Nucleotide Sequence of the 5′ Terminus of Satellite Tobacco Necrosis Virus Ribonucleic Acid†

David W. Leung,‡ Karen S. Browning, Joyce E. Heckman, Uttam L. RajBhandary, and John M. Clark, Jr.*

ABSTRACT: Treatment of the RNA of satellite tobacco necrosis virus (STNV) with phosphomonoesterase followed by heat denaturation and treatment with polynucleotide kinase in the presence of [γ-^{32}P]ATP yields a STNV [5′-^{32}P]RNA containing a homogeneous 5′ terminus. Analyses of this STNV [5′-^{32}P]RNA yield the sequence of the first 42 nucleotides from the 5′ terminus of STNV RNA. This nucleotide sequence contains the translation initiation AUG codon starting at position 30 from the 5′ terminus as indicated by match of subsequent nucleotides with the genetic code assignments for the N-terminal amino acids of STNV coat protein in the 5′-terminal sequence ppAGUAAAGACAGGAAACUU-UACUGACUAACAUGGCAAAACAAC. An interesting feature of this sequence is its potential to form a hairpin loop structure involving perfect Watson–Crick base pairing between the first seven nucleotides and nucleotides at positions 16–22.

The specific features of eucaryotic mRNAs which influence their binding to ribosomes are not known. Comparisons of the 5′-terminal noncoding sequences of several eucaryotic mRNAs reveal no readily identifiable common features other than the initiator AUG[1] codon and the 5′-terminal m^{7}G-(5′)ppp(5′)Np··· cap structure found in almost all eucaryotic mRNAs (Shatkin, 1976; Baralle & Brownlee, 1978). Yet closely related mRNAs with no apparent common features are translated in vivo with similar efficiency (Kozak & Shatkin, 1978; Hagenbüchle et al., 1978; Rose, 1978). Kozak & Shatkin (1977a) have used these observations and the fact that 40S ribosome initiation complexes protect both the 5′-terminal cap and the initiator AUG of a variety of reovirus mRNAs to propose that initial recognition of eucaryotic mRNAs involves 40S ribosomal interaction with the 5′-terminal cap and the initiator AUG. Such interaction with both the 5′-terminal cap and the initiator AUG is easy to visualize in mRNAs such as BMV RNA 4 (Dasgupta et al., 1975) and the VSV N, NS, and L mRNAs (Rose, 1977) where the initiator AUG is located only 10–14 nucleotides from the 5′ terminus. Similar interaction with both the 5′-terminal cap and the initiator AUG may exist with the reovirus mRNAs where the initiator AUGs are 20–40 nucleotides from the 5′ termini but may be brought close to the cap structures by secondary folding of the 5′-terminal regions of the mRNAs (Kozak & Shatkin, 1977b). However, other eucaryotic mRNAs have large numbers of nucleotides between their 5′-terminal capping groups and their initiator AUGs (Kop-er-Zwarthoff et al., 1977; Baralle, 1977; Richards et al., 1977; Haseltine et al., 1977) so that one must invoke extensive secondary and tertiary folding of these mRNAs to bring their 5′-terminal capping groups and the initiator AUGs into close proximity with each other.

Other workers have proposed that hybridization between eucaryotic mRNAs and some portion of the 3′ terminus of 18S rRNAs may also enhance the specific binding of eucaryotic mRNAs into initiation complexes. Baralle & Brownlee (1978) suggest such mRNA–18S rRNA interaction employs the initiator AUG of mRNAs. Hagenbüchle et al. (1978) suggest that an uncoded CUUPyPyG sequence preceding the initiator AUG interacts with a polypurine sequence present near the 3′ end of 18S rRNAs. Yet these theories lack extensive experimental proofs and suffer from the fact that some eucaryotic mRNAs lack the necessary CUUPyPyG sequence preceding their initiator AUG (Baralle & Brownlee, 1978).

Virtually all eucaryotic mRNAs contain a 5′-terminal capping group which appears to function during eucaryotic mRNA translation. Such capping groups increase the rates of mRNA binding to ribosomes and enhance the efficiency of eucaryotic mRNA translation both in vivo and in vitro (Shatkin, 1976; Lockard & Lane, 1979). However, STNV RNA, polio mRNA, EMC RNA, and cowpea mosaic virus RNA do not contain 5′-terminal cap structures (Wimmer et al., 1968; Nuss et al., 1975; Hewlett et al., 1976; Nomoto et al., 1977; Klootwijk et al., 1977). It is not known whether these eucaryotic mRNAs contain unique structural or sequence features which compensate for the absence of 5′-terminal cap groups.

This paper reports the nucleotide sequence of the 5′ terminus of satellite tobacco necrosis virus RNA. Two features of STNV RNA suggest that this eucaryotic mRNA may provide a useful link in our understanding of the structure–function relationship of eucaryotic mRNAs. First, STNV RNA is one of the few eucaryotic mRNAs that does not contain or employ a 5′-terminal capping group (Leung et al., 1976); STNV RNA may therefore contain features which compensate for this lack of a 5′-terminal capping group. Second, STNV RNA is a small monocistronic eucaryotic mRNA which correctly codes for STNV coat protein in both procaryotic and eucaryotic systems (Klein et al., 1972; Lundquist et al., 1972; Klein & Clark, 1973; Leung et al., 1976). Thus STNV RNA must contain the basic features recognized by both procaryotic and eucaryotic systems. This paper describes the use of Edman degradation to characterize the first 7 N-terminal amino acids of STNV coat protein and in vitro ^{32}P-labeling methods to characterize the first 42 nucleotides of the 5′ terminus of STNV RNA. A comparison of the nucleotide sequence

† From the Biochemistry Department, University of Illinois, Urbana, Illinois 61801 (D.W.L., K.S.B., and J.M.C.), and the Department of Biology, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139 (J.E.H. and U.L.R.). *Received December 19, 1978.* This work was supported by U.S. Public Health Service Grant GM 08647 to J.M.C. and U.S. Public Health Service Grant GM 17151 and American Cancer Society Grant NP114-C to U.L.R.

‡ Present address: Department of Biochemistry, University of British Columbia, Vancouver, B.C., Canada V6T 1W5.

[1] Abbreviations used: initiator AUG, an AUG codon serving as a translation initiation site; cap or capping group, any 5′-terminal m^{7}G-(5′)ppp(5′)Np··· group on mRNA; BMV, brome mosaic virus; VSV, vesticular stomatitis virus; EMC, encephalomyocarditis; AlMV, alfalfa mosaic virus; TMV, tobacco mosaic virus; Pth, phenylthiohydantoin; PEI-cellulose, cellulose pretreated with polyethylenimine; STNV, satellite tobacco necrosis virus.

0006-2960/79/0418-1361$01.00/0

predicted from the N-terminal amino acid sequence with the observed 5′-terminal nucleotide sequence establishes that the initiator AUG codon begins 30 nucleotides from the 5′ terminus of STNV RNA.

Experimental Procedures

Materials. Reagents for Edman degradation analyses were purchased for Pierce Chemical Co. STNV coat protein and STNV RNA were prepared from STNV isolated from germinating mung beans infected with the B strain of TNV and the SV1 strain of STNV (Clark & Klein, 1974). [γ-^{32}P]ATP was prepared by means of a glyceraldehyde-3-phosphate dehydrogenase (Calbiochem Co.) and 3-phosphoglycerate kinase (Calbiochem Co.) dependent exchange of [^{32}P]P_i into ATP (Maxam & Gilbert, 1977). RNase T_1 and RNase U_2 were purchased from Calbiochem Co. Nuclease P_1 and nucleoside 3,5-diphosphates were purchased from P-L Biochemicals, Inc. RNase Y was prepared from *Bacillus cereus* (Lockard et al., 1978). RNase Phy I was prepared from *Physarum polycephalum* (Philly et al., 1978). Bacterial alkaline phosphatase and T_4 polynucleotide kinase were purchased from Boerhinger Mannheim Co. UltraPure urea was purchased from Schwarz/Mann Co. Other reagent chemicals were as provided by Sigma-Aldrich Co. Polygram Cel 300 DEAE/HR-2/15 plastic backed DEAE-cellulose thin-layer plates (Brinkmann Co.) were used for homochromatographic analysis of the 5′-terminal homogeneity of STNV [5′-^{32}P]RNA. Uniplate Cell HR/Avicel/DEAE glass-backed DEAE-cellulose thin-layer plates (Analtech, Inc.) were used in "wandering spot" analyses. Polygram Cel 300 PEI-cellulose thin-layer plates (Brinkmann Co.) were used to characterize the 5′-terminal nucleotide of STNV [5′-^{32}P]RNA.

Edman Degradation of STNV Coat Protein. The N-terminal sequence of 7.5 mg of STNV coat protein was determined by automatic Edman degradation on a Beckman Model 890-C sequencer after initial derivatization with 4-sulfonophenyl isothiocyanate (Bailey et al., 1977). The resultant anilinothiazolinone derivatives from cycles 3–7 were converted to Pth-amino acids by treatment in 0.1 mL of 1.0 M HCl at 80 °C for 10 min prior to extraction into ethyl acetate and thin-layer chromatographic characterization (Jeppsson & Sjöquist, 1967; Li & Riggs, 1970). Pth-Arg (as indicated by the absence of any detectable Pth-amino acid in the ethyl acetate extraction of the seventh degradation cycle) was detected in the aqueous phase of the seventh cycle by the phenanthrenequinone dip test of Niall (1973).

Preparation of STNV [5′-^{32}P]RNA. One hundred micrograms of STNV RNA was incubated with 0.2 unit of bacterial alkaline phosphatase in 0.1 mL of 20 mM Tris-HCl, pH 8.0, at 37 °C for 30 min. The reaction mixture was then extracted with phenol and the STNV 5′-dephospho-RNA was recovered from the aqueous phase by ethanol precipitation (Palimiter, 1974). The final pellet of STNV 5′-dephospho-RNA was washed by resuspension in cold 95% ethanol (0.5 μg of RNA/μL), resedimentation (8000 g, 5 min, 1 °C), and lyophilization.

One hundred micrograms of STNV 5′-dephospho-RNA was dissolved in 25 μL of water and placed in a 55 °C water bath for 3 min before rapid cooling in an ice bath. To this solution were added 10 μL of 50 mM Tris-HCl, pH 8.0, 50 mM $MgCl_2$, 50 mM dithiothreitol, 10 μL of [γ-^{32}P]ATP (1.5 nmol, 1–2 mCi/nmol), 5 units of T_4 polynucleotide kinase, and H_2O up to 50 μL. This reaction was incubated at 37 °C for 30 min and then stopped by addition of 0.1 mL of 7 M urea, 0.1% xylene cyanole FF, 0.1% bromophenol blue. This solution was loaded onto 7 cm of a 1.5-mm deep, 30-cm long, 4% polyacrylamide slab gel preequilibrated with 50 mM Tris–borate, pH 8.3, 2 mM Na_2EDTA, 7 M urea (Lockard & RajBhandary, 1976). Electrophoresis was at 25 °C and 250–400 V for 12–24 h, i.e., until the xylene cyanole FF dye ran off the gel. Smaller scale (1–5% of total sample) analytical separations (Figure 1) were also performed in a similar manner. After radioautographic analysis (Du Pont Cronex 2DC film, 15 min, 25 °C), the zone of intact STNV [5′-^{32}P]RNA was cut from the slab gel and the STNV [5′-^{32}P]RNA was recovered by electrophoretic elution into dialysis bags (Gillum et al., 1975; Lockard & RajBhandary, 1976), precipitation of the calcium salt (Dolja et al., 1977), and centrifugation (8000*g*, 20 min, 1 °C). This product was further purified by solubilization in 20 mM Na_2EDTA, 0.1 M NaOAc, pH 6.0 (0.5 μg/μL), reprecipitation for 4–12 h with 2 volumes of −15 °C absolute ethanol, centrifugation (8000*g*, 20 min, 1 °C), further solubilization in cold 0.1 M NaOAc (0.5 μg/μL), and final precipitation for 4–12 h with 2 volumes of −15 °C absolute ethanol and centrifugation as before.

Analysis of the Homogeneity of the 5′ Terminus of STNV [5′-^{32}P]RNA. The 5′-terminal nucleotide of STNV [5′-^{32}P]RNA was determined by incubating a small aliquot of STNV [5′-^{32}P]RNA ($\sim$1000 cpm of ^{32}P) for 2 h at 37 °C with 0.05–0.1 unit of RNase T_2 in a 5-μL reaction containing 0.1 M NH_4OAc, pH 4.5. The entire reaction was then spotted onto a 4 × 20 cm strip of PEI-cellulose along with 50-μmol aliquots of pAp, pGp, pCp, and pUp standards and developed with 0.8 M $(NH_4)_2SO_4$ by thin-layer chromatography at 25 °C. The 5′-terminal nucleotide was identified by comparison of the migration of radioactive material with that of the UV-detectable standards.

The homogeneity of the 5′ terminus of STNV [5′-^{32}P]RNA was further established by separate incubations of small aliquots of STNV [5′-^{32}P]RNA ($\sim$1000 cpm of ^{32}P) with 0.005 unit of RNase T_1 and 5 μg of RNase A. These incubations were for 1 h at 37 °C in 5-μL reactions (no salts added). These reactions were spotted separately onto 4 × 20 cm DEAE-cellulose plates along with a RNase T_2 digested sample of STNV [5′-^{32}P]RNA (above) and subjected to homochromatography at 65 °C using a 0.05 M KOH homochromatography mix (Silberklang et al., 1977). The homogeneity of the resultant ^{32}P-labeled digestion products was evaluated by radioautography.

5′-Terminal Nucleotide Sequence Analysis of STNV [5′-^{32}P]RNA. "Wandering spot" analysis of STNV [5′-^{32}P]RNA employed 8-μL, 50 °C reactions containing 25 mM NH_4OAc, pH 5.3, 0.5 ng of nuclease P_1, and $\sim 10^5$ cpm of STNV [5′-^{32}P]RNA augmented with unlabled STNV RNA to contain 10 μg of total RNA. At 10 min, and again at 15 min, 4-μL reaction aliquots were mixed with 1 μL of 25 mM Na_2EDTA, pH 7.0, and the enzyme was inactivated by heating at 100 °C for 3 min. The reaction aliquots were then combined, lyophilized to dryness, taken up in a minimum volume of water ($\sim$2 μL), and analyzed by the two-dimensional "wandering spot" method (Silberklang et al., 1977) using electrophoresis in 5% acetic acid, 0.5% pyridine, 2 mM Na_2EDTA, 7 M urea, pH 3.5, on cellulose acetate in the first dimension, elution transfer to DEAE-cellulose (Volkaert et al., 1976), and second dimension resolution by 65 °C homochromatography with 0.01 M KOH homochromatography mix.

Polyacrylamide gel electrophoretic analyses of the ^{32}P-labeled partial digestion products formed from STNV [5′-^{32}P]RNA by specific RNases were as described by Koper-

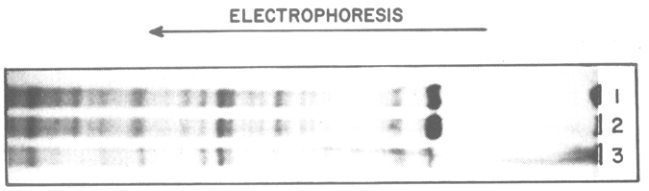

FIGURE 1: Radioautogram of a gel electrophoretic resolution of STNV [5′-^{32}P]RNA. Separate aliquots each containing 0.1 A_{260} unit of STNV RNA were dephosphorylated, heat denatured, ^{32}P labeled, and resolved as in Experimental Procedures. Track 1 (26 550 cpm in the major band) is an STNV RNA that was dephosphorylated but not heat denatured prior to labeling with ^{32}P. Track 2 (64 010 cpm in the major band, i.e., 30% of input STNV RNA converted to intact STNV [5′-^{32}P]RNA) represents the complete system. Track 3 (4860 cpm in the major band) is STNV RNA that was not treated with phosphomonoesterase prior to labeling with ^{32}P.

Zwarthoff et al. (1977) and Lockard et al. (1978).

Results

N-Terminal Amino Acid Sequence of STNV Coat Protein. Knowledge of the N-terminal amino acid sequence of a specific protein translated from a monocistronic mRNA allows the matching of nucleotide sequences predicted from the genetic code assignments of the N-terminal amino acids with observed nucleotides of the mRNA. Such matchings or overlap comparisons will characterize an AUG sequence as the initiator AUG of the mRNA. STNV RNA is a monocistronic mRNA coding only for STNV coat protein (Klein et al., 1972; Lundquist et al., 1972; Klein & Clark, 1973; Leung et al., 1976). Accordingly, we employed automated Edman degradation on STNV coat protein to increase our knowledge of the N-terminal sequence of STNV coat protein and to facilitate identification of the initiator AUG of STNV RNA.

Thin-layer chromatographic analyses of the Pth-amino acid derivatives derived from the organic phases of cycles 3 through 6 of an Edman degradation of STNV coat protein reveal a sequence Gln-Gln-Asn-Asn. The seventh amino acid from the N terminus is Arg as indicated by the lack of any detectable Pth-amino acid in the organic phase obtained in the procedure and a positive phenanthrenequinone dip test for Pth-Arg (Niall, 1973) from the aqueous phase obtained in this procedure. Combination of these data with the previously determined N-terminal sequence of STNV coat protein (Klein et al., 1972)

$$\text{Ala-Lys-}\begin{matrix}\text{Glu}\\\text{Gln}\end{matrix}\text{...}$$

establishes that the N-terminal sequence of STNV coat protein is Ala-Lys-Gln-Gln-Asn-Asn-Arg-····.

Characterization of STNV [5′-^{32}P]RNA. The preparation of a [5′-^{32}P]RNA for use in meaningful nucleotide sequence studies requires, in turn, (1) removal of any 5′-terminal polyphosphate or m^{7}G(5′)ppp constituents from the RNA; (2) ^{32}P labeling of the resultant 5′-hydroxyl-terminated RNA through use of polynucleotide kinase and [γ-^{32}P]ATP; (3) purification of the intact product [5′-^{32}P]RNA away from any 5′-^{32}P-terminated RNA fragments generated in the procedure; and (4) characterization of the homogeneity of the resultant intact [5′-^{32}P]RNA by characterization of various 5′-^{32}P-terminated nucleotides and oligonucleotides released from the [5′-^{32}P]RNA by exhaustive digestion with specific RNases (Lockard & RajBhandary, 1976; Silberklang et al., 1977; Koper-Zwarthoff et al., 1977). STNV RNA contains a 5′-terminal nucleotide sequence of AGU··· augmented with a 5′-terminal triphosphate, diphosphate, or monophosphate (Wimmer et al., 1968; Horst et al., 1971; Leung et al., 1976). We therefore treated STNV RNA with phosphomonoesterase and treated the product STNV 5′-dephospho-RNA (i.e., STNV 5′(OH)RNA) with [γ-^{32}P]ATP and T_4 polynucleotide kinase. As seen in Figure 1, such treatment yields a high percentage of the incorporated ^{32}P label in intact STNV [5′-^{32}P]RNA. Pretreatment of the STNV RNA with phosphomonoesterase enhances the incorporation of ^{32}P into the apparent intact STNV [5′-^{32}P]RNA by 5-fold. Interestingly, heat denaturation of STNV 5′-dephospho-RNA immediately before treatment with the kinase further enhances the incorporation of ^{32}P by an additional 2.5-fold, suggesting that the 5′ terminus of STNV RNA may be inaccessible due to some form of H-bonded secondary structure. Spermine and NaCl, at levels reported to enhance polynucleotide kinase activity (Lillehaug & Kleepe, 1975), inhibit the ^{32}P labeling of STNV 5′-dephospho-RNA by approximately 90%. Removal of the zone of the polyacrylamide slab gel that corresponds to the intact STNV [5′-^{32}P]RNA (Figure 1) followed by electrophoretic elution of the gel yields an alcohol precipitable STNV [5′-^{32}P]RNA. Exhaustive digestion of this STNV [5′-^{32}P]RNA with RNase T_2 yields [^{32}P]pAp as the only labeled product. Exhaustive digestions of the [^{32}P]RNA with RNase T_1 and RNase A yield 5′-^{32}P-labeled dinucleotide triphosphate and trinucleoside tetraphosphate, respectively, as the only labeled products. These results establish that the STNV [5′-^{32}P]RNA contains the correct and homogeneous 5′-terminal sequence of [^{32}P]pApGpUp···.

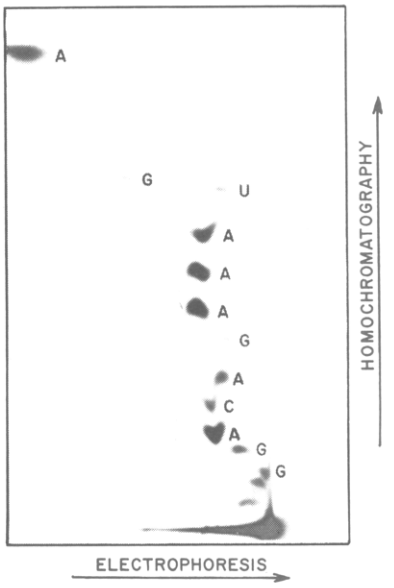

FIGURE 2: Radioautogram of a "wandering spot" analysis (Silberklang et al., 1977) of the 5′ terminus of STNV [5′-^{32}P]RNA.

Sequence Analysis of the 5′ Terminus of STNV [5′-^{32}P]RNA by Partial Digestion with Nuclease P_1. The availability of the STNV [5′-^{32}P]RNA containing a homogeneous 5′ terminus allows a sequence analysis of the first few nucleotides at the 5′ end of the STNV [5′-^{32}P]RNA by partial digestion of the STNV [5′-^{32}P]RNA with nuclease P_1 and subsequent "wandering spot", two-dimensional resolution (Silberklang et al., 1977). Figure 2 shows the results of such an analysis which characterizes the 5′-terminal 12 nucleotides of STNV [5′-^{32}P]RNA as AGUAAAGACAGG.

Further Sequence Analysis of STNV [5′-^{32}P]RNA by Polyacrylamide Gel Electrophoresis of Partial Digests. Polyacrylamide gel electrophoretic analysis of the ^{32}P-labeled oligonucleotides produced after partial digestions with specific nucleases (Donis-Keller et al., 1977; Simoncsits et al., 1977; Lockard et al., 1978) provides the most convenient way to characterize the sequence of additional nucleotides at the 5′

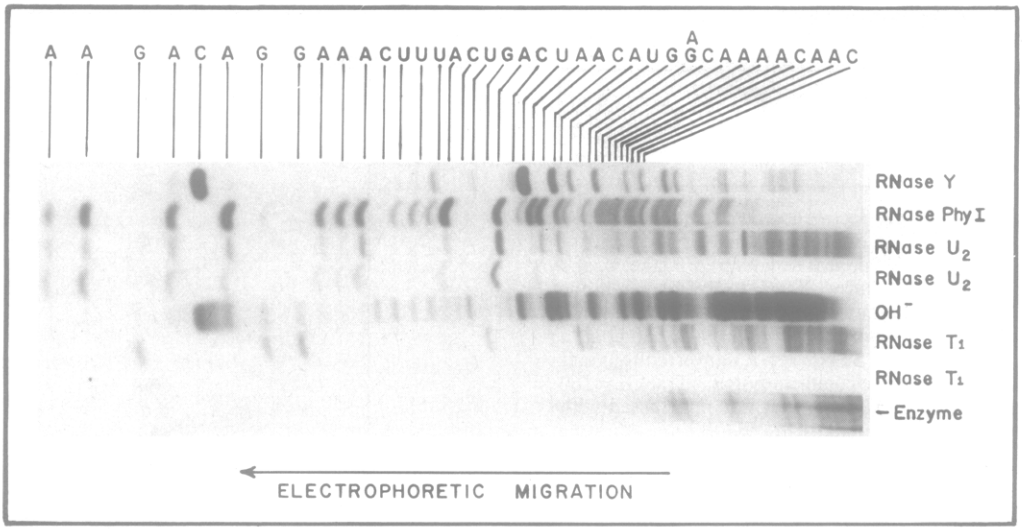

FIGURE 3: Radioautogram of the gel electrophoretic resolution of the ^{32}P-labeled partial digestion products derived from STNV [5′-^{32}P]RNA by specific chemical and enzymatic degradations (Lockard et al., 1978).

terminus of STNV [5′-^{32}P]RNA. Figure 3 shows the radioautographic results of such an analysis. In this particular experiment, the first four ^{32}P-labeled partial digestion products corresponding to the 5′-terminal sequence, AGUA, were run off of the gel in order to increase the resolution of the larger ^{32}P-labeled oligonucleotides. Also, the minus enzyme tract indicates a slight degree of degradation of the STNV [5′-^{32}P]RNA due to limited breakdown of the STNV [5′-^{32}P]-RNA after isolation. Even so, it is possible to use the nuclease specificities of the specific enzymes employed in this assay and overlap comparison with the "wandering spot" data (Figure 2) to read off the indicated 5′-terminal nucleotide sequence for STNV [5′-^{32}P]RNA from position 5 to position 42. The only ambiguity in these data occurs at position 33 which shows bands in both the T_1 and U_2 tracks indicating either a G or an A at this position. Repeat experiments similar to that of Figure 3 show this position to be a G.

Characterization of the Initiator AUG of STNV RNA. Translation of STNV RNA by eucaryotic systems involves incorporation of an initiator Met prior to assembly of the N-terminal amino acid sequence of STNV coat protein (Klein & Clark, 1973; Leung et al., 1976). The Ala-Lys-Gln-Gln-Asn-Asn-Arg sequence detected by Edman degradation analysis of STNV coat protein and the genetic code assignments for these amino acids therefore limit the nucleotide sequence possibilities of the translation initiation region of STNV RNA to (X signifies A, G, U, or C)

$$\text{AUGGC}\begin{matrix}\text{A}\\\text{G}\\\text{U}\\\text{C}\end{matrix}\text{AA}\begin{matrix}\text{A}\\\text{G}\end{matrix}\text{CA}\begin{matrix}\text{A}\\\text{G}\end{matrix}\text{CA}\begin{matrix}\text{A}\\\text{G}\end{matrix}\text{AA}\begin{matrix}\text{U}\\\text{C}\end{matrix}\text{AA}\begin{matrix}\text{U}\\\text{C}\end{matrix}\begin{matrix}\text{CGX}\\\text{AG}\begin{matrix}\text{A}\\\text{G}\end{matrix}\end{matrix}$$

Positions 30–32 of the sequence of Figure 3 contain the only AUG sequence detected in our characterization of the 5′-terminal sequence of STNV RNA. This AUG sequence is the initiator AUG of STNV RNA because the AUGGCAAAACAAC sequence of positions 30–42 of Figure 3 exactly matches a nucleotide sequence possibility predicted above for the first 13 nucleotides of the translation initiation region of STNV RNA.

Discussion

The Edman degradation data with STNV coat protein and the nucleotide sequence data of Figures 2 and 3 establish the nucleotide sequence and translation initiation character of the

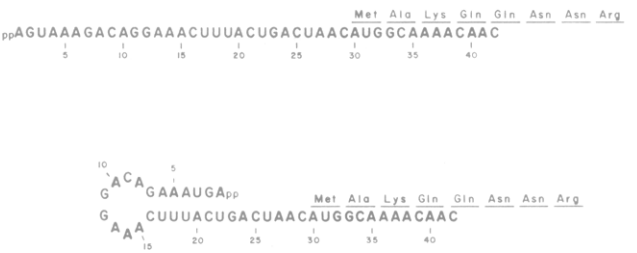

FIGURE 4: The 5′-terminal nucleotide sequence of STNV RNA.

first 42 nucleotides at the 5′ terminus of STNV RNA (Figure 4). Interestingly, the first seven nucleotides of this sequence are complementary to nucleotides 16–22 so as to form the indicated secondary structure. Theoretical evaluation of the free-energy change during formation of this hairpin loop (Tinoco et al., 1973) predicts $\Delta G°(25\ °C) = -5.8$ kcal/mol, indicating that this loop is stable under physiological conditions. The observed increase in polynucleotide kinase dependent ^{32}P labeling of STNV 5′-dephospho-RNA upon heat denaturation (Figure 1) and the observed inhibition of polynucleotide kinase dependent ^{32}P labeling by the secondary structure stabilizing agents, spermine and NaCl, further support the existence of this loop at the 5′ terminus of STNV RNA.

Our characterization of the 5′-terminal region of STNV RNA allows nucleotide sequence comparisons of STNV RNA with other eucaryotic mRNAs. The nucleotide sequence of the 5′ terminus and translation initiation site of STNV RNA (Figure 4) is similar to most eucaryotic mRNAs in that translation begins at the first AUG sequence in from the 5′ terminus. Interestingly, this initiator AUG is preceded by two out-of-phase terminator codons (UGA in positions 22–24, UAA in positions 26–28) that preclude any premature out-of-phase translation. The proposed secondary structure loop at the 5′ terminus of STNV RNA makes STNV RNA similar to BMV RNA 4 (Dasgupta et al., 1975) and the N, NS, and L mRNAs of VSV (Rose, 1978) in that the initiator AUG is close to the 5′ terminus of the RNA. STNV RNA can also be considered similar to the reovirus mRNAs (Kozak & Shatkin, 1977a, b) in that STNV RNA has a potential loop forming sequence between its 5′ terminus and its initiator AUG.

The 5′ terminus and associated translation initiation region of STNV RNA differ from other eucaryotic mRNAs in

several respects. STNV RNA does lack a 5′-terminal cap structure common to most eucaryotic mRNAs. STNV RNA is also different from several other eucaryotic mRNAs such as Rous sarcoma mRNA (Haseltine et al., 1977), rabbit globin β-chain mRNA (Baralle, 1977), and TMV RNA (Richards et al., 1977) in that these mRNAs feature large uncoded 5′-terminal regions with an apparent lack of secondary structure. STNV RNA is quite different from AlMV RNA 4 (Koper-Zwarthoff et al., 1977) in that, although STNV RNA and AlMV RNA 4 both contain a similar number of nucleotides in their noncoded 5′ termini, STNV RNA is pyrimidine-deficient and appears to contain secondary structure in this region, while AlMV RNA 4 is pyrimidine-rich and "relatively" unstructured in this noncoded region. However, one must recognize that it is quite possible that further studies of the secondary and tertiary structures of larger portions of all of these eucaryotic mRNAs may reveal common aspects of mRNA structure and function not evident in comparisons based upon the primary nucleotide sequences currently known.

The potential hairpin loop structure involving the 5′-terminal nucleotide of STNV RNA (Figure 4) is perhaps the most interesting feature of the STNV RNA nucleotide sequence reported here. Some bacterial viral mRNAs such as R17 RNA and MS2 RNA do feature potential loop structures that include their 5′-terminal nucleotide (Adams & Cory, 1970; DeWachter et al., 1971). Ovalbumin mRNA features a similar hairpin loop near its 5′ terminus (McReynolds et al., 1978). Yet, to date, no other eucaryotic mRNA is known to feature a potential 5′-terminal hairpin loop that includes the 5′-terminal nucleotide. STNV RNA is also the first uncapped eucaryotic mRNA whose 5′-terminal nucleotide sequence has been determined. It is therefore logical to question whether other uncapped eucaryotic mRNAs such as polio mRNA, EMC RNA, and cowpea mosaic virus RNA feature similar 5′-terminal secondary structures that include the 5′-terminal nucleotide.

The discovery of the potential hairpin loop at the exact 5′ terminus of STNV RNA also suggests several functional roles for this structure. First, it is possible that the 5′-terminal hairpin loop of STNV RNA acts as a substitute for the more common 5′-terminal cap structure of eucaryotic mRNAs during the formation of 40S ribosome translation initiation complexes. The formation of similar loop structures by reovirus mRNAs during such complex formations (Kozak & Shatkin, 1977a, b) dictates that loop structures do function in translation initiation in eucaryotes and therefore supports this hypothesis. Second, the 5′-terminal hairpin loop of STNV RNA serves to bring the 5′ terminus of STNV RNA closer to the initiator AUG. If joint ribosomal recognition of 5′ termini and initiator AUGs is necessary for eucaryotic mRNA function, the 5′-terminal hairpin loop of STNV RNA would favor such a process. Third, a hairpin loop at the exact 5′ terminus of STNV RNA may protect this uncapped mRNA from nucleolytic degradation by 5′-exonucleases. This theory is supported by the fact that double-stranded RNAs are more resistant to nucleases and by the observations (Shimotohno et al., 1977; Furuichi et al., 1977) that removal of 5′-terminal capping groups from mRNAs lacking a hairpin loop at their exact 5′ termini makes these mRNAs particularly susceptible to nucleolytic degradation. Fourth, the potential hairpin loop at the 5′ terminus of STNV RNA contains a feature which may account for the correct translation of STNV RNA in the procaryotic system. Specifically, the exposed portion of the hairpin loop contains the sequence AGGA (i.e., positions 10–13). This sequence is complementary to a pyrimidine-rich region at the 3′ end of the 16S rRNA of *E. coli* so as to facilitate mRNA:16S rRNA hybridization in the manner proposed for initiation complex formation by procaryotic systems (Shine & Dalgarno, 1974; Steitz & Jakes, 1975).

Acknowledgments

The authors acknowledge Dr. Thomas Baldwin for assistance with the Edman degradation procedure and Ronnie Barker's invaluable technical assistance.

References

Adams, J. M., & Cory, S. (1970) *Nature (London) 227*, 570.
Bailey, G. S., Gillett, D., Hill, D. F., & Peterson, G. B. (1977) *J. Biol. Chem. 252*, 2218.
Baralle, F. E. (1977) *Nature (London) 267*, 279.
Baralle, F. E., & Brownlee, G. G. (1978) *Nature (London) 274*, 84.
Clark, J. M., Jr., & Klein, W. H. (1974) *Methods Enzymol. 30*, 754.
Dasgupta, R., Shih, D. S., Saris, C., & Kaesberg, P. (1975) *Nature (London) 256*, 624.
DeWachter, R., Vandenberghe, A., Merregaert, J., Contreras, R., & Fiers, W. (1971) *Proc. Natl. Acad. Sci. U.S.A. 68*, 585.
Dolja, V. V., Negruk, V. I., Novikov, V. K., & Atabekov, J. G. (1977) *Anal. Biochem. 80*, 502.
Donis-Keller, H., Maxam, A. M., & Gilbert, W. (1977) *Nucleic Acid Res. 4*, 2527.
Furuichi, Y., Lafiandra, A., & Shatkin, A. J. (1977) *Nature (London) 266*, 235.
Gillum, A., Urquhart, N., Smith, M., & RajBhandary, U. L. (1975) *Cell 6*, 395.
Hagenbüchle, O., Santer, M., Steitz, J. A., & Mans, R. J. (1978) *Cell 13*, 551.
Haseltine, W. A., Maxam, A. M., & Gilbert, W. (1977) *Proc. Natl. Acad. Sci. U.S.A. 74*, 989.
Hewlett, M. J., Rose, J. K., & Baltimore, D. (1976) *Proc. Natl. Acad. Sci. U.S.A. 73*, 327
Horst, J., Fraenkel-Conrat, H., & Mandeles, S. (1971) *Biochemistry 10*, 4748.
Jeppsson, J., & Sjöquist, J. (1967) *Anal. Biochem. 18*, 264.
Klein, W. H., & Clark, J. M., Jr. (1973) *Biochemistry 12*, 1528.
Klein, W. H., Nolan, C., Lazar, J. M., & Clark, J. M., Jr. (1972) *Biochemistry 11*, 2009.
Klootwijk, J., Klein, I., Zabel, P., & van Kammen, A. (1977) *Cell 11*, 73.
Koper-Zwarthoff, E. C., Lockard, R. E., DeWeerd, B. A., RajBhandary, U. L., & Bol, J. F. (1977) *Proc. Natl. Acad. Sci. U.S.A. 74*, 5504.
Kozak, M., & Shatkin, A. J. (1977a) *J. Biol. Chem. 252*, 6895.
Kozak, M., & Shatkin, A. J. (1977b) *J. Mol. Biol. 112*, 75.
Kozak, M., & Shatkin, A. J. (1978) *Cell 13*, 201.
Leung, D. W., Gilbert, C. W., Smith, R. E., Sasavage, N. L., & Clark, J. M., Jr. (1976) *Biochemistry 15*, 4943.
Li, S. L., & Riggs, A. (1970) *J. Biol. Chem. 245*, 6149.
Lillehaug, J. R., & Kleppe, K. (1975) *Biochemistry 14*, 1225.
Lockard, R. E., & RajBhandary, U. L. (1976) *Cell 9*, 747.
Lockard, R. E., & Lane, C. D. (1978) *Nucleic Acid Res. 5*, 3237.
Lockard, R. E., DeWeerd, B. A., Heckman, J. E., MacGee, J., Tabor, M. W., & RajBhandary, U. L. (1978) *Nucleic Acid Res. 5*, 37.
Lundquist, R. E., Lazar, J. M., Klein, W. H., & Clark, J. M., Jr. (1972) *Biochemistry 11*, 2014.

Maxam, A. M., & Gilbert, W. (1977) *Proc. Natl. Acad. Sci. U.S.A. 74*, 560.
McReynolds, L., O'Malley, B. W., Nisbet, A. D., Fothergill, J. E., Girol, D., Fields, S., Robertson, M., & Brownlee, G. G. (1978) *Nature (London) 273*, 723.
Niall, H. D. (1973) *Methods Enzymol. 27*, 942.
Nomoto, A., Kitamura, N., Golini, F., & Wimmer, E. (1977) *Proc. Natl. Acad. Sci. U.S.A. 74*, 5345.
Nuss, D. L., Furuichi, Y., Koch, G., & Shatkin, A. J. (1975) *Cell 6*, 21.
Palmiter, R. D. (1974) *Biochemistry 13*, 3606.
Philly, D., Niemeyer, A., Schmidt, M., & Bargetzi, J. P. (1978) *J. Biol. Chem. 253*, 437.
Richards, K., Guilley, H., Jonard, G., & Keith, G. (1977) *Nature (London) 267*, 548.
Rose, J. K. (1977) *Proc. Natl. Acad. Sci. U.S.A. 74*, 3672.
Rose, J. K. (1978) *Cell 14*, 345.
Shatkin, A. J. (1976) *Cell 9*, 645.
Shimotohno, K., Kodama, Y., Hashimoto, J., & Miura, K.-I. (1977) *Proc. Natl. Acad. Sci. U.S.A. 74*, 2734.
Shine, J., & Dalgarno, L. (1974) *Proc. Natl. Acad. Sci. U.S.A. 71*, 1342.
Silberklang, M., Gillum, A. M., & RajBhandary, U. L. (1977) *Nucleic Acid Res. 4*, 4097.
Simoncsits, A., Brownlee, G. G., Brown, R. S., Rubin, J. R., & Guilley, H. (1977) *Nature (London) 269*, 833.
Steitz, J. A., & Jakes, K. (1975) *Proc. Natl. Acad. Sci. U.S.A. 72*, 4734.
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.
Volkaert, G., Min Jou, W., & Fiers, W. (1976) *Anal. Biochem. 72*, 433.
Wimmer, E., Chang, A. Y., Clark, J. M., Jr., & Reichmann, M. E. (1968) *J. Mol. Biol. 38*, 59.

Effect of Capping upon the mRNA Properties of Satellite Tobacco Necrosis Virus Ribonucleic Acid†

Robert E. Smith‡ and John M. Clark, Jr.*

ABSTRACT: The mRNA guanylyltransferase–mRNA methyltransferases of vaccinia virions can be used to introduce a 5′-terminal m^{7}G(5′)ppp(5′)Ap^m··· capping group onto the RNA of satellite tobacco necrosis virus (STNV RNA) to yield intact capped STNV RNA. Studies with an in vitro system from wheat germ and limiting quantities of capped and uncapped STNV RNA show that the rates and extents of formation of initiation complexes of protein synthesis by intact capped and uncapped STNV RNA are identical, suggesting that 5′-terminal cap groups cannot function in the translation of STNV RNA. Also, the cap analogue pm^{7}G equally inhibits the initiation and the translation of limiting quantities of both capped and uncapped STNV RNA. These contrasting observations suggest that the wheat germ system contains a pm^{7}G sensitive protein and that STNV RNA has a tertiary structure that restricts the function of an added 5′-terminal capping group. This theory is supported by observations that fragmented capped STNV RNA is better at forming initiation complexes than is equally fragmented uncapped STNV RNA.

The RNA of satellite tobacco necrosis virus (STNV)[1] is a monocistronic mRNA coding only for STNV coat protein in in vitro systems from both procaryotes and eucaryotes (Klein et al., 1972; Lundquist et al., 1972; Klein & Clark, 1973; Leung et al., 1976). In vivo STNV RNA lacks a 5′-terminal capping group (Wimmer et al., 1968; Horst et al., 1971; Leung et al., 1976) and does not acquire or require a 5′-terminal capping group during its translation by an in vitro system from wheat germ (Leung et al., 1976). The purified mRNA guanylyltransferase–mRNA methyltransferases of vaccinia virions introduce a m^{7}G(5′)ppp(5′)Ap^m···5′-terminal capping group onto STNV RNA (Moss, 1977). This property allows assay of the impact of the 5′-terminal capping group upon the translation of STNV RNA. This paper reports comparisons of the mRNA potential of STNV RNA with 5′-terminally capped STNV RNA in an in vitro system from wheat germ.

Experimental Procedures

Materials. STNV RNA, STNV [^{125}I]RNA, and cell-free S-30 extracts from wheat germ were prepared as previously described (Leung et al., 1976). P_1 nuclease and various 5′-terminal fragment standards of capped mRNAs were obtained from P-L Biochemicals, Inc. [α-^{32}P]GTP, a mixture of 15 U-^{14}C-labeled L-amino acids, carrier-free Na^{125}I, and *S*-adenosyl[*methyl*-^{3}H]methionine were purchased from Amersham/Searle. Unlabeled *S*-adenosyl-L-methionine (Sigma

† From the Biochemistry Department, University of Illinois, Urbana, Illinois 61801. *Received December 19, 1978.* This work was supported in part by U.S. Public Health Service Grant GM-08647.

‡ Present address: Roche Institute of Molecular Biology, Nutley, NJ 07110.

[1] Abbreviations used: STNV, satellite tobacco necrosis virus; m^{7}G(5′)ppp(5′)N, m^{7}G(5′)ppp(5′)Np, m^{7}G(5′)ppp(5′)A^m, and m^{7}G(5′)ppp(5′)Ap^m, 7-methylguanosine linked by its 5′ position through three phosphate anhydrides to the 5′ position of a nucleoside, a nucleoside 3′-phosphate, a 2′-*O*-methyladenosine, or a 2′-*O*-methyladenosine 3′-phosphate, respectively; capping groups, any 5′-terminal m^{7}G(5′)ppp(5′)Np··· group on RNA; capped RNA, any RNA containing a 5′-terminal m^{7}G(5′)ppp(5′)Np··· group; pm^{7}G, 7-methylguanosine 5′-phosphate; ATP and GTP, adenosine and guanosine 5′-triphosphates; PEI-cellulose, cellulose pretreated with polyethylenimine; KMH buffer, 0.1 M KCl, 3 mM $Mg(OAc)_2$, 2 mM dithiothreitol, 20 mM Hepes–K salt, pH 7.2; KMT buffer, 0.1 M KCl, 5 mM $Mg(OAc)_2$, 10 mM Tris–OAc, pH 7.5; SAM, *S*-adenosylmethionine.

0006-2960/79/0418-1366$01.00/0