5S rRNA sequences of representatives of the genera *Chlorobium*, *Prosthecochloris*, *Thermomicrobium*, *Cytophaga*, *Flavobacterium*, *Flexibacter* and *Saprospira* and a discussion of the evolution of eubacteria in general

HILDE VAN DEN EYNDE,¹ YVES VAN DE PEER,¹ JEROME PERRY² and RUPERT DE WACHTER¹*

¹ Departement Biochemie, Universiteit Antwerpen (UIA), Universiteitsplein 1, B-2610 Antwerpen, Belgium
² Department of Microbiology, North Carolina State University, Box 7615, Raleigh, NC 27695-7615, USA

(Received 31 July 1989; revised 24 October 1989; accepted 26 October 1989)

5S rRNA sequences were determined for the green sulphur bacteria *Chlorobium limicola*, *Chlorobium phaeobacteroides* and *Prosthecochloris aestuarii*, for *Thermomicrobium roseum*, which is a relative of the green non-sulphur bacteria, and for *Cytophaga aquatilis*, *Cytophaga heparina*, *Cytophaga johnsonae*, *Flavobacterium breve*, *Flexibacter sp.* and *Saprospira grandis*, organisms allotted to the phylum ‘Bacteroides–Cytophaga–Flavobacterium’ and relatives as determined by 16s rRNA analyses. By using a clustering algorithm a dendrogram was constructed from these sequences and from all other known eubacterial 5S RNA sequences. The dendrogram showed differences, as well as similarities, with respect to results obtained by 16S rRNA analyses. The 5S RNA sequences of green sulphur bacteria were closely related to one another, and to a cluster containing 5S RNA sequences from *Bacteroides* and its relatives, including *Cytophaga aquatilis*. 5S RNA sequences of all other representatives of the ‘Bacteroides–Cytophaga–Flavobacterium’ phylum as distinguished by 16S rRNA analysis failed to group with *Bacteroides* and related clusters. On the basis of 5S RNA sequences, *Thermomicrobium roseum* clustered with *Chloroflexus aurantiacus*, as was expected from 16S rRNA analysis.

Introduction

The order Chlorobiales (Gibbons & Murray, 1978) consists of the families Chlorobiaceae (green sulphur bacteria) and Chloroflexaceae (gliding, filamentous, non-sulphur green bacteria). The members of these families share a number of unique properties. For instance, the light-harvesting pigments are located in typical chlorosome structures (Cohen-Bazire et al., 1964; Staehelin et al., 1978), and members of both families contain monogalactosyl diglyceride, but not phosphatidylethanolamine in their lipids. Conversely, the families differ in numerous features, such as the structure of their photosynthetic reaction centres (Olson & Pierson, 1987) and their gliding ability. Furthermore, representatives of the Chlorobiaceae and Chloroflexaceae differ in their physiological and ecological properties. With the exception of *Chloroherpeton thalassium* (Gibson et al., 1984) and *Chloroherpeton limophilum* (Eichler & Pfennig, 1988), which demonstrate gliding motility, the Chlorobiaceae are non-motile, mesophilic, strictly anaerobic and obligately photo-autotrophic rods. In comparison, *Chloroflexus aurantiacus* comprises gliding, thermophilic, facultatively chemo- or phototrophic filaments (Pierson & Castenholz, 1974). Members of both families inhabit different ecological niches (Gorlenko, 1988). Moreover, analyses of 16S rRNA oligonucleotide catalogues (Gibson et al., 1985) and full sequences (Oyaizu et al., 1987; Woese, 1987) have indicated a lack of phylogenetic relationship between the two families (Gibson et al., 1985). Whereas the genera *Chlorobium* (Cb.), *Prosthecochloris* and *Chloroherpeton* form a tight phylogenetic unit, the thermophilic *Chloroflexus aurantiacus* is distantly related to two non-photosynthetic genera, the thermophilic *Thermomicrobium* and the mesophilic *Herpetosiphon* (Gibson et al., 1985; Oyaizu et al., 1987). Phenotypically, the latter arrangement is not immediately obvious (Oyaizu et al., 1987). Indeed, the possession of unusual long-chain diols (Pond et al., 1986) is the only phenotypic characteristic that links *Thermomicrobium roseum* to *Chloroflexus aurantiacus* (Stackebrandt et al., 1988a), whereas *Herpetosiphon giganteus* and *Chloroflexus aurantiacus* possess similar carotenoids (Kleining & Reichenbach, 1977). From 16S RNA analy-
ses, both families have been provisionally assigned the rank of 'phylum' (Woese et al., 1985). The lack of phylogenetic relationship between the two families is corroborated by lipopolysaccharide and peptidoglycan analyses (Jürgens et al., 1987; Meissner et al., 1987, 1988). The Chloroflexus–Thermomicrobium–Herpetosiphon phylum is one of the earliest branchings in eubacterial evolution (Oyaizu et al., 1987). As a consequence Chloroflexus, rather than the cyanobacteria, has been suggested to be the photosynthetic component of the stromatolites (Oyaizu et al., 1987). Another eubacterial phylum, termed ‘Bacteroides–Cytophaga–Flavobacterium’, made up of genera such as Bacteroides, Cytophaga (C.), Flexibacter, Saprospira, Spirocytophaga, Haliscenomonobacter, Fusobacterium and Flavobacterium (Flv.) (Paster et al., 1985; Weisburg et al., 1985) has been suggested to bear a somewhat closer relationship to the phylum of green sulphur bacteria than to any of the remaining eubacterial phyla (Woese, 1987).

In order to compare results obtained with 16S and 16S rRNA analyses, the SS RNA sequences of Ch. limicola, Ch. phaeobacteroides, Prosthecococaris aestuarii, Thermomicrobium roseum, Saprospira grandis, Flexibacter sp., Flc. breve, C. heparina and C. johnsonae were determined. Using weighted average pairgroup clustering, a phylogenetic tree was constructed from these sequences, together with some 282 other eubacterial SS RNA sequences, representative of eight of the ten major phyla distinguished by Woese (1987).

**Methods**

**Bacterial strains.** The following strains were used: Chlorobium limicola DSM 249, Ch. phaeobacteroides DSM 266T (T. type strain), Prosthecococaris aestuarii strain SK 413, Thermomicrobium roseum ATCC 27502T, Cytophaga aquatilis ATCC 29551T, C. heparina ATCC 13125T, C. johnsonae DSM 425, Flavobacterium breve ATCC 14234, Flexibacter sp. DSM 527 and Saprospira grandis ATCC 23119T. These were obtained from the American Type Culture Collection (ATCC), Rockville, Md, USA, and from the Deutsche Sammlung von Mikroorganismen (DSM), Braunschweig, FRG. Culture conditions were as follows. Ch. limicola, Ch. phaeobacteroides and P. aestuarii were grown using the method described by Pfennig & Trüper (1981). T. roseum was obtained from Dr Robert F. Ramsay, Department of Biochemistry, University of Nebraska, USA, and grown at 65 °C in a medium (pH 7.5) containing (1-1) tryptone (Difco), 1 g, yeast extract (Difco), 1 g, and mineral salts as described by Zarilla & Perry (1984). C. aquatilis, C. heparina and F. breve were grown aerobically at 26 °C in 8 g nutrient broth (Merck) l−1, pH 7.0. Flexibacter sp. was grown aerobically at 26 °C in medium (pH 7.2) containing (1-1) Casitone (Difco), 3 g, and CaCl2.2H2O, 1.36 g. The same culture conditions applied to C. johnsonae, with the exception that yeast extract (Difco; 1 g l−1) was added to the medium. S. grandis was grown according to Lewin (1972). Cultivated organisms were harvested by centrifugation and stored by freezing or freeze-drying until needed.

**Preparation of SS RNA.** Cells were ruptured using either glass beads and a cell mill (Willekens et al., 1986), or by grinding in a mortar with alumina (Van den Eynde et al., 1987). Sheared cells were then extracted with phenol (Willekens et al., 1986) in order to obtain the rRNA fraction. SS RNA was prepared by PAGE, as previously described (Fang et al., 1982). SS RNA was labelled at the 3'-terminus by cytidine 3',5-[5'-32P]bisphosphate (Amersham), purified on 8% (w/v) polyacrylamide gels, and subjected to partial chemical degradation (Peatie, 1979). Whenever necessary, partial enzymic degradation, as described previously (Dams et al., 1983), was used to confirm the nucleotide pattern. In the case of length heterogeneity, components of different chain length were sequenced separately. With Ch. limicola, Ch. phaeobacteroides, P. aestuarii, C. johnsonae, C. heparina, Fc. breve, S. grandis, Flexibacter sp. and T. roseum, electrophoresis of the partial degradation products on 8–20% (w/v) polyacrylamide gels, kept at a constant temperature of 64 °C, proved sufficient to elucidate the entire sequence. In the case of C. aquaticus, the identity of the 5'-terminal nucleotide was confirmed by gel electrophoresis of enzymically degraded (Ap)nA-extended, 5'-terminally labelled (Dams et al., 1983) SS RNA.

**Data analysis.** Dendrograms were constructed by using a weighted pairgroup clustering algorithm, as described previously (Huysmans & De Wachter, 1986; Dams et al., 1987). In short, a dissimilarity matrix that takes into account multiple and back mutations serves as the input for the tree-construction algorithm. A correction, to allow for the effects of unequal evolutionary rates along different eubacterial lineages, is incorporated. The standard deviations on the substitution terms of the computed dissimilarities are indicated by error flags on all branching points of the resulting dendrogram.

**Results and Discussion**

Part of an alignment in which the newly determined sequences were aligned with approximately 600 SS RNA sequences of eubacterial, archaeobacterial, eukaryotic and organelle nature is included in Fig. 1. The complete alignment is available from the authors. Boxes superimposed on the alignment indicate putative double-stranded areas, according to a well-established secondary structure proposal (De Wachter et al., 1982). Starting from an alignment containing the 303 eubacterial SS RNA sequences and supplemented with 10 cytoplasmic SS RNA sequences from red algae to permit correction of unequal evolutionary rates, a dendrogram was constructed (Fig. 2). The majority of the sequences included are found in the biannual compilation of Wolters & Erdmann (1988). The sequences are representative of eight of the ten major eubacterial phyla, as determined by 16S RNA analysis (Woese, 1987). When comparing the pattern deduced from SS RNA sequence information with that based on 16S RNA analyses, a number of similarities and discrepancies become apparent. For example, no relationship was detected between plant mitochondrial SS RNA sequences and the alpha subdivision of the recently installed (Stackebrandt et al., 1988b; formerly the purple bacteria and their relatives) class Proteobacteria. This might be an aberration of the tree-construction algorithm, due to the extremely fast rate at which mitochondrial genomes appear to be evolving.
1. Cytophaga heparina
2. Cytophaga johnsonae
3. Cytophaga aquatilis
4. Flavobacterium breve
5. Flexibacter sp.
6. Saprospira grandis
7. Thermomicrobium roseum
8. Chlorobium phaeobacteroides
9. Chlorobium limicola
10. Prosthecococcus aestuarii

Fig. 1. Alignment of 5S rRNA sequences determined in the present study. Blank alignment positions are necessary to accommodate nucleotides of 5S RNAs from other taxonomic groups. Length heterogeneity is indicated by terminal residues printed in lower-case letters. Boxes labelled A and A', B and B', and so on, enclose the complementary strands of helices A, B, etc., in the secondary-structure model of De Wachter et al. (1982). Bulges within helices are enclosed in nested boxes; bases belonging to odd base pairs (pairs other than A-U, G-C and G-U), intercalated between Watson–Crick and/or G-U pairs, are in parentheses.
Cyanobacteria and plastids (Van den Eynde et al., 1988) are found as a monophyletic group, in accordance with the picture based on 16S RNA data. With 5S RNA, however, the cluster appears to branch off earlier in time than with 16S RNA. Furthermore, *Campylobacter* spp., on the basis of their 5S RNA sequence, exhibit no specific relationship to any of the eubacterial clusters (Fig. 2). Conversely, on the basis of 16S RNA sequences, a slight relationship of *Campylobacter* to the class Proteobacteria has been proposed (Lau et al., 1987).

The cluster that branches off next in Fig. 2 contains the majority of the 5S RNA sequences determined here. The cluster is defined by sequences from the *Bacteroides-Cytophaga-Flavobacterium* group of organisms (Paster et al., 1985; Weisburg et al., 1985), and by sequences from green sulphur organisms (Gibson et al., 1985).
Eubacterial 5S RNA sequences and evolution

Fig. 3. (a) Detailed structure of the cluster containing the majority of the sequences determined in this study. Except for the sequences reported here, sequences were taken from Wolters & Erdmann (1988) or are unpublished (N. R. Pace, personal communication). (b) Detailed structure of the cluster termed 'miscellaneous' in Fig. 2. Except for the sequences reported here, sequences were taken from Wolters & Erdmann (1988), or are unpublished so far (H. Van den Eynde and others).

1985). A detailed picture of the cluster is presented in Fig. 3(a). In the cluster, 5S RNA sequences of Chlorobium and Prosthecochloris spp. are found closely related to one another, as was expected from 16S RNA analyses (Gibson et al., 1985). The 5S RNA sequences of Ch. limicola DSM 249, Ch. phaeobacteroides and Ch. vibrioforme (N. R. Pace, unpublished results), and of Ch. limicola (Chumakov, 1987) are closely related.

The 5S RNA sequence of Prosthecochloris aestuarii was recovered next. As a whole, the green sulphur bacteria are associated with a cluster containing Bacteroides fragilis, B. thetaiotaomicron, Porphyromonas gingivalis and B. veroralis. This tendency towards a grouping of Bacteroides and relatives with green sulphur bacteria has been observed previously, on the basis of both 16S RNA (Woese, 1987) and 5S RNA (Van den Eynde et al., 1989) analyses. With 16S RNA, however, the genus Fusobacterium is included in the 'Bacteroides'-like cluster, whereas with 5S RNA Fusobacterium is only peripherally related to the Bacteroides–green sulphur bacteria assemblage (Van den Eynde et al., 1989). Furthermore, with 16S RNA sequences, Flavobacterium breve, Flexibacter sp., S. grandis, C. heparina and C. johnsonae are recovered with the genus Bacteroides to form the so-called 'Bacteroides–Cytophaga–Flavobacterium' phylum. That this is not the case with 5S RNA is documented below. The 5S RNAs
of *Streptococcus lactis* subsp. *cremoris* and of *Mycoplasma hyopneumoniae* were recovered, atypically, close to the *Fusobacterium* cluster, whereas they were expected to group with mycoplasmata and lactobacilli 16S RNA sequences.

The 5S RNA sequences from two out of three representative species of the phylum of green non-sulphur bacteria and relatives, namely *Chloroflexus aurantiacus* (D. A. Stahl, unpublished results) and *Thermomicrobium roseum*, were found (Fig. 2) to bear a specific relationship to one another, in accordance with the picture derived from 16S RNA analyses (Oyaizu et al., 1987). The suggested ancient nature of the phylum of green non-sulphur bacteria and relatives (Oyaizu et al., 1987) was not corroborated by 5S RNA analysis. In fact, both the cyanobacteria—plastid cluster and the cluster containing the green sulphur bacteria branched off earlier than the *Chloroflexus* cluster (Fig. 2). The results presented in this paper thus fail to affirm the proposal (Oyaizu et al., 1987) of *Chloroflexus* being the photosynthetic component of the stromatolites. The 5S RNA of ‘*Herpetosiphon* strain’ Senghas Wie 2 (Van den Eynde et al., 1987), originating from exactly the same strain (Senghas & Lingens, 1985) as was used for 16S RNA analysis, was not, contrary to 16S RNA results, recovered in close association with *Chloroflexus* and *Thermomicrobium* 5S RNA. Instead, the ‘*Herpetosiphon* strain’ Senghas Wie 2 sequence grouped with a cluster termed ‘miscellaneous’ in Fig. 2, and presented in detail in Fig. 3(b). Apart from the ‘*Herpetosiphon* strain’ Senghas Wie 2 sequence and a number of sequences from low G + C content Firmicuta, sequences from organisms belonging to the Proteobacteria delta group, from *C. hepatica* and from *Flexibacter* sp. were scattered throughout this cluster. Thus, regarding the assignment of strain Senghas Wie 2 to the phylum of green sulphur bacteria and their relatives (Senghas & Lingens, 1985), and the connected suggestion for an assignment of the strain to the genus *Herpetosiphon* (Gibson et al., 1985; Woese et al., 1985), we urge caution. In any case, in the original paper of Senghas & Lingens (1985), the strain was set well apart from the genus *Herpetosiphon* on the basis of physiological differences and of differences in fatty acid composition and in DNA G + C content.

As for the Proteobacteria, in agreement with the findings based on 16S RNA, the subdivisions alpha, beta and gamma were recovered intact, and in close connection with one another. However, with 5S RNA analysis, the alpha subdivision was severed from the beta and gamma clusters by a group containing radio-resistant organisms such as *Deinococcus* and their relatives. We can see no evidence for the existence of a fourth subdivision, delta, of the class Proteobacteria (Woese, 1987; Stackebrandt et al., 1988b), since organisms of the delta subgroup (H. Van den Eynde and others, unpublished) that were examined had 5S RNA sequences that grouped with, or, in the case of the *Sorangium cellulosum* 5S RNA, close to the miscellaneous cluster (Figs 2 and 3b).

Rather more in accordance with results based on 16S RNA sequencing is the occurrence in Fig. 2 of a cluster harbouring the genera *Planctomyces*, *Pirellula*, *Gemmatia* and *Isosphaera*, which all lack peptidoglycan.

The phylum of Gram-positive organisms, or Firmicuta, a monophyletic grouping with four subdivisions according to 16S RNA analysis (Woese, 1987), showed a different arrangement on the basis of 5S RNA analysis, as can be seen from Fig. 2. The two most important subdivisions, namely Firmicuta with low, and Firmicuta with high DNA G + C contents, have a large number of representatives in the 5S RNA data bank. Results from 16S RNA analysis were corroborated to a certain extent by 5S RNA sequence data, although some of the 5S RNA sequences of Gram-positive organisms were recovered in unexpected places in the tree. Examples are the occurrence of some sequences of clostridia and bacilli, Firmicuta of low DNA G + C content, in the aforementioned ‘miscellaneous’ cluster (Fig. 3b), a cluster more closely related to the Firmicuta with high rather than to those with low DNA G + C content. Also unexpected, was the occurrence of the 5S RNA of *Flv. breve* and of *Saprospira grandis* among the Firmicuta with high DNA G + C content, and the occurrence of *Cytophaga johnsonae, Anaerorhabdus furcosus* (Van den Eynde et al., 1989) and *Verrucomicrobium spinosum* (Bomar & Stackebrandt, 1987) 5S RNAs with those of Firmicuta with a low genomic G + C content. 16S RNA studies have previously shown the relatedness of the Gram-negative genera *Selenomonas*, *Sporomusa* and *Megasphaera*, to the Firmicuta and relatives phylum (Stackebrandt et al., 1985). 16S RNA analysis showed these genera to be closely related, and to constitute one of the four subdivisions of the phylum (Woese, 1987). In contrast, several 5S RNA sequences from Gram-negative species, were scattered throughout the low and high G + C content subdivisions of the Gram-positive cluster, a clear discrepancy with results obtained by 16S RNA analysis. Moreover, since we chose exactly the same strains as were used for 16S RNA analysis (Paster et al., 1985), in the case of *C. johnsonae*, *C. hepatica*, *Flv. breve* and *S. grandis* 5S RNA, an association with the ‘*Bacteroides–Cytophaga–Flavobacterium*’ cluster, and not with the Gram-positive organisms, was most strongly expected. That this was not the case makes the observed discrepancy even more unusual. Moreover, 5S RNA analysis did not corroborate the ancient line of descent proposed for *Verrucomicrobium spinosum* by Albrecht et al. (1987).
We cannot offer a clearcut explanation for the differences revealed by analyses of 5S and 16S RNA sequences. The nature of the 5S RNA molecule, as opposed to the 16S RNA molecule, cannot, in our opinion be held responsible. Although 5S RNA is much smaller than 16S RNA (an average of 120 nucleotides for 5S RNA against an average of about 1550 nucleotides for eubacterial 16S RNA), 5S RNA is far more conserved than 16S RNA and has had to suffer less from extensive insertion- and/or deletion events. The alignment procedure with 5S RNA is thus much less prone to error than is the case with 16S RNA. Furthermore, it is not solely discrepancies that are encountered when results from 5S RNA sequence analysis are compared to those obtained using 16S RNA: many features of the 16S RNA evolutionary picture corroborated fairly well by analyses of 5S RNA sequences. The part played by unequal eubacterial evolutionary rates, is thought to be limited due to the application of a correction procedure for this phenomenon. In view of the conflicting results obtained with 5S RNA and 16S RNA, we believe that a renewed and thorough investigation of the phylogenetic interrelationships of the organisms assigned to the 'Bacteroides--Cytophaga--Flavobacterium' is needed. Questions can also be raised regarding the extent to which the use of different tree- construction algorithms influences the inferred evolutionary picture. The differing results obtained using 16S RNA and 5S RNA as molecular tools for studying eubacterial evolution suggests that the use of additional molecular clocks, such as 23S rRNA, may be required in order to arrive at a consensus picture of eubacterial evolution.

We thank Dr N. Pfennig for providing cells of Chlorobium limicola, Chlorobium phaeobacteroides and Prosthecochloris aestuarii, and Drs N. R. Pace and D. A. Stahl for providing unpublished data. Our research was subsidized in part by a grant from the 'Fonds voor Kollektief Fundamenteel Onderzoek'. Y. Vdp is holder of a grant from the 'Instituut tot aanmoediging van het Wetenschappelijk Onderzoek In Nijverheid en Landbouw'.

References


