Characterization of Distinct 5′-Terminal Cap Structures of Adenovirus Type 2 Early Messenger Ribonucleic Acid and KB Cell Messenger Ribonucleic Acid†

Shuichi Hashimoto,* Michael H. Pursley, William S. M. Wold, and Maurice Green‡

ABSTRACT: The "early" genes of the adenovirus 2 DNA genome (M_r 23 × 10^6) are organized in four regions (E1–E4)[1] and are transcribed from four major promoters, one in each region, plus possibly two minor promotors in E1, the transforming region. We have previously analyzed the 5′-terminal cap structures of adenovirus 2 (Ad2) early RNA labeled in vivo and purified by DEAE-Sepharose chromatography in 7 M urea, followed by paper electrophoresis. Evidence was obtained for a *minimum* of six distinct caps in Ad2 early mRNA. We have now developed an independent approach to characterize the caps of Ad2 early RNA and of KB cell RNA and have established the validity of this approach to characterize the 5′ termini of capped RNAs. The RNA 5′-terminal caps were labeled with 20–40% efficiency in vitro (after removal of the m^7G by oxidation and β-elimination) by using polynucleotide kinase and [γ-^{32}P]ATP. Ad2-specific RNA was selected by one or two cycles of hybridization with Ad2 DNA. RNA was digested with RNase T2, and −5 (cap 1: $^{32}pN^mpNp$) and −6 (cap 2: $^{32}pN^mpN^mpNp$) charge fractions were isolated by DEAE-Sephadex chromatography in 7 M urea. Each fraction was digested with nuclease P1 and the $^{32}pN^m$ nucleotides identified by 2-D-TLC on cellulose: pA^m, pm^6A^m, pG^m, pU^m, and pN^{*m} were found in Ad2 cap 1 and cap 2 fractions (pN^{*m} probably is an unknown uridylic acid derivative). KB RNA caps had these plus pC^m. Therefore, at least one Ad2 early mRNA has U^m as its 5′-terminal penultimate nucleoside, suggesting that eucaryotic mRNAs may initiate in a pyrimidine residue. In other experiments, the −5 and −6 fractions were resolved into individual cap derivatives by a new procedure, 2-D-TLC on PEI–cellulose. We resolved 10 cap 1 and 11 cap 2 spots in Ad2 RNA; each cap spot does not necessarily represent a different 5′ terminus, because (1) cap 1 spots could be cap 2 spots lacking ribose methylation in the penultimate nucleoside and (2) the methylation of bases (A and probably U) could be partial, which would give rise to different cap spots derived from the same 5′ terminus. To resolve these possibilities, we analyzed each spot by nuclease P1 digestion and 2-D-TLC; the same $^{32}pN^m$ nucleotides were detected as described above. Our results suggest that there are at least eight different cap structures in early Ad2 mRNA. Thus, the minimum number of 5′ termini in early Ad2 mRNA appears to be greater than the number of known early promoters. In analyses of KB cell RNA, we found that all but one cap 1 and two cap 2 structures in Ad2 mRNA were common to KB cell mRNA.

Productive infection of cultured human cells by human adenovirus 2 (Ad2) (Green, 1962; Green et al., 1970) occurs in two major stages of gene expression, early (before viral DNA synthesis begins at 6–7-h postinfection) and late [reviewed by Wold et al. (1978)]. Early and late Ad2 mRNAs are transcribed from viral DNA in the cell nucleus and are transported to cytoplasmic polyribosomes for translation (Thomas & Green, 1966; Green et al., 1970; Parsons et al., 1971; Wall et al., 1972). Early genes are transcribed and viral RNA transcripts are processed mainly or completely by cellular enzymes and mechanisms; thus, the transcription and processing of early RNAs provide a model for analyzing mRNA biogenesis and posttranscriptional regulation in eucaryotic cells (Parsons & Green, 1971).

The early genes are clustered in four regions, designated E1–E4, located at positions 1.5–11 (E1) and 78–86 (E3) on r strand and 62–68 (E2) and 91.5–99 (E4) on l strand [the genome map is 100 units long; r and l refer to rightward and leftward transcription [see Wold et al. (1978)]]. Recent reports using electron microscopic visualization of RNA–DNA hybrids, and S-1 nuclease–exonuclease VII mapping methods, have established that each early region codes a complex variety of partially overlapping "spliced" RNAs (Kitchingman et al., 1978; Chow et al., 1977; Berk & Sharp, 1978; Broker & Chow, 1979). How these mRNAs are generated is not known, but there is fairly good evidence that they may be processed from larger polycistronic nuclear RNA precursors derived from some or all of each early gene region. Two independent approaches, UV mapping (Berk & Sharp, 1977) and nascent RNA mapping (Evans et al., 1977), have indentified four major promotors, one at the 3′ side of each early region. Recent studies suggest that E1 contains one and possibly two additional promotors (Wilson et al., 1979; Sehgal et al., 1979). Ad-specific nuclear RNAs are about 20–25 S (Craig & Raskas, 1976; Bachenheimer, 1977), which is 15% to several-fold larger than cytoplasmic RNAs that range from 9 to 22 S (Craig et al., 1975; Büttner et al., 1976; Spector et al., 1978). Recent studies on E2 RNAs indicate that the primary transcript is a 28S RNA derived from mp 62–75 that is processed via splicing into a mRNA that lacks the intervening sequences at mp 75–69 and 68–66 (Goldenberg & Raskas, 1979).

Analysis of the 5′ termini of early RNA should provide insights into mechanisms of transcription, processing, and translation of both Ad2 early RNA and cellular RNAs. We previously reported (Hashimoto & Green, 1976) that in vivo labeled Ad2 early mRNAs have cap 1 (m^7GpppN^mpNp) and cap 2 ($m^7GpppN^mpN^mpNp$) structures in an average of one cap at the 5′ terminus per mRNA molecule. We recently

† From the Institute for Molecular Virology, Saint Louis University Medical School, St. Louis, Missouri 63110. *Received July 3, 1979.* This work was supported by Grant AI-01725.

‡ Holds Research Career Award 5K5-AI-04739 from the National Institutes of Health.

[1] Recent data suggest that there is a fifth early region expressed in relatively low abundance early after infection (Galos et al., 1979).

0006-2960/80/0419-0294$01.00/0 © 1980 American Chemical Society

found (Hashimoto & Green, 1979) that there are at least six distinct caps in Ad2 early mRNA, by partial separation of each cap 1 and cap 2 structure by one-dimensional DEAE-cellulose paper electrophoresis and analysis of the separated caps. In this communication, we extend our previous data, using an independent approach to label 5′ termini in vitro and to separate on a two-dimensional system and characterize the structures of each terminus. Our results indicate that there are *at least* eight different cap structures in early Ad2 mRNA. Most of these structures were also observed in KB cell mRNA. 5′-Terminal structure studies of Ad2 late mRNA using in vitro 5′-^{32}P-labeling techniques have been recently published (Ziff & Evans, 1978; Lockard et al., 1979).

Materials and Methods

Cell Culture, Virus Infection, and in Vivo Labeling of RNA. Exponentially growing KB cells in suspension were cultured in Eagle minimum essential medium (MEM) containing 5% horse serum. Cells were concentrated to (3–5) × 10^6 cells/mL and infected in MEM with 100 PFU[2]/cell Ad2. After 1-h adsorption, cells were diluted with 10 volumes of MEM containing 5% horse serum and 25 μg of CH per mL. At 2-h postinfection, cells were concentrated 10-fold, suspended in MEM (without horse serum) containing 25 μg of CH per mL, and labeled from 2- to 6-h postinfection with 10–20 μCi of [^{3}H]Urd ([5,6-^{3}H]uridine; 40–50 Ci/mmol; New England Nuclear) per mL. Uninfected KB cells were also labeled with [^{3}H]Urd (5–10 μCi/mL) for 4 h and with L-[*methyl*-^{3}H]-methionine (Hashimoto & Green, 1979).

Cell Fractionation and Isolation of Poly(A+) RNA. Polyribosomal RNA was isolated from the cytoplasmic fraction obtained from Ad2-infected or uninfected KB cells as described (Hashimoto & Green, 1979). Poly(A)-containing [poly(A+)] and poly(A−) RNA fractions were prepared by using a poly(U)–Sepharose column (0.5 × 3 cm). About 100–250 μg (2–5 OD_{260} units) of poly(A+) RNA were obtained from 1 L of Ad2-infected or uninfected cells [(3–5) × 10^5 cells/mL], with [^{3}H]Urd specific activities of (2–6) × 10^7 and (4–10) × 10^6 cpm/mg for infected and uninfected RNA, respectively.

Oxidation and β-Elimination of Poly(A+) RNA. The m^{7}G was removed from the RNA caps by oxidation using sodium periodate and β-elimination using aniline (Fraenkel-Conrat & Steinschneider, 1967; Miura et al., 1974). Poly(A+) RNA (500 μg) from Ad2-infected or uninfected KB cells was incubated in the dark at 0 °C for 1 h in 1 mL of 1 mM sodium periodate and 150 mM sodium acetate (pH 5.3). RNA was precipitated twice with ethanol and dissolved in 2 mL of freshly prepared 300 mM aniline and 10 mM acetic acid, adjusted to pH 5.0 with concentrated HCl. After 3 h at room temperature, RNA was precipitated twice with ethanol and dissolved in 1 mL of 50 mM Tris-HCl (pH 8.5) and 1 mM $MgCl_2$.

Removal of 5′-Terminal Phosphates with Alkaline Phosphatase. The three 5′-terminal phosphates were removed from RNA (200–500 μg) lacking m^{7}G by treatment with 4 μg of calf intestine alkaline phosphatase (Boehringer Mannheim Biochemicals) at 45 °C for 1 h (Lockard & RajBhandary, 1976) in 1 mL of 50 mM Tris-HCl (pH 8.5) and 1 mM $MgCl_2$. The RNA was extracted twice with phenol saturated with 20 mM Tris-HCl (pH 7.5) and precipitated with ethanol. In some experiments, oxidation, β-elimination, and phosphatase treatments were repeated. Under these conditions, the phosphatase treatment produced no small [^{3}H]Urd-labeled RNA fragments detectable by gel filtration on Sephadex G-200.

In Vitro Labeling of RNA 5′ Termini with ^{32}P. The RNAs, free of 5′-terminal m^{7}G and phosphates, were labeled by using polynucleotide kinase and high specific activity (100–500 Ci/mmol) [γ-^{32}P]ATP, prepared as described by Glynn & Chappell (1964). RNA (100–200 μg) dissolved in 0.4 mL of 50 mM Tris-HCl (pH 7.5), 10 mM $MgCl_2$, and 5 mM dithiothreitol was incubated with 10 units of polynucleotide kinase and 5–10 nmol of [γ-^{32}P]ATP at 37 °C for 1 h (Takanami, 1967). The 5′ termini labeled RNA was purified from excess [γ-^{32}P]ATP on Sephadex G-200 (superfine; Pharmacia) in 50 mM Tris-HCl (pH 7.5) and 100 mM NaCl. The 5′-^{32}P-labeled RNA eluting as a sharp peak in the excluded fractions (*minimum* molecular weight 2 × 10^5) was precipitated with ethanol. Incorporation of ^{32}P was approximately 2–4 pmol/μg of RNA. Because of internal nicks in RNA, only about one-fifth of the label was at termini (see Results).

Polynucleotide kinase was purified from *Escherichia coli* A19 (RNase I$^-$) infected with phage T4 *am*N82 × E1140 (DO mutant) (Takanami, 1967). *E. coli* and phage strains were gifts from Dr. Sugiura, National Institute of Genetics, Japan. The kinase preparation produced no smaller fragments, detectable by chromatography on Sephadex G-200, when incubated with [^{3}H]Urd-labeled poly(A+) RNA under conditions used for 5′-terminal labeling. Contamination of polynucleotide kinase and alkaline phosphatase with nuclease was also examined as follows. About 200 μg of tRNA was treated with phosphatase, extracted with phenol, and labeled with [γ-^{32}P]ATP and polynucleotide kinase. No fragmentation of tRNA was observed by chromatography on Sephadex G-50.

Hybridization–Purification of Ad2-Specific RNA. Ad2-specific RNA was purified by hybridization to Ad2 DNA on nitrocellulose filters (Hashimoto & Green, 1976). About 50 μg of 5′-^{32}P-labeled poly(A+) RNA was annealed for 20 h in 50% formamide at 37 °C with 50 μg of Ad2 DNA per filter. Four different control hybridizations were performed with 50-μg Ad2 DNA filters: (1) 50 μg of [^{3}H]Urd-labeled poly(A+) RNA from uninfected cells; (2) 50 μg of methyl-^{3}H-labeled poly(A+) RNA from uninfected cells; (3) 50 μg of [^{3}H]Urd-labeled poly(A+) RNA from uninfected cells mixed with 5–10 μg of unlabeled Ad2 early mRNA; (4) 50 μg of 5′-^{32}P-, [^{3}H]Urd-labeled poly(A+) RNA from uninfected cells.

RNase T2 Digestion and DEAE-Sephadex Column Chromatography. 5′-^{32}P-Labeled Ad2 early mRNA or KB cell mRNA plus 100 μg of carrier yeast tRNA were digested with 2 units of RNase T2 (Sankyo Co., Ltd.) in 0.5 mL of 50 mM sodium acetate (pH 4.5) at 37 °C for 15 h. The digests were chromatographed on a column of DEAE-Sephadex A-25 (0.4 × 30 cm) with oligonucleotide markers in 7 M urea–30 mM Tris-HCl (pH 7.6) (Hashimoto & Green, 1976). Elution was performed with a linear gradient of 0–0.4 M NaCl (70 mL each) in the 7 M urea buffer. The A_{260} of each fraction (1.2 mL) was measured, and the ^{3}H and ^{32}P radioactivity in 20–100-μL aliquots was determined. ^{32}pNp (−4 charge), ^{32}pN^{m}pNp (−5 charge), and ^{32}pN^{m}pN^{m}pNp (−6 charge) fractions were pooled and desalted by using a DEAE-Sephadex column (0.3 × 2.5 cm) (Hashimoto & Green, 1976).

Nuclease P1 Digestion and 2-D-TLC. The desalted and dried −4, −5, and −6 charge fractions were incubated in 10 μL of 20 mM sodium acetate (pH 5.3) containing 2 μg of nuclease P1 (Boehringer Mannheim Biochemicals) for 2 h at 37 °C.

[2] Abbreviations used: PFU, plaque-forming units; CH, cycloheximide; Me-^{3}H, L-[*methyl*-^{3}H]methionine; 2-D-TLC, two-dimensional thin-layer chromatography; other abbreviations are those recommended by the IUPAC-IUB.

Table I: Selection of 5′-^{32}P-Labeled Poly(A+) mRNA by Hybridization to Ad2 DNA

expt no.	labeled RNA	input		RNA hybridized		
		^{32}P (cpm)	^{3}H (cpm)	^{32}P (cpm)	^{3}H (cpm)	^{3}H (%)
1	poly(A+) (^{32}P, [^{3}H]Urd)[a] Ad2-infected KB cells	1.67×10^7	2.65×10^6	3.20×10^5	1.15×10^5	4.4 (3–6)[e]
2	poly(A+) (Me-^{3}H)[b] uninfected KB cells		4.37×10^4		60	0.14 (0.05–0.2)
3	poly(A+) ([^{3}H]Urd)[b] uninfected KB cells		1.51×10^5		80	0.05 (0.04–0.2)
4	poly(A+) ([^{3}H]Urd)[c] uninfected KB cells		1.51×10^5		400	0.26 (0.2–0.3)
5	Poly(A+) (^{32}P, [^{3}H]Urd)[d] uninfected KB cells	1.74×10^7	1.90×10^5	6.96×10^3	120	0.06 (0.05–0.2)

[a] Poly(A+) RNA from KB cells labeled with [^{3}H]Urd from 2–6-h postinfection with Ad2 in the presence of CH was labeled with [γ-^{32}P]-ATP after decapping as described under Materials and Methods. About 30–50 μg of the 5′-^{32}P-, [^{3}H]Urd-labeled RNA was annealed to Ad2 DNA (50 μg) immobilized on nitrocellulose filters in 1 mL of hybridization buffer at 37 °C for 20 h. [b] About 30–50 μg of RNA labeled with L-[*methyl*-^{3}H]methionine (13 Ci/mmol) was annealed to Ad2 DNA as described above. [c] Same as in experiment 3 except that hybridization was performed in the presence of unlabeled poly(A+) Ad2 mRNA (see Materials and Methods). [d] About 30–50 μg of 5′-^{32}P-, [^{3}H]Urd-labeled poly(A+) RNA was annealed to Ad2 DNA as described above. [e] The values in parentheses are the range in percent hybridized in three or four experiments.

The digest was mixed with pN (pA, pG, pU, pC) and/or pN^m markers (pA^m, pm^{6}A^m, pC^m, pU^m, pG^m), approximately 0.2 OD_{260} unit of each (P-L Biochemicals), and applied to a cellulose TLC plate (EM Laboratories, Inc.; 10 × 10 cm). 2-D-TLC was done by using isobutyric acid–0.5 N NH_4OH (5:3) for the first dimension, and 2-propanol–concentrated HCl–H_2O (70:15:15) for the second dimension (Nishimura, 1972). The TLC plates were exposed to X-ray film (Kodak; XR-5) for 10–20 h using intensifying screens. The marker spots detected with a UV lamp were eluted in 0.01 N HCl and identified by their UV spectra at pH 2 (Hashimoto & Muramatsu, 1973).

Separation of Different pN^{m}pNp and Different pN^{m}pN^{m}pNp Using 2-D-TLC on PEI–Cellulose and Identification of the 5′-Terminal Nucleotide of Each Cap. The −5 (pN^{m}pNp) and −6 (pN^{m}pN^{m}pNp) fractions were dissolved in 10 μL of water, applied to a poly(ethylenimine) (PEI) TLC plate (20 × 20 cm; Macherey-Nagel Co.), and separated by 2-D-TLC using as solvents 1.4 M lithium formate and 7 M urea (pH 3.5) for the first dimension and 0.8 M lithium chloride, 7 M urea, and 20 mM Tris-HCl (pH 8.0) for the second dimension (Mirzabehov & Griffin, 1972). The TLC plates were exposed to X-ray film for 1–2 days with intensifying screens. The spots identified by autoradiography were scraped into tubes, and the oligonucleotides were eluted with 2 M triethylamine bicarbonate. After removal of triethylamine bicarbonate by repeated evaporation, the sample was incubated in 10 μL of 20 mM sodium acetate (pH 5.3) containing 2 μg of nuclease P1 and 20 μg of carrier tRNA for 2 h at 37 °C, and the nucleotides were resolved by cellulose 2-D-TLC with pN^m markers, as described above. The [5′-^{32}P]nucleotides (^{32}pN^m) were identified by autoradiography and the pN^m markers with a UV lamp.

Results

Validity of the Approach to Label RNA 5′ Termini in Vitro. Before presenting the analyses of in vitro labeled 5′ termini of Ad2 early RNA and cellular RNA, we discuss control experiments to determine whether (1) all 5′ termini were labeled, (2) internal sites were labeled, (3) labeling altered the composition of the 5′ termini, and (4) Ad2-specific RNA was free of cellular RNA.

From 100 to 250 μg of poly(A+) RNA was obtained per L of KB cells [(3–5) × 10^5 cells/mL]. During in vitro labeling, 2–4 pmol of ^{32}P was incorporated per μg of RNA. Labeled RNA was routinely assayed to estimate how much ^{32}P was incorporated into (1) bonafide 5′ termini and (2) 5′ termini resulting from "nicks" in the RNA. This was achieved by digesting labeled RNA with nuclease P1 and analyzing the resulting nucleotides by 2-D-TLC and autoradiography. With bona fide termini, ^{32}P counts should be in 2′-O-methylated nucleotides, whereas with termini resulting from internal nicks the counts should be in unmethylated nucleotides. This analysis indicated that 0.4–0.8 pmol of ^{32}P was incorporated into bona fide 5′ termini per μg of Ad2 RNA, and 0.2–0.6 pmol was incorporated per μg of cellular RNA. Thus, with Ad2 early RNA, one of five molecules of ^{32}P incorporated was into bona fide termini and four of five were incorporated into nicked termini.

Given that 0.4–0.8 pmol of ^{32}P was incorporated into bona fide 5′ termini per μg of Ad2 RNA, the extent of labeling of 5′ termini can be estimated. Ad2-specific early RNA ranges from 12 to 22 S (polyacrylamide gel electrophoresis) and has an average size of 1500 nucleotides. Therefore, 1 μg of Ad2 early RNA is about 2 pmol, and the efficiency of labeling of 5′ termini was 20–40%.

After labeling, we purified Ad2-specific RNA by hybridization to Ad2 DNA on DNA filters. Four different control experiments were done to estimate the level of contamination of Ad2 RNA by cellular RNA. As summarized in Table I, about 5% of input [^{3}H]Urd-labeled RNA from Ad2-infected KB cells hybridized to Ad2 DNA. In contrast, less than 0.2% of [^{3}H]Urd-labeled (or methyl-^{3}H-labeled) RNA from uninfected cells was retained, presumably nonspecifically, by the filters. As another control, [^{3}H]Urd-labeled RNA from uninfected cells was annealed to Ad2 DNA in the presence of unlabeled Ad2 early mRNA; again 0.2–0.3% of labeled RNA was retained on filters. In other control experiments, less than 0.2% of [^{3}H]Urd and ^{32}P double-labeled RNA from uninfected cells was retained. These data indicate that cellular RNA and methylated components of cellular RNA represent at most 2–5% of the Ad2-specific RNA. This level of contamination does not affect our analysis of Ad2 early RNA, a conclusion verified by analysis of Ad2 RNA purified by two cycles of hybridization.

Purification and Characterization of 5′ Termini of Ad2 Early mRNA and Cellular mRNA. In vitro ^{32}P-labeled Ad2-specific mRNA and KB cell mRNA were digested with RNase T2 and then chromatographed on DEAE-Sephadex in 7 M urea and 30 mM Tris-HCl (pH 7.6). This column should separate in vitro labeled cap 1 (^{32}pN^{m}pNp; −5 charge) and cap 2 (^{32}pN^{m}pN^{m}pNp; −6 charge) moieties, as well as ^{32}pNp (−4 charge) derived from nicked molecules. The elution profile of a typical RNase T2 digest of Ad2 early RNA is shown in Figure 1. Similar results were obtained with KB cell RNA. The column clearly resolved ^{32}P-labeled fractions eluting with charges of −4, −5, and −6. Rechromatography of the −5 and −6 charged fractions indicated that each fraction was essentially pure. The pN^{m}pNp (−5) and pN^{m}pN^{m}pNp (−6) fractions eluted slightly sooner than the markers

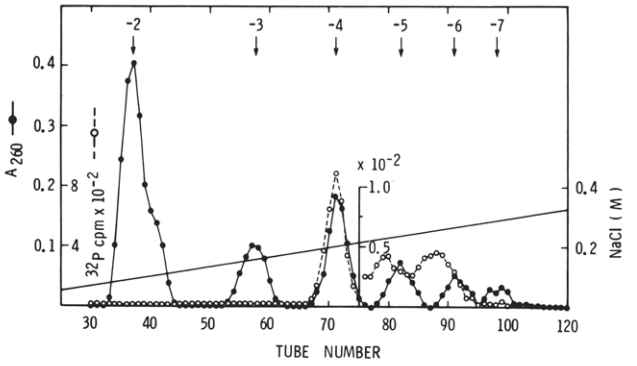

FIGURE 1: Chromatography of RNase T2 digest of 5′-^{32}P-labeled Ad2 early mRNA on a DEAE-Sephadex column in 7 M urea at pH 7.5. The enzyme digest was cochromatographed with 5 OD_{260} units of pancreatic RNase digest of yeast tRNA to provide oligonucleotide markers as described under Materials and Methods. Fractions 77–82 (–5 charge fraction) and 85–92 (–6 charge fraction) were pooled and used for subsequent analysis. In some experiments, the –4 charge fraction was also analyzed.

[NpNpNpNp (–5) and NpNpNpNpNp (–6)], as noted previously by Takanami (1967). Each of the three fractions was pooled, digested with nuclease P1, and mixed with nucleotide markers, and the ^{32}P-labeled nucleotides were identified by 2-D-TLC and autoradiography. When the –4 charge fraction from Figure 1 was analyzed in this manner, only unmethylated nucleotides (pA, pC, pG, pU) were observed (data not shown). Plates 1 and 3 of Figure 2 show the –5 charge fraction of Ad2 and KB cell RNA, respectively; plates 2 and 4 show the –6 charge fraction of Ad2 and KB cell RNA, respectively. The dotted circles indicate the pN and pN^m markers identified by UV absorbance. As expected, only 2′-O-methylated nucleotides were labeled with ^{32}P. Both –5 and –6 charged fractions of Ad2 early mRNA contained ^{32}P-labeled pA^m, pm^6A^m, pU^m, and pG^m. KB cell mRNA contained these nucleotides plus pC^m. With both Ad2 and KB cell mRNA, pA^m and pm^6A^m were the most abundant labeled nucleotides.

Other than these nucleotides, one or two unusual spots (referred to as pN^{*m}) that did not comigrate with markers were reproducibly observed. We previously detected an unidentified 2′-*O*-methyl nucleoside in cap structures of both Ad2 and KB cell mRNA using in vivo methyl-^{3}H-, ^{32}P-labeled RNA (Hashimoto & Green, 1979). The unknown nucleoside migrated between U^m and A^m by paper chromatography in 1-butanol–0.8 M boric acid–concentrated ammonium hydroxide (Hashimoto & Green, 1979) and was presumed to be a derivative of adenine or uridine (Hashimoto & Green, 1979). As seen in Figure 2, an unusual nucleotide spot was always observed at the left side of pG. This spot probably represents the unidentified nucleoside previously noted that migrated between U^m and A^m by paper chromatography, because all other nucleotides identified by 2-D-TLC correspond to known nucleosides previously identified by paper chromatography (Hashimoto & Green, 1979).

In some experiments, the unusual spot separated into two spots by 2-D-TLC (see plate 1 of Figure 2). This spot may

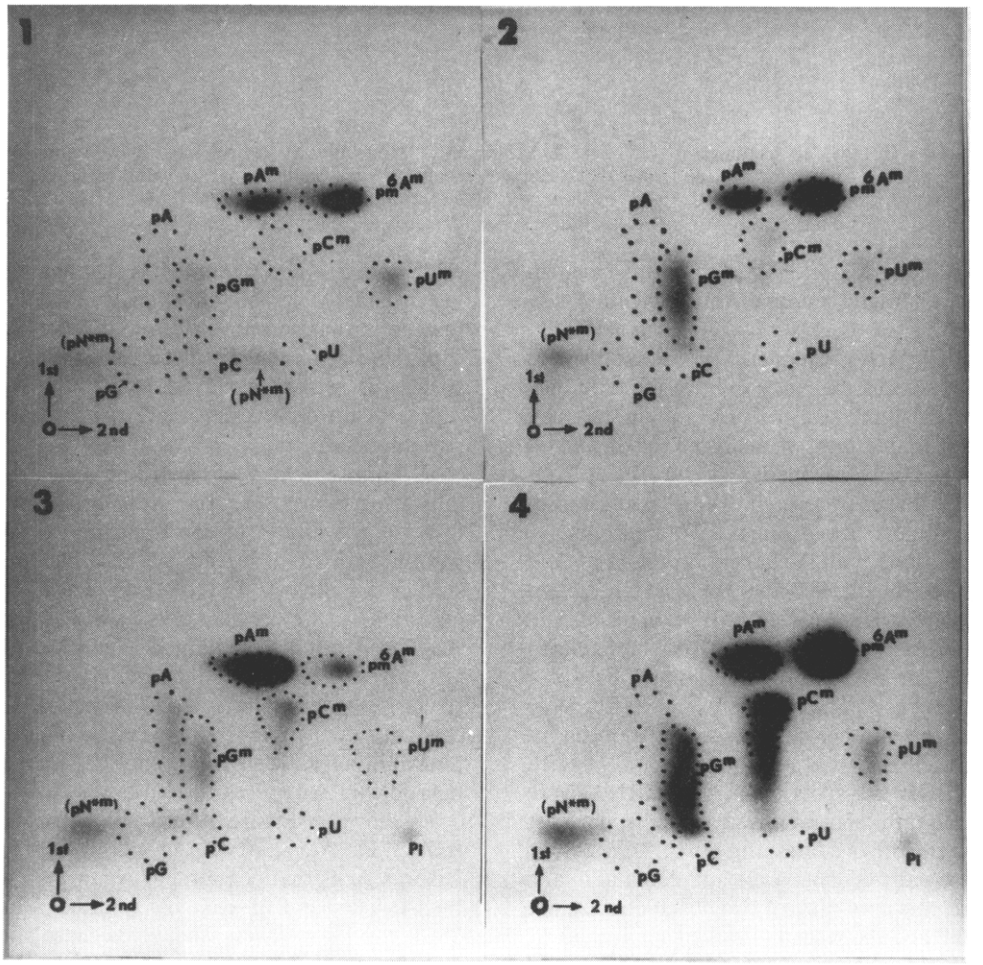

FIGURE 2: Two-dimensional thin-layer chromatography of nuclease P1 digests of the RNase T2 stable oligonucleotides derived from 5′ termini of mRNA. Each enzyme digest was cochromatographed with pN^m and/or pN markers (0.1–0.3 OD_{260} unit each) as described under Materials and Methods. Plates 1 and 2 indicate the digests of –5 and –6 charge fractions of Ad2 early mRNA, and plates 3 and 4 indicate –5 and –6 fractions of KB cell mRNA. Dotted circles indicate markers detected by a UV lamp.

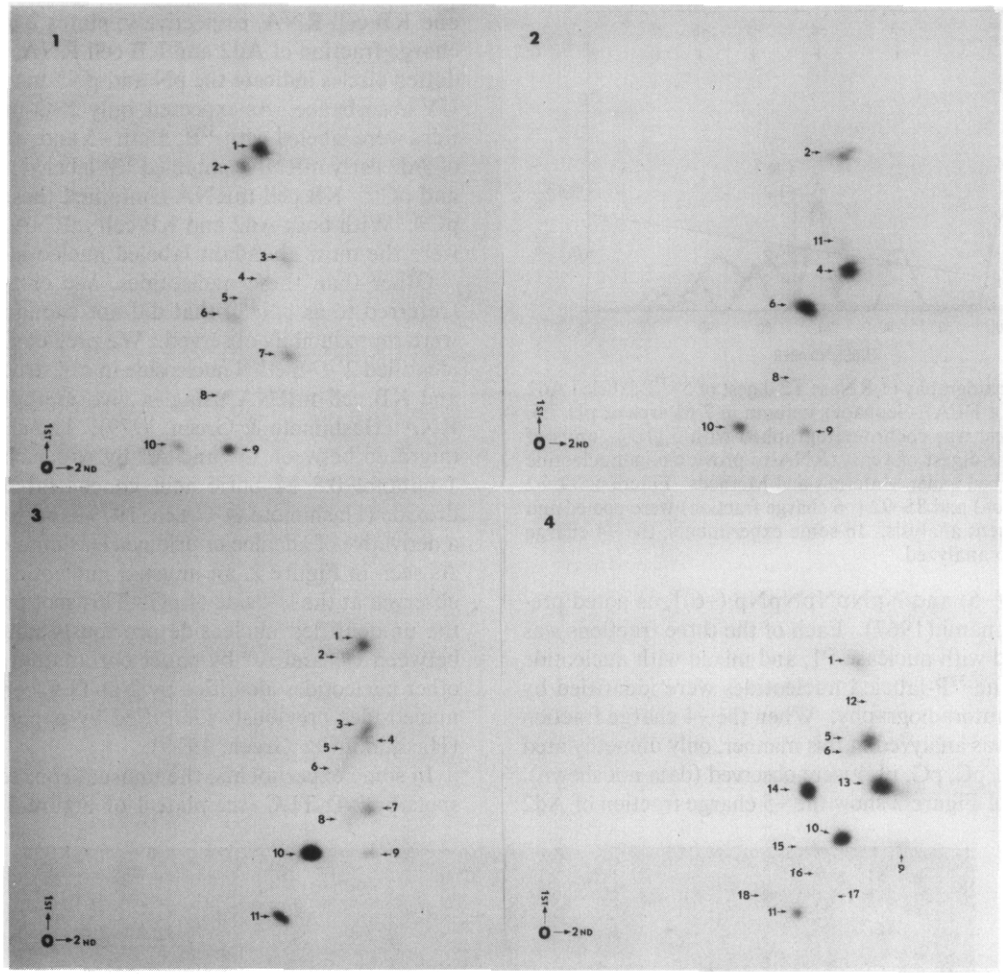

FIGURE 3: Autoradiograms of two-dimensional PEI–cellulose thin-layer chromatography of RNase T2 stable −5 and −6 fractions. The −5 (pN^mpNp) and −6 (pN^mpN^mpNp) fractions eluted from DEAE-Sephadex column were pooled separately and fractionated by PEI–cellulose thin-layer chromatography as described under Materials and Methods. The spot numbers correspond to those in Table II. Plate 1: −5 charge, Ad2 early mRNA. Plate 2: −5 charge, KB cell mRNA. Plate 3: −6 charge, Ad2 early mRNA. Plate 4: −6 charge, KB cell mRNA.

represent an unstable nucleotide, as observed for some modified nucleotides. Further characterization of this nucleotide is in progress.

The experiments described above identified the methylated constituents in mRNA caps, but they did not indicate how many distinct caps exist in Ad2 mRNA and in KB cell mRNA. To address this question, we analyzed the −5 and −6 charge fractions from DEAE-Sephadex (Figure 1) by 2-D-TLC on PEI–cellulose to separate each cap. The autoradiograms are shown in Figure 3 and illustrative diagrams are given in Figure 4. Plates 1 and 2 are the −5 fraction of Ad2 and cell RNA, respectively; plates 3 and 4 are the −6 fraction of Ad2 and KB cell RNA, respectively. All the spots of Ad2 mRNA and the major (dark) spots of cell mRNA were eluted and digested with nuclease P1, and the 5′-terminal nucleotide was identified by 2-D-TLC. The results in Table II provide a means to estimate the number of caps and in addition indicate sequence similarities of cap structures of viral and cell mRNA (Figure 4). Several spots from Ad2 mRNA were observed as closely migrating but well-resolved pairs, one spot containing pA^m and the other containing pm^6A^m. With cap 1 of Ad2 RNA, these pairs of spots are 1 and 2, 3 and 4, and 5 and 6 [spot 6 represents a mixture of two different caps, one (90%) containing pG^m and the second (10%) containing pA^m]. With cap 2 of Ad2 RNA, the paired spots are 1 and 2, 3 and 4, 5 and 6, and 7 and 8. The pairs are indicated by braces in Table II to suggest that they may represent the same cap sequences differing only in the extent of methylation. The evidence for this interpretation is that pm^6A^m migrates slightly faster than pA^m in both directions on PEI–cellulose. Assuming that pairs that contain pA^m and pm^6A^m represent the same cap, there are eight distinct cap 1 and eight distinct cap 2 structures in Ad2 early mRNA. However, if these pairs represent different caps, there are 11 cap 1 and 12 cap 2 structures.

Of interest, spot 7 of cap 1 and spot 9 of cap 2 had pU^m, indicating that at least one Ad2 early mRNA initiates with U^m. Spots 9 and 10 of cap 1 and spot 11 of cap 2 contained the unidentified nucleotide, pN^{*m}, which migrated to the left side of pG on 2-D-TLC (see Figure 2) after nuclease P1 digestion.

KB cell mRNA contained 7 cap 1 and 13 or 14 cap 2 structures (Table II), although only 4 or 5 cap 1 and cap 2 structures were major spots. Theoretically, 16 cap 1 and 64 cap 2 different sequences may exist, suggesting that the major 5′ termini of KB cell RNA may be limited to specific combinations of sequences.

As shown in Figure 2, unfractionated cap 1 structures of KB cell RNA contained pA^m, pG^m, pm^6A^m, pC^m, and pN^{*m}. Distinct spots that contain pm^6A^m or pC^m were not observed by 2-D-TLC on PEI–cellulose (plate 2 in Figure 3; Table II) of the cap 1 structures, probably because of the heterogeneous distribution of these structures. KB cell RNA differed from Ad2 RNA in that cap 1 had little pm^6A^m whereas cap 2 had large quantities of pm^6A^m (Figure 2). KB cell RNA also contained pU^m and pC^m, indicating that some cellular mRNAs

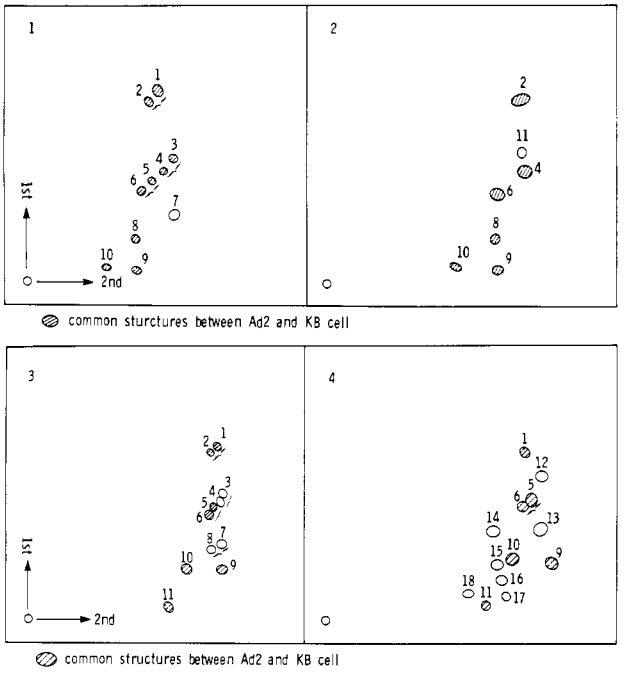

FIGURE 4: Diagram of Figure 3. The shadowed circles indicate the structures that were observed both in Ad2 early mRNA and in KB cell mRNA as determined by location of the spots and analysis of the 5′-terminal nucleotide (see the text). Plate 1: −5 charge, Ad2 early mRNA. Plate 2: −5 charge, KB cell mRNA. Plate 3: −6 charge, Ad2 early mRNA. Plate 4: −6 charge, KB cell mRNA.

Table II: Analysis of the 5′ Terminal Nucleotide in the Separated Cap Structures[a]

cap 1 (pN^mpNp)[c]			cap 2 (pN^mpN^mpNp)[c]		
spot[b] no.	Ad 2	KB cell	spot[b] no.	Ad 2	KB cell
1	pm^6A^m } [d]		1	pm^6A^m }	pm^6A^m
2	pA^m }	pA^m	2	pA^m }	–
3	pm^6A^m }	–	3	pm^6A^m }	–
4	pA^m }	pA^m	4	pA^m }	–
5	pm^6A^m }	–	5	pm^6A^m }	pm^6A^m
6[e]	pA^m } + pG^m	pA^m	6	pA^m }	NA^i
7	pU^m	–	7	pm^6A^m }	–
8	pG^m	pG^m	8	pA^m }	–
9[g]	pN^*m	pN^*m	9	pU^m	pU^m
10[g]	pN^*m	pN^*m	10[f]	pm^6A^m + pG^m	pm^6A^m
11	–[h]	pA^m	11[g]	pN^*m	pN^*m
			12	–	pC^m
			13	–	pC^m
			14	–	pA^m
			15	–	pG^m
			16	–	pG^m + pm^6A^m
			17	–	pG^m
			18	–	NA

[a] Separated cap 1 (pN^mpNp) and cap 2 (pN^mpN^mpNp) structures shown in Figure 3 were eluted, digested with nuclease P1, and analyzed on 2-D-TLC. [b] Spot numbers correspond to those in Figure 3. [c] Autoradiographs are shown in Figure 3. [d] The two spots shown by the braces probably contain the same sequence (see the text). [e] There are two different caps in this spot, one cap contributing 90% of the radioactivity of the spot as $^{32}pG^m$ and the other cap contributing 10% as $^{32}pA^m$. [f] There are two different caps in this spot, one cap contributing 70% of the radioactivity of the spot as $^{32}pG^m$ and the other cap contributing 30% as $^{32}pm^6A^m$. [g] After nuclease P1 digestion, the radioactivity was observed at the left side of pG on 2-D-TLC (see Figure 2). [h] The sequences indicated by the dash were not observed on the autoradiographs. [i] The spots indicated by NA were not analyzed.

initiate with pyrimidines. Ad2 mRNA contained one cap 1 structure (spot 7) and at least two cap 2 structures (spots 3 and 4 and spots 7 and 8) that were not observed in KB cell RNA (see Figure 4). This raises the possibility that the processing mechanisms of some Ad2 mRNAs may differ from that of KB cell mRNA.

Discussion

5′-Terminal Labeling of mRNA. As discussed under Results, although 2–4 pmol of ^{32}P was incorporated per μg of Ad2 early RNA, only 0.4–0.8 pmol was incorporated into bona fide 5′ termini. Each RNA molecule contained on the average of four nicks at which ^{32}P was incorporated into ^{32}pNp moieties if the termini were labeled in the same extent with bona fide termini. We do not know when these nicks became introduced into the RNA, but it was probably during the labeling procedure. Other workers have also observed degradation of single-stranded RNA during in vitro labeling (Miura et al., 1974; Lockard & RajBhandary, 1976; Ziff & Evans, 1978; Lockard et al., 1979).

Number of 5′ Termini in Ad2 Early mRNA. Nucleoside analyses of partially resolved cap 1 and cap 2 structures of in vivo labeled Ad2 early mRNA suggested a minimum of six cap 1 and seven cap 2 sequences (Hashimoto & Green, 1979). By the more sensitive in vitro approach described here, cap 1 and cap 2 fractions of Ad2 early mRNA were resolved into 10 and 11 spots, and the 5′-terminal nucleotide of each spot was identifed by nuclease P1 digestion and 2-D-TLC. Since spot 6 of cap 1 and spot 10 of cap 2 were mixtures of two sequences, a total of 11 cap 1 structures and 12 cap 2 structures were identified. However, each cap does not necessarily represent a distinct mRNA class for the following reasons: (1) cap 1 structures could represent cap 2 structures lacking methylation of the penultimate nucleoside, as reported for several purified mRNAs (Lockard & RajBhandary, 1976; Gelinas & Roberts, 1977); (2) our preliminary analyses detected several different RNase T1 generated 5′-terminal oligonucleotides derived from Ad2 early mRNA, each having both pA^m and pm^6A^m, suggesting that base methylation of A^m to m^6A^m may be partial (partial methylation of A^m to m^6A^m could result in cap doublets as observed in Figure 3); (3) since N^{*m} is probably a modified product of U^m, spots containing N^{*m} and U^m may represent the same cap. On the basis of R_f values on 2-D-TLC, pN^{*m} may be 5-(carboxymethylaminomethyl)uridine which has been detected in $tRNA^{Gly}$ from *Bacillus subtilis* (Murao & Ishikura, 1978). Based upon the above considerations, we conclude that there are a *minimum* of eight major 5′ termini in Ad2 early RNA. As indicated by the braces in Table II, several spots may represent the same cap that differs in the base methylation of adenine. Thus, with cap 1 structures, spots 1 and 2 may be one cap, spots 3 and 4 may be a second cap, spots 5 and 6 may be a third cap (containing pA^m and pm^6A^m), spot 6 that contains pG^m may be a fourth cap, spot 7 may be a fifth cap, spot 8 may be a sixth cap, spot 9 may be a seventh cap, and spot 10 may be an eighth cap.

Our data are relevant to how Ad2 early mRNAs are generated and to the structural relationships between the mRNAs. Electron microscopic studies have identified as many as 20 different spliced early mRNAs although some species are minor. Our data suggest that some of these mRNAs may have common 5′ termini. Several lines of evidence indicate that each early gene region may be transcribed into a large nuclear RNA precursor molecule that is processed into individual mRNAs and that is initiated at a major promoter immediately to the 3′ side of the gene region (Craig & Raskas, 1976; Goldenberg & Raskas, 1979). E1 apparently also contains second or third weaker promoters (Wilson et al., 1979; Sehgal et al., 1979). The majority of late Ad2 mRNAs contain a

5′-terminal leader sequence coded at position 16.4, near the promoter. If the synthesis of Ad2 early mRNA is similar to that of Ad2 late mRNA, then the early mRNAs should contain 5′-terminal leader sequences coded near the promoter. This model predicts the existence of four or five 5′ termini in Ad2 early mRNA. Since we observe *at least* eight termini, our data argue against this model in its most simple form. If promoter–proximal leaders exist in early mRNA, they may be heterogeneous at the immediate 5′ termini. An interesting possibility is that some of the 5′ termini that we detect, especially those that begin with U^m or N^{*m}, represent mRNAs that were derived by internal cleavage of the putative nuclear RNA precursor.

Existence of a Pyrimidine Nucleoside (U^m) at the 5′ Termini of Ad2 Early mRNA. The 5′ termini of more than 10 purified eucaryotic mRNA species have been characterized, and all were shown to have purine nucleosides as the penultimate nucleoside in the cap structures. In our analysis of cap structures from total cell mRNA, C^m and U^m were found as the penultimate nucleoside as indicated in Figure 2 and reported previously (Hashimoto & Green, 1979). Other workers have observed U^m and C^m in total unfractionated mRNAs of human and mouse cells (Furuichi et al., 1975; Cory & Adams, 1975; Wei et al., 1975; Perry et al., 1975). We show here that at least one Ad2 early mRNA has U^m as its 5′-terminal penultimate nucleoside, supporting the possibility that eucaryotic mRNAs may initiate in a pyrimidine residue.

The 5′ Termini of Cellular mRNA. We have identified six different 2′-*O*-methyl nucleosides (A^m, m^6A^m, C^m, U^m, G^m, and N^{*m}) as the penultimate base in 5′-terminal cap structures of KB cell mRNA. As discussed earlier, A^m and U^m could be undermethylated precursors to m^6A^m and N^{*m}, respectively. Assuming this to be the case, then 16 (4 × 4) different cap 1 structures (pN^mpNp) and 64 (4 × 4 × 4) different cap 2 structures (pN^mpN^mpNp) are theoretically possible. If the sequences of the 5′ termini can be random, then each of the 16 and 64 cap 1 and cap 2 structures probably should have appeared as spots with equal intensity on 2-D-TLC autoradiograms. In fact, we observed 7 cap 1 and 13 or 14 cap 2 structures, with 4 or 5 cap 1 and cap 2 spots being much more predominant than the others. Therefore, our data suggest a degree of specificity, either by transcription–initiation or by RNA processing, to the generation of cellular mRNA 5′ termini.

Acknowledgments

We thank Dr. H. Thornton for assistance with cell culture.

References

Bachenheimer, S. L. (1977) *J. Virol. 22*, 577–582.
Berk, A. J., & Sharp, P. A. (1977) *Cell 12*, 45–55.
Berk, A. J., & Sharp, P. A. (1978) *Cell 14*, 695–711.
Broker, T. R., & Chow, L. T. (1979) *ICN-UCLA Symp. Mol. Cell. Biol. 14*, 611–637.
Büttner, W., Veres-Molnar, Z., & Green, M. (1976) *J. Mol. Biol. 107*, 93–114.
Chow, L. T., Gelinas, R. E., Broker, T. R., & Roberts, R. J. (1977) *Cell 12*, 1–8.
Cory, S., & Adams, J. M. (1975) *J. Mol. Biol. 99*, 519–547.
Craig, E. A., & Raskas, H. J. (1976) *Cell 8*, 205–213.
Craig, E. A., McGrogan, M., Mulder, C., & Raskas, H. J. (1975) *J. Virol. 16*, 905–912.
Evans, R. M., Fraser, N. W., Ziff, E., Weber, J., Wilson, M., & Darnell, J. E. (1977) *Cell 12*, 733–739.
Fraenkel-Conrat, H., & Steinschneider, A. (1967) *Methods Enzymol. 12B*, 143–246.
Furuichi, Y., Morgan, M., Shatkin, J., Jelinek, W., Salditt-Georgieff, M., & Darnell, J. E. (1975) *Proc. Natl. Acad. Sci. U.S.A. 72*, 1904–1908.
Galos, G. S., Williams, J., Binger, M.-H., & Flint, S. V. (1979) *Cell 17*, 945–956.
Gelinas, R. E., & Roberts, R. J. (1977) *Cell 11*, 533–544.
Glynn, I. M., & Chappell, J. B. (1964) *Biochem. J. 90*, 147–149.
Goldenberg, C. J., & Raskas, H. J. (1979) *Cell 16*, 131–138.
Green, M. (1962) *Cold Spring Harbor Symp. Quant. Biol. 27*, 219–235.
Green, M., Parsons, J. T., Pina, M., Fujinaga, K., Caffier, H., & Landgraf-Leurs, I. (1970) *Cold Spring Harbor Symp. Quant. Biol. 35*, 803–818.
Hashimoto, S., & Muramatsu, M. (1973) *Eur. J. Biochem. 33*, 446–458.
Hashimoto, S., & Green, M. (1976) *J. Virol. 20*, 425–435.
Hashimoto, S., & Green, M. (1979) *Virology 94*, 254–272.
Kitchingman, G. R., Lai, S. P., & Westphal, H. (1978) *Proc. Natl. Acad. Sci. U.S.A. 74*, 4392–4395.
Lockard, R. E., & RajBhandary, U. L. (1976) *Cell 9*, 747–760.
Lockard, R. E., Berget, S. M., RajBhandary, U. L., & Sharp, P. A. (1979) *J. Biol. Chem. 254*, 587–590.
Mirzabehov, A. D., & Griffin, B. E. (1972) *J. Mol. Biol. 72*, 633–643.
Miura, K., Watanabe, K., & Sugiura, M. (1974) *J. Mol. Biol. 86*, 31–48.
Murao, K., & Ishikura, H. (1978) *Nucleic Acids Res., Spec. Publ. No. 5*, S333–S336.
Nishimura, S. (1972) *Prog. Nucleic Acid Res. Mol. Biol. 12*, 50–85.
Parsons, J. T., & Green, M. (1971) *Virology 45*, 154–162.
Parsons, J. T., Gardner, J., & Green, M. (1971) *Proc. Natl. Acad. Sci. U.S.A. 68*, 557–560.
Perry, R., Kelley, D. E., Friderici, K., & Rottman, F. (1975) *Cell 4*, 387–394.
Sehgal, P. B., Fraser, N. W., & Darnell, J. E. (1979) *Virology 94*, 185–191.
Spector, D. J., McGrogan, M., & Raskas, H. J. (1978) *J. Mol. Biol. 126*, 395–414.
Takanami, M. (1967) *J. Mol. Biol. 23*, 135–148.
Thomas, D., & Green, M. (1966) *Proc. Natl. Acad. Sci. U.S.A. 56*, 243–246.
Wall, R., Philipson, L., & Darnell, J. E. (1972) *Virology 50*, 27–34.
Wei, C.-M., Gershowitz, A., & Moss, B. (1975) *Cell 4*, 379–386.
Wilson, M. C., Fraser, N. W., & Darnell, J. E. (1979) *Virology 94*, 175–184.
Wold, W. S. M., Büttner, W., & Green, M. (1978) in *The Molecular Biology of Animal Viruses* (Nayak, D., Ed.) pp 673–768, Marcel Dekker, New York.
Ziff, E. B., & Evans, R. M. (1978) *Cell 15*, 1463–1475.