Identification of a Ribosome-Binding Site for a Leader Peptide Encoded by Rous Sarcoma Virus RNA

ROBERT B. PETERSEN, CHARLES H. HENSEL, AND PERRY B. HACKETT*

Department of Genetics and Cell Biology, University of Minnesota, St. Paul, Minnesota 55108

Received 21 February 1984/Accepted 1 June 1984

A new method for identifying ribosome-binding sites was developed to determine whether AUG codons in the 5'-terminal RNA sequence of Rous sarcoma virus were used to initiate protein synthesis. We found that when translation is inhibited, the major ribosome-binding site on Rous sarcoma virus RNA is at the 5'-proximal AUG codon, even though the primary translational product from this RNA, Pr76^{gag}, is encoded behind the fourth AUG codon 331 bases downstream from the observed initiation site. These results suggest that ribosomes can initiate translation on Rous sarcoma virus RNA at more than one site, thereby producing a seven-amino-acid peptide, as well as the *gag* gene polyprotein precursor of *M_r* 76,000.

Studies of initiation of protein synthesis in eucaryotes have indicated that the process is, for the most part, relatively simple. The scanning hypothesis (21) describes the process as follows. (i) A 40S ribosomal subunit locates the 5' end of an mRNA, which is generally capped with a 7-methyl-guanine residue. (ii) The 40S subunit, with a charged tRNA^{met} and initiation factors, scans the mRNA until it encounters an AUG codon to form an initiation complex. (iii) The 60S ribosomal subunit binds to this initiation complex to form a complete ribosome. (iv) Translation begins. (v) Translation continues up to a termination codon. More than 90% of eucaryotic mRNAs have translational properties that are in accord with this model (21). The scanning hypothesis also provides an explanation for the observation made by Jacobson and Baltimore (19) that eucaryotic mRNAs each direct the synthesis of a single protein. However, exceptions to both general features have been found. Several viral mRNAs and a few cellular mRNAs encode proteins that are initiated at sites that follow the 5'-proximal AUG codon (15, 16, 28, 30), and some mRNAs direct the synthesis of more than one protein (3, 20, 23). The mRNAs encoding the proteins of Rous sarcoma virus (RSV) are also exceptional with respect to the initiation of protein synthesis (14, 29, 46).

RSV mRNAs are derived from a 9.5-kilobase RNA transcript, designated RSV RNA, that is spliced to form two or three smaller mRNAs that have a common 5' leader sequence of 389 bases and common 3' ends of ca. 3 kilobases (14, 46). A comparison of the sequences of these mRNAs with the identified splicing sites and the known protein sequences of the gene products indicates that in the RSV strain SR-A the initiation codon for three of the four gene products is the fourth AUG codon from the 5' terminus (31, 44, 46). An AUG initiation codon shift, from the fourth to the fifth AUG codon, appears to be necessary for the synthesis of the *src* gene product (14, 29, 43).

There are two primary methods of determining the sites of initiation of protein synthesis: (i) comparison of a known protein sequence (or its possible precursors) with the corresponding nucleotide sequence (39) and (ii) ribosome-binding site analysis (26, 41). Protein analysis has shown that the major product of RSV RNA translations is Pr76^{gag}, the polyprotein product of the *gag* gene, which is initiated at the

fourth AUG codon, AUG(*gag*). Pr180^{gag/pol}, a fusion protein containing the *gag* and *pol* gene products, is also initiated at AUG(*gag*) but presumably is made from a spliced RNA product of RSV RNA (46, 49). However, two early studies indicated that ribosomes bind at sites within the first 100 bases of the leader sequence, ca. 300 bases upstream from AUG(*gag*) (5, 40). These results suggest that initiation of translation at AUG(*gag*) may occur via either of two modes. The first model proposes that the ribosome scans the entire RSV leader sequence of 374 bases, initiating or not at the AUG codons between the 5' cap and AUG(*gag*), before synthesizing Pr76^{gag}; this is a linear scanning mode. The alternative model proposes that ribosomes may translocate from one region of the RSV leader RNA sequence to AUG(*gag*) and thereby skip over sequences between the jump site and AUG(*gag*); this is a spatial scanning mode and may be mediated in part by the secondary structure that the leader RNA can assume.

As a first step in determining the mode of initiation of protein synthesis at AUG(*gag*), we analyzed the sites of ribosome binding during stalled translation by a procedure that is more accurate than the methods previously used. Our results suggest that ribosomes initiate translation at the 5'-proximal AUG codon and consequently that RSV RNA may encode both a small leader peptide of seven amino acid residues and the major polyprotein product Pr76^{gag}. These results are in accord with the linear scanning mode, but they do not rule out a ribosome translocation after the synthesis of the first leader peptide.

MATERIALS AND METHODS

Restriction endonucleases and T4 DNA ligase were from Bethesda Research Laboratories and New England Biolabs. RNase T1 for digestion of ribosome initiation complexes was from Boehringer Mannheim Biochemicals, and sequencing-grade RNase T1 and T4 polynucleotide kinase were from P-L Biochemicals. Anisomycin was obtained from Sigma Chemical Co. RSV RNA was obtained from chicken embryo fibroblasts infected with the SR-A strain of RSV as previously described (14).

Molecular cloning of RSV (strain SR-A) DNA. The 573-base-pair (bp) *Eco*RI-*Bam*HI fragment, isolated from p165 (obtained from W. DeLorbe)—a derivative of the full RSV proviral DNA clone pSRA2—was cloned into pUC8 (47) in the polycloning site. Either this clone or six subclones

* Corresponding author.

derived from further restrictions of the 573-bp insert with *Hae*III were used as sources of RSV-specific DNA. The four internal fragments with blunt ends from the *Hae*III cleavages were ligated into pUC8 opened at the *Hinc*II site; the 5'-terminal fragment was cloned into pUC8 cleaved with *Hinc*II and *Eco*RI, and the *gag* gene-specific fragment with *Hae*III-*Bam*HI termini was cloned into pUC8 cleaved with *Hinc*II and *Bam*HI. Isolation of plasmid DNAs was by the procedure of Ish-Horowitz and Burke (18) with slight modifications. Due to the very high number of pUC8 recombinant plasmids per cell, up to 5 µg of plasmids per ml of cells could be obtained without plasmid amplification.

Translation in vitro. The rabbit reticulocyte translational system was essentially as described by Pelham and Jackson (34) with some modifications (4). RSV virion RNA was added to a final concentration of 40 ng/µl. In experiments utilizing denatured template, the RNA was treated for 5 min in 10 mM methyl mercuric hydroxide (Alpha-Ventron) at room temperature (2). The translational products were separated on 9% polyacrylamide gels, with a 4.5% stacking gel, according to the procedure of Laemmli (25; see reference 4).

Isolation of ribosome initiation complexes. Ribosome initiation complexes were formed in translational reactions essentially as described above, except that [³⁵S]methionine was the only amino acid added to the translational mix. The final translation reaction conditions were 250 µM MgCl₂–90 mM KCl–9 mM dithiothreitol–9 mM creatin phosphate–40 µg of calf liver tRNA (Boehringer) per ml–15 µg of exogenous mRNA per ml in a final reaction volume of 350 µl. Before addition of the mRNA, the lysate was incubated for 5 min with anisomycin at a final concentration of 270 µM. After addition of the mRNA and incubation at 29°C for the times indicated in the text, unprotected RNA was digested by the addition of 100 U of RNase T1 (Boehringer) per ml, and the incubation was continued for another 10 min at 29°C. The reactions were diluted with 2 volumes of ice-cold gradient buffer (50 mM NaCl, 20 mM MgCl₂, 20 mM HEPES [*N*-2-hydroxyethylpiperazine-*N'*-2-ethanesulfonic acid], pH 7.4), and the monosomes were isolated from the rest of the translational components by centrifugation through 15 to 50% sucrose gradients in an SW41 rotor at 36,000 rpm at 4°C for 5.5 h as described by Rose (37). RNA was isolated from monosomes by digestion of the protein with proteinase K (Merck & Co., Inc.) at 350 µg/ml for 30 min at 37°C, followed by phenol-chloroform-isoamyl alcohol (1:1:0.02) extraction and ethanol precipitation.

Selection of RSV RNA fragments by hybridization. Either the plasmid containing the full 573-bp *Eco*RI-*Bam*HI insert or the six subclones, as described above, were linearized by cleavage with *Eco*RI, denatured, and covalently coupled to 1-cm² diazobenzoyloxymethyl-paper (DBM) disks as described by Goldberg et al. (12). The disks were prehybridized in competition buffer (50% formamide, 10 mM HEPES [pH 7.4], 1× Denhardt's solution, 1 mg of RNA from rabbit reticulocytes per ml, 100 µg of denatured herring sperm DNA per ml, 0.1% sodium dodecyl sulfate [wt/vol]) at 40°C for at least 3 h before use. The RNA fragments isolated from monosomes were labeled at their 5' ends by using T4 polynucleotide kinase (38). Briefly, 5 to 10 µg of isolated RNA was incubated in a 20-µl reaction containing 50 mM Tris (pH 7.4), 10 mM MgCl₂, 1 mM dithiothreitol, 0.5 to 1 mCi of [³²P]ATP (ca. 8,000 Ci/mmol; ICN Pharmaceuticals Inc.), and 10 U of T4 polynucleotide kinase. After incubation at 37°C for 30 min, the reaction was precipitated twice with 1 volume of 5 M NH₄Ac, 20 µg of tRNA (Boehringer), and 6 volumes of ethanol to remove the unincorporated ³²P and resuspended in competition buffer. The hybridization solu-

tion was heated in a boiling water bath for 5 min before addition to the DBM disks. Hybridization was at 40°C for 40 h. The disks were washed three times in 40% formamide–3× SSC (1× SSC is 0.15 M NaCl plus 0.015 M sodium citrate)–0.1% sodium dodecyl sulfate for 30 min each at the hybridization temperature, followed by three washes in 0.3× SSC–0.1% sodium dodecyl sulfate for 30 min at the hybridization temperature. The hybridized RNA fragments were eluted at 65°C for 15 min with freshly deionized formamide.

Polyacrylamide gel electrophoresis of RNA fragments. The hybrid-selected RNA fragments were precipitated in ethanol and resuspended in 95% formamide–25 mM Tris (pH 8.0)–10 mM EDTA–0.02% (wt/vol) bromphenol blue–0.02% (wt/vol) xylene cyanol. Before being loaded onto gels, the samples were heated in a boiling water bath for 5 min. The fragments were separated on 10% polyacrylamide gels (40 cm long, 17 cm wide, 0.3 mm thick) containing 0.3% bisacrylamide, 8.3 M urea, 100 mM Tris-borate (pH 8.3), and 1 mM EDTA. Electrophoresis was at 2 to 3 kV, to maintain a temperature between 40°C and 50°C, until the bromphenol blue marker migrated 35 cm.

Analysis of hybrid-selected ribosome-protected fragments. Ribosome-protected RSV oligoribonucleotides were localized in gels by autoradiography, excised, and eluted in 1.5-ml minifuge tubes containing 15 volumes of 0.5 M NH₄Ac (pH 7.5)–10 mM MgCl₂–0.1 mM EDTA–0.1% sodium dodecyl sulfate for 6 to 20 h at room temperature. After phenol-chloroform extraction of the supernatant, the recovered RNA fragments were precipitated by addition of 3 µg of tRNA (Boehringer)–0.1 volume of 2 M NH₄Ac (pH 7.0)–3 volumes of 95% ethanol. The eluted RNA fragments were partially digested with RNase T1 (8) in 10-µl reactions containing 7 M urea, 20 mM sodium citrate (pH 5.0), 1 mM EDTA, 0.05% (wt/vol) each of bromphenol blue and xylene cyanol, and 1 U of sequencing-grade RNase T1 (P-L Biochemicals); digestion was for 15 min at 50°C. The reactions were chilled on ice and immediately electrophoresed through 20% acrylamide–0.6% bisacrylamide gel in the same buffer and under the same conditions as before. The gels were wrapped in Saran Wrap for autoradiography at –80°C with Du Pont Cronex Lightning-Plus intensifying screens and Kodak X-Omat AR film.

RESULTS

The goal of the experiments described below was to determine the site of ribosome binding for the production of Pr76^{gag}. Our strategy was to hybridize ribosome-protected fragments of RSV RNA to cloned DNA homologous to the RSV leader RNA, the RNA between the cap site and AUG(*gag*). Figure 1A shows the arrangement of the RSV leader RNA with the three, presumably unused, open reading frames situated behind the three 5'-proximal AUG codons, designated AUG(1), AUG(2), and AUG(3), and the known site of initiation of Pr76^{gag} at nucleotide 372, AUG(*gag*) (44). We used a standard procedure for obtaining ribosome-protected fragments (37), but rather than using uniformly labeled mRNA, as is the custom in ribosome-binding assays, we labeled the ribosome-protected fragments at their 5' ends after isolation from monosomes. Using this procedure, we were able to characterize directly the whole ribosome-protected fragment. In addition, we stalled initiation of translation with anisomycin, an antibiotic that inhibits the elongation step of polypeptide synthesis (27), thus assuring that we only were stalling ribosomes that had productively initiated protein synthesis.

We tested the selectivity of this procedure by examining the ribosome-protected RNA fragments that were produced

FIG. 1. Map of the 5' end of RSV (strain SR-A) RNA. (A) Positions of the AUG codons in the leader region of the SR-A strain of RSV RNA. The AUG codons are designated by numbers above the map in the order of their proximity to the 5' end of the RNA, the cap site. The fourth AUG codon is AUG(*gag*). The shaded regions represent the open reading frames behind the AUG codons, and the numbers below indicate the number of encoded amino acids. The scale at the top indicates the nucleotide positions in the RSV RNA; the 5' cap is at position 1. (B) Subclones of the RSV RNA leader region produced by *Hae*III cleavages of the 573-bp *Eco*RI-*Bam*HI segment from RSV proviral DNA. The clones are aligned with the map shown in (A) above. The arrow indicates the position of the 5' end of the RSV RNA, the circles indicate the AUG codon positions, and the stippled portion indicates the tRNA^{Trp}-binding site. E, *Eco*RI site; H, *Hae*III site; B, *Bam*HI site. (C) Sequence of the major ribosome-binding site as determined from mapping (see Fig. 3 and 4). The lines below the map indicate the five major, RNase T1-generated fragments that bind to pSRA5'-2 (Fig. 3); the numbers to the right indicate the length of these oligonucleotides. AUG(1) and AUG(2) are indicated by the overlines.

from cell-free translational reactions directed by either RSV RNA or mRNA from rabbit reticulocytes. To determine whether the initial secondary structure of the mRNA might influence initiation of translation (33), the RSV RNA was either denatured with methylmercuric hydroxide (2) or left untreated before addition to the cell-free translational system. Additionally, we allowed stalled ribosome complexes to form for either 3 or 20 min to determine whether ribosomes might bind to different sites on the RNA as a function of incubation time.

For this experiment, RNA fragments resulting from RNase T1 digestion of monosomes were labeled with ³²P at their 5' ends and hybridized to DBM disks (12) containing pSRA5' DNA, a subclone with 573 bp from RSV proviral DNA (strain SR-A) (7) that has the complete 374-bp leader, 150 bases of the *gag* gene, and 49 bases of the U3 region of the proviral long terminal repeat. The entire sequence shown in Fig. 1B is contained in this plasmid. The oligoribonucleotides that bound to the DNA were eluted, precipitated, and then separated by electrophoresis (Fig. 2). The specificity of the selection by hybridization is shown by the multitude of intensely labeled fragments from the RSV RNA directed translations (lanes 3, 4, and 5), as compared with the lighter background fragments obtained from the translation directed by globin RNA (lane 2). Lane 1 shows the size distribution of the fragments before selection on the DBM disks; we estimate from the ratio of total counts in the DBM-selected RNAs relative to the nonhybridizing RNA that there was a 1,000-fold purification of the RSV sequences from the material shown in lane 1. The changes in the distributions of labeled fragments selected by hybridization (lanes 3, 4, and 5) relative to the fragment distribution of unselected RNA

(lane 1) demonstrate that the hybridizing RNA fragments comprise a specific subset of the initial ³²P-labeled oligoribonucleotides.

There were no detectable differences in the distribution and sizes of the oligoribonucleotides in lanes 3, 4, and 5 of Fig. 2, suggesting that ribosomes protect the same sites on RSV RNA regardless of prior denaturation (lanes 4 and 5) or the duration of the ribosome-binding assay (lanes 3 and 4). We repeated the above experiment by using aurin tricarboxylic acid to inhibit formation of initiation complexes to establish that bona fide ribosome protection produced the observed fragments, which hybridized to the RSV DNA on DBM disks, rather than spurious association of the RNA fragments with ribosomes. The major ribosome-protected RNA fragments seen in lanes 3 through 5 of Fig. 2 were not present when aurin tricarboxylic acid inhibited the formation of initiation complexes (data not shown), indicating that the bands seen in Fig. 2 resulted from ribosome protection at translational initiation sites.

We expected that there would be many ribosome-protected RSV RNA fragments, including those from 5'-proximal locations (as shown by others [5, 40]) and those coming from the neighborhood of AUG(*gag*), a known start site of protein synthesis. The data shown in Fig. 2 confirm this expectation. To reduce the complexity of the fragments under examination at any one time, the ribosome-protected fragments were selectively divided into classes according to their ability to hybridize to DNA complementary to specific regions of the RSV RNA. To achieve this, pSRA5' was digested with endonuclease *Hae*III, and the DNA fragments of 71 to 127 bases which were complementary to RSV RNA were subcloned into pUC8 (47). The subclones were linearized by a

FIG. 2. Distribution of ribosome-protected fragments selected by hybridization to pSRA5'. RNA fragments isolated from RNase T1-digested monosomes, produced in translational assays containing anisomycin, were labeled at their 5' ends and selected by hybridization to DBM disks containing covalently bound pSRA5' DNA. Lane 1, total RNA fragments before selection; lanes 2, 3, 4, and 5, fragments selected by hybridization from translational assays containing either globin mRNA, 20-min reaction (lane 2); RSV RNA, 20-min reaction (lane 3); RSV RNA, 3-min reaction (lane 4); or denatured RSV RNA, 3-min reaction (lane 5). The numbers to the left indicate the sizes of the prominent RSV-specific oligonucleotides in lanes 3, 4, and 5.

single *Eco*RI restriction endonuclease cleavage, denatured, and covalently attached to DBM disks. These subclones are designated pSRA5'-1 through pSRA5'-6; they divide the RSV leader into six nearly equal domains (Fig. 1B).

The results of one such ribosome protection assay with the subclones are shown in Fig. 3. Approximately 80% of the RSV-specific protected fragments that bound to the pSRA5' disk also bound to the pSRA5'-2 subcloned DNA. This indicated that ribosomes were translating sequences behind AUG(1) or AUG(2), or both. The sizes of the prominent oligonucleotides binding to pSRA5'-2 were 45, 47, 53, 57, and 60 bases (lane 3), as determined by using polyadenylate markers (lane 8). Faint RNA bands of ca. 82 and 89 nucleotides could be found hybridizing to pSRA5'-1 (lane 2), and RNA bands of 21, 22, 23, and 35 bases hybridized to pSRA5'-3 (lane 4), but no specific RNA fragments were found complementary to pSRA5'-4, pSRA5'-5, and pSRA5'-6 (lanes 5, 6, and 7). There was no detectable difference between the patterns obtained by using nondenatured and methyl mercuric hydroxide-denatured RSV RNA. Some

bands, which are seen in all of the lanes, are contaminants of the intensely labeled rRNA and tRNA fragments that can be removed only by rigorous washing of the filters and inclusion of a 10- to 50-fold excess of unlabeled rRNA and tRNA (unpublished data). Lane 1 shows the distribution of the labeled rRNA and tRNA fragments before selection by hybridization to the pSRA5' DNAs. The oligonucleotides that bind to all of the subclones, and thereby provide a weak background, can be found in lane 1 as prominent bands. We disregarded these fragments in the analyses described below.

By comparing the known sequence of the RSV RNA that could hybridize to pSRA5'-2 with the sizes of the protected RNase T1-generated oligonucleotides (RNase T1 cleaves after guanine residues), the possible origins for most of the fragments could be deduced. To identify precisely the prominent fragments, we prepared a RNase T1 map of the oligoribonucleotides by eluting them from the gel and partially digesting them with RNase T1 to produce a ladder of fragments (8), all labeled at their 5' ends. The results of this analysis are shown in Fig. 4. Lane 2 shows the complete RNase T1 map of the 60-base ribosome-protected fragment from pSRA5'-2, given as a sequence of T1 oligonucleotides that begins at residue 28 and continues to position 87. The sizes of the RNase T1-generated fragments are 8, 9, 10, 13, 16, 17, 20, 21, 25, 28, 36, 39, 45, 47, 53, 57, and 60 bases.

FIG. 3. Distribution of ribosome-protected RNA fragments produced from different regions of the RSV RNA leader sequence. Lane 1, RNA fragments before selection on the DBM disks; lanes 2 through 7, RNA fragments selected by hybridization to pSRA5'-1 through pSRA5'-6, respectively; lane 8, 32 P-labeled polyadenylate marker fragments. The scale on the right indicates the sizes of the oligonucleotides.

FIG. 4. Partial RNase T1 digestion products from isolated ribosome-protected RSV RNA fragments. Lanes 1, 3, 5, 7, and 9, Ribosome-protected fragments of 60, 57, 53, 47, and 45 bases, respectively. Lanes 2, 4, 6, 8, and 10, Partial RNase T1 digests of the 60-, 57-, 53-, 47-, and 45-base RSV RNA-protected fragments, respectively.

Lanes 4, 6, 8, and 10 show similar profiles of the RNase T1 fragments generated from partial cleavage of the 57-, 53-, 47-, and 45-base protected sequences isolated by hybridization to pSRA5'-2. These results demonstrate that the five major RNase T1 oligoribonucleotides that bind to pSRA5'-2 are related and that they are derived from the RSV RNA sequence that surrounds AUG(1) (see Fig. 1C). Thus, all of the major RSV RNA fragments have a common 5' end but differing 3' ends. This identification of the ribosome-protected fragments by RNase T1 mapping is based on the probability of obtaining a given spacing of guanine residues from random fragments; for the 60- and 45-base fragments the chances of such arrangements are ca. 10^{-15} and 10^{-11} , respectively.

Besides the major fragments that bind to pSRA5'-2, there are several other ribosome-protected oligonucleotides. The oligonucleotides of 82 and 89 nucleotides that bind to both pSRA5'-1 and pSRA5'-2 are fragments that appear to be extensions of the major protected pSRA5'-2 fragments; they span nucleotides 9 through 90 and 2 through 90, respectively. The 21-, 22-, 23-, and 25-base fragments that bind to pSRA5'-3, analyzed by the partial RNase T1 mapping procedure, probably come from a region between nucleotides 151 and 185. This might represent the effects of the secondary structure between the 3' ends of the protected fragments containing AUG(1) or a type of abortive binding by ribo-

somes to mRNA during protection studies, as has been observed by others (23). No observable RSV-specific fragments bound to pSRA5'-5 containing sequences complementary to those surrounding and including AUG(*gag*). Based on the intensities of the RSV-specific bands binding to pSRA5'-2, if any sequences around AUG(*gag*) were protected to 5 to 10% of that seen around AUG(1), they would have been detectable by our assay.

The lack of RNase T1 fragments binding to pSRA5'-5 and the presence of major fragments binding to pSRA5'-2 indicate that ribosomes were not binding to AUG(*gag*) under conditions where peptide elongation was inhibited at the beginning of translation. We translated RSV RNA without any inhibitors to demonstrate that our translational system primarily directs the synthesis of its known viral product, Pr76^{gag} (Fig. 5), as has been shown by many others (e.g., see reference 49). We compared synthesis of Pr76^{gag} as a function of time from untreated and methyl mercuric hydroxide-treated RSV RNA. The data in Fig. 5 demonstrate that under the conditions used in the ribosome protection experiments, our translational system produces Pr76^{gag} from both templates and that prior denaturation of the RSV RNA influences neither the timing of the onset of initiation nor the final products. The smaller peptides seen were most probably due to premature termination of translation (13).

DISCUSSION

In general, mRNAs that do not use the first AUG codon for initiation of translation, at least for the "major" protein product, have long leader RNAs preceding the initiator AUG but no uniquely defined ribosome-binding sites (9, 35). In contrast, most cellular mRNAs have short leaders of ca. 50 bases, use the 5'-proximal AUG codon as an initiation site, and have well-defined ribosome-binding sites (24). The lack of definitive ribosome-binding sites on many viral mRNAs, which have very long leader sequences, could be due either to variability in the positioning of the ribosome during the stalled initiation or to a high background of fragments coming from loci that, although protected by ribosomes, are not sites of initiation of protein synthesis. There are instances where two ribosomes or ribosomal subunits have

FIG. 5. Translation of untreated and methyl mercuric hydroxide-treated RSV RNA for various periods. Lanes 1 through 6, Methyl mercuric hydroxide-denatured RSV RNA translated for 2, 5, 10, 15, 20, and 30 min. Lanes 7 through 12, Nondenatured RSV RNA translated for the same periods as shown for lanes 1 through 6. Lane o, Products of the translational system without exogenously added mRNA after a 30-min incubation. The arrow marks the position of Pr76^{gag}.

protected regions of mRNA that do not have central AUG codons (1, 10, 23). This type of undefined binding may be due to a stacking up of ribosomes behind a preceding ribosome that is stalled at an AUG codon. In several detailed studies (26, 42) ribosome-binding sites that contain central AUG codons have been shown to be true sites of initiation of translation, whereas abortive sites do not have central AUG codons (see reference 23).

The ribosome-protected sequences that we have found on RSV RNA have a central AUG codon. This result extends the earlier findings that show there is a ribosome-binding site within the first 100 bases of the RSV RNA (5, 40), and it strongly suggests that AUG(1) is the site of productive ribosome binding for the synthesis of a seven-amino-acid peptide. However, the 80S ribosome-protected fragments we observe are 50 to 60 nucleotides in size, about twice the size found with most mRNAs (21). The 5' ends of the five major fragments begin at nucleotide 28, 13 to 15 bases upstream from AUG(1), but have various 3' termini. The 3' heterogeneity may be due to the secondary structure that RSV RNA can assume (see reference 6), which can inhibit RNase T1 digestion. By this interpretation, bases 28 through ca. 63 would normally be protected by the stalled ribosome, and the following 9 through 24 bases, from the protected fragments of 45 through 60 bases, would be protected by the secondary structure. The 57- and 60-base ribosome-protected fragments are probably not due to ribosomal subunits that have stacked up behind AUG(2), since these fragments barely contain this codon. Therefore, we conclude that AUG(1) is a normal site of initiation of translation. Our interpretation is based on (i) the known mode of action of anisomycin, blocking elongation by the fully initiated 80S ribosome and (ii) the appropriate distance of the AUG(1) codon from the 5' ends of the protected fragments. The data in Fig. 3 suggest that no more than 5 to 10% of the ribosomes, if any, are able to circumvent AUG(1) and bind to downstream AUG codons such as AUG(*gag*).

When ribosomes are not inhibited by anisomycin, they can reach AUG(*gag*) to initiate translation of the polypeptide Pr76^{gag} (Fig. 5). Pr76^{gag} should be the only protein product encoded by RSV RNA if the model proposed by Jacobson and Baltimore (19) is correct. The data we present above indicate that the seven-amino-acid peptide should be made in vitro since anisomycin blocks elongation of completely formed ribosomes that are actively engaged in protein synthesis. The leader peptide may be made in vivo as well (45). Initial analyses of small peptides synthesized in vitro indicate the presence of a product that comigrates on a high-pressure liquid chromatography C₁₈ reverse-phase column with a synthetic seven-amino-acid peptide marker identical to the leader peptide encoded behind AUG(1) (unpublished data).

Accordingly, the question arises whether all of the AUG codons in the leader RNA region can initiate protein synthesis. After splicing to produce mRNA^{sc}, AUG(*gag*) is no longer needed to initiate protein synthesis; in fact, if it were used, synthesis of pp60^{sc} would be nil according to the rules of the modified scanning hypothesis (22), since the splice on mRNA^{sc} places a translational termination codon 24 bases downstream from AUG(*gag*). Hughes et al. (S. Hughes, K. Mellstrom, E. Kosik, F. Tamanoi, and J. Brugge, Mol. Cell. Biol., in press) recently showed that AUG(*gag*) can be "activated" on mRNA^{sc} by specifically mutating the termination codon to allow readthrough into the sequences encoding pp60^{sc}. These results suggest that AUG(*gag*) may not be silent on mRNA^{sc} and that ribosomes may continue past

short open reading frames and initiate translation more than once on an mRNA, as suggested by our analysis of ribosome recognition of AUG(1) as a site of initiation of translation.

Three factors may bear on the selection by ribosomes of AUG codons for the initiation of translation: (i) proximity to the 5' end of the mRNA, (ii) the sequences that flank the AUG codon, and (iii) the secondary structure of the mRNA. Proximity of an AUG codon to the 5' cap clearly is important (21–23). The sequences flanking AUG codons are also important (22). Although the sequences flanking AUG codons 2 and 3 are suboptimal for their use as sites of initiation of translation, the sequence flanking AUG(1) is recognized as an initiation site on some mRNAs (24). The data shown in Fig. 3 do not support the proposition that AUG(1) is a leaky initiation site which allows some ribosomes to pass through to downstream initiation sites. Moreover, the data do not support the model of Darlix et al. (6), which implies that ribosomes can bypass AUG(1) and bind directly to AUG(*gag*). The secondary structure of mRNA is important in the regulation of translation in procaryotes (11, 17), and it may influence the efficiency of translational initiation at particular AUG codons in eucaryotic mRNAs (30, 32, 36, 48). Our results testing ribosome binding to and translation of methyl mercury-denatured and nondenatured RSV RNA (Fig. 2 and 5) do not unequivocally address this possibility, since renaturation of the RSV leader may occur before ribosome binding.

In conclusion, these data suggest that the scanning model for initiation of protein synthesis correctly predicts that ribosomes will initiate translation at the 5'-proximal AUG codon of RSV RNA. Consequently, it appears that RSV RNA encodes at least two protein products, a small leader peptide that is initiated at AUG(1) and the major product, Pr76^{gag}, that is initiated at AUG(*gag*). The question of how ribosomes identify AUG(*gag*) as an initiation codon remains unanswered. These results raise the possibility either that initiation at AUG(*gag*) requires the migration of a ribosome along the entire RSV RNA leader sequence, with initiations at AUG codons 1, 2, and 3, or that the ribosome can bypass one or more of the open reading frames between AUG(1) and AUG(*gag*).

ACKNOWLEDGMENTS

We thank Barbara Triplett for the generous gift of alkaline hydrolyzed polyadenylate RNA and Bill DeLorbe for plasmid p165. Walter Sauerbier and Barb Triplett provided helpful suggestions and discussions.

This work was supported by National Science Foundation grant PCM80-10975, an award from the University of Minnesota graduate school, and an American Cancer Society Institutional Research Grant (IN-13-V-59 and IN-13-V-47) to the University of Minnesota (R.B.P. and C.H.H.). R.B.P. and C.H.H. are supported by predoctoral training grant GM 07094 from the National Institutes of Health.

LITERATURE CITED

1. Ahlquist, P., R. Dasgupta, D. S. Shih, D. Zimmermann, and P. Kaesberg. 1979. Two-step binding of eukaryotic ribosomes to brome mosaic virus RNA3. *Nature (London)* **281**:277–282.
2. Bailey, J. M., and N. Davidson. 1976. Methylmercury as a reversible denaturing agent for agarose gel electrophoresis. *Anal. Biochem.* **70**:75–85.
3. Bos, J. L., L. J. Polder, R. Bernards, P. I. Schrier, P. J. van den Elsen, A. J. van der Eb, and H. van Ormondt. 1981. The 2.2-kb E1b mRNA of human Ad12 and Ad5 codes for two tumor antigens starting at different AUG triplets. *Cell* **27**:121–131.
4. Brandt, T. L., and P. B. Hackett. 1983. Characterization of messenger RNA by direct translation of mRNA from agarose

- gels. *Anal Biochem.* **135**:401–408.
5. Darlix, J.-L., P.-F. Spahr, P. A. Bromley, and J.-C. Jaton. 1979. In vitro, the major ribosome binding site on Rous sarcoma virus RNA does not contain the nucleotide sequence coding for the N-terminal amino acids of the *gag* gene product. *J. Virol.* **29**:597–611.
 6. Darlix, J.-L., M. Zuker, and P.-F. Spahr. 1982. Structure-function relationship of Rous sarcoma virus leader RNA. *Nucleic Acids Res.* **10**:5183–5196.
 7. DeLorbe, W. J., P. A. Luciw, H. M. Goodman, H. E. Varmus, and J. M. Bishop. 1980. Molecular cloning and characterization of avian sarcoma virus circular DNA molecules. *J. Virol.* **36**:50–61.
 8. Donis-Keller, H. 1979. Site specific enzymatic cleavage of RNA. *Nucleic Acids Res.* **7**:179–192.
 9. Dörner, A. J., L. F. Dörner, G. R. Larsen, E. Wimmer, and C. W. Anderson. 1982. Identification of the initiation site of poliovirus polypeptide synthesis. *J. Virol.* **42**:1017–1028.
 10. Filipowicz, W., and A.-L. Haenni. 1979. Binding of ribosomes to 5'-terminal leader sequences of eukaryotic messenger RNAs. *Proc. Natl. Acad. Sci. U.S.A.* **76**:3111–3115.
 11. Gold, L., D. Pribnow, T. Schneider, S. Shinedling, B. S. Singer, and G. Stormo. 1981. Translational initiation in prokaryotes. *Annu. Rev. Microbiol.* **35**:365–400.
 12. Goldberg, M. L., R. P. Lifton, G. R. Stark, and J. G. Williams. 1979. Isolation of specific RNAs using DNA covalently linked to diazobenzoyloxymethyl cellulose or paper. *Methods Enzymol.* **68**:206–220.
 13. Hackett, P. B., E. Egberts, and P. Traub. 1979. Translation of ascites and mengovirus RNA in fractionated cell-free systems from uninfected and mengovirus-infected Ehrlich ascites tumor cells. *Eur. J. Biochem.* **83**:341–352.
 14. Hackett, P. B., R. Swanstrom, H. E. Varmus, and J. M. Bishop. 1982. The leader sequence of the subgenomic mRNAs of Rous sarcoma virus is approximately 390 nucleotides. *J. Virol.* **41**:527–534.
 15. Hall, L., R. K. Craig, M. R. Edbrooke, and P. N. Campbell. 1982. Comparison of the nucleotide sequence of cloned human and guinea pig pre- α -lactalbumin cDNA with that of chick prelysozyme cDNA suggests evolution from a common ancestral gene. *Nucleic Acids Res.* **10**:3503–3515.
 16. Hendy, G. N., H. M. Kronenberg, J. T. Potts, Jr., and A. Rich. 1981. Nucleotide sequence of cloned cDNAs encoding human parathyroid hormone. *Proc. Natl. Acad. Sci. U.S.A.* **78**:7365–7369.
 17. Hercules, K., M. Schweiger, and W. Sauerbier. 1974. Cleavage by RNase III converts T3 and T7 early precursor RNA into translatable message. *Proc. Natl. Acad. Sci. U.S.A.* **71**:840–844.
 18. Ish-Horowicz, D., and J. F. Burke. 1981. Rapid and efficient cosmid cloning. *Nucleic Acids Res.* **9**:2989–2998.
 19. Jacobson, M. F., and D. Baltimore. 1968. Polypeptide cleavages in the formation of poliovirus proteins. *Proc. Natl. Acad. Sci. U.S.A.* **61**:77–85.
 20. Jay, G., S. Nomura, C. W. Anderson, and G. Khoury. 1981. Identification of the SV40 agnogene product: a DNA binding protein. *Nature (London)* **291**:346–349.
 21. Kozak, M. 1981. Mechanism of mRNA recognition by eukaryotic ribosomes during initiation of protein synthesis. *Curr. Top. Microbiol. Immunol.* **93**:81–123.
 22. Kozak, M. 1981. Possible role of flanking nucleotides in recognition of the AUG initiator codon by eukaryotic ribosomes. *Nucleic Acids Res.* **9**:5233–5252.
 23. Kozak, M. 1982. Analysis of ribosome binding sites from the s1 message of reovirus: initiation at the first and second AUG codons. *J. Mol. Biol.* **156**:807–820.
 24. Kozak, M. 1984. Compilation and analysis of sequences upstream from the translation start site in eukaryotic mRNAs. *Nucleic Acids Res.* **12**:857–872.
 25. Laemmli, U. K. 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature (London)* **227**:680–685.
 26. Legon, S. 1979. The binding of ribosomes to polyoma virus RNA. Possible role of the leader region in initiation site recognition. *J. Mol. Biol.* **134**:219–240.
 27. Lodish, H. F. 1971. Alpha and beta globin messenger ribonucleic acid: different amounts and rates of initiation of translation. *J. Biol. Chem.* **246**:7131–7138.
 28. Lomedico, P., and S. J. McAndrew. 1982. Eukaryotic ribosomes can recognize preproinsulin initiation codons irrespective of their positions relative to the 5' end of mRNA. *Nature (London)* **299**:221–226.
 29. Mardon, G., and H. E. Varmus. 1983. Frameshift and intragenic suppressor mutations in a Rous sarcoma provirus suggest *src* encodes two proteins. *Cell* **32**:871–879.
 30. Nomoto, A., T. Omata, H. Toyoda, S. Kuge, H. Horie, Y. Kataoka, Y. Genba, Y. Nakano, and N. Imura. 1982. Complete nucleotide sequence of the attenuated Sabin I strain genome. *Proc. Natl. Acad. Sci. U.S.A.* **79**:5793–5797.
 31. Palmiter, R. D., J. Gagon, V. M. Vogt, S. Ribley, and R. N. Eisenman. 1978. The NH₂-terminal sequence of the avian oncornavirus *gag* precursor polypeptide (Pr76^{gag}). *Virology* **91**:423–433.
 32. Pavlakis, G. N., R. E. Lockard, N. Vamvakopoulos, L. Rieser, U. L. RajBhandary, and J. N. Vournakis. 1980. Secondary structure of mouse and rabbit α - and β -globin mRNAs: differential accessibility of α and β initiator AUG codons towards nucleases. *Cell* **19**:91–102.
 33. Payvar, F., and R. T. Schimke. 1979. Methylmercury hydroxide enhancement of translation and transcription of ovalbumin and conalbumin mRNA's. *J. Biol. Chem.* **254**:7636–7642.
 34. Pelham, H. R. B., and R. J. Jackson. 1976. An efficient mRNA-dependent translational system from reticulocyte lysate. *Eur. J. Biochem.* **67**:247–256.
 35. Ravelonandro, M., T. Godefroy-Colburn, and L. Pinck. 1983. Structure of the 5'-terminal untranslated region of the genomic RNAs from two strains of alfalfa mosaic virus. *Nucleic Acids Res.* **11**:2815–2826.
 36. Ray, B. K., T. G. Brendler, S. Adya, S. Daniels-McQueen, J. K. Miller, J. W. B. Hershey, J. A. Grifo, W. C. Merrick, and R. E. Thach. 1983. Role of mRNA competition in regulating translation: further characterization of mRNA discriminatory initiation factors. *Proc. Natl. Acad. Sci. U.S.A.* **80**:663–667.
 37. Rose, J. K. 1978. Complete sequences of the ribosome recognition sites in vesicular stomatitis virus mRNAs: recognition by the 40S and 80S complexes. *Cell* **14**:345–353.
 38. Schroeder, H. W., C. D. Liarakos, R. C. Gupta, K. Randerath, and B. W. O'Malley. 1979. Ribosome binding site analysis of ovalbumin messenger ribonucleic acid. *Biochemistry* **18**:5798–5808.
 39. Sherman, F., G. McKnight, and J. W. Stewart. 1980. AUG is the only initiation codon in eukaryotes. *Biochim. Biophys. Acta* **609**:343–346.
 40. Shine, J., A. P. Czernilofsky, R. Friedrich, J. M. Bishop, and H. M. Goodman. 1977. Nucleotide sequence at the 5' terminus of the avian sarcoma virus genome. *Proc. Natl. Acad. Sci. U.S.A.* **74**:1473–1477.
 41. Steitz, J. A. 1969. Peptide chain initiation: nucleotide sequences of the three ribosomal binding sites in bacteriophage R17 RNA. *Nature (London)* **224**:957–964.
 42. Steitz, J. A. 1973. Specific recognition of non-initiator regions in RNA bacteriophage messengers by ribosomes of *Bacillus stearothermophilus*. *J. Mol. Biol.* **73**:1–16.
 43. Swanstrom, R., R. C. Parker, H. E. Varmus, and J. M. Bishop. 1983. Transduction of a cellular oncogene: the genesis of Rous sarcoma virus. *Proc. Natl. Acad. Sci. U.S.A.* **80**:2519–2523.
 44. Swanstrom, R., H. E. Varmus, and J. M. Bishop. 1982. Nucleotide sequence of the 5' noncoding region and part of the *gag* gene of Rous sarcoma virus. *J. Virol.* **41**:535–541.
 45. Thummel, C., R. Tjian, S.-L. Hu, and T. Grodzicker. 1983. Translational control of SV40 T antigen expressed from the adenovirus late promoter. *Cell* **33**:455–464.
 46. Varmus, H. E., and R. Swanstrom. 1982. Replication of retroviruses, p. 369–512. In R. Weiss, N. Teich, H. Varmus, and J. Coffin (ed.), *RNA tumor viruses*. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.

47. **Vieira, J., and J. Messing.** 1982. The pUC plasmids, an M13mp7-derived system for insertion mutagenesis and sequencing with synthetic universal primers. *Gene* **19**:259–268.
48. **Vogeli, G., H. Ohkubo, M. E. Sobel, Y. Yamada, I. Pastan, and B. deCrombrughe.** 1981. Structure of the promoter for chicken $\alpha 2$ type I collagen gene. *Proc. Natl. Acad. Sci. U.S.A.* **78**:5334–5338.
49. **Weiss, S. R., P. B. Hackett, H. Oppermann, A. Ullrich, L. Levintow, and J. M. Bishop.** 1978. Cell-free translation of avian sarcoma virus RNA: suppression of the *gag* termination does not augment synthesis of the joint *gag/pol* product. *Cell* **15**:607–614.