Amplification and Direct Sequence Analysis of the 23S rRNA Gene from Thermophilic Bacteria

BioTechniques 30:414-420 (February 2001)

Ashraf Ibrahim[1], Jacob Hofman-Bang[1], and Birgitte K. Ahring[1,2]

[1]The Technical University of Denmark, Lyngby, Denmark and [2]University of California, Los Angeles, CA, USA

ABSTRACT

We present a simplified and fast method to obtain high-quality sequences directly from PCRs without the traditional gel purification. We also report on an improved method to obtain sequence-quality PCR products from microorganisms that are difficult to lyse with no need for DNA extraction. The technique uses exonuclease I and shrimp alkaline phosphatase to degrade residual dNTPs and primers. Our technique is shown to work on both Gram-positive and Gram-negative bacteria.

INTRODUCTION

Gene amplification techniques integrated with direct sequence analysis protocols have established themselves as powerful tools in almost all biological disciplines. Microbiologists in particular have been blessed with the privileges of such techniques in studying microbial diversity, a concept of microbial ecology that has been greatly underestimated in traditional cultural methodology.

The ribosomal RNA genes, especially 16S rRNA, have paved the way for establishing phylogenetic relationships among living organisms, whereby a well-cemented evolutionary interrelationship among all living organisms is currently shaped on a solid basis (12). Moreover, the 16S rRNA gene, with more than 9000 entries in the databases (6), has provided researchers with a considerable number of gene probes for the rapid identification of microorganisms (1). However, with the continual identification of new microbial taxa, the 16S rRNA approach appears to be suffering from two major drawbacks. On the phylogenetic level, an overlapping of closely related microbial species seems to overshadow the potential of this gene as a reliable phylogenetic tool in the long run. More importantly, designing species- and subspecies-specific probes on the basis of 16S rRNA is becoming a more daunting task over time. Recently, the potential of the 23S rRNA gene as a phylogenetic and identification tool has started to attract the attention of more researchers (13). The inability to amplify the full gene by PCR, because of strong secondary structure and because the sequencing costs until recently have been too high, has hindered amplification and subsequent direct sequencing of the 23S rRNA gene. Yet, being almost twice the size of 16S, the 23S rRNA should statistically provide additional DNA polymorphism for probe development and more information for constructing phylogenetic trees. Commonly, 23S rRNA genes are amplified as three or even more PCR fragments (2,9), thus increasing both the time and cost of the experiment. Furthermore, sequence strategies developed for the 23S rRNA have relied mainly on cloning the gene, generating single-stranded templates, or direct sequencing of several partially amplified fragments.

In the present communication, we describe a protocol that allows (*i*) amplification of the complete 23S rRNA gene directly from microorganisms that are difficult to lyse without the need for DNA extraction and (*ii*) direct fluorescent sequencing of the amplified double-stranded fragment with no need for cloning or preparation of single-stranded template. A protocol for amplification and subsequent sequencing of the 23S rDNA from four different thermophilic bacteria is described to demonstrate the potential of the reported methodology.

MATERIALS AND METHODS

Bacterial Strains

Thermoanaerobacter mathranii (DSM no. 11426), *Thermoanaerobacter thermohydrosulfuricus* (DSM no. 567), *Thermoanaerobacter ethanolicus* (DSM no. 22466), and *Thermoanaer-*

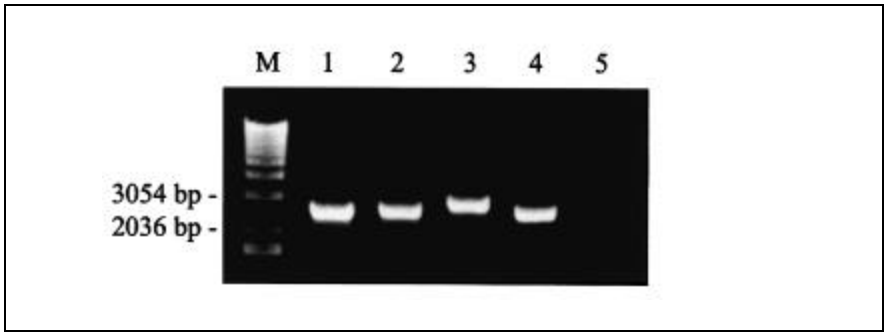

Figure 1. Amplification of the ribosomal DNA fragment encompassing the 16S–23S spacer region and the 23S rDNA. Lane M, 1-kb ladder (Life Technologies, Rockville, MD, USA). Lanes 1–4, 10% of the PCR product from *T. mathranii*, *T. thermohydrosulfuricus*, *T. ethanolicus*, and *T. thermocopriae*. Lane 5, negative control.

obacter thermocopriae (ATCC no. 51646) were obtained from our culture collection.

PCR

Frozen cultures of *Thermoanaerobacter* strains were first thawed to room temperature. Cells were then pelleted in a benchtop centrifuge and washed twice in 1 mL TE buffer. The pellet was resuspended in the same buffer, and the cell density was adjusted to 10^6 cells/mL. One microliter of the cell suspension was added to a PCR tube containing 5 µL 1.2× PCR buffer, prepared in our laboratory [10× buffer consists of 500 mM Tris-HCl, 100 mM KCl, 20 mM $MgCl_2$, 10% dimethyl sulfoxide (DMSO), 10 mg/mL bovine serum albumin, and 10% betaine] and 41 µL distilled water. The tubes were overlaid with 35 µL mineral oil and heated in a model 9700 thermal cycler (Applied Biosystems, Foster City, CA, USA) for 30 min. After this denaturation period, 1 µL dNTPs (10 mM) and 1 µL primers (50 µM) were added, followed by 0.4 µL (2 U) *Taq* DNA polymerase (Amersham Pharmacia Biotech, Little Chalfont, UK), keeping the thermal cycler at the high temperature. The amplification primers are listed in Table 1 and graphically outlined in Figure 3. The cycling continued with the following profile: one cycle at 94°C for 2 min, 10 cycles (each consisting of denaturation at 94°C for 1 min, annealing at 60°C for 30 s, and extension at 72°C for 2 min), 20 cycles (each consisting of denaturation at 94°C for 30 s, annealing at 50°C for 30 s, and extension at 72°C for 2 min), and a final extension at 72°C for 7 min. The tubes were immediately cooled to 4°C. The PCR temperature profile is designed to specifically amplify the gene of interest with high specificity in low amounts for 10 cycles followed by amplification for 20 cycles under normal conditions (i.e., the primer 2747R used in the initial PCR has a melting temperature lower than the annealing temperature used in the PCR for the first 10 cycles). Ten percent of the PCR product was separated on a 1% agarose gel and visualized by UV light after ethidium bromide staining.

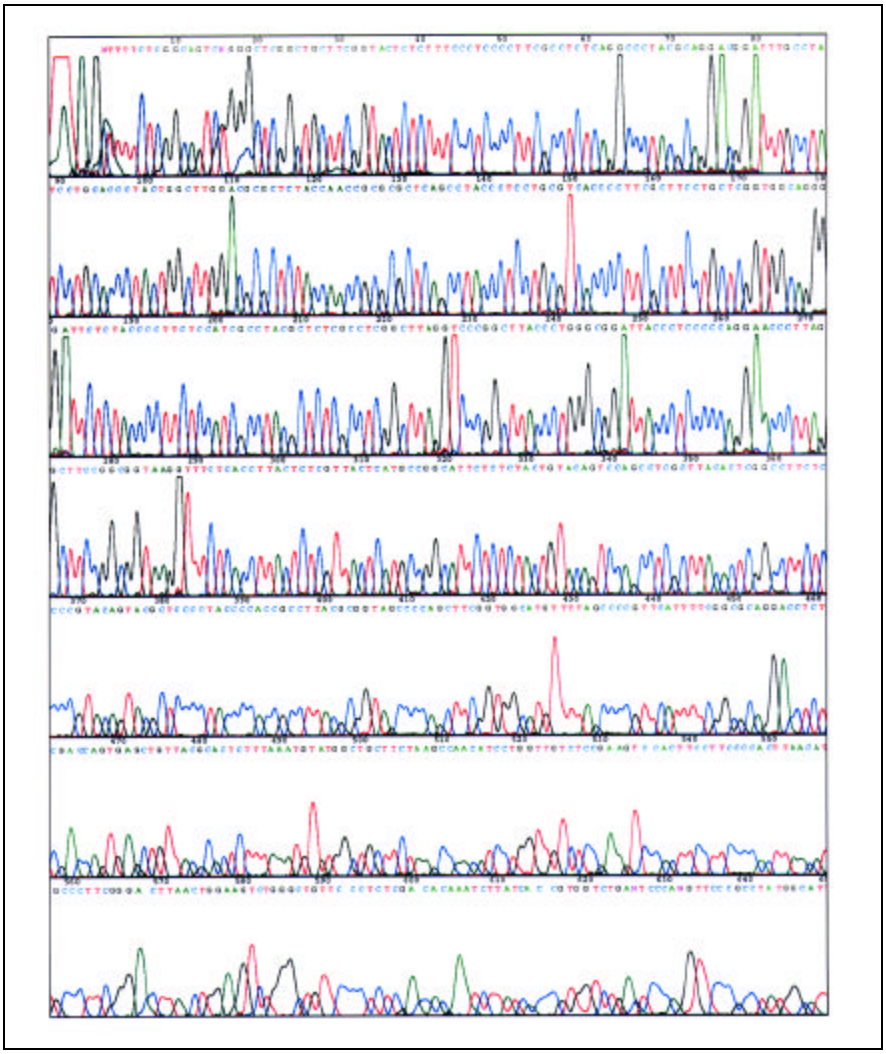

Figure 2. Electropherogram of ABI 377 automated sequencing of 23S rDNA gene from *T. mathranii* using primer 1091R High-quality sequence is obtained even after 650 bases with very low background.

Template Preparation for Sequencing

For each sequencing reaction, 3 µL PCR product were combined with 0.3 µL exonuclease I and 0.3 µL shrimp alkaline phosphatase (Amersham Pharmacia Biotech) in a 1.5-mL Eppendorf® tube. The reactants were incubated at 37°C for 15 min, followed by incubation at 80°C for 10 min to inactivate any residual enzyme.

Sequencing of 23S rDNA

Sequencing was carried out using the *Taq* Dyedeoxy Terminator Cycle Sequencing Kit (Applied Biosystems) according to the supplier's recommendations and ABI 377 automated sequencer with sequencing primers listed in Table 1 according to the protocol described in Table 2.

RESULTS AND DISCUSSION

The growth of many anaerobic microorganisms necessitates special enrichment and isolation requirements. Yet, the cell harvest is hardly adequate for further identification, even at the molecular level. Both the yield and quality of the extracted DNA are not usually suitable for PCR amplification and subsequent manipulation. Ruano et al. (8) have described a more efficient PCR amplification protocol, whereby DNA is subjected to longer denaturation period in 1.1× PCR buffer before the rest of the amplification mixture is added. In our attempt to develop a more robust ampli-

fication strategy for DNA amplification from anaerobic organisms, we decided to take the protocol even further and eliminate the entire DNA extraction procedure. A PCR buffer was formulated in our laboratory, and 10^3 cells were added to 1.2× amplification buffer. The mixture was heated for different periods ranging from 20 min to 1 h, followed by the addition of dNTPs, primers, and *Taq* DNA polymerase. Initial results showed that a 30-min initial denaturation followed by 30 cycles in a step-down PCR led to the amplification of a discrete fragment of the expected size. Still, to avoid the time-consuming methodology to obtain fresh microbial cell harvest, we have tested the protocol on frozen cultures of both Gram-positive anaerobes (*Thermoanaerobacter* species) and Gram-negative aerobes (data not shown; *Francisella tularensis* and *E. coli*). A single amplification product encompassing the 16S–23S spacer region and 23S rDNA was obtained (Figure 1). Negative photographs of the agarose gel showed no visible co-amplification products. Note the different size in Figure 1, lane 4 due to a longer 16S–23S spacer region.

To obtain successful and reproducible sequencing profiles from PCR products, the product must be thoroughly cleaned with no residual salts, agarose impurities, or co-amplification products. The most common technique to achieve this objective so far is agarose gel electrophoresis separation and subsequent purification of the sliced agarose fragment containing the product of interest. A more recent strategy is to rapidly pass the PCR through a purification column of a certain pore size, followed by elution of the product in distilled water or TE buffer. Both techniques are costly and still lead to inconsistent sequencing profile, mainly because of residual salts and agarose impurities. A recent purification technique reported by Amersham Pharmacia Biotech relies on the direct addition of two enzymes to the PCR whereby residual dNTPs and primers are completely degraded. However, an important precondition for this rather simple strategy is to obtain a single discrete PCR product during the amplification process. As the optimized PCR protocol described here satisfied this requirement, this enzymatic purification proved to be an optimal choice to further facilitate our sequencing strategy.

Table 1. Amplification and Sequencing Primers

Designation	Primer Sequence
1522F[a,b]	5′-TGCGGTGGATCACCTCCTT-3′
130R	5′-CCTTGCCCCATTCGG-3′
250F[c]	5′-AGTAG(C/T)GGCGAGCGAA-3′
1091R[c]	5′-(A/G)GTGAGCT(A/G)TTACGC-3′
1104F[c]	5′-(A/T)GCGTAA(C/T)AGCTCAC-3′
1410R[c]	5′-CTCGGCTTAGG(C/T)CCC-3′
1606R[c]	5′-C(C/T)ACCTGTG(A/T)CGGTTT-3′
1930R	5′-CGACAAGGAATTTCGCTAC-3′
2241R[c]	5′-ACCGCCCCAGT(A/C/T)AAACT-3′
2512F[c]	5′-CCTCGATGTCG(A/G)CTC-3′
2747R[a,b,c]	5′-G(C/T)TTAGATGC(C/T)TTC-3′

[a]Primers used for both amplifications of the target fragment.

[b]The primer is located at the 3′ end of the 16S rRNA gene. All other primers are within the 23S rRNA gene. All numberings are according to the *E. coli* numbering system (3). The number indicates the 5′ position, and "F" indicates forward (i.e., downstream) and "R" indicates reverse (i.e., upstream).

[c](...) indicates mixed nucleotides at one position.

The 23S rRNA gene, though containing both more phylogenetic information and polymorphism to design more specific probes than the 16S rRNA gene, has been far less investigated. This can be attributed to two principal reasons: (*i*) the strong secondary structure of the gene that hinders the proper extension after primer annealing, thus leading to short sequence reads and (*ii*) the inefficient amplification protocols described so far for 23S rDNA that entice researchers to the simpler option of 16S rRNA analysis. Partial amplification and subsequent direct sequencing (2,9), cloning of the partially amplified gene (11), or preparation of single-stranded template (5) have been the available options for scientists willing to analyze the 23S rRNA. These approaches are not only time consuming and costly but also discourage researchers from pursuing a more reliable candidate for a plethora of genetic information. So, it is not surprising that less than 200 full 23S ribosomal RNA genes are available in the databases compared to the thousands of 16S ribosomal RNA sequences. Using our described sequence protocol, it was possible to alleviate most of the inherent sequencing hurdles associated with this gene and possibly other difficult templates. Figure 2 shows one of the electropherograms of sequencing of the 23S rDNA gene. Read lengths of up to 640 bases (nucleotides 10–650) per reaction were obtained from both reverse and

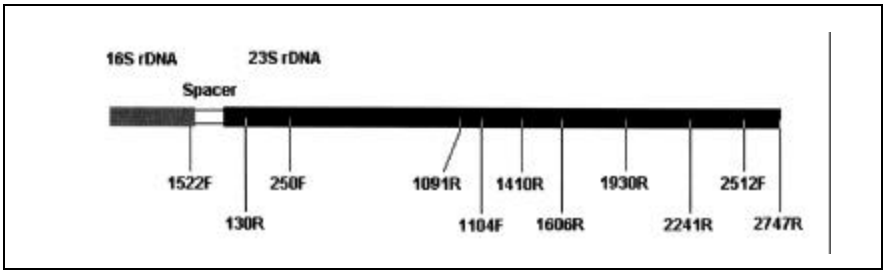

Figure 3. Schematic outline of the primer locations used in the direct sequencing of the 16S rDNA–Spacer–23S rDNA gene region. The names of the primers denotes the position in the gene when compared to the corresponding genes in *E. coli*, and "R" of "F" denote the orientation of the primers [i.e., reverse or forward (3)]. The gray region indicates the 3′ end of the 16S rDNA gene; the white region indicates the Spacer region, and the black region indicates almost all of the 23S rDNA region.

Table 2. Sequencing Protocol

1. For each sequencing reaction, combine the following in a 200-µL tube: 6 µL cleaned PCR product, 0.5 µL 1% betaine, and 4.5 µL distilled water.
2. Overlay with 15 µL mineral oil and incubate in a model 9700 thermal cycler (Applied Biosystems) at 94°C for 30 min.
3. Meanwhile, combine the following reagents (all provided in the kit except the primers) for each 10 sequence reactions: 40 µL 5× sequencing buffer, 10 µL dNTP mixture, 20 µL (5 µL each dye terminator), 10 µL AmpliTaq® DNA Polymerase FS, and 10 µL sequencing primer (3 pmol/µL).
4. After the 30-min denaturation cycle, add 9 µL premixture prepared in Step 3. Never allow the temperature to drop below 90°C.
5. Immediately start cycle sequencing as follows: 10 cycles each with the following profile; 96°C for 30 s, 50°C for 5 s, and 60°C for 6 min. 15 Cycles each with the following profile: 96°C for 30 s and 60°C for 4 min.
6. Rapidly cool the tubes to 4°C.
7. In clean 200-µL tubes combine 2 µL 3 M sodium acetate (pH 5.2) and 50 µL 95% ice-cold ethanol.
8. For each tube, add 19 µL sequencing reaction (avoiding the mineral oil). Mix gently and incubate on ice for 15 min.
9. Spin down for 15 min at 15000× *g*. Carefully aspirate the liquid against a light source and wash twice with 75% ice-cold ethanol.
10. Spin the tubes down in the same orientation so that the pellet is kept stable on the bottom side of the tube.
11. Aspirate any traces of ethanol and simply incubate the tubes wide open at 37°C for 5 min. This is more than sufficient to get a completely dry pellet.
12. Loading buffer, denaturation step, and electrophoresis are carried out according to the sequencer model.

forward primers with superior sequence quality. In addition, less than eight sequence primers were needed to achieve the full sequence of the 23S rDNA. Overall, a continuous stretch of approximately 2700 bases was obtained with the primers listed in Table 1. Note that the 16S–23S spacer region in certain bacterial species will not be possible to obtain because these loci are not linked.

Two of these primers were designed in our laboratory through the analysis of all available 23S sequences in the databases (4,6,10). Primer 130R was specifically designed to extend through the spacer region and thus enabled us to read the 5′ end of the gene starting from position 1. Primer 1410R closed a gap around position 1300 and eliminated the need for two additional primers. Only three forward primers were needed to resolve some sequence ambiguities.

In conclusion, the described amplification protocol, in which there is no need for any DNA extraction and regardless of the age of the microbial cultures, should facilitate the amplification of difficult templates exemplified here by the 23S rRNA gene. Because of the stringent conditions applied during the amplification process, there were no coamplification products, and the quick purification protocol was used to prepare the template. With more stringent conditions during fluorescent sequencing, including the use of betaine and a long presequencing denaturation period, it was possible to directly sequence almost the full 23S rDNA with far fewer sequencing primers and longer reads. This protocol should have the potential to handle other difficult-to-sequence genes and save time and cost of partial amplification, cloning, or preparation of single-stranded templates when individual primer selections and PCR conditions are optimized.

ACKNOWLEDGMENT

This work was supported by The Danish Biotechnology Program.

REFERENCES

1. **Alm, E.W., D.B. Oerther, N. Larsen, D.A. Stahl, and L. Raskin.** 1996. The oligonucleotide probe database. Appl. Environ. Microbiol. *62*:3557-3559.
2. **Ash, C. and M.D. Collins.** 1992. Comparative-analysis of 23S-ribosomal RNA gene-sequences of Bacillus-anthracis and emetic Bacillus-cereus determined by PCR-direct sequencing. FEMS Microbiol. Lett. *94*:75-80.
3. **Brosius, J., T.J. Dull, and H.F. Noller.** 1980. Complete nucleotide sequence of a 23S ribosomal RNA gene from *Escherichia coli*. Proc. Natl. Acad. Sci. USA *77*:201-204.
4. **De Rijk, P., E. Robbrecht, S. de Hoog, A. Caers, Y. Van de Peer, and R. De Wachter.** 1999. Database on the structure of large subunit ribosomal RNA. Nucleic Acids Res. *27*:174-178.
5. **Lawson, V.A., D.A. McPhee, and N.J. Deacon.** 1996. Elimination of sequence ambiguities by a single-step modification of a solid-phase, single-stranded sequencing protocol. BioTechniques *21*:356-358.
6. **Maidak, B.L., J.R. Cole, C.T. Parker, Jr., G.M. Garrity, N. Larsen, B. Li, T.G. Lilburn, M.J. McCaughey et al.** 1999. A new version of the RDP (Ribosomal Database Project). Nucleic Acids Res. *27*:171-173.
7. **Mauel, M.J., S.J. Giovannoni, and J.L. Fryer.** 1999. Phylogenetic analysis of *Piscirickettsia salmonis* by 16S, internal transcribed spacer (ITS) and 23S ribosomal DNA sequencing. Dis. Aquat. Organ. *35*:115-123.
8. **Ruano, G., E.M. Pagliaro, T.R. Schwartz, K. Lamy, D. Messina, R.E. Gaensslen, and H.C. Lee.** 1992. Heat-soaked PCR—an efficient method for DNA amplification with applications to forensic analysis. BioTechniques *13*:266-274.
9. **Sallen, B., A. Rajoharison, S. Desvarenne, F. Quinn, and C. Mabilat.** 1996. Comparative analysis of 16S and 23S rRNA sequences of Listeria species. Int. J. Syst. Bacteriol. *46*:669-674.
10. **Stoesser, G., M.A. Tuli, R. Lopez, and P. Sterk.** 1999. The EMBL Nucleotide Sequence Database. Nucleic Acids Res. *27*:18-24.
11. **Vancamp, G., S. Chapelle, and R. Dewachte.** 1993. Amplification and sequencing of variable regions in bacterial 23S ribosomal-RNA genes with conserved primer sequences. Curr. Microbiol. *27*:147-151.
12. **Woese, C.** 1998. The universal ancestor. Proc. Natl. Acad. Sci. USA *95*:6854-6859.
13. **Woo, T.H.S., L.D. Smythe, M.L. Symonds, M.A. Norris, M.F. Dohnt, and B.K.C. Patel.** 1997. Rapid distinction between Leptospiral interrogans and Leptospira biflexa by PCR amplification of 23S ribosomal DNA. FEMS Microbiol. Lett. *150*:9-18.

Received 20 March 2000; accepted 1 November 2000.

Address correspondence to

Professor B.K. Ahring
Department of Biotechnology
BLD 227
The Technical University of Denmark
2800 Lyngby, Denmark
e-mail: bka@imt.dtu.dk