Volume 4 Number 12 December 1977 Nucleic Acids Research

The 5'-termini of heterogeneous nuclear RNA: a comparison among molecules of different sizes and ages

U.Schibler and R.P.Perry

The Institute for Cancer Research, Fox Chase Cancer Center, Philadelphia, PA 19111, USA

Received 9 September 1977

ABSTRACT

The composition of the 5' polyphosphorylated and capped termini of pulse labeled hnRNA from mouse L cells was analyzed by two-dimensional electrophoresis. Purine tri- and diphosphates: pppG, pppA, ppG and ppA; as well as four varieties of cap structure: m^7GpppX^m, $X^m = G^m$, $(m^6)A^m$, C^m and U^m, were detected. With increasing labeling time the relative proportion of hnRNA molecules with tri- and diphosphorylated 5' ends decreases and the relative proportion of capped hnRNA increases, indicating that caps are metabolically more stable than the polyphosphate termini. About half of the hnRNA molecules that are labeled within 2 hr have capped ends. This finding, together with results of earlier kinetic and structural studies, implies that a relatively high proportion of the labeled hnRNA molecules are mRNA precursors. Large hnRNA molecules exhibit a higher proportion of capped ends and a lower proportion of 5'-triphosphate ends as compared to small hnRNA. Given the lower stability of triphosphate termini relative to caps, this result may mean that capping of some hnRNA molecules can occur before the completion of transcription.

INTRODUCTION

Heterogeneous nuclear RNA (hnRNA) is a diverse mixture of molecules representing primary gene transcripts and processed derivatives of these transcripts. A substantial portion of the hnRNA molecules contain messenger RNA sequences in various stages of processing (1). A valuable diagnostic feature for sorting out the different forms of hnRNA is the character of the 5'-terminus. For example, primary transcripts are terminated by 5'-purine triphosphates (pppA··· or pppG···), whereas processed molecules may have other types of 5'-termini, such as diphosphates (ppX···), monophosphates (pX···) and caps (m^7GpppX^m···) (2-4). The cap-containing sequences appear to be conserved, and eventually become the 5'-termini of mature cytoplasmic mRNA (5,6). A consideration of the nucleotide compositions of the various 5'-termini in ^{32}P pulse-labeled total hnRNA of mouse L cells led us to propose two general processing pathways (2). For one class of transcripts thepppA··· orpppG··· termini would presumably be converted directly to ppA··· or ppG···, and then to the respective A^m and G^m-containing caps, while for another class, the primary transcription products would be cleaved at internal positions that ultimately become the sites for

subsequent capping reactions. This latter class would include all of the precursors of mRNAs whose cap sequences contain pyrimidines in the X^m position.

Since the extent of processing may also be related to the size and age of the hnRNA molecules, further insight into the nature of the processing pathways might be gained if one knew how the relative amounts of the different types of termini varied with size and age of the hnRNA. For this reason we have quantitatively analyzed the 5' termini of hnRNA that was fractionated on the basis of molecular size and hnRNA that was derived from cells labeled for varying periods with radioactive RNA precursors. In order to perform these analyses on relatively small amounts of radioactive material, we developed a compact two-dimensional electrophoretic method, which enabled us to separate, on one small plate, the relevant 5'-terminal derivatives bearing similar net charges. Our results demonstrate that the tri- and diphosphorylated termini are metabolically less stable than the corresponding cap termini, and that the proportion of triphosphorylated termini relative to cap in large (>35S) hnRNA is considerably less than in small hnRNA. These findings are consistent with the view that the 5' portions of some primary transcripts are capped, in some cases possibly even before the completion of transcription.

MATERIALS AND METHODS

Isolation of labeled nuclear RNA

Labeling of mouse L cells with $^{32}PO_4$ or the four ^{3}H nucleosides was performed as described in the appropriate figure and table legends. In all protocols the cells were incubated with 80 ng/ml actinomycin D immediately prior to, and during the incubation with radioactive precursors, in order to selectively suppress the synthesis of ribosomal RNA. Under these conditions the synthesis and capping of hnRNA occurs, and capped mRNA appears on cytoplasmic polyribosomes; however, the rate of mRNA processing may be somewhat retarded (5). RNA was extracted from detergent washed nuclei (7) according to Scherrer (8), or, where indicated, according to Hames and Perry (9). The purified RNA was subjected to stringent denaturation conditions and fractionated on sucrose gradients as described previously (5).

Enzyme digestions and analytical procedures

RNA was digested with T2 ribonuclease (0.5 - 1 unit/A_{260} unit) at 37°C for 14 - 16 hr in 1 ml of 0.02 M Naacetate, pH 4.5. After the digestion mixture was brought to pH 7 with Tris base, and EDTA added to 1 mM, 10 μg of ribonuclease T1 and 20 μg of ribonuclease A were added, and the incubation continued for 2 hr at 37°C. The samples were then adjusted to 7 M urea, 20 mM Tris, and directly applied to DEAE-Sephadex (urea) columns (2). The material eluting with

a net charge of -5 was pooled, diluted about 7-fold with H_2O and applied to a 1 ml column of DEAE-Sephadex in a 3 ml plastic syringe. After washing the column with 100-150 ml of 0.3% triethylammonium carbonate (TEAC), ~pH 6, the nucleotide derivatives were eluted in the cold with 30% TEAC, pH 7.6. The elimination of salt and urea residues by this extensive washing procedure was necessary in order to make the samples suitable for fingerprinting procedures.

The eluted nucleotides were dried in an Evapo-mix (Buchler Instrument Co.), redissolved in H_2O, dried again, and the process repeated until no visible residue was discernible. The nucleotide derivatives were then dissolved in 200 µl H_2O, transferred to 1.5 ml conical Brinkman tubes, lyophilized, redissolved in 10 µl 20 mM ammonium acetate, pH 5.3, containing 30 µg penicillium (P1) nuclease, and incubated for 1 hr at 50°C. The appropriate marker nucleotides were added, the incubation mixture lyophilized, the residue dissolved in 6 - 10 µl of H_2O, and a portion applied to cellulose acetate strips as described below.

Nucleotide pyrophosphatase digestions, used for the estimation of specific activity ratios of β and α phosphates, were carried out for 2 hr at 37°C in 200 µl 20 mM Tris, pH 7.6, 1 mM $MgCl_2$ containing 50 µg nucleotide pyrophosphatase (Sigma Chem. Co.). The nucleotide pyrophosphatase contained no detectable phosphomonoesterase activity as judged by control incubations with ^{32}P labeled 5' mononucleotides. The reaction mixtures were lyophilized, dissolved in 3 - 20 µl H_2O, and electrophoresed together with ATP, ADP, AMP and adenosine markers on Whatman 3MM paper (6). The marker nucleotides were located under UV, and the radioactive nucleotides detected either by autoradiography or by cutting the paper into 1 cm strips which were directly counted in Liquifluor (New England Nuclear Corp.). Partial digestions of ^{32}P-ATP, used to estimate γ-phosphate specific activity, were done in 50 µl volumes in the presence of 4 µM ATP, 10^5 cpm ^3H-ATP and 0.8 unit inorganic pyrophosphatase (Boeringer). Fifteen µl aliquots of the reaction mixture were withdrawn at 15, 30 and 60 min, quenched with 3 µl 10 mM EDTA, and analyzed as described above.

Two-dimensional electrophoresis

Separation of 5' terminal structures was performed according to Cory and Adams (10), except that 20 x 20 cm DEAE cellulose thin layer plates (Analtech) were used in the second dimension rather than DEAE paper. This adaptation gave excellent resolution of the different types of polyphosphorylated 5' termini as small compact spots (Fig. 1), thus allowing maximum sensitivity of detection. Two to three µl of the P1 nuclease digest were applied as a 0.7 cm streak about 6 cm from the end of a 31 cm long cellulose acetate strip (Schleicher and Schuell Co.) that was moistened with 7 M urea, 5% pyridinium acetate, pH 3.5, 1 mM EDTA.

The nucleotide streak was then flanked by two spots of electrophoresis dye mixture (11), and electrophoresis carried out at 500 v for 20 min in a small tank containing varsol, and then at 2000 v until the fast moving (pink) dye reached the anode buffer. The cellulose-acetate strip was then briefly blotted with filter paper to remove most of the varsol and the nucleotides transferred onto the DEAE cellulose plate about 2 cm from the origin as described by Volckaert et al. (2). The plate was then immersed in methanol for 10 min to remove the urea, sprayed with 5% pyridinium acetate pH 3.5, 1 mM EDTA, and electrophoresed at 1000 v in a well-cooled (5-10°C) varsol tank until the blue dye moved about 8 cm. We found that with some batches of plates cracking or folding of the DEAE cellulose layers can occur during electrophoresis. However, this did not significantly decrease resolution of the nucleotides.

After the plates had been thoroughly dried with a fan the nucleotides were located under UV and by autoradiography. The areas containing radioactivity were scraped off the plate, collected by suction in a disposable plastic pipette tip plugged with some glass wool and cellulose powder (2), and eluted with 300 µl of 30% TEAC, pH 7.6. This eluate was either diluted five times with H_2O and directly counted in 15 ml of ACS scintillation fluid (Amersham) or lyophilized for experiments involving subsequent digestion with nucleotide pyrophosphatase.

RESULTS

Quantitative analysis of 5'-termini

^{32}P-labeled hnRNA, size fractionated so as to exclude the low molecular weight species of nuclear RNA, was digested with ribonucleases T2, T1 and A, and the digestion products separated by chromatography on DEAE-Sephadex in the presence of 7 M urea. The material eluting at about -5 charge, which consists of a mixture of the pppXp, ppXp and m^7GpppX^mpYp derivatives, was treated with P1 nuclease, and the products (pppX, ppX, m^7GpppX^m and pY) analyzed by two-dimensional electrophoresis. An example of such an electrophoretogram is shown in Fig. 1. Radioactive components co-migrating with markers of GTP, ATP, GDP, ADP and four species of cap core (m^7GpppX^m, X^m = G^m, m^6A^m or A^m, C^m and U^m) are routinely observed. The cap cores containing doubly methylated and singly methylated A co-migrate in this system. With some preparations faint radioactive spots were seen near, but not coincident with, the UDP and CDP markers. These components, which migrate slightly faster than the markers in the first dimension and slightly slower in the second dimension, probably correspond to the radioactive components previously considered to be UDP and CDP on the basis of a one-dimensional chromatographic analysis (2). Owing to their scarcity, we

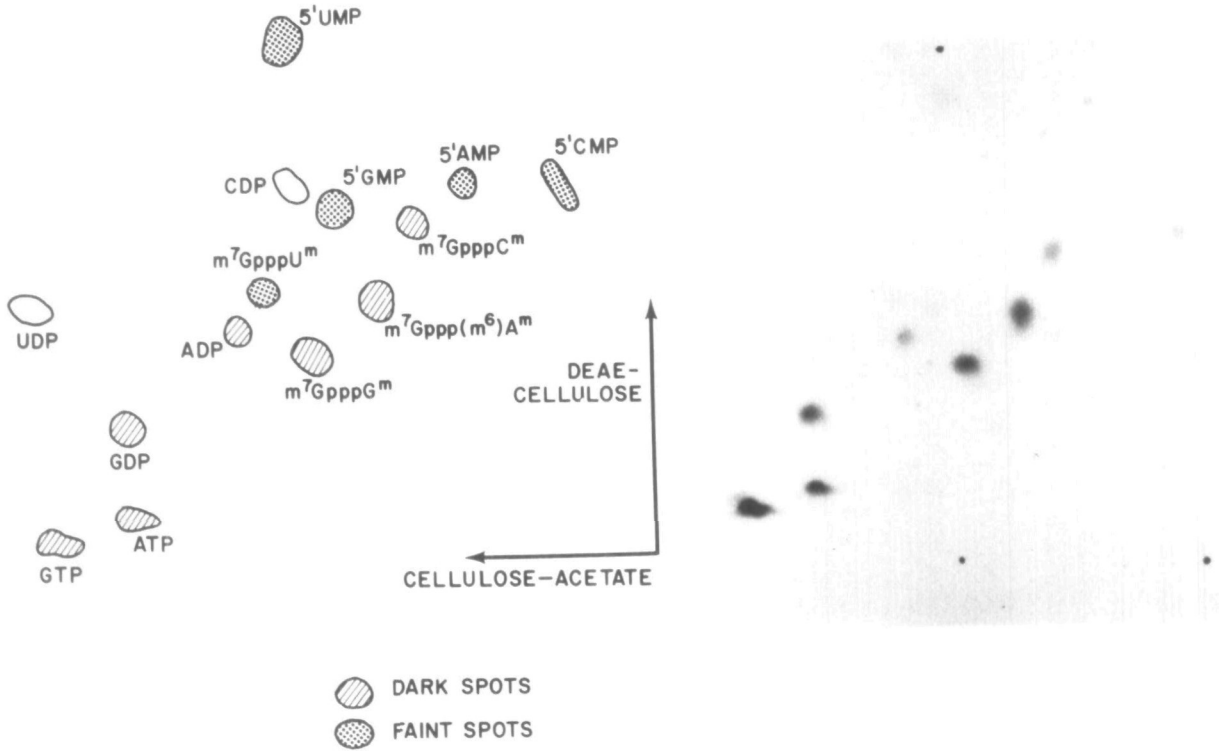

Fig. 1 Two-dimensional electrophoresis of 5' terminal derivatives. Left panel: positions of marker nucleotides as identified under UV; right panel: autoradiogram of a P1 digest of the -5 nucleotide derivatives from hnRNA labeled for 45 min with ^{32}P. Shaded and stippled markers correspond to the spots that are respectively dark and faint on the autoradiogram.

Nucleic Acids Research

have not been able to analyze these components further. However, given the superior resolution of the two-dimensional system, it seems reasonable to conclude that there is little, if any, hnRNA with an unmodified pyrimidine diphosphate at the 5'-termimus.

When the appropriate spots are scraped from the DEAE plate and counted the resultant data provide quantitative assays of the different 5'-terminal derivatives containing more than one phosphate. The proportion of radioactivity in each spot is a function of the mole fraction of the particular derivative and the specific activities of its labeled constituents. Since uniform labeling of the various precursor pools is not realized during the brief incubations used in the present experiments, an assessment of relative specific activities is required in order to convert radioactivity fractions to mole fractions. In the case of ^{32}P labeled hnRNA, this entails specific activity estimates for the α, β and γ phosphates of each of the four nucleotide species. For hnRNA labeled with ^{3}H nucleosides only the relative specific activity of each nucleotide species is required.

The relative specific activities of the α-^{32}P phosphates were determined by digesting parallel samples of the relevant hnRNA preparations with either RNAse T2 or P1 nuclease. The T2 digestion produces nucleotides containing the phosphates incorporated with the 3'-nearest neighbors. Since the nearest neighbors of a particular nucleotide species, N, are essentially random, the proportion of radioactivity, $R_N(T2)$, is equivalent to the molar proportion (M_N) of that species in hnRNA. The P1 digestion produces nucleotides containing the 5' (α) phosphates, so that the proportion of radioactivity, $R_N(P1)$, is equal to the product $(M_N)(S_N)$, where S_N represents the relative specific activity of the α-phosphate in nucleotide species N. Thus, for each nucleotide species, the ratio $R_N(P1)/R_N(T2)$ provides a measure of S_N. Similarly, a P1 digest of ^{3}H nucleoside-labeled hnRNA can be used together with the molar compositions determined above to calculate the relative specific activity of each tritiated nucleotide species.

The relative specific activities of the β phosphates were estimated by digesting either $\overset{\beta\alpha}{pp}X$ or $m^7G\overset{\alpha\beta\alpha}{ppp}X^m$ derivatives with nucleotide pyrophosphatase, and measuring the relative proportions of radioactivity in Pi and nucleotide monophosphates (Fig. 2). In all cases the preponderant fraction of radioactivity was released as Pi indicating a high ratio of β/α specific activity. For an estimate of the γ/β ratio, the $\overset{\gamma\beta\alpha}{pp}pA$ derivative was <u>partially</u> digested with nucleotide pyrophosphatase in the presence of ^{3}H-ATP and excess inorganic pyrophosphatase. Using the tritiated products to monitor molar yields of ATP, ADP and AMP, we estimated the γ/β (and β/α) specific activities from the corres-

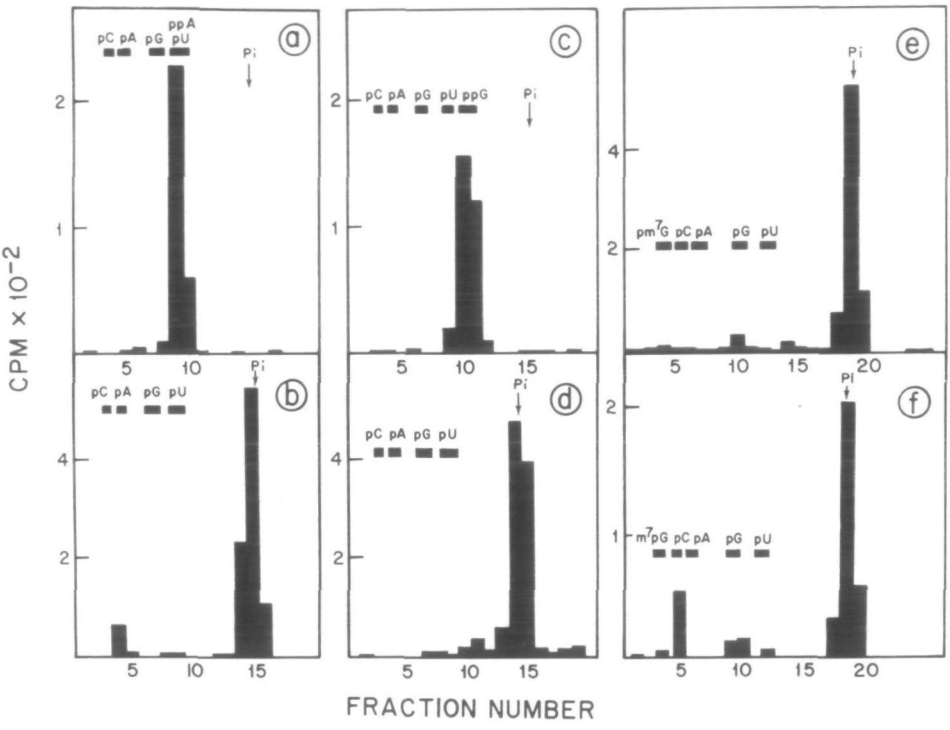

Fig. 2 Determination of the β/α specific activity in ppX and m^7GpppX^m derivatives. Nucleotide derivatives, recovered from DEAE-cellulose plates after two-dimensional electrophoresis, were analyzed by paper electrophoresis in the presence of appropriate UV markers with or without subsequent nucleotide pyrophosphatase (Npp-ase) digestion. The position of inorganic phosphate (Pi) was determined in a separate run with $^{32}PO_4$. (a) ppA, untreated; (b) ppA, pyrophosphatase-treated; (c) ppG, untreated; (d) ppG, pyrophosphatase-treated; (e) m^7GpppG^m, pyrophosphatase-treated; (f) m^7GpppC^m, pyrophosphatase-treated.

ponding ^{32}P products.

The results of these specific activity determinations are summarized in Table 1. For the ^{32}P-labeled hnRNA it is evident that almost all of the radioactivity in the terminal derivatives is contributed by the γ and β phosphates. However, since the relative specific activities of these phosphates are only roughly estimated by the methods employed here, calculations of molar ratios from the ^{32}P data are subject to considerable error. On the other hand, with hnRNA labeled with ^3H nucleosides, reasonably precise determinations of relative specific activities can be made, and since the magnitude of the S_N corrections is small, we would expect such data to provide a reliable estimate of molar ratios.

Table 1

The relative specific activities of α, β and γ phosphates and tritiated nucleosides in pulse-labeled hnRNA

Isotope-position	Duration of labeling (min)	Relative specific activity (S_N) G	A	C	U
^{32}P-α	30	0.12	0.82	1.5	1.4
β*		~10	~11	~8	n.d.
γ**		n.d.	~ 6	n.d.	n.d.
^{3}H nucleosides***	15	1.1	1.2	0.67	1.2
	30	1.6	1.7	0.63	0.9
	120	1.5	1.4	0.62	1.0

A molar base composition of 23% G, 23% A, 26% C and 28% U for 30 min labeled hnRNA, based on analyses of RNAse T2 digests of the 30 min ^{32}P-labeled hnRNA was used for these calculations.

* Estimated from β/α ratios of $\overset{\beta\alpha}{pp}X$ and $m^7Gp\overset{\beta\alpha}{ppp}X^m$. The β/α ratios for XDP and cap were essentially the same.

** Estimated from γ/β ratios of $\overset{\gamma\beta\alpha}{ppp}A$.

*** The relative specific activities refer to each particular labeling period. The increase in absolute specific activity with labeling time is not measured by this analysis.

n.d. = not determined

A comparison of 5'-termini in large and small hnRNA

For this experiment hnRNA from cells labeled for 30 min with $^{32}PO_4$ was extracted by the method of Hames (9), denatured by heating for 2 min at 70°C in 80% dimethylsulfoxide, and sedimented through a 15-30% sucrose gradient (Fig. 3a). Fractions estimated to be greater than and less than 35S were concentrated and resedimented through a second set of sucrose gradients (Fig. 3b,c). The >35S fraction was re-exposed to stringent denaturation conditions before sedimentation through the second sucrose gradient. The fractions corresponding to >35S and 16-35S were pooled, concentrated and digested with ribonucleases T2, T1 and A. Figure 4 illustrates the elution profiles on DEAE-Sephadex of the RNAse digests. There are very large peaks at -2 consisting of the internal nucleotides (Np), small peaks at -4 composed of 5'-terminal monophosphate derivatives (pXp) and peaks at -5 containing the tri- and diphosphate termini and the cap derivatives.

That our size fractionation procedure is highly effective in eliminating adventitious aggregation of hnRNA molecules (13) was verified by a measurement

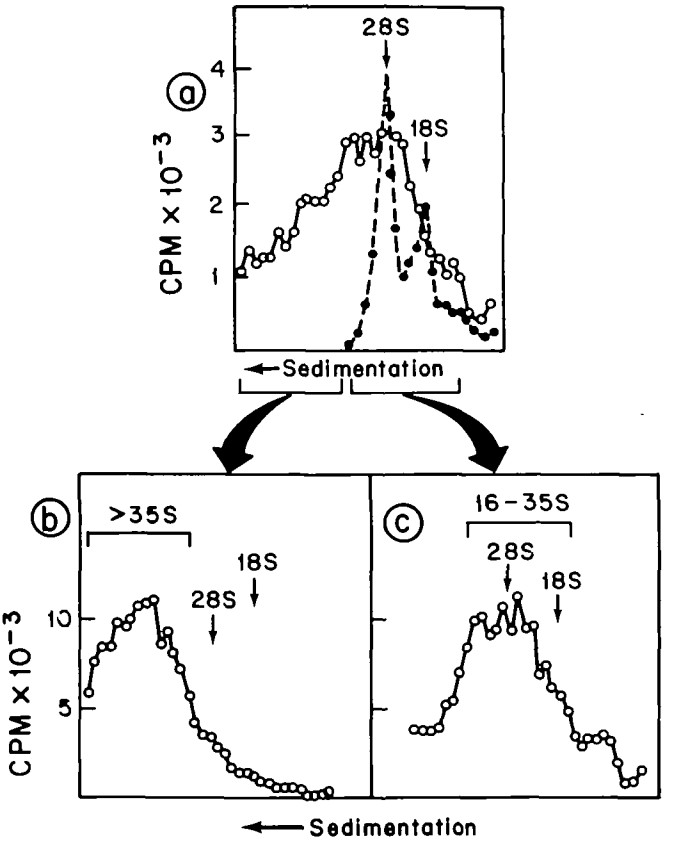

Fig. 3 Sucrose gradient fractionation of hnRNA. 2.2 liters of L cells at 3×10^5 cells/ml were pelleted, washed with 100 ml of growth medium lacking phosphate and suspended in 66 ml of phosphate lacking medium. After 45 min incubation at 37°C, actinomycin D was added to 80 ng/ml, and the cells incubated for an additional 15 min. Then, 25 mC of carrier-free $^{32}PO_4$ was added; the cells incubated for 30 min and harvested. hnRNA was extracted by the method of Hames (9), denatured in 80% dimethylsulfoxide and sedimented at 18,000 rpm for 17 hr at 22°C through a 32 ml 15-30% (w/w) sucrose gradient overlaid on a 2 ml cushion of 45% sucrose. Gradients and cushion were made with 10 mM Tris HCl buffer (pH 7.4) containing 0.1 M NaCl, 1 mM EDTA and 0.5% sodium dodecyl sulphate. Fractions sedimenting >32S (panel a) were pooled, concentrated by ethanol precipitation, denatured in 80% dimethylsulphoxide and submitted to a second sucrose gradient centrifugation under the conditions described above (panel b). Fractions corresponding to about 7-32S were pooled, concentrated and sedimented at 20,000 rpm for 15 hr at 22°C through a 32 ml 5-25% (w/w) sucrose gradient overlaid on a 2 ml cushion (panel c). The horizontal bars in panels b and c indicate the fractions used for subsequent analysis.

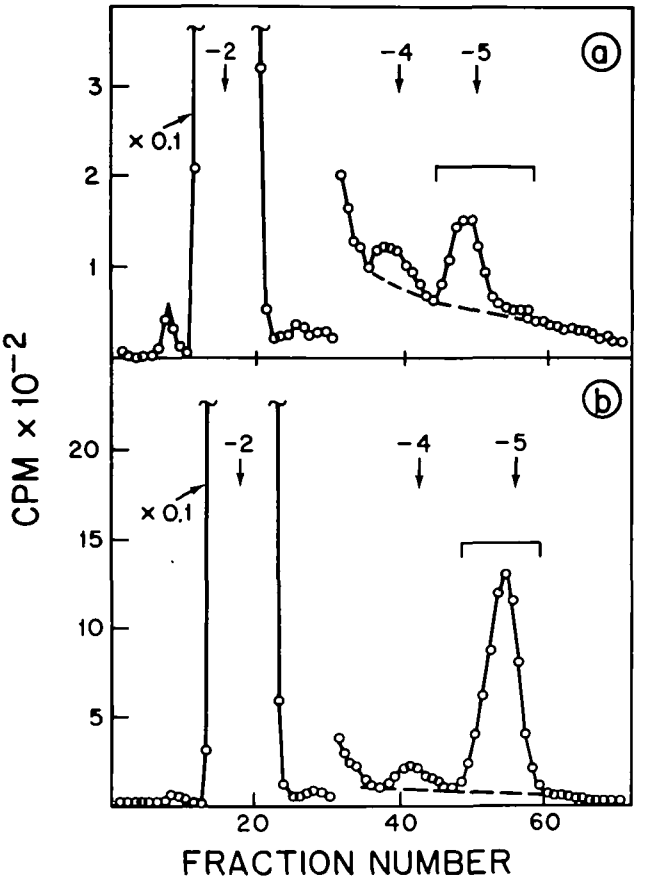

Fig. 4 Separation of nucleotide derivatives of hnRNA by DEAE-Sephadex chromatography. The hnRNA samples of Fig. 3 were digested with ribonucleases T2, T1 and A and chromatographed on DEAE-Sephadex columns. The material in the -5 peaks (horizontal bars) was used for subsequent analysis. The dashed lines under the -4 and -5 peaks indicate the base line corrections applied to the calculations of Table 2. a: >35S hnRNA; b: 16-35S hnRNA

of the chain lengths of the two hnRNA fractions. The average chain lengths, estimated from the relative proportions of radioactivity in the various peaks of Fig. 4, were approximately 9.7 kb for the >35S hnRNA fraction and 2.1 kb for the 16-35S hnRNA fraction (Table 2).

Examination of P1 digests of the material in the -5 peaks by two-dimensional electrophoresis indicated a marked difference in the distribution of termini in the two size fractions (Table 3). The proportion of radioactivity in triphosphate termini relative to cap termini is considerably lower in large (>35S)

Table 2

Relative amounts of radioactivity in internal nucleotides and 5' terminal derivatives for hnRNA of two different size classes

Size fraction of hnRNA	Total cpm in internal nucleotides (-2 peak) ($\times 10^{-6}$)	Total cpm in pppXp, ppXp and cap (-5 peak) ($\times 10^{-3}$)	Total cpm in pXp (-4 peak) ($\times 10^{-3}$)	Est. avg. chain length[a] (kb)
>35S	11.7	5.27	1.66	9.7
16-35S	15.9	66.5	6.30	2.1
Total (>16S)	27.6	71.8	7.96	5.4

Measurements from data of Fig. 4. The baseline contribution to the -5 peak was estimated as shown in figure and subtracted.

[a] Calculated as $(N_2/\alpha) \div [N_4/2_\alpha + N_5/(\alpha \cdot P_\alpha + \beta \cdot P_\beta + \gamma \cdot P_\gamma)]$ where N_2, N_4 and N_5 are the cpm in the -2, -4 and -5 peaks, respectively; α, β and γ are the relative specific activities of the corresponding phosphates (Table 1); and P_α, P_β and P_γ are the average number of respective phosphates in the derivatives comprising the -5 peak (calculated from the data of Table 3). Since the pX··· termini may be relatively pyrimidine rich (2), the appropriate specific activity value for the N_4 peak may be as much as 2.3_α rather than 2_α; if so, this would lead to a slight ($\leq$5%) underestimate of chain length.

Nucleic Acids Research

Table 3

A comparison of the 5'-termini of hnRNA molecules of different sizes

Size class of hnRNA	pppX (X:A+G)	ppX (X:A+G)	m^7GpppX^m (X^m:A^m+G^m)	m^7GpppX^m (X^m:C^m+U^m)	Tri + diphosphates / purine caps Radioactivity	Moles*
>35S	33	5	53	9	0.7	0.5
16-35S	67	7	21	5	3.6	2.3

The material from the -5 peaks of Fig. 3 was treated with P1 nuclease and analyzed by two-dimensional electrophoresis as illustrated in Fig. 1. The radioactivity in pppG, pppA, ppG, ppA and m^7GpppX^m [X^m = G^m, $(m^6)A^m$, C^m and U^m] was measured, and the combined radioactivity in each category of terminus expressed as a percentage of the total recovered radioactivity.

*Estimated from the data of Table 1 assuming the γ phosphate of GTP to be equal to that of ATP.

hnRNA as compared to small hnRNA. Although the molar ratios are calculated with approximate specific activity corrections (see above), we would not expect these corrections to differ significantly between the two hnRNA size classes. Thus, it is clear from these data that the larger hnRNA molecules have a relatively low probability of retaining the 5'-termini characteristic of primary transcription products.

The foregoing results may be indicative of some intrinsic difference between large and small transcription units. Alternatively, they might be related to a

Table 4

A comparison of the 5'-termini of hnRNA labeled for different periods with $^{32}PO_4$

Duration of labeling (min)	pppX (X:A+G)	ppX (X:A+G)	m^7GpppX^m (X^m:A^m+G^m)	m^7GpppX^m (X^m:C^m+U^m)	Tri + diphosphates / purine caps Radioactivity	Moles
30	64	13	18	5	4.3	3.0
45	49	19	27	5	2.6	-
90	49	11	33	7	2.0	-

L cells were concentrated to 10^7 cells/ml and incubated for 30 min in medium lacking phosphate and containing 80 ng/ml actinomycin D. One mCi/ml $^{32}PO_4$ was then added and the cells incubated for the times indicated. From each batch of cells the hnRNA >18S was isolated, digested with ribonucleases T2, T1 and A, and the digests chromatographed on DEAE-Sephadex columns. The material in the -5 peaks was analyzed and expressed as percentage of total radioactivity as described in Table 3.

difference in the relative metabolic stabilities of triphosphate and capped termini. If the elimination of triphosphate termini and the formation of relatively stable cap termini could occur during the growth of RNA chains, then the probability of there having been an alteration at the 5' terminus would increase with the length of the growing transcript. Such an effect would be detectable if growing chains and recently completed transcripts comprise a reasonably high proportion of the labeled hnRNA molecules. This is likely to be the case in this experiment since the specific activity of the nucleotide precursor pools is continually increasing during the 30 min incubation with $^{32}PO_4$. To investigate the relative stabilities of phosphorylated and capped ends, we examined the relative proportions of different 5' termini in hnRNA labeled for different periods of time.

Proportions of different 5' termini as a function of the duration of labeling

An analysis similar to that described above was made with total (>18S) hnRNA labeled for varying periods of time with ^{32}P. With increasing duration of labeling we observed an increase in the proportion of radioactivity in capped termini relative to that in tri- and diphosphate termini (Table 4), suggesting a significant alteration in the composition of 5'-termini as a function of labeling time.

To confirm this observation under conditions which permit a reliable estimate of molar ratios, we repeated this experiment using hnRNA from cells labeled with

Table 5

The molar percentage of 5'-termini of hnRNA labeled for different periods with ^3H nucleosides

Duration of labeling (min)	pppX (X:A+G)	ppX (X:A+G)	m^7GpppX^m X^m:$(m^6)A^m$+G^m	m^7GpppX^m (X^m:C^m+U^m)	Tri + diphosphates / purine caps
15	40	22	32	6	2.0
30	36	12	44	9	1.1
120	21	8	60	11	0.5

L cells were incubated for 30 min with 80 ng actinomycin D/ml and then concentrated to 3 - 5 x 10^6 cells/ml and incubated with 20 - 25 μc/ml each of 8-^3H guanosine, 2,8-^3H adenosine, 5-^3H cytidine and 5-^3H uridine for the times indicated. From each batch of cells hnRNA >18S was isolated and hydrolyzed with RNAses T2, T1 and A. The -5 derivatives were analyzed as described in Table 3. The percentage ^3H radioactivity in each spot was divided by the relative specific activity value appropriate for the particular nucleotide species (Table 1) to give the molar percentage. In the case of the cap cores the value corresponding to the particular X^m species was used: the tritium on the 8 position of guanine is lost when it is converted to m^7G (14).

Table 6

Proportions of the various types of 5'-termini of hnRNA as a function of labeling time

Duration of labeling (min)	Molar percentage of total 5' termini		
	Polyphosphate (pppXp and ppXp)	Monophosphate (pXp)	Cap (m^7GpppX^mpYp)
15	36	42	22
30	31	35	34
120	20	31	49

Data from the hnRNA samples analyzed in Table 5. The molar percentage of monophosphate termini relative to the other types of termini was calculated from the relative proportions of radioactivity in the -4 and -5 peaks of DEAE-Sephadex profiles, accounting for the two nucleotides contributed by cap structures to the -5 peak and for the specific activity differences among termini of different composition as given in Tables 1 and 5.

all four species of ^3H nucleosides. The results (Table 5) indicate that there is indeed a relative diminution in tri- and diphosphorylated ends relative to caps as the incubation interval is increased. Between 15 min and 2 hr the proportion of triphosphate ends drops from about 40 to 21% while the proportion of capped termini increases from 38 to 71%. In terms of purine-terminated molecules the ratio of tri- and diphosphorylated to capped ends decreases by a factor of four. These results are most readily interpreted in terms of a greater metabolic stability of capped ends as compared to di- and triphosphate termini.

The data from this experiment were also used to calculate the relative proportions of pX··· termini and to estimate the average chain length of the hnRNA molecules. The proportion of pX··· termini, like the proportion of polyphosphate termini, diminishes with increasing labeling time (Table 6). The average chain length, estimated from the ratios of moles of internal nucleotides to moles of total 5' termini, is 3 to 5 kb, which is roughly comparable to the calculated value for total ^{32}P-labeled hnRNA (Table 2). The fact that these average chain length estimates are not very different from the number average molecular weight calculated from the sedimentation behavior of hnRNA (1) indicates that we have accounted for most, if not all, of the 5'-termini in these studies. If an appreciable fraction of hnRNA had ends that were not detected by our analysis, e.g., 5'-OH, then our chain length estimates would have yielded abnormally high values; this is clearly not the case.

DISCUSSION

Our analysis of the composition of 5'termini of hnRNA by two-dimensional electrophoresis revealed the purine tri-and diphosphates: pppG, pppA, ppG and ppA; as well as four cap structures: m^7GpppX^m, X^m = G^m, m^6A^m or A^m, C^m and U^m. We also observed monophosphorylated termini, pX···, which, according to previous studies (2), contain pU··· as the predominant species. Contrary to our earlier interpretations of one-dimensional chromatographic analyses, there appear to be little or no hnRNA molecules bearing unmodified pyrimidine diphosphate termini. The presumed existence of such termini led us to propose a model in which all of the termini that are destined to be capped, including those formed by cleavage at internal sites of hnRNA, pass through a ppX··· intermediate. Although the present results eliminate an important piece of supportive evidence for this model, they do not necessarily negate it. Indeed, the only cellular (non-viral) cap-forming enzyme described to date appears to require a ppX··· terminus (15).

The studies of hnRNA labeled for increasing periods of time indicate a trend toward a higher proportion of capped ends and a lower proportion of tri-and diphosphorylated ends. This implies that caps are metabolically more stable than the pppX··· and ppX··· termini. This result is consistent with our previous proposal (2) that the 5' ends of some primary transcripts are directly converted to caps. On the other hand, we cannot exclude the possibility that capping occurs exclusively at internal cleavage sites, and that the putative 5'-leader sequences are either modified or absent in our hnRNA preparations. However, the recent discovery of an enzyme activity in mammalian cells which forms caps by condensation of GTP with diphosphate terminated RNA (16,17) lends considerable credence to the idea of direct capping of initial ends of primary transcripts.

In cells incubated for two hours with tritiated nucleosides approximately half of the labeled hnRNA molecules have capped ends. This rather high proportion of capped molecules agrees well with an estimation based on the amounts of internal m^6A per cap and per nucleotide. Previously we observed that there are about 3.6 moles of m^6A per mole of cap in the >16S hnRNA of cells labeled for either 15 or 60 min with ^3H methionine (5,6). A second type of determination, based on a two-dimensional chromatographic separation of m^6Ap and the four unmodified nucleotides in a ribonuclease digest of 30 min ^{32}P-labeled hnRNA, indicated that there are about four residues of m^6A per 10^4 nucleotides in >16S hnRNA (unpublished observation). Combining these values [$(4m^6A/10^4$ nucleotides) ÷ (3.6 m^6A/cap) = 1.1 cap/10^4 nucleotides], and taking 4 to 5 kb as the average chain length of hnRNA, we obtain 50 ± 6% for the proportion of capped hnRNA molecules. Since most or all hnRNA caps appear to constitute precursors of mRNA caps (5,6), this

result would seem to mean that there is a high proportion of pre-mRNA molecules in the hnRNA that is labeled in a 2 hr period. This conclusion is consistent with the substantial sequence homology observed between the mRNA and pulse-labeled hnRNA of mouse L cells (9). What about hnRNA molecules that lack mRNA sequences? On the basis of sequence complexity studies such molecules are generally considered to comprise a reasonably large fraction of the steady-state hnRNA population [cf. Davidson et al. (17)]. This notion could be reconciled with our findings if non mRNA-related molecules had a relatively high stability, in which case their contribution to pulse-labeled hnRNA fractions would predictably be small.

In hnRNA from cells labeled for 30 min with $^{32}PO_4$ the proportion of triphosphorylated termini relative to cap in large molecules is significantly lower than in small molecules. In view of the lower metabolic stability of the pppX termini, this observation could imply that some hnRNA molecules are subject to processing reactions before the completion of chain growth. If such were the case, the probability of having undergone a modification of the 5' terminus would be greater for large hnRNA than for small hnRNA, and one would expect the proportion of triphosphate termini to be less in the large hnRNA class.

In contrast to our finding, when poly $(A)^+$hnRNA from HeLa cells labeled for 3 hr with $^{32}PO_4$ was analyzed, the large molecules were observed to have the same proportion of capped 5'-termini as the small molecules (18). This discrepancy is probably attributable to the fact that the 30 min labeled hnRNA analyzed in our experiments tends to emphasize growing and recently completed transcripts, whereas the 3 hr labeled poly $(A)^+$hnRNA consists exclusively of completed chains that have subsequently been polyadenylated at their 3' termini. In fact, the result obtained with the poly $(A)^+$hnRNA tends to argue against the alternative explanation for our observation, namely, that the extent of cap formation and triphosphate degradation is related to the intrinsic size of the transcription units. If this were so, one would have expected to see similar differences in the relative extents of capping of large and small hnRNAs at all stages of processing. The possibility of cap formation on growing RNA chains was indicated by the finding that the guanylyltransferase of Reovirus can form a cap with the dinucleotide ppGpC as a substrate (19). The apparent coupling of transcription and cap formation in mammalian subcellular systems (16,20) is also consistent with this idea.

ACKNOWLEDGMENTS

This work was supported by a grant from the National Science Foundation (PCM77-06539), grants from the National Institutes of Health (CA-06927 and

RR-05539), and a fellowship from the Swiss Science Foundation.

REFERENCES

1 Perry, R. P., Bard, E., Hames, B. D., Kelley, D. E. and Schibler, U. (1976) in Progress in Nucleic Acid Research and Molecular Biology, Cohn, W. E. and Volkin, E., Eds., Vol. XIX, pp. 275-292, Academic Press, New York
2 Schibler, U. and Perry, R. P. (1976) Cell 9, 121-130
3 Schmincke, C. D., Herrmann, K. and Hausen, P. (1976) Proc. Natl. Acad. Sci. USA 73, 1994-1998
4 Bajszar, G., Samarina, O. P. and Georgiev, G. P. (1976) Cell 9, 323-332
5 Perry, R. P. and Kelley, D. E. (1976) Cell 8, 433-442
6 Schibler, U., Kelley, D. E. and Perry, R. P. (1977) J. Mol. Biol. 115,
7 Perry, R. P., Kelley, D. E. and LaTorre, J. (1974) J. Mol. Biol. 82, 315-331
8 Scherrer, K. (1969) in Fundamental Techniques in Virology, Habel, K. and Salzman, N. P., Eds., pp. 413-432, Academic Press, New York
9 Hames, B. D. and Perry, R. P. (1977) J. Mol. Biol. 109, 437-453
10 Cory, S. and Adams, J. M. (1975) J. Mol. Biol. 99, 519-547
11 Sanger, F. and Brownlee, G. G. (1967) in Methods in Enzymology, Grossman, L. and Moldave, K., Eds.,Vol. XII, pp. 361-381, Academic Press, New York
12 Volckaert, G., MinJou, W. and Fiers, W. (1976) Anal. Biochem. 72, 433-446
13 Federoff, N., Wellauer, P. and Wall, R. (1977) Cell 10, 597-610
14 Shatkin, A. J. (1976) Cell 9, 645-653
15 Wei, C.-M. and Moss, B. (1977) Proc. Natl. Acad. Sci. USA 74, 3758-3761
16 Groner, Y., Gilboa, E. and Aviv, H., manuscript submitted for publication
17 Davidson, E. H., Klein, W. H. and Britten, R. J. (1977) Develop. Biol. 55, 68-84
18 Salditt-Georgieff, M., Jelinek, W., Darnell, J. E., Furuichi, Y., Morgan, M. and Shatkin, A. (1976) Cell 7, 227-237
19 Furuichi, Y. and Shatkin, A. J. (1976) Proc. Natl. Acad. Sci. USA 73, 3448-3452
20 Winicov, I. and Perry, R. P. (1976) Biochemistry 15, 5039-5046