

Comparative analysis of ribonuclease P RNA of the planctomycetes

Margaret K. Butler and John A. Fuerst

Correspondence

John A. Fuerst

j.fuerst@mailbox.uq.edu.au

Department of Microbiology and Parasitology, School of Molecular and Microbial Sciences, University of Queensland, Brisbane, Queensland 4072, Australia

The planctomycetes, order *Planctomycetales*, are a distinct phylum of domain *Bacteria*. Genes encoding the RNA portion of ribonuclease P (RNase P) of some planctomycete members were sequenced and compared with existing database planctomycete sequences. *rnpB* gene sequences encoding RNase P RNA were generated by a conserved primer PCR strategy for *Planctomyces brasiliensis*, *Planctomyces limnophilus*, *Pirellula marina*, *Pirellula staleyi* strain ATCC 35122, *Isosphaera pallida*, one other *Isosphaera* strain, *Gemmata obscuriglobus* and three other strains of the *Gemmata* group. These sequences were aligned against reference bacterial sequences and secondary structures of corresponding RNase P RNAs deduced by a comparative approach. P12 helices were found to be highly variable in length, as were helices P16.1 and P19, when present. RNase P RNA secondary structures of *Gemmata* isolates were found to have unusual features relative to other planctomycetes, including a long P9 helix and an insert in the P13 helix not found in any other member of domain *Bacteria*. These unique features are consistent with other unusual properties of this genus, distinguishing it from other bacteria. Phylogenetic analyses indicate that relationships between planctomycetes derived from RNase P RNA are consistent with 16S rRNA-based analyses.

INTRODUCTION

Ribonuclease P (RNase P) is the most widespread ribonuclease (Condon & Putzer, 2002), found in all organisms so far examined (Altman *et al.*, 1989) except the bacterial species '*Aquifex aeolicus*' (Condon & Putzer, 2002), and also found within the subcellular compartments of eukaryotes known to synthesize tRNA (Frank & Pace, 1998). It is one of the key enzymes involved in the processing of tRNA, responsible for the generation of the 5' termini of mature tRNA (Altman, 1975), via a single endonucleolytic cleavage (Deutscher, 1984). Due to its universal existence, it is most likely an ancient enzyme, possibly descended from the postulated 'RNA world', prior to the development of protein-directed catalysis (Brown & Pace, 1991; Jeffares *et al.*, 1998). In bacteria, the RNase P molecule is composed of a small protein and a large RNA moiety (Pace & Brown, 1995), the latter of which is capable of *in vitro* catalysis (Guerrier-Takada *et al.*, 1983). RNase P RNA molecules in bacteria can generally be divided into two structurally separate classes: type A, the most common form, and type B, which is found in the low-G+C Gram-positive bacteria

(Haas *et al.*, 1996). Type A is distinguished by the presence of helix P6 and the absence of helix P5.1. A third intermediate type, type C, has been observed in the green non-sulfur bacteria (Haas & Brown, 1998). RNase P RNA has proven useful for phylogenetic analysis of the *Chlamydiales* (Herrmann *et al.*, 2000), *Prochlorococcus* (Schön *et al.*, 2002), the LL-2,6-diaminopimelic acid-containing actinomycetes (Yoon & Park, 2000), *Bartonella* (Pitulle *et al.*, 2002) and *Vibrio* core species (Maeda *et al.*, 2001), among others.

The phylum *Planctomycetes* (Garrity *et al.*, 2001) comprises a distinct group of the domain *Bacteria* on the basis of 16S rRNA analyses (Embley *et al.*, 1994; Fuerst, 1995; Hugenholtz *et al.*, 1998; Liesack *et al.*, 1992; Schlesner & Stackebrandt, 1986; Woese *et al.*, 1985). They may be one of the most deeply branching lineages within the domain *Bacteria* (Brochier & Philippe, 2002; Di Giulio, 2003). They display a number of unusual phenotypic features for bacteria, including peptidoglycan-less cell walls, short 5S rRNA and an unlinked *rrn* operon organization (Fuerst, 1995). Unique within the domain *Bacteria* is the compartmentalized cell structure shared by all planctomycetes investigated so far (Lindsay *et al.*, 2001). Most striking among compartmentalization of this type is the double-membrane-bounded nucleoid with analogy to the eukaryotic nucleus, displayed by isolates of the *Gemmata* group (Fuerst & Webb, 1991; Lindsay *et al.*, 2001; Wang *et al.*, 2002).

Published online ahead of print on 9 February 2004 as DOI 10.1099/ijms.0.03013-0.

The GenBank accession numbers for the RNase P RNA gene sequences determined in this study are AY371111–AY371120 and AY528422.

There are four genera so far described with members that have been cultured, and all are aerobic chemoheterotrophs – *Planctomyces*, *Pirellula*, *Gemmata* and *Isosphaera*. *Gemmata* and *Isosphaera* have been described only from a single species in each case, *Gemmata obscuriglobus* and *Isosphaera pallida* (Franzmann & Skerman, 1984; Giovannoni *et al.*, 1987). The diversity of the group extends beyond these to include at least three further planctomycete genera at *Candidatus* status (*Brocadia*, *Kuenenia* and *Scalindua*), described from enrichments and consisting of anaerobic ammonium-oxidizing autotrophs (Kuypers *et al.*, 2003; Schmid *et al.*, 2000; Strous *et al.*, 1999) with unique types of membranes and membrane lipids (Sinninghe Damsté *et al.*, 2002). The presence of planctomycete sequences in 16S rRNA clone libraries constructed from environmental samples shows that these organisms can be found in diverse habitats (Derakshani *et al.*, 2001; Vergin *et al.*, 1998).

The comparative study of the RNA portion of RNase P is valuable firstly for studying molecular evolution due to significant variation of primary and secondary structure among different phylogenetic groups. Secondly, it is useful for inferring the functional significance of different regions within the RNA molecule from the nature of the structural variation. Planctomycete RNase P RNA sequence data have only been available for *Pirellula staleyi*, *Planctomyces maris* (Haas & Brown, 1998) and some environmental clone sequences (Brown *et al.*, 1996; Harris *et al.*, 2001) and, more recently, the first complete genome sequence for a planctomycete has become available for a *Pirellula* strain (Glöckner *et al.*, 2003), but this does not adequately represent the diversity of even the cultured species, which extends to members of the genera *Gemmata* and *Isosphaera*. Comparison of planctomycete RNase P RNA sequences with those of other bacteria will enable us to determine whether the unique position of these organisms is reflected by any unique features of their RNase P RNA.

In this study, the sequence information on planctomycete RNase P RNA was expanded to include ten new sequences to cover all four genera for which pure cultures exist. The RNase P RNA sequences of the species studied were compared both within the phylum and with sequences of other bacteria by analysis of predicted secondary structure and by phylogenetics. This was done with a view to determining whether planctomycete RNase P RNA is consistent with the distinctive nature of this phylum phenotypically and phylogenetically and with planctomycete relationships based on other molecules.

METHODS

Nucleic acid purification. *Planctomyces limnophilus* ATCC 43296^T (5 days), *Pirellula staleyi* ATCC 27377^T (12 days), *Pirellula staleyi* ATCC 35122 (11 days), *Gemmata obscuriglobus* ACM 2246^T (4 days) and strains ACM 5151 (8 days), ACM 5153 (8 days), ACM 5154 (11 days) and ACM 5156 (8 days) were all grown on M1 agar (Schlesner, 1994) plates at 28 °C for the length of time indicated in parentheses. *Isosphaera pallida* IS1B^T was grown on IMC agar

(Giovannoni *et al.*, 1987) for 6 days at 42 °C in a CO₂-enriched atmosphere. *Pirellula marina* ACM 3344^T (11 days) and *Planctomyces brasiliensis* ATCC 49424^T (5 days) were grown on half-strength marine agar (Lindsay *et al.*, 1997) at 28 °C.

All DNAs, except those from *Isosphaera pallida* and ACM 5154 were extracted using DNAzol reagent (Life Technologies) with modifications to the manufacturer's instructions. Briefly, cells were scraped directly into 20 mM Tris/HCl (pH 8.0), 10 mM EDTA before being pelleted and gently resuspended in DNAzol reagent. Cell suspensions were incubated at room temperature for 10 min and then heated at 65 °C for 30 min with vortexing every 10 min. Suspension was pelleted and the supernatant put into a fresh tube, while cell lysis steps were repeated on the pellet. Supernatant was then centrifuged at 14 000 r.p.m. for 15 min and 0.5 ml –20 °C 100 % ethanol was added to the resulting supernatant. This was then incubated at –70 °C for 1 h before being pelleted. The resulting pellet was washed with –20 °C 95 % ethanol, resuspended in TE buffer with 20 µg RNase A ml⁻¹ and incubated at 37 °C for 30 min. DNA from *Isosphaera pallida* and ACM 5154 were extracted using DNeasy Tissue kit (Qiagen), according to the manufacturer's instructions for the isolation of genomic DNA from Gram-positive bacteria.

PCR amplification and sequence determination. A combination of primers was used for PCR amplification of the RNase P RNA gene. These were 59FBam (5'-CGGGATCCGIIAGGAAAGTCCII-GC-3'), incorporating target sequence positions 58–75 in *Escherichia coli*, 347REco (5'-CGGAATTCRTAAGCCGGRTTCTGT-3'), positions 347–362 in *E. coli*, 174R (5'-CTCTTACCSCACCNTTTTCA-CCCT-3'), positions 174–196 in *E. coli*, 59F (5'-GIIGAGGAAAG-TCCIIIGC-3', a modification of 59FBam) and 347R (5'-RTAAGC-CGGRTTCTGT-3', a modification of 347REco) (Brown *et al.*, 1996; Haas *et al.*, 1994). PCR was performed for each organism in reaction mixtures containing 50 mM KCl, 10 mM Tris/HCl, pH 8.3, 50 µM dNTPs, 5 ng each primer, 1.25 mM MgCl₂ and 100–200 ng DNA. For optimal amplification of *Planctomyces brasiliensis*, *Planctomyces limnophilus*, *Isosphaera pallida*, *Gemmata obscuriglobus*, ACM 5151 and ACM 5154, the PCR mixtures required the addition of 1 M betaine and, for *Gemmata obscuriglobus* and the *Gemmata*-like isolates ACM 5151, ACM 5153 and ACM 5156, PCR mixtures required 2.5 mM MgCl₂. Amplification was performed using an initial incubation of 10 min at 95 °C, followed by 35 cycles of 1 min at 95 °C, 1 min at 38 °C and 1 min at 72 °C, followed by a final incubation at 72 °C for 10 min. Resulting PCR products were cleaned using either the QIAEX II gel extraction kit (Qiagen) or the NucleoSpin Extract 2 in 1 kit (Macherey-Nagel), according to the manufacturers' instructions, and then sequenced, using the above primers and the primer 174F (5'-AGGGTGAAANGGTGSGGTAAGAG-3') (Brown *et al.*, 1996), via cycle sequencing with BigDye fluorescent dideoxynucleotide chain terminators (versions 2.0 or 3.1) followed by electrophoresis on an ABI automated sequencer (Australian Genome Research Facility). The resulting sequence represented the region between 76 and 346 of *E. coli* (Haas & Brown, 1998).

Sequence, structure and phylogenetic analysis. New sequences were aligned to selected available type A sequences, representing all possible groups in the domain *Bacteria*, from an alignment in the RNase P database (Brown, 1999) using [GeneDoc version 2.6](#) (Nicholas & Nicholas, 1997) and were manually edited. **Secondary structures were determined and drawn based on the *E. coli* structure deduced by Harris *et al.* (2001) and refined using the alignment and by comparison with other bacterial structures. The alignment was then edited again, based on the determined secondary structure.** The helix nomenclature used is that of Haas *et al.* (1994) and nucleotide position numbering is based on *E. coli* (Haas & Brown, 1998). The program MFOLD (Walter *et al.*, 1994; Zuker, 2003; <http://mfold.burnet.edu.au/>) was used to **determine specific folding of certain**

helices. The resulting alignment was edited to remove all equivalent column positions where a gap was present, leaving 248 positions for inclusion in phylogenetic analysis. Phylogenetic analysis was performed using the final corrected alignment and the phylogenetic package PHYLIP 3.6a3 (Felsenstein, 1989), used via <http://www.pasteur.fr/english.html>. The program DNADIST was used to generate a distance matrix for the neighbour-joining method, using the Kimura two-parameter model, which was then used to prepare a distance tree, using NEIGHBOR. The programs DNAPARS and fastDNAmI were used to create parsimony and maximum-likelihood trees, respectively. One thousand bootstrap resamplings were performed (100 for maximum-likelihood), with sequences being added in random order. Gene similarity was determined using the same alignment as used for phylogenetic analysis. This was performed using DNADIST with Similarity Table selected as the DNADIST option.

RESULTS AND DISCUSSION

Sequence determination and characteristics of the primary sequence

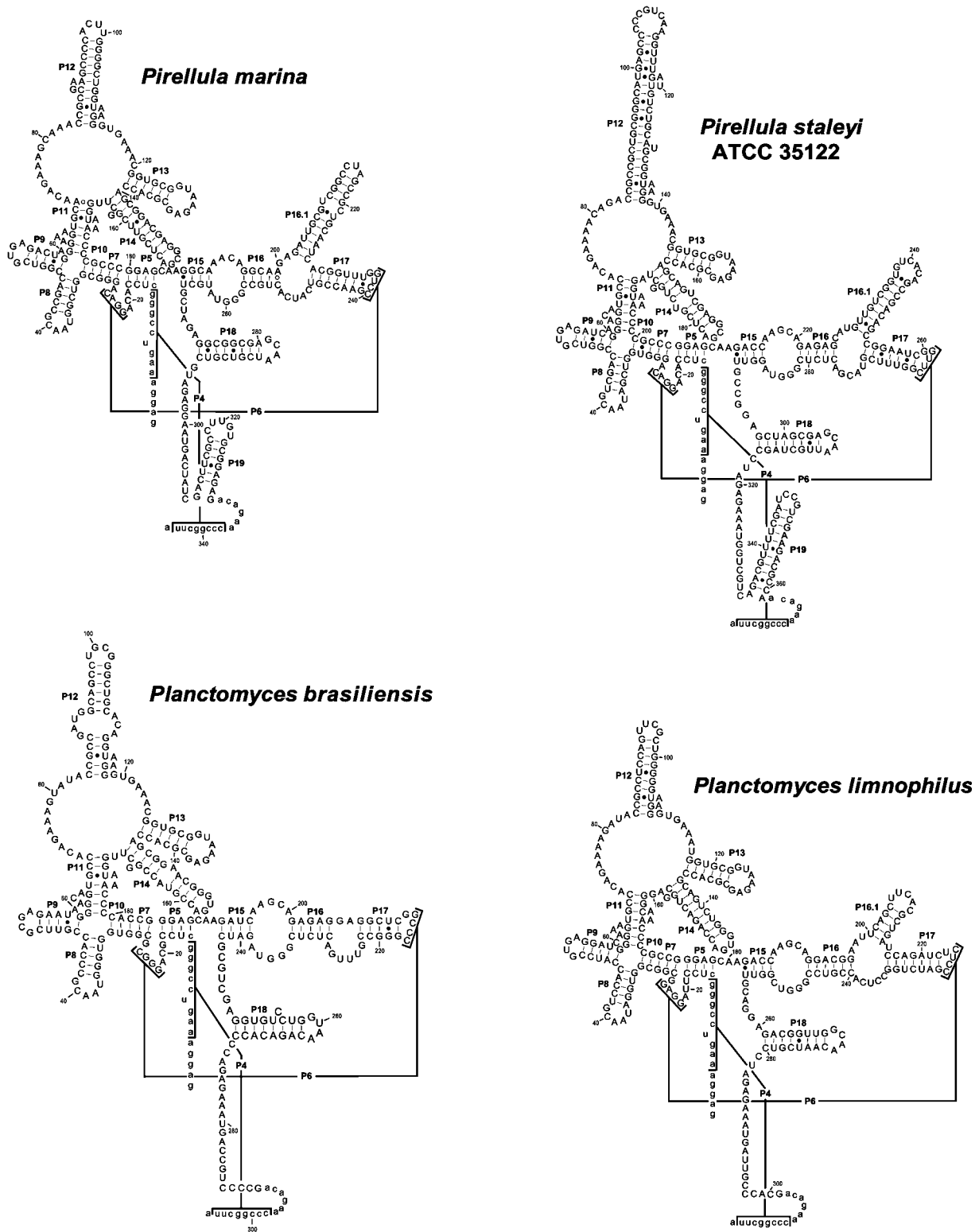
The RNase P RNA genes of a total of ten different isolates of the phylum *Planctomycetes* were sequenced for the first time in this study and as a result we now have a representative collection of sequences of members of all cultured species in this phylum. The isolates for which the RNase P RNA genes were sequenced were *Planctomyces brasiliensis*, *Planctomyces limnophilus*, *Pirellula marina*, *Pirellula staleyii* ATCC 35122, *Isosphaera pallida* and the *Isosphaera*-like isolate ACM 5154 and *Gemmata obscuriglobus* and the *Gemmata*-like isolates ACM 5151, ACM 5153 and ACM 5156. The partial sequences ranged in length from 306 to 482 bases. In addition, although a database sequence exists for the *Pirellula staleyii* strain ATCC 27377^T, it was re-sequenced to resolve database ambiguities, resulting in the revision of nucleotides at two positions, and this sequence was used in all further analyses. The sequences of the two strains of *Pirellula staleyii*, ATCC 27377^T and ATCC 35122, were found to be identical except for two bases, one each in the P14 helix and the J17/16 region. This is consistent with the confirmation of the very close phylogenetic relationship of ATCC 35122 to ATCC 27377^T and with the phenotypic similarities of these two strains (Butler *et al.*, 2002). The common core structure for type A bacterial RNase P RNA (Haas *et al.*, 1996) is present in all cultured and uncultured planctomycetes, including all conserved helices, nucleotide positions and the catalytic and specificity regions (Christian *et al.*, 2000; Green *et al.*, 1996). All invariant nucleotides are also present, except in the Baltic Sea isolate *Pirellula* sp. strain 1 (Glöckner *et al.*, 2003) and the related database clone sequences of SM-A06, SM-A61 and SM-A88 (Harris *et al.*, 2001), isolated from a salt marsh, which possess a U in position 248 of *E. coli* RNase P RNA. Among the planctomycetes, sequence similarities between homologous regions ranged from 70.2% (Lake Griffy sediment clone LGW#46 vs the *Gemmata*-like isolate ACM 5151) to 99.6% (clone SM-A61 vs clone SM-A88 and between the two *Pirellula staleyii* strains), with a mean of 78.8%. Within the individual genera themselves, *Pirellula*

marina displayed similarity values of 80.7 and 81.0% when compared to the two *Pirellula staleyii* strains, ATCC 27377 and ATCC 35122, respectively. The *Planctomyces* species displayed similarity values between themselves ranging from 78.2% (*Planctomyces brasiliensis* and *Planctomyces limnophilus*) to 85.1% (*Planctomyces brasiliensis* and *Planctomyces maris*), the two *Isosphaera* isolates displayed similarity values of 89.1% and the *Gemmata* isolates displayed similarity values between themselves ranging from 81.9% (ACM 5151 vs *Gemmata obscuriglobus* and ACM 5153) to 98.0% (ACM 5153 vs ACM 5156).

Analysis of existing database sequences, clones SM-A06, SM-A61 and SM-A88 (Harris *et al.*, 2001) and *Pirellula* sp. strain 1, has revealed the mutation of a base universally conserved in *Archaea* and *Bacteria* (Zahler *et al.*, 2003). Thought to be invariant, A248 has been replaced by a U in these sequences. This nucleotide change occurs in a conserved site considered to be important for function (Zahler *et al.*, 2003). An A→U mutation at this position (*E. coli* position 248) has been shown to be of particular significance to the binding affinity and catalysis relevant to pre-tRNA processing (Zahler *et al.*, 2003). The change in the base at this site, which is otherwise conserved throughout the *Bacteria*, is consistent with the divergent position of the planctomycetes.

Secondary structure analysis

The secondary structures of the ten newly sequenced planctomycete isolates were deduced and are shown in Figs 1 and 2. A summary of selected secondary structure features of all planctomycetes so far sequenced is presented in Table 1. As expected, the secondary structure of planctomycete RNase P RNA is more conserved within the phylum than the primary sequence. This is illustrated by the cultured species of *Pirellula*, which possess identical secondary structures and yet share only 80.7–81.0% similarity in their primary sequence. However, there are significant variations in the secondary structure that reflect the phylogenetic distinctions within the planctomycetes and of planctomycetes within *Bacteria*. The variability in the primary sequence of different species within the same genus is generally not reflected in the presence or absence of specific helices, for example in the case of *Pirellula staleyii* and *Pirellula marina* (Fig. 1). In most cases the same identifiable helices were present in cultured members of the same genus, although helices in some instances varied in length and primary sequence, for example P12 and P19 helices in *Pirellula staleyii* and *Pirellula marina* (Fig. 1). However, in the genus *Planctomyces*, the P16.1 helix is absent from the RNase P RNA of *Planctomyces brasiliensis*. Between genera the helices P16.1 and P19 varied with respect to their occurrence. The P19 helix is absent in *Planctomyces limnophilus* but present in all other isolates illustrated in Fig. 1 and also absent from *Planctomyces maris* and one of the clone sequences from Lake Griffy, LGA#6 (not shown). The helix P16.1 is absent from all *Gemmata* isolates (Fig. 2), *Planctomyces brasiliensis*, *Pirellula* sp. strain 1 and one Lake



Griffy clone sequence, LGW#46, but present in all other species. However, when present, both these regions are highly variable in both sequence and length. Also variable amongst species is the P12 helix, ranging in size from 26 bases in *Planctomyces limnophilus* (Fig. 1) up to 151 bases

in the *Gemmata*-like isolate ACM 5156 (Fig. 2), representing the longest P12 helix found so far in type A bacteria. In contrast is the secondary structure of helix P16, which is constant, consisting of 5 bp in all species except *Planctomyces limnophilus*, which has 6 bp.

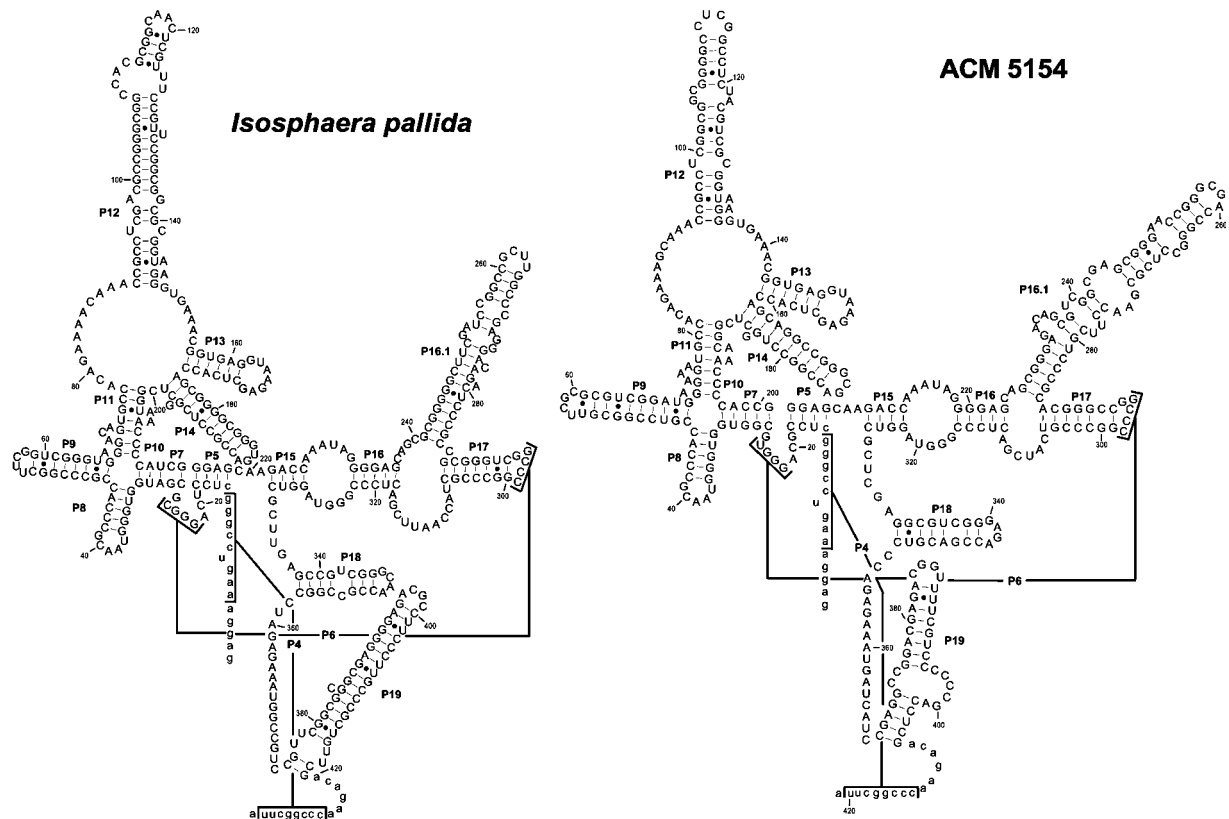


Fig. 1. Deduced secondary structures of RNase P RNA in *Pirellula marina*, *Pirellula staley* ATCC 35122, *Planctomyces brasiliensis*, *Planctomyces limnophilus*, *Isosphaera pallida* and the *Isosphaera*-like isolate ACM 5154. Primer sequences are in lower case.

Another region of interest is the P9 helix. As illustrated by *Pirellula* and *Planctomyces* species (Fig. 1), most planctomycetes have a P9 helix composed of 15 bases, in the form of a 5 bp region, a GNRA terminal tetraloop and a single base bulge. The *Isosphaera* strains *Isosphaera pallida* and ACM 5154 differ from this 5 bp P9 helix organization, in that the P9 helix of *Isosphaera pallida* is composed of 21 bases including a UUCG terminal tetraloop and that of ACM 5154 comprises 25 bases with a UCGC terminal tetraloop. As discussed in detail below, *Gemmata* isolates (Fig. 2) also have significant P9 helix variations. Also interesting is the J15/16 region, which consists of five bases in all species except *Pirellula marina* and the two *Isosphaera* isolates, in which it has a length of four and six bases, respectively.

Comparison of the newly deduced secondary structures (Figs 1 and 2) with each other and with previously available secondary structures has revealed a number of interesting features of planctomycete RNase P RNA molecules. The availability of the cultures ACM 5151, ACM 5153 and ACM 5156, known via 16S rRNA sequence analysis to cluster with the genus *Gemmata*, has revealed a shared insert in the P13 helix of the deduced RNase P RNA secondary

structure (Fig. 3), unique not only within the planctomycetes but also unique among secondary structures of any RNase P RNA available in the database for the domain *Bacteria*. When the P13 helix is present in bacterial sequences, it usually takes the form of a 6 bp helix ending in a loop of GUAAGAG or a close variation of this sequence. However, in all the *Gemmata* isolates sequenced thus far, including two other isolates whose results are not shown, this helix has contained an additional 4 or 5 bp helix, ending in a tetraloop, followed by the five base sequence UUAUU. This inserted helix, designated P13.1, is always located between the end of the base-paired region and the beginning of the signature loop sequence of P13. Whilst other *mpB* genes have been found to possess an insert within the P13 helix, including the proteobacteria clone sequence ESH7-4 (Brown *et al.*, 1996), which has a large, two-helix insert, none have been found to possess an insert in a similar position to the *Gemmata* isolates and none were found in all members of the same genus. The base-paired region of the P13.1 helix was found to be an identical 4 bp region in *Gemmata obscuriglobus* and all *Gemmata*-like isolates except ACM 5151, which has a 5 bp helix and also varied in sequence from the other isolates. The tetraloop region of the P13.1 helix was highly variable among

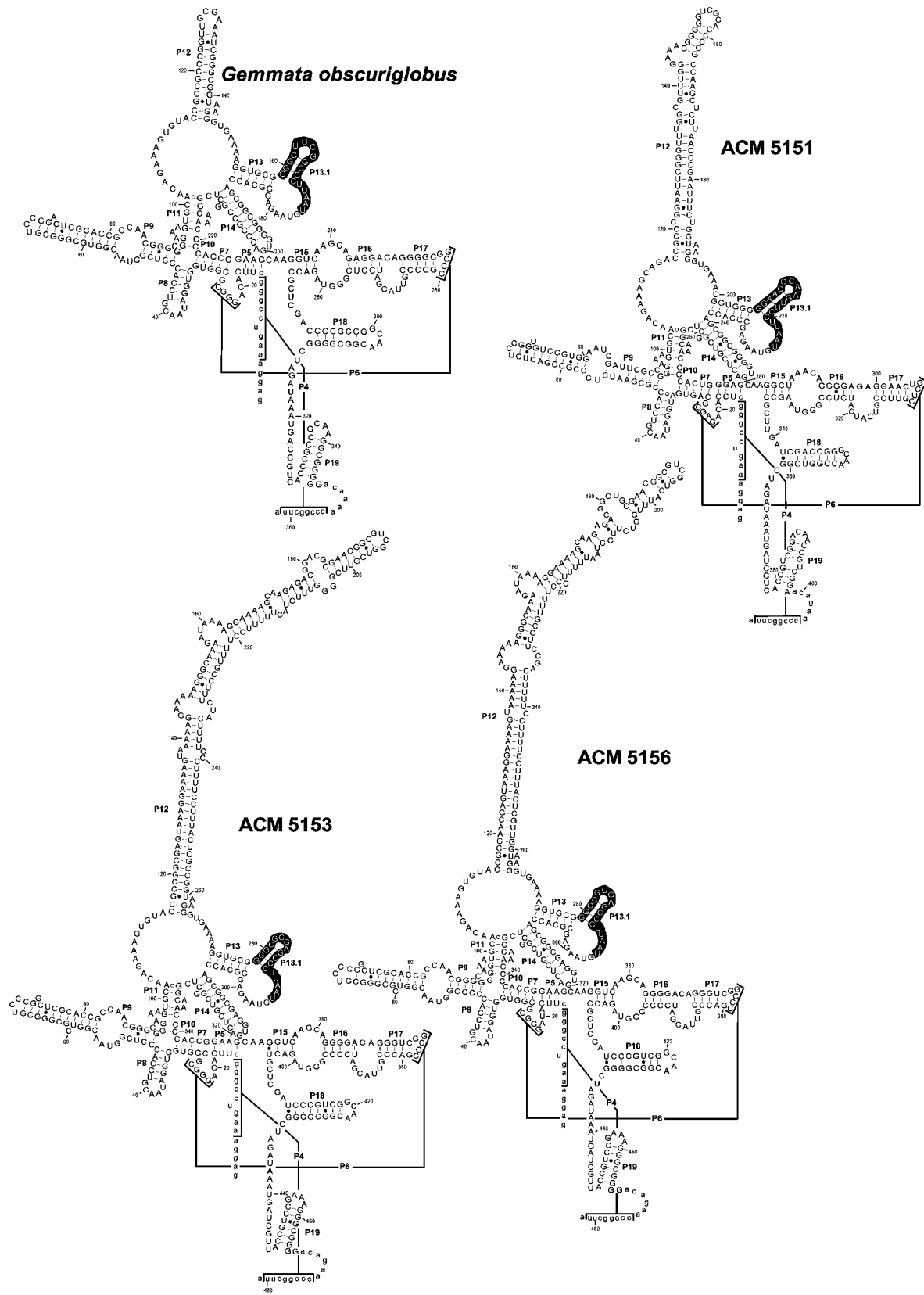


Fig. 2. Deduced secondary structures of RNase P RNA in *Gemmata obscuriglobus* and the *Gemmata*-like isolates ACM 5151, ACM 5153 and ACM 5156. Primer sequences are in lower case. The *Gemmata*-specific unique insert (including the P13.1 helix) is highlighted with a black background.

Table 1. Distribution of secondary structure features among planctomycete sequences studied

Secondary structure features include: P9, extended in length and lacking a terminal GNRA tetraloop in *Gemmata*-like and *Isosphaera*-like organisms; P12, variable in length, ACM 5156 possesses longest found so far in type A bacteria, ACM 5153 and ACM 5156 possess short tandem repeats; P13.1, tailed by the sequence UUAU, unique among organisms in the domain *Bacteria*; J15/16, constant in length in all planctomycetes except *Pirellula marina* and *Isosphaera*-like isolates; P16, constant in length in all planctomycetes except *Planctomyces limnophilus*; P16.1 and P19, variable in length. +, Presence of structural feature; –, absence of structural feature.

Organism	Secondary structure feature						
	P9	P12	P13.1	J15/16	P16	P16.1	P19
<i>Pirellula marina</i>	+	+	–	+	+	+	+
<i>Pirellula staleyi</i>	+	+	–	+	+	+	+
<i>Pirellula</i> sp. strain 1	+	+	–	+	+	–	+
SM-A06	+	+	–	+	+	+	+
SM-A61	+	+	–	+	+	+	+
SM-A88	+	+	–	+	+	+	+
LGA#6	+	+	–	+	+	+	–
LGA#8	+	+	–	+	+	+	+
LGW#46	+	+	–	+	+	–	+
LGW#116	+	+	–	+	+	+	+
<i>Planctomyces brasiliensis</i>	+	+	–	+	+	–	–
<i>Planctomyces limnophilus</i>	+	+	–	+	+	+	–
<i>Planctomyces maris</i>	+	+	–	+	+	+	–
<i>Isosphaera pallida</i>	+	+	–	+	+	+	+
ACM 5154	+	+	–	+	+	+	+
<i>Gemmata obscuriglobus</i>	+	+	+	+	+	–	+
ACM 5151	+	+	+	+	+	–	+
ACM 5153	+	+	+	+	+	–	+
ACM 5156	+	+	+	+	+	–	+

Gemmata sequences, though in all sequences except *Gemmata obscuriglobus*, the tetraloop of P13.1 is of the form GNRA. This suggests the possibility that this helix is involved in long-range interactions, since the interaction between GNRA tetraloops and the shallow groove of RNA helices is one of the most common long-range structural motifs (Massire *et al.*, 1997). Study of the covariation of the nucleotides within this helix supports the secondary structure deduction. In ACM 5151 there is covariation with the A:U pair at positions 209:218, which is a G:C pair in all other *Gemmata* isolates. The insert in the P13 helix thus appears phylogenetically significant, and is unique to the planctomycete genus *Gemmata* among organisms of the domain *Bacteria*. It thus may provide a useful taxonomic marker for this genus. Interestingly, there is also an insertion characteristic for all cultured *Gemmata* isolates present within their 16S rRNA gene sequences (Liesack & Stackebrandt, 1992).

Another phylogenetically significant feature and one forming a differential characteristic of *Gemmata* isolates in comparison to all other planctomycetes is the P9 helix (see Fig. 2), which is both longer in length and lacks a terminal tetraloop. All *Gemmata* isolates have a P9 helix consisting of 45–47 bases with a UCC loop at the distal end. This is

in contrast to the 15 base structures including a GNRA terminal tetraloop of most planctomycetes. The absence of a terminal GNRA tetraloop is highly significant in terms of functional interactions, since this GNRA tetraloop forms tertiary contact with base-pairs near the distal end of helix P1 (Harris *et al.*, 1997; Massire *et al.*, 1997).

The occurrence of only a triloop in *Gemmata* isolates, rather than a GNRA tetraloop in this helix, is rare among the domain *Bacteria*, occurring mainly in the spirochaetes, scattered unclassified clone sequences and one example within the phylum *Bacteroidetes* (formerly the *Cytophaga–Flavobacterium–Bacteroides* division) (Brown, 1999). Whilst not as rare, the presence of a non-GNRA tetraloop in *Isosphaera* isolates is also uncommon among the domain *Bacteria*, occurring in groups such as the phylum *Proteobacteria* and the groups mentioned above. The size of the P9 helix in *Gemmata* and *Isosphaera* isolates is also rare among the domain *Bacteria*, with relatively few sequences having P9 helices longer than 15 bases and fewer still having a P9 helix longer than the *Gemmata* helices.

Also notable in the *Gemmata* cluster is the presence of the longest P12 helix in RNase P RNA of bacterial type A known, in the *Gemmata*-like isolate ACM 5156 (Fig. 2).

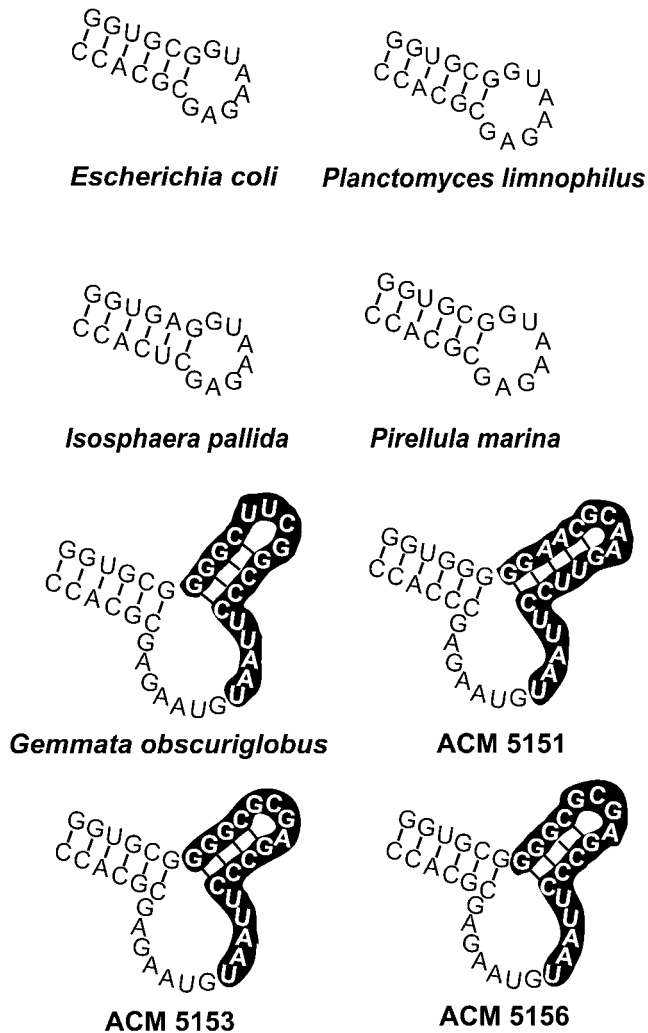


Fig. 3. Deduced secondary structure of the P13 helix region of representatives of all four genera of the planctomycetes, including *Gemmata obscuriglobus* and all *Gemmata*-like isolates, compared with the published secondary structure for the P13 helix of *E. coli* (Harris *et al.*, 2001). The unique insert in the *Gemmata* isolates is highlighted with a black background.

At 151 bases (ACM 5153 is only one base shorter) it is 40% of the length of the full *E. coli rnpB* gene, as opposed to the *E. coli* P12 helix which is only 9% of the length of the full RNase P RNA gene. It has been suggested that there may be tertiary interactions between the P13 stem-loop structure and P12 (Brown *et al.*, 1996; Massire *et al.*, 1998), although not clearly attributable to any sequence positions. However, it is possible that the insert in the P13 helix somehow acts to stabilize or otherwise interact with the long P12 helix in ACM 5156 and ACM 5153 (which differs in length from ACM 5156 by only one base). However, the presence of the P13 helix insert in *Gemmata obscuriglobus*, which does not possess a long P12 helix, being more similar in length to other planctomycete species, and the presence

of a relatively long P12 helix in the environmental clone sequence LGW#46 that lacks an insert in its P13 helix, complicates this argument. As a result of long P12 helices, the sequences of ACM 5153 and ACM 5156, at 481 and 482 bases in length, respectively, represent the longest *rnpB* genes sequenced so far in bacteria. Another feature of these two isolates was the presence of short tandem repeats, six or seven bases long. In both isolates, the repetitive sequence is 130-GGAAAA GUAAAA GGAAAA GGGCAAA GAUAAAA GGAAAA-169 and its complement when a base-paired region is present. These repeats suggest the possibility that these are analogous to the short tandemly repeated repetitive (STRR) sequences as observed in the P12 helix of RNase P RNA of heterocyst-forming cyanobacteria (Vioque, 1997). Such repeats and the consequent length variation in a helix may form through slippage mechanisms, such as slipped-strand mispairing, similar to those suggested to form a major mutational mechanism in variable regions of rRNA (Hancock & Vogler, 2000). Cyanobacteria and planctomycetes appear to be the only organisms among either the *Bacteria* or the *Archaea* that possess STRR repeats in their RNase P RNA. However, unlike the cyanobacteria, all planctomycetes contain the GGU sequence in the loop connecting the P15 and P16 helices, a motif believed to interact with the 3'-end RCCA sequence of tRNAs (Kirsebom & Svård, 1994; Oh & Pace, 1994).

The *Gemmata* cluster members are all known to possess a unique cell ultrastructure among bacteria, most dramatically in the form of a nucleoid bounded by a double membrane envelope forming a compartmentalized nuclear region, analogous to the eukaryote nucleus in some ways (Fuerst & Webb, 1991; Wang *et al.*, 2002). Such a structure might conceivably have implications for the ways in which precursors of tRNA were processed and therefore for RNase P RNA structure. The phylogenetic position and distinctiveness of the cell structural plan in the cultured *Gemmata* strains studied are consistent with the shared distinctive RNase P RNA structure in this group. The availability of RNase P RNA sequences for *Gemmata* strains makes possible the design of an oligonucleotide probe for use in subcellular localization of RNase P RNA within *Gemmata* cells, and determination of possible correlations with compartmentalized functions as well as structures within these nucleated cells. Investigation of compartmentalization functions in other planctomycetes is also facilitated by the availability of RNase P RNA sequences for all four cultured genera.

The wider phylum *Planctomycetes* representatives, including three other genera represented in pure culture, *Isosphaera*, *Planctomyces* and *Pirellula*, display great diversity in RNase P RNA secondary structure, though genus-level shared features are detectable. A primary sequence signature or feature of secondary structure shared by all planctomycetes is not detectable. The wide diversity of even secondary structure in planctomycetes, to some extent even at the

genus level, is illustrated by the absence of the P19 helix only in the genus *Planctomyces* and the variable length of the P12 helix even within just the *Gemmata* group. This is consistent with the 16S rRNA sequence heterogeneity and relatively large evolutionary distances within the planctomycetes as a phylum. It is possibly also consistent with an evolutionary mechanism involving the formation of indels or repeated repetitive sequence duplication via slippage rather than only nucleotide substitutions. The possibility of a mechanism for variation in length of the P12 helix in *Gemmata* is indicated by the occurrence of possible short tandem repeat sequences in ACM 5153 and ACM 5156.

Phylogenetic analysis

Phylogenetic analysis was performed using several different analysis tools with the complete set of planctomycete sequences and a small set of outgroup sequences, including *Verrucomicrobium spinosum*, *Chlamydia trachomatis* and *Chlamydophila psittaci*, organisms sometimes clustering with planctomycetes in a superphylum in 16S rRNA-based trees, and the members of two distinct, deeply branching phyla of *Bacteria*, *Thermotoga maritima* and *Thermodesulfotobacterium commune*. The topology of the subtree of planctomycete RNase P RNA sequences (Fig. 4) was consistent with that of 16S rRNA-based trees based on the members represented in both trees (Wang *et al.*, 2002),

suggesting that lateral gene transfer has not contributed markedly to RNase P RNA or 16S rRNA evolution within the phylum. For example, *Gemmata obscuriglobus* and the isolates ACM 5151, ACM 5153 and ACM 5156 [strains JW10-3f1, JW3-8s0 and JW9-3f1 in Wang *et al.* (2002), respectively], and *Isosphaera pallida* and ACM 5154 [strain CJuql1 in Wang *et al.* (2002)] clustered monophyletically with significant bootstrap support in both 16S rRNA and RNase P RNA-based trees in at least one of the analyses. Other nodes, whilst not supported with significant bootstrap values, were consistent with 16S rRNA-derived trees, including the cluster of *Pirellula staleyi* and *Pirellula marina*. Consistent with these conclusions, for strains clustering within the *Gemmata* and *Isosphaera* groups, the ranking of similarity values between homologous regions of RNase P RNA sequence is identical to that found for analogous values for 16S rRNA. Of particular interest in relation to the *Pirellula* species for which a complete genome sequence has recently been published (Glöckner *et al.*, 2003), *Pirellula* sp. strain 1, is its clustering with the three clone sequences SM-A06, SM-A61 and SM-A88. These four sequences cluster separately from *Pirellula marina* and *Pirellula staleyi*, with high bootstrap confidence. Consistent with this is the shared occurrence of several inserts and nucleotide substitutions in members of this cluster. Evidence from RNase P RNA may thus be consistent with the separation of this strain from other species of *Pirellula* or

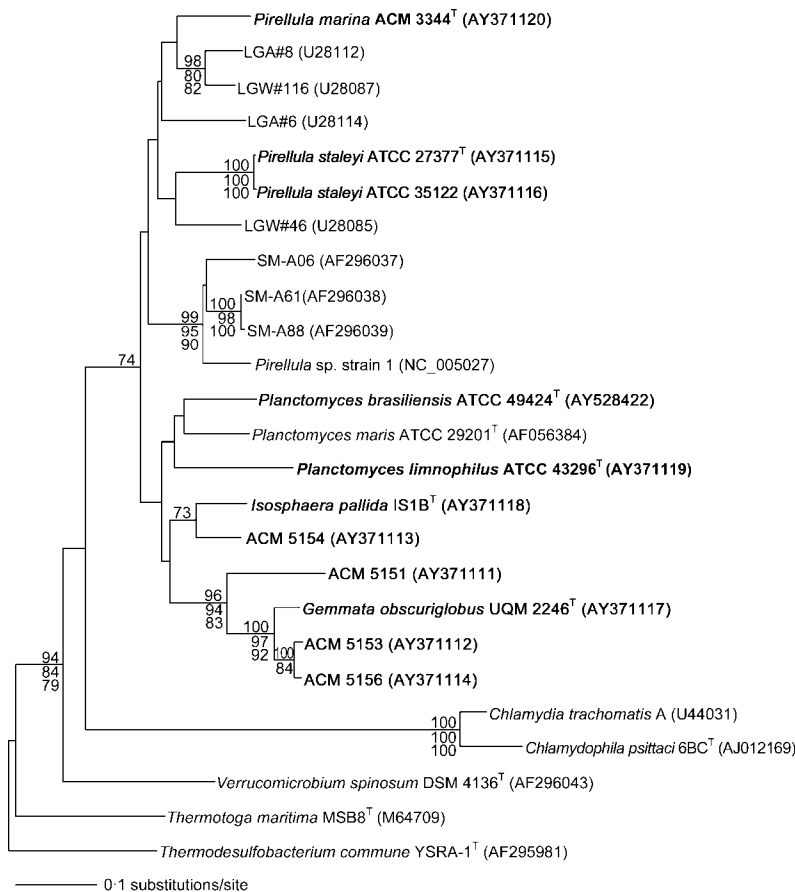


Fig. 4. Phylogenetic tree based on the aligned sequences of RNase P RNA showing members of the planctomycetes, compared with selected sequences. Newly sequenced planctomycetes from this study are shown in bold. The topology shown was obtained by the distance neighbour-joining method using the Kimura two-parameter model. Bootstrap values (obtained from 1000 resamples for distance and parsimony, 100 for maximum-likelihood) shown above the branch at individual nodes were derived from distance neighbour-joining analysis using Kimura two-parameter as a distance model and below the branch from maximum-parsimony analysis and maximum-likelihood analysis. Only those values with greater than 70% confidence are shown. The RNase P RNA sequence for *Pirellula* sp. strain 1 was derived from contig 22 of the database for this genome at NCBI. Scale bar, 10 substitutions per 100 nucleotides.

even from the genus *Pirellula*, as seems to be implied by the use of the tentative name '*Rhodopirellula baltica*' for this strain by the group generating the genome sequence (Glöckner *et al.*, 2003), and its formal proposal recently (Schlesner *et al.*, 2004).

The usefulness of RNase P RNA in differentiating planctomycetes is consistent with its utility for distinguishing members of the *Chlamydiales* (Herrmann *et al.*, 2000) and consistency between 16S rRNA phylogenies and those derived from RNase P RNA has previously been found in studies of the cyanobacteria (Vioque, 1997) and the *Chlamydiales* (Herrmann *et al.*, 2000). RNase P RNA gene sequences may prove useful for taxonomic investigations within planctomycete genera because of the relatively low similarity between homologous sequence regions relative to 16S rRNA, and the results of the present study are consistent with this view. The present study has been especially valuable for placing clone sequences from molecular ecology studies within the context of taxonomy of cultured planctomycete species, where this context was previously at low resolution only.

When a dataset including representatives of all phyla of domain *Bacteria* for which RNase P RNA sequences are available was used, phylogenetic analysis resulted in unstable tree topologies. This is not unexpected, since RNase P RNA is known to be most suitable for analyses at species or strain level because of its limited length and high variability (Pitulle *et al.*, 2002). The unstable behaviour of the large dataset is consistent with the view that RNase P RNA appears to be limited as a phylogenetic marker with respect to relationships between bacterial phyla. This effect may be reinforced by the loss of signal due to gaps introduced for alignment purposes, combined with the removal of gap positions.

An important group not represented in this study is the anammox planctomycetes, which occur only in bioreactors and not pure cultures so far, and differ from other planctomycetes in being chemoautotrophic and responsible for the process of anaerobic ammonium oxidation (a process unique to the anammox planctomycetes). The anammox genome projects now under way (e.g. <http://www.micrbiol.sci.kun.nl/tech/genomics.html>) will assist in revealing RNase P RNA gene sequences and one would expect even greater diversity among the planctomycete sequences to be revealed, since anammox organisms appear to be the deepest branching planctomycetes (Strous *et al.*, 1999).

Conclusion

In summary, features of the RNase P RNA of planctomycetes are consistent with the unique position of members of this phylum within the domain *Bacteria*. This is illustrated most dramatically by the unique insert in the P13 helix of *Gemmata* isolates. It is also illustrated by the replacement of the base A248, universally conserved in other *Bacteria* and *Archaea*, in *Pirellula* sp. strain 1 and

clone sequences clustering phylogenetically with it, and the occurrence of short tandem repeat sequences in the P12 helix of two *Gemmata*-like strains, otherwise only known to occur in the heterocyst-forming cyanobacteria. The phylogenetic analysis of RNase P RNA of the planctomycetes is also consistent with the distinctiveness of their RNase P RNA structures.

ACKNOWLEDGEMENTS

We wish to acknowledge the Australian Research Council for funding support of this research in the laboratory of J. A. F. and for support of a scholarship for M. K. B. We also thank R. Gutell, J. Brown and P. Hugenholtz for valuable discussion.

REFERENCES

- Altman, S. (1975). Biosynthesis of transfer RNA in *Escherichia coli*. *Cell* **4**, 21–29.
- Altman, S., Baer, M. F., Bartkiewicz, M., Gold, H., Guerrier-Takada, C., Kirsebom, L. A., Lumelsky, N. & Peck, K. (1989). Catalysis by the RNA subunit of RNase P – a minireview. *Gene* **82**, 63–64.
- Brochier, C. & Philippe, H. (2002). Phylogeny: a non-hyperthermophilic ancestor for Bacteria. *Nature* **417**, 244.
- Brown, J. W. (1999). The ribonuclease P database. *Nucleic Acids Res* **27**, 314.
- Brown, J. W. & Pace, N. R. (1991). Structure and evolution of ribonuclease P RNA. *Biochimie* **73**, 689–697.
- Brown, J. W., Nolan, J. M., Haas, E. S., Rubio, M. A. T., Major, F. & Pace, N. R. (1996). Comparative analysis of ribonuclease P RNA using gene sequences from natural microbial populations reveals tertiary structural elements. *Proc Natl Acad Sci U S A* **93**, 3001–3006.
- Butler, M. K., Wang, J., Webb, R. I. & Fuerst, J. A. (2002). Molecular and ultrastructural confirmation of classification of ATCC 35122 as a strain of *Pirellula staleyii*. *Int J Syst Evol Microbiol* **52**, 1663–1667.
- Christian, E. L., Kaye, N. M. & Harris, M. E. (2000). Helix P4 is a divalent metal ion binding site in the conserved core of the ribonuclease P ribozyme. *RNA* **6**, 511–519.
- Condon, C. & Putzer, H. (2002). The phylogenetic distribution of bacterial ribonucleases. *Nucleic Acids Res* **30**, 5339–5346.
- Derakshani, M., Lukow, T. & Liesack, W. (2001). Novel bacterial lineages at the (sub)division level as detected by signature nucleotide-targeted recovery of 16S rRNA genes from bulk soil and rice roots of flooded rice microcosms. *Appl Environ Microbiol* **67**, 623–631.
- Deutscher, M. P. (1984). Processing of tRNA in prokaryotes and eukaryotes. *CRC Crit Rev Biochem* **17**, 45–71.
- Di Giulio, M. (2003). The ancestor of the Bacteria domain was a hyperthermophile. *J Theor Biol* **224**, 277–283.
- Embley, T. M., Hirt, R. P. & Williams, D. M. (1994). Biodiversity at the molecular level: the domains, kingdoms and phyla of life. *Philos Trans R Soc London B Biol Sci* **345**, 21–33.
- Felsenstein, J. (1989). PHYLIP – phylogeny inference package (version 3.2). *Cladistics* **5**, 164–166.
- Frank, D. N. & Pace, N. R. (1998). Ribonuclease P: unity and diversity in a tRNA processing ribozyme. *Annu Rev Biochem* **67**, 153–180.

- Franzmann, P. D. & Skerman, V. B. D. (1984). *Gemmata obscuriglobus*, a new genus and species of the budding bacteria. *Antonie van Leeuwenhoek* **50**, 261–268.
- Fuerst, J. A. (1995). The planctomycetes: emerging models for microbial ecology, evolution and cell biology. *Microbiology* **141**, 1493–1506.
- Fuerst, J. A. & Webb, R. I. (1991). Membrane-bounded nucleoid in the eubacterium *Gemmata obscuriglobus*. *Proc Natl Acad Sci U S A* **88**, 8184–8188.
- Garrity, G. M., Winters, M. & Searles, D. B. (2001). Taxonomic outline of the procaryotic genera. In *Bergey's Manual of Systematic Bacteriology*, release 1.0, April 2001. New York: Springer. <http://dx.doi.org/10.1007/bergeysoutline>
- Giovannoni, S. J., Schabtach, E. & Castenholz, R. W. (1987). *Isosphaera pallida*, gen. and comb. nov., a gliding, budding eubacterium from hot springs. *Arch Microbiol* **147**, 276–284.
- Glöckner, F. O., Kube, M., Bauer, M. & 11 other authors (2003). Complete genome sequence of the marine planctomycete *Pirellula* sp. strain 1. *Proc Natl Acad Sci U S A* **100**, 8298–8303.
- Green, C. J., Rivera-León, R. & Vold, B. S. (1996). The catalytic core of RNase P. *Nucleic Acids Res* **24**, 1497–1503.
- Guerrier-Takada, C., Gardiner, K., Marsh, T., Pace, N. & Altman, S. (1983). The RNA moiety of ribonuclease P is the catalytic subunit of the enzyme. *Cell* **35**, 849–857.
- Haas, E. S. & Brown, J. W. (1998). Evolutionary variation in bacterial RNase P RNAs. *Nucleic Acids Res* **26**, 4093–4099.
- Haas, E. S., Brown, J. W., Pitulle, C. & Pace, N. R. (1994). Further perspective on the catalytic core and secondary structure of ribonuclease P RNA. *Proc Natl Acad Sci U S A* **91**, 2527–2531.
- Haas, E. S., Banta, A. B., Harris, J. K., Pace, N. R. & Brown, J. W. (1996). Structure and evolution of ribonuclease P RNA in Gram-positive bacteria. *Nucleic Acids Res* **24**, 4775–4782.
- Hancock, J. M. & Vogler, A. P. (2000). How slippage-derived sequences are incorporated into rRNA variable-region secondary structure: implications for phylogeny reconstruction. *Mol Phylogenet Evol* **14**, 336–374.
- Harris, M. E., Kazantsev, A. V., Chen, J.-L. & Pace, N. R. (1997). Analysis of the tertiary structure of the ribonuclease P ribozyme-substrate complex by site-specific photoaffinity crosslinking. *RNA* **3**, 561–576.
- Harris, J. K., Haas, E. S., Williams, D., Frank, D. N. & Brown, J. W. (2001). New insight into RNase P RNA structure from comparative analysis of the archaeal RNA. *RNA* **7**, 220–232.
- Herrmann, B., Pettersson, B., Everett, K. D. E., Mikkelsen, N. E. & Kirsebom, L. A. (2000). Characterization of the *rnpB* gene and RNase P RNA in the order *Chlamydiales*. *Int J Syst Evol Microbiol* **50**, 149–158.
- Hugenholtz, P., Goebel, B. M. & Pace, N. R. (1998). Impact of culture-independent studies on the emerging phylogenetic view of bacterial diversity. *J Bacteriol* **180**, 4765–4774.
- Jeffares, D. C., Poole, A. M. & Penny, D. (1998). Relics from the RNA world. *J Mol Evol* **46**, 18–36.
- Kirsebom, L. A. & Svärd, S. G. (1994). Base pairing between *Escherichia coli* RNase P RNA and its substrate. *EMBO J* **13**, 4870–4876.
- Kuypers, M. M. M., Sliemers, A. O., Lavik, G., Schmid, M., Jørgensen, B. B., Kuenen, J. G., Sinninghe Damsté, J. S., Strous, M. & Jetten, M. S. M. (2003). Anaerobic ammonium oxidation by anammox bacteria in the Black Sea. *Nature* **422**, 608–611.
- Liesack, W. & Stackebrandt, E. (1992). Occurrence of novel groups of the domain *Bacteria* as revealed by analysis of genetic material isolated from an Australian terrestrial environment. *J Bacteriol* **174**, 5072–5078.
- Liesack, W., Söller, R., Stewart, T., Haas, H., Giovannoni, S. & Stackebrandt, E. (1992). The influence of tachytelically (rapidly) evolving sequences on the topology of phylogenetic trees – intrafamily relationships and the phylogenetic position of *Planctomycetaceae* as revealed by comparative analysis of 16S ribosomal RNA sequences. *Syst Appl Microbiol* **15**, 357–362.
- Lindsay, M. R., Webb, R. I. & Fuerst, J. A. (1997). Pirellulosomes: a new type of membrane-bounded cell compartment in planctomycete bacteria of the genus *Pirellula*. *Microbiology* **143**, 739–748.
- Lindsay, M. R., Webb, R. I., Strous, M., Jetten, M. S. M., Butler, M. K., Forde, R. J. & Fuerst, J. A. (2001). Cell compartmentalisation in planctomycetes: novel types of structural organisation for the bacterial cell. *Arch Microbiol* **175**, 413–429.
- Maeda, T., Furushita, M., Hamamura, K. & Shiba, T. (2001). Structures of ribonuclease P RNAs of *Vibrio* core species. *FEMS Microbiol Lett* **198**, 141–146.
- Massire, C., Jaeger, L. & Westhof, E. (1997). Phylogenetic evidence for a new tertiary interaction in bacterial RNase P RNAs. *RNA* **3**, 553–556.
- Massire, C., Jaeger, L. & Westhof, E. (1998). Derivation of the three-dimensional architecture of bacterial ribonuclease P RNAs from comparative sequence analysis. *J Mol Biol* **279**, 773–793.
- Nicholas, K. B. & Nicholas, H. B., Jr (1997). GeneDoc: a tool for editing and annotating multiple sequence alignments. Distributed by the authors: <http://www.psc.edu/biomed/genedoc>
- Oh, B.-K. & Pace, N. R. (1994). Interaction of the 3'-end of tRNA with ribonuclease P RNA. *Nucleic Acids Res* **22**, 4087–4094.
- Pace, N. R. & Brown, J. W. (1995). Evolutionary perspective on the structure and function of ribonuclease P, a ribozyme. *J Bacteriol* **177**, 1919–1928.
- Pitulle, C., Strehse, C., Brown, J. W. & Breitschwerdt, E. B. (2002). Investigation of the phylogenetic relationships within the genus *Bartonella* based on comparative sequence analysis of the *rnpB* gene, 16S rDNA and 23S rDNA. *Int J Syst Evol Microbiol* **52**, 2075–2080.
- Schlesner, H. (1994). The development of media suitable for the microorganisms morphologically resembling *Planctomyces* spp., *Pirellula* spp., and other *Planctomycetales* from various aquatic habitats using dilute media. *Syst Appl Microbiol* **17**, 135–145.
- Schlesner, H. & Stackebrandt, E. (1986). Assignment of the genera *Planctomyces* and *Pirella* to a new family *Planctomycetaceae* fam. nov. and description of the order *Planctomycetales* ord. nov. *Syst Appl Microbiol* **8**, 174–176.
- Schlesner, H., Rendsmann, C., Tindall, B. J., Gade, D., Rabus, R., Pfeiffer, S. & Hirsch, P. (2004). Taxonomic heterogeneity within the *Planctomycetales* as derived by DNA–DNA hybridization, description of *Rhodopirellula baltica* gen. nov., sp. nov., transfer of *Pirellula marina* to the genus *Blastopirellula* gen. nov. as *Blastopirellula marina* comb. nov., and emended description of the genus *Pirellula*. *Int J Syst Evol Microbiol* Papers in press, <http://dx.doi.org/10.1099/ijs.0.63113-0>
- Schmid, M., Twachtman, U., Klein, M., Strous, M., Juretschko, S., Jetten, M., Metzger, J. W., Schleifer, K.-H. & Wagner, M. (2000). Molecular evidence for genus level diversity of bacteria capable of catalyzing anaerobic ammonium oxidation. *Syst Appl Microbiol* **23**, 93–106.
- Schön, A., Fingerhut, C. & Hess, W. R. (2002). Conserved and variable domains within divergent RNase P RNA gene sequences of *Prochlorococcus* strains. *Int J Syst Evol Microbiol* **52**, 1383–1389.
- Sinninghe Damsté, J. S., Strous, M., Rijpstra, W. I. C., Hopmans, E. C., Genevasen, J. A. J., van Duin, A. C. T., van Niftrik, L. A. &

- Jetten, M. S. M. (2002).** Linearly concatenated cyclobutane lipids form a dense bacterial membrane. *Nature* **419**, 708–712.
- Strous, M., Fuerst, J. A., Kramer, E. H. M., Logemann, S., Muyzer, G., van de Pas-Schoonen, K. T., Webb, R., Kuenen, J. G. & Jetten, M. S. M. (1999).** Missing lithotroph identified as new planctomycete. *Nature* **400**, 446–449.
- Vergin, K. L., Urbach, E., Stein, J. L., DeLong, E. F., Lanoil, B. D. & Giovannoni, S. J. (1998).** Screening of a fosmid library of marine environmental genomic DNA fragments reveals four clones related to members of the order *Planctomycetales*. *Appl Environ Microbiol* **64**, 3075–3078.
- Vioque, A. (1997).** The RNase P RNA from cyanobacteria: short tandemly repeated repetitive (STRR) sequences are present within the RNase P RNA gene in heterocyst-forming cyanobacteria. *Nucleic Acids Res* **25**, 3471–3477.
- Walter, A. E., Turner, D. H., Kim, J., Lyttle, M. H., Müller, P., Matthews, D. H. & Zuker, M. (1994).** Coaxial stacking of helices enhances binding of oligoribonucleotides and improves predictions of RNA folding. *Proc Natl Acad Sci U S A* **91**, 9218–9222.
- Wang, J., Jenkins, C., Webb, R. I. & Fuerst, J. A. (2002).** Isolation of *Gemmata*-like and *Isosphaera*-like planctomycete bacteria from soil and freshwater. *Appl Environ Microbiol* **68**, 417–422.
- Woese, C. R., Stackebrandt, E., Macke, T. J. & Fox, G. E. (1985).** A phylogenetic definition of the major eubacterial taxa. *Syst Appl Microbiol* **6**, 143–151.
- Yoon, J.-H. & Park, Y.-H. (2000).** Comparative sequence analyses of the ribonuclease P (RNase P) RNA genes from LL-2,6-diaminopimelic acid-containing actinomycetes. *Int J Syst Evol Microbiol* **50**, 2021–2029.
- Zahler, N. H., Christian, E. L. & Harris, M. E. (2003).** Recognition of the 5' leader of pre-tRNA substrates by the active site of ribonuclease P. *RNA* **9**, 734–745.
- Zuker, M. (2003).** Mfold web server for nucleic acid folding and hybridisation prediction. *Nucleic Acids Res* **31**, 3406–3415.