Volume 17 Number 12 1989 Nucleic Acids Research

Length heterogeneity in rat salivary gland $\alpha 2\mu$ globulin mRNAs: multiple splice-acceptors and polyadenylation sites

Feng Gao, Hideya Endo and Mikio Yamamoto*

Department of Biochemistry, Medical Institute of Bioregulation, Kyushu University, Higashi-ku Fukuoka, 812, Japan

Received February 27, 1989; Revised and Accepted April 28, 1989 EMBL accession no. X14552

ABSTRACT

Rat $\alpha 2\mu$ globulins are coded for by a family of about 25 structurally related genes, some of which are expressed in the male adult liver while the other subset seems to be active in several excretory organs, including salivary and lacrymal glands. To estimate the number and specificity of genes expressed in the salivary glands, we determined nucleotide sequences of 30 cDNA clones. At least two $\alpha 2\mu$ globulin genes are active and two thirds of mRNAs were shown to code for the peptide two amino acids shorter than the others. Unexpected observation was the intense length polymorphism in the 3′ non-coding 6th intron-7th exon regions presumably caused by alternative splice-acceptor selection. At least six acceptor sites were utilized and the longest type retained the entire 6th intronic sequence resulting in a formation of unusually longer product. A stable mRNA molecule of this type was demonstrated in salivary glands by Northern blotting probed with the 6th intron-specific fragment. Together with three independent polyadenylation sites, the rat salivary glands generate a diverse set of $\alpha 2\mu$ globulin mRNAs.

INTRODUCTION

The $\alpha 2\mu$ globulin is one of major proteins synthesized in the liver of mature male rats. The protein is secreted into the blood and excreted in the urine (1). The $\alpha 2\mu$ globulin is encoded by a multigene family of about 25 highly homologous members (2,3). The hepatic synthesis of $\alpha 2\mu$ globulin is under complex hormonal control with androgens, insulin, glucocorticoids, thyroid and growth hormones stimulating, and estrogens suppressing its biosynthesis (4−6). Some of these $\alpha 2\mu$ globulin genes are expressed in the liver, and different subsets of genes in salivary, lacrymal and preputial glands (3,7−9). Genes transcribed in salivary and lacrymal glands are considered to be the same as deduced from the similarity in restriction digestion pattern of cDNAs and in the 2D polyacrylamide gel pattern of in vitro translation products using mRNAs from these tissues (8). The developmental pattern of $\alpha 2\mu$ globulin gene expression is, however, different in the liver, lacrymal and salivary glands. While both liver and lacrymal gland levels exhibit a sexual dimorphism and are greatly reduced after hypophysectomy, salivary gland levels exhibit no sexual dimophism and are unaffected by hypophysectomy. Lacrymal gland $\alpha 2\mu$ globulin mRNA levels respond only to hydrocortisone among several hormones which affect hepatic levels (8). The lack of hormonal modulation of salivary gland $\alpha 2\mu$ globulin synthesis is not simply due to the lack of hormone responsiveness of this tissue. Submaxillary duct cells, a sole type of cells in which salivary $\alpha 2\mu$ globulin is localized (10), are the site of synthesis of androgen-responsive proteins such as epidermal growth factor (11). In addition, this cell type is glucocorticoid and thyroid hormone-responsive (11).

The major in vitro translation products using mRNAs from salivary and lacrymal glands are more acidic than are the hepatic mRNA products and are electrophoresed on 2D polyacrylamide gels as at least two variants (8). Although only one cDNA sequence has been reported for salivary type mRNA (7), electrophoretic protein variants indicate the presence of multiple mRNA species.

We isolated cDNA clones from the library made against the salivary mRNAs. From analysis of these sequences, we established the presence of two groups of sequences. These cDNA sequences varied in length due to the alternative usage of at least six splice-acceptor sites located in 6th intron-7th exon regions, together with three polyadenylation sites. The transcription of salivary type $\alpha 2\mu$ globulin genes thus produced a set of mRNAs differing diversely in 3′ noncoding regions.

MATERIALS AND METHODS

Construction of the rat salivary gland cDNA library

The cDNA library was constructed against poly(A)$^+$RNA from salivary glands of six-week old female Donryu rats, according to the method described by Gubler and Hoffman (12). EcoRI linker-ligated double stranded cDNAs were inserted into the λgt11 vector precut with EcoRI and pretreated with calf intestine alkaline phosphatase (Boehringer). Final population size was 2×10^6 and amplified once before screening.

Isolation of salivary type $\alpha 2\mu$ globulin cDNA clones

About one million clones of the cDNA library were screened with nearly full-length liver type cDNA clone pA108 (13) as a probe. Over 50 clones with positive signals were isolated and 30 clones which contained >1kbp inserts were selected for further analysis.

Nucleotide sequencing

EcoRI digested inserts of cDNA clones were subcloned into the plasmid vector pUC118 or pUC119 (14). After single-stranded DNAs were isolated with the aid of helper phages, the nucleotide sequences were determined by the dideoxy nucleotide chain-termination methods with 2′-deoxy-7-deazaguanosine triphosphate instead of dGTP (15) or by using a Sequenase sequencing kit (United States Biochemical Corp. Cleveland USA).

RNA blot analysis

Poly(A)$^+$RNA or total RNA was electrophoresed through a 1.5% agarose/ 6% (v/v) formaldehyde gel and blotted onto nitrocellulose filters, according to the method of Thomas(16). The papers were then baked at 80°C for 2hr, prehybridized and hybridized with the appropriate DNA fragments ^{32}P-labeled by random priming (17).

RESULTS

1. Structural features of salivary $\alpha 2\mu$ globulin mRNAs.

Screening of a cDNA clone library from salivary glands of six-week old female Donryu rats with a liver type $\alpha 2\mu$ globulin cDNA clone, pA108 (13) as a probe, we isolated 30 positive clones. Figure 1 summarizes the sequences determined for all of these clones. Several distinct features became apparent from the inspection of these sequences. Clones could be divided into two groups; one was homologous to that reported previously for the salivary type cDNA (7) and the other different in several positions by base substitution that scattered throughout the entire mRNA region, as indicated in figure 1. We called hereafter these as type 2 and type 1 mRNAs and genes, respectively. All base changes appeared in the coding region, except for the 3′-most one, were conservative in coding amino acids. However, the last one had a change at the first position of the codon and

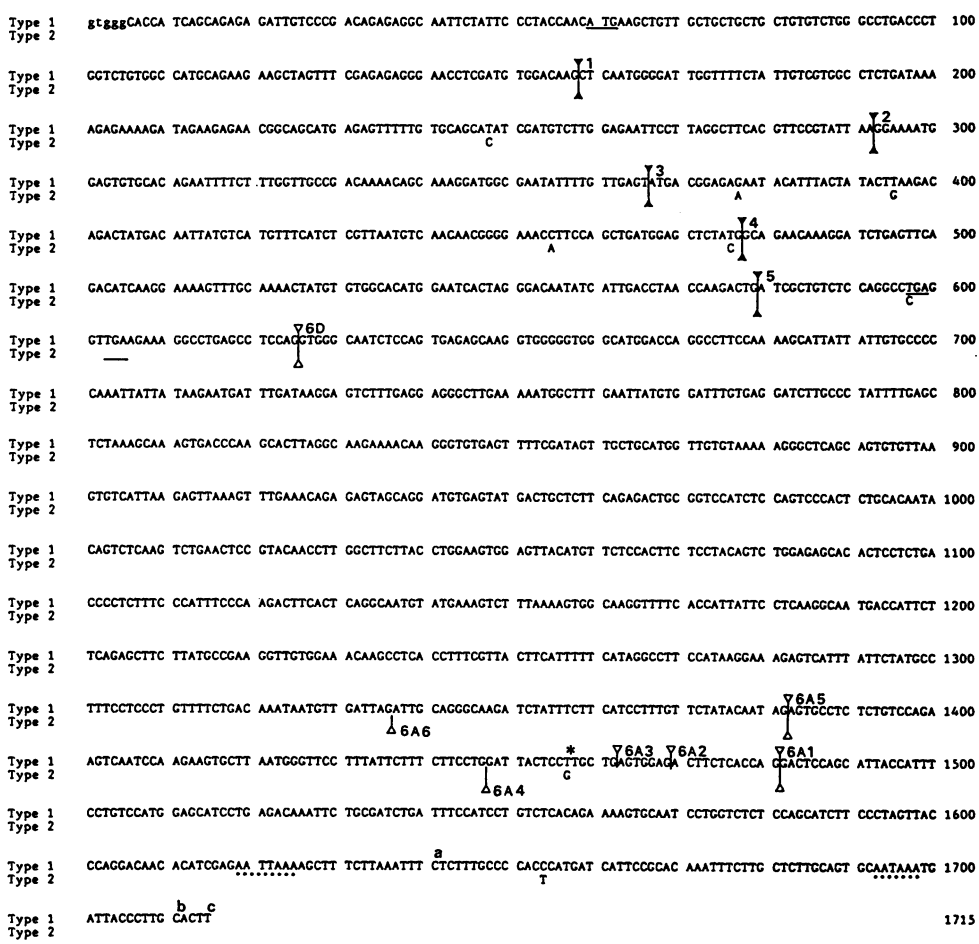

Fig.1. Complete structures of type 1 and type 2 cDNAs of rat salivary gland $\alpha 2\mu$ globulins. Nucleotide sequences of type 1 and type 2 cDNAs were aligned, where only mismatched bases were shown in type 2 cDNA. First 5 bases shown in lower case letters were tentatively taken from the genome sequence after S1 mapping analysis (unpublished data), since the cDNAs extended to the position shown by upper case letters. Translational start and stop signals are underlined. Symbols ▼ and ▲ (and ▽ and △) indicate the exon-intron junctions for type 1 and 2 genes, respectively. D and As denote the splice donor and acceptors, respectively, for example; 6D, donor site of 6th intron ; 6A4, acceptor site of 6th intron that operates for the production of class 4 mRNA in type 2 gene. Polyadenylation sites are shown with a, b and c, and putative polyadenylation signals are dot-underlined. The asterisk (*) corresponds to the position of a unique splice-acceptor of 6th intron operated in liver type genes (3). Since the sequence of class 0 mRNA from type 2 gene is not available at present, sequence stretch between positions 6D to 6A6 is applicable only for type 1 cDNA.

caused arginine to nonsense change, resulting in a size-reduction of the product peptide by two amino acid residues. Although exon 3 is known to show remarkable difference between liver type and salivary type cDNAs (7), two types of salivary cDNA were identical in sequences of this region. Among 30 clones analyzed 19 clones belonged to type 1 and 11 clones to type 2 mRNAs. Assuming that the cloning and translation efficiencies of type

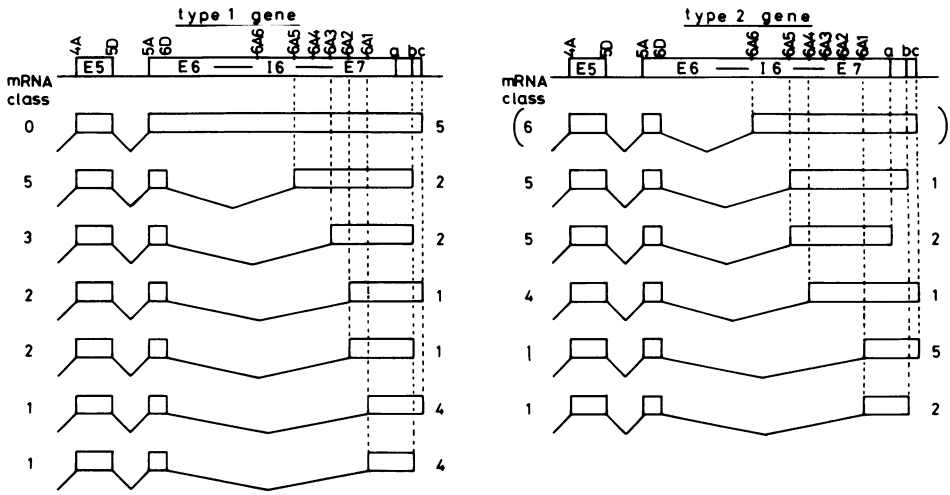

Fig.2. Schematic presentation of 3′ structures of type 1 and type 2 genes and mRNAs of rat salivary gland $\alpha 2\mu$ globulins. At the top, 3′ terminal regions of salivary type 1 and type 2 genes are depicted, where Es and Is indicate exons and introns, respectively, Ds and As indicate splice donors and acceptors, respectively. Multiple splice acceptors observed in the I6 are denoted as 6A6−6A1 as in Fig.1. Multiple polyadenylation sites are shown with lower case letters, a, b and c. Distances between these sites are not in scale. Under the maps of genes, structures of classes of mRNAs are presented. The arabic numerals at right side of each spliced mRNA species indicate the numbers of clones obtained that belonged to the classes illustrated at their left.

1 and type 2 mRNAs are similar, the major class of salivary $\alpha 2\mu$ globulin peptide is the one lacking two carboxy-terminal amino acid residues from that made in the liver.

The most striking feature was the size-difference in the 3′regions of mRNAs, as schematized in Fig.2. Class 0 mRNA was identified as a product of type 1 gene and contained the whole stretch of sequence corresponding to the 6th intron established in liver type gene (3 and our unpublished data). Class 6 mRNA was described by Laperche et al. (7). Although the sequence was consistent with that of type 2 gene, we did not identify this product in our cDNA isolates. Class 5 and class 1 mRNAs were commonly produced by both type 1 and type 2 genes, while class 4 mRNA was only from type 2 gene, and class 3 and class 2 mRNAs were, on the contrary, only from type 1 gene. MRNAs belonging to classes 0, 6, 5 and 4 contained considerable parts, or even all, of 6th intronic region, while those of classes 3, 2 and 1 lost parts of exon 7. No single clone was found to use the conventional 6th intron-7th exon junction, ...ttactcagTGCTG..., established in the liver type gene (3, * in Fig.1). All these classes of mRNAs, except for class 0 mRNAs, thus appeared to use a common splice donor site, identical to that used by the liver type gene (3), but to select splice acceptor sites in a highly divergent fashion. Among these, class 1 mRNAs, which lacked 5′ 24b stretch from the 7th exon, were the most frequent subtype as judged from cloning efficiencies. Polyadenylation was found to occur at three distinct locations (Fig.1 a, b and c). Major population of cDNAs had poly (A) tail at the same position (*c* in Fig.1) determined previously for liver cDNAs (3,13) and salivary cDNA (7), and at 4 bases upstream of this site (*b* in Fig.1). Frequencies were almost the same for these two (14 vs. 13, Fig.2). Polyadenylation at the position *b* was also observed in liver cDNAs, but far less frequently (unpublished results). Two cDNA clones of class

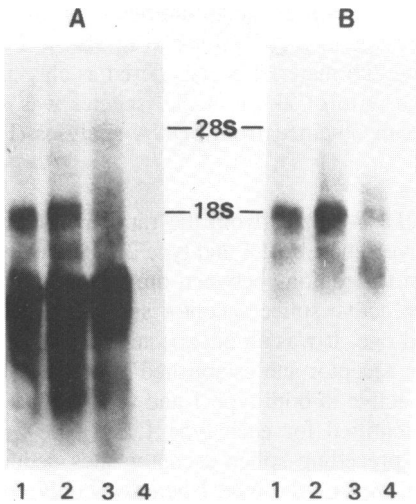

Fig.3. Northern blot analysis of $\alpha 2\mu$ globulin transcripts probed with fragments specific to the exons and the 6th intron. Poly $(A)^+$RNA ($3\mu g$: lanes 1 and 2) or total RNA ($10\mu g$: lanes 3 and 4) was subjected to electrophoresis under denaturing conditions, transferred to nitrocellulose paper and hybridized with the probes derived from exons (A; nt 264 EcoRI- nt 1626 HindIII fragment of class 1 cDNA) and intron (B; nt 673 AatI-nt 1267 AatI fragment of class 0 cDNA, Fig.1). (Lanes 1 and 2) salivary gland RNAs from 6 week old male and female, respectively; (lanes 3 and 4) liver RNAs from 6 week old male and female, respectively. Exposure time was 24hr at −80°C with an intensifying screen.

5 mRNA of type 2 gene (Fig.2) had a polyadenylation at about 70 bases upstream of the major site (position *a*), while the other clone of this class of mRNA terminated at the major poly (A) site. Polyadenylation at the position *a* might recognize the AATTAAA sequence that located 20b upstream of a poly(A) sequence as an alternative of the consensus polyadenylation signal, AATAAA. Since polyadenylation from the corresponding site has been reported to occur also in minor population of liver cDNAs (3), position *a* is utilized in both liver and salivary type genes though less efficiently. MRNAs of many classes from both types of genes appeared to use these poly (A) sites arbitrarily (Fig.2). From these observations polyadenylation site selection is likely to be random with much greater preference of positions *b* and *c* to *a*, and is independent from the splicing events.

2. *Existence of stable class 0 mRNAs in salivary glands*

We have identified unusually spliced class 0 mRNA by cDNA analysis, in which the entire 6th intronic sequence was retained while all other introns were spliced out normally. The number of clones of this class was five out of 30 clones isolated, indicating a considerable fraction among all $\alpha 2\mu$ globulin mRNAs in the salivary glands. To confirm the existence and fractional ratio, we performed Northern blotting analysis with the use of intron-less cDNA and 6th intron-specific (nt 673 AatI-nt 1267 AatI fragment) probes. As Fig.3 shows, when the exon-specific probe was used, two major signals were observed in both RNA samples from male and female salivary glands, one at around 2kb and the other at around 1.2kb (Fig.3A-1 and 2). The major 1.2kb band was remarkable but 2kb band was barely visible in a male liver RNA sample (Fig.3A-3). No signal was observed in a female liver sample (Fig.3A-4). When the RNA blot was screened with the intron-specific probe, only

2kb materials were hybridized in salivary gland samples (Fig.3B-1 and 2). A small amount of this species may be present in a male liver (Fig.3B-3). From the length and the hybridization specificity, the 2kb material is considered as class 0 mRNA. Hybridization intensities indicated that the ratio of 2kb to 1.2kb species was about 1 to 5. This ratio was consistent with the result obtained from cDNA analyses (Fig.2).

DISCUSSION

Sequencing $\alpha 2\mu$ globulin cDNA clones from the rat salivary glands, we elucidated the presence of at least two active genes (type 1 and type 2). Unusually heterogeneous splicing was shown to occur within the regions between 6th intron and 7th exon for both types of genes (Fig.1). Among 6 active splice acceptor sites identified, site 6A1 appeared to be the most frequently used one. It was an unexpected result that no one class of mRNA was found to use the splice acceptor site established in the liver type gene (3). All these six splice sites may not be active in both type 1 and type 2 genes, because cDNA clones using only 4 sites were identified for each type 1 and type 2 genes (Fig.1). Table 1 summarizes the sequences preceding splice acceptor sites deduced from longer cDNA sequences. The 6A2 site utilized by the type 1 gene was only poorly homologous to the consensus. This sequence motif appeared too remote to be spliced preferentially to the other sequence motifs. Unusual splice acceptor structures, ttgttaaagT and ttttaaaagG, were reported for a certain human endogenous retrovirus (18), and adenovirus E1A gene (19), respectively. Even more interesting were the putative sequences for 6A4 and 6A3, where the last two nucleotides were TGs instead of AGs, for which no exception has ever been reported (20). Class 3 and 4 mRNAs may be transcribed from other genes of otherwise identical sequence to type 1 and 2 genes but with altered 6A3 and 4 sequences. We think, however, that this possibility is less likely from the following reason. Among about 25 $\alpha 2\mu$ globulin genes, only one fifth are thought to be salivary type as revealed by differential hybridization using oligonucleotide probes specific to liver and salivary type gene sequences

Table 1

Splice sites	Splice acceptor sequences	genes operated (clone numbers obtained)	
(6A6	taatgttgattagA		type 2) (1)
6A5	ttctatacaatagA	type 1 (2)	type 2 (3)
6A4	ttctttcttcctgG		type 2 (1)
6A3	ttactccttgctgA	type 1 (2)	
6A2	ttgctgagtggagA	type 1 (2)	
6A1	gacttctcaccagG	type 1 (8)	type 2 (7)

Upper case letters indicate the first nucleotide of exons, and lower case letters show the sequence stretches preceding them (Fig.1). Since class 0 mRNA was not identified for type 2 gene, the sequence for the 6A6 site was tentatively taken from type 1 gene (parenthesis). Non-canonical dinucleotide sequences, ...tg, were underlined. See the text (Discussion section) for their possible usage.

resided in exon 3 (7, 13 and F.G. et al., unpublished data), indicating that salivary type genes comprised about 5 members. We have analyzed 6 independent salivary type genomic clones. These were categorized into 4 sequence groups, none of which coincided completely to the cDNA sequences shown in Fig. 1. Together with type 1 and type 2 genes, which are needed to be identified, at least 6 salivary type genes are thus considered to be present in the rat genome. Unless numbers of total family member(2) were exceedingly underestimated, possibility of the presence of other members of salivary type gene seems low. We therefore considered that 6A4 and 6A3 sites are likely to use non-canonical acceptor sequence ...TG instead of ...AG dinucleotide. Isolation and sequence analysis of salivary type 1 and type 2 genes are now in progress to determine the possibilities.

We reported alternative splicing in liver type $\alpha 2\mu$ globulin genes (13). In this case, however, splice donor site selection was the cause,rather than acceptor selection, although sequences around splice acceptor sites were highly homologous to those for salivary type genes (Fig.1). On the contrary, salivary type cDNA did not show donor site selection, despite of high homology in the 5′ parts of the 6th intron region with that of liver type gene. Tissue specificity in splice site selection between liver and salivary glands indicates that the recognition of donor site and acceptor site would be regulated independently and in a tissue specific manner.

Physical consequence of an alternative splicing in the 6th intron-7th exon regions of salivary type $\alpha 2\mu$ globulin genes appeared to be simply the generation of length heterogeneity of 3′ noncoding regions of the mRNAs, without affecting protein products. Attention has recently been focused on the function of the 3′ noncoding region of mRNA with respect to the mRNA stability (21−24). We have identified seven classes of mRNAs, many of which are commonly present in both type 1 and type 2 genes, and three polyadenylation sites that appeared to be used randomly. If so, as many as 42 species of mRNAs may be generated, although only 13 species were demonstrated in this study. The length heterogeneity occurred in the same region of liver mRNAs did not affect the stability (13).

Type 1 mRNAs were different from type 2 mRNAs at 6 positions by base substitution in the coding regions (Fig.1). Five of them resulted in no change in amino acid residues. This contrasted sharply to the results obtained from the comparison between liver-type and salivary-type mRNAs, where 20 non-conservative differences were observed with only 7 conservative changes (7). It should, however, be noted that in-frame CGA to TGA change occurred 2 amino acid residues upstream of the termination site of liver type proteins in type 1 mRNAs (Fig.1). As a result, type 1 mRNA products lack c-terminal arginine-glycine residues. Two isoelectric variants identified by in vitro translation using salivary poly(A)$^+$RNAs (8) may be explained by the difference between type 1 and type 2 mRNAs, since amino acid sequences of other parts are identical.

In salivary glands, a diverse set of $\alpha 2\mu$ globulin mRNAs are thus produced by multiple splicing and polyadenylations from at least two genes. Less obvious but similar phenomenon was also found in the liver type genes in a similar gene region (13), indicating a common transcription-splicing feature of $\alpha 2\mu$ globulin gene family. Physical and biological consequences of mRNA diversification is a matter of speculation at present. Since gene regions around 6th intron-7th exon of $\alpha 2\mu$ globulin are thus accommodated with many splice donor and acceptor sites that are used alternatively with variable efficiencies in vivo, they should provide interesting systems of natural source for in vitro splicing studies and for an analysis of tissue specificity in splicing reactions.

ACKNOWLEDGMENTS

We thank Dr.M.Furuichi, Mr.N.Kinoshita and Ms.N.Mochizuki for assistance. This work was supported in part by Grants-in-Aid from the Ministry of Education, Science and Culture, Japan.

*To whom correspondence should be addressed

REFERENCES

1. Roy,A.K. and Neuhaus,O.W. (1966) Proc.Natl.Exp.Biol.Med. 121, 894-899
2. Kurtz,D.T. (1981) J.Mol.Appl.Genet. 1, 29-38
3. Dolan,K.P., Unterman,R.D., MacLaughlin,M., Nakhasi,H.L., Lynch,K.R. and Feigelson,P. (1982) J.Biol.Chem.257,13527-13534
4. Feigelson,P. and Kurtz,D.T.(1978) Adv.Enzymol. 47, 275-310
5. Sippel,A.E.,Feigelson,P.and Roy,A.K. (1975) Biochemistry 14. 825-829
6. Kurtz,D.T.,Sippel,A.E.,Ansah-Yiadom,R.and Feigelson,P. (1976) J.Biol.Chem.251,3594-3598
7. Laperche,Y.,Lynch,K.R.,Dolan,K.P. and Feigelson,P.(1983) Cell 32, 453-460
8. Gubits,R.M., Lynch,K.R., Kulkarni,A.B., Dolan,K.P., Gresik,E.W., Hollander,P. and Feigelson,P. (1984) J.Biol. Chem. 259, 12803-12809
9. Murty,C.V.R., Sarkar,F.H., Mancini,M.A. and Roy,A.K. (1987) Endocrinology 121, 1000-1005
10. Antakly,T., Laperche,Y. and Feigelson,P. (1982) J.Histochem. Cytochem. 30, 1293-1296
11. Barka,T. (1980) J.Histochem.Cytochem. 28, 836-859
12. Gubler,U. and Hoffman,B.J. (1983) Gene 25, 263-269
13. Ichiyoshi,Y., Endo,H. and Yamamoto,M. (1987) Biochim.Biophys. Acta 910, 43-51
14. Vieira,J. and Messing,J. (1987) Methods Enzymol.153,3-11
15. Mizusawa,S. Nishimura,S. and Seela,F (1986) Nucl.Acids Res. 14, 1319-1324
16. Thomas,P.S (1980) Proc.Natl.Acad.Sci.USA 77, 5201-5205
17. Feinberg,A.P. and Vogelstein,B. (1983) Anal.Biochem.132, 6-13
18. Kato,N., Pfeifer-Ohlsson,S., Kato,M., Larsson,E., Rydnert,J., Ohlsson,R. and Cohen,M. (1987) J.Virol. 61, 2182-2191
19. Perricaudet,N., Akusjaervi,G., Virtanen,A. and Pettersson,U. (1979) Nature 281, 694-696
20. Krainer,A.R. and Maniatis,T. (1988) in Transcription and Splicing (Hames,B.D and Glover,D.M., eds), pp131-206, IRL Press., England
21. Meijlink,F., Curran,T., Miller,A.D.and Verma,I.M.(1985) Proc. Natl.Acad.Sci.USA 82, 4987-4991
22. Treisman,R. (1985) Cell 42, 889-902
23. Bergmann,I.E. and Brawerman,G. (1980) J.Mol.Biol. 139,439-454
24. Shaw,G. and Kamen,R. (1986) Cell 46, 659-667

This article, submitted on disc, has been automatically converted into this typeset format by the publisher.