Biochem. J. (1980) **192**, 489–498
Printed in Great Britain

Relative distribution of post-nuclear poly(A)-containing RNA abundance groups within the nuclear and post-nuclear polyadenylated and non-polyadenylated RNA populations of the lactating guinea-pig mammary gland

Ian C. BATHURST,*† Roger K. CRAIG,*‡ David G. HERRIES§ and Peter N. CAMPBELL*
**Courtauld Institute of Biochemistry, The Middlesex Hospital Medical School, London W1P 7PN, U.K., and §Department of Biochemistry, University of Leeds, 9 Hyde Terrace, Leeds LS2 9LS, U.K.*

(*Received 13 March 1980/Accepted 30 June 1980*)

1. RNA isolated from the post-nuclear supernatant of the lactating guinea-pig mammary gland was fractionated with oligo(dT)–cellulose into three populations; those that bound at 'low salt' [long poly(A) tracts, 78–32 nucleotides]; those that bound at 'high salt' [shorter poly(A) tracts, 48–21 nucleotides]; and those that did not bind [no poly(A) or short poly(A) tracts, <20 nucleotides]. Nuclear RNA was fractionated into two populations, those that bound in 'low salt' and those that did not bind. All the post-nuclear RNA fractions directed the synthesis of milk proteins in a Krebs II ascites cell-free system. 2. ^{3}H-labelled DNA complementary to the post-nuclear poly-(A)-containing RNA population (low-salt fraction) was fractionated into abundant (milk-protein mRNA), moderately abundant and scarce sequences. This complementary DNA was then used to investigate the distribution of the mRNA sequences in the different RNA populations. This showed that all sequences were present in polyadenylated and non-polyadenylated fractions, but that major quantitative differences were apparent. The abundant milk-protein mRNA sequences predominated in the 'low-salt' post-nuclear poly(A)-containing RNA fraction, whereas the moderately abundant sequences predominated in the non-polyadenylated post-nuclear RNA fraction. In total cellular RNA, those sequences deemed initially to be moderately abundant within the 'low-salt' poly(A)-containing RNA population were present at a concentration very similar to those of the abundant milk-protein mRNA (approx. 6×10^5 copies of each sequence/cell). Similarly, analysis of the nuclear RNA populations showed that the 'abundant' and so-called 'moderately abundant' sequences were present in essentially identical concentrations (2×10^3 copies of each sequence/cell). The majority of these (90–95%) were non-polyadenylated. 3. The results are discussed in terms of the post-transcriptional mechanisms involved in the regulation of gene expression in the lactating guinea-pig mammary gland.

With the exception in most instances of histone mRNA (see Ruderman & Pardue, 1978), the majority of well-characterized eukaryotic mRNA species occur in a form that is polyadenylated at the 3′-terminus. However, recent studies have demonstrated that, in addition to histone mRNA, mRNA-like sequences exist in a number of eukaryote cell types in an apparently non-polyadenylated form. This has been best demonstrated by comparative complexity analyses, using unique sequence DNA. These studies have shown that in mouse liver, and cultured mouse cells (Grady *et al.*, 1978), mouse brain (Van Ness *et al.*, 1979) and rat brain (Chikaraishi, 1979), mRNA sequences are subdivided into two distinct populations, one of which is polyadenylated and one which is not. Similar conclusions have also been drawn from pulse–chase experiments which demonstrate that between 30 and 80% of polyribosomal mRNA sequences are not polyadenylated (Milcarek *et al.*, 1974; Miller, 1978). Nemer *et al.* (1974) also provide evidence for a non-histone non-polyadenylated mRNA population in sea-urchin embryos.

Abbreviation used: cDNA, complementary DNA.

† Present address: Department of Biochemistry, University of Otago, Dunedin, New Zealand.
‡ To whom reprint requests should be addressed.

0306-3283/80/110489-10$01.50/1

However, the most abundant polyadenylated mRNA sequences appear to be present in both populations (Gray & Cashmore, 1976; Gedamu & Dixon, 1976; Kaufmann *et al.*, 1977; Hunter & Garrels, 1977; Ragg *et al.*, 1977; Ruderman & Pardue, 1977; Brandhorst *et al.*, 1979), as determined by cell-free protein synthesis. In some instances (Milcarek *et al.*, 1974; Kaufmann *et al.*, 1977) these conclusions have been substantiated by hybridization data, which demonstrate the presence of the abundant poly(A)-containing sequences within the non-polyadenylated population.

A similar analysis of nuclear transcripts has yet to be performed, though it is apparent that a large percentage of globin transcripts within the nucleus are not polyadenylated (Spohr *et al.*, 1976), even though these sequences exist in predominantly polyadenylated forms within the cytoplasm.

In the lactating rat and ewe mammary gland, up to 50% of the total milk-protein mRNA sequences do not bind to oligo(dT)–cellulose or poly(U)–Sepharose, as determined by cell-free protein synthesis (Houdebine *et al.*, 1974; Rosen *et al.*, 1975). In the lactating guinea-pig mammary gland the post-nuclear (cytoplasmic) poly(A)-containing RNA species occur in three distinct abundance populations (Craig *et al.*, 1979), of which the most abundant, the milk-protein mRNA sequences, represent in excess of 50% of the total polyadenylated sequences. In contrast, similar complexity analysis of the nuclear polyadenylated sequences (Bathurst *et al.*, 1980) showed that the milk-protein mRNA sequences, although relatively abundant, represent only 2% of the total nuclear polyadenylated sequences.

In the present paper we describe the distribution of the three post-nuclear poly(A)-containing RNA abundance groups between polyadenylated and non-polyadenylated RNA populations in total tissue, and in the nuclear and post-nuclear RNA fractions. The results demonstrate that over 95% of the nuclear transcripts under investigation are not polyadenylated, and that those sequences previously described as moderately abundant (Craig *et al.*, 1979), are present at a much higher concentration than was initially envisaged from studies using poly(A)-containing RNA alone.

Materials and methods

Materials

All materials, chemicals, enzymes and solvents were obtained from sources previously described (Craig *et al.*, 1976, 1979; Bathurst *et al.*, 1980). L-[4,5-^{3}H]Leucine (53 Ci/mmol), deoxy[5-^{3}H]cytidine 5′-triphosphate (22 Ci/mmol) and [5-^{3}H]uridine 5′-diphosphate (7–13.5 Ci/mmol) were purchased from The Radiochemical Centre, Amersham, Bucks., U.K. Avian-myeloblastosis-virus reverse transcriptase (lot no. G-678) was provided by Dr. J. W. Beard, Life Sciences Inc., St. Petersburg, FL 33707, U.S.A., and poly(A) standards of defined chain length were from Miles Laboratories, Slough, Bucks. SL2 4LY, U.K. Safeguards against RNA degradation by exogenous ribonucleases were as described previously (Craig *et al.*, 1976, 1979).

Preparation of post-nuclear RNA and fractionation on the basis of poly(A) content

Post-nuclear RNA was isolated from the lactating guinea-pig mammary gland 3–6 days *post partum* as described previously (Craig *et al.*, 1976). The bulk of the poly(A)-containing RNA population was then isolated by affinity chromatography on oligo(dT)–cellulose as described previously (Craig *et al.*, 1976). Material that did not bind under 'low-salt' ionic conditions (100 mM-NaCl) was recovered by ethanol precipitation, washed once in ethanol, dried under vacuum, and redissolved at 10–20 A_{260} units/ml in 10 mM-Tris/HCl, pH 7.6 at 20°C, containing 500 mM-KCl. An additional poly(A)-containing RNA fraction was then isolated from this by further chromatography on oligo(dT)–cellulose at room temperature after pre-equilibration of the column in the same buffer. Bound material ['high-salt' poly(A)-containing RNA] was eluted at 60°C with 10 mM-Tris/HCl, pH 7.5. To this was added 0.1 vol. of 2 M-NaCl, and the RNA was precipitated at −20°C after the addition of 2.5 vol. of ethanol. The RNA that did not bind to oligo(dT)–cellulose (the non-polyadenylated RNA) was precipitated at −20°C after the addition of 2.5 vol. of ethanol. Precipitated RNA was recovered by centrifugation at 10000 $g_{av.}$ for 10 min at −10°C, washed once in 70% (v/v) ethanol/40 mM-NaCl, once in ethanol, dried briefly under vacuum, dissolved in double-distilled water, and stored in batches at −70°C.

Preparation of nuclear RNA and fractionation on the basis of poly(A) content

Poly(A)-containing RNA was isolated by affinity chromatography on oligo(dT)–cellulose from nuclear RNA of the lactating guinea-pig mammary gland as described previously (Bathurst *et al.*, 1980). RNA that did not bind after two passages through oligo(dT)–cellulose at 100 mM-NaCl (the non-polyadenylated RNA) was recovered by ethanol precipitation, washed once in 70% ethanol/40 mM-NaCl, once in ethanol, dried briefly under vacuum, dissolved in double-distilled water and stored in batches at −70°C. Yields of poly(A)-containing RNA represent 0.2–0.3% of total nuclear RNA.

Preparation of total cellular RNA

Total cellular RNA was isolated from the lactating guinea-pig mammary gland by the method of

Hall *et al.* (1979), except that the CsCl-centrifugation step was omitted, and DNA was removed by digestion with ribonuclease-free deoxyribonuclease as described by Bathurst *et al.* (1980).

Synthesis of cDNA and fractionation into abundance groups

High-specific-radioactivity ^{3}H-labelled DNA complementary to the post-nuclear 'low-salt' poly(A)-containing RNA population of the lactating guinea-pig mammary gland was synthesized as described by Craig *et al.* (1979), by using [^{3}H]-dCTP as the radio-labelled deoxyribonucleoside triphosphate. This was fractionated into the abundant, moderately abundant and scarce populations, by differential hybridizations followed by separation of double- and single-stranded nucleic acid species on hydroxyapatite as described by Bathurst *et al.* (1980). The final cDNA preparations had a specific radioactivity of 3.37×10^7 d.p.m./μg.

Hybridization of RNA preparations with cDNA

Hybridizations were carried out in 0.24 M-sodium phosphate buffer, pH 6.8, containing 0.2% (w/v) sodium dodecyl sulphate and 1 mM-EDTA over an RNA concentration range of 1 μg/ml to 1 mg/ml (1 A_{260} unit $\equiv$ 40 μg of RNA) for periods of 10 s to 72 h. The extent of hybridization was determined by S_1-nuclease digestion precisely as described previously (Craig *et al.*, 1979). Results represent the best-line fit as determined by computer analysis (Craig *et al.*, 1979), except where hybridizations were adjudged not to have gone to completion (<60%). In such instances, the 'best fit' was assessed by eye.

Determinations of poly(A) content of RNA populations and of the size of poly(A) tracts

[^{3}H]Poly(U) was synthesized as described by Bishop *et al.* (1974), by using polyribonucleotide phosphorylase from *Micrococcus luteus*. Hybridization to poly(A)-containing RNA or poly(A) was also as described by Bishop *et al.* (1974).

The size distribution of poly(A) tracts within given RNA populations was determined essentially as described by Rosbash & Ford (1974), except that after digestion of the RNA populations by deoxyribonuclease and ribonucleases A and T_1 the incubation mixture was extracted once with phenol/chloroform (1:1, v/v) and the remaining poly(A) precipitated with 2.5 vol. of ethanol at −20°C after the addition of 0.1 vol. of 2 M-NaCl and 20 μg of *Escherichia coli* tRNA. The precipitate was recovered by centrifugation at 16000 $g_{av.}$ for 20 min at −10°C, washed once in 70% ethanol/40 mM-NaCl, once in ethanol, and then dried under vacuum. The samples were then dissolved in 50 μl of 10 mM-Tris/HCl, pH 7.8, containing 5 mM-NaH_2PO_4 and 2 mM-EDTA, then subjected to electrophoresis on 7.5%-polyacrylamide tube gels as described by Loening (1967) at 5 mA/tube for 120 min at 4°C. Poly(A) standards of 20, 45 and 90 average monomer length were run on parallel gels.

The gels were frozen at −20°C, then sliced into 3 mm sections. Each slice was then crushed in 2 ml of 30 mM-sodium citrate containing 0.3 M-NaCl, and the poly(A) eluted overnight at 4°C. Polyacrylamide fragments were removed by centrifugation at 13000 $g_{av.}$ for 15 min at 4°C, and the resulting supernatant assayed for poly(A) content in the presence of excess [^{3}H]poly(U) (Rosbash & Ford, 1974).

Cell-free protein synthesis and product analysis

mRNA-directed cell-free protein synthesis was carried out in a Krebs II ascites-cell system in the presence of crude reticulocyte initiation factors, with [^{3}H]leucine as the labelled amino acid as described previously (Craig *et al.*, 1976). Antibody-precipitation procedures and sodium dodecyl sulphate/polyacrylamide-gel-electrophoretic analysis of the [^{3}H]leucine-labelled products performed on vertical slab gels were as described previously (Craig *et al.*, 1976, 1979). The products were detected by fluorography (Bonner & Laskey, 1974).

RNA and DNA measurements

RNA was determined by the method of Mejbaum (1939) and DNA by the method of Burton (1956).

Results

Fractionation of post-nuclear RNA isolated from the lactating guinea-pig mammary gland on the basis of poly(A)-tract size, and the identification of active milk-protein mRNA within these fractions

Total post-nuclear RNA from lactating guinea-pig mammary gland was separated into three fractions: those RNA species that bound to oligo(dT)–cellulose at 100 mM-NaCl ('low-salt'), those that bound at 500 mM-KCl ('high salt'), and those that did not bind at 500 mM-KCl. The ability of total RNA and of each fractionated RNA population to direct protein synthesis was then determined in a Krebs II ascites-cell system supplemented by crude reticulocyte initiation factors. The results (Fig. 1) demonstrate that, although each fraction had the ability to direct the synthesis of milk proteins as determined by antibody precipitation of the products synthesized *in vitro*, over 30% of the translatable milk-protein mRNA did not bind to oligo(dT)–cellulose. Not unexpectedly, the fractions that bound to oligo(dT)–cellulose were most active in the Krebs II ascites system. However, whatever the source of mRNA activity, milk proteins represented greater than 85% of the total protein synthesized *in vitro*

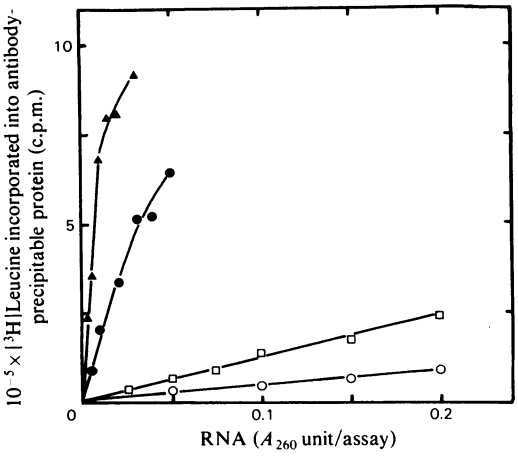

Fig. 1. *Identification of polyadenylated and non-polyadenylated milk-protein mRNA sequences from the lactating guinea-pig mammary gland by cell-free protein synthesis*

Total RNA was isolated from the post-nuclear supernatant of the lactating guinea-pig mammary gland, then fractionated on oligo(dT)–cellulose into three fractions: those RNA sequences that bound at 100mM-NaCl or at 500mM-KCl, and those that did not. Recoveries were 0.36, 0.18 and 70% of the starting material. Increasing amounts of RNA from each fraction, and also the total RNA before fractionation, were then assayed in a Krebs II ascites-cell-free system supplemented with crude reticulocyte initiation factors for milk-protein mRNA activity as determined by antibody precipitation of the proteins synthesized *in vitro* (see the Materials and methods section). ▲, 'Low-salt' (100mM-NaCl) poly(A)-containing RNA; ●, 'high-salt' (500mM-KCl) poly(A)-containing RNA; □, total RNA; ○, non-polyadenylated RNA.

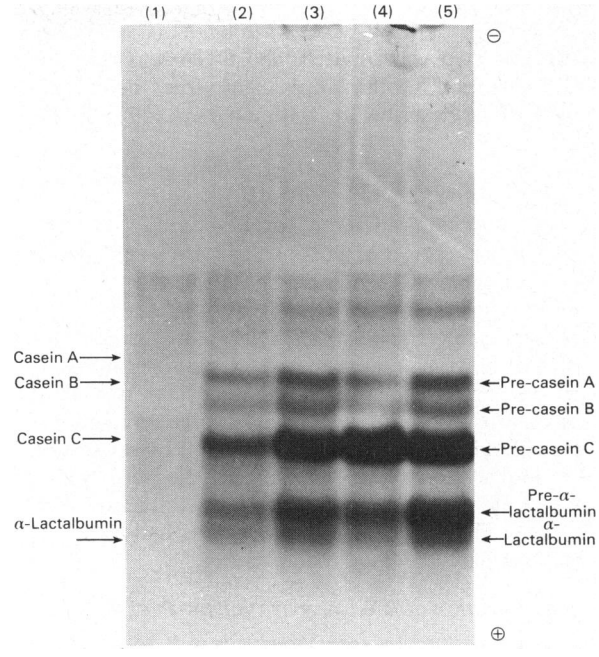

Fig. 2. *Sodium dodecyl sulphate/polyacrylamide-gel electrophoresis of milk proteins synthesized in a Krebs II ascites-cell-free system*

[^{3}H]Leucine-labelled milk proteins synthesized as described in Fig. 1 by addition of the post-nuclear RNA fractions to a Krebs II ascites-cell-free protein-synthesizing system, and then recovered by antibody precipitation, were separated on the basis of molecular weight by sodium dodecyl sulphate/polyacrylamide-slab-gel electrophoresis, and the ^{3}H-labelled milk proteins detected by fluorography. The relative positions of marker proteins, guinea-pig caseins A, B and C, and the whey protein α-lactalbumin (2μg each) as identified by Coomassie Blue staining are indicated. Lane (1), no added RNA; (2), non-polyadenylated RNA; (3), total post-nuclear RNA; (4), RNA binding to oligo(dT)–cellulose at 500mM-KCl; (5), RNA binding to oligo(dT)–cellulose at 100mM-NaCl.

[results not shown; see Craig *et al.* (1979) for discussion of this phenomenon]. Analysis of the antibody-precipitable material (Fig. 2) by sodium dodecyl sulphate/polyacrylamide-gel electrophoresis followed by fluorography confirmed that all fractions directed the synthesis of the three caseins and α-lactalbumin. Moreover, it was also apparent that pre-α-lactalbumin had been partially processed to α-lactalbumin (see Craig *et al.*, 1976).

The size of the poly(A) tracts present in the total post-nuclear RNA fraction and in the poly(A)-containing RNA fractions that bound at 100mM- and 500mM-salt was then determined after digestion of each RNA population with a combination of ribonucleases T_1 and A followed by polyacrylamide-gel electrophoresis of the undigested polyribonucleotides in parallel with poly(A) marker species of known size. This revealed a broad distribution of poly(A) tracts, ranging in size from 11 to 90 nucleotides within the total post-nuclear RNA fraction (Fig. 3). Analysis of those RNA sequences that bound to oligo(dT)–cellulose showed that, although increasing the salt concentration resulted in an increased binding efficiency of RNA species containing smaller poly(A) tracts ('low-salt', 78–32 nucleotides; 'high-salt', 48–21 nucleotides), RNA species containing 20 or less adenosine residues were not retained on oligo(dT)–cellulose. Thus in subsequent work all reference to non-polyadenylated RNA strictly refers to RNA preparations containing poly(A) tracts of less than 20 nucleotides.

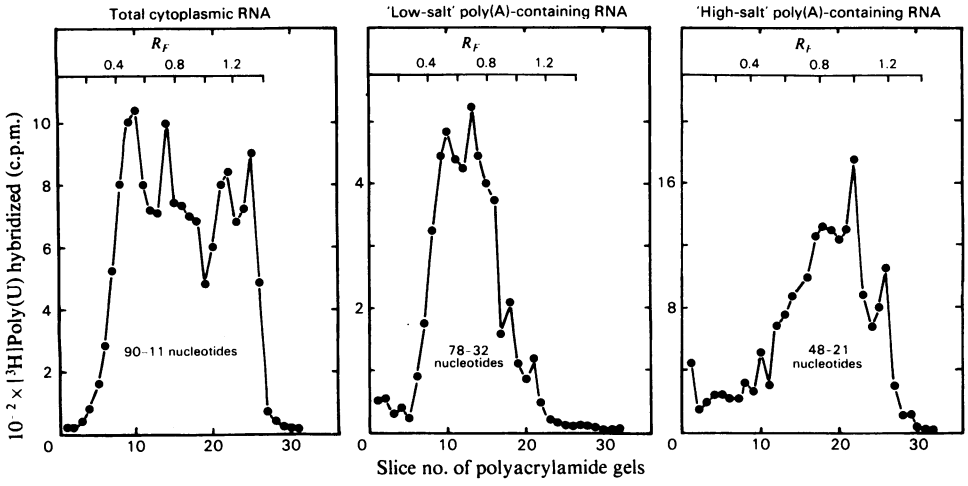

Fig. 3. *Analysis of the size distribution of poly(A) tracts within the post-nuclear RNA population of the lactating guinea-pig mammary gland*

Total post-nuclear RNA, and poly(A)-containing RNA which bound to oligo(dT)–cellulose during sequential fractionation at 100mM-NaCl and 500mM-KCl, were digested with ribonucleases A and T_1. The size of undigested poly(A)-containing material was then determined by electrophoresis on 7.5%-polyacrylamide tube gels with Bromophenol Blue as a reference marker. Poly(A) molecular-size markers of known monomer length were electrophoresed on parallel gels. The gels were sliced, the poly(A) was eluted, and the relative mobility of the poly(A) tracts determined by hybridization to [^{3}H]poly(U). The mobility of the poly(A) standards relative to Bromophenol Blue ($R_F = 1.0$) was as follows: $(rA)_{90(av.)}$, 0.36; $(rA)_{45(av.)}$, 0.69 $(rA)_{25(av.)}$, 0.97. The size distribution of the unknown poly(A) tracts was then determined from the relationship between relative mobility and logarithm of the number of nucleotides.

Similar analysis of the two nuclear RNA fractions showed that those sequences that bound at 'low-salt' contained poly(A) tracts ranging in size from 45 to 90 nucleotides, whereas those that did not bind contained poly(A) tracts predominantly 20–30 nucleotides in length (I. C. Bathurst & R. K. Craig, unpublished work). The latter we refer to as the non-polyadenylated nuclear RNA population.

Differential distribution of the abundant, moderately abundant and scarce RNA sequences found in post-nuclear ('low-salt') poly(A)-containing RNA within the polyadenylated and non-polyadenylated post-nuclear and nuclear RNA species

In order to quantify the total cellular distribution of the three mRNA abundance groups previously identified in the 'low-salt' post-nuclear poly(A)-containing RNA (Craig *et al.*, 1979), and in particular the distribution of these sequences between polyadenylated and non-polyadenylated RNA populations, ^{3}H-labelled DNA complementary to the 'low-salt' post-nuclear poly(A)-containing RNA was synthesized, and then fractionated into the three individual abundance classes.

The kinetics of hybridization of each cDNA fraction was then examined by using total cellular RNA as the driver species. From the resulting $R_0t_{\frac{1}{2}}$ values, the number of copies of each sequence within the different abundance groups was estimated as described by Roop *et al.* (1978). The results (Fig. 4, Table 1) demonstrate that the 'abundant' milk-protein mRNA population was indeed the most numerous population. Assuming that two sequences were present (see Craig *et al.*, 1979), then we estimate that each was present at a concentration of up to 10^5 copies per cell. However, it was apparent that within the total cellular RNA population, those sequences previously described as moderately abundant within the poly(A)-containing RNA of the post-nuclear fraction (Craig *et al.*, 1979; Bathurst *et al.*, 1980) were present in very much greater numbers (100-fold) than previously envisaged, at a concentration of approx. 6×10^4 copies/sequence per cell. Estimation of the concentration of scarce mRNA sequences was less precise, as the hybridization reaction had not reached completion by a R_0t of 1000 mol of nucleotide·litre^{-1}·s^{-1}. It seemed likely, however, that these were present at a concentration of 60–100 copies/sequence per cell.

To investigate these somewhat unexpected observations, we have examined on a quantitative basis the relative amount of each abundance group in the nucleus and cytoplasm (post-nuclear fraction), and also the distribution of each between the polyadenylated and non-polyadenylated RNA populations of both subcellular fractions.

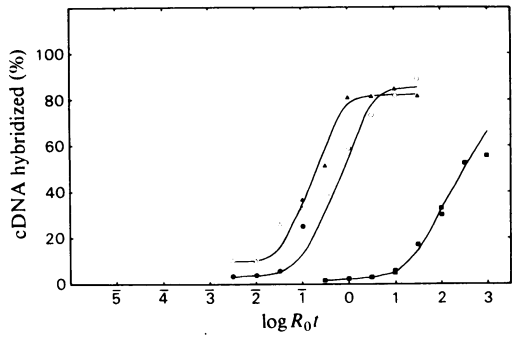

Fig. 4. *Analysis of the distribution of the abundant, moderately abundant and scarce 'low-salt' poly(A)-containing RNA sequences present in the post-nuclear supernatant of the lactating guinea-pig mammary gland, within the total cellular RNA population of the lactating guinea-pig mammary gland*

^{3}H-labelled cDNA was prepared against 'low-salt' poly(A)-containing RNA isolated from the post-nuclear RNA population of the lactating mammary gland. This was separated by differential hybridization and hydroxyapatite column chromatography (see the Materials and methods section) into three fractions representing the abundant, moderately abundant and scarce mRNA populations. Each fractionated cDNA population was then used to determine the distribution of abundant, moderately abundant and scarce mRNA sequences of the 'low-salt' poly(A)-containing RNA population within total cellular RNA of the lactating guinea-pig mammary gland: abundant cDNA at RNA concentrations of 50 μg/ml (△) and 1 mg/ml (▲); moderately abundant cDNA at RNA concentrations of 50 μg/ml (●) and 1 mg/ml (○); scarce cDNA sequences at RNA concentrations of 1 mg/ml (■).

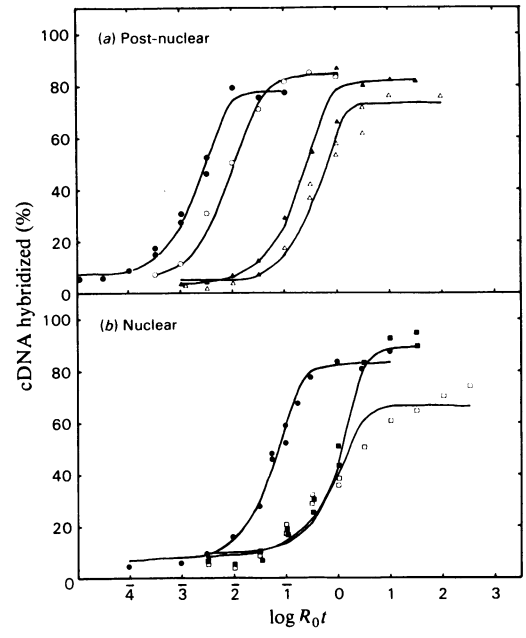

Fig. 5. *Distribution of the abundant 'low-salt' poly(A)-containing RNA sequences within the polyadenylated and non-polyadenylated RNA populations of the nuclear and post-nuclear RNA fractions of the lactating guinea-pig mammary gland*

Abundant cDNA was isolated as described in Fig. 4. This was then used to examine the distribution of the abundant milk-protein mRNA sequences within the nuclear and post-nuclear RNA populations of the lactating guinea-pig mammary gland. (*a*) Post-nuclear RNA. cDNA hybridized to: (i) 'low-salt' poly(A)-containing RNA (●) at RNA concentrations of 1 μg/ml and 2.5 μg/ml; (ii) 'high-salt' poly(A)-containing RNA (○) at 2.5 μg/ml and 50 μg/ml; (iii) total post-nuclear RNA (▲) at 2.5 μg/ml, 50 μg/ml and 1 mg/ml; (iv) non-polyadenylated RNA (△) at 2.5 μg/ml, 50 μg/ml and 1 mg/ml. (*b*) Nuclear RNA. cDNA hybridized to: (i) poly(A)-containing RNA (●) at 20 μg/ml and 200 μg/ml; (ii) total nuclear RNA (■) at 200 μg/ml and 1 mg/ml; (iii) non-polyadenylated RNA (□) at 200 μg/ml and 1 mg/ml.

Analysis of the abundant milk-protein mRNA sequences in this manner confirmed the distribution pattern within the post-nuclear fraction obtained by cell-free protein synthesis (Fig. 5*a*, Table 1). Thus about 55–60% of the sequences were present in the 'low-salt' poly(A)-containing RNA population, and 35–40% in the non-polyadenylated population, but only 3–5% were present in the 'high-salt' poly-(A)-containing RNA population.

Similar analysis of the abundant milk-protein mRNA sequences within the nuclear RNA population revealed a vastly different distribution (Fig. 5*b*). In this instance, of the 2200 copies/sequence identified in each nucleus (or 2–3% of the total cellular sequences), only 2–3% were polyadenylated. Furthermore, for reasons that we are unable to explain, although the $R_0t_{\frac{1}{2}}$ values were very similar (Table 1) whether total nuclear RNA or non-polyadenylated RNA was used as the driver species, we could not obtain greater than 70% protection of the cDNA from S_1-nuclease activity when non-polyadenylated RNA was the driver species.

Examination of the distribution within the nucleus of the moderately abundant 'low-salt' poly(A)-containing RNA population substantiated our findings described above. In this instance we could detect up to 2200 copies/sequence within the nucleus, of which 5–6% were polyadenylated (Fig. 6*b*, Table 1). Again, incomplete protection of the cDNA hybridization probe was obtained with non-polyadenylated RNA as the driver species.

Table 1. *Comparative distribution of the post-nuclear 'low-salt' poly(A)-containing RNA abundance groups within the polyadenylated and non-polyadenylated RNA populations of the nuclear and post-nuclear fractions of the lactating guinea-pig mammary gland*

cDNA representing the abundant, moderately abundant and scarce post-nuclear 'low-salt' poly(A)-containing RNA sequences was prepared by differential hybridization and subsequent chromatography on hydroxyapatite (see the Materials and methods section). Each was then hybridized in RNA excess to total nuclear and post-nuclear RNA preparations. Each preparation was performed on separate animals, all 3–5 days *post partum*. Calculations of the copy number per sequence for each RNA population was as described by Roop *et al.* (1978), and are based on (i) the assumption that the abundant, moderately abundant and scarce sequences consist of 2, 32 and 3242 species respectively, each with a R_0t pure of 1.31×10^{-3}, 2.76×10^{-2} and 2.74 mol of nucleotide·litre^{-1}·s^{-1} (Craig *et al.*, 1979) and (ii) the following measured parameters: DNA per cell, 5.8 pg; DNA per g of tissue (wet wt.), 2.5 mg; RNA per g of tissue (wet wt.), 10.7 mg; post-nuclear RNA per g of tissue (wet wt.), 9.7 mg; nuclear RNA per g of tissue (wet wt.), 1.0 mg.

Source of cDNA population	... Abundant (2)		Moderately abundant (32)		Scarce (3242)	
Source of RNA	$R_0t_{\frac{1}{2}}$ (mol·l^{-1}·s^{-1})	No. of copies/ sequence per cell	$R_0t_{\frac{1}{2}}$ (mol·l^{-1}·s^{-1})	No. of copies/ sequence per cell	$R_0t_{\frac{1}{2}}$ (mol·l^{-1}·s^{-1})	No. of copies/ sequence per cell
Total cellular	2.19×10^{-1}	100×10^3	4.55×10^{-1}	62.0×10^3	282*	60–100
Post-nuclear						
Total	2.88×10^{-1}	69.3×10^3	3.64×10^{-1}	70.3×10^3	—	—
Non-polyadenylated	5.73×10^{-1}	34.8×10^3	9.18×10^{-1}	27.8×10^3	—	—
'Low-salt' poly(A)-containing	2.08×10^{-3}	62.2×10^3	2.12×10^{-1}	7.85×10^2	9.04	18.5
'High-salt' poly(A)-containing	8.38×10^{-3}	4.76×10^3	4.98×10^{-2}	1.03×10^3	—	—
Nuclear						
Total	9.3×10^{-1}	2.23×10^3	2.1	1.26×10^3	118.2	22.5
Non-polyadenylated	9.55×10^{-1}	2.15×10^3	1.2	2.20×10^3	121.4	21.9
Poly(A)-containing	1.27×10^{-1}	44	6.06×10^{-2}	1.18×10^2	64.2	0.11

* Estimated $R_0t_{\frac{1}{2}}$ (see Fig. 4)

However, examination of the distribution of the moderately abundant sequences within the post-nuclear fraction demonstrated that, unlike the abundant milk-protein mRNA sequences, the majority of these sequences were associated with the non-polyadenylated RNA (Fig. 6*a*). Although in this instance quantitative estimations were not additive, possibly a reflection of the experimental approach (small variations in $R_0t_{\frac{1}{2}}$ estimations may give rise to apparently significant quantitative differences), it was apparent that the vast majority of these sequences were present in the non-polyadenylated fraction, possibly as much as 90–95%, as only 2×10^3 copies of each of the moderately abundant sequences could be identified within the two polyadenylated-RNA populations (Table 1). Moreover, the greater proportion of the polyadenylated sequences was associated with the 'high-salt' poly(A)-containing RNA population (Fig. 6*a*), those sequences with shorter poly(A) tracts.

The distribution of the scarce 'low-salt' poly(A)-containing RNA sequences within the nucleus again demonstrated that these sequences were predominantly non-polyadenylated. We estimate that in total about 20 copies per sequence were present per nucleus (Fig. 7, Table 1), but that polyadenylated scarce sequences were present in only one of every ten nuclei. Analysis of the distribution of these sequences within the post-nuclear fraction was less definitive, as only in one instance, the 'low-salt' poly(A)-containing RNA population, did hybridization go to completion (Fig. 7*a*). Determination of this population indicated the presence of up to 20 copies per sequence per cell (Table 1).

Discussion

In previous RNA-complexity studies we have used fractionated cDNA populations to investigate the relative distribution of poly(A)-containing mRNA sequences between nucleus and cytoplasm (Bathurst *et al.*, 1980). Here we have described the use of cDNA to determine on a quantitative basis the distribution of polyadenylated and non-polyadenylated RNA populations between nucleus and cytoplasm.

The distribution of the abundant milk-protein

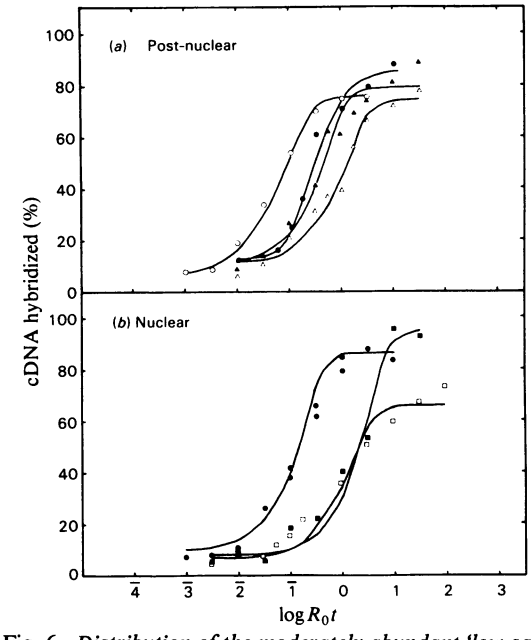

Fig. 6. *Distribution of the moderately abundant 'low-salt' poly(A)-containing RNA sequences within the polyadenylated and non-polyadenylated RNA populations of the nuclear and post-nuclear RNA fractions of the lactating guinea-pig mammary gland*

cDNA representative of moderately abundant sequences of the 'low-salt' poly(A)-containing RNA population was isolated as described in Fig. 4. This was then used to examine the distribution of the moderately abundant milk-protein mRNA sequences within the nuclear and post-nuclear RNA populations of the lactating guinea-pig mammary gland. (*a*) Post-nuclear RNA. cDNA hybridized to: (i) 'low-salt' poly(A)-containing RNA (●) at RNA concentrations of 20 μg/ml and 200 μg/ml; (ii) 'high-salt' poly(A)-containing RNA (○) at 20 μg/ml and 200 μg/ml; (iii) total post-nuclear RNA (▲) at 20 μg/ml and 200 μg/ml; (iv) non-polyadenylated RNA (△) at 20 μg/ml and 200 μg/ml. (*b*) Nuclear RNA. cDNA hybridized to: (i) poly(A)-containing RNA (●) at 20 μg/ml and 200 μg/ml; (ii) total nuclear RNA (■) at 200 μg/ml and 1 mg/ml; (iii) non-polyadenylated RNA (□) at 200 μg/ml and 1 mg/ml.

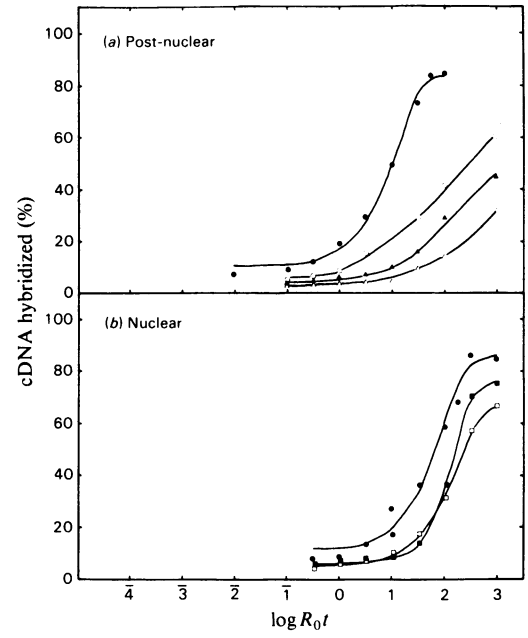

Fig. 7. *Distribution of the scarce 'low-salt' poly(A)-containing RNA sequences within the polyadenylated and non-polyadenylated RNA populations of the nuclear and post-nuclear RNA fractions of the lactating guinea-pig mammary gland*

cDNA representative of the scarce sequences of the low-salt poly(A)-containing RNA population was isolated as described in Fig. 4. This was then used to examine the distribution of the scarce milk-protein mRNA sequences within the nuclear and post-nuclear RNA populations of the lactating guinea-pig mammary gland. (*a*) Post-nuclear RNA. cDNA hybridized to: (i) 'low-salt' poly(A)-containing RNA (●); (ii) 'high-salt' poly(A)-containing RNA (○); (iii) total post-nuclear RNA (▲); (iv) non-polyadenylated RNA (△). (*b*) Nuclear RNA. cDNA hybridized to: (i) poly(A)-containing RNA (●); (ii) total nuclear RNA (■); (iii) non-polyadenylated RNA (□). All hybridizations were at RNA concentrations of 1 mg/ml.

mRNA in the post-nuclear fraction, as assessed by either cell-free protein synthesis or RNA-complexity analysis, is similar to the distribution reported by others for the rat (Rosen *et al.*, 1975) and for the ewe (Houdebine *et al.*, 1974) lactating mammary glands. It is apparent that bimorphic distribution between polyadenylated and non-polyadenylated post-nuclear RNA populations is a characteristic feature of the most abundant mRNA sequences in many systems (Gedamu & Dixon, 1976; Gray & Cashmore, 1976; Hunter & Garrels, 1977; Kaufmann *et al.*, 1977; Ruderman & Pardue, 1977; Milcarek, 1979). The reason for this distribution remains obscure, but may reflect a role for poly(A) in the stability of mRNA (see Nudel *et al.*, 1976; Huez *et al.*, 1978). Consequently, it is possible that non-polyadenylated milk-protein mRNA sequences, and those with shorter poly(A) tracts, may represent mRNA sequences at different stages of maturation compared with those with long poly(A) tracts. A shortening of the poly(A) region of mouse globin mRNA during maturation supports this supposition (Merkel *et al.*, 1976).

Such a hypothesis is attractive when we turn to a consideration of the number and distribution of the so-called moderately abundant sequences. These sequences are found in a polyadenylated state on free polyribosomes (Craig *et al.*, 1979), and probably represent mRNA sequences which direct the synthesis of enzymes and structural proteins required for lipid biosynthesis and the general maintenance of the lactogenic state of the mammary gland. *In toto* these sequences are predominantly non-polyadenylated, and may comprise a class of mRNA sequences which have a high turnover rate when compared with the predominantly polyadenylated milk-protein mRNA sequences.

Substantiation of this hypothesis requires kinetic analysis yet to be performed on our system. However, it has been demonstrated that in the rat mammary gland (Guyette *et al.*, 1979), the milk-protein mRNA sequences are preferentially more stable. A similar phenomenon has been reported for globin sequences in a mouse erythroleukaemia cell line (Lowenhaupt & Lingrel, 1978).

In spite of the incomplete examination of the scarce post-nuclear abundance population, it seems probable that this population is also bimorphic. Thus, of the estimated 60–100 copies/sequence per cell, only 20 copies/sequence could be accounted for in the nucleus, and a further 20 copies/sequence in the post-nuclear 'low-salt' poly(A)-containing RNA fraction. Consequently it seems likely that the remaining sequences will be non-polyadenylated within the post-nuclear fraction. This is consistent with observations with a HeLa-cell system (Milcarek, 1979), which demonstrated that a considerable proportion of the scarce abundance class readily hybridizes with the non-polyadenylated fraction.

Studies on the nuclear distribution of the post-nuclear poly(A)-containing RNA abundance classes substantiate our previous studies based on poly-(A)-containing RNA alone, and confirm that post-transcriptional mechanisms operate, and thereby account for the preferential accumulation of the abundant and moderately abundant sequences within the cytoplasm (see Bathurst *et al.*, 1980). Our observation that the vast proportion of all nuclear sequences contain short or no poly(A) tracts is similar to those of Spohr *et al.* (1976), who investigated the distribution of avian globin mRNA sequences in the nuclei of avian erythroblasts. Furthermore, our estimation of 2000 copies of each abundant nuclear sequence/cell is similar to estimates of ovalbumin mRNA sequences within oestrogen-stimulated tubular-gland-cell nuclei (Roop *et al.*, 1978), a sequence known to be transcribed at an elevated rate in the presence of oestrogen.

Overall our analyses of nuclear RNA are consistent with the view that, in the lactating guinea-pig mammary gland, the milk protein sequences and a separate set of about 30 sequences are transcribed at an elevated rate relative to the scarce sequences (see also Bathurst *et al.*, 1980). Each then undergoes differential post-transcriptional processing, resulting in one predominantly polyadenylated and one predominantly non-polyadenylated mRNA population.

Our observations imply that polyadenylation may be a late nuclear event. This assumes that our suppositions about the differential turnover of the mRNA sequences are correct, and that polyadenylation occurs before transport of processed sequences from the nucleus to the cytoplasm [see Houdebine (1976) for a different view]. This last assumption is substantiated to a limited extent by the observed incomplete protection of the cDNA probes by non-polyadenylated nuclear RNA. Whether this is due to the absence of certain mRNA sequences or parts of sequences has yet to be established. However, the situation within the non-polyadenylated nuclear RNA population is certainly complicated. We have based our analysis and subsequent quantification on single-component curves. Additional analysis using individual cDNA sequences, as opposed to multiple-component abundance groups, may well demonstrate that the situation within this fraction is indeed more intricate than we have portrayed. However, immaterial of the outcome, the overall conclusions outlined in the preceding paragraph remain unaltered.

Our studies demonstrate that the accumulation of those sequences present in a polyadenylated form in the lactating guinea-pig mammary gland is a function of both transcriptional and post-transcriptional regulation. We have yet to establish whether there exists in the mammary gland an additional mRNA population, exclusively non-polyadenylated, as has been described for several tissues and cell lines (Grady *et al.*, 1978; Chikaraishi, 1979; Milcarek, 1979; Van Ness *et al.*, 1979).

We are grateful to Mrs. D. McIlreavy and Mr. D. Parker for skilled technical assistance, and the Medical Research Council and the Wellcome Trust for supporting this work. Avian-myeloblastosis-virus reverse transcriptase was the gift of Dr. J. Beard and Dr. J. Gruber.

References

Bathurst, I. C., Craig, R. K., Herries, D. G. & Campbell, P. N. (1980) *Eur. J. Biochem.* **109**, 183–191

Bishop, J. O., Rosbash, M. & Evans, D. (1974) *J. Mol. Biol.* **85**, 75–86

Bonner, W. M. & Laskey, R. A. (1974) *Eur. J. Biochem.* **27**, 83–88

Brandhorst, B. P., Verma, D. P. S. & Fromson, D. (1979) *Dev. Biol.* **71**, 128–141
Burton, K. (1956) *Biochem. J.* **62**, 315–327
Chikaraishi, D. M. (1979) *Biochemistry* **18**, 3249–3256
Craig, R. K., Brown, P. A., Harrison, O. S., McIlreavy, D. & Campbell, P. N. (1976) *Biochem. J.* **160**, 57–74
Craig, R. K., Boulton, A. P., Harrison, O. S., Parker, D. & Campbell, P. N. (1979) *Biochem. J.* **181**, 737–756
Gedamu, L. & Dixon, G. H. (1976) *J. Biol. Chem.* **251**, 1455–1463
Grady, L. J., North, A. B. & Campbell, W. P. (1978) *Nucleic Acids Res.* **5**, 697–712
Gray, R. E. & Cashmore, A. R. (1976) *J. Mol. Biol.* **108**, 595–608
Guyette, W. A., Matusik, R. J. & Rosen, J. M. (1979) *Cell* **17**, 1013–1023
Hall, L., Craig, R. K. & Campbell, P. N. (1979) *Nature (London)* **277**, 54–56
Houdebine, L. M. (1976) *FEBS Lett.* **66**, 110–113
Houdebine, L. M., Gaye, P. & Favre, A. (1974) *Nucleic Acids Res.* **1**, 413–426
Huez, G., Marbaix, G., Gallwitz, D., Weinberg, E., Devos, R., Hubert, E. & Cleuter, Y. (1978) *Nature (London)* **271**, 572–573
Hunter, T. & Garrels, J. I. (1977) *Cell* **12**, 767–781
Kaufmann, Y., Milcarek, C., Berissi, H. & Penman, S. (1977) *Proc. Natl. Acad. Sci. U.S.A.* **74**, 4801–4805
Loening, U. E. (1967) *Biochem. J.* **102**, 251–257
Lowenhaupt, K. & Lingrel, J. B. (1978) *Cell* **14**, 337–344
Mejbaum, W. (1939) *Hoppe-Seyler's Z. Physiol. Chem.* **258**, 117–124
Merkel, C. G., Wood, T. G. & Lingrel, J. B. (1976) *J. Biol. Chem.* **251**, 5512–5515
Milcarek, C. (1979) *Eur. J. Biochem.* **102**, 467–476
Milcarek, C., Price, R. & Penman, S. (1974) *Cell* **3**, 1–10
Miller, L. (1978) *Dev. Biol.* **64**, 118–129
Nemer, M., Graham, M. & Dubroff, M. L. (1974) *J. Mol. Biol.* **89**, 435–454
Nudel, U., Soreq, H., Littauer, U. Z., Marbaix, G., Huez, G., Leclereq, M., Hubert, E. & Chantrenne, H. (1976) *Eur. J. Biochem.* **64**, 115–121
Ragg, H., Schroder, J. & Hahlbrock, K. (1977) *Biochim. Biophys. Acta* **474**, 226–233
Roop, D. R., Nordstrom, J. L., Tsai, S. Y., Tsai, M. J. & O'Malley, B. W. (1978) *Cell* **15**, 671–685
Rosbash, M. & Ford, P. J. (1974) *J. Mol. Biol.* **85**, 87–101
Rosen, J. M., Woo, S. L. C. & Comstock, J. P. (1975) *Biochemistry* **14**, 2895–2903
Ruderman, J. V. & Pardue, M. L. (1977) *Dev. Biol.* **60**, 48–68
Ruderman, J. V. & Pardue, M. L. (1978) *J. Biol. Chem.* **253**, 2018–2025
Spohr, G., Dettori, G. & Manzari, V. (1976) *Cell* **8**, 505–512
Van Ness, J., Maxwell, I. H. & Hahn, W. E. (1979) *Cell* **18**, 1341–1349