20. W. Richards, J. Farley, D. L. Alkon, *ibid.* **9**, 916 (1983); W. Richards, J. Farley, D. L. Alkon, *Behav. Neural Biol.*, in press.
21. J. Farley and G. Kern, *Anim. Learn. Behav.*, in press.
22. I. Lederhendler and D. L. Alkon, *Biol. Bull. Abstr. (Woods Hole, Mass.)* **165**, 528 (1983).
23. T. J. Crow and D. L. Alkon, *Science* **209**, 412 (1980); J. Farley and D. L. Alkon, *J. Neurophysiol.* **48**, 785 (1982).
24. A. West, E. Barnes, D. L. Alkon, *J. Neurophysiol.* **48**, 1243 (1982).
25. I. I. Lederhendler, Y. Goh, D. L. Alkon, *Soc. Neurosci. Abstr.* **8**, 824 (1982).
26. D. L. Alkon, I. Lederhendler, J. J. Shoukimas, *Science* **215**, 693 (1982).
27. D. L. Alkon, in *Primary Neural Substrates of Learning and Behavioral Change*, D. L. Alkon and J. Farley, Eds. (Cambridge Univ. Press, New York, 1984), p. 291; D. L. Alkon, *Sci. Am.* **249**, 70 (July 1983).
28. R. Forman, D. L. Alkon, M. Sakakibara, J. Harrigan, I. Lederhendler, J. Farley, *Soc. Neurosci. Abstr.* **10**, 121 (1984); J. Farley, M. Sakakibara, D. L. Alkon, *ibid.*, p. 270.
29. Y. Goh and D. L. Alkon, *J. Neurophysiol.* **52**, 156 (1984).
30. D. L. Alkon, *Science* **210**, 1375 (1980).
31. J. A. Connor and D. L. Alkon, *J. Neurophysiol.* **51**, 745 (1984).
32. D. L. Alkon, J. Shoukimas, E. Heldman, *Biophys. J.* **40**, 245 (1982).
33. J. Farley, W. G. Richards, L. J. Ling, E. Liman, D. L. Alkon, *Science* **221**, 1201 (1983).
34. J. J. Shoukimas and D. L. Alkon, *Soc. Neurosci. Abstr.* **6**, 17 (1980); D. L. Alkon, *Science* **205**, 810 (1979).
35. D. L. Alkon, J. Farley, M. Sakakibara, B. Hay, *Biophys. J.*, in press.
36. D. L. Alkon and M. Sakakibara, *Soc. Neurosci. Abstr.* **10**, 101 (1984).
37. J. T. Neary, T. Crow, D. L. Alkon, *Nature (London)* **293**, 658 (1981); D. Alkon *et al.*, *Biol. Bull. Abstr. (Woods Hole, Mass.)*, in press.
38. J. T. Neary and D. L. Alkon, *J. Biol. Chem.* **28**, 8979 (1983); J. T. Neary, in *Primary Neural Substrates of Learning and Behavioral Change*, D. L. Alkon and J. Farley, Eds. (Cambridge Univ. Press, New York, in press).
39. J. Acosta-Urquidi, D. L. Alkon, J. T. Neary, *Science* **224**, 1254 (1984).
40. A. McElearney and J. Farley, *Soc. Neurosci. Abstr.* **9**, 915 (1983).
41. M. Tabata and D. L. Alkon, *J. Neurophysiol.* **48**, 174 (1982).
42. E. Heldman, Y. Grossman, T. Jerussi, D. L. Alkon, *ibid.* **42**, 153 (1979).
43. M. Sakakibara, D. L. Alkon, I. Lederhendler, E. Heldman, *Soc. Neurosci. Abstr.* **10**, 950 (1984).
44. E. R. Kandel and J. H. Schwartz, *Science* **218**, 433 (1982).
45. J. S. Camarado, S. Siegelbaum, E. R. Kandel, in *Primary Neural Substrates of Learning and Behavioral Change*, D. L. Alkon and J. Farley, Eds. (Cambridge Univ. Press, New York, in press).
46. S. A. Siegelbaum, J. S. Camarado, E. R. Kandel, *Nature (London)* **299**, 413 (1982).
47. H. B. Kistler, Jr., R. D. Hawkins, J. Kuester, E. R. Kandel, J. H. Schwartz, *Soc. Neurosci. Abstr.* **9**, 915 (1983).
48. J. L. McGaugh, J. L. Martinez, Jr., R. A. Jensen, T. J. Hannan, B. J. Vasquez, R. B. Messing, K. C. Liang, C. B. Brewton, V. R. Spichler, in *Neurobiology of Learning and Memory*, H. Matthies, Ed. (Raven Press, New York, 1983).
49. E. F. Barrett and J. N. Barrett, *J. Physiol. (London)* **255**, 737 (1976).
50. R. Llinás and Y. Yarom, *ibid.* **315a**, 549 (1981); *ibid.*, **315b**, 569 (1981).
51. M. Segal and J. L. Barker, *J. Neurophysiol.* **51**, 1409 (1984); R. B. Clark and R. K. S. Wong, *Soc. Neurosci. Abstr.* **9**, 601 (1983).
52. P. Brehm, R. Eckert, D. Tillotson, *J. Physiol. (London)* **306**, 193 (1980); R. Eckert and D. Tillotson *ibid.* **314**, 265 (1981); R. Eckert and D. Ewald, *ibid.* **345**, 533 and 599 (1983); J. Chad, R. Eckert, D. Ewald, *ibid.* **347**, 279 (1984).
53. T. Bliss and T. Lømo, *ibid.* **232**, 334 (1973); P. Andersen, S. H. Sondberg, O. Sveen, H. Wigström, *Nature (London)* **266**, 736 (1977).
54. G. Lynch, J. Larson, S. Kelso, G. Barrionvero, F. Schottler, *Nature (London)* **305**, 719 (1983).
55. K. Krnjevic, M. E. Morris, R. J. Riefenstein, N. Ropert, *Can. J. Physiol. Pharmacol.* **60**, 1658 (1982).
56. R. R. Llinás, in *Handbook of Physiology*, Section 1, *The Nervous System*, vol. 2, *Motor Control*, V. B. Brooks, Ed. (American Physiological Society, Bethesda, Md., 1981), part 2, p. 831.
57. I. Gormezano, in *Experimental Methods and Instrumentation in Psychology*, J. B. Sidowski, Ed. (McGraw-Hill, New York, 1966), p. 385.
58. J. Disterhoft, D. Coulter, D. L. Alkon, *Biol. Bull. (Woods Hole, Mass.)*, in press.
59. G. Lynch, M. Browning, W. F. Bennett, *Fed. Proc. Fed. Am. Soc. Exp. Biol.* **38**, 2117 (1979); G. Lynch, S. Halpain, M. Baudry, *Brain Res.* **244**, 101 (1982).
59a. R. F. Thompson *et al.*, in *Conditioning: Representation of Involved Neural Functions*, C. D. Woody, Ed. (Plenum, New York, 1982), p. 115.
60. S. Gobel, W. M. Falls, G. J. Bennett, M. Abdelmoumene, H. Hyashi, E. Humphrey, *J. Comp. Neurol.* **194**, 781 (1980).
61. D. H. Hubel and T. N. Wiesel, *J. Neurophysiol.* **28**, 1041 (1965); T. N. Wiesel and D. H. Hubel, *Soc. Neurosci. Abstr.* **4**, 478 (1974); P. Rakic, *UCLA Forum Med. Sci.* **18**, 3 (1975); W. T. Greenough, J. M. Juraska, F. R. Volkmar, *Behav. Neural Biol.* **26**, 287 (1979).
62. Y. Naitoh and R. Eckert, in *Cilia and Flagella*, M. Sleigh, Ed. (Academic Press, New York, 1974), p. 305.
63. A. Pellionisz and R. Llinás, *Neuroscience* **4**, 323 (1979).
64. I thank J. Kuzirian for technical assistance and R. Golder and L. Golder for artwork.

RESEARCH ARTICLE

Role of the Conserved AAUAAA Sequence: Four AAUAAA Point Mutants Prevent Messenger RNA 3′ End Formation

Marvin Wickens and Pauline Stephenson

In bacteria, messenger RNA's (mRNA's) generally are primary transcripts and are exact replicas of DNA sequences in the genome. In eukaryotes, however, a primary transcript becomes an mRNA only after a collection of physical modifications: capping, cleavage to form a new 3′ terminus, polyadenylation, splicing, base methylation, and the transport of mRNA from nucleus to cytoplasm (*1–3*). In principle, each maturation step provides a means of regulating mRNA formation. This article focuses on defining the sequences within a primary transcript that are necessary for three of these steps: cleavage, polyadenylation, and transport.

Interest in these three steps stems first from their universality (*1–3*). Almost all cellular and viral mRNA precursors that have been examined are cleaved to generate a new 3′ terminus to which polyadenylic acid [poly(A)] is added (*1–4*) [for an exception, see (*5, 6*)]. Only mRNA's with a mature 3′ terminus are transported to the cytoplasm. Cleavage involves at least one endonucleolytic scission (*7*). The site of this scission is unknown; it may coincide with the polyadenylation site, or may lie downstream. Similarly, the protein or nucleic acid factors that catalyze the reaction have not yet been characterized.

Interest in these maturation steps also stems from their potential to determine how much mRNA, and what type of mRNA, a cell contains. For example, the same gene can produce two or more mRNA's with different 3′ termini in different cell types (*8–10*) or at different stages of viral infection (*2*).

Formation of SV40 (simian virus 40) mRNA's, like that of most mRNA's, involves posttranscriptional processing. This small double-stranded DNA virus genome is transcribed into two families of mRNA's in monkey cells (Fig. 1A). All mRNA's that accumulate late in viral infection (late mRNA's) are cleaved (*11*) and polyadenylated at the same site (*12*).

Sequences involved in cleavage, polyadenylation, and transport of mRNA in animal cells have not yet been completely defined, but must include the highly conserved AAUAAA sequence located 6 to 26 bases before the polyadenylation site of nearly all animal cell mRNA's (*13*). Fitzgerald and Shenk (*10*) demonstrated that deletion of this sequence

Marvin Wickens and Pauline Stephenson are members of the Department of Biochemistry, College of Agriculture and Life Sciences, University of Wisconsin–Madison, Madison 53706. Part of this work was done by Marvin Wickens while he was visiting the M.R.C. Molecular Biology Laboratory, Hills Road, Cambridge, England.

Abstract. *A small region (220 bases) of SV40 sequence information—141 bases before the polyadenylation site and 79 beyond—are sufficient for cleavage of an messenger RNA precursor (that is, the formation of a mature 3′ terminus), the addition of polyadenylic acid, and the transport of messenger RNA from the nucleus to the cytoplasm. These 220 bases include a highly conserved sequence—AAUAAA (A, adenine; U, uracil). Four point mutations in this sequence—AACAAA, AAUUAA, AAUACA, and AAUGAA (C, cytosine; G, guanine)—prevent cleavage.*

plus ten adjacent nucleotides prevents the formation of stable mRNA. Montell *et al.* (*14*) have reported that the AAUAAA (A, adenine; U, uracil) sequence of the adenovirus E1A gene is necessary for cleavage.

In this article, we present our analysis of the expression of recombinant SV40/pBR322 DNA's after they have been injected into the nucleus of *Xenopus* oocytes. Oocytes provide a valuable assay system for RNA processing because (i) DNA, RNA, or protein can be injected into the nucleus or cytoplasm; (ii) a wide spectrum of maturation steps are carried out; and (iii) abnormal transcripts often are stable (*5, 15–17*). Moreover, frog oocytes efficiently process RNA precursors from other species and cell types. For example, oocytes injected with SV40 DNA produce late region transcripts that are cleaved, polyadenylated, spliced, transported to the cytoplasm (*18*), and translated (*19*).

A small region of SV40 encodes 3′ end formation. Prokaryotic DNA templates, like pBR322, are transcribed by RNA polymerase II after injection into the nucleus of frog oocytes (*20*). In fact, functional mRNA can be produced from recombinant plasmids in which a eukaryotic gene, lacking its own promoter, is inserted into pBR322 (*21*). In the following experiments we exploited this finding. Small fragments of SV40 DNA, lacking a promoter, were inserted into pBR322. These pBR322/SV40 recombinants were then injected into oocytes. RNA polymerase II initiates transcription in pBR322 sequences, elongates through the SV40 insert, and continues into the pBR322 sequences beyond. We then determined whether the SV40 sequences present were sufficient for processing by analyzing whether such "fused" pBR322/SV40 primary transcripts are processed even in the absence of most of the normal mRNA sequence.

The construction of a pBR322/SV40 plasmid (*22*) is summarized in Fig. 1A. The SV40 fragment contains 141 bases before and 802 bases after the polyadenylation site of late mRNA's; the plasmid is therefore designated pSVL−141/+802. The C (cytosine) preceding poly(A) is designated −1. The 141 residues before the polyadenylation site contain the entire 3′ untranslated region of late mRNA's, including the AAUAAA sequence.

Transcripts of pSVL−141/+802 are accurately and efficiently cleaved in oocytes (Fig. 1C). RNA purified (*23*) from oocytes injected with pSVL−141/+802 or with SV40 DNA (*24*) was analyzed by S1 mapping (*25, 26, 27*). The labeled DNA fragment that was used as probe spans the 3′ terminus of late mRNA's (Fig. 1C). RNA that has been cleaved (that is, processed RNA) should protect a 143 base fragment from nuclease S1 digestion, whereas RNA that has not been cleaved (that is, unprocessed RNA) should protect a 240-base fragment (Fig. 1C). RNA from oocytes that had been injected either with SV40 DNA (Fig. 1C, lane 11) or with pSVL−141/+802 DNA

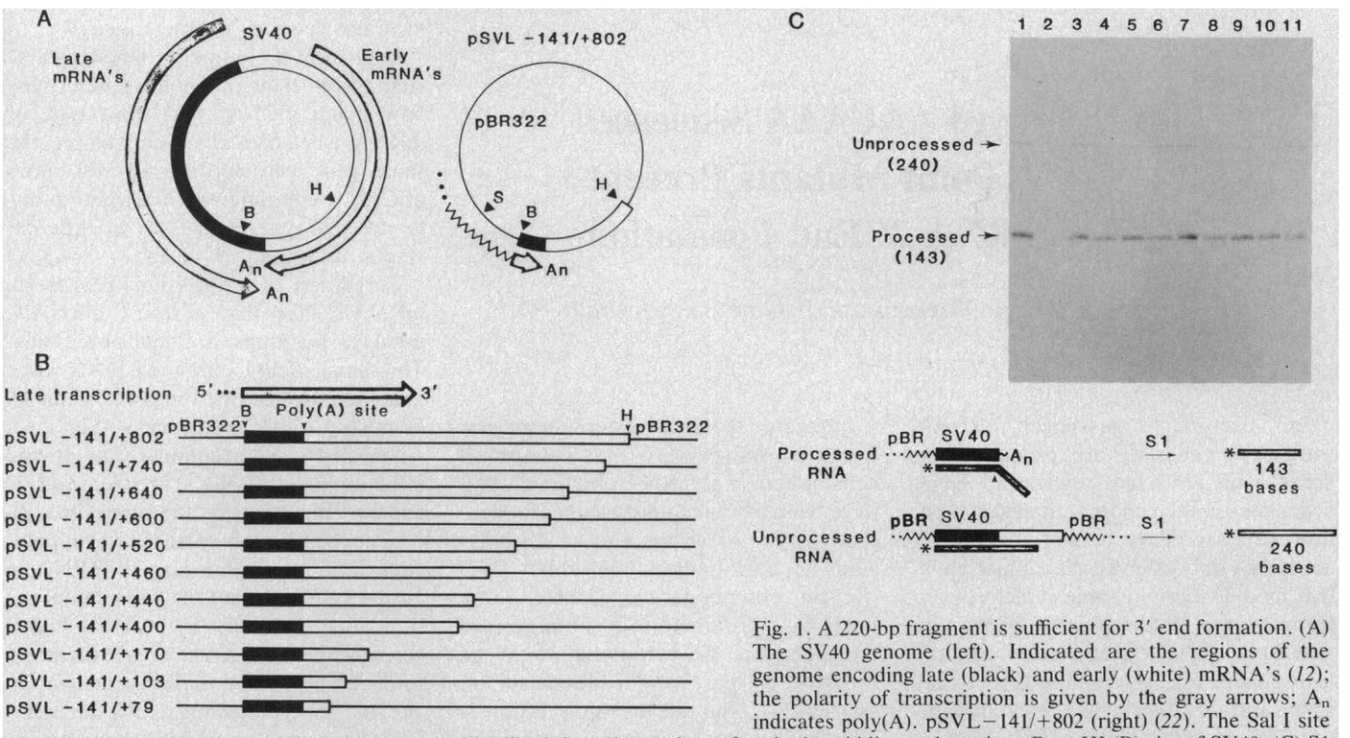

Fig. 1. A 220-bp fragment is sufficient for 3′ end formation. (A) The SV40 genome (left). Indicated are the regions of the genome encoding late (black) and early (white) mRNA's (*12*); the polarity of transcription is given by the gray arrows; A_n indicates poly(A). pSVL−141/+802 (right) (*22*). The Sal I site of pBR322 is designated by S. (B) Deletion plasmids. The leftward boundary of each plasmid lies at the unique Bam HI (B) site of SV40. (C) S1 mapping (*23, 24*). RNA was hybridized to a single-stranded, 3′ end-labeled Bam HI–Bcl I fragment of SV40 (*25, 26*). RNA was obtained from oocytes injected with pSV−141/+802 (lane 1), wild-type SV40 DNA (lane 11), or with the deletion templates pSVL−141/+740 (lane 3), −141/+600 (lane 4), −141/+520 (lane 5), −141/+440 (lane 6), −141/+400 (lane 7), −141/+170 (lane 8), −141/+103 (lane 9), and −141/+79 (lane 10). The RNA used in lane 2 was prepared from uninjected oocytes. The structure of RNA-DNA hybrids formed between RNA (wide line) and the probe DNA (thin gray line). S1 cleavage sites are indicated by arrowheads, and the labeled end of the DNA indicated with an asterisk. These notations are used in all subsequent figures.

(lane 1) yield the same result; two fragments are protected, one consisting of 240 bases and the other of 143 bases, in a molar ratio of 1 to 6 (determined by microdensitometry). As we expected, RNA from uninjected oocytes did not protect any probe (Fig. 1C, lane 2). We conclude that primary transcripts containing only 943 bases of SV40 sequence are cleaved as efficiently and as accurately as primary transcripts of intact SV40 DNA. (We define efficiency as the fraction of accumulated RNA that is processed.)

To further delimit the sequences required for processing, we constructed (*22*) a series of plasmids bearing deletions of the 3′ flanking sequence present in pSVL−141/+802 (Fig. 1B). RNA's prepared from oocytes injected with each deletion DNA were analyzed by S1 mapping (Fig. 1C, lanes 3 to 10). Each template directs cleavage with the same efficiency and accuracy as wild-type SV40 DNA (lane 11). The smallest region of SV40 DNA tested in Fig. 1, present in pSVL−141/+79, is 220 bases long. It contains 141 bases before and 79 bases after the polyadenylation site.

Cleavage of RNA's transcribed from these templates is due to the SV40 sequence they contain. We constructed plasmids in which the pSVL−141/+79 fragment is inverted with respect to pBR322. The SV40 late-region transcripts of these plasmids necessarily contain different pBR322 sequences, since these transcripts initiate on the opposite strand of pBR322. Nonetheless, they are cleaved as efficiently as transcripts of pSVL−141/+79 (data not shown).

Polyadenylation and nucleocytoplasmic partitioning. To determine whether transcripts of pSVL−141/+79 are polyadenylated, RNA that had been extracted from oocytes injected with pSVL−141/+79 was fractionated into polyadenylated and nonpolyadenylated components (Fig. 2A). These two RNA fractions were analyzed by S1 mapping. Nonpolyadenylated RNA protects a fragment of 220 bases, corresponding to uncleaved RNA, but does not protect the 143-base fragment that corresponds to cleaved RNA. Polyadenylated RNA protects primarily the 143-base fragment, and only a very small amount of the 220-base fragment. Thus, cleaved RNA is polyadenylated and uncleaved RNA is not.

Transcripts of pSVL−141/+79 are properly partitioned between nucleus and cytoplasm (Fig. 2B). Oocytes were injected with pSVL−141/+79 (oocytes 1 and 2) or with pSVL−141/+802 (oocytes 3 and 4). After 24 hours, nuclei and cytoplasms were isolated manually (*24*). RNA was prepared from the separated nucleus and cytoplasm of single oocytes. The level of processed and unprocessed RNA in each fraction was determined by S1 mapping.

Both templates yield the same result as wild-type SV40 (*18*): nuclear RNA protects both the 220 and 143-base fragments, while cytoplasmic RNA protects only the 143-base fragment (Fig. 2B). One obvious explanation of these data is that only those RNA's that have been cleaved enter the cytoplasm. An alternative possibility (*28*), which we have not eliminated, is that all transcripts emerge into the cytoplasm but that unprocessed RNA's are rapidly degraded.

Isolation of point mutations in AAUAAA. The highly conserved sequence AAUAAA is contained within the 220-base-pair SV40 fragment present in pSVL−141/+79. To isolate point mutations in the AAUAAA sequence, a 16-base-pair fragment of the SV40 late region, containing AAUAAA, was cloned into the β-galactosidase gene of M13 mp8 (Fig. 3, top). In this clone, Eco RI and Hind III linkers flank the insert; they were introduced to simplify subsequent manipulations. The translational reading frame of β-galactosidase is preserved; ribosomes initiating translation at the AUG of the *lac* Z gene are still in frame after passing through the SV40 insert. However, cells containing the

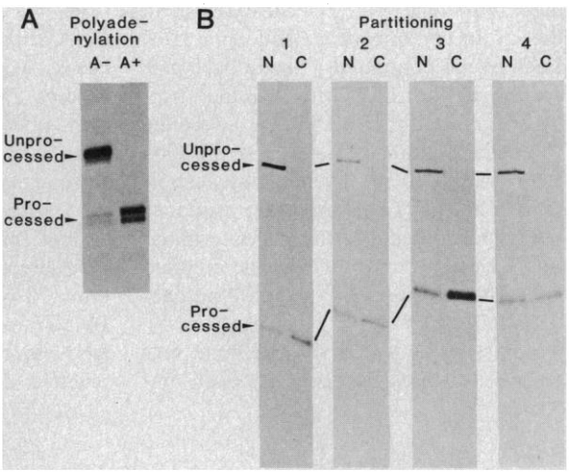

Fig. 2. Polyadenylation and transport. (A) Nonpolyadenylated (A^-) and polyadenylated (A^+) RNA from oocytes injected with pSVL−141/+79 was subjected to S1 mapping with the same method and probe described in Fig. 1. RNA that had not been cleaved (unprocessed) protects 220 bases of the probe (−141 to +79). (B) RNA was obtained from the nucleus (N) or cytoplasm (C) of oocytes injected with pSVL−141/+79 (oocytes 1 and 2) or pSVL −141/+802 (oocytes 3 and 4). S1 mapping as in Fig. 1 but with a probe extending from −141 to only +79.

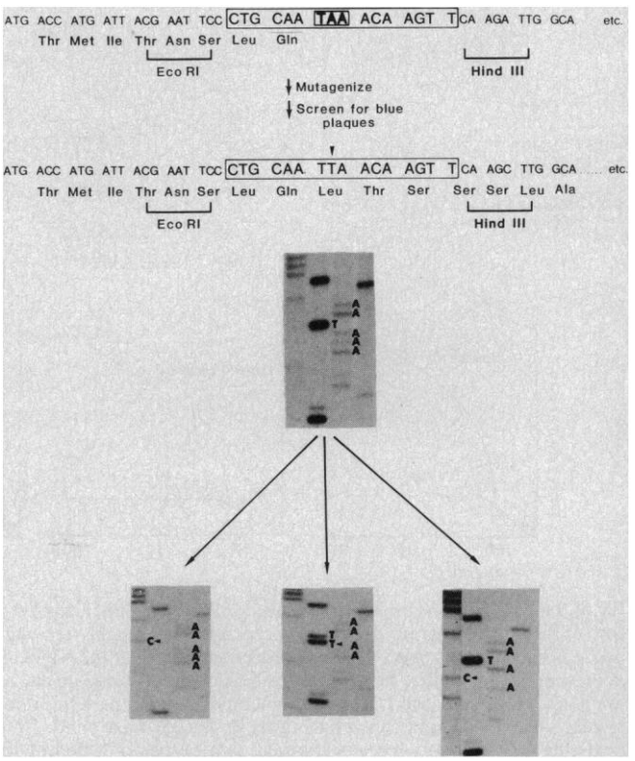

Fig. 3. Isolation of point mutations in AAUAAA (*40, 41*).

phage produce no β-galactosidase because the UAA termination codon which by chance is found in the AAUAAA sequence is in frame with respect to translation from the *lac* AUG (G, guanine). As a result, plaques of this phage are clear on plates containing the indicator dye, BCIG (*29*). Point mutations that convert the UAA codon to a missense codon should permit ribosomes to read through the SV40 insert and synthesize functional β-galactosidase. Plaques of such AAUAAA mutants therefore should be blue on plates containing the indicator dye.

Phage carrying the wild-type AATAAA sequence were mutagenized with ultraviolet light and plated onto an indicator strain on plates containing BCIG. Of roughly 50,000 plaques examined, four were blue. The SV40 region in these four genomes was sequenced (*30*). Each phage contained a single base pair change in the TAA (T, thymine) triplet, such that the AATAAA sequence was converted to AACAAA, AATACA, AATGAA, and AATTAA (Fig. 3, bottom). The AATGAA mutation produces blue plaques because the UGA termination codon is suppressed by the endogenous UGA suppressor of wild-type bacteria (*31*).

To assay the effect of these point mutations on mRNA processing, each mutant insert was transferred to a plasmid, pSVL1. This plasmid contains the same region of SV40 as pSVL−141/+103, but, in addition, includes single-copy Eco RI and Hind III linkers to either side of the AAUAAA sequence. The linkers lie at the same positions as in the M13 clones.

Point mutations in AAUAAA prevent cleavage. To determine the effect of AAUAAA point mutations on mRNA processing, oocytes were injected either with pSVL1 or with a pSVL1 derivative containing AACAAA, AATTAA, AATACA, or AATGAA in place of AATAAA (designated, for example, pSVL1: AACAAA). RNA was prepared at various times after injection, ranging from 15 to 300 minutes. Processing efficiency was determined by S1 mapping (Fig. 4). Cleaved and polyadenylated RNA should protect a fragment of 159 bases, while unprocessed RNA should protect 257 bases. As controls, RNA from oocytes injected with each DNA plus α-amanitin, or with buffer only, were analyzed in parallel.

1) *Wild-type template.* RNA transcribed from the wild-type template was first detectable after 15 minutes. At that time, 70 percent of it was unprocessed. By 60 minutes, the level of unprocessed RNA increased by a factor of 5 and reached a plateau value that did not change over the next 4 hours. In contrast, the level of processed RNA increased steadily over the entire incubation period. As a result, the ratio of processed to unprocessed RNA gradually increased during the incubation period, from 0.5 at 15 minutes to 8 after 5 hours.

Two aspects of these data merit discussion. First, whereas unprocessed RNA reaches a plateau after about 60 minutes, processed RNA accumulates gradually from 45 minutes to 5 hours. We conclude that the unprocessed molecules detected by S1 mapping are cleaved to form processed RNA and are not inert with respect to processing. This interpretation is consistent with the observed nucleocytoplasmic partitioning of transcripts (Fig. 2). Second, the 10- to 15-minute delay between the accumulation of unprocessed and processed RNA's suggests that, on average, transcripts are cleaved after 10 minutes.

In principle, the Eco RI and Hind III linkers in pSVL1 might have affected processing themselves, and so might have complicated analysis of mutants. However, in a group of control experiments we found that these foreign sequences do not reduce the efficiency or accuracy of cleavage. The details of these control experiments are presented in Fig. 5.

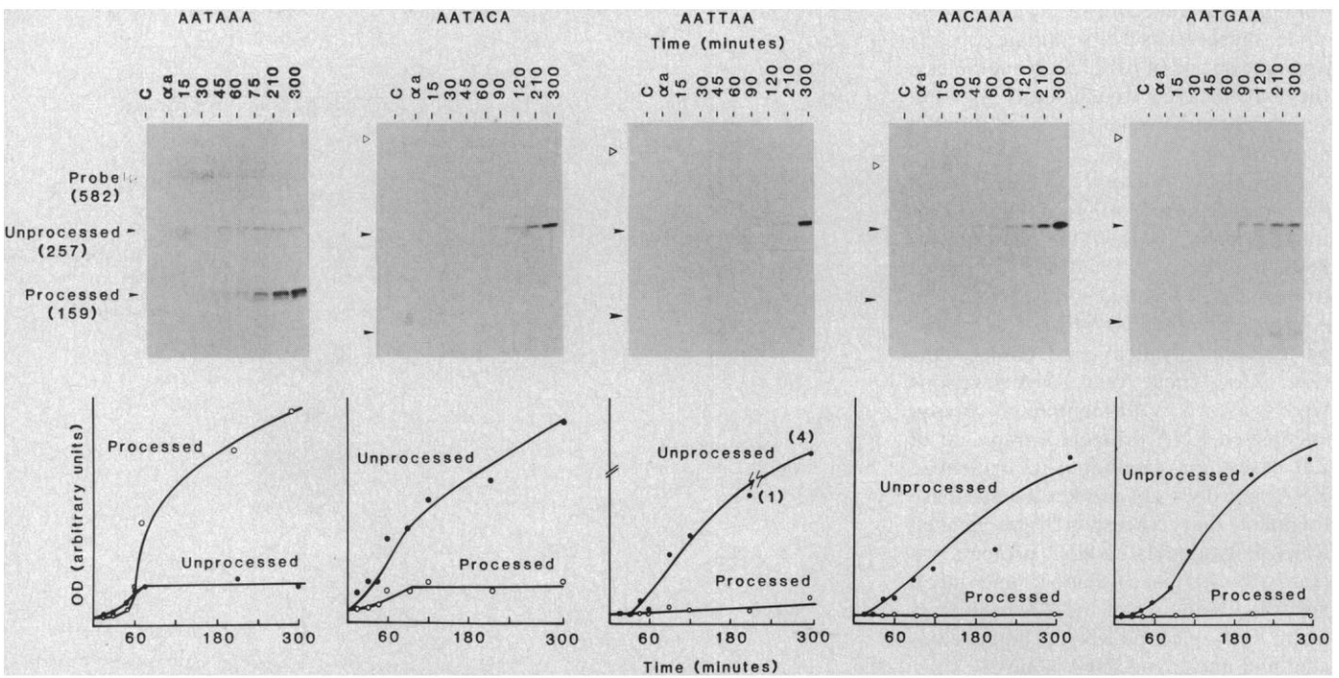

Fig. 4. Point mutations in AAUAAA affect cleavage. The time after injection (in minutes) at which the RNA was prepared is indicated above each lane of the autoradiogram. RNA was prepared from roughly ten oocytes per time point. S1 mapping was carried out as diagrammed in Fig. 6. The probe is homologous to pSVL1 over the entire length of the SV40 insert (259 bases), then diverges in sequence for its last 225 bases (see diagram at bottom of Fig. 6). The first two lanes in each autoradiogram are controls: C lanes: control, uninjected oocytes; α a lanes: oocytes injected with α-amanitin (5 μg/ml) and DNA simultaneously. Lower panels: intensity of bands corresponding to processed and unprocessed RNA (quantitated by laser microdensitometry) versus time; ●, the 257-base band; ○, the 159-base band. The intensity scale is not the same from one template to the next since experiments were performed with probes of different specific activities.

2) *Mutant templates*. Each point mutation in AAUAAA dramatically affects processing (Fig. 4). At time points when most transcripts of the wild-type template were processed, most of the transcripts of the mutant templates were not. Thus, 5 hours after injection, when the ratio of processed to unprocessed RNA in oocytes injected with pSVL1:AATAAA is 8, that ratio is only 0.13 for oocytes injected with pSVL1:AATACA, 0.04 for oocytes injected with pSVL1:AATTAA, and less than 0.01 (undetectable) for oocytes injected with pSVL1:AACAAA or pSVL1:AATGAA. The AATGAA transcripts yield a faint band of 140 ± 5 nucleotides, which is artifactual (*32*).

The small amount of 159-base fragment protected by RNA from pSVL1:AATACA and pSVL1:AATTAA is

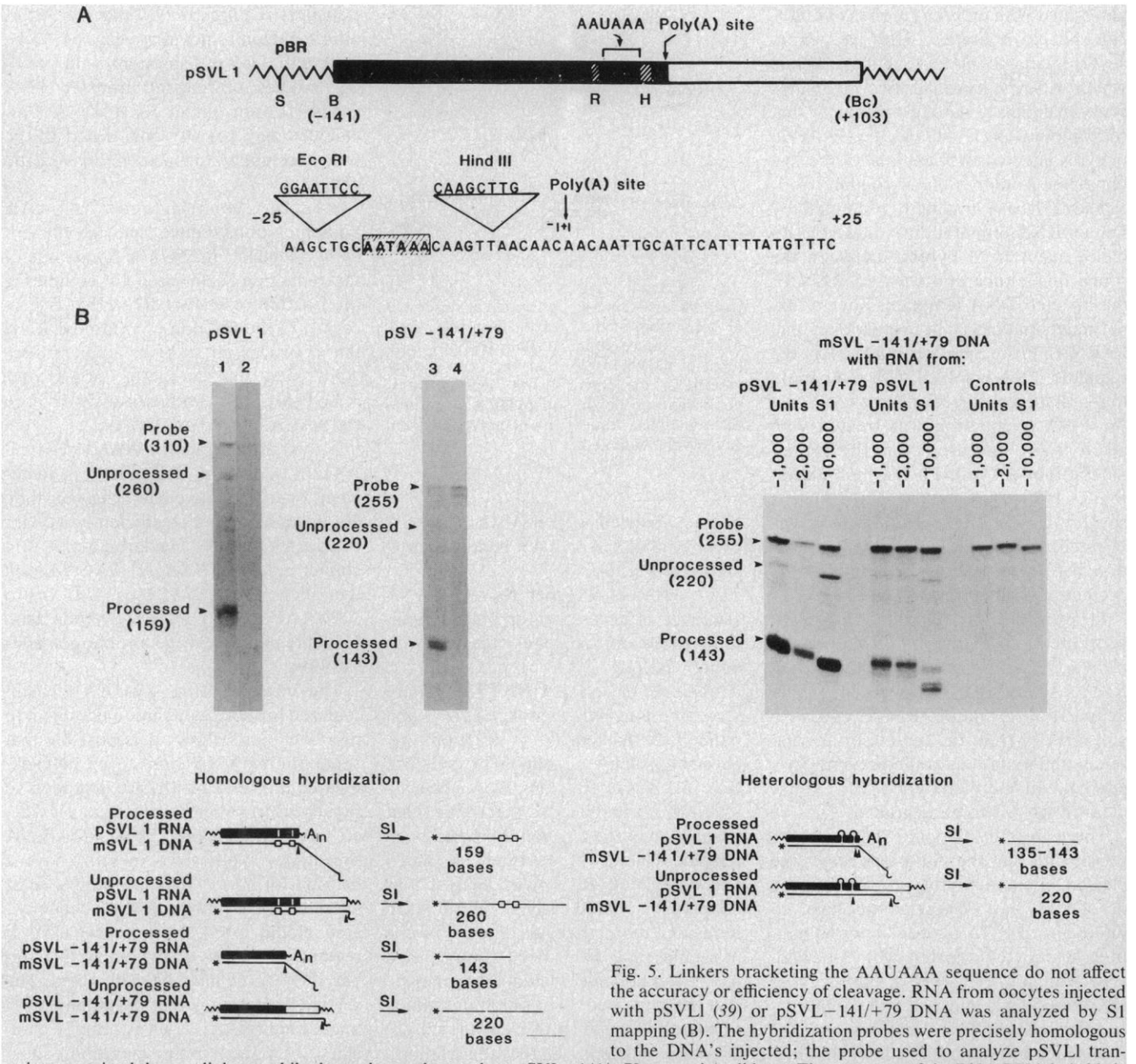

Fig. 5. Linkers bracketing the AAUAAA sequence do not affect the accuracy or efficiency of cleavage. RNA from oocytes injected with pSVL1 (*39*) or pSVL−141/+79 DNA was analyzed by S1 mapping (B). The hybridization probes were precisely homologous to the DNA's injected; the probe used to analyze pSVL1 transcripts contained the two linkers, while the probe used to analyze pSVL−141/+79 transcripts did not. The structure of the RNA-DNA hybrids is diagrammed at the bottom of the figure (thin black line, probe DNA having SV40 sequence; white boxes interrupting the black line, synthetic linkers; the wavy lines indicate a region of M13 present in the probe but missing from the RNA; the RNA's are diagrammed as in Fig. 1). Transcripts of pSVL1 are processed as efficiently as transcripts of pSVL−141/+79 (B, lanes 1 to 4). The pSVL1 RNA protects fragments of 159 and 260 bases (lane 1); a small amount of undigested probe also is visible (310 bases). Similarly, pSVL−141/+79 RNA protects fragments of 143 and 220 bases (lane 3). RNA from uninjected oocytes does not protect either probe from digestion (lanes 2 and 4). Processed pSVL1 RNA protects a fragment 16 bases longer than that protected by pSVL−141/+79 RNA because it contains two 8-bp linkers. The ratio of processed to unprocessed RNA and the amount of RNA per oocyte is the same with both templates. To determine more precisely whether the 3′ end of pSVL1 RNA's was the same as that of pSVL−141/+79 RNA's, the following experiment was performed. (C) The mSVL−141/+79 DNA probe (which does not contain linkers) was hybridized to RNA from oocytes that had been injected with pSVL1 or pSVL−141/+103. The structure of the hybrids between pSVL1 RNA (which contains linkers) and the probe (which does not) is diagrammed. Hybrids were then treated with low S1 concentrations (1000 or 2000 units per milliliter) to avoid cleavage at the end of the RNA-DNA duplex. At low S1 concentrations pSVL1 and pSVL−141/+103 RNA's protect fragments of exactly the same length, whereas RNA from uninjected oocytes do not protect any probe. At a high S1 concentration (10,000 units per milliliter), pSVL1 and pSVL−141/+103 RNA's still each protect an identical 143-base fragment; but, in addition, pSVL1 RNA protects a heterogeneous group of fragments ranging from 135 to 140 bases. This heterogeneity almost certainly results from S1 digestion of a transiently single-stranded structure at the discontinuity in the RNA-DNA duplex.

barely detectable in the exposure shown in Fig. 4. However, the data in Fig. 6 confirm that a low level of accurate cleavage does occur with these templates.

The absolute amounts of RNA in oocytes injected with different templates cannot be determined from Fig. 4, since the hybridizations and S1 analyses were performed with different probes of different specific activities. However, from seven independent experiments (*33*) in which RNA's from all five templates were analyzed in parallel, we know that the amount of SV40-specific RNA in oocytes injected with any one of the five templates is approximately equal.

The 257-base fragment protected by mutant RNA preparations could, in principle, result from hybrids between the probe and either unprocessed RNA or the injected DNA template. Several experiments indicate that protection is due to RNA. Protection is abolished if the template DNA is mixed with α-amanitin (5 μg/ml) before injection (lanes αa), or if the RNA preparation was treated with alkali or ribonuclease before the hybridization (not shown). Furthermore, addition of 100 ng of template DNA (that is, roughly ten times the amount that was injected) to a single hybridization reaction before hybridization results in no protection of the probe.

In principle, these mutations might exert their effect either by preventing the cleavage reaction or by causing mutant RNA's to be rapidly degraded. Rapid degradation of mutant RNA's could result directly from the mutation; alternatively, the mutations could prevent polyadenylation and thereby indirectly cause cleaved RNA's to be degraded.

Three lines of evidence led us to the conclusion that the mutations affect the cleavage reaction, not polyadenylation or stability. (i) Whereas the level of unprocessed RNA plateaus after 60 minutes in oocytes injected with the wild-type template, the level of unprocessed RNA in oocytes injected with the mutant templates continued to increase over the full incubation period. We infer that mutant unprocessed RNA's accumulate because they are not cleaved. (ii) If mutant RNA's were rapidly degraded, then oocytes injected with mutant templates should accumulate less total RNA (processed plus unprocessed) than oocytes injected with the wild-type template. However, to the contrary, oocytes injected with either wild-type or mutant templates accumulate the same amount of RNA (*33*). (iii) The point mutants do not prevent polyadenylation, since the small amount of cleaved RNA found in

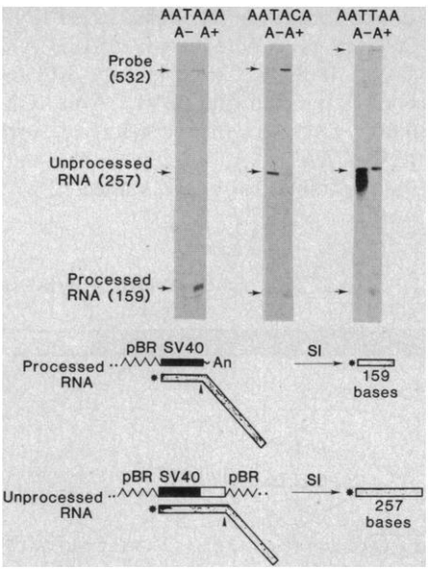

Fig. 6. Polyadenylation of mutant RNA's. S1 nuclease analysis was performed as described at the bottom of the figure. The probe was a 532-base single-stranded Bam HI/Sal I fragment of pMW115 (*26*) that had been 3′ end-labeled at the Bam HI site. One oocytes' equivalent of nonpolyadenylated RNA and three oocytes' equivalent of polyadenylated RNA were analyzed.

oocytes injected with pSVL1:AATACA or with pSVL1:AATTAA is polyadenylated (see Fig. 6).

As a test of whether the AAUAAA sequence is necessary for polyadenylation, RNA from oocytes injected with either pSVL1 (wild type), pSVL1:AATACA or pSVL1:AATTAA were separated into polyadenylated (A+) and nonpolyadenylated (A−) fractions by chromatography on oligo(dT) cellulose (Fig. 6). These fractions were then assayed by S1 mapping to determine what proportion of the cleaved RNA had been polyadenylated. Three times as much polyadenylated RNA was analyzed as nonpolyadenylated RNA. With each template, more than 90 percent of the molecules that have been cleaved also have been polyadenylated. We conclude that the same two point mutations which reduce the efficiency of cleavage do not affect the efficiency of polyadenylation.

Cleavage, polyadenylation, and transport encoded by a local region of the mRNA precursor. Transcripts containing only 220 bases of SV40 sequence are cleaved, polyadenylated, and transported, even though they lack more than 3000 bases present in SV40 pre-mRNA's (Figs. 1 and 2). These three processing steps therefore are independent of initiation at the normal site, of splicing of the SV40 late mRNA introns, and of termination at a specific sequence in the SV40 3′ flanking region.

These data do not bear on whether capping is required for cleavage or polyadenylation. Polymerase II synthesizes the pBR322/SV40 transcripts (Fig. 4), and capping may be coupled to initiation by this polymerase (*34*).

Transport of RNA's that have been cleaved and polyadenylated, but not spliced, may indicate that, in the oocyte, transport requires only a mature 3′ terminus. Intron removal appears not to be essential. This is consistent with previous studies of injected transfer RNA genes (*17*) and intact SV40 DNA (*18*). Oocytes may possess only a part of the transport apparatus present in somatic cells.

AAUAAA point mutants. AAUAAA point mutations reduce cleavage efficiency in oocytes (Fig. 4). The magnitude of the reduction, calculated by comparing the fraction of mutant RNA that is processed to the fraction of wild-type RNA that is processed, is by a factor of 8 for AAUACA, by a factor of 25 for AAUUAA, and by a factor of 50 or more for AACAAA and AAUGAA.

Although cleavage of AAUUAA and AAUACA mRNA precursors is inefficient, once they have been cleaved, they are polyadenylated as efficiently as with AAUAAA (Fig. 6). Similarly, a mutation converting AAUAAA to AAGAAA in the adenovirus E1A gene (*14*) or to AAUAAG in the human α-globulin gene (*34a*) prevents cleavage but not polyadenylation.

The inference that AAUAAA is not required for polyadenylation is subject to two limitations. First, it is possible that some mutant RNA's fail to be polyadenylated but, as a result, are degraded so rapidly as to never be detected. However, nonpolyadenylated SV40 late mRNA precursors synthesized in vitro have a half-life of more than 10 hours after injection (not shown), as does deadenylated globin mRNA (20 hours) (*35*). If mutant transcripts were comparably stable, they would have been detected. The second limitation is that polyadenylation might require AAUAAA yet not be abolished by a single point mutation. Thus, strictly speaking, our results demonstrate that the requirement, if any, for AAUAAA in polyadenylation is less stringent than the requirement for AAUAAA in cleavage.

If the AAUAAA sequence is not directly involved in polyadenylation, then what does encode polyadenylation? Two possibilities exist. Another sequence, as yet unidentified, could be involved. Alternatively, as discussed below, polyadenylation could require the interaction of the cleavage factor with AAUAAA.

For instance, the polyadenylation activity might recognize the cleavage factor bound near an RNA 3′ hydroxyl, or cleavage and polyadenylation activities might be physically associated.

Neither our mutants nor the mutant of Montell *et al.* (*14*) cause inefficient polyadenylation of cleaved RNA's. Similarly, cleaved but nonpolyadenylated transcripts are not detectable in isolated nuclei (*36*). In general, it appears that any cleaved molecule is efficiently polyadenylated. This suggests that polyadenylation rapidly follows cleavage, and that, as Manley has proposed (*4, 36*), the two reactions may be coupled. However, since synthetic pre-mRNA's that are not cleaved but contain AAUAAA are efficiently polyadenylated in vitro (*4*), cleavage in itself must not be required for polyadenylation in vitro.

Although the conservation of the AAUAAA sequence is dramatic, natural variants do occur. The two mutant sequences that were found to produce detectable levels of cleaved RNA are also found in natural genes. From a computer-assisted comparison of the 3′ terminal sequences of 134 mRNA's from vertebrates (*37*), we derived the following consensus sequence:

$$A_{98}A_{91}U_{100}A_{99}A_{99}A_{98}$$

where the subscripts indicate the percentage of mRNA's containing that base. (This consensus sequence excludes histone mRNA's.) The only variations of AAUAAA that were observed were AUUAAA (12 percent), AAUUAA (2 percent), AAUACA (2 percent), AAUAAU (2 percent), AAUAAC (1 percent) and CAUAAA (1 percent). Thus AAUUAA and AAUACA, which direct cleavage in oocytes, albeit inefficiently, are found in natural mRNA precursors.

Although the AAUAAA sequence is necessary for the cleavage of mRNA precursors, it clearly is not sufficient: AAUAAA is found in the protein coding region of mRNA's where it does not direct 3′ end formation (*4, 10*). Thus at least one other feature of the precursor mRNA must be necessary. This additional element could be a simple linear sequence or a secondary structure. The latter possibility seems more likely, both because RNA processing enzymes generally recognize the conformation of RNA precursors (*38, 39*) and because no other sequence near the 3′ end of mRNA's is dramatically conserved.

Whatever signals other than AAUAAA are involved in cleavage must reside in the −141/+79 domain of SV40 late mRNA precursors. Studies analagous to those described should reveal those additional signals and establish whether the processing apparatus recognizes primary or secondary structure.

References and Notes

1. J. E. Darnell, Jr., *Nature (London)* **297**, 365 (1982).
2. J. R. Nevins, *Cell* **28**, 1 (1982).
3. ———, *Annu. Rev. Biochem.* **52**, 441 (1983).
4. J. L. Manley, *Cell* **33**, 595 (1983).
5. Histone mRNA's are exceptional. They are cleaved [C. Birchmeier, D. Schumperli, G. Sconzo, M. L. Birnstiel, *Proc. Natl. Acad. Sci. U.S.A.* **81**, 1057 (1984)] but probably are not polyadenylated (*6*).
6. C. C. Hentschel and M. L. Birnstiel, *Cell* **25**, 301 (1981).
7. J. R. Nevins and J. E. Darnell, Jr., *ibid.* **15**, 1477 (1978).
8. F. W. Alt, A. L. M. Bothwell, S. Knapp, E. Siden, E. Mather, M. Koshland, D. Baltimore, *ibid.* **20**, 293 (1980); J. Rogers, L. Hood, R. Wall, *ibid.*, p. 303.
9. S. G. Amara, V. Jonas, M. G. Rosenfeld, E. S. Ong, R. M. Evans, *Nature (London)* **298**, 240 (1982).
10. M. Fitzgerald and T. Shenk, *Cell* **24**, 251 (1981).
11. J. P. Ford and M. Hsu, *J. Virol.* **28**, 795 (1978).
12. J. Tooze, Ed., *DNA Tumor Viruses* (Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y., ed. 2, 1981).
13. N. J. Proudfoot and G. G. Brownlee, *Nature (London)* **263**, 211 (1976).
14. C. Montell, E. F. Fisher, M. H. Caruthers, A. J. Berk, *ibid.* **305**, 600 (1983).
15. D. A. Melton and R. Cortese, *Cell* **18**, 1165 (1979).
16. T. J. Miller and J. E. Mertz, *Mol. Cell. Biol.* **2**, 1595 (1982).
17. M. P. Wickens and R. A. Laskey, *Genetic Engineering* (Academic Press, London, 1981), vol. 1, p. 103.
18. M. P. Wickens and J. B. Gurdon, *J. Mol. Biol.* **163**, 1 (1983).
19. E. M. DeRobertis and J. E. Mertz, *Cell* **12**, 175 (1977).
20. J. E. Mertz and J. B. Gurdon, *Proc. Natl. Acad. Sci. U.S.A.* **74**, 1502 (1977).
21. M. P. Wickens, S. Woo, B. W. O'Malley, J. B. Gurdon, *Nature (London)* **285**, 628 (1980).
22. The pSVL−141/+802 fragment was constructed by cloning the Bam HI (position 2533) to Hind III (position 3476) fragment of SV40 into the Bam HI and Hind III sites of pBR322. [Numbering according to Tooze (*12*)]. Deletion plasmids were constructed by Bal 31 mutagenesis. Hind III–cleaved pSVL−141/+802 DNA was treated with Bal 31 exonuclease. Hind III linkers were added and the Bam HI/Hind III fragments cloned into pBR322.
23. RNA was prepared as described (*18, 42*).
24. Oocyte culture and micromanipulation has been described (*18, 41*). All injected DNA's were prepared by CsCl buoyant density centrifugation (*43*).
25. S1 mapping has been described (*18, 27*). Msp I pBR322 fragments served as length markers.
26. Single-stranded hybridization probes were prepared by one of two methods. Method 1 (all figures except Fig. 5). A 3′ end-labeled DNA restriction fragment was digested with a restriction enzyme and denatured, and the desired strand was then purified by electrophoresis on a strand-separating gel (*44*). Method 2 (Fig. 5) was developed by D. Leonard Bentley. A DNA sequencing primer is annealed to a M13 DNA carrying the region to be assayed and is extended in the presence of [^{32}P]deoxynucleoside triphosphates. A restriction enzyme is then added that cleaves the DNA only once, beyond the region of interest. DNA is then precipitated and purified by electrophoresis through a sequencing gel (*29, 45*).
27. A. J. Berk and P. A. Sharp, *Cell* **12**, 721 (1977).
28. T. J. Miller, D. L. Stephens, J. E. Mertz, *Mol. Cell. Biol.* **2**, 1581 (1982).
29. M13 mSVL−22/−7:AAUAAA were irradiated with ultraviolet light, as were cells of strain D91 (from S. Brenner, Cambridge University). D91 carries a lac deletion, and an Hfr. Irradiated phage and cells were mixed, and after 10 hours at 37°C, a phage stock was prepared. These phage were then plated onto *Escherichia coli* strain JM101 on plates containing BCIG and IPTG. About 10,000 plaques were screened per plate.
30. F. Sanger, S. Nickleu, A. R. Coulson, *Proc. Natl. Acad. Sci. U.S.A.* **74**, 5463 (1977).
31. J. R. Roth, *J. Bacteriol.* **102**, 467 (1970).
32. AAUGAA transcripts yield a faint band of 140 ± 5 nucleotides, corresponding to the one base mismatch between the probe and the RNA. The RNA protecting this fragment was not polyadenylated. We therefore assumed that the band was artifactual, resulting from S1 cleavage at or near the mismatched base. The alternative possibility is that AAUGAA transcripts were cleaved inefficiently to generate nonpolyadenylated 3′ termini 20 ± 5 bases upstream of the pSVL1 polyadenylation site.
33. We analyzed, in parallel, oocytes injected with each template by S1 mapping. The experiment was repeated seven times, with incubations from 5 to 36 hours. RNA was quantified by microdensitometry. With values of wild-type RNA considered equal to 1.0, the values for mutant templates were for AATACA, 1.0; for AATTAA, 0.7; for AACAAA, 1.2; for AATGAA, 0.8.
34. E. B. Ziff and R. M. Evans, *Cell* **13**, 1463 (1978).

34a. D. R. Hicks *et al.*, *Nature (London)* **306**, 398 (1983).

35. G. Marbaix *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **72**, 3065 (1975).
36. J. L. Manley, P. A. Sharp, M. L. Gefter, *J. Mol. Biol.* **159**, 581 (1982).
37. 3′terminal sequences of all vertebrate mRNA's contained in the NIH library of sequences were compared. AAUAAA sequences more than 30 bases from the 3′ terminus were not considered. Variant AAUAAA sequences were aligned with AAUAAA at the U residue. AAUUAAA was classified as AUUAAA, not AAUUAA.
38. S. Altman, *Cell* **23**, 3 (1981).
39. M. I. Baldi, G. Mattoccia, E. P. Tochini-Valentini, *ibid.* **35**, 109 (1983).
40. The M13 clone used in generating point mutations (mSVL−22/−7) was constructed by transferring the single-copy Eco RI and Hind III fragment of pMW107 to Eco RI/Hind III-cleaved M13 mp8.
41. pSVL1 was constructed as follows. The Bam HI (position 2533) to Bcl I (position 2770) fragment of SV40 was cloned into the Bam HI site of a pBR322 derivative from which the Eco RI and Hind III sites had been deleted, generating plasmid pSVL−141/+103 (insert oriented as in pSVL−141/+802). A Hind III linker was added to the Hpa I site (position 2666) of pSVL−141/+103, generating pMW121. The 775-base-pair Alu I/Alu I fragment of the SV40 insert in pSVL−141/+802 (positions 2651 to 3426) was joined to Eco RI linkers and cleaved with Mbo I. The resulting Eco RI (Alu I)/Mbo I fragment (positions 2651 to 2770) was inserted into Eco RI/Bam HI pBR322 to generate pMW104. Hind III linkers were added to the Hpa I site in pMW104, generating pMW107. The Alu I fragment of pSVL−141/+802, which contains the Bam HI site, was joined to Eco RI linkers and cloned into Eco RI–cleaved pMW107. The Bam HI to Hind III fragment of this plasmid was then transferred to the Bam HI and Hind III sites of a pBR322 lacking an Eco RI site, generating pMW115. The larger Hind III/Pst I fragment of pMW115 was joined to the smaller Hind III/Pst I fragment of pSVL−141/+103 (construction above). This plasmid is pSVL1 (Fig. 5).
42. H. Aviv and P. Leder, *Proc. Natl. Acad. Sci. U.S.A.* **69**, 1408 (1972).
43. R. Radloff, W. Bauer, J. Vinograd, *ibid.* **64**, 1514 (1967).
44. T. Maniatis, E. F. Fritsch, J. Sambrook, *Molecular Cloning* (Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y., 1982).
45. A. M. Maxam and W. Gilbert, *Methods Enzymol.* **65**, 499 (1980).

46. We thank Kevin Struhl, Angus Lamond, Barbara Meyer, and Sydney Brenner for advice during the early stages of this work; Ron Laskey and John Gurdon for research space, advice, encouragement, and support; Ruth Stancombe for technical assistance; and J. Dahlberg, W. Reznikoff, and P. Anderson for comments on the manuscript. Supported by NIH grant R-R01-GM31892-02.

1 June 1984; accepted 4 September 1984