Nucleic Acids Research

Characterization of the major altered leader sequence of late mRNA induced by SV40 deletion mutant dl-1811

Guy Haegeman, Dirk Iserentant, Dirk Gheysen and Walter Fiers

Laboratory of Molecular Biology, State University of Ghent, Ledeganckstraat, 35 Ghent, Belgium

ABSTRACT

dl-1811 is a viable SV40 mutant with a 40 base pair dele-
tion that includes the major wild-type capping site of late mRNA
at map position 0.72. The late viral mRNAs induced by dl-1811
have now been further characterized by inversed S1-mapping ana-
lysis. The S1-resistant, ^{32}P-labeled RNA fragments derived from
the leader region were examined by fingerprinting and by analy-
sis of RNase T2-generated 5'-terminal cap structures. The results
show that most if not all of the mutant leader fragments analyzed
have their 5' terminus to the left of the dl-1811 deletion site,
i.e., closer to the origin of DNA replication. The major altered
leader fragment is a continuous transcript from the DNA in the
region 0.716 to 0.761 map unit and its 5' terminus has been pre-
cisely mapped at nucleotide L290. The observation that the cap
structure of the major altered leader is only a minor cap species
in wild-type late RNA suggests regulation in the use of different
capping sites in SV40.

INTRODUCTION

During the late stage of SV40 infection of monkey kidney
cells, i.e., after the onset of viral DNA replication, two cyto-
plasmic late mRNA species are synthesized sedimenting at 16S and
19S (1, 2). Both species have been mapped on the genome (3, 4),
and more recently it has been found that these viral mRNA mole-
cules are composite structures comprising an untranslated "leader"
segment and a coding sequence or "body". The sequence is not co-
linear with the DNA: each mRNA has been formed by at least one
splicing event (5-8). The most abundant leader fragment is a se-
quence of 202(±1) nucleotides that is transcribed from the DNA
region between 0.72 and 0.76 map unit and which is subsequently
spliced 42(±1) nucleotides before the body of the VP1 gene (9-
12; the uncertainty is because the exact position of splicing
is ambiguous). Alternative leader fragments may also occur in

late viral RNA and these actually appear to constitute the most
abundant leader segments in the 19S species (7, 13, 14).

Procedures have been described for creating viable SV40
deletion mutants in the late region of the genome (15-18). The
deletions have been characterized by restriction mapping analysis;
in some cases the exact size and location of the deletion has
been precisely determined by nucleotide sequence analysis (18-20).
It is remarkable that in some of these mutants the major late
capping site, which maps at position 0.722 (21), is included in
the deletion (18, 20). Although in some other systems cap forma-
tion seems to be directly linked with initiation of RNA trans-
cription (22-26), deletion of the major capping site in SV40
does not result in loss of viability but at most only results in
a decreased plaque size (18, 27).

Previously we have shown that deletion of the major capping
site in the mutant d1-1811 results in the appearance of a variety
of alternative capped 5' termini in the corresponding late mRNA
(27). This suggests the presence of a number of alternative lea-
der segments. Indeed, altered 16S and 19S mRNAs have been obser-
ved after infection with mutants that lack DNA segments of diffe-
rent size and at slightly different positions compared to the
d1-1811 deletion (28). In this paper we report a further charac-
terization of the d1-1811 late mRNA both by fingerprint analysis
(9, 29) and by the S1 mapping procedure (30), followed again by
fingerprinting and by cap analysis of the S1-resistant fragments.
This enabled us to determine at the nucleotide level the major
alternative leader segment present in the d1-1811 late mRNA and
to identify the two related cap structures $^{7m}Gppp^mAmpG$ and
$^{7m}Gppp^mAmpGmpA$. These major terminal structures constitute about
40% of the total d1-1811 late mRNA cap content (27). Furthermore,
the latter terminal trinucleotide sequence A-G-A permitted us
to pinpoint unambiguously the 5' end of the alternative leader
segment.

METHODS

Isolation and purification of ^{32}P-labeled cytoplasmic late
viral mRNA from CV-1 monkey kidney cells infected by either wild-
type (WT) SV40 or by d1-1811, fingerprint analysis of the puri-

fied RNA, and isolation and structure determination of the appro-
priate 5'-terminal capped structures have been described in de-
tail previously (9, 27, 31). The Berk and Sharp technique (30)
for mapping mRNA sequences on a genome, for pinpointing the 5'
end of leader fragments and for positioning spliced-out regions
has been applied in these studies in an inversed way, i.e., by
using ^{32}P-labeled RNA and unlabeled DNA. The S1-resistant DNA·RNA
hybrids were isolated by polyacrylamide gel electrophoresis and
used for further characterization of the mRNA leader fragments
by fingerprinting techniques.

Poly(A)-containing viral RNA was selected by a single hy-
bridization step onto Sepharose-bound SV40 DNA, precipitated in
the presence of carrier yeast ribosomal RNA, and dissolved in
80% deionized formamide, 0.4 M NaCl, 0.001 M EDTA and 0.04 M
Pipes buffer, pH 6.4, along with the appropriate SV40 DNA frag-
ment. The latter was obtained by digestion on a preparative scale
of viral DNA isolated by Hirt extraction of infected cells, or
was derived from plasmids consisting of segments of SV40 DNA in-
tegrated in pBR322 (details given below). In a typical experiment
this solution (final vol. 60 µl) was heated for 15 min. at 85°C
to completely denature the DNA and was then kept at 52°C for 3 h
to allow formation of DNA·RNA hybrids. After this hybridization
step the solution was diluted in 900 µl ice-cold S1 buffer con-
taining 0.025 M NaAc, 0.25 M NaCl and 0.0045 M ZnAc, pH 4.4.
150 µl S1 nuclease (1 unit/µl) was added and the digestion was
carried out for 30 min. at 37°C. The mixture was extracted with
phenol, and the S1-resistant hybrids were precipitated, and frac-
tionated on a polyacrylamide gel run in 0.09 M Tris-borate,
0.0025 M EDTA, pH 8.4. The gel bands containing the DNA·RNA hy-
brids were visualized by autoradiography and were eluted with
1 M NaCl. Each fragment was redissolved in 5 µl 0.02 M Tris-Cl,
0.002 M EDTA, pH 7.4, heated for 10 min. at 85°C and quickly
chilled on ice to keep the hybrids denatured. 1 µl of T1 ribo-
nuclease (10 units/µl) was added and the solution was incubated
for 1 h at 37°C. The digest was subsequently analyzed by minifin-
gerprinting (29). To isolate the 5'-terminal cap structures of
S1-resistant RNA fragments, DNA·RNA hybrids were eluted and de-
natured as described above and digested with ribonuclease T2 and

bacterial alkaline phosphatase as reported previously (31).

RESULTS

Characterization of dl-1811 late mRNA by minifingerprinting

Twice-hybridized ^{32}P-labeled dl-1811 late mRNA was digested with T1 ribonuclease and a two-dimensional fingerprint was made. The dl-1811-specific T1 pattern is shown in Figure 1. The profile is homologous to the WT RNA fingerprint except for only one additional spot which occurs in about molar amount in the dl-1811
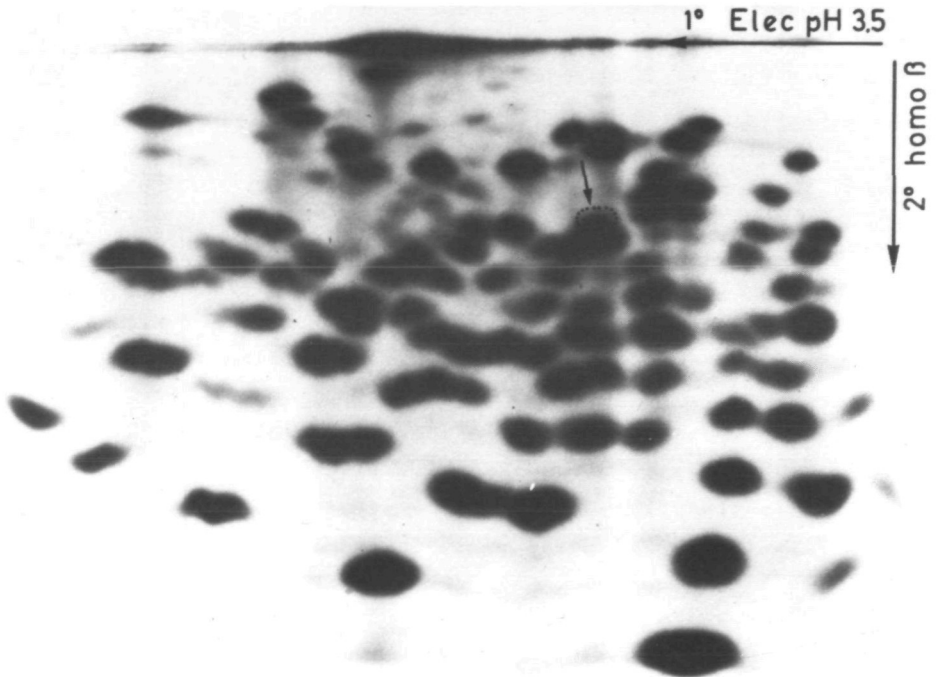
Fig. 1 : T1-ribonuclease fingerprint of dl-1811 late mRNA. Two-dimensional fractionation on PEI-cellulose thin-layer plates (20 cm x 20 cm) was carried out according to the procedure of Volckaert et al. (29) of a T1 hydrolysate of total cytoplasmic late mRNA of dl-1811. The arrow indicates the additional spot (T1 oligonucleotide 243) present in the dl-1811 pattern which represents the unique product with composition $(U_2, C_2, A-C, A-A-C)G$.

pattern. This additional product (T1 oligonucleotide 243; cf.
legend to Figure 5) was analyzed with ribonuclease A (not shown)
and its double-digestion products $(U_x, C_y, A-C, A-A-C)G$ correspond
to the sequence U-A-C-C-U-A-A-C-C-G, which is the predicted T1
oligonucleotide on the basis of the SV40 DNA nucleotide sequence
(32) and the extent of the dl-1811 deletion (27). The first Up
residue of this oligonucleotide (at L296) is part of the KpnI
restriction site at map position 0.716 while the pG residue is
derived from beyond the deletion (at L345) and corresponds to the
position next to the HpaII site at 0.726 map unit. The presence
of this oligonucleotide in about molar quantity (the resolution
was not sufficient to allow rigorous quantitation) suggests that
most if not all of the dl-1811 leader fragments extend into the
5' side of the deletion site and that the variety of cap struc-
tures associated with dl-1811 late mRNA (27) is predominantly
derived from an area of DNA preceding the position of the major
WT capping site (i.e., from a region closer to the origin of DNA
replication around 0.67 map unit). The preceding characteristic
T1 oligonucleotides, UUCUUUCCG (N° 530) and CCUCAG (N° 131) (cf.
Figure 5), were not revealed as independent spots in the T1 fin-
gerprint of total dl-1811 late mRNA, as they were only partially
resolved from related sequence isomers. However, their presence,
although in variable amounts, was revealed by pancreatic analysis
of the appropriate mixed T1-oligonucleotide spots. Other unique
T1 oligonucleotides located more upstream (i.e. towards map posi-
tion 0.67) were not readily detectable in the fingerprint of
total dl-1811 mRNA and were considered to be appreciably submolar
compared to the relative abundance of the normal WT leader oligo-
nucleotide sequences.

Isolation of the major WT leader segment by inversed S1 mapping

In order to specifically isolate the leader fragments of the
viral RNA, an inversed S1-mapping procedure was used (see Methods
section). Total poly(A)-containing RNA was first hybridized to
immobilized SV40 DNA in order to decrease the background of the
gel pattern; this step is almost essential when subsequent mapping
of the oligonucleotides of the protected RNA fragment is intended.
The SV40-specific RNA was then hybridized to the Hind III restric-

tion fragment C (0.646 to 0.860 map unit) and treated with S1 nu-
clease according to the Berk and Sharp (30) procedure. The gel
pattern obtained with WT SV40 RNA is shown in Figure 2a. One
major band (N° 5) is apparent indicating the presence of a major
late leader fragment, whose size and T1-fingerprint is in com-
plete agreement with previous results derived from oligonucleotide
mapping (9, 10) and reverse transcription (11, 12) of late viral
RNA. Slower migrating bands were also analyzed by T1 mapping and
were shown to contain RNA sequences protected by the Hind II +
III fragment D (i.e., part of the coding portion of the 19S mRNA)
and, furthermore, RNA fragments including the major leader seg-
ment in addition to sequences corresponding to the Hind-D frag-
ment. This means that besides the major WT SV40 leader sequence
common to both 16S and 19S RNA but mainly derived from the former
(9-13), some specific RNA fragments were picked up which were
derived exclusively from the 19S mRNA. Minor leader fragments
shorter than the major 202(±1) nucleotide long sequence and cor-
responding to those observed by reverse transcription (13, 14)
may have been lost in the background smear of the gel.

Isolation and characterization of the major dl-1811 leader se-
quence by inversed S1 mapping

   As the mutant dl-1811 late mRNA starts with a wide variety
of 5'-terminal cap structures (27), a great diversity of leader
fragments was expected. The major 5'-terminal leader fragments
have now been further characterized by reversed S1-mapping. To
circumvent the additional complexity of obtaining RNA segments
protected entirely or partially by coding portions of the DNA
during hybridization, the dl-1811 DNA segment comprising Hind II
+ III restriction fragments C and L (0.646 to 0.756 map unit) was
inserted into plasmid pBR322 under CIII physical containment con-
ditions. The resulting plasmid, pSCL12, was then used as a leader-
specific probe in the S1-mapping experiments with dl-1811 mRNA
(a detailed description of the construction of pSCL12 will be
published elsewhere). Figure 2b shows the polyacrylamide gel of
^{32}P-labeled dl-1811 late mRNA after hybridization to pSCL12 DNA
and subsequent S1 nuclease digestion. A number of successive
bands of nearly the same size may be seen; however, one particu-
lar species, N° 3, seems to be more abundant compared to the ad-
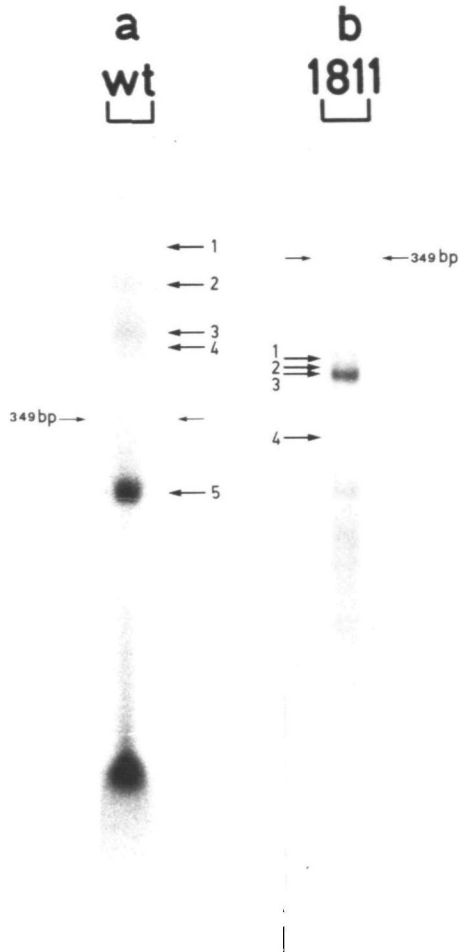
Fig. 2 : Fractionation by electrophoresis of ^{32}P-labeled S1-
resistant RNA·DNA hybrids.
(a) Separation on a 4% polyacrylamide gel of WT SV40 late mRNA
hybridized in 60 λ hybrid buffer (see Methods section) onto
3 μg of SV40 Hind-C DNA in the presence of 60 μg carrier
plasmid DNA and 30 μg carrier yeast RNA. (b) Separation on a
12% polyacrylamide gel of dl-1811 late mRNA hybridized in 20 λ
hybrid buffer onto 25 μg PstI-cleaved pSCL12 plasmid DNA
(which contains the SV40 Hind II + III fragments C + L) in the
presence of 12 μg carrier yeast RNA. Note that the S1 digestion
conditions (see Methods section) have been adapted to the quan-
tity of nucleic acid material present in the initial hybridi-
zation step. In this particular case, for example, 50 units of
S1 nuclease were used in 300 μl S1 buffer. The size marker used
is a S1-treated MS2 RNA·DNA hybrid of chain length 349 basepairs.

jacent compounds. All the bands were eluted and characterized by
T1 fingerprinting. Several of them (N° 1, 2 and 3) were also
analyzed with respect to cap content. Figure 3a and 3b shows the
fingerprint and corresponding oligonucleotide identification,
respectively, of the major band N° 3. Beginning at the 3' termi-
nus of the leader fragment and going in the 5' direction (see
Figure 5), all expected oligonucleotides, including the afore-
mentioned T1 oligonucleotide 243 which spans the deletion site,
are present on the fingerprint. However, oligonucleotide 131,
which precedes oligonucleotide 243 by only 3 nucleotides, was
definitely absent from band N° 3, as were all the remaining T1
products extending further in the 5' direction. These results
allow the localization of the 5' end of the major mutant leader
to a rather narrow area neighboring the KpnI restriction site at
position 0.716. Total degradation of another aliquot of this
leader segment RNA with T2 ribonuclease and bacterial alkaline
phosphatase (Figure 4) revealed two prominent cap structures
identified by position as $^{7m}Gppp^{m}AmpG$ and $^{7m}Gppp^{m}AmpGmpA$. This

a   **b**
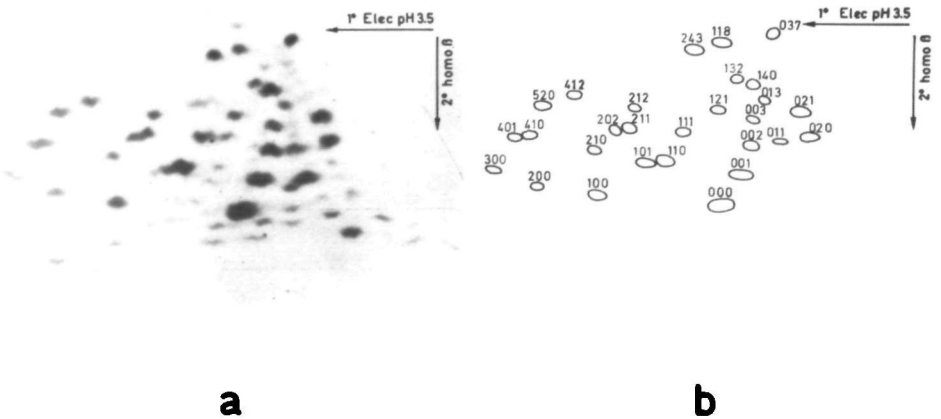
Fig. 3 : Characterization of the major leader fragment of dl-1811
          late mRNA.
(a) A portion of the S1-resistant RNA·DNA hybrid band N° 3 in
Figure 2b was digested with T1 ribonuclease (see Methods section)
and fractionated in two dimensions on PEI-cellulose thin-layer
plates (20 cm x 20 cm). (b) Schematic diagram identifying the
oligonucleotide spots in (a). A standard nomenclature of three
digits indicating the number of uridine, cytidine and adenosine
residues, respectively, is used for identifying the T1-spots (38).
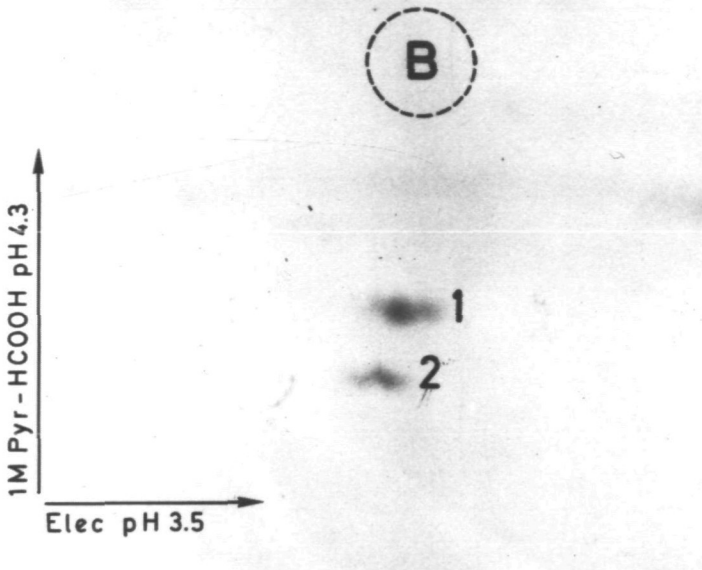
Fig. 4 : Isolation and identification of the cap structures from the major dl-1811 leader fragment.
Another aliquot of the S1-resistant band N° 3 (see Figure 2b) was treated with T2 ribonuclease and bacterial alkaline phosphatase and separated by electrophoresis at pH 3.4 in the first dimension and chromatography in 1 M pyridine-formic acid (pH 4.3) in the second dimension (31). B denotes the position of the blue dye marker xylene cyanol FF. The identity of the capped products $^{7m}Gppp^{m}AmpG$ (cap I) and $^{7m}Gppp^{m}AmpGmpA$ (cap II) was deduced from their specific position on the map relative to the blue dye marker; these spots correspond to spots 5 and 7 in Figure 3 of ref. 27, which describes the characterization of all the different dl-1811 capped termini.

couple of cap I and corresponding cap II structures is indeed identical to the major 5' cap terminus of total dl-1811 late mRNA as previously characterized (27).

Based on the trinucleotide sequence A-G-A, we concluded from these results that the capped nucleotide of the major dl-1811 leader fragment must correspond to position L290 of the late region of the genome (Figure 5; nucleotides are numbered as described in ref. 20). L290 lies upstream from the major WT capping

```
              310            320            330            340
         ┌─────────────────────────────────────────────────────┐
         │A A G T T C C T C T T T C A G A G G T T*A T T T C A G G C C A T G G T G C T G C│
         │T T C A A G G A G A A A G T C T C C A A T A A A G T C C G G T A C C A C G A C G│
         └─ ─ ─ ─ ─ ─ ─ ─ ─ ─ ─ ─ ─ ─ ─ ─ ─ ─ ─ ─ ─ ─ ─ ─ ─ ─ ─ ─ ─ ─ ─ ─ ─ ─ ─ ─ ─ ─ ┘

  260          270            280            290            300
  A C A C A T T C C A C A G C T G G T T C T T T C C G C C T C A G A A G G T A C C T A A C C G C C G G
  T G T G T A A G G T G T C G A C C A A G A A A G G C G G A G T C T T C C A T G G A T T G G C G G C C
 (0.709)              └──────┘                            └──────┘        └──────┘
                       Pvu II                              Kpn I          Hpa II

                                           ᵐG_ppp ᵐA_m G_m A A G G U A C C U A A C C G C C G G
                                                              └─────────┘ └─────────┘
                                                                  243          020

  350          360            370            380            390
  C U G U C A C G C C A G G C C U C C G U U A A G G U U C G U A G G U C A U G G A C U G A A A G U A A
  └───┘ └─────┘ └─────┘ └───────┘ └───────┘ └───────┘ └───────┘ └───────┘ └───────┘
   110    121     021     140       202       210       101       211       111       003

  400          410            420            430            440
  A A A A A C A G C U C A A C G C C U U U U U G U G U U U G U U U U A G A G C U U U U G C U G C A A U
  └─────────┘ └─────────┘ └─────────────┘ └─────────┘ └───────┘ └─────────┘ └───────┘
     118          132           520          100 300     401       001 410     110

  450          460            470            480            490
  U U U G U G A A G G G G A A G A U A C U G U U G A C G G G A A A C G C A A A A A A A C C A G A A A G G
  └─────┘ └─────┘ └─────┘ └───────┘ └───────┘ └───────┘ └───────┘ └─────────────────┘ └───────┘
   412    100 002    002     212       200 011     013            037                   003

  500 ↓        510            520
  U U A A C U G A A A A A C C A G A A A G U U A A C U G...
```

<u>Fig. 5</u> : Nucleotide sequence of part of the Hind II + III restric-
          tion fragment C from the late region of dl-1811 DNA (27)
          and of the corresponding 5'-terminal sequence of the
          major altered leader fragment of the viral late mRNA.
Nucleotide numbers are based on a total WT SV40 sequence of 5243
basepairs according to the system of Fiers et al. (32) subse-
quently corrected for the late region (20). The recognition
sequences of relevant restriction endonucleases are enclosed in
brackets. The 40 basepair dl-1811 deletion is shown in the box
and the site of the deletion, between nucleotides L304 and L344,
is indicated (20, 27). The position of the major WT capping site
(21) is marked with an asterisk.
        The major leader fragment of dl-1811 late mRNA starts at WT
position L290 and is colinear with the DNA sequence up to nucleo-
tide L526(±1)(map position 0.761), where a donor splicing signal
is encountered. The T1 oligonucleotides are indicated below the
RNA sequence and they are named in a code which gives the base
composition of the respective products (see legend to Figure 3).
These oligonucleotides may be seen on the T1 fingerprint in Fig-
ure 3. The arrow marks the end of the DNA fragment used for hy-
bridization, i.e., the Hind II cutting site at position 0.756.

site, and the mutant leader sequence is colinear with the mutant

DNA sequence (27) from L290 up to the position where splicing

occurs (at 0.761 map unit, i.e., nucleotide L526 ± 1). This means

that the length of the major mutant leader sequence (Figure 2b,

band 3) amounts to 197(±1) nucleotides.

     Analysis of the other bands revealed the presence of addi-
tional A-cap structures characterized previously (cf. ref. 27)
and the occurrence of specific Tl oligonucleotides derived from
the DNA region preceding the aforementioned major dl-1811 capping
site (evidence not shown). However, as there was too little radio-
activity present in these minor S1-resistant bands, unambiguous
characterization proved difficult and only visual inspection of
the two-dimensional fractionation patterns was possible. From
these data we nevertheless were able to conclude that all the
alternative leader fragments analysed (5 bands) extend further in
the 5' direction than the major mutant leader species (i.e., the
5' terminus was located closer to the origin of DNA replication).
However, considering the variety of trinucleotides which could
correspond to the 5'-cap sequences (27), we were not able to pin-
point the 5' terminus of these additional leader segments.

DISCUSSION
     The most prominent leader sequence in dl-1811 late mRNA
was characterized both by fingerprinting and by cap structure
analysis of the material purified by the Berk and Sharp S1-
procedure (30). Under the experimental conditions used, the
RNA·DNA hybrids were found to be fairly stable and showed little
if any breathing at the ends. Indeed, even after extensive di-
gestion with S1 nuclease the 3'-terminal Tl products were pre-
sent in molar amount on the fingerprint. The same holds for the
5'-terminal cap structures, which did not disturb hybridization
and in fact allowed isolation of capped termini from S1-digested
RNA·DNA hybrids. Indeed, we found that the major leader fragment
in the mutant mRNA starts at nucleotide L290 (map position 0.716)
and is colinear with the dl-1811 DNA sequence up to the major
donor splicing site of WT RNA at nucleotide L526(±1) (map posi-
tion 0.761). One consequence of the deletion in dl-1811 is that
the coding potential of the open reading frame present in the
WT leader fragment is lost: the presumptive initiation codon of
this "agnogene" is comprised within the dl-1811 deletion (cf.
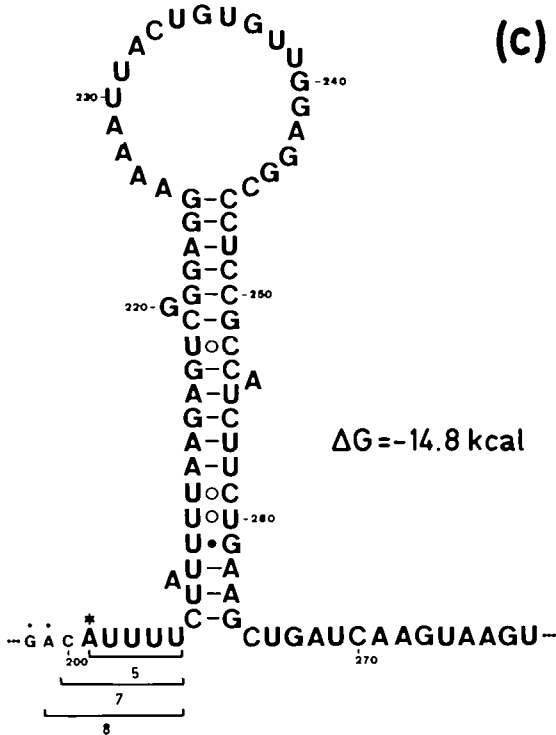Figure 5) and the reading frame is not restored in the mutant

leader segments by any preceding initiation codon. It follows
that the potential product of the agnogene, if it is indeed pro-
duced by the WT virus, is not essential for viability under
tissue-culture conditions. Indeed, the non-essential nature of
this region of the genome is well known (15-18). The slower growth
rate of dl-1811, as manifested by the smaller size of the plaques
(27), may perhaps be due to a slightly decreased level of late
mRNA synthesis and/or its different 5'-terminal structure. The
donor splicing site in the leader sequence, viz the tetranucleo-
tide A-G-G-U corresponding to nucleotides L293-296 (map position
0.716) (13, 14), is present in the major mutant mRNA, but appar-
ently the intra-leader splicing event has not occurred, because
the RNA started only a few bases earlier. Possibly adjacent cap-
ping and/or methylation distorts the splicing recognition signal
and prevents the RNA from being spliced at this position, due,
perhaps, to absence of a proper secondary structure (33). In con-
trast, mutant leader fragments which begin much more upstream do
undergo splicing at this position, as has been tentatively de-

(a)

$\Delta G = -24$ kcal

(b)

$\Delta G = -26.8$ kcal

(c)

$$\Delta G = -14.8\ kcal$$

(RNA secondary structure diagram, polyoma virus late mRNA)

```
              C U G U G
           A           U U
          U               G -240
  230- U                   G
        A                   A
        A                 G G
         A  A    G-C   C
            G-C
            A-U
            G-C
      220-G G-C -250
            C-G
            U○C
            G-C  A
            A-U
            G-C
            A-U
            A-U
            U○C
            U○U -260
            U•G
          A U-A
            U-A
            C-G
  --G A C *A U U U U   C U G A U C A  A G U A A G U --
          200                       270
              5
              7
              8
```

Fig. 6 : Proposed secondary structure models for the RNA se-
quence corresponding to the 5'-terminal major leader fragment(s)
and the preceding, possibly transcribed, nucleotide sequence of
(a) WT SV40 late mRNA, (b) dl-1811 late mRNA, and (c) polyoma
virus late mRNA. The SV40 nucleotides are numbered as in Figure 5
and hence are identical for both WT and dl-1811 DNA (and cor-
responding RNA). The 40 nucleotide deletion is indicated by an
arrow in (b). Polyoma nucleotide numbers were taken from Flavell
et al. (35). Free energies of the hairpin structures were calcu-
lated according to the stability rules of Tinoco et al. (39).
The small contribution made by a G·U basepair to the stability of
a helix is indicated by a dot. The symbol ○ between two bases
indicates that they occur opposite each other in the helix but
that the union does not add to the stability of the structure
(40). The positions of the major capping sites are marked by an
asterisk.
Note that in all the examples the major capping site precedes
a stable (G·C)-rich hairpin structure by exactly 5 nucleotides.
However, in the WT RNA sequence, the major dl-1811 capping site
(marked by 1811[*]) is part of the stem of a hairpin loop. The
boxed AUG sequence in (a) is the initiation codon of the conjec-
tural agnogene.

duced from visual interpretation of T1 fingerprints of the alter-
native dl-1811 minor leader segments.

It is not unexpected that the most prominent cap I and II
structure found on dl-1811 late mRNA corresponds to the 5' termini
of the fragment found in the major band of Sl-resistant leader
RNA. It may be noted that this particular capped end is only a
minor component in the diversity of WT SV40 cap structures (13,
34, and our unpublished data) and more particularly it seems that
only one 19S leader species but no 16S species starts at this
position (13). Favoring of the capping site at nucleotide L290
in dl-1811 as compared to the relative use of this position in
WT SV40 suggests that the appearance of different caps is subject
to regulation or depends on local configuration of the mRNA.
Indeed, the relative abundance of different caps may be directly
dependent on structural or functional requirements. A secondary
structure model for part of the WT SV40 leader sequence (Figure 6)
shows that the major WT capping site is located at the base of a
stable hairpin, whereas nucleotide L290 is fully involved in base-
pairing in the stem of the preceding loop. Also, some minor WT
capping sites may be found in single-stranded regions preceding
similar (G·C)-rich loop structures (our unpublished results).
Remarkably the structure model for the dl-1811 RNA sequence shows
that now the position of the major mutant capping site at nucleo-
tide L290 lies free and at exactly the same distance (5 nucleo-
tides) from an analogous (G·C)-rich hairpin. In the case of
polyoma virus, which is distantly related to SV40, several cap
structures in the late mRNA corresponding to consecutive purine
positions in the DNA sequence have been identified (35). These
may conceivably be derived by a stuttering effect of the polymer-
ase II during initiation of transcription (35). The variety of
capping sites in SV40 late RNA and especially their relative
broad distribution over the DNA sequence suggests a more compli-
cated mechanism for their formation than simple stuttering. Pos-
sibly, the RNA polymerase starts at many different positions in
a defined region and only those RNA molecules with a hairpin near
their 5' end become stabilized against rapid degradation and are
conserved. Alternatively, initiation of RNA synthesis may occur
more proximal to the origin of DNA synthesis and processing

enzymes may subsequently recognize the proposed hairpin loops
(Figure 6), whereupon caps may be formed after further phosphory-
lation at the processed 5' ends, analogous to, e.g., the capping
system of vaccinia virus (36, 37). It is of interest that also in
polyoma virus mRNA a loop structure following (and specifying ?)
the capping site can be drawn. However, this structure is less
stable (and less G·C-rich) than the proposed SV40 hairpins and
may perhaps be responsible for rather imprecise processing,
thereby giving rise to multiple caps around the same position
(35). More direct experimental approaches are needed to dis-
tinguish unambiguously between these two possibilities for
generating the diversity of capped 5'-termini in the papova
virus late mRNAs.

REFERENCES

1.   Weinberg, R., Warnaar, O. and Winocour E. (1972) J. Virol.
      10, 193-201.
2.   Fareed, G. and Davoli, D. (1977) Annual Rev. Biochem. 46,
      471-522.
3.   May, E., Kopecka, H. and May, P. (1975) Nucl. Acids Res.
      2, 1995-2005.
4.   Khoury, G., Carter, B., Ferdinand, F., Howley, P., Brown,
      M. and Martin, M. (1976) J. Virol. 17, 832-840.
5.   Aloni, Y., Dhar, R., Laub, O., Horowitz, M. and Khoury, G.
      (1977) Proc. Nat. Acad. Sci. U.S.A. 74, 3686-3690.
6.   Hsu, M. and Ford, J. (1977) Proc. Nat. Acad. Sci. U.S.A.
      74, 4982-4985.
7.   Lai, C., Dhar, R. and Khoury, G. (1978) Cell 14, 971-982.
8.   Lavi, S. and Groner, Y. (1977) Proc. Nat. Acad. Sci. U.S.A.
      74, 5323-5327.
9.   Haegeman, G. and Fiers, W. (1978) Nature 273, 70-73.
10.   Celma, M., Dhar, R., Pan, J. and Weissman, S. (1977) Nucl.
      Acids Res. 4, 2549-2559.
11.   Bina-Stein, M., Thoren, M., Salzman, N. and Thompson, J.
      (1979) Proc. Nat. Acad. Sci. U.S.A. 76, 731-735.

12. Ghosh, P., Reddy, V., Swinscoe, J., Choudary, P., Lebowitz, P. and Weissman, S. (1978) J. Biol. Chem. 253, 3643-3647.
13. Ghosh, P., Reddy, V., Scinscoe, J., Lebowitz, P. and Weissman, S. (1978) J. Mol. Biol. 126, 813-846.
14. Reddy, V., Ghosh, P., Lebowitz, P. and Weissman, S. (1978) Nucl. Acids Res. 5, 4195-4213.
15. Mertz, J. and Berg, P. (1974) Proc. Nat. Acad. Sci. U.S.A. 71, 4879-4883.
16. Shenk, T., Carbon, J. and Berg, P. (1976) J. Virol. 18, 664-671.
17. Cole, C., Landers, T., Goff, S., Manteuil-Brutlag, S. and Berg, P. (1977) Virology 24, 277-294.
18. Subramanian, K. (1979) Proc. Nat. Acad. Sci. U.S.A. 76, 2556-2560.
19. Contreras, R., Cole, C., Berg, P. and Fiers, W. (1979) J. Virol. 29, 789-793.
20. Van Heuverswyn, H. and Fiers, W. (1979) Eur. J. Biochem., in press.
21. Haegeman, G. and Fiers, W. (1978) Nucl. Acids Res. 5, 2359-2371.
22. Furiuchi, Y., Muthukrishnan, S., Tomasz, J. and Shatkin, A. (1976) J. Biol. Chem. 251, 5043-5053.
23. Furiuchi, Y. (1978) Proc. Nat. Acad. Sci. U.S.A. 75, 1086-1090.
24. Groner, Y. and Hurwitz, J. (1975) Proc. Nat. Acad. Sci. U.S.A., 72, 2930-2934.
25. Wei, C. and Moss, B. (1977) Proc. Nat. Acad. Sci. U.S.A. 74, 3758-3761.
26. Winicov, I. and Perry, R. (1976) Biochemistry 15, 5039-5046.
27. Haegeman, G., Van Heuverswyn, H., Gheysen, D. and Fiers, W. (1979) J. Virol. 31, 484-493.
28. Villareal, L., White, R. and Berg, P. (1979) J. Virol. 29, 209-219.
29. Volckaert, G., Min Jou, W. and Fiers, W. (1976) Anal. Biochem. 72, 433-446.
30. Berk, A. and Sharp, P. (1977) Cell, 12, 721-732.
31. Haegeman, G. and Fiers, W. (1978) J. Virol. 25, 824-830.
32. 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.
33. Iserentant, D. and Fiers, W. (1979) manuscript in preparation.
34. Canaani, D., Kahana, C., Mukamel, A. and Groner, Y. (1979) Proc. Nat. Acad. Sci. U.S.A. 76, 3078-3082.
35. Flavell, A., Cowie, A., Arrand, J. and Kamen, R. (1979) J. Virol. in press.
36. Spencer, E., Loring, D., Hurwitz, J. and Monroy, G. (1978) Proc. Nat. Acad. Sci. U.S.A. 75, 4793-4797.
37. Moss, B., Gershowitz, A., Wei, C. and Boone, R. (1976) Virology 72, 341-351.
38. De Wachter, R., Merregaert, J., Vandenberghe, A., Contreras, R. and Fiers, W. (1971) Eur. J. Biochem. 22, 400-414.
39. Tinoco, I., Borer, P., Dengler, B., Levine, M., Uhlenbeck, O., Crothers, D. and Gralla, J. (1973) Nature New Biol. 246, 40-41.
40. Iserentant, D. and Fiers, W. (1979) Eur. J.Biochem. in press.