Proc. Nat. Acad. Sci. USA
Vol. 71, No. 10, pp. 3979–3983, October 1974

The 5′-Terminal Nucleotide Sequences of the Double-Stranded RNA of Human Reovirus

(modified 5′ termini)

KIN-ICHIRO MIURA, KUMIKO WATANABE, MASAHIRO SUGIURA, AND AARON J. SHATKIN*

National Institute of Genetics, Mishima, 411 Japan; and * Roche Institute of Molecular Biology, Nutley, New Jersey 07110

Communicated by B. L. Horecker, July 29, 1974

ABSTRACT **The 5′-terminal nucleotide sequences of human reovirus double-stranded RNA were determined after labeling the RNA with [^{32}P]phosphate by polynucleotide kinase. The 5′ termini were labeled to only a limited extent prior to sequential oxidation, β-elimination, and phosphomonoesterase treatment, indicating that the terminal phosphates were in a modified, blocked configuration. Each genome segment, after removing the blocking group, contained the same two 5′-terminal sequences: GpApUp in one chain and G*pCp in the other. G*p is a derivative of guanylic acid, probably 2′-*O*-methyl-Gp, which renders the 5′-terminal sequence resistant to hydrolysis by alkali. The results indicate that the transcription of reovirus double-stranded RNA starts from the 3′ end complementary to the G*pCp-terminal, resulting in the synthesis of single-stranded mRNA carrying the same 5′ sequence as the G*pCp-chain. The presence of a modified nucleotide at the 5′ terminus of the strand complementary to the mRNA template is a feature common to another double-stranded RNA virus, cytoplasmic polyhedrosis virus.**

The double-stranded RNA genome of cytoplasmic polyhedrosis (CP) virus from silkworm consists of 10 segments of different size (1). All 10 RNA segments have the same 5′-terminal structures: one strand of each duplex starts with guanylic acid and the other with a modified adenylic acid (2). CP virus carries transcriptase in its coat protein (3, 4), and synthesizes mRNA by copying only one chain of each double-stranded RNA segment from the end where the modified nucleotide exists in the complementary strand (2, 5, 6). Reoviruses also contain 10 double-stranded RNA segments (1, 7) and a transcriptase (8, 9). It was, therefore, of interest to determine if structures similar to those in the RNA of CP virus are also present in the reovirus genome.

Experiments with uniformly ^{32}P-labeled reovirus RNA indicated that the 5′ termini contain ppGpPyp- (10). However, the 5′-terminal structure appears to be more complicated because native reovirus RNA, with or without phosphomonoesterase pretreatment, was scarcely labeled with [^{32}P]phosphate by the polynucleotide kinase procedure. A similar finding was observed with CP virus RNA (2), but after sequential oxidation, β-elimination, and phosphomonoesterase digestion, both 5′ ends of the CP virus RNA were successfully labeled with [^{32}P]phosphate. These treatments were also necessary to obtain reovirus double-stranded RNA labeled specifically at the 5′ termini with [^{32}P]phosphate. In this report, the 5′-terminal nucleotide sequences of reovirus RNA are described, and the possible relationship between their structure and template function in mRNA transcription is discussed.

Abbreviations: CP, cytoplasmic polyhedrosis; Pu and Py, unspecified purine and pyrimidine nucleosides, respectively.

MATERIALS AND METHODS

The procedure for the multiplication and purification of human reovirus type 3 and phenol extraction of viral RNA have been described previously (10). The viral RNA was passed through a column of Sephadex G-100 to separate the double-stranded RNA from the adenine-rich single-stranded oligonucleotides (10). To prepare the double-stranded RNA for labeling with [^{32}P]phosphate by polynucleotide kinase, two cycles of oxidation with periodate, β-elimination with aniline, and phosphomonoesterase treatment were necessary (2). Methods for the 5′-terminal nucleotide sequence analysis of the [5′-^{32}P]RNA and other procedures for the characterization of RNA have been published previously (2).

RESULTS

Labeling of RNA at the 5′ termini

Double-stranded reovirus RNA was labeled with ^{32}P at the 5′ termini with polynucleotide kinase as described for CP virus RNA (2). The 5′-labeling efficiency was calculated to be about 80%, and the 10 ^{32}P-labeled reovirus genome segments were indistinguishable from native reovirus RNA (7) by electrophoretic analysis in a 3% polyacrylamide gel (Fig. 1). Extensive internal scissions in the polynucleotide chains of the double-stranded RNA did not occur during these treatments, as shown by the absence of low-molecular-weight RNA in the sedimentation profile of denatured ^{32}P-labeled RNA.

Analysis of the 5′-terminal nucleotide sequences

An alkaline hydrolysate of the [5′-^{32}P]RNA of reovirus yielded two radioactive peaks corresponding to ^{32}pXp (net charge: −4) and ^{32}pYpZp (net charge: −5) when analyzed by DEAE-cellulose–7 M urea chromatography (Fig. 2a). Similar results were also obtained by digestion with ribonuclease T_2, which leads to complete hydrolysis of unmodified RNA to nucleoside-3′-phosphates (Fig. 2b). Dowex-1 chromatography of alkaline and T_2 digests (Fig. 3a and b) yielded a radioactive component coinciding with authentic pGp. By two-dimensional paper chromatography, two major radioactive spots were also detected (Fig. 4a). One of them is completely superimposed on authentic pGp, but the other does not correspond to any known mononucleoside diphosphate or the ATP that was used as the polynucleotide kinase substrate. It is concluded from these results that the ^{32}pXp

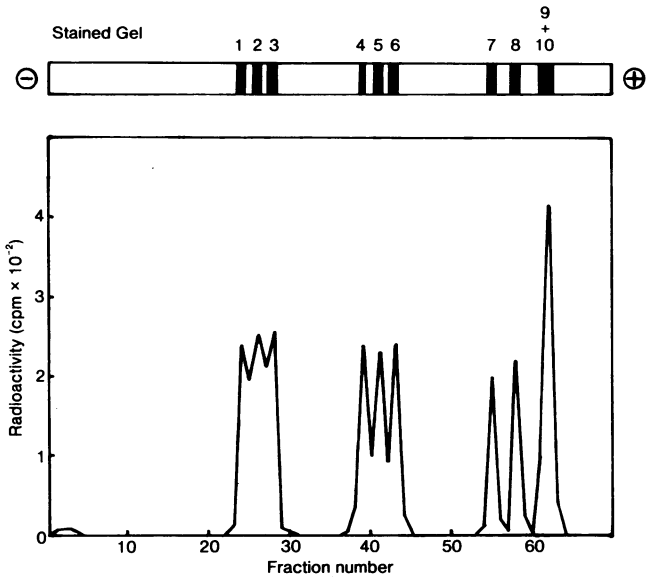

FIG. 1. Polyacrylamide gel electrophoresis of ^{32}P-labeled reovirus RNA. 5′-^{32}P-labeled RNA (0.5 A_{260} units) was applied to a 3% polyacrylamide gel column (0.7 cm × 14 cm) and electrophoresed at 6 mA for 14 hr. The gel was stained with acridine orange and then analyzed for radioactivity.

in the DEAE-cellulose elution profile corresponds to ^{32}pGp and that the other radioactive component is a dinucleoside triphosphate which contains an alkaline-resistant nucleotide at the 5′ end. In the alkaline and ribonuclease T_2 hydrolysates, a small amount of ^{32}pAp was also observed (Fig. 3a and b and Fig. 4a). Although the origin of this component is not clear, it appears to be a contamination, possibly from residual viral oligonucleotides (10), since reduced and variable levels were observed in the digests of individual RNA segments.

A nuclease from *Penicillium citrinum* hydrolyzes denatured nucleic acids to a mixture of 5′ nucleotides, even at linkages containing an alkaline-resistant nucleotide (2). Digestion of 5′-^{32}P-labeled reovirus RNA with this enzyme gave only radioactive 5′ nucleotides of net charge −2 by DEAE-cellulose chromatography (Fig. 2c). On Dowex-1 chromatography, most of the radioactivity eluted in one heterogeneous peak which is partially superimposed with pG (Fig. 3c). ^{32}pG and a modified derivative (^{32}pG*) appear to be the main radioactive components in the *Penicillium* nuclease digest of the 5′-^{32}P-labeled reovirus RNA. These two components were separated by two-dimensional paper chromatography (Fig. 4b). The spot of pG was radioactive, and the other highly radioactive spot was located near, but not completely superimposed with, pC. This component is probably a derivative of pG, not pC, judging from (a) its behavior in other chromatographic systems and (b) the results with pancreatic ribonuclease digests described below. A small amount of ^{32}pA (less than 10% of the total radioactivity) is also observed in the *Penicillium* nuclease digest, again presumably contaminating materials.

Ribonuclease T_1, which splits RNA at the 3′ phosphate of guanylic acid residues, hydrolyzed the 5′-^{32}P-labeled reovirus RNA to yield ^{32}pGp and some longer ^{32}P-labeled nucleotides (Fig. 2d).

The 5′-terminally labeled reovirus RNA was digested with pancreatic ribonuclease A, which splits RNA at the 3′ phosphate of pyrimidine nucleotides. As shown in Fig. 5a, the radioactive peaks appeared as dinucleoside triphosphate ^{32}pPupPyp (net charge: −5) and trinucleoside tetraphosphate ^{32}pPupPupPyp (net charge: −6) in a DEAE-cellulose–urea column. When the 5′-^{32}P-labeled reovirus RNA was digested with both pancreatic ribonuclease A and ribonuclease T_1, ^{32}pGp (net charge: −4) and 5′-^{32}P-dinucleoside triphosphate

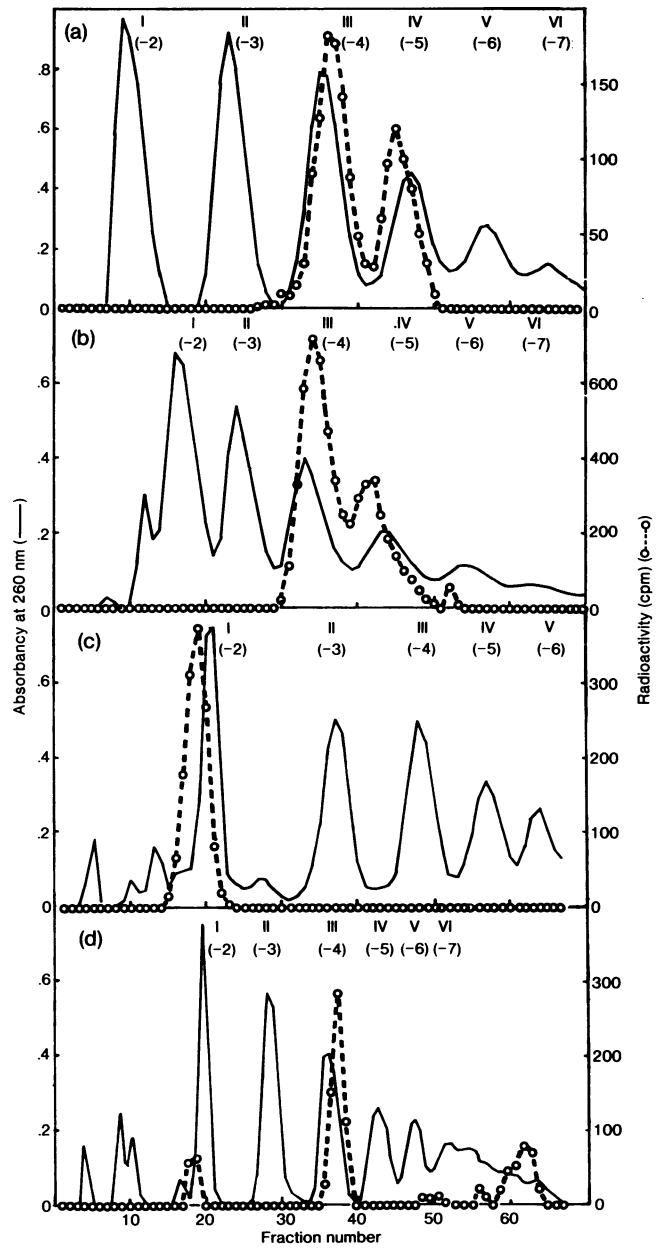

FIG. 2. DEAE-urea column chromatography of hydrolyzed 5′-^{32}P-labeled reovirus RNA. (a) alkaline digest, (b) ribonuclease T_2 digest, (c) *Penicillium* nuclease digest, (d) ribonuclease T_1 digest. Digests were chromatographed on a DEAE-cellulose column (0.6 cm × 40 cm) with the pancreatic ribonuclease digest of ribosomal RNA included as a marker. Elution was carried out with a linear gradient of NaCl from 0 M (100 ml) to 0.4 M (100 ml) in 7 M urea–0.02 M Tris·HCl (pH 7.6–8.0). *Roman numerals* over peaks indicate the chain length of nucleotides, and *numerals in parentheses* indicate net charge. Experimental details are the same as in ref. 2.

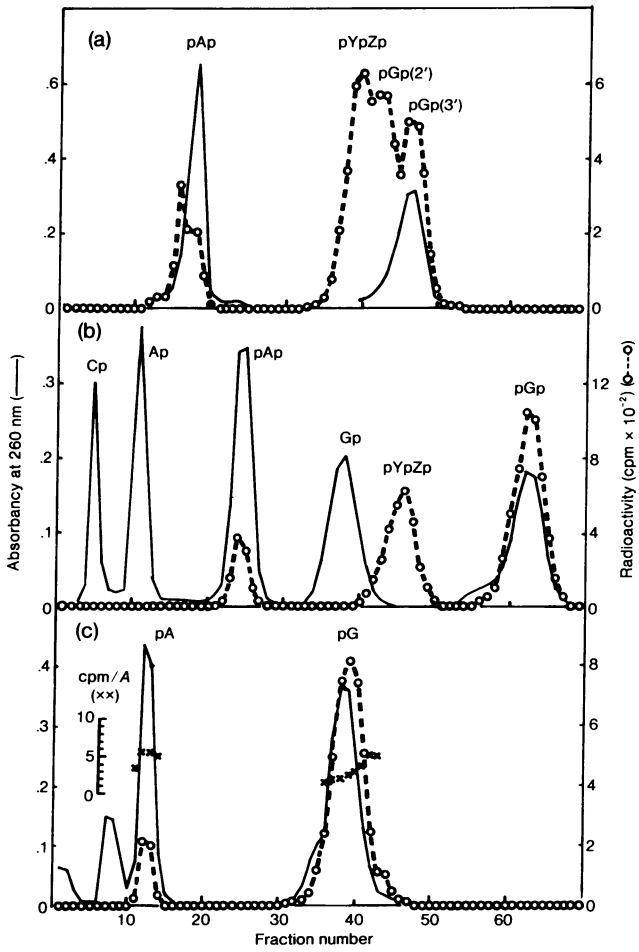

FIG. 3. Dowex-1 chromatography of hydrolyzed 5′-^{32}P-labeled reovirus RNA. (a) alkaline digest, (b) ribonuclease T_2 digest, (c) *Penicillium* nuclease digest. Digests were chromatographed on a Dowex-1 column (Cl^- form, 0.6 cm × 25 cm) with adenosine 3′,5′-diphosphate and guanosine 3′,5′-diphosphate (Boehringer Mannheim Co.) as authentic markers. Elution was carried out with a linear gradient from 0 to 0.25 M NaCl (50 ml each) in 0.01 N HCl. In (b), absorbancy includes measurement of nucleoside 3′-monophosphates.

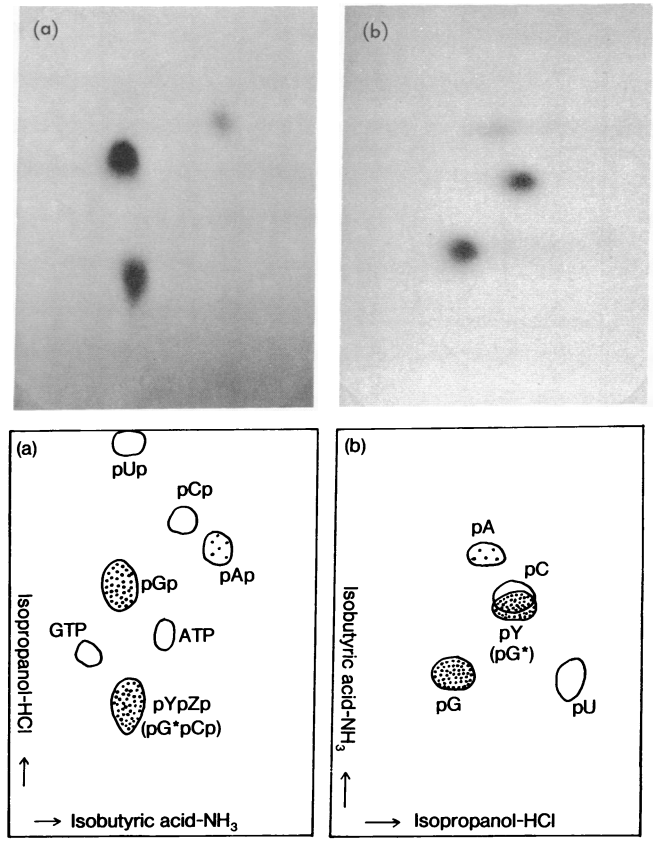

FIG. 4. Radioautograph of the two-dimensional paper chromatogram of the hydrolysis products of 5′-^{32}P-labeled reovirus RNA. (a) ribonuclease T_2 digest, (b) *Penicillium* nuclease digest. In the key diagrams (*lower panel*), the positions of the markers, identified by ultraviolet absorption, are shown by a *solid line* and of the radioactivity by the *dotted areas*.

(net charge: −5) were obtained (Fig. 5b). Considering the results with the ribonuclease T_1 digest (Fig. 2d), a long nucleotide in a T_1 digest would start with ^{32}pG*pPyp. Therefore, the dinucleotide in the pancreatic ribonuclease digest probably is ^{32}pG*pPyp, and the trinucleotide is ^{32}pGpPupPyp. To confirm this, the radioactive dinucleotide and trinucleotide in the pancreatic ribonuclease A digest were separated by DEAE-cellulose–urea chromatography, desalted, hydrolyzed with alkali and with *Penicillium* nuclease, respectively, and analyzed by DEAE-cellulose–urea and Dowex-1 column chromatography. The dinucleotide, ^{32}pPupPyp, was not split by alkali, but ^{32}pG* was obtained from it by treatment with *Penicillium* nuclease. The radioactive component in the ^{32}pPupPupPyp was split by alkali to give ^{32}pPup, which was identified as ^{32}pGp; i.e., the trinucleotide is ^{32}pGpPupPyp. The oligonucleotides, ^{32}pG*pPyp and ^{32}pGpPupPyp, obtained from the pancreatic ribonuclease A digest were then analyzed by Dowex-1 column chromatography. The elution positions of dinucleotides and trinucleotides were determined by comparing relative positions with added markers according to Takanami (11, 12). Fig. 6 shows that the dinucleotide corresponds to pGpCp and the trinucleotide to p(Ap,Gp)Up. The dinucleotide contains the alkali-resistant derivative of guanylic acid at the 5′ side, i.e., it is ^{32}pG*pCp. Since the trinucleotide carries guanylic acid at the 5′-terminal, it is identified as ^{32}pGpApUp. Small amounts of radioactivity are also present in the positions of pApUp and pGpGpCp.

Analysis of the separated RNA segments

Reovirus genome RNA can be separated into 10 distinct segments by electrophoresis in a 3% polyacrylamide gel (1, 7). The 5′-terminal nucleotides of the separated segments were analyzed as follows. ^{32}P-Labeled RNA was applied to a 3% gel and after electrophoresis stained with acridine orange (Fig. 1). Individual stained bands were cut out with a razor blade, crushed with a Teflon pestle, and incubated with 1 N KOH at 37° for 18 hr. After neutralization with concentrated $HClO_4$ and removal of the $KClO_4$ precipitate by centrifugation, the supernatant fluid containing essentially all of the ^{32}P-labeled nucleotides was chromatographed on a DEAE-cellulose–urea column. Samples from each of the 10 genome segments after alkaline digestion showed the same elution pattern as the whole mixture of genome segments (compare Figs. 2a and 5c). Two main radioactive peaks were detected: ^{32}pXp = ^{32}pGp and ^{32}pYpZp = ^{32}pG*pCp. The results in-

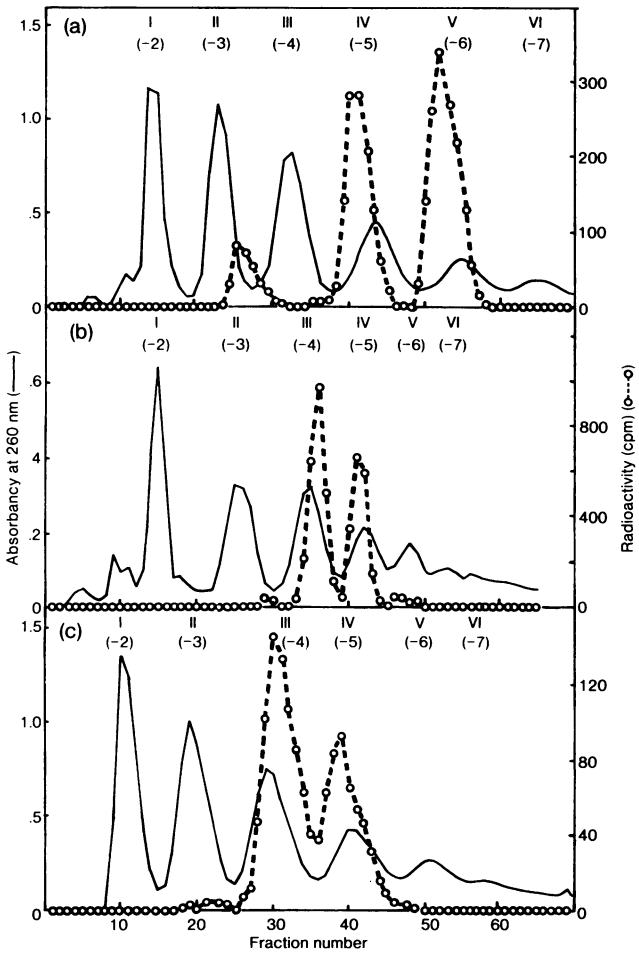

FIG. 5. DEAE-urea column chromatography of digested 5'-^{32}P-labeled reovirus RNA. (a) pancreatic ribonuclease A digest, (b) pancreatic ribonuclease A plus ribonuclease T_1 digest, (c) alkaline digest of a separated reovirus genome segment (band no. 1 in the gel shown in Fig. 1). Conditions were the same as those in Fig. 1.

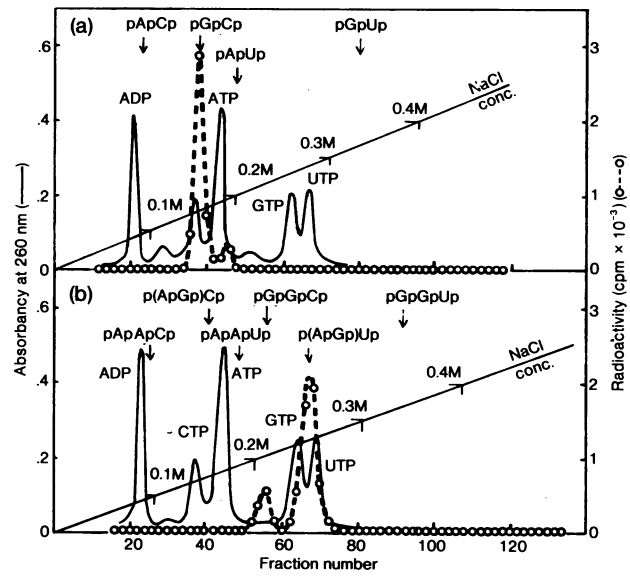

FIG. 6. Dowex-1 column chromatography of the dinucleotides and trinucleotides in the pancreatic ribonuclease A digest. The 5'-terminal ^{32}P-labeled dinucleoside triphosphate (a) and trinucleoside tetraphosphate (b) fractions in the pancreatic ribonuclease A digest, separated by DEAE-urea column chromatography as shown in Fig. 5a, were desalted, and chromatographed on a Dowex-1 column with CDP, ADP, CTP, GTP, and UTP as markers. The *arrows* in the figure indicate the peak positions of dinucleoside triphosphate and trinucleoside tetraphosphate estimated relatively by the positions of the added markers according to references 11 and 12. The column size was 0.55 cm × 20 cm, and elution was carried out with a linear gradient of 0–0.4 M NaCl (200 ml each) in 0.01 N HCl.

dicate that every RNA genome segment in reovirus contains two common sequences: ^{32}pGpApUp... and ^{32}pG*-pCp.... Although the ratio of the amounts of ^{32}pGp and ^{32}pG*pCp is constant for each individual segment, the former is 1.5 times greater than the latter. This may reflect a difference in the ability of the two 5′ termini to accept [^{32}P]phosphate.

DISCUSSION

Human reovirus type 3 contains 10 segments of double-stranded RNA (1, 7). Analysis of the 5'-terminal nucleotides of each of the 10 genome segments separated by gel electrophoresis showed that each has two common types of 5'-terminal structures: one strand of the double-stranded RNA has ^{32}pGp at the 5′ end and the other has ^{32}pG*pCp-. There was always 1.5 times more ^{32}pGp than ^{32}pG*pCp in alkali digests of each RNA segment. This ratio would depend on the efficiency of ^{32}P-labeling at the two different 5′ termini. Although alkaline digests of either total double-stranded RNA or separated genome segments contained some minor components including pAp, it is suggested that each segment of reovirus double-stranded RNA contains two kinds of 5′ terminal sequences: GAU and G*C-.

Reovirus double-stranded RNA accepted only small amounts of [^{32}P]phosphate by the polynucleotide kinase procedure, either with or without pretreatment of the RNA by phosphomonoesterase prior to phosphorylation. After the serial operations, which consisted of periodate oxidation, β-elimination by aniline treatment, and removal of phosphate by phosphomonoesterase digestion, there was extensive incorporation of [^{32}P]phosphate into the 5′ termini, indicating that reovirus RNA resembles CP virus double-stranded RNA (2).

There is evidence that completely double-stranded DNA can be labeled with [^{32}P]phosphate at the 5′ termini by polynucleotide kinase (13, 14). The results of studies with the single strand specific nuclease S_1 indicate that the 5′ and 3′ termini of reovirus genome RNA are base-paired (S. Muthukrishnan and A. J. Shatkin, unpublished results). Thus, the limited labeling of reovirus RNA by the polynucleotide kinase technique before oxidation suggests that the 5′ termini are masked by a constituent that prevents removal of the 5′ phosphates by phosphomonoesterase prior to oxidation and β-elimination.

One of the 5′ termini of reovirus RNA contains an alkali-resistant guanylic acid derivative (G*p) which presumably is modified by methylation at the 2′ position of ribose, as found for reovirus mRNA (15). A single radioactive peak (−5 charge) was observed previously (10) in an alkaline digest of uniformly ^{32}P-labeled reovirus RNA analyzed by DEAE-cellulose–urea column chromatography. This peak is probably a mixture of ppGp and XppG*pCp (both −5 charge), where

X denotes a blocking group, i.e., a nucleoside or sugar, that is removed during the pretreatment before the 5′-labeling with polynucleotide kinase. The presence of a similar structure, XppAmpGpUp, at the 5′ terminus in CP virus mRNA and in one chain of the double-stranded RNA of CP virus genome has also been found (Y. Furuichi and K. Miura, unpublished results). A 5′,5′-pyrophosphate linkage, $m_3^{2,2,7}$G5′p-5′pAmp, has been reported recently for the 5′ termini of several low-molecular-weight RNAs from Novikoff hepatoma ascites cells (16). The 5′-terminal masking of reovirus (and CP virus) RNA may be due to the presence of a similar structure. Reovirus genome RNA uniformly labeled with [^{32}P]phosphate also appears to consist of a plus strand containing a methylated 5′ end with an alkali-stable blocking group (XppG*pCp) and an unmethylated minus strand which yields an unblocked 5′-terminal ppGp upon alkaline hydrolysis (N. L. Chow and A. J. Shatkin, unpublished results). Further studies will be necessary to elucidate the nature of the viral RNA 5′ termini.

Reovirus cores prepared by chymotrypsin digestion of purified virions contain a transcriptase which copies one of the strands of the genome RNA (8, 9). The 5′-terminal structure of the resulting mRNA mixture was ppGpPyp- (17). Nichols *et al.* (18) analyzed a long initiated sequence in the small size class of mRNA and obtained 5′-(p)ppGpCp.... Our results indicate that the 5′-terminal sequences of all 10 double-stranded RNA segments of reovirus genome are G*pCp-... and GpApUp-..., the former corresponding to the mRNA strand (15). Analyses of reovirus genome RNA labeled specifically at the 3′ termini with [^{3}H]borohydride are consistent with a base-paired structure containing 3′-Cp-Gp-... (the template strand) and 3′-Cp-Up-... (the mRNA strand) (S. Muthukrishnan and A. J. Shatkin, unpublished results). The modification of a nucleotide residue located in viral genome RNA at the end where initiation of transcription occurs suggests that the modified residue may be required as a signal for mRNA synthesis. In view of the finding that reovirus genome RNA is synthesized by formation of a complementary (−) strand on a template mRNA (+) strand (19), another attractive, possible role of the modification is the conversion of single-stranded mRNA to template RNA in reovirus-infected cells, a process which may occur during a specific time interval of the viral replicative cycle. Thus, the initiation site for transcription of the reovirus double-stranded genome RNA is distinguished not only by nucleotide sequence, but also by modification of the 5′-terminal nucleotide. This same relationship between genome structure and transcription is found in CP virus of insects (2, 6), another double-stranded RNA-containing virus.

K.M. and A.J.S. are indebted to Dr. Norman P. Salzman of the National Institute of Allergy and Infectious Diseases, National Institutes of Health, for facilities at the early stage of this work. We thank Dr. Yasuhiro Furuichi for helpful discussion. Part of this work was supported by grants from the Japanese Ministry of Education and from Naito Foundation.

1. Fujii-Kawata, I., Miura, K. & Fuke, M. (1970) *J. Mol. Biol.* **51**, 247–253.
2. Miura, K., Watanabe, K. & Sugiura, M. (1974) *J. Mol. Biol.* **86**, 31–48.
3. Lewandowski, L. J., Kalmakoff, J. & Tanada, Y. (1969) *J. Virol.* **4**, 857–865.
4. Shimotohno, K. & Miura, K. (1973) *J. Biochem.* **74**, 117–125.
5. Shimotohno, K. & Miura, K. (1973) *Virology* **53**, 283–286.
6. Shimotohno, K. & Miura, K. (1974) *J. Mol. Biol.* **86**, 21–30.
7. Shatkin, A. J., Sipe, J. D. & Loh, P. (1968) *J. Virol.* **2**, 986–991.
8. Shatkin, A. J. & Sipe, J. D. (1968) *Proc. Nat. Acad. Sci. USA* **61**, 1462–1469.
9. Borsa, J. & Graham, A. F. (1968) *Biochem. Biophys. Res. Commun.* **33**, 895–901.
10. Banerjee, A. K. & Shatkin, A. J. (1971) *J. Mol. Biol.* **61**, 643–653.
11. Takanami, M. (1967) *J. Mol. Biol.* **23**, 135–148.
12. Takanami, M. (1967) *Seikagaku* (*Biochemistry*, in Japanese) **39**, 455–462.
13. Richardson, C. C. (1965) *Proc. Nat. Acad. Sci. USA* **54**, 158–165.
14. Richardson, C. C. (1966) *J. Mol. Biol.* **15**, 49–61.
15. Shatkin, A. J. (1974) *Proc. Nat. Acad. Sci. USA* **71**, 3204–3207.
16. Ro-Choi, T. S., Reddy, R., Choi, Y. C., Raj, N. B. & Henning, D. (1974) *Fed. Proc.* **33**, 1548.
17. Banerjee, A. K., Ward, R. & Shatkin, A. J. (1971) *Nature* **230**, 169–172.
18. Nichols, J. L., Hay, A. J. & Joklik, W. K. (1972) *Nature* **235**, 105–107.
19. Schonberg, M., Silverstein, S. C., Levin, D. H. & Acs, G. (1971) *Proc. Nat. Acad. Sci. USA* **68**, 505–508.