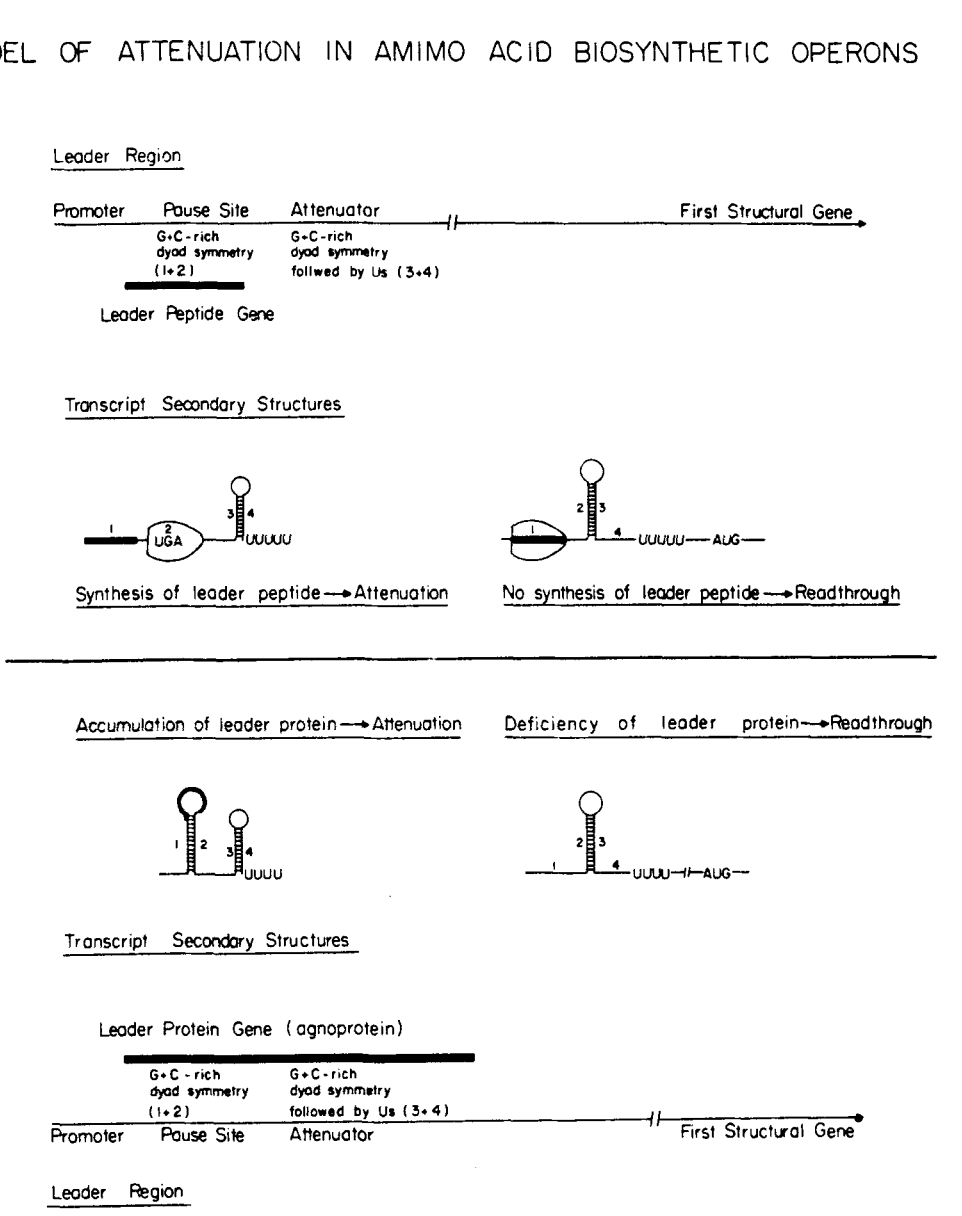

FIGURE 28. Similarities shared by the model of attenuation in amino acid biosynthetic operons of bacteria and the model of attenuation in SV40 late operon. For details see text. (Adapted from Kolter, R. and Yanofsky, C., *Ann. Rev. Genet.*, 16, 113, 1982. With permission.)

system or a widespread mechanism operating in animal viruses and cells, we have analyzed the pattern of transcription of the autonomously replicating parvovirus, minute virus of mice (MVM).

X. TRANSCRIPTION OF MINUTE VIRUS OF MICE, AN AUTONOMOUS PARVOVIRUS, MAY BE REGULATED BY ATTENUATION

A. Introduction to the MVM System

The autonomously replicating parvoviruses are the smallest viruses known and have, therefore, been used as a model system for studies of DNA replication[164-167] and RNA transcription[168,169] in animal cells. For comprehensive review articles see the book published by The Cold Spring Harbor Laboratory.[170] The infectious virions are icosahedral particles, with diameters of 18 to 30 mm, and contain linear single-stranded DNA of about 1.5×10^6 daltons.[171] The autonomously replicating parvoviruses package only the minus-coding strand of the viral DNA (v-strand) in contrast to the adeno-associated parvovirus, a defective subgroup which packages both DNA strands.[165,171]

Minute virus of mice (MVM) is one of the group of autonomously replicating parvoviruses of rodents that includes H-1, H-3, and Kilham rat virus.[172] These viruses share a high degree of nucleotide sequence homology and can complement each other for many functions.[173-176]

Initial studies on the transcription of the DNA from MVM and H-1 suggested that the genomes of these viruses encode only a single transcription unit with a promoter near mu 4.[168,169] However, recent studies have demonstrated that three major transcripts are produced from two overlapping transcription units with separate promoters positioned near the left end (4 mu) and the middle (39 mu) of both MVM and H-1 genomes.[177,178] The complete nucleotide sequences of MVM and H-1 have also recently been determined.[178,179]

In a recent study[180] we have mapped the initiation sites of the two overlapping transcription units of MVM and have determined the activities of their promoters. In addition, we have found that attenuation mechanism is functioning in vitro at the left promoter-proximal region yielding 142 nt RNA. The DNA sequences at the attenuation site are strikingly similar to those of the termination signal in prokaryotes and of the SV40 attenuator. We propose that RNA transcription in MVM may be regulated by attenuation. Below is a description of a few representative experiments.

B. Transcription of MVM May Be Regulated by Attenuation

As defined by Yanofksy,[4] attenuation is a mechanism whereby a site within the operon causes pausing or premature termination of transcription. This site which acts conditionally to allow either pausing and termination or transcriptional readthrough is termed attenuator site.[4] If attenuation regulates MVM transcription one expects to find in vivo short RNA of a discrete length. Unfortunately, in eukaryotes primary RNA transcripts and nonfunctional RNAs are extremely unstable and it is difficult to identify them. Thus, to date, termination sites for any pre-mRNA have not been identified.[181] As an approach to overcome this difficulty isolated nuclei and various transcriptional systems are being used. One of them is VTC which has been shown in SV40 to represent at least 90% of the viral transcriptional complexes that generate viral RNA in intact nuclei and, presumably, in vivo.[182-186] We, therefore, assume that the VTC system can represent the pattern of in vivo transcription. Thus, as a result of attenuation it is expected to find on the VTC a sharp decrease in the number of nascent RNA beyond an attenuation site. To verify this possibility, VTC was extracted from the infected cells at 24 hr postinfection and the in vivo preinitiated nascent RNA was elongated in vitro for 5 min in the presence of 1 μM (α-^{32}P)-UTP. The nascent RNA is elongated by about 20 nt under these conditions,[94,185,186] allowing determination of their in vivo distribution along the DNA template.[83] The (^{32}P) RNA was then purified and hybridized to a blot of Eco RI restricted clone of MVM in pBR322. Just before elution of the RNA, the blot was cut (Figure 29E) so that the fragment carrying the left promoter at 4 mu (fragment "c" — spanning residues 1/1088) was separated from the other two fragments "a" and "b", carrying the rest of the MVM genome, including the right promoter at 39

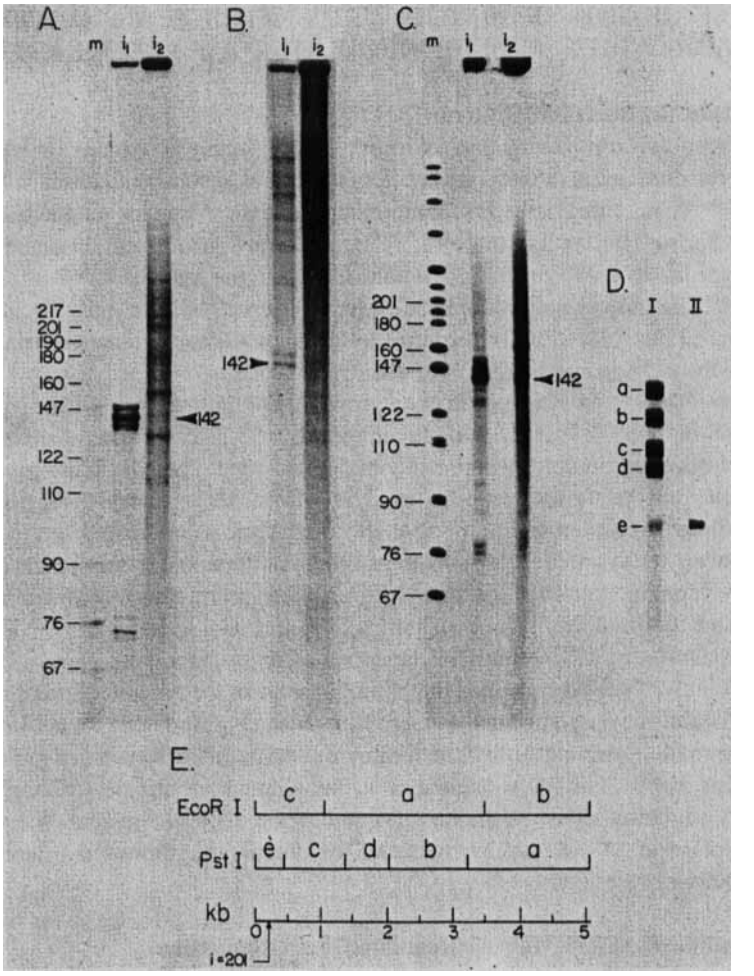

FIGURE 29. Size analysis of viral RNA synthesized from the two promoters of MVM. (A) In vitro transcription of VTC in the presence of 1 μM (α-^{32}P)-UTP was allowed to proceed for 5 min; (B) in vitro transcription of VTC in the presence of 10 μ*M*, (α-^{32}P)-UTP allowed to proceed for 10 min; (C) in vitro transcription of VTC in the presence of 10 μ*M* UTP and 1 μ*M* (α-^{32}P)-CTP allowed to proceed for 5 min. RNAs were extracted and viral RNAs were selected by hybridization to and elution from Eco RI cleaved MVM DNA blotted on nitrocellulose (see text); (i_1) RNA transcripts eluted from the Eco RI "c" fragment; (i_2) RNA transcripts eluted from the Eco RI "a" + "b" fragments. The ^{32}P-labeled viral RNAs were subjected to electrophoresis on 10% acrylamide gel; (m) size markers; (D, I) hybridization of ^{32}P-MVM DNA labeled by nick translation with Pst I fragments of MVM DNA; (D, II) the RNA in the 142-nt band shown in A(i_1) was eluted and hybridized to a Pst I cleaved MVM DNA; (E) restriction map of MVM DNA cut with Eco RI and Pst I.[179] Note that MVM DNA inserted in pBR 322 was cleaved with Eco RI and RF DNA was cleaved with Pst I. Only the viral DNA fragments are represented. (From Ben-Asher, E. and Aloni, Y., *J. Virol.*, 52, 266, 1984. With permission.)

mu. Thus, nascent RNA of up to about 900 nt initiated at the left promoter were separated from nascent RNA molecules initiated at the right promoter. The eluted RNA was analyzed by gel electrophoresis under denaturing conditions. The results in Figure 29A show that a discrete major band is eluted from the fragment containing the left promoter (i_1) with an apparent length of 142 nt. Two minor bands above and below the major band are also

recognizable. However, almost no longer viral RNA transcripts are revealed, as indicated by the low level of radioactivity above the major bands. This observation is consistent with a mechanism of in vivo pausing or premature termination at a site located 142 nt downstream from the left initiation site. The low level of long nascent RNA is not an artifact caused by the loss of long RNA molecules during preparation of the VTC, because heterologous size RNA was eluted from the fragments containing the right promoter, as indicated by the occurrence of radioactivity along the gel as well as at the origin of the gel (i_2, in Figure 29A). Moreover, Figure 29B (i_1) shows that when in vitro incubation was carried out for 10 min in the presence of 10 μM (α-^{32}P) UTP, the 142-nt band remained major, at least on a molar basis, but long RNA hybridizing to Eco RI fragment "c" appeared. It is noteworthy that with this UTP concentration the rate of synthesis is about 100 nt/min (our unpublished results). The level of the 142-nt band depended, therefore, both on the UTP concentration and on the in vitro incubation time. Note also that under these conditions more and longer RNA hybridized to the Eco RI fragments "a" and "b" carrying the right promoter (i_2).

Several groups have reported that RNA polymerase transcribing certain templates in vitro pauses transiently at particular sites on the DNA, resulting in a nonuniform rate of chain elongation, and thereby generating discrete size classes of RNAs detectable as bands on acrylamide gels.[187,188] At least in some cases such pauses in transit are apparently related to the use of low concentration of nucleoside triphosphates. We have, therefore, tested whether the transcription block which leads to the production of the 142-nt band is alleviated and if other bands appear by reducing the concentration of CTP to 1 μM and increasing the concentration of UTP to 10 μM. The results of this experiment are shown in Figure 29C (i_1). In comparison to Figure 29A (i_1) it is apparent that two bands remained, the 142-nt band and the minor band above it. However, the minor band below the 142-nt band of Figure 29A (i_1) disappeared and a new minor band of about 150 nt appeared. Furthermore, the minor bands of about 70 nt which appeared in Figure 29A (i_1) disappeared and new minor bands of about 80 nt appeared. Based on these results it is concluded that the appearance of the 142-nt band is not due to limited UTP concentration in the reaction mixture, whereas the production of the minor bands depends, at least in part, on the limited concentration of the particular nucleotide.

To delineate the region of the genome which serves as the template for the 142-nt band the RNA was eluted from the 142-nt band shown in Figure 29A (i_1) and hybridized to a blot of Pst I-cut MVM DNA (Figure 29E) followed by RNase treatment of the blot. Figure 29D (I) shows that whereas uniformly labeled (^{32}P) MVM DNA hybridized extensively with fragments a to d, less ^{32}P-DNA hybridized with fragment e due to its smaller size. Figure 29D (II) shows that the RNA of the major band hybridized exclusively to fragment e. These results establish that the RNA in the major band is transcribed from a region of the genome spanning residues 201 to 411.[179]

C. Mapping of the 142-nt RNA

In order to localize the initiation site of the 142-nt RNA, Eco RV-cleaved VTC was prepared, divided into two fractions, and incubated in vitro under "pulse" and "pulse chase" conditions. Labeled RNAs complementary to Eco RI fragment c (see Figure 29) were purified from the two fractions and analyzed by gel electrophoresis under denaturing conditions. Figure 30 I (P) shows that following a "pulse" a major band of 142 nt and two minor bands, similar to those shown in Figure 29A (i_1), were produced. However, in contrast to Figure 29A (i_1) a new band of 182 nt now appeared. This band contains runoff transcripts at the Eco RV cleavage site. Figure 30 I (C) shows that following the "chase" the 182-nt band became major, while the intensity of the 142-nt band was considerably reduced. These results indicate that following the chase the radioactivity in the 142-nt band ran off at the Eco RV cleavage site. As Eco RV cuts MVM DNA at residue 383 (see Figure 30 III) we

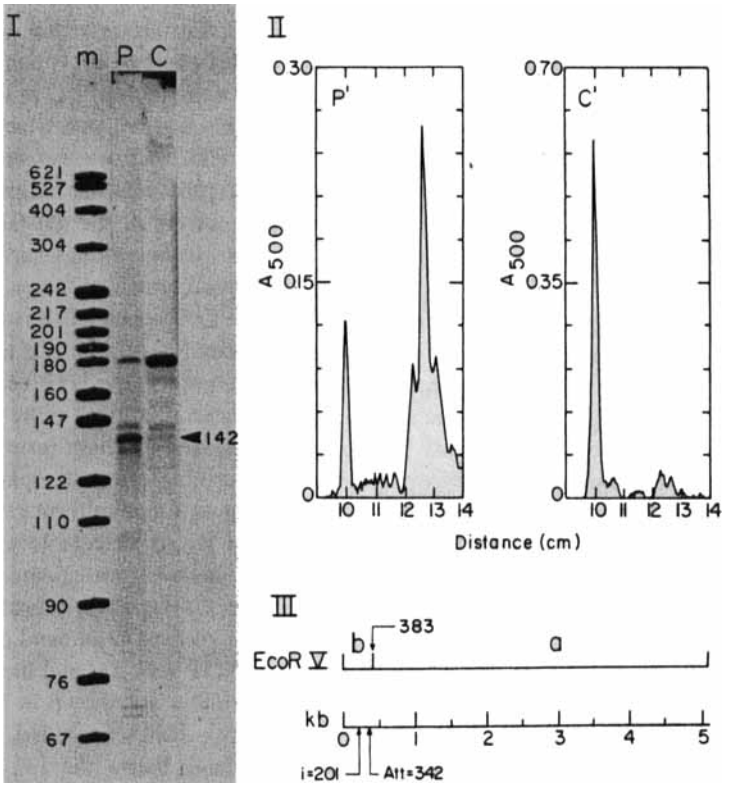

FIGURE 30. Mapping of the 5′ end of the 142-nt RNA. (I) VTC digested with Eco RV was incubated in vitro in the presence of 1 μM (α-^{32}P) UTP for 5 min (P) or incubated for 5 min as in P followed by 30 min incubation in the presence of 200 μM UTP (C), (^{32}P) RNA was selected by hybridization to and elution from Eco RI fragment "c" (see Figure 29). The viral RNA was subjected to gel electrophoresis. m — Size markers; (II) phototracing of the gel in (I) by a scanner (Beckman DU-8). The absorbance at wavelength 500 nm vs. distance from the origin of the gel was plotted. P′ and C′ correspond to P and C in 1; (III) restriction map of Eco RV-cut MVM DNA indicating the residues numbers (v-strand) of the Eco RV cleavage site[179] and the left initiation (i) and attenuation (Att.) sites. (From Ben-Asher, E. and Aloni, Y., *J. Virol.*, 52, 266, 1984. With permission.)

conclude that the initiation site of the RNA in the 142-nt band is at residue 201 (=383-182). According to these calculations in vitro pausing or premature termination occurs 142 nt downstream from the left initiation site, namely, at residue 342.

A scan of the autoradiogram reveals that while most of the radioactivity in the 142-nt band was chased, about 10% of it was unchaseable (Figure 30 II-C′). In vitro termination of transcription at this specific site may involve simply the cessation of chain elongation by the enzyme-DNA complex, namely, "pausing", or it may be accompanied by the actual release of the RNA chain. It is worth noting that in the *trp* attenuator the termination event in vitro does not include release of the transcript from the template.[187]

D. The 142-nt RNA Is Produced in Isolated Nuclei

In order to exclude the possibility that the 142-nt RNA is synthesized only by a special class of VTC we analyzed the viral RNA synthesized in isolated nuclei. Similar to VTC RNA synthesis in isolated nuclei depends on nucleoside triphosphates (NTPs) and thus

nuclear RNA can be highly labeled with high-specific activity NTPs. There are several studies suggesting that in vitro chain elongation in isolated nuclei is an accurate reflection of a brief label inside cells.[181] Therefore, isolated nuclei are widely used to prepare nascent RNA for the measurement of differential transcripts of specific genes and transcription rate. In isolated nuclei, previously initiated RNA polymerase II molecules elongate growing chains by less than 500 nt in 10 to 20 min.[181]

In the present experiment nuclei were prepared with a detergent-free hypotonic buffer[164] from uninfected and MVM-infected cells at 24 hr postinfection. The in vivo preinitiated nascent RNA was elongated in vitro for 5 min in the presence of either 10 or 100 μM (α-^{32}P) UTP. The (^{32}P) RNA was then purified and the viral RNA synthesized from the two promoters was selected by hybridization to restriction fragments as detailed in Figure 29. The eluted RNA was analyzed by gel electrophoresis under denaturing conditions. Similar to the results of Figure 29 the results of this experiment show that a discrete major band is eluted from the fragment containing the left promoter (i_1) with an apparent length of 142 nt either when the nuclei were incubated with 10 (Figure 31A) or 100 μM (α-^{32}P) UTP (Figure 31B). This band is absent in RNA hybridizing to the fragment carrying the right promoter (i_2). No radioactivity hybridized to either fragment when RNA of uninfected cells was analyzed (results not shown). The synthesis of the 142-nt RNA in isolated nuclei may indicate that it is also synthesized in the infected cells. The 142-nt RNA could be produced either as a result of pausing of the RNA polymerase or transcription termination. It is noteworthy that in nuclei prepared from SV40 infected cells a cellular factor and RNA secondary structure were found to be essential elements of transcription termination at the SV40 attenuator.[98]

Pause and termination sites within the operon are used to set the rate of transcription of a gene or an operon below that attainable by polymerase-promoter interaction alone. Both of these mechanisms are considered as a mechanism of attenuation.[4] We, therefore, designated the 142-nt RNA as attenuated RNA.

E. The Attenuated RNA Terminates at a Transcription Termination Signal

Examination of the attenuated RNA for the presence of a potential secondary structure that might relate to the transcription termination event revealed that the DNA region 142 nt downstream from the left initiation site is AT-rich and is immediately preceded by a region that exhibits dyad symmetry. Figure 32 shows a proposed secondary structure of the RNA transcribed from this region. A stem-and-loop structure followed by U-residues is evident. This secondary structure resembles the transcription termination signal in prokaryotes[4,8,9,69,74,76,77] and the attenuator in SV40.[91-93]

It is noteworthy that the RNA transcribed from the attenuator region of MVM can be folded into two alternative conformations which are mutually exclusive. These are deignated (1, 2 + 3) attenuation and (1 + 2, 3) readthrough in Figure 33A and 33B, respectively. It seems that the presence of such unique structures at the promoter-proximal region allows one at least to speculate that it functions in vivo as the recognition site for protein or proteins which influence the frequency of transcription. We suggest that in the nucleus the conformation (1, 2 + 3) serves as a typical transcription termination structure that has been implicated in the process of premature termination.[91-93] Conversely, conformation (1 + 2, 3) leads to readthrough and to the production of mRNA. It is also interesting to note that in the cytoplasm the same alternative conformations can be formed at the 5′ end of the viral mRNA. Consequently, in the (1, 2 + 3) conformation the AUG initiating codon of VP_1 or of another still unidentified protein at residues 261 to 263[179] is free for ribosome binding. Conversely, in the alternative conformation (1 + 2, 3) the same AUG is sequestered in the stem (Figure 33B) and apparently is not free for ribosome binding.[24,25,28,30] Analogous to our model for SV40 (see Figure 24) we speculate that modulation between the two conformations could be a fundamental element in a feedback control mechanism that regulates in

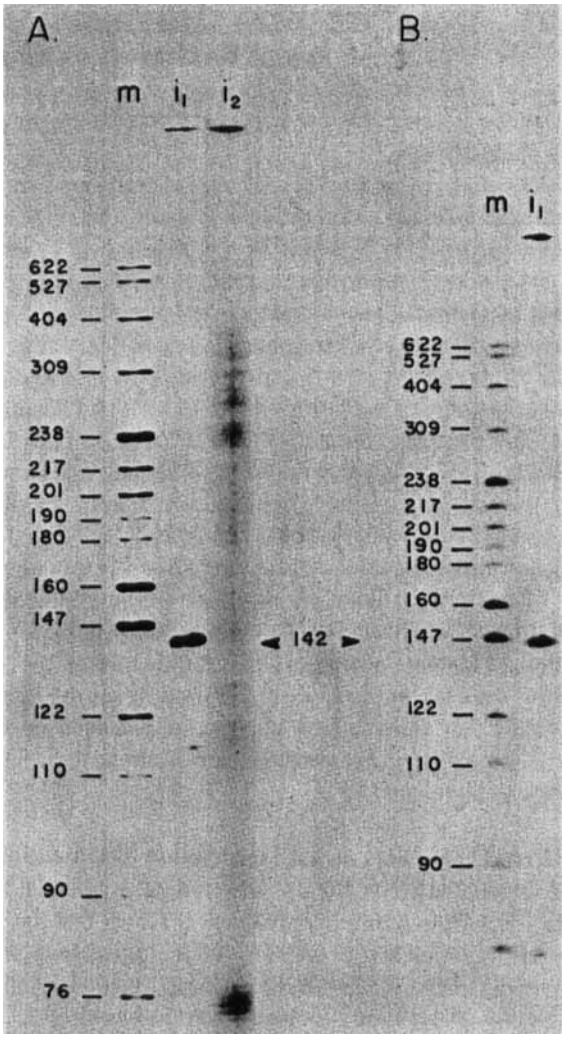

FIGURE 31. Size analyses of viral RNA synthesized in isolated nuclei. In vitro transcription in isolated nuclei in the presence of 10 μ*M* (α-^{32}P)-UTP in (A) or 100 μ*M* (α-^{32}P)-UTP in (B) was allowed to proceed for 5 min. RNAs were extracted and viral RNAs were selected by hybridization to and elution from a blot of Eco RI cleaved MVM DNA as in Figure 29. (i_1) RNA transcripts eluted from Eco RI "c" fragment; (i_2) RNA transcripts eluted from Eco RI "a + b" fragments. The ^{32}P-labeled viral RNAs were subjected to gel electrophoresis. (From Ben-Asher, E. and Aloni, Y., *J. Virol.*, 52, 266, 1984. With permission.)

the nucleus the amount of mRNA synthesized and in the cytoplasm the amount of the protein produced. The details of the model should await the characterization of the various proteins encoded by the MVM genome and the position of their initiating codons.

As far as the physiological significance of the attenuation mechanism in MVM is concerned, we suggest that this mechanism enables a communication between the two major promoters which is essential for a coordinated synthesis of the viral proteins. Thus, we

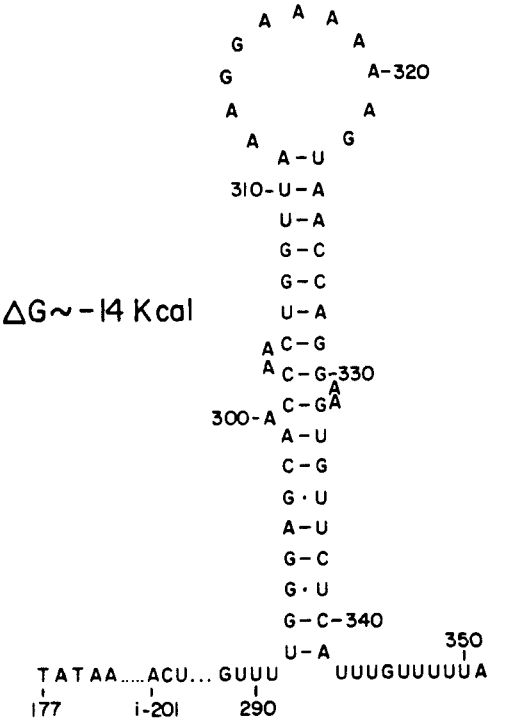

FIGURE 32. Secondary structure at the 3′ end of the attenuated RNA of MVM. The ΔG was calculated as described.[211] Nucleotide residue numbers are from Astell et al.[179] (From Ben-Asher, E. and Aloni, Y., *J. Virol.*, 52, 266, 1984. With permission.)

speculate that a protein encoded by the right transcription unit (VP_2 or another protein) acts as an attenuator or antiattenuator factor, most likely by a mechanism that stabilizes one of the alternative RNA conformations (see Figure 33). The present study is, thus, in support of our previous suggestion for SV40 that attenuation and modulation of RNA secondary structure in a feedback mechanism regulates viral gene expression.[91,92]

F. Is Attenuation a Prevalent Control Mechanism in Animal Viruses?

As an approach to answer this question we have scrutinized the DNA sequences of promoter-proximal regions of five animal viruses: SV40 and BK of the papova virus group and MVM, H-1, and KRV of the parvovirus group. These analyses revealed the existence of dyad symmetries which, in the resulting RNAs, have the potential of forming stem-and-loop structures that are immediately followed by U-residues. Figure 34 represents our proposed transcription termination signals in these viruses (Figure 34B) as compared to those of the attenuation sites of the amino acid biosynthetic operons in bacteria (Figure 34A).[9] It is apparent that in comparison to the prokaryotic termination signals (Figure 34A) the stem-and-loop structures of the animal viruses (Figure 34B) are less stable, have an A-rich loop, and include unpaired bases. Also, it is interesting to note that all the stems contain the tetranucleotide CCAG in the proximity of the unpaired bases (Figure 34B shadowed regions). It is conceivable that all these features provide the recognition specificity for viral attenuator and/or antiattenuator factors, or for a cellular termination and antitermination factors which take part in the transcription termination process.[91,98,136,137]

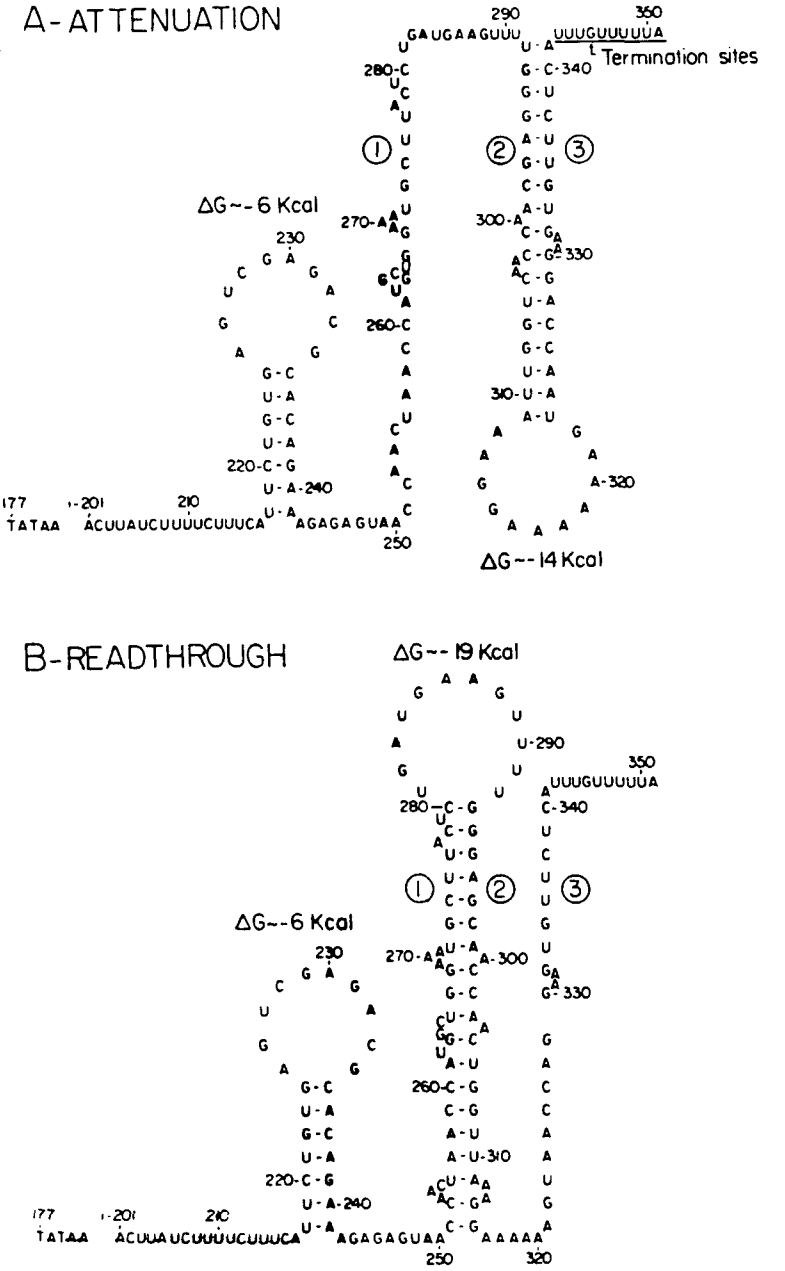

FIGURE 33. Schema of alternative conformations in the attenuated RNA and at the 5′ end of the viral mRNA initiated at residue 201 ± 5. (A) Attenuation conformation showing a typical terminal signal in 2 + 3. The initiation codon at residues 261 to 263 of the mRNA[179] is free for ribosome binding; (B) readthrough conformation showing sequestration of initiation codon at residues 261 to 263 of the mRNA. Residue numbers are from Astell et al.[179] and the ΔG was calculated as described in Tinoco et al.[211] (From Ben-Asher, E. and Aloni, Y., *J. Virol.*, 52, 266, 1984. With permission.)

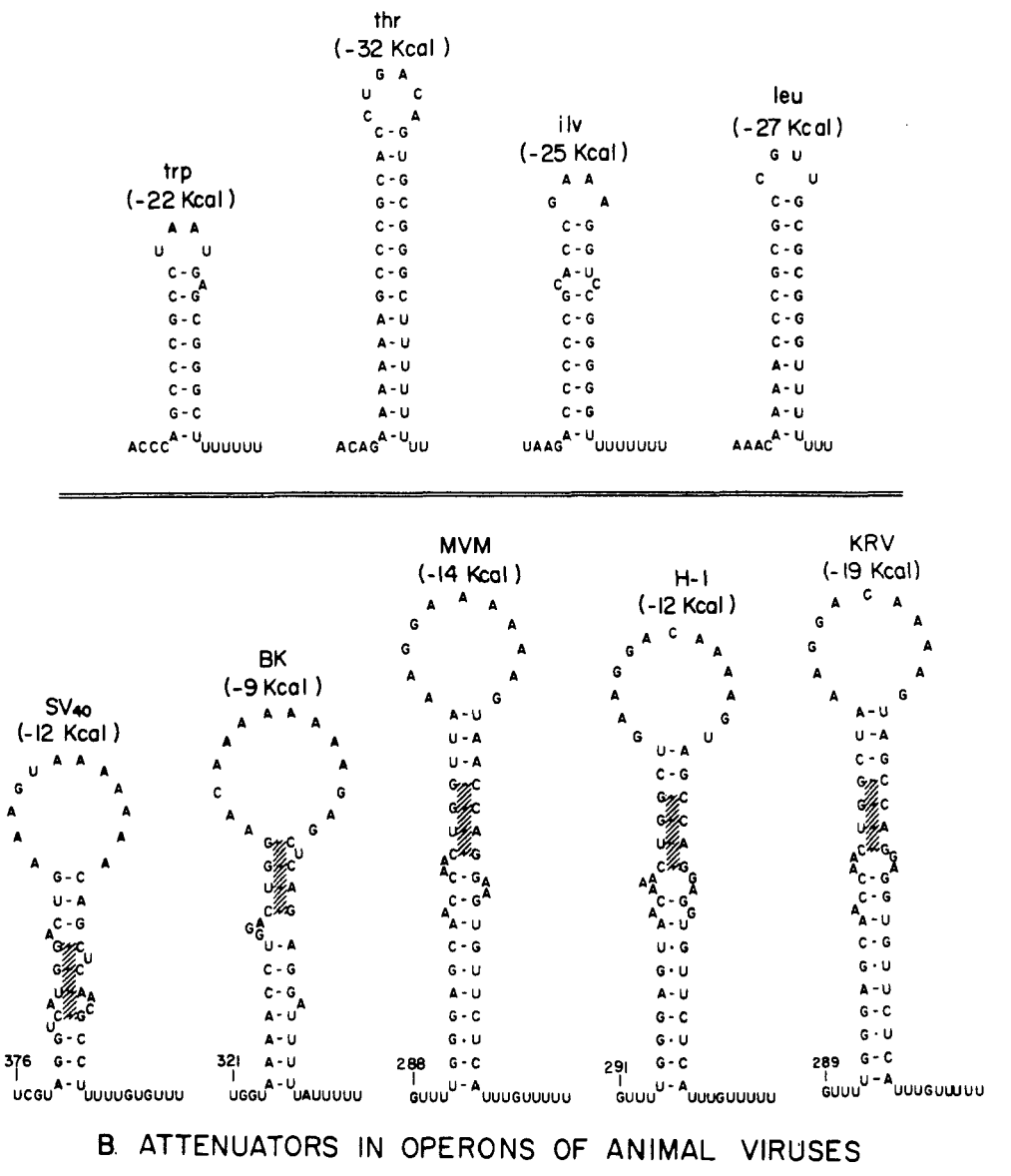

FIGURE 34. Comparison between proposed secondary structures of the attenuators of the amino acid biosynthetic operons of bacteria and those of animal viruses. Those of bacteria were adopted from Kolter and Yanofsky.[9] The structure of the attenuators of animal viruses suggested here are based on the nucleotide sequence of SV40 and BK,[11] MVM,[179] H-1,[178] and KRV.[173] In (B) the shadowed regions represent the tetranucleotides CCAG present in the stems of the attenuators of the animal viruses.

XI. ATTENUATION OF TRANSCRIPTION IN OTHER EUKARYOTIC SYSTEMS

There are several reports which show that attenuation of transcription occurs in other eukaryotic systems and it is probably a more general phenomenon. In adenovirus 2 (Ad2) the large late transcription unit extends from map unit 16.4 to the right for 25 kb to about map unit 99. It constitutes the major transcription unit functioning late after infection.[11] It

was observed that short promoter proximal RNA from the major late transcription unit of adenovirus is accumulated late after infection in HeLa cells. The concentration of nascent RNA chains extending to the right was approximately fivefold higher in the DNA fragment containing the promoter, than from regions further downstream.[80,81] Sucrose gradient sedimentation analyses[81] and polyacrylamide denaturing gel electrophoresis analyses[80] showed that the short promoter proximal RNA consists of multiple discrete species.[80] Furthermore, RNA fingerprint analyses showed that the short RNA chains are capped and correspond in sequence to the first 350 nucleotides of the major late Ad2 transcription unit. These observations indicate that there is a significant premature RNA chain termination. Many of these observations have been independently confirmed[189,190] and have been extended to a second late Ad2-transcription unit.

The Ad2 mature late mRNAs have tripartite leaders covalently attached to their 5′ ends. The biogenesis of these mosaic RNAs starts with the synthesis of a very large transcript from the major late adenovirus promoter. Subsequently, three leader segments are covalently joined to the mRNA body by elimination of three intron sequences. The tripartite leaders lack coding frames.[191] The most frequently occurring leader is the tripartite leader, but there also exists a leader with an additional segment which is called the i-leader.[192] The i-leader contains an open reading frame for translation which codes for an hypothetical 15.9-kd polypeptide. This polypeptide has been found by in vitro translation.[193] It has been shown that the i-leader is predominantly expressed late after adenovirus infection, and somewhat suppressed the translation of downstream genes.[28,193] The Ad2 mRNA with the i-leader resembles in some respects the late mRNA of SV40 which contains the coding frame for the agnoprotein in its leader region. It is possible that attenuation of transcription in Ad2 is coupled to translation as has been suggested for SV40.

Promoter proximal Ad2-specific RNA derived from several early transcription units has also been detected. Thus, several of the early Ad2 transcription units appear to be subject to attenuation of transcription.[190]

A different type of transcription termination control, which may be responsible for the early-late switch, the same as was found in bacteriophages,[2,6] has been found to occur in Ad2 virus. During late viral infection transcription from the major late promoter terminates mainly at the end of the genome. In sharp contrast during early infection the same promoter is utilized, but transcripts terminate near the middle of the genome, producing the L_1 transcripts only. This prevents the expression of distal exons.[194,195] It is interesting to note that similar to the attenuators of SV40 and MVM (see Figures 23 and 33) two alternative conformations can be formed at the 3′ end of the L_1 transcripts (see Figure 35). One conformation contains a stem-and-loop structure followed by a run of Us. This structure has been suggested to function as a terminator of RNA polymerase II.[91,92,180] It is possible that the two alternative conformations with additional factors are involved in the mechanism of the early-late switch.

Other viruses may have attenuation-like mechanisms. According to the results of Montandon and Acheson,[196] it is possible that the late polyoma transcription unit is also subject to transcription attenuation. During transcription of vesicular stomatitis virus (VSV), short leader transcripts are accumulated[197] and transcription decreases in an ordered manner at or near the junctions of neighboring genes. This results in a cumulative reduction in distal gene expression. It was suggested that there are protein factors which cause release of the leader RNA in VSV transcription and that there is another factor which causes readthrough of transcription.[197,198] In retroviruses it has been shown that the LTR contains sequences that code for the initiation and termination of RNA synthesis. In this case the promoter region is immediately adjacent to the termination signal. The termination signal must be ignored by RNA polymerase II, during transcription of the LTR at the 5′ end of the RNA. However, the polymerase has to recognize such signals at the 3′ end of the transcript which is 5400

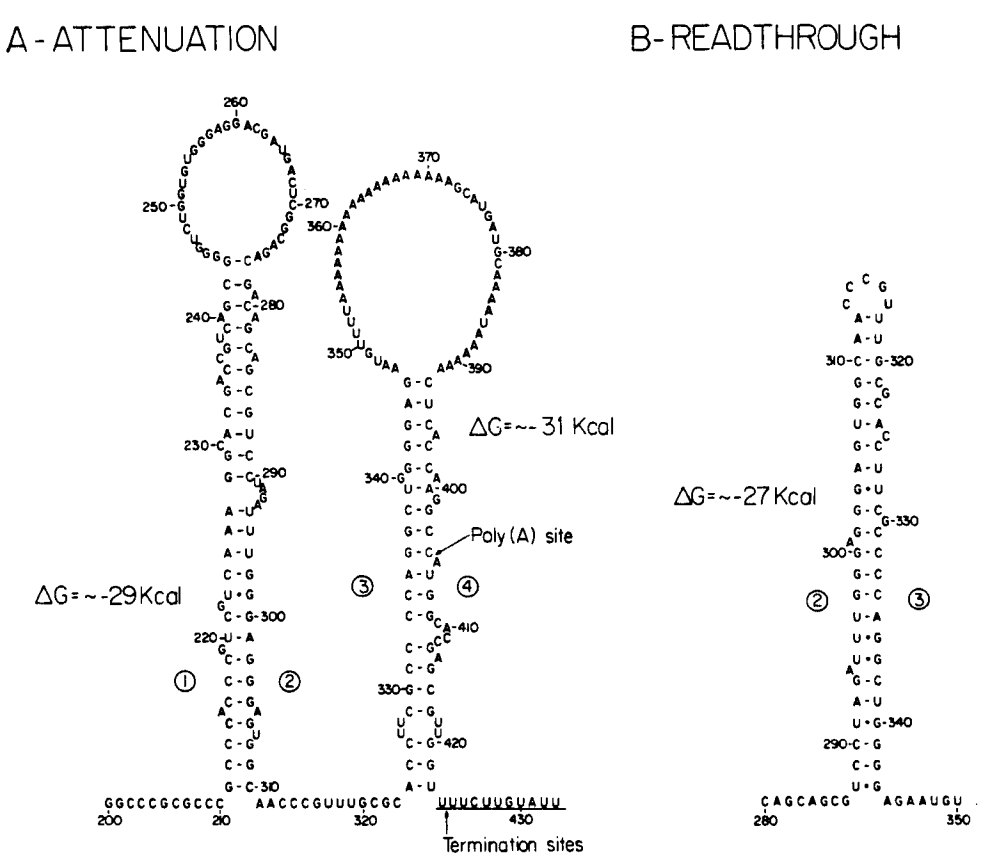

FIGURE 35. Schema of alternative conformations surrounding the poly(A) addition site of the L_1 family mRNAs of adenovirus type 2. The sequence and residue numbers are from Le Moullec et al.[212]

nucleotides downstream. A model has been suggested which proposes that RNA secondary structure is influential in determining readthrough at the 5′ end of the transcript and termination at 3′ end.[199]

Attenuation of transcription in specific eukaryotic nonviral genes has also been documented. It was found that RNA polymerase molecules are located preferentially in the 5′ end of the β-globin genes of mature chicken erythrocytes. In the β-globin genes of immature erythrocytes they are apparently evenly distributed.[200] Since adult globin genes are no longer transcribed in mature chicken erythrocytes, it was predicted that an attenuation-like mechanism plays a role during differentiation by controlling the expression of β-globin genes. Thus, in some genes which have the potential to be transcribed in a given tissue, RNA polymerase molecules would be allowed to initiate transcription but not to elongate, resulting in clustering of polymerase molecules around the 5′ end of the genes.[200] There are many eukaryotic genes which are expressed in a controlled low amount during the normal cell cycle. This kind of control can be achieved by attenuation of transcription. Especially of interest are cellular oncogenes which may have a role during normal growth and development.[201] In differentiated mature cells they are kept at a low level of expression. However, if this fine tuning control of oncogenes expression is released the cells will have the potentiality to be transformed.

It has recently been suggested that an attenuation mechanism operates in yeast which is similar to the amino acid operons in bacteria.[202] Upstream from the third enzyme of the leucine biosynthetic pathway (LEU2) in yeast, there is a sequence which could pontentially

code for a 23-residue peptide, 6 residues of which include leucine. Moreover, sequences preceding LEU2 are capable of forming alternative base-paired secondary structures in the corresponding transcript. One of the secondary structures contains a hairpin stem-and-loop structure which is followed by a run of Us. Such a structure may cause transcription termination. The presence of such a hairpin structure in the RNA prevents the formation of a competitive and overlapping structure. Thus, it leaves the ribosome binding site free for translation of the leader peptide. It was suggested that the leader peptide serves as a feedback mediator between the cytoplasm and the nucleus for the regulation of LEU2 gene expression. Such a mechanism is similar to what has been suggested for the role of the agnoprotein in SV40, in the regulation of production of the capsid proteins. According to the model for LEU2 regulation, the leader peptide is translated in the cytoplasm and then transported into the nucleus. In the nucleus the leader peptide stabilizes the secondary hairpin structure at the putative attenuator sites and leads to premature termination of transcription. The concentration of the leader peptide may reflect the intracellular pool of leucine through the leucine codons in the peptide sequence. Thus, only when leucine pool is high will enough of the leader peptide be synthesized to affect attenuation.[202]

In prokaryotes, there is a considerable volume of evidence which confirms the involvement of the modulation of mRNA secondary structure in the regulation of gene expression. The early work of Lodish,[203] Steiz,[204] and Fiers et al.[205] strongly suggests that secondary structure acts to regulate protein synthesis in bacteriophages, f2, R17, and MS2 systems. Insertant and Fiers[206] used a series of recombinant plasmids containing the cro gene of λ phage to show a correlation between the level of expression of the cro protein and the secondary structure of the 5′ regions of the mRNA coded for by the respective plasmids. They argue that initiation of translation in prokaryotes is influenced by the position of the initiator AUG in the secondary structure. This means that the accessibility of the AUG is more important than the accessibility of the Shine and Dalgarno sequence.[206] There are also proteins which modulate RNA structure and autoregulate their expression. Among them are the coat proteins of coliphages R17 and MS2 and the single-stranded DNA binding protein coded for by coliphage T4 gene 32. These proteins repress their own synthesis by binding near the AUG codon.[28] Several of ribosomal proteins from E. coli regulate their own expression by binding to the RNA. One of the ribosomal proteins, L4, regulates the transcription of the S10 ribosomal protein operon by stabilizing the stem-and-loop structure in the leader and, therefore, enhancing premature termination of transcription.[207]

Work on the ribosome binding sites of eukaryotes in BMV mRNA, ovalbumin mRNA, alfa mosaic virus mRNA, and globin mRNA indicates that base pairing may be involved in the size and sequences of the 5′ region which are protected by the 40S and 80S ribosome initiation complexes.[20] Kozak[28] has demonstrated that disruption of secondary structure by chemical means can alter the formation of the 40S ribosomal initiation complexes in reovirus mRNA. Direct measurements by physical techniques and enzymatic means demonstrate that eukaryotic mRNA has considerable specific secondary structure.[30] It was shown by S_1 and T_1 digestion that there are extensive secondary structures at the 5′ end of globin mRNAs.[30,208] The initiator AUG regions of the mouse and rabbit α mRNAs are not susceptible to S_1 and T_2 digestion, whereas the AUG regions of the respective β mRNAs are highly susceptible. These results may partially explain why initiation of protein synthesis with β-globin mRNA in reticulocytes occurs at a rate 30 to 40% faster than with α-globin mRNA.[30] Recently, it was found that the 5′ exon (exon 1) of the c-myc inhibits expression of this oncogene in normal cells.[66] It was suggested that this inhibition is at the level of translation. Examination of the cy-myc sequence reveals a region of exon 1 to have high complementary to region of exon 2. A stem-and-loop secondary structure for the c-myc mRNA may, therefore, be proposed. In this structure the initiator AUG is located within the loop. It was proposed that this stem-and-loop structure inhibits translation of c-myc.[209]

XII. EPILOG

In the long term it should be possible to establish whether attenuation in SV40 MVM and other viruses represents specialized viral control mechanism or, as seems more likely, whether the viruses are mimicking a control mechanism operative in normal eukaryotic cells. The suggested mechanisms described in the models presented in this review article open up approaches to the investigation of attenuation and mRNA modulation as a possible mechanism whereby eukaryotes may regulate transcription in a variety of different circumstances. We hope that testing some of the predictions presented in our models will confirm or rule out particular aspects, and in doing so will improve our understanding of how viruses and the eukaryotic cell regulate their gene expression.

ACKNOWLEDGMENTS

We wish to thank all our colleagues who contributed to the studies presented in this review article: Hagit Skolnik-David, Edna Ben-Asher, Ron Pruzan, Pierre Pfeiffer, Rivka Abulafia, Edward B. Jakobovits, Mark Kessler, Orna Resnekov, David King, Eyal Ben-Gal, Ruchama Leiserowitz, Orgad Laub, and Avri Ben-Ze'ev, and, also Mrs. Melvine Stein for expert preparation of the manuscript. This research was supported by USPHS Research Grant CA 14995 and, in part, by the Norman V. Wechsler Endowment Fund and the Leo and Julia Forcheimer Center for Molecular Genetics.

REFERENCES

1. **Jacob, F. and Monod, J.,** Genetic regulatory mechanisms in the synthesis of proteins, *J. Mol. Biol.*, 3, 318, 1961.
2. **Adhya, S. and Gottesman, M.,** Control of transcription termination, *Ann. Rev. Biochem.*, 47, 967, 1978.
3. **Rosenberg, M. and Court, D.,** Regulatory sequences involved in the promotion and termination of RNA transcription, *Ann. Rev. Genet.*, 13, 319, 1979.
4. **Yanofsky, C.,** Attenuation in the control of expression of bacterial operons, *Nature (London)*, 289, 751, 1981.
5. **Greenblatt, J.,** Regulation of transcription termination by the N gene protein of bacteriophage λ, *Cell*, 24, 8, 1981.
6. **Ward, D. F. and Gottesman, M. E.,** Suppression of transcription termination by phage lambda, *Science*, 216, 946, 1982.
7. **Roberts, J. W.,** Termination factor for RNA synthesis, *Nature (London)*, 224, 1168, 1969.
8. **Crawford, I. P. and Stauffer, G. V.,** Regulation of tryptophan biosynthesis, *Ann. Rev. Biochem.*, 49, 163, 1980.
9. **Kolter, R. and Yanofsky, C.,** Attenuation in amino acid biosynthetic operons, *Ann. Rev. Genet.*, 16, 113, 1982.
10. **Remington, J. A.,** Attenuation as a general mechanism for the regulation of differential gene transcription in eucaryotes, *FEBS Lett.*, 100, 225, 1979.
11. **Tooze, J.,** *DNA Tumor Viruses*, 2nd ed., Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y., 1981.
12. **Lebowitz, P. and Weissman, S. M.,** Organization and transcription of the Simian virus 40 genome, *Curr. Top. Microb. Immunol.*, 87, 43, 1979.
13. **Aloni, Y.,** Biogenesis and characterization of SV40 and polyoma RNAs in productively infected cells, *Cold Spring Harbor Symp. Quant. Biol.*, 39, 165, 1974.
14. **Aloni, Y., Bratosin, S., Dhar, R., Laub, O., Horowitz, M., and Khoury, G.,** Splicing of SV40 mRNA: a novel mechanism for the regulation of gene expression in mammalian cells, *Cold Spring Harbor Symp. Quant. Biol.*, 43, 559, 1978.

15. **Gidoni, D., Kahana, G., Cannani, D., and Groner, Y.,** Specific in vitro initiation of transcription of simian virus 40 early and late genes occurs at the various cap nucleotides including cytidine, *Proc. Natl. Acad. Sci. U.S.A.*, 78, 2174, 1981.
16. **Griffith, J. D., Dieckmann, M., and Berg, P.,** Electron microscope localization of a protein bound near the origin of simian virus 40 DNA replication, *J. Virol.*, 15, 167, 1975.
17. **Reddy, V. B., Thimmappaya, B., Dhar, R., Subramanian, K. N., Zain, B. S., Pan, J., Celma, M. L., Ghosh, P. K., and Weissman, S. M.,** The genome of Simian virus 40, *Science*, 200, 494, 1978.
18. **Fiers, W., Contreras, R., Haegeman, G., Rogiers, R., Van de Voorde, A., Van Heuverswyn, H., Van Herreweghe, J., Van Volkaert, G., and Ysebaert, M.,** The total nucleotide sequence of SV40 DNA, *Nature (London)*, 273, 113, 1978.
19. **Davidson, B. L., Egly, J. M., Mulvhill, E. R., and Chambon, P.,** Formation of stable preinitiation complexes between eukaryotic class transcription factors and promoter sequences, *Nature (London)*, 301, 680, 1983.
20. **Dynan, W. S. and Tijian, R.,** Isolation of transcription factors that discriminate between different promoters recognized by RNA polymerase II, *Cell*, 32, 669, 1983.
21. **Fromm, M. and Berg, P.,** Deletion mapping of DNA regions required for simian virus 40 early region promoter function in vivo, *J. Mol. Appl. Gen.*, 1, 457, 1982.
22. **Parker, C. S. and Topol, J.,** Drosphila RNA polymerase II transcription factor contain a promoter region-specific DNA binding activity, *Cell*, 36, 357, 1984.
23. **Samuels, M., Fire, A., and Sharp, P. A.,** Separation and characterization of factors mediating accurate transcription by RNA polymerase II, *J. Biol. Chem.*, 257, 14419, 1982.
24. **Kozak, M.,** How do eucaryotic ribosomes select initiation regions in messenger RNA?, *Cell*, 15, 1109, 1978.
25. **Kozak, M.,** Evaluation of the "scanning model" for initiation of protein synthesis in eucaryotes, *Cell*, 22, 7, 1980.
26. **Kozak, M.,** Possible role of flanking nucleotides in recognition of the AUG initiator codon by eukaryotic ribosomes, *Nucl. Acids Res.*, 9, 5233, 1981.
27. **Kozak, M.,** Comparison of initiation of protein synthesis in procaryotes, eucaryotes and organelles, *Microbiol. Rev.*, 47, 1, 1983.
28. **Kozak, M.,** Point mutations close to the AUG initiator codon affect the efficiency of translation of rat preproinsulin in vivo, *Nature (London)*, 308, 241, 1984.
29. **Yanofsky, C.,** Prokaryotic mechanism in eukaryotes, *Nature (London)*, 302, 751, 1983.
30. **Vary, C. P. H. and Vournakis, J. N.,** Secondary structure of eukaryotic messenger RNA, in *Biochemical Society Symp.*, No. 47, Phelps, C. F. and Arenstein, H. R. V., Eds., Biochemical Society, London, 1982.
31. **Shatkin, A. J.,** A closer look at the 5′ end of mRNA in relation to initiation, in *Protein Biosynthesis in Eukaryotes*, Perez-Bercoff, E., Ed., Plenum Press, New York, 1982.
32. **Mathis, D. J. and Chambon, P.,** The SV40 early region TATA box is required for accurate in vitro initiation of transcription, *Nature (London)*, 290, 310, 1981.
33. **Wasylyk, B., Wasylyk, C., Matthes, H., Wintzerith, M., and Chambon, P.,** Transcription from the simian virus 40 early-early and late-early overlapping promoters in the absence of DNA replication, *EMBO J.*, 2, 1605, 1983.
34. **Gruss, P., Dhar, R., and Khoury, G.,** Simian virus 40 tandem repeated sequences as an element of the early promoter, *Proc. Natl. Acad. Sci. U.S.A.*, 78, 943, 1981.
35. **Berk, A. J. and Sharp, P. A.,** Spliced early mRNAs of simian virus 40, *Proc. Natl. Acad. Sci. U.S.A.*, 75, 1274, 1978.
36. **Reddy, V. B., Ghosh, P. K., Lebowitz, P., Piatak, M., and Weissman, S. M.,** Genomic localization of the 3′ and 5′ termini and the two major splices in mRNA from transformed and lytically infected cells, *J. Virol.*, 30, 279, 1979.
37. **Mark, D. F. and Berg, P.,** A third splice site in SV40 early mRNA, *Cold Spring Harbor Symp. Quant. Biol.*, 44, 55, 1980.
38. **Tegtmeyer, P., Schwartz, M., Collins, J. K., and Rundell, K.,** Regulation of tumor antigen synthesis by simian virus 40 gene A, *J. Virol.*, 16, 168,
39. **Reed, S. I., Stark, G. R., and Alwine, J. C.,** Autoregulation of simian virus 40 gene by T antigen, *Proc. Natl. Acad. Sci. U.S.A.*, 73, 3083, 1976.
40. **Khoury, G. and May, E.,** Regulation of early and late simian virus 40 transcription: overproduction of early viral RNA in the absence of functional T antigen, *J. Virol.*, 23, 167, 1977.
41. **Laub, O. and Aloni, Y.,** Transcription of simian virus 40. V. Regulation of simian virus 40 gene expression, *J. Virol.*, 16, 1171, 1975.
42. **Rio, D., Robbins, A., Myers, R., and Tijan, R.,** Regulation of SV40 early transcription in vitro by a purified T antigen, *Proc. Natl. Acad. Sci. U.S.A.*, 77, 5706, 1980.

43. **Myers, R. M., Rio, D. C., Robbins, A. K., and Tjian, R.,** SV40 gene expression is modulated by the cooperative binding of T-antigen to DNA, *Cell*, 25, 373, 1981.
44. **Tjian, R.,** T antigen binding and the control of SV40 gene expression, *Cell*, 26, 1, 1981.
45. **Gutai, M. W. and Nathans, D.,** Evolutionary variants of simian virus 40: cellular DNA sequences and sequences at recombinant joints of substituted variants, *J. Mol. Biol.*, 126, 275, 1978.
46. **Keller, J. M. and Alwine, J. C.,** Activation of the SV40 late promoter: direct effects of T antigen in the absence of viral DNA replication, *Cell*, 36, 381, 1984.
47. **Ghosh, P. K., Reddy, V. B., Swinscoe, J., Lebowitz, P., and Weissman, S. M.,** Heterogeneity and 5′ terminal structure of the late RNAs of simian virus 40, *J. Mol. Biol.*, 126, 813, 1978.
48. **Piatak, M., Subramanian, K. N., Roy, P., and Weissman, S. M.,** Late messenger RNA production by viable simian virus 40 mutants with deletions in the leader region, *J. Mol. Biol.*, 153, 589, 1981.
49. **Vershavsky, A. J., Sundin, O. H., and Bohm, M. J.,** A stretch of "late" SV40 viral DNA about 400 bp long which includes the origin of replication is specifically exposed in SV40 minichromosomes, *Cell*, 16, 453, 1979.
50. **Scott, W. A. and Wigmore, D. J.,** Sites on simian virus 40 chromatin which are preferentially cleaved by endonucleases, *Cell*, 15, 1511, 1978.
51. **Jakobovits, E. B., Bratosin, S., and Aloni, Y.,** A nucleosome free region in SV40 minichromosome, *Nature (London)*, 285, 263, 1980.
52. **Jakobovits, E. B., Saragosti, S., Yaniv, M., and Aloni, Y.,** *E. coli* RNA polymerase mimics SV40 in vivo transcription when the template in vitro is viral nucleoprotein, *Proc. Natl. Acad. Sci. U.S.A.*, 77, 6556, 1980.
53. **Jakobovits, E. B., Bratosin, S., and Aloni, Y.,** Formation of a nucleosome-free region in SV40 minichromosomes is dependent upon a restricted segment of DNA, *Virology*, 120, 340, 1982.
54. **Saragosti, S., Moyne, G., and Yaniv, M.,** Absence of nucleosomes in a fraction of SV40 chromatin between the origin of replication and the region coding for late leader RNA, *Cell*, 20, 65, 1980.
55. **Reddy, V. B., Ghosh, P. K., Lebowitz, P., and Weissman, S. M.,** Gaps and duplicated sequences in the leaders of SV40 16S RNA, *Nucl. Acids Res.*, 5, 4195, 1978.
56. **Piatak, M., Ghosh, P. K., Norkin, L. C., and Weissman, S. M.,** Sequence locating the 5′ ends of the major simian virus 40 late mRNA forms, *J. Virol.*, p. 503, 1983.
57. **Barkan, A. and Mertz, J.,** The number of ribosomes on simian virus 40 late 16S mRNA is determined in part by the nucleotide sequence of its leader, *Mol. Cell. Biol.*, 4, 813, 1984.
58. **Alwine, J. C. and Khoury, G.,** Simian virus 40 associated small RNA: mapping on the simian virus 40 genome and characterization of its synthesis, *J. Virol.*, 36, 701, 1980.
59. **Alwine, J. C., Dhar, R., and Khoury, G.,** A small RNA induced late in simian virus 40 lytic infection can associate with early viral mRNAs, *Proc. Natl. Acad. Sci. U.S.A.*, 77, 1379, 1980.
60. **Alwine, J. C.,** Hybrid selection of small RNAs by using simian virus 40 DNA: evidence that the simian virus 40-associated small RNAs is synthesized by specific cleavage from large viral transcripts, *J. Virol.*, 43, 987, 1982.
61. **Derynck, R., Leung, D. W., Gray, P. W., and Goeddel, D. V.,** Human interferon γ is encoded by a single class of mRNA, *Nucl. Acids Res.*, 10, 3605, 1982.
62. **Holmgren, R., Croces, V., Morimoto, R., Blackman, R., and Messelson, M.,** Sequence homologies in the 5′ regions of four Drosophila heat-shock genes, *Proc. Natl. Acad. Sci. U.S.A.*, 78, 3775, 1981.
63. **Gallwitz, D., Perrin, F., and Seidel, R.,** The actin gene in yeast *Saccharomyces cerevisiae:* 5′ and 3′ end maping, flanking and putative regulatory sequences, *Nucl. Acids Res.*, 9, 6339, 1981.
64. **Ingolia, T. D. and Craig, E. A.,** Primary sequence of the 5′-flanking regions of the Drosophila heat shock genes in chromosome subdivision 67B, *Nucl. Acids Res.*, 9, 1627, 1981.
65. **Shen, L. P., Pictet, R., and Rutter, W. J.,** Human somatostatin I: sequence of the cDNA, *Proc. Natl. Acad. Sci. U.S.A.*, 79, 4575, 1982.
66. **Stanton, L. W., Watt, R., and Marcu, K. B.,** Translocation breakage and truncated transcripts of c-myc oncogene in murine plasmacytomas, *Nature (London)*, 303, 401, 1983.
67. **Rosenberg, M. and Paterson, B. M.,** Efficient cap-dependent translation of polycistronic prokaryotic mRNAs is restricted to the first gene in the operon, *Nature (London)*, 279, 696, 1979.
68. **Parviakis, G. N., Lockard, P. E., Vanvakopulos, N., Riser, L., Raj Bhandary, U., and Vournakis, J.,** Secondary structure of mouse and rabbit α- and β-globin mRNAs: differential accessibility of α and β initiator AUG codons towards nucleases, *Cell*, 19, 91, 1980.
69. **Biro, P. A. and Weissman, S. M.,** Recognition and control sequences in nucleic acids, in *Molecular Genetics*, Vol. 3, Taylor, J. H., Ed., Academic Press, New York, 1979, 177.
70. **Grosschedl, R. and Birnstiel, M. L.,** Spacer DNA sequences upstream of the T-A-T-A-A-A-T-A sequence are essential for promotion of H_2A histone gene transcription in vivo, *Proc. Natl. Acad. Sci. U.S.A.*, 77, 7102, 1980.

71. **Benoist, C. and Chambon, P.,** In vivo sequence requirements of the SV40 early promoter region, *Nature (London)*, 290, 304, 1981.
72. **Dierks, P., Van Ooyen, A., Mantei, N., and Weissman, C.,** DNA sequences preceding the rabbit-β-globin gene are required for formation in mouse L cells of β-globin RNA with the correct 5′ terminus, *Proc. Natl. Acad. Sci. U.S.A.*, 78, 1411, 1981.
73. **Gruss, P., Dhar, R., and Khoury, G.,** Simian virus 40 tandem repeated sequences as an element of the early promoter, *Proc. Natl. Acad. Sci. U.S.A.*, 78, 943, 1981.
74. **Holmes, W., Platt, T., and Rosenberg, M.,** Termination of transcription in *E. coli*, *Cell*, 32, 1029, 1983.
75. **Gilbert, W.,** Starting and stopping sequences for the RNA polymerase, in *RNA Polymerase*, Losick, R. and Chamberlin, M., Eds., Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y., 1976, 193.
76. **Galluppi, G. R. and Richardson, J. P.,** ATP-induced changes in the binding of RNA synthesis termination protein rho to RNA, *J. Mol. Biol.*, 138, 513, 1980.
77. **Gottesman, M. E., Adhya, S., and Das, A.,** Transcription antitermination by bacteriophage γ N gene product, *J. Mol. Biol.*, 140, 57, 1980.
78. **Tamm, I. and Kikuchi, T.,** Early termination of heterogeneous nuclear RNA transcripts in mammalian cells: accentuation by 5,6-dichloro-1-β-D ribofuranosylbenzimidazole, *Proc. Natl. Acad. Sci. U.S.A.*, 76, 5750, 1979.
79. **Salditt-Georgieff, M., Harpold, M., Chen-Kiang, S., and Darnell, J. E., Jr.,** The addition of 5′ cap structures occurs early in hnRNA synthesis and prematurely terminated molecules are capped, *Cell*, 19, 68, 1980.
80. **Fraser, N. W., Shegal, P. B., and Darnell, J. E., Jr.,** Multiple discrete sites of premature RNA chain termination late in adenovirus-2 infection: enhancement by 5,6-dichloro-1-β-D ribofuranosylbenzimidazole, *Proc. Natl. Acad. Sci. U.S.A.*, 76, 2571, 1979.
81. **Evans, R., Weber, J., Ziff, E., and Darnell, J. E., Jr.,** Premature termination during adenovirus transcription, *Nature (London)*, 278, 367, 1979.
82. **Tamm, I. and Sehgal, P. B.,** Effects of drugs on cell nucleus, in *Bristol-Myers Symp.: Effect of Drugs on the Cell Nucleus*, Busch. H., Ed., Academic Press, New York, 1979, 251.
83. **Laub, O., Bratosin, S., Horowitz, M., and Aloni, Y.,** The initiation of transcription of SV40 DNA at late time after infection, *Virology*, 92, 310, 1979.
84. **Laub, O., Jakobovits, E. B., and Aloni, Y.,** 5,6-Dichloro-1-β-D-ribofuranosylbenzimidazole (DRB) enhances premature termination of late SV40 transcription, *Proc. Natl. Acad. Sci. U.S.A.*, 77, 3207, 1980.
85. **Dreyer, C. and Hausen, P.,** Inhibition of mammalian RNA polymerase by 5,6-dichloro-ribofuranosylbenzimidazole (DRB) and DRB triphosphate, *Nucl. Acids Res.*, 5, 3332, 1978.
86. **Canaani, D., Kahana, C., Mukamel, A., and Groner, Y.,** Sequences heterogeneity at the 5′ termini of late simian virus 40 19S and 16S mRNAs, *Proc. Natl. Acad. Sci. U.S.A.*, 76, 3078, 1979.
87. **Contreras, R. and Fiers, W.,** Initiation of transcription by RNA polymerase II in permeable, SV40 infected or noninfected CV1 cells. Evidence for multiple promoters of SV40 late transcription, *Nucl. Acids Res.*, 9, 215, 1981.
88. **Griffin, B. E.,** *DNA Tumor Viruses*, Tooze, J., Ed., Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y., 1980, 61.
89. **Garber, E. A., Seidman, M. M., and Levine, A. J.,** Intracellular SV40 nucleoprotein complexes: synthesis and encapsidation, *Virology*, 107, 389, 1980.
90. **Brady, J. V., Winston, V. D., and Conigli, R. A.,** Characterization of a DNA-protein complex and capsomere subunits derived from polyoma virus by treatment with ethyleneglycol-B:S-N,N′ -tetraacetic acid and dithiothreitol, *J. Virol.*, 27, 193, 1978.
91. **Hay, N., Skolnik-David, H., and Aloni, Y.,** Attenuation in the control of SV40 gene expression, *Cell*, 29, 183, 1982.
92. **Aloni, Y. and Hay, N.,** Attenuation and modulation of mRNA secondary structure in a feedback control system regulating SV40 gene expression, *Mol. Biol. Rep.*, 9, 91, 1983.
93. **Skolnik-David, H., Hay, N., and Aloni, Y.,** Site of premature termination of late transcription of simian virus 40 DNA: enhancement by 5,6-dichloro-1-β-D-ribofuranosylbenzimidazole, *Proc. Natl. Acad. Sci. U.S.A.*, 79, 2743, 1982.
94. **Skolnik-David, H. and Aloni, Y.,** Pausing of RNA polymerase molecules during in vivo transcription of the SV40 leader region, *EMBO J.*, 2, 179, 1983.
95. **Pfeiffer, P., Hay, N., Pruzan, R., Jakobovits, E. B., and Aloni, Y.,** In vitro premature termination in SV40 late transcription, *EMBO J.*, 2, 185, 1983.
96. **Aloni, Y., Hay, N., Skolnik-David, H., Pfeiffer, P., Abulafia, R., Pruzan, R., Ben-Asher, E., Jakobovits, E. B., Laub, O., and Ben-Ze'ev, A.,** Attenuation in the control of gene expression in animal viruses, in *Developments in Molecular Virology IV*, Kohn, A. and Fuchs, P., Eds., Martinus Nijhoff Publishers, Boston, 1983, 1.

97. **Abulafia, R., Ben-Ze'ev, A., Hay, N., and Aloni, Y.,** The control of late SV40 transcription by the attenuation mechanism and transcriptionally active ternary complexes are associated with the nuclear matrix, *J. Mol. Biol.*, 172, 467, 1984.
98. **Hay, N. and Aloni, Y.,** Attenuation in SV40 as a mechanism of transcription termination by RNA polymerase B, *Nucl. Acids Res.*, 12, 1401, 1984.
99. **Sanger, F., Brownlee, G. G., and Barrel, B. G.,** A two-dimensional fractionation procedure for radioactive nucleotides, *J. Mol. Biol.*, 13, 373, 1965.
100. **Saucier, J. M. and Wang, J. C.,** Alteration of the DNA helix by *E. coli* RNA polymerase, *Nature (London) New Biol.*, 239, 167, 1972.
101. **Hsieh, T. S. and Wang, J. C.,** Physicochemical studies on interactions between DNA and RNA polymerase. Ultraviolet absorption measurements, *Nucl. Acids Res.*, 5, 3337, 1978.
102. **Winkler, M. E. and Yanofsky, C.,** Pausing of RNA polymerase during in vitro transcription of the tryptophan operon leader region, *Biochemistry*, 20, 3738, 1981.
103. **Farnham, P. J. and Platt, T.,** Rho-independent termination: dyad symmetry in DNA causes RNA polymerase to pause during transcription in vitro, *Nucl. Acids Res.*, 9, 563, 1981.
104. **Lycan, D. and Danna, K.,** Characterization of the 5′ termini of purified nascent simian virus 40 late transcripts, *J. Virol.*, 54, 264, 1983.
105. **Shaper, J. H., Pardoll, D. M., Kaufman, S. A., Barrck, E. R., Vogelstein, B., and Coffey, D.,** The relationship of the nuclear matrix to cellular structure and function, *Adv. Anz. Regul.*, 17, 213, 1979.
106. **Berezney, R.,** Dynamic properties of the nuclear matrix, in *The Cell Nucleus*, Vol. 7, Busch, H., Ed., Academic Press, New York, 1979, 413.
107. **Agutter, P. S. and Richardson, J. C.,** Nuclear non-chromatin proteinaceous structures: their role in the organization and function of the interphase nucleus, *J. Cell Sci.*, 44, 395, 1980.
108. **Faiferman, J. and Pogo, A. O.,** Isolation of a nuclear ribonucleoprotein network that contains heterogeneous RNA and is bound to the nuclear envelope, *Biochemistry*, 14, 3808, 1975.
109. **Herman, R., Weymouth, L., and Penman, S.,** Heterogeneous nuclear RNA-protein fibers in chromatin-depleted nuclei, *J. Cell Biol.*, 78, 663, 1978.
110. **Miller, T. E., Huang, C.-Y., and Pogo, A. D.,** Rat liver nuclear skeleton and ribonucleoprotein complexes containing hnRNA, *J. Cell Biol.*, 76, 675, 1978.
111. **Herlan, G., Eckert, W. A., Kaffenberger, W., and Wunderlich, F.,** Isolation and characterization of an RNA-containing nuclear matrix from tetrahymena macronuclei, *Biochemistry*, 18, 1782, 1979.
112. **Long, B. H., Huang, C. Y., and Pogo, A. O.,** Isolation and characterization of the nuclear matrix in Friend erythroleukemia cells: chromatin and hnRNA interaction with the nuclear matrix, *Cell*, 18, 1079, 1979.
113. **Mariman, E. C. M., Van Eekelen, C. A. G., Reinders, R. J., Berns, A. J. M., and Van Venrooij, W. J.,** Adenoviral heterogeneous nuclear RNA is associated with the host nuclear matrix during splicing, *J. Mol. Biol.*, 154, 103, 1982.
114. **Ross, D. A., Yen, R. W., and Chae, C. B.,** Association of globin ribonucleic acid and its precursors with the chicken erythroblast nuclear matrix, *Biochemistry*, 21, 764, 1982.
115. **Ben-Ze'ev, A., Abulafia, R., and Aloni, Y.,** SV40 virions and viral RNA metabolism are associated with cellular substructures, *EMBO J.*, 1, 1225, 1982.
116. **Ben-Ze'ev, A. and Aloni, Y.,** Synthesis and processing of SV40 RNA are associated with the nuclear matrix and are not followed by the accumulation of low molecular weight RNA processing products, *Virology*, 125, 475, 1983.
117. **Van Eekelen, C. A. G. and Van Venrooij, W. J.,** hnRNA and its attachment to a nuclear protein matrix, *J. Cell Biol.*, 88, 554, 1981.
118. **Jakobovits, E. B. and Aloni, Y.,** Isolation and characterization of various forms of simian virus 40 dna-protein complexes, *Virology*, 102, 107, 19 .
119. **Felsenfeld, G.,** Chromatin, *Nature (London)*, 271, 115, 1978.
120. **Korenberg, R. D. A.,** Structure of chromatin, *Annu. Rev. Biochem.*, 46 931, 1977.
121. **Chambon, P.,** The molecular biology of the eukaryotic genome is coming of age, *Cold Spring Harbor Symp. Quant. Biol.*, 42, 1209, 1977.
122. **Cremisi, C.,** The appearance of DNase I hypersensitive sites at the 5′ end of the late SV40 genes is correlated with the transcriptional switch, *Nucl. Acids Res.*, 9, 5949, 1981.
123. **Groudine, M., Eisenman, R., and Weintraub, H.,** Chromatin structure of endogenous retroviral genes and activation by an inhibitor of DNA methylation, *Nature (London)*, 292, 311, 1981.
124. **Herbormel, P., Saragosti, S., Blangy, D., and Yaniv, M.,** Fine structure of the origin-proximal DNAase-I hypersensitive region in wild-type and EC mutant polyoma, *Cell*, 25, 651, 1981.
125. **Keene, M. A., Corces, V., Lowenhaupt, K., and Elgin, S. C. R.,** DNAase I hypersensitive sites in Drosophila chromatin occur at the 5′ ends of regions of transcription, *Proc. Natl. Acad. Sci. U.S.A.*, 78, 143, 1981.

126. **Stadler, J., Larsen, A., Engel, J. D., Dolan, M., Groudine, M., and Weintraub, H.,** Tissue-specific DNA cleavages in the globin chromatin domain introduced by DNAase I, *Cell*, 20, 451, 1980.
127. **Mathis, D., Oudet, P., and Chambon, P.,** Structure of transcribing chromatin, *Prog. Nucl. Acids Res. Mol. Biol.*, 24, 1, 1980.
128. **Weintraub, H., Larsen, A., and Groudine, M.,** α-Globin-gene switching during the development of chicken embryos: expression and chromosome structure, *Cell*, 24, 333, 1981.
129. **Gariglio, P., Llopis, R., Oudet, P., and Chambon, P.,** The template of the isolated native simian virus 40 transcriptional complexes is a minichromosome, *J. Mol. Biol.*, 31, 75, 1979.
130. **Horowitz, M., Laub, O., Bratosin, S., and Aloni, Y.,** Splicing of SV40 late mRNA is a posttranscriptional process, *Nature (London)*, 257, 558, 1978.
131. **Beard, P. and Nyfeler, K.,** Transcription of simian virus 40 chromosomes in an extract of HeLa cells, *EMBO J.*, 1, 9, 1982.
132. **Weintraub, H.,** Recognition of specific DNA sequences in eukaryotic chromosome, *Nucl. Acids Res.*, 8, 4745, 1980.
133. **Aloni, Y.,** Splicing of viral mRNAs, *Prog. Nucl. Acid Res. Mol. Biol.*, 25, 1, 1981.
134. **Fey, G. and Hirt, B.,** Fingerprints of polyoma virus proteins and mouse histones, *Cold Spring Harbor Symp. Quant. Biol.*, 39, 235, 1974.
135. **Meinke, W., Hall, M. R., and Goldstein, D. A.,** Proteins in intracellular simian virus 40 nucleoprotein complexes: comparison with simian virus 40 core proteins, *J. Virol.*, 15, 439, 1975.
136. **Leer, Y., Tiryaki, D., and Westergaard, O.,** Termination of transcription in nucleoli isolated from Tetrahymena, *Proc. Natl. Acad. Sci. U.S.A.*, 76, 5563, 1979.
137. **Stunnenberg, H. G. and Birnstiel, M. L.,** Bioassay for components regulating eukaryotic gene expression: a chromosomal factor involved in the generation of histone mRNA 3′ termini, *Proc. Natl. Acad. Sci. U.S.A.*, 79, 6201, 1982.
138. **Birchmeier, C., Schumperli, D., Sconzo, G., and Birnstiel, M. L.,** 3′ Editing of mRNAs: sequence requirements and involvement of a 60-nucleotide RNA in maturation of histone mRNA precursors, *Proc. Natl. Acad. Sci. U.S.A.*, 81, 1057, 1984.
139. **Adhya, S., Sarkar, P., Valenzuela, D., and Maitra, U.,** 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. U.S.A.*, 76, 1613, 1979.
140. **Lee, F. and Yanofsky, C.,** Transcriptional termination at the trp operon attenuators of *Escherichia coli* and *Salmonella typhimurium:* RNA secondary structure and regulation of termination, *Proc. Natl. Acad. Sci. U.S.A.*, 74, 4365, 1977.
141. **Neff, N. F. and Chamberlin, M. J.,** Termination of transcription by *Escherichia coli* RNA polymerase in vitro is affected by ribonucleoside triphosphate base analogs, *J. Biol. Chem.*, 253, 2455, 1978.
142. **Farnham, P. J. and Platt, T.,** A model for transcription termination suggested by studies on the trp attenuator in vitro using base analogs, *Cell*, 20, 739, 1980.
143. **Platt, T.,** Termination of transcription and its regulation in the tryptophan operon of *E. coli*, *Cell*, 24, 10, 1981.
144. **Farnham, P. J., Greenblatt, J., and Platt, T.,** Effect of NusA protein on transcription termination in the tryptophan operon of *Escherichia coli*, *Cell*, 29, 945, 1982.
145. **Hatfield, G. W., Sharp, Y. A., and Rosenberg, M.,** Pausing and termination of human RNA polymerase II transcription at a procaryotic terminator, *Mol. Cell. Biol.*, 3, 1687, 1983.
146. **Murphy, J. T., Burgess, R. R. Dahlberg, J. E., and Lund, E.,** Transcription of a gene for human U_1 small nuclear RNA, *Cell*, 29, 265, 1982.
147. **Salditt-Georgieff, M. and Darnell, J. E., Jr.,** A precise termination site in the mouse β^{major}-globin transcription unit, *Proc. Natl. Acad. Sci. U.S.A.*, 80, 4694, 1983.
148. **Dubin, D. T., Montoya, J., Timko, K. D., and Attardi, G.,** Sequence analysis and precise mapping of the 3′ ends of HeLa cell mitochondrial ribosomal RNAs, *J. Mol. Biol.*, 157, 1, 1982.
149. **Biro, P. A. and Weissman, S. M.,** *Molecular Genetics*, Part 3, Taylor, J. H., Ed., Academic Press, New York, 1979, 177.
150. **Birchmeier, C., Grosschedl, R., and Birnstiel, M. L.,** Generation of authentic 3′ termini of an H_2A mRNA in vivo is dependent on a short inverted DNA repeat and on spacer sequences, *Cell*, 28, 739, 1982.
151. **Zaret, K. S. and Herman, F.,** DNA sequences required for efficient transcription termination in yeast, *Cell*, 28, 563, 1982.
152. **Bogenhagen, D. F. and Brown, D.,** Nucleotide sequences in xenopous 5S DNA required for transcription termination, *Cell*, 24, 261, 1981.
153. **Bendig, M. M. and Hentschel, C. C.,** Transcription of sea urchin histone genes in HeLa cells, *Nucl. Acids Res.*, 11, 2337, 1983.
154. **Wickens, M. P. and Gurdon, J. B.,** Posttranscriptional processing of simian virus 40 late transcripts in injected frog oocytes, *J. Mol. Biol.*, 163, 1, 1983.

155. **Jay, G., Nomura, S., Anderson, C. W., and Khoury, G.,** Identification of the SV40 agnogene product: a DNA binding protein, *Nature (London)*, 291, 346, 1981.
156. **Carbon, J., Shenk, T. E., and Berg, P.,** Biochemical procedure for production of small deletions in simian virus 40 DNA, *Proc. Natl. Acad. Sci. U.S.A.*, 72, 1392, 1975.
157. **Shenk, T. E., Carbon, Y., and Berg, P.,** Construction and analysis of viable deletion mutants of simian virus 40, *J. Virol.*, 18, 664, 1976.
158. **Barkan, A. and Mertz, J. E.,** DNA sequence analysis of simian virus 40 mutants with deletions mapping in the leader region of the late viral mRNAs: mutants with deletions similar in size and position exhibit varied phenotypes, *J. Virol.*, 37, 730, 1981.
159. **Mertz, J. E., Murphy, A., and Barkan, A.,** Mutants deleted in the agnogene of simian virus 40 define a new complementation group, *J. Virol.*, 45, 36, 1983.
160. **Margolskee, R. F. and Nathans, D.,** Suppression of a VP_1 mutant of simian virus 40 by missense mutations in serine codons of the viral agnogene, *J. Virol.*, 48, 405, 1983.
161. **Nomura, S., Khoury, G., and Jay, G.,** Subcellular localization of the simian virus 40 agnoprotein, *J. Virol.*, 45, 428, 1983.
162. **Hay, N., Kessler, M., and Aloni, Y.,** SV40 deletion mutant (d1861) with agnoprotein shortened by four amino acids, *Virology*, 137, 160, 1984.
163. **Thimmappaya, B., Reddy, V. B., Dhar, R., Celma, M., Subramanian, K. N., Zain, B. S., Pan, J., and Weissman, S. M.,** The structure of genes, intergenic sequences, and mRNA from SV40 virus, *Cold Spring Harbor Symp. Quant. Biol.*, 42, 449, 1977.
164. **Ben-Asher, E., Bratosin, S., and Aloni, Y.,** Intracellular DNA of the parvovirus minute virus of mice is organized in a minichromosome structure, *J. Virol.*, 41, 1044, 1982.
165. **Berns, K. I. and Hauswirth, W. W.,** Parvovirus DNA structure and replication, in *Replication of Mammalian Parvoviruses*, Tattersall, P., Ed., Cold spring Harbor Laboratory, Cold Spring Harbor, N.Y., 1978, 13.
166. **Bratosin, S., Laub, O., Tal, Y., and Aloni, Y.,** Mechanism of circularization of linear DNAs: circular parvovirus MVM DNA is formed by a "noose" sliding in a "lasso"-like DNA structure, *Proc. Natl. Acad. Sci. U.S.A.*, 76, 4289, 1979.
167. **Tattersall, P. and Ward, D. C.,** The rolling hairpin: a model for the replication of parvovirus and linear chromosomal DNA, *Nature (London)*, 263, 106, 1976.
168. **Green, M. R., Lebovitz, R. M., and Roeder, R. G.,** Expression of autonomous parvovirus H1 genome: evidence for a single transcriptional unit and multiple spliced polyadenylated transcripts, *Cell*, 17, 967, 1979.
169. **Tal, J., Ron, D., Tattersall, P., Bratosin, S., and Aloni, Y.,** About 30% of minute virus of mice RNA is spliced out following polyadenylation, *Nature (London)*, 279, 649, 1979.
170. **Ward, D. C. and Tattersall, P., Eds.,** *Replication of Mammalian Parvoviruses*, Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y., 1978.
171. **Tattersall, P. and Ward, D. C.,** The parvoviruses — an introduction, in *Replication of Mammalian Parvoviruses*, Ward, D. C. and Tattersall, P., Eds., Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y., 1978.
172. **Siegl, G.,** The parvoviruses, in *Virology Monographs*, Goard, S. and Hallauer, C., Eds., Springer-Verlag, New York, 1976.
173. **Astell, C., Smith, M., Chow, M., and Ward, D.,** Structure of the 3′ hairpin termini of four rodent parvovirus genomes: nucleotide sequence homology of origins of DNA replication, *Cell*, 17, 691, 1979.
174. **McMaster, G. K., Beard, P., Engers, H. D., and Hirt, B.,** Characterization of an immunosuppressive parvovirus related to the minute virus of mice, *J. Virol.*, 38, 317, 1981.
175. **Rhode, S. L., III,** Replication process of the parvovirus H-1. X. Isolation of mutant defective in replicative form DNA replication, *J. Virol.*, 25, 215, 1978.
176. **Rhode, S. L., III,** Complementation for replicative form DNA replication of a deletion mutant of H-1 by various parvoviruses, *J. Virol.*, 42, 1118, 1982.
177. **Pintel, D., Dadachanji, D., Astell, C., and Ward, D. C.,** The genome of minute virus of mice, an autonomous parvovirus, encodes two overlapping transcription units, *Nucl. Acids Res.*, 11, 1019, 1983.
178. **Rhode, S. L., III and Paradiso, P. R.,** Parvovirus genome: nucleotide sequence of H-1 and mapping of its genes by hybrid-arrested translation, *J. Virol.*, 45, 173, 1983.
179. **Astell, C. R., Thomson, M., Merchlinsky, M., and Ward, D. C.,** The complete DNA sequence of minute virus of mice, an autonomous parvovirus, *Nucl. Acids Res.*, 11, 999, 1983.
180. **Ben-Asher, E. and Aloni, Y.,** Transcription of minute virus of mice, an autonomous parvovirus, may be regulated by attenuation, *J. Virol.*, 52, 266, 1984.
181. **Darnell, J. E., Jr.,** Variety in the level of gene control in eukaryotic cells, *Nature (London)*, 297, 365, 1982.

182. **Birkenmeier, E., Radonovich, M. F., Shani, M., and Salzman, N. P.,** The SV40 DNA template for transcription of late mRNA in viral nucleoprotein complexes, *Cell*, 11, 495, 1977.
183. **Gariglio, P. and Mousset, S.,** Isolation and partial characterization of a nuclear RNA polymerase-SV40 DNA complex, *FEBS Lett.*, 56, 149, 1975.
184. **Gariglio, P. and Mousset, S.,** Characterization of a soluble simian virus 40 transcription complex, *Eur. J. Biochem.*, 76, 583, 1977.
185. **Laub, O. and Aloni, Y.,** VII SV40-DNA RNA polymerase complex isolated from productively infected cells transcribed in vitro, *Virology*, 75, 346, 1976.
186. **Shani, M., Birkenmeier, E., May, E., and Salzman, N.,** Properties of simian virus 40 transcriptional intermediates isolated from nuclei of permissive cells, *J. Virol.*, 23, 20, 1977.
187. **Lee, F., Squires, C. L., Squires, C., and Yanofsky, C.,** Termination of transcription in vitro in the *Escherichia coli* tryptophan operon leader region, *J. Mol. Biol.*, 103, 383, 1976.
188. **Maizels, N. M.,** The nucleotide sequence of the lactose messenger ribonucleic acid transcribed from the UV5 promoter mutant of *Escherichia coli*, *Proc. Natl. Acad. Sci. U.S.A.*, 70, 3585, 1973.
189. **Vennström, B., Persson, H., Petterson, U., and Philipson, L.,** A DRB (5,6,dichloro-β-D-ribofuranosylbenzimidazole) resistant adenovirus mRNA, *Nucl. Acids Res.*, 7, 1405, 1979.
190. **Tamm, I., Sehgal, P. B., Lamb, R. A., and Goldberg, A. R.,** *Antiviral Drugs and Interferon: The Molecular Basis of Their Activity*, Becker, Y., Ed., Martinus Nijhoff Press, The Hague, Netherlands, 1983.
191. **Akushärvi, G. and Petterson, U.,** Sequence analysis of adenovirus DNA: complete nucleotide sequence of the spliced 5′ noncoding region of adenovirus 2 hexon messenger RNA, *Cell*, 16, 841, 1979.
192. **Chow, L. T. and Broker, T.,** The spliced structures of adenovirus 2 fiber message and the other late mRNAs, *Cell*, 15, 497, 1978.
193. **Virtanen, A., Aleström, P., Persson, H., Katze, M. G., and Petterson, U.,** An adenovirus agnogene, *Nucl. Acids Res.*, 10, 2539, 1982.
194. **Nevis, J. R. and Wilson, M. D.,** Regulation of adenovirus-2-gene expression at the level of transcriptional termination and RNA processing, *Nature (London)*, 290, 113, 1981.
195. **Akusjärvi, G. and Persson, H.,** Control of RNA splicing and termination in the major control adenovirus transcription unit, *Nature (London)*, 292, 420, 1981.
196. **Montandon, P. E. and Acheson, N. H.,** Synthesis of prematurely terminated late transcripts of polyoma virus DNA is resistant to inhibition by 5,6-dichloro-1-β-D-ribofurnaosylbenzimidazole, *J. Gen. Virol.*, 59, 367, 1982.
197. **Testa, D., Chanda, P. K., and Banergee, A. K.,** Unique mode of transcription in vitro by vesicular stomatitis virus, *Cell*, 21, 267, 1980.
198. **Ball, L. A. and Wertz, G. W.,** VSV RNA synthesis: how can you be positive?, *Cell*, 26, 143, 1981.
199. **Benz, E. W., Wydro, R. M., Nadal-Ginard, B., and Dina, D.,** Moloney murine sarcoma proviral DNA is a transcriptional unit, *Nature (London)*, 288, 665, 1980.
200. **Gariglio, P., Bellard, M., and Chambon, P.,** Clustering of RNA polymerase B molecules in the 5′ moiety of the adult β-globin gene of hen erythrocytes, *Nucl. Acids Res.*, 9, 2589, 1981.
201. **Bishop, M. J.,** Cancer genes come of age, *Cell*, 32, 1018, 1983.
202. **Andreadis, A., Hsu, Y.-P., Kohlhaw, G. B., and Schimmel, P.,** Nucleotide sequence of yeast LEU 2 shows 5′-noncoding region has sequences congate to leucine, *Cell*, 31, 319, 1982.
203. **Lodish, H. F.,** Secondary structure of bacteriophage f2 ribonucleic acid and the initiation of in vitro protein biosynthesis, *J. Mol. Biol.*, 50, 689, 1970.
204. **Steitz, J. A.,** Polypeptide chain initiation: nucleotide sequences of three ribosomal binding sites in bacteriophage R17 RNA, *Nature (London)*, 224, 957, 1969.
205. **Fiers, W., Contreras, R., Duerinck, F., Haegeman, G., Insertant, D., Merregaert, J., Min Jou, W., Molemans, F., Raeymaekers, A., Van den Berghe, A., Volckaert, C., and Ysebaert, M.,** Complete nucleotide sequence of bacteriophage MS2 RNA: primary and secondary structure of the replicase gene, *Nature (London)*, 260, 500, 1976.
206. **Insertant, D. and Fiers, W.,** Secondary structure of RNA and efficiency of translation initiation, *Gene*, 9, 1, 1980.
207. **Lindahl, L., Archer, R., and Zengel, J. M.,** Transcription of the S10 ribosomal protein operon is regulated by an attenuator in the leader, *Cell*, 33, 241, 1983.
208. **Baralle, F. E.,** The functional significance of leader and trailer sequences in eukaryotic mRNAs, *Int. Rev. Cytol.*, 31, 71, 1983.
209. **Saito, H., Hayday, A. C., Wiman, K., Hayward, W. S., and Tonegawa, S.,** Activation of the c-myc by translocation: a model for translational control, *Proc. Natl. Acad. Sci. U.S.A.*, 80, 7478, 1983.
210. **Smith, S. O.,** Recovery of DNA from gels, *Methods Enzymol.*, 65, 371, 1980.
211. **Tinoco, I., Borer, P. N., Dengler, B., Levine, M. D., Uhlenbeck, O., Crothers, D., and Gralia, J.,** Improved estimation of secondary structure in ribonucleic acids, *Nature (London) New Biol.*, 246, 40, 1973.

212. **Le Moullec, J. M., Akusjarvi, G., Stalhandske, P., Pettersson, U., Chambraud, B., Gilardi, P., Nasri, M., and Perricaudet, M.,** Polyadenylic acid addition sites in the adenovirus type 2 major late transcription unit, *J. Virol.*, 48, 127, 1983.
213. **Hay, N. and Aloni, Y.,** unpublished results.
214. **Mertz, J.,** personal communication.
215. **Ben-Ze'ev, A., Skolnik-David, H., and Aloni, Y.,** unpublished results.
216. **Pruzan, R., Ben-Ze'ev, A., and Aloni, Y.,** unpublished results.
217. **Hay, N. and Aloni, Y.,** *Mol. Cell. Biol.*, 1985, in press.
218. **King, D. and Aloni, Y.,** in preparation.