Volume 9 Number 2 1981 Nucleic Acids Research

Initiation of transcription by RNA polymerase II in permeable, SV40-infected or noninfected, CV1 cells; evidence for multiple promoters of SV40 late transcription

Roland Contreras and Walter Fiers

Laboratory of Molecular Biology, State University of Ghent, Ledeganckstraat 35, B-9000 Ghent, Belgium

Received 27 November 1980

ABSTRACT

CV1 cells were made permeable by treatment with lysolecithin and incubated in a transcription mixture containing ribonucleoside triphosphates including ATP or GTP ^{32}P-labeled either in the α or β position. 5'-terminal cap structures ($^{7m}G^{\gamma\beta\alpha}pppX$) on newly synthesized RNA were analyzed by digestion with nuclease P1 or with ribonuclease T2/bacterial alkaline phosphatase. Cap structures obtained after labeling with α-^{32}P-GTP show that the ^{32}P is found only adjacent to the ^{7m}G residue (i.e., in the γ position) and adjacent to the penultimate Gm or G nucleotide (i.e., in the α position). Analysis of RNA synthesized in the presence of β-^{32}P-ATP, however, shows GpppA cap structures which are labeled only in the β position. In the presence of β-^{32}P-GTP, only GpppG structures are labeled; these findings exclude the hypothesis that caps are synthesized from GTP and a monophosphate 5'-terminal RNA molecule. The results imply that the initial transcripts are used for cap formation, which indicates that the large majority (if not all) of capping sites correspond to initiation sites for transcription. In cells infected with wild-type SV40 the distribution of virus-specific caps is similar when labeled either with β-^{32}P-ATP or with α-^{32}P-GTP or with ^{32}P-phosphate. Thus, evidence is presented that heterogeneity of the cap structures in late SV40 is a consequence of independent initiation events and not of processing of a primary transcript followed by capping of the 5' ends generated.

INTRODUCTION

Molecular characterization of a transcription unit on a known genome sequence requires, as a first datum, knowledge of the initiation point of RNA synthesis. This point can be deduced by comparing the genome sequence with the 5' end structure of initial transcripts. In eukaryotic cells the 5' terminus of mRNA usually occurs in the form of a cap structure with the general formula $^{7m}GpppX_m$ (1). Capped eukaryotic viral RNAs so far have been found to contain almost exclusively a purine nucleotide in the 5'-penultimate position. Cap structures are formed very early in hnRNA synthesis (2) and

some experiments with polyhedrosis virus mRNA synthesis even suggest that in this system capping is the initial step in RNA formation (3). Thus far it has not been possible to obtain and characterize noncapped RNA polymerase II primary transcripts under in vivo conditions. By ultraviolet mapping (4), pulse labeling (5,6) and electron microscope analysis (7), it was shown that in the adenovirus 2 major late transcription unit the initiation point of transcription and the capping site are identical or at least in very close proximity.

In SV40, analysis of late transcription products by reverse transcription (8) and direct examination of the cap structures (9,10) has indicated a heterogeneity of 5' termini. These termini have been localized on the SV40 genome between 0.68 and 0.74 map unit, which corresponds to a stretch of approximately 300 base pairs. 5'-terminal heterogeneity in SV40 late mRNA could reflect different transcription initiation events, involving an unusual type of promoter activity, or it could result from unequal 5'-terminal processing of a primary transcript with subsequent capping by a phosphorylation and a guanylylation reaction (11). Thus far the most prominent capping mechanism involves the transfer of a GMP residue from GTP to a 5'-di- or triphosphate-containing RNA molecule. This mechanism operates in the capping of vaccinia virus RNA (12,13), reovirus RNA (14,15), and cytoplasmic polyhedrosis virus RNA (3), and it has been observed in HeLa cell nuclei (16,17), in rat liver nuclei (18) and in L-cell nuclei (19). A second mechanism, in which a GDP from the capping GTP is transferred to the 5'-terminal pN nucleotide of the RNA chain, was postulated for vesicular stomatitis virus mRNA synthesis (20) and for the cellular mRNA pyrimidine-containing caps (21). Thus, by analyzing the cap structures synthesized in the presence of ATP or GTP ^{32}P-labeled in the β position it should be possible to demonstrate which capping mechanism is active in CV1 cells and also to distinguish which pathway is responsible for the heterogeneity found in the SV40 cap pattern.

To enable the entry of α- or β-^{32}P-XTPs or nucleoside triphosphate analogs, cells were made permeable with lysolecithin. It has been shown that initiation of RNA synthesis of class II genes can take place in isolated nuclei (22,23), although nuclei often were found to be deficient in initiation by RNA polymerase II (24-26). Lysolecithin treatment (27) is a milder approach than the rather unphysiological procedure of nuclei preparation; it avoids the possible loss of factors necessary for correct initiation of class II transcripts and provides a transcription system which better approximates an in vivo system. We show that RNA synthesized under these conditions carry

β-labeled GpppG or GpppA type caps when the radiolabeled precursor is β-^{32}P-GTP or β-^{32}P-ATP, respectively. After infection of cells with SV40 and analysis of caps derived from SV40-specific RNA, it was found that the heterogeneity previously observed at the 5' terminus of in vivo mRNAs (9,10,28) is retained when the RNA is synthesized in the presence of β-^{32}P-ATP. This means that cap sites correspond to RNA initiation events.

MATERIALS AND METHODS

Growth and permeabilization of cells

CV1 monkey cells (Bio-Cult, Paisley, Scotland) were grown in roller bottles to subconfluency and either made permeable (as described in ref. 27) or infected with 60 PFU of SV40 (strain 776) per cell. At 40-44 hrs after infection the cells were collected by treatment with trypsin and made permeable. The permeability of the cells was monitored by staining with 0.1% trypan blue.

Transcription, isolation and analysis of RNA

The transcription mixture was added to the collected (1 min at 800xg), permeable cells within 8 minutes after addition of the lysolecithin and incubated at 30°C for various periods. The transcription conditions were as described by Mory and Gefter (23), except that the S-adenosylmethionine (SAM) concentration was raised to 2.5 mM. The concentration of the cold nucleoside triphosphates was also as described by Mory and Gefter (23), and the ^{32}P-labeled nucleoside triphosphates (The Radiochemical Centre, Amersham, England; specific activity between 100 and 400 Ci/mmol) were kept at concentrations between 2.5 and 10 μM. In the preparative experiments the cells were incubated in the transcription mixture at a maximal concentration of 5 x 10^7 per ml. The kinetics of the transcription were followed as described by Smith et al. (22). The cells were lysed by addition of 0.2 M EDTA (to 20 mM) and 10% sarkosyl (to 1%). The mixture was digested with predigested proteinase K (Merck, Darmstadt, F.R.G.) at 0.5 mg/ml for 30 min at 37°C and extracted three times with phenol-chloroform (1/1). After precipitation of the nucleic acid from 1 M NaCl the DNA was digested with pancreatic DNase (Worthington Biochemicals Corp., Freehold, N.J.) at 0.25 mg/ml for 60 min at 37°C. The pancreatic DNase was degraded by a second digestion with proteinase K. After threefold phenol-chloroform extraction and precipitation, the α- or β-^{32}P-XTPs were removed by passage over a Sephadex G50 column (0.6 x 20 cm). The RNA was

precipitated and analyzed by digestion with nuclease P1 or T2 as described by Haegeman and Fiers (29), or hybridized to SV40 DNA immobilized on nitro-cellulose filters (Millipore Corp., Bedford, Mass.)(30), or on DBM-paper (Schleicher and Schuell Inc., Keene, N.H.)(31) in 4 x SSC, 50% formamide at 37°C with a DNA/RNA ratio of approximately 10/1. The hybridized RNA was eluted for 4 min with water at 80°C (22). The purity of the SV40-specific RNA was estimated as better than 90% from the ribonuclease T1 fingerprints. Digestion of the eluted RNA with nuclease P1, T1 and T2 was described by Haegeman and Fiers (29). Detection, elution and further analysis of the 5'-terminal structures was also as described by Haegeman and Fiers (29).

RESULTS

General characteristics of transcription in cells treated with lysolecithin

The procedure used by Miller, Castellot and Pardee (27) to render Chinese hamster ovary cells in suspension culture permeable could be used without modification to make CV1 cells isolated from monolayers by trypsin-EDTA treatment permeable. In initial experiments monolayer cells were successfully rendered permeable, but this system was abandoned in favor of cell suspensions because large volumes of transcription mixture were required to cover a small number of cells and because the cells detached during transcription. Using the transcription mixture described by Miller, Castellot and Pardee (27), it was not possible to obtain a linear incorporation of ribotriphosphates for more than 15 min. However, as seen in Fig. 1, a linear incorporation of rXTPs into RNA for longer than 60 min was obtained with the transcription mixture described by Mory and Gefter (23). This improvement may be due to the introduction of an ATP-generating system, since a rather fast degradation of ATP was observed. A linear increase in RNA synthesis for more than 60 min strongly suggests that initiation of RNA synthesis occurs under these labeling conditions.

Furthermore, as can be seen in Fig. 1, substitution of ATP by γ-S-ATP resulted in stimulation of RNA synthesis; we interpret this result as being due to a slower rate of nucleoside triphosphate breakdown. The kinetics of the RNA synthesis are dependent on the temperature. Most of the transcriptions were at 30°C and at a cell concentration of 10^7 cells/ml or lower. RNA synthesis was inhibited about 40% by α-amanitin concentrations between 0.5 and 2.5 µg/ml. Approximately the same level of inhibition was found for permeable mouse L-cells, for monolayer cells or cells after trypsinization, and for SV40-infected or noninfected cells. This type of inhibition suggests that

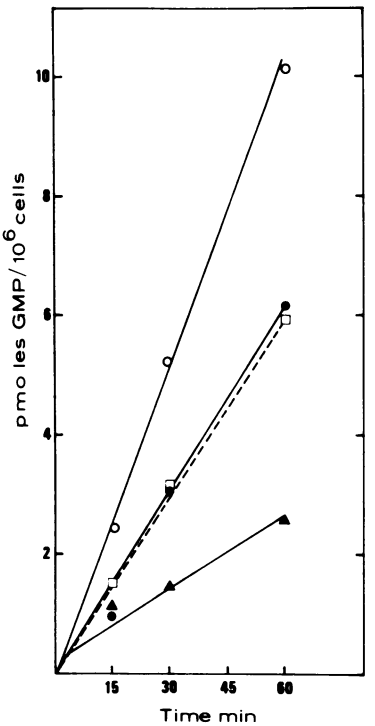


Fig. 1. Rate of RNA synthesis in permeabilized CV1 cells. CV1 cells were grown to subconfluency on a Roux flask, washed with phosphate-buffered saline, collected by treatment with trypsin and made permeable as described in Materials and Methods. The cells were divided into four equal aliquots and incubated, respectively, in the standard transcription mixture (□--□), in the same but with 1 mM ATP replaced by 1 mM γ-S-ATP (○—○) or 1 mM pCH_2ppA (▲—▲), or in a mixture which contained 0.5 mM SAH (●—●) instead of SAM. The kinetics were monitored by spotting samples on DE81 filter paper and washing the XTPs with 0.5 M KH_2PO_4 (22). The concentration of the α-^{32}P-GTP (300 Ci/mmol) was 2.5 µM.

about 40% of the RNA synthesized in this system is dependent on RNA polymerase II (32). The RNA synthesis is also sensitive to actinomycin D: 98% inhibition at 5 µg/ml, 65% at 1 µg/ml and 22% at 0.1 µg/ml were found.

Cap synthesis in CV1 cells labeled with α-^{32}P-GTP and α-^{32}P-ATP

CV1 cells were grown to confluency in roller bottles, treated with trypsin and made permeable as described in Materials and Methods. RNA was synthesized in the presence of 2.5 µM α-^{32}P-GTP and 1.5 mM SAM. The RNA was extracted, digested with nuclease P1 and bacterial alkaline phosphatase and

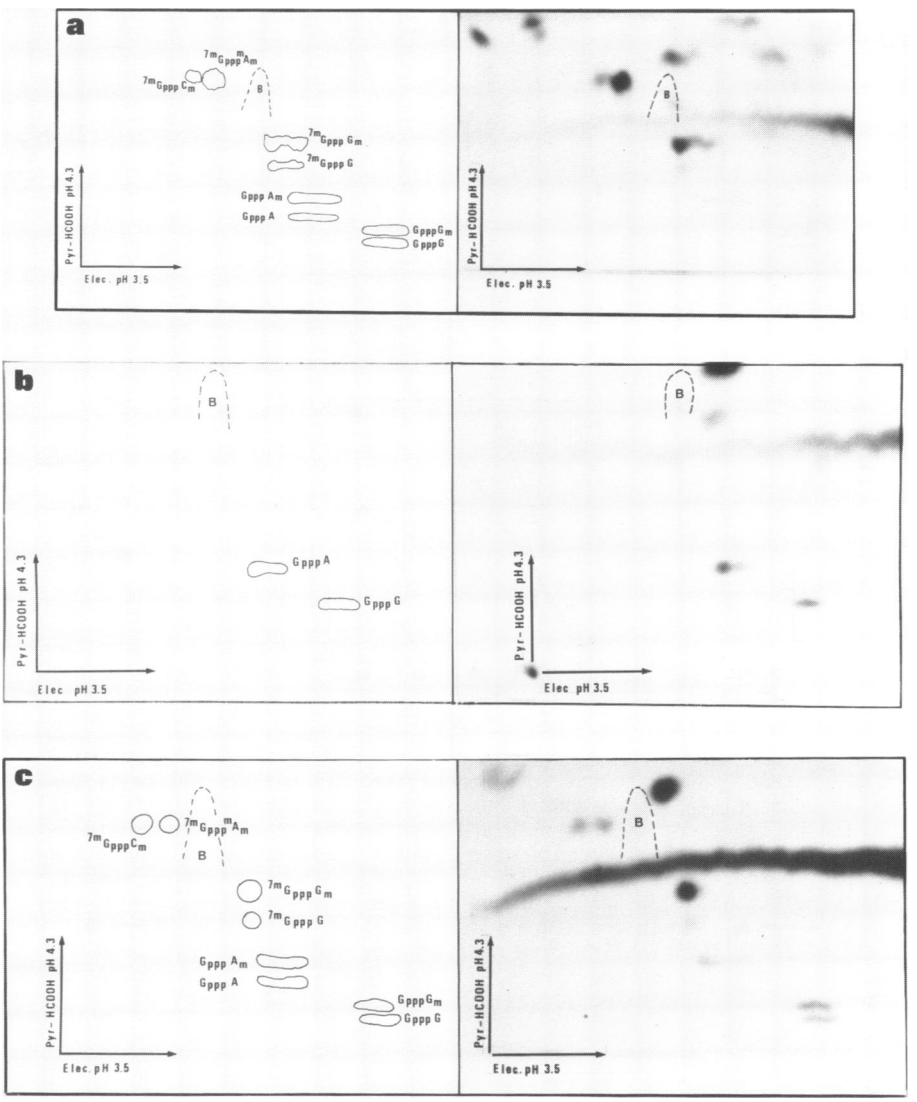
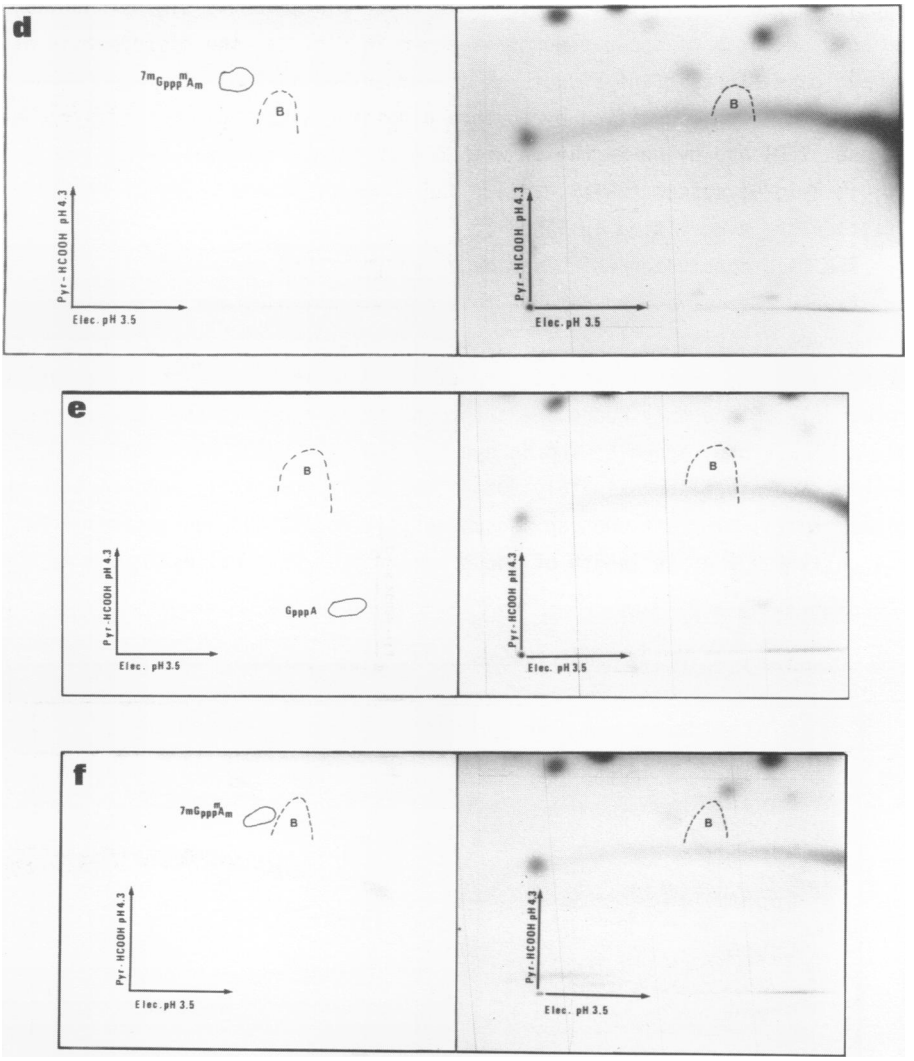


Fig. 2. Nuclease P1/bacterial alkaline phosphatase analyses of CV1 total cellular RNA labeled in the presence of α-32P-GTP or α-32P-ATP. Noninfected CV1 cells were grown in a roller bottle to subconfluency, treated with trypsin and made permeable (27). In these experiments the cells were not removed from the permeabilization mixture, but added directly to the transcription mixture containing either (a) 2.5 μM α-32P-GTP (specific activity ~400 Ci/mmol), (b) α-32P-GTP and 0.5 mM SAH instead of SAM, (c) α-32P-GTP and 1 mM pCH_2ppA instead of ATP, (d) 2 μM α-32P-ATP (specific activity, 320 Ci/mmol), (e) 2 μM α-32P-ATP and 0.5 mM SAH, (f) 2 μM α-32P-ATP and 1 mM



pCH_2ppG. After 60 min of RNA synthesis at 30°C the total RNA was isolated, digested with nuclease P1 and bacterial alkaline phosphatase and fingerprinted as described by Haegeman and Fiers (29). Identification of the P1-resistant spots was also carried out by comparing with authentic optical density markers (P-L Biochemicals Inc., Milwaukee, Wis.). The results of the tobacco acid pyrophosphatase digests of the P1 caps are shown in Fig. 3. The horizontal line is the tailing of the 32Pi which does not show on the photographs.

fractionated on a polyethyleneimine thin-layer plate in two dimensions (29). The autoradiogram of the cap cores is shown in Fig. 2a; the distribution of the nuclease P1-resistant structures is summarized in Table 1. The radioactive spots were identified by further digestion with tobacco acid pyrophosphatase (TAP) and by co-migration with commercial 5'-terminal markers on a two-dimensional system (29)(Fig. 3). The mononucleotides found after TAP digestion are also listed in Table 1.

The main conclusion which can be drawn from this experiment is that cap formation occurs at a good rate in this transcription system. At least 0.02% of the radioactivity is found in the cap structures (this value is presumably an underestimate since it is not corrected for losses during the fingerprint procedure). Since only guanosine triphosphate is added as the labeled precursor, at least one capped 5' end is synthesized per 20,000 nucleotides incorporated, and since approximately 50% of the RNA synthesis is dependent on RNA polymerase II, at least one cap is synthesized for 10,000 nucleotides. In HeLa cells the average length of the transcription unit was estimated as 5000

Table 1. Distribution of α-^{32}P-GTP labeled cap cores

Cap structure	Percent of total[a]	TAP digest[b]
^{7m}GpppAm[c]	46.8	mpG[e]
GpppAm[c]	9.1	pG
GpppA	6.5	pG
^{7m}GpppGm	20.4	mpG, pGm
^{7m}GpppG	3.0	mpG, pG
GpppGm	1.6	pGm, pG
GpppG	1.6	pG
^{7m}GpppCm[d]	11.0	mpG

a. Refers to total moles of cap structure.
b. The TAP digests are shown in Fig. 3.
c. We have not determined the extent of methylation of the 6-amino position of the A residue.
d. The cap ^{7m}GpppCm was identified by two-dimensional cofractionation with the authentic optical density marker.
e. mpG stands for 7-methyl-pG; pGm stands for 2'-methyl-pG.

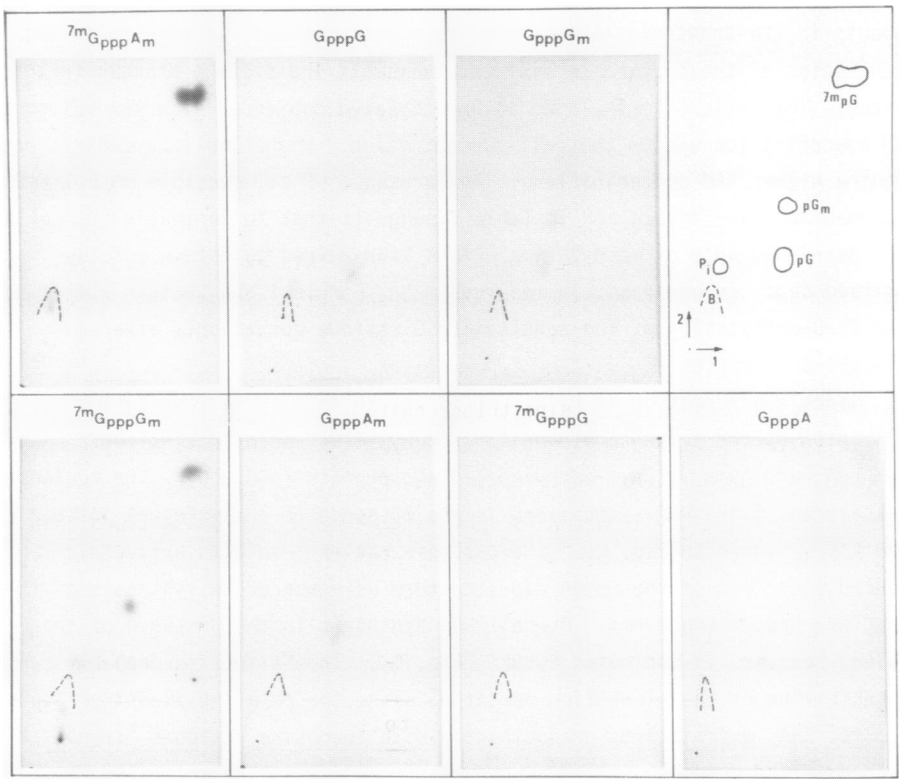


Fig. 3. Characterization of cap cores by analysis with tobacco acid phosphatase. Cap cores obtained from α-32P-GTP labeled CV1 RNA (Fig. 2) were eluted from the polyethyleneimine thin-layer plate with 2 M triethylamine bicarbonate, pH 10, and digested with 0.8 unit TAP (BRL Inc., Rockville, Md) in 10 µl 0.05 M NaAc, pH 5.0, 0.01 M β-mercaptoethanol during 2 hr at 37°C. The digests were applied to PEI plates (6.6 x 10 cm) and fractionated in the first (short) dimension with 15% formic acid and in the second (long) dimension with 0.5 M HCOOH/pyridine, pH 4.3 (60). Optical density markers of 7-methyl-pG, pG and pGm were cochromatographed and detected by ultraviolet illumination. It should be noted that the small amount (<10%) of 32Pi could result from a contaminating phosphatase activity in the pyrophosphatase preparation.

nucleotides (33). Assuming an identical average length for the transcription unit in permeabilized CV1 cells, 50% of the synthesized RNA will have a newly-formed cap structure.

The nuclease P1 analysis of RNA synthesized in the presence of S-adenosyl-homocysteine is shown in Fig. 2b. Only completely unmethylated cap cores were

found. Since these unmethylated blocked termini are also found in variable amounts in the transcription mixture which contains SAM, it appears that the methylation of the blocked termini is a somewhat inefficient process in this system. The ratio of methylated to unmethylated cap cores is dependent on the SAM concentration and on the cell concentration. Higher cell concentrations require higher SAM concentrations. The presence of all possible methylated intermediates in Fig. 2a and in Table 1 suggests that in permeable CV1 cells no ordered sequence of methylation events is required to obtain a fully modified cap. In contrast, in reovirus mRNA, Furuichi and Shatkin (14) found that 2'-O-methylation of the penultimate G residue occurs only after 7-methyl-G formation.

Since cap formation is rather efficient in this system, we tested the inhibition by the triphosphate analogs adenylyl(β,γ-methylene)-diphosphonate (pCH_2ppA) and guanylyl(β,γ-methylene)-diphosphonate (pCH_2ppG). The nuclease P1 analysis of the RNA synthesized in the presence of 1 mM pCH_2ppA instead of 1 mM ATP is shown in Fig. 2c. The cap core pattern exhibits a fivefold reduced synthesis of the GpppA cap structure as compared to synthesis of the GpppC and GpppG structures. Overall RNA synthesis in the presence of the analog, however, is inhibited by 52% (Fig. 1). This effect is probably caused by inhibition of the elongation reaction, since the relative amount of SV40-specific RNA, which starts predominantly with adenosine residues, is not changed in the presence of the pCH_2ppA analog (data not shown). The remaining synthesis of GpppA cap structures can easily be explained by the ATP pool which is present in the cells. Conversely, when α-^{32}P-ATP is used as labeled precursor in the presence of pCH_2ppG, no inhibition of the GpppA-type caps (e.g., $^{7m}Gppp^{m}A_m$) is observed (Fig. 2f). Thus, these experiments are circumstantial evidence for the model in which only 5'-pyrophosphorylated primary transcripts are guanylylated to form a cap structure.

The experiment shown in Fig. 2d and 2e, in which α-^{32}P-ATP was used as labeled precursor, confirm the identity of the GpppA cap structures and show that no detectable exchange of the α-phosphate of the ATP occurs.

Cap synthesis in CV1 cells labeled with β-^{32}P-ATP and β-^{32}P-GTP

The 5' end of newly initiated RNA can be labeled with nucleoside triphosphates carrying ^{32}P in the β position (25,34). The nuclease P1 digestion pattern of RNA synthesized in CV1 cells in the presence of β-^{32}P-GTP is shown in Fig. 4a. All the spots observed on the chromatogram were identified by co-migration with authentic optical density markers and the cap structures were

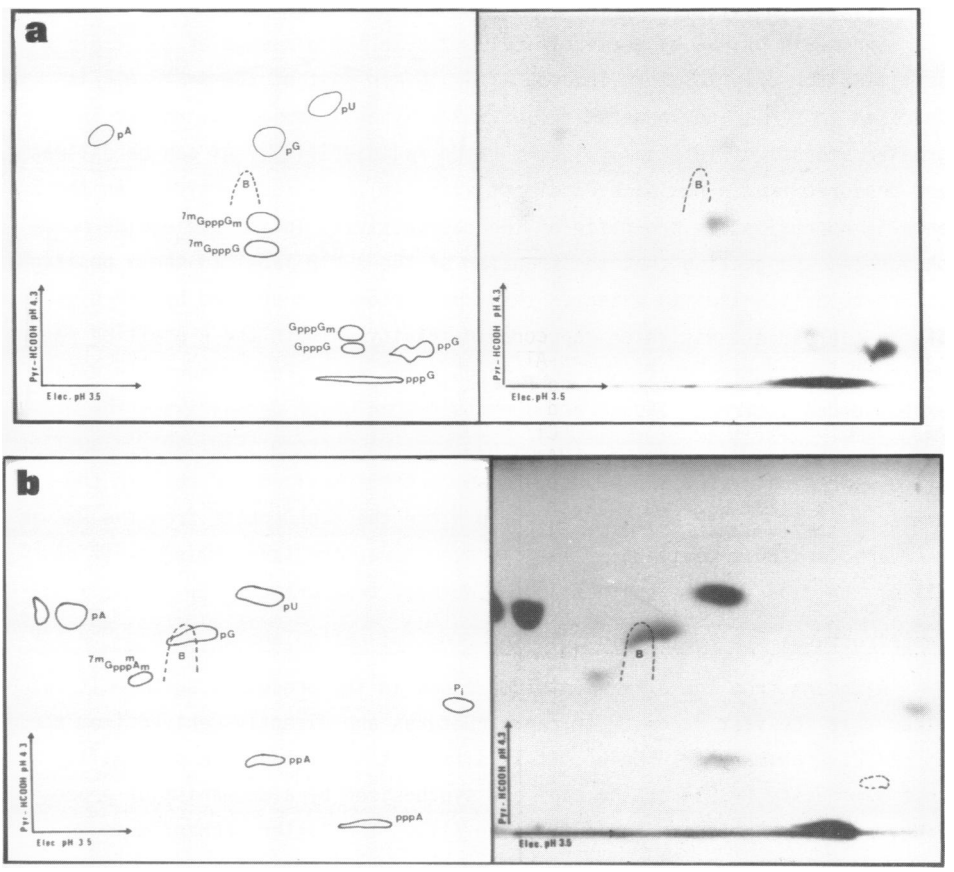
further characterized by TAP digestion (Fig. 4c). The P1 digestion pattern shows only GpppG-derived cap structures and no traces of GpppA structures, even after overexposure of the chromatogram, which would detect less than 5% contaminating components. The absence of label in the GpppA caps allows the conclusion that the β-phosphate in the GpppA cap structures originates from the RNA, and there is no reason to assume another mechanism of cap synthesis for GpppG caps than for GpppA caps. The TAP digest of the cap structures releases only ^{32}Pi, suggesting that exchange of ^{32}Pi from the β to the α position occurs only to a very limited extent (Fig. 4c). Reincorporation of ^{32}Pi formed as a degradation product of the β-^{32}P-GTP was reduced by addition of 1 mM sodium phosphate to the transcription mixture and by limiting the transcription time to 30 min. Since the amount of radioactivity found in the mononucleotides is lower than the amount found in the cap structures (Fig. 4a), the degree of exchange is very low (<0.1%), assuming that RNA of normal length (>1000 nucleotides) is synthesized.

Synthesis of RNA by permeable CV1 cells in the presence of β-^{32}P-ATP was permitted for only 10 min. Indeed, after 10 min most of the β-^{32}P-ATP is degraded to ^{32}Pi, and longer times of RNA synthesis would only result in greater amounts of internally incorporated radioactivity. As can be estimated and measured from a nuclease P1 digest (Fig. 4b), the radioactivity in the 5' ends is approximately one-fifth of the radioactivity in the nucleoside monophosphates, suggesting that the transfer of the β-^{32}P label to the α position occurs to only a limited extent. This conclusion is supported by TAP digestion of the cap structures, which shows radioactivity only in the β position and no counts in the position of ^{7m}pG, ^{m}pAm, or pA (Fig. 4c). Furthermore, if extensive exchange had occurred, cap structures other than GpppA derivatives (for example ^{7m}GpppG$_m$) would have appeared in the nuclease P1 pattern (Fig. 4b). Thus, since approximately equal amounts of caps are synthesized per cell in the presence of β-^{32}P-GTP and α-^{32}P-GTP and since the β phosphate from the RNA is recovered in the β position of the cap structure, the large majority if not all of the caps must be synthesized on primary transcripts.

Cap structures synthesized in SV40-infected CV1 cells labeled with β-^{32}P-ATP

Although from the experiments described in the preceding section it is clear that the vast majority of cap structures are directly derived from the 5' end of a primary transcript, it is difficult to exclude the possibility that a minority (<10%) of the caps are synthesized by a mechanism of processing, phosphorylation and guanylylation (11). The latter mechanism is an

especially attractive hypothesis in the case of SV40 mRNA, as it allows to readily explain the remarkable heterogeneity in cap structures (9,10,28,35). To test whether the heterogeneous pattern of SV40 capping arises from independent initiation events or from a unique initiation event followed by processing and capping at different sites, SV40-infected cells were harvested at 40 hrs post-infection by treatment with trypsin, made permeable, and labeled with β-^{32}P-ATP. The SV40-specific RNA was isolated by hybridization to immobilized SV40 DNA and analyzed by consecutive ribonuclease T2 and bacterial alkaline phosphatase digestion, and fingerprinted (Fig. 5a). The spots were identified by comparison with a pattern obtained by uniform labeling (G. Haegeman, unpublished results) or by labeling with α-^{32}P-GTP (Fig. 5c). The pattern obtained after labeling with β-^{32}P-ATP is obviously heterogeneous and shows (by visual interpretation) a distribution which is



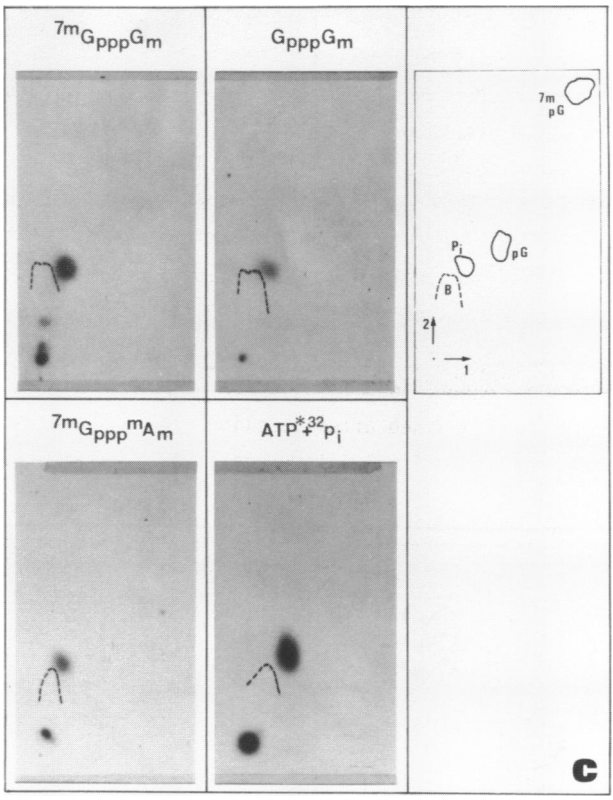
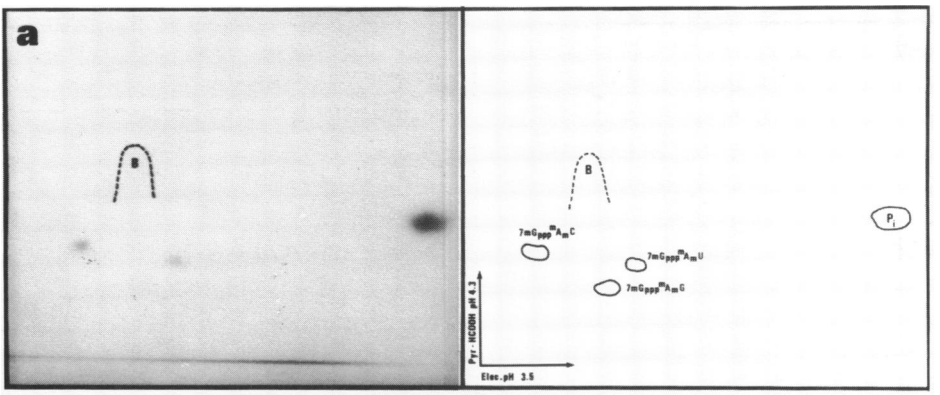
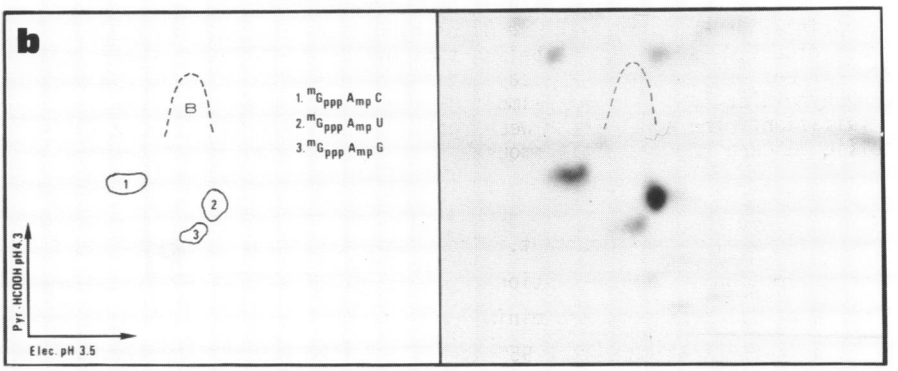
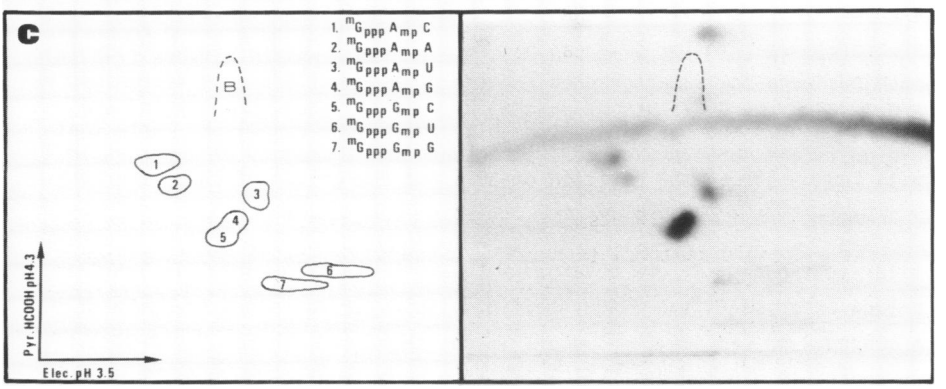


Fig. 4. Characterization of cap cores obtained from noninfected CV1 cells labeled with β-32P-GTP or β-32P-ATP. The total RNA from noninfected CV1 cells was digested with nuclease P1 and fingerprinted (29). The labeling times were reduced to 30 min and 10 min for β-32P-GTP (2.5 μM; 460 Ci/mmol) (a) and β-32P-ATP (2.5 μM; 400 Ci/mmol)(b), respectively. Since the analysis was carried out on the total RNA, the GTP, GDP, ATP and ADP 5'-terminal ends are probably from RNA polymerase I and RNA polymerase III products. The cap cores were eluted, digested with tobacco acid pyrophosphatase and processed as described in the legend to Fig. 3 (c). Optical density markers (7-methyl-pG and pG) were cochromatographed and the position of 32Pi was determined by chromatography of a mixture of γ-32P-ATP, which stays at the origin, and 32Pi. B denotes the position of the blue dye xylene cyanol FF. Note that the digests are not overdigested, since very small amounts of nondigested cap is found between the position of 32Pi and the origin of chromatography.

similar to the cap pattern obtained with α-^{32}P-GTP (Fig. 5b). In long labeling experiments the SV40-specific cytoplasmic RNA shows predominantly the ^{7m}Gppp^{m}A$_m$pU cap (29). However, in our short (60 min for α-^{32}P-GTP) or very short (10 min for β-^{32}P-ATP) labeling experiments, the ^{7m}GpppA$_m$pC cap


a
B
B
$^{7m}G_{ppp}{}^{m}A_{m}C$
$^{7m}G_{ppp}{}^{m}A_{m}U$
$^{7m}G_{ppp}{}^{m}A_{m}G$
P_i
Pyr. HCOOH pH 4.3
Elec. pH 3.5


b
B
1. $^{m}G_{ppp}A_{mp}C$
2. $^{m}G_{ppp}A_{mp}U$
3. $^{m}G_{ppp}A_{mp}G$
1
2
3
Pyr. HCOOH pH 4.3
Elec. pH 3.5


c
B
1. $^{m}G_{ppp}A_{mp}C$
2. $^{m}G_{ppp}A_{mp}A$
3. $^{m}G_{ppp}A_{mp}U$
4. $^{m}G_{ppp}A_{mp}G$
5. $^{m}G_{ppp}G_{mp}C$
6. $^{m}G_{ppp}G_{mp}U$
7. $^{m}G_{ppp}G_{mp}G$
1
2
3
4
5
6
7
Pyr. HCOOH pH 4.3
Elec. pH 3.5

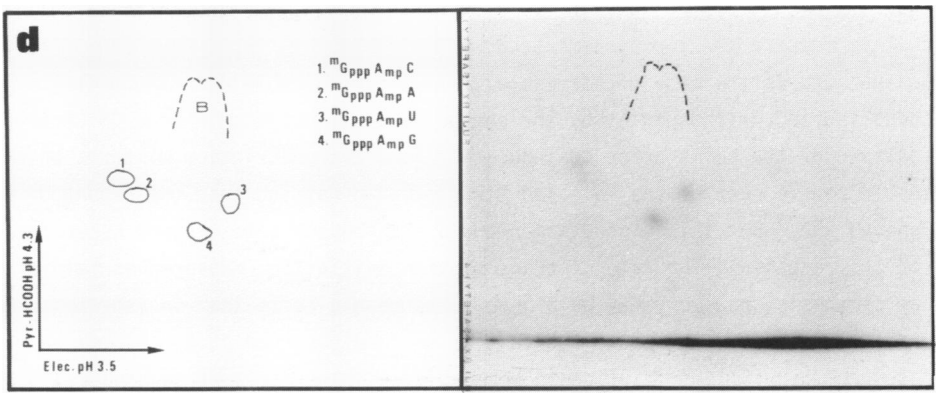


Fig. 5. Ribonuclease T2 caps from SV40-infected (a,b) and noninfected (c,d) CV1 cells labeled with α-32P-GTP (b,c) or with β-32P-ATP (a,d). The isolated RNA was digested with ribonuclease T2 and bacterial alkaline phosphatase and fingerprinted on PEI plates (c and d) or hybridized against SV40 DNA and then further digested (a and b). Labeling time with α-32P-GTP was 60 min and with β-32P-ATP was only 10 min. The cap pattern of the SV40-specific RNA after a second hybridization round was almost exactly the same as after one round of hybridization. The three lower weak spots in the α-32P-GTP labeled SV40-specific pattern (b) are small amounts of 7-unmethylated cap structures.

becomes relatively more important. Presumably the latter situation is a better reflection of the initiation ratio, since effects of RNA stability and preferential transport are minimized.

It is known that more than 95% of SV40 mRNA molecules have A-type caps. Since the caps are labeled with β-^{32}P-ATP and we have shown above that there is virtually no exchange of label, it follows that these caps must correspond to initial transcripts (it may also be noted that in this same experiment the G-type caps of the cellular mRNA were not labeled). Hence we conclude that the heterogeneity in the SV40 cap pattern is produced by a series of independent RNA initiation events.

DISCUSSION

Transcription experiments in permeable cells would be expected to be efficient and yield faithful transcripts because the cells retain a high proportion (75%)(27) of their protein content and should therefore presumably have at their disposal all factors necessary for these processes. The amount of RNA synthesized by permeable cells, however, is not significantly higher

than the amount synthesized in isolated nuclei (36,37), although it is difficult to compare the two systems because different cell types have been used and because in the case of permeable cells a considerable dilution of the added labeled triphosphates by the endogeneous XTP pool could occur. Centrifugation of the cells prior to incubation in the transcription mixture, which should remove part of the released XTP pool, did not increase the amount of labeled RNA. But this manipulation could also remove some factors necessary for the reaction. The rate of transcription initiation, measured as rate of cap synthesis, however, may be higher in permeable cells than in isolated nuclei. Indeed, compared to the initiation rate observed at the Ad2 major late promoter (37), initiation in permeable cells is at least twenty-fold higher. Part of the explanation for the observed difference can be found in the large length (25 kilobases) of the Ad2 late transcription unit and part in a deficient methylation system in isolated nuclei. Indeed, when we used the SAM concentrations which were used by Manley et al. (37), often low levels of methylation of the cap structures were observed.

The faithfulness of the transcription system is supported by three arguments: first, the relative amount of SV40-specific RNA (2.8%) is comparable to values obtained from labeling experiments *in vivo* (38); second, the cap pattern obtained from wild-type SV40 RNA is similar to the cap pattern obtained *in vivo* after uniform labeling of infected cells (9,28); third, the different cap pattern obtained with dl-1811, an SV40 deletion mutant which lacks the major late wild-type capping site, again is similar to the *in vivo* pattern (data not shown)(10). It has to be mentioned, however, that methylation of the blocked termini never goes to completion, which would influence the ribonuclease T2/BAP cap pattern to a small extent if the methylation rate is different for different cap structures.

The experiments in which cap synthesis was monitored by incubation with α-^{32}P-GTP (and α-^{32}P-ATP) allow the following conclusions. After labeling with α-^{32}P-GTP, equal amounts of radioactivity were found in the α and γ positions of the GpppG cap structures. This result suggests that the guanylylation reaction occurs on newly synthesized RNA and not on existing pools of previously synthesized RNA. This is in agreement with the facts that nucleoside diphosphates or triphosphates are nearly undetectable in hydrolysates of polymerase II transcripts and that the cap synthesis *in vivo* is a very early event in the process of RNA synthesis (2). Furthermore, these labeling experiments showed that the methylation reaction was completely inhibited by S-adenosylhomocysteine and that the synthesis of the GpppA cap was drastically

and specifically reduced when adenosine triphosphate was replaced by pCH_2ppA. This observation suggests the possibility to isolate noncapped RNA polymerase II primary transcripts. Thus, although capping occurs very early in transcription, it is probably not an obligatory event (39), as observed in one viral transcription system (3).

A direct relationship between cap formation and initiation of transcription was shown by synthesizing RNA in the presence of β-^{32}P-ATP or β-^{32}P-GTP. Analysis of the cap structures reveals that the β phosphate in the cap is derived from the 5'-terminal polyphosphate of the RNA. Indeed, in the presence of β-^{32}P-ATP, only GpppA cap cores were labeled, and in the presence of β-^{32}P-GTP, only GpppG cap cores were labeled. Similar results have been obtained with less physiological systems like HeLa nuclear homogenates (40), a capping system from rat liver nuclei (18), and the viral capping system from vaccinia virus (41). Our results do not exclude entirely an alternative pathway for cap synthesis in which processed RNA intermediates are phosphorylated by a putative polyribonucleotide kinase and further guanylylated (11), and this mechanism may look a priori very attractive to explain the heterogeneity in the SV40 late and early cap patterns. Indeed, if the major cap site corresponded to the initiation site of transcription, then the minor cap sites could possibly originate by processing. If this were the case the SV40-specific RNA would show only one β-^{32}P-ATP labeled cap structure. However, the labeling experiments with the wild-type virus as well as with the mutant dl-1811 show heterogeneous cap patterns which are similar to the in vivo cap patterns. Thus it follows that the heterogeneity has to be explained by a series of different initiation events.

A repeated consensus sequence in the SV40 late promoter region

Since it was clear from our results with SV40 late RNA that different initiation events occur over a DNA region of more than 200 nucleotides (8,10), the regions preceding the more principal cap sites (positions 0.695, 0.709 and 0.720; ref. 8) were compared for nucleotide sequence homology. As shown in Fig. 6, these three presumed late promoter regions in SV40 DNA have sequence homology of 73.3% (without any special lining up) between positions -60 and -74 from the cap site. In addition, there is 60% homology between the early and the late promoter sequences and 53.3% homology between the late SV40 sequences and the sea urchin H2A histone gene promoter (42) when small gaps are introduced. This homology includes the consensus sequence between position -70 and -80 which has been found for several different promoter regions (43) and which seems to modulate the rate of transcription (44). Several

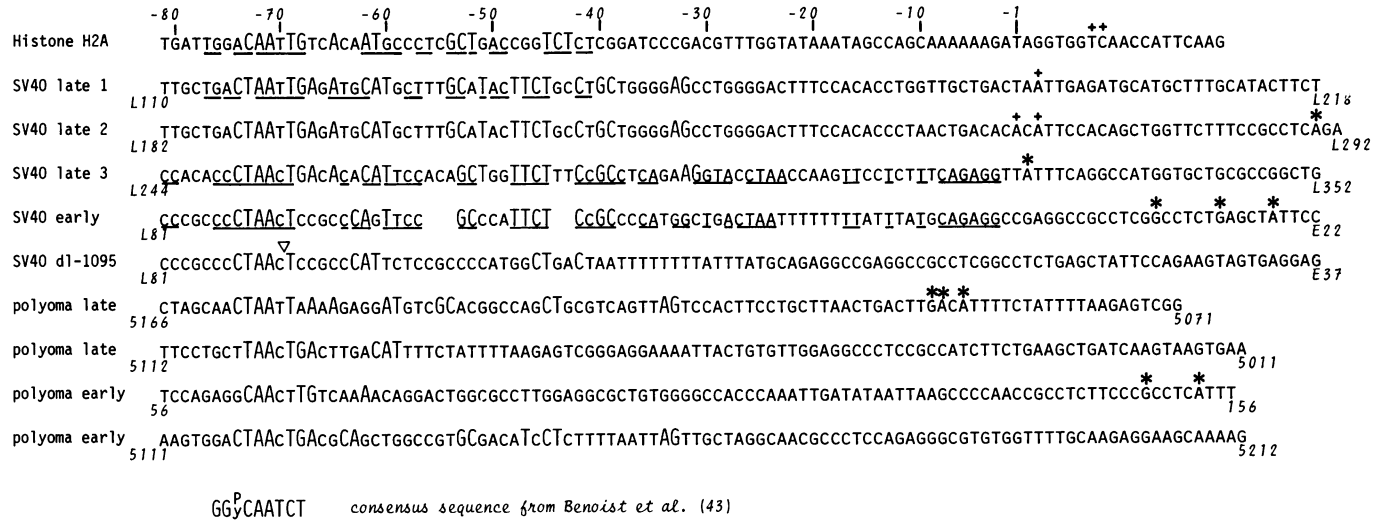


Fig. 6. Nucleotide sequence homology between the cap-prelude regions of SV40 and polyoma virus late and early genes and the sea urchin H2A histone gene. The SV40 sequences are taken from Fiers et al. (54) and Van Heuverswyn and Fiers (55); the polyoma sequences are from Soeda et al. (56); the H2A sequence is from Busslinger et al. (42). The sequence of the early SV40 deletion mutant 1095 is from Di Maio and Nathans (45); the position of the deletion is indicated by a delta symbol. The number of the first and last nucleotide of each SV40 and polyoma sequence is given below the sequence; nucleotides are also numbered (top) in the minus (5') direction from the major late SV40 cap site (late promoter 3). Well characterized cap sites in SV40 (8,9,57) and in polyoma (58, and R. Kamen personal communication) are marked with an asterisk; less precisely mapped cap sites (10,42,59) are marked with a plus sign. The main homology is indicated in large type; it compares with the consensus sequence around the corresponding position in other eukaryotic prelude regions (43). It should be noted that possibly other promoter-like sequences may be present in the control regions both in SV40 and in polyoma. A second type of homology, between the third late promoter and the first early promoter of SV40, and between the H2A promoter and the first late SV40 promoter, is indicated by underlining. Note that the H2A and the first early polyoma consensus sequence are nearly identical, CAATTGTCA and CAACTTGTCA, respectively.

characterized SV40 mutants have deletions between the consensus sequence and the Hogness-Goldberg box (45); only the mutant dl-1095 lacks sequences immediately to the right of the consensus sequence, and it is remarkable that the deletion is made in such a way that the original nucleotide sequence in this area is restored. When the sequences preceding the known and presumed cap sites in polyoma virus are compared (Fig. 6), a consensus sequence similar to that in SV40 is found around position -70. As is the case with the SV40 late promoters, the polyoma late promoters overlap over a short distance. A more striking overlap, however, is found between the presumed second early promoter of polyoma and the first late promoter. If the second early promoter is functional, it would mean that the region of polyoma DNA between about 5111 and 5166 is able to promote RNA synthesis in two directions, the early promoter being activated only late in infection (after the onset of DNA replication).

It is thus very attractive to explain the 5'-terminal heterogeneity in late SV40 mRNAs by the presence of several promoter signals. This model, however, would account for only part of the heterogeneity observed. The remaining heterogeneity most likely can be attributed to the absence of a Hogness-Goldberg box (36,44,46). Indeed, it has been pointed out that the absence of a Hogness-Goldberg box results in heterogeneous initiation around one particular site (36). Furthermore, it is possible that late SV40 transcription is complicated by being linked in some way to DNA replication (38, 47) and that the large-T antigen, which binds to a region of SV40 DNA in the vicinity of the origin of DNA replication, is directly or indirectly a possible positive effector of late transcription (48,49). Prokaryotic examples in which binding of proteins close to or within a promoter region exerts a positive effect on the rate of transcription include the lac promoter system (50) and the bacteriophage λ P_{prm}/P_R control region (51). The SV40 early promoter (Fig. 6) partially overlaps the T2 and T3 binding sites of T-antigen, as determined by Tjian (52)(the T1 binding site occurs further in the early direction). However, Shalloway et al. (53), by binding T-antigen to a restriction fragment containing the origin region and testing for resistance to exonuclease III digestion, located the border of the T-antigen binding site on the late side of the origin at position L118, which overlaps the beginning of the SV40 late promoter 1 (Fig. 6). It is perhaps worth mentioning that the third late promoter (corresponding to the major late SV40 cap site) shows sequence homology of ~59% (small gap formation was allowed) with the early promoter region over a stretch of 80 nucleotides. The potential significance of this homology with respect to late transcription has yet to

be determined.

ACKNOWLEDGEMENTS

We are most grateful to José Van der Heyden for his expert help with culturing of cells and viruses. We thank Dirk Gheysen for a gift of SV40 DNA and for his stimulating discussions during the course of this work. William Kuziel is acknowledged for his help with the manuscript. R.C. is Bevoegdverklaard Navorser of the Belgian Nationaal Fonds voor Wetenschappelijk Onderzoek. The research was supported by the Kankerfonds of the Algemene Spaar- en Lijfrentekas and by the Geconcerteerde Akties of the Belgian Ministry of Science.

REFERENCES

1. Shatkin, A. J. (1976) Cell, 9, 645-653.
2. Salditt-Georgieff, M., Harpold, M., Chen-Kiang, S. and Darnell, J. E. (1980) Cell, 19, 69-78.
3. Furuichi, Y. (1978) Proc. Nat. Acad. Sci. U.S.A. 75, 1086-1090.
4. Goldberg, S., Weber, J. and Darnell, J. E. (1977) Cell, 10, 617-621.
5. Bachenheimer, S. and Darnell, J. E. (1975) Proc. Nat. Acad. Sci. U.S.A. 72, 4445-4449.
6. Ziff, E. B. and Evans, R. M. (1978) Cell, 15, 1463-1475.
7. Berget, S. M., Moore, C. and Sharp, P. A. (1977) Proc. Nat. Acad. Sci. U.S.A. 74, 3171-3175.
8. Ghosh, P. K., Reddy, V. B., Swinscoe, J., Lebowitz, P. and Weissman, S. M. (1978) J. Mol. Biol. 126, 813-846.
9. Haegeman, G. and Fiers, W. (1978) Nucl. Acids Res. 5, 2359-2371.
10. Haegeman, G., Van Heuverswyn, H., Gheysen, D. and Fiers, W. (1979) J. Virol. 31, 484-493.
11. Spencer, E., Loring, D., Hurwitz, J. and Monroy, G. (1978) Proc. Nat. Acad. Sci. U.S.A. 75, 4793-4797.
12. Moss, B., Gershowitz, A., Wei, C. and Boone, R. (1976) Virology, 72, 341-351.
13. Moss, B., McAir, A., Ensinger, M. J., Boone, R. F. and Wei, C. M. (1976) Prog. Nucl. Acid Res. Mol. Biol. 19, 63-81.
14. Furuichi, Y. and Shatkin, A. J. (1976) Proc. Nat. Acad. Sci. U.S.A. 73, 3448-3452.
15. Furuichi, Y. and Shatkin, A. J. (1977) Virology, 77, 566-578.
16. Groner, Y. and Hurwitz, J. (1975) Proc. Nat. Acad. Sci. U.S.A. 72, 2930-2934.
17. Wei, C. M. and Moss, B. (1977) Proc. Nat. Acad. Sci. U.S.A. 74, 3758-3761.
18. Mizumoto, K. and Lipman, F. (1979) Proc. Nat. Acad. Sci. U.S.A. 76, 4961-4965.
19. Winicov, I. and Perry, R. P. (1976) Biochemistry, 15, 5039-5046.
20. Abraham, G., Rhodes, D. P. and Banerjee, A. K. (1975) Nature, 255, 37-40.
21. Schibler, U. and Perry, R. P. (1976) Cell, 9, 121-130.
22. Smith, M. M., Reeve, A. E. and Huang, R. C. (1978) Cell, 15, 615-626.
23. Mory, Y. Y. and Gefter, M. L. (1978) Nucl. Acids Res. 5, 3899-3912.
24. Udvardy, A. and Seifart, K. H. (1976) Eur. J. Biochem. 62, 353-363.

25. Gilboa, E., Soreq, H. and Aviv, H. (1977) Eur. J. Biochem. 77, 393-400.
26. Patton, J. T., Stout, E. R. and Bates, R. C. (1979) J. Virol. 30, 917-922.
27. Miller, M. R., Castellot, J. J., Jr. and Pardee, A. B. (1978). Biochemistry, 17, 1073-1080.
28. Canaani, D., Kahana, C., Mukamel, A. and Groner, Y. (1979) Proc. Nat. Acad. Sci. U.S.A. 76, 3078-3082.
29. Haegeman, G. and Fiers, W. (1978) J. Virol. 25, 824-830.
30. Raskas, H. J. and Green, M. (1971) in Methods in Virology, vol. 5, K. Maramorosch and H. Koprowski, eds. pp. 247-269 (New York: Academic Press).
31. Alwine, J. C., Kemp, D. J. and Stark, G. R. (1977) Proc. Nat. Acad. Sci. U.S.A. 74, 5350-5454.
32. Roeder, R. G. (1976) in RNA Polymerase, R. Losick and M. Chamberlin, eds. (New York: Cold Spring Harbor Laboratory).
33. Darnell, J. E. (1979) Prog. Nucl. Acid Res. Mol. Biol. 22, 327-353.
34. Vennström, B., Pettersson, U. and Philipson, L. (1978) Nucl. Acids Res. 5, 205-219.
35. Haegeman, G. and Fiers, W. (1980) J. Virol. 35, 955-961.
36. Baker, C. C. and Ziff, E. B. (1980) Cold Spring Harbor Symp. Quant. Biol. 44, in press.
37. Manley, J. L., Sharp, P. A. and Gefter, M. L. (1979) Proc. Nat. Acad. Sci. U.S.A. 76, 160-164.
38. Acheson, N. H. (1976) Cell, 8, 1-12.
39. Weil, P. A., Luse, D. S., Segall, J. and Roeder, R. G. (1979) Cell, 18, 469-484.
40. Groner, Y., Gilboa, E. and Aviv, H. (1978) Biochemistry, 17, 977-982.
41. Monroy, G., Spencer, E. and Hurwitz, J. (1978) J. Biol. Chem. 253, 4490-4498.
42. Busslinger, M., Portmann, R., Irminger, J. C. and Birnstiel, M. L. (1980) Nucl. Acids Res. 8, 957-977.
43. Benoist, C., O'Hare, K., Breatnach, R. and Chambon, P. (1980) Nucl. Acids Res. 8, 127-142.
44. Grosschedl, R. and Birnstiel, M. L. (1980) Proc. Nat. Acad. Sci.U.S.A. 77, 1432-1436.
45. Di Maio, D. and Nathans, D. (1980) J. Mol. Biol. in press.
46. Flavell, R. A. (1980) Nature, 285, 356-357.
47. Tooze, J. ed. (1973) The Molecular Biology of Tumor Viruses (New York: Cold Spring Harbor Laboratory).
48. Ferdinant, F.-J., Brown, M. and Khoury, G. (1977) Proc. Nat. Acad. Sci. U.S.A. 74, 5443-5447.
49. Parker, B. A. and Stark, G. R. (1979) J. Virol. 31, 360-369.
50. Dickson, R. C., Abelson, J., Barnes, W. M. and Reznikoff, W. S. (1975) Science, 187, 27-35.
51. Ptashne, M., Jeffrey, A., Johnson, A. D., Maurer, R., Meyer, B. J., Pabo, C. O., Roberts, T. M. and Sauer, R. T. (1980) Cell, 19, 1-11.
52. Tjian, R. (1978) Cell, 13, 165-179.
53. Shalloway, D., Kleinberger, T. and Livingston, D. H. (1980) Cell, 20, 411-422.
54. Fiers, W., Contreras, R., Haegeman, G., Rogiers, R., Van de Voorde, A., Van Heuverswyn, H., Van Herreweghe, J., Volckaert, G. and Ysebaert, M. (1978) Nature, 273, 113-120.
55. Van Heuverswyn, H. and Fiers, W. (1979) Eur. J. Biochem. 100, 51-60.
56. Soeda, E., Arrand, J. R., Smolar, N., Walsh, J. E. and Griffin, B. E. (1980) Nature, 283, 445-453.
57. Thompson, J. A., Radonovich, M. F. and Salzman, N. P. (1979) J. Virol. 31, 437-446.

58. Flavell, A. J., Cowie, A., Arrand, J. R. and Kamen, R. (1980) J. Virol. 33, 902-908.
59. Grosschedl, R. and Birnstiel, M. L. (1980) Proc. Nat. Acad. Sci. in press.
60. Volckaert, G. and Fiers, W. (1977) Anal. Biochem. 83, 228-239.