Both yeast W double-stranded RNA and its single-stranded form 20S RNA are linear

N.Rodriguez-Cousíño and R.Esteban*
Instituto de Microbiologia Bioquímica, Consejo Superior de Investigaciones Científicas/Universidad de Salamanca, Salamanca 37071, Spain

Received February 25, 1992; Revised and Accepted April 24, 1992

GenBank accession no. M63893

ABSTRACT
Most yeast strains carry a cytoplasmic double-stranded RNA (dsRNA) molecule called W, of 2.5 kb in size. We have cloned and sequenced most of W genome (1), and we proposed that W (+) strands were identical to 20S RNA, a single-stranded RNA (ssRNA) species, whose copy number is highly induced under stress conditions. Recently it was proposed that 20S RNA was circular (2). In this paper, however, we demonstrate that both W dsRNA and 20S RNA are linear. Linearity of W dsRNA is shown by the stoichiometric labelling of both strands of W with 32P-pCp and T4 RNA ligase. The last 3' end nucleotide of both strands is about 70 to 80% C and 20 to 30% A. Linearity of 20S RNA is directly demonstrated by a site-specific cleavage of 20S RNA with RNase H, using an oligodeoxynucleotide complementary to an internal site of 20S RNA. The cleavage produced not one but two RNA fragments expected from the linearity of 20S RNA.

INTRODUCTION
Circular single-stranded RNA genomes have been identified as pathogens of plants and, at least in one case, of animals. They belong, mainly, to four groups: i) Viroids; ii) Virusoids; iii) Encapsidated linear satellite RNAs, which also exist as non-encapsidated circular RNA forms, and iv) Hepatitis delta virus (HDV) (3). All of them seem to replicate by a rolling circle mechanism that involves, with some variations, a self-cleavage and self-ligation step to produce unit-length monomer circles from concatameric minus or plus strands. None of them encodes an RNA-dependent RNA polymerase. Other variants of circular RNAs are a Neurospora mitochondrial plasmid transcript called VS (4,5), and the transcript of a nuclear satellite DNA of a newt (6). Recently, it was reported another circular single-stranded RNA species in S. cerevisiae (2). This species, called 20S RNA, was about 2.5 kb in size, and it had been identified previously as a species which can be induced under the nitrogen starvation conditions that also induce sporulation (7,8). 20S RNA circularity was proposed based upon three lines of evidence: i) Electron microscopy, ii) Abnormal mobility on a two dimensional gel, and iii) The inaccessibility of 5' and 3' ends for labelling with T4 polynucleotide kinase and T4 RNA ligase respectively (2). 20S RNA cDNA clones were obtained by random priming, and the nucleotide sequence of most of the genome was determined (9); however, even though it was proposed that 20S RNA was circular, the authors were unable to clone across the gap in the 20S RNA sequence. At the same time 20S RNA nucleotide sequence was published, we reported, independently, the nucleotide sequence of W (1). W is a low copy number double-stranded RNA, 2.5 kb in size, present in many laboratory yeast strains (10). One of W strands (the (+) strand) encodes its own putative RNA polymerase (1). We learned that W (+) strand had the same nucleotide sequence as 20S RNA (1,9). We also showed that W (+) strands and 20S RNA had the same electrophoretic mobility in glyoxal-denaturing agarose gels and on an acrylamide strand separation gel, suggesting that W (+) strands and 20S RNA were identical. Based upon these lines of evidence, and on some preliminary data about W dsRNA linearity, we proposed that both molecules were linear.

In this paper we report that both W dsRNA and 20S RNA are linear. W dsRNA linearity is based upon two lines of evidence: i) Both strands can be stoichiometrically labelled at their 3' ends with 32P-pCp and T4 RNA ligase, indicating that they have free 3' end -OH groups, and ii) Direct RNA sequencing of denatured W dsRNA with reverse transcriptase suggests that there exists a free end on the template. 20S RNA linearity is based on the results obtained by site-specific RNAse H cleavage. The RNase H treatment of 20S RNA that had been annealed with a synthetic oligodeoxynucleotide, originated two RNA fragments of discrete sizes as expected from its linear nature. If 20S RNA was circular, the cleavage reaction should originate only one RNA product.

MATERIALS AND METHODS

Strains and media
Yeast strains used to obtain W dsRNA and 20S RNA were 37-4C (a leu1 karl-1 T, W, 23S RNA, 20S RNA) and RE458 (a ski2-2 L-BC, W, 20S RNA). Media were as previously described (11). Escherichia coli MV1190 (bio-Rad) was used for the propagation of plasmids and as the host for the M13 helper phage R408 (Stratagene) to obtain ssDNA.

* To whom correspondence should be addressed
dsRNA purification
T and W dsRNAs were purified by cellulose CF-11 chromatography as previously described (12). W dsRNA was freed from T dsRNA by electroelution from native agarose gels.

20S RNA preparation
Yeast strain RE458 grown to late log phase in YPAD medium (1% yeast extract, 2% peptone, 0.04% adenine sulphate, and 2% glucose) was transferred to 1% potassium acetate, pH 7.0, and incubated at 30°C for 16h (13). Cells were harvested and 20S RNA was purified as described (1). All buffers and solutions used to manipulate 20S RNA were prepared in diethylpyrocarbonate-treated H2O to avoid RNase contamination.

W cDNA clones
W cDNA clones were obtained and sequenced as described (1). They carry different portions of W cDNA sequence in the SmaI site of the Bluescript-KS+ vector (Stratagene). Some of the independently-isolated clones or subclones of those previously sequenced were used for run-off transcription to synthesize 32P-labelled T7 or T3 RNA probes. Plasmid pNR6 contains 614 bases of W cDNA sequence from base 633 (numbered from the known W 5' end) to base 1246. pNR11 contains a 1252-bp insert of W cDNA from bp 1246 to 2497. Plasmid p19 has a 1030-bp W cDNA insert between bp 1466 and bp 2495. Plasmid pW28 is one of the independently-isolated clones, and contains W sequences from base 142 to 429. Plasmid pNR14 was constructed from different W cDNA clones and contains almost the full length W cDNA sequence (from base 6 to base 2497). The known W cDNA sequence is, at present, 2505 base pairs.

Other nucleic acid manipulations
Plasmid purifications, restriction enzyme digestions, DNA ligations, and transformations were as described (14). ssRNA transcripts were made in vitro using T7 or T3 RNA polymerases following the recommendations of the supplier (U.S. Biochemical Corp.). Before the transcription reaction, plasmid DNA was digested with appropriate restriction enzymes to obtain discrete DNA template fragments for run-off synthesis by T7 or T3 RNA polymerases. Strand specific probes for Northern blot hybridization experiments were synthesized from some of the plasmids mentioned above by this method using [a-32P]UTP in the transcription reaction. RNA sequencing was done as described (15). We used synthetic oligodeoxynucleotide OII that contains W (+) strand sequences from nt 2426 to 2442 as primer, and denatured W dsRNA as template for reverse transcription.

3' end labelling
Both ends of W dsRNA (1 µg) were labelled with 32P-pCp and T4 RNA ligase in a 30 µl reaction mixture. The conditions were as suggested by the manufacturer (B.R.L. Inc.). W dsRNA from strain 37-4C was first separated from T dsRNA on a native agarose gel and electroeluted. W was further purified by filtration through an Elutip column (Schleicher & Schuell). We estimated that more than 90% of both W 3' ends were labelled with 32P-pCp (see text). The labelled strands were separated in an acrylamide strand separation gel (see below) and used to determine the 3' end terminal nucleotides.

Strand separation gels
Separation of the two W dsRNA strands in a 5% polyacrylamide gel was carried out as described (16), with slight modifications (17). RNA was denatured at 90°C for 1 min in the presence of 7 M urea instead of 30% dimethyl sulfoxide. The samples were quickly chilled on ice and applied to the gel. Both strands were separated at 8 V/cm for 18–20h. 32P-labelled W strand-specific probes were used to determine their polarity. The faster moving band was the (+) strand and the slower moving band the (–) strand (1).

Terminal analysis of W 3' ends
W 3' end terminal nucleotides were analyzed by thin layer chromatography. Both strands of 32P-labelled W dsRNA, separated as described above, were extracted from the acrylamide gel with 0.5 M ammonium acetate/1 mM EDTA at 37°C overnight. About 6000 cpm of each strand were digested with 3 U of RNase T2 (B.R.L. Inc.) in 25 mM Na citrate pH 5.0 for 4h at 37°C. The labelled terminal Np nucleotides were separated in a cellulose plate with 1 M LiCl, and detected by autoradiography; 30 nmoles of each four Up, Cp, Ap and Gp were used as standards; the radioactivity in each spot was measured by removing the cellulose from the plate and counting it in a liquid scintillation counter.

Northern blot hybridization
We performed two types of Northern blot analysis. In some experiments samples were first glyoxal denatured, separated in an agarose gel as described (14) and then the gel was blotted onto a Nytran sheet (Schleicher & Schuell) overnight, air dried, and baked under vacuum at 80°C for 1–2h. In others, 7 M urea denaturing acrylamide gels were used; six percent acrylamide, 7 M urea gels were prepared in TBE buffer and run at high voltage (20 V/cm) to maintain high temperature (55–60°C) for 5–6 hours. RNA samples were denatured in 7 M urea at 90°C for 1min, and loaded immediately. After the electrophoresis was completed, the gel was washed in H2O:methanol:acetic acid (8:2:1) to remove urea, stained with 0.5 µg/ml ethidium bromide, and photographed. The RNA was electrotransferred to a Nytran membrane (Schleicher and Schuell) in TBE buffer overnight, and hybridized with W (+)-strand-specific probes. Hybridization was done as described (18) with 32P-labelled ssRNA transcripts made in vitro with T7 or T3 RNA polymerases as mentioned above.

RNase H treatment
Total nucleic acids prepared from strain RE458 grown in conditions that induce 20S RNA accumulation, were mixed with
oligodeoxynucleotide OV (5' CGTGAGTCTCCCATTGCAGAGCACC 3'), complementary to W (+) strand sequence between nt 915 to 937, in 10 µl TE buffer. The concentration of oligonucleotide OV was 20 molar excess respect to 20S RNA; the mixture was heat-denatured for 90 seconds at 100°C (19). Samples were allowed to cool slowly over a period of 40 min until the temperature was below 35°C, and 40-µl of incubation buffer (47.5 mM Tris.HCl pH 7.9, 5 mM MgCl₂, 1.25 mM DTT) were added (20). The reaction was started by adding 0.5 U of RNase H (Boheringer Mannheim) and the digestion was continued for 30 min at 37°C. After the reaction was completed samples were extracted once with phenol:chloroform (1:1), once with chloroform:isoamyl alcohol (24:1), and precipitated with ethanol. RNA samples were either glyoxal-denatured and loaded onto an agarose denaturing gel, or dissolved in 7 M urea, 1 mM EDTA, 0.035% xylene cyanol and bromophenol blue, denatured at 90°C for 1 min, and separated in a denaturing acrylamide gel as mentioned above. After blotting, the membranes were hybridized as described above. The crude extract used contained W dsRNA in addition to 20S RNA; however, the molecular ratio of W dsRNA to 20 S RNA was less than 1/100 under the conditions used for 20S RNA accumulation (unpublished results). Therefore, the fragments detected by W (+) strand-specific probes in Fig. 4 were mostly due to 20S RNA. The presence of W dsRNA in the same sample was detected with a W (−) strand-specific probe made from plasmid pN6, but only after prolonged exposure.

RESULTS

W has free -OH 3' ends

Purified W dsRNA was efficiently labelled with 32P-pCp and T4 RNA ligase. When the labelled product was analyzed in an acrylamide strand separation gel, we found two species which incorporated radioactivity with equal intensity (Fig. 1, lane 2). These species corresponded to the two strands of W, as revealed by ethidium bromide staining of the same gel (lane 1). The polarity of the strands has been established previously; the faster moving band is the (+) strand and the slower band the (−) strand (1). There were no other species which incorporated radioactivity. We estimated the efficiency of labelling by measuring TCA-precipitable radioactivity. One µg of W dsRNA (or 1.2 pmol of free ends) incorporated 5 to 6% of the total radioactivity (60 µCi of 32P-pCp; 20 pmol) into TCA-insoluble material, indicating that more than 90% of W molecules were labelled at both ends. Therefore, both strands of W dsRNA have free 3' end -OH groups accessible to T4 RNA ligase for stoichiometric labelling with 32P-pCp.

W (+) strand 3' end sequence

We directly determined the W (+) strand 3' end by reverse transcription, using the primer OII (5' GGGCCGATGGCGGAGACT 3') that contained W (+) strand sequence from nt 2426 to 2442 and denatured W dsRNA as template. In this way the nucleotide sequence of the last W (+) strand 21 nucleotides was determined (Fig. 2). In the same gel we sequenced an independently-isolated W cDNA clone, plasmid p19, which contains W sequences from bp 1266 to 2495, as a reference. We chose this plasmid because it had the W sequences most proximal to the 3' end of the (+) strand among our W cDNA clones. Direct

Fig. 2. Direct RNA sequencing of W (+) strand 3' end. W dsRNA purified from strain 37-4C was denatured with methylmercuric hydroxide (14) and annealed with the 17-mer oligodeoxynucleotide primer indicated at the bottom of the figure. This primer was complementary to a region 53 nt upstream of the known W 3' end cDNA sequence. RNA sequencing reactions were done with reverse transcriptase, and the products were analyzed on a 8% acrylamide strand separation gel. The sequence of W 3' end (+) strand was extended 10 extra nucleotides by this method and the extra sequence is underlined at the bottom of the figure.

Fig. 3. Terminal nucleotides of W (+) and (−) strand 3' ends. 3' end labelled W (+) and (−) strands were purified from an acrylamide strand separation gel and digested with T2 nuclease. The terminal labelled Np mononucleotides were separated in a cellulose plate with 1 M LiCl and detected by autoradiography. 30 nmoles of Np standards were run in the same plate, and detected by absorption at 260 nm. The positions of Np standards are indicated. (+) and (−) indicate W (+) and (−) strands.
RNA sequencing of denatured W dsRNA, thus, allowed us to extend the known W (+) strand sequence 10 extra nucleotides (1). Interestingly the synthesis of cDNA in the four lanes corresponding to the four dideoxynucleotides ended at the same point, and there was no more synthesis beyond it (Fig. 2). This might reflect some strong secondary structure on the template that blocked the advance of the reverse transcriptase, but it is more likely due to the existence of a free end on the template, since as shown above, these free 3' ends actually exist on W (+) and (-) strands. Therefore this result is consistent with W dsRNA being linear. We, however, could not establish the terminal nucleotide with this procedure, and so we analyzed W 3' ends composition by thin layer chromatography (TLC).

The 3' ends of W (+) and (-) strands are either C or A

We next analyzed the terminal nucleotides of W (+) and (-) strand 3' ends by TLC, and we found them to be a mixture of cytidine and adenine (Fig. 3). Purified W strands, that had been labelled at their 3' ends with 32P-pCp, were digested with T2 nuclease and the products were separated by thin layer chromatography with Np standards. Two radioactive spots corresponding to Cp and Ap were found (Fig. 3). This result eliminates the possibility that the 3' end -OH groups were created artificially by random nicking of W dsRNA. The radioactivity present in each spot was measured by removing the cellulose from the plate and counting it in a liquid scintillation counter. Eighty percent of radioactivity was in the form of Cp for W (+) strands and 70% in the case of W (-) strands. The rest (20% for (+) strands and 30% for (-) strands) was Ap. This result may reflect the existence of heterogeneity in W 3' ends, or alternatively it may be due to a non-encoded posttranscriptional addition of an extra adenine to the newly synthesized RNA, which is often seen in many dsRNA and (+) ssRNA viral genomes (21,22). L-A dsRNA, another dsRNA present in most yeast strains, has been shown to have either a paired C or an unpaired A at its 3' ends (23,24). The sequence of W (+) strand 3' end thus could be ...CC-OH or ...CCA-OH, quite similar to the 3' end terminal sequences found in other (+) strand or double-stranded RNA viruses, such as the RNA Coliphage's (...CCCA-OH) (25,26), or to the RNA 3' ends (...CCA-OH). We also found the same terminal composition in T dsRNA (another dsRNA present in the same host S. cerevisiae (10)); about 85-90% of both T 3' ends were C and only a minor 10-15% were A (Luis M. Esteban and R. Esteban, unpublished results), consistent with close relationship between T and W (27).

20S RNA is linear

As shown above, W dsRNA is a linear molecule and has free -OH groups at both (+) and (-) strand 3' ends. Since W (+) strands comigrate with 20S RNA in strand separation acrylamide gels and in denaturing agarose gels, it is likely that 20S RNA is also linear (1). However, recently, it was proposed that 20S RNA was circular (2). Therefore we designed an experiment to solve this discrepancy. The outline of the experiment is shown in Fig. 4-1. 20S RNA was annealed with the synthetic DNA oligodeoxynucleotide OV, complementary to W (+) strand sequence from nt 915 to 937, and then digested with RNase H (see MATERIALS AND METHODS). If 20S RNA was circular, the cleavage would produce only one linear, but full length (2.5 kb in size) molecule. On the other hand, if 20S RNA was linear, the cleavage should produce two fragments (1.6 kb and 0.9 kb

Fig. 4. RNase H treatment of 20S RNA. 1. Diagram of two possible 20S RNA configurations (linear and circular), and the results expected from RNase H digestion. 20S RNA was annealed with the synthetic DNA oligodeoxynucleotide OV (indicated as olig. V) and digested with RNase H. If 20S RNA was linear two RNA fragments of about 0.9 and 1.6 kb would be originated; if 20S RNA was circular only one 2.5 kb RNA molecule would be obtained. 2. Diagram of the W cDNA, the site where oligodeoxynucleotide OV hybridizes and the regions which probes a, b, and c recognize. 3. Proof that 20S RNA is linear. RNase H digested 20S RNA was separated either in a glyoxal denaturing agarose gel (A to C) or in a 7 M urea denaturing 6% acrylamide gel (D). Then RNA was blotted onto nylon membranes and hybridized with the probes a, b, or c shown in 2. Lane C; original untreated crude extract which contained 20S RNA. Lane 2; RNase H treated crude extract. Lane 1; same as lane 2 except that RNase H was omitted from the reaction. Note that in A, B, and C the same nylon membrane was used to hybridize with probes a, b, and c respectively. In A and D we used the same probe (a). Molecular markers were L-A (4.6 kb) and M1 (1.8 kb) dsRNAs.
DISCUSSION

In this paper we showed that both W dsRNA and 20S RNA are linear. W dsRNA linearity came from T4 RNA ligase experiments. Both strands of W dsRNA were stoichiometrically labelled with $^{32}$P-pCp, indicating that those strands contained free -OH groups at their 3' ends. The last 3' end nucleotide labelled by $^{32}$P-pCp was 80% C and 20% A for the (+) strand and 70% C and 30% A for the (-) strand. Consistently, direct RNA sequencing of denatured W dsRNA with reverse transcriptase suggested that there existed a free end on the template very close to the known end of W (+) strand sequence. 20S RNA linearity was directly shown by a site-directed RNase H cleavage of 20S RNA. The cleavage produced two discrete RNA fragments with the sizes expected from the linear nature of 20S RNA. Since the known sequences of W (+) strand and 20S RNA are the same, the results shown in this paper suggest that W (+) strand and 20S RNA are identical. In fact, both RNAs comigrate not only in strand separation gels (1), but also in a denaturing gel (7 M urea acrylamide gel (Fig. 4-3, D)), we ruled out the possibility that any purification procedure. Furthermore, since the mocked experiment in the absence of RNase H gave the same 20S RNA linearity, we confirmed this we used two additional probes (Fig. 4-2). As expected probe b and probe c hybridized with the 3' end-proximal 1.6 kb larger fragment (Fig. 4-3, B) and the 5' end-proximal 0.9 kb smaller fragment (Fig. 4-3, C), respectively. Since we used a crude extract as source of 20S RNA, we ruled out the possibility of artificial cleavage associated with any purification procedure. Furthermore, since the mocked experiment in the absence of RNase H gave the same 20S RNA mobility as that of the original, untreated sample, not only in a denaturing agarose gel, but also in a denaturing (7 M urea) acrylamide gel (Fig. 4-3, D), we ruled out the possibility that linear 20S RNA molecules were created during the experiment by self-cleavage due to the conditions used for RNase H treatment. From these results we concluded that 20S RNA is linear.

Previously, 20S RNA was proposed to be circular (2). The proposal was based in three lines of evidence, i) It was claimed that 50% of 20S RNA molecules were circular when examined in Electron microscopy. In the procedure employed, however, the authors crosslinked 20S RNA with the T4 gene 32 protein to spread the RNA. Therefore, the circular molecules observed may be an artifact caused by the procedure used. Alternatively, in the case of Infectious Pancreatic Necrosis Virus (IPNV), proteins covalently attached to the 5' ends of the dsRNA genome could circularize the linear dsRNA and this circular form disappeared after proteinase K treatment (28). No such controls were done for 20S RNA. ii) It was claimed that 20S RNA showed abnormal mobility in a two dimensional gel, specifically in an acrylamide gel under denaturing conditions. As shown in Fig. 4, however, 20S RNA did not show any abnormal mobility in a denaturing agarose gel or in a 7 M urea acrylamide gel. Furthermore, linear T7 or T3 run-off transcripts from plasmid pNR14, which contained W sequences from nt 6 to 2497 and few extra nucleotides from the vector, comigrated with 20S RNA in a 7 M urea 6% polyacrylamide gel (N. Rodriguez-Cousino and R. Esteban, unpublished results). Therefore, it is likely that the 20S RNA or RNA markers used were not properly denatured in their two dimensional gel electrophoresis (2). iii) The third line of evidence was the inability of labelling 20S RNA 5' and 3' ends with T4 polynucleotide kinase and T4 RNA ligase, respectively. However, it has been well known the difficulty of labelling free 3' end -OH groups of ssRNA with T4 RNA ligase (29). In this context, it should be noted that the last 24 nucleotides of W (+) strand 3' end can be folded forming a stem-loop structure with a free energy of $-25$ Kcal/mol (30). This strong secondary structure, thus, may affect the labelling reaction with T4 RNA ligase. Respect to the 5' end, since we can not rule out the possibility of a small protein moiety attached to it, this could explain the inability of labelling 20S RNA with T4 polynucleotide kinase. Therefore, the circularity of 20S RNA was not proposed based on solid evidence.

Previously, we proposed two possible replication pathways for 20S RNA and W dsRNA (1). None of these models contains a rolling circle mechanism as suggested by Matsumoto and Wickner (9), because we considered that these RNAs were linear. Rather, one of them resembles more RNA coliphages' model, such as Q5's; that is, 20S RNA is the genomic RNA and W dsRNA is a by-product of 20S RNA replication cycle originated by the annealing of 20S RNA and its complementary strand. In this context it should be noted that W's putative RNA polymerase consensus sequence is related most closely to those of coliphages' RNA polymerases. The other model proposes that both W dsRNA and 20S RNA are constituents of the same replication cycle; that is, W dsRNA is transcribed to make 20S RNA and then 20S RNA replicates to produce new W dsRNA. In either case, their replication mechanism should be quite different from those of other encapsidated dsRNA viruses such as L-A or L-BC found in the same host, Saccharomyces cerevisiae, since neither W dsRNA nor 20S RNA seem to be encapsidated into viral particles (10,31).

Recently we cloned and sequenced T dsRNA, another dsRNA found in yeast (27). Like W, T encodes its own putative RNA polymerase. T's polymerase shares amino acid identities with W's polymerase and this identity extends beyond the region well conserved among RNA-dependent RNA polymerases. This suggests their common origin. T dsRNA also has its (+) ssRNA form, 23S RNA. Like 20S RNA, the 23S RNA copy number can be induced under stress conditions such as heat-shock or starvation. Thus we proposed that 20S RNA share with their ssRNA counterparts constitute a distinct RNA family of Saccharomyces cerevisiae. The fact that T 3' ends were also labelled stoichiometrically with $^{32}$P-pCp and T4 RNA ligase (thus T dsRNA is linear) and that the composition of the last 3' end nucleotides was similar to that of W's, are in good agreement with our proposal of the new RNA family in yeast.
ACKNOWLEDGMENTS

This research was supported by Grant BIO88-0247 from The Comisión Interministerial de Ciencia y Tecnología (CICYT), Spain. Nieves Rodriguez-Cousiño is the recipient of a fellowship from the University of Salamanca, Spain.

REFERENCES