Phylogenetic Analysis of a Bacterial Aerobic Degrader of Azo Dyes

MEERA GOVINDASWAMI,¹ THOMAS M. SCHMIDT,² DAVID C. WHITE,³ AND JOHN C. LOPER^{1*}

Department of Molecular Genetics, Biochemistry and Microbiology, University of Cincinnati Medical Center, Cincinnati, Ohio 45267-0524¹; Department of Microbiology, Miami University, Oxford, Ohio 45056²; and Center for Environmental Biotechnology, University of Tennessee, Knoxville, Tennessee 37932³

Received 22 February 1993/Accepted 7 July 1993

Eubacterial consensus oligonucleotide primers were used to amplify by polymerase chain reaction the nearly full-length 16S rRNA gene of isolate C7, a gram-negative rod capable of aerobic degradation of azo dyes. The DNA product was cloned and sequenced. Phylogenetic analysis based upon this DNA sequence places C7 within the α subdivision of proteobacteria, most closely related to *Caulobacter subvibrioides*. The phospholipid fatty acid pattern resembles that of caulobacters, with monounsaturated 16- and 18-carbon fatty acids predominating. C7 is unusual in having a monounsaturated branched fatty acid in the phospholipids and exclusively 2-hydroxy fatty acids in the lipid-extracted residue. This organism is of potential use in bioreactors operated for azo dye degradation.

Bacterial isolate C7 has been described in earlier reports as a gram-negative rod capable of oxygen-insensitive azo bond cleavage of dyes such as acid orange 7 and acid red 151 during aerobic growth in glucose-enriched minimal medium (28). This organism is a potential candidate for incorporation into experimental bioreactors operated for azo dye degradation. We have pursued a better understanding of C7 through its taxonomic characterization. However, attempts at identification with the API NFT test strip (Analytab, Plainview, N.Y.), with the Biolog Identification System (Biolog, Inc., Hayward, Calif.), and by fatty acid analysis profile (Microbial ID, Inc., Newark, Del.) were all inconclusive (27).

In recent years, 16S rRNA sequence comparison has been used as a powerful tool for establishing phylogenetic and evolutionary relationships among organisms (14, 16, 32). An approach presently employed in many laboratories uses the polymerase chain reaction (25) to obtain 16S rRNA-specific DNA for sequence analysis. With this approach, the goal of this investigation was to determine the 16S rDNA sequence of C7 in order to assess its phylogenetic position. Examination of the pattern of phospholipid fatty acids (PLFA) and lipid-extracted residue hydroxy fatty acids (OHFA) was also utilized to further characterize this gram-negative rod.

Strain C7 was cultured by shaking at 30°C in Vogel-Bonner minimal E-medium (7) supplemented with 25 μg of filter-sterilized acid orange 7 or acid red 151 per ml. Cells in exponential growth were pelleted by centrifugation at 12,100 $\times g$ for 5 min at 4°C. Genomic DNA was isolated by standard methods and was gel purified (15).

The nearly full-length $16\bar{S}$ rRNA gene was amplified by polymerase chain reaction with oligodeoxynucleotide primers designed to anneal to conserved regions of eubacterial 16S rRNA. The forward primer corresponded to nucleotide positions 8 to 27 of *Escherichia coli* 16S rRNA (5'-AGAGTTTGATCCTGGCTCAG-3'), and the reverse primer was rP2 (29) (5'-TACGGCTACCTTGTTACGAC-3'). Primer rP2 corresponds to the complement of nucleotide positions 1513 to 1494 in the *E. coli* sequence except that nucleotide 1508 was changed from T to C. Polymerase chain reaction

These conditions reproducibly gave a single, full-length product of approximately 1.5 kb, while the control containing no added template DNA yielded no amplification product. To reduce the likelihood of misidentifying errors incorporated during amplification because of the low fidelity of Taq polymerase, samples from six amplifications were pooled and used for cloning. With this practice an error at any given position in one amplification would have to occur at the same position in all six amplifications for it to be misidentified. The purified polymerase chain reaction product was treated with Klenow DNA polymerase (New England BioLabs) and ligated as a blunt-end fragment into SmaI (New England BioLabs)-digested Bluescript KS(-) (Stratagene) by standard protocols (15). The ligation mixture was transformed into E. coli DH5a cells and plated onto LB agar plates containing 100 µg of ampicillin per ml (13). Plasmid DNA was prepared from 10 positive recombinant transformants (2, 15), pooled, and used as template for direct dideoxy sequencing with ³²P, using Sequenase V.2 (United States Biochemical) according to the protocol recommended by the manufacturer for sequencing double-stranded plasmid DNA. The cloned C7 16S rDNA was sequenced on both strands by using primers complementary to conserved regions of the E. coli 16S rRNA sequence. Ninety percent of the sequence was determined on both strands; the remaining was determined from replicate sequencing reactions using one strand.

amplification was carried out in a 100- μ l mixture containing 10 mM Tris-HCl (pH 8.3), 1.5 mM MgCl₂, 50 mM KCl, 100 μ g of gelatin per ml, 200 μ M dATP, 200 μ M dCTP, 200 μ M dGTP, 250 nM each oligonucleotide primer, 2.5 U of *Taq* DNA polymerase (Perkin-Elmer Cetus), and 100 ng of purified C7 chromosomal DNA. The mixture was subjected to 30 cycles in a Perkin-Elmer DNA Thermal cycler, with each cycle consisting of denaturation at 94°C for 1 min, primer annealing at 60°C for 1 min, and chain extension at 72°C for 2 min; 7 min more was added to the chain extension step on the last cycle. Following amplification, the DNA mixture was extracted with phenol-chloroform (1:1) and with chloroform, and the DNA was precipitated with ethanol in the presence of sodium acetate and resuspended in 100 μ l of Tris-EDTA.

^{*} Corresponding author.



Proteobacteria

FIG. 1. Phylogenetic relationships of selected proteobacteria from 16S rRNA sequences. The tree was rooted with *Clostridium perfringens* (not shown).

The C7 16S rRNA sequence was aligned manually with the sequences of other 16S rRNAs on the basis of regions of conserved sequences and secondary structures. All reference 16S rRNA sequences were obtained from either GenBank (Los Alamos, N.Mex.) or the ribosomal data base project (21). Regions of sequences that could not be aligned unambiguously were omitted from the phylogenetic analysis. Regions included in this analysis were positions 28 to 68, 101 to 180, 220 to 451, 481 to 836, 850 to 1002, 1037 to 1128, 1144 to 1439, and 1460 to 1489 (*E. coli* numbers).

The evolutionary distance between each pair of sequences was calculated, and a least-squares method was used to infer the phylogenetic tree most consistent with the pairwise distance estimate (20). The helices between bases (*E. coli* numbers) 69 and 100, 180 and 220, 455 and 480, and 997 and 1044 were either truncated to shorter loops (4), deleted, or altered in secondary structure; these features strongly suggest that C7 is a member of the α subdivision of the proteobacteria (12, 33). Analysis of signature sequences (32) is also consistent with placement of C7 within the α subdivision of the proteobacteria.

A phylogenetic tree (Fig. 1) constructed by the distance matrix method (19) displays C7's position in relation to other members of the proteobacteria, using sequences obtained from GenBank (1) or the 16S rRNA data base project (21). The statistical significance of the phylogenetic tree was tested by bootstrap analysis (9). The relationship shown is robust; bootstrap analysis with 100 replicates places C7 within the α group of proteobacteria, and most closely related to *Caulobacter subvibrioides*, in all 100 bootstrap trees. The overall secondary structure of C7 16S rRNA showed no unusual base pairing, structural deviations, or variation in length compared with other members of this subdivision of proteobacteria. Initial phylogenetic analysis showed C7 to be most closely related to *Aquaspirillum magnetotacticum* (10), with a sequence similarity of 85.4%, while our analysis using a more recent data base detected one more closely related organism, *C. subvibrioides*, with a sequence similarity of 92.8% (Fig. 1). A sequence similarity comparison for these and other selected bacteria appears in Table 1.

Although the evolutionary origin of C7 most closely coincides with those of C. subvibrioides and A. magnetotacticum, C7 differs from both of these type species in several physiological traits. C. subvibrioides isolates have been found in water and in soil, but laboratory cultures have requirements for unknown organic growth factors that are not satisfied by mixtures of B vitamins, amino acids, and purine and pyrimidine bases (23). A. magnetotacticum is an obligate microaerophile and grows best under an atmosphere of 1 to 3% oxygen in a simple, chemically defined medium. It is unable to grow anaerobically or under an air atmosphere (3). C7 is not a microaerophile and grows well in defined synthetic media under high-oxygen conditions. It has been observed that C7 grows much better on minimal medium than it does in more-complex media such as tryptic soy broth. Thus, C7 does not require special growth factors and

 TABLE 1. Similarities between 16S rRNA sequences of C7 and selected bacteria

Organism compared with C7	Similarity determined with:	
	Positions considered in phylogenetic analysis	All positions
Escherichia coli	0.826	0.797
Biejerinckia indica	0.879	0.859
Caulobacter subvibrioides	0.942	0.928
Aquaspirillum magnetotacticum	0.879	0.854
Rhodopseudomonas marina	0.881	0.860
Agrobacterium tumefaciens	0.884	0.861
Rochalimaea quintana	0.888	0.860
Ancylobacter aquaticus	0.887	0.865
Rhodopseudomonas palustris	0.868	0.835
Rhodopseudomonas acidophila	0.868	0.840
Methylosinus trichosporium	0.844	0.817
Clostridium perfringens	0.806	0.774
Pseudomonas aeruginosa	0.822	0.788
Chromatium vinosum	0.835	0.799
Methylophilus methylotrophus	0.799	0.761
Nitrosomonas europaea	0.820	0.779
Pseudomonas testosteroni	0.811	0.780
Rhodocyclus gelatinosus	0.809	0.772
Erythrobacter longus	0.854	0.820
Rhodobacter capsulatus	0.858	0.847
Methylobacterium sp.	0.844	0.817

may be slightly inhibited in the presence of complex organic compounds.

In a recent study Stahl et al. (26) showed through phylogenetic analysis that the caulobacters examined are all affiliated with the alpha proteobacteria but that there was an early divergence into two primary lines of descent, one composed exclusively of marine caulobacters and the other made up primarily of freshwater species. The only exception to this pattern was the species *C. subvibrioides*, which was seen to be only peripherally related to the greater caulobacter assemblage. This implies an early origin of the caulobacter phenotype within the alpha subdivision. Since C7 is most closely related to *C. subvibrioides*, C7 could also have an early origin within the alpha proteobacteria.

For lipid analysis, approximately 20 to 25 mg (dry weight) of bacterial cells was extracted in a chloroform-methanol, single-phase solvent system modified to include phosphate buffer (30). The total lipid extract was fractionated on silicic acid columns into neutral, glyco-, and phospholipids (22). The polar lipid fraction recovered in methanol was transesterified with mild alkaline methanolysis (30), and the methyl esters were separated, quantified, and tentatively identified after injection by capillary gas chromatography (GC) using a 50-m HP-1 (nonpolar methyl silicone) column under the conditions described previously (24). The PLFA structures were verified by GC-mass spectrometry (MS) with the same GC column for Hewlett-Packard 5996A GC-MS as described previously (24). Mono- and dienoic PLFA double-bond positions and conformations were determined by GC-MS analysis of the dimethyl disulfide adducts (17). Cyclopropyl-PLFA ring positions were determined by GC-MS after hydrogenation (11). Equivalent chain length was calculated as described previously (6).

The lipid-extracted residue OHFA were recovered from the lipid A fraction (8, 18). This material was found in the upper phase of the extraction. The upper phase was concentrated in a rotary evaporator in vacuo and hydrolyzed in 1 N

TABLE 2. Analysis of PLFA and lipid-extracted residue OHFA of C7

Fatty acid ^a	% ^b
PLFA	
16:1w7c	
16:1w5c	1.37 (0.31)
cy17:0	
17:0	
	0.87 (0.13)
	0.37 (0.06)
br19:1	
cy19:0	0.10 (0.08)
Total	
Residue OHFA	
Total	

^a Fatty acid designations are total number of carbon atoms:number of double bonds, with the position of the double bond closest to aliphatic end (w) of the molecule (aliphatic for methyl). Configurations of the double bonds are indicated as *cis* (c) or *trans* (t). For example, 16:1w7c is a PLFA with a total of 16 carbons with one double bond 7 carbons from the methyl end in the *cis* configuration. Branched (br) fatty acids are designated *iso* (i) or *anteiso* (a) (see text) if the methyl branch is one or two carbons from the w end (e.g., (i15:0) or by the position from the methyl end of the molecule (e.g., 10Me16:0). Cyclopropyl (cy) fatty acids are designated by the total number of carbons (e.g., cy.17:0), and the position of the hydroxyl from the carboxyl end of the fatty acid is indicated as a prefix (e.g., 2-OH16:0).

^b Mean, with standard deviation in parentheses (n = 2).

HCl for 2 to 4 h at 100°C. The residue was then reextracted with two portions of methanol and three portions of chloroform so that the final volume contained chloroform-methanol-1 M HCl (5:2:3, vol/vol/vol). After 24 h of storage at room temperature, the preparation was centrifuged, and the chloroform phase was recovered, evaporated to dryness, and methylated with "magic" methanol (methanol-chloro-form-concentrated HCl, [10:1:1, vol/vol/vol]) (16a) at 100°C for 1 h. The methylated OHFA were recovered in the upper phase following extraction (with three washes) in hexanechloroform (4:1, vol/vol). The volatile solvent was then removed with a stream of nitrogen, and the OHFA were purified on thin-layer plates (precleaned with hexane-ethylether [1:1]). Plates were developed in the same solvent, and the OHFA band was recovered after location with a standard that was run simultaneously. Before GC-MS analysis, the hydroxyl groups were derivatized with N,O-bis(trimethylsilyl)trifluoroacetamide (Pierce, Rockford, Ill.).

The PLFA of C7 have 16:0, 16:1w7c, and 18:1w7c constituting 94% of the total (Table 2). This is a rather common pattern, generated by the anaerobic desaturase pathway and shared by Caulobacter species, among others. However, the presence of the unusual br19:1 at nearly 3 mol% in the PLFA clearly differentiates this organism from the Caulobacter species that have been examined (31). None of the 14 species or strains of Caulobacter or the two species of Asticcacaulis or Prosthecomicrobium show 19-carbon PLFA, none show PLFA longer than 17 carbon atoms, and, except for Caulobacter vibrioides GB-G and Caulobacter intermedius CB63, these strains have only traces of branched PLFA (GB-G shows 6% a15:0 and CB63 shows 5% a15:0 [5]). C. subvibrioides subsp. albus CB88 contains only traces (<1%) of a15:0 and i16:0 as branched PLFA (5). Of two Hyphomicrobium species, one has 7% hydroxy-PLFA, which C7 does not have, and the other has no branched PLFA (31). Moreover,

C7 is extraordinary in having only 2-OHFA acids in the lipid-extracted residue (Table 2). Other gram-negative bacteria have 3-OHFA or both 3-OHFA and 2-OHFA in their lipopolysaccharide (LPS).

Caulobacters and other bacteria of similar morphology represent an extremely diverse phylogeny as assessed by 16S rRNA-based phylogenetic analysis (26). Of the five major groups defined by 16S rRNA, species in three groups show divergences from the PLFA pattern of C7. PLFA patterns of two groups of marine caulobacters have not been examined. On the basis of 16S rRNA analysis, one of these clusters with *Pseudomonas diminuta*. Fifty-two percent of the LPS OHFA of *P. diminuta* was shown to be 3-OH12:0, a compound which is not found in C7. Thus, C7 shows no PLFA or LPS OHFA patterns in common with the analyzed caulobacters, and there are no organisms with C7's characteristics in the major lipid structure compendium (31).

C7 had been tentatively identified as *Pseudomonas paucimobilis* by the API NFT test strip (Analytab) (28) and was compared with the reference strain *P. paucimobilis* ATCC 29837 by use of the Biolog Identification System (Biolog, Inc.) and by fatty acid analysis profile (Microbial ID, Inc.). Both systems identified the ATCC strain as *P. paucimobilis*, but neither identified C7. (The Microbial ID analysis came up with a match with *Pseudomonas saccharophila* with a low index value of 0.049, 1.0 being a perfect match. ATCC 29837 did not decolorize acid orange 7 in aerobic liquid culture (27).

Our data provide the first phylogenetic classification of a potentially useful aerobic azo dye-degrading microorganism, isolate C7, by 16S rDNA and lipid analysis. The results indicate that C7 most probably is a previously unrecognized species and not a variant of a known species. Isolation and characterization of microorganisms showing unique functional properties from natural environments and 16S rRNA analysis of these bacteria are likely to lead to the continued identification of novel, previously uncharacterized species, providing further insights into the diversity and complexity of the microbial world. Isolate C7 is another organism for which lipid analysis has revealed unique properties of taxonomic relevance.

Nucleotide sequence accession number. The sequence results from this study are available from GenBank (accession no. L22759).

This research was supported in part by USEPA grant R816118 (to J.C.L.) and by grant ES04908 from the National Institutes of Environmental Health Sciences (to J.C.L.). The phylogenetic analyses were supported in part by USEPA cooperative agreement no. CR 820061010 (to T.M.S.). The lipid analyses was supported in part by grant DE-FG05-90ER50988, in the Subsurface Science Program, Office of Health and Environment, Department of Energy (to D.C.W.).

REFERENCES

- 1. Bilofsky, H. S., and C. Burks. 1988. The GenBank genetic sequence data bank. Nucleic Acids Res. 16:1861-1863.
- Birnboim, H. C., and J. Doly. 1979. A rapid alkaline extraction procedure for screening recombinant plasmid DNA. Nucleic Acids Res. 7:1513–1518.
- 3. Blakemore, R. P., D. Maratea, and R. S. Wolfe. 1979. Isolation and pure culture of a fresh water magnetic spirillum in chemically defined medium. J. Bacteriol. 140:720-729.
- Brosius, J., J. L. Palmer, J. P. Kennedy, and H. F. Noller. 1978. Complete nucleotide sequence of a 16S ribosomal RNA gene from *Escherichia coli*. Proc. Natl. Acad. Sci. USA 75:4801– 4805.
- 5. Carter, R. N., and J. M. Schmidt. 1976. Fatty acid composition

NOTES 6065

of selected prosthecate bacteria. Arch. Microbiol. 110:91–94.

- 6. Christie, W. W. 1989. Gas chromatography and lipids. The Oily Press, Ayr, Scotland.
- 7. Davis, R. W., D. Botstein, and J. R. Roth. 1980. A manual for genetic engineering—advanced bacterial genetics. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
- Edlund, A., P. D. Nichols, R. Roffey, and D. C. White. 1985. Extractable and lipopolysaccharide fatty acid and hydroxy acid profiles from *Desulfovibrio* species. J. Lipid Res. 26:982–988.
- 9. Felsenstein, J. 1985. Confidence limits on phylogenies: an approach using the bootstrap. Evolution 39:783-791.
- Govindaswami, M., and J. C. Loper. 1992. 16S rRNA probes of bacterial isolate C7, an aerobic degrader of azo dyes, abstr. Q151, p. 360. Abstr. 92nd Gen. Meet. Am. Soc. Microbiol. 1992. American Society for Microbiology, Washington, D.C.
- Guckert, J. B., C. P. Antworth, P. D. Nichols, and D. C. White. 1985. Phospholipid, ester-linked fatty acid profiles as reproducible assays for changes in prokaryotic community structure of estuarine sediments. FEMS Microbiol. Ecol. 31:147–158.
- Gutell, R. R., B. Weiser, C. R. Woese, and H. F. Noller. 1985. Comparative anatomy of 16S-like ribosomal RNA. Prog. Nucleic Acid Res. Mol. Biol. 32:155–216.
- Hanahan, D. 1983. Studies on transformation of *Escherichia coli* with plasmids. J. Mol. Biol. 166:557–580.
- Lane, D. L., B. Pace, G. J. Olsen, D. A. Stahl, M. L. Sogin, and N. R. Pace. 1985. Rapid determination of 16S ribosomal RNA sequences for phylogenetic analysis. Proc. Natl. Acad. Sci. USA 82:6955-6959.
- 15. Maniatis, T., E. F. Fritsch, and J. Sambrook. 1982. Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
- Martinez-Murcia, A. J., and M. D. Collins. 1990. A phylogenetic analysis of the genus *Leuconostoc* based on reverse transcriptase sequencing of 16S rRNA. FEMS Microbiol. Lett. 70:73-84.
- 16a. Mayberry, W. Personal communication.
- Nichols, P. D., J. B. Guckert, and D. C. White. 1986. Determination of monounsaturated fatty acid double-bond position and geometry for microbial monocultures and complex consortia by capillary GC-MS of their dimethyl disulphide adducts. J. Microbiol. Methods 5:49-55.
- Nichols, P. D., G. A. Smith, C. P. Antworth, R. S. Hanson, and D. C. White. 1985. Phospholipid and lipopolysaccharide normal and hydroxy fatty acids as potential signatures for the methaneoxidizing bacteria. FEMS Microbiol. Ecol. 31:327-335.
- Olsen, G. J. 1987. The earliest phylogenetic branchings: comparing rRNA-based evolutionary trees inferred with various techniques. Cold Spring Harbor Symp. Quant. Biol. 52:825-838.
- Olsen, G. J. 1988. Phylogenetic analysis using ribosomal RNA. Methods Enzymol. 164:793–812.
- Olsen, G. J., N. Larsen, and C. R. Woese. 1991. The ribosomal RNA database project. Nucleic Acids Res. 19(Suppl.):2017– 2021.
- Parker, J. H., G. A. Smith, H. L. Fredrickson, J. R. Vestal, and D. C. White. 1982. Sensitive assay based on hydroxy-fatty acids from lipopolysaccharide lipid A from gram-negative bacteria in sediments. Appl. Environ. Microbiol. 44:1170-1177.
- Poindexter, J. S. 1964. Genus Caulobacter 1964, 289^{AL}, p. 1938. In J. T. Staley, M. P. Bryant, N. Pfennig, and J. G. Holt (ed.), Bergey's manual of systematic bacteriology, vol. 3. The Williams & Wilkins Co, Baltimore.
- 24. Ringelberg, D. B., J. D. Davis, G. A. Smith, S. M. Pfiffner, P. D. Nichols, J. B. Nickels, J. M. Hensen, J. T. Wilson, M. Yates, D. H. Kampbell, H. W. Reed, T. T. Stocksdale, and D. C. White. 1988. Validation of signature polarlipid fatty acid biomarkers for alkane-utilizing bacteria in soils and subsurface aquifer materials. FEMS Microbiol. Ecol. 62:39–50.
- Saiki, R. K., D. H. Gelfand, S. Stoffel, S. J. Scharf, R. Higuchi, G. T. Horn, K. B. Mullis, and H. A. Erlich. 1988. Primerdirected enzymatic amplification of DNA with a thermostable DNA polymerase. Science 239:487–491.
- 26. Stahl, D. A., R. Key, B. Flesher, and J. Smit. 1992. The phylogeny of marine and freshwater caulobacters reflects their

habitat. J. Bacteriol. 174:2193-2198.

- 27. Tabor, M. W., M. Carsiotis, D. Feldhake, E. Holder, P. V. Scarpino, S. T. Giolando, and J. C. Loper. Aerobic degradation of sulfonated azo dyes by a gram-negative rod, isolate C7. Submitted for publication.
- Tabor, M. W., M. Carsiotis, E. Holder, P. V. Scarpino, S. T. Giolando, and J. C. Loper. 1991. Aerobic degradation of azo dyes by a *Pseudomonas paucimobilis*, isolate C7, abstr. Q169, p. 304. Abstr. 91st Gen. Meet. Am. Soc. Microbiol. 1991. American Society for Microbiology, Washington, D.C.
- American Society for Microbiology, Washington, D.C.
 29. Weisburg, W. G., S. M. Barns, D. A. Pelletier, and D. J. Lane. 1991. 16S ribosomal DNA amplification for phylogenetic study.

J. Bacteriol. 173:697-703.

- White, D. C., W. M. Davis, J. S. Nickels, J. D. King, and R. J. Bobbie. 1979. Determination of the sedimentary microbial biomass by extractable lipid phosphate. Oecologia 40:51-62.
- Wilkinson, S. G. 1988. Gram-negative bacteria, p. 299-488. In C. Ratledge and S. G. Wilkinson (ed.), Microbial lipids, vol. 1. Academic Press, Inc., New York.
- Woese, C. R. 1987. Bacterial evolution. Microbiol. Rev. 51:221– 271.
- Woese, C. R., R. Gutell, R. Gupta, and H. R. Noller. 1983. Detailed analysis of the higher-order structure of 16S-like ribosomal ribonucleic acids. Microbiol. Rev. 47:621–669.