5′-Conformation of Capped Alfalfa Mosaic Virus Ribonucleic Acid 4 May Reflect Its Independence of the Cap Structure or of Cap-Binding Protein for Efficient Translation†

Lee Gehrke,* Philip E. Auron, Gary J. Quigley, Alexander Rich, and Nahum Sonenberg

ABSTRACT: Most eukaryotic mRNAs are characterized by the presence of a 5′-terminal cap structure (m^7GpppN), and removal of the cap or translation of capped mRNAs in the presence of cap analogues (m^7G) results in most cases in a significant decrease in the translational efficiency of the mRNAs. One way of explaining the importance of the 5′-cap is that cap-binding proteins recognize the cap structure, destabilize the mRNA secondary structure, and thus allow the 40S ribosomal subunit to bind to the mRNA [Sonenberg, N., Guertin, D., Cleveland, D., & Trachsel, H. (1981) *Cell (Cambridge, Mass.) 27*, 563–572]. Our data and those of others indicate that the translational efficiency of alfalfa mosaic virus RNA 4 (AMV-4 RNA), a naturally capped RNA, is not affected significantly by cap analogues or by removal of the cap. In order to examine the potential relationship between the function of the cap structure and secondary structure at the 5′-mRNA terminus, partial enzymatic digestion of capped AMV-4 RNA with single strand specific and double strand specific nucleases has been performed, and the experimental data have been compared with computer-generated models of AMV-4 secondary structure. In addition, the in vitro translatability of AMV-4 has been examined as a function of increasing potassium concentration, conditions that are likely to increase mRNA secondary structure. The nuclease-digestion results demonstrate that under native ionic conditions, the 5′-terminus of AMV-4 RNA is predominantly single stranded, although computer modeling and double-strand nuclease digestions indicate that the 5′-terminus can form weak base pairs with internal regions of the molecule. Binding of ribosomes to AMV-4 in vitro is relatively resistant to increased potassium concentration, again indicating that the 5′-terminus of AMV-4 RNA does not form stable base pairs. The data are consistent with a model wherein those mRNAs with little potential for formation of stable base pairs in the 5′-terminal region preceding the initiation codon have a decreased dependence upon the cap structure and cap-binding protein for efficient translation.

Initiation of eukaryotic protein synthesis is a complex phenomenon, and the initiation event is a probable site of significant regulation of eukaryotic gene expression (Lodish, 1976). Among the salient features of initiation of protein synthesis in eukaryotes are the requirement for ATP hydrolysis and the presence of a cap structure (m^7GpppN) at th 5′-end of most mRNAs[1] (Grunberg-Manago, 1980; Shatkin, 1976).

Numerous studies have indicated that the 5′-terminal cap structure facilitates initiation of protein synthesis in eukaryotes [for reviews, see Banerjee (1980) and Shatkin (1976)]. In addition, cap binding proteins that interact specifically with the cap structure have been identified and purified from rabbit reticulocytes (Sonenberg et al., 1978, 1979; Sonenberg, 1981; Tahara et al., 1981). However, it has not yet been established how the cap structure functions or how it might regulate the initiation step of protein synthesis.

It has been demonstrated that the binding of ribosomes to irreversibly denatured, inosine-substituted reovirus mRNA is less dependent on both the cap structure and ATP than the binding to native reovirus mRNA (Kozak, 1980a,b; Morgan & Shatkin, 1980). Evidence has been obtained for the existence of a protein factor in initiation-factor preparations that may be responsible for an RNA secondary-structure melting activity (Ilan & Ilan, 1977), and more recent findings suggested the involvement of a cap-recognition element in an ATP-dependent unfolding of mRNA secondary structure (Sonenberg et al., 1981; Lee et al., 1983). Furthermore, there is evidence that cowpea mosaic virus RNA, which is naturally uncapped, is less dependent on ATP for initiation-complex formation as compared to capped mRNAs (Jackson, 1982). These data suggest a model in which the cap structure might mediate the destabilization of mRNA secondary structure in order to facilitate ribosome binding.

In light of these observations, it is significant that the translation of alfalfa mosaic virus RNA 4 (AMV-4), a naturally capped mRNA, appears to be almost independent of the cap structure for translation (Van Vloten-Doting et al., 1977; Sonenberg et al., 1981). Alfalfa mosaic virus (AMV) is a simple aphid-transmitted RNA-containing plant virus that is characterized by a tripartite genome consisting of 2.5×10^6 dalton of RNA (Van Vloten-Doting & Jaspers, 1977). The AMV genome consists of three primary subgenomic RNAs and, in addition, RNA 4 (AMV-4), which is a subfragment of RNA 3. AMV-4 codes for the coat protein, and although

† From the Harvard–Massachusetts Institute of Technology Division of Health Sciences and Technology and the Department of Anatomy, Harvard Medical School, Cambridge, Massachusetts 02139 (L.G.), the Department of Biology, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139 (P.E.A., G.J.Q., and A.R.), and the Department of Biochemistry and McGill Cancer Center, McGill University, Montreal, Quebec, Canada H3G 1Y6 (N.S.). *Received April 25, 1983.* This research was supported by grants from the National Institutes of Health and the American Cancer Society to A.R. and by grants from the National Cancer Institute and the Medical Research Council of Canada to N.S. N.S. is the recipient of a Terry Fox Cancer Research Scientist Award of the National Cancer Institute of Canada. L.G. was supported in part by a postdoctoral fellowship from the National Institutes of Health.

* Address correspondence to this author at the Division of Health Sciences and Technology, Massachusetts Institute of Technology, Cambridge, MA 02139.

[1] Abbreviations: AMV-4, alfalfa mosaic virus RNA 4; mRNA, messenger RNA; Hepes, *N*-(2-hydroxyethyl)piperazine-*N*′-2-ethanesulfonic acid; Tris, tris(hydroxymethyl)aminomethane; AdoMet, *S*-adenosylmethionine; EDTA, ethylenediaminetetraacetic acid; DTT, dithiothreitol; SDS, sodium dodecyl sulfate; NMR, nuclear magnetic resonance; mRNP, messenger ribonucleoprotein particle.

0006-2960/83/0422-5157$01.50/0

it is not replicated, AMV-4 or its translated protein product is required for infectivity (Bol & Van Vloten-Doting, 1973). The 5′-ends of all AMV RNAs are capped by $m^7GpppGp$ (Pinck, 1975), and Van Vloten-Doting et al. (1977) reported that AMV-4 RNA translates more efficiently in vitro than the other subgenomic RNAs. In addition, other translational properties of AMV-4 correlate with a non-base-paired 5′-end structure. For example, translation of AMV-4 RNA in wheat germ extracts is relatively resistant to inhibition by high salt concentrations as compared to other eukaryotic mRNAs (Herson et al., 1979). Inhibition of ribosome binding by high salt concentrations could be due to the more stable secondary structure that can be formed under these conditions (Sonenberg et al., 1982). Furthermore, AMV-4 RNA is translatable in extracts prepared from polio virus infected HeLa cells in which translation of other capped mRNAs is severely restricted (Sonenberg et al., 1982), probably due to inactivation of a cap-binding protein (Etchison et al., 1982).

In an attempt to correlate the 5′-proximal mRNA secondary structure of AMV-4 with translation-initiation parameters, we analyzed initiation-complex formation in wheat germ extracts with respect to inhibition by cap analogues and sensitivity to high ionic strength. To determine mRNA secondary structure, capped AMV-4 RNA has been analyzed by partial enzymatic digestion under native conditions (Wurst et al., 1978; Wrede et al., 1979). The experimental data indicate that the 5′-nontranslated region of AMV-4 RNA is single stranded under native experimental conditions, and computer analyses indicate that the 5′-terminus has little potential for forming stable base pairs with other regions of the molecule. The combination of enzymatic structure-mapping data and computer analysis of the AMV-4 RNA sequence provide direct evidence consistent with an interpretation in which the absence of 5′-secondary structure in this mRNA reduces its requirement for the cap structure and cap-binding protein.

Experimental Procedures

Enzymes and Reagents. Ribonuclease T1, *Physarum* M (Phy M), and nuclease mung bean 1 (MB1) were obtained from P-L Biochemicals. Calf intestinal alkaline phosphatase, S1 nuclease, and T4 polynucleotide kinase were purchased from Boehringer Mannheim. Tobacco acid pyrophosphatase and the vaccinia guanylyltransferase preparation containing methyltransferase activity were obtained from Bethesda Research Laboratories. Ribonuclease V1 was purified from cobra venom as previously described (Auron et al., 1982). Sodium periodate and reagent-grade aniline were purchased from Fisher Scientific. P1 nuclease was obtained from Sankyo Biochemicals.

Ribosome Binding. [^{3}H]Methyl-labeled reovirus or AMV-4 RNA in the amount specified was incubated for 10 min at 25 °C in 50 μL of an S23 wheat germ extract containing 20 mM Hepes (pH 7.6), 10 μM each of 20 amino acids, 2 mM dithiothreitol, 1 mM ATP, 0.2 mM GTP, 5 mM creatine phosphate, 4 μg of creatine phosphokinase, 3 mM $Mg(OAc)_2$, 2 mM sparsomycin to inhibit protein synthesis, and 75 mM KCl unless otherwise indicated. Initiation-complex formation was analyzed in glycerol gradients as described by Sonenberg & Shatkin (1977). Centrifugation was for 90 min at 48 000 rpm and 4 °C in an SW 50.1 rotor.

Preparation of [^{3}H]Methyl-Labeled mRNA. Reovirus mRNA ^{3}H-labeled in both methyl groups of the cap structure ($m^7GpppGm$) was synthesized with viral cores in the presence of [^{3}H]methyl-AdoMet (sp act. 70 Ci/mmol; New England Nuclear) as described by Muthukrishnan et al. (1976) to a specific activity of ~40 000 cpm/μg. [^{3}H]Methyl-labeled bromouridine-substituted reovirus mRNA was prepared according to Kozak (1980a) to a specific activity of ~15 000 cpm/μg. AMV-4 RNA, ^{3}H-labeled in the 2′-*O*-methyl group of the 5′-penultimate nucleotide, was prepared according to Moss (1977) by incubating AMV-4 RNA with [^{3}H]methyl-AdoMet in the presence of methyltransferase purified from vaccinia virus according to Ensinger et al. (1975). The specific activity of the [^{3}H]methyl-labeled AMV-4 RNA obtained was 12 000 cpm/μg.

Chemical Removal of Cap and Replacement with Vaccinia Guanylyltransferase. Alfalfa mosaic virus RNA 4 (AMV-4) was a generous gift from J. Bol (State University of Leiden, Leiden, The Netherlands). The RNA was radioactively labeled by chemical removal of the inverted m^7G (Fraenkel-Conrat & Steinschneider, 1967), followed by enzymatic replacement with [α-^{32}P]GTP in the presence of vaccinia guanylyltransferase (Plotch et al., 1979). Thirty micrograms of AMV-4 RNA was solubilized in a solution containing 100 mM sodium acetate, pH 5.3, 1 mM EDTA, and a final concentration of 0.2 mM fresh sodium periodate. The solution was incubated at room temperature for 1 h, and then the oxidized RNA was ethanol precipitated, washed extensively with 70% ethanol, and dried. The AMV-4 RNA was solubilized in 25 μL of a solution containing 0.2 M sodium acetate, pH 5.3, and 2 mM EDTA. The aniline solution was prepared by combining 4.6 mL of water, 0.27 mL of freshly distilled aniline, and 10 μL of concentrated HCl. Twenty-five microliters of the aniline solution was added to the solubilized AMV-4 RNA, and the solution was incubated at room temperature for 3 h. Following the ethanol precipitation and the washing, the decapped AMV-4 RNA was radioactively labeled. Five hundred microcuries of [α-^{32}P]GTP (sp act. >3000 Ci/mmol; New England Nuclear) was lyophilized to dryness and solubilized in a solution containing (final concentrations) 50 mM Tris-HCl, pH 8.3, 1.3 mM $MgCl_2$, 2.5 mM DTT, 20 μM AdoMet, 20 μM nonradioactive GTP, and 3 units of vaccinia guanylyltransferase (Bethesda Research Laboratories). The solubilized GTP was transferred to the tube containing decapped AMV-4 RNA, and the mix was incubated for 30 min at 37 °C. Following the incubation, 50 μL of cold water was added, and the solution was passed through a 1-mL column of Sephadex G-50 to remove unincorporated label. Cap-labeled AMV-4 RNA was then ethanol precipitated, sedimented, and washed with 70% ethanol.

Purification of Labeled AMV-4 RNA by Preparative Gel Electrophoresis. Full-length radioactively labeled AMV-4 RNA was isolated by preparative gel electrophoresis (Bethesda Research Laboratories Model 1100PG). A 3.6% polyacrylamide gel (1-cm diameter × 3.5-cm length) was cast in 35 mM NaH_2PO_4, 36 mM Tris, and 2 mM EDTA, pH 7.4–7.6. Electrode buffer was identical with the exception that SDS was added to 0.2%. Labeled AMV-4 RNA was solubilized in 50 μL of sample buffer (35 mM NaH_2PO_4, 36 mM Tris base, 2 mM EDTA, 10% glycerol) and heated at 55 °C for 5 min. The solution was quick cooled on ice and layered onto the gel. Electrophoresis was at 10-mA constant current, and the flow rate of the elution buffer was 5 mL/h. One-milliliter fractions were collected, and full-length AMV-4 RNA eluted after approximately 6 h. Fractions were pooled, ethanol precipitated, sedimented, and washed several times with 70% ethanol. The position of the radioactive label was verified by hydrolyzing the RNA with tobacco acid pyrophosphatase, calf intestinal alkaline phosphatase, and P1 nuclease, followed by analysis of the digestion products by thin-layer chromatography (Gupta & Randerath, 1979).

Partial Digestion of AMV-4 RNA. AMV-4 RNA labeled with ^{32}P in the 5′-m^{7}G cap (6000–7000 cpm per reaction) was used for all reactions. For the T1 and Phy M digestions, 0.1 A_{260} unit of RNA was utilized, while the native S1, MB1 and V1 reactions contained 0.05 A_{260} unit of RNA. Yeast tRNA^{Phe} (Boehringer) was used to adjust RNA concentrations. An equal volume of gel-sample preparation buffer (7.9 M urea, 0.05% xylene cyanol FF/bromophenol blue, 10% glycerol) was added to all native digestions following termination of the reaction. Digestions done under denaturing conditions contained urea and dye and were loaded directly onto polyacrylamide gels.

Native digestions with S1 nuclease were done in 2.5 μL of 25 mM NaOAc, pH 4.5, 5 mM $MgCl_2$ and 50 mM KCl (S1 buffer). A 0.5-μL aliquot of S1 nuclease solution [5 units/μL in 20 mM NaOAc, pH 4.5–1 mM $Zn(OAc)_2$] was added to the solution, which was then incubated at 37 °C for 1–2 min. The reactions were terminated by rapid freezing in dry ice.

Nuclease S1 digestions under denaturing conditions utilized 0.1 A_{260} unit of mRNA solubilized in 2.5 mM NaOAc, pH 4.5, 10 mM EDTA, and 8 M urea. The reaction was incubated at 60 °C for 30 min. Formamide "ladders" were generated by solubilizing a similar quantity of labeled AMV-4 RNA in 2.5 μL of deionized formamide; the solution was then placed in a 90 °C water bath for 15 min.

Digestions employing ribonuclease T1 and Phy M under denaturing conditions were carried out in 2.5 μL of a solution containing 20 mM sodium citrate, pH 5.0, 1 mM EDTA, 0.04% xylene cyanol FF/bromophenol blue, 0.75 mg/mL yeast tRNA^{Phe}, and 7.9 M urea. A total of 0.5–1 μL of enzyme (1 unit/μL in water) was added, and the reaction mixtures were incubated at 55 °C for 10 min.

Digestion of 5′-^{32}P-labeled AMV-4 RNA with ribonuclease V1 was done in 2.5 μL of native buffer [50 mM sodium cacodylate, pH 6.7–5 mM $MgCl_2$ to which was added 1–2 μL of the pooled M_r 27 000 fractions described previously (Auron et al., 1982)]. The incubation temperatures are indicated in the figure legends, and the reactions were terminated by rapid freezing in dry ice.

Ribonuclease MB1 reactions were prepared by dissolving 5′-^{32}P-labeled AMV-4 RNA in 2.5 μL of native buffer. One microliter of a 1:100 dilution (in native buffer) of commercial enzyme was added to each reaction. Incubation was at 37 °C for 0.5–5 min and was terminated by rapid freezing.

Computer-Generated Dot Matrix Analysis. The matrices shown in Figure 5 are similar in principal to those described by Maizel & Lenk (1981) and Pustell & Kafatos (1982). The dots constituting the diagonals in the matrix represent potential base pairs, including G-U pairs, which are part of a contiguous string of some preset minimum length, and the relative free energies of which are less than a preselected stacking/pairing energy value. The stacking/pairing energy is a simple sum of the relative free-energy contributions of the individual base pairs in which G-C pairs contribute −3 units, A-T pairs −2 units, and G-U pairs 0 units. These relative energy values roughly parallel the average energies expected from stem energy calculations according to the methodology described by Tinoco et al. (1973). The use of the length and energy filtering provides an increased contrast relative to that seen if all potential base pairs were displayed. A more detailed description of the matrix program will be presented elsewhere.

Results

Alfalfa mosaic virus RNA 4 was radioactively labeled by introducing a [^3H]methyl group into the 2′-ribose position of the 5′-penultimate base with the vaccinia virus coded methyltransferase. Labeled AMV-4 RNA was used in initiation-complex formation assays to determine the cap dependence of translation initiation in wheat germ extracts. Figure 1A shows that 73% of input AMV-4 RNA added to a wheat germ extract was bound to 80S ribosomes. When the cap analogue m^{7}GDP at a concentration of 0.8 mM was included in the mixture, binding was reduced to 52% of that of the control. This partial inhibition of AMV-4 RNA binding by m^{7}GDP is in contrast to the drastic effect exerted by m^{7}GDP on reovirus mRNA ribosome binding, which is reduced to 16% of that of the control (Figure 1B). The relative cap-independent binding of AMV-4 RNA to ribosomes might be due to the lack of stable secondary structure at the 5′-region of the mRNA, on the basis of the finding that ribosome binding of inosine-substituted reovirus mRNA, which has a relaxed secondary structure, is less inhibited by m^{7}GDP as compared to native reovirus mRNA (Morgan & Shatkin, 1980). This finding led to the suggestion that the cap structure might mediate the unfolding of mRNA secondary structure to facilitate ribosome binding.

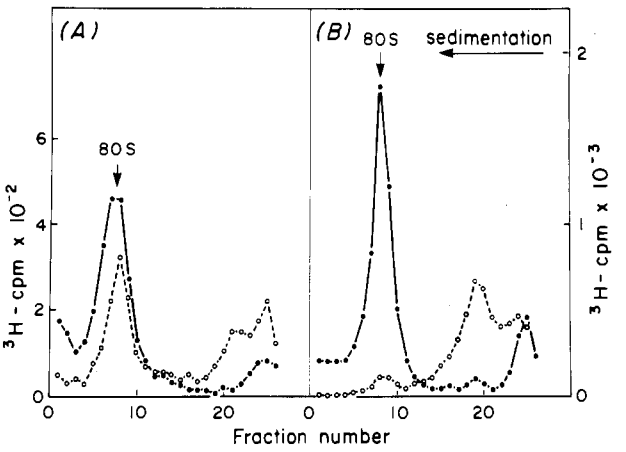

FIGURE 1: Effect of m^{7}GDP on ribosome binding of reovirus and AMV-4 RNAs. (A) AMV-4 RNA (4900 cpm, ~0.4 μg) or (B) reovirus mRNA (20 000 cpm, ~0.4 μg) was incubated in 50 μL of a wheat germ extract at 25 °C for 10 min in the absence (●) or presence (○) of 0.8 mM m^{7}GDP. Initiation complexes formed were analyzed on glycerol gradients as described under Experimental Procedures.

Studies by Herson et al. (1979) have shown that translation of AMV-4 RNA in wheat germ extracts is resistant to inhibition by high salt concentrations as compared to rabbit globin and several other mRNAs. Therefore, it was of importance to determine whether this insensitivity to inhibition by high salt concentration is exhibited at the level of initiation of protein synthesis. Figure 2 shows the ribosome binding profiles of AMV-4 RNA in wheat germ extracts as a function of increasing K^+ concentrations. Binding was only slightly inhibited when the K^+ concentration was increased from 45 (panel A) to 125 mM (panel C; 79% and 72% of input mRNA was bound at the two concentrations, respectively). However, in sharp contrast to the resistance of AMV-4 ribosome binding to inhibition by high salt concentrations, binding of bromouridine-substituted reovirus mRNA, which possesses enhanced secondary structure as compared to native reovirus mRNA (Kozak, 1980a), was inhibited by 80% when the K^+ concentration was increased from 45 (panel D) to 125 mM (panel F), in agreement with previous results (Lee et al., 1983). Thus, the binding profile of AMV-4 to ribosomes at different K^+ concentrations resembles that of inosine-substituted reovirus mRNA (Sonenberg et al., 1982). It has been proposed that inhibition of reovirus mRNA ribosome binding at high salt concentrations is due to the enhanced secondary structure of

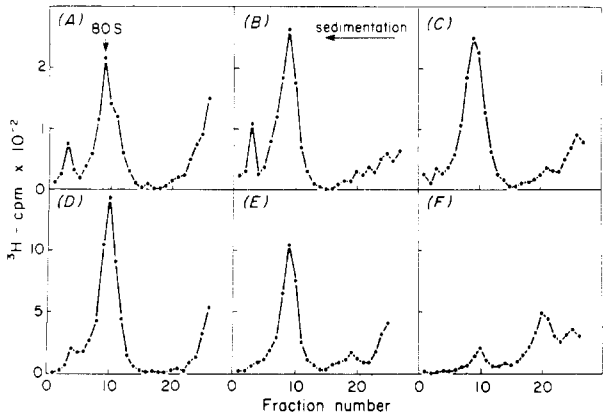

FIGURE 2: Effect of K^+ concentration on binding of AMV-4 RNA and bromouridine-substituted reovirus mRNA to ribosomes. [^{3}H]-Methyl-labeled AMV-4 RNA (2500 cpm, ~0.2 μg, panels A–C) or [^{3}H]methyl-labeled bromouridine-substituted reovirus mRNA (12000 cpm, 0.8 μg, panels D–F) was incubated in a wheat germ extract as described in Figure 1 and under Experimental Procedures. Potassium concentrations were adjusted by the addition of KOAc to the reaction mixtures. The final concentration of K^+ including 45 mM KCl contributed by the wheat germ extract was as follows: (A and D) 45 mM; (B and E) 75 mM; (C and F) 125 mM. The percentage of input mRNA bound to ribosomes was the following: (A) 79%; (B) 76%; (C) 72%; (D) 73%; (E) 60%; (F) 17%.

the mRNA (Lee et al., 1983). Since inosine-substituted reovirus mRNA does not contain extensive secondary structure, it is expected to be less sensitive to inhibition by high salt concentrations. Hence, the ability of AMV-4 RNA to efficiently bind ribosomes at high salt concentrations is consistent with relaxed secondary strucutre.

The enzymatic structure-mapping experiments were designed to analyze the conformation of the 5′-terminus of the molecule, encompassing nucleotides 1–120. Although ribonuclease T1 has been used traditionally for analysis of RNA structure, the absence of guanosine residues in the 5′-nontranslated region required the use of other enzymes. Nuclease S1 was chosen because of its specificity for single-strand RNA; however, the pH optimum of this enzyme is 4.5, somewhat out of the physiological range. We also used mung bean nuclease (MB1), which is active at pH 6.7. Nucleases S1 and MB1 both attack nonpaired nucleotides; however, they do not cleave identically, perhaps because of their interaction with the three-dimensional structure of RNA (Vournakis et al., 1981). Both nucleases S1 and MB1 have a preference for cleaving A and U residues over C and G residues (Pavlakis et al., 1979; Laskowski, 1980). As a means of demonstrating that the digestion patterns reflect structure and not simply base specificity, AMV-4 RNA was cleaved over a wide range of enzyme activities (Figure 3), and the banding patterns were analyzed for significant evidence of secondary cleavages, which occur if new cleavage sites are generated as a result of primary "hits". Conclusions regarding AMV-4 structure were derived from those gel patterns that did not change with increasing enzyme activity.

Ribonucleases T1 (Figure 3, lanes 2 and 3) and Phy M (Figure 3B, lanes 4 and 5) were used under denaturing conditions to generate digestion patterns that unequivocally assigned the nucleotide numbering. Position 1 (Figure 3) represents m^7 GpppGp for T1 and Phy M digestions, while in the S1, V1, and MB1 reactions, position 1 represents m^7GpppG_{OH}. Because the fragments generated by T1 and Phy M have 5′-OH groups, while those generated by S1, V1, and MB1 possess 5′-phosphates, two types of digestion ladders were analyzed. Hydrolysis of mRNA with formamide (Figure 3B, lane 6) yields a ladder with 5′-OH groups, while hydrolysis with S1 under denaturing conditions (Figure 3B, lane 7) yields a ladder with 5′-phosphates. The differences in migration are greatest when low molecular weight fragments are compared (positions 1–18). This behavior is consistent with the fact that 3′-phosphate fragments have a higher charge-to-mass ratio than S1-derived fragments (Auron et al., 1982).

The digestion patterns generated with single strand specific and double strand specific nucleases under our nondenaturing conditions indicate that the 5′-terminus of AMV-4 RNA assumes a predominantly non-base-paired configuration. The S1 reactions (Figure 3A, lanes 5 and 6; Figure 3B, lanes 8–11) were done over a 15-fold range of enzyme activity as judged by the incubation time and the enzyme-to-substrate ratio. The digestion patterns were examined semiquantitatively by visually analyzing both the density of the band representing undigested material at the top of the polyacrylamide gels and also the changes in the patterns of bands appearing with greater concentrations of enzyme and length of incubation time. The patterns in part A, lanes 5 and 6, and part B, lanes 8–11, of Figure 3 represent a 14-fold increase, respectively, in enzyme activity. The kinetics of enzymatic hydrolysis favor large molecules as primary targets under conditions of low enzyme activity, generating large fragments that migrate slowly in the polyacrylamide gel. However, even when the enzyme activity is low (Figure 3A, lanes 5, 6, 9, and 10; Figure 3B, lanes 8 and 16–19), nucleotide positions 3–14 are relatively strong targets for the single strand specific enzymes. These fragments are interpreted as being the result of primary enzymatic cleavages because the great majority of labeled RNA migrates as undigested material and also because with increasing enzyme-to-substrate ratios, there is no apparent loss of high molecular weight bands with an accompanying accumulation of radioactivity in low molecular weight fragments.

The cleavage pattern generated by the cobra venom V1 double strand specific ribonuclease is shown in part A, lanes 7 and 8, and in part B, lanes 12–15, of Figure 3. There is a region of relatively strong cleavages centered on positions 10 and 11, and there are also two sites (positions 34 and 36) immediately preceding the initiation codon (nucleotides 37–39) (Figure 3A, lanes 7 and 8). In general, the ratio of enzyme concentration and incubation time to substrate was much greater in the reaction mixtures analyzed in part B than that in part A of Figure 3. (For example, the V1 ribonuclease concentration and length of incubation were 30 times greater in the reaction shown in part B, lane 15, as that in part A, lane 7 of Figure 3.) With the exception of the cleavage sites at positions 9–12, nucleotides 1–30 did not provide a good substrate for the V1 ribonuclease, and the quantity of enzyme and length of incubation had to be increased significantly in order to visualize the cleavage sites as bands on the polyacrylamide gels (Figure 3B, lanes 12–15).

Figure 4 shows a summary of the nuclease cleavage sites and their relative intensities. The data indicated that in addition to the strong single strand cleavage sites at positions 3–25, there were also reproducible low-intensity V1 cleavages in the same region. Two interpretations of this result are that (1) the cobra venom V1 ribonuclease is not completely specific for double-strand regions or (2) positions 3–25 assume multiple conformations with some equilibrium among them. A computer-generated dot matrix analysis was used to test the potential of the 5′-terminus for forming stable base pairs with other regions of the entire AMV-4 RNA molecule. The computer analysis is performed by defining a matrix in which each nucleotide of the AMV-4 sequence was compared with

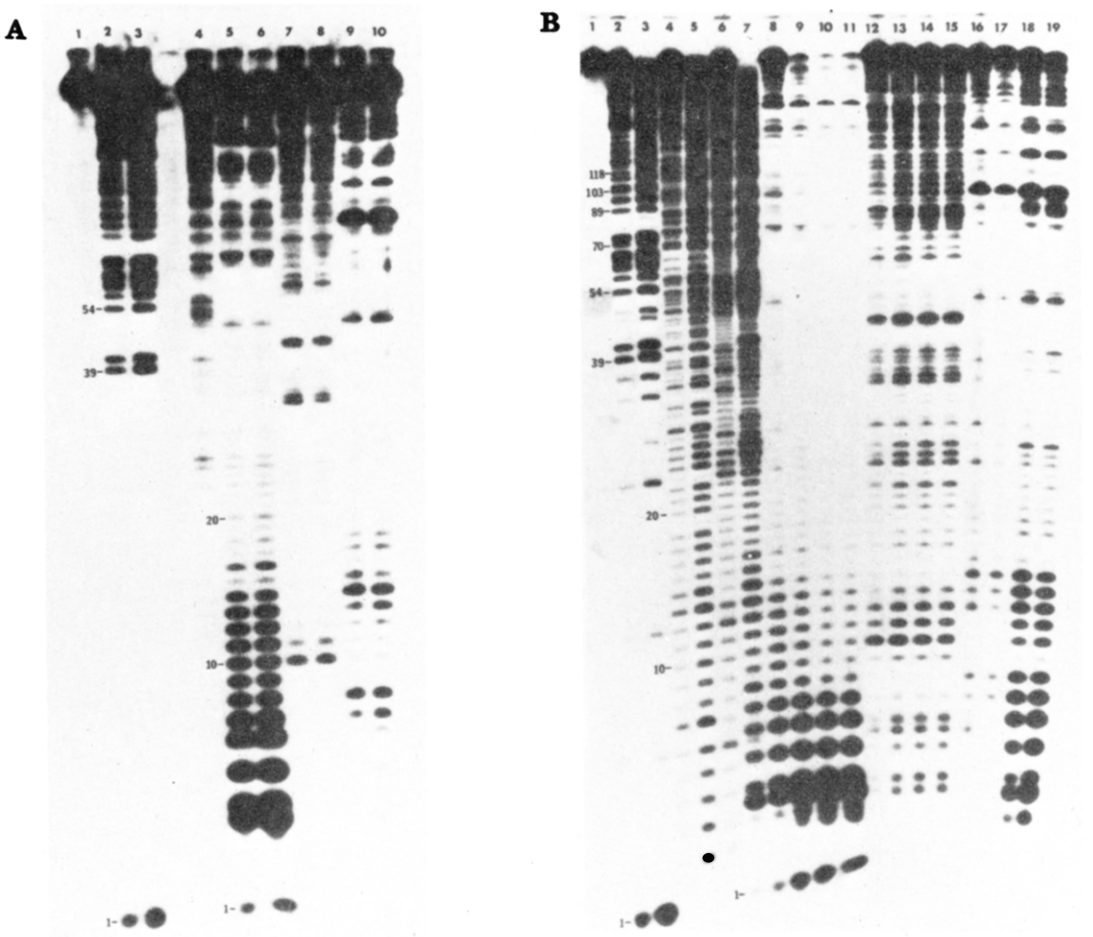

FIGURE 3: (A) Electrophoretic analysis of partial enzymatic digests of cap-labeled AMV-4 RNA. The polyacrylamide gel was 0.3 mm thick and was 15% polyacrylamide–8 M urea in TBE buffer (90 mM Tris, 2.5 mM EDTA, 90 mM boric acid). Enzymatic reaction buffer conditions are described in the text. (1) Minus enzyme. (2 and 3) T1 ribonuclease; 0.5 and 1.0 unit of enzyme, respectively, per 0.1 A_{260} unit of RNA was added. Incubation was at 60 °C for 10 min. (4) S1 ladder, underdigested. (5 and 6) S1 nuclease, native. 2.5 units of enzyme per 0.05 A_{260} unit of RNA was added. Incubation was at 37 °C for 1 and 2 min, respectively. (7 and 8) V1 ribonuclease, native. The (sulfoethyl)cellulose column fraction that exhibited the greatest double-strand nuclease activity was used throughout. The fraction was diluted 1:100 with native buffer, and 0.5 μL of diluted enzyme per 0.05 A_{260} unit of RNA was added. Incubation was at 25 °C for 10 and 60 s, respectively. (9 and 10) Mung bean nuclease, native. Stock MB1 (P-L Biochemicals) was diluted 1:100 with native buffer. 0.5 μL of diluted enzyme was added per 0.05 A_{260} unit of RNA, and the mixtures were incubated at 25 °C for 30 and 60 s, respectively. (B) Electrophoretic analysis of partial enzymatic digests of cap-labeled AMV-4 RNA. The gel is identical with that shown in Figure 1. (1) Minus enzyme. (2 and 3) T1 ribonuclease, denaturing conditions. 1.25 and 2.5 units of enzyme, respectively, was added per 0.1 A_{260} unit of RNA, followed by incubation at 55 °C for 10 min. (4 and 5) Phy M nuclease, denaturing conditions. 2.5 and 5.0 units, respectively, of enzyme were added per 0.1 A_{260} unit of RNA. The reaction mixtures were incubated at 55 °C for 10 min. (6) Formamide ladder. 0.1 A_{260} unit of RNA was solubilized in 2.5 μL of deionized formamide, and the solution was incubated at 90 °C for 10 min. (7) S1 ladder, denaturing conditions. 1.25 units of S1 nuclease was added per 0.05 A_{260} unit of RNA to a reaction containing 7 M urea. Incubation was at 60 °C for 20 min. (8–11) S1 nuclease, native. 1.25 units of enzyme was added per 0.05 A_{260} unit of RNA. Incubation was at 37 °C for 1, 5, 10, and 15 min, respectively. (12–15) V1 ribonuclease, native. The (sulfoethyl)cellulose column fraction (Auron et al., 1982) was diluted 1:5 with native buffer, and 0.5 μL of diluted enzyme was added per 0.05 A_{260} unit of RNA. Incubation was at 37 °C for 1, 5, 10, and 15 min, respectively. (16–19) Mung bean nuclease, native. Stock enzyme (P-L Biochemicals) was diluted 1:100 with native buffer, and 0.5 μL of diluted enzyme was added per 0.05 A_{260} unit of RNA. Incubation was at 37 °C for 0.5, 1, 5, and 10 min.

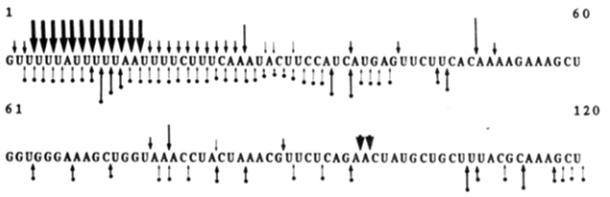

FIGURE 4: Summary figure illustrating the experimental single-strand and double-strand nuclease cleavage sites at the 5′-end of AMV-4 RNA. The intensity of the cleavages is indicated by both the length and breadth of arrows, that is, the longer and broader the arrow, the greater the intensity of cleavage. Single strand cleavage sites (nucleases S1 and MB1) are located above the sequence; double strand nuclease sites (V1) are located below the sequence and are shown by arrows whose tails end in closed circles.

all other nucleotides and analyzed for potential base pairing. In this way, all regions of potential secondary structure are identified, and base-paired stems are represented by diagonals. A reflection axis is drawn from the upper left corner to the lower right corner of each matrix plot. This reflection axis separates pairs of symmetrical diagonals; each diagonal in a pair represents one strand of a complementary pair of strands.

Figure 5A represents a filtered-matrix secondary-structure analysis in which only those duplex regions characterized by six or more base pairs and a relative free-energy value of less than or equal to −9 are shown. In Figure 5B two matrix analyses are presented. The upper right half has been filtered at a minimum accepted pairing length of 4 and a maximum energy value of −9. The lower left half has been filtered at a minimum accepting pairing length of 8 and a maximum relative energy filter of −9. Multiple diagonal elements along a given horizontal or vertical line drawn across the matrix cannot simultaneously occur in a given molecule. The true

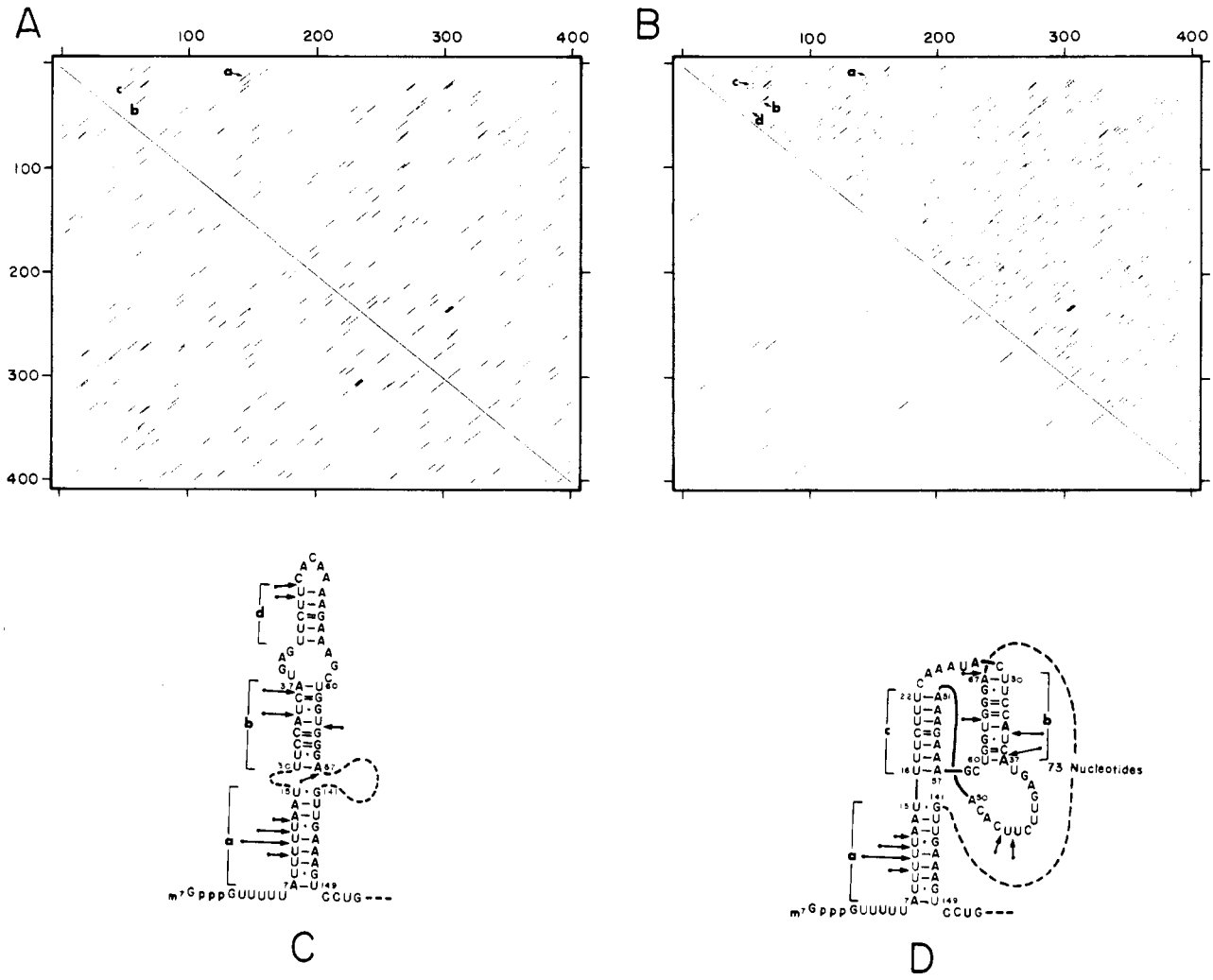

FIGURE 5: Dot matrix analysis illustrating potential secondary structures at the 5′-end of AMV-4 RNA. Only positions 1–400 are shown. (A) Matrix program output generated by analyzing AMV-4 RNA with the length filter set at 6 and the energy filter set at −9. (B) A composite matrix of AMV-4 RNA consisting of two half-matrices positioned on either side of the reflection diagonal. The lower left half is filtered with the length filter set at 8 and the energy filter set at −9, while the upper right of the figure was generated with the length filter at 4 and the energy filter at −9. (C and D) Representations of potential structures predicted by the matrix analyses shown in (A) and (B). Lower case letter designations denote base-paired regions corresponding to the diagonals labeled in (A) and (B).

molecular folding may be represented by only one diagonal, or alternatively, the molecules may sample several of the possibilities, particularly if none is significantly more stable than the others.

The dot matrix data indicate that, indeed, it is theoretically possible for nucleotides in the 5′-terminal region of the mRNA to form alternative base pairs, and this pairing results in trade-off structures, which might represent structural heterogeneity. The experimental evidence supports two such structures, which are shown in Figure 5C,D. These two structures share common base pairs (stems a and b) and, in addition, contain regions that can form trade-off structures (stems c and d). The calculated free energies (Tinoco et al., 1973) of these regions are −3.6 kcal/mol (stem a), −9.4 kcal/mol (stem b), −9.2 kcal/mol (stem c), and −6.8 kcal/mol (stem d). NMR melting-temperature studies conducted on model oligoribonucleotides of similar size (Alkema et al., 1981) indicate that such duplexes should have T_m values of less than 30 °C. Therefore, under our experimental conditions, fewer than half the molecules, on average, should contain such duplex structures.

Several other trade-off structures were identified in the matrix; however, they were not supported by the experimental data. Nevertheless, the fact that weak, but detectable, V1 digestion occurs throughout the 30 5′-proximal nucleotides of AMV-4 RNA suggests that some of these alternate structures might contribute to the overall equilibrium disposition of paired nucleotides. Our interpretation, therefore, is that the V1 ribonuclease is a highly specific and sensitive probe for duplex regions in AMV-4 RNA. The experimental and theoretical data indicate that the 5′-untranslated region of AMV-4 RNA is principally single stranded under native conditions and that unstable paired complexes may also form transiently with this U-rich region.

Discussion

The possible role of the cap structure and of cap-binding protein in modulating the efficiency of mRNA translation has been suggested (Muthukrishnan et al., 1976; Sonenberg et al., 1980, 1981), and it has been proposed that the cap-binding protein may act to destabilize secondary structure of mRNA and thereby facilitate ribosome binding (Sonenberg, 1981; Sonenberg et al., 1981). Because AMV-4 RNA seems to function with relative efficiency in the presence of cap analogues (Van Vloten-Doting et al., 1977; Sonenberg et al., 1981; Figure 1), we decided to theoretically model the 5′-structure of AMV-4 RNA and carry out experiments to determine if the native conformation reflects the apparent translational

independence of the cap or cap-binding protein.

The nucleotide sequence of AMV-4 RNA has been reported (Brederode et al., 1980), and computer programs (Zuker & Stiegler, 1981; G. J. Quigley and P. E. Auron, unpublished results) were used to analyze the AMV-4 RNA nucleotide sequence and attempt to predict the secondary structure formed by the molecule. Subgenomic plant viral RNAs are commonly characterized by 5′-U+A-rich sequences that have been called "Butler boxes" (Goelet et al., 1982). In AMV-4, the 5′-nontranslated region is 36 nucleotides long and is 81% A+U. Because of the large number of uridine residues (56%), it seemed unlikely that the 5′-untranslated region could form hairpin loops, and the computer programs support this conclusion (Figure 5). The experimental approach to analysis of mRNA secondary structure has involved partial enzymatic digestion of AMV-4 RNA under native conditions. The substrate for structure-mapping experiments was radioactively labeled, capped AMV-4 RNA. The labeling protocol involved first removing the 5′-terminal m^7G by periodate oxidation and β-elimination, followed by regeneration of the cap structure by incubating the RNA with [α-^{32}P]GTP, vaccinia guanylyltransferase, and *S*-adenosylmethionine. Muthukrishnan et al. (1978) have shown that the periodate/aniline treatment does not cause significant mRNA damage and that translatable mRNA is regenerated by recapping β-eliminated mRNA. Our data strongly indicate that the labeled AMV-4 molecules were intact and that the labeling process caused minimal damage to the RNA (Figure 3; unpublished data). To remove minor degradation products generated during the labeling procedure, AMV-4 RNA was always purified by preparative gel electrophoresis prior to the structure-mapping reactions. Approximately 75% of the total incorporated label migrated with intact AMV-4 RNA during preparative electrophoresis.

Ribonuclease V1 is isolated from cobra venom (Vassilenko & Ryte, 1975; Auron et al., 1982), and the enzyme has a very high degree of specificity for base-paired RNA with some affinity for stacked, but non-base-paired bases at the ends of helices (Auron et al., 1982). The experimental data collected here on AMV-4 RNA showed that, distal to position 26, the single-strand cleavage patterns (S1 and MB1) are basically complementary to the double-strand (V1) ribonuclease pattern (Figures 3 and 4). These data support the conclusion that the 5′-terminus of AMV-4 RNA is principally single stranded. We did, however, note that very weak V1 cleavages, as compared to the intensity of the S1 patterns, were reproducibly observed at positions 2–26 and these sites overlapped with the region that is very strongly cleaved by S1 and MB1 nucleases. By combining results from computer modeling and experimental enzymatic structure mapping, we conclude that the 5′-terminus of AMV-4 RNA is principally single stranded but that the region retains the potential for assuming multiple conformations, including a population of molecules that are base paired in the 5′-nontranslated region. The first evidence for potential duplex regions at the 5′-end was derived from the results of the V1 enzymatic digestions. The great majority of the AMV-4 RNA molecules were not cleaved in the untranslated region, and it was necessary to increase the V1 enzyme activity considerably in order to visualize the sites in positions 1–25 (Figure 3B). As compared to the S1 patterns, the cleavages are clearly very weak, and it was possible that the V1 enzyme was not completely specific for double-stranded RNA. Our data do not support the possibility of limited specificity by V1 ribonuclease; on the contrary, on the basis of the agreement of the enzymatic cleavage patterns and the computer analysis of potential secondary structure, the results indicate that V1 ribonuclease is an extremely specific and sensitive probe for locating secondary structure. Digestion of AMV-4 RNA with S1 and V1 nucleases at sites other than the 5′-proximal region reveals a complimentary pattern of cleavages in which V1 and S1 do not seem to attack the same nucleotide positions (Figure 3). The dot matrix analysis (shown in part in Figure 5) reveals all potential regions of secondary structure in AMV-4 RNA, including G-U base pairs. Indeed, the computer model showed that there are several regions within the first 25 nucleotides that can form weak intramolecular duplexes. One of these regions (stem a) corresponds exactly to a region of relatively strong (compared to other V1 sites) V1 cleavage sites (Figures 3 and 4). The data indicate that two potential structures (Figures 5C,D), in combination with numerous shorter base-paired regions, could provide a substrate for V1 nuclease. The existence of weak, easily melted trade-off secondary structures in this part of the molecule is consistent with the preponderance of weak V1 cleavages superimposed upon the strong S1 cleavages.

The mechanisms that underly translational efficiency in eukaryotes are not understood. There is no evidence in eukaryotes of a sequence homology or a conservation of the 5′-end in the noncoding regions that might constitute a ribosome binding site (Baralle & Brownlee, 1978). Messenger RNA secondary structure has been implicated in facilitating ribosome binding and initiation of protein synthesis (Kozak, 1980a; Lee et al., 1983). The role of proteins in modulating RNA structure has also been suggested and is now being tested directly. For example, Douthwaite et al. (1982) have recently used structure-mapping methods to analyze structural changes in *Escherichia coli* 5S RNA that are induced by ribosomal proteins L18 and L25. Their results indicate that the RNA–protein interaction resulted in RNA conformational changes both within and outside the binding domain of the proteins.

Cordell et al. (1982) have recently reported that expression of two nonallelic rat insulin genes in a transplantable insulinoma is regulated at the level of initiation of translation. In this system, translation of rat insulin 1 mRNA is 10-fold greater than translation of rat insulin 2 mRNA; however, the ratio of expression is nearly 1:1 if the mRNAs are translated in the presence of a vaccinia capping enzyme extract. Thus, the data are supportive of the idea of significant in vivo regulatory mechanisms involving mRNA, the cap structure, and perhaps mRNA structure. Structure mapping of mRNAs and mRNPs whose efficient translation is strongly dependent upon the integrity of the 5′-cap is required for further understanding of posttranscriptional regulation of eukaryotic gene expression.

Acknowledgments

We thank John Bol of the University of Leiden for the generous gift of AMV-4 RNA.

References

Alkema, D., Bell, R. A., Hader, P. A., & Neilson, T. (1981) in *Biomolecular Stereodynamics* (Sarma, R. H., Ed.) Vol. I, Adenine Press, New York.

Auron, P. E., Weber, L. D., & Rich, A. (1982) *Biochemistry 21*, 4700–4706.

Banerjee, A. K. (1980) *Bacteriol. Rev. 44*, 175–205.

Baralle, F. E., & Brownlee, G. G. (1978) *Nature (London) 274*, 84–87.

Bol, J. F., & Van Vloten-Doting, L. (1973) *Virology 51*, 102–108.

Brederode, F. T., Koper-Zwarthoff, E. C., & Bol, J. F. (1980) *Nucleic Acids Res. 8*, 2213–2223.

Cordell, B., Diamond, D., Smith, S., Punter, J., Schone, H. H., & Goodman, H. (1982) *Cell (Cambridge, Mass.) 31*, 531–542.

Douthwaite, S., Christensen, A., & Garrett, R. A. (1982) *Biochemistry 21*, 2313–2320.

Ensinger, M. J., Martin, S. A., Paoletti, E., & Moss, B. (1975) *Proc. Natl. Acad. Sci. U.S.A. 72*, 2525–2529.

Etchison, D., Milburn, S. C., Edery, I., Sonenberg, N., & Hershey, J. W. B. (1982) *J. Biol. Chem. 257*, 14806–14810.

Fraenkel-Conrat, H., & Steinschneider, A. (1967) *Methods Enzymol. 12B*, 243–246.

Goelet, P., Lomonossoff, G. P., Butler, P. J. G., Akam, M. E., Gait, M. J., & Karn, J. (1982) *Proc. Natl. Acad. Sci. U.S.A. 79*, 5818–5822.

Grunberg-Manago, M. (1980) in *Ribosomes: Structure, Function and Genetics* (Chamblis, G., Craven, G. R., Davie, J., Davis, K., Kahan, L., & Nomura, M., Eds.) pp 445–477, University Park Press, Baltimore, MD.

Gupta, R. C., & Randerath, K. (1979) *Nucleic Acids Res. 6*, 3443–3458.

Herson, D., Schmidt, A., Seal, S. N., Marcus, A., & Van Vloten-Doting, L. (1979) *J. Biol. Chem. 254*, 8245–8249.

Ilan, J., & Ilan, J. (1977) *Proc. Natl. Acad. Sci. U.S.A. 74*, 2325–2329.

Jackson, R. J. (1982) in *Protein Biosynthesis in Eukaryotes* (Perez-Bercoff, R., Ed.) pp 362–418, Plenum Press, New York.

Kozak, M. (1980a) *Cell (Cambridge, Mass.) 22*, 459–467.

Kozak, M. (1980b) *Cell (Cambridge, Mass.) 19*, 79–90.

Laskowski, M., Sr. (1980) *Methods Enzymol. 65*, 263–276.

Lee, K. A. W., Guertin, D., & Sonenberg, N. (1983) *J. Biol. Chem. 258*, 707–710.

Lodish, H. F. (1976) *Annu. Rev. Biochem. 45*, 39–72.

Maizel, J. V., & Lenk, R. P. (1981) *Proc. Natl. Acad. Sci. U.S.A. 78*, 7665–7669.

Morgan, M. A., & Shatkin, A. J. (1980) *Biochemistry 19*, 5960–5966.

Moss, B. (1977) *Biochem. Biophys. Res. Commun. 74*, 374–383.

Muthukrishnan, S., Morgan, M., Banerjee, A. K., & Shatkin, A. J. (1976) *Biochemistry 15*, 5761–5768.

Muthukrishnan, S., Moss, B., Cooper, J. A., & Maxwell, E. S. (1978) *J. Biol. Chem. 253*, 1710–1715.

Pavlakis, G. N., Jordan, B. R., Wurst, R. M., & Vournakis, J. N. (1979) *Nucleic Acids Res. 7*, 2213–2238.

Pinck, L. (1975) *FEBS Lett. 29*, 24.

Plotch, S. J., Bouloy, M., & Krug, R. M. (1979) *Proc. Natl. Acad. Sci. U.S.A. 76*, 1618–1622.

Pustell, J., & Kafatos, F. C. (1982) *Nucleic Acids Res. 10*, 4764–4782.

Shatkin, A. J. (1976) *Cell (Cambridge, Mass.) 9*, 645–653.

Sonenberg, N. (1981) *Nucleic Acids Res. 9*, 1643–1656.

Sonenberg, N., & Shatkin, A. J. (1977) *Proc. Natl. Acad. Sci. U.S.A. 74*, 4288–4292.

Sonenberg, N., Merrick, W. C., Morgan, M. A., & Shatkin, A. J. (1978) *Proc. Natl. Acad. Sci. U.S.A. 75*, 4843–4847.

Sonenberg, N., Rupprecht, K. M., Hecht, S. M., & Shatkin, A. J. (1979) *Proc. Natl. Acad. Sci. U.S.A. 76*, 4345–4349.

Sonenberg, N., Trachsel, H., Hecht, S., & Shatkin, A. J. (1980) *Nature (London) 285*, 4345–4349.

Sonenberg, N., Guertin, D., Cleveland, D., & Trachsel, H. (1981) *Cell (Cambridge, Mass.) 27*, 563–572.

Sonenberg, N., Guertin, D., & Lee, K. A. W. (1982) *Mol. Cell. Biol. 2*, 1633–1638.

Tahara, S. M., Morgan, M. A., & Shatkin, A. J. (1981) *J. Biol. Chem. 256*, 7691–7694.

Tinoco, I., Borer, P. N., Dengler, B., Levine, M. D., Uhlenbeck, O. C., Crothers, D. M., & Gralla, J. (1973) *Nature (London), New Biol. 246*, 40–41.

Van Vloten-Doting, L., & Jaspers, E. M. J. (1977) in *Comprehensive Virology* (Fraenkel-Conrat, H., & Wagner, R. R., Eds.) Vol. 11, Plenum Press, New York.

Van Vloten-Doting, L., Bol, J., Neeleman, L., Rutgers, T., Van Dalen, D., Castel, A., Bosch, L., Marbaix, G., Huez, G., Hubert, E., & Cleuter, Y. (1977) in *NATO Advanced Studies on Nucleic Acids and Protein Synthesis in Plants* (Bogorad, L., & Weil, M., Eds.) p 387, Plenum Press, New York.

Vassilenko, S. K., & Ryte, V. K. (1975) *Biokhimiya (Moscow) 40*, 578–583.

Vournakis, J. N., Celantano, J., Finn, M., Lockard, R. E., Mitra, T., Pavlakis, G., Troutt, A., van den Berg, M., & Wurst, R. M. (1981) in *Gene Amplification and Analysis* (Chirkjian, J. G., & Papas, T., Eds.) Vol. I, pp 267–298, Elsevier/North-Holland, New York.

Wrede, P., Wurst, R., Vournakis, J., & Rich, A. (1979) *J. Biol. Chem. 254*, 9608–9616.

Wurst, R. M., Vournakis, J. N., & Maxam, A. M. (1978) *Biochemistry 17*, 4493–4499.

Zuker, M., & Stiegler, P. (1981) *Nucleic Acids Res. 9*, 133–148.