Specific in vitro initiation of transcription of simian virus 40 early and late genes occurs at the various cap nucleotides including cytidine

([β-32P]NTP/cap sites/staggered transcriptional initiation/RNA polymerase II/initiator CTP)

DAVID GIDONI, CHAIM KAHANA, DAN CANAAN, AND YORAM GRONER

Department of Virology, Weizmann Institute of Science, Rehovot, Israel

Communicated by Michael Sela, December 29, 1980

ABSTRACT High specific activity [β-32P]ATP and [β-32P]CTP were used to study in vitro transcriptional initiation and subsequent capping of simian virus 40 (SV40) early and late RNAs. More than 40% of the capped SV40 RNA synthesized in vitro was also polyadenylated. With [β-32P]ATP, only adenosine-containing caps were labeled and the incorporated radioactive phosphate was found exclusively in the β position. Cap digestion patterns showed extensive qualitative and quantitative similarities between these 32P-labeled caps and caps labeled in vivo [Canaani, D., Kahana, C., Mukamel, A., & Groner, Y. (1979) Proc. Natl. Acad. Sci. USA 76, 3078–3082]. With [β-32P]CTP, only early SV40 RNA was labeled, consistent with the absence of cytosine-containing caps in late transcripts. The [β-32P]CTP-labeled cap was identified as m'GpppCmpU, which was previously identified as the major cap of in vivo-labeled early SV40 mRNA (Kahana, C., Gidoni, D., Canaani, D., & Groner, Y. (1981) J. Virol. 37, 7–16). This experiment provides biochemical evidence for eukaryotic RNA polymerase II initiation of transcription with CTP. The data imply that, on SV40 DNA, RNA polymerase II initiates transcription at multiple nucleotide sequences and capping occurs at the initiator nucleotide.

Eukaryotic cellular and viral mRNA contains a cap structure that is unique to the 5′-terminal of the mRNA molecule (1). This cap therefore can serve as a convenient and unambiguous marker for identification of the 5′-terminal sequences of mRNAs. In vitro transcription experiments and analyses of in vivo-labeled RNA (2–11), as well as direct biochemical approaches (12–14), have indicated that, very early after initiation, caps may be attached to the triphosphate termini resulting from transcriptional initiation events. Unexpectedly, it was found recently that some virus-specific mRNAs—such as late polyoma (5, 15), early and late simian virus 40 (SV40) (6, 16–18, §), and early adenovirus type 2 (19–21)—contain multiple 5′-terminal caps. The biological significance of this 5′-terminal heterogeneity of mRNA is still not known. The multiple caps could evolve via two alternative mechanisms: (i) transcriptional initiation at multiple sites (i.e., each capped 5′ end represents an initially transcribed sequence to which m'PpG was attached), and (ii) processing of primary transcripts (i.e., transcription began at one point and the multiple caps were generated by cleavage of the pre-mRNA at a number of sites that were capped subsequently).

To distinguish between these two possibilities, we carried out in vitro transcription experiments with [β-32P]ATP and [β-32P]CTP as radioactive precursors. This approach was previously used by us to demonstrate that, in isolated nuclei, capping is coupled to initiation of transcription, and RNA polymerase.

II primary transcripts are the substrate for the nuclear capping enzyme (13). Nuclear caps are formed by condensation of the α phosphate of GTP with the 5′-terminal β-phosphate of RNA (12); thus, the β-32P originating from initiated termini will be included in the caps.

A procedure for the synthesis of carrier-free β-32P-labeled nucleoside triphosphates has been developed (22), and we have used such high specific activity [β-32P]NTPs to label SV40 caps in vitro. We found that the multiple caps of SV40 mRNAs are generated by the first mechanism—namely, on SV40 DNA, RNA polymerase II initiates transcription at multiple nucleotide sequences including cytidine. These results have been reported in brief.6

MATERIALS AND METHODS

In Vitro Labeling and Isolation of SV40 RNA. Nuclei were prepared from SV40-infected BSC-1 cells 30–36 hr after infection by using methods described for the isolation of nuclei from HeLa cells (2, 13). Reaction mixtures (0.5–1 ml) contained 30 mM Hepes-KOH (pH 7.9), 150 mM KCl, 6 mM MgCl2, 5 mM dithiothreitol, 0.1 mM EDTA, 0.5 mM KPO4, 5 mM NaF, 10% (vol/vol) glycerol, 5 mM S-adenosylmethionine, unlabeled nucleoside triphosphates at 2 mM each, 50–100 μM [β-32P]ATP or [β-32P]CTP (200 Ci/mmol; 1 Ci = 3.7 × 1010 becquerels), and 108 SV40-infected BSC-1 nuclei per ml. Incubation was at 30°C for 60–90 min. Reactions were stopped and RNA was isolated as described (2, 13).

Cells were made permeable according to Miller et al. (23). SV40-infected BSC-1 cells in monolayer cultures were treated with lysolecithin (250 μg/ml) (Sigma) at 4°C for 1 min. The lysolecithin was removed, and transcription mixture was added. The permeability of the cells was monitored in one of the cultures and was always >90% as deduced from the uptake of trypan blue (23). The incubation mixture was essentially as described by Samal et al. (24) and contained 150 mM sucrose, 35 mM Hepes-KOH (pH 7.4), 80 mM KCl, 10 mM MgCl2, 5 mM dithiothreitol, 0.5 mM KPO4, 0.5 mM CaCl2, 10 mM creatine phosphate, 6 mM NaF, 0.5 mM S-adenosylmethionine, tRNA at 100 μg/ml, 0.3 mM unlabeled nucleoside triphosphates and 50 μM [β-32P]ATP or [β-32P]CTP (2000–3000 Ci/mmol).

Both the efficiency and the fidelity of transcriptional initiation

Abbreviations: SV40, simian virus 40; A-cap, adenosine-containing cap; II, caps containing two 2′-O-methylated nucleotides; C-cap, cap containing methylated cytosine at the penultimate position.

* To whom reprint requests should be addressed.


The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "advertisement" in accordance with 18 U. S. C. §1734 solely to indicate this fact.
RESULTS AND DISCUSSION

Initiation of Late SV40 Transcripts in Vitro; Incorporation of [\(^{32}\)P]ATP into Late SV40 Caps. It has been shown (6) that adenosine is the major penultimate nucleotide among the multiple late SV40 caps. Therefore, [\(^{32}\)P]ATP was selected as the radioactive precursor in the in vitro reaction. RNA was synthesized by using SV40-infected isolated nuclei or permeable cells. Labeled RNA was purified and SV40-specific transcripts were selected by hybridization to SV40 DNA bound to Sepharose. When SV40-infected permeable cells were used, >40% of SV40-specific radioactive activity was in mRNA molecules that were polyadenylated as determined by binding to oligo(dT)-cellulose in 0.125 m M NaCl. This result emphasizes the fidelity and efficiency of the permeable-cell system and thus its usefulness for in vitro studies of gene expression (23, 24).

\(^{32}\)P-Labeled SV40 RNA was treated with alkaline phosphatase to remove uncapped terminal phosphates and analyzed by hybridization to immobilized SV40 DNA fragments obtained by digestion with \(Bgl\ I\), \(Hpa\ I\), and \(Hae\ II\) (Fig. 1). SV40 RNA synthesized under similar conditions but in the presence of \([\alpha^{32}\)P]UTP was used to demonstrate that both early and late genes were transcribed; this RNA gave proportional hybridization to fragments representing the entire early and late regions. When \(^{32}\)P-labeled SV40 RNA was hybridized to the same DNA fragments, radioactivity was detected only in fragment D (0.67–0.76 map units), a region where the multiple late caps have been mapped (6, 18).

\(^{32}\)P-ATP-Labeled SV40 RNA was also hybridized to blots containing separated strands of SV40 DNA fragments obtained by digestion with \(Bgl\ I\) and \(Hpa\ I\) (Fig. 1). All the radioactive activity was found in the L strand of fragment D (0.67–0.76 map units), verifying the late polarity of the labeled transcript (not shown).

To isolate and identify the alkaline phosphatase-resistant \(^{32}\)P radioactivity that hybridized to fragment D, SV40 RNA from a parallel experiment was extensively digested with RNase T2 (which hydrolyzes every phosphodiester bond in RNA except those containing 2'-O-methylated residues and thus leaves only the cap intact). After treatment with alkaline phosphatase, cap structures were fractionated by two-dimensional electrophoresis/chromatography. The patterns obtained from the \(^{32}\)P-ATP-labeled SV40 RNA made in vitro (Fig. 2B) were compared with those of late SV40 mRNA labeled in vivo (Fig. 2A) (6). The similarities between the two patterns will become apparent from the analysis of the \(^{32}\)P-ATP-labeled oligonucleotides described below. There were two exceptions: first, only the adenosine-containing cap species (A-caps) were labeled in vitro with \(^{32}\)P-ATP; second, the proportion of cap structures containing two 2'-O-methylated nucleotides (cap II) was higher in the in vivo-labeled RNA.

The identity of the in vitro caps was confirmed by chromatography and secondary digestion with nuclease P1, (Fig. 3; Table I). The relative abundance of the in vitro synthesized caps resembled the in vivo situation: the two major in vivo cap I structures, m\(\text{GgcppAmpC}\) (16%) and m\(\text{GgcppAmpG}\) (47%), were the most prominent caps labeled in vitro (spots 2 and 4); the in vivo less abundant m\(\text{GgcppAmpG}\) (6%) was also minor in vitro (spot 5). The only cap II structure detected in vitro was m\(\text{GgcppAmpUmp}\) (spot 6), which was the most abundant in vitro cap II. Two cap core structures, not detected in vivo, were found in vitro: m\(\text{GgcppA}\) (spot 1) and its m\(\text{A}\)-containing derivative m\(\text{GgcppAmp}\) (spot 3). This and the decreased amount of cap II structures probably reflect lower methylating activity of the in vitro system which may result from the relative instability of the methyl donor S-adenosylmethionine. Spots 7 and 8 were identified as open ring derivatives of A-cap species and probably were generated during isolation (Fig. 3). Note that the opening of the imidazole ring of m\(\text{G}\) results in loss of a positive charge and thus a faster migration in paper electrophoresis at pH 3.5 (Fig. 3A).

In order to identify the penultimate nucleotide in the caps and to establish unequivocally the position of the \(^{32}\)P in the labeled caps, the major spots were further analyzed (Fig. 4). After desalting, the RNase T2-resistant cap m\(\text{GgcppAmpU}\) (spot 4 from Fig. 2B) was treated with nuclease P1 (which cleaves phosphodiester bonds in RNA including those containing 2'-O-methylated residues) and analyzed by paper electrophoresis (Fig. 4A) and by thin-layer chromatography with a solvent sys-
tem that separates the three cap species (Fig. 3B). All radioactivity moved as one spot comigrating with m'GpppAm marker. This material was eluted and digested partially or completely with nucleotide pyrophosphatase (which cleaves the pyrophosphate bridge in cap structures). Fig. 4, lane B, shows the partial cleavage yielded m'GDP, ADP, and P1; upon further incubation, all the radioactivity was converted to P1 (lane C), without the appearance of either m'G or AMP. These results imply that the ^32P was indeed at the β position within an A-cap structure—m'Gp[32P]pAm—which proves that the multiple SV40 late caps result from transcriptional initiation events: each of the different SV40 caps represents 5'-triphosphate terminus (pppA-β) of initiated RNA molecule which was capped by condensation of pG residue to form GpppA-β.

The results of these experiments are consistent with our previous observation that, in isolated nuclei, capping is coupled to initiation of transcription, and RNA polymerase II primary transcripts are the substrate for the nuclear capping enzyme (13). The similarity in the relative abundance between the various caps made in vitro to those made in vitro signifies that SV40 late transcription was faithfully initiated in the in vitro system.

**Initiation of Early SV40 Transcription in Vitro; Incorporation of [β-^32P]CTP into Early SV40 Caps.** In contrast to late SV40 caps that contain 95% A-caps, the early mRNA contains a high proportion of C-caps (i.e., contain methylated cytosine at the penultimate position) (17, 18). In fact, m'GpppCmpU is the most abundant early SV40 cap. Therefore, [β-^32P]CTP was used as the radioactive precursor to study the initiation and subsequent capping of early SV40 transcripts in vitro. Fig. 5 shows hybridization of [α-^32P]UTP-labeled (lane A) and [β-^32P]CTP-labeled (lane B) alkaline phosphatase-treated SV40 RNA to blots containing five DNA fragments obtained by digestion of SV40 DNA with EcoRI, Bgl II, Bam HI, Hpa I, and Hae II. β-^32P-labeled RNA hybridized only to the early fragment A (0.37–0.67 map units), although both the early and late regions were transcribed (Fig. 5, lane A). These results are consistent with the absence of C-caps in late SV40 RNA (6) and with the mapping data that place the 5' termini of early mRNA at coordinate 0.67 (26–28).

The digestion pattern of the [β-^32P]CTP-labeled caps was compared with that of in vitro-labeled early caps (Fig. 6). Five spots were labeled in vitro with [β-^32P]CTP, two prominent ones (1 and 2) and three minor ones. They were further analyzed by chromatography on polyethyleneimine plates and, after digestion with nuclease P1, by thin-layer chromatography (Fig. 3; Table 1). Spot 2 was identified as the early cap m'GpppCmpU and spot 1, as its cap core derivative, m'GpppC. Spot 3 migrated like cap II structure of material in spot 2, and spots 4 and 5 migrated like open-ring derivatives of spots 1 and 2 but, due to insufficient radioactivity, they were not analyzed further. We concluded from these experiments that, akin to late transcrip-

---

**Table 1. Analysis of RNase T2-resistant caps derived from SV40 RNA labeled in vitro with [β-^32P]ATP and [β-^32P]CTP**

<table>
<thead>
<tr>
<th>Spot</th>
<th>Nuclease P1 products</th>
<th>Assigned cap</th>
<th>Relative abundance, %</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>In vitro</td>
</tr>
<tr>
<td>Fig. 2B</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>m'GpppA</td>
<td>m'GpppA</td>
<td>5</td>
</tr>
<tr>
<td>2</td>
<td>m'GpppAm</td>
<td>m'GpppAmC</td>
<td>11</td>
</tr>
<tr>
<td>3</td>
<td>m'GpppAm</td>
<td>m'GpppA</td>
<td>8</td>
</tr>
<tr>
<td>4</td>
<td>m'GpppAm</td>
<td>m'GpppAmU</td>
<td>67</td>
</tr>
<tr>
<td>5</td>
<td>m'GpppAm</td>
<td>m'GpppAmG</td>
<td>4</td>
</tr>
<tr>
<td>6</td>
<td>m'GpppAm</td>
<td>m'GpppAmUmpU</td>
<td>3</td>
</tr>
<tr>
<td>7</td>
<td>mGpppAm</td>
<td>mGpppAmX</td>
<td>1</td>
</tr>
<tr>
<td>8</td>
<td>mGpppAm</td>
<td>mGpppAmY</td>
<td>1</td>
</tr>
</tbody>
</table>

| Fig. 6B |                     |              |          |          |
| 1    | mGpppC              | mGpppC      | 5        |          |
| 2    | mGpppCm             | mGpppCmpU   | 3        |          |
| 3    | mGpppCm             | mGpppCmpX   | 8        |          |
| 4    | mGpppCm             | mGpppCmpC   | 5        |          |
| 5    | mGpppCm             | mGpppCmpC   | 5        |          |

Spots were eluted with 2 M NH4HCO3 (pH 7.5) and desalted, and a portion was analyzed by chromatography on polyethyleneimine plates with appropriate markers (not shown). The rest of the sample was digested with nuclease P1 and analyzed by both paper electrophoresis and thin-layer chromatography, as in Fig. 3. m'GpppAm and mGpppAm are caps containing an open ring structure; these markers were prepared from m'GpppAm and mGpppCm by incubation with 10 M NH4OH for 60 min at 40°C. X and Y represent unidentified nucleotides.

*Data from Canaani et al. (6).
tion, the early caps represent transcriptional initiation points. This implies that RNA polymerase II initiates transcription on SV40 DNA at many different nucleotide sequences. In the late genes, only purines are used as initiators and in more than 90% of the cases RNA polymerase starts with adenosine followed by pyrimidine, reminiscent of prokaryotic transcriptional

**Fig. 3.** Analysis of nuclease P1-digested β-32P-labeled caps. Radioactive spots were eluted from the polyethyleneimine plates, desalted, digested with nuclease P1, and analyzed. (A) Paper electrophoresis at pH 3.5. Lanes 1–5 correspond to material eluted from spots 1, 2, 5, 7, and 8 in Fig. 2B. (B) Thin-layer chromatography with isobutyric acid/NH₄OH, 10:6 (vol/vol). Lanes: 1, material eluted from spot 4 (Fig. 2B); 2, same as in 1 but incubated with 1 M NH₄OH for 5 min at 37°C before it was spotted on plate; 3, material eluted from spot 2 (Fig. 2B); 4, material eluted from spot 8 (Fig. 2B). *[m⁷G]pppAm is a cap marker containing an open ring structure and was prepared from m⁷GpppAm by incubation with 10 M NH₄OH for 60 min at 40°C.

**Fig. 4.** Electrophoretic analysis of products generated by enzymatic digestion of Spot 4 in Fig. 2B. Lanes: A, nuclease P1 digestion of the RNase T2-resistant cap 4; B, partial nucleotide pyrophosphatase digestion of cap eluted from A; C, complete nucleotide pyrophosphatase digestion of cap eluted from A.

**Fig. 5.** Hybridization of in vitro synthesized 32P-labeled SV40 RNA to blots containing DNA fragments obtained by digestion of SV40 DNA with BglI, EcoRI, BamI, HpaI, and HaeII. SV40 RNA labeled in vitro in isolated nuclei with [γ-32P]UTP (lane A) or [β-32P]CTP (lane B) was treated and hybridized as described in the legend to Fig. 1.
initiation (29). As for the early region, our data provide biochemical evidence that eukaryotic RNA polymerase II initiates transcription with CTP, as was recently described for the prokaryotic enzyme (30–32). This observation may also suggest that other pyrimidine caps in cellular (1) and viral (20) mRNAs are initially transcribed sequences rather than processed capped termini.

Multiple caps were reported for late polyoma (5, 15) and early and late SV40 mRNAs (6, 16–18, 29). As initiation multiple caps in cellular and viral RNA is a feature of the transcription system, we used this as a tool to identify m7GpppAmpU RNA. We synthesized a [m7GpppAmpU] RNA as described in Fig. 2. We then measured the presence of this RNA in various cell lines.

We thank M. Revel for his encouragement and support and A. Mukamel for her excellent technical assistance. Support for this work was provided by funds from the United States-Israel Binational Foundation (Jerusalem, Israel).