Family- and Genus-Level 16S rRNA-Targeted Oligonucleotide Probes for Ecological Studies of Methanotrophic Bacteria

JAY GULLEDGE,† AZEEM AHMAD, PAUL A. STEUDLER, WILLIAM J. POMERANTZ, AND COLLEEN M. CAVANAUGH

The Biological Laboratories, Harvard University, Cambridge, Massachusetts 02138, and The Ecosystems Center, Marine Biological Laboratory, Woods Hole, Massachusetts 02543

Received 9 April 2001/Accepted 17 July 2001

Methanotrophic bacteria play a major role in the global carbon cycle, degrade xenobiotic pollutants, and have the potential for a variety of biotechnological applications. To facilitate ecological studies of these important organisms, we developed a suite of oligonucleotide probes for quantitative analysis of methanotroph-specific 16S rRNA from environmental samples. Two probes target methanotrophs in the family *Methylcoccaseae* (type II methanotrophs) as a group. No oligonucleotide signatures that distinguish between the two genera in this family, *Methylcocca* and *Methylosinus*, were identified. Two other probes target, as a single group, a majority of the known methanotrophs belonging to the family *Methylcoccaseae* (type I/X methanotrophs). The remaining probes target members of individual genera of the *Methylcoccaseae*, including *Methyllobacter*, *Methylomonas*, *Methylomicrobium*, *Methylococcus*, and *Methylocaldum*. One of the family-level probes also covers all methanotrophic endosymbionts of marine mollusks for which 16S rRNA sequences have been published. The two known species of the newly described genus *Methylosarcina* gen. nov. are covered by a probe that otherwise targets only members of the closely related genus *Methylomicrobium*. None of the probes covers strains of the newly proposed genera *Methylocaldum* and “*Methylothermus*,” which are polyphyletic with respect to the recognized methanotrophic families. Empirically determined midpoint dissociation temperatures were 49 to 57°C for all probes. In dot blot screening against RNA from positive- and negative-control strains, the probes were specific to their intended targets. The broad coverage and high degree of specificity of this new suite of probes will provide more detailed, quantitative information about the community structure of methanotrophs in environmental samples than was previously available.

MATERIALS AND METHODS

Bacterial cultures. The reference cultures used in this study were obtained from various sources, as indicated, and are available from either the National Collection of Industrial and Marine Bacteria (NCIMB, Aberdeen, United Kingdom) or the American Type Culture Collection (ATCC, Manassas, Va.).
sequence strains include Methylophilus trichosporium OB3b (NCIMB 11131) and Methylococcus capsulatus Bath (NCIMB 11132) (both provided by J. C. Murrell), Methylohalobacterunix A45 (nonex tant culture; genomic DNA provided by A. A. DiSpirito), Methylibiromycobium album BGS (NCIMB 11123; provided by G. M. King), Methylostomax rivalis (NCIMB 11193) and Methylostomax methanica SI (NCIMB 11130) (both provided by J. D. Semrau), Methylospallarcinac (NCIMB 11192; purchased from NCIBM), Cualobacter crescentus CB15A (ATCC 19009; provided by J. S. Pointon), and Escherichia coli UT1K1LH7 (ATCC 11775; from laboratory stock culture).

All methanotrophs were grown at 30°C with Methylococcus capsulatus Bath and Methylohalobacterunix grilae, which were grown at 45°C, in nitrate mineral salts medium with CH4 and CO2 at an initial headspace mixing ratio of 45:5:50 (CH4, to CO2, to air) (35). E. coli was grown in Luria-Bertani broth under standard conditions (53), and C. crescentus CB15A was grown in PYCM medium (27) at room temperature.

Sequencing of 16S rRNA genes. Because ambiguous and missing bases in several of the sequences available from GenBank hindered sequence comparisons, we resequenced the 16S rRNA genes of Methylohalobacterunix NCIMB 11913, Methylophilusluter NCIMB 11914, Methylostomax methanica SI NCIMB 11130, and Methylostomax marinus strain A45. Nearly complete (1,450-bp) sequences were obtained for both the sense and antisense strands of the 16S rRNA gene using 5′% Long Ranger gel and an ABI PRISM DNA sequence (41).

Selection of reference sequences. Probes were designed based on reference 16S rRNA sequences available from GenBank (6) and the Ribosomal Database Project (RDP-II) (42), as well as as resequencing of key laboratory strains (see Table 1 and Fig. 3). BLAST (GenBank) and Probe Match (RDP-II) database searches were used to assess the potential breadth and specificity of the probe sequences. The reference sequences were aligned with the probe sequences to determine the apparent range of coverage of the candidate probes relative to the abundance and diversity of known methanotrophs. The 16S rRNA sequences specified by accession numbers in Fig. 2 and 3 represent all those available in the databases for confirmed methanotrophic isolates at the time of analysis. With the exception of the methanotrophic endosymbionts of marine mollusks (see below), we did not include sequences obtained from cultures that had not been characterized phenotypically or that were obtained by PCR amplification of environmental DNA.

Only cultured isolates with published, genus-level phylogenetic data were assigned genus designations in Fig. 2 and 3. Three general groups of confirmed methanotrophs were placed under “other α-methanotrophs” (Fig. 2) or “other γ-methanotrophs” (Fig. 3); (i) strains clearly belonging to the α- or γ-methanotrophs, but lacking or having dubious generic affiliations because of insufficient phylogenetic and taxonomic information (for example, “Methylostomax methanica” strain S1Z is clearly a γ-methanotroph [57] but has not been characterized at the genus level); (ii) isolates validly assigned to the genera Methylohydroilla, Methylohydroilla, and Methylibiromycobium, for which we did not design genus-level probes because there were only one or two known representatives of each genus; and (iii) the methanotrophic endosymbionts of marine mollusks, which lack generic descriptions. Although uncultured marine mollusk endosymbionts are included because they are of interest to microbial ecologists and evolutionary biologists and because there is strong phenotypic and phylogenetic evidence that they are γ-methanotrophs (19, 20, 25, 28).

All available methanotroph 16S rRNA sequences that met the criteria given above were included in our analysis, regardless of sequence quality. However, a number of sequences appeared to be affected by common sequencing errors, including transposition of bases and duplicated or omitted bases. Some errors could be confirmed because they violated the integrity of the secondary structure of the 16S rRNA molecule, but others could not because they occurred in unpaired loop positions. Because sequence errors make designing genus-level probes very difficult, we developed specific criteria for disregarding unexpected mismatches between a probe and a target sequence. We deemed destabilization of secondary structure sufficient grounds for disregarding mismatches. Additionally, we considered any two of the following criteria sufficient: (i) the mismatch occurs in a low-quality sequence as indicated by ambiguous bases in >0.5 of the positions in the entire sequence; (ii) the mismatch results from an ambiguous or missing base in the probe target region; (iii) multiple sequences for the same strain disagree in the mismatch position, and the higher-quality sequence, as indicated by percent ambiguity, matches the probe; (iv) a multiple alignment of all the sequences representing the target group shows that the mismatch is not representative of the target group; (v) the mismatch occurs in a highly conserved position of the 16S rRNA molecule; (vi) the mismatch is consistent with a common sequencing error, such as the transposition of two bases or the repetition of the same base, that disagrees with several other related sequences.

Oligonucleotide probe design. The oligonucleotide probes developed and/or optimized in this study are listed in Table 1. The matching position in probe designation represents the forward position of the homologous base in the E. coli 16S rRNA gene. By use of the SEQLAB sequence editor in the Wisconsin Package (Genetics Computer Group, Madison, Wis.), 16S rRNA sequences (>1,300 bp) were aligned initially using the PILEUP function within the editor and then adjusted manually with secondary-structure considerations as described previously (2). With the help of computer-generated consensus sequences, the alignments were scanned visually for signature sequences of 18 to 30 nucleotides that distinguished methanotrophs at the family or genus level. Candidate oligonucleotide sequences were then examined for specificity using the basic BLAST search and Probe Match functions of GenBank and the RDP-II, respectively (5, 42). Except as described below, only sequences exhibiting high specificity for methanotrophs and retrieving a majority of the sequences in their target groups were pursued further.

T <sub>D</sub> determination and specificity testing. Oligodeoxynucleotides were synthesized commercially (DNAgency, Malvern, Pa.). Each probe was characterized by empirical determination of its midpoint dissociation temperature (T <sub>D</sub>) using a serial washing procedure with progressively higher temperatures in a PCR thermal cycler as described by Gulledge and Cavanaugh (32). All T <sub>D</sub> curves were determined using triplicate blots for both positive and negative controls (see Fig. 1). The ability of each probe to distinguish between positive and negative controls was screened in Northern dot blot hybridization assays, as described below, using total RNA from reference cultures representing target strains as positive controls and total RNA from reference cultures representing nontarget strains with 1- or 2-base mismatches as negative controls. In all but two cases, a strain with a single-base mismatch with the probe was used as a negative control (Table 1).

Because no nontarget organisms had fewer than two mismatches with probes Am445 were identified, an organism with two mismatches was used as a negative control. Also, because no potential control organisms with fewer than four mismatches to probe Mcd77 were identified, we designed a probe with a single mismatch at position 15 to serve as a negative control (Table 1).

RNA extraction from bacterial cultures. Pure cultures were grown to late-log phase in 40 ml of liquid growth medium and centrifuged at 5,000 × g for 10 min at 4°C. Total RNA was extracted selectively from cell pellets using the FastPrep bead beater system with the FastRNA Blue kit according to the manufacturer’s protocol (Bio 101, La Jolla, Calif.). Cells were beaten in the FP120 bead beater with 25 to 40 s at a speed of 6 m/s. After extraction and centrifugation, the RNA pellets were air dried, resuspended in diethyl pyrocarbonate-treated H2O, and stored at −80°C.

DNA dot blotting and hybridization. Northern dot blots were prepared from RNA extracts as described previously (48) using a Minifold I Microsample Filtration Manifold (Schleicher & Schuell, Keene, N.H.). Blots were prepared with 100 ng of 16S rRNA per dot to be blotted, assuming that 16S rRNA represented 27% of total RNA (47), as described previously (49).

Oligonucleotide probes were labeled enzymatically with 32P (49), and hybridization assays were carried out as described previously (48). Labeled oligonucleotides were hybridized to the dot blots overnight at 30°C, finishing with two 30-min rinses at the appropriate T <sub>D</sub> for each probe (Table 1). Oligonucleotide labeling of the dot blots was analyzed by radiodensitometry using a BAS-MS 2025 imaging plate and a Fujix 2000 PhosphorImager, with MacBAS, version 2.5, image analysis software (Fuji Medical Systems, Stamford, Conn.).

Nucleotide sequence accession numbers. The new sequences of the 16S rRNA genes of Methylohalobacterunix NCIMB 11913, Methylophilusluter NCIMB 11914, Methylostomax methanica SI NCIMB 11130, and Methylostomax marinus strain A45 have been deposited in GenBank (accession numbers AF304194 to AF304197).

RESULTS AND DISCUSSION

Overview. In recent years, interest in the physiology, ecology, and evolution of methanotrophs has intensified, and there is high demand for tools to facilitate quantitative studies of in situ methanotroph community structure (21, 34, 46, 50). Our objectives were to develop phylogenetic oligonucleotide probes for analysis of methanotrophs at the family and genus levels and to optimize the probes for use in quantitative hybridization through empirical determination of their T <sub>D</sub> under standard hybridization conditions.

Visual comparison of aligned 16S rRNA reference sequences initially revealed 36 potential probe sequences for fur-
Table 1. Oligonucleotide probes targeting methanotrophic bacteria

<table>
<thead>
<tr>
<th>Probe name</th>
<th>Target group(s)</th>
<th>Probe sequence (5'→3')</th>
<th>Abbreviation</th>
<th>Positive control(s)</th>
<th>Negative control(s) (no. of mismatches)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Am445 S-F-</td>
<td>Methylosinus</td>
<td>TCCGGTACCGTCATTATCGTCCC</td>
<td>Mtr-0445-a-A</td>
<td>CB15A (2)</td>
<td>57</td>
</tr>
<tr>
<td>Am976 S-F-</td>
<td>Methylophilus</td>
<td>AGTT</td>
<td>Mtr-0976-a-A</td>
<td>CB15A (1)</td>
<td>51</td>
</tr>
<tr>
<td>Gm633 S-G-</td>
<td>Methylosinus</td>
<td>AGGTACCTGACGCTGTAATTGTG</td>
<td>Mtr-0667-a-A</td>
<td>Methylomonas rubra</td>
<td>50</td>
</tr>
<tr>
<td>Gm705 S-F-</td>
<td>Methylophilus</td>
<td>TCTGGTAAGGTTC</td>
<td>Mtr-0633-a-A</td>
<td>Methylomonas rubra</td>
<td>50</td>
</tr>
<tr>
<td>Mmb1007 S-G-</td>
<td>Methylophilus</td>
<td>ATCTCTCACAG</td>
<td>Mmb-0482-a-A</td>
<td>Methylomonas rubra</td>
<td>51</td>
</tr>
<tr>
<td>Mlm732 S-G-</td>
<td>Methylophilus</td>
<td>GTTT</td>
<td>Mlm-0732-a-A</td>
<td>Methylomonas rubra</td>
<td>57</td>
</tr>
<tr>
<td>Mlc123 S-G-</td>
<td>Methylococcus</td>
<td>ACAAGGCAGATTCCTACG</td>
<td>Mlc-1436-a-A</td>
<td>Methylomonas rubra</td>
<td>50</td>
</tr>
<tr>
<td>Mmc1436 S-G-</td>
<td>Methylococcus</td>
<td>A</td>
<td>Mmc-1436-a-A</td>
<td>Methylomonas rubra</td>
<td>50</td>
</tr>
<tr>
<td>Mcd-0077-a-A</td>
<td>Methylocaldum</td>
<td>GCCACCCACCGGTT</td>
<td>Mcd-0077-a-A</td>
<td>Methylomonas rubra</td>
<td>56</td>
</tr>
</tbody>
</table>

**Probes designed for discrimination of specific genera.**

**Notes:**
- Boldfaced letters in probe sequences correspond to positions of mismatch in the corresponding negative controls.
- Probe numbers denote forward homologous position in *E. coli* probe database (3).
- See Fig. 2 and 3 for detailed coverage information.
- Previously published as PCR primer Mb1007 (44).
- Because no negative-control organism with one or two mismatches was available, probe Mcd77 was modified by 1 base (A to T at position 15) and hybridized to VKM-14L under the conditions prescribed for Mcd77. To provide a conservative control, the mismatch disrupted an AT pair, which bonds more weakly than a GC pair.

**Probe coverage for α-methanotrophs.** Two family-level probes, Am445 and Am976, perfectly match the 16S rRNA sequences of nearly all known α-methanotrophs (Fig. 2), including some novel strains recently isolated from landfill soils (59) and lake sediments (22). *Methylocella palustris* strain K7, a novel acidoophilic methanotroph isolated recently from a northern peat bog and the only cultured representative of its genus (24), was the only α-methanotroph whose 16S rRNA sequence was not covered by either probe. Because these probes do not distinguish between the *Methylosinus* and *Methylocystis* genera, they can detect α-methanotrophs only as a group. No oligonucleotide signatures that distinguish between these two genera were identified.

**Probe coverage for γ-methanotrophs.** For γ-methanotrophs we identified both family- and genus-level probes. Together,
two family-level probes (Gm633 and Gm705) covered 82% of the available sequence. The two family-level probes (Gm633 and Gm705) covered 82% of the 16S rRNA molecule. If the two apparent insertions are disregarded, the sequence matches Mlb482 perfectly. Probes Mmb482 and Mmb1007 each matched all available Methylocrobium sequences. Mmb1007 also covered both strains of the newly described genus Methylosarcina, which are closely related to Methylocaldum spp. (58). Three other probes covered all of the recognized Methylomonas isolates. Representatives of this genus fell into two groups that differ by an A versus a C at position 746 (E. coli numbering). We designed two probes (Mlm732a and Mlm732b) to distinguish between the two subgroups. Mlm506 covered the broadest coverage of Methylomonas spp., but all representatives of the genus were covered only when the three Mlm probes were combined.

Three probes covered all representatives of the two recognized thermophilic genera, Methylococcus and Methylocaldum. Mlc123 and Mlc1436 each matched all Methylococcus sequences available. PCR primers corresponding to these two probes might be ideal for specific amplification of nearly complete (~1,300-bp) 16S rRNA genes from Methylococcus strains in environmental samples. Probe Mcd77 covered the three recognized strains of the recently described genus Methylocaldum. The target region was unique, and a Probe Match analysis retrieved no sequences with fewer than four mismatches from non-Methylocaldum species.

The complete suite of \( \gamma \)-methanotroph probes covered 97% of the strains listed in Fig. 3; only two sequences were not covered. One is that of Methylomonas \( \gamma \)-methanica strain 81Z, cultures of which are no longer extant and whose affiliation with the genus Methylomonas was never verified (J. P. Bowman and P. N. Green, personal communication). Because this sequence is of low overall quality (3.3% ambiguity), one or more of the indicated mismatches could be incorrect. The other organism not covered by the probes is a novel thermophilic methanotroph, \( \text{Methylothermus} \) sp. strain HB. Because it is the only known \( \gamma \)-methanotroph that is polyphyletic with respect to the family Methylococaceae (8), this result was expected.

Probes are intended to quantify 16S rRNA from specific microbial populations against a background of many unknown populations in environmental samples. The probes must discriminate against unknown, nontarget 16S rRNA that may have a difference of only 1 base from the intended target. The primary factor for achieving stringent specificity and quantitative hybridization of 16S rRNA from environmental samples is accurate determination of the melting characteristics of the probe-target duplex. Hence, empirical determination of the \( T_m \) is essential (32, 54). We have optimized the probes presented here for stringent discrimination against nontarget RNA and also for quantitative hybridization by empirically determining the \( T_m \) for each probe.

The \( T_m \)s of the probes ranged from 49 to 57°C (Table 1). When Northern blots were hybridized overnight and then washed at the appropriate \( T_m \), target and nontarget rRNAs were visually distinguishable on blots and yielded quantitatively distinct results when analyzed using a scintillation counter (Fig. 1) or a phosphorimager (data not shown). These results verify that the use of known concentrations of reference rRNA as standards will permit quantitative analysis of environmental rRNA possessing the target sequence, as demonstrated previously (49, 54).
Probes Am445, Mmb1007, Mlm482, Mlm732b, Mlc123, and Mcd77 each exhibited at least two base mismatches against any nonmethanotroph sequence, whereas probes Gm633, Gm705, Mlb482, Mmb1007, and Mlc1436 each exhibited at least one base mismatch with any nonmethanotroph sequence. Probes Mlb662, Mmb482, and Mmb1007 matched sequences from one to four \(\delta\text{-}H\text{H}253\)-methanotrophs outside their respective target genera (Fig. 3). Although we consider this problem to be minor, these probes could yield ambiguous results for fine-scale descriptions of \(\delta\text{-}H\text{H}253\)-methanotroph communities. All other genus-level probes were specific to their intended target genera. The \(\alpha\)- and \(\gamma\)-methanotroph probes had no cross-family hybridization potential.

Two probes, Am976 and Mlb662, present the more serious problem of complementing 16S rRNA sequences of some non-methanotrophic bacteria. They have been retained despite this weakness for two reasons. First, they are needed to ensure complete coverage of their target groups, in combination with other probes, when broad-spectrum probing is desired. Second, they were deemed particularly useful for certain experi-

FIG. 3. Range of strain coverage for oligonucleotide probes targeting \(\gamma\)-methanotrophs (Gm). Mic, Methylococcus; Mcd, Methylocaldum. % Ambiguity, percentage of positions within the entire sequence that indicate ambiguous bases, shown as an index of overall sequence quality. The unpublished 16S rRNA sequence for \textit{Methylomonas methanica} strain S1 (marked with a star in the “GenBank accession number” column) is available as RDP sequence Mlm.metha1 (C. R. Woese, 1991). Under “Probes,” solid fill indicates identity between the probe and a target sequence; diagonal hatching indicates identity between the probe and a nontarget \(\delta\text{-}H\text{H}253\)-methanotroph strain; cross-hatching indicates identity between the probe and a nonmethanotroph; numbers are numbers of mismatches between the probe sequence and the corresponding 16S rRNA sequence. Where a number is shown in white on a solid background, the apparent mismatches were disregarded based on criteria outlined in Materials and Methods. An open diamond denotes the occurrence of one or more ambiguous bases in the probe target region that are consistent with the probe sequence. For example, if the probe has an A corresponding to a Y (IUPAC ambiguity code for C or T), then the possible T is consistent with the probe sequence. ns, no sequence available in probe target region.
Cultures, use as PCR primers in cases where amplification approaches, such as monitoring of CH₄ enrichment, are to be sequenced for identification, or analysis of community composition in environmental samples where the non-target organisms with which the probes hybridize should be minor components of the community. For instance, because marine *Cycloclasticus* spp. were the only nonmethanotrophs that matched Mlb662 (Fig. 3), this probe might be appropriate for probing nonmarine samples.

**Probing the database.** The GenBank database contains thousands of bacterial 16S rRNA gene sequences from cultures and environmental clones (6). Hence, “probing” this database should provide a powerful assessment of a probe’s ability to select specifically for methanotroph sequences against a background of myriad nonmethanotroph sequences. We subjected each probe sequence to a basic BLAST search (5) and examined sequences retrieved with an identical match. Only sequences identified as 16S rRNA genes were considered. The organism identifications were based solely on information provided in the accession records or in publications cited therein.

Eleven of 14 probes retrieved only sequences that were identified as methanotrophs (Table 2). Probes Am976, Mlb662, and Mlc1436 retrieved a number of sequences representing a narrow range of nonmethanotrophic taxa. The first two of these probes matched environmentally restricted taxa, such as obligate pathogens (*Afipia* spp.) and obligate marine bacteria (*Cycloclasticus* spp.). If used strategically, therefore, these probes are likely to be useful for studying methanotroph communities. From the data in Table 2, it would be premature to conclude that Mlc1436 is nonspecific. All but one of the non-methanotroph sequences retrieved by this probe were nearly identical clones of putative β-Proteobacteria from an activated sludge reactor. However, no cultured organisms belonging to the β-Proteobacteria were retrieved, and no published data were cited in the accession records to confirm the phylogenetic position of these environmental clones. Overall, the data in Table 2 suggest that at least 11 and possibly 12 of the probes presented here are highly specific to methanotrophic bacteria and that the two clearly nonspecific probes should hybridize to a phylogenetically limited range of nonmethanotrophs with restricted environmental distributions.

**Summary and conclusions.** The breadth and specificity of the probes reported here are unprecedented, providing 97% coverage of the 87 methanotroph 16S rRNA sequences examined (Fig. 2 and 3). Several new methanotroph genera that have been proposed recently following the isolation of novel strains are covered. Of the three strains apparently not covered by the probes, one is no longer extant and the available 16S rRNA sequence is of low overall quality, bringing into question whether the indicated probe mismatches are correct. The other two strains (*Methylocella palustris* sp. strain K and *Methylothermus* sp. strain HB) that did not match any probe are polyphyletic with respect to the *Methylcystaceae* and *Methylcoccaceae*, thus reflecting the high specificity of the probes to the phylogenetic clades they were designed to target. Initial results from studies with several soils indicate that the probes are effective for studying methanotroph communities in soil (unpublished data), perhaps the most difficult substrate on which to perform quantitative hybridization assays (4). Hence, all of the methanotroph taxa that have become well known through years of laboratory studies, as well as several recently described taxa, can now be studied at both the family and genus levels in environmental samples by using the probes reported here.

**ACKNOWLEDGMENTS**

J. Gulledge and A. Ahmad contributed equally to this work. We gratefully acknowledge the following individuals: M. Polz for training and helpful discussions on designing oligonucleotide probes; A. J. Auaman, A. M. Costello, and M. E. Lidstrom for updated 16S rRNA sequences and a protocol for extracting nucleic acids from training and helpful discussions on designing oligonucleotide probes; A. J. Auaman, A. M. Costello, and M. E. Lidstrom for updated 16S rRNA sequences and a protocol for extracting nucleic acids from bacteria; G. M. King, R. Knowles, J. C. Murrell, J. S. Poin dexter, and J. D. Semrau for providing reference cultures; and A. A. DiSpirito for reference genomic DNA.

This work was supported by the U.S. National Science Foundation (award DEB9708092) and was initiated while J. Gulledge was a DOE-Energy Biosciences Research Fellow of the Life Sciences Research Foundation.

**REFERENCES**


