Yeast Positive-stranded Virus-like RNA Replicons

20 S AND 23 S RNA TERMINAL NUCLEOTIDE SEQUENCES AND 3' END SECONDARY STRUCTURES RESEMBLE THOSE OF RNA COLIPHAGES *

(Received for publication, November 20, 1997, and in revised form, May 21, 1998)

Nieves Rodríguez-Cousiño, Alicia Solórzano‡, Tsutomu Fujimura, and Rosa Esteban§

From the Departamento de Microbiología y Genética / Instituto de Microbiología Bioquímica, Universidad de Salamanca / Consejo Superior de Investigaciones Científicas, Salamanca 37007, Spain

Saccharomyces cerevisiae strains carry single-stranded RNAs called 20 S RNA and 23 S RNA. These RNAs and their double-stranded counterparts, W and T dsRNAs, have been cloned and sequenced. A few nucleotides at both ends, however, remain unknown. These RNAs do not encode coat proteins but their own RNA-dependent RNA polymerases that share a high degree of conservation to each other. The polymerases are also similar to the replicases of RNA coliphages, such as Qβ. Here we have determined the nucleotide sequences of W and T dsRNAs at both ends using reverse transcriptase polymerase chain reaction-generated cDNA clones. We confirmed the terminal sequences by primer-extension and RNase protection experiments. Furthermore, these analyses demonstrated that W and T dsRNAs and their single-stranded RNA counterparts (i) are linear molecules, (ii) have identical nucleotide sequences at their ends, and (iii) have no poly(A) tails at their 3' ends. Both 20 S and 23 S RNAs have GGGGC at the 5' ends and the complementary 5'-nucleotide sequence, GCCCC-5', at their 3' ends. S1 and V1 secondary structure-mapping of the 3' ends of 20 S and 23 S RNAs shows the presence of a stem-loop structure that partially overlaps with the conserved 3' end sequence. Nucleotide sequences and stem-loop structures similar to those described here have been found at the 3' ends of RNA coliphages. These data, together with the similarity of the RNA-dependent RNA polymerases encoded among these RNAs and RNA coliphages, suggest that 20 S and 23 S RNAs are plus-strand single-stranded virus-like RNA replicons in yeast.

Many fungi carry viruses (mycoviruses), most often with double-stranded RNAs (dsRNAs) as genomes. Some of these viruses confer phenotypic changes in the host, but many others are maintained without any special properties associated. All of them are intracellular parasites with no extracellular stage. Transmission is mainly vertical or through mating or hyphal anastomosis.

Yeast strains of Saccharomyces cerevisiae have been described to carry at least 5 types of double-stranded RNAs, L-A, L-BC, M, W, and T (1). L-A, L-BC, and M are encapsidated into isometric viral particles. W and T are not encapsidated into viral coats (2). W (2.5 kilobases) and T (2.9 kilobases) have been cloned and sequenced almost entirely (3, 4). Both RNAs code for proteins with domains conserved among RNA-dependent RNA polymerases (RDRPs) of RNA viruses (5–8). The protein encoded by W (±) strands (p91) and the protein encoded by T (±) strands (p104) share a high degree of homology that extends beyond the RDRP consensus motifs, indicating a close evolutionary relationship between these RNAs (Fig. 1). Comparison with other RDRPs suggests that these polymerases are more similar to the RNA coliphage replicases than to RDRPs from dsRNA viruses, including those present in the same host, namely L-A and L-BC viruses (3, 4, 9–11).

All strains carrying W dsRNA also carry a single-stranded RNA called 20 S RNA, and all strains carrying T also have a single-stranded RNA called 23 S RNA. 20 S RNA and 23 S RNA have been proposed to be identical to the W and T (±) strands, respectively (3, 4, 12). 20 S RNA and 23 S RNA copy number is highly induced under stress conditions such as growth under nitrogen starvation (4, 13), reaching up to 100,000 copies/cell. 20 S and 23 S RNAs are not encapsidated into viral particles (14, 15) but are associated with their own RNA polymerases, forming ribonucleoprotein complexes (15, 16). Recently we have shown that the p91/20 S RNA complexes have in vitro RNA polymerase activity that synthesizes 20 S RNA (17). p104/23 S RNA complexes have similar activity. Since cis-acting signals at the ends of the RNA viral genomes often play critical roles in the template specificity of viral RNA polymerases (18–21), we decided to determine the nucleotide sequences at the ends of W and T dsRNAs.

Here we report the cloning and analysis of the nucleotide sequences at the 5' and 3' ends of W and T dsRNAs. Both (±) strands have conserved 5' end GGGGC and 3' end GCCCC-5' secondary structure-mapping of the 3' ends confirm that not only 20 S RNA and 23 S RNA share similar sequences at their ends but 3' end secondary structures as well. These sequences and secondary structures are similar to those found at the 3' ends of the genomic RNAs in (+) strand single-stranded RNA coliphages. Based on the available data we believe that 20 S and 23 S RNAs are similar to positive-stranded RNA viruses.
**MATERIALS AND METHODS**

*Yeast Strains—* Yeast strain used was strain 37-4C (*Saccharomyces cerevisiae*).

**Participation of RNAs**—W and T dsRNAs from strain 37-4C were purified by CF-11 cellulose chromatography as described previously (22). Then, W and T dsRNAs were separated on an agarose gel, electroeluted from the gel, and further passed through Elutip columns (Schleicher & Schuell). 20 S RNA and 23 S RNA were purified from strain 37-4C grown under induction conditions as described (3). Briefly, cells were grown for 48 h to stationary phase, washed, and incubated in the presence of 1% potassium acetate for 14–16 h to achieve induction of 20 S RNA and 23 S RNA (23). Total nucleic acids were separated on 1.5% agarose gels, and 23 S RNA or 20 S RNA were electroeluted from the gel, excised, once with phenol:chloroform, and precipitated with ethanol. To isolate the + and − strands of W dsRNA, the dsRNA was first denatured in the presence of 7 M urea at 90 °C for 1 min and then loaded onto a 5% polyacrylamide strand separation gel (24). Both strands were located by ethidium bromide staining, excised from the gel, and purified.

### 3' Rapid Amplification of cDNA Ends (RACE) — To clone the 3' ends of (+) and (−) strands of W and T dsRNAs, we used 3' RACE (25). The 3' ends of W and T dsRNAs were first A-tailed using poly(A) polymerase (Life Technologies, Inc.) in a buffer that contained 50 mM Tris-HCl, pH 7.9, 10 mM MgCl₂, 2.5 mM MnCl₂, 125 mM NaCl, 0.25 mM ATP, 0.25 μg/μl bovine serum albumin. The poly(A)-tailed RNAs were then denatured with CH₃OH (hydroxymethyl mercuric hydroxide) as described (26), and cDNA was synthesized using oligonucleotide Bam-dT16 (5'-CCGATCCATTTTTTTTTTTTTTT-3') and Superscript reverse transcriptase (Life Technologies, Inc.). PCR amplification of the cDNAs was carried out using oligo Bam-dT16 and one of four oligonucleotides that are complementary to the internal regions of W or T dsRNAs located at 250 and 200 nt from the end of the known sequences. The four internal primers used were PG6 (291-5'-CGGATCCGCC-GTCG-3') and NR2 (175-5'-GCCGCGTCCACCCCGTA-3') for the (+) and (−) strands of W dsRNA and NR23 (2716-5'-TTCTACGGCTGCCTCC-3') and NR22 (208-5'-GGATGCACGTCGTTAAAGC-3') for the (+) and (−) strands of T dsRNA, respectively. Amplification was carried out using Taq DNA polymerase (Promega) for 30 cycles at 95 °C (90 s), 55 °C (30 s), and 72 °C (30 s), with 200 μM of dNTPs, 50 mM KCl, 10 mM MgCl₂, 1.5 mM MgCl₂, 10 μM of each (−) primers, and 1.0 U of Taq DNA polymerase. The labeled 31P-containing (5′) primer (8000 cpm) was first precipitated in the reaction buffer for 10 min at 37 °C and then digested with Nuclease S1 (Life Technologies, Inc.) or RNase V1 (Amersham). Nuclease S1 digestion was performed in a reaction mixture (6 μl) containing 30 mM sodium acetate, pH 4.6, 1 mM zinc acetate, 5% glycerol, 80 mM NaCl, 0.5 μg of tRNA, and 0.2, 2, or 10 units of S1 nuclease (26). RNase V1 digestion was done in a buffer (6 μl) that contained 30 mM sodium citrate, 0.5 μg of RNA, 0.5 μg of tRNA, and 0.009 or 0.018 units of RNase V1. Nuclease treatments were done at 37 °C for 10 min, and the reactions were stopped by the addition of 1 μl of 100 mM EDTA and the same volume of loading buffer (10 μl, 1.5 mM EDTA, 0.05% xylene cyanol, and 0.05% bromphenol blue). To generate a sequence ladder, alkaline hydrolysis was carried out at 90 °C for 5 min in a 6-μl reaction volume that contained 50 mM NaOH, 0.5 g of NaCl, 0.5 μg of tRNA, and twice the amount of labeled RNA used for the enzymatic digestions. The cleaved products were analyzed on a 7% urea 20% or 10% polyacrylamide gels.

### General Procedures—Plasmid DNA was sequenced by the dideoxy chain termination method (29) using T7 DNA polymerase (Amersham). *In vitro* run-off RNA transcription by T7 or T3 RNA polymerases (Promega) was carried out using plasmids linearized with appropriate restriction enzymes. Then the DNA template was digested with 1 μg of DNase I (Promega) for 15 min at 37 °C. RNA secondary structure prediction was done using the RNAfold program (20). RNA cDNA, oligonucleotides, and labeled RNA were isolated from the EMBL data bank.

### RESULTS

**Cloning of W and T dsRNAs Ends**—Previously we cloned and sequenced random primer-generated cDNAs from W and T dsRNAs. In these works we obtained 2505- and 2871-base pair nucleotide sequences for W and T dsRNAs, respectively. As judged from the mobilities of denaturing acrylamide gels of *in vitro* transcribed cDNAs from these sequences, we estimated that our cDNA sequences lacked only a few nucleotides at the ends of these RNAs. To understand the replication mechanism of these RNAs, however, it is essential to know the exact nucleotide sequences of these molecules, especially at both ends.

To clone the ends of both RNAs, we used the method called 3' RACE (rapid amplification of cDNA ends) (25). W or T dsRNAs fused to the T7 RNA polymerase promoter and the Smal site of plBluescript-KS+ vector. Run-off transcription of Smal-digested pALI17 by T7 RNA polymerase, therefore, gave transcripts that have the entire sequence of 20 S RNA with the correct 5' ends. The products were digested out at 90 °C for 7 min in a 6-μl reaction mixture that contained 50 mM Tris-HCl, 10 mM MgCl₂, 0.5 μg of RNA, and 500 U of RNase T1 under high salt, and the protected RNA fragments were separated on a 7% urea 6% polyacrylamide gel and detected by autoradiography.

**Secondary Structure Analysis—** W or T dsRNAs corresponding to the 3' ends of W and T (+) strands were generated by run-off transcription with T3 RNA polymerase from Pou-II-digested pW3-2 and Eaug-digested pT3-8, respectively. The 32P-labeled probes were separated in a denaturing acrylamide gel and purified from the gel as mentioned above. The probe was then annealed with 20 S RNA or W (+) strands or with 23 S RNA or T (+) strands, depending on the probe used. The RNA hybrids were digested with RNase T1 or RNase T1 under high salt, and the protected RNA fragments were separated on a 7% urea 6% polyacrylamide gel and detected by autoradiography.
Newly acquired nucleotide sequences in the reverse transcription-PCR-generated clones are underlined. The number of independent clones in each case are indicated in parenthesis. Initiation and termination codons for p91 (20S RNA) or p104 (23S RNA) are in boldface.

| 5' ends      | 3' ends                                                                 |
|--------------|-------------------------------------------------------------------------|
| W (+) strands | GGCGCGUGAUCCCAUG...                                                     |
| T (+) strands | GGCGCGUG...                                                             |
|             | GGCGCGUG...                                                             |

### TABLE I

5' and 3' end sequences of W and T (+) strands

In the primer extension experiments, purified (+) or (−) strands of W dsRNA or denatured T dsRNA were mixed with a 5' end-labeled oligo primer that was complementary to the 5' end region of the RNA. Then the labeled primer was extended toward the 5' end of the template by reverse transcription. As shown in Fig. 2A, lane 1, the primer complementary to the 5' end region of W (+) strand was fully extended and terminated as a single band at the position corresponding to the last 3' end C of the W (−) strand. This result clearly indicates that the 5' end sequence of W (+) strand shown in Table I is correct and has no preceding extra Ts. When the T (+) strands were examined similarly by primer extension, we obtained again a single band of the extended primer terminating at the position corresponding to the last 3' end C of the W (−) strand. This result again indicates that the 5' end sequence of T (+) strands shown in Table I is correct and has no extra nucleotide sequences at the 5' end. We also examined the (−) strands of W and T dsRNAs and obtained the same results; that is, the first nucleotides of the 5' ends of W and T (−) strands are G and there are no extra Ts attached to them (not shown). It should be pointed out that although the cloning of each end of W and T dsRNA by 3' RACE was manipulated at their 3' ends, the primer extension experiments shown in Fig. 2 directly analyzed the 5' end of each RNA strand. The fact that these two independent but complementary experiments gave consistent results strongly suggests that our cloned sequences are correct and represent the real W and T end sequences.

Logically, however, the possibility still remains that there exists a nonbase-pairing poly(A) tail at the 3' end of W or T (+)
strands. To rule out this possibility, we undertook a second experimental approach; RNase protection experiments. We made uniformly labeled RNA in vitro that had the nucleotide sequence from base 2514 to 2288 (numbering refers to the (+) strand sequence) of the W (−) strand attached to the 5′ upstream poly(T) sequence (Fig. 3A). This probe therefore can hybridize to the 3′ end region of W (+) strand. If W (+) strands have poly(A) tails at their 3′ ends, a part of the poly(T) sequence of the probe complementary to the poly(A) tail should be protected from RNase digestion. As shown in Fig. 3A, a part of the probe corresponding to W (−) strand from base 2514 to 2288 was fully protected, but the 5′ end poly(T) sequence was completely digested with the RNases (lanes 4 and 5). When the 3′ end of the T (+) strand was examined using a similar probe, a portion of the probe corresponding to the T (−) strand sequence from base 2891 to 2812 was fully protected, but again, the adjacent upstream poly(T) sequence was completely digested (Fig. 3B, lanes 4 and 5). These results, therefore, (i) confirm the correctness of our 3′ end nucleotide sequences of W and T (+) strands and (ii) clearly rule out the possibility that the W and T (+) strands have non-base paring poly(A) tails at their 3′ ends. Altogether, the results from primer extension analysis and RNase protection experiments indicate that the 3′ end sequences of W and T dsRNAs obtained by 3′ RACE are genuine, and that there are no non-base paring poly(A) tails at their 3′ ends. Thus we have now established the complete nucleotide sequences of W and T dsRNAs.

20 S and 23 S RNAs Are Identical to the (+) Strands of W and T dsRNAs, Respectively—Previously we proposed that 20 S and 23 S RNAs are identical to the (+) strands of W and T dsRNA, respectively, based on the following evidence. (i) All the strains carrying W and T dsRNA also harbor 20 S and 23 S RNA, respectively, and vice versa. (ii) Specific probes for the (+) strands of W (or T) dsRNA hybridize with 20 S (or 23 S) RNA. (iii) 20 S (or 23 S) RNA was indistinguishable from the (+) strands of W (or T) dsRNA in denaturing and strand separation acrylamide gels. (iv) The known nucleotide sequence of 20 S RNA (2479 base pairs) (12) is identical to the W (+) strand sequence from base 13 to base 2491.

Since we have obtained the complete sequences of W and T dsRNAs, we asked whether 20 S and 23 S RNA have the same corresponding sequences at their ends. The 5′ end of 20 S RNA was examined by primer extension with the same oligonucleotide used for W (+) strands. As shown in Fig. 2A, lane 2, the primer was extended and terminated as a single band at the same position where the 5′ end of the W (+) strand terminates. This result indicates that 20 S RNA has the same primer binding site at the same distance from the 5′ end as the W (+) strand. When the 5′ end of 23 S RNA was analyzed similarly with the primer used for the T (+) strands, we obtained the same result; that is, the primer was terminated as a single band at the same position corresponding to the 5′ end terminus of the T (+) strand (Fig. 2B, lane 2). Therefore, the primer extension analysis indicates that the 5′ ends of 20 S RNA and 23 S RNA are indistinguishable from those of W and T (+) strands, respectively. The 3′ end regions of 20 S and 23 S RNAs were analyzed by RNase protection experiments with the same RNA probes used for the 3′ ends of W and T (+) strands, respectively. As shown in Fig. 3A, lane 3, the W (+) strand-specific probe was protected by 20 S RNA from RNase digestion in the region complementary to the 3′ end of the W (+) strand. This result clearly indicates that the 3′ end region of 20 S RNA is identical to that of the W (+) strand. Since the upstream poly(T) sequence of the probe was, however, completely digested by RNase, it also indicates that 20 S RNA has no poly(A) tail at its 3′ end. When the 3′ end of 23 S RNA was examined similarly with the T (+) strand-specific probe, the probe was protected by 23 S RNA in the region complementary to the 3′ end of the T (+) strand (Fig. 3B, lane 3). Again, the upstream poly(T) sequence of the probe was completely digested. Therefore this result indicates that the 3′ end of 23 S RNA is identical to that of the T (+) strand and has no poly(A) tail attached to it. All the results from primer extension and RNase protection experiments, together with our previous observations, thus
**Fig. 4. Possible secondary structures of 20 S and 23 S RNAs 5' ends.**
The AUG initiation codons for p91 and p104 are boxed.

**Fig. 5. Nuclease S1 and V1 mapping of W (+) strand 3' end.**

**(A)** S1 mapping. 3' end-labeled W (+) strands (lanes 8 to 10) or WHindIII transcript (lanes 3 to 5) were digested with various amounts of nuclease S1 (0.2, 2, or 10 units (U)) or mocked-treated (lane C) and separated on a 7 M urea 20% acrylamide gel. Cleaved products were detected by autoradiography. Arrows indicate the positions of the two most sensitive regions to S1 digestion, designated as loop I and loop II. The size of the cleaved products was estimated from an RNA ladder obtained by alkaline hydrolysis of the samples (lane -OH). The larger RNA fragments from the W (+) strands were resolved in a 10% acrylamide gel and shown on the upper right panel. Diagrams of the W (+) strand and the WHindIII transcripts are shown at the bottom of the autoradiograms. WHindIII transcripts contain the 3' end fragment of W (+) strand from nt 2288 to 2514 and the upstream vector-derived sequence (open square).

**(B)** V1 mapping. The same samples as in A were digested with 0.009 or 0.018 units of RNase V1 (lanes 3 and 4 for the WHindIII transcript or lanes 9 and 10 for W (+) strands) or mocked-treated (lane C). The cleaved products were separated on a 7 M urea 20% acrylamide gel and detected by autoradiography. For comparison, we analyzed the same samples digested with 2 units of nuclease S1 in the same gel (lanes 6 and 12). Arrows indicate positions highly sensitive to V1 digestion (numbered from the 3' end). The positions of loop I and loop II from the S1 digestion are also indicated.
we tried to label purified 20 S RNA with [32P]pCp and T4 RNA ligase. To analyze the 3' end secondary structure of 20 S RNA because of its smaller size. There are two regions close to the 3' end clearly sensitive to S1 digestion, namely the loop parts of stem-loops I and II. Loop I extends from nt 18 to 21, and loop II extends from nt 62 to 66 from the 3' end. These two stem-loops are present in the most optimal conformations for 20 S RNA and the WHindIII transcript predicted with the RNAFOLD program (30) (Fig. 7). Furthermore, there is a third region sensitive to S1 nuclease, located around nt 92, that corresponds to a single-stranded region between stem-loops II and III (Fig. 5A, upper right panel and Fig. 7). We confirmed the presence of these stem-loop structures by V1 nuclease digestion (Fig. 5B). The preference of V1 to digest double-stranded regions gives a pattern of partially digested products essentially complementary to that of S1 digestion (compare lane 4 and lane 6 in Fig. 5B; summarized in Fig. 7).

To determine the secondary structure of 23 S RNA 3' end region, we tried to label purified 23 S RNA with [32P]pCp and

indicate that 20 S and 23 S RNAs are identical to the (+) strands of W and T dsRNAs, respectively.

W and T dsRNAs and 20 S and 23 S RNAs Are Linear—W and T dsRNAs can be labeled stoichiometrically at their 3' ends with [32P]pCp and T4 RNA ligase, indicating that these molecules have free OH groups at their 3' termini (24, 31). In addition, as shown in Fig. 2, all the reverse-transcribed products from W and T dsRNA by primer extension terminated at the positions corresponding to the 5' termini of W and T dsRNA, and there is no more synthesis beyond them. These results clearly indicate that W and T dsRNAs are linear molecules.

The linearity of 20 S RNA was demonstrated previously by a site-directed single cleavage with RNase H, which produced two fragments as expected from its linear nature (24). As demonstrated in Figs. 2 and 3, 20 S RNA and 23 S RNA have the 5' and 3' termini identical to those of the W and T (+) strands, respectively. Therefore, these results indicate that not only 20 S RNA but also 23 S RNA are linear molecules.

Characteristics of 20 S and 23 S RNA Termini—20 S and 23 S RNAs share the 5 nt sequence GGGGC at the 5' ends and the 5 nt GCCCC-OH at the 3' ends, thus confirming their close relationship. Interestingly, these 5-nt sequences are complementary. They can potentially form panhandle structures. It also means that the template strands for (+) and (−) strand synthesis have the same 5-nt sequences (GCCCC-OH) at their 3' ends from which the new strand synthesis begins. Computer-generated secondary structure prediction suggests that 20 S and 23 S RNAs have very similar secondary structures at their 5' ends (Fig. 4). In both RNAs, the leading sequences upstream of the first AUG initiation codons are very short, 12 nt in 20 S RNA and only 6 nt in 23 S RNA and rich in GC content, and these AUG codons are located within a region of a strong secondary structure (Fig. 4). Both RNAs share nucleotides at −2 (C), −1 (C), and +5 (A) positions with respect to the AUG codon. Host mRNAs have much longer leader sequences (usually about 50 nt long), and these sequences are rich in adenine and void of significant secondary structure (32). The consensus sequence around the initiation codons is AAU/AUAGUCU, and less expressed mRNAs have more G-rich and shorter leader sequences. Therefore, the leader sequences of 20 S and 23 S RNAs have characteristics of less expressed yeast mRNAs, and these features may contribute to the low abundance of the encoded p91 and p104 proteins in the host cell. On the contrary, these features may contribute to the low abundance of the encoded p91 and p104 proteins in the host cell. On the contrary, these features may contribute to the low abundance of the encoded p91 and p104 proteins in the host cell.
T4 RNA ligase. Again we observed poor labeling of 23 S RNA. We also tried to isolate labeled T strands from pCp-labeled T dsRNA in a strand separation acrylamide gel. Because of poor separation of both strands in the gel we could not obtain labeled T strands pure enough to be analyzed. Therefore we expressed in vitro a small fragment (TSpeI) of 23 S RNA with the correct 3' end and labeled its 3' end with [32P]pCp and T4 RNA ligase. This fragment, according to the RNAFOLD program prediction, contained the 23 S RNA 3' end terminal domain that can form three stem-loop structures (Fig. 7) similar to those present in 20 S RNA 3' end. The same prediction was obtained with the entire 23 S RNA nucleotide sequence.

The labeled fragment was partially digested with S1 or V1 nuclease and analyzed in a sequencing gel. As shown in Fig. 6 the patterns of digestions clearly identified two 3' end stem-loop structures (I and II), consistent with the computer predicted secondary structure of 23 S RNA 3' end (Fig. 7).

**DISCUSSION**

In this paper we have cloned and sequenced both ends of W and T dsRNA. Together with the previously known sequences, we have now established the complete nucleotide sequences of these RNAs. The genuineness of the end sequences obtained was demonstrated by primer extension and RNase protection experiments. Furthermore, these experiments clearly indicated that both ends of 20 S and 23 S RNAs are identical to those found in RNA coliphages. The terminal 3 C residues shared among 20 S and 23 S RNAs and RNA coliphages are underlined. The non-templated A residues at the 3' ends are indicated by parenthesis. Coliphages belong to the following groups: MS2 (group I), GA (group II), Qβ (group III), and SP (group IV).

**FIG. 7.** Secondary structures of 20 S and 23 S RNAs 3' ends. Computer-predicted secondary structures of 20 S and 23 S RNAs 3' ends were obtained by the RNAFOLD program (30). In both cases there are three putative stem-loop structures. Results of nuclease mapping of the 20 S RNA and 23 S RNA fragments shown in Figs. 5 and 6 are consistent with the presence of stem-loops I and II in both RNAs. Positions sensitive to S1 and V1 digestions are indicated by triangles and arrows, respectively.

**FIG. 8.** 20 S RNA and 23 S RNA have 3' end sequences and secondary structures similar to those found in RNA coliphages. The terminal 3 C residues shared among 20 S and 23 S RNAs and RNA coliphages are underlined. The non-templated A residues at the 3' ends are indicated by parenthesis. Coliphages belong to the following groups: MS2 (group I), GA (group II), Qβ (group III), and SP (group IV).
those of W and T (+) strands, respectively, thus confirming our proposal that 20 S and 23 S RNAs are single-stranded (+) forms of W and T dsRNA.

Interestingly, 20 S and 23 S RNA have no poly(A) tails at their 3’ ends, in contrast to some minus strand RNA viruses that synthesize poly(A)-tailed mRNAs distinct from their genomic complementary strand RNAs (33). We do not know whether 20 S and 23 S RNAs are 5’ end-capped. It is extremely difficult to label the 5’ end of 20 S and 23 S RNA with [γ-32P]ATP and T4 polynucleotide kinase. Since we have encountered a similar difficulty in labeling the 5’ end of 20 S RNA, even made in vitro by T7 RNA polymerase with the correct termini, this difficulty apparently comes from the inaccessibility of the enzyme to the 5’ ends, perhaps because of the large RNA size rather than from a capped structure at the 5’ end. So far we have not noticed any consensus motifs in p91 and p104 implicated in viral capping enzymes (34). Considering the cytoplasmic localization of 20 S and 23 S RNAs within the cell, it is likely that these RNAs are not capped at their 5’ ends. At any rate, it should be clarified.

20 S and 23 S RNAs apparently belong to the same RNA family of viral origin. They share many similarities, such as their genome organization, RNA polymerase sequences, and inducibility under nitrogen starvation conditions among others. When the nucleotide sequences at both ends are compared, their resemblance is again quite evident. Both RNAs share the same 5-nt sequence (GGGGC) at their 5’ ends and also its complementary 5-nt sequence (GCCCU-OH) at their 3’ ends. The latter 3’ end sequence partially overlaps with a strong secondary structure, stem-loop I in 20 S RNA (Fig. 7), and probably also in 23 S RNA (Fig. 7). When we searched the 3’ end sequences of viral RNAs in data banks, we found that (+) strand single-stranded RNA coliphages have similar sequences and secondary structures at the 3’ ends of their genomic RNAs (Fig. 8). Previously we had noticed that the consensus sequences for RNA-dependent RNA polymerases found in p91 and p104 are most closely related to those of the replicases of RNA coliphages (3, 4, 9, 10). These resemblances, especially in their 3’ end RNA structures, might be conserved during their evolution, possibly from the beginning of the RNA world as suggested by Maizels and Weiner (35) in their genomic tag hypothesis. Alternatively, these polymerases might require similar chemical natures in their RNA templates. The polymerase senses the stem-loop structure as a signal of the 3’ end of the template, and the successive C residues at the 3’ end provide chemical stability in the initiation complex. At any rate, these resemblances strongly suggest that 20 S and 23 S RNAs are similar to (+) strand single-stranded RNA viruses. The available data fit to this framework are: (i) 20 S and 23 S RNAs are of positive sense; (ii) the (−) strands are much less, usually less than 1% of the (+) strands; (iii) both 20 S and 23 S RNAs are linear molecules and have no poly(A) tails. Therefore their reproduction proceeds through end-to-end (−) strand synthesis and then from these (−) strands, the (+) strand RNAs are regenerated by end-to-end transcription. This replication scheme is consistent with our previous observation of in vitro 20 S RNA synthesis. Crude extracts prepared from induced cells synthesized mostly 20 S RNA and a small amount of the − strands. In time course measurements, we detected nascent molecules that grew in size to full unit-length 20 S RNA during the incubation (17). W and T dsRNAs can be referred as replicative forms of 20 S and 23 S RNAs. However it should be mentioned that they could be formed artifically during their isolation or generated in vivo as dead end products. In the Q8 system, it is well documented that the double-stranded forms are inactive as templates (36). Another feature of 20 S and 23 S RNAs shared with coliphages is the existence of a single nontemplated A residue at their 3’ ends (Fig. 8). Based on the analysis of the 3’ end terminal nucleotides done by thin layer chromatography, we found that about 80–90% of W or T (+) strands are C residues, and 10–20% are A residues (24, 31). The same results were obtained when the 20 S and 23 S RNA 3’-terminal nucleotides were analyzed, thus confirming the identity of W (+) strands with 20 S RNA and T (+) strands with 23 S RNA.

Apparently, it is controversial to refer to 20 S and 23 S RNAs as (+) strand RNA viruses, since these RNAs do not encode coat proteins and are not encapsidated into viral particles. However, we believe that this controversy could be reconciled if we take their biological context into consideration. So far all known mycoviruses have no extracellular pathway of transmission, except for some opportunistic events (37). It means that these viruses do not need elaborated machineries for exit and re-entry into new host cells. For example, yeast Ty retroelements have much simpler genome organization compared with those of retroviruses (38), and their genomic RNAs are encapsidated into particles that are equivalent to the inner cores of the higher eukaryote counterparts. Yeast totoivirus L-A encodes only two proteins (coat protein and RNA polymerase), and the RNA genome is encapsidated into a particle that corresponds to the inner core of reoviruses. In both cases they lack envelope and outer capsid. Therefore, if we placed the RNA coliphages into yeast, the gene necessary for cell lysis and the functions of coat protein that may be involved in exit and re-entry would become dispensable. In the cases of Ty and L-A, however, coat (or capsid) proteins also provide other functions. They still protect the encapsidated genome from the nuclease-rich intra-cellular environment (39). They also provide an environment where their genome and replication machinery are intimately concentrated. This compartmentalization may promote efficient and faithful reproduction of their progeny (or provirus). Concerning those functions, we believe 20 S and 23 S RNAs have evolved uniquely. As mentioned, 20 S and 23 S RNAs are not encapsidated into particles, but they form ribonucleoprotein complexes with their cognate RNA polymerases. This complex formation thus substitutes the necessity of compartmentalization by coat proteins. The polymerase in the complex sees the high local concentration of the 3’ ends of the bound template. The RNAs themselves might have evolved in such a way that they can form stable secondary or tertiary structures relatively resistant to cellular RNases, even without a protective protein coat. Such extreme cases may be seen in the small RNA plant pathogens viroids, which form highly organized rod-shaped structures by extensive intramolecular hydrogen bondings. The high GC contents in 20 S RNA (58.3%) and 23 S RNA (59.0%), which are equivalent to those of viroids (53–60%) (40), surely stabilize such structures. On the other hand, the cellular mRNAs and L-A dsRNA virus in the same host have much lower GC contents (38–40 and 45.7%, respectively). In addition, the ribonucleoprotein complex formation would further stabilize and protect the RNA, thus substituting the necessity of a protein coat to protect the genomic RNA. Our preliminary results indicate that the complex formation greatly enhances the RNA stability in vitro. The lack of infectious cycle, therefore, allowed 20 S and 23 S RNAs to simplify their genomes to the extremity. They encode

---

2 N. Rodríguez-Cousínó, T. Fujimura, and R. Esteban, unpublished results.
3 N. Rodríguez-Cousínó, A. Solórzano, R. Esteban, and T. Fujimura, manuscript in preparation.
4 T. Fujimura and R. Esteban, unpublished data.
only their RNA polymerases. Their simple appearance in the genome organization is, however, deceptive. They have to reproduce their progenies like other viruses and live peacefully in the host. Now all the information necessary for these activities is concentrated into only one gene product and perhaps the RNA genome structure itself.

REFERENCES

1. Wickner, R. B. (1996) Microbiol. Rev. 60, 250–265
2. Wesselowski, M., and Wickner, R. B. (1984) Mol. Cell. Biol. 4, 181–187
3. Rodríguez-Cousíno, N., Esteban, L. M., and Esteban, R. (1991) J. Biol. Chem. 266, 12772–12778
4. Esteban, L. M., Rodríguez-Cousíno, N., and Esteban, R. (1992) J. Biol. Chem. 267, 10874–10881
5. Kamer, G., and Argos, P. (1984) Nucleic Acids Res. 12, 7269–7282
6. Argos, P. (1988) Nucleic Acids Res. 16, 9909–9916
7. Poch, G., Sauvaget, I., Delarue, M., and Tordo, N. (1989) EMBO J. 8, 3867–3874
8. Koonin, E. V. (1991) J. Gen. Virol. 72, 2197–2206
9. Esteban, R., Rodríguez-Cousíno, N., and Esteban, L. M. (1993) Prog. Nucleic Acid Res. Mol. Biol. 46, 155–182
10. Koonin, E. V. (1992) Semin. Virol. 3, 327–339
11. Bruenn, J. A. (1993) Nucleic Acids Res. 21, 5667–5669
12. Matsumoto, Y., and Wickner, R. B. (1991) J. Biol. Chem. 266, 12779–12783
13. Matsumoto, Y., Fishel, R., and Wickner, R. B. (1990) Proc. Natl. Acad. Sci. U. S. A. 87, 7628–7632
14. Widmer, W. R., Matsumoto, Y., and Wickner, R. B. (1991) Mol. Cell. Biol. 11, 2905–2908
15. Esteban, L. M., Fujimura, T., Garcia-Cuellar, M. P., and Esteban, R. (1994) J. Biol. Chem. 269, 29771–29777
16. Garcia-Cuellar, M. P.; Esteban, L. M., Fujimura, T., Rodríguez-Cousíno, N., and Esteban, R. (1995) J. Biol. Chem. 270, 20084–20089
17. Garcia-Cuellar, M. P., Esteban, R., and Fujimura, T. (1997) RNA 3, 27–36
18. Miller, W. A., Rajarski, J. J., Drcher, T. W., and Hall, T. C. (1986) J. Mol. Biol. 187, 537–546
19. Esteban, R., Fujimura, T., and Wickner, R. B. (1989) EMBO J. 8, 947–954
20. Buck, K. W. (1996) Adv. Virus Res. 47, 159–251
21. Patton, J. T., Wentz, M., Xiao, J., and Ramig, R. F. (1996) J. Virol. 70, 3961–3971
22. Toh-He, A., Guerry, P., and Wickner, R. B. (1978) J. Bacteriol. 136, 1002–1007
23. Wejksnora, P., and Haber, J. E. (1978) J. Bacteriol. 134, 246–260
24. Rodríguez-Cousíno, N., and Esteban, R. (1992) Nucleic Acids Res. 20, 2761–2766
25. Frohman, M. A. (1990) in PCR Protocols: A Guide to Methods and Applications (Innis, M. A., Gelfand, D. H., Sninsky, J. J., and White, T. J., eds) pp. 28–38, Academic Press, Inc., New York
26. Maniatis, T., Fritsch, E. F., and Sambrook, J. (1982) Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY
27. Sambrook, J., Fritsch, E. F., and Maniatis, T. (1989) Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY
28. Hannig, E. M., and Leibowitz, M. J. (1985) Nucleic Acids Res. 13, 4379–4400
29. Sanger, F., Nicklen, S., and Coulson, A. R. (1977) Proc. Natl. Acad. Sci. U. S. A. 74, 5463–5467
30. Zuck, M., and Stiegler, P. (1981) Nucleic Acids Res. 9, 133–148
31. Esteban, L. M. (1993) Doctoral dissertation, University of Salamanca, Spain
32. Cigan, A. M., and Donahue, T. F. (1987) Gene 59, 1–18
33. Lamb, R. A., and Krug, R. M. (1996) in Fundamental Virology (Field, B. N., Knipe, D. M., and Howly, P. M., eds) 3rd Ed., pp. 605–647, Lippincott-Raven Publishers, Philadelphia
34. Koonin, E. V. (1993) J. Gen. Virol. 74, 733–740
35. Maizels, N., and Weiner, A. M. (1993) in The RNA World (Gesteland, R. F., and Atkins, J. F., eds) pp. 577–602, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY
36. Priano, C., Kramer, F. R., and Mills, D. R. (1987) Cold Spring Harbor Symp. Quant. Biol. 52, 321–330
37. El-Sherbeini, M., and Bostian, K. A. (1987) Proc. Natl. Acad. Sci. U. S. A. 84, 4293–4297
38. Beeke, J. D., and Sandmeyer, S. B. (1991) in The Molecular and Cellular Biology of the Yeast Saccharomyces cerevisiae (Broach, J. R., Pringle, J. R., and Jones, E. W., eds) pp. 193–261, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY
39. Fujimura, T., Esteban, R., Esteban, L. M., and Wickner, R. B. (1990) Cell 62, 819–828
40. Flores, R. (1995) in Virus Taxonomy. Sixth report of the International Committee on Taxonomy of Viruses (Murphy, F. A., Fauquet, C. M., Bishop, D. H. L., Ghabrial, S. A., Jarvis, A. W., Martelli, G. P., Mayo, M. A., and Summers, M. D., eds) pp. 495–497, Springer-Verlag, Vienna