Subcellular localization of RNAs in transfected cells: role of sequences at the 5' terminus

Aixa Alfonso-Pizarro, David P. Carlson and Jeffrey Ross

McArdle Laboratory for Cancer Research, 450 North Randall Avenue, Madison, WI 53706, USA

Received 10 August 1984; Revised and Accepted 22 October 1984

ABSTRACT

We have investigated the intracellular location of RNAs transcribed from transfected DNA. COS cells transfected with a clone containing the human adult β globin gene contain three classes of globin RNAs. Their 3' termini and splice sites are indistinguishable from those of mature reticulocyte β globin mRNA, and they are polyadenylated. However, as determined by S₁ mapping, their 5' sequences are different. The 5' terminus of one is the same as that of mature β globin mRNA (+1, cap site). The presumed 5' terminus of the second is located 30 nucleotides downstream from the cap site (+30). The third class contains additional nucleotides transcribed from sequences located 5' to the cap site (5' upstream RNA). The 5' upstream RNA molecules are restricted to the nucleus and are more stable than heterogeneous nuclear RNA. The +30 and +1 RNAs are located primarily in the cytoplasm. The data support the notion that nucleotide sequences and/or secondary modifications in the 5' region determine if an RNA is to be transported.

INTRODUCTION

The transfection of cloned genetic material into cultured cells has provided a powerful tool for investigating transcription (for review and references see 1-3), control of transcription (4-10), RNA processing (11-14), and the function of genes that alter cellular phenotypes (15-16). In many cases, pseudogenes and genes with naturally occurring mutations behave in a predictable fashion, at least qualitatively, when transfected into cells (12-14, 17-23). In other cases, the steady-state level of total RNA produced by a mutant gene may be inappropriately high. For example, HeLa cells transfected with either a normal or a β thalassemic globin gene may contain equivalent amounts of globin RNA, even though erythroid cells from the thalassemic patient contain little β globin mRNA (24). The relative levels of normal vs. mutant RNA may also vary with the recipient cell line. HeLa cells transfected with a normal β -globin gene or a β^0 -thalassemia allele containing a nonsense triplet at codon 60 (25) accumulate equivalent amounts of globin RNAs (unpublished observations). COS cells transfected

with the same genes contain five-fold more normal than thalassemic RNA. Even in COS cells, however, the thalassemic RNA is at least ten-fold more stable than it is in thalassemic erythroblasts (26; unpublished observations).

In an effort to resolve this discrepancy, we have investigated the structure and intracellular location of globin RNAs produced from a transfected normal human β globin gene in COS cells. Three classes of globin RNAs are observed. One is indistinguishable from β globin mRNA; as determined by S_1 mapping, its 5' terminus is at the mRNA cap site, designated +1, and its 3' terminus is at the normal polyadenylation site. A second class generates an S_1 mapping pattern consistent with either a 5' terminus or a 3' splice site (27) located approximately 30 nt downstream from the mRNA cap site (+30 RNA). The 5' region of the most abundant class of RNA is heterogeneous and includes sequences upstream from the cap site (upstream RNA). As determined by S_1 mapping, the splice sites and 3' terminus of most of the upstream RNA molecules and of the +1 and +30 RNAs are the same as those of authentic globin mRNA, and the RNAs are polyadenylated. Therefore, most of these RNAs differ only in their 5' regions. Similar observations have already been made by others using a variety of genes in long- and short-term transfection assays (9,27-29). Our novel observation is that the upstream RNAs are restricted to the nucleus and are relatively stable, as compared with total heterogeneous nuclear RNA. In contrast, the +1 and +30 RNAs are detected in the cytoplasm, and the +1 mRNA is located primarily in polysomes. We conclude that the level of total RNA produced by a transfected gene does not necessarily reflect mRNA production by that gene, because, at least in COS cells, most of the RNA molecules contain upstream sequences and are not functional messengers. More importantly, these results imply that primary sequence or secondary modifications in the 5' region determine if an RNA molecule will be transported to the cytoplasm.

MATERIALS AND METHODS

Cell culture and transfection:

The cloned COS-M6 cell line obtained from Dr. J. Mertz was maintained at 37°C in 5% CO₂ in Dulbecco's modified medium (DMEM) supplemented with high glucose and 10% FCS (growth medium). The transfection method was essentially as described (71). In brief, 1-3 x 10⁷ cells in suspension were exposed to 2 ml of a calcium phosphate-DNA coprecipitate (DNA at 25 µg/ml)

for 15 minutes. The suspension was then diluted to 50 ml with growth medium containing 6.25 mM CaCl_2 and 0.05X HeBS (1X HeBS = 21 mM Hepes, 0.137 M NaCl, 0.75 mM Na_2HPO_4 , 4.5 mM KCl, 5.5 mM dextrose, pH 7.12). Ten ml of the suspended cells were plated in tissue culture dishes (100 x 15 mm) and incubated for 3.5 hours at 37°C, 5% CO_2 . The cells were then exposed to 25% glycerol in DMEM for 1 minute, after which they were washed with serum-free DMEM, overlaid with fresh growth medium, and returned to 37°C.

Purification of nuclei, polysomes, and RNA:

Cells were collected at 47-48 hours post-transfection and rinsed with cold TBS (10 mM Tris, pH 7.5, 0.15 M NaCl, 1.5 mM MgCl_2). Total cell RNA was purified as previously described (45). Polysomal RNA was prepared at 4°C as follows: $1-2 \times 10^7$ cells were washed once in TBS and were resuspended in 8 volumes of RSB (10 mM NaCl, 10 mM Tris, pH 7.4, 1.5 mM MgCl_2). Cells were then disrupted with ten strokes of a tight-fitting teflon homogenizer. Nuclei were separated from cytoplasm by centrifugation for 10 min at 900 rpm in a PRJ centrifuge (IEC). Nuclear RNA was prepared from the pellet by the same method as for total cell RNA. The low-speed supernatant (cytoplasm) was adjusted to 0.1 M NaCl and centrifuged at 11,000 rpm for 20 minutes in the Sorvall HB4 rotor. The high-speed supernatant was divided into two portions, one of which was made 20 mM in EDTA. The samples were then layered on top of 1.5 ml cushions of 1 M sucrose with or without 20 mM EDTA in a buffer containing 0.1 M NaCl, 10 mM Tris, pH 7.5, 1.5 mM MgCl_2 , 2 mM DTT. Each fraction was centrifuged for 2.5 hours at 36,000 rpm in an SW60 rotor at 0°C. Polyribosomal (untreated) and RNP (EDTA-treated) pellets were resuspended in 1X NETS buffer (0.1 M NaCl, 5 mM EDTA, 10 mM Tris, pH 7.4, 0.5% sodium dodecyl sulfate) and phenol extracted, and the RNA was precipitated with ethanol. The ultracentrifuge supernatants were adjusted to 0.5% SDS and then extracted with phenol and precipitated with ethanol. All RNAs were resuspended in 0.05 M NaCl, 0.1 mM EDTA, 1 mM Tris, pH 7.4. Total cell RNA was fractionated on oligo-dT cellulose columns as described (72).

S₁ analyses of RNAs:

5' End-labeled DNA probes were prepared from cloned human genomic β globin DNA (30) digested with appropriate restriction enzymes. Terminal phosphates were removed with bacterial alkaline phosphatase (BRL). The fragments were then 5' end-labeled with γ - ^{32}P -ATP (ICN) and T_4 polynucleotide kinase (Promega-Biotec). The 3' end-labeled fragments were prepared with M. luteus DNA polymerase and the appropriate $\alpha^{32}\text{P}$ -deoxynucleotide

(Amersham, 3000 Ci/mole). All but the 5' end-labeled Hae III probe (Figure 2) were made strand specific by cleavage with a second enzyme and purification in an agarose gel.

Heat-denatured probes were hybridized to RNAs in buffer containing 80% formamide for 12-23 hours at 49°C and further processed as described (33,73-74).

RESULTS

Heterogeneous globin RNAs in transfected COS cells

The clone used in these experiments contains 3.6 kb of human DNA, including the complete human β globin structural gene (30), plus 1480 nt of 5' flanking and 548 nt of 3' flanking DNA (Figure 1; 31). A portion of the SV40 genome, including the replication origin and the 72 nt repeat enhancer sequence, is placed 3' to the human genomic insert.

To identify globin RNAs from COS cells transfected with this clone, total cell RNA was hybridized to a 5'-³²P-labeled Hae III DNA fragment (Figure 2). The reaction mix was treated with S1 nuclease; S1 resistant DNA fragments were electrophoresed in an 8.3 M urea-8% polyacrylamide gel and were identified by autoradiography. Three globin RNA-specific bands were observed (Figure 2, lane 1). One corresponds to an RNA whose 5' terminus is indistinguishable from that of normal β globin mRNA (+1, cap site RNA; Fig. 2, lane 2). A second corresponds to an RNA with a 5' terminus or 3' splice site approximately 30 nucleotides 3' to the mRNA cap site (+30 downstream RNA). The third band corresponds to multiple RNAs with sequences located upstream from the mRNA cap site (upstream RNA; see Fig. 3). The larger bands (Fig. 2, lane 1) probably result from hybridization of pBR322 RNA to a contaminating 587 nt pBR322 ³²P-DNA fragment; these bands cannot contain globin sequences, because they are larger than the human sequence in the 629 nt fragment.

These data confirm those of many investigators (Discussion) and demonstrate that a transfected gene may produce different RNAs, some with normal and others with so-called "aberrant" 5' termini. Assuming that each class of RNA hybridizes with the same efficiency, scanning densitometry of this and of similar gels from three additional experiments reveals that the upstream, +1, and +30 RNAs account for 57%, 25%, and 18% of the total cell globin RNA, respectively. In HeLa cells transfected with the same DNA, the 5' terminus of approximately 90% of the globin RNA molecules is at +1 (data not shown). Recent experiments indicate that the production of upstream RNA

Figure 1. Structure of plasmid DCQ21 DNA used for transfections. The ampicillin-resistant plasmid contains a 3.6 kb fragment of human DNA including the complete human β -globin gene (large open boxes represent exons I-III, small open boxes represent intervening sequences 1 and 2) plus 1480 bp of 5' flanking and 548 bp of 3' flanking DNA (thick, dark lines). The 795 bp Bam HI fragment (sequences -1460 to -2255; 31) has been deleted from the human 5' flanking region. The remainder of the plasmid contains in clockwise order: a 6 bp Cla I-Hind III linker (pBR322 sequences 23 to 29; 75) flanking each side of a 419 bp SV40 fragment (SV40 sequences 5089 to 265; 76) that contains the 72 bp repeats and origin of replication (hatched box); a 774 bp fragment of the SV40 late region (SV40 sequences 964 to 190; 76) (hatched box); a pBR322 fragment (sequences 2065 to 23; 75) containing the bacterial origin of replication and the β -lactamase gene (thin line); a 59 bp Msp I-Pst I linker (pBR322 sequences 3549 to 3608; 75). The SV40 insert contains a duplicated region that has no known enhancer, splice site, or control sequences. It is present in this clone only because it was easier to make the construct with the duplication than without it. The early (E) and late (L) strands of SV40 DNA are noted, and +1 and poly A designate the globin cap and termination sites, respectively.

is enhanced if the recipient cells contain SV40 T antigen (as do COS cells) and if the transfected DNA includes an SV40 origin of replication (27).

The structure of the 5' region of the upstream RNAs was analyzed in greater detail by S_1 mapping with a DNA probe ^{32}P -5'-end-labeled at the Bal I site, located at -76, or 76 nt upstream from the mRNA cap site. Many protected fragments are observed (Figure 3, lane 3). Additional bands appear in lanes 3 and 4 with longer exposures (data not shown). The three most prominent bands correspond to RNAs with termini approximately 515, 825, and 1460 nt upstream from the mRNA cap site. Since the human portion of the transfected DNA extends only to -1460, the RNA that generates the -1460 band probably extends into pBR322 sequences. The major upstream RNA (-172) observed in human bone marrow (lane 1) and reticulocytes (32-33) is not detectable in transfected cells (lane 3). These data indicate that there are multiple upstream RNAs containing 200 to 1400 or more additional nucleotides 5' to the normal +1 site. We have not determined if these bands correspond to 5' termini or 3' splice sites in the RNAs, nor do we know

Figure 2. S_1 analysis of globin RNAs with the 5'- ^{32}P HaeIII probe. 2.5×10^7 Cells were transfected with pDCQ21 and harvested 48 hours thereafter. RNAs were hybridized with a 5'- ^{32}P HaeIII probe (6.2×10^4 cpm per reaction) for 23 hours. The human genomic portion of the probe extends from the HaeIII site (+132) in exon I to the BalI site at position -76 (+1 = cap site; see diagram). The full-length probe contains 629 nucleotides, the remaining sequence being derived from pBR322. The gel was exposed for 18 hours with a Lightning Plus screen. Lane 1: 10 μg total transfected cell RNA. Lane 2: 500 pg human adult reticulocyte RNA. Lane 3: 6 μg *E. coli* t-RNA. Marker (M): HaeII digest of pBR322. Sizes are noted in nucleotides on the right. The diagram outlines the probe and hybridization results. The jagged line denotes pBR322 DNA in the probe. Numbers in parentheses denote the nucleotides in the probe and protected fragments. +1 Refers to the cap site and -76 to the Bal I site upstream of the β -globin gene. The open box corresponds to exon sequences.

where transcription of these RNAs initiates (see Discussion).

Upstream RNAs in the nucleus

Since the three classes of RNA differ so at their 5' termini, it seemed reasonable to investigate the possibility that they are metabolized differently by the cell. For example, normal mRNA molecules might enter polysomes more readily than upstream or downstream RNAs. Transfected cells were harvested and disrupted in the absence of detergents, and nuclear and cytoplasmic extracts were prepared; the cytoplasm was then centrifuged to separate polysome and supernatant fractions. Some of the cytoplasm was centrifuged in buffer containing 20 mM EDTA, to disaggregate polysomes and thereby to demonstrate that ribosomal subunits and mRNP do not pellet through the sucrose cushion. A 5' labeled Hae III probe was used to assay the RNA in each fraction. Seventy percent of the +1 RNA molecules is detected in the cytoplasm (Fig. 4., lanes 2, 5, and 6), and 60% of these is located on polysomes that are disrupted by EDTA (Figure 4, lanes 3 and 5).

Figure 3. S_1 analysis of upstream RNAs. 2×10^7 Cells were transfected with pDCQ21 and total and fractionated RNA isolated 47 hours thereafter. Total and nuclear RNA corresponding to an equal number of cells was hybridized with a strand-specific, gel purified 2200 bp 5'- 32 P-BalI probe (2.8×10^3 cpm per reaction) for 12 hours. The probe contains approximately 1400 nt of human sequence from the 5' flanking beta globin gene region, the remaining 800 nt being derived from pBR322. The probe was labeled at the -76 position and was made strand-specific by secondary digestion with Pst I. S_1 -resistant fragments were separated in an 8% acrylamide-8.3M urea sequencing gel. The gel was exposed for 4 days with a Quanta III intensifying screen. Lane 1: 50 μ g adult human bone marrow RNA. Lane 2: 50 μ g *E. coli* tRNA. Lane 3: 50 μ g total RNA from transfected cells. Lane 4: 14 μ g nuclear RNA from transfected cells. Lane 5: Input probe. Lane 6: HaeII digest of pBR322. Sizes in nucleotides are noted on the right. The diagram outlines the probe and hybridization results (see Fig. 2 legend). The major bands protected by the probe with bone marrow RNA (-172) and transfected cell RNA (-515, -825 and -1460) are indicated with capital letters to the left. Approximate upstream positions relative to the cap site are indicated by the numbers in parentheses on the right of the diagram.

Figure 4. S_1 analysis of globin RNAs from transfected cell cytoplasm and nuclei. 2×10^7 Cells were transfected with pDCQ21 and harvested 47 hours thereafter. The cells were fractionated into nuclei, cytoplasm, and polysomes as described in Materials and Methods. RNA was hybridized with a $5'$ - ^{32}P HaeIII probe (4.8×10^4 cpm/reaction) for 16 hours. See Figure 2 for description of the probe. RNA for each reaction corresponded to that from 2×10^5 cells. The gel was exposed for 3 days with a Lightning Plus intensifying screen. Lane 1: 5 μg total RNA. Lane 2: 1.4 μg nuclear RNA. Lane 3: 120 ng RNA from the pellet of EDTA-treated cytoplasm. Lane 4: 1.2 μg supernatant RNA from EDTA-treated cytoplasm. Lane 5: 500 ng RNA from the polysomal pellet (no EDTA). Lane 6: 970 ng supernatant RNA from the untreated cytoplasm (no EDTA). Lane 7: 450 pg human adult reticulocyte RNA. Lane 8: Probe. The sizes of marker DNA fragments (Hae II digest of pBR322) are indicated on the right.

In contrast, greater than 95% of the upstream RNAs is restricted to the nucleus (Fig. 4, lane 2). These results imply that globin RNAs containing a sufficient number of nucleotides upstream from the +1 cap site are transported to the cytoplasm inefficiently, if at all.

Splice sites and 3' terminus of globin RNAs in transfected COS cells

In order to test this conclusion, it was necessary to analyze in more detail the structural features of the three RNA classes. Are they 3' terminated, polyadenylated, and spliced in the same manner as authentic globin mRNA? If so, then the additional nucleotides must play a role in restricting transport. Six ^{32}P -labeled DNA probes were used in S_1 mapping assays to compare the structures of transfected cell RNAs with authentic globin mRNA. A $3'$ - ^{32}P -labeled Nco I probe identifies the 5' splice site of the first intervening sequence. S_1 analysis with reticulocyte RNA produces

Figure 5. S_1 analysis of splicing sites in IVS-1. 1×10^7 Cells were transfected with pDCQ21 and harvested 48 hours thereafter. Nuclear and cytoplasmic fractions were prepared. RNAs were hybridized for 19 hours (1×10^4 cpm per reaction). S_1 resistant fragments were separated in an 8% acrylamide gel, which was exposed for 24 hours with a Quanta III intensifying screen. Lanes 1: 5 μ g nuclear RNA. Lanes 2: 5 μ g cytoplasmic RNA. Lanes 3: 1.8 μ g total poly A⁺ RNA. Lanes 4: 800 pg adult human reticulocyte RNA. Lanes 5: 6 μ g *E. coli* tRNA. Panel A. Hybridization with a 1345 bp 3'-³²P NcoI probe, which had been made strand-specific by secondary digestion with EcoRI. Panel B. Hybridization with a 2099 bp 5'-³²P BamHI probe, the human genomic portion of which extends from the BamHI site in exon II (+481) to the BalI site (-76). The probe was made strand-specific by secondary digestion with Pst I. The diagram outlines the probes and hybridization results. The jagged line denotes pBR322 sequences.

fragments of 91 ± 2 nt, consistent with correct splicing of the first and second mRNA coding blocks (Fig. 5, panel A, lane 4). The same fragments were generated with nuclear, cytoplasmic, and total poly A⁺ RNAs from transfected cells (Fig. 5, panel A, lanes 1-3). The 440 nt band observed

Figure 6. S_1 analysis of upstream nuclear RNA that retains IVS 1. The same RNAs analyzed in Fig. 5 were used in this experiment. They were hybridized with a 3'- ^{32}P -Bal I probe (10^4 cpm per reaction) for 19 hr. S_1 -resistant fragments were separated in a 3% acrylamide-8.3M urea gel. The film was exposed for 1.5 days with a Lightning Plus intensifying screen. Lane 1: 5 μg nuclear RNA; Lane 2: 5 μg cytoplasmic RNA; Lane 3: 3 μg total poly A⁺ RNA; Lane 4: 120 pg adult reticulocyte RNA; Lane 5: 6 μg *E. coli* tRNA; Lane 6: input probe. M: Hae II digest of pBR322. Sizes in nucleotides are noted on the right. The diagram outlines the probe and hybridization results. The probe comes from pDCQ21 digested with Bal I, 3'-end labelled, and then digested with Hind III. It contains all the β -globin coding sequences and the 3' flanking region.

with nuclear and total poly A⁺ RNAs corresponds to RNA that retains IVS 1 but is spliced at the 5' site of IVS 2 (Fig. 5, panel A, lanes 1 and 3; see below). A 5'- ^{32}P -labeled Bam HI probe was used to assay the 3' splice site of the first intervening sequence. Each fraction protected a 209 nt DNA fragment (Fig. 5, panel B, lanes 1-3; see 18 for explanation of why there are two bands in some lanes). This band is generated with properly spliced RNA. The 480-550 nt fragment (lanes 1 and 3) corresponds to nuclear RNA that retains IVS 1, consistent with the results with the Nco I probe. As determined by scanning densitometry of lanes 1 and 3 in panels A and B, the IVS 1-containing RNA accounts for approximately 25% of the total nuclear globin RNA.

It was important to confirm this value, because globin RNAs with IVS 1 and (probably) a normal 5' terminus are retained in the nucleus (12,18). If IVS 1 was intact in most of the nucleus-restricted, upstream RNAs, then it might be the IVS, not the upstream sequences, that are responsible for the failure to transport. Nuclear RNA was annealed with a 3'-end-labeled Bal I probe. The data indicate that 30% of the nucleus-restricted upstream RNA molecules retain IVS 1 (Fig. 6, lane 1). The structure and metabolism of these RNAs have not been investigated further. We conclude from experiments with these probes that IVS 1 is excised properly from 70-75% of the nuclear upstream RNA molecules. This experiment also confirms the observation that little, if any, of the upstream RNA is transported (Fig. 6, lane 2).

Hybridization and S1 treatment of reticulocyte and transfected cell RNAs with a 3'-³²P-labeled Acc I probe produces a 212 nt band (Fig. 7, panel A, lanes 1-3). This fragment corresponds to an RNA that is spliced at the normal 5' site of the second IVS. Hybridization to a 5'-³²P-labeled Eco RI probe generates a 53 ± 2 nt band, indicating normal splicing at the 3' end of IVS 2 (Fig. 7, panel B, lanes 1-3; see also 12). The conclusion from these experiments (Figs. 5-7) is that the majority of the nuclear RNA molecules is spliced in the same manner as authentic globin mRNA.

To investigate the 3' terminus, RNAs were assayed with a 3'-³²P-labeled Eco RI DNA fragment. The 209 nt band observed with nuclear and cytoplasmic RNAs was the same as that generated with authentic reticulocyte globin mRNA (Figure 7, panel C, lanes 1-3). The smaller band observed with transfected cell RNAs (lanes 1 and 2) was visible also in the reticulocyte lane after a longer exposure. The intensity of the 1200 nt band (nuclear RNA, lane 1) is only 5% of that of the 209 nt band. This result indicates that the 3' terminus of at least 95% of the nuclear globin RNA molecules is indistinguishable from that of normal β globin mRNA. Moreover, most of the upstream globin RNA is polyadenylated (Fig. 5, lanes 3).

Turnover of upstream RNAs

The data of Figures 3 - 7 could be interpreted in three ways: 1. The upstream RNA is not transported from the nucleus; the +1 mRNA is transported; both RNAs are relatively stable. 2. The upstream RNAs are synthesized at high levels but are degraded so rapidly that they never have the opportunity to be transported to the cytoplasm; the +1 mRNA is stable and is transported. Interpretation 2 is consistent with the S1 mapping experiments, which measure only steady-state RNA levels. 3. The upstream RNAs are transported to the cytoplasm and are degraded there so rapidly that

Figure 7. S_1 analysis of splicing sites in IVS 2 and of the termination site. The RNAs used in these assays were the same as those in Figure 5. The probes were as follows: Panel A: ^{32}P -3'-labeled *AccI*, 1100 nt, extending from the *AccI* site in coding block II to the *EcoRI* site in coding block III. Panel B: ^{32}P -5'-labeled *EcoRI*, 3739 nt, extending from the *EcoRI* site in coding block III into pBR322 sequences. The probe was made strand-specific by secondary digestion with *MspI*. Panel C: ^{32}P -3'-labeled *EcoRI*, 1465 nt, extending from the *EcoRI* site in coding block III into pBR322 sequences. The human genomic portion of the probe includes sequences from the *EcoRI* site to the 3'-flanking *MspI* site (+2154). Panels A and C: The RNAs were hybridized for 21 hours (1.9×10^4 cpm per reaction). S_1 -resistant fragments were separated in an 8% acrylamide gel, which was exposed for 3 days with a Quanta III intensifying screen. Lane 1: 10 μg nuclear RNA. Lane 2: 10 μg cytoplasmic RNA. Lane 3: 230 pg human adult

reticulocyte RNA. Lane 4: 6 μ g *E. coli* tRNA. Panel B: Hybridization for 20 hours (1×10^4 cpm/per reaction). S_1 -resistant fragments were separated in an 8% acrylamide gel, which was exposed for 31 hours with a Quanta III intensifying screen. Lane 1: 5 μ g nuclear RNA. Lane 2: 5 μ g cytoplasmic RNA. Lane 3: 800 pg adult human reticulocyte RNA. Lane 4: 6 μ g *E. coli* tRNA.

they fail to accumulate to detectable levels.

To distinguish between the first two possibilities, transfected cells were cultured for six hours in the absence or presence of 15 μ g/ml actinomycin D, to block de novo RNA synthesis. Labeling of cells with 3 H-uridine revealed that this level of drug reduces RNA synthesis by 95% (unpublished observations). Cells were fractionated, and nuclear and cytoplasmic RNAs were assayed by S_1 mapping. The amount of upstream RNA should be low or undetectable after a six hour chase if the RNA is degraded rapidly. The results indicate that the upstream RNAs are present at easily detectable levels in cells treated for six hours with actinomycin (Fig. 8, lane 3). As determined by scanning densitometry of this autoradiogram, the half-lives of the upstream and +1 RNAs in the nucleus are 2.5 and 4 hr, respectively. Although more complete characterization of RNA metabolism requires kinetic (pulse-chase) experiments, this simple chase protocol does indicate that the nuclear upstream RNAs are not degraded so rapidly as to preclude their being transported. Even if the transport time were as long as 2 hr in these cells (see Discussion), at least a portion of the RNA (half-life of 2-3 hr.) should reach the cytoplasm, if it were capable of being transported. Therefore, these results support the hypothesis that the upstream RNAs are located in the nucleus because they are not transported. It is unclear why the nuclear half-life of the +1 mRNAs is so long under these conditions. One possibility is that the nuclei are contaminated with small amounts of cytoplasm. Scanning densitometry reveals that the cytoplasm:nucleus ratio of +1 mRNA is 7:3 (Fig. 4, lanes 2, 5, and 6), suggesting that there is some contamination. Another possibility is that actinomycin D retards the transport of globin mRNA in these cells (see 34). Finally, these data do not exclude the possibility that the upstream RNAs are transported and rapidly degraded in the cytoplasm.

DISCUSSION

The major finding of these experiments is that globin RNAs containing additional nucleotides upstream from the +1 cap site remain in the nucleus. This conclusion is based on experiments comparing the intracellular location

Figure 8. Actinomycin D chase of transfected cell globin RNAs. 1×10^7 Cells were transfected with pDCQ21. 47 hours after transfection the cells were labeled for one hour with ^3H -uridine (30 Ci/ml) to monitor the effectiveness of the chase (see text). After labeling, half of the cells were harvested. The other half received fresh medium with actinomycin D at 15 g/ml. These cells were incubated at 37°C for 6 hours and then harvested. Nuclear and cytoplasmic RNA fractions were prepared and hybridized to the $5'$ - ^{32}P -labeled HaeIII probe (Fig. 2) for 14 hr, and the reactions were treated with S_1 nuclease and electrophoresed. The film was exposed for 7 days. Lane 1: 5 μg nuclear RNA, time zero. Lane 2: 5 μg cytoplasmic RNA, time zero. Lane 3: 5 μg nuclear RNA, time 6 hours. Lane 4: 5 μg cytoplasmic RNA, time 6 hours. Lane 5: 340 μg human adult reticulocyte RNA. Lane 6: 6 μg *E. coli* tRNA.

of globin upstream RNA versus globin mRNA (+1) molecules. The upstream RNAs do not have the normal mRNA cap site nucleotides at their $5'$ termini, but neither do the +30 RNAs, which are transported to the cytoplasm. Therefore, we favor the notion that the additional nucleotides are directly responsible for confining the upstream RNAs to the nucleus. The possibility that these RNAs are transported and then rapidly degraded in the cytoplasm has not been excluded. The actinomycin D data provide some assurance that the upstream RNAs are confined because they are not transported, rather than because they are synthesized and degraded rapidly. It may be that actinomycin chase experiments do not yield authentic RNA half-life values (35). However, numerous studies have revealed that the half-life of heterogeneous nuclear RNA (hnRNA) in actinomycin-treated cells is one hour or less (reviewed in 36; for examples, see 37-40). Therefore, even if the half-life of the

globin upstream RNAs is not 2.5 hr., it is clearly longer than that of the average hnRNA molecule (in actinomycin-treated cells) and longer than the 20 minutes or so required to transport newly synthesized mRNA to the cytoplasm (40-43). Furthermore, there are no documented examples of specific RNAs that are stabilized selectively by actinomycin.

The data illustrate a limitation of the transfection system, at least with COS cells: the amount of total RNA may not indicate the amount of translated mRNA. COS cells produce abundant globin RNA, but relatively little globin mRNA, under these conditions. The production of abundant upstream RNA in COS cells transfected with SV40 ori-containing plasmids seems to occur consistently with enhancer-dependent genes (9,27,44). Upstream and downstream RNAs are also detected in other cells, such as NIH 3T6 fibroblasts (3).

Little is known about the crucial question of why some gene products are transported to the cytoplasm while others are restricted to the nucleus. The basic observation is that conversion of nuclear precursors into cytoplasmic mRNA occurs rapidly and efficiently (34,45-50), but most hnRNA molecules are restricted to and turn over in the nucleus (reviewed in 36 and 51). Special mechanisms for controlling processing and/or transport of large classes of RNA may function in certain circumstances, for example, in virus-infected cells (52-53), in cells cultured in medium deficient in nutrients (54-57), and in embryonic development (58). These mechanisms probably differ from those that operate in normal somatic cells, in which more is known about what is not essential for transport than what is. Splicing is not obligatory for transport (12,18,59-62). Some polyadenylated RNA turns over in the nucleus (63), and some non-adenylated RNAs (tRNA, rRNA, histone mRNA) are transported to the cytoplasm, indicating that polyadenylation is not a sufficient signal for transport. It is unlikely that a methylated cap is required, because ribosomal and tRNAs are uncapped but are efficiently transported. On the other hand, ribosomal and mRNAs might utilize distinct transport systems, and mRNA transport might require a combination of signals, such as capping, methylation, and polyadenylation. Lack of any one of these might block transport.

Experiments measuring transport of endogenous RNA from isolated nuclei have been difficult to interpret (64-65). Recent approaches that exploit defined nucleic acids injected into the oocyte cytoplasm or germinal vesicle provide a more defined system (66-70). In fact, precursor tRNAs with additional 5' terminal nucleotides are not transported out of the oocyte

germinal vesicle until the 5' nucleotides are cleaved (66). The results presented here are consistent with the tRNA experiment. We have not completely mapped the upstream globin sequences, nor are we certain if all of them are derived from human DNA (Fig. 3). In view of the relatively low levels of upstream RNAs generated by plasmid DCQ21 (Fig. 3, lane 2), there is no straightforward method for determining unequivocally the structure of all of their 5' ends. In any event, the heterogeneity would make it difficult to define those sequences that block transport, even if their structures were known. A more logical approach might exploit constructs containing a strong promoter located upstream from the cap site, such that only one or a few RNAs with a fixed number of upstream nucleotides would be generated in each transfection experiment. It should then be possible to determine more directly what upstream sequences or modifications block transport. It would be particularly interesting to know if specific sequences in nucleus-restricted RNAs are recognized by specific nuclear proteins.

ACKNOWLEDGMENTS

We are grateful to Dr. Janet Mertz for giving us COS cells, to Dr. Tom Maniatis for the human beta globin gene, to Jill Rojek for growing plasmids, and to Mary Jo Gilden for typing the manuscript. This work was funded by grants CA-07175 and CA-23076 from the NCI.

REFERENCES

1. Breathnach, R., and Chambon, P. (1981). *Ann. Rev. Biochem.* 50, 349-383.
2. Banerji, J., Rusconi, S., and Schaffner, W. (1981). *Cell* 27, 299-308.
3. Dierks, P., van Ooyen, A., Cochran, M. D., Dobkin, C., Reiser, J., and Weissmann, C. (1983). *Cell* 32, 695-706.
4. Lee, F., Mulligan, R., Berg, P., and Ringold, G. (1981). *Nature* 294, 228-232.
5. Huang, A. L., Ostrowski, M. C., Berard, D., and Hager, G. L. (1981). *Cell* 27, 245-255.
6. Doehmer, J., Barinaga, M., Vale, W., Rosenfeld, M. G., Verma, I. M., and Evans, R. M. (1982). *Proc. Nat. Acad. Sci. USA* 79, 2268-2272.
7. Schlosser, C. A., Steglich, C., de Wet, J. R., and Scheffler, I. E. (1981). *Proc. Nat. Acad. Sci. USA* 78, 1119-1123.
8. Corces, V., Pellicer, A., Axel, R., and Meselson, M. (1981). *Proc. Nat. Acad. Sci. USA* 78, 7038-7042.
9. Pelham, H. R. B. (1982). *Cell* 30, 517-528.
10. Chao, M. V., Mellon, P., Charnay, P., Maniatis, T., and Axel, R. (1983). *Cell* 32, 483-493.
11. Chu, G., and Sharp, P. A. (1981). *Nature* 289, 378-382.
12. Busslinger, M., Moschonas, N., and Flavell, R. A. (1981). *Cell* 27, 289-298.
13. Fukumaki, Y., Ghosh, P. K., Benz, E. J., Reddy, V. B., Lebowitz, P., Forget, B. G., and Weissman, S. M. (1982). *Cell* 28, 585-593.

14. Treisman, R., Orkin, S. H., and Maniatis, T. (1983). *Nature* 302, 591-596.
15. Cooper, G. M. (1982). *Science* 217, 801-806.
16. Duesberg, P. H. (1983). *Nature* 304, 219-226.
17. Felber, B. K., Orkin, S. H., and Hamer, D. H. (1982). *Cell* 29, 895-902.
18. Treisman, R., Proudfoot, N. J., Shander, M., and Maniatis, T. (1982). *Cell* 29, 903-911.
19. Goldsmith, M. E., Humphries, R. K., Ley, T., Cline, A., Kantor, J. A., and Nienhuis, A. W. (1983). *Proc. Nat. Acad. Sci. USA* 80, 2318-2322.
20. Orkin, S. H., Sexton, J. P., Goff, S. C., and Kazazian, H. H. (1983). *J. Biol. Chem.* 258, 7249-7251.
21. Orkin, S. H., Sexton, J. P., Cheng, T., Goff, S. C., Giardina, P. J. V., Lee, J. I., and Kazazian, H. H. (1983). *Nucl. Acids Res.* 11, 4727-4734.
22. Dobkin, C., Pergolizzi, R. G., Bahre, P., and Bank, A. (1983). *Proc. Nat. Acad. Sci. USA* 80, 1184-1188.
23. Whitelaw, E., and Proudfoot, N. J. (1983). *Nucl. Acids Res.* 11, 7717-7733.
24. Moschonas, N., de Boer, E., Grosveld, F. G., Dahl, H. H. M., Wright, S., Shewmaker, C. K., and Flavell, R. A. (1981). *Nucl. Acids Res.* 9, 4391-4401.
25. Kinniburgh, A. J., Maquat, L. E., Schedl, J., Rachmilewitz, E., and Ross, J. (1982). *Nucl. Acids Res.* 10, 5421-5427.
26. Maquat, L. E., Kinniburgh, A. J., Rachmilewitz, E. A., and Ross, J. (1981). *Cell* 27, 543-553.
27. Treisman, R., Green, M. R., and Maniatis, T. (1983). *Proc. Nat. Acad. Sci. USA* 80, 7428-7432.
28. Wold, B., Wigler, M., Lacy, E., Maniatis, T., Silverstein, S., and Axel, R. (1979). *Proc. Nat. Acad. Sci. USA* 76, 5684-5688.
29. Breathnach, R., Mantei, N., and Chambon, P. (1980). *Proc. Natl. Acad. Sci. USA* 77, 740-744.
30. Lawn, R. M., Efstratiadis, A., O'Connell, C., and Maniatis, T. (1980). *Cell* 21, 647-651.
31. Poncz, M., Schwartz, E., Ballantine, M., and Surrey, S. (1983). *J. Biol. Chem.* 258, 11599-11609.
32. Ley, T. J., and Nienhuis, A. W. (1983). *Biochem. Biophys. Res. Commun.* 112, 1041-1048.
33. Carlson, D. P., and Ross, J. (1983). *Cell* 34, 857-864.
34. Bastos, R. N., and Aviv, H. (1977). *Cell* 11, 641-650.
35. Singer, H. H., and Penman, S. (1973). *J. Mol. Biol.* 78, 321-334.
36. Lewin, B. (1975). *Cell* 4, 11-20.
37. Scherrer, K., Marcaud, L., Zajdela, F., London, I. M., and Gros, F. (1966). *Proc. Nat. Acad. Sci. USA* 56, 1571-1578.
38. Warner, J. R., Soeiro, R., Birnboim, H. C., Girard, M., and Darnell, J. E. (1966). *J. Mol. Biol.* 19, 349-361.
39. Bramwell, M. E., and Harris, H. (1967). *Biochem. J.* 103, 816-830.
40. Penman, S., Vesco, C., and Penman, M. (1968). *J. Mol. Biol.* 34, 49-69.
41. Raskas, H. J., and Okubo, C. K. (1971). *J. Cell Biol.* 49, 438-449.
42. Jelinek, W., Adesnik, M., Salditt, M., Sheiness, D., Wall, R., Molloy, G., Philipson, L., and Darnell, J. E. (1973). *J. Mol. Biol.* 75, 515-532.
43. Perry, R. P., Kelley, D. E., and La Torre, J. (1974). *J. Mol. Biol.* 82, 315-331.
44. Humphries, K., Ley, T., Turner, P., Moulton, A.D., and Nienhuis, A.W. (1982) *Cell* 30, 173-183.
45. Ross, J. (1976). *J. Mol. Biol.* 106, 403-420.

46. Curtis, P. J., Mantei, N., van den Berg, J., and Weissmann, C. (1977). *Proc. Nat. Acad. Sci. USA* 74, 3184-3188.
47. Kwan, S.-P., Wood, T. G., and Lingrel, J. B. (1977). *Proc. Nat. Acad. Sci. USA* 74, 178-182.
48. Gilmore-Hebert, M., and Wall, R. (1978). *Proc. Nat. Acad. Sci. USA* 75, 342-345.
49. Crawford, R. J., and Wells, J. R. E. (1978). *Biochemistry* 17, 1591-1596.
50. Harpold, M. M., Dobner, P. R., Evans, R., Bancroft, F. C., and Darnell, J. E. (1979). *Nucl. Acids Res.* 6, 3133-3144.
51. Lewin, B. (1975). *Cell* 4, 77-93.
52. Therwath, A., and Scherrer, K. (1978). *Proc. Nat. Acad. Sci. USA* 75, 3776-3780.
53. Beltz, G. A., and Flint, S. J. (1979). *J. Mol. Biol.* 131, 353-373.
54. Johnson, L. F., Williams, J. G., Abelson, H. T., Green, H., and Penman, S. (1975). *Cell* 4, 69-75.
55. Bramwell, M. E., and Handmaker, S. D. (1975). *Exptl. Cell Res.* 91, 358-364.
56. Mangiarotti, G., Zuker, C., Chisholm, R. L., and Lodish, H. F. (1983). *Mol. Cell. Biol.* 3, 1511-1517.
57. Collins, M. L., Wu, J.-S. R., Santiago, C. L., Hendrickson, S. L., and Johnson, L. F. (1983). *Mol. Cell. Biol.* 3, 1792-1802.
58. Chan, L.-N. L. (1976). *Nature* 261, 157-159.
59. Schochetman, G., and Perry, R. P. (1972). *J. Mol. Biol.* 63, 591-596.
60. Klessig, D. F., and Chow, L. T. (1980). *J. Mol. Biol.* 139, 221-242.
61. Ghosh, P. K., Roy, P., Barkan, A., Mertz, J. E., Weissman, S. M., and Lebowitz, P. (1981). *Proc. Nat. Acad. Sci. USA* 78, 1386-1390.
62. Gruss, P., Efstratiadis, A., Karathanasis, S., Konig, M., and Khoury, G. (1981). *Proc. Nat. Acad. Sci. USA* 78, 6091-6095.
63. La Torre, J., and Perry, R. P. (1973). *Biochim. Biophys. Acta* 335, 93-101.
64. Raskas, H. J. (1971). *Nature New Biol.* 233, 134-136.
65. Stuart, S. E., Rottman, F. M., and Patterson, R. J. (1975). *Biochem. Biophys. Res. Commun.* 62, 439-447.
66. Melton, D. A., and Cortese, R. (1979). *Cell* 18, 1165-1172.
67. Zasloff, M., Rosenberg, M., and Santos, T. (1982). *Nature* 300, 81-84.
68. De Robertis, E. M., Lienhard, S., and Parisot, R. F. (1982). *Nature* 295, 572-577.
69. Wickens, M. P., and Gurdon, J. B. (1983). *J. Mol. Biol.* 163, 1-26.
70. Zasloff, M. (1983). *Proc. Nat. Acad. Sci. USA* 80, 6436-6440.
71. Chu, G., and Sharp, P. A. (1981) *Gene* 13, 197-202.
72. Ross, J., and Pizarro, A. (1983). *J. Mol. Biol.* 167, 607-617.
73. Berk, A. J., and Sharp, P. A. (1977). *Cell* 12, 721-732.
74. Weaver, R. F., and Weissmann, C. (1979). *Nucl. Acids Res.* 7, 1175-1193.
75. Sutcliffe, J. G. (1979). *Cold Spring Harbor Symp. Quant. Biol.* 43, 77-90.
76. Reddy, V. B., Thimmappaya, B., Dhar, R., Subramanian, K. N., Zain, B. S., Pan, J., Ghosh, P. K., Celma, M. L., and Weissman, S. M. (1978). *Science* 200, 494-502.