JOURNAL OF BACTERIOLOGY, Aug. 1994, p. 4597–4609
0021-9193/94/$04.00+0

Vol. 176, No. 15

Phylogenetic and Molecular Characterization of a 23S rRNA Gene Positions the Genus *Campylobacter* in the Epsilon Subdivision of the *Proteobacteria* and Shows that the Presence of Transcribed Spacers Is Common in *Campylobacter* spp.†

TREVOR J. TRUST,[1]* SUSAN M. LOGAN,[1] CORINNE E. GUSTAFSON,[1] PAUL J. ROMANIUK,[1] NAM W. KIM,[2] VOON L. CHAN,[2] MARK A. RAGAN,[3] PATRICIA GUERRY,[4] AND ROBIN R. GUTELL[5]

Department of Biochemistry and Microbiology, University of Victoria, Victoria, British Columbia V8W 3P6,[1] Department of Microbiology, University of Toronto, Toronto, Ontario M5S 1A8,[2] and Institute for Marine Biosciences, National Research Council of Canada, Halifax, Nova Scotia B3H 3Z1,[3] Canada; Enteric Diseases Program, Naval Medical Research Institute, Bethesda, Maryland 20814[4]; and MCD Biology, University of Colorado, Boulder, Colorado 80309[5]

Received 28 January 1994/Accepted 21 May 1994

The nucleotide sequence of a 23S rRNA gene of *Campylobacter coli* VC167 was determined. The primary sequence of the *C. coli* 23S rRNA was deduced, and a secondary-structure model was constructed. Comparison with *Escherichia coli* 23S rRNA showed a major difference in the *C. coli* rRNA at approximately position 1170 (*E. coli* numbering) in the form of an extra sequence block approximately 147 bp long. PCR analysis of 31 other strains of *C. coli* and *C. jejuni* showed that 69% carried a transcribed spacer of either ca. 147 or ca. 37 bp. Comparison of all sequenced *Campylobacter* transcribed spacers showed that the *Campylobacter* inserts were related in sequence and percent G+C content. All *Campylobacter* strains carrying transcribed spacers in their 23S rRNA genes produced fragmented 23S rRNAs. Other strains which produced unfragmented 23S rRNAs did not appear to carry transcribed spacers at this position in their 23S rRNA genes. At the 1850 region (*E. coli* numbering), *Campylobacter* 23S rRNA displayed a base pairing signature most like that of the beta and gamma subdivisions of the class *Proteobacteria*, but in the 270 region, *Campylobacter* 23S rRNA displayed a helix signature which distinguished it from the alpha, beta, and gamma subdivisions. Phylogenetic analysis comparing *C. coli* VC167 23S rRNA and a *C. jejuni* TGH9011 (ATCC 43431) 23S rRNA with 53 other completely sequenced (eu)bacterial 23S rRNAs showed that the two campylobacters form a sister group to the alpha, beta, and gamma proteobacterial 23S rRNAs, a positioning consistent with the idea that the genus *Campylobacter* belongs to the epsilon subdivision of the class *Proteobacteria*.

The members of the *Campylobacter* group of organisms are important pathogens in human and veterinary medicine. This group includes *Campylobacter*, *Arcobacter*, *Helicobacter*, *Wolinella*, and "*Flexispira*" spp. and has been argued to constitute a separate bacterial lineage within the class *Proteobacteria* (62, 63). The most recently published phylogenetic tree based on small-subunit rRNA sequences has the group positioned in the delta and epsilon subdivisions of the *Proteobacteria* (purple bacteria) (39). There are 11 *Campylobacter* species (9, 56), and the roles of several of these species in diseases of humans and animals have been well described (43). Typically, these small, nutritionally fastidious, microaerophilic, gram-negative spiral bacteria colonize moist body surfaces, including the human gingival cavity, the small intestinal mucosa, and the vaginal mucosa of bovines and this colonization leads to disease. The two most important human pathogens in the genus are *Campylobacter jejuni* and *C. coli*, which cause enteritis.

In keeping with their small size and limited nutritional properties, the campylobacters have a small genome (23, 37, 57). At 1.7 to 1.8 Mb, the *C. jejuni* and *C. coli* genomes are only 36% of the size of the *Escherichia coli* chromosome (53) and are also slightly smaller than the genomes of such organisms as *Streptococcus sanguis* (2.3 Mb) (2) and *Haemophilus influenzae* (2.3 Mb) (4). Consistent with their small genome size and their slow growth rate, the campylobacters also appear to have a reduced number of transcriptional units or operons encoding their rRNAs. While *E. coli* has seven rRNA genomic loci (26), *C. jejuni* has only three copies of the 16S and 23S rRNA genes (23, 24, 27, 37, 57). Unlike the characteristic continuous 16S-23S (eu)bacterial rRNA operon structure seen in *E. coli* (35), several studies have suggested that in at least one of the three *C. jejuni* rRNA operons, the 16S and 23S genes are located separately (23, 27, 37, 57). However, recent fine mapping of the three rRNA gene loci of one of these *C. jejuni* strains has shown that all three operons possess a contiguous 16S-23S structure (25). The sequence and organization of one of these contiguous rRNA operons has recently been described (GenBank accession number Z29326).

Of the *Campylobacter* rRNAs, the 16S species has received the most attention. For example, partial 16S rRNA sequences have been used in a number of phylogenetic studies of the *Campylobacter* group of organisms (21, 30, 42, 50, 54, 58) and in the development of rapid identification methodology (48, 49, 63). The 16S rRNA genes of *C. jejuni*, *C. coli*, and a number of other campylobacters have also been sequenced (7, 9, 55, 63). The sequence of a 5S rRNA gene of *C. jejuni* has also been reported (30). In contrast, the *Campylobacter* 23S rRNA species has received little attention. Recently, however, partial 23S rRNA gene sequences, together with complete 16S rRNA

* Corresponding author. Phone: (604) 721-7079. Fax: (604) 721 8855. Electronic mail address: ttrust@sol.uvic.ca.

† National Research Council of Canada report no. 34899.

sequences, were used to investigate the phylogeny of the genus *Campylobacter* (60) and PCR amplification of 23S rRNA gene fragments was used to discriminate among thermophilic *Campylobacter* species (10). In several of the campylobacters tested, internal transcribed spacers were found in the 23S rRNA genes yielding fragmented 23S rRNA. In this study, we cloned and sequenced a gene encoding 23S rRNA from *C. coli* VC167, and here we report the sequence of that gene and the deduced primary and secondary structures of the 23S rRNA encoded by the gene. We show that the gene carries a transcribed spacer of approximately 147 bp which results in the production of fragmented 23S rRNA, and we demonstrate that this is a common phenomenon in *C. coli* and *C. jejuni* because many strains carry transcribed spacers of this approximate length or shorter ones and also produce fragmented 23S rRNAs. In addition, we report on the phylogenetic position of the genus *Campylobacter*, based on the presence of phylogenetic signatures at positions 270 and 1850 (*E. coli* numbering [36]) of the *Campylobacter* 23S rRNA, and on a phylogenetic analysis which was performed by using this *C. coli* 23S rRNA, a *C. jejuni* 23S rRNA sequence (GenBank accession number Z29326), and 53 other available complete (eu)bacterial 23S rRNA sequences.

MATERIALS AND METHODS

Bacterial strains and growth conditions. The *Campylobacter* strains used in this study were from the culture collection of T. J. Trust and are listed in Table 1. Campylobacters were grown at 37°C for 24 h on Mueller-Hinton agar (Difco Laboratories, Detroit, Mich.) under an atmosphere containing 5% oxygen and 10% carbon dioxide. *E. coli* JM109 was used for transfections with M13 phage. *E. coli* DH5 was the host for plasmid cloning experiments and for control 23S rRNA size comparisons. *E. coli* was grown aerobically at 37°C for 24 h on Mueller-Hinton agar.

DNA methods. Total DNA extractions from *C. coli* were done by the method of Hull et al. (22). Plasmid DNAs were purified as previously described (31). Restriction enzymes were purchased from Boehringer Mannheim Biochemicals (Indianapolis, Ind.) or Pharmacia (Uppsala, Sweden) and used under the conditions recommended by the supplier. Sequencing of DNA containing the 23S rRNA gene of *C. coli* VC167 was performed manually with M13 single-stranded template DNA or alkali-denatured double-stranded plasmid DNA by using ^{35}S-dATP (Dupont Research Products, Boston, Mass.) and Sequenase (United States Biochemical Corp., Cleveland, Ohio) as specified by the manufacturers. The templates used were as indicated in Fig. 1 or in Results. The primers used were either the commercially available universal primer or custom primers synthesized at approximately 200- to 250-bp intervals on both strands. On other occasions, sequencing reactions were carried out with dye-labeled terminators (Applied Biosystems, Inc. [ABI], Foster City, Calif.) and a Perkin Elmer Cetus model 480 thermal cycler by following the protocol and cycling parameters suggested by ABI. Results of completed sequencing reactions were electrophoresed on a 6% acrylamide–8.3 M urea sequencing gel for 14 h at 2,200 V with an ABI Automated DNA Sequencer. Sequence data were analyzed with the ABI SeqEd program. Synthesis of oligonucleotides was done with an ABI 391 PCRMate oligonucleotide synthesizer.

Oligonucleotide hybridizations. Oligonucleotides were labeled with ATP as previously described (31). Southern blotting and colony blotting were performed with nitrocellulose as described by Sambrook et al. (51). Hybridization was performed at 37°C in 6× SSC (1× SSC is 0.15 M NaCl plus 0.015 M sodium citrate)–20 mM $NaPO_4$–2 mM EDTA–5× Denhardt's solution–0.1× sodium dodecyl sulfate–100 mg of calf thymus DNA per ml. Wash conditions were 37°C in 0.5× SSC for probe pg8 and Romaniuk probe D (49) and 1× SSC for Romaniuk probe G (49).

TABLE 1. Occurrence of transcribed spacers in 23S RNA genes and fragmented 23S rRNAs of *C. coli* and *C. jejuni*

Strain	LIO serogroup	PCR product length (bp)	Insert	23S RNA cleavage
C. coli VC20	8	310	+	Yes
C. coli VC143	8	310	+	Yes
C. coli VC144	8	310	+	Yes
C. jejuni VC156	8	310	+	Yes
C. jejuni VC159	8	310	+	Yes
C. coli VC167	8	310	+	Yes
C. coli VC189	8	310	+	Yes
C. jejuni VC152	8/29	310	+	Yes
C. jejuni VC157	8/29	167	−	No
C. coli VC236	29	310	+	Yes
C. coli VC168	29	167	−	No
C. coli VC234	29/55	167	−	No
C. coli VC235	55	167	−	No
C. jejuni VC87	1	167	−	No
C. jejuni VC83	4	167	−	No
C. jejuni VC207	4	167	−	No
C. jejuni VC88	5	310	+	Yes
C. jejuni 81176	5	310	+	Yes
C. jejuni VC84	6	200	+	Yes
C. jejuni VC209	6	200	+	Yes
C. jejuni VC222	6	200	+	Yes
C. jejuni 81116	6	200	+	Yes
C. jejuni VC74	11	167	−	No
C. jejuni VC91	11	310	+	Yes
C. jejuni VC228	11	167	−	No
C. coli VC92	12	310	+	Yes
C. coli VC229	12	310	+	Yes
C. jejuni VC94	15	200	+	Yes
C. jejuni VC227	15	200	+	Yes
C. jejuni VC185	15	167	−	No
C. jejuni VC220	15	200	+	Yes
C. jejuni VC104	19	200	+	Yes

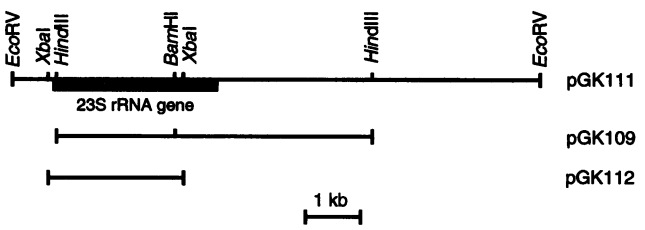


FIG. 1. Schematic representation of plasmid pGK111, which carries *C. coli* VC167 DNA containing a 23S rRNA gene cloned into the *Eco*RV site of pBR322. The *Hin*dIII fragment was subcloned into pBR322 to yield pGK109. For DNA sequence analysis of this DNA, the two *Hin*dIII-*Bam*HI fragments were further subcloned into M13mp18 and M13mp19. The *Xba*I fragment from pGK111 was subcloned into pUC19 to give pGK112, and this plasmid was used as a double-stranded template for sequencing.

PCR amplification of DNA. PCR amplification of chromosomal DNA was performed with capillary tubes in a 1605 Air Thermo-Cycler from Idaho Technology (Idaho Falls, Idaho) essentially as described by Gustafson et al. (14). The samples were subjected to 35 cycles of 1 s of denaturation at 94°C, 1 s of primer annealing at 51°C, and extension at 71°C for 40 s. PCR-generated DNA products were detected by gel electrophoresis on 1.0% agarose. DNA fragments were visualized by UV fluorescence after ethidium bromide staining. A PCR control was routinely included with every primer pair. This control reaction contained water instead of DNA to detect unwanted DNA contaminants.

RNA extraction and gel electrophoresis. RNA was prepared as described by Burgin et al. (3). Cells grown for 24 h were washed off Mueller-Hinton agar plates with Tris HCl (50 mM, pH 8.0). Following centrifugation at 4°C, the cell pellet was resuspended in 3 ml of Tris HCl (50 mM pH 8.0)–2 mM $MgCl_2$–1 mg of lysozyme per ml–10 μg of RNase-free DNase 1. The suspension was frozen in a dry ice bath and thawed at 37°C for 1 min a total of three times. Sodium dodecyl sulfate was added to a final concentration of 1%, and EDTA was added to a concentration of 20 mM. The lysate was then extracted with phenol (equilibrated in 10 mM Tris HCl [pH 8.0]–1 mM EDTA–100 mM NaCl) four times, and the RNA was recovered by ethanol precipitation of the aqueous phase. RNA species were analyzed by electrophoresis through 1.2% agarose–1% formaldehyde gels (45) and visualized by ethidium bromide staining.

Secondary-structure and phylogenetic analyses. The secondary structure was drawn in a format based on that of the *E. coli* 23S rRNA model, which may be regarded as the reference standard (18, 36). The methods used to fold this secondary structure have been discussed previously (15–17). For phylogenetic analysis, the 23S rRNAs of *C. coli* VC167 and *C. jejuni* TGH9011 (ATCC 43431) (GenBank accession number Z29326) were aligned with 53 other completely sequenced (eu)bacterial 23S rRNAs on the basis of secondary and higher-order structures, yielding a full matrix of 55 species by 3,180 positions. From this, ambiguously aligned regions and sparsely populated columns were removed to yield a more conservative matrix of 55 species by 2,409 positions. For parsimony analysis, a variant of the latter was used in which fourth and subsequent gap positions were coded as unknown data (?) to reduce bias arising from long, shared gaps. Trees were inferred by using PHYLIP 3.5c (11) on Sun 10/30 and 10/51 workstations, with global rearrangements (G option) and a randomized order of sequence input (J option). Distances were calculated under Felsenstein's maximum-likelihood model of sequence change. Fitch-Margoliash distance, neighbor-joining, and parsimony trees were bootstrapped 100 times. Maximum-likelihood analysis was done with fastDNAml (38) with 53-level rearrangement and randomized sequence input but, for reasons of computational complexity, was not bootstrapped. In all analyses, *Thermotoga maritima* 23S rRNA was set as the out-group. Trees were displayed by using TreeTool, version 1.0 (31a; available from the Ribosomal Database Project [29]).

Nucleotide sequence accession number. The nucleotide sequence of the *C. coli* VC167 rRNA gene has been deposited in GenBank under accession number U09611.

RESULTS

Cloning and sequencing of the *C. coli* VC167 23S rRNA gene. Construction of plasmid pGK109, which contained most of a 23S gene from *C. coli* VC167, as shown in Fig. 1, has already been described (13). To clone the complete 23S gene, an oligonucleotide termed pg8 (5′-TTACCTATATAGGAGCGA-3′) was synthesized and was shown by Southern blot analysis to hybridize to *C. coli* VC167 DNA but not to DNA from the *E. coli* DH5 host (data not shown). *C. coli* VC167 DNA was digested with *Eco*RV and size fractionated on sucrose gradients. Fractions of approximately 8 to 10 kb were ligated into *Eco*RV-digested pBR322 and transformed into DH5. Following colony hybridization to pg8, a plasmid termed pGK111, containing an approximately 9-kb insert which contained a 5.6-kb *Hin*dIII fragment equivalent to pGK109 was identified. Moreover, pGK111 hybridized to a probe specific for the 3′ end of the 16S rRNA genes of *C. coli* and *C. jejuni* (Romaniuk and Trust [49] probe D, whose sequence is ACTCGAGAG CATGAAGC) but not to a probe specific for the 5′ end of the 16S gene (Romaniuk and Trust [49] probe G, whose sequence is TATGACGCTTAACTGGT), indicating that part of the 16S gene was also found on pGK111, in the 5′ *Eco*RV-*Xba*I fragment shown in Fig. 1. A 2.5-kb *Xba*I fragment which contained the 5′ end of the 23S gene was subcloned from pGK111 into pUC19 and termed pGK112.

For DNA sequence analysis, pGK109 was digested with *Hin*dIII and *Bam*HI and the resulting 2.1- and 3.5-kb fragments were subcloned into both M13mp18 and M13mp19. These clones were then used as templates in single-stranded sequencing reactions primed with oligonucleotide primers synthesized on both strands at approximately 200- to 250-bp intervals. The 5′ end of the 23S gene was sequenced from pGK112 directly as a double-stranded template, also by using synthetic custom primers.

Primary and secondary structures of *C. coli* VC167 23S rRNA. The complete sequence of the cloned *C. coli* VC167 23S rRNA gene is shown in Fig. 2. The gene is 3,057 bases long with a 46.2% G+C content, compared with the 29 to 35% G+C content reported for the *C. coli* genome (1, 40). The length of the predicted primary sequence of the *C. coli* VC167 23S rRNA is in keeping with the range of lengths (approximately 2,883 to 3,022 bases) of other (eu)bacterial 23S rRNAs in the data bank. The 29.9% A and 18.9% C contents of the *C. coli* VC169 23S rRNA molecule are consistent with the ranges of 21.8 to 29.3% A and 19 to 27.2% C displayed by other (eu)bacterial 23S rRNAs. However, the 27.2% G (range, 30.1 to 36.6%) and 24% U (range, 14.4 to 21.2%) contents of the *C. coli* molecule are slightly lower and higher, respectively, than those of other 23S rRNAs. In terms of sequence similarity, *C. coli* rRNA has 99.2% sequence identity to *C. jejuni* TGH9011 23S rRNA, 71% sequence identity to *E. coli* 23S rRNA, and approximately 68.0 to 77.0% identity to the 23S rRNA sequences of other (eu)bacteria in the data bank. Of the 17 nucleotide changes between the *C. coli* and *C. jejuni* sequences (Fig. 3), 7 are located in the 270 region (*E. coli* numbering; Fig. 3A). The *C. coli* VC167 23S rRNA sequence was then arranged into a secondary-structure model. As shown in Fig. 3, this model also had overall similarity to other 23S-like rRNAs and specific structural similarity to (eu)bacterial 23S rRNAs.

Identification of an intervening sequence in the *C. coli* VC167 23S rRNA gene. A prominent difference in the *C. coli* VC167 23S rRNA structure from the *E. coli* and *C. jejuni* TGH9011 (GenBank accession number Z29326) 23S rRNA secondary structures occurs at position 1193 (ca. position 1170 on the basis of *E. coli* numbering [Fig. 3A]). The deduced *C. coli* 23S rRNA sequence contains an extra sequence block approximately 147 bp long which is not present in the *E. coli* or *C. jejuni* sequences (Fig. 2 and 3A). Similar sequence blocks have been shown to be present at this position in the 23S rRNA genes from a variety of bacteria (3, 20, 44, 47, 52, 60) and chloroplasts (59), and these extra sequence blocks in the 23S

rRNA genes have been termed intervening sequences (3) or transcribed spacers (60).

Distribution of intervening sequences in *C. coli* and *C. jejuni* 23S rRNA genes. PCR analysis was performed to determine whether the 23S rRNA genes of other strains of *C. coli* and *C. jejuni* also carry transcribed spacers in the same region of the gene. In the case of *C. coli* VC167, forward primer 1162 (*C. coli* numbering; 5′-TGCGCGGAAAATATAACGGGGCTA-3′; nucleotides −56 to −33 relative to the transcribed spacer) and reverse primer CG1425 (*C. coli* numbering; 5′-CTCAACTTA ATTATCGCTACTCAT-3′; nucleotides +111 to +88 relative to the intervening sequence) amplified a ca. 310-bp sequence which carried the ca. 147-bp transcribed spacer (see Fig. 6, lane 5). The results in Table 1 show that when these primers were used with DNA containing the 23S rRNA genes of 31 other *Campylobacter* strains, a ca. 310-bp fragment was amplified from 13 other strains, including strains belonging to LIO serogroups 5, 8, 8/29, 11, 12, and 29. In eight strains belonging to LIO serogroups 6, 15, and 19, a ca. 200-bp amplicon was obtained, which is consistent with the presence of a shorter transcribed spacer of ca. 37 bp. In the other 10 strains tested, a ca. 167-bp amplicon was obtained, which is consistent with the absence of a transcribed spacer in this region of the gene. These strains belonged to LIO serogroups 1, 4, 8/29, 11, 15, 29, 29/55, and 55. Figure 4 shows two examples of these ca. 167-bp amplicons obtained from the 23S rRNA genes of *C. jejuni* VC157 (lane 1) and *C. jejuni* VC74 (lane 2). Also shown in Fig. 4 are examples of the ca. 200-bp amplicons obtained from *C. jejuni* VC84 (lane 3) and *C. jejuni* VC227 (lane 4), as well as the ca. 310-bp amplicon from *C. coli* VC92 (lane 6). DNA sequence analysis of the transcribed spacers from *C. coli* VC92 and *C. jejuni* 81176 indicated that under conditions for maximum matching (34), they had 85 and 63% homology, respectively, to the transcribed spacer from *C. coli* VC167 (data not shown). The truncated transcribed spacers from *C. jejuni* VC84 and *C. jejuni* 81116 also shared identity with the 5′ end of the VC167 transcribed spacer sequence (data not shown).

Table 2 shows the predicted 23S rRNA sequences reported for the 1170 region helices of other *Campylobacter* (60), *Salmonella* (3), and *Yersinia* (52) strains compared to that of *C. coli* VC167. While the sequences of a number of the *C. coli* and *C. jejuni* strains are clearly similar to that of *E. coli*, giving rise to a short hairpin loop, *Campylobacter* strains which carry a transcribed spacer at this region of their 23S rRNAs have a much longer hairpin loop sequence. In terms of sequence identity, the transcribed-spacer-containing *Campylobacter* sequences appear to be related and to be distinct from the *Salmonella* and *Yersinia* sequences. Under conditions necessary for maximum matching (34), the two most closely related transcribed spacer sequences were *C. coli* VC167 and *C. jejuni* subsp. *doylei* (60) at 83%. Overall, the transcribed spacer sequences from different *Campylobacter* species (60) were >53% identical to each other. The similarity between the *Campylobacter* transcribed spacers was also apparent in their base compositions, with the various sequenced *Campylobacter* transcribed spacers (60) displaying a mean G+C content of 18.1%. This differed significantly from the mean 47% G+C content of the *Salmonella* and *Yersinia* transcribed spacers shown in Table 2. It should also be noted that the 18.9% G+C content of the ca. 147-bp transcribed spacer from *C. coli* VC167 differed significantly from the 46.1% overall G+C content of the VC167 23S rRNA gene sequenced here.

Demonstration of fragmented 23S rRNA. Analysis of the *C. coli* and *C. jejuni* 23S rRNAs by denaturing gel electrophoresis showed that while strains which appeared not to contain a transcribed spacer in their 23S rRNA genes and displayed discrete 23S rRNAs of approximately 2.9 kb, the rRNAs prepared from *C. coli* VC167 displayed two smaller, distinct rRNA fragments with approximate sizes of 1.7 and 1.2 kb along with the 1.6-kb 16S RNA (Table 1). Figure 5 shows the two distinct, smaller 23S rRNA fragments in representative strains carrying either the longer 147-bp or the shorter 37-bp transcribed spacer compared to the unfragmented, approximately 2.9-kb 23S rRNAs from *C. jejuni* VC157 and VC228.

Secondary structure and phylogenetic signatures. The secondary structures of the *C. coli* VC167 and *C. jejuni* TGH9011 23S rRNAs (GenBank accession number Z29326) were constructed by comparative methods (16, 19), and the overall structures of these two rRNAs are virtually identical, except at the region where the intervening sequence is located (GenBank accession number Z29326 and Fig. 3). The extra 147-bp transcribed spacer increases the length of the stem-loop structure at position 1193 of the *C. coli* VC167 23S rRNA. In the 270 region (*E. coli* numbering; Fig. 3A), the *C. coli* and *C. jejuni* rRNAs display a structural signature involving a helix (marked a in Fig. 3A). The sequences of *E. coli* and other representative *Proteobacteria* (e.g., subclass alpha member *Rhodobacter capsulatus*, subclass beta member *Pseudomonas cepacia*, and subclass gamma member *E. coli*) do not contain this helix, while all other (eu)bacteria, chloroplasts, and archaea sequences (e.g., the gram-positive bacterium, *Bacillus subtilis*, the spirochete *Leptospira interrogans*, and the cyanobacterium *Anacystis nidulans*) in the data bank (18) contain this helix-insertion. A comparison of representative sequences of this region of 23S rRNA is shown in Fig. 6. In the 1850 region (*E. coli* numbering; Fig. 3B), however, the genus *Campylobacter* is clearly similar to the beta and gamma *Proteobacteria* (Fig. 7 and Table 3). In the juxtapositions 1855-1887, 1856-1886, 1858-1884, and 1859-1883, the genus *Campylobacter* favors U/C, U/U, A/G, and U/U pairings, respectively, as do the beta and gamma *Proteobacteria*. These base pairings at these positions are rarely, if ever, present in the 23S rRNAs from the alpha *Proteobacteria*, cyanobacterial plastids, or bacteria belonging to other (eu)bacterial divisions.

Phylogenetic analysis. The 23S RNAs of *C. coli* VC167 and *C. jejuni* TGH9011 (GenBank accession number Z29326) were aligned with sequences from all other available (eu)bacterial 23S rRNAs, and phylogenetic trees were inferred from full-length and more conservative matrices by the Fitch-Margoliash distance, neighbor-joining, parsimony, and maximum-likelihood methods. The bootstrapped neighbor-joining result from the more conservative matrix is shown in Fig. 8 because it is representative and because simulations have shown the neighbor-joining method to be relatively successful in recovering correct topologies (reference 5 and references therein).

A number of lineages diverging more basically in the tree (Fig. 8) are resolved poorly or not at all, while bootstrap support is excellent for seven preterminal groupings, including

FIG. 2. Nucleotide sequence of *C. coli* VC167 DNA containing a 23S rRNA gene. The nucleotides of the 23S rRNA gene are in boldface, numbers of nucleotides relative to the start of this gene are on the left, a transcribed spacer is in underlined italics, and the targets for oligonucleotides 1162 (5′ to the transcribed spacer) and 1425 (3′ to the transcribed spacer) used to detect transcribed spacers in other campylobacters are underlined.

1

-76 ATCTTTTAGG TTGGCGGGGT AGGTCTTTAG CGTGCTTCTT AGTCAAGCTT TGCCTTGACG CTAAAGAAGG TAAAAA**AGGT**

5 **AAGCTACTAA GAGCGAATGG TGGATGCCTT GACTGGTAAA GGCGATGAAG GACGTACTAG ACTGCGATAA GCTACGGGGA**

85 **GCTGTCAAGA AGCTTTGATC CGTAGATTTC CGAATGGGGC AACCCAATGT ATAGAGATAT ACATTACCTA TATAGGAGCG**

165 **AACGAGGGGA ATTGAAACAT CTTAGTACCC TCAGGAAAAG AAATCAATAG AGATTGCGTC AGTAGCGGCG AGCGAAAGCG**

245 **CAAGAGGGCA AACCCAGTGC TTGCACTGGG GGTTGTAGGA CTGCAATGTG CAAGAGGTGA GTTTAGCAGA ACATTCTGGA**

325 **AAGTATAGCC ATAGAGGGTG ATAGTCCGTA TGCGAAAACA AACCTTAGCT AGCAGTATCC TGAGTAGGGC GGGACACGAG**

405 **AAATCCTGTC TGAAGCTGGG TCGACCACGA TCCAACCCTA AATACTAATA CCAGATCGAT AGTGCACAAG TACCGTGAGG**

485 **GAAAGGTGAA AAGAACTGAG GTGATCAGAG TGAAATAGAA CCTGAAACCA TTTGCTTACA ATCATTCAGA GCCCTATGTA**

565 **GCAATACAGG GTGATGGACT GCCTTTTGCA TAATGAGCCT GCGAGTTGTG GTGTCTGGCA AGGTTAAGCA AACGCGAAGC**

645 **CGTAGCGAAA GCGAGTCTGA ATAGGGCGCT TAGTCAGATG CTGCAGACCC GAAACGAAGT GATCTATCCA TGAGCAAGTT**

725 **GAAGCTAGTG TAAGAACTAG TGGAGGACTG AACCCATAGG CGTTGAAAAG CCCCAGGGTA GACTTGTGGA TAGGGGTGAA**

805 **AGGCCAATCA AACTTCGTGA TAGCTGGTTC TCTCCGAAAT ATATTTAGGT ATAGCGTTGT GTCGTAATAT AAGGGGGTAG**

885 **AGCACTGAAT GGGCTAGGGC ATACACCAAT GTACCAAACC CTATCAAACT CCGAATACCT TATATGTAAT CACAGCAGTC**

965 **AGGCGGCGAG TGATAAAATC CGTCGTCAAG AGGGAAACAA CCCAGACTAC CAGCTAAGGT CCCTAAATCT TACTTAAGTG**

1045 **GAAAACGATG TGAAGTTACT TAAACAACCA GGAGGTTGGC TTAGAAGCAG CCATCCTTTA AAGAAAGCGT AATAGCTCAC**

1125 **TGGTCTAGTG ATTTTGCGCG GAAAATATAA CGGGGCTAAA GTAAGTACCG AAGCTGTAGA CTTGCGCACA *ACTTAGATTA***

1205 ***TTTAAGTTTA GAATATGAGA AACTAAGTTA TATGTTTAGT TATATTTTTA CTGATTTTTA TAGAGTAAAG ATAGAAATAA***

1285 ***AACTTAGTAA AATCAGTAAA AATATTCTTA GACTAAAGTT AAGTAGTTTA AGTTGT*GTGC AAGTGGTAGG AGAGCGTTCT**

1365 **ATTTGCGTCG AAGGTATACC GGTAAGGAGT GCTGGAGCGA ATAGAAGTGA GCATGCAGGC ATGAGTAGCG ATAATTAATG**

1445 **TGAG**AATCAT TAACGCCGTA AACCCAAGGT TTCCTACGCG ATGCTCGTCA TCGTAGGGTT AGTCGGGTCC TAAGTCGAGT

1525 **CCGAAAGGGG TAGACGATGG CAAATTGGTT AATATTCCAA TACCAACATT AGTCGTCGCG ATGGAAGGAC GCTTAGGGCT**

1605 **AAGGGGGCTA GCGGATGGAA GTGCTAGTCT AAGGTCGTAG GAGGTTATAC AGGCAAATCC GTATAACAAT ACTCCGAGAA**

1685 **CTGAAAGGCT TTTTGAAGTC TTCGGATGGA TAGAAGAACC CCTGATGCCA TCGAGCCAAG AAAAGTTTCT AAGTTTAGCT**

1765 **AATGTTGCCC GTACCGTAAA CCGACACAGG TGGGTGGGAT GAGTATTCTA AGGCGCGTGG AAGAACTCTC TTTAAGGAAC**

1845 **TCTGCAAAAT AGCACCGTAT CTTCGGTATA AGGTGTGGTT CGCTTCGTAT TAGGATTTAC TCCGAAAGCG AAGAAACTTA**

1925 **CAACAAAGAG TCCCTCCCGA CTGTTTACCA AAAACACAGC ACTCTGCTAA CTCGTAAGAG GATGTATAGG GTGTGACGCC**

2005 **TGCCCGGTGC TCGAAGGTTA ATTGATGGGG TTAGCATTAG CGAAGCTCTT GATCGAAGCC CGAGTAAACG GCGGCCGTAA**

2085 **CTATAACGGT CCTAAGGTAG CGAAATTCCT TGTCGGTTAA ATACCGACCT GCATGAATGG CGTAACGAGA TGGGAGCTGT**

2165 **CTCAAAGAGG GATCCAGTGA AATTGTAGTG GAGGTGAAAA TTCCTCCTAC CCGCGGCAAG ACGGAAAGAC CCCGTGGACC**

2245 **TTTACTACAG CTTGACACTG CTATTTGGAT AAGAATGTGC AGGATAGGTG GGAGGCTTTG AGTATATGAC GCCAGTTGTA**

2325 **TATGAGCCAT TGTTGAGATA CCACTCTTTC TTATTTGGGT AGCTAACCAG CTTGAGTTAT CCTCAAGTGG GACAATGTCT**

2405 **GGTGGGTAGT TTGACTGGGG CGGTCGCCTC CCAAATAATA ACGGAGGCTT ACAAAGGTTG GCTCAGAACG GTTGGAAATC**

2485 **GTTCGTAGAG TATAAAGGTA TAAGCCAGCT TAACTGCAAG ACATACAAGT CAAGCAGAGA CGAAAGTCGG TCTTAGTGAT**

2565 **CCGGTGGTTC TGTGTGGAAG GGCCATCGCT CAAAGGATAA AAGGTACCCC GGGGATAACA GGCTGATCTC CCCCAAGAGC**

2645 **TCACATCGAC GGGGAGGTTT GGCACCTCGA TGTCGGCTCA TCGCATCCTG GGGCTGGAGC AGGTCCCAAG GGTATGGCTG**

2725 **TTCGCCATTT AAAGCGGTAC GCGAGCTGGG TTCAGAACGT CGTGAGACAG TTCGGTCCCT ATCTGCCGTG GGCGTAAGAA**

2805 **GATTGAAGAG ATTTGACCCT AGTACGAGAG GACCGGGTTG AACAAACCAC TGGTGTAGCT GTTGTTCTGC CAAGAGCATC**

2885 **GCAGCGTAGC TAAGTTTGGA AAGGATAAAC GCTGAAAGCA TCTAAGCGTG AAGCCAACTC TAAGATGAAT CTTCTCTAAG**

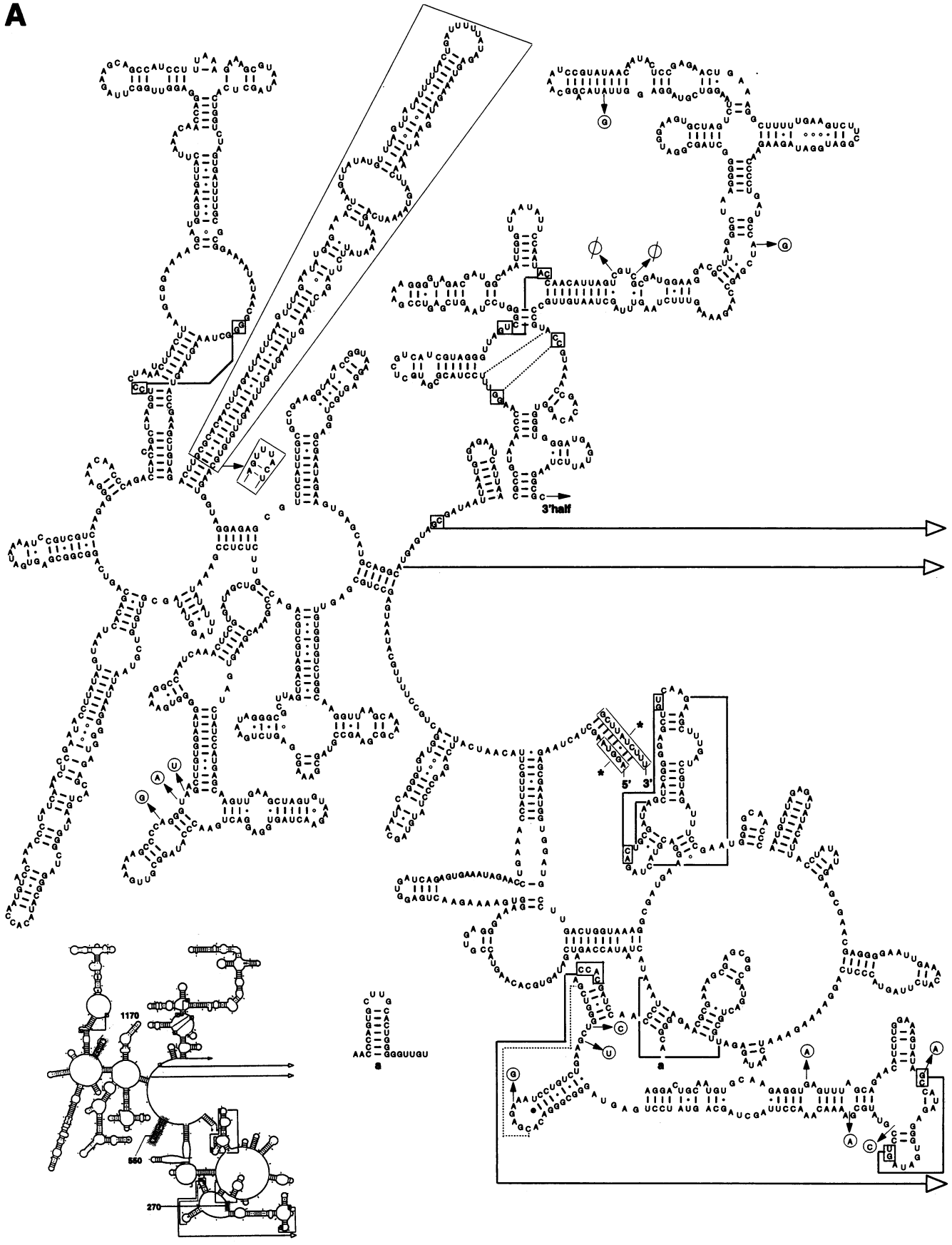
2965 **CTCTCTAGAA GACTACTAGT TTGATAGGCT GGGTGTGTAA TGGATGAAAG TCCTTTAGCT GACCAGTACT AATAGAGCGT**

3045 **TTGGCTTATC TTT**AATAAAG CATCATTCTT GTTAAGGTTT TTAAGAAGAC TTTGAATATA GATAATATTT AGAGTTTAAT

3125 AGAAATCTTT CAAGTAGAGT TTGTATTAGA ACTTGCTCTT AACATTGTTT TTTAAGTATT CTAAAATAAA AACTTATCAA

3205 AGATAAAAGA TAAGAAAAGA AGAAAGAGAA TAAAGAAAAA AGTAAAGGAA TAAAAGATTA AGTTTTATTC TTAAATTCAA

3285 TCTTTCAAAG AATATTTAAA TAACAATGTC CGTGATTATA CAGATGTGGA AACGCCTTGC TCCATCCCGA ACCAAGAAGC


A
3'half
5'
3'
1170
550
270

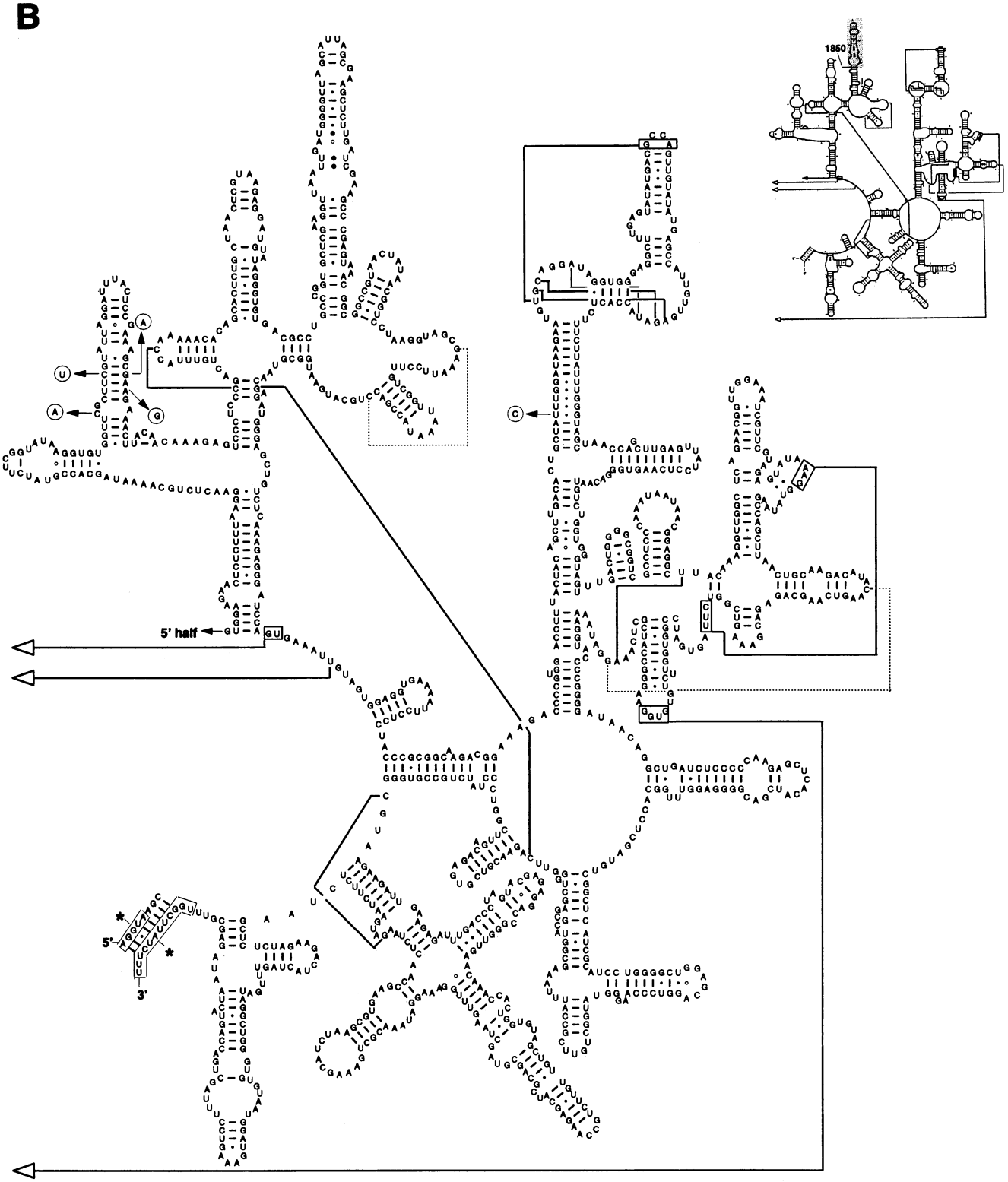


FIG. 3. Structural model for the 5′ (A) and 3′ (B) halves of *C. coli* VC167 23S RNA. Canonical (C · G, G · C, etc.) base pairs are connected by lines, G/U pairs are connected by dots, A/G type pairs are connected by open circles, and other noncanonical pairings are connected by filled circles. Proposed tertiary interactions are connected by thicker (and longer) solid lines, and more tentative assignments are connected by thinner, dashed lines. The helix in the 270 region which distinguishes *Campylobacter* 23S rRNA from alpha, beta, and gamma proteobacterial 23S rRNAs is labeled a. The potential helix in the 1170 region formed as a result of the presence of the transcribed spacer in the 23S rRNA gene is boxed, as is the corresponding region of the sequence in the nontranscribed spacer containing 23S rRNA of *C. jejuni* TGH9011 (GenBank accession number Z29326). The nucleotide changes and deletions (Ø) contained in the *C. jejuni* TGH9011 23S rRNA sequence are arrowed. Also shown are schematic representations of the 5′ (A, lower left diagram) and 3′ (B, upper right diagram) halves of *E. coli* 5′ 23S RNA showing the relative positions of structural regions identified in the text.

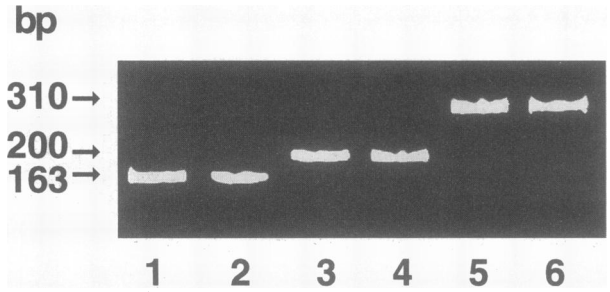


FIG. 4. Agarose gel showing different sizes of PCR products amplified from 23S rRNA genes of *C. jejuni* VC157 (lane 1), VC74 (lane 2), VC84 (lane 3), and VC227 (lane 4) and *C. coli* VC167 and VC92 (lane 6) by using primers 1162 and 1425. Size markers are on the left.

plastids and cyanobacteria (100% bootstrap); mycoplasmas (100%); the genera *Flexibacterium*, *Flavobacterium*, and *Chlorobium* (100%); two groups of gram-positive and relative bacteria (low G+C content [82%] and high G+C content [100%]); alpha, beta, and gamma *Proteobacteria*; and the genus *Pirellula*. The *Proteobacteria* group is always exclusive, holophyletic, and supported in 100% of bootstrap analyses.

In most of our trees (e.g., Fig. 8), rRNAs of the two campylobacteria formed a sister group to the alpha, beta, and gamma proteobacterial rRNAs, although bootstrap support was lower (32 to 68%). The maximum-likelihood tree showed the same relationship (data not shown). This positioning is consistent with the idea that the genus *Campylobacter* represents a different (e.g., epsilon) subdivision of *Proteobacteria*. In a few trees inferred from the full-sequence matrix, the two campylobacterial rRNAs fail to associate with the proteobacterial grouping, but bootstrapping support for alternative groupings is poor (15 to 46%).

Although the 23S tree shown in Fig. 8 clearly displays seven strongly supported groupings, it should be noted that the tree loses the resolution necessary to securely link these groupings. The apparent splitting of the gram-positive phylogenetic group into three groups (high and low G+C content groups and mycoplasmas) on the 23S tree probably reflects this loss of resolution, coupled with probable G+C content effects on topology. Loss of resolution is also presumably why the *Leptospira* sequence has broken free from the genus *Borrelia* and drifted into the low-G+C positive group on the 23S tree. With these exceptions, however, and taking into account the better 16S sequence data density, the fact that 16S rRNA is a better molecule for deep branchings, as well as the better established rooting of the 16S tree, the 23S-derived tree shows considerable overall similarity to the recently published 16S-derived tree (39).

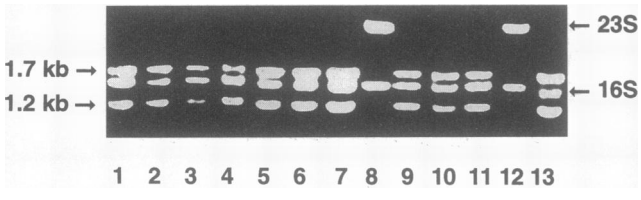


FIG. 5. Agarose formaldehyde gel electrophoretic analysis of rRNAs isolated from *Campylobacter* species and stained with ethidium bromide. The approximately 1.6-kb 16S and 2.9-kb nonfragmented 23S rRNA species are indicated on the right, and the 1.7- and 1.2-kb 23S rRNA fragments are indicated on the left. Lanes: 1, *C. coli* VC20; 2, *C. jejuni* VC88; 3, *C. jejuni* VC94; 4, *C. jejuni* VC104; 5, *C. jejuni* VC167; 6, *C. coli* VC143; 7, *C. coli* VC144; 8, *C. jejuni* VC152; 9, *C. jejuni* VC157; 10, *C. jejuni* VC209; 11, *C. jejuni* VC222; 12, *C. jejuni* VC228; 12, *C. jejuni* VC229.

DISCUSSION

The large-subunit rRNA species of (eu)bacteria has received much less attention than the 16S rRNA species. Recent compilations have reported nearly 1,400 published complete (eu)bacterial 16S rRNA sequences (6) compared with 70 complete 23S rRNA sequences (18), including the 2 *Campylobacter* sequences reported in this and another study (GenBank accession number Z29326). Here we have presented the first complete sequence of a 23S rRNA gene from *C. coli*, along with a prediction of the secondary structure of the rRNA encoded by this gene. In terms of overall size, nucleotide composition, and primary and secondary structures, this *C. coli* 23S rRNA displays considerable overall similarity to other (eu)bacterial 23S sequences. In addition, all tertiary interactions proposed on the basis of comparative methods (16, 19) are consistent with the *C. coli* primary and secondary structures proposed here. The particular copy of the 23S gene sequenced appeared to be located in a typical prokaryotic ribosomal operon. For example, with oligonucleotides specific for *Campylobacter* 16S RNA, a 16S RNA gene was demonstrated to lie 5′ to the 23S gene and this 16S rRNA gene appeared to be transcribed in the same direction as the 23S rRNA gene because clone pGK111, which hybridized to an oligonucleotide specific for the 3′ end of the 16S rRNA gene, did not hybridize to a 5′-specific 16S rRNA probe. The 5′ portion of the 16S coding region must therefore extend beyond the 5′ boundaries of clone pGK111.

The 23S rRNA gene of *C. coli* VC167 contained a 147-bp transcribed spacer which resulted in fragmented 23S rRNA. Although 23S rRNAs occur as contiguous molecules in most bacteria, the occurrence of fragmented 23S RNA in some bacteria was first reported in 1970 (32, 46). Burgin et al. (3) subsequently showed that in some strains of salmonellae, transcribed spacers can form helical structures and expand the helices in the 500 and 1170 regions (*E. coli* numbering; Fig. 3A) of premature 23S rRNAs. These transcribed spacers are removed during maturation of the rRNAs, and after processing, the 23S rRNAs remain fragmented within the ribosome. Skurnik and Toivanen (52) reported corresponding transcribed spacers at the 1170 region of the 23S rRNA genes of *Y. enterocolitica* which also resulted in fragmented 23S rRNA.

The occurrence of a 23S rRNA gene containing a transcribed spacer and the presence of fragmented 23S rRNA was not confined to *C. coli* VC167 but appeared to be a common phenomenon in *C. coli* and *C. jejuni*. Of the 32 strains of *C. coli* and *C. jejuni* examined here, 69% exhibited both properties. In addition, Van Camp et al. (60) found internal transcribed spacers in the 23S rRNA genes of 5 of 17 campylobacters examined, including two of the four strains of *C. jejuni* tested, in both *C. fetus* strains tested, and in one of two strains of *C. upsaliensis*. These spacers varied in length from ca. 120 to about 180 bp compared with the ca. 37- and ca. 147-bp transcribed spacer lengths observed in this study. The various *Campylobacter* inserts appeared to be related, in terms of both nucleotide sequence and base content, consistent with a common original source.

The 23S rRNA fragmentation patterns we observed suggested cleavage at similar, if not identical, positions in all of the strains examined, and the 1.7- and 1.2-kb sizes of the 23S rRNA fragments were consistent with cleavage at approxi-

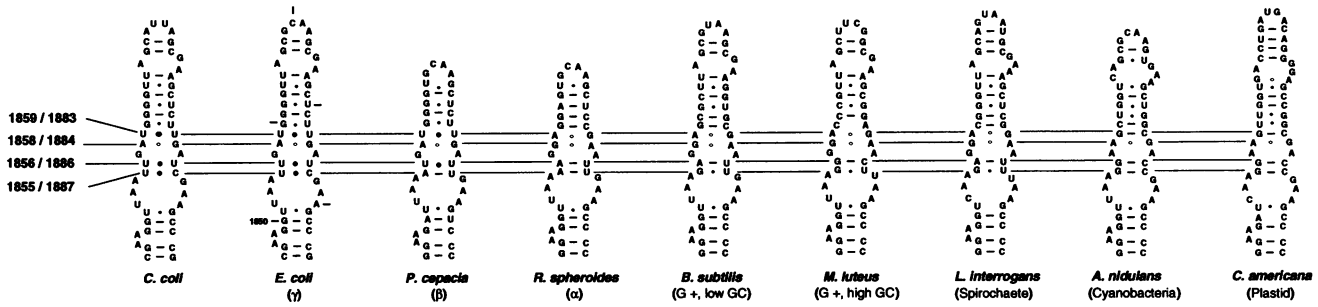


FIG. 6. Comparative diagrams showing the absence of the helix structure found in the 270 region (*E. coli* numbering; Fig. 3A) of *C. coli* VC167 23S rRNA in the corresponding region (arrowed) of *R. capsulatus* (alpha), *P. cepacia* (beta), and *E. coli* (gamma) 23S rRNAs and the presence of a similar helix structure in *B. subtilis* (gram positive [G +], low G+C content), *L. interrogans* (spirochete), and *A. nidulans* (cyanobacterium).

mately position 1170 (*E. coli* numbering) in helix 45, one of the most variable helices in bacterial 23S rRNA (60). The 1170 position is also at a position homologous to an expansion segment in eucaryotes (41). The five internal spacers reported by Van Camp et al. (60) were also located in this helix, as were transcribed spacers of *Salmonella* (3), *Yersinia* (52), *Actinobacillus* (20), *Leptospira* (52), and *Chlamydomonas* chloroplasts (59). In the genus *Salmonella*, the occurrence of intervening sequences appears to be sporadic because even within a particular *Salmonella* strain, some of the multiple rRNA operons have transcribed spacers while other operons have none (3). In the genus *Yersinia*, some individual strains also appear to have some 23S gene copies carrying transcribed spacers and other 23S gene copies lacking transcribed spacers, but in other strains there are transcribed spacers in all copies of the 23S rRNA genes (52). In the transcribed-spacer-carrying *Campylobacter* strains studied here, the apparent absence of intact 23S RNA suggests that all expressed copies of the 23S rRNA gene in a given strain carry a transcribed spacer. The function of the transcribed spacers remains obscure. Gray and Schnare (12) have argued that transcribed spacers were an early acquisition, while Pace and Burgin (41) have argued that they were acquired relatively recently and have no purpose in the cell. Whatever the case, however, the ability of *C. coli* to exhibit normal growth indicates that the fragmented 23S RNAs retain their functionality, presumably through secondary and tertiary interactions. In terms of origin, a number of the transcribed spacers found in helix 45 of *Salmonella* and *Yersinia* 23S rRNAs have >90% sequence identity, and Skurnik and Toivanen (52) have suggested the possibility that they have

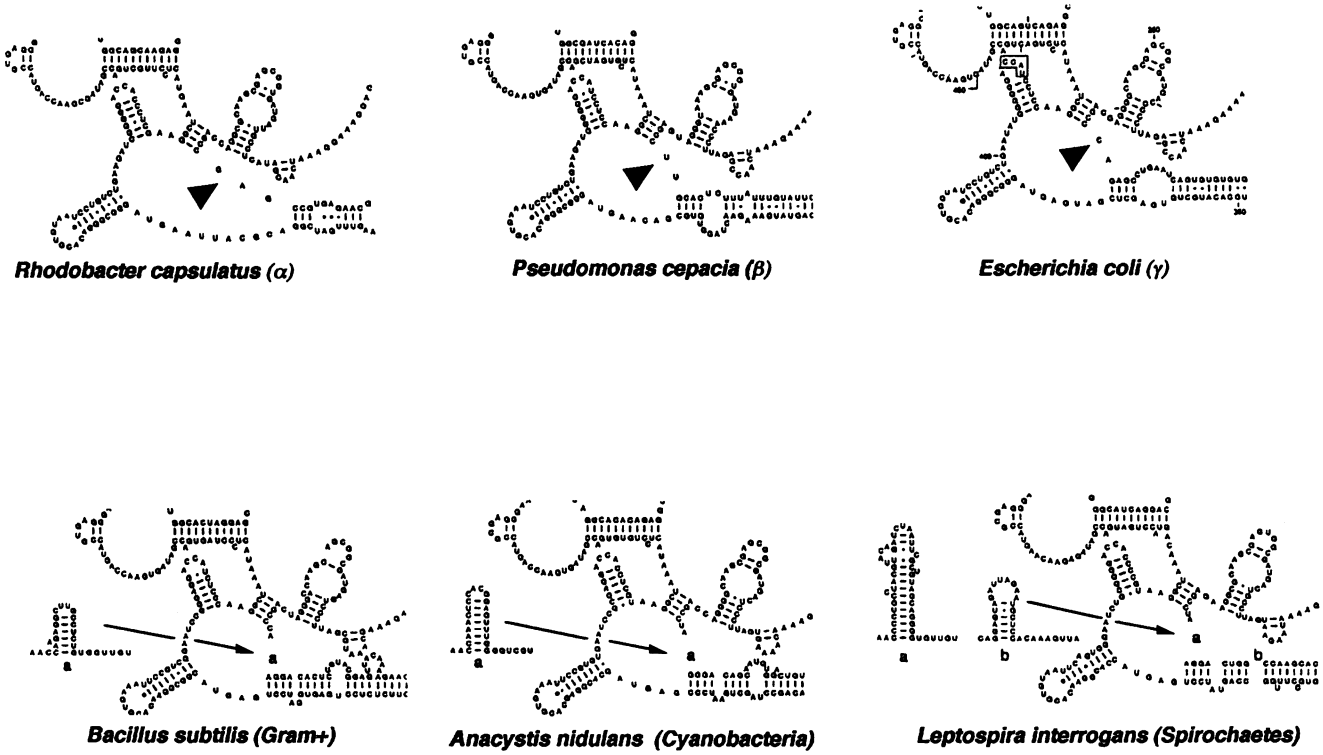


FIG. 7. Comparative diagrams showing noncanonical and canonical base pairings in the 1870 region (based on *E. coli* numbering) of the 23S rRNAs representative of a phylogenetically diverse variety of (eu)bacteria and a plastid (*Conophilus americana*). The positions of base pairs 1855 and 1887, 1856 and 1886, 1858 and 1884, and 1859 and 1883 are indicated, as are the species.

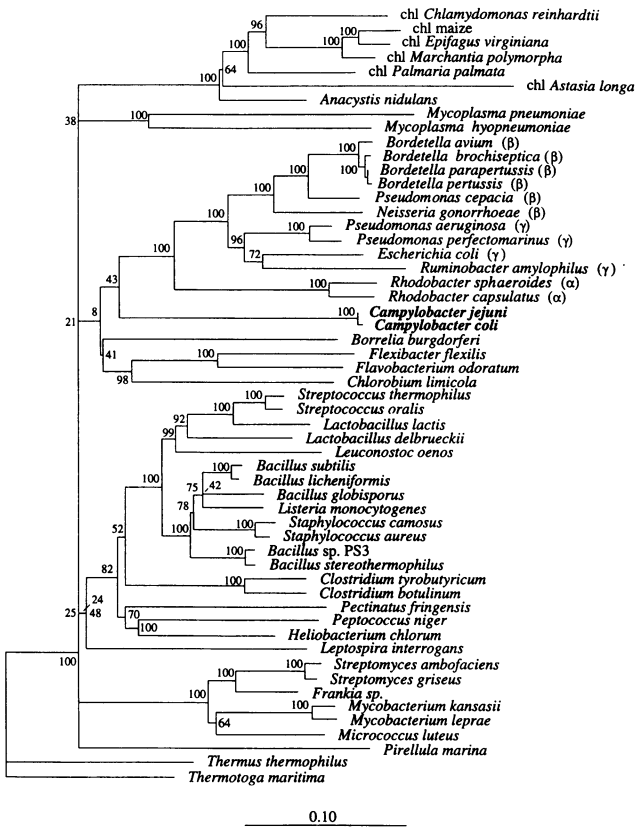


FIG. 8. Phylogenetic analysis of *Campylobacter* 23S rRNAs by the method of neighbor joining. The 23S rRNA of *T. maritima* (order *Thermogales*) served as the out-group. The tree was computed from a data set of 2,409 semiconserved positions spanning the entire large-subunit rRNA molecule. Branch length on the horizontal axis represents the evolutionary distance between nodes. The number at each node indicates the percentage of times that the monophyletic group to the right occurred in 100 bootstrap samples. Subdivisions of the class *Proteobacteria* are in parentheses. chl, chloroplast.

a source in common. The *Campylobacter* transcribed spacers are different in sequence and base composition and appear to have been derived from a source different from that of the transcribed spacers of these organisms.

Within the *Campylobacter* group, Eyers et al. (10) have recently identified a series of oligonucleotide primers in the 1060-to-1935 region of the 23S rRNA gene (*E. coli* numbering) and used it to discriminate among thermophilic *Campylobacter* strains by PCR. In terms of phylogeny within the genus *Campylobacter*, Van Camp et al. (60) have also shown that the topology of trees based on complete 16S rRNA and partial 23S rRNA sequences (based on fragments of 875 or 975 bp located between helices 43 and 69 [8]) are the same, except for *C. sputorum*. Moreover, the results obtained by DNA-rRNA hybridization (62) provided evidence in favor of the phylogeny predicted by the 23S rRNA tree. While we did not consider the question of the phylogenetic position of *C. coli* relative to other campylobacters, we did examine the phylogenetic position of the organism relative to members of the class *Proteobacteria* and several other major (eu)bacterial classes from the perspective of the 23S rRNA sequence.

The phylogenetic position of the genus *Campylobacter* has been contentious despite several detailed analyses with partial 16S RNA sequence data. Romaniuk et al. (50) suggested that

TABLE 2. Sequence of 1170 helix region (*E. coli* numbering) from *C. coli*, *C. jejuni*, and *C. fetus* 23S rRNAs compared to sequences of 23S rRNAs from *S. arizonae*, *S. typhimurium*, and *Y. enterocolitica* strains carrying transcribed spacers

Organism (ref.)	5′ half	Sequence of 1170 helix hairpin loop	3′ half
E. coli	cagcgacgc	uuau	gcguugguug
C. coli 6440 (60)	acuuag	uuua	cuaagu
C. coli 7535 (60)	acuuagu	uuu	acuaagu
C. coli VC167	acuugcgca	caacuuagauauuuuaaguuuagaauauggaaacuaaguauauguuuaguauauuuuuacugauuuuuuauagaguaaa gauagaauaaaaacuuaguaaa aucaguaaaaauauuucuuagacuaaaguuaaguuuaaguuuaaguug	ugugcaaguu
C. jejuni 6692 (60)	acuuag	uuua	cuaagu
C. jejuni 6444 (60)	acuuag	uuua	cuaagu
C. jejuni subsp. *doylei* (60)	auuugcacacaa	cuagauuauuuaaguuuagaauauuaaagaacuaaguuauauuuuuacugauuuuuuauagaguaaagauauuaaacuu aguaaaacagaaaaauauucuuagacuaaaguuaaguuuaaguuuaag	uuguugugcaagu
C. fetus (60)	acuugcacuuaa	uucuaauuauaaauuuggcuaugagcuuuuaaguuauauuucucauaauuuaauuugcacucauauaaauuauaauuaugaga aaauaguuuagcuaaaauuagaaguaauuagaauua	agugcaagu
S. arizonae RF908 (3)	cagcgaacg	uaucacuuaaagcgauucagcggaugacgauugacggagcgaagcgacgucaaagcguuuuacaaagucgagucggguuaa ggaua	cguucguug
S. typhimurium (3)	cagcgaacg	uaucacccaagacaacuuuacggaguugacgauugacggagcgaagcgacgucaaagcguucauuaaagucgaguuggcuu agggaua	cguucguug
Y. enterocolitica O:1 (52)	cagcgaacg	uaucacuuaaaacgcuuacgugacgcaaagccggcauguucaagccacacuaaacguugaguuggccggugugcugacaa acgaacaggcguuuuaaggua	cgttcgtt

TABLE 3. Frequency of base pairings at positions 1855 and 1887, 1856 and 1886, 1858 and 1884, and 1859 and 1883 (based on *E. coli* numbering)[a] relative to phylogenetic classification

(Eu)bacterial phylum or subdivision	Frequency of base pairings indicated at following positions:																
	U-1855–C-1877				U-1856–C-1886				U-1858–C-1884			U-1859–C-1883					
	U/C	A/U	G/U	G·C	U/U	A/U	G·C	C·G	A/G	G/A	A/A	U/U	A/G	G·C	G/U	A/U	Other
Proteobacteria																	
Alpha	5	1			6				6			6					
Beta	2	1		4	7				7			5	2				
Gamma		5				4	1			5			5				
Campylobacters	11				11				11			11					
Gram-positive bacteria																	
Low-G+C group, others		6	22	7		20	13	2		33	2		33				2
High G+C group			6			2	4			6			6				
Bacteroids-flavobacteria			2			2				2			2				1
Green-sulfur bacteria				1			1			1			1				
Radioresistant micrococci				1			1			1			1				
Thermotoga maritima				1			1			1			1				
Planctomyces sp.				1			1		1					1			
Spirochetes and relatives			2			2				2			−2				
Cyanobacteria				1			1		1					1			
Plastids		5	3	26			34		30		4			27	5	2	

[a] A compilation of sequences, structures, organism names, and references is presented in the annual Nucleic Acids Research database issue (18; R. R. Gutell, unpublished collection).

the true campylobacters, including *C. coli*, appear to belong to a previously undefined phylogenetic branch, on the basis of a signature sequence analysis and comparison of sequence data with partial sequences of representatives of 8 of the 10 major phylogenetic branches of the (eu)bacteria. Lau et al. (30) subsequently suggested that while there is some evidence of affinity between an expanded division containing the purple photosynthetic bacteria and their relatives and the *Campylobacter* cluster, the *Campylobacter* cluster is not specifically positioned in either the alpha, beta, or gamma subdivision of the group. In a subsequent, expanded, partial-16S rRNA study involving all species in the genus *Campylobacter*, Thompson et al. (58) found three separate *Campylobacter* 16S rRNA identity groups similar to one another at a level of 68%, suggesting that they belong together on the same phylogenetic branch. Further, the three *Campylobacter* groups had a 49.3% overall 16S rRNA sequence similarity to representatives of the alpha, beta, and gamma phylogenetic branches of the class *Proteobacteria*, a depth of divergence greater than that among the three branches themselves and consistent with the conclusions reached by Romaniuk et al. (50). Stackebrandt et al. (54) took the position that overall there was a meaningful affinity between the campylobacters and members of the class *Proteobacteria*, and Vandamme et al. (61, 62) concluded that while the campylobacters clearly constitute a sixth rRNA superfamily within the gram-negative bacteria and are equally removed from all subclasses within the class *Proteobacteria*, it seems appropriate to consider rRNA superfamily V1 a fifth subclass of the class *Proteobacteria*, as deduced from a suggestion of Murray et al. (33).

Recently evolutionary analyses by Lane et al. (28) have indicated that the root of the *Thiovulum-Campylobacter* group falls very close to the root of the class *Proteobacteria* and, for consistency of usage, included the group in the class *Proteobacteria*. Depending on the exact placement of the roots, these workers suggested that the genus *Campylobacter* and the closely related *Thiovulum* group of organisms can be viewed either as a deep offshoot of the delta subdivision of the class *Proteobacteria* or as an independent subdivision. However, Lane et al. (28) further noted that the campylobacters differ from the alpha, beta, gamma, and delta proteobacteria on the basis of a structural signature in the 184-to-219 region of 16S rRNA and suggested that this signature argues for placement of the campylobacters as a distinct division or phylum of the (eu)bacteria. Unfortunately, no 23S sequence data are available for members of the delta group, which contains nonchemolithotropic bacteria, e.g., *Bdellovibrio* spp., myxobacteria, and the majority of sulfur- and sulfate-reducing (eu)bacteria, e.g., *Desulfovibrio* spp., *Desulfococcus* spp., etc. However, the phylogenetic tree analysis based on the 23S rRNAs of all available (eu)bacterial sequences derived here placed the genus *Campylobacter* as a sister subdivision to the alpha, beta, and gamma subdivisions of the class *Proteobacteria*. Structural signatures appear to confirm such a positioning. In the 1850 region, *Campylobacter* 23S rRNA shows similarities to the purple-beta and gamma subdivisions, consistent with positioning in the class *Proteobacteria*. However, *Campylobacter* 23S rRNA also contains a helix as a phylogenetic signature in the 270-bp region which is lacking in representative sequences from alpha, beta, and gamma *Proteobacteria* but present in other (eu)bacteria, emphasizing the distinctness of the genus *Campylobacter* from these other *Proteobacteria* at the phylogenetic level. Interestingly, an unpublished 23S rRNA sequence from *Stigmatella aurantiaca* (64), a member of the delta group of *Proteobacteria* also contains the characteristic helix insertion at the 270 region. However, in keeping with the positioning of the genus *Campylobacter* relative to the delta *Proteobacteria* suggested by Lane et al. (28), the genus *Campylobacter* appears to be positioned in the epsilon subdivision of the class *Proteobacteria*.

ACKNOWLEDGMENTS

This work was supported in part by grants from the Medical Research Council of Canada to T.J.T. and V.L.S. and from the NIH (GM 48207) to R.R.G. and by U.S. Navy Research and Development Command Research Unit 61102A3M161102BS13 AK111. R.R.G. and M.R. are associate members of the Program in Evolutionary Biology of the Canadian Institute of Advanced Research. The Keck foundation is gratefully acknowledged for support of R.R.G. at the University of Colorado, Boulder.

Carl Woese is thanked for sharing *Stigmatella* 23S rRNA sequence information.

REFERENCES

1. **Belland, R. J., and T. J. Trust.** 1982. Deoxyribonucleic acid relatedness between thermophilic members of the genus *Campylobacter*. J. Gen. Microbiol. **128**:2515–2522.
2. **Bourgeois, P. L., M. Mata, and P. Ritzenthaler.** 1989. Genome comparison of Lactococcus strains by pulsed-field gel electrophoresis. FEMS Microbiol. Lett. **59**:65–70.
3. **Burgin, A. B., K. Parodos, D. J. Lane, and N. R. Pace.** 1990. The excision of intervening sequences from *Salmonella* 23S ribosomal RNA. Cell **60**:405–414.
4. **Butler, P. D., and E. R. Moxon.** 1990. A physical map of the genome of *Haemophilus influenzae* type b. J. Gen. Microbiol. **136**:2333–2342.
5. **DeBry, R. W.** 1992. The consistency of several phylogeny-inference methods under varying evolutionary rates. Mol. Biol. Evol. **9**:537–551.
6. **De Rijk, P., J. M. Neefs, Y. Van de Peer, and R. D. Wachter.** 1992. Compilation of small ribosomal subunit RNA sequences. Nucleic Acids Res. **20**(Suppl.):2075–2089.
7. **Eaton, K. A., F. E. Dewhirst, M. J. Radin, J. G. Fox, B. J. Paster, S. Krakowka, and D. R. Morgan.** 1993. *Helicobacter acinonyx* sp. nov., a new species of *Helicobacter* isolated from cheetahs with gastritis. Int. J. Syst. Bacteriol. **43**:99–106.
8. **Egebjerg, J., N. Larsen, and R. A. Garrett.** 1990. Structural map of 23S rRNA, p. 168–179. *In* W. E. Hill, P. B. Moore, A. Dahlberg, D. Schlessinger, R. A. Garrett, and J. R. Warner (ed.), The ribosome. Structure, function, and evolution. American Society for Microbiology, Washington, D.C.
9. **Etoh, Y., F. E. Dewhirst, B. J. Paster, A. Yamamoto, and N. Goto.** 1993. *Campylobacter showae* sp. nov., isolated from the human oral cavity. Int. J. Syst. Bacteriol. **43**:631–639.
10. **Eyers, M., S. Chapelle, G. V. Camp, H. Goossens, and R. D. Wachter.** 1993. Discrimination among thermophilic *Campylobacter* species by polymerase chain reaction amplification of 23S rRNA gene fragments. J. Clin. Microbiol. **31**:3340–3343.
11. **Felsenstein, J.** 1989. PHYLIP—phylogeny inference package (version 3.2). Cladistics **5**:164–166.
12. **Gray, M. W., and M. N. Schnare.** 1990. Evolution of the modular structure of rRNA, p. 589–597. *In* W. E. Hill, P. B. Moore, A. Dahlberg, D. Schlessinger, R. A. Garrett, and J. R. Warner (ed.), The ribosome. Structure, function, and evolution. American Society for Microbiology, Washington, D.C.
13. **Guerry, P., S. M. Logan, and T. J. Trust.** 1988. Genomic rearrangements associated with antigenic variation in *Campylobacter coli*. J. Bacteriol. **170**:316–319.
14. **Gustafson, C. E., C. J. Thomas, and T. J. Trust.** 1992. Detection of *Aeromonas salmonicida* from fish by using polymerase chain reaction amplification of the virulence surface array protein gene. Appl. Environ. Microbiol. **58**:3816–3825.
15. **Gutell, R. R.** 1992. Evolutionary characteristics of 16S and 23S rRNA structures, p. 243–309. *In* H. Hartman and K. Matsuno (ed.), The origin and evolution of the cell. World Scientific, Singapore.
16. **Gutell, R. R., N. Larsen, and C. R. Woese.** 1994. Lessons from an evolving RNA: 16S and 23S rRNA structures from a comparative perspective. Microbiol. Rev. **58**:10–26.
17. **Gutell, R. R., A. Power, G. Z. Hertz, E. J. Putz, and G. D. Stormo.** 1992. Identifying constraints on the higher-order structure of RNA: continued development and application of comparative sequence analysis methods. Nucleic Acids Res. **20**:5785–5795.
18. **Gutell, R. R., M. N. Schnare, and M. W. Gray.** 1993. A compilation of large subunit (23S- and 23S-like) ribosomal RNA structures. Nucleic Acids Res. **21**(Suppl.):3055–3074.
19. **Gutell, R. R., and C. R. Woese.** 1990. Higher order structural elements in ribosomal RNAs: pseudoknots and the use of noncanonical pairs. Proc. Natl. Acad. Sci. USA **87**:663–667.
20. **Haraszthy, V. I., G. J. Sunday, L. A. Bobek, T. S. Motley, H. Preus, and J. J. Zambon.** 1992. Identification and analysis of the gap region in the 23S ribosomal RNA from *Actinobacillus actinomycetemcomitans*. J. Dent. Res. **71**:1561–1568.
21. **Hooknikanne, J., M. L. Solin, T. U. Kosunen, and M. Kaartinen.** 1991. Comparison of partial-16S rRNA sequences of different *Helicobacter pylori* strains, *Helicobacter mustelae* and a gastric *Campylobacter*-like organism (GCLO). Syst. Appl. Microbiol. **14**:270–274.
22. **Hull, R. A., R. E. Gill, P. Hsu, B. H. Minshew, and S. Falkow.** 1981. Construction and expression of recombinant plasmids encoding type 1 or D-mannose-resistant pili from a urinary tract infection *Escherichia coli* isolate. Infect. Immun. **33**:933–938.
23. **Kim, N. W., H. Bingham, R. Khawaja, H. Louie, E. Hani, K. Neote, and V. L. Chan.** 1992. Physical map of *Campylobacter jejuni* TGH9011 and localization of 10 genetic markers by use of pulsed-field gel electrophoresis. J. Bacteriol. **174**:3494–3498.
24. **Kim, N. W., and V. L. Chan.** 1989. Isolation and characterization of the ribosomal RNA genes of *Campylobacter jejuni*. Curr. Microbiol. **19**:247–252.
25. **Kim, N. W., R. Lombardi, H. Bingham, E. Hani, H. Louie, D. Ng, and V. L. Chan.** 1993. Fine mapping of the three rRNA operons on the updated genomic map of *Campylobacter jejuni* TGH9011 (ATCC 43431). J. Bacteriol. **175**:7468–7470.
26. **Krawiec, S., and M. Riley.** 1990. Organization of the bacterial chromosome. Microbiol. Rev. **54**:502–539.
27. **Labigne-Roussel, A., P. Courcoux, and L. Tompkins.** 1988. Gene disruption as a feasible approach for mutagenesis of *Campylobacter jejuni*. J. Bacteriol. **170**:1704–1708.
28. **Lane, D. J., A. P. Harrison, D. Stahl, B. Pace, S. J. Giovannoni, G. J. Olsen, and N. R. Pace.** 1992. Evolutionary relationships among sulfur- and iron-oxidizing eubacteria. J. Bacteriol. **174**:269–278.
29. **Larsen, N., G. J. Olsen, B. L. Maidak, M. J. McCaughey, R. Overbeek, T. J. Mackie, T. L. Marsh, and C. R. Woese.** 1993. The ribosomal database project. Nucleic Acids Res. **21**:3021–3023.
30. **Lau, P. P., B. DeBrunner-Vossbrinck, B. Dunn, K. Motto, M. T. MacDonell, D. M. Rollins, C. J. Pillidge, R. B. Hespell, R. R. Colwell, M. L. Sogin, and G. E. Fox.** 1987. Phylogenetic diversity and position of the genus *Campylobacter*. Syst. Appl. Microbiol. **9**:231–238.
31. **Logan, S. M., T. J. Trust, and P. Guerry.** 1989. Evidence for posttranslational modification and gene duplication of *Campylobacter* flagellin. J. Bacteriol. **171**:3031–3038.

31a. **Maciukenas, M.** Unpublished data.

32. **Marrs, B., and S. Kaplan.** 1970. 23S precursor ribosomal RNA of *Rhodopseudomonas sphaeroides*. J. Mol. Biol. **49**:297–317.
33. **Murray, R. G. E., D. J. Brenner, R. R. Colwell, P. De Vos, M. Goodfellow, P. A. D. Grimont, N. Pfennig, E. Stackebrandt, and G. A. Zavarzin.** 1990. Report of the ad hoc committee on approaches to taxonomy within the Proteobacteria. Int. J. Syst. Bacteriol. **40**:213–215.
34. **Needleman, S. B., and C. D. Wunsch.** 1970. A general method applicable to the search for similarities in the amino acid sequence of two proteins. J. Mol. Biol. **48**:443–453.
35. **Noller, H. F.** 1984. Structure of ribosomal RNA. Annu. Rev. Biochem. **53**:119–162.
36. **Noller, H. F., J. Kop, V. Wheaton, J. Brosius, R. R. Gutell, A. M. Kopylov, F. Dohme, and W. Herr.** 1981. Secondary structure model for 23S ribosomal RNA. Nucleic Acids Res. **9**:6167–6189.
37. **Nuijten, P. J. M., C. Bartels, N. M. C. Bleuminkpluym, W. Gaastra, and B. A. M. Vanderzeijst.** 1990. Size and physical map of the *Campylobacter jejuni* chromosome. Nucleic Acids Res. **18**:6211–6214.
38. **Olsen, G. J., H. Matsuda, R. Hagstrom, and R. Overbeek.** 1993. fastDNAml: a tool for construction of phylogenetic trees of DNA sequences using maximum likelihood. Comput. Appl. Biosci. **10**:41–48.

39. **Olsen, G. J., and C. R. Woese.** 1994. The winds of (evolutionary) change: breathing new life into microbiology. J. Bacteriol. **176:**1–6.
40. **Owen, R. J., and S. Leaper.** 1981. Base composition, size and nucleotide sequence similarities of genome deoxyribonucleic acids of the genus *Campylobacter*. FEMS Microbiol. Lett. **12:**395–400.
41. **Pace, N. R., and A. B. Burgin.** 1990. Processing and evolution of the rRNAs, p. 417–425. *In* W. E. Hill, P. B. Moore, A. Dahlberg, D. Schlessinger, R. A. Garrett, and J. R. Warner (ed.), The ribosome. Structure, function, and evolution. American Society for Microbiology, Washington, D.C.
42. **Paster, B. J., and F. E. Dewhirst.** 1988. Phylogeny of campylobacters, wolinellas, *Bacteroides gracilis*, and *Bacteroides ureolyticus* by 16S ribosomal ribonucleic acid sequencing. Int. J. Syst. Bacteriol. **38:**56–62.
43. **Penner, J. L.** 1988. The genus *Campylobacter*: a decade of progress. Clin. Microbiol. Rev. **1:**157–172.
44. **Ralph, D., and M. McClelland.** 1993. Intervening sequence with conserved open reading frame in eubacterial 23S rRNA genes. Proc. Natl. Acad. Sci. USA **90:**6864–6868.
45. **Reddy, K. J., and M. Gilman.** 1988. Preparation and analysis of RNA, p. 4.0.1–4.4.7. *In* F. M. Ausubel, R. Brent, R. E. Kingston, D. D. Moore, J. G. Seidman, J. A. Smith, and K. Struhl (ed.), Current protocols in molecular biology. Greene Publishing Associates and Wiley-Interscience, New York.
46. **Robinson, A., and J. Sykes.** 1971. A study of the atypical ribosomal RNA components of *Rhodopseudomonas sphaeroides*. Biochim. Biophys. Acta **238:**99–115.
47. **Roller, C., W. Ludwig, and K. H. Schleifer.** 1992. Gram-positive bacteria with a high DNA G+C content are characterized by a common insertion within their 23S rRNA genes. J. Gen. Microbiol. **138:**1167–1175.
48. **Romaniuk, P. J., and T. J. Trust.** 1987. Identification of *Campylobacter* species by Southern genomic hybridization of genomic DNA using an oligonucleotide probe for 16S rRNA genes. FEMS Microbiol. Lett. **43:**331–335.
49. **Romaniuk, P. J., and T. J. Trust.** 1989. Rapid identification of *Campylobacter* species using oligonucleotide probes to 16S ribosomal RNA. Mol. Cell. Probes **3:**133–142.
50. **Romaniuk, P. J., B. Zoltowska, T. J. Trust, D. J. Lane, G. J. Olsen, N. R. Pace, and D. A. Stahl.** 1987. *Campylobacter pylori*, the spiral bacterium associated with human gastritis, is not a true *Campylobacter* sp. J. Bacteriol. **169:**2137–2141.
51. **Sambrook, J., E. F. Fritsch, and T. Maniatis.** 1989. Molecular cloning: a laboratory manual, 2nd ed. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
52. **Skurnik, M., and P. Toivanen.** 1991. Intervening sequences (IVSs) in the 23S ribosomal RNA genes of pathogenic *Yersinia enterocolitica* strains. The IVSs in *Y. enterocolitica* and *Salmonella typhimurium* have a common origin. Mol. Microbiol. **5:**585–593.
53. **Smith, C. L., J. G. Econome, A. Schutt, S. Klco, and C. R. Cantor.** 1987. A physical map of the *Escherichia coli* K-12 genome. Science **236:**1448–1453.
54. **Stackebrandt, E., V. Fowler, H. Mill, and A. Kröger.** 1987. 16S rRNA analysis and the phylogenetic position of *Wolinella succinogenes*. FEMS Microbiol. Lett. **40:**269–272.
55. **Stanley, J., D. Linton, A. P. Burnens, F. E. Dewhirst, R. J. Owen, A. Porter, S. L. On, and M. Costas.** 1993. *Helicobacter canis* sp. nov., a new species from dogs: an integrated study of phenotype and genotype. J. Gen. Microbiol. **139:**2495–2504.
56. **Taylor, D. E.** 1992. Genetics of *Campylobacter* and *Helicobacter*. Annu. Rev. Microbiol. **46:**35–64.
57. **Taylor, D. E., M. Eaton, W. Yan, and N. Chang.** 1992. Genome maps of *Campylobacter jejuni* and *Campylobacter coli*. J. Bacteriol. **174:**2332–2337.
58. **Thompson, L. M., III, R. M. Smibert, J. L. Johnson, and N. R. Krieg.** 1988. Phylogenetic study of the genus *Campylobacter*. Int. J. Syst. Bacteriol. **38:**190–200.
59. **Turmel, M., R. R. Gutell, J.-P. Mercier, C. Otis, and C. Lemieux.** 1993. Analysis of the chloroplast large subunit ribosomal RNA gene from 17 *Chlamydomonas* taxa. Three internal transcribed spacers and 12 group I intron insertion sites. J. Mol. Biol. **232:**446–467.
60. **Van Camp, G., Y. Van de Peer, J. Neefs, P. Vandamme, and R. De Wachter.** 1993. Presence of internal transcribed spacers in the 16S and 23S rRNA genes of *Campylobacter*. Syst. Appl. Microbiol. **16:**361–368.
61. **Vandamme, P., and J. De Ley.** 1991. Proposal for a new family, *Campylobacteraceae*. Int. J. Syst. Bacteriol. **41:**451–455.
62. **Vandamme, P., E. Falsen, R. Rossau, B. Hoste, P. Segers, R. Tytgat, and J. De Ley.** 1991. Revision of *Campylobacter*, *Helicobacter*, and *Wolinella* taxonomy: emendation of generic descriptions and proposal of *Arcobacter* gen. nov. Int. J. Syst. Bacteriol. **41:**88–103.
63. **Wesley, I. V., R. D. Wesley, M. Cardella, F. E. Dewhirst, and B. J. Paster.** 1991. Oligodeoxynucleotide probes for *Campylobacter fetus* and *Campylobacter hyointestinalis* based on 16S rRNA sequences. J. Clin. Microbiol. **29:**1812–1817.
64. **Woese, C. R.** Unpublished data.