NifH and NifD phylogenies: an evolutionary basis for understanding nitrogen fixation capabilities of methanotrophic bacteria

Svetlana N. Dedysh,1 Peter Ricke2 and Werner Liesack2

1Institute of Microbiology, Russian Academy of Sciences, Moscow 117312, Russia
2Max-Planck-Institut für terrestrische Mikrobiologie, Karl-von-Frisch-Str., D-35043 Marburg, Germany

The ability to utilize dinitrogen as a nitrogen source is an important phenotypic trait in most currently known methanotrophic bacteria (MB). This trait is especially important for acidophilic MB, which inhabit acidic oligotrophic environments, highly depleted in available nitrogen compounds. Phylogenetically, acidophilic MB are most closely related to heterotrophic dinitrogen-fixing bacteria of the genus *Beijerinckia*. To further explore the phylogenetic linkage between these metabolically different organisms, the sequences of *nifH* and *nifD* gene fragments from acidophilic MB of the genera *Methylocella* and *Methylocapsa*, and from representatives of *Beijerinckia*, were determined. For reference, *nifH* and *nifD* sequences were also obtained from some type II MB of the alphaproteobacterial *Methylosinus/Methylocystis* group and from gammaproteobacterial type I MB. The trees constructed for the inferred amino acid sequences of *nifH* and *nifD* were highly congruent. The phylogenetic relationships among MB in the NifH and NifD trees also agreed well with the corresponding 16S rRNA-based phylogeny, except for two distinctive features. First, different methods used for phylogenetic analysis grouped the NifH and NifD sequences of strains of the gammaproteobacterial MB *Methylococcus capsulatus* within a clade mainly characterized by Alphaproteobacteria, including acidophilic MB and type II MB of the *Methylosinus/Methylocystis* group. From this and other genomic data from *Methylococcus capsulatus* Bath, it is proposed that an ancient event of lateral gene transfer was responsible for this aberrant branching. Second, the identity values of NifH and NifD sequences between *Methylocapsa acidiphila* B2 and representatives of *Beijerinckia* were clearly higher (98-5 and 96-6%, respectively) than would be expected from their 16S rRNA-based relationships. Possibly, these two bacteria originated from a common acidophilic dinitrogen-fixing ancestor, and were subject to similar evolutionary pressure with regard to nitrogen acquisition. This interpretation is corroborated by the observation that, in contrast to most other diazotrophs, *M. acidiphila* B2 and *Beijerinckia* spp. are capable of active growth on nitrogen-free media under fully aerobic conditions.

INTRODUCTION

Currently recognized acidophilic methanotrophic bacteria (MB) are represented by two genera, *Methylocella* and *Methylocapsa*, which belong to the Alphaproteobacteria (Dedysh et al., 2000, 2002). Similar to the traditionally known alphaproteobacterial MB, *Methylosinus* and *Methylocystis*, acidophilic MB utilize the serine pathway for carbon assimilation. However, they possess several unique morphological and physiological characteristics, and, based on 16S rRNA phylogeny, *Methylocella* and *Methylocapsa* are evolutionarily more closely related to acidophilic heterotrophic bacteria of the genus *Beijerinckia* than to the *Methylosinus/Methylocystis* group.

Acidophilic MB inhabit diverse acidic environments, including wetlands and upland soils of the boreal zone and tundra (Dedysh et al., 2001, 2003, 2004; Dunfield et al., 2003). These environments, especially *Sphagnum* peat bogs, are characterized by extremely low concentrations of available nitrogen compounds (Richardson et al., 1978; Mitsch & Gosselink, 1986). Thus, the ability to utilize N₂ as a nitrogen source is a necessary phenotypic trait for...
microbial inhabitants of these ecosystems, including acidophilic MB. Both Methylocella and Methylocapsa are capable of dinitrogen fixation. However, members of these two genera behave in a different way when grown in liquid nitrogen-free media. Under high partial pressures of oxygen, representatives of Methylocella usually show quite poor growth in nitrogen-free media. This is also typical for other dinitrogen-fixing MB such as Methylococcus (Murrell & Dalton, 1983; Dedysh et al., 2000; Auman et al., 2001). Decreasing dissolved oxygen tension can enhance both growth and acetylene reduction activity of these MB. In contrast, Methylocapsa is capable of exponential growth in liquid nitrogen-free media under both aerobic and microaerobic conditions (Dedysh et al., 2002). This growth capability of Methylocapsa is very similar to that of Beijerinckia, which was one of the first bacteria described as being capable of fixing dinitrogen effectively (Alston, 1936; Starkey & De, 1939; Becking, 1999).

This study aimed to compare the molecular genetic basis for the nitrogen fixation capabilities of acidophilic MB with those of phylogenetically related heterotrophic nitrogen fixers. Our investigation was based on the comparative sequence analysis of the nitrogenase and acetyl-CoA reductase genes, which encode dinitrogen reductase and the α-subunit of dinitrogenase: the key components of the nitrogenase enzyme complex. Public-domain databases currently contain about 2000 nifH and 300 nifD sequences. However, nifH sequences for MB have become available only recently, after the corresponding PCR protocols had been established (Auman et al., 2001; Boulygina et al., 2002), and no studies have been done on nifD genes in these bacteria. Thus, our study offers the first comparison of MB 16S rRNA phylogeny with phylogenies based on two different structural genes of nitrogenase. Moreover, the results obtained for growth experiments of representative MB and Beijerinckia spp. on nitrogen-free media agreed well with the principal conclusions drawn from the phylogenetic trees constructed for NifH and NifD.

### METHODS

#### Bacterial strains and growth conditions. Bacterial strains used in this study are listed in Table 1. The set of acidophilic MB included four type strains (Methylocapsa acidiphila B2, Methylocella palustris K, Methylocella silvestris BL2 and Methylocella tundrae T4), three strains of Methylocella palustris isolated either from the Sphagnum peatlands of north-eastern Germany (strain H4) or from the tundra wetlands of Russia (strains Ch3 and Y5), and two strains of Methylocella tundrae (Y1 and Ch1) isolated from two different tundra wetland sites in northern Russia and described by Dedysh.

### Table 1. Bacterial strains used in this study

<table>
<thead>
<tr>
<th>Group</th>
<th>Strain</th>
<th>Accession no. in culture collection or source of isolation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Acidophilic MB</td>
<td>Methylocapsa acidiphila B2</td>
<td>DSM 13967(^T), <em>Sphagnum</em> peat bog, western Siberia, Russia</td>
</tr>
<tr>
<td></td>
<td>Methylocella palustris K</td>
<td>ATCC 700799(^T), <em>Sphagnum</em> peat bog, western Siberia, Russia</td>
</tr>
<tr>
<td></td>
<td>Methylocella palustris Ch3</td>
<td>Tundra wetland, Chukotka, northern Russia</td>
</tr>
<tr>
<td></td>
<td>Methylocella palustris Y5</td>
<td>Tundra wetland, Yugorsk, northern Russia</td>
</tr>
<tr>
<td></td>
<td>Methylocella silvestris BL2</td>
<td>Peat bog lake Kleine Fuchskuhle, northern Germany</td>
</tr>
<tr>
<td></td>
<td>Methylocella palustris H4</td>
<td>DSM 15510(^T), acidic forest soil, Marburg, Germany</td>
</tr>
<tr>
<td></td>
<td>Methylocella silvestris BL2</td>
<td>DSM 15673(^T), tundra wetland, Yorkska, northern Russia</td>
</tr>
<tr>
<td></td>
<td>Methylocella tundrae T4</td>
<td>Tundra wetland, Yugorsk, northern Russia</td>
</tr>
<tr>
<td></td>
<td>Methylocella tundrae Ch1</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Methylococcus capsulatus Texas</td>
<td>NCIMB 11853(^T)</td>
</tr>
<tr>
<td></td>
<td>Methylobacter luteus</td>
<td>NCIMB 11914(^T)</td>
</tr>
<tr>
<td></td>
<td>Methylocalicium album</td>
<td>NCIMB 11123(^T)</td>
</tr>
<tr>
<td>Type I MB</td>
<td>Methylosinus trichosporum OB3b</td>
<td>ATCC 35070(^T)</td>
</tr>
<tr>
<td></td>
<td>Methylosinus trichosporum SM6</td>
<td>River Saale, Wichmar, Germany*</td>
</tr>
<tr>
<td></td>
<td>Methylosinus sporium SM27a</td>
<td>River Saale, Wichmar, Germany*</td>
</tr>
<tr>
<td></td>
<td>Methylosinus sporium 44/2</td>
<td>Spring bog, Rügen, Germany*</td>
</tr>
<tr>
<td></td>
<td>Methylocystis echnoides</td>
<td>IMET 10491(^T)</td>
</tr>
<tr>
<td></td>
<td>Methylocystis sp. H9a</td>
<td>Lake Haussee, Feldberg, Germany*</td>
</tr>
<tr>
<td>Type II MB</td>
<td>Beijerinckia indica subsp. indica</td>
<td>ATCC 9039(^T)</td>
</tr>
<tr>
<td></td>
<td>Beijerinckia indica subsp. lacticogenes</td>
<td>DSM 1719(^T)</td>
</tr>
<tr>
<td></td>
<td>Beijerinckia mobilis</td>
<td>DSM 2326(^T)</td>
</tr>
<tr>
<td></td>
<td>Beijerinckia dextrii subsp. dextrii</td>
<td>DSM 2328(^T)</td>
</tr>
<tr>
<td></td>
<td>Beijerinckia dextrii subsp. venezuelae</td>
<td>DSM 2329(^T)</td>
</tr>
<tr>
<td></td>
<td>Rhodoblastus acidophilus</td>
<td>ATCC 25092(^T)</td>
</tr>
<tr>
<td></td>
<td>Azorhizobium cauliformans</td>
<td>DSM 5975(^T)</td>
</tr>
<tr>
<td></td>
<td>Bradyrhizobium japonicum</td>
<td>DSM 30131(^T)</td>
</tr>
</tbody>
</table>

*Strains described by Heyer *et al.* (2002).
et al. (2004). The set of reference organisms included nine strains of type I and type II MB, and eight type strains of dinitrogen-fixing heterotrophic bacteria.

*Methylocapsa acidiphila* B2 was grown on half-strength (1:2) liquid nitrogen-free M1 mineral medium (Dedysh et al., 2002). The same medium, in agar form, supplemented with 250 mg KNO₃ l⁻¹, was used for cultivation of all strains of *Methylocella*. Other MB were cultivated on basic mineral salts (MS) medium (Whittenbury et al., 1970). This medium was either supplemented with 1 g KNO₃ l⁻¹, to give nitrate mineral salts (NMS) medium, or was used without addition of nitrate to give nitrate-free mineral salts (NFMS) medium. All MB were cultivated under a gas headspace containing 20 % (v/v) methane. The liquid cultures were shaken at 120 r.p.m. and incubated at 24 °C, with the exception of *Methylococcus capsulatus* Texas, which was incubated at 37 °C. *Beijerinckia* strains were cultivated on half-strength liquid nitrogen-free M1 mineral medium supplemented with 0-2 % (w/v) glucose. *Azorhizobium caulindans, Bradyrhizobium japonicum* and *Rhodoblastus acidophilus* were grown on media recommended by the DSMZ (Deutsche Sammlung von Mikroorganismen und Zellkulturen GmbH) and ATCC, respectively.

**Growth experiments.** Cultures of *Methylocapsa acidiphila* B2, type I MB *Methylobacter luteus*, type II MBs *Methylocystis echnoides* IMET 10491 and *Methylosinus trichosporium* OB3b, and the heterotrophic bacterium *Beijerinckia indica* subs. lactiogenes were grown in 120 ml serum flasks containing 30 ml of the appropriate nitrogen-free medium. Half-strength M1 medium was used in the case of *Methylocapsa* and *Beijerinckia*, and NFMS medium in the case of *Methylobacter, Methylocystis* and *Methylosinus*. Growth was monitored by measuring OD₆₀₀. Serum flasks containing 20 ml NFMS medium, with a range of PO₂ values from 0-005 to 0-2 bar, were inoculated with the MB cells obtained by cultivation in nitrogen-free media. The required PO₂ in the headspace of the flasks was obtained by purging flasks with N₂ after inoculation and then injecting O₂ via a syringe. Methane (10 %, v/v) was added to all flasks, which were then incubated at 24 °C on a shaker at 120 r.p.m.. At intervals of 2-4 h, the flasks were examined for growth by measuring OD₆₀₀ and the concentration of oxygen in the headspace. The latter was determined using a Shimadzu 8A gas chromatograph fitted with a thermal conductivity detector and a Molecular Sieve 5Å column. The specific growth rate calculated for each incubation period was followed by 35 cycles consisting of denaturation for 30 s at 94 °C, followed by 35 cycles consisting of denaturation at 92 °C for 30 s, primer annealing at 55 °C for 30 s, and elongation at 72 °C for 30 s. The final elongation step was extended to 5 min.

Alternatively, the forward primer nifD-f (5'-GGGGTGCGCCTA-YGCCCG-3') could be used in combination with nifD-r to amplify nifD gene fragments, yielding a 1130 bp fragment that started at position 227 of the *Bradyrhizobium japonicum* USDA 110 nifD gene (GenBank X01045). This assay represented a confirmatory test for the presence of nifD and covered the possibility that nifH and nifD did not cluster contiguously within the same operon.

All PCR reaction mixtures contained 0-5 μl template DNA, 5 μl 10x reaction buffer, 1-5 mM MgCl₂, 200 μM of each dNTP, 0-3 mM of each primer, and 1-25 U Taq DNA polymerase (Promega). Amplification was carried out in a total volume of 50 μl in 0-2 ml reaction tubes. Based on a previous report on the occurrence of nifH genes among MB (Auman et al., 2001), *Methylosinus trichosporium* OB3b and *Methylomicrobium album* were chosen as positive and negative controls, respectively. The 16S RNA, nifH and nifD gene amplicons were purified using QIAquick spin columns (Qiagen) and sequenced on an ABI Prism 377 DNA sequencer using BigDye terminator chemistry, as specified by the manufacturer (PE Applied Biosystems).

**Data analysis.** Based on sequence information deposited in public-domain databases or generated in the course of this study, we established databases for nifH and nifD sequences using the ARB program package (http://www.arb-home.de). The databases were checked manually for correct alignment. The alignments were used for phylogenetic analyses without making changes to possible errors in the public-domain nifH and nifD sequences or the inferred amino acid sequences. The newly obtained 16S rRNA gene sequences were added to a database of about 24000 nearly complete 16S rRNA sequences. Phylogenetic analyses were carried out using PHYLIP (Felsenstein, 1989) and Tree Puzzle (Strimmer & von Haeseler, 1996). Multiple trees were constructed for both 16S RNA genes and amino acid sequences inferred from nifH and nifD, using distance-based (neighbour-joining, NJ) and maximum-likelihood methods, including DNAML (16S rRNA), ProtML (NifH, NifD), and Tree Puzzle (16S rRNA, NifH, NifD), in combination with various models of evolution, such as JC (16S rRNA), PAM, JTT and WAG (NifH, NifD). The final tree topologies shown for NifH and NifD are consensus trees that take into consideration only those branch points which were confirmed by each of the treeing methods and models of evolution applied. In addition, the significance levels of interior branch points obtained in NJ analysis were determined by bootstrap analyses (1000 data resamplings). The overall identity values between pairs of 16S rRNA gene sequences and of inferred amino acid (NifH, NifD) sequences were calculated using the appropriate tools of ARB.

Codon usage tables for *Methylococcus capsulatus* Bath, as well as alphaproteobacterial reference organisms such as *Bradyrhizobium japonicum*, were obtained from the Kazusa codon usage database (http://www.kazusa.or.jp/codon/) (Nakamura et al., 2000). Alternatively, a table for *Methylococcus capsulatus* Bath was computed based on 57 coding sequences (>21 000 codons) obtained from GenBank. Deviations of *Methylococcus capsulatus* nifH/nifD codon usage from the mean species-specific values deposited in the codon usage database were computed using the graphical codon usage analyser (http://gcua. schoedl.de/) and compared to those of genes encoding particulate methane monoxygenase (pmoCA). Using the same on-line tool, the
codon usage patterns of various alphaproteobacterial organisms were compared with those computed for *Methylococcus capsulatus*.

**RESULTS**

**Phylogeny of acidophilic MB based on the 16S rRNA gene**

Comparative sequence analysis of the 16S rRNA gene showed that acidophilic MB of the genera *Methylocella* and *Methylocapsa*, along with acidophilic heterotrophic bacteria of the genus *Beijerinckia*, form a monophyletic cluster within the *Alphaproteobacteria* (Fig. 1). Trees constructed with different methods, including NJ and maximum-likelihood (DNAmI, Tree Puzzle), provided support for the coherence of this cluster (data not shown).

Eight strains of *Methylocella* used in this study represented three currently known species of this genus, *Methylocella palustris*, *Methylocella silvestris*, and *Methylocella tundræa*, and formed a defined subgroup within the *Methylocella-Methylocapsa-Beijerinckia* cluster. The strains of *Methylocella palustris* and *Methylocella tundræa* were isolated from acidic *Sphagnum*-dominated peatlands of the boreal zone and the tundra (Dedysh et al., 2000, 2004), while the third species, *Methylocella silvestris*, was obtained from acidic forest soil (Dunfield et al., 2003). Despite the fact that the strains of each species of *Methylocella* were obtained from
Beijerinckia spp. are acidophilic, heterotrophic, dinitrogen-fixing bacteria. The four recognized species, Beijerinckia indica, Beijerinckia mobilis, Beijerinckia derxii, and Beijerinckia fluminensis, were taxonomically described half a century ago. However, the 16S rRNA gene sequence of Beijerinckia indica subsp. indica (Woese et al., 1984) has until now been the only Beijerinckia sequence available in public-domain databases. We have determined the 16S rRNA gene sequences for four other representatives of this genus: Beijerinckia indica subsp. lacticogenes, B. mobilis, B. derxii subsp. derxii and B. derxii subsp. venezuelae. The 16S rRNA gene sequences of these species formed a phylogenetically distinct and tight cluster, and thus confirmed the assignment of these bacteria to the genus Beijerinckia. The sequence identities among the Beijerinckia spp. were in the range from 98.3 to 99.8% (Fig. 1), while the identities between members of the genus Beijerinckia and the acidophilic MB ranged from 96.0 to 97.3%.

**Growth on nitrogen-free media**

To compare the nitrogen-fixation ability of different MB in the absence of combined nitrogen, the growth on nitrogen-free media of the acidophilic MB of the acidophilic MB in the absence of combined nitrogen, the growth on nitrogen-free media of the acidophilic MB ranged from 96 to 97%. For example, although strains K, Y5 and H4 of Methylocella palustris were isolated from western Siberia, the tundra zone of Russia, and northern Germany, respectively, they exhibited identical full-length 16S rRNA gene sequences. The genus Methylocapsa is currently represented by only one species, Methylocapsa acidiphila, and by the single strain B2. This might be due to the fact that these MB do not grow on agar media (Dedysh et al., 2002), and are therefore very difficult to isolate.

**Fig. 2.** Growth dynamics of three MB species (a) and of Beijerinckia indica subsp. lacticogenes (b) in batch cultures on nitrogen-free mineral media at high headspace pO2 (0–18–0.19 bar). ■, Methylocapsa acidiphila B2; ▲, Methylocystis echinoides IMET 10491; ●, Methylobacter luteus. The pO2 in the Methylocystis echinoides and Methylobacter luteus cultures was re-established during the experiment at 0.01 bar (arrows).

IMET 10491 and Methylobacter luteus responded to the reduction in pO2 with immediate growth. However, only a relatively small increase of the OD600 of both cultures was achieved, since O2 was rapidly consumed by growing cells. Therefore, the injection of O2 into the headspace of the flasks (pO2 value of up to 0.01 bar) was repeated several times, and the cultures of the two MB species always showed the same growth response. Using this approach, we were able to obtain batch cultures of Methylobacter luteus and Methylocystis echinoides IMET 10491 of a relatively high OD600, up to 0.35–0.45, on NFMS medium. In addition, we applied the same strategy to establish a batch culture of another type II MB, Methylosinus trichosporium OB3b, on NFMS medium, since this bacterium had already been used as a model organism to study the nitrogen-fixation capability of MB (Murrell & Dalton, 1983; Auman et al., 2001). The cells from these three cultures were used to inoculate a number of flasks with a range of pO2 values from 0.005 to 0.2 bar. This was done to assess the specific growth rate of different MB species on nitrogen-free medium as a function of headspace pO2 (Fig. 3). The type I MB, Methylobacter luteus, was able to develop only at pO2 values below 0.02 bar. In contrast, the two type II MB, Methylosinus trichosporium OB3b and Methylocystis echinoides IMET 10491, grew well under a wide range of headspace pO2 values, up to 0.15–0.17 bar. However, growth of the type II MB at relatively high oxygen concentrations was possible only after multiple culture transfers.
in NFMS medium. The highest specific growth rates on nitrogen-free medium, for both type I and type II MB, were recorded at low pO₂ values, below 0.02 bar.

**PCR amplification of nifH and nifD**

Two different protocols for nifH gene amplification from DNA of MB have been described. The protocol of Auman et al. (2001) was based on degenerate nifH primers (nifH-f and nifH-r), designed by Zehr & McReynolds (1989). PCR with these primers yielded a nifH gene fragment of approximately 360 bp. Boulygina et al. (2002) designed another pair of nifH-targeted primers (F1 and R6), which enabled the retrieval of an approximately 450 bp nifH gene fragment. In this study, we used a combination of the forward primer F1 and the reverse primer nifH-r, which yielded a 453 bp PCR product. Amplification of nifH using this primer set showed consistent results for all MB and other related bacteria tested.

This work represents the first attempt to amplify nifD from the DNA of MB. Two pairs of newly designed primers enabled the retrieval of either an approximately 1130 bp fragment or an approximately 1900 bp fragment. The latter fragment, which consisted of the 3’ region of nifH plus the almost complete nifD gene, was obtained from all the strains tested (Fig. 4b). Successful retrieval of these amplicons was possible only if nifH and nifD clustered contiguously within the same operon, suggesting that the structural organization of the nif genes in these strains is similar to that reported in most dinitrogen-fixing Alphaproteobacteria and Gammaproteobacteria (Zehr et al., 2003). For comparison, previously reported assays for nifD gene amplification yielded PCR products of only 450 to 830 bp (Ueda et al., 1995b; Minerdi et al., 2001; Parker et al., 2002).

Until now, the database of nifH sequences from MB was limited to 26 sequences, of which only five sequences were derived from the type strains of recognized species. This study contributes another 25 nifH and 15 nifD sequences. This includes 17 nifH and 11 nifD sequences from MB, of which 16 were obtained from the type strains of MB species.

**NifH- and NifD-based phylogenies of MB and related heterotrophic bacteria**

**General aspects.** The topologies of the NifH- and NifD-based trees constructed in this study for MB and some other representatives of the Alphaproteobacteria and Gammaproteobacteria were very similar (Fig. 4). It should be noted that only 137 deduced amino acid residues were used for construction of the NifH tree, while the NifD tree is based on 356 deduced amino acid residues. Thus, the high congruity between the two trees suggests that the relatively short NifH sequence stretch provides sufficient phylogenetic information to reconstruct meaningful Nif-based trees for diazotrophic bacteria. In general, the clusters formed by the NifH and NifD sequences corresponded well with rRNA phylogeny. However, a few of the NifH and NifD sequences previously obtained from non-MB showed examples of an aberrant grouping. These were representatives of the Betaproteobacteria, namely Herbaspirillum seropedicae., and two strains of Burkholderia, free-living Burkholderia fungorum and symbiotic 'Candidatus Glomeribacter gigas'. The NifH and NifD sequences of both species were affiliated with alphaproteobacterial-like Nif sequence types (compare Fig. 4a with 4b).
In accordance with the 16S rRNA gene-based phylogeny, the NifH and NifD sequences from acidophilic MB and from the Methylosinus/Methylocystis group were affiliated with NifH and NifD sequences from other dinitrogen-fixing Alphaproteobacteria. In both NifH and NifD trees, these organisms formed distinct clusters composed of sequences from (i) Methylocella, (ii) Beijerinckia/Methylocapsa acidiphila, and (iii) Methylosinus/Methylocystis.

The NifH sequences from all eight representatives of the genus Methylocella grouped together (Fig. 4a), and the identity values between them varied between 98-5 and 100%. The NifH fragments obtained from strains of the two peat-inhabiting species, Methylocella palustris and Methylocella tundracea, displayed complete sequence identity, while the NifH fragment from the upland soil-inhabiting species, Methylocella silvestris, was slightly divergent from the other sequence types of this cluster (1-5% sequence divergence). The nifH sequence-divergence among strains of Methylocella was slightly higher on the nucleic acid level (2-7 to 5-7%) than on the inferred amino acid level. The NifD sequences obtained from different species of Methylocella also formed a coherent cluster. The intracluster identity values ranged from 96-2 to 98-0%.

The NifH and NifD sequence identities within representatives of the genus Beijerinckia ranged from 97-7 to 100%,

Fig. 4. Consensus trees of partial NifH sequences constructed based on 137 deduced amino acid sites (a) and partial NifD sequences constructed based on 356 deduced amino acid sites (b). The two trees show acidophilic MB in relation to Beijerinckia, the Methylosinus/Methylocystis group, gammaproteobacterial MB, and other representative dinitrogen-fixing members of the Alphaproteobacteria and Gamma-proteobacteria. The assignment of bacterial taxa to either Alphaproteobacteria (white background) or Gammaproteobacteria (grey-shadowed) follows the 16S rRNA-based phylogeny, except for those marked *, which are Betaproteobacteria. Only branch points supported by both neighbour-joining and maximum-likelihood (ProtML, Tree Puzzle) methods are shown. The numbers at the nodes indicate the percentage recovery in 1000 bootstrap resamplings. Micro-organisms for which sequence data were obtained in this study are indicated in bold. The GenBank accession numbers of the NifH and NifD reference sequences used for construction of the respective trees are given. Scale bars, 0-1 substitution per amino acid position.
and 98·5 to 99·1 %, respectively. Unexpectedly, the NifH fragment from Methylocapsa acidiphila B2 exhibited much higher identity values to the corresponding NifH fragments of Beijerinckia spp. (98·0–98·5 %) than to those of Methylocella spp. (90·8 %). In fact, the NifH fragment from Methylocapsa acidiphila B2 differed from the corresponding NifH fragments of the two subspecies of Beijerinckia indica by only two amino acid residues. A similar observation was made for NifD: the NifD fragment of strain B2 displayed sequence identities of 96·5–96·7 % to NifD fragments of Beijerinckia spp., and only 90·8–92·3 % sequence identities to NifD fragments of Methylocella spp.

The representatives of the Methylosinus/Methylocystis group were more divergent from each other than were the acidophilic bacteria. These organisms were separated in the NifH tree into two distinct lineages, characterized either by Methylosinus spp. or by Methylocystis spp., while their NifD sequences were clustered together into a common lineage. The identity values between NifH sequences of Methylosinus spp. and Methylocystis spp. ranged from 94·6 to 100 %, while the corresponding NifD values varied from 95·0 to 99·7 %. However, the strains selected for this study represent only a small proportion of the currently known strains of Methylosinus and Methylocystis (Heyer et al., 2002), and thus our investigation might underestimate NifH and NifD divergence within this group.

Gammaproteobacterial MB. The NifH sequences from members of the MB genera Methylomonas and Methylobacter were assigned to lineages formed by NifH sequences of other Gammaproteobacteria. The identity values between the NifH sequences of alphaproteobacterial MB and those of Methylomonas and Methylobacter did not exceed 81·5 %. The identity values of NifD sequences from alphaproteobacterial MB to the corresponding fragment of Methylobacter luteus did not exceed 80·2 %, and thus provided further evidence for a clear separation of type I and type II MB in Nif-based phylogenies.

However, a striking disagreement between the tree topologies from, on the one hand, the 16S rRNA gene and, on the other hand, NifH and NifD, was identified for strains of the gammaproteobacterial MB Methylococcus capsulatus. The NifH and NifD sequences were grouped within clusters, characterized by alphaproteobacterial-like NifH and NifD sequence types, respectively. This disagreement was originally reported for NifH sequences from strains 114 and 115 of Methylococcus capsulatus (Boulygina et al., 2002), while the taxonomic affiliation of these two strains was supported by partial 16S rRNA gene sequence data. To verify this aberrant grouping for other strains of Methylococcus capsulatus, we obtained sequence data for both the nearly complete 16S rRNA gene and partial nifH and nifD genes of Methylococcus capsulatus Texas (NCIMB 11853T). In addition, we have exploited the data available for the Methylococcus capsulatus Bath genome (GenBank NC 002977). The NifH sequences of Methylococcus capsulatus Texas and Methylococcus capsulatus Bath were identical to those of strains 114 and 115. The identity values of the four NifH sequences of Methylococcus capsulatus strains to those of the Methylocella and Methylosinus/Methylocystis groups were in the range 94·7 to 96·2 % and 93·2 to 95·4 %, respectively. The corresponding identity values calculated for NifD of Methylococcus capsulatus to those of the Methylocella and Methylosinus/Methylocystis groups ranged from 90·0 to 91·2 % and 90·5 to 91·4 %, respectively, while the identity to NifD of Methylobacter luteus was only 77·8 %.

Consideration of environmental NifH sequences

NifH sequences from acidophilic MB were used to perform a BLAST search against the GenBank database in order to identify closely related environmental clone sequences. A large number of NifH sequences retrieved from forest and agricultural soils of the boreal zone (Bürgmann et al., 2004) displayed high identity values to the NifH sequences obtained in our study. The clone b1-W1 (accession no. AY196397) was retrieved from acidic deciduous-forest soil of pH 4·6, and exhibited 99·1 % sequence identity to NifH from both Beijerinckia indica and Methylocapsa acidiphila. In this case, NifH-based phylogeny does not allow a reliable assignment of environmental nifH sequences into one of these two distinct acidophilic physiotypes. The sequence of the clone c1-HW3 (AY196440), which had been obtained from another acidic forest soil of pH 5·5, exhibited complete sequence identity to NifH from Beijerinckia mobilis and two subspecies of Beijerinckia dextra. A group of environmental NifH sequences, including the clones b1-W16, b1-W12, f1-HW5, f1-PA10, f1-PA2, f1-PA4, f1-PA1, b1-GA6, b1-GA7, f1-HA1, and F1-HA2 (accession nos AY196383, AY196398, AY196454, AY196459, AY196460, AY196462, AY196458, AY196395, AY196396, AY196445, and AY196446, respectively) exhibited 99·1 % sequence identity to NifH sequences of Methylocella palustris and Methylocella tundrae. These clone sequences were retrieved from two acidic (pH 4·6 and 5·5) and one neutral (pH 6·8) forest soil, and from two agricultural soils (pH 7·2 and 7·5). The detection of these NifH sequences in soils of different acidity is not surprising, since Methylocella species are capable of growing in a pH range from 4·2 to 7·5. Two other environmental NifH sequences retrieved from deciduous forest soils of pH 6·8 and 5·5, clones b1-HA3-1 and f1-HW6 (AY196373 and AY196455), were closely related (98·2 % sequence identity) to NifH sequences from representatives of Methylocella and Methylocystis.

A BLAST search against the GenBank database was also performed using NifD sequences obtained in the study. This, however, did not yield any closely matching results, which might be due to the very limited number of environmental NifD clone sequences deposited in public-domain databases.
DISCUSSION

The ability to fix atmospheric dinitrogen is widespread in both Archaea and Bacteria (Postgate, 1987; Young, 1992; Martinez-Romero, 2000). The enzyme responsible for nitrogen fixation, nitrogenase, shows a high degree of conservation of structure and function across wide phylogenetic ranges (Dean & Jacobson, 1992). Nitrogenase is composed of two components, component I (dinitrogenase, or Fe-Mo protein), an $\alpha_3\beta_2$ tetramer encoded by the nifD and nifK genes, and component II (dinitrogenase reductase, or Fe protein), a homodimer encoded by the nifH gene (Fani et al., 1999; Zani et al., 2003). The structural organization of the nif genes shows remarkable differences among nitrogen-fixing organisms (Zehr et al., 2003). In most representatives of the Alphaproteobacteria and Gammaproteobacteria, nif genes are transcribed as a single transcriptional unit (nifHDK operon). However, in some Alphaproteobacteria, such as slow-growing rhizobia, the genes encoding nitrogenase are separated into two operons, so that the genes encoding dinitrogenase (nifDK) are transcribed from a promoter different from that of the structural gene for dinitrogenase reductase (nifH) (Yun & Szalay, 1984).

The cultivation-independent retrieval of nif sequences from different habitats has become a widely used approach to analyse the diversity of dinitrogen-fixing bacteria in ecosystems (Ueda et al., 1995a, b; Zehr et al., 1995, 1998; Widmer et al., 1999; Zani et al., 2000; Lovell et al., 2001; Rösch et al., 2002). However, with very few exceptions (Ueda et al., 1995b), most of the environmental studies so far have focused on the retrieval and comparative sequence analysis of nifH. As a consequence, the currently existing public-domain databases are deficient with respect to nifD, while the rapidly expanding nifH dataset is dominated by sequences from micro-organisms which have not been cultivated. The lack of a sufficient number of nif gene sequences from cultured micro-organisms makes the interpretation of results obtained in environmental studies difficult and limits the clarity of the conclusions that can be drawn from phylogenetic treeing analyses. Our study contributes a representative set of nifH and nifD sequences from MB. These organisms are widely distributed in terrestrial and aquatic environments, and represent an important component of microbial assemblages in many nitrogen-depleted habitats. In addition, MB are known to differ in their nitrogen-fixation capabilities. This makes them an attractive target for comparative analysis of 16S rRNA- and nif-based phylogenies.

In general, nifH gene-based phylogenies have been shown to be in good agreement with the phylogenetic relationships derived from 16S rRNA gene sequences (Hennecke et al., 1985; Young, 1992; Ueda et al., 1995a; Zehr et al., 1995). Comparison of phylogenies constructed for nifH and 16S rRNA genes from the same set of cultivated organisms showed little evidence of lateral gene transfer (Zehr et al., 2003). However, a few cases of incongruent nif and 16S rRNA phylogenies have also been reported, and some of these have been solidly supported by sequence data obtained for the whole nif operon. One example is the root-associated, dinitrogen-fixing species, Herbaspirillum seropedicae. Treeing analyses grouped the NiFH and NiFD sequences of this member of the Betaproteobacteria into a major clade, characterized mainly by Alphaproteobacteria (compare Fig. 4a with 4b) (Machado et al., 1996). Other examples are the two nitrogen-fixing betaproteobacterial strains of Burkholderia, the free-living Burkholderia fungorum (whole-genome project NZ_AA00000001), and the symbiotic 'Candidatus Glomeribacter gigas', which is a Burkholderia endosymbiont of the arbuscular mycorrhizal fungus Gigaspora margarita (Minerdi et al., 2001). The NiFH and NiFD sequences of these organisms exhibit the same aberrant affiliation to alphaproteobacterial-like Nif sequence types as observed in Herbaspirillum seropedicae.

The genus Burkholderia represents an extremely heterogeneous group, including soil bacteria, plant-growth-promoting rhizobacteria, and human and plant pathogens. Thus, the number of Burkholderia spp. that exhibit an aberrant phylogeny of their nifHD gene sequences is still unknown, due to the limited number of strains for which data on both 16S rRNA and nif genes are available. Rösch et al. (2002) assessed the microbial diversity in acid forest soil by parallel characterization of environmental 16S rRNA and nifH gene sequences, using primer sets considered to be universal for nifH. It had been shown in previous studies that the nifH-targeted primers used by Rösch et al. (2002) amplify the corresponding gene fragment in a wide range of dinitrogen-fixing organisms, including members of the Proteobacteria, Firmicutes (representatives of both the Actinobacteria and the Bacillus/Clostridium group), and cyanobacteria. The 16S rRNA gene-based approach revealed an enormous bacterial diversity, while the nifH sequences were assigned mainly to a single cluster characterized by Alphaproteobacteria. One possible explanation for this finding is that many of the micro-organisms present did not possess nifH, or possessed nifH genes that were not targeted by the primer set used. Another hypothesis is that these micro-organisms have acquired alphaproteobacterial nifH gene-sequences by lateral gene transfer. The high prevalence of alphaproteobacterial-like nifH gene-sequence types has also been reported in other environmental studies (Widmer et al., 1999; Lovell et al., 2001). In addition, most of the nifH sequences detected in the environment could not be assigned to any known dinitrogen-fixing bacterium. Thus the evolutionary significance of the lateral transfer of nif genes, as derived from the comparison of 16S rRNA gene- and Nif-based phylogenies, still remains unclear, because of the limited database of nifH and nifD sequences available for cultured bacteria.

However, the assignment of NiFH and NiFD sequences from strains of the gammaproteobacterial Methylococcus capsulatus to a major clade, characterized mainly by Alphaproteobacteria, is an additional example of disagreement...
between the phylogenies constructed for the 16S rRNA gene and inferred Nif sequences. This finding provides further support for the assumption that the lateral transfer of alphaproteobacterial nif gene-sequence types might have distributed effective nitrogen-fixation capabilities among diverse micro-organisms. Further analysis of the sequence data of the Methylococcus capsulatus Bath genome (GenBank NC 002977) confirmed that this MB has single copies of nifH and nifD. The NifH and NifD polypeptides are parts of the two components of nitrogenase, so that convergent evolution in both nifH and nifD of Methylococcus strains is unlikely and the phenomenon can more plausibly be explained by lateral gene transfer. However, the test for differences in codon usage between nif and pmo genes in Methylococcus capsulatus Bath failed to reveal any significant discrepancies. We also compared the codon-usage patterns calculated for Methylococcus capsulatus Bath with those of various diazotrophic Alphaproteobacteria, such as Bradyrhizobium japonicum. These patterns revealed a relatively high degree of similarity. A major determinant of codon usage is the genomic G+C content (Snyder & Champness, 2003). This content was very similar in Methylococcus capsulatus Bath and those Alphaproteobacteria used for comparison. As a consequence, even an evolutionarily recent incorporation of alphaproteobacterial-like nif genes into the Methylococcus genome by lateral gene-transfer might not be detectable by comparative analysis of codon-usage patterns. In addition, such an event of lateral nifH gene transfer could be ancient, so that the molecular traces of the event might have faded to the extent that they are undetectable.

Since the seminal studies of Whittenbury et al. (1970), it has been recognized that, although there are two basic groupings of MB, type I and type II, Methylococcus is an unusual genus because it shares properties of both types (Whittenbury, 1981; Whittenbury & Dalton, 1981). It has therefore been termed type X. The G+C content of DNA from representatives of Methylococcus (59–66 mol%) is closer to the corresponding values of type II MB (61–67 mol%) than to those of type I MB (49–60 mol%). However, all other phylogenetic assignments that have been made on the basis of comparative sequence analysis of the 16S rRNA gene, and genes encoding methane mono-oxygenase or methanol dehydrogenase, have consistently placed Methylococcus together with other representatives of gammaproteobacterial MB. Thus, the nif sequences provide the first molecular basis to highlight some of the physiological similarities observed between Methylococcus spp. and type II MB.

Interestingly, the ability to fix dinitrogen has been considered one of the major characteristics that distinguish representatives of the genus Methylococcus from other gammaproteobacterial MB. Murrell & Dalton (1983) used a combined approach, examining growth under a wide range of PO2 values and using the acetylene reduction test to assess the dinitrogen-fixation capability of MB. This study revealed that only alphaproteobacterial MB and Methylococcus capsulatus grow well in nitrogen-free liquid medium and actively reduce acetylene. Except for Methylococcus spp., gammaproteobacterial MB showed no growth in nitrogen-free medium, even if they were cultivated under microaerobic conditions. The findings of Murrell & Dalton (1983) correspond well with the results of our NifH- and NifD-based trees, in which all alphaproteobacterial MB plus Methylococcus capsulatus form a common cluster, clearly distinct from the gammaproteobacterial MB Methylobomonas and Methylobacter. Auman et al. (2001) demonstrated acetylene reduction activity in representatives of Methylobomonas and Methylobacter. However, growth of these organisms on nitrogen-free media has not been observed, nor has any strain ever been isolated using nitrogen-free enrichment conditions.

Thus, our investigation provides the first evidence that, besides Methylococcus, at least some representatives of type I MB are capable of growth in nitrogen-free media, though under a very limited range of conditions. It should be noted that in the experimental study of Murrell & Dalton (1983), a number of batch cultures of type I MB were also incubated under low PO2 (0.01–0.02 bar) to achieve growth on nitrogen-free media. However, as pointed out above, the OD increase under these conditions is usually very small because the culture becomes limited by O2, and the OD soon declines back to zero. This transient growth cannot be detected if OD is measured at the beginning of the experiment and then at the end of a few days, as done by Murrell & Dalton (1983). A batch culture of type I MB of reasonably high OD was obtained in our experiments on nitrogen-free medium only by repeated injection of small amounts of oxygen to the headspace of the flasks. Alternatively, such a culture might be obtained using chemostat conditions, with O2 as the growth-limiting substrate. This approach has been successfully applied to establish chemostat cultures of Methylosinus trichosporium OB3b and Methylococcus capsulatus Bath, and to show that nitrogen fixation in type II MB is less sensitive to PO2 than in Methylococcus (Murrell & Dalton, 1983). The culture of Methylosinus trichosporium OB3b tolerated a dissolved oxygen tension (DOT) of 22 μM before wash-out occurred, while the maximum DOT tolerated by Methylococcus capsulatus Bath was 6 μM. However, routinely used aerobic batch-cultures contain about 250 μM of dissolved oxygen. Thus, it is not surprising that no type I or type II MB strain has ever been isolated using nitrogen-free enrichment conditions, and that growth on nitrogen-free media has never been demonstrated for these organisms. The sensitivity to O2 of diazotrophic growth of MB might also explain why the maximum population density and activity of these bacteria in nitrogen-depleted environments always corresponds to sites with relatively low PO2 values, such as the chemocline zone in water reservoirs and the layer below the water table in wetlands (Rudd et al., 1976; Sundh et al., 1994; Krumholz et al., 1995). This pattern of MB distribution is usually explained by the availability of both
methane and oxygen. However, the autonomy in nitrogen acquisition, which can be realized only under a limited DOT range, is another explanation of the strict spatial localization of MB in natural environments.

The acidophilic heterotroph *Beijerinckia*, and the acidophilic MB *Methylocapsa*, represent the opposite case, where the genetic potential for dinitrogen fixation is realized under a wide range of pO2 values. All species of *Beijerinckia* examined in this study were capable of growth on nitrogen-free media under fully aerobic conditions (results shown for *Beijerinckia indica* subsp. *lactigenes* only). In contrast to some other well-studied free-living diazotrophic bacteria, such as *Azotobacter*, the mechanisms of dinitrogenase protection from oxygen in *Beijerinckia* have never been investigated. Protection of dinitrogenase from oxygen in *Azotobacter* occurs mainly through two mechanisms: a high respiratory activity that removes oxygen at the cell surface, and formation of alginate, which increases the viscosity of the culture broth and reduces the oxygen transfer rate from the gas phase to the aqueous phase (Sabra et al., 2000). Like *Azotobacter*, strains of *Beijerinckia* also produce large amounts of polysaccharide material when grown in nitrogen-free media, so that the cultures become highly viscous. The cells of *Methylocapsa acidiphila* B2 are also covered by an extracellular polysaccharide matrix (Dedysh et al., 2002), but this polysaccharide is not released into the medium. Thus, at present we cannot explain the mechanism of dinitrogenase protection in this acidophilic MB.

The almost complete identity of both NifH and NifD sequences in *Beijerinckia* and *Methylocapsa* demands some speculation about the evolutionary mechanisms that led to this phenomenon. One explanation would be that the two genera originated from a common acidophilic dinitrogen-fixing ancestor and subsequently experienced similar evolutionary selection pressures with regard to nitrogen acquisition. The *nifH* gene fragments of *Methylocapsa acidiphila* B2 and *Beijerinckia indica* subsp. *indica* exhibit a sequence divergence of 9.9%. The *nifD* gene fragments of these organisms display the same level of sequence divergence. However, most nucleotide sites that differ between the corresponding *nif* gene fragments in *Methylocapsa* and *Beijerinckia* are synonymous nucleotide substitutions and thus do not change the amino acid sequence. Apparently, selection pressure has maintained the primary structure of nitrogenase in these bacteria to enable their survival in nitrogen-depleted conditions.

The 16S rRNA gene-based phylogeny suggests, however, that members of the genus *Methylocella* also originated from the same common ancestor as *Methylocapsa* and *Beijerinckia*. Thus, although *Methylocella* and *Methylocapsa* have been isolated from the same acidic *Sphagnum* peat-bog environment, the pattern of *nif* gene evolution implies that *Methylocella* spp. possess a different niche-specialization for nitrogen from that of *Methylocapsa acidiphila* and *Beijerinckia* spp. (compare Figs 1 and 4). In contrast to *Methylocapsa acidiphila*, the representatives of the genus *Methylocella* can utilize, in addition to inorganic nitrogen compounds, methyamines and some amino acids (Dunfield et al., 2003; Dedysh et al., 2004). Thus, the ability to fix dinitrogen is not a strictly obligatory trait for *Methylocella* spp., and the *nif* genes in these bacteria might have evolved in a different way.

**ACKNOWLEDGEMENTS**

This research was supported in part by the Russian Fund of Basic Research (grant no. 02-04-04003), the Program PCB of the Russian Academy of Sciences, the Deutsche Forschungsgemeinschaft [436 RUS 113/543/0-2(R)], and by the Competence Network Göttingen ‘Genome research on bacteria’ (GenoMik), financed by the German Federal Ministry of Education and Research (BMBF).

**REFERENCES**


