Launching of the Yeast 20 S RNA Narnavirus by Expressing the Genomic or Antigenomic Viral RNA in Vivo*

Received for publication, June 16, 2005, and in revised form, July 25, 2005 Published, JBC Papers in Press, July 27, 2005 DOI 10.1074/jbc.M506546200

Rosa Esteban 1, Lorena Vega 2, and Tsutomu Fujimura 3

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

20 S RNA virus is a persistent positive strand RNA virus found in Saccharomycyes cerevisiae. The viral genome encodes only its RNA polymerase, p91, and resides in the cytoplasm in the form of a ribonucleoprotein complex with p91. We succeeded in generating 20 S RNA virus in vivo by expressing, from a vector, genomic strands fused at the 3'-ends to the hepatitis delta virus antigenomic ribozyme. Using this launching system, we analyzed 3'-cis-signals present in the genomic strand for replication. The viral genome has five-nucleotide inverted repeats at both termini (5'-GGGGC...GCCCC-OH). The fifth G from the 3'-end was dispensable for replication, whereas the third and fourth Cs were essential. The 3'-terminal and penultimate Cs could be eliminated or modified to other nucleotides; however, the generated viruses recovered these terminal Cs. Furthermore, extra nucleotides added at the viral 3'-end were eliminated in the launched viruses. Therefore, 20 S RNA virus has a mechanism(s) to maintain the correct size and sequence of the viral 3'-end. This may contribute to its persistent infection in yeast. We also succeeded in generating 20 S RNA virus similarly from antigenomic strands provided active p91 was supplied from a second vector in trans. Again, a cluster of four Cs at the 3'-end in the antigenomic strand was essential for replication. In this work, we also present the first conclusive evidence that 20 S and 23 S RNA viruses are independent replicons.

Positive strand RNA viruses encode RNA-dependent RNA polymerases in their genomes and utilize them to synthesize viral RNA in conjunction with other viral or host proteins (1). The antigenic (negative) strand RNA is an intermediate of replication and serves as a template to synthesize progeny positive strands. Because viral replication takes place within the host cell and because there are plenty of cellular RNAs in these cells, the viruses must find efficiently not only the positive strands but also the negative strands during replication. Therefore, both the positive and negative strands bear cis-acting signals for replication, including those necessary to interact with the polymerase machinery. To analyze and characterize these signals, it is essential to develop appropriate in vivo or in vitro systems in which changes in the signals can be tested.

20 S and 23 S RNAs are positive strand RNA viruses found in Saccharomycyes cerevisiae and belong to the genus Narnavirus (2). These viruses were initially identified as RNA species induced in cells under nitrogen starvation conditions (3, 4). Most laboratory strains carry 20 S RNA, and fewer strains harbor 23 S RNA along with 20 S RNA. They are compatible in the same host cell. Because all known strains that carry 23 S RNA also carry 20 S RNA, it has not been clear whether 23 S RNA virus can replicate without 20 S RNA virus. Typical of fungal viruses, they are not infectious and have no extracellular pathways of transmission. They are transmitted horizontally through mating or vertically from mother to daughter cells. No curing methods to eliminate these viruses from yeast are known so far. Furthermore, 20 S and 23 S RNA viruses do not confer any phenotypic changes on the host. This makes their genetic manipulation difficult.

The 20 S and 23 S RNA genomes are small (2514 and 2891 nucleotides, respectively), and each RNA encodes only a single protein: a 91-kDa protein (p91) and a 104-kDa (p104), respectively (5–8). Both proteins contain four amino acid motifs well conserved among RNA-dependent RNA polymerases (9). When yeast cells are grown at 37°C, the cells accumulate double-stranded RNAs called W and T (10), the double-stranded forms of 20 S and 23 S RNA genomes, respectively (8). They are not replication intermediates, but by-products (11). Because these viruses do not have genes for capsid proteins, their genomes are not encapsidated into intracellular viral particles (12–14). Instead, the RNAs form ribonucleoprotein complexes with their cognate RNA-dependent RNA polymerases at a 1:1 stoichiometry and reside in the cytoplasm (15). The viral genomes lack poly(A) tails at the 3'-ends and have perhaps no 5' cap structures (8), thus resembling degradation intermediates of mRNAs. How these viruses can survive as persistent viruses in the host cytoplasm without their RNA genomes being digested by exonucleases involved in mRNA degradation is therefore interesting.

Recently, we succeeded in generating 23 S RNA virus in vivo from a vector containing the entire cDNA sequence of the viral genome (16). Using this launching system, we began reverse genetics to investigate cis-acting signals for replication in 23 S RNA virus. 20 S and 23 S RNA genomes share five-nucleotide inverted repeats at the 5' and 3' termini (5'-GGGGC...GCCCC-OH) (8). The 23 S RNA genome contains a bipartite cis-signal in the 3'-region that consists of the cluster of the terminal four Cs and a mismatched pair of purines present in a stem structure adjacent to the 3'-end (17). Although the 3'-terminal and penultimate Cs are dispensable for launching 23 S RNA virus, the generated viruses recover these Cs. This indicates that the virus has an efficient 3'-terminal repair mechanism(s). Subsequently, we found that the bipartite 3'-cis-signal for replication is also essential for formation of ribonucleoprotein complexes in vivo with its RNA-dependent RNA polymerase, p104 (18). This indicates the importance of complex formation for 23 S RNA virus and suggests that p104 protects the viral 3'-ends from degradation by binding to the 3'-cis-signal.

In this work, we describe the generation of 20 S RNA virus in vivo from a vector, a system similar to the one developed previously for 23 S RNA virus (16). In contrast to the previous work, however, we suc-

* This work was supported in part by Grants BMC2001-1065 and BFU2004-00373 from the Spanish Ministry of Education and Science. The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

1 To whom correspondence should be addressed: Inst. de Microbiología Bioquímica, CSIC/Universidad de Salamanca, Avda. del Campo Charro s/n, Salamanca 37007, Spain. Tel.: 34-923-120-673; Fax: 34-923-224-876; E-mail: mrosa@gugu.usal.es.

2 Recipient of a fellowship from the Autonomous Regional Government of Castilla and León (Spain) and the European Fund for Regional Development.

3 Recipient of a contract from the Spanish Research Program "Ramón y Cajal."
ceeding in producing 20 S RNA virus not only from positive strands but also from negative strands if, in the latter case, active p91 was provided in trans from a second vector. Using these systems, we modified the viral 3′-ends and found that the clusters of the 3′-terminal four Cs in both the positive and negative strands are 3′-cis-signals for replication. We also provide, for the first time, conclusive evidence that 23 S RNA virus does not require 20 S RNA virus for replication; thus, they are independent replicons.

**EXPERIMENTAL PROCEDURES**

**Strains and Media**—20 S RNA negative strains 2928-4 and 2928-5 were obtained in this work (see Fig. 1) from strain 2928 L-A-o (a ura3 trp1 his3, 20 S RNA, 23 S RNA-o, L-A-o) (16). These strains were used to analyze launching of 20 S RNA virus from plasmids. Cells were grown in rich YPAD medium (1% yeast extract, 2% peptone, 0.04% adenine sulfate, and 2% glucose) or synthetic medium deprived of tryptophan, uracil, or both (19). Nitrogen starvation in 1% potassium acetate was performed as described previously (4).

**Northern Hybridization**—Cells were broken with glass beads (15), and RNA was extracted from cell lysates once with phenol and twice with phenol/chloroform and precipitated with ethanol. RNA was separated on an agarose gel, blotted onto a neutral nylon membrane (Hybond-N, Amersham Biosciences), and hybridized with a 32P-labeled 20 S RNA or 23 S RNA positive or negative strand-specific probe (20). We analyzed 5–10 independent transformants in each experiment. For the sake of simplicity, only a representative of each experiment is presented in the figures.

**Plasmids**—The standard 20 S RNA virus-launching plasmid pRE740 contains the entire 20 S RNA cDNA sequence (2514 bp) downstream of the PGK1 promoter in a pBluescript KS (+) vector. In the case of negative strand 3′-ends, we used oligonucleotide RE157 and oligonucleotide RE233 (5′-GCTGAGGGAGATCCGATC-3′) and SuperScript II RNase H− reverse transcriptase (Invitrogen). After cDNA synthesis, RNA was digested with RNase A, and the unincorporated primer was eliminated using a Sephadex G-50 mini spin column (Worthington). cDNA containing the 3′-ends of 20 S RNA positive strands was PCR-amplified for 30 cycles using Taq polymerase (Promega) and primers RE157 (5′-GACTCGAGGGAGATCCGATC-3′) and PG6 (5′-CGAATCTGTCGAGATCCGATC-3′). Amplified fragments from positive strand 3′-ends were digested with HindIII and BamHI and ligated into the pBlueScript KS+ vector. In the case of negative strand 3′-ends, PCR products were digested with SalI and BamHI and ligated into the pBlueScript KS+ vector.

**RESULTS**

### 20 S and 23 S RNA Viruses Are Independent Replicons

23 S RNA virus can be generated in vivo from an expression plasmid containing the entire viral cDNA sequence. We wished to establish a similar launching system for 20 S RNA virus to carry out reverse genetics. However, most laboratory strains harbor 20 S RNA virus. During the course of 23 S RNA virus-launching experiments, we happened to obtain cells that had lost endogenous 20 S RNA virus. Strain 2928 L-A-o was transformed with the standard 23 S RNA-launching plasmid pRE637 (16). When transformants were analyzed, all of them generated 23 S RNA virus from the plasmid, and one of them (transformant 1) had apparently lost its endogenous 20 S RNA virus (Fig. 1A). We eliminated the launching plasmid from transformant 1 by isolating single colonies on nonselective YPAD agar plates. When the plasmid-cured colonies were analyzed, we found two types of segregants; the majority of colonies (10 of 12) contained only 23 S RNA virus generated from the plasmid, and the rest harbored neither 20 S nor 23 S RNA virus. Both types of segregants retained the genetic markers of the original strain. The segregants carrying 23 S RNA virus maintained the virus stably and did not produce virus-free colonies anymore. This suggests that the loss of endogenous 20 S RNA virus preceded the generation of 23 S RNA virus from the launching plasmid.

We confirmed the curing of 20 S RNA virus in these segregants by three criteria. First, we did not detect 20 S RNA by Northern blotting (Fig. 1B). Even when the gels were overloaded, we found no signal for 20 S RNA. The conditions used should have allowed us to detect 20 S RNA if the virus were present at as low as one copy/cell. Second, we could not amplify 20 S RNA cDNA from the lysates of these cells by RT-PCR (data not shown). Finally, as will be shown below, we could generate 20 S RNA virus tagged with a silent mutation from a launching plasmid using these 20 S RNA-negative segregants as hosts. We confirmed that the virus retained the mutation and was not the result of amplification of any residual endogenous 20 S RNA virus. Generated 20 S RNA virus could be maintained stably and was also induced normally under nitrogen starvation conditions in these cells. Therefore, the 20 S RNA-negative segregants have no genetic defects to harbor 20 S RNA virus. We designated segregants 4 and 5 shown in Fig. 1A as 2928-4 and 2928-5, respectively. These segregants showed the same doubling time as the original strain, and we noticed no phenotypic changes in them.

Strain 2928-5 carries only 23 S RNA virus derived from plasmid pRE637, but no 20 S RNA virus. Strain 2928-4 harbors none of the narnaviruses and can be used successfully as a host to generate 23 S RNA virus from the 23
S RNA-launching plasmid (17). Although it has been reported that some industrial yeast strains (Saccharomyces diastaticus) appear to harbor only 23 S RNA-like molecules (24), these results are the first conclusive evidence that 23 S RNA virus does not require 20 S RNA virus for replication. Therefore, 20 S and 23 S RNA viruses are independent replicons.

**Launching of 20 S RNA Virus from Expression Plasmids—** We constructed a launching plasmid (pRE740) to generate 20 S RNA virus in yeast by inserting the entire 20 S RNA cDNA sequence (2514 bp) downstream of the constitutive PGK1 promoter (Fig. 2A). Positive strands of 20 S RNA can be transcribed from the promoter. To generate transcripts in vivo with the precise viral 3′-end sequence at the 3′ termini, an 82-nucleotide HDV antigenomic ribozyme cDNA sequence (25) was directly fused to the 3′-end of the 20 S RNA sequence. The plasmid has the TRP1 gene as a selective marker. We transformed strains 2928-4 and 2928-5 with pRE740. Transformants were transferred to 1% potassium acetate to induce 20 S RNA, and the RNA was extracted from the induced cells. By Northern hybridization, we detected single-stranded 20 S RNA positive strands in all of the transformants. The amounts of positive strands among them were similar (Fig. 2B, lanes 1–4). We also found small amounts of W, the double-stranded form of 20 S RNA, in all transformants. Furthermore, we detected single-stranded 20 S RNA negative strands with a negative strand-specific probe (data not shown).
These results indicate that 20 S RNA negative strands are synthesized from the positive strand transcripts in vivo and suggest the generation of 20 S RNA virus from the plasmid. To confirm this, we took two experimental approaches. First, we cured the plasmid from the transformants by isolating single colonies in nonselective YPAD agar plates. More than 70% of the plasmid-cured colonies retained 20 S RNA virus. Once generated, the virus replicated autonomously and could be maintained stably in the cells for >100 generations (so far examined) in the absence of the plasmid. The generated virus could be induced under nitrogen starvation conditions, and the viral RNA could be seen directly by ethidium bromide staining on agarose gels. The amount of induced RNA and the positive/negative strand ratios were similar to those found in endogenous 20 S RNA virus. We did not notice any effect of 23 S RNA virus on the efficiency of 20 S RNA virus generation (Fig. 2B, compare lanes 1 and 2 with lanes 3 and 4). Second, we introduced a unique SmaI site into the 20 S RNA cDNA sequence in pRE740 by changing the T at position 1476 to G (numbered from the 20 S RNA 5′-end) (Fig. 2C). This marking did not alter the amino acid sequence of the encoded p91 protein.

We transformed strains 2928-4 and 2928-5 with the SmaI-tagged plasmid. Transformants from both strains generated 20 S RNA virus. After curing the plasmid, RNA was extracted from both strains, and an 842-bp 20 S RNA cDNA fragment encompassing the SmaI site was amplified by RT-PCR. We also amplified the 842-bp fragment from endogenous 20 S RNA virus present in the original 2928 L-A-o strain as a control. As shown in Fig. 2C (lower panel), the amplified cDNA fragment from generated 20 S RNA virus in strain 2928-4 (lane 1) or 2928-5 (lane 2) was completely digested with SmaI, whereas the control from endogenous 20 S RNA virus was fully resistant to the enzyme (lane C).

Therefore, these results demonstrate that 20 S RNA virus was generated from the launching plasmid and also confirm that the host strains 2928-4 and 2928-5 do not carry any endogenous 20 S RNA virus.

**p91 Is Essential**—20 S RNA encodes a single protein, p91. p91 has the four amino acid motifs well conserved among RNA-dependent RNA polymerases. When one of the motifs (60GDD62) was changed to EF062 in the launching plasmid, the modified vector failed to produce 20 S RNA virus (Fig. 2B, lane 5), indicating that active p91 is essential for replication. The presence of 23 S RNA virus in the host could not rescue the modified plasmid to generate 20 S RNA virus (data not shown). This indicates that p104 encoded by 23 S RNA cannot substitute p91 for replication of 20 S RNA virus and is consistent with the fact that 20 S and 23 S RNA viruses are independent replicons. p91 and p104 share an eight-amino acid sequence (R(V/I)CGDLI) surrounding the GDD motif with only one mismatch. When the mismatched amino acid (Val659) in p91 was replaced with Ile, thus making the stretch identical to the one found in p104, the plasmid with the modified p91 sequence produced 20 S RNA virus without any deleterious effects (Fig. 2B, lane 6). Previously, we found that 23 S RNA virus can also tolerate the reverse modification (Ile to Val) in the eight-amino acid sequence in p104 (16).

**A Cluster of Four Cs at the 3′-End of the Positive Strand Is Essential for Replication**—The transcription termination site for the FLP gene of the 2μ plasmid is located 0.7 kb downstream of the 20 S RNA genome in the vector (26). When the ribozyme GGG sequence 3′ to the cleavage site was substituted with AAA in the standard launching plasmid pRE740, the modified plasmid failed to generate 20 S RNA virus (data not shown). Therefore, it is important to generate transcripts in vivo with the precise viral 3′ termini for virus launching. 20 S RNA has five-nucleotide inverted repeats at both termini (5′-GGGCG...GCCCG-3′). We examined the role of the 3′-terminal nucleotides of the 20 S RNA positive strand in replication by modifying each nucleotide with A in the vector. As shown in Fig. 3, changing the 3′-terminal or penultimate C did not affect the generation of 20 S RNA virus. On the contrary, a modification at the third or fourth C did not produce the virus. Therefore, the third and fourth Cs from the 3′-end are essential for replication. In good agreement, deletions of up to two Cs (but not three) generated 20 S RNA virus (Fig. 3) (data not shown). We were interested in the 3′-terminal sequences of 20 S RNA viruses generated from these modified vectors. We cured the launching plasmids from the cells, and the viral RNAs were isolated. The 3′-terminal sequences of these RNAs were then amplified by 3′-RACE, and 8–10 independently isolated clones were sequenced. As shown in TABLE ONE, the substitution of the terminal (pRE757) or penultimate (pRE758) C with A at the viral 3′-end was corrected to the wild-type C in the generated viruses. Furthermore, viruses launched from 20 S RNA cDNA lacking two Cs at the 3′-end (pRE759) recovered these nucleotides. Therefore, these results indicate that 20 S RNA virus, like 23 S RNA virus, has an efficient 3′-end repair mechanism(s) in vivo. The major transcription start site in the launching plasmid is located at position −42, relative to the 5′ terminus.

**Launching of 20 S RNA Virus from Positive/Negative Strands**—To validate the significant role of nucleotides at the 3′-end of the positive strand for replication by ribozyme cleavage, we introduced a unique SmaI site into the positive strand transcripts in pRE759 by modifying each nucleotide with A and C. Transformants from both strains generated 20 S RNA virus. After curing the plasmid, RNA was extracted from both strains, and 20 S RNA was detected by Northern blotting using a positive strand-specific probe. dsRNA, double-stranded RNA.
of the 20 S RNA genome. These extra nucleotides were not present in the generated viruses as judged from the 3′-sequences of the negative strands (TABLE ONE). We also added extra nucleotides between the 3′-end of 20 S RNA and the riboyme sequence and examined their effects on virus launching. As shown in Fig. 3, the addition of a single U (lane 6), C (lane 7), or A (data not shown) did not affect the generation of 20 S RNA virus. However, the addition of a single G reduced virus launching (lane 8). Likewise, the addition of three Gs (lane 10), but not three Cs (lane 9), severely affected 20 S RNA virus generation. Extra Gs at the 3′-end reduced the efficiency of virus launching by severalfold as judged by counting virus-positive and virus-negative colonies after curing the plasmids. Once the plasmids were cured, however, the amount of 20 S RNA in the cells was similar to that generated from the wild-type cDNA without extra nucleotides at the 3′-end. We isolated viral RNA from plasmid-cured cells and analyzed the 3′terminal sequences by 3′-RACE. We sequenced 17 independently isolated clones. All of them retained the compensatory mutation 35C→G at the 3′ side of the stem. Of these 17, 11 clones had the full 3′-terminal sequence ( … GCCCCC-OH), and the rest possessed the 3′ termini truncated by one nucleotide ( … GCCCC_OH; four clones) or by two nucleotides ( … GCCC_OH; two clones). Therefore, these results show that no revertants were generated during virus launching. Furthermore, as the majority (11 of 17) retained the full 3′-terminal sequence, the sixth G from the 3′-end did not substitute for the modified nucleotide at position 5 in replication.

| Launching plasmida | Clones | Positive strandb | Clones | Negative strandb |
|--------------------|--------|------------------|--------|------------------|
| pRE757 (…GCCCA-OH) | 8      | 7_GCCCG-OH       | 3      | 3_GCCCG-OH       |
| pRE758 (…GCCCA-OH) | 8      | 7_GCCCG-OH       | 3      | 2_GCCCG-OH       |
| pRE759 (…GCC__-O H) | 8      | 7_GCCCG-OH       | 2      | 1_GCCCG-OH       |
| pRE794 (…GCCCCU-O H) | 9      | 7_GCCCG-OH       | ND     | ND               |
| pRE800 (…GCCCCG-O H) | 10     | 7_GCCCG-OH       | ND     | ND               |
| pRE795 (…GCCCCCC-O H) | 10     | 7_GCCCG-OH       | ND     | ND               |
| pRE793 (…GCCCGGG-O H) | 10     | 8_GCCCG-OH       | ND     | ND               |

* Modifications of or extra nucleotides added at the 3′-end of the 20 S RNA positive strand in the launching vectors are underlined.
* Extra sequences found in the viral 3′-ends cloned are indicated by parentheses.
* Not determined.

The Fifth G from the 3′-End in the Positive Strand Is Dispensable for Replication—When the fifth G from the 3′-end was changed to C (5G→C, numbered from the 3′-end in the vector), the modified plasmid failed to generate 20 S RNA virus (Fig. 4, lane 1), thus suggesting that this nucleotide is essential for replication. Alternatively, because the fifth G is part of a structure adjacent to the 3′-end, the inability of this plasmid to generate the virus may be a secondary effect due to a perturbation of the stem structure caused by the modification. Because the corresponding nucleotide in the 23 S RNA genome 3′-end is dispensable for its replication (17), we were interested in this result and tried to distinguish these two possibilities. To restore the stem structure in the 5G→C mutant, we introduced a second mutation, 35C→G, at the other side of the stem (lane 3). Because this compensatory mutation resulted in the change of one amino acid (Thr223 to Ser) at the C terminus of p91, we made two more constructs as controls. One construct contained the single mutation 36A→U with the wild-type stem and wild-type … GCCCC-OH terminus and, as shown in lane 4, produced 20 S RNA virus. Thus, p91 can tolerate the amino acid change T823S for its activities. The second control contained the single mutation 35C→G with the wild-type … GCCCC-OH terminus. This construct, as shown in lane 2, failed to generate 20 S RNA virus. Therefore, these control experiments suggest that the fifth nucleotide needs to be hydrogen-bonded. Consistently, the compensatory mutation 35C→G restored the ability to generate virus on the 5G→C mutant (lane 3). This result clearly indicates that, as in the case of 23 S RNA virus, the fifth G in 20 S RNA virus is dispensable for replication. To rule out the possibility that this double mutant generated replicable revertants during launching, we isolated viral RNA after curing the plasmid and determined the 3′-terminal sequences by 3′-RACE. We sequenced 17 independently isolated clones. All of them retained the compensatory mutation 35C→G at the 3′-side of the stem. Of these 17, 11 clones had the full 3′-terminal sequence ( … GCCCCC-OH), and the rest possessed the 3′ termini truncated by one nucleotide ( … GCCCC_OH; four clones) or by two nucleotides ( … GCCC_OH; two clones). Therefore, these results show that no revertants were generated during virus launching. Furthermore, as the majority (11 of 17) retained the full 3′-terminal sequence, the sixth G from the 3′-end did not substitute for the modified nucleotide at position 5 in replication.
FIGURE 5. Generation of 20 S RNA virus from negative strands in a two-vector system. A, the template plasmid (pRE805) has URA3 as a selective marker and contains the entire 20 S RNA cDNA sequence downstream of the PGK1 promoter. The orientation of 20 S RNA is reversed so that the negative strand of the viral genome can be transcribed from the promoter. The viral genome was tagged by a silent mutation (1476T→G, numbered from the 5′-end of the positive strand) to create a unique SmaI site. The HDV antigenomic ribozyme (R) is directly fused to the 3′-end of the negative strand. The polymerase plasmid (pRE787) has a TRP1 selective marker and is a derivative of the standard launching plasmid pRE740 with a 4C→3A mutation at the 3′-end of the positive strand genome (indicated by the asterisk). Transcripts from this plasmid can be translated into wild-type p91 but could not generate 20 S RNA virus because of the 4C→3A mutation (see Fig. 3C, lane 4). The 20 S RNA genome in the polymerase plasmid has no SmaI site.

B, cells were transformed with either the template plasmid (pRE805) or the polymerase plasmid (pRE787) alone or together. As a control, cells were also transformed with the standard launching plasmid pRE740.

RNA was extracted from transformants, and 20 S RNA was detected by Northern blotting using a positive strand-specific probe as described in the legend to Fig. 1A.

dsRNA, double-stranded RNA.

C, proof that the virus was generated from negative strands is shown. RNA was prepared from 20 S RNA-generated cells from which both plasmids had been cured, and an 842-bp cDNA fragment encompassing the SmaI site was obtained by RT-PCR (lanes 2 and 4). As a control, we also amplified the cDNA fragment from RNA extracted from 2928 L-A-o cells carrying its endogenous 20 S RNA virus (lanes 1 and 3). The cDNA fragments, either undigested (lanes 1 and 2) or digested with SmaI (lanes 3 and 4), were separated on an agarose gel. Ethidium bromide staining of the gel is shown.
Launching of 20 S RNA Virus from Positive/Negative Strands

A

(-) strand transcripts 3' ends

\[ C\ldots GCCC\cdots OH \]

B

\[ \begin{array}{c}
-\text{strand specific probe} \\
W\text{dsRNA} \\
20\text{S RNA}\ldots
\end{array} \]

\[ \begin{array}{cccccc}
1 & 2 & 3 & 4 & 5 \\
\cdot & \cdot & \cdot & \cdot & \cdot
\end{array} \]

C

\[ \begin{array}{cccccc}
\text{W dsRNA} & \cdot & \cdot & \cdot & \cdot & \cdot
\end{array} \]

\[ \begin{array}{cccccc}
\text{20 S RNA} & \cdot & \cdot & \cdot & \cdot & \cdot
\end{array} \]

The A, B, and C diagrams show the expected 3' ends at high frequencies (50% in some cases). Positive strands cloned from the same RNA preparations had an extra G at the 3' ends, but at much lower frequencies. At present, we do not know the biological significance of this observation.

DISCUSSION

In this work, we have described the generation of 20 S RNA virus from a vector by expressing the viral RNA in vivo. The RNA can be either the positive or negative strand of the 20 S RNA genome. In the latter case, active p91 needs to be provided in trans from a second vector. Once 20 S RNA virus is generated, however, the virus does not require these vectors any longer and can be propagated autonomously. Because 20 S RNA virus does not confer any phenotypic changes on the host, it has been impossible to perform comprehensive genetic analysis on this virus. Therefore, these launching systems have allowed us to carry out reverse genetics to determine and analyze cis-acting signals necessary for replication present not only in the positive strand but also in the negative strand of the viral genome.

Host Strains—During the course of experiments launching 23 S RNA virus from a vector, we happened to obtain 20 S RNA-cured cells. Interference in viral replication by defective interfering particles or by other viruses is well known among DNA and RNA viruses. In yeast, L-A and L-BC double-stranded RNA viruses can be eliminated from the cells by transcribing L-A and L-BC positive strand RNAs, respectively, from a vector (27, 28). Although 20 S and 23 S RNA viruses are compatible in the same host and do not exclude each other, their copy numbers in logarithmically growing cells are extremely low compared with those under induction conditions (nitrogen starvation). Therefore, overexpression of 23 S RNA or its gene product (p104) from a vector in growing cells may have taken up a hypothetical limiting cellular factor shared by 20 S and 23 S RNA viruses necessary for replication or for stable transmission to daughter cells. We have not yet determined the frequency of exclusion, but it appears to be quite low.
p91 Is Essential—When the GDD sequence in p91 was changed to EF in the vector, the modified launching plasmid failed to generate 20 S RNA virus. Crystallographic studies have demonstrated that RNA-dependent RNA polymerases of RNA viruses, like DNA-dependent DNA or RNA polymerases or reverse transcriptases, form tertiary structures resembling a right hand (29–31). The four amino acid motifs well conserved among RNA-dependent RNA polymerases, including the GDD motif, are located within a subdomain called the palm. The first aspartic acid residue of the GDD motif coordinates two Mg$^{2+}$ ions, which are essential for the catalytic activity in polymerization. Therefore, the failure of the EFD mutant to generate 20 S RNA virus is due to destruction of the catalytic site of RNA polymerization in p91. Short stretches of amino acids flanking the GDD motif form a β-hairpin structure with GDD in the hairpin loop. p91 and p104 polymerases share the same eight-amino acid sequence (R(V/I)CGDDLI) surrounding the GDD motif with one mismatch. When the mismatched Val in p91 was changed to Ile, thus making the eight-amino acid stretch identical to the one in p104, p91 tolerated this change, and 20 S RNA virus was generated from the modified vector. Previously, we found that 23 S RNA virus also tolerates the reverse change (Ile to Val) in the same stretch in p104 (16). Because these aliphatic amino acids are favored for β-sheet formation, these exchanges apparently do not significantly affect the overall structure of each polymerase. These results indicate the potential usefulness of the launching systems of 20 S and 23 S RNA viruses for structure-function analysis of these polymerases. Furthermore, as p91 and p104 also form ribonucleoprotein complexes with their cognate viral RNAs, these launching systems will be powerful tools for assigning and analyzing domains with distinct functions in these polymerases.

Positive Strand 3′-cis-Signal—In this work, we developed two launching systems to generate 20 S RNA virus in vivo: one from the genomic strands and the other from the antigenomic strands. The former system is similar to the one established previously to generate 23 S RNA virus (16). Using this launching system, we analyzed 3′-cis-signals for replication present in the 20 S RNA positive strand. 20 S and 23 S RNA viruses share the same five-nucleotide inverted repeats at both termini (5′-GGGCC...GCCCC-OH) (Table 2). Our results indicate that the third and fourth Cs from the 3′-end of the positive strand are essential for replication, whereas the terminal and penultimate Cs are dispensable, at least for virus launching. In the latter cases, however, we found that the viruses regained the wild-type four Cs. Thus, 20 S RNA virus, like 23 S RNA virus, has an efficient repair mechanism(s) for the viral 3′-ends, and the stretch of four Cs at the 3′-end of the 20 S RNA positive strand is essential for replication. In contrast to these terminal four Cs, the fifth G is not essential; however, it needs to be hydrogen-bonded with the nucleotide at the other side of the stem (Fig. 4). There are two probable explanations for this. 20 S and 23 S RNA viruses are closely related but independent replicons. Therefore, their replication machineries have to discriminate each other and recognize their own RNA template for replication. 23 S RNA virus has a bipartite 3′-cis-signal that consists of a stretch of four Cs at the 3′-end and a mismatched pair of purines in a stem-loop structure adjacent to the 3′-end (Fig. 7). In 23 S RNA virus, the fourth C from the 3′-end is hydrogen-bonded and located at the lower edge of the stem structure. Therefore, the replication machinery of 20 S RNA virus may recognize a stretch of single-stranded four Cs at the 3′-end flanked by a stem structure at the 5′-side without a spacer as a cis-signal for replication. Alternatively, 20 S RNA virus may also have a bipartite 3′-cis-signal with the second site yet to be located in the stem-loop structure adjacent to the 3′-end. Thus, elimination of the hydrogen bonding at the lower edge of the stem may alter replication of the virus.

### TABLE TWO

| Template plasmid<sup>a</sup> | Clones | Negative strand<sup>b</sup> | Clones | Positive strand<sup>b</sup> |
|-----------------------------|--------|-----------------------------|--------|-----------------------------|
| pLOR2 (..GCCCA-OH)         | 7      | 3..GCCCC-OH                 | 6      | 4..GCCCC-OH                 |
|                             |        | 3..GCCCC(G)-OH              |        | 1..GCCCC(G)-OH              |
|                             |        | 1..GCC-OH                   |        | 1..GCC-OH                   |
| pLOR4 (..GCCCC-OH)         | 12     | 6..GCCCC-OH                 | 10     | 7..GCCCC-OH                 |
|                             |        | 6..GCCCC(G)-OH              |        | 2..GCCCC(G)-OH              |
|                             |        | 1..GCC-OH                   |        | 1..GCC-OH                   |

<sup>a</sup> Modifications introduced at the 3′-end of the 20 S RNA negative strand in the template plasmid are underlined.
<sup>b</sup> Extra sequences found in the viral 3′-ends cloned are indicated in parentheses.

### FIGURE 7.

Comparison of the 3′-terminal secondary structures in the positive and negative strands of 20 S and 23 S RNA viruses with the top half-domain of tRNA. The non-templated As at the viral 3′ termini are indicated by parentheses. The 3′-cis-acting signals for replication identified in this work present in the positive and negative strands of 20 S RNA virus are boxed. The bipartite 3′-cis-signal present in the positive strands of 23 S RNA virus is also boxed and circled. The top half-domain of eukaryotic tRNA<sup>Thr</sup> was derived from Sprinzl et al. (43). Y, R, and N stand for pyrimidine, purine, and any base, respectively.
the bipartite signal by shortening the distance between the two sites or changing their topological configuration. 23 S RNA virus did not tolerate a change of adding or eliminating 2 base pairs at the lower stem that separates both sites.

Launching of 20 S RNA Virus from Antigenomic Strands—In the second launching system, we used two vectors to separately express the 20 S RNA antigenomic strands and the polymerase. Either vector alone could not generate the virus. However, when these two plasmids were introduced together into the same host, the cells generated 20 S RNA virus. That the virus was not generated by recombination between these two plasmids or between their transcripts is illustrated by the following evidence. First, transcripts from these two vectors had opposite polarities. Therefore, the 20 S RNA wild-type genome cannot be produced by recombination between them. Second, the high efficiency of virus generation (as efficient as that from a single vector expressing positive strands) suggests that the virus was not generated through rare events such as RNA-RNA recombination. Third, when the essential GDD motif of p91 in the polymerase plasmid was changed to EFD, the modified plasmid failed to assist virus generation. This indicates the importance of an active p91 protein rather than its DNA or RNA sequence in the polymerase plasmid for virus generation. Fourth, when the ribozyme GGG sequence 3' to the cleavage site was changed to AAA in the template plasmid, thus destroying its cleavage activity, the modified template plasmid failed to generate 20 S RNA virus. This indicates that the negative strand transcripts need to be processed by the ribozyme to produce correct 3'-ends at the RNA termini. Finally, we found that the viruses retained the silent mutation (1476T→G) tagged in the template plasmid. This indicates that their RNA genome derived from the negative strands transcribed from the template plasmid. This also eliminates the possibility that the 4C→A mutation in the polymerase plasmid was rescued by recombination between the two plasmids. Because the GDD motif and the silent mutation were separated by only 85 nucleotides in the viral genome, such recombination should also efficiently rescue the EFD mutation in the polymerase plasmid.

Using this two-vector system, we found that the third and fourth Cs from the 3'-end of the negative strand are essential for replication. The 3'-terminal and penultimate Cs were dispensable for virus launching; however, the generated viruses regained the wild-type Cs at these positions. Therefore, the cluster of four consecutive Cs at the 3'-end of the negative strand is again a 3'-cis-signal for replication or part of it.

3'-End Repair and Viral Persistence—20 S RNA virus can repair mutations introduced at the terminal or penultimate position from the 3'-end or deletion of up to two nucleotides from the 3'-end in the positive or negative strand. Likewise, the addition of three extra Cs to the 3'-end of the positive strand did not affect virus launching, and the virus generated possessed the wild-type termini without these extra nucleotides. Therefore, 20 S RNA virus has a mechanism(s) to maintain not only the correct sequence but also the correct length of the viral 3'-ends. Interestingly, the addition of three extra Gs reduced the efficiency of virus generation, although the virus generated did not retain these extra Gs at the 3'-end. We observed previously that the addition of a row of eight Gs to the 5'-end of the 23 S RNA cDNA genome reduces the generation of 23 S RNA virus from a launching vector (16). However, the addition of eight Cs does not have such an effect. It is known that oligo G tracts inhibit the progression of 5'- and 3'-exonucleases in yeast (32, 33). This is thus consistent with the idea that the 5'- and 3'-non-viral sequences in the transcripts after ribozyme cleavage were eliminated by host exonucleases. The launching vectors expressing 20 S RNA positive and negative strands contained non-viral sequences of 42 and 41 nucleotides, respectively, between the major transcription start site and the 5'-end of the viral genome. Although these extra sequences are almost 10 nucleotides longer than the one present in the 23 S RNA-launching vector, we found that launching of 20 S RNA virus from these vectors was as efficient as (or better than) the generation of 23 S RNA virus observed previously. On the other hand, the addition of one or three extra Gs to the 3'-end greatly reduced the generation of 20 S RNA virus. These results correlate with the fact that the SK11/XRN1 5'-exonuclease is more active than the 3'-exonuclease (exosome) in mRNA degradation in yeast (34).

A mutation introduced at the 3'-terminal or penultimate position could be mended by the polymerase machinery. RNA polymerases are known to incorporate mismatched nucleotides at high frequency. Alternatively, the machinery could synthesize oligonucleotides in a template-independent manner and use those oligonucleotides as a primer to initiate complementary strand synthesis (35). Host enzymes might also be involved in the 3'-end repair. Previously, we pointed out the similarity between the 3'-end stem-loop structures present in the positive and negative strands of 23 S RNA virus and the top half-domain of tRNA (17). The acceptor stem and T stem of tRNA stack on each other coaxially and form a stem-loop structure with a single-stranded 3'-tail referred as the top half of tRNA (Fig. 7). The structure consists of a long stem of 12–13 bp with a seven-nucleotide TVC loop. The top half-domain contains the determinants necessary for specific interactions with tRNA-related enzymes, including the CCA-adding enzyme (36, 37). The positive strand of 20 S RNA virus also possesses a stem-loop structure with a tail of four single-stranded Cs at the 3'-end that resembles the top half-domain of tRNA. This molecule could be a substrate of the CCA-adding enzyme (Fig. 7). A mutation introduced at the 3'-end of the positive strand may be excised by an exonuclease. It is known that RNase T, a non-processive 3'-exonuclease, is responsible for the end turnover of tRNA in Escherichia coli. The progression of the cleavage by this enzyme is blocked by consecutive Cs (38). Then, the shortened 3'-end would be filled to the wild-type sequence by the CCA-adding enzyme. Consistent with this possibility is the evidence that 20–30% of 20 S RNA positive and negative strands possessed a non-templated A at their 3'-ends when W double-stranded RNA was analyzed after pCp labeling (39). The negative strand of 20 S RNA has a similar secondary structure at the 3'-end but with a much shorter stem (8 bp). This raises the possibility that the negative strand is not a good substrate for the CCA-adding enzyme. In this work, we observed that the negative strands of 20 S RNA viruses generated from a vector possessed, at a high frequency, a non-templated G at the 3'-ends (TABLE TWO). Because endogenous 20 S RNA virus also possesses this G in the negative strands at high frequency, it is not a phenomenon related to virus generation from a vector. In E. coli, it is known that poly(A) polymerase I and polynucleotide phosphorylase can partially compensate for the absence of the CCA-adding enzyme in the repair of the 3'-terminal sequence of tRNA (40, 41). Therefore, the extra G at the 3'-end could be the manifestation that an enzyme other than the CCA-adding enzyme is involved in the repair of the 3'-end of the negative strand and that the virus somehow tolerated this G during its replication. Alternatively, because most of the negative strands in the lysates are present in replication complexes engaged in positive strand synthesis (11), this G may be involved in the replication process.

The positive strands of 20 S and 23 S RNA viruses form ribonucleoprotein complexes with their cognate RNA-dependent RNA poly-

---

5 R. Esteban, M. Ramirez-Garrastacho, and T. Fujimura, unpublished data.

6 L. Vega, T. Fujimura, and R. Esteban, unpublished data.
merases at a 1:1 stoichiometry and reside in the host cytoplasm. These RNAs have no poly(A) tails and perhaps no 5’-cap structures, thus resembling degradation intermediates of mRNAs. We expect them to be vulnerable to the exonucleases involved in mRNA turnover. This raises the possibility that the formation of ribonucleoprotein complexes is essential for virus persistence by protecting these RNA genomes in the host cytoplasm. The positive strand of 23 S RNA virus possesses the bipartite 3’-cis-acting signal for replication (Fig. 7). Previously, we found that the bipartite 3’-signal is also important for the formation of a complex between the 23 S RNA genome and its RNA-dependent RNA polymerase, p104 (18). The mismatched pair of purines and the third and fourth Cs from the 3’-end are essential for this activity. A mutation at the penultimate position significantly reduced the activity, whereas a modification at the 3’-terminal C did not affect the formation of the complex. These results suggest that p104 directly interacts with the third and fourth Cs from the 3’-end and protects the viral RNA from exonucleases. The terminal and penultimate nucleotides at the 3’-end may be accessible to the 3’-repair machinery. As demonstrated in this work, the stretch of four Cs at the 3’-end in the 20 S RNA positive strand is also a cis-acting signal for replication. Based on analogy to 23 S RNA virus, we expect that p91 interacts with the third and fourth Cs and protects the viral 3’-end from exonucleases. Because the negative strands are replication intermediates, they also need to be protected in the cytoplasm.

The negative strands of 20 S RNA virus also form complexes with p91, as anti-p91 antiserum can immunoprecipitate the negative strands from cell lysates (42). It is likely that the stretch of four Cs at the 3’-end in the negative strand is also involved in the formation of complexes with p91.

The 20 S RNA genome contains only 12 nucleotides in the 3’-non-coding region (Fig. 3). This makes the comprehensive analysis of the 3’-cis-acting signals (for replication or for complex formation) quite difficult without modifying the amino acid sequence of p91. As shown in this work, however, p91 can act in trans to generate 20 S RNA virus from two vectors. Therefore, it will be possible to develop an in vivo system to analyze these cis-acting signals by separately expressing the viral RNA and the polymerase p91 from two vectors. These experiments will provide information essential to understanding the molecular basis of narnavirus replication and its persistence in yeast.

REFERENCES

1. Buck, K. W. (1996) Adv. Virus Res. 47, 159–251
2. Wickner, R. B., Esteban, R., and Hillman, B. I. (2000) in Viral Taxonomy: Seventh Report of the International Committee on Taxonomy of Viruses (Van Regenmortel, M. H. V., Fauquet, C., Bishop, D. H. L., Carstens, E., Estes, M., Lemon, S. M., Maniloff, J., Mayo, M., McGeon, D., Pringle, C., and Wickner, R. B., eds) pp. 651–656, Academic Press, Inc., New York
3. Kadowaki, K., and Halvorson, H. O. (1971) J. Bacteriol. 105, 826–830
4. Weiksnora, P. J., and Haber, J. E. (1978) J. Bacteriol. 134, 246–260
5. Rodríguez-Cousiño, N., Esteban, L. M., and Esteban, R. (1991) J. Biol. Chem. 266, 12772–12778
6. Matsumoto, Y., and Wickner, R. B. (1991) J. Biol. Chem. 266, 12779–12783
7. Esteban, L. M., Rodríguez-Cousiño, N., and Esteban, R. (1992) J. Biol. Chem. 267, 10874–10881
8. Rodríguez-Cousiño, N., Solórzano, A., Fujimura, T., and Esteban, R. (1998) J. Biol. Chem. 273, 20363–20371
9. Esteban, R., Rodríguez-Cousiño, N., and Esteban, L. M. (1993) Prog. Nucleic Acid Res. Mol. Biol. 46, 155–182
10. Wesolowski, M., and Wickner, R. B. (1984) Mol. Cell. Biol. 4, 181–187
11. Fujimura, T., Solórzano, A., and Esteban, R. (2005) J. Biol. Chem. 280, 7398–7406
12. Widner, W., Matsumoto, Y., and Wickner, R. B. (1991) Mol. Cell. Biol. 11, 2905–2908
13. Esteban, L. M., Fujimura, T., García-Cuáller, M. P., and Esteban, R. (1994) J. Biol. Chem. 269, 29771–29777
14. García-Cuáller, M. P., Esteban, L. M., Fujimura, T., Rodríguez-Cousiño, N., and Esteban, R. (1995) J. Biol. Chem. 270, 20084–20089
15. Solórzano, A., Rodríguez-Cousiño, N., Esteban, L. M., and Fujimura, T. (2000) J. Biol. Chem. 275, 26248–26255
16. Esteban, R., and Fujimura, T. (2003) Proc. Natl. Acad. Sci. U. S. A. 100, 2568–2573
17. Fujimura, T., and Esteban, R. (2004) J. Biol. Chem. 279, 13215–13223
18. Fujimura, T., and Esteban, R. (2004) J. Biol. Chem. 279, 44219–44228
19. Wickner, R. B. (1980) Cell 21, 217–226
20. Fujimura, T., Esteban, R., Lemos, T. M., and Wickner, R. B. (1990) Cell 62, 819–828
21. Wickner, R. B., Icho, T., Fujimura, T., and Widner, W. R. (1991) J. Virol. 65, 155–161
22. Esteban, R., Fujimura, T., and Wickner, R. B. (1989) EMBO J. 8, 947–954
23. Maniatis, T., Fritsch, E. F., and Sambrook, J. (1982) Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY
24. López, V., Gil, R., Carbonell, J. V., and Navarro, A. (2002) Yeast 19, 545–552
25. Perotta, A. T., and Been, M. D. (1991) Nature 350, 434–436
26. Hartley, J. L., and Donelson, J. E. (1980) Nature 286, 860–865
27. Valle, R. P., and Wickner, R. B. (1993) J. Virol. 67, 2764–2771
28. Yao, W., Muqtaidir, K., and Brunn, J. A. (1995) J. Virol. 69, 1917–1919
29. Bressanelli, S., Tomei, L., Roussel, A., Incitti, I., Vitale, R. L., Mathieu, M., De Francesco, R., and Rey, F. A. (1999) Proc. Natl. Acad. Sci. U. S. A. 96, 13034–13039
30. Butcher, S. J., Grimes, J. M., Makeyev, E. V., Bamford, D. H., and Stuart, D. I. (2001) Nature 410, 235–240
31. Ng, K. C., Cheryn, M. M., Vazquez, A. L., Machin, A., Alonso, J. M., Parra, F., and James, M. N. (2002) J. Biol. Chem. 277, 1381–1387
32. Verkhen, P., and Rané, H. A. (1992) Mol. Cell. Biol. 12, 2986–2996
33. Decker, C. I., and Parker, R. (1993) Genes Dev. 7, 1632–1643
34. Jacobs Anderson, J. S., and Parker, R. (1998) EMBO J. 17, 1497–1506
35. Guan, H., and Simon, A. E. (2000) Proc. Natl. Acad. Sci. U. S. A. 97, 12451–12456
36. Shi, P. Y., Maizels, N., and Weiner, A. M. (1998) EMBO J. 17, 3197–3206
37. Shi, P. Y., Weiner, A. M., and Maizels, N. (1998) RNA (N. Y.) 4, 276–284
38. Zuo, Y., and Deutscher, M. P. (2002) J. Biol. Chem. 277, 29654–29661
39. Rodríguez-Cousiño, N., and Esteban, R. (1992) Nucleic Acids Res. 20, 2761–2766
40. Reuten, N. B., Zhou, Z., and Deutscher, M. P. (1997) J. Biol. Chem. 272, 33255–33259
41. Mohanty, B. K., and Kushner, S. R. (2000) Proc. Natl. Acad. Sci. U. S. A. 97, 11966–11971
42. García-Cuáller, M. P. (1996) In Vitro Replication of 20 S RNA Virus from Saccharomyces cerevisiae. Doctoral dissertation, University of Salamanca, Salamanca, Spain
43. Sprinzl, M., Hartmann, T., Weber, J., Blank, J., and Zeidler, R. (1989) Nucleic Acids Res. 17, r1–r172