Volume 9 Number 10 1981 Nucleic Acids Research

Multiple polyadenylation sites in a mouse α-amylase gene

Mario Tosi, Richard A.Young*, Otto Hagenbüchle and Ueli Schibler

Swiss Institute for Experimental Cancer Research, Avenue des Boveresses, 1066 Epalinges, Switzerland

Received 22 January 1981

ABSTRACT

Two α-amylase mRNAs which differ in the length of their 3' non-translated region accumulate in the cytoplasm in both mouse liver and salivary gland tissues. The two species in each tissue are transcribed from the same gene (Amy-1^A). The minor species is approximately 20-fold less abundant than the major species and contains 237 additional nucleotides preceding the poly(A) tract. Sequence analysis of genomic DNA shows that these extra 237 nucleotides are specified by sequences contiguous to those shared by the two mRNAs. These data demonstrate that transcription can proceed through the major polyadenylation site and that alternative polyadenylation sites are used in the Amy-1^A gene. Sequences which trail the two polyadenylation sites exhibit extensive homology and might therefore be involved in polyadenylation or transcription termination.

INTRODUCTION

Polyadenylation is a nearly universal mRNA processing event in eucaryotic cells. It occurs early in mRNA biosynthesis and is characterized by the addition of approximately 200 adenosine residues to what is destined to be the 3' terminus of the mRNA (1-3). The demonstration that RNA polymerase II terminates transcription downstream of polyadenylation sites in the adenovirus 2 major late transcription unit as well as in the adenovirus 2 early regions 2 and 4 suggests that poly(A) addition requires endonucleolytic cleavage of the transcripts (4-6). Whether this is also true for cellular genes or whether the sites of polyadenylation and transcription termination are coincident has not yet been established.

We previously reported the nucleotide sequence of the major α-amylase mRNAs which accumulate in the salivary gland and in the liver of adult mice (7). These mRNAs are specified by the gene Amy-1^A in a tissue specific fashion : they share a common translated region but

differ in 5' terminal nontranslated leaders (8). Here we show that Amy-1^A transcripts are polyadenylated at two sites in both tissues.

MATERIALS AND METHODS

Preparation of RNAs. Three to four month old mice of the strain A/J were used in all experiments. The extraction of polyadenylated cytoplasmic RNAs from salivary gland and liver has been described (9). Polyadenylated nuclear RNA was prepared as described by Perry et al (10). In experiments where nuclear and cytoplasmic RNAs were compared, the two cell fractions were obtained from the same homogenate.

Molecular cloning. To obtain a liver α-amylase cDNA clone, double stranded cDNA containing sequences for sucrose gradient enriched polyadenylated RNA was fractionated electrophoretically and cDNA between 1.5 and 2.0 kb was inserted into the PstI site of pBR322 by use of the (dG)n·(dC)n tailing procedure (11). The conditions for transformation and colony hybridization have been described (9). Genomic sequences were purified by molecular cloning as described (8). All work involving recombinant DNA was carried out according to the NIH guide-lines for recombinant DNA (4-11-1979). P3 physical and EK1 biological containment was used.

Hybridization techniques. Size fractionation of RNA on methylmercury-hydroxide agarose gels (12) as well as RNA transfer to diazobenzyloxymethyl-cellulose (DBM)- paper and its hybridization with nicktranslated DNA probes (13) were performed as previously described (9). S1 nuclease mapping was according to Berk and Sharp (14), except that RNA-DNA hybridization was for 4 hr at 50^{o}C in a total volume of 40 µl.

Sequence analysis. Restriction mapping and DNA sequencing were performed according to Smith and Birnstiel (15) and Maxam and Gilbert (16), respectively.

RESULTS

Amy-1^A transcripts can be polyadenylated at multiple sites

To compare the α-amylase mRNAs which accumulate in the salivary gland and in the liver, we have cloned cDNA made against these mRNAs. The isolation and sequence analysis of the salivary gland α-amylase cDNA clone pMSal04 was recently described (9, 17). To screen a cDNA

library made against size fractionated liver polyadenylated RNA, the labeled cDNA insert of pMSl04 was used as a probe ; this cDNA is highly homologous to liver α-amylase mRNA as shown by thermal melting and S1 nuclease digestion of cDNA/mRNA hybrids (9). Among 3,000 liver cDNA clones, one positive clone was recovered (α-amylase mRNAs account for only 0.02% of the cytoplasmic polyadenylated RNA in this tissue). This clone, termed pMLa52, was subjected to restriction endonuclease mapping and sequence analysis.

The restriction endonuclease cleavage map determined for pMLa52 is identical to that determined previously for pMSal04 (the salivary gland α-amylase cDNA clone) with two exceptions (Fig. 1). First, pMLa52 cDNA lacks sequences which encode approximately 350 base pairs of 5' terminal mRNA sequences found in pMSal04. Second, unlike pMSal04, pMLa52 contains approximately 250 additional base pairs of DNA derived from the mRNA 3' terminus. DNA sequence analysis confirms these conclusions ; pMLa52 and pMSal04 cDNA sequences are identical in the comparable 1.3 kb shared by both clones (7), but pMLa52 cDNA contains an additional 237 nucleotides preceding the 3' terminal poly(A) tail, sequences which are absent in pMSal04 cDNA (Fig. 1, C). Since it has been demonstrated that sequences shared by pMSal04 and pMLa52 are specified by a single gene(8), the presence of the additional 3' terminal 237 residues in pMLa52 cDNA indicates that polyadenylation can occur at more than one site in Amy-1^A transcripts.

Major and minor α-amylase mRNAs accumulate both in liver and salivary gland

To ascertain whether the additional 237 nucleotides in pMLa52 are specific only to α-amylase mRNA in the liver, the 1.8 kb EcoRI fragment of pMLa52 containing these sequences (Fig. 1) was hybridized to either liver or salivary gland mRNA and the hybrids were examined by S1 nuclease digestion. Figure 2 shows that regardless of the source of the mRNA, the same 760 and 1020 nucleotide S1-nuclease resistant cDNA fragments were observed. Thus, two different α-amylase mRNAs accumulate in both tissues. The sizes of the two S1 nuclease resistant fragments are predicted by 3' terminal cDNA sequences found in pMSal04 (760 nucleotides) and pMLa52 (1020 nucleotides). As determined by densitometry, the 760 nucleotide fragment is approximately 20-fold more abundant than the 1020 nucleotide species whether liver or salivary gland RNA is used

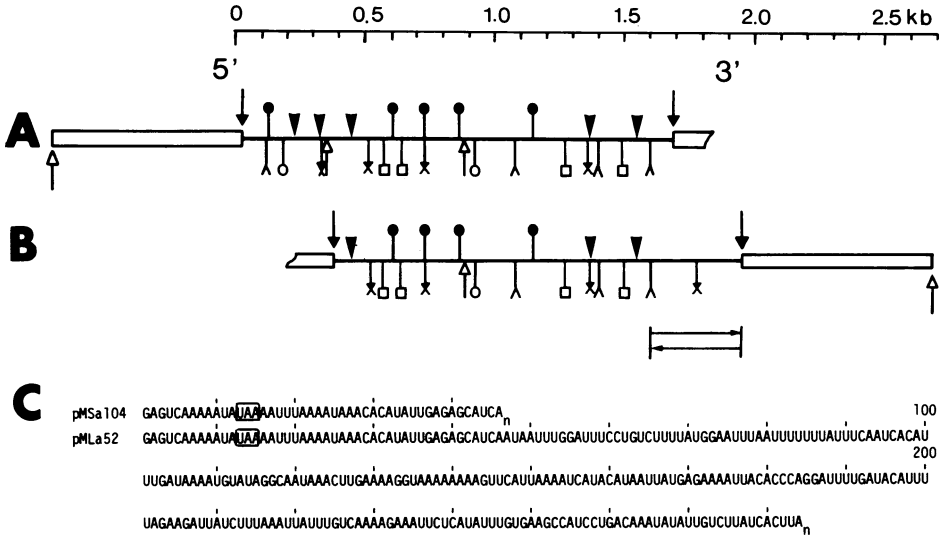

Figure 1. The α-amylase cDNA restriction cleavage sites and 3' terminal mRNA sequences derived from the cDNA. (A) and (B) : restriction endonuclease cleavage sites were determined as described by Smith and Birnstiel (15) for the cDNA insert of pMSa104 (A) and pMLa52 (B). The cDNA insert is designated by a line, pBR322 vector is shown as a box. Sites indicated are : PstI (↓), EcoRI (↑), AluI (▼), HaeIII (↑), HhaI (↓), HinfI (↓), TaqI (↓), and EcoRII (↓). Horizontal arrows indicate the direction and extent of DNA sequence analysis reported here. (C) mRNA sequences elucidated from the cDNA preceding the poly(A) tract are shown ; those from pMSa104 cDNA are reproduced from Ref. (17). The numbering begins with the HinfI DNA recognition sequence which occurs at 1.6 kb in (A) and (B). The α-amylase polypeptide terminator triplet UAA (17) is boxed.

(lanes 4 and 5 of Fig. 2). We conclude that pMLa52 contains sequences for the minor α-amylase mRNA, that pMSa104 contains sequences for the 20-fold more abundant major species, and that both mRNAs accumulate in the liver and in the salivary gland to similar ratios.

Because the minor α-amylase mRNA species contains the entire sequence of the major species, the possibility of a precursor-product relationship between the two molecules had to be considered. If such a relationship exists, one would expect that the presence of the minor species in the cytoplasmic RNA fraction reflects nuclear contamination, since no cytoplasmic pre-mRNA has been detected for cellular mRNAs. To

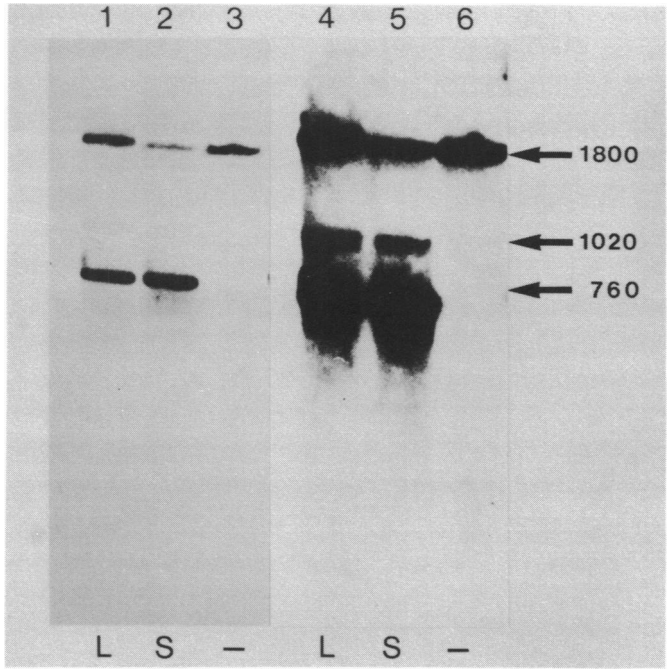

Figure 2. Examination of liver and salivary gland α-amylase mRNA 3' termini by S1 nuclease digestion of mRNA/cDNA hybrids. 2 μg of the 1.8 kb EcoRI DNA fragment of pMLa52 (see Fig. 1, B) were hybridized to 50 μg liver or 2 μg salivary gland cytoplasmic polyadenylated RNA. After S1 digestion, electrophoresis in a 2% alkaline agarose gel and transfer to a nitrocellulose filter, the S1 resistant fragments were visualized by hybridization to the labeled 1.8 kb EcoRI DNA fragment described above. Lanes 1 and 2 show the products from digestion in the presence of liver (L) and salivary gland (S) RNA, respectively. Lane 3 is a control in which no RNA was added to the cDNA. Lanes 4, 5 and 6 are a 20-fold longer exposure of lanes 1, 2 and 3, respectively. Numbers are in nucleotides. The 1800 nucleotide band represents undigested 1.8 kb EcoRI DNA fragments.

test the degree of nuclear contamination in our cytoplasmic RNA, we compared polyadenylated cytoplasmic and nuclear RNAs from the salivary gland by size fractionation on denaturing agarose gel. Figure 3, A shows that high molecular weight nuclear RNAs, including α-amylase pre-mRNA, are absent from the cytoplasmic fraction. Moreover, the ratios of minor to major species in nuclear and cytoplasmic fractions as determined by S1 nuclease mapping are similar (Fig. 3, B). Thus, nuclear contamina-

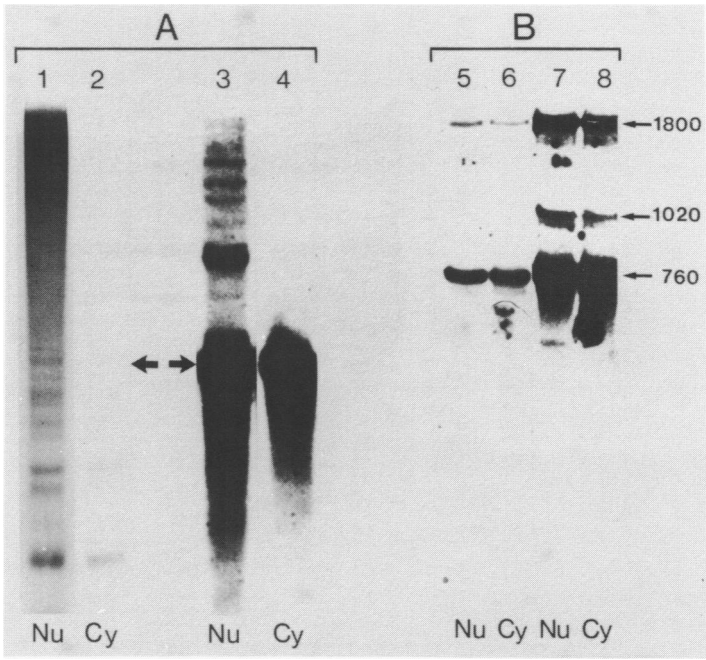

Figure 3. Nuclear and cytoplasmic RNA from the salivary gland. (A) 10 µg nuclear or 1 µg cytoplasmic polyadenylated RNA were electrophoresed on a 1.2% methylmercuryhydroxide agarose gel, transferred to DBM-paper and hybridized with the nicktranslated pMSal04 cDNA insert. Lanes 1 and 2 : ethydium bromide stained nuclear and cytoplasmic RNA respectively ; lanes 3 and 4 : hybridization of a DBM transfer of lanes 1 and 2 with labeled α-amylase cDNA. Arrows indicate the position of mature α-amylase mRNA.(B) S1 nuclease mapping of nuclear or cytoplasmic RNA from salivary gland. 2 µg of the 1.8 kb EcoRI fragment from pMLa52 containing the 3' terminal sequences were hybridized to 20 µg of nuclear poly(A)-containing RNA (lane 5) or to 2 µg of cytoplasmic poly(A)-containing RNA (lane 6). S1 resistant hybrids were detected as described in Fig. 2. Lanes 7 and 8 : 20 times longer exposure of the same autoradiograph.

tion in the cytoplasmic fraction cannot account for the presence of the minor α-amylase mRNA ; rather this RNA is a cytoplasmic species.

Genomic sequences of the Amy-1^{A} polyadenylation sites

To examine the Amy-1^{A} sequences which specify the polyadenylated 3' termini of major and minor α-amylase mRNAs, a mouse parotid gene bank (8) was screened by hybridization to the 1.75 kb HinfI DNA fragment of pMLa52 which contains 3' terminal mRNA sequences (Fig. 1). One

positive genomic clone, designated λchrαalA2(henceforth called λA2),was selected for further analysis. The sequences which specify the two polyadenylated 3' termini of the α-amylase mRNAs were located by Southern hybridization of λA2 DNA restriction fragments with the HinfI cDNA probe used in the genomic screen. The 1.5 kb PstI-EcoRI DNA fragment shown in Figure 4 was found to contain the probe sequences (data not shown). This fragment was mapped in greater detail with the endonucleases HinfI, HaeIII, EcoRII, MboI and TaqI (Fig. 4, A). The arrangement of TaqI, HinfI and EcoRII restriction cleavage sites in the center of this map is unique and identical to that found for the 3' terminal portion of pMLa52 cDNA (Fig. 1, B). Therefore, we used the strategy outlined in Figure 4, A to elucidate DNA sequences which specify the 3' termini of both major and minor α-amylase mRNAs (Fig. 4, B). The results demonstrate that genomic sequences which specify the 3' terminus of the major species are contiguous to those which specify the 237 additional 3' residues of the minor α-amylase mRNA.

DISCUSSION

The data presented here show that two α-amylase mRNAs which differ by 237 nucleotides in the length of their 3' terminal nontranslated region accumulate in the cytoplasm of both mouse salivary gland and liver. The DNA which specifies these additional 237 nucleotides is contiguous to that encoding the shared mRNA sequences. Thus the gene from which these RNAs are transcribed, Amy-1^A (8), contains two polyadenylation sites located 237 base pairs apart.

Two alternative models for polyadenylation of transcripts at multiple sites in the Amy-1^A gene can be postulated : 1) transcription termination occurs downstream of both polyadenylation sites and polyadenylation is determined by selective RNA processing. This has been shown for the adenovirus 2 early and late transcription units (4-6). 2) transcription termination occurs at or near the first polyadenylation site, but some level of readthrough exists to permit poly(A) addition at the second site. In either case, RNA polymerase II can transcribe through the major polyadenylation site in the Amy-1^A gene. The ratios of the two mRNA species appear to be identical in both salivary gland and liver, where the abundancy of total α-amylase mRNA

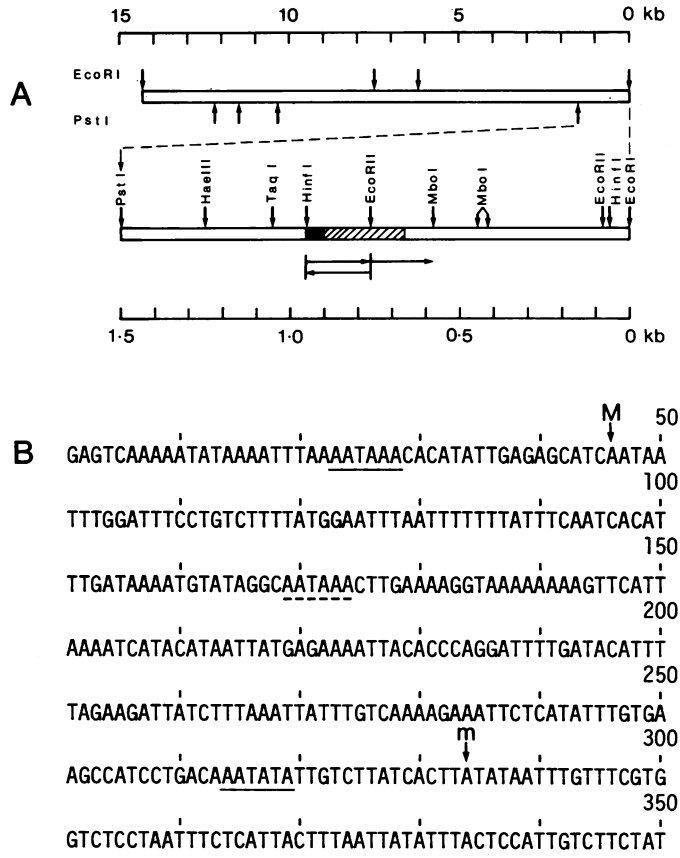

Figure 4. Examination of Amy-1^{A} DNA flanking the polyadenylation sites. (A) The open box represents λA2 genomic insert DNA. The 1.5 kb PstI-EcoRI DNA fragment is expanded below the complete insert. Vertical arrows indicate restriction endonuclease cleavage sites ; horizontal arrows designate the direction and extent of DNA sequence information obtained from this DNA. The blank box indicates exon sequences for the latter 46 nucleotides shared by the major and minor α-amylase mRNAs. The hatched box represents sequences which specify the 237 3' terminal nucleotides which precede the poly(A) tract in the minor species. (B) The DNA sequence determined using the strategy outlined in (A). Numbering begins with the HinfI recognition sequence at 0.95 kb (lower scale) in (A). The position of the initial adenosine in the mRNA poly(A) tract is indicated by an arrow ; (M) designates the major polyadenylation site, (m) designates the minor one. The AATAAA sequence (or AATATA) preceding polyadenylation sites is underlined. An AATAAA present in this sequence is underscored by a dotted line ; we do not detect a mRNA which is polyadenylated in this region.

differs by two orders of magnitude (9). This indicates that the sequences around the two polyadenylation sites may play a crucial role in determining the frequency of polyadenylation events at each site.

Multiple polyadenylation sites occur in other cellular genes. Early et al. (18) have described two polyadenylation sites in the mouse μ immunoglobulin heavy chain gene, where they may be involved in controlling the production of secreted or membrane-bound antibodies. Several mouse dihydrofolate reductase mRNA species differing in the length of 3' nontranslated regions could also result from polyadenylation at different sites in transcripts from the same gene (19).

The sequence elements of eucaryotic genes which signal polyadenylation are ill-defined. The hexanucleotide AAUAAA (20), or a variation thereof (17,21) precedes the poly(A) tract by 11 to 26 nucleotides in most eucaryotic mRNAs, including the major and minor Amy-1^{A} mRNAs (Fig. 4). It appears to play a role in polyadenylation, since variants of SV40 which carry a tandem duplication of this hexanucleotide produce late RNA species in which polyadenylation has occurred after each of the AAUAAA sequences (T. Shenk, personal communication). However, this hexanucleotide alone is insufficient, because it also occurs within gene sequences which do not serve as polyadenylation sites (22, 23). Thus, it is noteworthy that in Amy-1^{A} an AATAAA sequence (at position 119-124 in Fig. 4) is apparently not associated with a polyadenylation site, since we did not detect any α-amylase mRNA 3' end in this region. An additional feature of most polyadenylation sites examined thus far is the ability of genomic sequences to specify the initial adenosine residue of the poly(A) tract (18, 21, 23, 24, 25); both major and minor Amy-1^{A}polyadenylation sites contain an A in the appropriate position (Fig. 5). Finally, the DNA segments immediately following the two Amy-1^{A} polyadenylation sites show extensive homology over a stretch of about 18 nucleotides (Fig. 5). Conceivably, these homologous sequences may play a role in polyadenylation or termination of Amy-1^{A} transcripts. The fact that no such homologous DNA segment follows the hexanucleotide AATAAA at position 119-124 (Fig. 4), which does not seem to be involved in polyadenylation, is consistent with this idea. However, no striking similarity can be observed between the sequences trailing polyadenylation sites in a variety of other genes (24).

The analysis of RNA products of Amy-1^{A} genes in which sequences

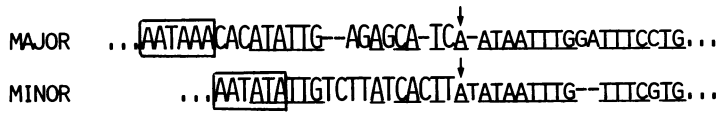

Figure 5. Amy-1^{A} sequences flanking polyadenylation sites for the major and minor species. Exon sequences are in large type, flanking sequences are in small type. The initial adenosine residue at the mRNA poly(A) tract is designated by an arrow. The two sequences were aligned to maximize blocks of homology ; homologous sequences are underlined. The AATAAA sequences (AATATA for the minor species) are boxed.

around the alternative polyadenylation sites have been altered should reveal sequence elements which are essential for polyadenylation.

ACKNOWLEDGEMENTS

We are grateful to P. Wellauer for help with the initial characterization of the pMLa52 clone and for critical reading of the manuscript. We thank B. Hirt for interest and support throughout the work. We are indebted to L. Korn and to P. Beard for valuable comments. R. Bovey and A. C. Pittet provided excellent technical assistance. This work was supported by a grant of the Swiss National Science Foundation (No. 3.148.77).

*Present address: Department of Biochemistry
Stanford University School of Medicine
Stanford, California 94305, USA

REFERENCES

1. Brawerman, G. (1974) Ann. Rev. Biochem. 43, 621-642
2. Perry, R. P. (1976) Ann. Rev. Biochem. 45, 605-629
3. Darnell, J. E. (1979) in Proceedings of the 11th Miami Winter Symposium "From Gene to Protein : Information transfer in normal and abnormal cells",pp. 207-228, Academic Press, New York
4. Nevins, J. R. and Darnell, J. E. (1978) Cell 15, 1477-1493
5. Frazer, N. W., Nevins, J. R., Ziff, E. and Darnell, J. E. (1979) J. Mol. Biol. 129, 643-656
6. Nevins, J. R., Blanchard, J. M. and Darnell, J. E. (1980) J. Mol. Biol. 144, 377-386
7. Hagenbüchle, O., Tosi, M., Schibler, U., Bovey, R., Wellauer, P., and Young, R. A. (1981) Nature, submitted
8. Young, R. A., Hagenbüchle, O. and Schibler, U. (1981) Cell, in press

9. Schibler, U., Tosi, M., Pittet, A. C., Fabiani, L. and Wellauer, P. K. (1980) J. Mol. Biol. 142, 93-116
10. Perry, R. P., Kelley, D. E. and Schibler, U. (1979) Proc. Natl. Acad. Sci. USA 76, 3678-3682
11. Chang, A. C. Y., Nunberg, J. H., Kaufmann, R. J., Erlich, H. A., Schimke, R. T. and Cohen,S.N.(1978) Nature 275, 617-623
12. Bailey, J. M. and Davidson, N. (1976) Anal. Biochem. 70, 75-85
13. Alwine, J. C., Kemp, D. J. and Stark, G. R. (1977) Proc. Natl. Acad. Sci. USA 74, 5350-5354
14. Berk, A. J. and Sharp, P. A. (1977) Cell 12, 721-732
15. Smith, H. O. and Birnstiel, M. L. (1976) Nucl. Acids Res. 3, 2387-2398
16. Maxam, A. M. and Gilbert, W. (1980) in Methods in Enzymology 65 : Nucleic Acids Part I, L. Grossman and K. Moldave, eds, pp. 499-560, Academic Press, New York
17. Hagenbüchle, O., Bovey, R. and Young, R. A. (1980) Cell 21, 179-187
18. Early, P., Rogers, J., Davis, M., Calame, K., Bond, M., Wall, R. and Hood, L. (1980) Cell 20, 313-319
19. Setzer, D. R., Mc Grogan, M., Nunberg, J. H. and Schimke, R. T. (1980) Cell 22, 361-370
20. Proudfoot, N. J. and Brownlee, G. G. (1976) Nature 263, 211-214
21. Nunberg, J. H., Kaufman, R. J., Chang, A. C. Y., Cohen, S.N., and Schimke, R. T. (1980) Cell 19, 355-364
22. Mc Reynolds, L., O'Malley, B. W., Nisbet, A. D., Fothergill, J. E., Givol, D., Fields, S., Robertson, M. and Brownlee, G. G. (1978) Nature 273, 723-728
23. Perricaudet, M., Le Moullec, J. M., Tiollais, P. and Pettersson, U. (1980) Nature 288, 174-176
24. Benoist, C., O'Hare, K., Breathnach, R. and Chambon, P. (1980) Nucl. Acids Res. 8, 127-142
25. Aleström, P., Akusjärvi, G., Perricaudet, M., Mathews, M. B., Klessig, D. F. and Pettersson, U. (1980) Cell 19, 671-681