Proc. Natl. Acad. Sci. USA
Vol. 78, No. 4, pp. 2283–2286, April 1981
Biochemistry

Mouse β-globin and adenovirus-2 major late transcripts are initiated at the cap site *in vitro*

[RNA polymerase II (RNA nucleotidyltransferase II)/ribonucleoside [β-^{32}P]triphosphates/capping mechanism]

OTTO HAGENBÜCHLE AND UELI SCHIBLER

Swiss Institute for Experimental Cancer Research, Chemin des Boveresses, CH-1066 Epalinges, Switzerland

Communicated by Robert Palese Perry, January 22, 1981

ABSTRACT **Initiation of transcription on major late DNA of adenovirus serotype 2 and the β-globin gene of the mouse was studied by using ribonucleoside [β-^{32}P]triphosphates as labeled precursors *in vitro*. The 5' termini of discrete runoff transcripts of both templates were preferentially labeled. The β phosphate of the cap structure was labeled by [β-^{32}P]ATP but not [β-^{32}P]GTP, indicating that the β phosphate of the cap is contributed by the initiating nucleoside triphosphate (ATP in both cases). Therefore, capping occurs on the initial nucleotide of these transcripts.**

Most eukaryotic cellular and viral mRNAs contain a 5'-terminal cap structure of the type m^{7}GpppNmp... (1). Whether cap formation occurs precisely at the sites of transcription initiation is not known; so far, the possibility that the 5' termini of some transcripts are generated by a rapid processing event has not been ruled out. Recently several groups have reported specific transcription of DNA with RNA polymerase II *in vitro* (2, 3). The transcripts that accumulate in these systems have capped 5' termini identical to those of their *in vivo* counterparts (2–4).

To determine whether the cap site is coincident with the site of transcription initiation, we have examined the transcription products from major late DNA† of adenovirus serotype 2 (Ad-2) and the mouse β-globin gene, using ribonucleoside [β-^{32}P] triphosphates as substrates. Our results indicate that capping occurs on the first nucleotide of these transcripts.

MATERIALS AND METHODS

HeLa Cell Extracts. HeLa human cells were grown in suspension culture in Eagle's minimal essential medium supplemental with 5% fetal calf serum to a density of 5×10^5 cells per ml. Cell lysates were prepared according to the method described by Manley *et al.* (3) except that the high-speed pellet was resuspended in 1/20 the vol of the supernatant.

DNA. Recombinant DNA was handled according to the National Institutes of Health guidelines (EK1, P3). A plasmid (pMSmaF) containing the *Sma* I F fragment of Ad-2 DNA with the major late cap site, inserted into pBR313 (2) was obtained from J. Segall and R. Roeder (Washington University). A plasmid (M-chrβG) containing the mouse β-globin gene inserted into the *Eco*RI site of pBR322 (5) was obtained from P. Dierks and C. Weissmann (University of Zürich). Plasmid DNA was prepared as described earlier (6), and samples of DNA that had been digested with restriction endonucleases were purified by phenol extraction and two ethanol precipitations. DNA was resuspended in 10 mM Tris·HCl, pH 8.0/0.1 mM EDTA at a concentration of 250 μg/ml.

***In Vitro* Transcription Reactions and RNA Purification.** Standard 100-μl reaction mixtures contained 12 mM Hepes at pH 7.9, 60 mM KCl, 7 mM $MgCl_2$, 0.02 mM EDTA, 1.3 mM dithiothreitol, 10% (vol/vol) glycerol, 10 mM creatine phosphate, 50 μM of each unlabeled ribonucleoside triphosphate, 5 μM ribonucleoside [α-or β-^{32}P]triphosphate (Radiochemical Centre, Amersham, England; 400 Ci/mmol; 1 Ci = 3.7×10^{10} becquerels), 60 μl of extract, and 5.0 μg of DNA.

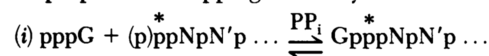
Reaction mixtures were incubated at 30°C for 60 min. After the addition of 400 μl of a solution containing 20 mM Tris·HCl at pH 7.0, 10 mM EDTA, and 0.5% sodium dodecyl sulfate, the samples were extracted three times with an equal volume of phenol/chloroform (7). Nucleic acids were precipitated with ethanol in the presence of 30 μg of carrier tRNA and the pellets were redissolved in 10 mM Tris·HCl, pH 7.0/1 mM EDTA/0.2% sodium dodecyl sulfate. After addition of NH_4OAc to 1 M, the nucleic acids were again precipitated with ethanol; a final ethanol precipitation was done from 0.3 M NaOAc. The pellet was dissolved in H_2O and stored at −70°C.

Analysis and Purification of RNA Transcripts by Gel Electrophoresis. For analytical purposes, 1% of the RNA obtained from a standard reaction mixture was glyoxylated and electrophoresed on 1.5% agarose gels as described by McMaster and Carmichael (8). Nucleic acid was denatured with glyoxal at 50°C for 5 min. Gels were dried and autoradiograms were obtained by exposure to Kodak XR-5 film at −70°C using a Du Pont Cronex Lightning Plus intensifying screen. For sequencing work, the remaining sample was run on agarose gels containing methylmercury hydroxide (9). After the gels had been washed in 1 M NH_4OAc and the film had been exposed, the appropriate bands were excised and RNA was eluted electrophoretically into dialysis bags (10). To eliminate any residual mercury bound to the RNA, the transcripts were precipitated from 1 M NH_4OAc with ethanol in the presence of 20 μg of carrier tRNA.

Analysis of Oligonucleotides. Eluted transcripts were digested with RNase T1 and the products were analyzed by two-dimensional separation as described by Platt and Yanofsky (11). Oligonucleotides from [α-^{32}P]UTP-labeled transcripts were identified by nearest-neighbor analysis by using RNase T2 as described (12). The cap cores in β-^{32}P-labeled oligonucleotides were analyzed by using nuclease P1 digestion and two-dimensional chromatography on cellulose thin layers (13, 14).

RESULTS

Experimental Strategy. To investigate whether the cap site corresponds to the site of transcription initiation, the capping mechanism must be understood. Two alternative models have been proposed for capping of eukaryotic mRNAs:

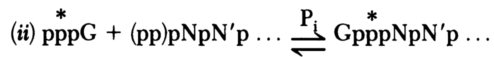


(*i*) pppG + (p)p$\overset{*}{\text{p}}$NpN'p ... $\overset{PP_i}{\rightleftharpoons}$ Gp$\overset{*}{\text{p}}$pNpN'p ...

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

Abbreviation: Ad-2, adenovirus serotype 2.

† "Major late DNA" refers to the Ad-2 *Sma* I F DNA fragment containing the major late cap site.



$(ii)\ pp\overset{*}{p}G + (pp)pNpN'p \ldots \underset{}{\overset{P_i}{\rightleftharpoons}} Gp\overset{*}{p}pNpN'p \ldots$

In the first model the β-phosphate (denoted by an asterisk) of the 5'-terminal residue (the N position) is retained during capping. This is the most straightforward mechanism postulated for capping of polyphosphorylated termini on primary transcripts and is involved in the capping of vaccinia virus (15, 16) and reovirus transcripts (17). In the second model the β-phosphate of the GTP cap donor is conserved. This alternative mechanism has been suggested for vesicular stomatitis virus (18); conceivably it could play a role in capping monophosphorylated 5' termini of RNA cleavage products. To discriminate between these two models we used either [β-^{32}P]ATP or [β-^{32}P]GTP as substrate for *in vitro* transcription of Ad-2 major late and mouse β-globin DNA in a HeLa cell extract (3). Because an A residue is present in the N position of both transcripts (19–23), one can determine whether the β-phosphate of the cap structure synthesized *in vitro* is contributed by the GTP cap donor or the penultimate nucleotide, A.

Analysis of Ad-2 Major Late and Mouse β-Globin Transcripts Labeled with Ribonucleoside [β-^{32}P]Triphosphates *in Vitro*. To simplify the assay for the initiation of specific transcription at the Ad-2 major late and mouse β-globin cap sites, we analyzed runoff transcripts made from appropriate DNA restriction fragments. DNA restriction fragments were prepared as templates by digesting cloned Ad-2 DNA (2) and mouse β-globin gene DNA (5, 24) with *Sma* I and *Bam*HI, respectively. Transcripts of 560 and 468 nucleotides, respectively are synthesized *in vitro* from these DNA restriction fragments (2–4). Fig. 1 shows that labeled Ad-2 and globin RNA fragments of the expected size were prominent after transcription in the presence of either [α-^{32}P]UTP (lanes 1 and 4) or [β-^{32}P]ATP (lanes 2 and 5). The labeled RNAs larger than the expected runoff transcripts were not further characterized. Their transcription must have started at sequences other than the cap sites. Manley *et al.* (3) have shown previously that only transcripts from cap-containing DNA restriction fragments were blocked at their 5' termini.

To examine the products of RNA synthesis using the Ad-2 DNA template, gel-purified runoff transcripts were subjected to two-dimensional analysis after RNase T1 digestion. The large T1 oligonucleotides from Ad-2 transcripts labeled with [α-^{32}P]UTP, [β-^{32}P]ATP, or [β-^{32}P]GTP are shown in Fig. 2 *A*, *B*, and *C*, respectively. Oligonucleotides labeled with [α-^{32}P]UTP (Fig. 2*A*) were identified by nearest-neighbor analysis after

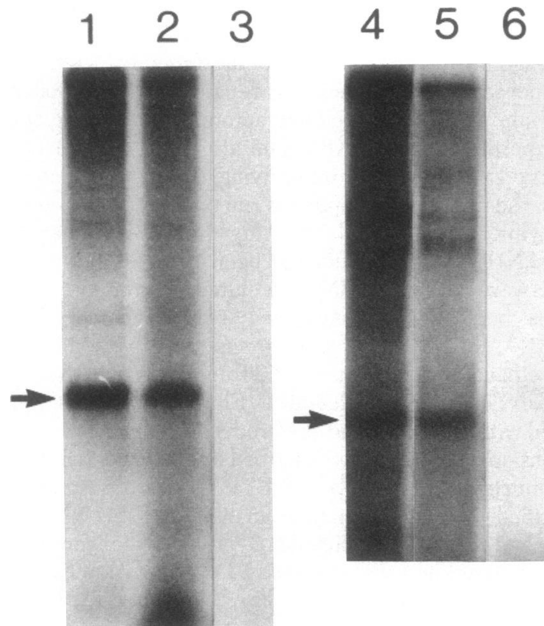


FIG. 1. Electrophoretic analysis of *in vitro* transcripts. Autoradiograms of denatured transcripts separated on 1.5% agarose gels are shown. RNA was synthesized in 100-μl reaction mixtures and either the template DNA or the radioactive ribonucleoside triphosphate was varied as follows: lanes 1 and 2, Ad-2 *Sma*-F-pBR313 template cut with *Sma* I and [α-^{32}P]UTP (lane 1) or [β-^{32}P]ATP (lane 2); lane 3, no DNA and [α-^{32}P]UTP; lanes 4 and 5, *Bam*HI-digested M-chrβG and [α-^{32}P]UTP (lane 4) or [β-^{32}P]ATP (lane 5); lane 6, no DNA and [α-^{32}P]UTP. Equal cpm were loaded on each lane. The arrows indicate the position of the 560- and 468-nucleotide transcripts, which originate from the Ad-2 (lanes 1 and 2) and M-chrβG DNA (lanes 4 and 5), respectively.

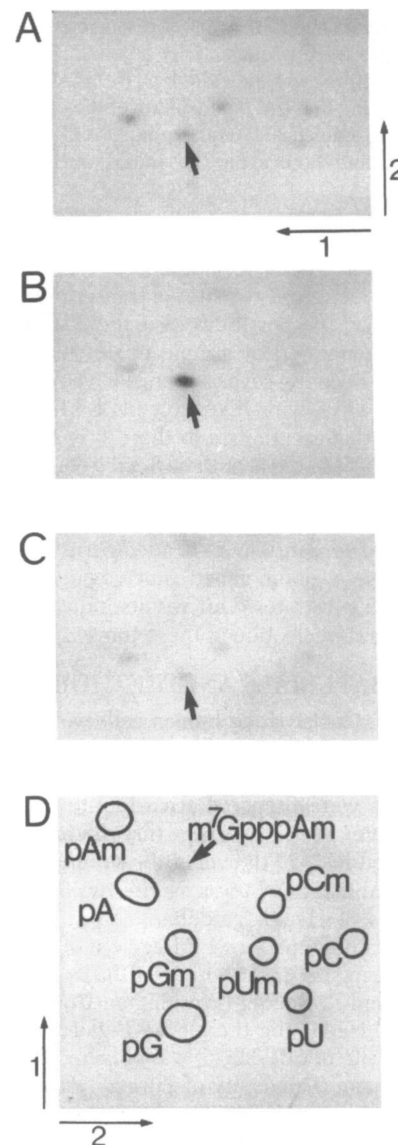


FIG. 2. Analysis of specific Ad-2 major late transcripts. Autoradiograms of RNase T1 oligonucleotide two-dimensional separation of gel-purified specific Ad-2 transcripts labeled with [α-^{32}P]UTP (*A*), [β-^{32}P]ATP (*B*), or [β-^{32}P]GTP (*C*) are shown. First (1) and second (2) dimensions are shown by arrows in *A*. Only the large T1 oligonucleotides are shown, and the position of the oligonucleotide containing the cap is indicated by a thick arrow. (*D*) Autoradiogram of a total nuclease P1 digest of the [β-^{32}P]ATP-labeled, cap-containing oligonucleotide (from *B*) after two-dimensional chromatographic separation.

digestion with RNase T2 (data not shown); the spot indicated by an arrow contained the 5′-terminal capped undecanucleotide [m^{7}GpppAmpCpUpCpUpCpUpUpCpCpGp (19, 21, 23)]. This oligonucleotide was preferentially labeled with [β-^{32}P]ATP (Fig. 2*B*). The other T1 oligonucleotides observed in Fig. 2*A* were also faintly labeled, suggesting that some β-phosphate groups were transferred to the α position during incubation in the transcription reaction.

To localize the ^{32}P label within the cap-containing oligonucleotide, this was further digested with nuclease P1 and the products were separated by two-dimensional chromatography. As seen in Fig. 2*D*, most of the label comigrated with the cap core m^{7}GpppAm. A faint radioactive spot comigrating with pU could also be detected. The mechanism behind the preferential labeling of pU, which also occurs in the other oligonucleotide shown in Fig. 2*B* (data not shown), is not clear. Because the pA and pG residues derived from internal positions were not labeled (Figs. 2*D* and 4*C*), we assume that the radioactivity associated with the cap core m^{7}GpppAm resides in the β phosphate. No preferential labeling of the capped oligonucleotide was observed with [β-^{32}P]GTP as the radioactive precursor (Fig. 2*C*).

To determine whether the preferential labeling of caps with [β-^{32}P]ATP versus [β-^{32}P]GTP was due to rapid turnover of the latter, their stabilities were compared. Samples were withdrawn from the *in vitro* reaction mixtures at different time points and analyzed by one-dimensional chromatography on polyethyleneimine-cellulose sheets (25). Fig. 3 demonstrates that [β-^{32}P]GTP was considerably more stable than [β-^{32}P]ATP during the 1-hr labeling period. The differential incorporation of β phosphates into the 5′ caps of transcripts was thus not the result of any relative instability of [β-^{32}P]GTP compared to [β-^{32}P]ATP. We conclude therefore that the β phosphate of the cap is derived from the penultimate nucleotide, A.

Two-dimensional analysis of the gel-purified β-globin runoff transcript is presented in Fig. 4. As for Ad-2 RNA, the capped globin T1 oligonucleotide [m^{7}GpppAmpCpApUpUpUpGp (4, 20, 22)] was preferentially labeled with [β-^{32}P]ATP (Fig. 4*B*) but not [β-^{32}P]GTP (Fig. 4*C*). Digestion of the [β-^{32}P]ATP-labeled oligonucleotide with nuclease P1 showed that most of the label resides in the cap core (Fig. 4*D*), again indicating that the β phosphate of the cap is derived from the penultimate nucleotide, A.

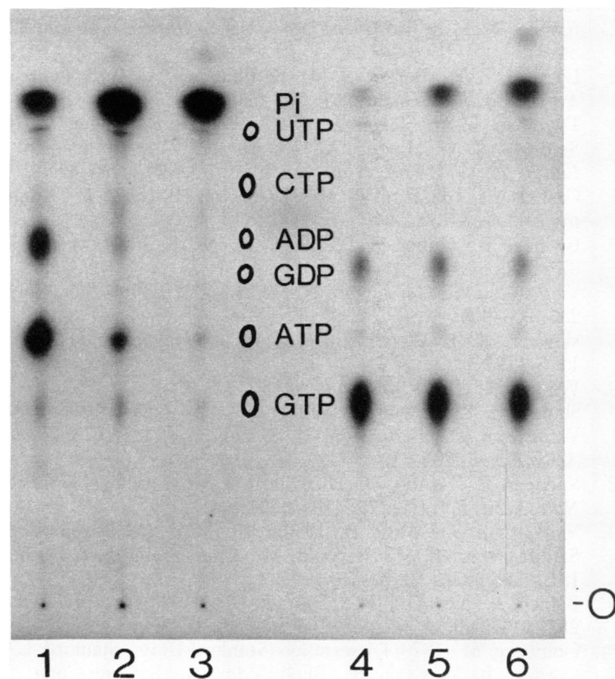


FIG. 3. Stability of [β-^{32}P]ATP and [β-^{32}P]GTP during RNA synthesis *in vitro*. Autoradiogram of a one-dimensional separation of samples taken at different times during transcription *in vitro* and spotted in parallel with marker nucleotides onto a polyethyleneimine-cellulose thin-layer chromatography sheet. Chromatography was in 0.75 M KH_2PO_4, pH 3.5 (25). Shown is 0.1% of the transcription reaction mixture incubated in the presence of [β-^{32}P]ATP for 2 min (lane 1), 15 min (lane 2), and 60 min (lane 3), or 0.1% of the reaction mixture containing [β-^{32}P]GTP incubated for 2 min (lane 4), 15 min (lane 5), and 60 min (lane 6).

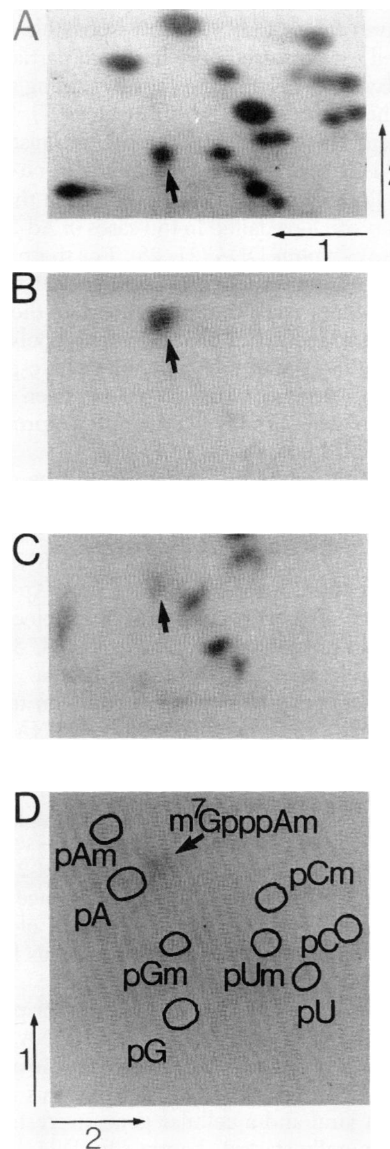


FIG. 4. Analysis of mouse β-globin transcripts. Autoradiograms of T1 oligonucleotide two-dimensional separation of gel-purified specific M-chrβG transcripts labeled with [β-^{32}P]UTP (*A*), [β-^{32}P]ATP (*B*), or [β-^{32}P]GTP (*C*). The two dimensions are indicated by arrows 1 and 2 in *A*. The arrows within the chromatograms designate the cap-containing oligonucleotides. Only the large T1 oligonucleotides are shown; the second-dimension chromatography was twice as far in *B* and *C* as that in *A*. The autoradiogram of *C* was exposed 10 times longer than *B*. (*D*) Autoradiogram of a total nuclease P1 digest of the [β-^{32}P]ATP-labeled, cap-containing oligonucleotide (shown in Fig. 3*B*) after two-dimensional separation.

DISCUSSION

The initiation of synthesis and capping of Ad-2 major late DNA and mouse β-globin gene transcripts has been examined *in vitro*. Both RNAs contain a penultimate A residue *in vivo* and *in vitro* (2, 3, 19–23). We find that the β phosphate of this A residue is incorporated into the RNA and is conserved during capping *in vitro*. The simplest explanation for these results is that capping occurs by GMP transfer to the initial nucleotide of the transcripts. Several lines of evidence support this conclusion: (*i*) The results of *in vitro* experiments using reovirus (17) and vaccinia virus (15, 16) capping enzymes are consistent with the GMP transfer mechanism. (*ii*) The β phosphate of ATP was incorporated into 5'-terminal cap structures of RNA polymerase II transcripts in a subcellular system of HeLa cells (26). (*iii*) A RNA guanylyltransferase that may be responsible for capping in the HeLa cell extracts used here has been partially purified from these cells (27). This enzyme caps by adding a GMP residue to 5' disphosphorylated RNA substrates (27). (*iv*) Finally, the coincidence of the cap site and the site of transcription initiation *in vitro* is supported by experiments *in vivo*; attempts to find nuclear RNA sequences transcribed from the DNA upstream of the cap site have failed in the cases of Ad-2 major late DNA and mouse β-globin DNA (21, 28). For these reasons we consider unlikely the alternative possibility that the β phosphate of an ATP donor is used to phosphorylate the 5' termini of RNA cleavage products. Cellular 5'-hydroxyl polynucleotide kinase activities (29–31) and a 5'-monophosphate polynucleotide kinase from vaccinia virus (32) have been described. Whether these enzymes are involved in mRNA processing has not yet been established.

The fact that only capped 5' termini could be detected in relatively short runoff transcripts indicates that capping occurs shortly after transcription initiation in the HeLa cell extract. Consistent with this notion, the purified HeLa RNA guanylyltransferase can accept substrates as short as ppApAp (33). Moreover, the presence of short capped HnRNA molecules in nuclear extracts from pulse-labeled HeLa cells (34, 35) suggests that transcripts are capped shortly after initiation.

A number of sequence elements are common to the DNA upstream of the cap site in most eukaryotic mRNA genes; the results presented here implicate these sequences in transcription initiation as well as capping. Two elements have been noted. TpApTpApApApTpA or similar sequences precede the cap site by 23 ± 1 base pairs (36); removal of these sequences diminishes accurate *in vitro* transcription of Ad-2 major late DNA and mouse β-globin DNA (37, 38). Sequences related to CpCpApApT are observed 60–90 base pairs 5' of the TpApTpApApApTpA box in most genes examined thus far (39, 40); their importance is less well defined.

In summary, we conclude that *in vitro* transcription of Ad-2 major late DNA and mouse β-globin gene DNA is initiated at the cap site, the N position of which becomes the penultimate nucleotide after GMP transfer. The fact that this mechanism applies to both a viral and a cellular gene suggests that transcription may generally start at the cap site. Whether capping events might also occur at sites other than those generated by transcription initiation has yet to be determined.

We thank Drs. P. K. Wellauer, R. A. Young, G. McMaster, and P. Beard for critical reading of the manuscript; R. Bovey for technical assistance; Drs. P. Dierks, C. Weissmann, P. Leder, J. Segall, R. Roeder, and G. Siegel for gifts of materials; and S. Cherpillod for typing the manuscript. This work was supported by a grant from the Swiss National Science Foundation (no. 3.148.77).

1. Shatkin, A. J. (1976) *Cell* **9**, 645–653.
2. Weil, P. A., Luse, D. S., Segall, J. & Roeder, R. G. (1979) *Cell* **18**, 469–484.
3. Manley, J. L., Fire, A., Cano, A., Sharp, P. A. & Gefter, M. L. (1980) *Proc. Natl. Acad. Sci. USA* **77**, 3855–3859.
4. Luse, D. S. & Roeder, R. G. (1980) *Cell* **20**, 691–699.
5. Van den Berg, J., Van Ooyen, A., Mantei, N., Schamböck, A., Grossveld, G., Flavell, R. A. & Weissmann, C. (1978) *Nature (London)* **276**, 37–44.
6. Schibler, U., Tosi, M., Pittet, A. C., Fabiani, L. & Wellauer, P. K. (1980) *J. Mol. Biol.* **142**, 93–116.
7. Perry, R. P., La Torre, J., Kelley, D. E. & Greenberg, J. R. (1972) *Biochim. Biophys. Acta* **262**, 220–226.
8. McMaster, G. K. & Carmichael, G. C. (1977) *Proc. Natl. Acad. Sci. USA* **74**, 4835–4838.
9. Bailey, J. M. & Davidson, N. (1976) *Anal. Biochem.* **70**, 75–85.
10. Galibert, F., Sedat, J. & Ziff, E. (1974) *J. Mol. Biol.* **87**, 377–407.
11. Platt, T. & Yanofsky, C. (1975) *Proc. Natl. Acad. Sci. USA* **72**, 2399–2403.
12. Sprague, K. U., Hagenbüchle, O. & Zuniga, M. C. (1977) *Cell* **11**, 561–570.
13. Marcu, K. B., Schibler, U. & Perry, R. P. (1978) *J. Mol. Biol.* **120**, 381–400.
14. Hagenbüchle, O., Bovey, R. & Young, R. (1980) *Cell* **21**, 179–187.
15. Ensinger, M. J., Martin, S. A., Paoletti, E. & Moss, B. (1975) *Proc. Natl. Acad. Sci. USA* **72**, 2525–2529.
16. Moss, B., Gershowitz, A., Wei, C. H. & Boone, R. (1976) *Virology* **72**, 341–351.
17. Furuichi, Y. & Shatkin, A. J. (1976) *Proc. Natl. Acad. Sci. USA* **73**, 3448–3452.
18. Abraham, G., Rhode, D. P. & Banerjee, A. K. (1975) *Cell* **5**, 51–58.
19. Gelinas, R. E. & Roberts, R. J. (1977) *Cell* **11**, 533–544.
20. Konkel, D. A., Tilghman, S. M. & Leder, P. (1978) *Cell* **15**, 1125–1132.
21. Ziff, E. B. & Evans, R. M. (1978) *Cell* **15**, 1463–1475.
22. Baralle, F. E. & Brownlee, G. G. (1978) *Nature (London)* **274**, 84–87.
23. Lockard, R. E., Berget, S. M., RajBhandary, U. L. & Sharp, P. A. (1979) *J. Biol. Chem.* **254**, 587–590.
24. Tilghman, S. M., Tiemeier, D. C., Polsky, F., Edgel, M. H., Seidman, J. G., Leder, A., Enquist, L. W., Norman, B. & Leder, P. (1977) *Proc. Natl. Acad. Sci. USA* **74**, 4406–4410.
25. Cashel, M., Lazzarini, R. A. & Kalbacher, B. (1969) *J. Chromatogr.* **40**, 103–109.
26. Groner, Y., Gilboa, E. & Aviv, H. (1978) *Biochemistry* **17**, 977–982.
27. Venkatesan, S., Gershowitz, A. & Moss, B. (1980) *J. Biol. Chem.* **255**, 903–908.
28. Weaver, X. & Weissmann, C. (1979) *Nucleic Acids Res.* **7**, 1175–1193.
29. Winicov, I. (1977) *Biochemistry* **16**, 4233–4237.
30. Shuman, S. & Hurwitz, J. (1979) *J. Biol. Chem.* **254**, 10396–10404.
31. Vennström, B., Pettersson, U. & Philipson, L. (1978) *Nucleic Acids Res.* **5**, 205–219.
32. Spencer, E., Loring, D., Hurwitz, J. & Monroy, G. (1978) *Proc. Natl. Acad. Sci. USA* **75**, 4793–4797.
33. Venkatesan, S. & Moss, B. (1980) *J. Biol. Chem.* **255**, 2825–2842.
34. Salditt-Georgieff, M., Harpold, M., Chen-Kiang, S. & Darnell, J. E. (1980) *Cell* **19**, 69–78.
35. Babich, A., Nevins, J. R. & Darnell, J. E. (1980) *Nature (London)* **287**, 246–248.
36. Goldberg, M. (1979) Dissertation (Stanford Univ., Stanford, CA).
37. Corden, J., Wasylyk, B., Buchwalder, A., Sarsone-Corsi, P., Kedinger, C. & Chambon, P. (1980) *Science* **209**, 1406–1413.
38. Proudfoot, N. J., Shander, M. H. M., Manley, J. L., Gefter, M. L. & Maniatis, T. (1980) *Science* **209**, 1329–1336.
39. Benoist, C., O'Hare, K., Brethnach, R. & Chambon, P. (1980) *Nucleic Acids Res.* **8**, 127–142.
40. Efstratiadis, A., Posakony, J. W., Maniatis, T., Lawn, R. M., O'Connell, C., Sprite, R. A., DeRiel, J. K., Forget, B. G., Weissman, S. M., Slightom, J. L., Blechl, A. E., Smithies, O., Baralle, F. E., Shoulders, C. C. & Proudfoot, N. J. (1980) *Cell* **21**, 653–668.