Volume 10 Number 14 1982 Nucleic Acids Research

# Cloning and characterisation of the abundant cytoplasmic 7S RNA from mouse cells

Allan Balmain, Robert Krumlauf, J.Keith Vass and George D.Birnie

Beatson Institute for Cancer Research, Wolfson Laboratory for Molecular Pathology, Garscube Estate, Switchback Road, Bearsden, Glasgow G61 1BD, UK

Received 16 April 1982; Revised and Accepted 18 June 1982

ABSTRACT

A cDNA library has been prepared from mouse embryo small RNAs and screened for the presence of clones complementary to the highly abundant cytoplasmic 7S RNA. One clone (pA6) was selected which hybridised exclusively with 7S RNA on a Northern blot prepared from cytoplasmic RNA run on high resolution polyacrylamide/urea gels. Sequence analysis of this clone has shown that at least 65 nucleotides at the 5' end of 7S RNA are extensively homologous with the highly repeated mouse B1 family. Heterologous hybridisations between the cloned mouse 7S sequence and RNAs prepared from rat, human and chick cells have shown that the non-B1 part of the 7S RNA molecule has been highly conserved during recent eucaryotic evolution. There are multiple copies of 7S RNA genes in the genomes of mouse, human, rat and chick cells, but substantial differences exist in copy number and genomic organisation in these organisms.

## INTRODUCTION

Eucaryotic cells contain a discrete number of small RNA molecules which are not related to the known ribosomal RNA species or tRNAs. These RNAs have a size distribution of about 100-300 nucleotides and exhibit the common properties of abundance, stability and a high degree of sequence conservation during recent eucaryotic evolution (1-4). The ubiquitous nature and high evolutionary conservation of these species suggests that they perform some important cellular function, the elucidation of which has proved to be extremely elusive.

Information collated from a number of studies over the past decade has led to the classification of the small RNAs into two groups, depending upon the nature of their 5' termini and the polymerase involved in their transcription (1, 5). The small nuclear RNAs (snRNAs) have capped 5' termini and are now generally thought to be transcribed by RNA polymerase II (6, 7). This has been confirmed, at least for the U1 snRNA, by the use of a cloned probe for the chick U1 gene (8). This class of small RNAs has been the focus



0305-1048/82/1014-4259$2.00/0

of much recent attention because of the proposed involvement of U1 and U2 sequences in RNA processing, in particular the removal of introns from pre-mRNA sequences (9-11).

The other class of small RNAs includes the species K, L and M and the 4.5S RNAs which are common to many mammalian cells (1, 12). All of these RNAs possess triphosphorylated 5' termini and are transcribed by RNA polymerase III (13, 14). The biological role of these RNA species is completely obscure, although the association which has been detected between 4.5S RNA and polyadenylated RNA in rodent cells has led to speculation that it may be involved in some stage of hnRNA processing (15, 16).

The most abundant RNA in this class is the cytoplasmic L RNA (also known and hereafter referred to as 7S RNA). Although 7S RNA was initially discovered as an integral component of oncornavirus particles it was subsequently identified as a major RNA species in mammalian and avian cells (3, 17). Until recently, technical limitations have largely prevented progress in the elucidation of the cellular function of 7S RNA. However, there has been a resurgence of interest in this topic with the advent of cloning technology and the demonstration that 7S RNA from human cells is partially homologous to a highly repeated 300 b.p. sequence element in human DNA (18). This element is the basic repeating unit in a family of related sequences which are present appoximately 300.000 times in the human genome. The Alu family, named after the Alu I restriction site present in most of its members (19) constitutes the dominant interspersed repeated sequence in the genomes of human and rodent cells (20, 21). As much as 25% of hnRNA in mammalian cells is complementary to the Alu sequence and can be isolated in the form of double stranded RNA duplexes (22, 23). It is possible that the ultimate solutions to the questions of 'function' for small RNAs and interspersed repeat sequences such as the Alu family may be intimately linked.

In this article, we address a number of questions related to the characterisation of 7S RNA in mouse cells. A cloned copy of 7S RNA has been isolated from a cDNA library prepared from small RNA molecules. Sequence analysis has been used to illuminate the relationship between 7S RNA and the highly repeated Alu family, while cross-species hybridisations have shown that although 7S RNA is highly conserved in avian, rodent and human cells, major changes have occurred in its genomic organisation during eucaryotic evolution.

## MATERIALS AND METHODS

All restriction endonucleases were purchased from Uniscience and were used under the conditions recommended by the manufacturers. Reverse transcriptase was obtained from the Division of Cancer Cause and Prevention, National Cancer Institute, Bethesda. DNA polymerase (Klenow fragment) was provided by Boehringer and T4 ligase was supplied by Biolabs. Sequencing reagents were obtained from sources described by Maxam and Gilbert (24). Guanidinium thiocyanate was from Fluka, and oligo(dT)-cellulose (T3 grade) was from Collaborative Research Inc. Thiol-sepharose was prepared from Sepharose CL4B (Pharmacia) as previously described (25). Nick translations were carried out using a nick-translation kit and $\alpha$-($^{32}$P)-dCTP supplied by the Radiochemical Centre, Amersham, England, who also provided ($^{3}$H)-uridine (55 Ci-mmole). Nitrocellulose was supplied by Sartorius.

### Isolation of polysomes from Friend erythroleukemia cells

The cell culture conditions for Friend cells (clone M2) have been described previously (26). When required, cells were labelled by addition of $^{3}$H-uridine (2mCi/litre culture medium). Cells were harvested in mid-log phase of growth and washed with cold phosphate buffered saline. Polysomes were isolated from the post-mitrochondrial supernatant of NP40-lysed Friend cells by pelleting through buffered 2M sucrose (27). Polysomes were further purified on 15-40% sucrose gradients. Only fractions sedimenting faster than 100S were used for the extraction of RNA (28). This procedure has been shown to provide polysomes which are free of detectable contamination with nuclear ribonucleoprotein particles (27, 28).

### Isolation of RNA

Total cellular RNA was isolated by a modification of the method of Ullrich et al (29). Briefly, freshly harvested cells were resuspended in solution containing 5 M guanidinium thiocyanate, 50 mM Tris-HCl, pH 7.0, 50 mM EDTA and 5% 2-mercaptoethanol. The solution was strongly homogenised to shear the DNA and was then layered over 1.2 ml cushions of 5.7M CsCl, 50 mM EDTA, pH 7.0. The gradients were centrifuged at 35,000 rpm and 15°C for 24 hours in a 6 x 4.2 ml swing-out rotor (International). The pelleted RNA was dissolved in sterile water and reprecipitated twice from 0.3 M sodium acetate, pH 7.0 with 3 volumes of ethanol at -20°.

Polysomal RNA was isolated using CsCl as previously described (28). Total cytoplasmic RNA was isolated from the post-nuclear supernatant of NP-40 lysed Friend cells (27). The supernatant was extracted with an equal volume of phenol/chloroform (1:1, v:v) saturated with NETS buffer (100 mM NaCl, 1 mM

EDTA, 10 mM Tris-HCl, pH 7.4, 0.5% sodium lauryl sulphate (SLS)) until no interface remained. The aqueous phase was extracted with chloroform and the RNA precipitated with ethanol.

The poly(A)$^+$ RNA was removed by two passages over oligo (dT)-cellulose in binding buffer (0.5 M NaCl, 10 mM Tris-HCl, pH 7.5, 0.1% w/v sarcosine, 1 mM EDTA) followed by elution with water (30). The poly(A)$^-$ RNA was precipitated with ethanol and stored at -20°C.

Construction of a cDNA library from small poly(A)$^-$ RNA

Polysomal poly(A)$^-$ RNA from 14-day mouse embryos (NIH/Swiss) was fractionated by sedimentation through a 15-30% sucrose gradient. The fractions containing 7-12S RNAs were further purified by centrifugation through a second gradient and RNA sedimenting at 7-12S was isolated. 20 µg of this RNA was transcribed in a total volume of 1 ml containing 1 mM dNTPs, 1 mCi $^3$H-dCTP (20 Ci/mmole), 50 mM Tris-HCl pH 8.2, 10 mM Mg acetate, 10 mM dithiothreitol (DTT) 40 mM NaCl, 200 µg calf thymus primers (prepared according to 31), 100 µg actinomycin D and 60 units of AMV reverse transcriptase. The mixture was incubated for 2 hrs at 37°C and the reaction terminated by two chloroform extractions. The aqueous layer was then adjusted to 0.3 M NaOH and incubated at 37°C for 30 min. The mixture was neutralized by the addition of 3 M Na acetate and yeast tRNA carrier was added to 10 µg/ml. The sample was run over a Sephadex G50 column and the excluded peak concentrated by ethanol precipitation.

Second-strand synthesis was performed in 2 ml containing 1 mM dNTPs, 30 mM Tris-HCl, pH 7.5, 5 mM $MgCl_2$, 1 mM 2-mercaptoethanol, 16.8 µg cDNA, and 200 units of E.coli DNA polymerase (Klenow fragment) at 12°C for 7 hrs. To produce flush ends the mix was adjusted to 30 mM Na acetate pH 4.6, 50 mM NaCl, 1 mM $ZnSO_4$ and 150 units of S1 nuclease in 3 ml and incubated at 37°C for 3 hrs. The double-stranded cDNA was harvested by ethanol precipitation following the addition of yeast tRNA carrier (to 20 µg/ml). These cDNAs were inserted into the phosphatased plasmid vector pAT153 by blunt-end ligation to the repaired Bam HI site and used to transform E.coli HB101 as previously described (32). Transformants were picked and inoculated into 96-well microtitre plates containing L-broth with 100 µg/ml ampicillin and glycerol, then stored at -20°C according to Gergen et al (33).

Selection of clones complementary to abundant, non-ribosomal RNAs

Filter replicas of the recombinant colonies in the microtitre plates were prepared on nitrocellulose filters and hybridised according to Grunstein and Hogness (34). One set of filters was hybridised with $^{32}$P-cDNA

transcribed from purified 18 and 28S rRNA. A duplicate set of filters was hybridised with $^{32}$P-cDNA from the 7-12S poly(A)$^+$ RNA used to prepare the cDNA library. Upon comparison, clones strongly positive with the small RNA probe and negative with the ribosomal probe were selected and subjected to a second round of hybridisation. Colonies which gave a positive signal with only the small RNA cDNA probe both times were picked and plasmid DNA was isolated on a small scale (2 ml cultures) from each. The plasmids were labelled with ($^{32}$P)-dCTP by nick translation and hybridized to total mouse RNA blots to further eliminate t-RNA, 5S and 5.8S rRNA species. Those probes hybridizing to non-ribosomal RNA species were selected for further analysis.

DNA sequencing

DNA sequencing was carried out using the methods of Maxam and Gilbert (24). Briefly, pA6 was cut by Bam H1 and 5' end labelled using polynucleotide kinase and ($\gamma$-$^{32}$P) ATP. The insert was purified by agarose gel electrophoresis and strand separated on a 6% polyacrylamide gel. The fragment was sufficiently short to be sequenced from both strands.

Gel electrophoresis and blotting

RNA was fractionated on 8% polyacrylamide gels containing 7M urea using sample and running buffers exactly as described by Maniatis and Efstratiadis (35). For analytical separations, gels normally measured 200 x 180 x 1.5 mm and were run at 200 volts for 3-4 hours or 40 volts overnight. The RNA was visualised with ethidium bromide (2 mg/l water, for 10 minutes) and transferred to nitrocellulose in 20 x SSC without any further treatment. In some cases (see text), RNA was fractionated in 1% agarose gels containing 2.2 M formaldehyde in MOPS buffer (20 mM Na-morpholinopropanesulfonic acid, 5 mM Na acetate, 1 mM EDTA). The RNA was denatured in MOPS buffer containing 50% formamide and 2.2 M formaldehyde at 55° for 15 minutes before loading the gel. Agarose gels were run at 100 volts for 5-6 hours or 35 volts overnight. The RNA was transferred directly to nitrocellulose in 20 x SSC without staining or further treatment. The sizes of RNAs which hybridised with ($^{32}$P)-labelled probes were estimated by comparison with the mobilities of RNA size markers (E.coli 23S and 16S rRNAs, eucaryotic 18S and 28S rRNAs and yeast tRNA) run in parallel and stained with acridine orange (36).

DNA was isolated by the method of Gross-Bellard et al (37), restricted and run on 1% agarose gels in 50 mM Tris, 20 mM sodium acetate, 1 mM EDTA, pH 8.0 for 16-18 hours at 40 volts. Transfer to nitrocellulose after staining with ethidium bromide was carried out according to Wahl et al (38).

Blot hybridisation

RNA or DNA blots on nitrocellulose were pre-hybridised for at least 12 hours at 43° and then hybridised for 20-24 hours with ($^{32}$P)-labelled probes in buffer containing 50% formamide and 10% dextran sulphate as described by Wahl et al (38). Unless otherwise stated in the text, the final washing of blots was carried out in 0.1 x SSC, 0.1% (w/v) SLS at 60°C.

Hybrid-selection using mercurated plasmid DNA

A total of 50 µg pA6 DNA was precipitated with ethanol to remove traces of chloride ion and mercurated essentially as described by Longacre and Mach (39). The DNA was dissolved in water (250 µl) to which was added an equal volume of buffer containing 100 mM sodium acetate, pH 6.0 and 40 mM mercuric acetate. The reaction mix was incubated for 18 hours at 50°. EDTA was added to a final concentration of 50 mM and the mercurated DNA was purified by exclusion on Sephadex G50 and ethanol precipitation.

Mercurated plasmid DNA (10 µg) was dissolved in 20 µl of hybridisation buffer (50% formamide, 20 mM Tris-HCl, pH 7.8, 0.3 M NaCl, 1 mM EDTA) together with 15 µg of cytoplasmic RNA labelled with $^{3}$H-uridine to a specific activity of 20,000 cmp/µg. The hybridisation buffer included 2-mercaptoethanol (2 mM) as ligand for the mercury groups on the DNA (25). The hybridisation mix was denatured at 70° for 10 minutes in a siliconised glass capillary, then incubated at 37° for 18 hrs. The capillary was flushed out with 1 ml NETS buffer. A thiol sepharose column (1 ml) was activated with 500 mM Tris, pH 8.0, 50 mM dithiothreitol and extensively washed with NETS buffer as previously described (25). The column was then washed for at least 1 h. at 60° with 97% formamide in TNE (20 mM Tris-HCl, pH 7.8, 300 mM NaCl, 1 mM EDTA) before application of the sample (1-2 ml/hour) at 20°. The unbound RNA was washed off the column with 10 volumes of NETS buffer and 5 volumes of 20% formamide in TNE before elution of the hybridised RNA with 97% formamide in TNE at 65°. At most 3% of the input RNA was isolated in the bound fraction. Both bound and unbound fractions were precipitated with 3 volumes of ethanol in the presence of 10 µg/ml yeast tRNA carrier.

## RESULTS

### Screening of a cDNA library for clones complementary to 7S RNA

A cDNA library was prepared from mouse embryo small RNA molecules as described under Materials and Methods. We adopted a screening strategy which would enable us to select clones complementary to abundant, non-ribosomal small RNAs. A number of clones were obtained which gave a strong positive

signal when hybridised with a cDNA probe prepared from small RNAs by random-primed reverse transcription, but were negative when screened with ribosomal cDNA. On the basis of preliminary data, one of these recombinants, designated pA6, was selected for further characterisation.

It has previously been shown by others that 7S RNA is an abundant cytoplasmic species in mammalian cells (18, 40, 41). Figure 1A clearly confirms that a band corresponding in mobility to 7S is the predominant poly(A)$^-$ cytoplasmic RNA after the known ribosomal and tRNA species. When a Northern blot of this RNA was hybridised with nick-translated DNA prepared from pA6, the only discrete band observed on the autoradiograph coincided with 7S RNA (Figure 1B). While this evidence was compatible with complementarity between pA6 and 7S RNA it was by no means conclusive: a number of other small, non-ribosomal RNA species, for example histone mRNAs, could have been selected by the same screening criteria. We therefore carried out a hybridisation-selection experiment designed to demonstrate complementarity between pA6 and the major 7S RNA species. pA6 DNA was mercurated and hybridised in excess with cytoplasmic RNA. The resulting DNA-RNA hybrids were isolated by thiol-sepharose chromatography and the purified RNA eluted in 97% formamide. Both the flow-through from the column and the purified RNA which hybridised with pA6 were precipitated using yeast tRNA

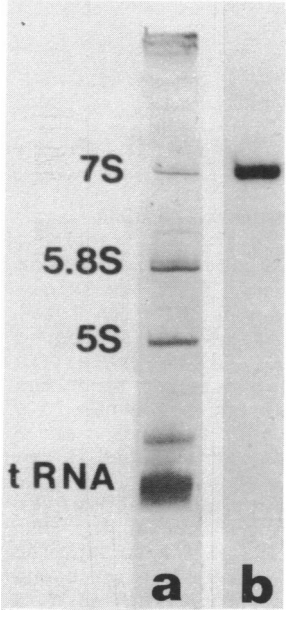


Figure 1. Demonstration of complementarity between 7S RNA and pA6 DNA. (a) shows the cytoplasmic poly(A)$^-$ RNA species from mouse Friend cells after separation on an 8% polyacrylamide/7M urea gel. (b) shows an autoradiograph of a Northern blot of the RNA shown in (a) after hybridisation with nick-translated pA6 DNA.

carrier and run on an 8% polyacrylamide/urea gel. The stained gel in Figure 2A shows that the 7S RNA was substantially depleted in the unbound fraction, but clearly visible in the bound component. This result indicates that pA6 has removed most of the 7S RNA from solution and must therefore be complementary to the major 7S RNA species. The additional lower molecular weight band which was observed in the eluted RNA (lower arrow, Figure 2A) is

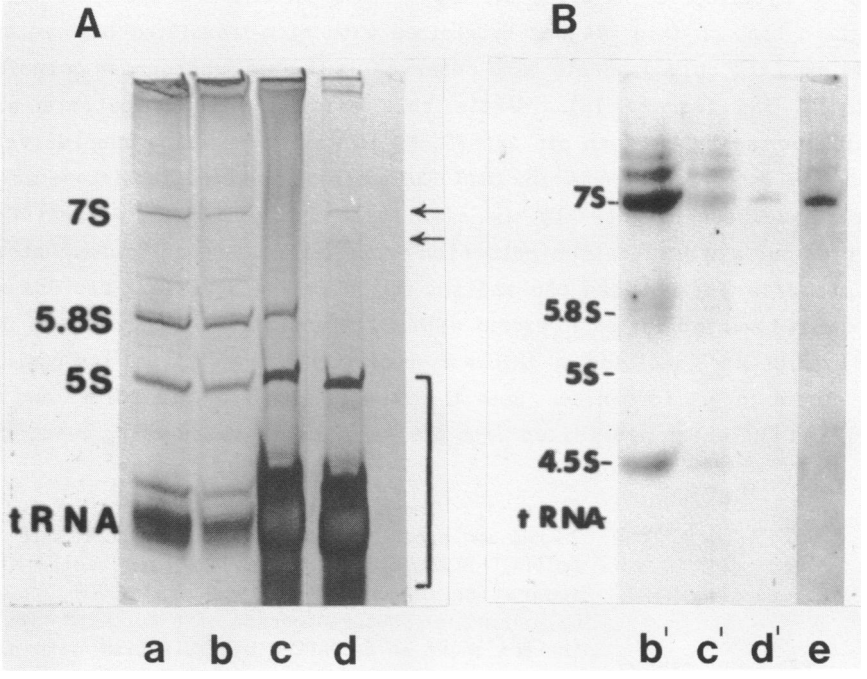


Figure 2. Hybridisation-selection of 7S RNA using mercurated pA6 DNA.
(A) A Polyacrylamide/urea gel was run on the following RNA fractions: (a) cytoplasmic poly(A)$^-$ RNA. (b) total cytoplasmic RNA used for hybridisation with mercurated pA6 DNA. (c) flow-through from the thiol-sepharose column after hybridisation of mercurated pA6 DNA with total cytoplasmic RNA. The 7S RNA is substantially depleted in this unbound fraction in comparison with the input RNA shown in (b). (d) RNA present in the bound fraction from the thiol sepharose column. Two distinct bands can be seen (arrows), one of which corresponds to 7S RNA and the other to a possible breakdown product. The strong 5S and tRNA bands in lanes c and d (denoted by the square bracket) are from the yeast tRNA carrier used in the precipitations.
(B) Slots b', c' and d' show an autoradiograph obtained after hybridisation of a nick-translated B1 probe with a Northern blot of the RNAs shown in slots b, c and d respectively of Figure 2A. The 7S RNA is depleted but still detectable in the unbound fraction (c') and enriched in the bound fraction (d'). Slot e shows a longer exposure of the autoradiograph in d'.

probably a specific breakdown product of 7S RNA since it hybridised with pA6 DNA but is not present in freshly prepared cytoplasmic RNA.

Further confirmation that pA6 is a 7S cDNA clone came from hybridisation experiments involving a member of the highly repeated mouse B1 family. Weiner (18) has demonstrated partial complementarity between a limited region of 7S RNA from HeLa cells and a cloned member of the human Alu family. In the mouse genome, the Alu-equivalent is a 125 b.p. sequence represented at least 100,000 times and collectively known as the B1 family (21). One member of this family has recently been shown to be situated about 2.5 kb from the 3' end of the mouse $\beta^{maj}$ globin gene (42). A specific probe for the B1 sequence was therefore conveniently obtained by excising the appropriate fragment from a BglII/HpaI digest of cloned genomic $\beta^{maj}$ globin DNA. When a $^{32}$P-labelled B1 probe was hybridised with a blot of cytoplasmic poly(A)$^-$ RNA, the strongest signal was observed at the position of 7S RNA (Figure 2B, slot b'). Moreover, the 7S RNA which had been selected with pA6 was capable of hybridising with the B1 probe, as shown by Figure 2B (slots d' and e). This evidence, together with the hybridisation data in Figure 1, supports the identification of pA6 as a 7S cDNA clone.

The autoradiograph in Figure 2B (slot b') shows that a series of cytoplasmic RNAs of slightly higher molecular weight than 7S also have homology with a B1 probe. Longer exposure of this autoradiograph enables the detection of at least five distinct RNA bands (data not shown). The pattern of hybridisation cannot be attributed to non-specific aggregation of the RNA, since this would also have been observed with the pA6 probe shown in Figure 1. It is possible that at least two of these RNAs may correspond to the species K and M which are present in both nucleus and cytoplasm and are also transcribed by RNA polymerase III (1, 13, 14, 41).

The nucleotide sequence of pA6 DNA

The insert cloned in pA6 contains about 190 of the 280 nucleotides represented in 7S RNA. Its sequence was determined by standard sequencing techniques and is presented in Figure 3. The salient feature of the sequence is the presence of a truncated B1 element at one end of the molecule. The first 65 nucleotides of the cloned cDNA are over 80% homologous with a specific region from bases 19-84 of the published B1 consensus sequence (21). This accords with the previous report of homology between human 7S RNA and the Alu repeat (18) and with the hybridisation data presented above. Li et al (43) have recently used RNA sequencing techniques to determine the sequence of rat 7S RNA which is also presented in Figure 3. A remarkable degree of

```
                          GCCTGTAGTT CCAGCTACTC GGGAGGCTGA GACAGGAGGA  Mouse 7S
pppGCCGGG CGGTGCGCAC ----------C ---------- ---------- ----------  Rat   7S
CCGGGCATGGTGGTGC  AT ----T--A-C --- - ---- -------A-- -G----C---  Mouse B1

                ****** ******
TCGCTTGAGT CCAAGAGTTC TGGGCTGTAG TGCGCTATGC CGATCGGGTG TCCGCACTAA  Mouse 7S
---------- ---G------ ---------- ---------- ---------- ----------  Rat   7S
-TT-- ---- T-G--   -  ---T--                                       Mouse B1
              /   \
             GCCAGC

                                                      .
GTTCGGCAAC AATATGGTGA CCTCCCGGGA GCGGGGGACC ACCANGTTGC CTAAGGAGGG  Mouse 7S
--------T- ---------- -------A     -------- ----G----- ----------  Rat   7S

GTGAAC  GA CCAGG CGGA A                                            Mouse 7S
------CG-C -----T---- -                                            Rat   7S
```

Figure 3. The nucleotide sequence of pA6 DNA.
The sequence of the cloned insert in pA6 is shown together with the homologous region of the B1 consensus sequence (21), and the rat 7S RNA sequence (43). A potential polymerase III recognition site in the 7S sequence is denoted by asterisks.

homology exists between the mouse and rat sequences. The first 160 nucleotides of the mouse sequence show only 5 base changes and a 4 nucleotide insertion with respect to the rat 7S RNA. The slightly lower homology over the 20 bases at the 3' end of the mouse 7S cDNA clone may be attributable to the use of random primers for the reverse transcription of the 7S RNA. The nucleotide substitutions shown in this region may therefore not be a reflection of genuine sequence differences between mouse and rat 7S RNAs.

The finding that pA6 contains a stretch of 65 bases with extensive homology to the B1 consensus sequence was initially surprising, since in a series of hybridisation experiments (data not shown) we had failed to detect any significant cross-hybridisation between a B1 probe and the cloned insert in pA6. This lack of cross-hybridisation is also evident from a comparison of Figures 1 and 2 which shows that pA6 hybridises only with one specific 7S RNA whereas the B1 probe has homology with a variety of RNAs of different molecular weights. We conclude that the relatively short hybrids which would be formed between pA6 and the B1 sequence are not sufficiently stable under the stringency conditions used for these hybridisation reactions.

Previous studies have shown that 7S RNA is transcribed by RNA polymerase III (13, 14). Comparison of the pA6 sequence with that of other known polymerase III transcripts such as the 4.5S RNA (16, 44) or the virus-associated small RNAs from adenovirus (45) or Epstein Barr virus (46) reveals the presence of a 12-nucleotide stretch which could function as a polymerase III recognition site (47). This sequence (denoted by asterisks in Figure 3) is within the B1 part of the molecule in the 5' half of 7S RNA, and accords

with data on other polymerase III transcripts in which the recognition site is also located within the RNA coding sequence (48-50).

Evolutionary conservation of the 7S RNA sequence in avian, rodent and human cells

Oligonucleotide fingerprinting has been used to show that avian and murine 7S RNA sequences are very similar (3), suggesting that this RNA species is involved in some important cellular function which has been subject to evolutionary conservation. The data shown in Figure 4 confirm this observation and extend it to a comparison of the RNAs in mouse, rat, avian and human cells. The overall pattern of small RNAs in these species was very similar (Figure 4A), although the chick "7S" RNA appeared to have a slightly higher molcular weight. The striking feature of the autoradiograph shown in Figure 4B is the extraordinary sequence conservation of 7S RNA. The pA6 probe hybridised equally well with the 7S band in all four species, despite extensive washing of the blot at 60° in 0.1 x SSC. Further washing at 60° in the absence of salt did however preferentially denature the hybrids formed with chick RNA (data not shown), suggesting that this species has a slightly more divergent sequence than the others. Preliminary evidence shows that the conservation of the sequence extends to the repetitive component of the 7S RNA, based on hybridisation of a mouse B1 probe with a similar spectrum of RNAs. The 7S RNA in all species hybridised, whereas the 4.5S RNA was only detected in rodent cells (51). This accords with the demonstration of Alu-type sequences in the genomes of all of these species (52).

The genomic organisation of 7S RNA sequences

The availability of a cloned 7S sequence which did not significantly hybridise with the highly repeated B1 element in mouse cells enabled us to begin a study of the organisation of 7S genes in eucaryotic cells. Southern blots were prepared from EcoRI-digests of genomic DNA from mouse, rat, avian and human cells and hybridised with nick-translated pA6 DNA. It was immediately obvious from the patterns shown in Figure 5 that major changes in the copy number and organisation of 7S genes have occurred during recent eucaryotic evolution. In an EcoRI digest of mouse DNA, about 8 bands of roughly equal intensity are observed. On the basis of the sequence data shown in Figure 3, a number of restriction endonucleases were selected which cut the 7S coding sequence at either one or two specific sites. Information obtained from genomic digests using these enzymes indicated that the 7S genes are probably not in a simple tandem array, but have a dispersed arrangement in the mouse genome (data not shown). In the rat genome, about 15 bands are

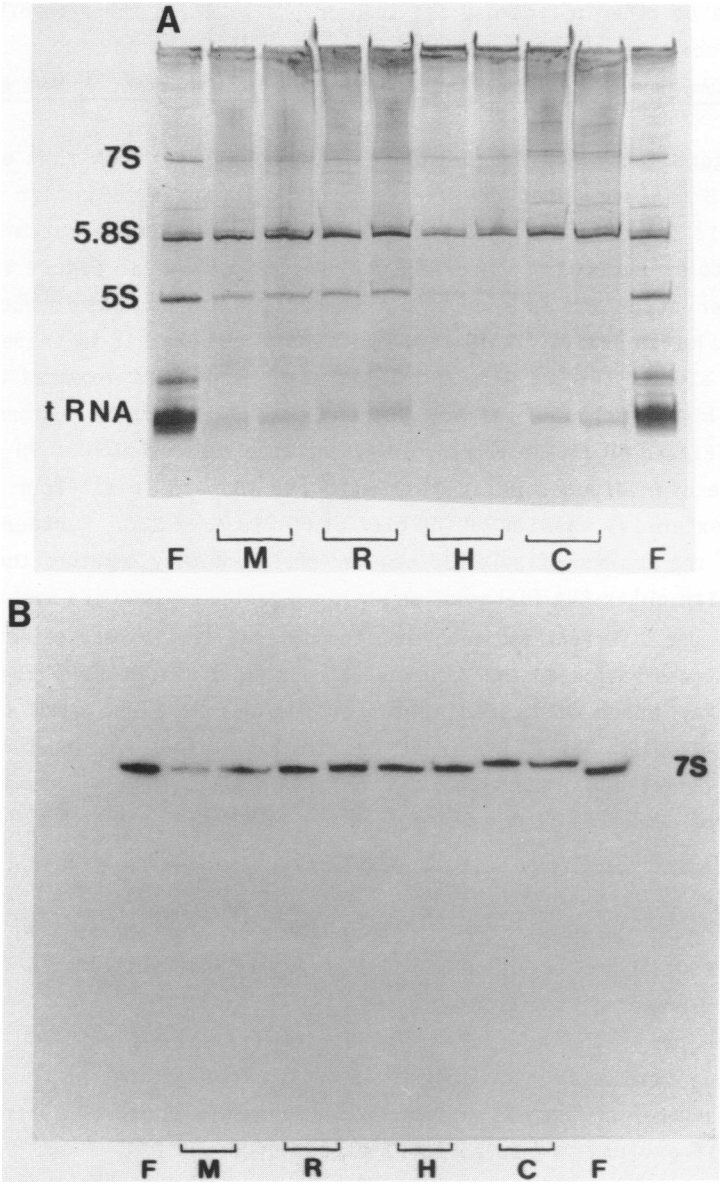


Figure 4. Evolutionary conservation of 7S RNA.

Total cellular RNA from mouse Friend cells (M), rat hepatoma cells (R), human fibroblasts (H) and chick embryos (C) was prepared and run on an 8% polyacrylamide/urea gel.

(A) stained gel of 15 and 30 µg amounts of each of these RNAs. The two outside lanes (F) each contain 30 µg of Friend cell poly(A)$^-$ cytoplasmic RNA.

(B) autoradiograph of the RNA shown in (A) after transfer to nitrocellulose and hybridisation with nick-translated pA6 DNA.

seen, some of which are of higher intensity and may contain multiple copies of the gene. The chick genome has about 4 EcoRI fragments containing 7S sequences, whereas neurospora DNA did not appear to contain any sequences capable of hybridising with the pA6 probe.

The most surprising aspect of this experiment was the apparent increase in copy number in the human genome. The smear of hybridisation observed is not an artefact specific to the DNA of the cell line used in these particular experiments: the pattern shown has subsequently been confirmed with DNA from three different human cell lines and from normal human peripheral blood.

## DISCUSSION

### The structure of 7S RNA

We have isolated and characterised a cDNA clone which is complementary to about 190 of the 280 nucleotides in 7S RNA. The 5' half of the RNA

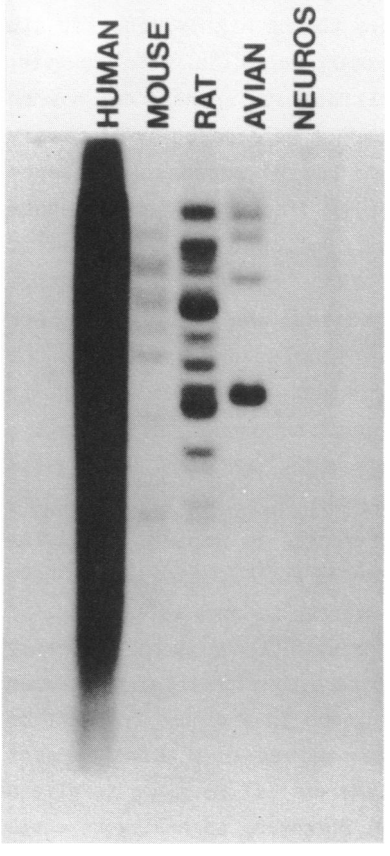


Figure 5. Genomic organisation of 7S RNA genes in eucaryotic cells. High molecular weight DNA (20 μg) from human fibroblasts, mouse embryos, rat liver, chick embryos and neurospora cells was restricted with EcoRI, run on a 1% agarose gel and transferred to nitrocellulose. The Figure shows the autoradiograph obtained after hybridisation of this blot with nick-translated pA6 DNA. The lowest band in the lane containing the avian DNA is due to the presence of a procaryotic contaminant in the DNA, as shown by hybridisation of an identical blot with the nick-translated plasmid pAT 153.

contains a stretch of 65 bases with extensive homology to the published consensus sequence of the mouse B1 family (21). On the basis of our hybridisation data, however, we conclude that an additional part of the B1 sequence must be present in mouse 7S RNA which has not been cloned in pA6: a B1 probe hybridises very strongly with 7S RNA but, under the same stringency conditions, very weakly with pA6 DNA. This conclusion is supported by the RNA sequencing data of Li et al (43) who have shown that rat 7S RNA contains sequences complementary to the Alu family at both 5' and 3' ends. The isolation and sequencing of genomic clones complementary to mouse 7S RNA, presently in progress in this laboratory, should further illuminate the relationship between the mouse and rat sequences.

Comparison of 7S RNA with other small RNA species

Some striking similarities exist between 7S RNA and a series of 4.5S RNAs found in rodents (15, 51). Both species have triphosphorylated 5' termini, are unmethylated and are transcribed by RNA polymerase III (13, 16). In addition, both are partially complementary to the highly repeated Alu family sequences in the genomes of human and rodent cells (18, 53). However, in spite of these obvious analogies, some critical differences can now be identified. First, although the 4.5S RNAs cannot be encoded by B1 family members, they exhibit extensive homology to the B1 consensus sequence throughout the whole length of the molecule (53). The 7S RNA, on the other hand, has a truncated B1 sequence at its 5' end, followed by a region of at least 120 bases with no homology to 4.5S RNA. Second, the relative proportions of 7S and 4.5S RNAs found in the nucleus and cytoplasm differ considerably: the vast majority of 7S RNA molecules are located in the cytoplasm whereas a substantial proportion (as much as 50%) of 4.5S RNA is found in the nucleus (15, and A.B. and R.K., unpublished results). Finally, while 7S RNA is equally abundant in mouse, rat, chick and human cells (this report), it has been reported that the 4.5S RNAs are only detectable in mice, rats or hamsters (51). These fundamental differences in sequence, cellular location and species of origin indicate that the 7S and 4.5S RNAs cannot have identical functions, although their roles may overlap to some extent.

Another RNA species which merits comparison with 7S RNA is that recently described by Haynes and Jelinek (53). These authors have previously reported a type 2 Alu-equivalent dispersed repeated sequence in a cloned fragment of Chinese hamster DNA (44). Subsequent studies showed that this fragment served as a template for transcription by polymerase III *in vitro* to give a series of discrete low molecular weight RNAs. Moreover, an analogous class

of RNAs has been found to exist *in vivo* (53). An interesting parallel between this class of RNA molecules and 7S RNA is that both have partial Alu-type sequences at their 5' ends. However, the points of overlap between the B1 and non-B1 parts of each molecule are different: 7S has at least 65 nucleotides homologous with residues 19-88 of the B1 sequence, whereas the region present in RNAs transcribed from type 2 Alu sequences contains residues 47 through 107 (Figure 6). The non-B1 parts of the molecules exhibit no obvious homology apart from a stretch of 18 nucleotides (indicated by boxed area in Figure 6) which has a counterpart with 75% sequence overlap in the type 2 Alu DNA. An additional difference between the two types of sequence is their relative copy numbers in rodent genomes. Haynes et al (44) have reported that the non-B1 part of the type 2 sequence is itself highly repeated in the Chinese hamster genome. Our studies with the cloned 7S cDNA indicate a maximum copy number of about 8-10 in the mouse genome, with slightly more being present in the rat.

An intriguing aspect of the comparison between 7S RNA and the Alu-type 2 sequence is that both contain potential polymerase III recognition sites at or very close to the junction between the B1 and non-B1 parts of the molecules. In the case of the Alu type 2 sequence, a 12 nucleotide stretch with very close homology to the consensus sequence described by Fowlkes and Shenk (47) actually spans the junction-point (see Figure 6). Similarly, an insertion of 2 T residues and a deletion of 5 nucleotides in what is normally a highly conserved area of the B1 sequence generates a new potential polymerase III recognition site in 7S RNA which ends within 3 nucleotides of the junction-point. Another similarity between the 7S sequence and the CHO Alu-type 2 sequence is that the short nucleotide stretch CCAGCCTGG, which is highly conserved in Alu-type repeats in human, rodent and chick genomes (52) has been rearranged either by insertion or deletion of a few nucleotides (Figure 6B). With regard to the overall structure of 7S RNA and CHO Alu type 2 RNAs, it is possible that transposition events involving complete or partial Alu type elements (44) might have given rise to new combinations of sequences, some of which become transcriptionally active and are stabilised within the genome.

Genomic organisation of 7S RNA sequences

The availability of a 7S cDNA clone which does not cross-hybridise with the highly repeated B1 sequences in the mouse genome has enabled us to show that the 7S RNA is encoded by a multigene family. An analogy might be drawn to the similar dispersed multigene families reported for the snRNAs (2, 8,

A

```
 GCCTGTAGTT CCAGCTACT CGGGAGGCTGA GACAGGAGGA TCGCTTGAGT        7S
(19)                            -- -G-T----AGA-G---C---G-      CHO Type II

     ***  **   * ******
 CCAAGAG  TT   C TGGGCTGTAG TGCGCTATGC CGATCGGGTG TCCGCACTAA   7S
 T------CCACCAA- --  --- -- CCA-AGG-C- T--GTTCAAT ---          CHO Type II
                                      ********** ***
                          (88)                (107)

 GTTCGGCAAC AATATGGTGA CCTCCCGGGA GCGGGGGACC ACCAAGTTGC        7S
    -A----- C-C------G -TCATAACA- TCTATAATGA GATCT-G---        CHO Type II


 CTAAGGAGGG GTGAACGACC AGGCGGAA                                7S
 -CTCTTCT-- TGTGCA--TA TATAT-G-                                CHO Type II
```

B

Comparison of a highly conserved 9 nucleotide sequence

```
C C A G C C T G G     B1 MOUSE
- - - - - - - - -     CHO TYPE I
- - - - - - - - -     ALU (HUMAN)
- - - - - - - - -     ALU (CHICK)
- - - C - - - - C     CHO TYPE II
        △
       A A
- [       ] - - -     7S
    △
   T T
```

Figure 6.

(A) Comparison of 7S sequence with CHO-Alu type 2 sequence. Potential polymerase III recognition sites are denoted by asterisks. Homologous nucleotides are indicated by dashes. The numbers in brackets refer to the numbers of residues in the B1 consensus sequence or the CHO Alu type 1 sequence (21, 44). pA6 DNA is homologous with residues 19-88 of the mouse B1 consensus sequence, and the CHO Alu type 2 sequence (44, 53) is homologous with residues 47 through 107. The boxed area denotes a region of homology between the non-B1 parts of the 7S cDNA clone and the CHO Alu type 2 sequence.

(B) Modifications to a highly conserved 9 nucleotide sequence in 7S and CHO Alu type 2 sequences. The 7S sequence has a deletion of 5 nucleotides and an insertion of two T residues into the 9-nucleotide stretch which is highly conserved in Alu-type ubiquitous repeats (52). The equivalent area of the CHO Alu type 2 sequence has two base changes and an insertion of two A residues.

54, 55). Many of the individual loci complementary to the snRNAs are in fact pseudogenes which have diverged during evolution and are not colinear with the RNA species (54-56). It is possible that a similar situation exists for the 7S RNA genes, particularly since detailed analysis using different restriction endonucleases which cleave the 7S cDNA clone at specific sites

has shown that not all of these restriction sites are shared by the 7S RNA genes in the mouse genome (R.K. and A.B., unpublished results). It is however interesting to note that RNA sequencing techniques have shown that some heterogeneity exists at the 5' and 3' ends of rat 7S RNA (43), suggesting that several of the gene-containing fragments in the rodent genomes may be transcribed. The vast increase in hybridisation signal which we have observed between the mouse 7S cDNA clone and the human genome is intriguing, and may be interpreted in terms of either a dramatic amplification of the 7S genes in human DNA or cross-hybridisation between the cloned mouse sequence and a sub-set of the highly repeated Alu family. Further investigation of the structure of 7S RNA-genes and pseudogenes, presently in progress in this laboratory, will be necessary to clarify this point.

Possible functions of 7S RNA

7S RNA is the most abundant cytoplasmic RNA in eucaryotic cells which has not yet been assigned any particular function. A prerequisite for the analysis of the function of an RNA species is the knowledge of its structure and its location within the cell. With regard to the 7S RNA, this has been the subject of some controversy. While there is a general consensus that 7S RNA is a predominantly cytoplasmic species, an initial report by Walker et al (57) that it is associated with polyribosomes in mouse and human cells was not confirmed in subsequent studies (40). More recently, Gunning et al (58) attributed this discrepancy to the previous authors' use of different methods for polysome isolation and concluded that 7S RNA is indeed involved in a functional association with polyribosomes. The latter studies were however based on estimation of 7S RNA concentrations in various cellular compartments by scanning of gel profiles. The availability of a cloned 7S sequence will enable us to resolve these discrepancies by accurately determining the concentration and localisation of 7S RNA within the cell. We also hope to identify and analyse specific ribonucleoprotein particles containing 7S RNA. These investigations should indicate whether the function of 7S RNA lies in the areas of RNA processing, nucleo-cytoplasmic transport of mRNA species or translational control.

Note added. After submission of this manuscript, a report appeared by Ullu and Melli on the isolation of cDNA clones complementary to human 7S RNA (Nucleic Acids Res. 10, 2209-2223 (1982)).

ACKNOWLEDGEMENTS

We thank Mr. John McVey for his help in the preparation and screening of the cDNA library. Excellent technical assistance was provided by Gail Cole and Lesley Frew. Friend cell cultures were kindly supplied by Mr. J. Sommerville. Reverse transcriptase was obtained from the Division of Cancer Cause and Prevention, National Cancer Institute. We thank Dr. Harris Busch for communication of the rat 7S RNA sequence prior to publication. The Beatson Institute is supported by grants from the Cancer Research Campaign and the Medical Research Council.

REFERENCES

1. Zieve, G.W. (1981). Cell 25, 296-297.
2. Wise, J.A. and Weiner, A. (1980). Cell 22, 109-118.
3. Erikson, E., Erikson, R.L., Henry, B. and Pace, N.R. (1973). Virology 53, 40-46.
4. Branlant, C., Krol, A., Ebel, J.P., Lazar, E., Gallinaro, H., Jacob, M., Sri-Widada, J. and Jeanteur, P. (1980). Nucl. Acids Res. 8, 4143-4154.
5. Lerner, M.R. and Steitz, J.A. (1981). Cell 25, 298-300.
6. Gram-Jensen, E., Larsen, P.H. and Frederickson, S. (1979). Nucl. Acids Res. 6, 321-330.
7. Eliceiri, G.L. (1980). J. Cell Physiol. 102, 199-207.
8. Roop, D.R., Kristo, P., Stumph, W.E., Tsai, M.J. and O'Malley, B.W. (1981). Cell 23, 671-680.
9. Rogers, J. and Wall, R. (1980). Proc. Natl. Acad. Sci. USA 77, 1877-1879.
10. Lerner, M.R., Boyle, J.A., Mount, S.M., Wolin, S.L. and Steitz, J. (1980). Nature 283, 220-224.
11. Ohshima, Y., Itch, M., Okada, N. and Miyata, T. (1981). Proc. Natl. Acad. Sci. USA 78, 4471-4474.
12. Ro-Choi, T.S., Reddy, R., Henning, D., Takano, T., Taylor, C.W. and Busch, H. (1972). J. Biol. Chem. 247, 3205-3222.
13. Zieve, G., Benecke, B.-J. and Penman, S. (1977). Biochemistry 16, 4520-4525.
14. Reichel, R. and Benecke, B.J. (1980). Nucl. Acids Res. 8, 225-234.
15. Jelinek, W.R. and Leinwand, L. (1978). Cell 15, 205-214.
16. Harada, F. and Kato, N. (1980). Nucl. Acids Res. 8, 1273-1285.
17. Bishop, J.M. (1978). Ann. Rev. Biochem. 47, 35-88.
18. Weiner, A.M. (1980). Cell 22, 209-218.
19. Houck, C.M., Rinehart, F.P. and Schmid, C.W. (1979). J. Mol. Biol. 132, 289-306.
20. Jelinek, W.R., Toomey, T.P., Leinwand, L., Duncan, C.H., Biro, P.A., Choudary, P.V., Weissman, S.M., Rubin, C.M., Houck, C.M., Deininger, P.L. and Schmid. C.W. (1980). Proc. Natl. Acad. Sci. USA 77, 1398-1402.
21. Krayev, A.S., Kramerov, D.A., Skryabin, K.G., Ryskov, A.P., Baev, A.A. and Georgiev, G.P. (1980). Nucl. Acids Res. 8, 1201-1215.
22. Jelinek, W., Evans, R., Wilson, M., Salditt-Georgiev, M. and Darnell, J.E. (1978). Biochemistry 17, 2776-2783.
23. Kramerov, D.A., Grigoryan, A.A., Ryskov, A.P. and Georgiev, G.P. (1979). Nucl. Acids Res. 6, 697-713.
24. Maxam, A.M. and Gilbert, W. (1979). Methods in Enzymol. 65, 499-560.
25. Brown, T.D.K. and Balmain, A. (1979). Nucl. Acids Res. 7, 2357-2368.
26. Gilmour, R.S., Harrison, P.R., Windass, J.D., Affara, N. and Paul, J.

(1974). Cell Different. 3, 9-22.
27. Birnie, G.D., MacPhail, E., Young, B.D., Getz, M.J. and Paul, J. (1974). Cell Different. 3, 221-232.
28. Young. B.D., Birnie, G.D. and Paul, J. (1976). Biochemistry 15, 2823-2829.
29. Ullrich, A., Shine, J., Chirgwin, J., Pictet, R., Tischer, E., Rutter, W.J. and Goodman, H. (1977). Science 196, 1313-1319.
30. Balmain, A., Minty, A. and Birnie, G.D. (1980). Nucl. Acids Res. 8, 1643-1659.
31. Taylor, J.M., Illmensee, R. and Summers, J. (1976). Biochim. et Biophys. Acta 442, 324-330.
32. Affara, N.A., Goldfarb, P.S., Vass, K., Lyons, A. and Harrison, P.R. (1981). Nucl. Acids Res. 9, 3061-3074.
33. Gergen, J.P., Stern, R.H. and Wensink, P.C. 91979). Nucl. Acids Res. 7, 2115-2136.
34. Grunstein, M. and Hogness, D.S. (1975). Proc. Natl. Acad. Sci. USA 72, 3961-3965.
35. Maniatis, T. and Efstratiadis, H. (1980). Methods in Enzymol. 65, 299-305.
36. Carmichael, G.G. and McMaster, G.K. 91980). Methods in Enzymol. 65, 380-391.
37. Gross-Bellard, M., Oudet, P. and Chambon, P. (1973). Eur. J. Biochem. 36, 32-38.
38. Wahl, G.M., Stern, M., and Stark, G. (1979). Proc. Natl. Acad. Sci. USA 76, 3683-3687.
39. Longacre, S. and Mach, B. (1979). Methods in Enzymol. 65, 192-199.
40. Zieve, G. and Penman, S. (1976). Cell 8, 19-31.
41. Gurney, T. and Eliceiri, G.L. (1980). J. Cell Biol. 87, 398-403.
42. Coggins, L.W., Vass, J.K., Stinson, M.A., Lanyon, W.G. and Paul, J. (1982). Gene 17, 113-116.
43. Li, W.-Y., Reddy, R., Henning, D., Epstein, P. and Busch, H. (1982). J. Biol. Chem., in press.
44. Haynes, S.R., Toomey, R.P., Leinwand, L. and Jelinek, W.R. (1981). Mol. Cell Biol. 1, 573-583.
45. Akusjarvi, G., Mathews, M.B., Andersson, P., Vennstrom, B. and Petterson, U. (1980). Proc. Natl. Acad. Sci. USA 77, 2424-2428.
46. Rosa, M.D., Gottlieb, E., Lerner, M.R. and Steitz, J. (1981). Mol. and Cell Biol. 1, 785-796.
47. Fowlkes, D.M. and Shenk, J. (1980). Cell 22, 405-413.
48. Sakonju, S., Bogenhagen, D.F. and Brown, D.D. (1980). Cell 19, 13-25.
49. Bogenhagen, D.F., Sakonju, S. and Brown, D.D. (1980). Cell 19, 27-35.
50. Ford, P.J. (1980). Nature 287, 109-110.
51. Harada, F., Kato, N. and Hoshino, H. (1979). Nucl. Acids Res. 7, 909-916.
52. Stumph, W.E., Kristo, P., Tsai, M.-J. and O'Malley, B.W. (1981). Nucl. Acids Res. 9, 5383-5397.
53. Haynes, S.R. and Jelinek, W.R. (1980). Proc. Natl. Acad. Sci. USA 74, 6130-6134.
54. Denison, R.A., Van Arsdell, S.W., Bernstein, L.B. and Weiner, A.M. (1981). Proc. Natl. Acad. Sci. USA 78, 810-814.
55. Ohshima, Y., Okada, N., Tani, T., Itoh, Y. and Itoh, M. (1981). Nucl. Acids Res. 9, 5145-5158.
56. Manser, T. and Gesteland, R.F. (1981). J. Mol. Appl. Gen. 1, 117-125.
57. Walker, T.A., Pace, N.R., Erikson, R.L., Erikson, E. and Behr, F. (1974). Proc. Natl. Acad. Sci. USA 71, 3390-3394.
58. Gunning, P.W., Beguin, P., Shooter, E.M., Austin, L. and Jeffrey, P.L. (1981). J. Biol. Chem. 256, 6670-6675.