JOURNAL OF BACTERIOLOGY, Jan. 1994, p. 514–517 Vol. 176, No. 2
0021-9193/94/$04.00+0
Copyright © 1994, American Society for Microbiology

Strain Identification and 5S rRNA Gene Characterization of the Hyperthermophilic Archaebacterium *Sulfolobus acidocaldarius*

PETER DUROVIC,[1] ULRIKE KUTAY,[1]† CHRISTA SCHLEPER,[2] AND PATRICK P. DENNIS[1,2]*

Department of Biochemistry and Molecular Biology and Canadian Institute for Advanced Research, University of British Columbia, Vancouver, British Columbia V6T 1Z3, Canada,[1] and Max Planck Institut für Biochemie, D-82152 München, Germany[2]

Received 13 July 1993/Accepted 3 November 1993

A commonly used laboratory *Sulfolobus* strain has been unambiguously identified as *Sulfolobus acidocaldarius* DSM639. The 5S rRNA gene from this strain was cloned and sequenced. It differs at 17 of 124 positions from the identical 5S rRNA sequences from *Sulfolobus solfataricus* and a strain apparently misidentified as *S. acidocaldarius*. Analysis of the transcripts from the 5S rRNA gene failed to identify any precursor extending a significant distance beyond the 5′ or 3′ boundary of the 5S rRNA-coding sequence. This result suggests that the primary transcript of the 5S rRNA gene corresponds in length (within 1 or 2 nucleotides) to the mature 5S rRNA sequence found in 50S ribosomal subunits.

Members of the genus *Sulfolobus* are aerobic or facultatively aerobic representatives of the sulfur-dependent hyperthermophilic archaebacteria (16). They can grow autotrophically on elemental sulfur or heterotropically on organic compounds. Because of their phylogenetic position within the archaebacterial kingdom, their hyperthermophilic growth characteristics, and their relative ease of cultivation, these organisms have been subjected to intense molecular and physiological characterization. Three prototype species have been used in most studies: *Sulfolobus solfataricus* PI (DSM1616), *Sulfolobus acidocaldarius* (DSM639), and *Sulfolobus shibatae* B12 (DSM5389).

The genomes of these organisms contain single-copy rRNA genes. The 16S and 23S rRNA genes are genetically linked and cotranscribed to produce a primary transcript approximately 5 kb long (2a, 3, 4, 9). The 5S rRNA gene is not linked to the 16S and 23S rRNA genes and is transcribed separately. It has been reported that the 5S rRNA sequences from the two species *S. solfataricus* and *S. acidocaldarius* are identical, whereas the *S. shibatae* 5S rRNA sequence differs from the *S. solfataricus* and *S. acidocaldarius* 5S rRNA sequences at 4 of 124 positions (1, 12, 15). The 16S rRNA gene from an organism identified as *S. solfataricus* has been cloned and sequenced (9). We have used this same strain (obtained from N. Pace, Indiana University) to initiate studies on transcription, nucleolytic processing, and assembly of the rRNAs into ribosomal particles. By the criteria of 16S and 23S rRNA gene sequences, our laboratory strain is identical to that employed by Olsen et al. (9). We suggest that their identification of the strain as *S. solfataricus* was probably incorrect.

Strain identification. There is considerable confusion about the species identification of the *Sulfolobus* strains employed by ourselves and others. To resolve this ambiguity, we have compared by Southern hybridization genomic DNA from our laboratory strain with genomic DNA from authentic *S. acidocaldarius* (DSM639). Figure 1 illustrates the results of one Southern hybridization with an 859-bp *Sau*3AI fragment isolated from a library of our laboratory strain as the probe. This fragment contains 849 bp of sequence 5′ to the start of the 16S rRNA gene and the first 10 bp of the 16S rRNA gene. This probe hybridizes to fragments of identical size from parallel digestions of genomic DNA from the two strains. Similar

FIG. 1. Southern hybridization analysis of genomic DNAs. Genomic DNA was prepared from our laboratory strain and from authentic *S. acidocaldarius* (DSM639; lanes denoted by ●). The DNAs were digested as follows: *Bam*HI and *Xba*I (lanes A and B), *Hin*dII and *Kpn*I (lanes C and D), *Pst*I and *Nsi*I (lanes E and F), *Sma*I (lanes G and H), and *Sau*3AI (lanes I and J). Following agarose gel electrophoresis and transfer to a nylon membrane, the DNAs were probed with an 859-bp *Sau*3AI fragment labeled by the random primer method. The probe fragment was originally isolated from a genomic library prepared from our laboratory strain. The probe consists of 849 bp of 5′ gene sequence flanking the 16S rRNA gene and the first 10 nucleotides of the 16S rRNA-coding sequence. The approximate sizes (in kilobase pairs) of the genomic DNA fragments are indicated on the left, and the positions of molecular length standards (in kilobase pairs) are indicated on the right. The *Sma*I enzyme cleaves genomic DNA very infrequently because of the high AT DNA content. The large fragment complementary to the probe transfers poorly to the nylon membrane. Experimental procedures used were described by Maniatis et al. (7).

* Corresponding author. Mailing address: Department of Biochemistry and Molecular Biology, University of British Columbia, 2146 Health Sciences Mall, Vancouver, B.C., V6T 1Z3, Canada. Phone: (604) 822-5975. Fax: (604) 822-5227. Electronic mail address: pdp1@unixg.ubc.ca.

† Present address: Max Delbrück Center for Molecular Medicine, 0-1115 Berlin, Germany.

A.

```
                   20                  40                  60                  80                 100
PstI               |                   |                   |                   |                   |
CTGCAGCTAAGCTAGAAGGAGAAGAGACAGGTGAGATAGCAGAAGGTACTGGAGTAGCTATGGGAGACCCAGGTCCTGAAAAGATAGCTATAGAAAGAGT  100
CGCTGTTAAGTATAATATACCAATAGACGCAGTCATAGTAAAAATGAGCATGGAAGAGGCAATAACTGCAATGCCTCAGAAAGTTTATGAAGCAGCAGAT  200
AAAGTGGTGGAAATGATCAAGAGAATGATAATGGAAAGAACTAAACCAGGAGACACGATAATACTTGTAGGAGTAGGAAATACAGTAGGTGTAGCACAAT  300
GAGTCAGCAAGTACAACCCCCTCAAATTATAACTATCTTAAAATGTCTCAAGTGTAATTATTCCACAGAGAGACAATTCAGAGAAGGAGATTTTGTTACC  400
                           SacI
AAAGTTGAGGGAGCTTGTCCTACAGATGGAGCTCCTTTATACATTTGGGGGATATATGCCAAAAACGTAGAGCAAAAAAACACTAATCAGAAGTATTAAA  500
          ●●●●●●                    ●●●●●●
TAAGATACTTTTTTAAGCAACTGTATCCTATATTATGTTGCCCACCCGGTCATAGTGAGCGGGTAACACCCGGACTCGTTTCGAACCCGGAAGTTAAGCC  600
                 BamHI              5S rRNA
GCTCACGTCAGAGGGGCCGTGGGATCCGAGAGGGCCCGCAGCCTCTCTGAGCTGGGATGGGCGTTCTATTTAATTTATAATCCATTATTAAGATAAAGTA  700
TTTCAAATGAAAAAAGTATTCGATGAAATTCACGGTTACATCGAACTTAACGATATGGAAACGAAGATAATGGACTCACCTGTCTTTCAGCGACTCAGAA  800
                                         NsiI
GAATAAGACAAACTAGTTTAGCCTTTTTAGTCTATCCTGGCGCAATGCATACTAGATTTAGCCATTCTTTAGGTACATTCCACCTTGCTACTAGAGTAGG  900
TAATAGGCTTATAAATGAAGGTCATATAAACCATGATGAAGCTCTACTTCTTAGACTTGCCTCCCTTCTTCATGATATAGGACAATTTCCATTTAGCCAT 1000
                  XbaI
TCTCTTGAGCCTATATACATATCTAGA
```

B.

C.

Position	Sso	Ssh	Sac
1	-	C●	G●
11	U	C●	U
14	C	C	U●
25	C	C	U●
39	A	A	G●
70	U	U	C●
71	A	G●	A
73	U	U	A●
84	A	A	G●
85	U	U	A●
86	A	A	U●
90	U	U	A●
95	A	A	G●
96	U	U	C●
105	C	C	U●
107	A	A	U●
110	A	A	G●
123	U	C●	C●
124	U	U	G●

D.

```
Sso 5S DNA   5'..ACTAAGCTGGGATGGGTTTT
oligo oPD1   3'  TGATTCGACCCTACCCAAAA
mismatches        ●  ●            ●●
Sac 5S DNA   5'..TCTGAGCTGGGATGGGCGTT
                    |          |
                   110        120
```

FIG. 2. Nucleotide sequence of the 5S rRNA gene and predicted secondary structure of the 5S rRNA. (A) A 1,027-nucleotide *Pst*I-*Xba*I fragment from a λGEM11 library isolate was subcloned into plasmid pGEM-3ZF(+) and sequenced (13). The 5S rRNA gene sequence (underline) and major restriction sites (overlines) are indicated. The positions of putative box A and box B promoter elements are indicated (●). (B) The predicted secondary structure of the *S. acidocaldarius* 5S rRNA is illustrated. Positions of nucleotide substitutions which differ from the nucleotides in the published *S. solfataricus* and *S. shibatae* 5S rRNA sequences are indicated (*). (C) Nucleotide substitution positions and the substituted bases in the respective sequences are listed. Species abbreviations are as follows: Sso, *S. solfataricus*; Ssh, *S. shibatae*; and Sac, *S. acidocaldarius*. Nucleotides that differ from those in the *S. solfataricus* 5S sequence are indicated (●). (D) The sequence of the oligonucleotide (oligo) probe oPD1 and its complementarity to the 5S rRNA gene sequences from *S. solfataricus* (Sso) and *S. acidocaldarius* (Sac) are indicated. The mismatches to the *S. acidocaldarius* sequence are indicated (●).

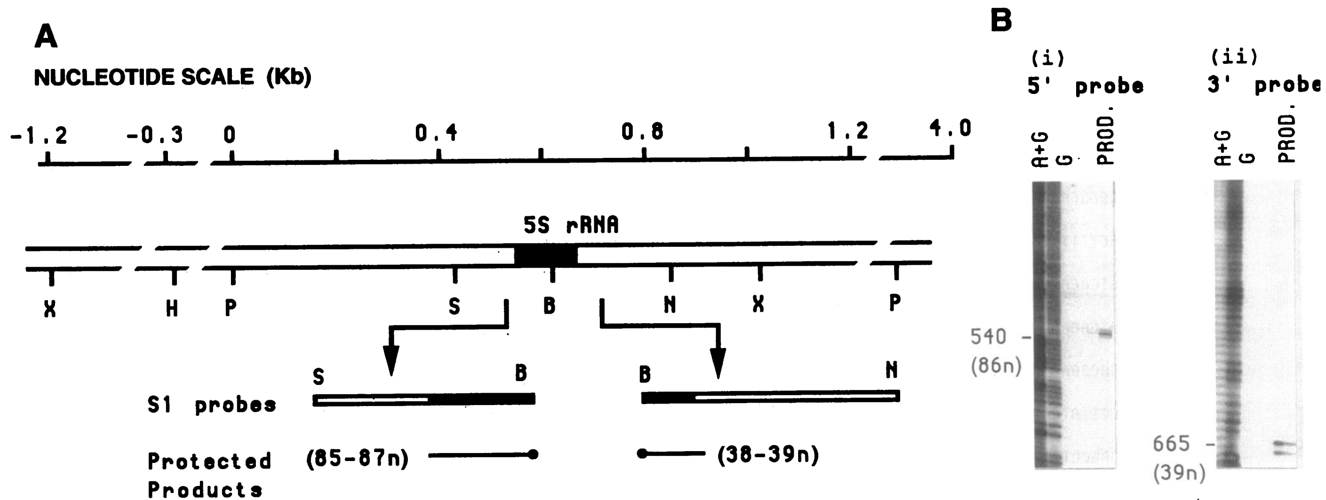

FIG. 3. Nuclease S1 mapping of the 5S rRNA gene transcript. Protection of end-labeled DNA fragments (about 30 ng; 5×10^4 cpm) from S1 nuclease digestion by total cellular RNA (10 μg; isolated from mid-log-phase cells) was used to locate the positions of the 5′ and 3′ ends of the 5S rRNA transcripts (2). (A) The genomic restriction map of the region containing the 5S rRNA gene is illustrated. Restriction enzyme sites and their positions within the sequence presented in Fig. 2 are as follows: X, *Xba*I (position 1022); H, *Hin*dIII; P, *Pst*I; S, *Sac*I (409); B, *Bam*HI (632); N, *Nsi*I (845). The 5S rRNA-coding region (solid box) is indicated. The *Sac*I-*Bam*HI and *Bam*HI-*Nsi*I restriction fragments used to locate the 5′ and 3′ transcript ends, respectively, are schematically indicated along with the nuclease-resistant products. (B) Autoradiograms illustrating (i) the protected products (PROD.) with the *Sac*I-*Bam*HI fragment 5′ end labeled on the minus strand at position 626, and (ii) the protected products with the *Bam*HI and *Bam*HI-*Nsi*I fragment 3′ end labeled on the minus strand at position 626. The molecular length ladders were generated by the A and A+G reactions of Maxam and Gilbert (8). The length (in nucleotides [n]) of the major protected fragment and the end position within the nucleotide sequence of Fig. 2 are indicated.

results were obtained with a number of different DNA probes and a number of different restriction enzymes. If the two DNAs were from different species, we would have expected to see at least some restriction site polymorphism. We conclude that our laboratory strain is identical to authentic *S. acidocaldarius* (DSM639).

This conclusion is supported by the work of Kurosawa and Itoh (6), who recently published the 16S rRNA sequence for *S. acidocaldarius* ATCC 33909. Their sequence is 99.9% identical to the 16S rRNA sequence we determined and to the 16S rRNA sequence reported by Olsen et al. (9). In addition, they partially sequenced the 16S rRNA gene from *S. solfataricus* DSM1616 and demonstrated that it was only 91% identical to the 16S rRNA sequences from *S. acidocaldarius* ATCC 3309 and DSM639.

5S rRNA gene. It has been reported that the 5S rRNA sequences from the two species *S. acidocaldarius* and *S. solfataricus* are identical (1, 15). An attempt to clone the 5S rRNA gene sequence from our laboratory strain of *S. acidocaldarius* with a 20-mer oligonucleotide sequence (oPD1) complementary to the 3′ end of the published 5S rRNA sequences was unsuccessful; the only probe-related sequence in the genome was shown to lack the adjacent *Bam*HI restriction site (predicted from the published 5S rRNA sequences) and to exhibit no sequence similarity to 5S rRNA outside the probe binding site. In addition, primer extension analysis with the oligonucleotide using total cellular RNA as template failed to yield any detectable extension products.

An attempt to clone the 5S rRNA gene using agarose gel-purified 5S rRNA end labeled with pCp and RNA ligase (5) was successful. The probe hybridized to unique sequence DNA and was used to identify a phage clone from a λGEM11 genomic library. The 5S rRNA gene was subcloned as an 1,137-nucleotide *Pst*I-*Xba*I fragment, and its nucleotide sequence was determined (Fig. 2A). By phylogenetic comparison, the 124-nucleotide 5S rRNA-encoding gene was located from positions 540 to 663. A partial genomic restriction map and the location of the 5S gene are illustrated in Fig. 3.

The 5S rRNA gene sequence exhibits 17 nucleotide differences from the published *S. solfataricus* and "*S. acidocaldarius*" sequences and 18 nucleotide differences from the related *S. shibatae* sequence (Fig. 2B to D). All of these sequences can be folded into the secondary structure typical of 5S rRNA. The positions of nucleotide substitution between the three sequences do not significantly alter the secondary structure of the 5S rRNA molecule. When substitutions occur in helical regions, they are most often compensatory; substitutions in loops or bulges tend to elongate the duplex region in the *S. acidocaldarius* structure. Comparison of the oPD1 20-mer sequence with the experimentally determined 5S gene sequence indicates that there were 4 mismatched base pairs. One mismatch is the 3′-terminal T residue of oPD1, which explains the failure of the oligonucleotide to prime a product in the primer extension assay (see above).

The strain used by Stahl et al. (15) for 5S rRNA sequence determination was obtained from T. Langworthy (University of South Dakota, Vermillion) and was believed to be *S. acidocaldarius*. Our 5S sequence indicates that this strain was apparently not *S. acidocaldarius* but was almost certainly, on the basis of the 5S rRNA sequence reported, *S. solfataricus*.

Mapping of 5S rRNA gene transcripts. The positions of putative transcription initiation and termination sites and of mature rRNA end sites have been localized by a combination of primer extension assays (data not shown) and S1 nuclease protection assays (2, 7). To characterize the 5′ end of the 5S rRNA gene transcript, a 194-nucleotide *Sac*I-*Bam*HI fragment was 5′ end labeled on the minus strand at nucleotide position 626, within the 5S rRNA gene (Fig. 3). The product protected from S1 nuclease digestion by total cellular RNA was 86 to 88 nucleotides long and corresponds to a 5′ transcript end site at or near nucleotide position 540. No longer protection products were detected. The 3′ end of the 5S rRNA gene transcript was

located by using a 224-nucleotide *Bam*HI-*Nsi*I fragment, 3′ end labeled on the minus strand at nucleotide position 626. The S1 nuclease-resistant products were 38 or 39 nucleotides long and correspond to a 3′ transcript end at or near nucleotide position 663. Again, no longer protection products were evident.

By phylogenetic alignment, the mature 5S rRNA-coding sequence extends from nucleotide positions 540 to 663. These positions are at or immediately adjacent to 5S transcript end sites mapped by S1 nuclease protection. Because longer protection products were not observed, the 5S rRNA gene transcript is probably not processed at either the 5′ or 3′ end but rather is equal in length to the mature form that is assembled directly into 50S subunits. The G residue at position 600 is located within a promoter-like box B element that is preceded at an appropriate distance by a box A element. A related promoter containing the same conserved sequence elements and driving the expression of the single 16S-23S rRNA operon in *S. shibatae* has been extensively characterized by site-directed mutagenesis and in vitro transcription analysis (10). The T-rich sequence immediately downstream of the 5S rRNA-coding sequence is characteristic of archaebacterial transcription termination signals (11, 14).

This work was supported by grants from the Medical Research Council of Canada (to P.P.D.). P.P.D. is a Fellow of the Canadian Institute for Advanced Research.

REFERENCES

1. **Dams, E., P. Londei, P. Cammarano, A. Vandenbeighe, and R. DeWachter.** 1983. Sequence of the 5S rRNAs of the thermoacidophilic archaebacterium *Sulfolobus solfataricus* (*Calderiella acidophila*) and the thermophilic eubacteria *Bacillus acidocaldarius* and *Thermus aquaticus*. Nucleic Acids Res. **11:**4667–4676.
2. **Dennis, P.** 1985. Multiple promoters for the transcription of the ribosomal RNA gene cluster in *Halobacterium cutirubrum*. J. Mol. Biol. **186:**457–461.

2a. **Durovic, P., and P. P. Dennis.** Submitted for publication.

3. **Grogan, D., P. Palm, and W. Zillig.** 1990. Isolate B12, which harbors a viruslike element represents a new species of the archaebacterial genus *Sulfolobus*, *Sulfolobus shibatae*, sp. nov. Arch. Microbiol. **154:**594–599.
4. **Kjems, J., and R. Garrett.** 1990. Secondary structural elements exclusive to the sequences flanking ribosomal RNAs lend support to the monophyletic nature of the archaebacteria. J. Mol. Evol. **31:**25–32.
5. **Krupp, G., and H. Gross.** 1983. Sequence analysis of *in vivo* ^{32}P-labeled RNA, p. 11–58. *In* P. F. Agris and R. A. Kopper (ed.), The modified nucleosides of transfer RNAII. Alan R. Liss, Inc., New York.
6. **Kurosawa, N., and Y. Itoh.** 1993. Nucleotide sequence of the 16S rRNA from the thermoacidophilic archaea *Sulfolobus acidocaldarius* ATCC 33909. Nucleic Acids Res. **21:**357.
7. **Maniatis, T., E. F. Fritsch, and J. Sambrook.** 1982. Molecular cloning: a laboratory handbook. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
8. **Maxam, A., and W. Gilbert.** 1980. Sequencing end-labeled DNA with base-specific chemical cleavages. Methods Enzymol. **65:**499–560.
9. **Olsen, G., N. Pace, M. Nuell, B. Kaine, R. Gupta, and C. Woese.** 1985. Nucleotide sequence of the 16S rRNA gene from the thermoacidophilic archaebacterium *Sulfolobus solfataricus*. J. Mol. Evol. **22:**301–307.
10. **Reiter, W., U. Hadepohl, and W. Zillig.** 1990. Mutational analysis of an archaebacterial promoter: essential role of a TATA box for transcription efficiency and start site selection in vitro. Proc. Natl. Acad. Sci. USA **87:**9509–9513.
11. **Reiter, W., P. Palm, and W. Zillig.** 1988. Transcription termination in the archaebacterium *Sulfolobus*: signal structures and linkage to transcription initiation. Nucleic Acids Res. **16:**2245–2459.
12. **Reiter, W. D., P. Palm, W. Voos, J. Kaniecki, B. Grampp, W. Schulz, and W. Zillig.** 1987. Putative promoter elements for the ribosomal RNA genes of thermoacidophilic archaebacterium *Sulfolobus* sp. strain B12. Nucleic Acids Res. **15:**5581–5595.
13. **Sanger, F., S. Nicklen, and A. R. Coulsen.** 1977. DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. USA **74:**5463–5467.
14. **Shimmin, L., and P. Dennis.** 1989. Characterization of the L11, L1, L10 and L12 equivalent ribosomal protein gene cluster of the halophilic archaebacterium *Halobacterium cutirubrum*. EMBO J. **8:**1225–1235.
15. **Stahl, D., K. Leuhrsen, C. Woese, and N. Pace.** 1981. An unusual 5S RNA from *Sulfolobus acidocaldarius* and its implications for a general 5S rRNA structure. Nucleic Acids Res. **9:**6129–6137.
16. **Stetter, K., A. Segerer, W. Zillig, G. Huber, G. Fiala, R. Huber, and H. Konig.** 1986. Extremely thermophilic sulfur metabolizing archaebacteria. Syst. Appl. Microbiol. **7:**393–397.