JOURNAL OF VIROLOGY, Nov. 1982, p. 538–543
0022-538X/82/110538-06$02.00/0

Vol. 44, No. 2

Homologous Terminal Sequences in the Double-Stranded RNA Genome Segments of Cytoplasmic Polyhedrosis Virus of the Silkworm *Bombyx mori*

YOSHIYUKI KUCHINO,[1] SUSUMU NISHIMURA,[1] ROBERT E. SMITH,[2] AND YASUHIRO FURUICHI[2]*

National Cancer Center Research Institute, Tsukiji 5-1-1, Chuo-ku, Tokyo, Japan,[1] and Roche Institute of Molecular Biology, Nutley, New Jersey 07110[2]

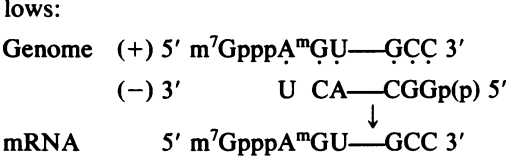
Received 24 June 1982/Accepted 6 August 1982

The 3'-terminal regions (20 to 32 residues) of the genome double-stranded RNA (dsRNA) segments of cytoplasmic polyhedrosis virus were sequenced. The dsRNAs, which were labeled at their 3' termini by incubation with [5'-^{32}P]pCp and T4 RNA ligase, were denatured and resolved into the plus and minus strands by agarose-urea gel electrophoresis. Ten single-stranded RNAs thus obtained from the five dsRNA segments IV, V, VIII, IX, and X were sequenced by postlabeling methods. Common 3'-terminal sequences, -GUUAGCC and -UUACU, were found in the plus and minus strands, respectively, of all five dsRNA segments. However, adjacent sequences diverged and were considerably variable. The homologous sequences found in the 3' end may be important recognition signals for viral RNA polymerases and for assembly of the genome segments.

Cytoplasmic polyhedrosis virus (CPV), which infects the silkworm *Bombyx mori*, is the prototype of insect viruses that have genomes consisting of 10 segments of double-stranded (ds) RNA (20). The molecular weights of these dsRNA segments range from 0.35×10^6 to 2.55×10^6 (6), and they are considered to code for viral-specific polypeptides as monocistronic genes, like human reovirus (19). CPV contains dsRNA-dependent RNA polymerase that transcribes the duplex genome RNA to form mRNA either in the infected animal or under appropriate conditions in vitro (16, 23, 24). In addition to RNA polymerase, CPV contains other enzymes required for the formation of complete mRNA with a capped 5' terminus, m^7GpppA^mpG- (12). These enzymes are nucleotide phosphohydrolase, mRNA methyltransferases, and mRNA guanylyltransferase. It was previously shown that CPV mRNA synthesis is unique in that (i) *S*-adenosylmethionine, a methyl donor, stimulates more than 70-fold the initiation of transcription (7); (ii) the action of *S*-adenosylmethionine is due to its allosteric effect on the virion RNA polymerase, resulting in a lower K_m for ATP in the initiation step of RNA synthesis (9); (iii) the allosteric conformational change of RNA polymerase is mediated in a cooperative manner with the methyltransferase in the CPV transcription complex (26); and (iv) initiation is perhaps coupled with the process of capping (8).

The terminal sequences of the dsRNA segments have been of significant interest in relation to understanding mechanisms of transcription initiation and assembly of genome segments. Furuichi and Miura (10) and Lewandowski and Leppla (17) have found previously that CPV genome RNAs contain approximately an equal amount of U and C as 3' termini by analyzing the 3'-terminal nucleosides of the dsRNA segments labeled by periodate oxidation followed by reduction with [^{3}H]borohydride. Furuichi and Miura subsequently found that each of the genome segments contain common 3' termini, PyC and PyU (Py, pyrimidine) for plus and minus strands, respectively (11). They also identified m^7GpppA^m as the common cap structure on the plus strand of all CPV dsRNA segments (12). The 5'-terminal sequences of CPV dsRNA were determined to be pA^mGU and pGGC by analyzing 5'-^{32}P-labeled CPV RNAs (K. Miura, Y. Furuichi, K. Shimotohno, T. Urushibara, K. Watanabe, and M. Sugiura, Abstr. Colloque on *In Vitro* Transcription and Translation of Virus Genomes, 1975, vol. 47, p. 153–160). From these results, the terminal trinucleotide sequences common to all genome segments of CPV have been proposed, assuming

that base pairing starts at the ends of plus and minus strands. In addition, transcription of the genome segments was thought to occur as follows:



It is important to know how far the common terminal structures of CPV dsRNA segments extend into the internal region, since the common 3'-terminal regions are considered to be recognition sites for viral transcription and genome replication. We recently found that CPV dsRNAs labeled at the 3' termini with [^{32}P]pCp and T4 RNA ligase can be separated into the plus and minus strands when the denatured RNAs are fractionated by agarose–6 M urea gel electrophoresis (24, 25). This finding made it possible for us to determine further nucleotide sequences at the 3' end of both plus and minus strands of each dsRNA segment. In this communication, we report the sequences at the 3' end of dsRNA segments IV, V, VIII, IX, and X, RNAs chosen because their complementary strands were most completely separated (25). The results indicate that CPV genome segments contain the homologous 3' sequences up to -GUUAGCC for plus strands and -UUACU for minus strands. The sequences located internal to these common regions were found to be considerably different from each other.

MATERIALS AND METHODS

Isolation of CPV dsRNA. CPV was purified from the infected midguts of silkworm *Bombyx mori* as described previously (24). The genome dsRNAs were isolated by phenol extraction from the purified viruses.

Labeling of the 3' termini of CPV dsRNAs by [5'-^{32}P]pCp by using T4 RNA ligase. Reaction mixtures (50 μl) contained 100 pmol of [^{32}P]pCp (specific activity, 3,000 Ci/mmol; Amersham Corp.), 24 μg of CPV dsRNAs equivalent to 33 pmol of 3' termini, 300 pmol of ATP, 6.3 U of T4 RNA ligase (P-L Biochemicals), and 0.5 μg of bovine serum albumin in 50 mM *N*-2-hydroxyethylpiperazine-*N*-2-ethanesulfonic acid (HEPES buffer) (pH 7.5), 20 mM $MgCl_2$, and 10% (vol/vol) dimethyl sulfoxide (4). After incubation for 4 h at 5 to 6°C, the RNAs were recovered by phenol extraction and filtered through a column of Sephadex G-100 (0.7 by 25 cm) to remove [^{32}P]pCp. Under these conditions, approximately 70% of the dsRNA termini were labeled (25).

Separation of CPV dsRNA segments. ^{32}P-labeled CPV dsRNAs (20 μg; specific activity, 10^7 dpm/μg) were applied to the 5% polyacrylamide slab gel (thickness, 1.5 mm; height, 30 cm; width, 15 cm). Electrophoresis was performed at 4°C for 18 to 20 h at 25 to 30 mA with circulation buffer containing 37 mM Tris-phosphate (pH 7.4)–5 mM EDTA. Ten individual dsRNA segments were resolved, and each band containing ^{32}P-labeled RNA was extracted from the excised gel by electrophoresis according to the method described by Yang et al. (27).

Separation of plus and minus strands from CPV dsRNAs. Each dsRNA segment was mixed with 100 μg of *Escherichia coli* rRNA and denatured by incubation at 50°C for 30 min in 5 mM citrate buffer (pH 3.5) containing 80 to 90% dimethyl sulfoxide. The denatured RNA was precipitated by addition of 0.2 M sodium acetate and 3 volumes of cold ethanol. The RNA precipitate was dissolved in 100 μl of 5 mM citrate buffer (pH 3.5) containing 7 M urea and loaded on a 1.75% agarose gel containing 25 mM citrate buffer (pH 3.5) and 6 M urea (thickness, 1.5 mm; length, 40 cm; width, 20 cm). Electrophoresis was performed at 4°C for 18 to 24 h at 30 mA to separate the complementary strands. To purify further the plus and minus strands thus obtained, the portion of the slice containing each RNA strand was excised, embedded in a second agarose gel prepared as before, and further fractionated by electrophoresis. This second electrophoresis was particularly necessary for segment IX (see Fig. 2, lane D). The purified RNAs were extracted from the gel, mixed with 100 μg of ribosomal RNA, and precipitated by the addition of 2 volumes of cold ethanol.

Sequence analysis of 3'-labeled CPV RNAs. For sequencing of 3'-end-labeled CPV RNAs, the chemical degradation method developed by Peattie (22) was mainly used. The sequences obtained by the above method were confirmed by the wandering spot method as described by DeWachter and Fiers (3). ^{32}P-labeled RNA partially hydrolyzed by alkali was fractionated by two-dimensional polyacrylamide gel electrophoresis. 3'-Terminal nucleotides of plus and minus RNAs were determined by two-dimensional thin-layer cellulose chromatography of the RNase T2 digest of each RNA (14). The chemicals and enzymes used in these experiments were as previously described (14, 15).

RESULTS

Separation of 3'-labeled CPV dsRNA segments. A mixture of 3'-labeled CPV dsRNA segments was resolved into nine bands by 5% polyacrylamide gel electrophoresis as described above (Fig. 1A). As previously reported (6), segments II and III were not separated and comigrated as a single band from the gel origin. The dsRNA segments were labeled to a similar extent by an RNA ligase reaction, using [5'-^{32}P]pCp as a donor of ^{32}P. In addition, both 3' termini of plus and minus strands of each dsRNA segment were equally labeled, as shown by nearest-neighbor analysis of the mixture of segments (Fig. 1B). These results indicate that ligation of [5'-^{32}P]pCp by RNA ligase is unaffected by the molecular weight of the dsRNA or by the 5' cap structure present in plus strands in an unpaired configuration, possibly protruding from duplex molecules.

Separation of the complementary strands of CPV dsRNAs. CPV dsRNA segments IV, V,

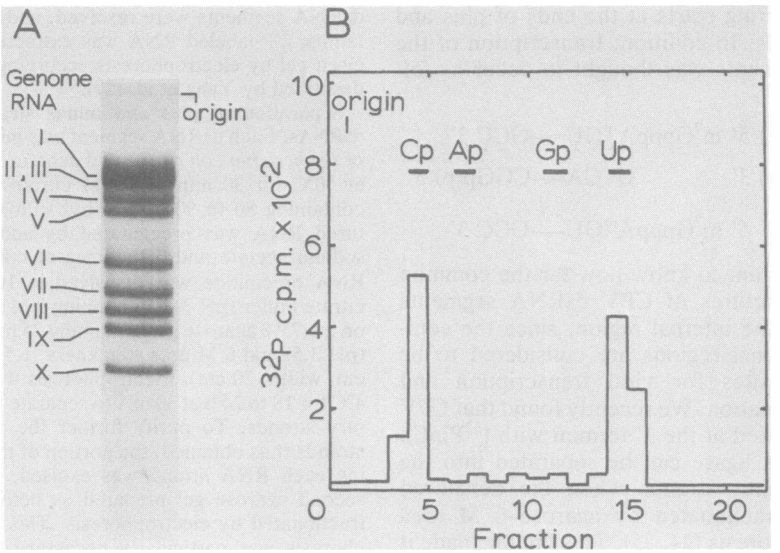


FIG. 1. (A) Separation of 3'-labeled CPV dsRNA segments by polyacrylamide gel electrophoresis. (B) Nearest-neighbor transfer analysis of 3'-terminal-labeled CPV genome dsRNA. A mixture of 10 CPV dsRNA segments was heat denatured by boiling for 2 min and rapid chilling in ice water. The denatured RNAs were digested by RNase T2, and the resulting ^{32}P-labeled nucleotides were analyzed by high-voltage paper electrophoresis at pH 3.5.

VIII, IX, and X thus isolated were denatured by incubation with dimethyl sulfoxide and further fractionated by agarose gel electrophoresis. As shown in Fig. 2, the complementary minus and plus strands (same sequence as viral mRNA) were well separated. Segment IX, because of overexposure in autoradiography, does not appear to be separated (Fig. 2D), but it is resolved into two sharp bands which are seen in the less exposed film. For sequencing, to avoid cross-contamination, the upper portion of the minus strand and the lower portion of the plus strand

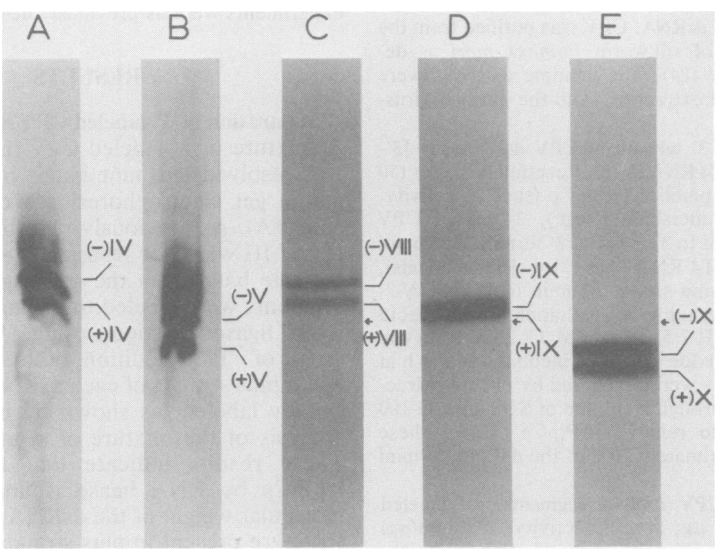


FIG. 2. Separation of complementary strands from isolated CPV dsRNA segments by agarose-urea gel electrophoresis. Lane A, segment IV; lane B, segment V; lane C, segment VIII; lane D, segment IX; lane E, segment X. The plus and minus strands are indicated by (+) and (−), respectively. Arrows represent position of xylene cyanol.

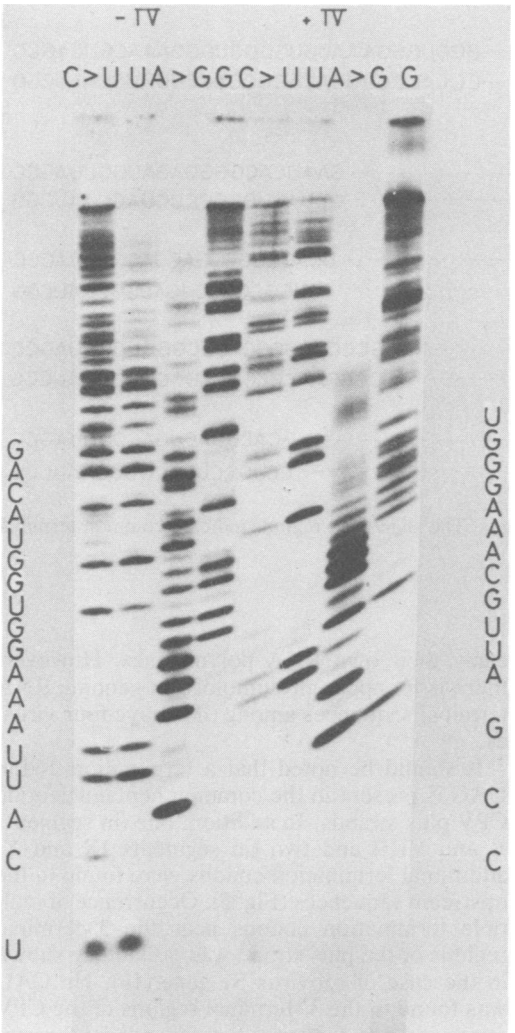


FIG. 3. Autoradiograms of 3'-terminal-labeled plus and minus strands of dsRNA segment IV, sequenced by the procedure of Peattie (22). Gel electrophoresis was carried out in a 20% polyacrylamide–7 M urea gel (20 by 40 by 0.6 cm) for 9 h at 700 V.

were excised, and RNAs were extracted. It should be noted that the plus strands of CPV segments migrated faster than the cognate minus strands in all cases, suggesting a common structural difference between plus and minus strands. It is possible that plus strands possess more G-rich sequences than their complements, as discussed previously (25).

Nucleotide sequences of 3'-terminal regions of plus and minus strands of CPV. The 3'-terminal sequences of 3'-end-labeled single-stranded RNAs were determined by two independent postlabeling procedures. More sequence information was obtained by the rapid read-off procedure described by Peattie (22). Figure 3 shows autoradiographs of 3'-end-labeled plus and minus strands of segment IV sequenced by this method. Nucleotide residues up to position 32 for the plus strand and 31 for the minus strand could be easily determined by this method. To confirm the sequences, the wandering-spot method was also carried out (Fig. 4). Results obtained by the two methods were found to be quite consistent. However, use of the two methods is necessary for unambiguous determination of long nucleotide sequences, because identification of some residues (particularly cytidine) by the rapid read-off procedure is not always clear. 3'-Terminal nucleotides were determined to be cytidine (plus strand) and uridine (minus strand) by analyzing RNase T2 complete digest

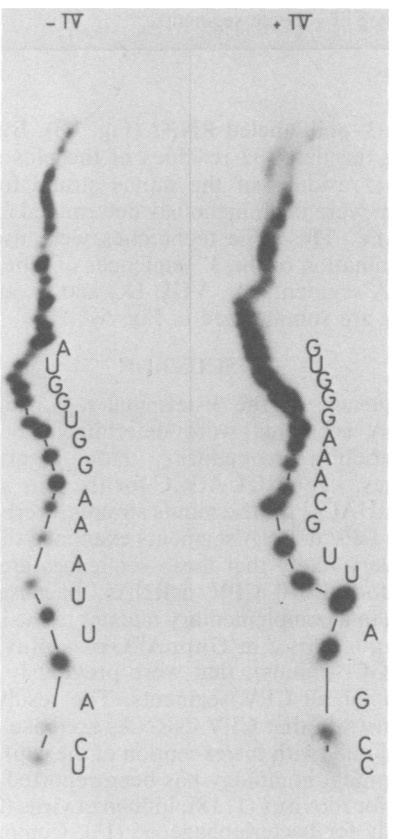


FIG. 4. Wandering-spot analysis of 3'-terminal-labeled plus and minus strands of dsRNA segment IV. Partial alkaline digest of the RNA was fractionated by two-dimensional polyacrylamide gel electrophoresis with 10% polyacrylamide gel (pH 3.5) as the first dimension (from left to right) and 20% polyacrylamide–7 M urea gel as the second dimension (from bottom to top).

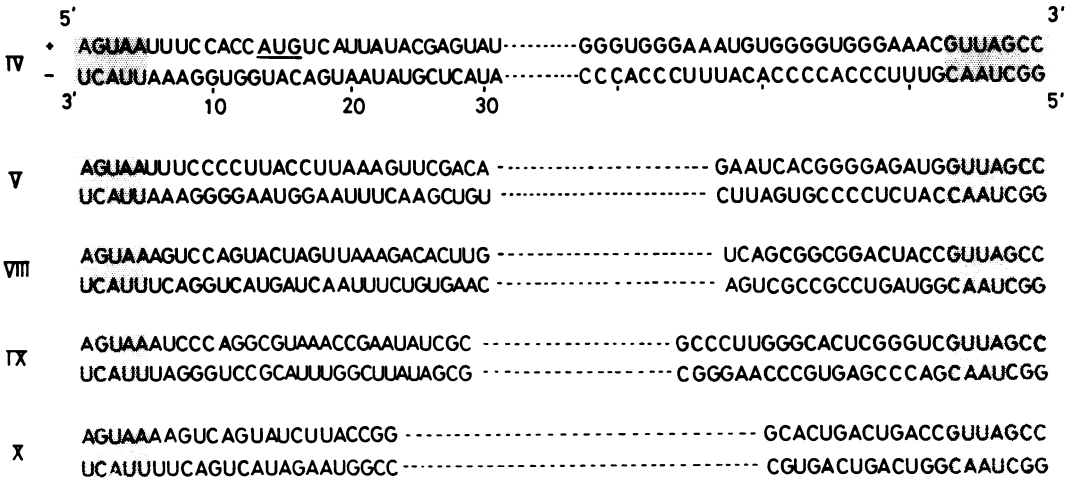


FIG. 5. 3'-Terminal sequences of CPV dsRNA segments. The shadowed regions indicate common terminal sequence of genome segments.

of the 3'-end-labeled RNAs (Fig. 1B). By combining the data, 32 residues of the plus strand and 31 residues of the minus strand from 3' termini were unambiguously determined for segment IV. The same techniques were used for determination of the 3' sequences of other CPV dsRNA segments, V, VIII, IX, and X, and the results are summarized in Fig. 5.

DISCUSSION

Sequences at the 3'-terminal region of CPV dsRNA segments were determined by using postlabeling procedures. Homologous sequences, i.e., -GUUAGCC for the plus strands and -UUACU for the minus strands, were found in five CPV dsRNA segments examined thus far. It is very likely that these sequences are common for all 10 CPV dsRNAs, because they match in a complementary manner to the limited 5' sequences, m^7GpppA^mGU- (plus) and (p)pGGC- (minus), that were previously determined for all CPV segments. The results also demonstrate that CPV dsRNAs are base paired end-to-end (with the exception of the cap). Similar terminal homology has been reported previously for reovirus (1, 18), influenza virus (2), and recently for bacteriophage ϕ6 (13). Common sequences found in reovirus genome segments (m^7GpppG^mCUA———UCAUC) are shorter than those of CPV dsRNA segments, as reported in this communication. In the cases of influenza virus and bacteriophage ϕ6, much longer common sequences were present at both termini. The existence of conserved terminal sequence in both ends is probably a general phenomenon for viruses that have segmented RNA genomes and carry their own RNA polymerases. However, there is no apparent homology in genome RNA terminal sequences among the above four viruses.

It should be noted that a termination codon UAG is present in the common heptanucleotide CPV plus strands. In addition, one (in segments V and VIII) and two (in segments IX and X) additional termination codons were found in the upstream sequences (Fig. 5). Occurrence of multiple termination codons near the 3'-terminal regions of the plus strand was previously shown in the case of reovirus S1 gene (18). No CAU was found in the 3'-terminal regions of the CPV minus strands sequenced so far, except for segment IV (the 14th to 16th nucleotide residues from

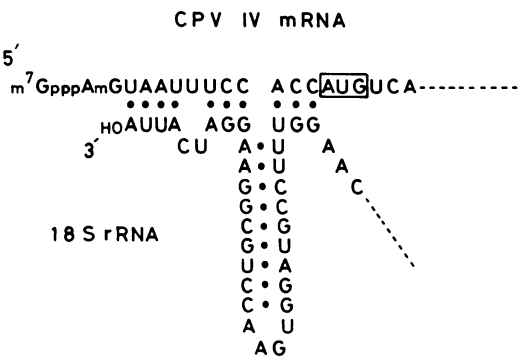


FIG. 6. Complementary sequence at 5'-terminal region of plus strand of segment IV and 18S rRNA. The sequence was deduced from the sequence of the 3' region of the minus strand.

the 3' terminus), indicating that the AUG initiation codon in the plus (mRNA) strand is farther from the 5' end for these segments.

The sequence AAUAAA that is thought to be a signal for polyadenylation (5) was not found in any of the plus strands of CPV. In view of the fact that CPV mRNAs are not polyadenylated, the absence of AAUAAA is not surprising.

The existence of direct repeats was found in 3' termini of plus strands IV and X and of minus strand V, namely, GGGUGGGAAA in segment IV, ACUGAC in an overlapping manner in segment X, and GGAAU in segment V. Another interesting feature of the sequence of CPV plus strand IV RNA is a sequence complementary to mammalian 18S rRNA found before an AUG codon (Fig. 6). This suggests that the AUG codon may be used for initiation of protein synthesis. Similar homology was previously found in several mammalian mRNAs, including reovirus mRNA (21).

LITERATURE CITED

1. **Darzynkiewicz, E., and A. J. Shatkin.** 1980. Assignment of reovirus mRNA ribosome binding sites to virion genome segments by nucleotide sequence analyses. Nucleic Acids Res. **8**:337–350.
2. **Desselberger, U., V. R. Racaniello, J. J. Zazra, and P. Palese.** 1980. The 3' and 5'-terminal sequences of influenza A, B and C virus RNA segments are highly conserved and show partial inverted complementarity. Gene **8**:315–328.
3. **DeWachter, R. D., and W. Fiers.** 1972. Preparative two-dimensional polyacrylamide gel electrophoresis of ^{32}P-labeled RNA. Anal. Biochem. **49**:184–197.
4. **England, T. E., and O. C. Uhlenbeck.** 1978. 3'-Terminal labeling of RNA with T4 RNA ligase. Nature (London) **275**:560–561.
5. **Fitzgerald, M., and T. Shenk.** 1981. The sequence 5'-AAUAAA-3' forms part of the recognition site for polyadenylation of late SV40 mRNAs. Cell **24**:251–260.
6. **Fujii-Kawata, I., K.-I. Miura, and M. Fuke.** 1979. Segments of genomes of viruses containing double-stranded ribonucleic acid. J. Mol. Biol. **51**:247–253.
7. **Furuichi, Y.** 1974. "Methylation-coupled" transcription by virus-associated transcriptase of cytoplasmic polyhedrosis virus containing double-stranded RNA. Nucleic Acids Res. **1**:809–822.
8. **Furuichi, Y.** 1978. "Pre-transcriptional capping" in the biosynthesis of cytoplasmic polyhedrosis virus messenger RNA. Proc. Natl. Acad. Sci. U.S.A. **75**:1086–1090.
9. **Furuichi, Y.** 1981. Allosteric stimulatory effect of *S*-adenosylmethionine on the RNA polymerase in cytoplasmic polyhedrosis virus—a model for the positive control of eukaryotic transcription. J. Biol. Chem. **256**:483–493.
10. **Furuichi, Y., and K.-I. Miura.** 1972. The 3'-termini of the genome RNA segment of silkworm cytoplasmic polyhedrosis virus. J. Mol. Biol. **64**:619–632.
11. **Furuichi, Y., and K.-I. Miura.** 1973. Identity of the 3'-terminal sequences in ten genome segments of silkworm cytoplasmic polyhedrosis virus. Virology **55**:418–425.
12. **Furuichi, Y., and K.-I. Miura.** 1975. A blocked structure at the 5'-terminus of mRNA of cytoplasmic polyhedrosis virus. Nature (London) **253**:374–375.
13. **Iba, H., T. Watanabe, Y. Emori, and Y. Okada.** 1982. Three double-stranded RNA genome segments of bacteriophage ϕ6 have homologous terminal sequences. FEBS Lett. **141**:111–115.
14. **Kuchino, Y., M. Kato, H. Sugisaki, and S. Nishimura.** 1979. Nucleotide sequence of starfish initiator tRNA. Nucleic Acids Res. **6**:3459–3469.
15. **Kuchino, Y., S. Watanabe, F. Harada, and S. Nishimura.** 1980. Primary structure of AUA-specific isoleucine transfer ribonucleic acid from *Escherichia coli*. Biochemistry **29**:2085–2089.
16. **Lewandowski, L. J., J. Kalmakoff, and Y. Tanada.** 1969. Characterization of a ribonucleic acid polymerase activity associated with purified cytoplasmic polyhedrosis virus of the silkworm *Bombyx mori*. J. Virol. **4**:857–865.
17. **Lewandowski, L. J., and S. H. Leppla.** 1972. Comparison of the 3' termini of discrete segments of the double-stranded ribonucleic acid genomes of cytoplasmic polyhedrosis virus, wound tumor virus, and reovirus. J. Virol. **10**:965–968.
18. **Li, J. K.-K., J. D. Keene, P. O. Sheible, and W. K. Joklik.** 1980. Nature of the 3'-terminal sequences of the plus and minus strands of the S1 genome of reovirus serotype 1, 2 and 3. Virology **105**:41–51.
19. **McCrae, M., and W. K. Joklik.** 1978. The nature of the polypeptide encoded by each of the 10 double-stranded RNA segments of reovirus type 3. Virology **89**:578–593.
20. **Miura, K., I. Fujii, T. Sakaki, M. Fuke, and S. Kawase.** 1969. Double-stranded ribonucleic acid from cytoplasmic polyhedrosis virus of the silkworm. J. Virol. **2**:1211–1222.
21. **Nakashima, K., E. Darzynkiewicz, and A. J. Shatkin.** 1980. Proximity of mRNA 5'-region and 18S rRNA in eukaryotic initiation complexes. Nature (London) **286**:226–230.
22. **Peattie, D. A.** 1979. Direct chemical method for sequencing RNA. Proc. Natl. Acad. Sci. U.S.A. **76**:1760–1764.
23. **Shimotohno, K., and K. Miura.** 1973. Transcription of double-stranded RNA in cytoplasmic polyhedrosis virus *in vitro*. Virology **53**:283–286.
24. **Smith, R. E., and Y. Furuichi.** 1980. Gene mapping of cytoplasmic polyhedrosis virus of silkworm by the full-length mRNA prepared under optimized conditions of transcription *in vitro*. Virology **103**:279–290.
25. **Smith, R. E., M. A. Morgan, and Y. Furuichi.** 1981. Separation of the plus and minus strands of cytoplasmic polyhedrosis virus and human reovirus double-stranded RNAs by gel electrophoresis. Nucleic Acids Res. **9**:5269–5286.
26. **Wertheimer, A. M., S.-Y. Chen, R. T. Borchardt, and Y. Furuichi.** 1980. S-adenosylmethionine and its analogs: structural features correlated with synthesis and methylation of mRNAs of cytoplasmic polyhedrosis virus. J. Biol. Chem. **255**:5924–5930.
27. **Yang, R. C. A., J. Lis, and R. Wu.** 1979. Elution of DNA from agarose gels after electrophoresis. Methods Enzymol. **68**:176–182.